Volume 329, number 1,2, 116–120 FEBS 12875 August 1993

The complete amino acid sequence of a thrombin-like enzyme/gyroxin analogue from venom of the bushmaster snake (*Lachesis muta muta*)

Arinos Magalhaes[a], Beatriz Campos Brasil Da Fonseca[a], Carlos Ribeiro Diniz[a], John Gilroy[a] and Michael Richardson[b]

[a]*Centro de Pesquisa e Desenvolvimento, Fundacão Ezequiel Dias, CP 026, Rua Conde Pereira Carneiro 80, CEP 30510-010 Belo Horizonte (MG), Brazil* and [b]*Department of Biological Sciences, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK*

Received 24 June 1993

The complete amino acid sequence of a thrombin-like enzyme with gyroxin activity isolated from the venom of the bushmaster snake *Lachesis muta muta* was determined by automated and DABITC/PITC microsequencing of the intact protein; fragments derived from it by separate cleavages with cyanogen bromide, iodosobenzoic acid and hydroxylamine; and peptides resulting from enzymatic digestions with trypsin, pepsin, chymotrypsin, and elastase. The protein, which is composed of 228 residues, contains four putative sites of N-linked glycosylation and exhibits significant sequence similarities with other serine proteases reported from snake venoms.

Snake venom; Amino acid sequence; Thrombin-like enzyme; Gyroxin analogue; *Lachesis muta muta*

1. INTRODUCTION

The bushmaster snake (*Lachesis muta muta*), a pit viper reaching lengths of 3 m, occurs in forested regions of equatorial Central and South America. *Lachesis* is believed to be ancestral to *Agkistrodon* and *Crotalus* [1].

The venom of this crotalid snake which is notable for its haemorrhagic, proteolytic and blood-clotting activities [2–5], contains a thrombin-like enzyme which has been purified by a combination of gel filtration on Sephadex G100 and affinity chromatography on Sepharose-agmatin [5]. The enzyme was shown to be a glycoprotein composed of a single polypeptide chain which existed in several isoelectric forms with pIs ranging from 3.1 to 5.0, but the asialoenzyme focused as a narrow band at pH 8.7. SDS-PAGE analysis gave a single broad band with an apparent M_r, of 41–47kDa. The enzyme hydrolyzed synthetic substrates with a specificity similar to trypsin, was competitively inhibited by amidines and guanidines, and irreversibly inhibited by diisopropylfluorophosphate. The enzyme cleaved only fibrinopeptide A from fibrinogen, did not activate factor XIII and was devoid of kallikrein-like activity [5]. When injected into the tail veins of mice at levels of 0.015–0.130 μg/g mouse, the thrombin-like enzyme induced temporary episodes of opisthotonos and rapid rolling around the long axis of the animals [5,6].

Correspondence address: M. Richardson, Department of Biological Sciences, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, U.K. Fax: (44) (91) 374-3741.

Thrombin-like enzymes which cause similar gyratory activity in test animals have been reported from a number of other snake venoms. These include the so-called gyroxin from *Crotalus durissus terrificus*, crotalase from *Crotalus adamanteus*, and ancrod from *Agkistrodon rhodostoma* [7–9].

We now report the complete primary structure of this neurotoxic enzyme and gyroxin analogue from *L. muta muta* and its sequence similarities with some other serine proteases from snake venoms.

2. MATERIALS AND METHODS

2.1. *Purification of thrombin-like enzyme*

The methods used for the collection and storage of venom and the purification of the thrombin-like enzyme by gel filtration on Sephadex G-100 and affinity chromatography on Sepharose-agmatine were as described in [5].

2.2. *Determination of amino acid sequence*

2.2.1. Reduction and S-carboxymethylation

The protein was reduced and S-carboxymethylated as described in [10].

