

Structure and Function of Snake Venom Metalloproteinase Family

武谷, 浩之

<https://doi.org/10.11501/3063849>

出版情報：九州大学，1992，博士（理学），論文博士
バージョン：
権利関係：



Structure and Function of Snake Venom
Metalloproteinase Family

武 谷 浩 之

①

Structure and Function of Snake Venom Metalloproteinase Family

Hiroyuki Takeya

Department of Biology, Faculty of Science,
Kyushu University, Japan

1992

CONTENTS

ABBREVIATIONS	-----	3
PREFACE	-----	4
PART I : The High Molecular Mass Hemorrhagic Protein, HR1B, Isolated from the Venom of <i>Trimeresurus flavoviridis</i>	-----	6
PART II : Coagulation Factor X Activating Enzyme from Russell's Viper Venom (RVV-X) : A NOVEL METALLOPROTEINASE WITH DISINTEGRIN (PLATELET AGGREGATION INHIBITOR)-LIKE AND C-TYPE LECTIN- LIKE DOMAIN	-----	72
CONCLUSION	-----	116
ACKNOWLEDGMENTS	-----	118

ABBREVIATIONS

HR2a, hemorrhagic principle 2a;

Ht-d, hemorrhagic toxin d;

T., *Trimeresurus*;

C., *Crotalus*;

GPIIb/IIIa, platelet glycoprotein IIb/IIIa;

SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis;

HPLC, high performance liquid chromatography;

CNBr, cyanogen bromide;

PTH, phenylthiohydantoin;

RVV-X, the factor X activating enzyme from Russell's viper venom;

RVV-V, the factor V activating enzyme from Russell's viper venom;

Gla, γ -carboxyglutamic acid;

LC1, light chain 1;

IX/X-bp, factor IX/factor X-binding protein.

PREFACE

Most snake venoms can be classified into several fundamental groups according to the main pathophysiological effects they manifest. Venoms of snakes belonging to families *Elapidae* (cobras, kraits, mambas, etc.) and *Hydrophiidae* (sea snakes) are highly neurotoxic and produce flaccid paralysis and respiratory failure in animals. Neurotoxins from these snake venoms have been extensively studied by many investigators. In contrast, *Viperidae* (viper) and *Crotalidae* (pit viper) venoms produce striking local effects, consisting of hemorrhage, necrosis, and edema, and often induce marked alterations of blood coagulation system as well. Among these pathological effects, hemorrhage is a most common occurrence in a victim bitten by crotalid and viperid snakes and various components, such as hemorrhagic factors and metalloproteinases, involved in these venoms have been isolated and characterized. These factors cause localized hemorrhage by direct actions on the blood vessel wall. Electron microscopic studies indicate that erythrocytes are leaked in a one-by-one fashion through widened inter-endothelial gaps when capillaries are exposed to these hemorrhagic proteins. The enzymes may disrupt the pericellular basement membrane through a proteolytic activity and with subsequent damage to the integrity of the vessel wall after which hemorrhage occurs. On the other hand, crotalid and viperid venoms contain many non-hemorrhagic metalloproteinases which act as procoagulants having very strict substrate specificities such as Russell's viper venom factor X activator (RVV-X).

In the 1970 era, four hemorrhagic metalloproteinases HR1A, HR1B, HR2a, and HR2b, and one non-hemorrhagic H₂-proteinase have been purified from the venom of *Trimeresurus (T) flavoviridis*. In the

earlier reports, we established the primary structures of HR2a (202 residues) and H₂-proteinase (201 residues), and other hemorrhagic metalloproteinase HT-2 (202 residues), isolated from the venom of *Crotalus ruber ruber*. All of these enzymes are very similar in sequence to each other and contain the putative active site sequence His-Glu-X-X-His that is homologous to the active sites of thermolysin and several other related bacterial enzymes as well as mammalian zinc metalloproteinases. However, since no significant sequence similarity beyond this region is found with any other known metalloproteinases, the snake venom enzymes belong to newly identified metalloproteinase subfamily. Among the venom metalloproteinases, the high molecular mass (Mr 60,000) hemorrhagic protein HR1B expresses 10 times higher hemorrhagic activities than does HR2a, thereby indicating major lethal factors in *T. flavoviridis* venom. In addition, HR1 (mixture of HR1A and HR1B) as well as the crude venom inhibit ADP-stimulated platelet aggregation.

In Part I of this thesis, the amino acid sequence of HR1B has been determined to explain these characteristic features of HR1B. In Part II, the entire amino acid sequence of RVV-X has been determined in order to elucidate the molecular mechanism of the RVV-X-catalyzed factor X activation, in particular how RVV-X specifically recognizes factor X.

PART I

The High Molecular Mass Hemorrhagic Protein, HR1B, Isolated from the Venom of *Trimeresurus flavoviridis*

SUMMARY

Hemorrhage is a common occurrence in a victim bitten by crotalid and viperid snakes, and hemorrhagic components in these various venoms have been isolated and characterized. Previously, we have shown that a low molecular weight hemorrhagic protein (HR2a, 202 amino acid residues) isolated from the venom of *Trimeresurus flavoviridis* is a member of new subfamily of metalloproteinases. We now report the complete amino acid sequence of a high molecular mass hemorrhagic protein isolated from the same venom. This protein, HR1B, is a mosaic protein composed of 416 residues containing 4 Asn-linked oligosaccharide chains. The amino-terminal half (residues 1-203) of HR1B contains a metalloproteinase domain, the sequence of which is 62 % identical to that of HR2a and 52 % identical to that of Ht-d isolated from the *Crotalus atrox* venom. The most interesting finding is that the middle region (residues 204-300) of HR1B shows a striking similarity to disintegrins, Arg-Gly-Asp-containing platelet aggregation inhibitors, recently found in several viper venoms. Interestingly, however, this region of HR1B does not contain the Arg-Gly-Asp-sequence which is known to be a putative binding site in the disintegrins for the platelet fibrinogen receptor, glycoprotein IIb/IIIa complex. We also found that the carboxy-terminal region (residues 213-336) of the middle part of HR1B shows 30 % identity to residues 1543-1656 of von Willebrand factor, and that the remaining region at the carboxyl-terminal end is unique and has a cysteine-rich sequence. These results suggest that the middle portion of HR1B, which shows structural similarities to the disintegrins and von Willebrand factor, may be important in synergistically stimulating hemorrhagic activity in the NH₂-terminal metalloproteinase domain.

INTRODUCTION

The neurotoxic effects of cobra and sea snake venoms are now well understood and purified neurotoxins serve as tools for the investigation of mammalian neuronal systems (1). In contrast, the local effects of crotalid and viperid venoms, toxins which cause severe hemorrhage, necrosis and edema, are poorly understood at the molecular level. Among factors responsible for these effects, hemorrhagic proteins have been highly purified and characterized (2). These proteins cause localized hemorrhage by direct actions on the blood vessel wall. Electron microscopic studies have shown that erythrocytes are broken in a one-by-one fashion through widened inter-endothelial gaps when capillaries are exposed to these hemorrhagic proteins (3, 4). Moreover, hemorrhagic proteins degrade basement membrane preparations (5) as well as isolated components, including type IV collagen, laminin, nidogen, and fibronectin (6). These findings suggest that hemorrhagic proteins may disrupt the pericellular basement membrane through proteolytic activity and with subsequent damage to the integrity of the vessel wall after which hemorrhage occurs (6, 7). The purified venom proteins should provide an appropriate tool for elucidating mechanisms related to the oozing of erythrocytes and plasma proteins from microcirculatory systems.

In 1989 Shannon et al. (6) and our group (7) reported the primary structures of low molecular weight hemorrhagic proteins (Mr 23,000), named hemorrhagic principle 2a (HR2a) and hemorrhagic toxin d (Ht-d), isolated from the venom of *Trimeresurus flavoviridis* and *Crotalus atrox*, respectively. The structures revealed that both are members of a newly identified subfamily of

metalloproteinases which have peptide segments similar to part of the zinc-chelating site and one of the catalytic residues of thermolysin. However, there is no significant sequence similarity to thermolysin or any other known metalloproteinases, except for the zinc chelating sequence.

High molecular weight (Mr 60,000-90,000) hemorrhagic proteins have also been isolated from the venom of various species of snakes including *Agkistrodon halys blomhoffii* (8, 9), *C. atrox* (10), and *C. horridus horridus* (11), all of which have potent hemorrhagic activity. Two high molecular weight hemorrhagic proteins, named HR1A and HR1B and isolated from the venom of *T. flavoviridis* (12, 13), are closely related, if not identical, immunologically. They have molecular weights of 60,000. HR1A and HR1B give LD₅₀ values of 7.2 and 4.9 mg/mouse, respectively, and show 10 times higher hemorrhagic activities than does HR2a, thereby indicating that they are major lethal factors in *T. flavoviridis* venom (13). In addition to these pathological functions, HR1 (a mixture of HR1A and HR1B) as well as the crude venom inhibit ADP-stimulated platelet aggregation (14). HR1 has the potential to inhibit 60 % of platelet aggregation at a concentration of 6 ng/ml, in platelet rich plasma. *In situ* microscopic observations also reveal that hemorrhagic proteins present in venom induce hemorrhage with little or no formation of white thrombi at the site of the hemorrhage (3, 4).

In ongoing work to elucidate the structure and function relationships of hemorrhagic components, we determined the complete amino acid sequence of HR1B and compared the findings with the low molecular mass hemorrhagic protein, HR2a. We obtained evidence that HR1B is a mosaic protein consisting of an

amino-terminal metalloproteinase domain, a large noncatalytic middle segment with a trigramin-like structure (15-17) and a unique carboxyl-terminal Cys-rich domain. This is apparently the first report of the entire sequence of a high molecular mass hemorrhagic protein detected in crotalid and viperid venoms.

MATERIALS AND METHODS

Materials ——— The sources of materials used were as follows: *Achromobacter lyticus* lysyl endopeptidase, 4-vinylpyridine and succinic anhydride from Wako Pure Chemical Industries, Osaka; endoproteinase Asp-N from Boehringer-Mannheim Biochemica, FRG; L-pyroglutamylpeptidase from Suntory, Osaka; trypsin treated with *N*-tosyl-phenylalanyl chloromethyl ketone and α -chymotrypsin from Worthington Biochemical, Freehold, N. J.; arginylendopeptidase from Takara Shuzo Co. Ltd., Kyoto; cyanogen bromide, *o*-iodosobenzoic acid and Cosmosil (3C18, 5C4) columns from Nacalai Tesque, Inc., Kyoto; TSK ODS-120T and Phenyl-5pw RP columns from Toyo Soda Manufacturing Co. Ltd., Osaka; a Vydac 214TP5415 column from The Separations Group, CA; reagents for gas-phase sequencer from Applied Biosystems, CA. All other chemicals were of analytical grade or of the highest quality commercially available.

Purification of HR1B ——— HR1B was highly purified from *T. flavoviridis* venom, as described (13). Briefly, the purification method consisted of three steps, gel filtration on Sephadex G-100, ion-exchange chromatography on DEAE-Sephadex A-50, and gel filtration on Sephadex G-200 superfine. The purified preparation gave a single band on SDS-PAGE under reduced conditions and the apparent molecular mass was estimated to be approximately Mr 60,000 (data not shown).

S-Pyridylethylation of HR1B ——— HR1B (200 nmol) was incubated in 6 ml of 6 M guanidine hydrochloride containing 0.5 M Tris-HCl, 10 mM EDTA, and 40 mM dithiothreitol, pH 8.6, under nitrogen for 3 h at 50 °C. 4-vinylpyridine (1.2 nmol) was then

added and the mixture was further incubated in the dark for 2.5 h at room temperature. The excess reagents were removed by dialysis against distilled water.

CNBr Digestion ——— Pe-HR1B (100 nmol) in 2 ml of 70 % formic acid was treated with CNBr, a 100-fold molar excess over methionine residues under nitrogen, and incubated at room temperature for 20 h in the dark. The reaction was terminated by lyophilization.

Subdigestion with Lysyl Endopeptidase ——— CNBr fragments, M1, M5, and M6 (10 nmol each) were subdigested with lysyl endopeptidase. The peptides in 50 mM Tris-HCl buffer, pH 9.0, containing 2 M urea were digested with an enzyme/substrate ratio of 1/50 (mol/mol) at 37 °C for 18 h.

Subdigestion with Endoproteinase Asp-N ——— CNBr fragments, M1, M5, and M6 (6 nmol each) were subdigested with endoproteinase Asp-N. The peptides in 50 mM Tris-HCl buffer, pH 8.0, containing 1 M urea were digested with an enzyme/substrate ratio of 1/150 (mol/mol) at 37 °C for 18 h.

Fragmentation with o-Iodosobenzoic Acid ——— CNBr fragment M1 was cleaved at tryptophanyl peptide bonds. The peptide (6.5 nmol) was incubated in 0.2 ml of 80 % acetic acid containing 4 M guanidine hydrochloride and 4.5 mM o-iodosobenzoic acid for 24 h at room temperature, in the dark (18).

Tryptic Digestion ——— Pe-HR1B was succinylated to protect the cleavage of lysyl peptide bonds before tryptic digestion. The Pe-protein (50 nmol) in 3 ml of 0.1 M NaHCO₃, pH 8.3, containing 8 M urea was treated with succinic anhydride, a 3-fold molar excess over lysine residues. After addition of succinic anhydride, pH of the mixture dropped; it was adjusted to

by adding NaOH. The solution was allowed to stand for 1.5 h at room temperature. The O-succinyl groups were removed by 1 M hydroxylamine, pH 8.5, for 3 h. The succinylated pe-HR1B in 0.75 ml of 50 mM Tris-HCl buffer, pH 7.5, containing 4 M urea was digested by trypsin with an enzyme/substrate ratio of 1/150 (mol/mol) at 37 °C for 23 h.

Digestion with Arginylendopeptidase ——— Pe-HR1B (50 nmol) in 0.9 ml of 50 mM Tris-HCl buffer, pH 8.0, containing 2 M urea was digested with an enzyme/substrate ratio of 1/150 (mol/mol) at 37 °C for 6 h.

Peptide Purification ——— Peptides were purified by reversed-phase HPLC in 0.1 % trifluoroacetic acid with acetonitrile gradient elution at a flow rate of 0.5 ml per min. Details are the same as described (19).

Amino Acid Analysis and Sequence Determination ——— The amino acid analysis of the pe-protein was performed by ion-exchange chromatography in a Hitachi model L-8500 high speed amino acid analyzer after hydrolysis with 5.7 M HCl containing 0.2 % phenol at 110 °C for 24, 48, and 72 h by the method of Spackman et al. (20). Tryptophan was determined by hydrolysis in 3 N mercaptoethanesulfonic acid (21). The CNBr fragments and arginylendopeptidase fragments were analyzed using a Hitachi L-8500 amino acid analyzer after hydrolysis with 5.7 M HCl containing 0.2 % phenol at 110 °C for 24 h. The tryptic peptides and subdigested peptides were analyzed using reversed-phase HPLC of phenylthiocarbamoyl derivatives (22), using a Waters PICO-TAG system, after hydrolysis with 5.7 M HCl containing 1 % phenol at 110 °C for 20 h. Automated sequence analyses were performed with an Applied Biosystems 477A protein sequencer, as described by

Hewick et al (23), with an Applied Biosystems model 120A PTH analyzer.

Determination of the Carboxyl-Terminal Amino Acid

The carboxyl-terminal amino acid of HR1B was determined by the vapor-phase hydrazinolysis method (24). Highly dried 10 nmol of protein was treated with vaporized hydrazine at 90 °C for 3 h. The hydrazinolysate was then treated with benzaldehyde. The amino acid released from the carboxyl-terminus of the protein was identified with a Hitachi model L-8500 amino acid analyzer.

Nomenclature of the Peptides ——— Peptides are designated by a serial number prefixed by a letter. The letters indicated the type of digestion: M, CNBr; K, lysyl endopeptidase; D, endoproteinase Asp-N; R, arginylendopeptidase; ST, tryptic digestion of succinylated protein; W, o-iodosobenzoic acid. The numbers in the peptide designation do not correspond to the order of their elution in HPLC, but rather to their positions in the protein sequence, starting from the N-terminus.

RESULTS

Amino Acid Composition and Sequence Analysis of HR1B — The amino acid composition of HR1B determined from the sequence as well as that obtained by amino acid analysis is shown in Table I. Based on the analysis, HR1B contained glucosamine but not galactosamine. HR1B had a high content of cysteine. Sequence analyses of intact protein (1 nmol) and Pe-protein (2 nmol) revealed no PTH derivatives up to 5-cycles of Edman degradation, thereby indicating that the amino-terminus is blocked, as are those of HR2a (7) and Ht-d (6). By treatment of intact HR1B with pyroglutamylpeptidase, the amino-terminal pyroglutamyl residue was released to give HR1B PG and the resultant protein HR1B PG was sequenced (Fig. 1). As with findings with HR2a (7) and Ht-d (6), sequence analysis of HR1B PG revealed the presence of a variant of the sequence lacking glutamine at position 2. Vapor-phase hydrazinolysis of HR1B yielded 0.11 mol of alanine, 0.07 mol of lysine and 0.13 mol of serine per mol of protein (uncorrected), hence, there is probably microheterogeneity at the carboxyl-terminus of HR1B.

CNBr Peptides — The entire amino acid sequence of HR1B was determined primarily from a set of methionyl-cleavage fragments (Fig. 1). The CNBr digest of pe-HR1B was separated by reversed-phase HPLC on a column of Vydac 214TP5415 (Fig. 2). Peptides M4-5 and M5 were separated by rechromatography on a column of ODS-120T (Fig. 3). The minor peptide M7" was separated from M7' by rechromatography as shown in Fig. 4. The amino acid compositions and sequences of obtained peptides are shown in Tables II and III, and Fig. 1. M5-DP was a peptide which was

partially hydrolyzed at Asp-Pro bond in 70 % formic acid during digestion. M4-5 and M3-4-5 were yielded by incomplete cleavage at a Met-Ile bond (residues 170-171), and a Met-Ser bond (residues 167-168). Similar but not identical peptides, M7, M7' and M7'' did not contain homoserine in their compositions, thereby indicating that these peptides originated from the carboxyl-terminal end. Amino acid sequences of M7, M7' and M7'' were complete, showing that M7' contained an extended sequence of Tyr-Lys and M7'' had a more extended sequence of Tyr-Lys-Ser from the carboxyl-terminus of M7. As described above, the result obtained on hydrazinolysis of intact protein indicated three carboxyl-terminal residues, Ala, Lys and Ser. Thus, the carboxyl-terminal ragged ends of HR1B, such as, -Ala, -Ala-Tyr-Lys and -Ala-Tyr-Lys-Ser, were confirmed. The amino acid sequences of CNBr peptides, M1, M5 and M6 were determined completely from two overlapping sets of peptides generated on digestion with lysyl endopeptidase and endoproteinase Asp-N.

Subdigestion of M1 ——— Subdigests of M1 with lysyl endopeptidase yielded 8 peptides (Fig. 5). The amino acid compositions and sequences of the peptides are shown in Tables III and IV, and Fig. 1. The amino acid composition and blocked amino-terminus indicated that M1K1 originated from the amino-terminal end. After digestion with pyroglutamylpeptidase, the resultant peptide M1K1 PG could be sequenced (Tables III and IV, and Fig. 1). Although the entire HR1B had the amino-terminal ragged end, sequence analysis of M1K1 PG did not reveal the presence of a variant peptide with glutamine at position 2. One possible explanation is that the glutamine residue newly appearing at the amino-terminal end of this variant peptide by the

pyroglutamylpeptidase treatment was cyclized again and this newly cyclized glutamine was further digested by pyroglutamylpeptidase and could be sequenced only from the arginine at position 3. M1K6 and M1K6' gave the same amino acid compositions (Table IV). However, M1K6' could not be sequenced, suggesting that cyclization of the amino-terminal glutamine of M1K6 might have occurred during digestion or HPLC. Subdigests of M1 with endoproteinase Asp-N yielded 6 peptides (Fig. 6). Amino acid compositions of these six peptides are shown in Table V. As shown in Table III and Fig. 1, M1D2 overlapped M1K2, M1K3, M1K4, and M1K5. M1D4 and M1D6 overlapped M1K5, M1K6 and M1K7. M1W1, generated by cleavage at tryptophanyl residues, confirmed the carboxyl-terminal sequence of M1D2. Hence, a continuous sequence of M1 was obtained.

Subdigestion of M5 ——— Subdigests of M5 with lysyl endopeptidase were separated by reversed-phase HPLC (Figs. 7 and 8). The total of 6 peptides were subjected to amino acid and sequence analyses (Tables III and VI, and Fig. 1). Twelve peptides were obtained from subdigests of M5 with endoproteinase Asp-N (Figs. 9-11), and their amino acid and sequence data are shown in Tables III and VII and Fig. 1. M5D7 overlapped M5K3, M5K4 and M5K5. M5D12 filled the carboxyl-terminal portion of M5K5.

Subdigestion of M6 ——— Two subdigestions of M6 were performed by cleavage with lysyl endopeptidase and endoproteinase Asp-N. The resulting fragments were separated by reversed-phase HPLC (Figs. 12-15). Peptides obtained were subjected to structural analyses (Tables III, VIII, and IX and Fig. 1). M6K4 and M6K5 overlapped by M6D6 provided the remainder of the sequence of M6.

Cleavage at Arginyl Residues ——— Alignment of the methionyl-cleavage fragments was obtained by analyses of the peptides isolated from cleavages at the arginyl residues. Pe-HR1B was cleaved by trypsin at arginyl residues after *N*-succinylation. The peptides were separated by reversed-phase HPLC as shown in Fig. 16. The amino acid compositions of the peptides are shown in Table X. The peak just before ST1 seemed to be an uncleaved peptide of ST1 and ST2, as determined by the amino acid composition (data not shown) and the blocked amino terminus. The amino acid sequence of ST2 overlapped M1 and M2 (Table III, and Fig. 1). The sequences of ST4 and ST4-5 overlapped M5K5 and M5D12. The sequence of ST5 overlapped M5D12, M5K6 and M6. Since ST6 overlapped M6K5 and M7, M6 and M7 were linked. The overlap between M2 and M3 was achieved by the sequence analysis of R1, one of the fragments of Pe-HR1B digested with arginylendopeptidase (Fig. 17 and Table XI).

Carbohydrate-Linked Asparagine Residues ——— The PTH derivatives at positions 73, 181, 327 and 380, were not identified by sequence analysis. These are very likely carbohydrate-linked Asn residues since they are followed by a -X-Thr/Ser sequence, a consensus signal sequence for the attachment of carbohydrate to asparagine and because the composition analyses of the fragments containing these residues show the presence of glucosamines in each fragment. The reasonably high yields for Thr and Ser residues in the entire sequence and composition analyses of all the fragments, including intact protein, indicate the absence of O-linked carbohydrate chains in HR1B.

Summary of Sequence Analyses of HR1B ——— Ninety-seven percent of the total sequence was determined by two or three

independent analyses using different peptides, as shown in Fig. 1. HR1B was composed of 416 residues with 4 Asn-linked sugar chains at positions 73, 181, 327 and 380. HR1B has been reported to contain neutral sugar, amino sugar and sialic acid accounting for 17-18 % on a total weight basis (13) and corresponding to 10-11 kilodaltons. The molecular weight for the polypeptide portion of the largest isoform was calculated to be 46,478, thus the molecular weight was approximately 57,000 with the addition of the four carbohydrate chains. This number differed slightly from the reported Mr of 60,000 (13), as estimated by SDS-PAGE. In all likelihood, this difference was due to abnormal migration of glycoproteins on SDS-PAGE (25). The net charge of the polypeptide portion of HR1B was calculated to be +6. Since the isoelectric point of HR1B was determined to be 4.4 (13), the sialic acid residues of the complex carbohydrate moieties must account for the acidic isoelectric point.

Titration of the Free SH-Group ——— Since structural analyses of HR1B gave a total of 35 cysteine residues, as indicated in Table 1 and Fig. 1, the content of the free SH-group was determined by the method of Ellman (26) and incorporation of 4-vinylpyridine in the absence and presence of a denaturant. The values determined by the two methods with Ellman's reagent (5,5'-dithiobis (2-nitrobenzoic acid)) and with S-pyridylethylation were in good agreement, showing that about 1 mol of SH-group per mol of HR1B was titrated only in the presence of denaturant (Table XII). HR1B probably has a free cysteine residue present in the molecule, which is in striking contrast to the case of a low molecular mass of the hemorrhagic protein HR2a.

DISCUSSION

We reported the complete amino acid sequence and the disulfide bridge locations of the hemorrhagic (HR2a) and non-hemorrhagic (H₂-proteinase) metalloproteinases (7, 27), isolated from snake venom used as the source of HR1B. In the same year, the amino acid sequence of Ht-d, a hemorrhagic metalloproteinase from the venom of *C. atrox*, was reported by other investigators (6). These are very similar metalloproteinases consisting of 201-203 residues. As shown in Fig. 18A, the sequence of the amino-terminal half (203 residues) of HR1B represents a metalloproteinase structure similar to the low molecular weight metalloproteinases, HR2a, H₂-proteinase and Ht-d. This metalloproteinase domain of HR1B contains an amino acid sequence (residues 143-147) consisting of the putative zinc-chelating (His-143, 147) and catalytic (Glu-144) residues (Fig. 18A), identified by homology to the zinc-chelating sequence of thermolysin (6, 7, 27). This domain of HR1B has a slightly higher level of identity to the sequences of HR2a (62 %) and H₂-proteinase (62 %) than to that of Ht-d (52 %), probably due to snake species differences. On the other hand, when the sequence identities of the three proteins from the same species are compared, HR2a and H₂-proteinase are more closely related (74 % identity) than are HR2a/H₂ proteinase and HR1B, thereby suggesting that HR1B diverged before the divergence of HR2a (hemorrhagic) and H₂-proteinase (non-hemorrhagic).

The most characteristic feature found in the metalloproteinase domain of HR1B is the presence of two sugar chains (Fig. 18A). One linked to Asn-73 is in a rather

hydrophobic region (residues 40-80) of HR1B, as predicted by the method of Kyte and Doolittle (28). Since the corresponding regions of the three other molecules, HR2a, H₂-proteinase and Ht-d, are also hydrophobic, only HR1B is likely to have hydrophilicity in this region. Although H₂-proteinase also has a potential carbohydrate binding sequence at Asn-73 (Fig. 18A), no carbohydrate moiety was detected (27). The other sugar chain located in HR1B is linked with Asn-181, at a position where Asp residues are followed by the -X-Ser sequence in the three other molecules.

As we reported, HR2a and H₂-proteinase have similar primary structures although the latter does not have hemorrhagic activity (27). We compared the three hemorrhagic metalloproteinases (HR1B, HR2a and Ht-d) and one non-hemorrhagic H₂-proteinase (Fig. 18A), to determine the structural element related to hemorrhaging. However, we found no clear differences between the hemorrhagic and non-hemorrhagic metalloproteinases. While there are few residues in common, at least six are shared by hemorrhagic but not by non-hemorrhagic proteinases (boxed residues in Fig. 18A). Such conserved residues may possibly be related to the induction of hemorrhagic activity. Further investigations are ongoing to evaluate the significance of these specific residues in relation to the substrate specificity of hemorrhagic proteinases.

Platelet aggregation inhibitory peptides, termed trigramin, echistatin, applaggin and bitistatin, have been isolated from the venom of *T. gramineus*, *Echis carinatus*, *Agkistrodon piscivorus piscivorus* and *Bitis arietans*, respectively, and the amino acid sequences have been determined (15-17, 29-32). These peptides are homologous cysteine rich peptides consisting of 49-83 residues. A

highly conserved structural feature in these peptides is the tripeptide unit Arg-Gly-Asp (see Fig. 18B), a sequence that appears in a variety of adhesive protein ligands, such as fibronectin, vitronectin, fibrinogen, and von Willebrand factor, and contributes to their interaction with specific membrane receptors, called integrins (33, 34). These peptides, now known as the disintegrin family (35) are competitive inhibitors of fibrinogen and von Willebrand factor binding to the activated platelet membrane integrin, glycoprotein IIb/IIIa (GPIIb/IIIa) complex (16, 17). Trigramin is about 500 times more potent than a tetra- or hexapeptide containing Arg-Gly-Asp-sequence in inhibiting the binding of fibrinogen to the GPIIb/IIIa (16). It has also been reported that HR1B is a potent inhibitor of platelet aggregation (14). Here, we propose the structural elements for this activity based on the sequence similarity. As shown in Fig. 18B, the cysteine-rich middle portion (residues 204-300) following the metalloproteinase domain of HR1B has a close similarity to the disintegrins. The sequence similarity of HR1B is scored 54 % for trigramin, 34% for echistatin, 57 % for applaggin, and 60 % for bitistatin. Bitistatin most closely resembles the cysteine-rich middle portion of HR1B, in terms of not only the highest degree of sequence identity but also in other respects. Bitistatin is the largest disintegrin (83 residues) to have been isolated and the amino-terminal 10 residues unique to bitistatin are identical to the corresponding residues of HR1B, except for the conservative substitution of Leu to Ile. Two cysteine residues at positions 214 and 233 in HR1B are also present in bitistatin, but not in other disintegrins. It is noteworthy, however, that HR1B has one

additional unique cysteine residue at position 276 not present in bitistatin.

The disintegrins bind to the GPIIb/IIIa complex through the Arg-Gly-Asp-sequence (16, 17). However, the Arg-Gly-Asp-sequence is absent in the HR1B molecule and is replaced by Glu-Ser-Glu-Cys (Fig. 18B). Several lines of evidence suggest that platelet aggregation inhibitory activity resides within the disintegrin-like domain of HR1B, despite the absence of the Arg-Gly-Asp-sequence. First, it is suggested that a particular conformation of disintegrin is a prerequisite for its binding to the GPIIb/IIIa complex, since the chemically reduced sample does not inhibit fibrinogen and von Willebrand factor binding to activated platelets (16, 17, 29). Second, in experiments using synthetic analogues of echistatin, it was noted that the Arg residue in the Arg-Gly-Asp-sequence plays a less important role in the binding of echistatin to activated platelets than does that of the Arg-Gly-Asp-containing tetrapeptide (29). Finally, the tyrosylpentadecapeptide (Tyr-Gly-Gln-His-His-Leu-Gly-Gly-Ala-Lys-Gln-Ala-Gly-Asp-Val) corresponding to the carboxyl-terminal sequence of the fibrinogen γ chain blocks fibrinogen binding and trigramin binding to platelets (16), thereby suggesting that the Arg-Gly-Asp-sequence present in the snake venom disintegrin family is not the only sequence binding to platelets. It remains to be determined which structural segment of HR1B is essential for the inhibition of platelet aggregation. Synthetic fragments or fragments derived from native HR1B should prove useful in such cases. A peptide termed CM-2 has been isolated from crude venom used for the source of bitistatin (36). CM-2 shows sequence identity to disintegrin and interestingly has the sequence of Arg-

Gly-Asn in place of the Arg-Gly-Asp-sequence of disintegrin. Its platelet aggregation inhibitory activity has not yet been documented.

It is also of interest that the sequence of the disintegrin-like domain, including the amino-terminal part of the unknown cysteine-rich region in HR1B, shows a weak but statistically significant similarity to part of the sequence of human von Willebrand factor (Fig. 19). Residues 213-336 of HR1B show 30 % identity to residues 1543-1656 of von Willebrand factor, for which the probability is less than 1.5×10^{-6} that the sequence similarity arose by chance. The von Willebrand factor participates in the initial reactions of hemostasis by mediating the adhesion of platelets to the subendothelium as well as the aggregation of platelets at the sites of vascular injury (37). Therefore, these structural similarities may be important in discerning the hemorrhagic function of HR1B, since both proteins have the same targets, e.g. platelets and the subendothelium.

