Journal of Chromatography, 562 (1991) 421–434
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 5577

Liquid secondary ion mass spectrometry applied to structural confirmation of enzymically prepared C-terminal-truncated derivatives of recombinant hirudin

M. MAFTOUH, C. PONTHUS, A. TUONG and C. PICARD*

Sanofi Recherche, 195 Route d'Espagne, 31036 Toulouse Cédex (France)

ABSTRACT

The thrombin-specific inhibitor, hirudin variant rHV_2-Lys 47 (rHirudin), is a 65-amino acid polypeptide produced by recombinant DNA technology in yeast. Previous studies have shown that the acidic C-terminal segment of hirudin is susceptible to enzymic degradation. To address the question of C-terminal-truncated forms of the protein in terms of by-products or metabolites, well-defined reference compounds are needed. We prepared nine derivatives by carboxypeptidase Y digestion of rHirudin followed by a two-step chromatographic purification. Liquid secondary ion mass spectrometric measurements performed on peptides collected after reversed-phase high-performance liquid chromatography showed three pure forms (1–64, 1–63 and 1–56) and three mixtures of two forms each (1–62 + 1–61, 1–58 + 1–57 and 1–55 + 1–54), which were readily distinguished from one another by their mass spectra. Further purification of these co-eluted samples was achieved by ion-exchange chromatography and their structures were confirmed by liquid secondary ion mass spectrometry. Preliminary studies conducted on intact rHirudin indicated that this is an excellent analytical tool for mass measurements of hirudin-related proteins. Indeed, it allowed rapid (within 10–15 min), precise (0.50 a.m.u. relative to expected value), reproducible (mean MH^+ = 6907.64 ± 0.42 a.m.u.), sensitive (up to 500 ng, *i.e.* 72 pmol) and specific measurement of the quasi-molecular ion (MH^+) of the protein, and was thus readily applicable to the analysis of several derivatives.

INTRODUCTION

Hirudin variant rHV_2-Lys-47 (rHirudin) is a 65-amino acid polypeptide (mol. wt. 6906.5) produced by recombinant DNA technology in the yeast, *Saccharomyces cerevisiae* [1]. It differs from one of the reported natural hirudin variants secreted by the leech, *Hirudo medicinalis*, by absence of sulphate on Tyr-63 and has a Lys-47 instead of Asn-47 (Fig. 1). rHirudin and native iso-hirudins are the most potent and highly specific inhibitors of thrombin known [2–4]. Pharmacolog-

[a] The pharmaceutical development of recombinant hirudin variant rHV_2Lys 47 is being pursued within a partnership agreement between Transgène (16 Rue H.-Regnault, 92411 Courbevoie Cedex, France) and Sanofi (40 Avenue Georges-V, 75008 Paris, France).

0378-4347/91/S03.50

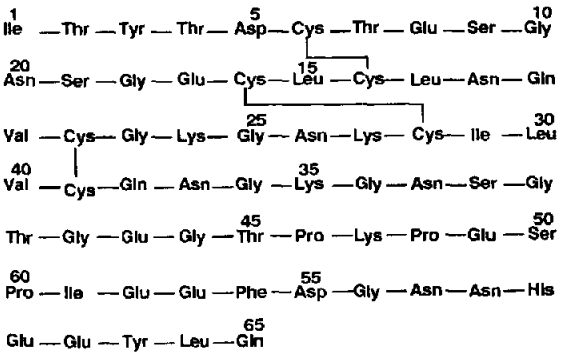

Fig. 1. Primary structure of recombinant hirudin variant rHV2-Lys-47, according to Harvey *et al.* [6], with residue 47 as Lys instead of Asn as reported.

ical studies of rHirudin in animals have demonstrated its *in vivo* anti-thrombotic efficacy [5] and offer promise for hirudin as a pharmaceutical thrombosis therapy.