2.2.2. Chemical cleavages

Cleavage at methionine residues was achieved by treatment of a sample (3 mg) of the reduced and *S*-alkylated protein with cyanogen bromide (100-fold molar excess) in 2 ml of 70% formic acid under nitrogen for 48 h at 20°C. After lyophilization the residue was dissolved in 1 ml of 0.1 M ammonium bicarbonate containing 6 M guanidine-HCl and fractionated on a column (1 × 200 cm) of Biogel P-6 in 0.1 M ammonium bicarbonate. Each fraction was further purified by reverse phase HPLC on a column (25 cm × 4.6 mm) of Vydac C18 (218TP54, Technicol Stockport) using a linear gradient of 0–70% acetonitrile in 0.1% (v/v) aqueous trifluoroacetic acid.

 Published by Elsevier Science Publishers B.V.

Cleavage after tryptophan was obtained by treating a further sample (5 mg) with 15 mg of *o*-iodosobenzoic acid in 2 ml of 80% acetic acid containing 4 M guanidine-HCl and 10 mol% *p*-cresol under nitrogen at 22°C for 24 h in the dark as described in [11]. The products were desalted and partially fractionated on a column (1 × 100 cm) of Bio-gel P30 in 70% formic acid before being further purified by RP-HPLC as described above.

Cleavage of an asparagine–glycine peptide bond was carried out using 2 M hydroxylamine-HCl in 6 M guanidine-HCl maintained at pH 9.0 by the addition of 4.5 M LiOH during 4 h at 23°C as described in [11]. Purification of the C-terminal fragment required for sequencing was achieved by RP-HPLC.

2.2.3. Enzymatic cleavages

Samples (50–100 nmol) of the reduced and *S*-alkylated protein were digested separately with trypsin (2% w/w enzyme/substrate in 0.1 M ammonium bicarbonate, pH 8.2, 4 h at 37°C), pepsin (1% w/w enzyme/substrate in 5% formic acid, 2 h at 25°C), elastase (2% w/w enzyme/substrate in 0.1 M ammonium bicarbonate, pH 8.2, 3 h at 37°C), and chymotrypsin (2% w/w enzyme/substrate in 0.2 M *N*-ethylmorpholine-HCl, pH 8.5, 2 h at 37°C). The peptides produced were purified by RP-HPLC.

2.2.4. Sequence determination

The N-terminal sequence of the intact reduced and *S*-alkylated protein (5 nmol) was determined by automatic sequencing in an Applied Biosystems Model 477A pulsed liquid phase Sequencer/Model 120A PTH HPLC analyzer system using a standard Edman degradation programme. The larger fragments resulting from the chemical cleavages with cyanogen bromide, iodosobenzoic acid and hydroxylamine were sequenced by the automated means using 2–8 nmol amounts. The smaller peptides obtained from the enzymatic digestions were subjected to the manual method of 4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate (DABITC)/phenylisothiocyanate (PITC) double coupling microsequencing [12] using 10–25 nmol.

2.2.5. Amino acid analysis

Samples of the reduced and *S*-carboxymethylated protein and the various peptides were hydrolyzed with 5.6 M HCl containing 0.02% (v/v) cresol at 108°C for 24 h. The amino acids in the hydrolyzates were derivatized with PITC and analyzed by HPLC using the Waters Pico-Tag method [13]. The presence of free cysteine residues in the native protein was examined in 8 M urea by the method of Ellman [14,15].

2.2.6. Sequence comparisons

The amino acid sequence of the thrombin-like enzyme/gyroxin analogue was compared with those of other proteins stored in the PIR-International Protein Sequence Database (1990) by computer analysis using the FASTA/RDF2 programmes [16]. Sequence alignments were also obtained by using the CLUSTAL program [17].

3. RESULTS AND DISCUSSION

The amino acid sequence of the thrombin-like enzyme/gyroxin analogue from the venom of *Lachesis muta muta* determined by both automated and manual sequencing methods is shown in Fig. 1.