Trigramin has been noted to inhibit the adhesion of melanoma cells to fibronectin (38). In this regard, the disintegrin-like domain of HR1B may be important in localizing the HR1B molecule to the peripheral blood vessel, especially to integrins located in the basement membrane attachment site of the endothelial cells, when the metalloproteinase domain of HR1B degrades the basement membrane. This localization of HR1B to the subendothelium, in addition to its platelet aggregation inhibitory activity, may be one reason why HR1B expresses 10 times higher hemorrhagic activity than does HR2a (13).

The carboxyl-terminal region (residues 301-416) of HR1B contains numerous cysteine residues and two sugar chains (Fig.

18B). This region is unique and not homologous to any known protein sequence. It seems likely that this region may also play an important role in inhibiting platelet aggregation and the cell adhesion inhibitory activities of HR1B, since the sequence Gln-Glu-Asp-Val (positions 360-363) found in this region resembles the Arg-Glu-Asp-Val-sequence corresponding to the second cell attachment site in fibronectin (39).

The well known matrix-degrading metalloproteinases are members of the collagenase family. They have large noncatalytic hemopexin-like segments attached to the carboxyl-terminal ends of the metalloproteinase domains (40-45), as noted with human interstitial collagenase (40), human stromelysin (41, 42), and rat stromelysin (transin (43)). Type IV collagenase (gelatinase) additionally contains a fibronectin-like collagen binding domain (45). However, in comparison with these mosaic proteinases, there is no significant structural similarity to HR1B. Further experiments are required to determine whether the noncatalytic segment of HR1B is essential for the biological function of this enzyme, and/or for regulating proteolytic activity.

REFERENCES

1. Karlsson, E. (1979) in *Handbook of Experimental Pharmacology* Vol 52, *Snake Venoms* (Lee, C.-Y., ed) pp. 159-219, Springer-Verlag, Berlin, Heidelberg, New York
2. Tu, A.T. (1988) in *Hemostasis and Animal Venoms* (Pirkle, H., and Markland, F.S., Jr., eds) pp. 425-443, Marcel Dekker, New York
3. Tsuchiya, M., Ohshio, C., Ohashi, M., Ohsaka, A., Suzuki, K., and Fujishiro, Y. (1974) in *Platelets, Thrombosis, and Inhibitors* (Didisheim, P., ed) pp. 439-446, F. K. Schattauer Verlag, Stuttgart, New York
4. Ohsaka, A. (1979) in *Handbook of Experimental Pharmacology* (Lee, C.-Y., ed) Vol. 52, pp. 480-546, Springer-Verlag, Berlin
5. Ohsaka, A., Just, M., and Habermann, E. (1973) *Biochim. Biophys. Acta* **323**, 415-428
6. Shannon, J.D., Baramova, E.N., Bjarnason, J.B., and Fox, J.W. (1989) *J. Biol. Chem.* **264**, 11575-11583
7. Miyata, T., Takeya, H., Ozeki, Y., Arakawa, M., Tokunaga, F., Iwanaga, S., and Omori-Satoh, T. (1989) *J. Biochem. (Tokyo)* **105**, 847-853
8. Ohshima, G., Iwanaga, S., and Suzuki, T. (1968) *J. Biochem. (Tokyo)* **64**, 215-225
9. Ohshima, G., Omori-Satoh, T., Iwanaga, S., and Suzuki, T. (1972) *J. Biochem. (Tokyo)* **72**, 1483-1494
10. Bjarnason, J.B., and Tu, A.T. (1978) *Biochemistry* **17**, 3395-3404
11. Civello, D.J., Duong, H.L., and Geren, C.R. (1983) *Biochemistry* **22**, 749-755
12. Omori-Satoh, T., and Ohsaka, A. (1970) *Biochim. Biophys. Acta* **207**, 432-444

13. Omori-Satoh, T., and Sadahiro, S. (1979) *Biochim. Biophys. Acta* **580**, 392-404
14. Yamanaka, M., Matsuda, M., Isobe, J., Ohsaka, A., Takahasi, T., and Omori-Satoh, T. (1974) in *Platelets, Thrombosis, and Inhibitors* (Didisheim, P., ed) pp. 335-344, F. K. Schattauer Verlag, Stuttgart, New York
15. Ouyang, C., and Huang, T.-F. (1983) *Biochim. Biophys. Acta* **757**, 332-341
16. Huang, T.-F., Holt, J.C., Lukasiewicz, H., and Niewiarowski, S. (1987) *J. Biol. Chem.* **262**, 16157-16163
17. Huang, T.-F., Holt, J.C., Kirby, E.P., and Niewiarowski, S. (1989) *Biochemistry* **28**, 661-666
18. Mahoney, W.C., and Hermodson, M.A. (1979) *Biochemistry* **18**, 3810-3814
19. Takeya, H., Kawabata, S., Nakagawa, K., Yamamichi, Y., Miyata, T., Iwanaga, S., Takao, T., and Shimonishi, Y. (1988) *J. Biol. Chem.* **263**, 14868-14877
20. Spackman, D.H., Stein, W.H., and Moore, S. (1958) *Anal. Chem.* **30**, 1190-1206
21. Penke, B., Ferenczi, R., and Kovacs, K. (1974) *Anal. Biochem.* **60**, 45-50
22. Bidlingmeyer, B.A., Cohen, S.A., and Tarvin, T.L. (1984) *J. Chromatogr.* **336**, 93-104
23. Hewick, R.M., Hunkapiller, M.W., Hood, L.E., and Dreyer, W.J. (1981) *J. Biol. Chem.* **256**, 7990-7997
24. Yamamoto, A., Toda, H., and Sakiyama, F. (1989) *J. Biochem. (Tokyo)* **106**, 552-554
25. Segret, J.P. and Jackson, R.L. (1972) *Methods Enzymol.* **28**, 54-63
26. Ellman, G.L. (1969) *Arch. Biochem. Biophys.* **82**, 70-77

27. Takeya, H., Arakawa, M., Miyata, T., Iwanaga, S., and Omori-Satoh, T. (1989) *J. Biochem. (Tokyo)* **106**, 151-157
28. Kyte, J., and Doolittle, R.F. (1982) *J. Mol. Biol.* **157**, 105-132
29. Gan, Z.-R., Gould, R.J., Jacobs, J.W., Friedman, P.A., and Polokoff, M.A. (1988) *J. Biol. Chem.* **263**, 19827-19832
30. Garsky, V.M., Lumma, P.K., Freidinger, R.M., Pitzenberger, S.M., Randall, W.C., Veber, D.F., Gould, R.J., and Friedman, P.A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4022-4026
31. 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. USA* **86**, 8050-8054
32. Shebuski, R.J., Ramjit, D.R., Bencen, G.H., and Polokoff, M.A. (1989) *J. Biol. Chem.* **264**, 21550-21556
33. Ruoslahti, E., and Pierschbacher, M.D. (1987) *Science* **238**, 491-497
34. Hynes, R.O. (1987) *Cell* **48**, 549-554
35. Niewiarowski, S., Huang, T.-F., Rucinski, B., Cook, J.J., Williams, J.A., Musial, J., Edmunds, L.H., Jr., Gould, R.J., Bush, L., Shebuski, R., and Friedman, P.A. (1989) *Thrombosis and Haemostasis* **62**, 319
36. Joubert, F.J., Haylett, T., Strydom, D.J., and Taljaard, N. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **263**, 1087-1096
37. Titani, K., Kumar, S., Takio, K., Ericsson, L.H., Wade, R.D., Ashida, K., Walsh, K.A., Chopek, M.W., Sadler, J.E., and Fujikawa, K. (1986) *Biochemistry* **25**, 3171-3184
38. Knudsen, K.A., Tuszynski, G.P., Huang, T.-F., and Niewiarowski, S. (1988) *Exp. Cell Res.* **179**, 42-49
39. Humphries, M.J., Akiyama, S.K., Komoriya, A., Olden, K., and Yamada, K.M. (1986) *J. Cell Biol.* **103**, 2637-2647

40. Goldberg, G.I., Wilhelm, S.M., Kronberger, A., Bauer, E.A., Grant, G.A., and Eisen, A.Z. (1986) *J. Biol. Chem.* **261**, 6600-6605
41. Whitham, S.E., Murphy, G., Angel, P., Rahmsdorf, H.-J., Smith, B.J., Lyons, A., Harris, T.J.R., Reynolds, J.J., Herrlich, P., and Docherty, A.J.P. (1986) *Biochem. J.* **240**, 913-916
42. Wilhelm, S.M., Collier, L.E., Kronberger, A., Eisen, A.Z. Marmer, B.L., Grant, G.A., Bauer, E.A., and Goldberg, G.I., (1987) *Proc. Natl. Acad. Sci. USA* **86**, 6725-6729
43. Matrisian, L.M., Leroy, P., Ruhlmann, C., Gesnel, M.-C., and Breathnach, R. (1986) *Mol. Cell. Biol.* **6**, 1679-1686
44. Muller, D., Quantin, B., Gesnel, M.-C., Millon-Collard, R., Abecassis, J., and Breathnach, R. (1988) *Biochem. J.* **253**, 187-192
45. Collier, I.E., Wilhelm, S.M., Eisen, A.Z. Marmer, B.L., Grant, G.A., Seltzer, J.L., Kronberger, A., He, C., Bauer, E.A., and Goldberg, G.I., (1988) *J. Biol. Chem.* **263**, 6579-6587

Table I
Amino acid composition of HR1B

Amino acid	Analysis ^a	Sequence
Residues/molecule		
Asp	53.5	28
Asn		26
Thr	18.8 ^b	19
Ser	26.5 ^b	28
Glu	39.1	22
Gln		17
Pro	16.8	20
Gly	27.4	26
Ala	22.0	21
1/2 Cys	34.3 ^c	35
Val	22.8 ^d	23
Met	6.2 ^d	7
Ile	20.5 ^d	23
Leu	25.0	25
Tyr	19.4	21
Phe	15.2	15
Lys	26.2	27
His	13.0	13
Trp	2.9 ^e	4
Arg	15.9	16
GlcNH ₂ ^f	+	
Total		416

^a Average values obtained from 24, 48, and 72 h hydrolyses with 5.7 N HCl.

^b Extrapolated values to zero time.

^c Determined as cysteic acid after performic acid oxidation.

^d Taken from 72 h hydrolysis.

^e Obtained from 24 h hydrolysis with 3N mercaptoethanesulfonic acid.

^f Glucosamine.

Table II

Amino acid compositions of cyanogen bromide peptides derived from Pe-HR1B^a

Amino acid	M1	M2	M3	M3-4-5	M4-5	M5	M5-DP	M6	M7	M7'	M7''
— Residues/molecule —											
Asp	13.1 (13)	3.2 (3)	2.9 (3)	21.2 (19)	16.9 (16)	16.0 (16)	17.5 (15)	14.1 (14)	5.0 (5)	5.1 (5)	4.9 (5)
Thr	7.5 (8)	1.1 (1)	0.9 (1)	7.2 (8)	6.4 (7)	6.3 (7)	6.4 (7)	1.0 (1)	1.0 (1)	1.2 (1)	0.9 (1)
Ser	6.6 (7)	2.0 (2)	1.3 (1)	13.2 (15)	11.8 (14)	11.0 (13)	10.4 (12)	1.9 (2)	1.1 (1)	1.2 (1)	1.6 (2)
Glu	9.4 (9)	2.2 (2)		19.6 (19)	19.5 (19)	18.5 (19)	18.8 (19)	7.3 (7)	2.2 (2)	2.0 (2)	2.0 (2)
Pro	2.2 (2)	2.1 (2)	2.3 (2)	14.1 (14)	10.8 (12)	10.7 (11)	10.5 (11)	2.2 (2)			
Gly	5.4 (5)	1.4 (1)	4.7 (5)	13.4 (14)	9.5 (9)	9.7 (9)	10.8 (9)	4.3 (4)	2.2 (2)	2.1 (2)	1.9 (2)
Ala	5.9 (6)	1.1 (1)		9.3 (9)	9.1 (9)	8.7 (9)	9.0 (9)	3.2 (3)	2.0 (2)	2.1 (2)	1.9 (2)
Val ^b	8.8 (9)	4.7 (5)	0.5 (0)	3.9 (3)	4.0 (3)	3.2 (3)	3.2 (3)	2.3 (2)	4.1 (4)	4.0 (4)	3.8 (4)
Met ^b	0.6 (1)	1.6 (2)	0.8 (1)	1.9 (3)	0.8 (2)	1.1 (1)	0.7 (1)	0.6 (1)			
Ile	11.2 (12)	2.7 (3)	1.8 (2)	6.5 (7)	4.5 (5)	4.4 (5)	4.2 (4)	1.0 (1)			
Leu	12.8 (13)	1.2 (1)	1.1 (1)	9.7 (8)	7.2 (7)	7.1 (7)	7.2 (7)	3.0 (3)			
Tyr	4.9 (5)	1.1 (1)		7.3 (6)	6.3 (6)	6.1 (6)	7.1 (6)	7.8 (8)		1.0 (1)	0.8 (1)
Phe	3.0 (3)		1.1 (1)	6.8 (5)	4.8 (4)	4.6 (4)	5.5 (4)	6.8 (7)			
Lys	8.9 (9)	1.1 (1)		8.1 (6)	6.6 (6)	6.2 (6)	6.9 (6)	8.0 (8)	2.0 (2)	3.0 (3)	2.6 (3)
His	3.6 ^d (4)	0.8 (1)	1.8 (2)	4.5 (6)	3.8 (4)	3.3 (4)	3.5 (4)	0.8 (1)	1.1 (1)	1.1 (1)	1.0 (1)
Trp ^c	n.d. ^d (3)			n.d. ^d (1)	n.d. ^d (1)	n.d. ^d (1)	n.d. ^d (1)				
Pec ^c	1.1 (1)	0.9 (1)	2.6 (3)	18.7 (23)	17.8 (20)	17.3 (20)	17.7 (20)	6.5 (7)	2.7 (3)	2.4 (3)	2.4 (3)
Arg	7.0 (7)	1.1 (1)		3.7 (5)	4.9 (5)	4.8 (5)	5.0 (5)	2.0 (2)	1.0 (1)	1.0 (1)	1.0 (1)
GlcNH ₂ ^e	+			+	+	+	+	+			
Total	117	28	22	171	149	146	143	73	24	26	27
Position	1-117	118-145	146-167	146-316	168-316	171-316	174-316	317-389	390-413	390-415	390-416
Yield (%)	44	24	48	6	6	23	8	54	27	16	5

^aValues in parentheses are from the sequence data. ^bDetermined as homoserine plus homoserine lactone.
^cS-pyridylethylcysteine. ^dNot determined. ^eGlucosamine.

Table III

Summary of amino acid sequences of fragments derived from HR1B

No.	Amino acid	Yield of PTH amino acid pmol	
1	Glu		
2	Gln		31.4HR1B
3	Arg	1.2M1K1	3.0 PG
4	Phe	86.7 PG	38.2
5	Pro	34.0	26.6
6	Arg	1.6	3.3
7	Arg	4.0	7.2
8	Tyr	14.5	11.0
9	Ile	27.5	32.2
10	Lys	16.8	21.6
12	Leu	94.4M1K2	31.9
12	Ala	91.3	19.5
12	Ile	74.4	34.6
14	Val	86.2	24.6
15	Val	98.0	40.5
16	Asp	38.0	25.8
17	His	18.8	13.1
18	Gly	56.5	71.8
19	Ile	51.4	71.1
20	Val	54.0	63.0
21	Thr	14.6	30.1
22	Lys	6.1	37.3
23	His	5.5M1K3	12.6
24	His	3.7	16.5
25	Gly	70.1	38.4
26	Asn	54.1	27.6
27	Leu	86.8	39.3
28	Lys	35.3	15.2
29	Lys		23.0
30	Ile		38.0
31	Arg		10.6
32	Lys		12.8
33	Trp	100.4M1K4	10.8
34	Ile	176.5	29.3
35	Tyr	81.8	16.9
36	Gln	49.6	16.8
37	Leu	64.5	24.6
38	Val	79.9	15.4
39	Asn	79.7	13.7
40	Thr	39.1	6.4
41	Ile	70.5	17.1
42	Asn	72.5	15.3
43	Asn	134.4	19.6
44	Ile	94.6	17.9
45	Tyr	61.2	8.5
46	Arg	25.4	9.9
47	Ser	14.4	2.6
48	Leu	39.8	10.2
49	Asn	33.7	9.1
50	Ile	27.8	10.5
51	Leu	35.8	10.7
52	Val	10.1	6.3
53	Ala	20.2	6.1
54	Leu	27.0	10.3
55	Val	16.2	6.7
56	Tyr	12.4	4.0
57	Leu	21.7	10.2
58	Glu	6.9	1.4
59	Ile	11.5	7.9
60	Trp	9.0	-

Table III

(continued-1)

No.	Amino acid	Yield of PTH amino acid pmol		
61	Ser	3.9(M1K4)	1.7(M1D2)	73.1M1W1
62	Lys		1.2	105.1
63	Gln		2.6	146.6
64	Asn		4.3	128.3
65	Lys		1.4	135.2
66	Ile	97.7M1K5	1.0	178.4
67	Thr	26.7	1.1	86.6
68	Val	67.9	3.6	96.6
69	Gln	47.4		112.3
70	Ser	7.0		35.9
71	Ala	70.9		88.9
72	Ser	10.5		33.7
73	Asn	-		-
74	Val	42.7		39.4
75	Thr	12.0		42.1
76	Leu	85.2		71.8
77	Asp	19.1		31.0
78	Leu	78.3		63.9
79	Phe	27.0		55.1
80	Gly	35.1		47.9
81	Asp	-	542.0M1D4	13.2
82	Trp	9.8	242.0	
83	Arg	-	27.2	
84	Glu	6.9	278.7	
85	Ser	-	212.3	
86	Val	6.6	509.2	
87	Leu	19.9	516.1	
88	Leu	24.5	755.6	
89	Lys		388.4	
90	Gln	24.3M1K6	413.9	
91	Arg	9.4	239.9	
92	Ser	7.6	142.3	71.9ST2
93	His	4.3	-	9.2
94	Asp	12.6	94.9M1D5	49.9
95	Pec	n.d.	n.d.	n.d.
96	Ala	13.2	230.6	156.7
97	Gln	18.5	168.9	100.7
98	Leu	20.6	162.4	105.7
99	Leu	32.6	179.9	167.0
100	Thr	2.6	26.2	73.0
101	Thr	3.5	33.9	59.8
102	Ile	11.1	30.1	92.5
103	Asp	4.4		26.1
104	Phe	5.6		80.0
105	Asp	4.9	57.9M1D6	47.9
106	Gly	9.2	65.6	70.6
107	Pro	3.0	52.9	52.9
108	Thr	1.5	9.6	30.3
109	Ile	7.7	29.4	51.7
110	Gly	3.0	38.7	48.2
111	Lys	0.5	28.9	-
112	Ala	504.9M1K7	48.6	54.0
113	Tyr	205.0	25.8	28.6
115	Thr	113.9	1.4	28.9
115	Ala	436.4	37.8	40.2
116	Ser	48.6	4.8	19.3
117	Met	n.d.a)	-	24.0
118	Pec		468.5M2	n.d.
119	Asp		265.2	18.8
120	Pro		468.7	46.2

(continued-2)

Table III

No.	Amino acid	Yield of PTH amino acid pmol			
121	Lys	611.9(M2)			
122	Arg	102.8			
123	Ser	88.9			35.9R1
124	Val	544.1			144.3
125	Gly	411.9			141.7
126	Ile	386.9			147.5
127	Val	526.5			133.4
128	Gln	297.7			110.8
129	Asp	173.4			50.1
130	Tyr	304.4			72.2
131	Ser	35.3			19.5
132	Pro	124.2			61.8
133	Ile	100.2			65.1
134	Asn	70.0			44.8
135	Leu	181.8			99.1
136	Val	74.1			47.4
137	Val	113.2			70.6
138	Ala	50.1			51.0
139	Val	78.2			55.3
140	Ile	41.6			49.6
141	Met	n.d. ^{a)}			35.1
142	Thr	23.5			9.6
143	His	9.0			7.3
144	Glu	7.9			19.7
145	Met	-			15.0
116	Gly	317.6M3	183.1M3-4-5		31.3
147	His	45.3	11.0		5.2
148	Asn	365.9	127.6		19.3
149	Leu	972.5	213.4		32.3
150	Gly	603.3	182.3		18.0
151	Ile	619.9	215.2		20.9
152	Pro	179.0	98.8		22.5
153	His	45.4	12.9		3.9
154	Asp	98.1	62.1		
155	Gly	226.9	81.6		
156	Asn	73.6	30.6		
157	Ser	29.5	61.3		
158	Pec	n.d.	n.d.		
159	Thr	42.1	45.1		
160	Pec	n.d.	n.d.		
161	Gly	99.8	75.5		
162	Gly	110.2	94.4		
163	Phe	58.9	90.6		
164	Pro	39.5	64.3		
165	Pec	n.d.	n.d.		
166	Ile	28.2	51.7		
167	Met	n.d. ^{a)}	-		
168	Ser	12.4M4-5	28.3		
169	Pro	95.0	84.5		
170	Met	25.5			
171	Ile	40.4		208.8M5	
172	Ser	10.4		53.1	
173	Asp	17.2		119.1	
174	Pro	49.8	62.4M5-DP	141.9	
175	Pro	47.7	49.5	155.6	
176	Ser	3.5	8.1	22.7	
177	Glu	58.6M5D1	13.3	40.2	54.3
178	Leu	328.7	25.9	52.0	148.9
179	Phe	235.0	21.0	48.7	68.3
180	Ser	21.2	9.6	5.6	12.5

Table III

(continued-3)

No. Amino acid		Yield of PTH amino acid pmol			
		- (M5D1)	- (M4-5)	- (M5-DP)	- (M5)
181	Asn				
182	Pec	n.d.	n.d.	n.d.	n.d.
183	Ser	20.2	4.3	11.8	7.5
184	Lys	139.5	11.6	46.6	30.6
185	Ala 468.8M5K1	68.7	23.8	38.8	42.0
186	Tyr 312.5	68.5	13.6	26.1	41.6
187	Tyr 368.4	46.1	14.1	41.2	38.5
188	Gln 300.8		15.1	22.6	39.7
189	Thr 229.4		7.9	12.2	7.3
190	Phe 262.2		15.3	21.3	42.5
191	Leu 237.8		26.8	34.5	67.3
192	Thr 157.2			4.5	6.2
193	Asp 138.6	572.6K5D3		25.0	12.9
194	His 118.8	204.0		5.0	5.2
195	Lys 79.0	689.0		18.6	14.6
196	Pro	384.0		9.0	30.1
197	Gln	264.5		12.1	12.4
198	Pec	n.d.		n.d.	n.d.
199	Ile	405.1		7.2	16.8
200	Leu	672.3		8.6	29.1
201	Asn	202.9		11.9	8.9
202	Ala	549.6		11.1	21.7
203	Pro	256.2		12.2	28.0
204	Ser	40.5		2.7	2.3
205	Lys	209.1		3.3	5.8
206	Thr 13.2M5K3	17.0		2.2	3.3
207	Asp 23.7			6.1	12.3
208	Ile 125.6			5.7	10.0
209	Val 80.5			3.1	9.4
210	Ser 16.8			2.0	2.6
211	Pro 80.5			7.4	22.8
212	Pro 85.1			7.2	22.5
213	Val 47.9			3.3	10.1
214	Pec 32.6			n.d.	n.d.
215	Gly 47.3			8.3	13.6
216	Asn 33.8			4.6	
217	Glu 25.3			6.8	
218	Leu 96.4			2.5	
219	Leu 164.0			3.6	
220	Glu 27.8				
221	Ala 34.2				
222	Gly 33.7				
223	Glu 23.8	91.5M5D5			
224	Glu 32.3	146.1			
225	Pec 14.1	n.d.			
226	Asp 12.4	150.0M5D6			
227	Pec 21.0	n.d.			
228	Gly 21.9	185.0			
229	Ser 3.0	38.4			
230	Pro 12.1	135.9			
231	Glu 7.3	79.6			
232	Asn 6.6	74.3			
233	Pec 6.5	n.d.			
234	Gln 7.0	82.9			
235	Tyr 4.3	73.7			
236	Gln 14.8	99.2			
237	Pec 9.8	n.d.			
238	Pec 10.0	n.d.			
239	Asp 3.9	82.3M5D7			
240	Ala 6.3	340.4			

(continued-4)

Table III

No.	Amino acid	Yield of PTH amino acid pmol	
241	Ala	12.7(M5K3)	328.8(M5D7)
242	Ser	0.9	14.6
243	Pec	4.5	n.d.
244	Lys	1.0	174.1
245	Leu	706.2M5K4	312.2
246	His	66.4	30.9
247	Ser	61.9	10.3
248	Trp	257.5	52.4
249	Val	263.3	80.5
250	Lys	152.6	54.8
251	Pec	59.6M5K5	n.d.
252	Glu	48.6	16.8
253	Ser	13.5	6.7
254	Gly	68.4	
255	Glu	39.3	
256	Pec	101.3	
257	Pec	100.3	
258	Asp	40.6	55.1M5D8
259	Gln	36.9	170.6
260	Pec	74.0	176.2
261	Arg	26.7	30.8
262	Phe	56.6	326.2
263	Arg	31.9	79.0
264	Thr	9.1	187.3
265	Ala	36.7	334.6
266	Gly	29.0	215.2
267	Thr	6.9	140.1
268	Glu	18.5	146.1
269	Pec	34.3	220.6
270	Arg	11.9	72.3
271	Ala	14.1	125.2
272	Ala	26.6	80.5
273	Glu	13.0	
274	Ser	3.0	
275	Glu	9.7	
276	Pec	15.9	
277	Asp	9.4	162.5M5D10
278	Ile	8.2	253.8
279	Pro	8.5	200.0
280	Glu	4.3	106.6
281	Ser	1.0	35.3
282	Pec	7.4	n.d.
283	Thr	-	51.9
284	Gly	6.4	114.8
285	Gln	3.3	89.8
286	Ser	0.8	24.3
287	Ala	3.8	27.1
288	Asp	-	51.0
289	Pec	3.3	n.d.
290	Pro	3.4	50.2
291	Thr		31.5
292	Asp		32.6
293	Arg		16.6
294	Phe		37.7
295	His		4.0
296	Arg		16.5
297	Asn		27.6ST5
298	Gly		71.8
299	Gln		58.6
300	Pro		58.1

Table III

(continued-5)

No.	Amino acid	Yield of PTH amino acid pmol		
301	Pec	n.d.(M5D12)	59.8(ST5)	- (ST4-5)
302	Leu	62.2	61.9	12.7
303	Tyr	25.7	34.5	15.7
304	Asn	20.8	37.3	
305	His	6.1	2.8	
306	Gly	34.2	39.8	
307	Tyr	16.7	24.9	
308	Pec	n.d.	27.1	
309	Tyr	16.3	33.6	
310	Asn	7.6	12.3	
311	Gly	11.9	33.1	
312	Lys	6.8	-	
313	Pec 321.0M5K6		22.2	
314	Pro 244.9		9.5	
315	Ile 210.8		7.9	
316	Met n.d.a)		2.2	
317	Phe 798.5M6K1	276.0 M6	13.6	
318	Tyr 604.9	91.0	8.2	
319	Gln 381.8	186.1	6.9	
320	Pec 290.2	n.d.	18.2	
321	Tyr 579.0	143.1	-	
322	Phe 571.7	201.4	14.9	
323	Leu 540.7	170.1	11.9	
324	Phe 540.0	225.1	-	
325	Gly 420.9	134.3	14.9	
326	Ser 139.9	30.8		
327	Asn -	-		
328	Ala 297.2	215.3		
329	Thr 225.1	40.3		
330	Val 265.9	124.3		
331	Ala 292.1	101.4		
332	Glu 120.9	68.7		
333	Asp 63.7	45.5M6D2 7.1		
334	Asp 127.9	66.7 7.3		
335	Pec 95.4	72.2 n.d.		
336	Phe 93.9	132.8 29.7		
337	Asn 110.7	86.3 77.7		
338	Asn 122.1	70.0 86.4		
339	Asn 146.5	75.6 31.0		
340	Lys 16.9	89.2 69.8		
341	Lys	85.0 60.6		
342	Gly	16.5 25.6		
343	Asp	8.1		
344	Lys	57.0		
345	Tyr 487.4M6K2	31.1		
346	Phe 622.5	49.7		
347	Tyr 467.4	24.1		
348	Pec n.d.	n.d.		
349	Arg 191.9	47.9		
350	Lys 328.2	21.8		
351	Glu	8.3		
352	Asn	12.2		
353	Glu	8.4		
354	Lys	29.5		
355	Tyr 121.8M6K3	15.1		
356	Ile 149.8	17.5		
357	Pro 136.8	24.6		
358	Pec n.d.	n.d.		
359	Ala 118.3	14.5		
360	Gln 90.6	14.6		

(continued-6)

Table III

No.	Amino acid	Yield of PTH amino acid pmol			
361	Glu 52.7(M6K3)	9.1(M6)			
362	Asp 95.0	4.8			
363	Val 68.3	7.1			
364	Lys 127.9	6.8			
365	Pec n.d.M6K4	n.d.			
366	Gly 440.7	8.8			
367	Arg 157.7	-			
368	Leu 451.1	16.5		453.3ST6	
369	Phe 459.6	3.1		448.8	
370	Pec n.d.	n.d.		n.d.	
371	Asp 172.3 147.4M6D6			87.3	
372	Asn 217.5 244.1			182.0	
373	Lys 158.6 350.8			-	
374	Lys 312.9M6K5 375.6			-	
375	Tyr 291.9 259.6			228.4	
376	Pro 352.7 287.2			191.1	
377	Pec n.d. n.d.			n.d.	
378	His 50.4 60.4			17.1	
379	Tyr 259.0 97.6			122.2	
380	Asn -	-		-	
381	Tyr 240.7 186.2			176.8	
382	Ser 45.7 31.9			53.4	
383	Glu 129.8 30.7			56.3	
384	Asp 105.7			52.9	
385	Leu 316.8			84.4	
386	Asp 132.4			78.1	
387	Phe 148.3			71.2	
388	Gly 109.9			58.0	
389	Met n.d. ^{a)}			40.5	
390	Val	574.9M7	379.4M7'	466.1M7"	57.1
391	Asp	63.6	194.8	139.9	41.9
392	His	24.1	45.4	29.2	5.6
393	Gly	255.3	179.8	156.3	30.7
394	Thr	44.0	27.4	123.6	41.0
395	Lys	132.3	149.9	214.8	-
396	Pec	n.d.	168.5	133.9	n.d.
397	Ala	132.3	141.5	244.6	14.0
398	Asp	51.9	87.7	45.7	10.6
399	Gly	104.0	105.8	104.9	24.4
400	Lys	130.7	95.4	107.9	-
401	Val	156.4	92.5	139.0	11.0
402	Pec	n.d.	127.5	113.0	n.d.
403	Ser	18.2	19.2	41.6	5.4
404	Asn	57.8	52.2	91.0	5.9
405	Arg	33.7	43.2	24.6	1.7
406	Gln	64.2	52.4	92.4	
407	Pec	n.d.	71.8	86.5	
408	Val	106.8	43.2	51.6	
409	Asp	35.3	39.2	22.9	
410	Val	93.5	55.3	63.6	
411	Asn	36.1	38.7	39.3	
412	Glu	22.8	28.2	39.1	
413	Ala	12.5	37.9	54.8	
414	Tyr		26.7	36.3	
415	Lys		7.7	27.7	
416	Ser			9.9	

n.d.: not quantitatively determined. -: not identified. ^aIdentified as homoserine. The peptide nomenclature is described under "Materials and Methods" and "Results".