All hirudins possess an N-terminal compact domain held by three disulphide bridges and a highly acidic C-terminal portion [6–10]. It has been demonstrated that this C-terminal region is involved in the binding of hirudin to thrombin [9,11]. Previous studies have shown that the acidic C-terminal segment of hirudin is freely accessible to enzymic degradation [11,12]. To address the question of C-terminal-truncated forms of the protein in terms of by-products during fermentation processes or as metabolites in rHirudin-treated animals and humans, well-defined standard compounds are needed.

We report here (i) the preparation of nine C-terminal-truncated forms of rHirudin obtained by carboxypeptidase Y (CPY) digestion of the protein, (ii) their two-step purification by reversed-phase (RP) and ion-exchange high-performance liquid chromatography (IE-HPLC) and (iii) their structural characterization by amino acid composition and liquid secondary ion mass spectrometric (LSIMS) analyses.

EXPERIMENTAL

Chemicals

Recombinant hirudin rHV_2-Lys-47, produced in the yeast *S. cerevisiae*, was obtained from Sanofi Elf Bio-Recherches (Labège-Innopole, France). CPY from yeast was obtained from Boehringer (Mannheim, F.R.G.). HPLC-grade S acetonitrile was obtained from Rathburn (Walkerburn, U.K.). Sequencer-grade trifluoroacetic acid (TFA) was obtained from Pierce (Rockford, IL, U.S.A.). Aristar-grade formic acid was obtained from BDH (Poole, U.K.). High-grade reagents used for amino acid composition analyses, including hydrochloric acid (6 *M*

HCl), phenol, phenylisothiocyanate (PITC), standard amino acids, triethylamine, acetic acid, ethanol and sodium acetate trihydrate were those recommended in the Waters PicoTag amino acid analysis manual (Milford, MA, U.S.A.). All other chemicals were of reagent or chromatographic grade. The aqueous solutions were all made with high-purity water from a Milli-Q purification system (Millipore, Bedford, MA, U.S.A.).

Digestion of rHirudin with CPY

Time-course of CPY digestion. rHirudin (200 μg, 29 nmol) was digested with CPY (5 μg in 5 μl water; enzyme-to-substrate ratio, 1:40, w/w) in 400 μl of 0.01 *M* phosphate buffer (pH 5.5) at 37°C with shaking. At appropriate times between 10 s and 4 h, the reaction was quenched with an equal volume of 0.2% TFA. After centrifugation at 14 000 *g*, *ca.* 25% of each sample was analysed by RP-HPLC.

Preparation of C-terminal-truncated forms of rHirudin. The derivatives generated by CPY digestion (five-fold scale-up: 1 mg of hirudin) were purified by chromatography, and the peaks were collected using a Model 201 fraction collector (Gilson Medical Electronics, Villiers-Le-Bel, France). After drying in a Gyrovap vacuum-centrifuge (VA Howe, London, U.K.), the samples were submitted to a second RP-HPLC purification step.

Chromatography

RP-HPLC. The time-course of CPY digestion and purification of generated compounds were monitored by RP-HPLC. The chromatograph used (Waters) consisted of a 600 E quaternary-solvent pump, a cooled (5°C) 710B WISP autosampler and a 990+ photodiode array detector. Separations were performed on a Synchropak RP-4 column (250 × 4.1 mm I.D., 5 μm particle size, 300 Å) from Synchrom (Linden, IN, U.S.A.). The eluent consisted of two solvents: (A) 0.1% TFA in water, and (B) 0.1% TFA in acetonitrile–water (50:50, v/v). The linear gradient used was 20% to 30% B in 40 min followed by 30% to 60% B in 40 min, at a flow-rate of 1 ml/min.