The molecular weight calculated from the sequence is 25,656 which is much lower than the values of 41–47 kDa estimated by SDS-PAGE [5]. However, this is not surprising since it was previously demonstrated by neuraminidase treatment that the protein is extensively glycosylated [5], and four putative sites of N-linked glycosylation, at residues 45, 81, 145 and 224 were revealed during the present work by blank cycles of sequencing in these positions where an ASx was indicated by the results of amino acid analysis and the positions were followed by the sequence –X–Ser/Thr. The other thrombin-like enzymes from snake venoms show considerable variation in their carbohydrate contents, with ancrod the α-fibrinogenase from the Malayan pit viper (*Agkistrodon rhodostoma*) containing five putative sites of N-linked glycosylation [18], whilst there are two identified sites in batroxobin from *Bothrops atrax moojeni* [19,20], but flavoxobin from the habu snake (*Trimeresurus flavoviridis*) does not contain carbohydrate [21].

The sequence of the first 25 amino acids at the N-terminal was found to be identical with that reported previously for the gyroxin analogue from *L. muta muta* [6]. Alignments of the amino acid sequence of the *Lachesis muta muta* thrombin-like enzyme with those of other proteins from snake venom are shown in Fig. 2. Comparison of complete sequences reveals that the *Lachesis* protein is most similar with the thrombin-like coagulant enzymes flavoxobin (61% sequence identity) [21], batroxobin (62% identity) [19,20] and ancrod (63% identity) [18]. It is noteworthy that all of these enzymes like the *Lachesis* protein release only fibrinopeptide A in the conversion of fibrinogen to fibrin [5,9]. On the other hand, the *Lachesis* protein has only 29% sequence identity with bovine thrombin [22] which releases both fibrinopeptides A and B from fibrinogen and activates factor XIII.

It can also be seen that the *Lachesis* protein has 57% homology with a factor V cleaving serine esterase from the venom of the Russell's viper (*Vipera russelli*) [23], 65% sequence identity amongst the 122 amino acids which can be compared with the known fragments of the thrombin-like crotalase from the eastern diamondback rattlesnake (*Crotalus adamanteus*) [24,25], 71% identity with the 77 amino acids known for catroxobin from *Crotalus atrox* [26], 78% homology in a comparison with the N-terminal 36 amino acids of a kallikrein-like enzyme from the prairie rattlesnake (*Crotalus viridis viridis*) [27], 80% identity in the first 30 residues with the gyroxin from the South American rattlesnake (*Crotalus durissus terrificus*) [6] and 58% similarity with the N-terminus of the gabonase from *Bitis gabonica* [28].

There is also some homology evident with the other serine proteinases, bovine trypsin (37%) [29] and human kallikrein (33%) [30]. This homology facilitates the identification of the probable catalytic triad of amino acids in the *Lachesis* thrombin-like protein as His^{43}, Asp^{88} and Ser^{177}. As might be expected these amino acids occur in regions of the sequences which are more highly conserved.

As no free cysteine residue was detected in the *Lachesis* protein by the Ellman reaction, it may be presumed that the 12 cysteine residues in the sequence are all involved in the formation of six disulphide bridges.