Table IV

Amino acid compositions of lysyl endopeptidase peptides derived from M1^a

Amino acid	M1K1	M1K1 PG	M1K2	M1K3	M1K4	M1K5	M1K6	M1K6'	M1K7
Residues/molecule									
Asp			1.0 (1)	1.3 (1)	4.6 (4)	2.4 (3)	3.1 (3)	2.7 (3)	
Glu	1.9 (2)				2.2 (2)	1.9 (2)	2.0 (2)	1.9 (2)	
Ser					2.5 (2)	3.1 (3)	1.0 (1)	1.1 (1)	1.0 (1)
Gly			1.3 (1)	1.2 (1)		1.6 (1)	2.1 (2)	2.0 (2)	
His			0.9 (1)	2.2 (2)			1.0 (1)	1.0 (1)	
Arg	3.1 (3)	3.2 (3)			1.4 (1)	1.0 (1)	1.1 (1)	1.0 (1)	
Thr			0.9 (1)		1.2 (1)	1.7 (2)	2.7 (3)	2.3 (3)	1.0 (1)
Ala			1.0 (1)		0.7 (1)	1.0 (1)	1.1 (1)	1.0 (1)	2.2 (2)
Pro	1.0 (1)	1.0 (1)				0.5 (0)	1.2 (1)	1.1 (1)	
Tyr	0.8 (1)	1.0 (1)			2.6 (3)				0.9 (1)
Val _b			1.8 (3)		1.9 (3)	1.6 (3)			
Met _c									0.8 (1)
Pec _c							0.8 (1)	0.9 (1)	
Ile	1.0 (1)	1.0 (1)	1.2 (2)		4.2 (5)	0.8 (1)	1.8 (2)	1.8 (2)	
Leu			1.0 (1)	0.8 (1)	3.9 (5)	2.4 (4)	2.1 (2)	1.9 (2)	
Phe	1.0 (1)	1.0 (1)				0.7 (1)	1.0 (1)	1.0 (1)	
Trp					n.d. ^d (2)	n.d. ^d (1)			
Lys	0.8 (1)	1.0 (1)	0.9 (1)	0.9 (1)	1.2 (1)	0.7 (1)	0.8 (1)	0.9 (1)	
GlcNH ₂ ^e						+			
Total	10	8	12	6	30	24	22	22	6
Position	1-10	3-10	11-22	23-28	33-62	66-89	90-111	90-111	112-117
Yield (%)	13	n.d.	11	11	7	7	8	22	25

^aValues in parentheses are from the sequence data. ^bDetermined as homoserine plus homoserine lactone. ^cS-pyridylethylcysteine. ^dNot determined. ^eGlucosamine.

Table V

Amino acid compositions of endoproteinase Asp-N peptides derived from M1^a

Amino acid	M1D1	M1D2	M1D3	M1D4	M1D5	M1D6
— residues/molecule —						
Asp		5.9 (8)	1.0 (1)	1.4 (1)	0.9 (1)	1.0 (1)
Glu	1.7 (2)	4.4 (4)		2.4 (2)	1.0 (1)	
Ser		4.5 (4)		1.8 (2)		0.9 (1)
Gly		3.3 (2)	1.0 (1)			2.0 (2)
His		2.1 (3)		0.9 (1)		
Arg	1.9 (3)	2.8 (2)		2.0 (2)		
Thr	0.5 (0)	3.4 (4)			1.5 (2)	1.7 (2)
Ala	1.2 (1)	2.9 (2)			1.2 (1)	2.4 (2)
Pro	0.8 (1)	1.1 (0)				0.9 (1)
Tyr	0.9 (1)	2.7 (3)				1.0 (1)
Val ^b	1.6 (2)	6.4 (6)		0.8 (1)		
Met ^c						0.9 (1)
Pec ^c					0.9 (1)	
Ile	1.8 (2)	8.7 (8)			1.2 (1)	1.4 (1)
Leu	1.4 (1)	7.9 (7)	1.0 (1)	1.7 (2)	2.3 (2)	
Phe	0.9 (1)	0.9 ^d (0)	1.0 (1)			
Trp		n.d. ^d (2)		n.d. ^d (1)		
Lys ^e	0.8 (1)	5.2 (6)		1.3 (1)		1.0 (1)
GlcNH ₂		+				
Total	15	61	4	13	9	13
Position	1-15	16-76	77-80	81-93	94-102	105-117
Yield(%)	8	18	20	9	12	21

^aValues in parentheses are from the sequence data. ^bDetermined as homoserine plus homoserine lactone. ^cS-pyridylethylcysteine. ^dNot determined. ^eGlucosamine.

Table VI

Amino acid compositions of lysyl endopeptidase peptides derived from M5^a

Amino acid	M5K1	M5K2	M5K3	M5K4	M5K5	M5K6
—— residues/molecule ——						
Asp	1.1 (1)	1.0 (1)	5.1 (5)		7.0 (7)	
Glu	1.3 (1)	1.2 (1)	6.8 (7)		8.5 (9)	
Ser		1.2 (1)	2.8 (3)	1.0 (1)	3.7 (4)	
Gly			3.2 (3)		5.7 (6)	
His	0.9 (1)			1.0 (1)	1.8 (2)	
Arg					4.7 (5)	
Thr	1.7 (2)		0.9 (1)		3.7 (4)	
Ala	0.9 (1)	0.8 (1)	2.8 (3)		3.9 (4)	
Pro		1.6 (2)	3.2 (3)		3.3 (3)	0.9 (1)
Tyr	1.6 (2)		1.0 (1)		2.9 (3)	
Val ^b			1.6 (2)	1.0 (1)		
Met ^c						1.0 (1)
Pec ^c		1.0 (1)	6.1 (7)		6.4 (10)	1.0 (1)
Ile		0.7 (1)	0.4 (1)		1.3 (1)	1.1 (1)
Leu	1.0 (1)	0.9 (1)	2.0 (2)	1.1 (1)	1.4 (1)	
Phe	0.9 (1)				2.0 (2)	
Trp				n.d. ^d (1)		
Lys	0.9 (1)	1.0 (1)	1.1 (1)	1.0 (1)	1.5 (1)	
Total	11	10	39	6	62	4
Position	185-195	196-205	206-244	245-250	251-312	313-316
Yield(%)	45	72	34	37	7	33

^aValues in parentheses are from the sequence data. ^bDetermined as homoserine plus homoserine lactone. ^cS-pyridylethylcysteine. ^dNot determined.

Table VII

Amino acid compositions of endoproteinase Asp-N peptides derived from M5^a

Amino acid	M5D1	M5D2	M5D3	M5D4	M5D5	M5D6	M5D7	M5D8
— residues/molecule —								
Asp	1.3 (1)		2.0 (2)	1.9 (2)		1.9 (2)	1.3 (1)	1.0 (1)
Glu	1.1 (1)	1.2 (1)	1.2 (1)	1.2 (1)	2.0 (2)	2.5 (3)	0.9 (1)	2.0 (2)
Ser	1.8 (2)		1.0 (1)	1.1 (1)		1.4 (1)	2.2 (3)	
Gly				1.1 (1)		1.0 (1)	0.7 (1)	0.9 (1)
His			1.0 (1)				0.8 (1)	
Arg								2.6 (3)
Thr		1.8 (2)	0.9 (1)					1.6 (2)
Ala	0.8 (1)		1.0 (1)			0.6 (0)	2.0 (2)	2.2 (3)
Pro			1.8 (2)	1.7 (2)		1.3 (1)		
Tyr	1.4 (2)					1.1 (1)		
Val ^b				1.7 (2)			0.7 (1)	
Met ^c								
Pec ^c	0.9 (1)		0.7 (1)	0.1 (1)	1.0 (1)	2.8 (4)	1.5 (2)	1.9 (2)
Ile			0.9 (1)	0.6 (1)		0.5 (0)		
Leu	1.0 (1)	1.1 (1)	1.0 (1)	2.1 (2)			1.2 (1)	
Phe	0.7 (1)	1.1 (1)						0.8 (1)
Trp							n.d. ^d (1)	
Lys ^e	1.1 (1)		1.7 (2)				1.7 (2)	
GlcNH ₂ ^e +								
Total	11	5	14	13	3	13	16	15
Position	177-187	188-192	193-206	207-219	223-225	226-238	239-254	258-272
Yield(%)	6	19	19	27	43	26	11	30

^aValues in parentheses are from the sequence data. ^bDetermined as homoserine plus homoserine lactone. ^cS-pyridylethylcysteine. ^dNot determined. ^eGlucosamine.

Table VII

(continued)

Amino acid	M5D9	M5D10	M5D11	M5D12
——— residues/molecule ———				
Asp		1.1 (1)	1.1 (1)	3.6 (4)
Glu	0.9 (1)	1.7 (2)		1.1 (1)
Ser		1.7 (2)		
Gly		1.1 (1)		2.5 (3)
His				2.0 (2)
Arg				2.0 (2)
Thr		0.8 (1)	1.0 (1)	
Ala		0.9 (1)		
Pro		0.9 (1)	1.0 (1)	2.0 (2)
Tyr				2.2 (3)
Val ^b				
Met ^c				0.8 (1)
Pec ^c	1.1 (1)	0.6 (1)	0.7 (1)	1.7 (3)
Ile		0.9 (1)		0.9 (1)
Leu				1.1 (1)
Phe				0.8 (1)
Trp				
Lys ^e				1.1 (1)
GlcNH ₂				
Total	2	11	4	25
Position	275-276	277-287	288-291	292-316
Yield(%)	75	12	26	6

Table VIII

Amino acid compositions of lysyl endopeptidase peptides derived from M6^a

Amino acid	M6K1	M6K2	M6K3	M6K4	M6K5
— residues/molecule —					
Asp	5.6 (6)		1.0 (1)	2.2 (2)	2.6 (3)
Glu	2.2 (2)		2.0 (2)		1.5 (1)
Ser	1.0 (1)				1.0 (1)
Gly	1.4 (1)			1.0 (1)	1.1 (1)
His					1.0 (1)
Arg		1.0 (1)		1.0 (1)	
Thr	1.0 (1)				
Ala	2.1 (2)		0.9 (1)		
Pro			1.0 (1)		1.1 (1)
Tyr	2.1 (2)	2.0 (2)	0.9 (1)		2.6 (3)
Val ^b	1.1 (1)		1.0 (1)		
Met ^b					0.8 (1)
Pec ^c	1.6 (2)	0.9 (1)	1.4 (1)	1.6 (2)	1.0 (1)
Ile			0.8 (1)		
Leu	1.1 (1)			1.0 (1)	1.0 (1)
Phe	3.8 (4)	1.0 (1)		1.0 (1)	1.1 (1)
Trp					
Lys ^d	1.0 (1)	1.0 (1)	0.9 (1)	0.9 (1)	1.0 (1)
GlcNH ₂	+				+
Total	24	6	10	9	16
Position	317-340	345-350	355-364	365-373	374-389
Yield (%)	15	15	20	14	29

^aValues in parentheses are from the sequence data.

^bDetermined as homoserine plus homoserine lactone. ^cS-pyridylethylcysteine. ^dGlucosamine.

Table IX

Amino acid compositions of endoproteinase Asp-N peptides derived from M6^a

Amino acid	M6D1	M6D2	M6D3	M6D4	M6D5	M6D6	M6D7
— residues/molecule —							
Asp	1.1 (1)	3.2 (5)	1.3 (1)	0.9 (1)	1.2 (1)	3.2 (3)	1.3 (1)
Glu	2.1 (2)			4.0 (4)		1.0 (1)	
Ser	0.9 (1)					0.6 (1)	
Gly	1.3 (1)	1.4 (1)			1.1 (1)		1.0 (1)
His						1.0 (1)	
Arg			1.0 (1)		1.1 (1)		
Thr	0.9 (1)						
Ala	2.2 (2)			1.1 (1)			
Pro				1.1 (1)		1.0 (1)	
Tyr	1.9 (2)		1.8 (2)	1.2 (1)		2.5 (3)	
Val ^b	1.0 (1)				0.9 (1)		
Met ^c							0.7 (1)
Pec ^c	0.7 (1)	1.0 (1)	0.8 (1)	0.8 (1)	1.7 (2)	0.9 (1)	
Ile				1.2 (1)			
Leu	1.2 (1)				1.0 (1)		
Phe	3.0 (3)	1.2 (1)	1.0 (1)		1.0 (1)		1.0 (1)
Trp							
Lys ^d		2.1 (2)	1.7 (2)	1.0 (1)	0.9 (1)	1.8 (2)	
GlcNH ₂	+						
Total	16	10	8	11	9	13	4
Position	317-332	333-342	343-350	351-361	362-370	371-383	384-387
Yield(%)	16	9	13	11	17	28	11

^aValues in parentheses are from the sequence data. ^bDetermined as homoserine plus homoserine lactone. ^cS-pyridylethylcysteine. ^dGlucosamine.

Table X

Amino acid compositions of tryptic peptides derived from succinylated Pe-HR1B^a

Amino acid	ST1	ST2	ST3	ST4	ST5	ST4-5	ST6
— residues/molecule —							
Asp	9.2 (10)	18.7 (19)		2.8 (3)	11.9 (12)	14.9 (15)	6.2 (8)
Glu	8.3 (8)	18.5 (16)	0.8 (1)	4.4 (4)	6.5 (7)	9.2 (11)	2.1 (1)
Ser	4.1 (5)	12.6 (16)		2.3 (3)	2.3 (1)	3.4 (4)	1.8 (2)
Gly	3.0 (3)	10.6 (12)	1.0 (1)	2.3 (1)	5.9 (6)	7.3 (7)	3.2 (3)
His	3.0 (3)	5.9 (6)		1.0 (1)	2.3 (1)	2.7 (2)	2.1 (2)
Arg	7.1 (7)	2.4 (2)	1.4 (2)	2.0 (2)	2.7 (2)	4.0 (4)	1.4 (1)
Thr	4.2 (4)	7.6 (9)	1.8 (2)	1.9 (2)	2.1 (1)	3.2 (3)	1.1 (1)
Ala	2.7 (3)	7.8 (9)	1.3 (1)	2.8 (3)	3.6 (3)	5.2 (6)	1.1 (1)
Pro	1.2 (1)	13.7 (13)		2.0 (2)	3.1 (3)	4.5 (5)	1.1 (1)
Tyr	3.5 (4)	5.7 (5)	0.5 (0)	0.7 (0)	6.2 (8)	7.7 (8)	3.2 (3)
Val	6.5 (9)	6.4 (8)			2.0 (2)	2.5 (2)	2.0 (2)
Met ^b		5.2 (5)			1.4 (1)	1.8 (1)	0.7 (1)
Pec ^b	0.7 (0)	18.1 (18)	0.5 (1)	2.5 (3)	8.2 (8)	11.3 (11)	4.0 (4)
Ile	7.6 (10)	8.3 (10)		0.9 (1)	1.7 (2)	2.3 (3)	
Leu	9.4 (11)	10.0 (10)			2.6 (2)	3.2 (2)	2.1 (2)
Phe	2.0 (2)	4.1 (4)	0.9 (1)	0.9 (1)	4.5 (5)	5.7 (6)	1.9 (2)
Trp	n.d. ^c (3)	n.d. (1)					
Lys ^d	6.4 (8)	5.9 (7)			5.7 (7)	6.5 (7)	3.0 (4)
GlcNH ₂	+	+			+	+	+
Total	91	170	9	26	71	97	38
Position	1-91	92-261	262-270	271-296	297-367	271-367	368-405
Yield(%)	23	16	6	8	11	10	4

^aValues in parentheses are from the sequence data. ^bS-pyridylethylcysteine.
^cNot determined. ^dGlucosamine.

Table XI

Amino acid composition of
arginylendopeptidase peptide R1
derived from Pe-HR1B^a

Amino acid	R1
— residues/molecule —	
ASp	14.6 (15)
Thr	5.3 (5)
Ser	11.7 (14)
Glu	14.6 (15)
Pro	10.7 (11)
Gly	9.9 (10)
Ala	6.2 (6)
Val	7.2 (8)
Met	4.1 (4)
Ile	7.0 (8)
Leu	8.0 (8)
Tyr	3.9 (4)
Phe	3.9 (4)
Lys	5.0 (5)
His	5.2 (5)
Trp	n.d. ^b (1)
Pec ^c	14.6 (16)
Arg	2.3 (2)
GlcNH ₂ ^d	+
Total	141
Position	123-263
Yield(%)	6

^aValues in parentheses are from the sequence data. ^bNot determined.

^cS-pyridylethylcysteine.

^dGlucosamine.

Table XII

Reactivity of the free sulfhydryl group in HR1B

Treatment	mol/molecule
<u>DTNB</u>	
None	0.1
6M Guanidine	0.9
<u>4-vinylpyridine</u>	
6M Guanidine, 10 mM EDTA ¹	0.7

LEGEND FOR FIGURE

Fig. 1. **Complete amino acid sequence of HR1B.** Residues determined by Edman degradation are written in upper case letters below the summarized sequence. Residues written in lower case letters indicate tentative identifications deduced from the amino acid compositions of the peptides. Dashes indicate unidentified residues. The peptide nomenclature is described under "Materials and Methods" and "Results". ● indicate sugar chains linked to Asn residue. <E denotes 5-pyrrolidone-2-carboxylic acid.

Fig. 2. **HPLC of CNBr peptides derived from Pe-HR1B.** A part of the digest was subjected to HPLC as described in Materials and Methods. A Vydac 214TP5415 (0.46 x 15 cm) column was used. Peptides M4-5 and M5, and M7' and M7" were separated by rechromatography (see Figs. 3 and 4). Peptides M5-DP and M3-4-5 which had some contaminating peptides were further purified using a Cosmosil 3C18 (0.46 x 10 cm) column (chromatograms not shown).

Fig. 3. **Separation of M4-5 and M5 by HPLC.** Peptides M4-5 and M5 found in one peak (Fig. 2) were further separated by HPLC. A TSK ODS-120T (0.46 x 25 cm) column was used.

Fig. 4. **Separation of M7' and M7" by HPLC.** Peptides M7' and M7" found in one peak (Fig. 2) were separated by HPLC. A TSK ODS-120T (0.46 x 25 cm) column was used.

Fig. 5. **HPLC of subdigest of M1 with lysyl endopeptidase.** A part of the digest was subjected to HPLC as described in

and Methods. A TSK ODS-120T (0.46 x 15 cm) column was used.

Fig. 6. HPLC of subdigest of M1 with endoproteinase Asp-

N. A part of the digest was subjected to HPLC as described in Materials and Methods. A Cosmosil 3C18 (0.46 x 10 cm) column was used.

Fig. 7. HPLC of subdigest of M5 with lysyl endopeptidase.

A part of the digest was subjected to HPLC as described in Materials and Methods. A TSK ODS-120T (0.46 x 15 cm) column was used. Peptides M5K4 and M5K5 were separated by rechromatography (see Fig. 8).

Fig. 8. Separation of M5K4 and M5K5 by HPLC. Peptides M5K4 and M5K5 found in one peak (Fig. 7) were further separated by HPLC. A TSK ODS-120T (0.46 x 25 cm) column was used.

Fig. 9. HPLC of subdigest of M5 with endoproteinase Asp-

N. A part of the digest was subjected to HPLC as described in Materials and Methods. A Cosmosil 3C18 (0.46 x 10 cm) column was used. Peptides M5D6 and M5D10, and M5D1, M5D7 and M5D12 were separated by rechromatography (see Figs. 10 and 11).

Fig. 10. Separation of M5D6 and M5D10 by HPLC. Peptides

M5D6 and M5D10 found in one peak (Fig. 9) were further separated by HPLC. A TSK ODS-120T (0.46 x 25 cm) column was used.

Fig. 11. **Separation of M5D1, M5D7 and M5D12 by HPLC.** Peptides M5D1, M5D7 and M5D12 found in one peak (Fig. 9) were further separated by HPLC. A Cosmosil 3C18 (0.46 x 10 cm) column was used.

Fig. 12. **HPLC of subdigest of M6 with lysyl endopeptidase.** A part of the digest was subjected to HPLC as described in Materials and Methods. A TSK ODS-120T (0.46 x 15 cm) column was used. Peptides M6K2 and M6K4 were separated by rechromatography (see Fig. 13).

Fig. 13. **Separation of M6K2 and M6K4 by HPLC.** Peptides M6K2 and M6K4 found in one peak (Fig. 12) were further separated by HPLC. A TSK ODS-120T (0.46 x 25 cm) column was used.

Fig. 14. **HPLC of subdigest with of M6 endoproteinase Asp-N.** A part of the digest was subjected to HPLC as described in Materials and Methods. A Cosmosil 3C18 (0.46 x 10 cm) column was used. Peptides M6D2 and M6D7 were separated by rechromatography (see Fig. 15).

Fig. 15. **Separation of M6D2 and M6D7 by HPLC.** Peptides M6D2 and M6D7 found in one peak (Fig. 14) were further separated by HPLC. A TSK ODS-120T (0.46 x 25 cm) column was used.

Fig. 16. **HPLC of tryptic peptides derived from succinylated Pe-HR1B.** A part of the digest was subjected to HPLC as described in Materials and Methods. A Cosmosil 5C4 (0.46 x 15 cm) column was used. ST4-5 and ST5 were contaminated and

using a Cosmosil 3C18 (0.46 x 10 cm) column (chromatograms not shown).

Fig. 17. **HPLC of digest of Pe-HR1B with arginylendopeptidase.** A part of the digest was subjected to HPLC as described in Materials and Methods. A TSK ODS-120T (0.46 x 15 cm) column was used.

Fig. 18. **Structural relationships of HR1B with other proteins.** A, Sequence comparison of the amino-terminal region of HR1B with HR2a, Ht-d, and H₂-proteinase. HR2a (7), Ht-d (6), and H₂-proteinase (27) are low molecular mass metalloproteinases. HR1B, HR2a, and H₂-proteinase were isolated from the venom of *T. flavoviridis*, and Ht-d from *C. atrox*. Identical residues shared by all proteins are shaded. The residues shared by hemorrhagic metalloproteinases, HR1B, HR2a, and Ht-d and not shared by non-hemorrhagic H₂-proteinase are boxed. The putative zinc ligands and active site are indicated by ▲ and Δ, respectively.

● indicate sugar chains linked to Asn residues of HR1B. <E denotes 5-pyrrolidone-2-carboxylic acid. B, Comparison and alignment of the carboxyl-terminal portion of HR1B with several disintegrins. Trigramin (17), Echistatin (29), Applaggin (31), and Bitistatin (32) are all viper venom platelet aggregation inhibitors, called disintegrins (35). Cysteine residues are boxed. The Arg-Gly-Asp-sequence common to disintegrins are shaded.

Fig. 19. **Comparison and alignment of the sequences of HR1B and human von Willebrand factor (vWF).** Identical residues are boxed.

10 20 30 40 50 60 70 80 90 100 110
 <EQRFPRRYIKLAIIVVDHGIVTKHHGNLKKIRKWIYQLVNTINNIYRSLNILVALVYLEIWSKQNKITVQSASNTLDLFGDWRESVLLKQRSHDCAQLLTTIDFDGPTIG
 HR1B PG
 QRFPRRYIKLAIIVVDHGIVTKHH
 M1K2 LAIVVDHGIVTK M1K4 WIYQLVNTINNIYRSLNILVALVYLEIWSK M1K6 QRSHDCAQLLTTIDFDGPTIG
 M1K1 PG M1K3 HHGNLK M1K5 ITVQSAS-VTLDLFG-W-E-VLL M1D5 DCAQLLTTI
 RFPRRYIK M1D2 DHGIVTKHHGNLKKIRKWIYQLVNTINNIYRSLNILVALVYLEI-SKQNKITV M1D4 DWRESVLLKQRSh M1D6 DGPTIG
 M1W1 SKQNKITVQSAS-VTLDLFGD ST2 SHDCAQLLTTIDFDGPTIG

120 130 140 150 160 170 190 200 210 220
 KAYTASMCDPKRSVGIVQDYSPI NLVVAVIMTHEMGNLGI PHDGNSTCGGFPCIMSPMISDPPSELSNCSKAYYQTF LTHDKPQCILNAPSKTDIVSPPVCGNELLE
 M2 (M1K6) CDPKRSVGIVQDYSPI NLVVAVIMTHEM M4-5 SPMISDPPSELSNCSKAYYQTF L
 K M3 GHNLGIPH DGNSTCGGFPCIM M5-DP PPSELSNCSKAYYQTF LTHDKPQCILNAPSKTDIVSPPVCGNELLE
 M1K7 AYTASM M5 ISDPPSELSNCSKAYYQTF LTHDKPQCILNAPSKTDIVSPPVCG
 (M1D6) KAYTASm M3-4-5 GHNLGIPH DGNSTCGGFPCI-SP M5K1 AYYQTF LTHDK M5K3 TDIVSPPVCGNELLE
 (ST2) -AYTASMDP R1 SVGIVQDYSPI NLVVAVIMTHEMGNLGI PH M5D1 ELFSnCSKAYY M5D3 DHKPQCILNAPSKT

230 240 250 260 270 280 290 300 310 320
 AGEECDGSPENCQYQCCDAASCKLH SWVKCESGECCDQCRFR TAGTECRAAESECDIPESCTGQSADCP TDRFHRNGQPCLYNHGYCYNGKCPIMFYQC YFLFGSNATV
 (M5K3) AGEECDGSPENCQYQCCDAASCK M5K5 CESGECCDQCRFR TAGTECRAAESECDIPESC-GQSA-CP M6 FYQC YFLFGS-ATV
 M5D6 M5K4 LHSWVK M5D8 DQCRFR TAGTECRAA M5D10 DIPESCTGQSA M5D12 DRFHRNGQPCLYNHGYCYNGK M5K6 CPIM
 DCGSPENCQYQCC M5D5 EEC M5D7 DAASCKLH SWVKCESg ST3 AAES ECDIPESCTGQSADCP TDRFHR M6K1 FYQC YFLFGS-ATV
 ST4 ST5 NGQPCLYNHGYCYNG-CPIMFYQC-FL-G
 ST4-5 AAES ECDIPESCTGQSADCP TDRFHR---P-LY

340 350 360 370 390 400 410 416
 AEDDCFN NNKKGDKYFYCRKENEKYIPCAQEDVKCGR LFC DNKKYPCHYnYSEDLDFGMVDHGT KCADGKVC SNRQCVDVNEAYKS
 (M6) AEDDCFN NNKKGDKYFYCRKENEKYIPCAQEDVKCG-LFC M6K5 VDHGT KCADGKVC SNRQCVDVNEA
 M6K2 YFYCRK M6K4 KYPCHYnYSEDLDFGM M7 VDHGT KCADGKVC SNRQCVDVNEAYK
 (M6K1) AEDDCFN NNK M6K3 YIPCAQEDVK M6D6 DNKKYPCHYnYSE M7' VDHGT KCADGKVC SNRQCVDVNEAYKS
 M6D2 DDCFN NNKKG ST6 LFC DN--YPCHY-YSEDLDFGMVDHGT-CADG-VCSNR
 M7'' VDHGT KCADGKVC SNRQCVDVNEAYKS