IE-HPLC. The C-terminal-truncated forms, which co-eluted in the RP-HPLC mode, were repurified by anion-exchange HPLC. The chromatograph used (Waters) comprised two 510 pumps controlled by a 680 gradient programmer, a manual U6K injector, a 441 fixed-wavelength detector (214-nm filter) and a 740 data integrator. Separations were carried out on an anion-exchange ProPac PA1 precolumn (50 × 4.0 mm I.D.) and column (250 × 4.0 mm I.D.) from Dionex (Sunnyvale, CA, U.S.A.). Both were filled with 0.2-μm resin microbeads bonded to a 10-μm polymeric particles. Isocratic elution over 10 min at 0% B was followed by a linear gradient from 0% to 50% B in 50 min, at a flow-rate of 1.3 ml/min. Solvent A was 20 m*M* Tris (pH 7.0) and solvent B was 0.5 *M* sodium chloride in A.

Amino acid composition analysis

rHirudin and its related C-terminal-truncated forms (10 μg (1.45 nmol) of hirudin-equivalent based on RP-HPLC UV absorbance) were dissolved in 50 μl of 0.1% TFA and transferred to 50 × 6 mm I.D. Pyrex hydrolysis tubes (Corning, NY, U.S.A.) previously cleaned by heating at 600°C.

Amino acid analysis, except for cysteine/cystine, was performed essentially as previously described [13] using a Waters PicoTag workstation. Samples were hydrolysed by 6 *M* HCl vapour containing 0.1% phenol at 110°C during 24 h, and derivatized with PITC. Phenylthiocarbamyl amino acid derivatives were then analysed using an application-specific PicoTag RP-HPLC column (150 × 3.9 mm I.D., 4 μm particle size).

Cysteine/cystine analysis as cysteic acid was conducted by subjecting protein samples to performic acid oxidation [14] prior to standard amino acid analysis as described above. Performic acid was prepared by mixing 98% formic acid and 30% hydrogen peroxide (19:1, v/v) and left to stand for 2 h at 25°C. Then 10 μl of the mixture were added to the dry protein samples (10 μg), which were kept for 20 min at 25°C. Reagent was removed by successive drying, solubilization in 50 μl of water, and drying again under vacuum.

Mass spectrometry

LSIMS analysis was performed in the positive-ion mode using a ZAB-2E double-focusing mass spectrometer (VG Analytical, Manchester, U.K.) equipped with a caesium-ion gun operating at 2 μA current of a 35-kV energy beam. The accelerating voltage was 8 kV. Mass spectra were recorded with a VAXStation 2000, using an Opus 2000 data system. All scans were acquired in the continuum mode and then summed, smoothed and mass-assigned by the data system. A magnetic scan from 8000 to 4300 a.m.u. at a rate of 20 s/decade was used, and the spectra were recorded at 800 resolution. The peptides were dissolved in 5% aqueous acetic acid to a concentration of 5 μg/μl. 1-Thioglycerol containing 1% TFA was used as matrix.

RESULTS AND DISCUSSION

Production of C-terminal-truncated forms of rHirudin by CPY hydrolysis

We used CPY instead of other carboxypeptidases because of its ability to stepwise remove most amino acid residues from the C-terminus of polypeptide chains. In addition, its optimum pH for proteolytic activity (pH 5.5–6.5) was compatible with the chemical stability of rHirudin. The rate at which each amino acid is released depends both on the primary structure of the polypeptide substrate and the nature of the C-terminal residue [15], and unfortunately, there are no universal reaction conditions that can be applied to all proteins. Therefore, it appeared necessary to perform a preliminary kinetic experiment on rHirudin to determine the optimal conditions by which the various truncated forms could be