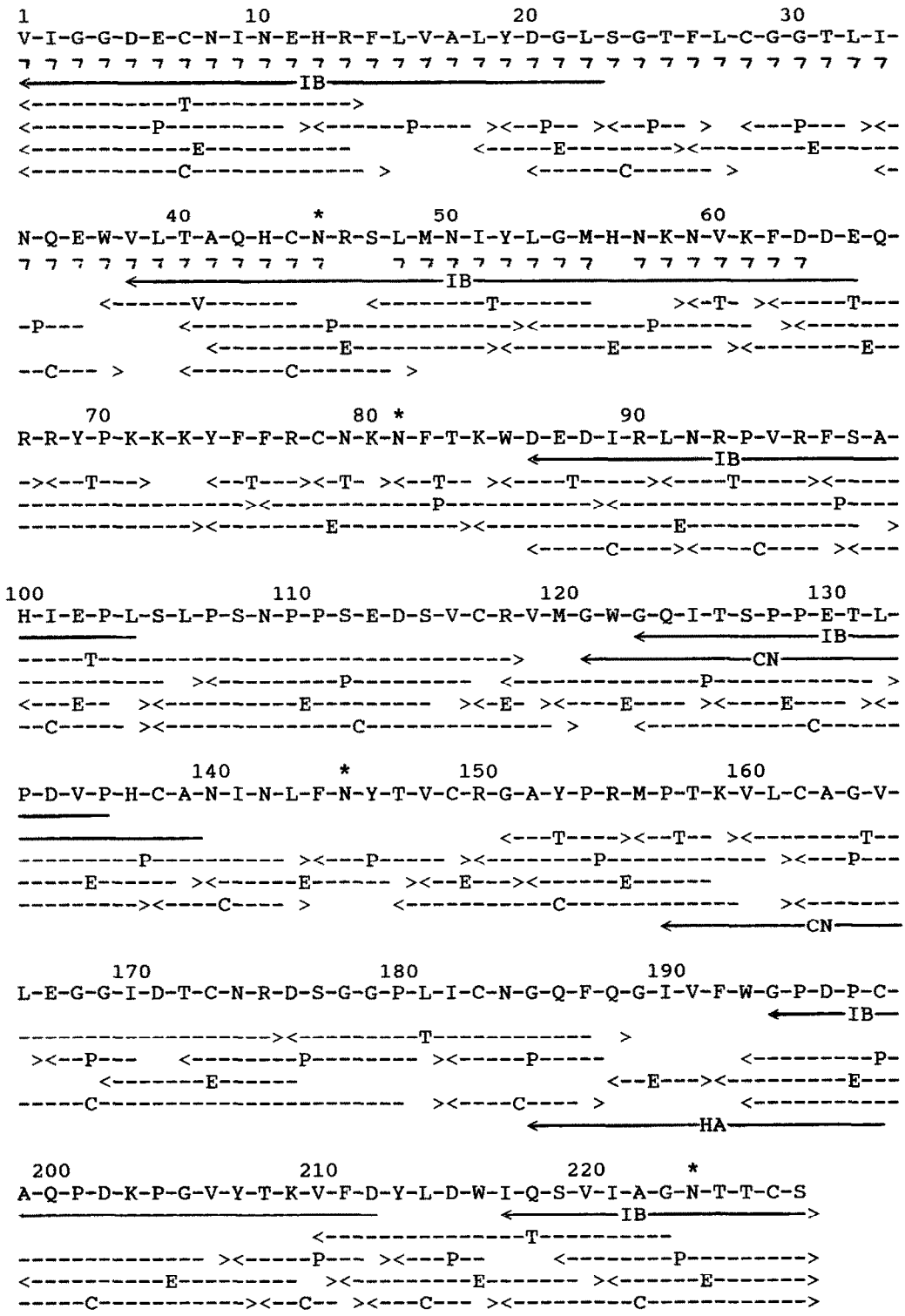

Fig. 1. The amino acid sequence of the thrombin-like/gyroxin analogue from the venom of the bushmaster snake (*Lachesis muta muta*). Arrows (—) indicate residues determined by automated degradation of the intact *S*-reduced and carboxymethylated protein. Solid lines indicate fragments derived from chemical cleavages with iodosobenzoic acid (IB), cyanogen bromide (CN) or hydroxylamine (HA) which were sequenced by the automatic method. Dotted lines indicate peptides from enzymatic digestions with trypsin (T), pepsin (P), elastase (E) and chymotrypsin (C) which were sequenced by the manual DABITC/PITC method. Asterisks (*) show sites of N-linked glycosylation.

```
                 10        20         30        40                     50
1  Lmm   VIGGDECNINEHRFLVALYDGLSG-TFLCGGTLINQEWVLTAQHCNRS-------LM-----NI
2  Anc   VIGGDECNINEHRFLVAVYEGTNW-TFICGGVLIHPEWVITAEHCARR-------RM-----NL
3  Fla   VIGGDECNINEHPFLVALYDAWSG-RFLCGGTLINPEWVLTAAHCDS-----KNFKM-----K-
4  Bat   VIGGDECDINEHPFL-AF-MYYSP-RYFCGMTLINQEWVLTAAHCNRR------F-M-----RI
5  Rvv   VVGGDECNINEHPFLVALYTSTSS-TIHCGGALINREWVLTAAHCDRR-------NI-----RI
6  Bth   IVEGQDAEVGLSPWQVMLFRKSPQ-ELLCGASLISDRWVLTAAHCLLYPPWBKNFTVDDLLVRI
7  Btr   IVGGYTCGANTVPYQVSLQ---SG-YHFCGGSLINSQWVVSAAHC-----YKSGIGV-----R-
8  Hka   IVGGTNSSWGEWPWQVSLQVKLTAQRHLCGGSLIGHQWVLTAAHCFDG--LPLQDVW-----RI
9  Cro   VIGGDECNINEHRFLVALYDYWXQ-XFL
10 Cat   VVGGDECNINEHRSLVAIFV--STE-FDCGGDLINVEWVLTAAHC
11 Gab   VVGGAECKIDGHRCLALLY
12 Ckl   VIGGDECNINEHRFLVALYDP-DG--FLSGGILLNEE
13 Gyr   VIGGDECNINEHNFLVALYEYWS-ESFLCGG
            *                         *  *    **  * **

          60        70         80            90        100       110
1  YLGMHNKNVKFDDEQRRYPKKKYFFRC-NK-NFTKWDEDI---RLNRPVRFSAHIEPLSLPSNPPS---
2  VFGMHRHSEKFDDEQERYPKKRYFIRC-NKT-RTSWDEDIMLIRLNKPVNNSEHIAPLSLPSNPPI---
3  -LGAHSQKVLNEDEQIRNPKEKFI--CPNKKNTEVLDKDIMLIKLDSPVSYSEHIAPLSLPSSPPS---
4  HLGKHAGSVANYDEVVRYPKEKFI--CPNKKKNVITDKDIMLIRLDRPVKNSEHIAPLSLPSNPPS---