Fig. 1

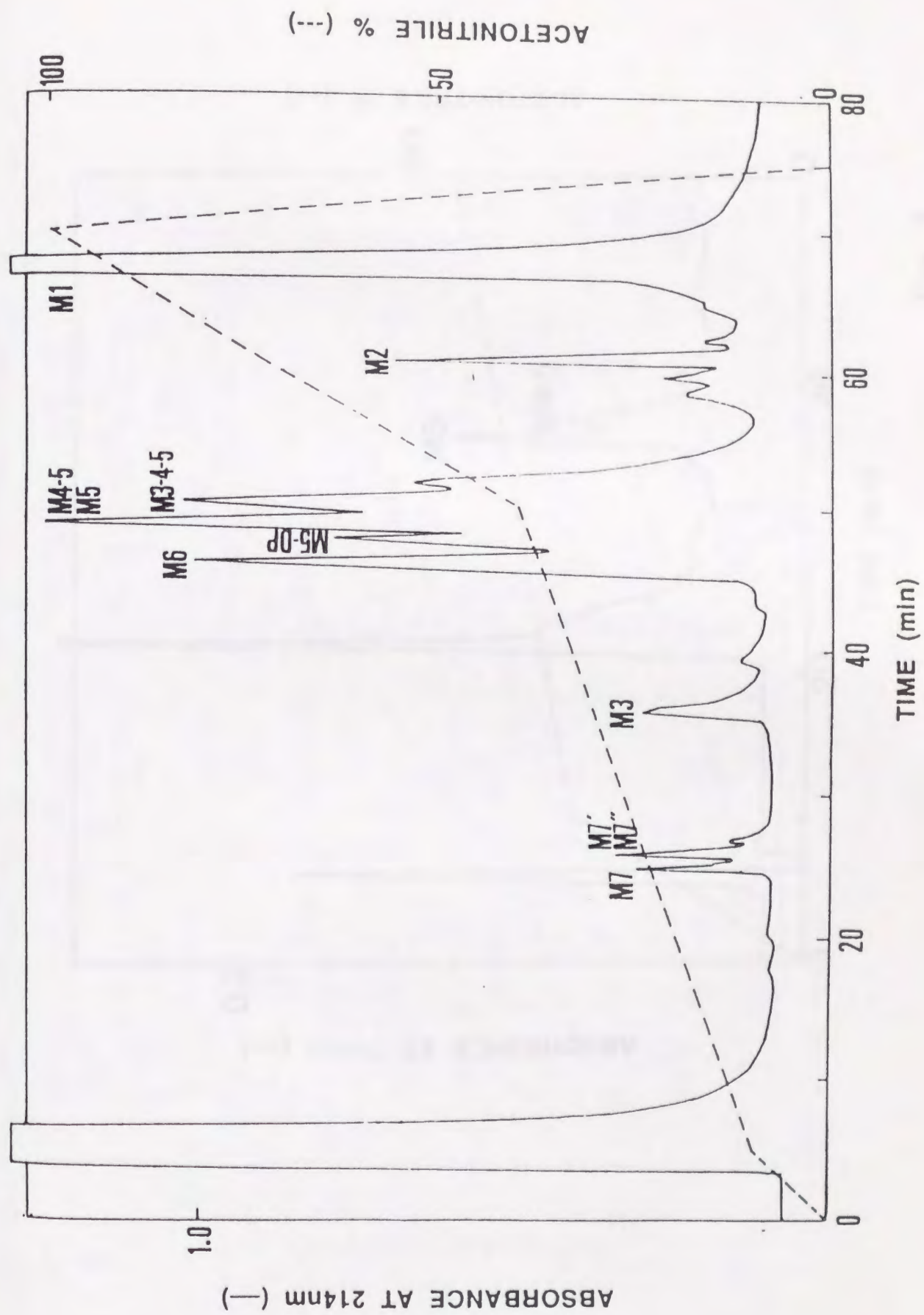


Fig. 2

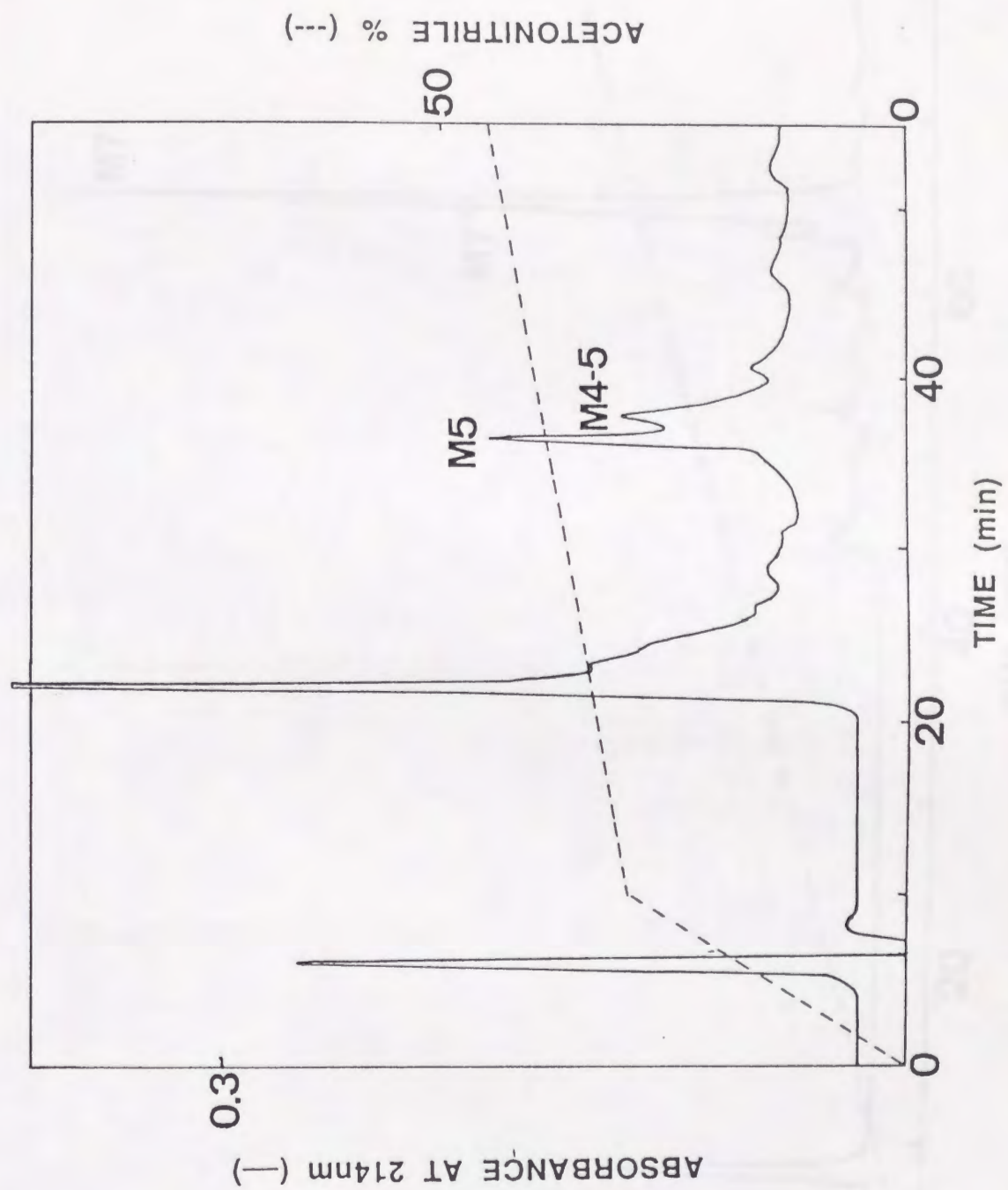


Fig. 3

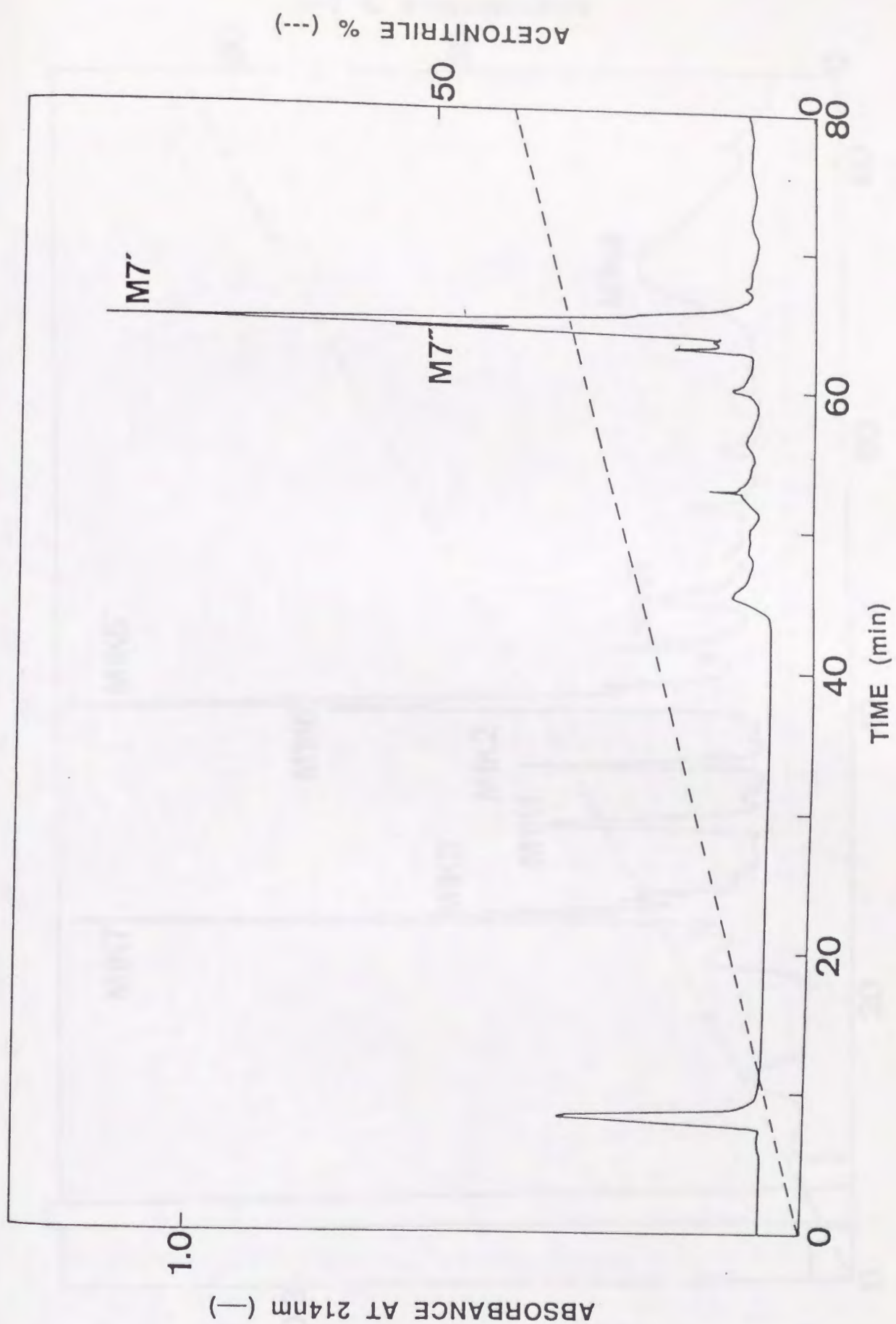


Fig. 4

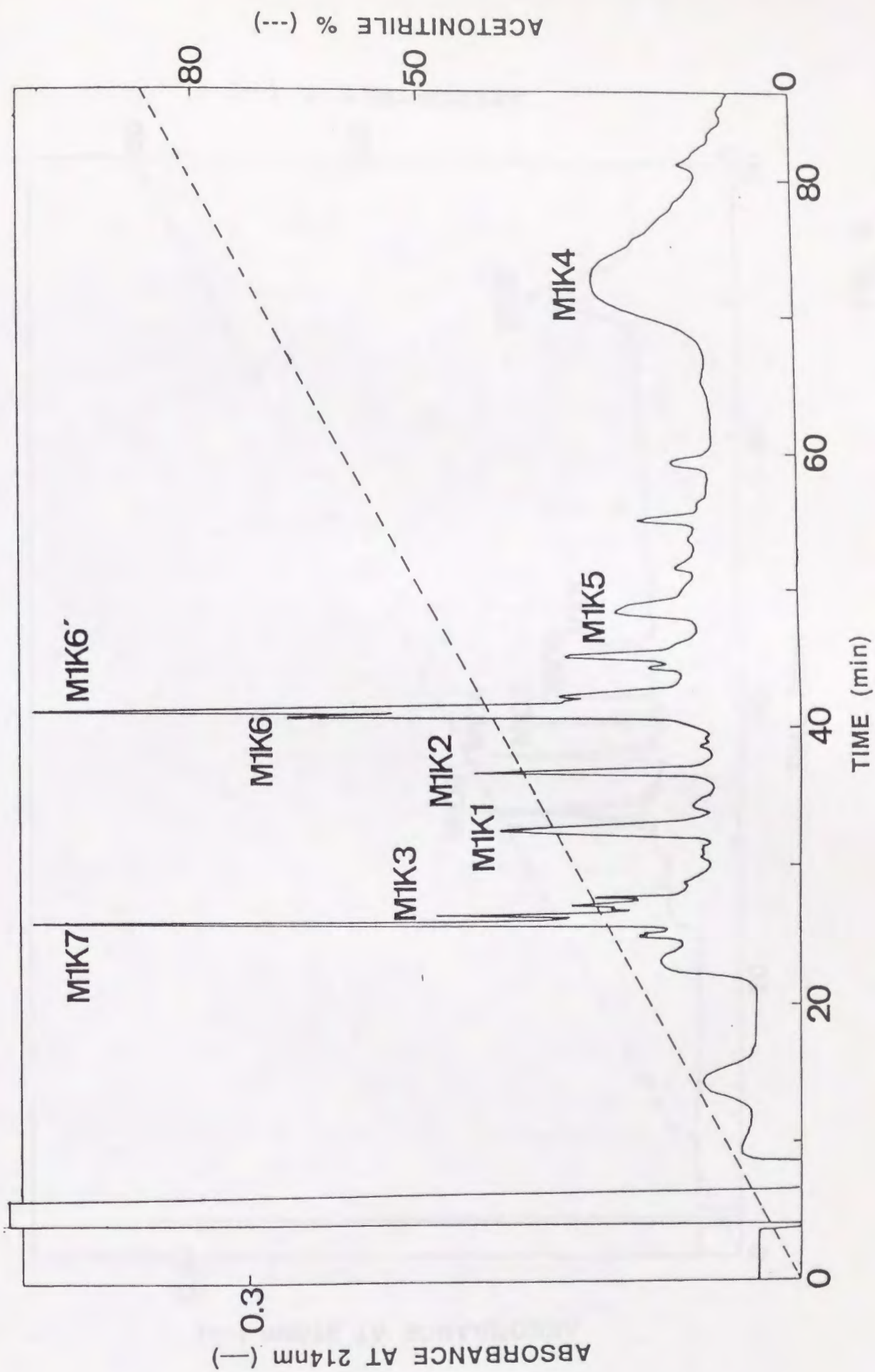


Fig. 5

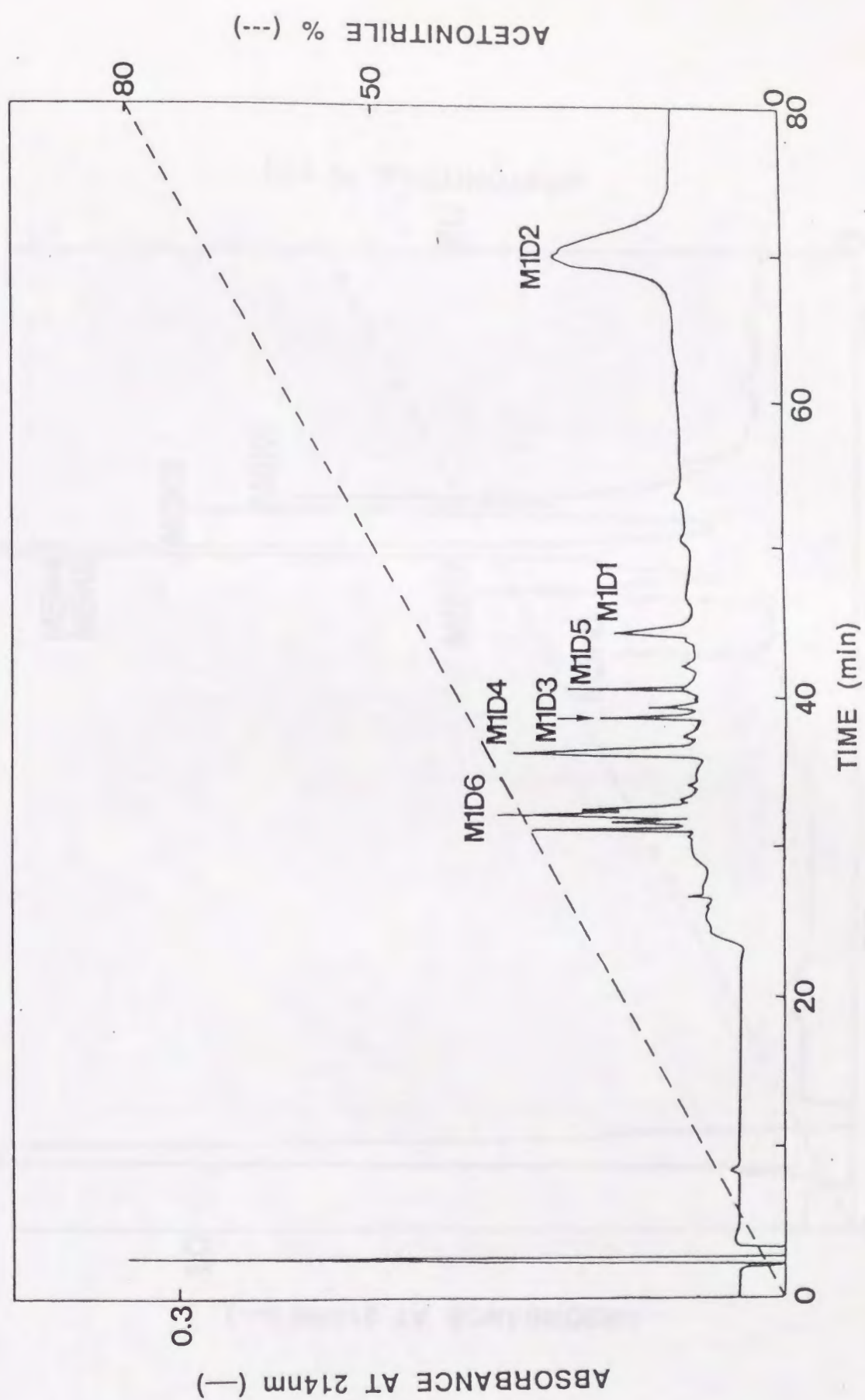


Fig. 6

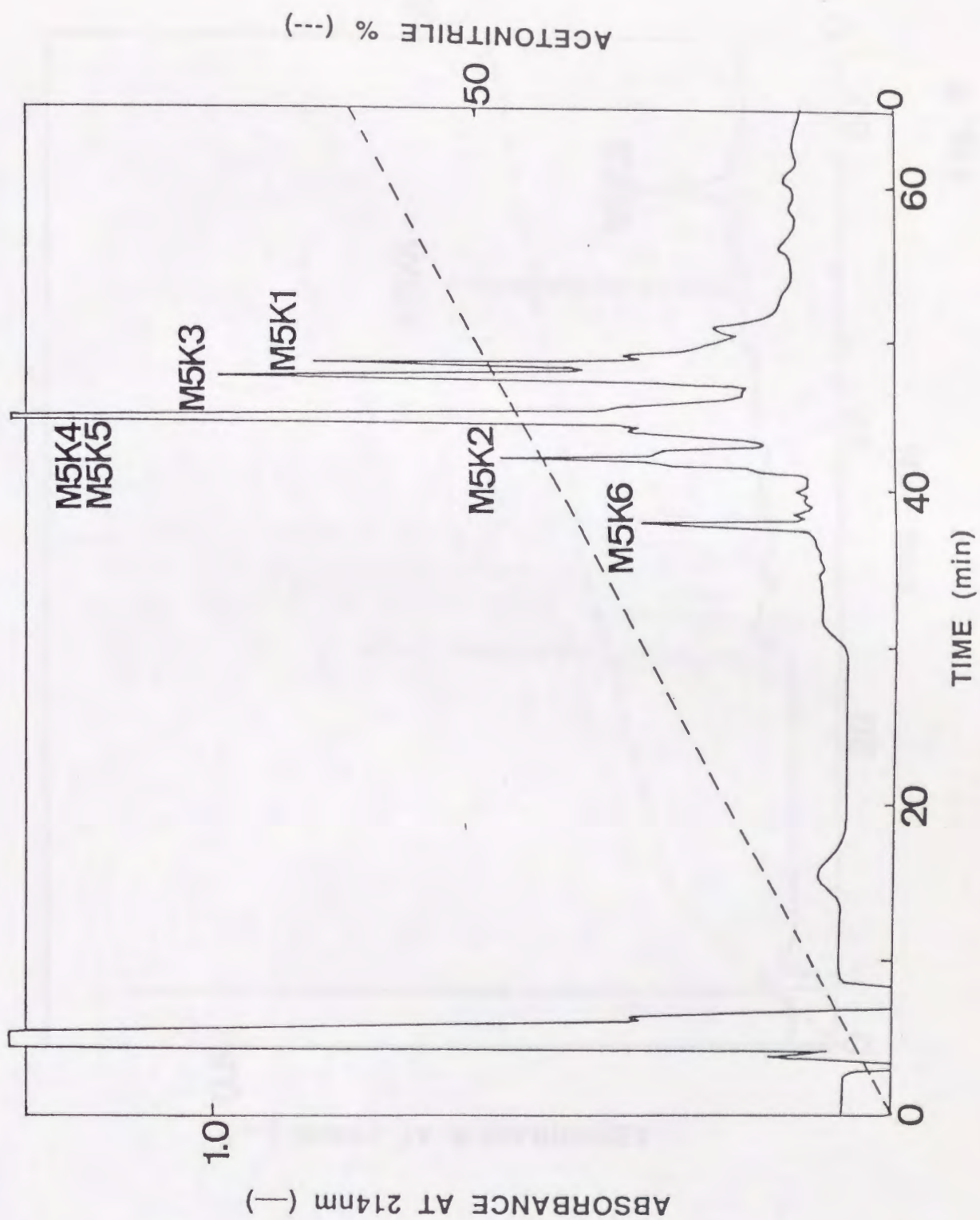


Fig. 7

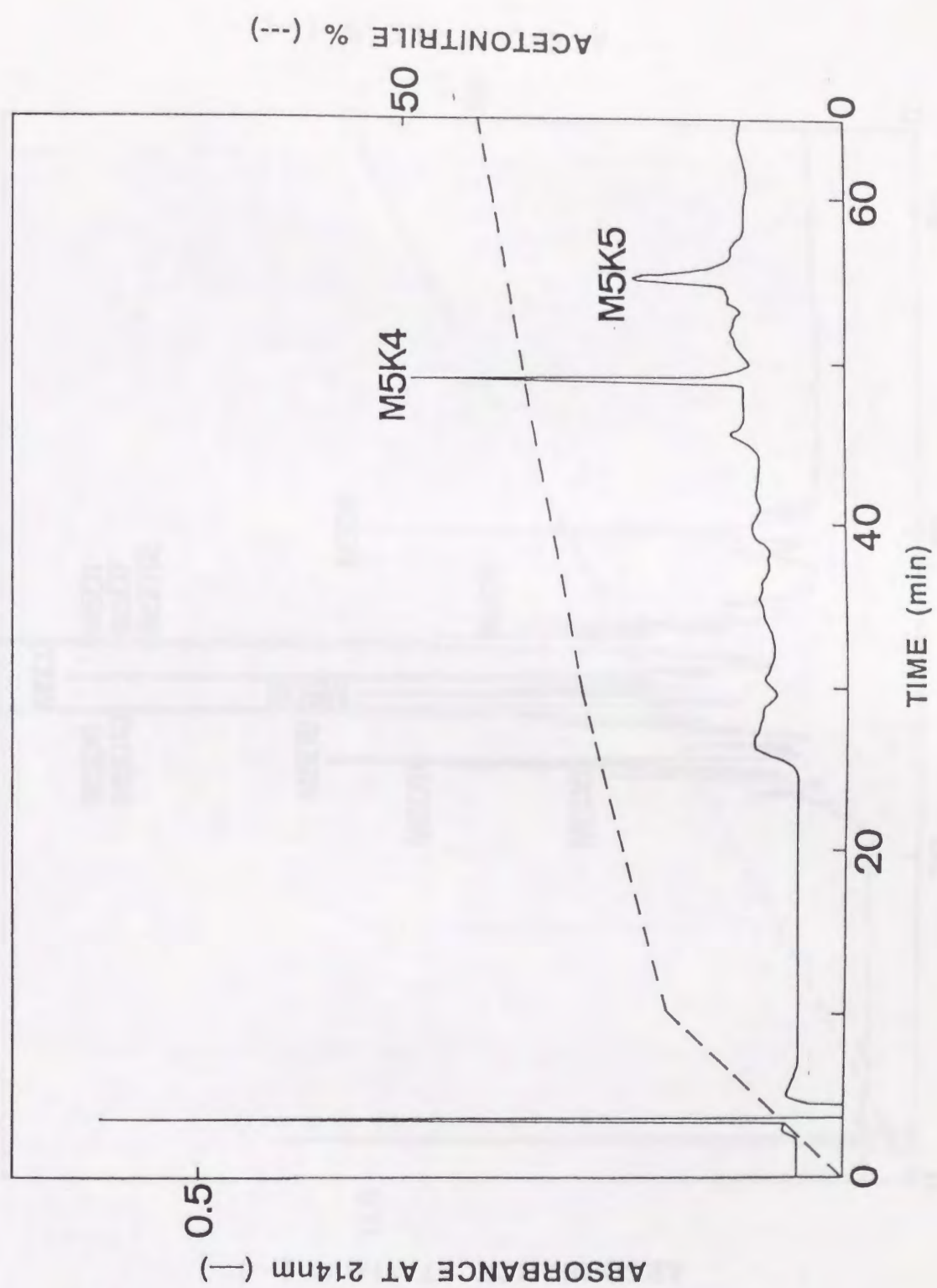


Fig. 8

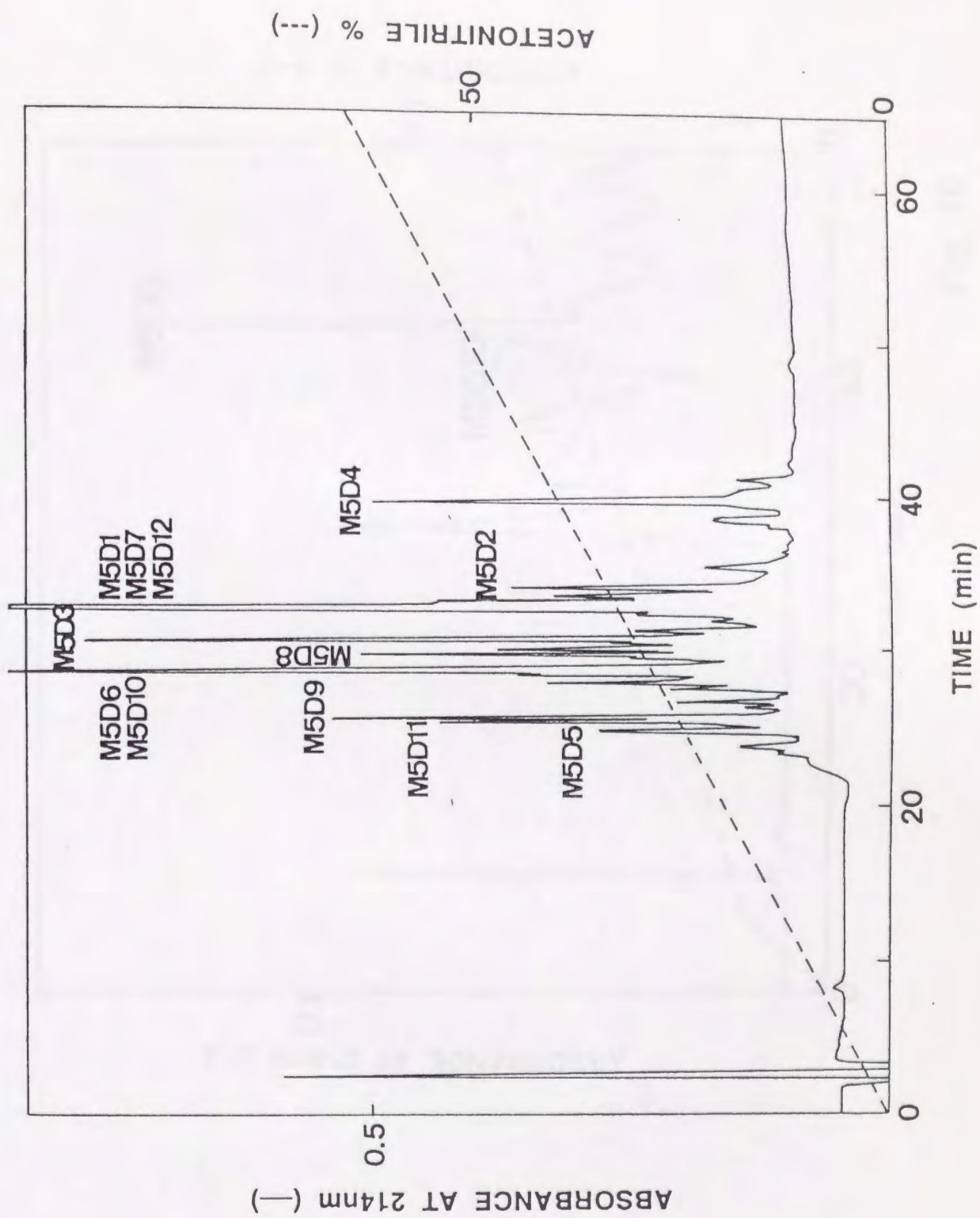


Fig. 9

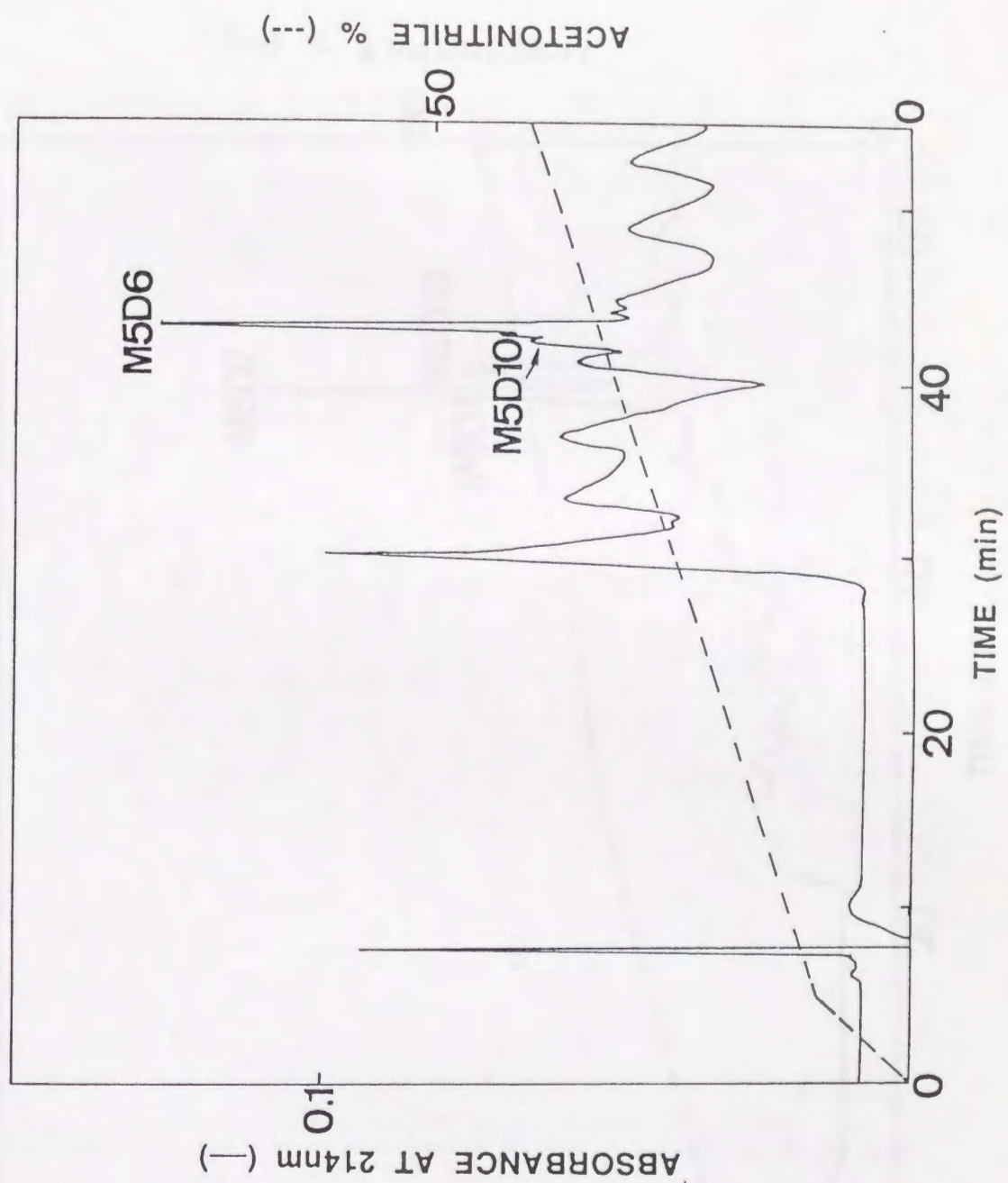


Fig. 10

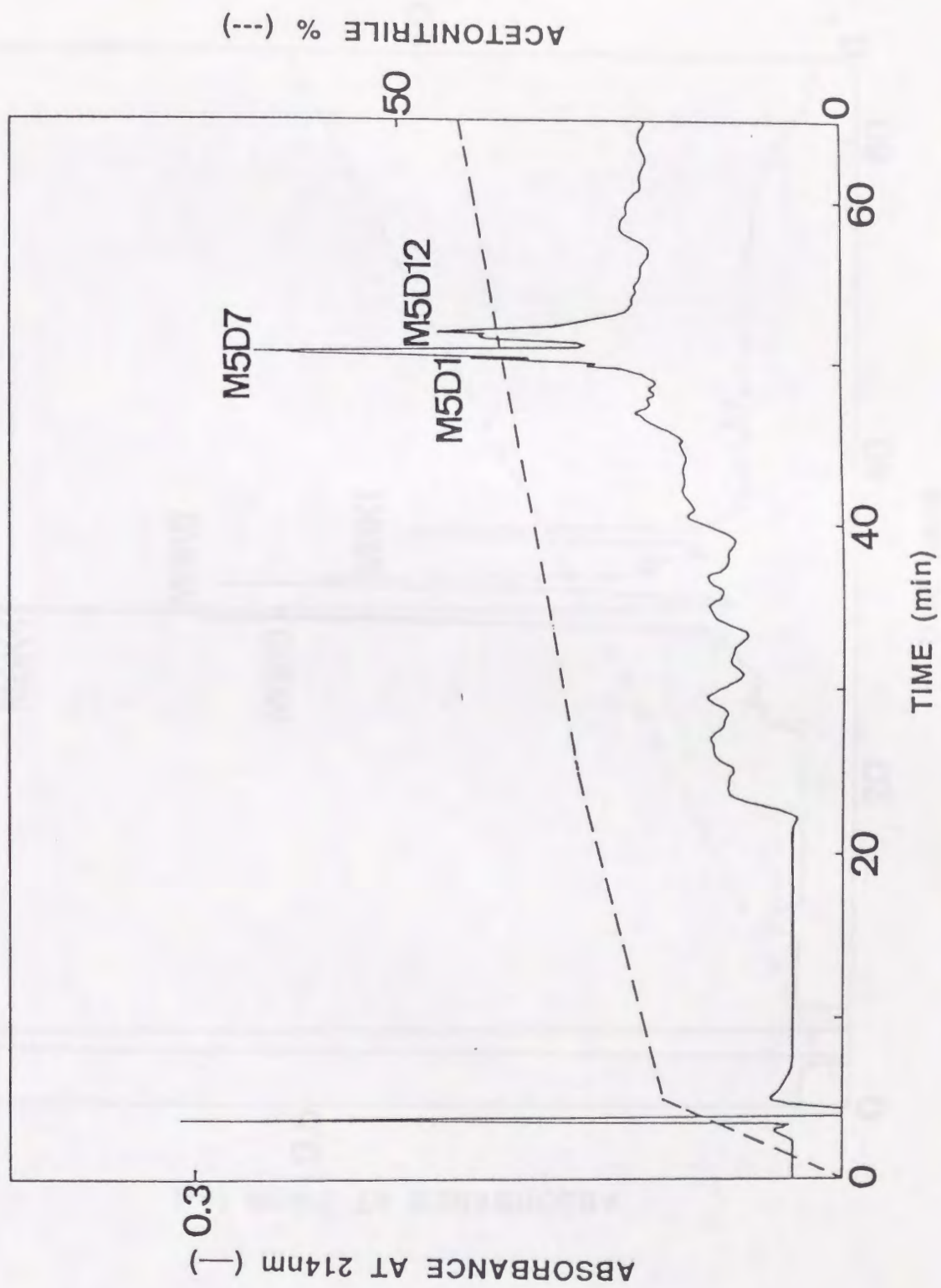


Fig. 11

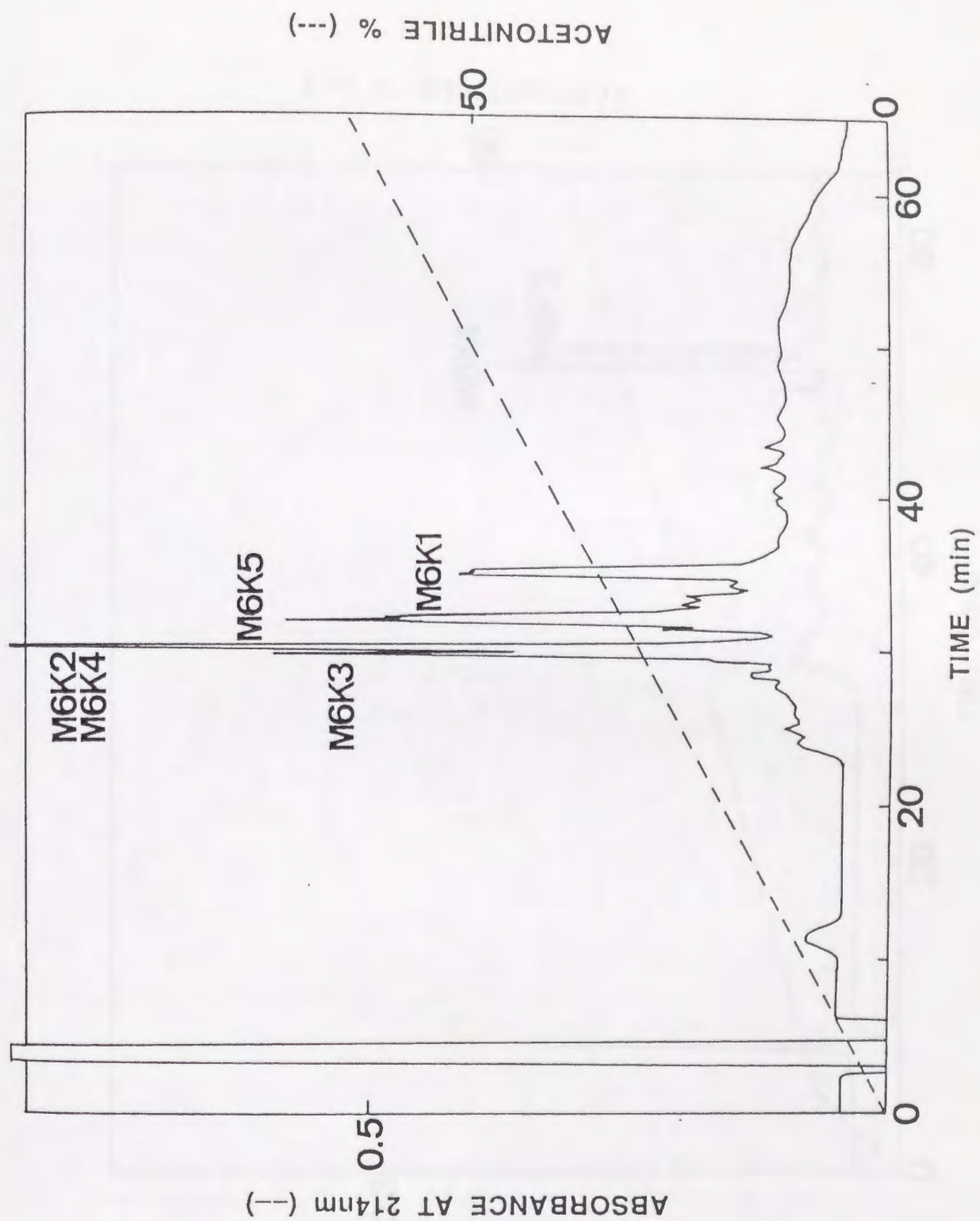


Fig. 12

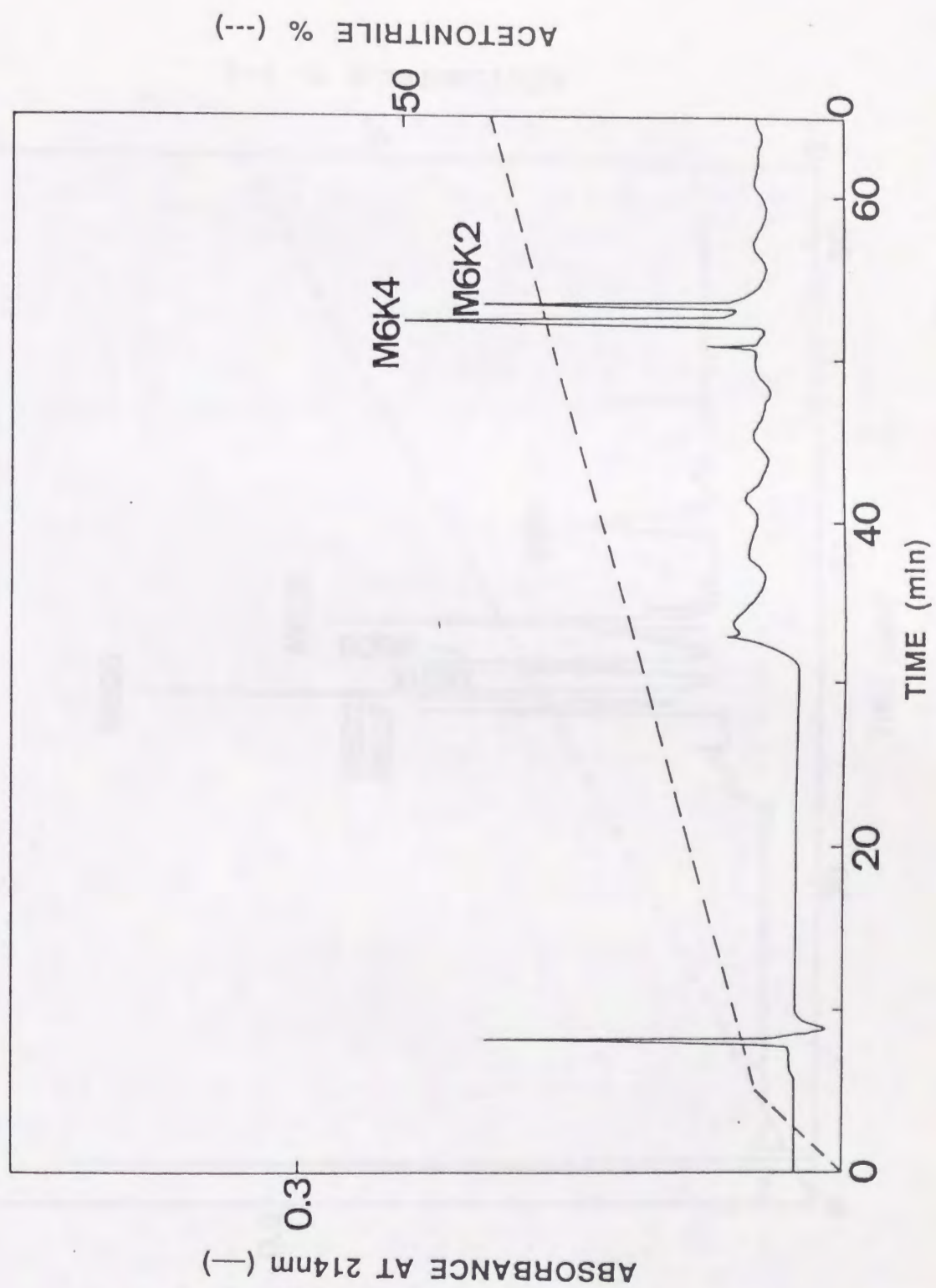


Fig. 13

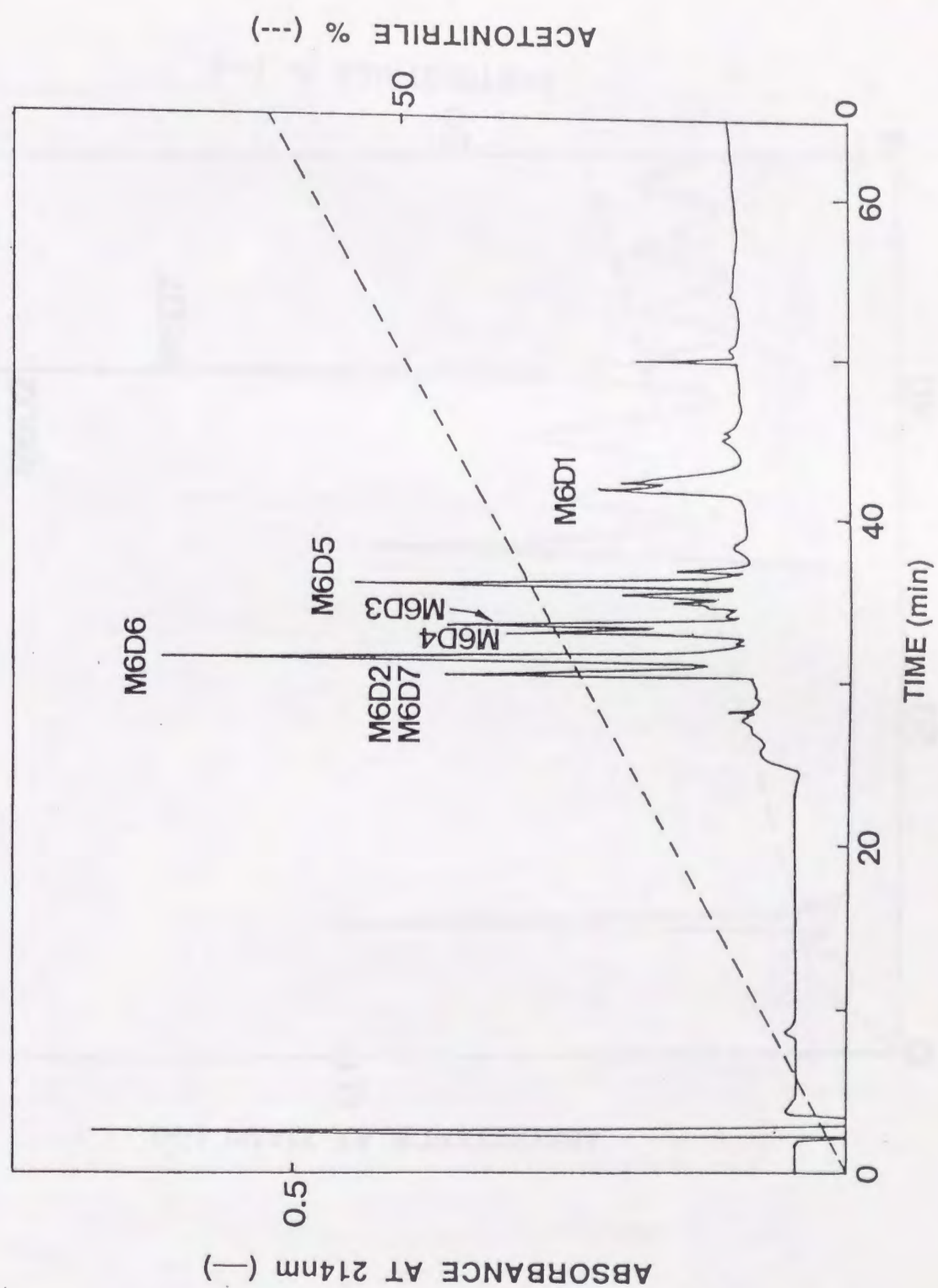


Fig. 14

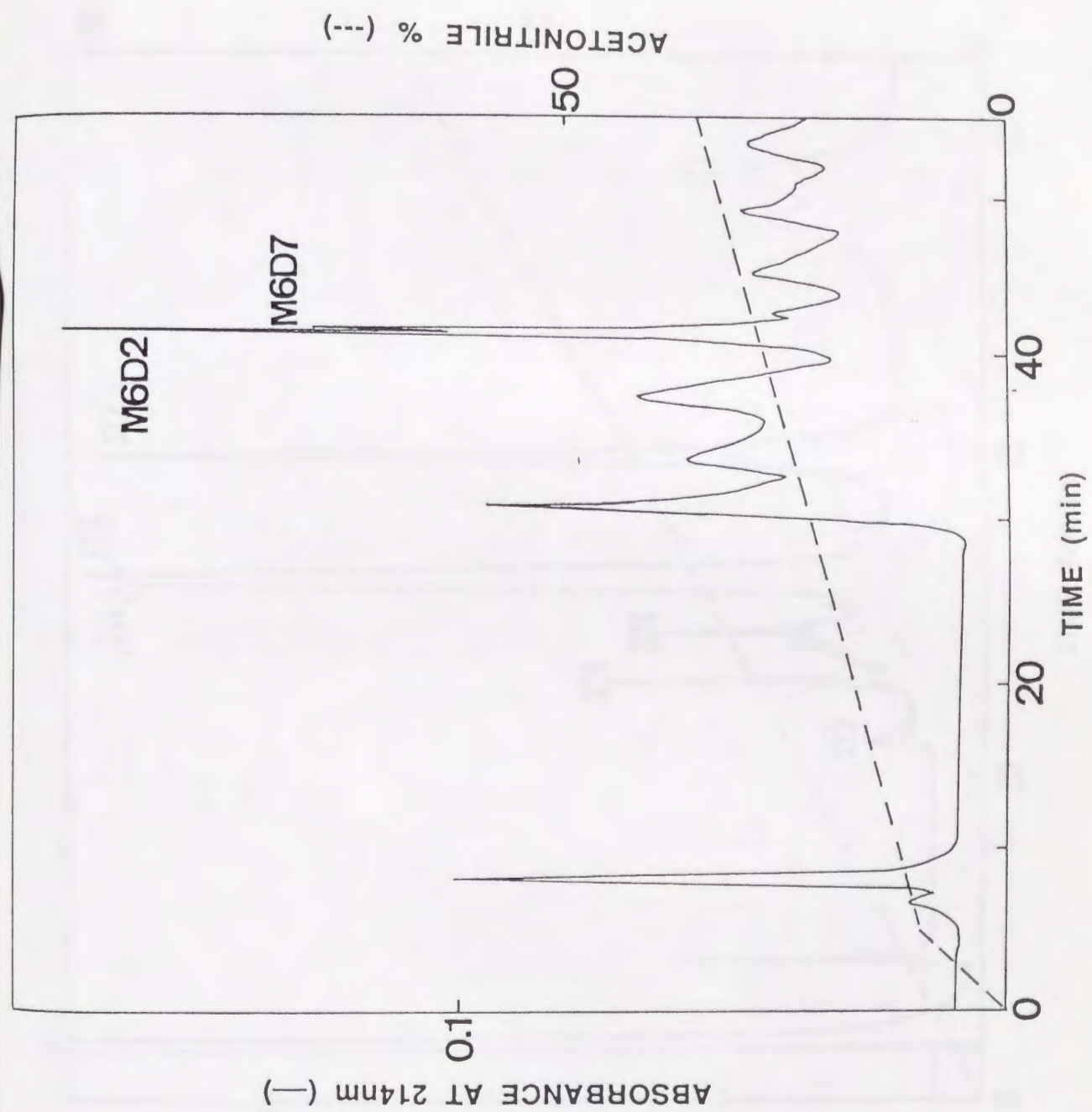


Fig. 15

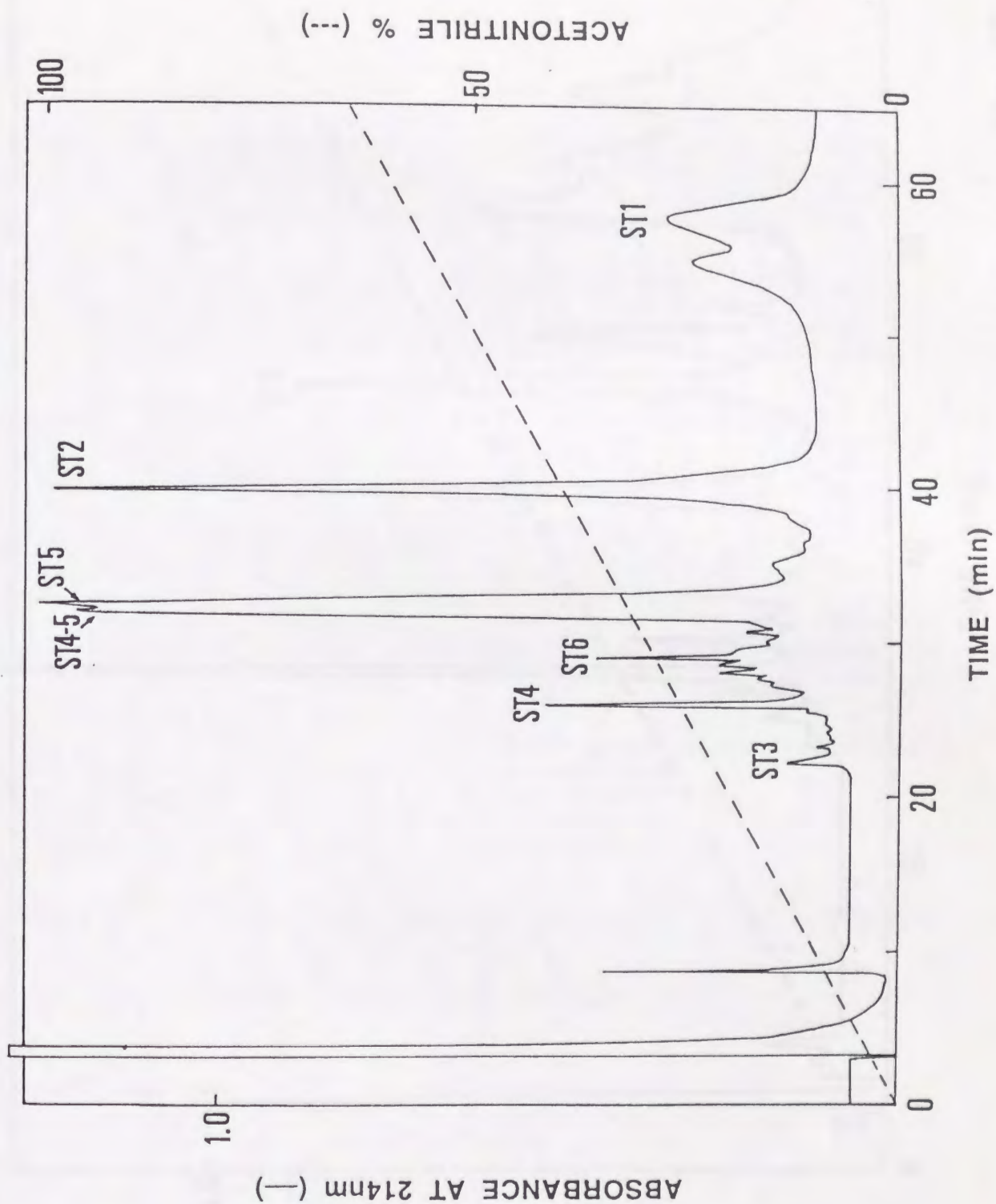
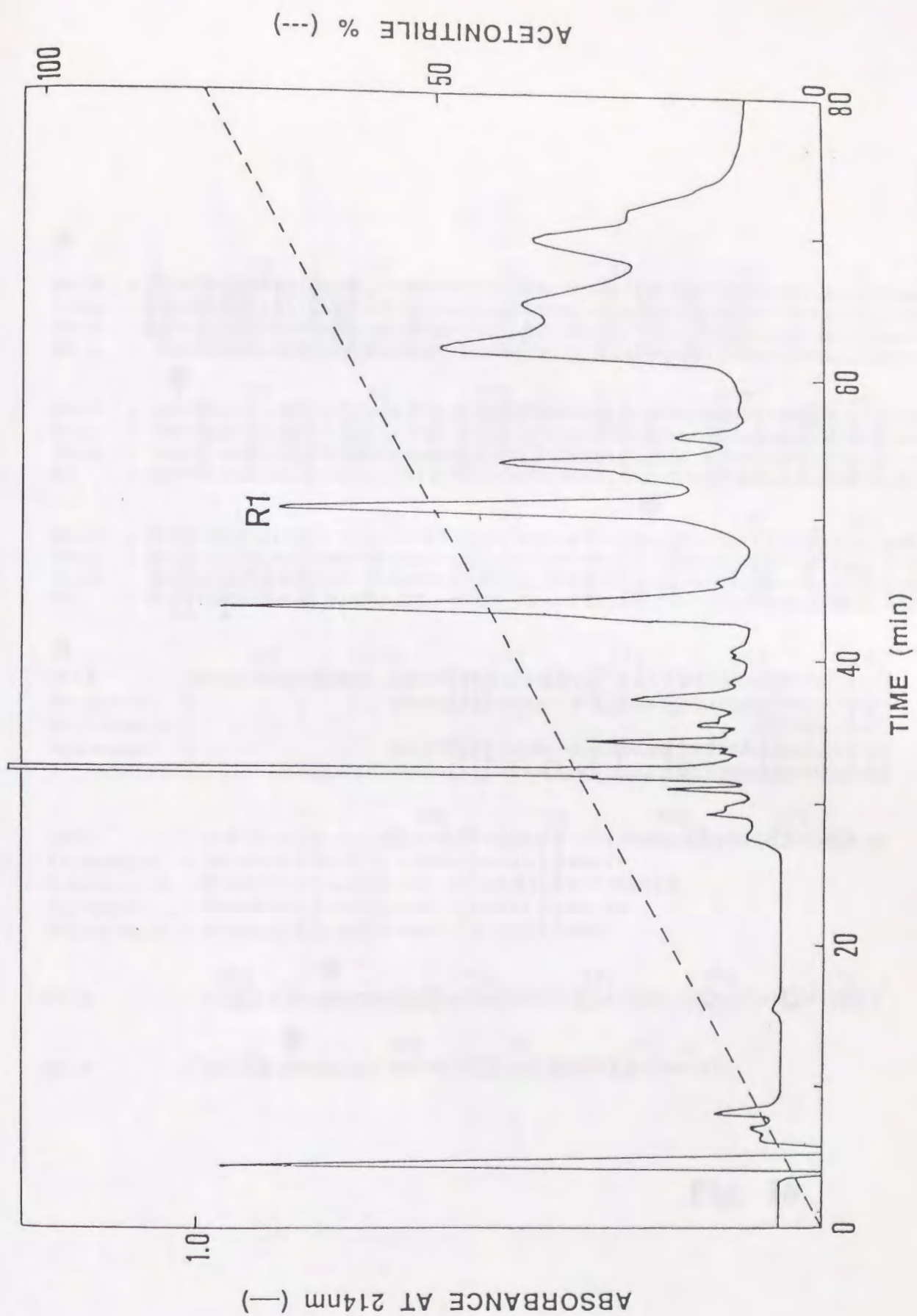


Fig. 16

Fig. 17



A

1 10 20 30 40 50 60 70
 HR1B : <EQRFPRRYIKLAIVVDHGIVTKHHGNLKKIRKWIYQLVNTINNIYRSLNLLVALVYEIWSKQNKITVOS
 HR2a : <EQRFQORYIELAIVVDHGMITYSSNFKKIRKRVHQMVMNNINEMYRPLNIAITLSLLDVSEKDLITMOA
 Ht-d : <EQNLQORYIELVVDADHRVFMKYNSDLNTIRTRVHEIVNFINGFYRSLNIHVSLTDLEIWSNEDQINLOS
 H2 : <ERFPQORYIELAIVVDHGMYYKNQNSDKIKRVHQMVMNHINEMYRPLNIAISLNLQIWSKKDLITVKS
 80 90 100 110 120 130 140
 HR1B : ASNVTLDLFGDWRESVLLKQSRSHDCAQLLTIDFDGPTIGKAYTASMCDPKRSVGIQVODYSPINLVVAVI
 HR2a : VAPTARLEGDWRETVLLKQKDHDAQLLTIDINFTGNTIGWAYMGGMCAKNSVGIQVODHSSNVFMVAVT
 Ht-d : ASSDTLNAAFAEWRETDLNLRKSHDCAQLLTIDELDEETLGLAPLGTMCDPKLSIGIIVODHSPINLLMGVT
 H2 : ASNVILESEGNWRETVLLKQONNDCAHLLTATNLNDNTIGLAYKKGMCPKLSVGLVQDYSPNVFMVAVT
 150 160 170 180 190 200
 HR1B : MTHEMGHNLGIPHDG-NSCTCGGFPCIMSPMISDPPSELEFSNCSKAYYQTFELTDHKPQCIINAP
 HR2a : MTHEIGHNLGMEHDDKDKCKCEA--CIMSAVISDKPSKLFSDCSKDYQTFELTNSKPOCIINAP
 Ht-d : MAHELGHNLGMEHDKD-CLRGASLCIMRPGTLKGRSYEFSDDSMHYYERFLKQYKPOCILNKE
 H2 : MTHELGHNLGMEHDDKDKCKCEA--CIMS DVISDKPSKLFSDCSKNDYQTFELTKYNPQCIINAP
 ▲▲

B

210 220 230 240 250 260
 HR1B : SKTDIVSPPVGNELLEAGEECDCGSPENQYQCCDAASCKLHSWVKCESGECCDQC
 Trigramin : EAGEDCCGSPAN--P-CDAATCKLIPGAQCGEGLCCDQC
 Echistatin : ECESGPCCRNCC
 Applaggin : EAGEECDCGSPEN--P-CDAATCKLRPGAQCAEGLCCDQC
 Bitistatin : SPPVGNELLEAGEECDCGSPANCCDCCNAATCKLTPGSCCNHGECCDQC
 270 280 290 300 310
 HR1B : RFRTAGTECRAAESECDIPESCTGQSADCTDRFHRNGQPCLYNHGYCYNGKCPIM
 Trigramin : SFIEEGTVCRIRAGD-DLDDYCNGRSAGCPRNPFH
 Echistatin : KFLKEGTICKRAAGD-DMDDYCNKGTDCPRNPHKGPAT
 Applaggin : KFMKEGTVCRAAGD-DVNDYCNGISAGCPRNPFH
 Bitistatin : KFKKARTVCRIAGD-WNDDYCTGKSSDCPWNH
 320 330 340 350 360 370
 HR1B : FYQCYFLFGSNATVAEDDCFNNNKKGDKYFYCRKENEKYIPCAQEDVKCGRLECDN
 380 390 400 410
 HR1B : KKYPCHYNSEDLDLFGMVDHGTCKADGKVCNRCQVDVNEAYKS

Fig. 18

213 230 250 270
 HR1B: VCGNELLEAGEECDCGSPENCQYQC-CDAASCKLHSWVKCESGECCDQCRFRTAGIECRAA-ESE
 vWF : LCEVARLRQNADQC-C-PE--YECMCDPVSCDLPPVPHCERGL--OPTLTNPG-ECRPNFTCA
 1543 1560 1580
 290 310 336
 CDIPESCTGQSA-DCPTDRFH--RNGQPCLYNHGYCYNGKCPIMFYQCYF-LFGSNATVAEDDCF
 CRK-ECKRVSPSPCPHRLPTLRKTOCC--DEYECA-CNCVNSTVSCPLGYLASTAT---NDCG
 1600 1620 1640 1656

Fig. 19

PART II

Coagulation Factor X Activating Enzyme from Russell's Viper Venom (RVV-X) : A NOVEL METALLOPROTEINASE WITH DISINTEGRIN (PLATELET AGGREGATION INHIBITOR)-LIKE AND C-TYPE LECTIN-LIKE DOMAINS

SUMMARY

We determined the complete amino acid sequence of RVV-X, the blood coagulation factor X activating enzyme, isolated from Russell's viper venom and studied the structure-function relationship. RVV-X (Mr 79,000) consisted of a disulfide-bonded two-chain glycoprotein with a heavy chain of Mr 59,000 and a light chain of heterogeneous Mr 18,000 or 21,000. These chains were separated by gel filtration after reduction and S-pyridylethylation. Two heterogeneous light chains, designated LC1 (Mr 18,000) and LC2 (Mr 21,000), were separated by reversed phase high performance liquid chromatography, and the isolated major component LC1 was used for sequence analysis. The heavy chain consisted of 427 residues containing four asparagine-linked oligosaccharides, and its entire sequence was similar to that of the high molecular mass hemorrhagic protein, HR1B, isolated from the venom of *Trimeresurus flavoviridis* (Takeya, H., Oda, K., Miyata, T., Omori-Satoh, T., and Iwanaga, S. (1990) *J. Biol. Chem.* 265, 16068-16073). The heavy chain contained three distinct domains, metalloproteinase, disintegrin (platelet aggregation inhibitor)-like and unknown cysteine-rich domains. On the other hand, the light chain LC1 consisted of 123 amino acid residues containing one asparagine-linked oligosaccharide and showed a sequence homology similar to those found in the so-called C-type (Ca^{2+} -dependent) lectin. Therefore, RVV-X is a novel metalloproteinase containing a mosaic structure with disintegrin-like, cysteine-rich, and C-type lectin-like domains. RVV-X potently inhibited collagen- and ADP-stimulated platelet aggregations, probably via its disintegrin-like domain, although

this domain does not contain the Arg-Gly-Asp sequence which is conserved in various venom disintegrins and which is thought to be one of the interaction sites for platelet integrins. Our findings also indicate that snake venom factor IX/factor X-binding protein with a C-type lectin structure (Atoda, H., Hyuga, M., and Morita, T. (1991) *J. Biol. Chem.* 266, 14903-14911) inhibits the RVV-X-catalyzed factor X activation, hence, the light chain of RVV-X probably participates in recognizing some portion of the zymogen factor X.