5  KLGMHSKNIRNEDEQIRVPRGKYF--CLNTKFPNGLDKDIMLIRLRRPVTYSTHIAPVSLPSRSRG---
6  --GKHSRTRYERKVEKISMLDKIYIH-PRYNWKENLDRDIALLKLKRPIELSDYIHPVCLPDKQTAAKL
7  -LGQDNINVVEGNQQFISASKSIV-H-PSYN-SNTLNNDIMLIKLKSAASLNSRVASISLPTSCAS---
8  YSGILNLSDITKDTPFSQIKEIII-H-QNYKVSEG-NHDIALIKLQAPLNYTEFQKPICLPSKGDTST-
9        RSVQFDKEQQR                   DKDIMLIRLNKPVSYSEHIAPLSLPSSPPI
     *                                   **    *              **

           120         130        140         150         160
1  EDSVC--RVMGWGQITS--PPET-LPDVPHCANINL--FNYTVCRGAYPRM-P-TKVLCAGVL--EGG-
2  VGSDC--RVMGWGSINR--RIDV-LSDEPRCANINL--HNFTMCHGLFRKMPKKGRVLCAGDL--RGR-
3  VGSVC--RIMGWGSITP--VEET-FPDVPHCANINL--LDDVECKPGYPELLPEYRTLCAGVL--QGG-
4  VGSVC--RIMGWGAITT--SEDT-YPDVPHCANINL--FNNTVCREAYNGL-P-AKTLCAGVL--QGG-
5  VGSRC--RIMGWGKIST--TEDT-YPDVPHCTNIFI--VKHKWCEPLYPWVPADSRTLCAGIL--KGG-
6  LHAGFKGRVTGWGNRRETWTTSVAEVQPSVLQVVNLPLVERPVCKASTRIRIT-NDMFCAGYKPGEGKR
7  AGTQC--LISGWGNTKS--SGTS-YPDVLKCLKA--PILSNSSCKSAYPGQITSN-MFCAGYL--EGG-
8  IYTNC--WVTGWGFSKE--KGEI-QNILQKV-NI--PLVTNEECQKRYQDYKITQRMVCAGYK--EGG-
9  VGSVC--RAMGWGQTTS--PQET-LPDVPHCANINL--LDYEVC
10                                                         TLCAGI-P-EGG-
             ***                              *              ***     *

            180             190       200       210       220
1  IDTCNRDSGGPL--IC--NGQFQ--GIVFWGPDPCAQPDKPGVYTKVFDYLDWIQSVIAGNTT--CS
2  RDSCNSDSGGPL--IC--NEELH--GIVARGPNPCAQPNKPALYTSIYNYRDWVNNVIAGNAT--CSP
3  IDTCGFDSGTPL--IC--NGQFQ--GIVYIGSHPCGQSRKPGIYTKVFDYNAWIQSIIAGNTAATCLP
4  IDTCGGDSGGPL--IC--NGQFQ--GILSWGSDPCAEPRKPAFYTKVFDYLPWIQSIIAGNKTATC-P
5  RDTCHGDSGGPL--IC--NGQIQ--GIVAGGSEPCGQHLKPAVYTKVFDYNNWIQNIIAGNRTVTCPP
6  GDACEGDSGGPFVMKSPYNNRWYQMGIVSWG-EGCDRNGKYGFYTHVFRLKKWIQKVIDRLGS
7  KDSCQGDSGGPV--VC--SGKLQ--GIVSWG-SGCAQKNKPGVYTKVCNYVSWIKQTIASN
8  KDACKGDSGGPL--VCKHNGMWRLVGITSWG-EGCARREQPGVYTKVAEYMDWILEKTQSSDGKAQMQSPA
9                                    CDCKEKYFDCWNTFKED
10 LDTCGGDSGGPL--IC--DGKPD--GITS
    * *  *** *              **   *   *                 *
```

Fig. 2. Comparison of the amino acid sequences of (1) the thrombin-like enzyme/gyroxin analogue from the bushmaster snake (*Lachesis muta muta*) (Lmm), with (2) ancrod (Anc) from the Malayan pit viper (*Agkistrodon rhodostoma*) [18], (3) the flavoxobin (Fla) from the habu snake (*Trimeresurus flavoviridis*) [21], (4) batroxobin (Bat) from *Bothrops atrox moojeni* [19,20], (5) a factor V cleaving serine esterase (Rvv) from the Russell's viper (*Vipera russelli*) [23], (6) bovine thrombin (Bth) [22], (7) bovine trypsin (Btr) [29], (8) human kallikrein (Hka) [30], (9) fragments of crotalase (Cro) from the eastern diamondback rattlesnake (*Crotalus adamanteus*) [24,25], (10) fragments of catroxobin (Cat) from *Crotalus atrox* [26], (11) N-terminal sequence of the gabonase from *Bitis gabonica* [28], (12) N-terminal of a kallikrein-like enzyme (Ckl) from the prairie rattlesnake (*Crotalus viridis viridis*) [27] and (13) gyroxin from the South American rattlesnake (*Crotalus durissus terrificus*) [6]. Gaps (-) are inserted in the sequences to maximize homology. Potential sites of N-linked glycosylation are underlined. Asterisks (*) show residues which are invariant in all sequences.

Acknowledgements: We are grateful for financial support from the Brazilian CNPQ (Conselho Nacional Desenvolvimento Cientifico e Tecnologico), FAPEMIG (Fundacao de Amparo a Pesquisa do Estado de Minas Gerais) and the University of Durham. We also thank Dr. M. Cordeiro, Miss A. Valentim and Miss S.G. de Figueiredo for valuable technical assistance.