INTRODUCTION

Snake venoms have various proteins and enzymes which affect mammalian blood coagulation and fibrinolytic systems (1). These components have a strict specificity in their contacts with blood coagulation factors (2-5). Russell's viper (*Vipera russelli*) venom contains two well known proteases, designated RVV-V and RVV-X, both of which induce coagulation of mammalian plasma. We reported earlier that RVV-V, which specifically activates factor V by limited proteolysis, consists of 236 amino acid residues and shows a sequence homology similar to that of trypsin-like serine proteinases, especially batroxobin, a thrombin-like enzyme isolated from the venom of *Bothrops atrox* (6).

RVV-X, a potent activator of factor X, is a well characterized proteinase (7) which activates specifically factor X as a result of a single cleavage at the same internal Arg-Ile bond in factor X as do factors IXa and VIIa (8,9). The RVV-X-catalyzed factor X activation, however, is not inhibited by diisopropyl fluorophosphate and phenylmethylsulfonyl fluoride but is inhibited by EDTA, thereby suggesting that RVV-X is not a serine proteinase but a metalloproteinase (10,11). Indeed, RVV-X contains nonexchangeable 1 mol of Ca^{2+} and 0.7 mol of Zn^{2+} essential for the proteolytic activity (12). Unlike vitamin K-dependent clotting factors, RVV-X does not require a negatively charged surface such as phospholipids for factor X activation, but it does require exogenous Ca^{2+} and the amino-terminal Gla-domain of factor X for enhanced activation (13). This reversibly bound Ca^{2+} is not essential for the proteolytic activity of RVV-X, since RVV-X is

able to hydrolyze apoprotein AI of human high-density lipoprotein, even in the absence of exogenous Ca^{2+} (12).

RVV-X is a glycoprotein containing 13% carbohydrate with an apparent Mr of 79,000. It is composed of two disulfide bonded chains, a heavy chain of Mr 59,000 and heterogeneous light chains with Mr 18,000 and 21,000. In the present study, we determined the entire amino acid sequence of RVV-X, in order to elucidate the molecular mechanism of the RVV-X-catalyzed factor X activation, in particular how RVV-X specifically recognizes factor X. Our interest in the structure of RVV-X is because it is a unique metalloproteinase with respect to strict substrate specificity (7). Since we identified the entire amino acid sequences of several hemorrhagic and non-hemorrhagic metalloproteinases isolated from snake venoms and as these are new members of the metalloproteinase subfamily (14-17), it was also of interest to compare the covalent structures between RVV-X and the venom metalloproteinases.

MATERIALS AND METHODS

Materials ——— The sources of materials used were as follows: The venom of *Vipera russelli siamensis* from the Japan Snake Institute, Gunma; *Achromobacter lyticus* lysyl endopeptidase, and 4-vinylpyridine from Wako Pure Chemical Industries, Osaka; endoproteinase Asp-N from Boehringer-Mannheim Biochemica, Mannheim; trypsin treated with N-tosyl-phenylalanyl chloromethyl ketone and α -chymotrypsin from Worthington Biochemical, Freehold, N. J.; *Staphylococcus aureus* V8 protease from Miles Laboratories, Elkhart, Ind.; cyanogen bromide, and Cosmosil 5C4-300 column from Nacalai Tesque, Inc., Kyoto; Phenyl-5pw RP, G3000SW and G2000SW columns from Tosoh, Osaka; Chemcosorb 5-ODS-H column from Chemco Scientific Co., Ltd., Osaka; Vydac 214TP5415 column from The Separations Group, CA; μ Bondasphere C8 column from Nihon Waters Ltd., Tokyo; YMC A-302 S-5 ODS column from Yamamura Kagaku, Kyoto; reagents for gas-phase sequencer from Applied Biosystems, CA. All other chemicals were of analytical grade or of the highest quality commercially available. Factor IX/factor X-binding protein (IX/X-bp) purified from the venom of *Trimeresurus flavoviridis* (18) was a gift of Dr. Hideko Atoda, Meiji College of Pharmacy.

Purification of RVV-X ——— RVV-X was isolated from the venom of *Vipera russelli siamensis* by minor modifications of published methods (11), consisting of two steps, gel filtration on Sephadex G-75 and ion-exchange chromatography on DEAE-Sephadex CL-6B.

SDS-PAGE ——— SDS-PAGE was carried out by the method of Laemmli (19). The reference protein kit for molecular weight estimate was obtained from Pharmacia Fine Chemicals, Uppsala.

Preparation of Heavy and Light Chains ——— The purified RVV-X was reduced and *S*-pyridylethylated as previously described (17). The resulting product was subjected to gel filtration HPLC on a G3000SW column equilibrated with 0.1 M sodium phosphate buffer, pH 6.0, containing 6 M guanidine hydrochloride and 1 mM EDTA. The light chain fraction was subjected to reversed phase HPLC on a Vydac 214TP5415 column in 0.1% trifluoroacetic acid with acetonitrile gradient elution.

CNBr and Proteolytic Digestion ——— The sample was cleaved with CNBr as previously described (17). The sample in 50 mM Tris-HCl buffer, pH 8.0, containing 4 M urea was digested with lysyl endopeptidase (E/S=1/50, mol/mol) at 37 °C for 12 h. The samples in 50 mM Tris-HCl buffer, pH 8.0, containing 2 M urea were digested with endoproteinase Asp-N (E/S=1/50, mol/mol) for 13h, trypsin (E/S=1/20, mol/mol) for 12 h, or chymotrypsin (E/S=1/50, mol/mol) for 6 h at 37 °C. The samples in 0.1M NH_4HCO_3 buffer, pH 8.0, containing 2 M urea were digested with *Staphylococcus aureus* V8 protease (E/S=1/70, mol/mol) at 37 °C for 17 h. For limited proteolysis, non-reduced RVV-X in 50 mM Tris-HCl buffer, pH 7.5, containing 0.1 M NaCl and 10 mM EDTA was digested with trypsin (E/S=1/50, mol/mol) at 37 °C for 3 h.

Peptide Purification ——— The peptides generated after CNBr cleavage of Pe-heavy chain were separated by reversed phase HPLC after gel filtration HPLC on a column of G2000SW, equilibrated with 0.1 M phosphate buffer, pH 6.0, containing 6 M guanidine

hydrochloride and 1 mM EDTA. All other peptides were purified by reversed phase HPLC as previously described (17).

Amino Acid Analysis and Sequence Determination ——— The amino acid analyses of the Pe-proteins were performed by ion-exchange chromatography in a Hitachi model L-8500 high speed amino acid analyzer after hydrolysis with 5.7 M HCl containing 0.2 % phenol at 110 °C for 24, 48, and 72 h by the method of Spackman et al. (20). Tryptophan was determined by hydrolysis in 3 M mercaptoethanesulfonic acid (21). The CNBr fragments derived from Pe-heavy chain were analyzed using a Hitachi L-8500 amino acid analyzer after hydrolysis with 5.7 M HCl containing 0.2 % phenol at 110 °C for 24 h. Other peptides were analyzed using reversed-phase HPLC of phenylthiocarbamoyl derivatives (22), using a Waters PICO-TAG system, after hydrolysis with 5.7 M HCl containing 1 % phenol at 110 °C for 20 h. Automated sequence analyses were performed with an Applied Biosystems 477A protein sequencer, as described by Hewick et al. (23), with an Applied Biosystems model 120A PTH analyzer.

Determination of the Carboxyl-Terminal Amino Acid ——— The carboxyl-terminal amino acid was determined by the vapor-phase hydrazinolysis method as previously described (16,24).

Platelet Aggregation Assay ——— Platelet aggregation assay was performed in human platelet-rich plasma (PRP) according to the methods of Moroi (25). Fifty ml of whole human blood (9 parts) from a healthy donor was collected in 3.8 % (wt/vol) sodium citrate (1 part). Blood was centrifuged (1,800 rpm for 5 min) at room temperature and PRP was collected. Platelet-poor plasma (PPP) was prepared from the remaining blood by centrifugation (3,000 rpm for 10 min) at room temperature. PRP plus solution of

RVV-X in Tris-HCl buffer saline, pH 7.5 or buffer alone was incubated for 3 min in a Chrono-log Aggregometer at 37 °C. An aggregating agent (collagen, 1 µg/ml or ADP, 10 µM) was added and the light transmittance was recorded.

Effect of IX/X-bp on the RVV-X-mediated Activation of Factor X ——— Factor X (3.6 µM) was incubated with the various concentrations of IX/X-bp (0 - 7.2 µM) in 50 mM Tris-HCl, pH 8.0, containing 10 mM CaCl₂ for 1 min at 25 °C. Then, RVV-X (3.6 nM) was added. After 1 min, 10 µl aliquots of the reaction mixture were added to the substrate solution, which consisted of 600 µl of 50 mM Tris-HCl containing 0.1 mg/ml bovine serum albumin, 2 mM EDTA and 0.1 M NaCl, and 10 µl of 5 mM N^α-benzyloxycarbonyl-L-pyroglutamyl-Gly-Arg-4-methylcoumaryl-7-amide in *N,N*-dimethylformamide. The released 7-amino-4-methylcoumarin was measured fluorometrically with excitation at 380 nm and emission at 440 nm.

Nomenclature of the Peptides ——— Peptides are designated by a serial number prefixed by a letter. The letters indicated the type of digestion: M, CNBr; K, lysyl endopeptidase; D, endoproteinase Asp-N; T, trypsin; E, V8 protease; C, chymotrypsin. The numbers in the peptide designation do not correspond to the order of their elution in HPLC, but rather to their positions in the protein sequence, starting from the amino-terminus.

RESULTS

Preparation of Heavy and Light chains and Their Amino Acid

Compositions ——— Highly purified RVV-X was reduced and S-pyridylethylated (Pe), and subsequently subjected to gel filtration (Fig. 1). As shown in the inset of Fig. 1, the SDS-PAGE analyses of two peaks revealed that the former peak (peak HC) contains a homogeneous heavy chain (Mr 59,000) and the latter peak (peak LC) contains heterogeneous light chains (Mr 18,000 and 21,000) with a doublet. These two heterogeneous light chains were further separated by reversed phase HPLC (Fig. 2). Highly purified two light chains (the inset of Fig. 2) were designated as LC1 and LC2. In the present study, the major component, LC1, was selected as being representative and was used for sequence analyses. The amino acid compositions of heavy and light chains are shown in Table 1. Although LC1 and LC2 have similar amino acid compositions, their peptide maps after digestions with lysyl endopeptidase differed (data not shown), indicating a polymorphism with several residues. The heavy and light chains contained glucosamine, but not galactosamine (Table 1).

Sequence Analyses of Heavy Chain ——— The automated Edman degradation of Pe-heavy chain established the amino-terminal 31 residues starting with the sequence of Leu-Val-Ser-Thr-Ser-Ala- as shown in Fig. 3. In addition to this sequence, we could identify the minor amino-terminal sequences Val-Ser-Thr-Ala- and Ser-Thr-Ala- starting at positions 2 and 3, respectively. The Pe-heavy chain was cleaved with CNBr and the resulting peptides were separated by gel filtration (Fig. 4). M5 and M12, and M3, M4, and M6 were further separated by reversed phase HPLC on a column of Cosmosil 5C4 300 (data not shown). The amino acid compositions and amino-terminal

sequences of the isolated peptides are shown in Table II and Fig. 3. M2-3 and M7-8 were yielded by incomplete cleavage at a Met-Thr bond (residues 90-91) and a Met-Ser bond (residues 169-170), respectively. To overlap the CNBr fragments, the Pe-heavy chain was digested, separately, with lysyl endopeptidase and endoproteinase Asp-N. The digest with lysyl endopeptidase was separated by reversed phase HPLC on a column of Cosmosil 5C4 300 (data not shown). The amino acid compositions and amino-terminal sequences of the isolated peptides are shown in Table III and Fig. 3. The digest with endoproteinase Asp-N was separated by reversed phase HPLC on a column of μ Bondasphere C8 and several fractions containing more than two peptides were further subjected to rechromatography on a column of Chemcosorb 5-ODS-H. The amino acid compositions of these peptides were determined (data not shown) and the amino-terminal sequences of 12 peptides were analyzed (Fig. 3). To complete the sequence of Pe-heavy chain, large fragments, M7-8, M9 and K4 were subdigested with chymotrypsin, lysyl endopeptidase, and trypsin, respectively, and the resulting peptides were isolated by reversed phase HPLC. The amino acid compositions of these peptides were determined (data not shown) and the amino-terminal sequences of overlapping peptides were analyzed (Fig. 3). The heavy chain consisted of 427 amino acid residues with 4 asparagine-linked sugar chains at positions 28, 69, 163 and 183 (Fig. 3).

Sequence Analyses of Light Chain (LC1) — The amino-terminal sequence analysis of Pe-LC1 established the first 38 residues with the exception of several unidentified residues. Pe-LC1 was primarily digested with lysyl endopeptidase and the resulting peptides were separated by reversed phase HPLC (Fig. 5). The amino acid compositions and amino-terminal sequences of the isolated

peptides are shown in Table IV and Fig. 6. Large fragment K5 was further subdigested with V8 protease, and the peptide K5E3 isolated by reversed phase HPLC was used to complete the sequence of K5. To overlap these peptides, Pe-LC1 was cleaved with CNBr. The resulting peptides were separated by reversed phase HPLC on a column of YMC A-302 ODS after gel filtration on a column of G2000SW (data not shown). The amino acid compositions and amino-terminal sequences of the isolated peptides are shown in Table V and Fig. 6. M3-4 was yielded by incomplete cleavage at a Met-Ile bond (residues 101-102) in the same manner as previously observed (16). The carboxyl-terminal amino acid of LC1 was determined to be phenylalanine residue by vapor-phase hydrazinolysis method (0.17 mol/mol of protein, uncorrected). The light chain LC1 consisted of 123 amino acid residues containing an asparagine-linked sugar chain at position 24 (Fig. 6).

Carbohydrate-linked Asparagine Residues ——— The PTH derivatives at position 24 in the light chain LC1 and at positions 28, 69, 163 and 183 in the heavy chain, were not identified by sequence analysis. These are very likely carbohydrate-linked Asn residues, since they have a potential -X-Thr/Ser sequence, which is known as a consensus signal sequence for the attachment of carbohydrate to asparagine. In fact, the composition analyses of the fragments containing these residues indicated the presence of glucosamines in each fragment. The reasonably high yields for Thr and Ser residues in the entire sequence and the composition analyses of all the fragments, including intact protein, indicated no O-linked carbohydrate chains in RVV-X.

Titration of the Free SH-Group ——— The concentration of a free SH-group was determined in the presence of 6M guanidine hydrochloride and 10 mM EDTA by the method of Ellman (26) and also

incorporation of 4-vinylpyridine in the absence and presence of the denaturant containing 6M guanidine hydrochloride and 10 mM EDTA. No free SH-group could be detected in the non-reduced RVV-X by the two methods with Ellman's reagent (5,5'-dithiobis (2-nitrobenzoic acid)) and with S-pyridylethylation.

Platelet Aggregation Inhibitory Activity of RVV-X ——— Since the middle portion of the heavy chain (residues 212-301) showed a remarkable sequence similarity to those of venom platelet aggregation inhibitors, disintegrins (see Fig. 7), we tested the platelet aggregation inhibitory activity of RVV-X. As shown in Fig. 8, RVV-X dose-dependently inhibited collagen- and ADP- induced platelet aggregations. The percentages of inhibition on collagen- and ADP- induced platelet aggregations were 44 % and 85 %, respectively, at a concentration of 310 nM.

Effect of IX/X-bp on the RVV-X-mediated Activation of Factor X ——— We also examined the effect of IX/X-bp on the RVV-X-mediated activation of factor X, since the amino acid sequence of the light chain LC1 is similar to that of IX/X-bp (Fig. 9). As shown in Fig. 10, IX/X-bp inhibited the activation of factor X, in a dose-dependent manner. The 86 % inhibition was observed in the presence of an equal molar ratio of IX/X-bp to factor X.

DISCUSSION

We reported the primary structures of several hemorrhagic and non-hemorrhagic metalloproteinases isolated from the venoms of *Trimeresurus flavoviridis* (14-16) and *Crotalus ruber ruber* (17), the objective being to elucidate the molecular mechanism of hemorrhage, the major symptoms following the bite of crotalid and viperid snakes. These metalloproteinases belong to a new metalloproteinase subfamily, all of which contain a zinc-chelation sequence His-Glu-X-X-His but have no significant sequence homology with any known metalloproteinases beyond this region (17). The most exciting finding in this series of studies is that high molecular mass hemorrhagic metalloproteinase HR1B (Mr 60,000) contains the cysteine-rich carboxyl-terminal half (residues 204-416) which includes the disintegrin-like structure, in addition to the amino-terminal half metalloproteinase domain (residues 1-203) similar to low molecular mass metalloproteinases in sequences (16). Disintegrins isolated from snake venoms are potent platelet aggregation inhibitors and are cysteine-rich and Arg-Gly-Asp (RGD)-containing polypeptides (total 48-84 residues) (see Refs. 27 and 28 for review).

As shown in Fig. 7, the entire amino acid sequence of the RVV-X-heavy chain resembles that of HR1B (53 % sequence identity). The sequence of its amino-terminal half (residues 1-205) is also similar to those of the low molecular mass metalloproteinases HR2a, HT-2, Ht-d (29) and H2-proteinase (40-43 % sequence identity), forming a metalloproteinase domain of RVV-X (Fig. 7A). Like others, this domain of RVV-X contains the His-Glu-X-X-His-sequence (residues 145-149), in which His¹⁴⁵ and His¹⁴⁹, and Glu¹⁴⁶ are believed to be two of the zinc-chelating ligands and one of the catalytic residues, as

of the zinc-chelating ligands and one of the catalytic residues, as discussed elsewhere (17). It is of interest that HR1B, HR2a, HT-2 and Ht-d cause a localized hemorrhage, but H₂-proteinase does not, even though all of them have a similar structure. In this regard, we first focused on structural elements associated with the hemorrhagic activity. Since the purified RVV-X did not express hemorrhagic activity (data not shown), we compared the amino acid sequences of two non-hemorrhagic metalloproteinases RVV-X and H₂-proteinase with those of hemorrhagic metalloproteinases HR1B, HR2a, HT-2, and Ht-d. There are five amino acid residues shared by hemorrhagic metalloproteinases but not by non-hemorrhagic RVV-X and H₂-proteinase (boxed residues in Fig. 7A). These residues might be closely associated with the hemorrhagic activity.

The middle region of the RVV-X-heavy chain (residues 212-301) shows a high percentage of sequence identity to the RGD-sequence containing disintegrins, scored 59 % for trigramin (30), 38 % for echistatin (31), 60 % for bitistatin (32) and 55 % for barbourin (33). Although RVV-X as well as the disintegrin family is able to inhibit platelet aggregation (Fig. 8), the RGD-sequence is absent and is substituted for by Arg-Asp-Glu in the corresponding region (Fig. 7). Similar results have been found in HR1B with a Glu-Ser-Glu substitution (16). The platelet aggregation inhibitory activity of RVV-X apparently resides within the disintegrin-like domain, based on the same notion as that discussed in the case of HR1B (see Ref. 16). However, there is a possibility that platelet aggregation inhibitory activity could be due to some of the proteolytic activity or the lectin like activity, since we have used the whole RVV-X for that experiment (Fig. 4). All of twenty-five disintegrins so far isolated contain the RGD-sequence, which is believed to be an

interaction site for several integrins, with one exception, barbourin containing Lys-Gly-Asp instead (33). The substitution in barbourin is suggested to impart integrin specificity to barbourin, since it has been isolated as a specific antagonist for glycoprotein IIb/IIIa, while other RGD-containing disintegrins interact with various integrins of β_1 and β_3 families, such as the vitronectin receptor(s) and fibronectin receptor. The integrin specificities of RVV-X and HR1B remain to be determined. Equally important in future experiments is to account for the *in vivo* relevance of this platelet aggregation inhibitory effect of RVV-X.

The carboxyl-terminal region of RVV-X-heavy chain (residues 302-427) following the disintegrin domain contains numerous cysteine residues. Computerized homology searches revealed that the cysteine-rich region of HR1B is the sole structure homologous to this region, the function of this region is unclear.

Interestingly, the light chain (LC1) of RVV-X shows a striking similarity with C-type lectin-like structure (Ca^{2+} -dependent carbohydrate recognition domain). Numerous proteins isolated from snakes and mammals were seen to contain this structure (34), as shown in Fig. 9. The sequence similarity between RVV-X-light chain (LC1) and rattlesnake lectin (35), phospholipase A2 inhibitor A subunits (36) and factor IX/factor X-binding protein A chain (37), mouse lymphocyte homing receptor (38), rat Kupffer cell receptor (39), rat mannose-binding protein A (40), human lymphocyte IgE receptor (41), human tetranectin (42) and human pulmonary surfactant associated protein (43) is 27, 13, 33, 17, 22, 10, 22, 15 and 10%, respectively. Factor IX/factor X-binding protein (IX/X-bp) is the most interesting among these proteins having the C-type lectin structure, since IX/X-bp and RVV-X share several key

characteristics, in addition to sequence similarity. IX/X-bp has been purified as an anticoagulant protein and it binds specifically to factors IX and X, but not to other vitamin K-dependent coagulation factors (18). This functional activity of IX/X-bp as well as RVV-X is dependent on Ca^{2+} and the amino-terminal Gla-domains of factors IX and X (13,18). IX/X-bp is able to inhibit the RVV-X-catalyzed factor X activation, in a dose dependent manner (Fig. 10), thereby indicating that IX/X-bp occupies a RVV-X-recognized site on factor X. Taken together, it is suggested that the light chain of RVV-X might play an important role in the binding of RVV-X with factor X, and that it recognizes the Gla-domain of factor X. There are two potential roles of the exogenous Ca^{2+} , which is required for the factor X activating activity but not for the proteolytic activity of RVV-X: First, the C-type lectin-like light chain of RVV-X requires Ca^{2+} for its function, and second, Ca^{2+} induces conformational change of the Gla-domain (44), an event which might be a prerequisite for recognition by RVV-X and/or for concomitant conformational change of factor X susceptible to RVV-X.

Since no free SH-group was detected in RVV-X, all the cysteine residues in RVV-X must be linked with disulfide bridges. Six of seven half-cystines in the RVV-X-light chain (LC1) are conserved in several C-type lectin-like structures in which the disulfide bridge locations have been determined, such as rattlesnake lectin (35), acorn banacle lectin (45), sea urchin echinoidin (46) and human tetranectin (42). Based on these data, three intrachain disulfide bridges of RVV-X-LC1 could be assigned by homology as follows: Cys⁴-Cys¹⁵, Cys³²-Cys¹²¹, and Cys⁹⁸-Cys¹¹³. Therefore, the remaining Cys⁷⁷ might be involved in an interchain disulfide bridge between the heavy and the light chains. Since 34 out of 37 half-cystines in

the RVV-X-heavy chain are located at the same positions as those of HR1B (16), three other cysteine residues Cys²⁷, Cys⁶³ and Cys³⁸⁹ are unique in the RVV-X-heavy chain (Fig. 7). Among them, Cys³⁸⁹ is suggested to be the half-cystine participating to form an interchain disulfide bridge between the heavy and the light chains, since trypsin cleaves the non-reduced heavy chain at several peptide bonds only in the amino-terminal half of the heavy chain, under the conditions described in Materials and Methods, and since the remaining carboxyl-terminal half starting at residue 209 of the heavy chain is still bridged with the light chain (data not shown). In this regard, the amino-terminal metalloproteinase domain composing RVV-X is not linked through disulfide bridges to any other domains located in the carboxyl-terminal portion. Six of eight half-cystines in this domain are conserved in HR2a and H₂-proteinase, in which the disulfide bridge locations have been determined (14,15). Therefore, three intra-chain disulfide bridges of this domain could be assigned by homology as follows: Cys¹²⁰-Cys²⁰⁰, Cys¹⁶⁰-Cys¹⁸⁴, and Cys¹⁶²-Cys¹⁶⁷, and the remaining Cys²⁷ and Cys⁶³ might also form an intrachain disulfide bridge in the metalloproteinase domain. Regarding the disulfide bonding patterns of the disintegrin-like domain in RVV-X, it has not been established. However, all the positions of cysteine residues found in the disintegrin-like domain of RVV-X except for Cys²⁷⁸ are the same as those of kistrin (47), in which the disulfide locations have recently been deduced from 2D NMR studies (48).

In Fig. 11, the gross structure of RVV-X is illustrated with other known members of snake venom metalloproteinase family. The putative precursor protein of trigramin is also a member of this family¹. These findings show the structural and evolutionary

relationship among these proteins, although each protein has diversified functional activity. These mosaic structures are also analogous to those of mammalian blood coagulation serine proteinases (51) and matrix metalloproteinases (50). While a common domain with apparently distinct functions can be defined among these proteinases, additional domains unique to individual proteinases will have specific functions. It is presumed that RVV-X evolutionally acquired substrate specificity as a result of assembly of metalloproteinase domain and C-type lectin-like domain.

FOOTNOTE

¹Neeper and Jacobson recently reported the sequence of cDNA for trigramin, but they did not translate a large nucleotide sequence, which is present upstream from the coding region for 72 residues mature trigramin (49). We translated the nucleotide sequence of this region into the corresponding amino acid residues and found that such a region represents a metalloproteinase structure similar to the venom metalloproteinase family (Fig. 12). The trigramin precursor contains an additional 172 residues between the signal sequence and metalloproteinase domain. This portion might be a propeptide found in inactive pro-forms of mammalian matrix metalloproteinases (50).

REFERENCES

1. Iwanaga, S., and Suzuki, T. (1979) in *Handbook of Experimental Pharmacology* (Lee, C.-Y., ed), Vol. 52, pp. 61-158, Springer-Verlag, New York
2. Morita, T., and Iwanaga, S. (1981) *Methods Enzymol.* **80**, 303-311
3. Kisiel, W., and Canfield, W.M. (1981) *Methods Enzymol.* **80**, 275-285
4. Pirkle, H., and Stocker, K. (1991) *Thromb. Haemostas.* **65**, 444-450
5. Rosing, J., and Tans, G. (1991) *Thromb. Haemostas.* **65**, 627-630
6. Tokunaga, F., Nagasawa, K., Tamura, S., Miyata, T., Iwanaga, S., and Kisiel, W. (1988) *J. Biol. Chem.* **263**, 17471-17481
7. Furie, B.C., and Furie, B. (1976) *Methods Enzymol.* **45**, 191-205
8. Fujikawa, K., Legaz, M.E., and Davie, E.W. (1972) *Biochemistry* **11**, 4892-4899
9. Di Scipio, R.G., Hermodoson, M.A., and Davie, E.W. (1977) *Biochemistry* **16**, 5253-5260
10. Jackson, C.M., Gordon, J.G., and Hanahan, D.J. (1971) *Biochem. Biophys. Acta.* **252**, 255-261
11. Kisiel, W., Hermodoson, M.A., and Davie, E.W. (1976) *Biochemistry* **15**, 4901-4906
12. Amphlett, G.W., Byrne, R., and Castellino, F.J. (1982) *Biochemistry* **21**, 125-132
13. Skogen, W.F., Bushong, D.S., Johnson, A.E., and Cox, A.C. (1983) *Biochem. Biophys. Res. Commun.* **111**, 14-20
14. Miyata, T., Takeya, H., Ozeki, Y., Arakawa, M., Tokunaga, F., Iwanaga, S., and Omori-Satoh, T. (1989) *J. Biochem. (Tokyo)* **105**, 847-853

15. Takeya, H., Arakawa, M., Miyata, T., Iwanaga, S., and Omori-Satoh, T. (1989) *J. Biochem. (Tokyo)* **106**, 151-157
16. Takeya, H., Oda, K., Miyata, T., Omori-Satoh, T., and Iwanaga, S. (1990) *J. Biol. Chem.* **265**, 16068-16073
17. Takeya, H., Onikura, A., Nikai, T., Sugihara, H., and Iwanaga, S. (1990) *J. Biochem. (Tokyo)* **108**, 711-719
18. Atoda, H., and Morita, T. (1989) *J. Biochem. (Tokyo)* **106**, 808-813
19. Laemmli, U.K. (1970) *Nature* **227**, 680-685
20. Spackman, D.H., Stein, W.H., and Moore, S. (1958) *Anal. Chem.* **30**, 1190-1206
21. Penke, B., Ferenczi, R., and Kovacs, K. (1974) *Anal. Biochem.* **60**, 45-50
22. Bidlingmeyer, B.A., Cohen, S.A., and Tarvin, T.L. (1984) *J. Chromatogr.* **336**, 93-104
23. Hewick, R.M., Hunkapiller, M.W., Hood, L.E., and Dreyer, W.J. (1981) *J. Biol. Chem.* **256**, 7990-7997
24. Yamamoto, A., Toda, H., and Sakiyama, F. (1989) *J. Biochem. (Tokyo)* **106**, 552-554
25. Moroi, M., Jung, S.M., Okuma, M., and Shinmyozu, K. (1989) *J. Clin. Invest.* **84**, 1440-1445
26. Ellman, G.L. (1959) *Arch. Biochem. Biophys.* **82**, 70-77
27. 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
28. Teng, C.-M., and Huang, T.-F. (1991) *Thromb. Haemostas.* **65**, 624-626
29. Shannon, J.D., Baramova, E.N., Bjarnason, J.B., and Fox, J.W. (1989) *J. Biol. Chem.* **264**, 11575-11583

30. Hung, T.-F., Holt, J.C., Kirby, E.P., and Niewiarowski, S.
(1989) *Biochemistry* **28**, 661-666
31. Gan, Z.-R., Gould, R.J., Jacobs, J.W., Friedman, P.A., and
Polokoff, M.A. (1988) *J. Biol. Chem.* **263**, 19827-19832
32. Shebuski, R.J., Ramjit, D.R., Bencen, G.H., and Polokoff, M.A.
(1989) *J. Biol. Chem.* **264**, 21550-21556
33. Scarborough, R.M., Rose, J.W., Hsu, M.A., Phillips, D.R., Fried
V.A., Campbell, A.M., Hannizzi, L., and Charo, I.F. (1991) *J.*
Biol. Chem. **266**, 9359-9362
34. Drickamer, K. (1988) *J. Biol. Chem.* **263**, 9557-9560
35. Hirabayashi, J., Kusunoki, T., and Kasai, K. (1991) *J. Biol.*
Chem. **266**, 2320-2326
36. Inoue, S., Kogaki, H., Ikeda, K., Samejima, Y., and Omori-Satoh,
T. (1991) *J. Biol. Chem.* **266**, 1001-1007
37. Atoda, H., Hyuga, M., and Morita, T. (1991) *J. Biol. Chem.* **266**,
14903-14911
38. Lasky, L.A., Singer, M.S., Yednock, T.A., Dowbenko, D., Fennie,
C., Rodriguez, H., Nguyen, T., Stachel, S., and Rosen, S.D.
(1989) *Cell* **56**, 1045-1055
39. Hoyle, G.W., and Hill, R.L. (1988) *J. Biol. Chem.* **263**, 7487-
7492
40. Drickamer, K., Dordal, M.S., and Reynolds, L. (1986) *J. Biol.*
Chem. **261**, 6878-6887
41. Kikutani, H., Inui, S., Sato, R., Barsumian, E.L., Owaki, H.,
Yamasaki, K., Kaisho, T., Uchibayashi, N., Hardy, R.R., Hirano,
T., Tsunasawa, S., Sakiyama, F., Suemura, M., and Kishimoto, T.
(1986) *Cell* **47**, 657-665
42. Fuhlendorff, J., Clemmensen, I., and Magnusson, S. (1987)
Biochemistry **26**, 6757-6764

43. Floros, J., Steinbrink, R., Jacobs, K., Phelps, D., Kriz, R., Recny, M., Sultzman, L., Jones, S., Taeusch, H.W., Frank, H.A., and Fritsch, E.F. (1986) *J. Biol. Chem.* **261**, 9029-9033
44. Higashi, S., Kawabata, S., Nishimura, H., Funasaki, H., Ohyama, S., Miyamoto, S., Funatsu, A., and Iwanaga, S. (1990) *J. Biochem. (Tokyo)* **108**, 654-662
45. Muramoto, K., and Kamiya, H. (1986) *Biochem. Biophys. Acta.* **874**, 285-295
46. Giga, Y., Ikai, A., and Takahashi, K. (1987) *J. Biol. Chem.* **262**, 6197-6203
47. Dennis, M.S., Henzel, W.J., Pitti, R.M., Lipari, M.T., Napier, M.A., Deisher, T.A., Bunting, S., and Lazarus, R.A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2471-2475
48. Adler, M., Lazarus, R.A., Dennis, M.S., and Wagner, G. (1991) *Science* **253**, 445-448
49. Neeper, M.P., and Jacobson, M.A. (1990) *Nucleic Acids Research* **18**, 4255
50. Van Wart, H.E., and Birkedal-Hansen, H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5578-5582
51. Davie, E.W., Fujikawa, K., and Kisiel, W. (1991) *Biochemistry* **30**, 10363-10370

Table I

Amino acid compositions of RVV-X-heavy (HC),
and -light chains (LC1 and LC2^a)

Amino acid	HC	Sequence	LC1	Sequence	LC2
——— Residues/molecule ———					
Asp	62.8	63	12.0	12	11.4
Thr ^b	20.5	21	5.2	5	3.9
Ser ^b	30.7	29	9.4	9	11.5
Glu	40.1	40	15.8	16	16.5
Pro	20.1	20	3.6	3	4.3
Gly	26.9	27	6.6	6	5.4
Ala	23.1	22	5.4	5	4.7
Val ^c	21.4	22	7.6	7	7.5
Met	10.7	11	5.2	4	4.4
Ile ^c	20.9	22	5.2	5	4.2
Leu	26.5	25	10.0	10	8.6
Tyr	14.1	14	5.0	5	5.2
Phe	14.9	14	5.0	5	8.0
Lys	25.6	26	11.1	11	9.3
His	10.7	10	3.7	4	3.5
Trp ^d	3.4	3	3.5	6	2.6
Pec ^e	34.2	37	5.5	7	5.2
Arg	20.1	21	3.3	3	6.7
GlcNH ₂ ^f	+		+		+
Total	426.7	427	123.1	123	122.9

^aCalculated from average values obtained from 24, 48 and 72 h hydrolysates.