REFERENCES

[1] Brattstrom, B.H. (1964) Trans. San Diego Soc. Nat. Hist. 13, 185–207.
[2] Deutsch, H.F. and Diniz, C.R. (1955) J. Biol. Chem. 216, 17–26.
[3] Magalhaes, A., Oliveira, G.J. and Diniz, C.R. (1981) Toxicon 19, 279–294.
[4] Sanchez, E.F., Magalhaes, A. and Diniz, C.R. (1987) Toxicon 25, 611–619.
[5] Silveira, A.M.V., Magalhaes, A., Diniz, C.R. and De Oliveira, E.B. (1989) Int. J. Biochem. 21, 863–871.
[6] Da Silva, N.J., Aird, S.D., Seebert, C. and Kaiser, I.I. (1989) Toxicon 27, 763–771.
[7] Barrabin, H., Martiarena, J.L., Vidal, J.C. and Barrio, A. (1978) in: Toxins: Animal, Plant and Microbial (Rosemberg, P. ed.) Pergamon Press, Oxford/New York.
[8] Alexander, G., Grothusen, J., Zepeda, H. and Schwartzman, R.J. (1988) Toxicon 26, 953–960.
[9] Pirkle, H. and Theodor, I. (1992) in: Handbook of Natural Toxins (Tu, A.T. ed.) vol. 5, pp. 225–252, Dekker, New York.
[10] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[11] Aitken, A., Geisow, M.J., Findlay, J.B.C., Holmes, C. and Yarwood, A. (1989) in: Protein Sequencing: a Practical Approach (Findlay, J.B.C. and Geisow, M.J. eds.) pp. 43–68, IRL Press, Oxford.
[12] Chang, J.Y., Brauer, D. and Wittman-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[13] Millipore Waters Chromatography Corp. (1984) Operators Manual No. 88140.
[14] Ellman, G.L. (1959) Arch. Biochem. Biophys. 82, 70–77.
[15] Kortt, A.A. and Liu, T.Y. (1973) Biochemistry 12, 320–327.
[16] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[17] Higgins, D.G. and Sharp, P.M. (1989) Computer Applications Biol. Sci. 5, 151–153.
[18] Burkhart, W., Smith, G.F.H., Su, J.L., Parikh, I. and Levine, H. (1992) FEBS Lett. 297, 297–301.
[19] Itoh, N., Tanaka, N., Mihashi, S. and Yamashina, I. (1987) J. Biol. Chem. 262, 3132–3135.
[20] Itoh, N., Tanaka, N., Funakoshi, T., Mihashi, S. and Yamashina, I. (1988) J. Biol. Chem. 263, 7628–7631.
[21] Shieh, T.C., Kawabata, S., Kihara, H., Ohno, M. and Iwanaga, S. (1988) J. Biochem. 103, 596–605.
[22] Magnusson, S. (1971) in: The Enzymes (Boyer, P.D. ed.) vol. 3, pp. 277–321, Academic Press, New York.
[23] Tokunaga, F., Nagasawa, K., Tamura, S., Miyata, T., Iwanaga, S. and Kisiel, W. (1988) J. Biol. Chem. 263, 17471–17481.
[24] Pirkle, H., Markland, F.S., Theodor, I., Baumgartner, R., Bajwa, S.S. and Kirakossian, H. (1981) Biochem. Biophys. Res. Commun. 99, 715–721.
[25] Pirkle, H., Markland, F.S. and Theodor, I. (1983) Fed. Proc. 42, 1993.
[26] Pirkle, H., Theodor, I. and Lopez, R. (1989) Thromb. Res. 56, 159–168.
[27] Komori, Y., Nakai, T. and Sugihara, H. (1988) Biochim. Biophys. Acta 967, 92–102.
[28] Pirkle, H., Theodor, I., Miyada, D. and Simmons, G. (1986) J. Biol. Chem. 261, 8830–8835.
[29] Kauffman, D.L. (1965) J. Mol. Biol. 12, 929–932.
[30] Chung, D.W., Fujikawa, K., McMullen, B.A. and Davie, E.W. (1986) Biochemistry 25, 2410–2417.