^bExtrapolation to 0 h.

^cTaken from 72 h hydrolysis.

^dTaken from 24 h hydrolysis with 3 M mercaptoethanesulfonic acid.

^eS-pyridylethylcysteine.

^fGlucosamine.

Table II

Amino acid compositions of cyanogen bromide peptides derived from RVV-X-heavy chain^a

Amino acid	M2-3	M3	M4	M5	M6	M7-8	M9	M12
Residues/molecule								
Asp	14.1 (15)	3.2 (3)	2.3 (2)	2.0 (2)	1.1 (1)	22.5 (23)	10.0 (11)	6.7 (6)
Glu	6.2 (5)			5.0 (5)	1.1 (1)	18.3 (18)	5.1 (5)	2.8 (3)
Ser	6.5 (7)	0.9 (1)		1.1 (1)	0.9 (1)	9.1 (10)	4.7 (5)	0.9 (1)
Gly	3.1 (2)		1.8 (2)	1.3 (1)	1.0 (1)	12.9 (12)	5.5 (5)	2.4 (3)
His	0.9 (2)	0.9 (1)			1.9 (2)	1.6 (3)		
Arg	2.6 (3)	1.0 (1)	1.0 (1)	1.7 (2)		8.0 (8)	4.9 (6)	0.8 (1)
Thr	7.1 (8)	2.0 (2)	2.0 (2)	1.0 (1)		4.9 (5)		1.8 (4)
Ala	5.9 (6)	0.9 (1)	0.9 (1)	1.9 (2)	1.0 (1)	7.5 (7)	2.9 (3)	0.9 (1)
Pro	2.3 (1)					16.1 (14)	3.8 (4)	1.0 (1)
Tyr	1.4 (1)			1.0 (1)		7.0 (7)	3.0 (3)	1.3 (1)
Val	5.3 (4)			2.9 (3)		6.9 (6)	2.8 (2)	3.9 (4)
Met ^b	1.1 (2)	0.6 (1)	0.7 (1)	0.4 (1)	0.7 (1)	1.7 (2)	0.9 (1)	
Pec ^c	1.3 (2)			0.6 (1)		17.3 (23)	4.8 (7)	2.9 (3)
Ile	6.1 (6)		0.8 (1)	1.5 (2)		6.6 (7)	2.1 (2)	0.4 (0)
Leu	7.7 (7)	1.6 (2)	2.7 (3)		2.0 (2)	8.1 (7)	4.1 (4)	
Phe	4.0 (4)	0.9 (1)	1.9 (2)	1.0 (1)		2.4 (2)	2.8 (3)	
Trp	n.d. ^d (2)					n.d. ^e (1)		
Lys	3.2 (4)	0.9 (1)		0.7 (1)		9.5 (10)	5.0 (6)	3.1 (3)
GlcNH ₂ ^e +						+		
Total	81	14	15	24	10	165	67	31
Position	24-104	91-104	105-119	120-143	144-153	154-318	319-385	397-427

^aValues in parentheses are taken from the sequence data.^bDetermined as homoserine.^cS-Pyridylethylcysteine.^dNot determined.^eGlucosamine.

Table III

Amino acid compositions of lysyl endopeptidase peptides derived from RVV-X-heavy chain^a

Amino acid	K1	K4	K5	K6	K7	K8	K10	K11	K12	K14	K15	K20	K21	K25
Residues/molecule														
Asp	1.5 (1)	9.0 (10)	8.2 (7)	2.4 (2)	4.0 (4)	3.2 (2)	1.2 (1)	7.0 (6)	2.1 (1)	3.6 (4)	7.5 (6)	2.8 (3)	2.7 (3)	2.3 (2)
Glu	1.1 (1)	4.8 (5)	5.7 (5)	1.6 (1)	1.2 (1)	1.9 (1)		6.0 (6)	2.2 (2)	5.7 (7)	2.6 (2)		1.3 (1)	1.5 (2)
Ser	1.6 (2)	5.1 (5)	2.2 (2)	1.1 (1)	3.4 (5)	2.0 (2)		1.6 (2)	1.3 (0)	1.2 (1)	2.0 (2)	0.8 (1)	1.1 (1)	0.8 (1)
Gly		3.9 (2)	3.3 (3)	1.8 (2)	0.4 (0)	0.7 (0)		2.8 (3)	3.3 (3)	2.7 (3)	2.7 (3)	1.0 (1)	0.4 (0)	1.3 (1)
His		1.2 (1)	0.9 (1)	2.1 (3)		0.9 (1)				0.5 (1)			0.8 (2)	
Arg		2.9 (3)	3.2 (3)	0.7 (0)	0.4 (0)	1.8 (2)	0.8 (1)	0.9 (0)		2.3 (4)	3.0 (3)	1.6 (2)	0.6 (0)	1.0 (1)
Thr	0.9 (1)	4.5 (4)	3.4 (3)	1.0 (1)		1.2 (1)		1.3 (1)		2.2 (3)	0.5 (0)			3.1 (3)
Ala	0.7 (1)	4.3 (3)	3.2 (3)	1.7 (2)		0.5 (0)		2.7 (3)	1.0 (1)	2.1 (3)	1.8 (2)			1.3 (1)
Pro	0.5 (0)	1.7 (1)	0.8 (0)		1.8 (2)		2.0 (2)	3.7 (4)	1.4 (1)	1.9 (2)	2.3 (2)	1.3 (1)	1.2 (1)	
Tyr		1.2 (1)	1.1 (1)	1.0 (1)			2.7 (3)	0.6 (0)	1.0 (1)		3.0 (2)		1.0 (1)	1.0 (1)
Val	0.7 (1)	2.5 (4)	1.6 (2)	0.7 (1)	1.4 (2)			1.6 (2)	0.4 (0)	1.4 (2)	0.9 (1)			2.0 (2)
Met		0.8 (1)	1.1 (2)	0.2 (2)	0.4 (1)						0.5 (1)		0.4 (2)	
Pec ^b	0.4 (0)	0.7 (1)	0.9 (1)		3.2 (3)	1.9 (1)	0.9 (1)	7.5 (7)	2.6 (4)	2.9 (4)	4.7 (4)	1.4 (2)	1.9 (2)	0.7 (1)
Ile		3.3 (6)	1.7 (2)	0.6 (1)	0.8 (1)	1.1 (1)	0.9 (1)	1.4 (2)			2.0 (2)			
Leu	1.0 (1)	5.5 (5)	5.9 (5)	2.3 (2)	1.3 (1)	2.5 (2)	1.1 (1)	0.5 (0)	2.0 (2)	1.7 (1)	1.8 (1)	2.0 (2)	0.4 (0)	
Phe	0.8 (1)	2.5 (3)	4.0 (4)			1.0 (1)	1.0 (1)				1.3 (1)	1.1 (1)		
Trp		n.d. ^c (2)						n.d. ^d (1)						
Lys	1.0 (1)	1.6 (1)	1.2 (1)	1.0 (1)	1.0 (1)	1.1 (1)	1.0 (1)	1.2 (1)	1.9 (2)	1.1 (1)	1.1 (1)	1.0 (1)	1.0 (1)	
GlcNH ₂ ^d		+			+	+								
Total	10	58	45	20	21	18	9	38	17	36	33	14	14	15
Position	1-10	36-93	94-138	139-158	159-179	180-197	200-208	209-246	247-263	266-301	302-334	367-380	381-394	413-427

^aValues in parentheses are taken from the sequence data.^bS-Pyridylethylcysteine.^cNot determined.^dGlucosamine.

Table IV

Amino acid compositions of lysyl endopeptidase peptides derived from RVV-X-light chain (LC1)^a

Amino acid	K1	K2	K3	K4	K5'	K5	K6	K7	K8	K10'	K10
	Residues/molecule										
Asp	1.2 (1)	1.9 (2)	1.6 (2)		3.1 (3)	2.4 (3)	2.6 (3)			0.8 (1)	1.2 (1)
Glu	2.1 (2)		1.2 (1)	2.2 (2)	6.9 (6)	5.6 (6)	1.0 (1)		3.1 (3)	0.6 (0)	
Ser	1.6 (2)				6.1 (4)	4.7 (4)	0.8 (1)		0.9 (1)	1.2 (1)	0.9 (1)
Gly	0.9 (1)	1.0 (1)			3.4 (3)	2.6 (3)	1.1 (1)			0.4 (0)	
His	0.9 (1)				2.2 (2)	1.8 (2)			0.9 (1)		
Arg					1.7 (1)	1.1 (1)	2.1 (2)				
Thr			1.0 (1)	0.9 (1)	1.1 (1)	0.9 (1)			1.0 (1)	1.3 (1)	1.0 (1)
Ala			0.9 (1)		1.1 (1)	0.9 (1)			1.8 (2)	1.2 (1)	0.9 (1)
Pro	1.0 (1)				1.5 (1)	1.1 (1)				1.1 (1)	1.0 (1)
Tyr	2.0 (2)				1.4 (1)	1.1 (1)		1.0 (1)	1.0 (1)	0.5 (0)	
Val	1.0 (1)				2.4 (3)	1.9 (3)	0.9 (1)			1.7 (2)	1.5 (2)
Met					0.7 (1)	0.8 (1)	0.7 (1)		0.6 (1)	0.7 (1)	0.7 (1)
Pec ^b	2.2 (2)			0.7 (1)			0.6 (1)		1.0 (1)	0.9 (2)	2.2 (2)
Ile					2.0 (2)	1.6 (2)			1.9 (2)	1.2 (1)	0.9 (1)
Leu	1.9 (2)	1.0 (1)			5.0 (5)	4.1 (5)			2.0 (2)	0.6 (0)	
Phe		1.0 (1)		1.0 (1)	1.2 (1)	0.9 (1)				2.1 (2)	1.0 (1)
Trp	n.d. ^c (1)		n.d. (1)		n.d.(2)	n.d.(2)	n.d.(1)				
Lys	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	2.0 (2)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)
GlcNH ₂ ^d			+								
Total	17	6	7	6	39	38	13	2	16	14	13
Position	1-17	18-23	24-30	31-36	37-75	38-75	76-88	89-90	91-106	110-123	110-122

^aValues in parentheses are taken from the sequence data.^bS-Pyridylethylcysteine.^cNot determined.^dGlucosamine.

Table V

Amino acid compositions of cyanogen bromide peptides derived from RVV-X-light chain (LC1)^a

Amino acid	M1	M2	M3	M4	M3-4
Residues/molecule					
Asp	8.0 (8)	1.1 (1)	2.0 (2)	1.2 (1)	2.7 (3)
Glu	10.4 (11)		3.0 (3)	2.0 (2)	5.4 (5)
Ser	4.4 (6)		1.7 (2)	1.3 (1)	3.5 (3)
Gly	4.8 (5)		1.2 (1)	0.8 (0)	3.9 (1)
His	2.8 (3)		0.4 (0)	1.0 (1)	1.0 (1)
Arg	1.1 (1)	1.0 (1)	0.9 (1)		1.5 (1)
Thr	2.9 (3)			1.9 (2)	1.7 (2)
Ala	1.9 (2)		1.8 (2)	1.0 (1)	3.3 (3)
Pro	2.0 (2)			1.0 (1)	1.0 (1)
Tyr	2.9 (3)		2.0 (2)		1.8 (2)
Val	3.2 (4)		1.0 (1)	1.6 (2)	2.4 (3)
Met ^b	0.6 (1)	0.9 (1)	0.8 (1)	0.6 (1)	0.2 (2)
Pec ^c	2.9 (3)	0.7 (1)	0.7 (1)	1.3 (2)	1.3 (3)
Ile	2.0 (2)		1.0 (1)	1.7 (2)	2.5 (3)
Leu	8.1 (8)		2.2 (2)		2.7 (2)
Phe	3.0 (3)			1.9 (2)	1.9 (2)
Trp	n.d. ^d (3)	n.d. (1)	n.d. (1)	n.d. (1)	n.d. (2)
Lys	5.3 (5)	1.2 (1)	1.9 (2)	2.4 (3)	3.5 (5)
GlcNH ₂ ^e +					
Total	73	6	22	22	44
Position	1-73	74-79	80-101	102-123	80-123

^aValues in parentheses are taken from the sequence data.

^bDetermined as homoserine.

^cS-Pyridylethylcysteine.

^dNot determined.

^eGlucosamine.

FIG. 1. **Separation of RVV-X-heavy (HC) and light chains (LC).** *S*-pyridylethylated RVV-X was subjected to gel filtration HPLC on G3000SW equilibrated with 0.1 M sodium phosphate buffer, pH 6.0, containing 6 M guanidine hydrochloride and 1 mM EDTA. The inset is 12.5 % SDS-PAGE analysis of the isolated HC and LC. The gels were stained with Coomassie Brilliant Blue R-250.

FIG. 2. **Separation of RVV-X-light chains, LC1 and LC2, by reversed phase HPLC.** The fraction LC shown in Fig. 1 was applied to a column of Vydac 214TP5415. The proteins were eluted at room temperature with a linear gradient of acetonitrile containing 0.1 % TFA at a flow rate of 0.5 ml per min. The inset is 12.5 % SDS-PAGE analysis of the isolated LC1 and LC2. The gels were stained with Coomassie Brilliant Blue R-250.

FIG. 3. **The complete amino acid sequence of RVV-X-heavy chain.** Amino acid residues are given in single letter code. Residues determined by Edman degradation are given below the summarized sequence. Dashes indicate unidentified residues. *N*-linked sugar chains are shown by ●. *Pe*-heavy chain, *S*-pyridylethylated heavy chain; *M*, CNBr-cleaved peptides; *K*, lysyl endopeptidase-digested peptides; *D*, endoproteinase Asp-N-digested peptides; *T*, tryptic peptides; *C*, chymotryptic peptides.

FIG. 4. **Separation of CNBr fragments derived from *Pe*-heavy chain on a TSK G2000SW column.** The same conditions as described for Fig. 1 were used. Fragments M5 and M12, and M3, M4, and M6 were

separated by reversed phase HPLC on a column of Cosmosil 5C4-300 (data not shown).

FIG. 5. Separation of lysyl endopeptidase-digested peptides derived from Pe-LC1 by reversed-phase HPLC. The digest was applied to a column of μ Bondasphere C8. A flow rate is 0.2 ml per min. Other conditions are the same as those described for FIG. 2.

FIG. 6. The complete amino acid sequence of RVV-X-light chain (LC1). Amino acid residues are given in single letter code. Residues determined by Edman degradation are given below the summarized sequence. Dashes indicate unidentified residues. N-linked sugar chain is shown by ●. Pe-LC1, S-pyridylethylated LC1; K, lysyl endopeptidase-digested peptides; M, CNBr-cleaved peptides; E, V8 protease-digested peptides.

FIG. 7. Structural comparisons of RVV-X-heavy chain with other proteins. The entire amino acid sequence of RVV-X-heavy chain (427 residues) is similar to that of HR1B (416 residues). HR1B is a high molecular mass hemorrhagic metalloproteinase isolated from the venom of *Trimeresurus flavoviridis* (16). A, Sequence comparison of the amino-terminal regions of RVV-X-heavy chain and HR1B with HR2a, HT-2, hemorrhagic toxin d (Ht-d) and H₂-proteinase. HR2a (14), HT-2 (17), Ht-d (29) and H₂-proteinase (15) are low molecular mass metalloproteinases isolated from the venoms of *T. flavoviridis*, *C. ruber ruber*, *Crotalus atrox*, and *T. flavoviridis*, respectively. The residues shared by hemorrhagic metalloproteinases HR1B, HR2a, HT-2, and Ht-d and not shared by non-hemorrhagic H₂-proteinases and RVV-X

are boxed. Residues conserved in all the proteins are shown at the bottom. The putative zinc ligands and active site are indicated by ▲ and Δ, respectively. B, Comparison and alignment of the carboxyl-terminal regions of RVV-X-heavy chain and HR1B with several disintegrins. Trigramin (30), echistatin (31), bitistatin (32) and barbourin (33) are all viper venom platelet aggregation inhibitors, called disintegrins, isolated from the venoms of *Trimeresurus gramineus*, *Echis carinatus*, *Bitis arietance*, and *Sistrurus m. barbouri*, respectively. Residues conserved in all the proteins are shown at the bottom. Arg-Gly-Asp sequences are boxed.

FIG. 8. **Inhibition of platelet aggregation by RVV-X.** PRP was incubated with Tris-HCl buffer saline, pH 7.5 or various concentrations of RVV-X at 37 °C for 3 min before addition of collagen or ADP. Other details are described under "Materials and Methods".

FIG. 9. **Alignment of the amino acid sequence of RVV-X-light chain (LC1) and those of C-type lectin structures.** Proteins on upper 4 lanes are originated from snakes, and proteins on lower 6 lanes are originated from mammals. RSL, rattlesnake lectin from the venom of *Crotalus atrox* (35); PLIA, phospholipase A2 inhibitor A from blood plasma of *Trimeresurus flavoviridis* (36); IX/XBPA, factor IX/factor X-binding protein A chain from the venom of *T. flavoviridis* (37); MLHR, mouse lymphocyte homing receptor (38); RKCR, rat Kupffer cell receptor (39); RMBPA, rat mannose-binding protein A (40); HFCER, human lymphocyte IgE receptor (41); HTET, human tetranectin (42); HPSAP, human pulmonary surfactant associated protein (43). Residues

conserved in more than seven proteins are boxed. Dashes are inserted to maximize similarity.

FIG. 10. Inhibition of the RVV-X-catalyzed factor X activation by IX/X-bp. Factor X (3.6 μ M) was activated by RVV-X (3.6 nM) in the presence of various concentrations of IX/X-bp (0 -7.2 μ M). The factor Xa activity generated was measured using N ^{α} -benzyloxycarbonyl-L-pyroglutamyl-Gly-Arg-4-methylcoumaryl-7-amide as a substrate. The activity in the absence of IX/X-bp was set as 100 %.

FIG. 11. Gross structures of snake venom metalloproteinases. a, Low molecular mass metalloproteinase including HR2a, HT-2, Ht-d and H2-proteinase (14, 17, 29 and 15). b, Precursor protein of trigramin deduced from its cDNA sequence (49). c, High molecular mass metalloproteinase HR1B (16). d, RVV-X. Identified and potential N-linked sugar chains are shown by ● and ○. The locations of conserved HEXXH sequence are indicated.

FIG. 12. The amino acid sequence deduced from the previously reported cDNA sequence encoding trigramin. Nucliotide sequence (upper) is taken from Ref. 49 and amino acid sequence is deduced (lower). Putative signal sequence is double-underlined. The potential attachment sites for N-linked sugar chains are indicated by ●. The region that are homologous to the metalloproteinase domain is indicated. The zinc-chelating HEXXH sequence is boxed. The amino acid sequence of mature trigramin is underlined.

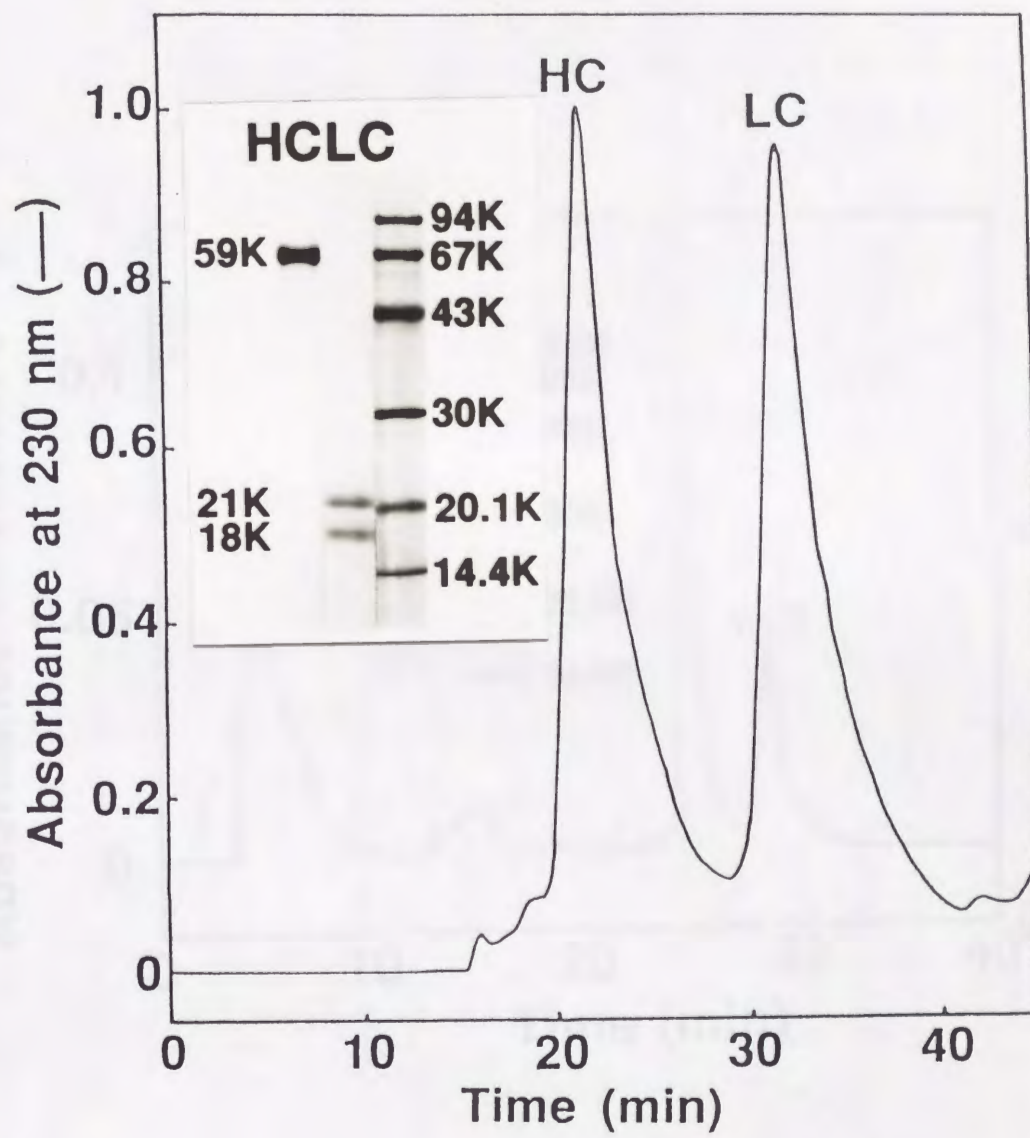


Fig. 1

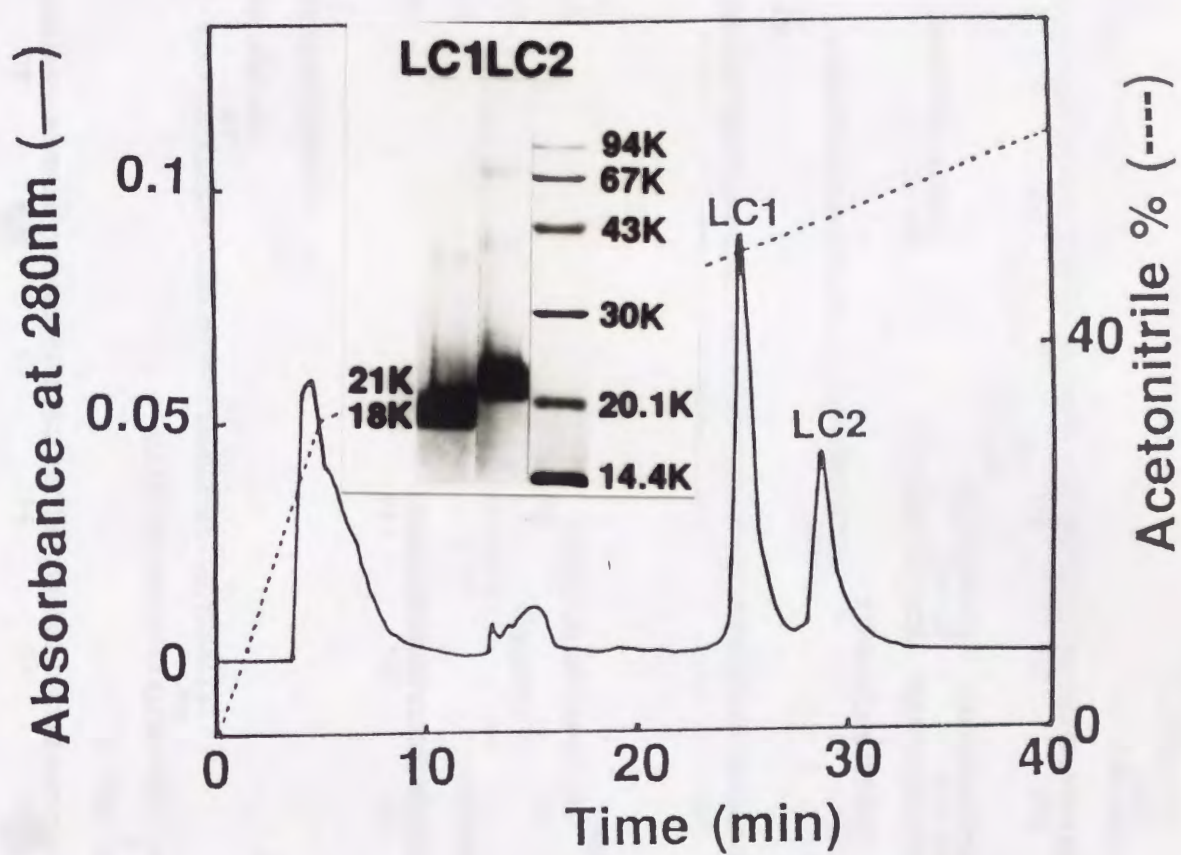


Fig. 2

10 20 30 40 50 60 70 80 90 100 110
 LVSTSAQFNKIFIELVIVDITMALKCNSTATNTKIYEIVNSANEIFNPLNIHVTLIGVEFWCDRLINVTSSADETLNSFGGEWRASDLMTKSHDNALLFTDMRFDLNT
 Pe-heavy chain
 LVSTSAQFNKIFIELVIVDITMALKC-STA-->
 M2-3
 ALKC-STATNTKIYEIVNSANEIFNPLNIHV-->
 K1
 LVSTSAQFNK
 K2
 IFIELVIVDI-MALK
 K4
 IYEIVNSANEIFNPLNIHVTLIGVEFWCDRLI-VTSSADETLNSFG-->
 D5
 DETLNSFGGEWRAS
 K4T2
 DLI-VTSSADETLNSFGGEWR
 K4T3
 ASDLMTR
 120 130 140 150 160 170 180 190 200 210 220
 LGITFLAGMCQAYRSVEIVQEQQNRNFKTAVIMAHELSHNLGMYHDGKNCICNDSSCVMSPVLSQPSKLFSCSIHDYQRYLTRYKPKCIFNPPLRKDIVSPPVCGNEI
 M5
 CQAYRSVEIVQEQQNRNFKTAV--
 LGITFLAGm
 M7-8
 SPVLSQPSKLFSCSIHDYQRYLTRYKPKC-->
 K11
 DIVSPPVCGNEI
 M6
 --ELSHNLG-
 K7
 Y-DGKNCIC-DSSCV-SPVL-->
 K8
 LFS-CSIHQRYLTRYK
 CIFNPPLRK
 K6
 TAVIMAHELSHNLGMYHDGK
 D13
 DYQRYLTRYKPKCIFNPPLRK
 (K5)
 LGITFLAGMCQAYRSVEIVQEQQNRNFK
 (D9)
 LGITFLAGMCQAYRSVEIVQ-->
 M7-8C1
 KPKCIF
 D14
 DIVSPPVCGNEI
 M7-8C2
 NPPLRKDIVSPPVCGNEI
 230 240 250 260 270 280 290 300 310 320 330
 WEEGEECDGSPANCQNPCCDAATCKLKPGAECGNGLCYQCKIKTAGTVCRARDECDVPEHCTGQSAECPRDQLQNGKPCQNNRGYCYNGDCPIMRNQCISLFGSRA
 M9
 RNQCISLFGSRA
 (K11)
 WEEGEECDGSPANCQNPCCDAA---
 D15
 DCGSPANCQNPCC
 K12
 LKPGAECGNGLCYQCK
 K14
 TAGTVCRARDECDVPEHCTGQSAECPRDQLQNGK
 K15
 PCQNNRGYCYNGDCPIMRNQCISLFGSRA
 (D14)
 WEEGEEC
 D16
 DAATCKLKPGAECGNGLC-YQCKIKTAGTVCR-AR
 M7-8C7
 SAECPRDQLQNGKPC
 (M7-8C2)
 WEEGEECDGSP-->
 M7-8C4
 M7-8C5
 M7-8C3
 KLKPGAECGNGLCYQCKIKTAGTVCR-AR
 NCQNPCCDAATC
 M7-8C6
 -RARDECDVPEHCTGQ
 340 350 360 370 380 390 400 410 420
 NVAKDSCFQENLKGSYGYCRKENGRIKIPCAPQDVKCGRLFCNNSPRNKPNPCNMHYSMDQHKGMVDPGTCEDGKVCNNKRQCVDVNTAYQSTTG
 (M9)
 NVAKDSCFQENLKGSYGYCRKEN-RKIPCAPQ-->
 (K15)
 NVAK
 M9K5
 IPCAPQDVK
 K21
 NPCNMHYSMDQHK
 K25
 RQCVDVNTAYQSTTG
 D21
 DSCFQENLKGSYGYCRKENGRIKIPCAP-
 K20
 CGRLFCNNSPRNK
 D23-24
 DQHKGMVDPGTC-
 D26
 DVNTAYQSTTG
 (D20)
 NVAK
 (D20,)
 NVAK
 D22
 DVKCGRLFCNNSPRNKPN-NMHYSMD
 D25
 DGKVCNNKRQCV

Fig. 3

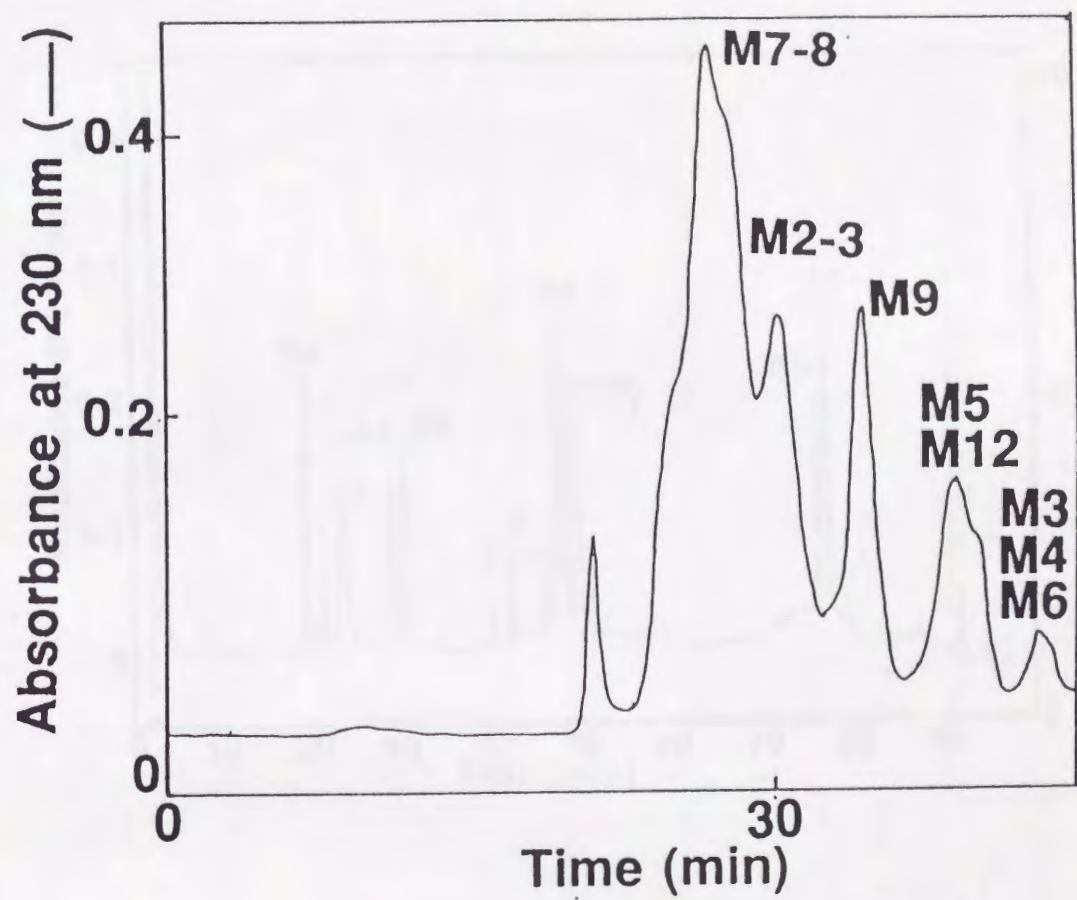


Fig. 4

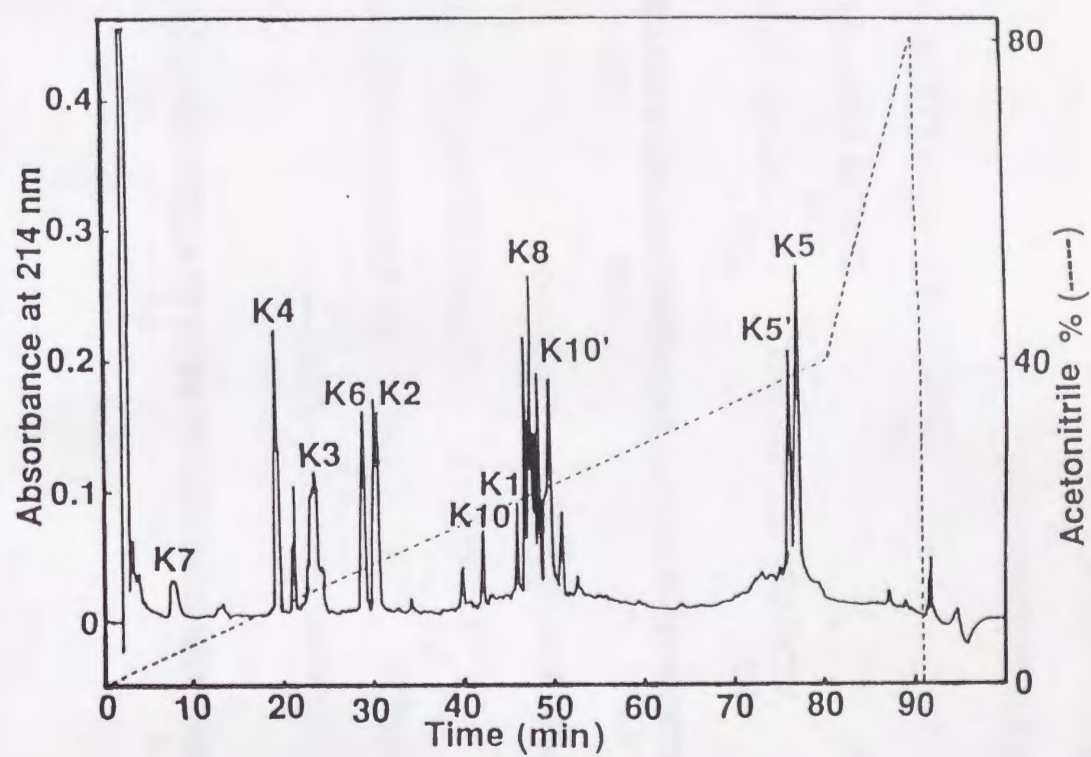


Fig. 5

10 20 30 40 50 60 70
 VLDCPSGWLSYEQHCYKGFNDLK-NWTDAEKFCTEQKKGSHLVSLHSREEEEFVVNLISENLEYPATWIGL

Pe-LC1

VLDCPSGWLSYEQHCYKGFNDLK-WTDAEKFCT-QKKG-->

K2 K4 K5
 GFNDLK FCTEQK GSHLVSLHSREEEEFVVNLISENLEYPAT-IGL
 K3 K5'
 -WTDAEK KGSHLVSLHSREE--FVVNLIS-NLE-PA-->

M1

VLDCPSGWLSYEQHCYKGFNDLK--TDAEKFC--QKKG-->

K5E3
 NLEYPATWIGL

80 90 100 110 120
 GNMWKDCRMEWSDRGNVKYKALAEESYCLIMITHEKEWKSMTCNFIAPVVCKF

(K5)

GNM--

K8

ALAEESYCLIMITHEK

K10'

SMTCNFIAPVVCKF

K6

DCRMEWSDRGNVK

K10

SMTCNFIAPVVCK

M2

WKDCR-

M4

ITHEKEWKS-TCNFIAPVVCK-

M3

EWSD-GNVKYKALAEESYCL--

M3-4

EWSDRGNVKYKALAEESY-LIMIT-->

(K5E3)

GNMWK

Fig. 6

[illegible]

RVVXH : TSSADETLNSFGWWRASDLMTKRSKSHDNALLFTDMRFDLNTLGITFLAGMCQAYRSVEIVQEQGNRNFKTA
 HR1B : QSASNVTLDLFGDWRESVLLKQRSHDCALLTTDFDGPITIGKAYTASMCDFPKRSVGIVQDYSPINLVVA
 HR2a : QAVAPTTARLFGDWRETVLLKQKDHHDHACLLTDINFTGNTIGWAYMGGMCNAKNSVGIVKDHSSNVFMVA
 HT-2 : CSSAKNTLHSFGWWRKSVLLNRKDKHDHNAALLTAIVLDDYTLGLAYLNSMCHPRNSVGLIQDHSPINLLMG
 Ht-d : QSASSDTLNAFAEWRETDLLNRKSHDNACLLTALDELDEETLGLAPLGTMCDPKLSIGIVQDHSPINLLMG
 H2 : KSASNVTLSEFGNWRETVLLKQNNDCAHLLTATNLNDNTIGLAYKKGMCPKLSVGLVQDYSPNVFMVA
 Consensus : -----T-----F-----WR-----L-----D-A-L-T-----T-G-----MC-----S-----

	150	160	170	180	190	200
RVVXH	:	VIMAHELSHNLGMYHDGKN-CICNDSSCVMSPVLSDQPSKLF	SNC	SIHDYQRYLTRYKPKCIFNPPLRKD		
HR1B	:	VIMTHEMGHNLGIPHDG-NSCTCGGFPCIMSPMISDPPSE	LF	SNC	KAIYQTF	FLTDHKPQCILNAPSKTD
HR2a	:	VTMTHEIGHNLGMEHDDKDKCKCEA--CIMSAVISDKPSK	LF	SDCSKDY	YQTF	FLTNSKPPQCIINAP
HT-2	:	VTMAHELGHNLGMEHDGKD-CLRGASLCIMRPGLTGPRSYE	FSD	ASMRYYQK	FLDQYK	PQCILNKP
Ht-d	:	VTMAHELGHNLGMEHDGKD-CLRGASLCIMRPGLTGPRSYE	FSD	SMHYYERF	FLQYK	PQCILNKP
H2	:	VTMTHLGHNLGMEHDDKDKCKCEA--CIMS	DISDKPSKLF	SDCSKNDYQTF	FLTKYNP	PQCILNAP
Consensus	:	V-M-HE--HNLG--HD----	C-M-----S--FS--S--Y--L----	P-CI-N-P----		

	210	220	230	240	250	260	270
RVVXH	:	IVSP	PVCGNEI	WEEGEE	CDGSPAN	CNPCCDAATCKLK	PGAECGNGLCCYQCKIKTAGTVCRRARDEC
HRlB	:	IVSP	PVCGNELLE	EAGEE	CDGSPENCQYQCCDAASCKLH	SWVKCESGECCDQCRFRTAGTECRAAESEC	
Trigramin	:		EAGED	CDGSPAN--P-CCDAATCKL	IPGAQC	GGLCCDQCSFIEEGTVCRI	ARGD
Echistatin	:					ECESGPCCRNCKFLKEGTICKR	ARGD
Bitistatin	:	SPPV	CGNKILEQ	GEDCDGSPANCQDQC	CNAATCKLTPGSQCNHGECCDQCKFKKARTVCRI	ARGD	
Barbourin	:		EAGEE	CDGSPEN--P-CCDAATCKLR	PGAQCADGLCCDQCRFMKKGTVC	RVAKGD	
Consensus	:	IVSP	PVCGN---	E-GE-CDGSP-N----	CC-AA-CKL-----	C--G-CC--C-----	T-C--A-----

	280	290	300	310	320	330	340
RVVXH	:	DVPEHCTGQSAECPDQLQONGKPCQNNRGYCYNGDCPIMRNQCISLFGSRANVAKDSCFQENLKGSYY					
HR1B	:	DIPESCTGQSADCPTDRFHRNGQPCLYNHGYCYNGKCPIMFYQCYFLFGSNATVAEDDCFNNNKKGDKY					
Trigramin	:	DLDDYCNGRSAGCPRNPFH					
Echistatin	:	DMDDYCNGKTCDCPRNPHKGPAT					
Bitistatin	:	WNDDYCTGKSSNCPWNH					
Barbourin	:	WNDDTCTGQSADCPRNGLYG					
Consensus	:	-----C-G-----CP-----PC--N-GYCYNG-CPIM--OC--LFGS-A-VA-D-CF--N-KG--Y					

		350	360	370	380	390	400	410																																																														
RVVXH	:	G	Y	C	R	K	E	N	G	R	K	I	P	C	A	P	Q	D	V	K	C	G	R	L	F	C	L	N	N	S	P	R	N	K	N	P	C	N	M	H	Y	S	C	M	D	Q	H	K	G	M	V	D	P	G	T	K	C	E	D	G	K	V	C	N	N	K	R	Q	C	V
HR1B	:	F	Y	C	R	K	E	N	E	K	Y	I	P	C	A	Q	E	D	V	K	C	G	R	L	F	C	D	N	K	K	-	-	-	-	Y	P	C	H	Y	N	Y	S	-	E	D	L	D	F	G	M	V	D	H	G	T	K	C	A	D	G	K	V	C	S	N	-	R	Q	C	V
Consensus	:	-	Y	C	R	K	E	N	-	-	-	I	P	C	A	-	-	D	V	K	C	G	R	L	F	C	-	N	-	-	-	-	-	-	P	C	-	-	-	Y	S	-	D	-	-	G	M	V	D	-	G	T	K	C	-	D	G	K	V	C	-	N	-	R	Q	C	V			

	420
RVVXH	: DVNTAYQSTTG
HR1B	: DVNEAYKS
Consensus	: DVN-AY-S---

110

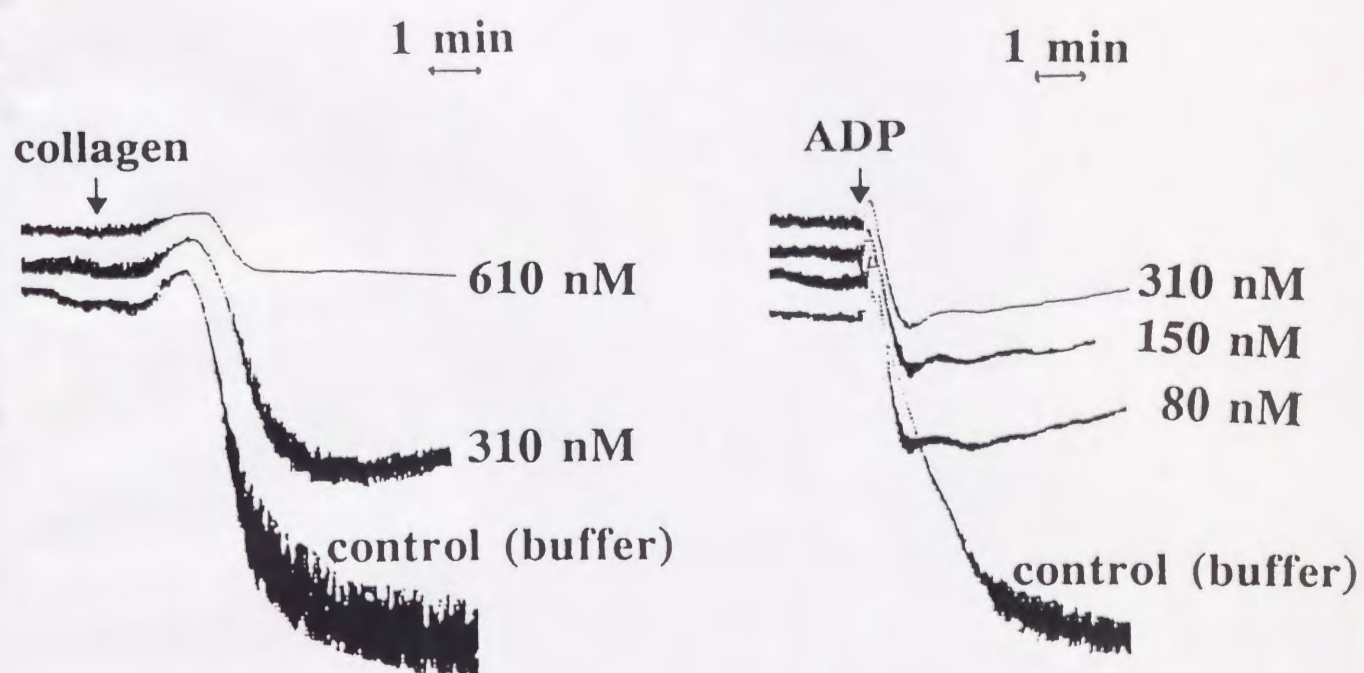


Fig. 8

	10	20	30	40	50	60	70
RVVXLC1 (1-123)	:VLDCPSGWLSYEQHCYKGFNDIKNWTDAEKEFCTEQKKGSH---	LVSLHSRE	EEEFVVNLISENLEYPAT--	WIGL			
RSL (1-133)	:NNCPDPLWLMNGLCYKIFNQIKTWEDAEMFCRKYKPGCH---	LASFHRYG	ESLEIAEYISDYHKGQENV	WIGL			
PLIA (33-143)	:TVNNARSFGSGSERLYVSNKEIKTFEPLKEICEEA-GGHIPSPQLENQNKAF-ASV-LERHNKAAYL-----	VVGD					
IX/XBPA (1-129)	:DCLSGWSSYEGHCYKAFEKYKTWEDAERVCTEQAKGAH---	LVSIESSG	EADFVAQLVTQNMKRLDFYI	WIGL			
MLHR (1-119)	:	WTYHYSEKPMNWNARKFKCKEN-YTD----	LVAIENKR	EIEYLENTLPKSPYYY---	WIGL		
RKCR (409-508)	:LQLIMQDWKYFNGKFYKFSRDKKSWHEAENFCVSQ--GAH---	LASVTSQE	EQAFVLVQITNAVDH-----	WIGL			
RMBPA (97-219)	:KLHAFSMGKKSGKKFFVTNHERMPFSKVKALCSE-LRGTVAIPKNAEENKAIQE---	VAKTSAFLGITDEV--	TEGQ				
HFCER (160-284)	:CNTCPEKWINFQRKCYFYGKGTQWVHARYACDD-MEQQ-----	LVSISHSPE	EQDFLTKHASHTGS-----	WIGL			
HTET (47-178)	:QTVCLKGTKVHMK-CFLAFTQTKTFHEASEDCIS-RGGT----	LGTPQTGS	ENDALYEYLRQSVGNEAEI	WIGL			
HPSAP (124-248)	:ALSLQGSIMTVGEKVFSSNGQSITFDAIQEACAR-AGGRIAV--PRNPEEN--EA--IASFVKKYNTYAYVGLTECP						

	80	90	110	120
RVVXLC1	:G-----NMWKD--CRME-----	WSDRGNVKKYKALAE-----	SYCLIMITHEKE-----	WKSMTCNFIAPVV--CKF
RSL	:RDKKKDFSWEWTDTRSDTYLT---	WDKNQ--PDHYQNK---	EFCELVSLTGYRL--	WNLQVCESKDAFL--CQC
PLIA	:SAN-----FTN-----	WAAGQ--PNEAD-----	GTCVKADTHGS-----	WHSASCDENLLVV--CEF
IX/XBPA	:RVQGK-VKQCNSEWSDGSSVSYEN--	WIEAES-----	KTCLGLEKETDFRK--	WVNIYCGQONPFV--CEA
MLHR	:RKIGKM--WTWV-GTNKTLTKEAENWGAGE--	PNNKSK---	EDCWEIYIKRERDSGKWNDDACHKRKAAL--	CYT
RKCR	:TDQGTGNWRWVDGTPFDYVQSRREWRKGQ--	PDNWRHNGNEREDCVHLQRM-----	WNDMACGTAYNWV--	CKK
RMBPA	:FMYVTGGRLTYSN-----	WKKDE--PNDHGSG--	EDCVTIVDNL-----	WNDISQASHTAV--CEF
HFCER	:RNLDLKGEFIWVDGSHVDYSN---	WAPGE--PFSRSQG--	EDCVMMRGSGR-----	WDAFCDRKLGAWV--CDR
HTET	:--NDMA-AEGTWVDMTGARIAYKN--	WETEITAQPDGGKT---	ENCAVLSGAANGK--	WFDKFRDQLPYI--COF
HPSAP	:SPGDFRYS DGTPVNYTN-----	WYRGE--PAGRGK--	EDCVEMYTDGQ-----	WDRNCLYSRLTI--CEF

Fig. 9

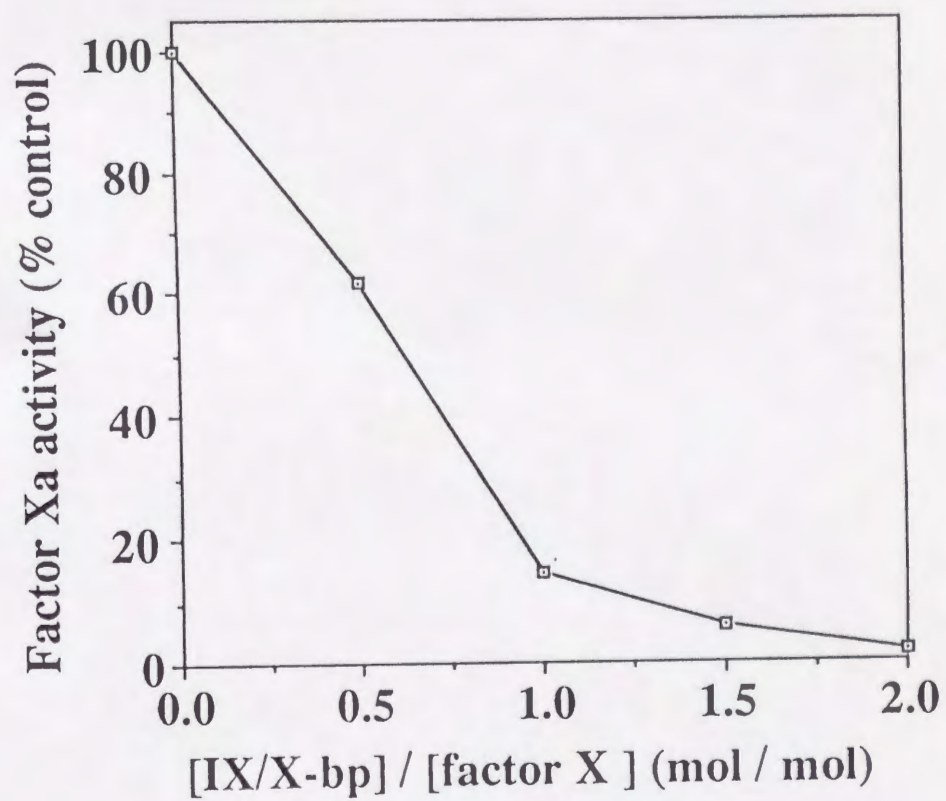


Fig. 10

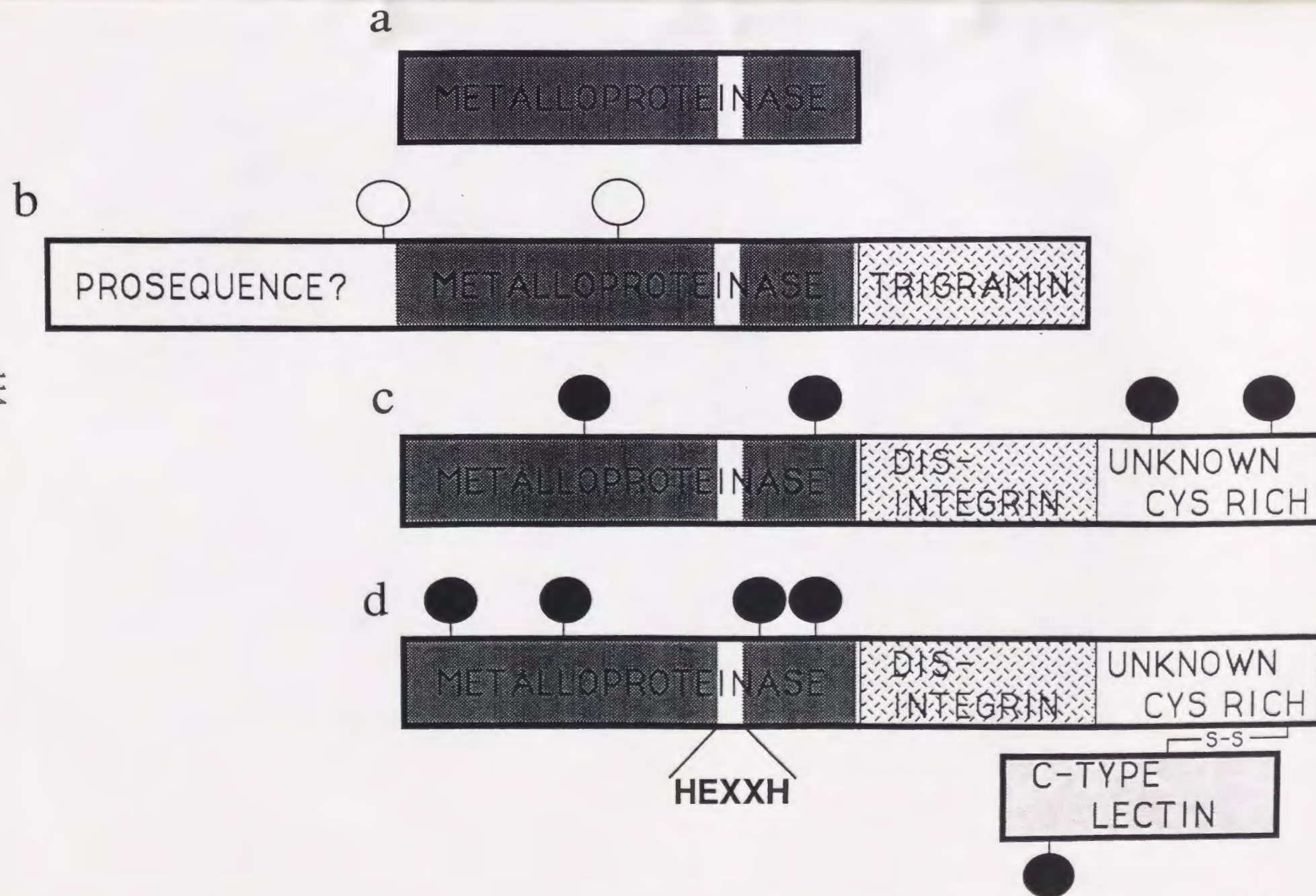


Fig. 11

CONCLUSION

Previously, we have shown that a low molecular weight hemorrhagic protein (HR2a, 202 amino acid residues) isolated from venom of *Trimeresurus flavoviridis* is a member of new subfamily of metalloproteinases. In the present study, the complete amino acid sequences of a high molecular mass hemorrhagic protein HR1B isolated from the same venom and RVV-X, the blood coagulation factor X activating enzyme, isolated from Russell's viper venom have been determined. Moreover, the structure-function relationship of these enzymes have been studied. HR1B, is a mosaic protein composed of 416 amino acid residues containing four Asn-linked oligosaccharide chains. The amino-terminal half (residues 1-203) of HR1B contains a metalloproteinase domain, the sequence of which is 62 % identical with that of HR2a and 52 % identical with that of Ht-d isolated from the *Crotalus atrox* venom. The most interesting finding is that the middle region (residues 204-300) of HR1B shows a striking similarity to disintegrins, Arg-Gly-Asp-containing platelet aggregation inhibitors, recently found in several viper venoms. This result suggests that the middle portion of HR1B may be important to stimulate synergistically hemorrhagic activity with the amino-terminal metalloproteinase domain. Interestingly, however, this region of HR1B does not contain the Arg-Gly-Asp-sequence which is known to be a putative binding site in the disintegrin to the platelet fibrinogen receptor, glycoprotein IIb/IIIa complex.

On the other hand, RVV-X (Mr 79,000) consisted of a disulfide-bonded two-chain glycoprotein with a heavy chain of Mr 59,000 and a light chain of heterogeneous Mr 18,000 (LC1) or

21,000 (LC2). The heavy chain had 427 residues containing four asparagine-linked oligosaccharides, and its entire sequence was similar to that of HR1B. The heavy chain contained three distinct domains, metalloproteinase, disintegrin (platelet aggregation inhibitor)-like and unknown cysteine-rich domains. The light chain LC1 consisted of 123 amino acid residues containing one asparagine-linked oligosaccharide and showed a sequence homology similar to those found in the so-called C-type (Ca^{2+} -dependent) lectin. Therefore, RVV-X is a novel metalloproteinase containing a mosaic structure with disintegrin-like, cysteine-rich, and C-type lectin-like domains. RVV-X potently inhibited collagen- and ADP-stimulated platelet aggregations, probably via its disintegrin-like domain, although this domain does not contain the Arg-Gly-Asp sequence which is conserved in various venom disintegrins and which is thought to be one of the interaction sites for platelet integrins. Our findings also indicate that snake venom factor IX/factor X-binding protein with a C-type lectin structure inhibits the RVV-X-catalyzed factor X activation, hence, the light chain of RVV-X probably participates in recognizing some portion of the zymogen factor X.

These findings indicate the structural and evolutionary relationship among these proteins, although each protein has diversified functional activity. These mosaic structures are also analogous to those of mammalian blood clotting serine proteinases and matrix metalloproteinases. While a common domain with apparently distinct functions can be defined among these proteinases, additional domains unique to individual proteinases may have specific functions.

ACKNOWLEDGMENTS

The study described in this thesis has been carried out under the guidance of Professor Sadaaki Iwanaga of Kyushu University. This thesis could not have been written without the cooperation, encouragement, and contributions of many people. I am grateful to Professor Sadaaki Iwanaga for his considerate guidance with magnanimity throughout this work. I am also grateful to Drs. Toshiyuki Miyata, Tamotsu Omori-Satoh, Takashi Morita, Yukari Saisaka, and Masaaki Moroi for helpful advise, discussion and continuous encouragement and to Miss Kyoko Oda, and Messrs. Soh-ichiro Kawada and Shinji Nishida for their cooperation. I am indebted to Mrs. Satsuki Hori-Kajiyama, and Misses Yuko Nishina and Sawako Oyama for amino acid and sequence analyses, to Mrs. Nobuko Ueno-Hashimoto and Itsuko Edamitsu for secretarial assistance, and Miss Mariko Ohara for reading the manuscripts. I wish to thank all the members of the laboratory of Professor Sadaaki Iwanaga.