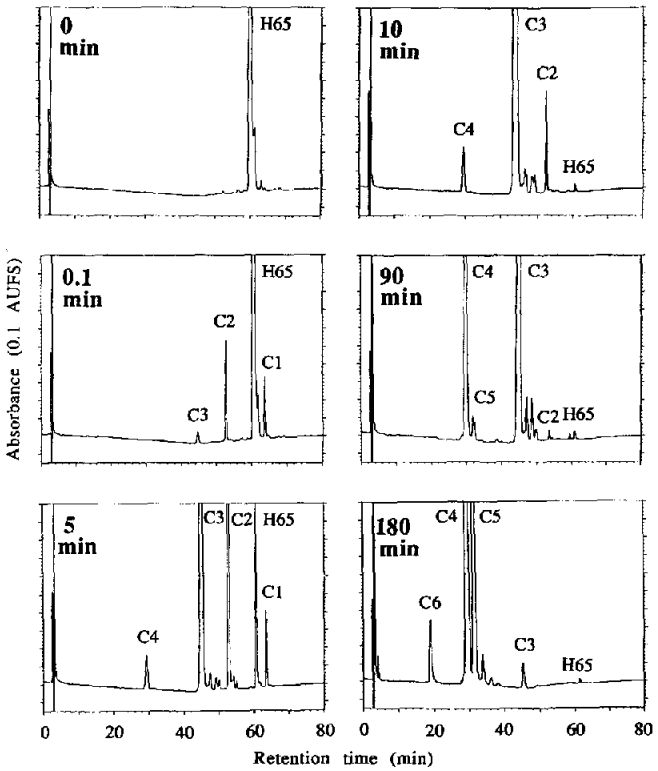

Fig. 2. RP-HPLC monitoring of the time-course of CPY digestion of rHirudin. Representative chromatograms obtained before addition of the enzyme (time 0) and after 0.1, 5, 10, 90 and 180 min of incubation. For chromatographic conditions, see Experimental.

produced. Fifteen sampling intervals (from 0.1 to 240 min) were taken to monitor the rate of release of the amino acids. Fig. 2 shows the RP-HPLC monitoring of the time-course of the cleavage of rHirudin by CPY for six representative sampling times. Altogether, we noted the appearance of six major peaks (C1 to C6) in addition to the starting rHirudin.

From this kinetic experiment, a few samplings corresponding to the maximum recovery of each of the six major peptides were retained for micropreparative chromatography. The time-points chosen were those indicated in Fig. 2.

Structural characterization of C-terminal-truncated forms of rHirudin

Amino acid analysis. Aliquots of the six RP-HPLC-purified peptides were subjected to amino acid composition analysis. Table I shows the resulting amino acid composition data. For several of the compounds the results did not fit the anticipated step-by-step release of single amino acids from native 1–65 rHirudin (predicted 1–64 to 1–59 C-terminal-truncated forms). Unexpectedly, the amino acid

TABLE I

AMINO ACID COMPOSITION ANALYSES[a] OF PEPTIDES GENERATED BY CPY DIGESTION OF rHIRUDIN

RP-HPLC peptides	H65	C1 (H64)[b]	C2 (H63)	C3 (H62)	C4 (H61)	C5 (H60)	C6 (H59)
Cya[c]	5.8(6)[d]	5.5	4.8	4.4	5.3	6.8	5.8
Asp + Asn	9.4(9)	8.8	10.0	9.9	10.2	10.1	9.8[9]
Glu + Gln	10.5(11)	9.7[10][e]	9.7[10]	9.5[10]	7.5[9][e]	6.9[8][e]	6.4[8]
Ser	4.2(4)	4.3	4.3	4.3	4.4	4.5	4.4
Gly	10.6(10)	10.1	11.1	10.7	11.0	11.1	10.9
His	1.0(1)	0.9	1.0	1.0	1.0	1.0	1.0
Arg	n.f.(0)[f]	n.f.	n.f.	n.f.	n.f.	n.f.	n.f.
Thr	5.7(5)	5.3	5.4	5.8	6.0	6.0	5.9
Ala	n.f.(0)	n.f.	n.f.	n.f.	n.f.	n.f.	n.f.
Pro	3.3(3)	3.0	3.9	3.4	2.3[3]	2.3[3]	2.2[2][e]
Tyr	2.3(2)	2.2	2.4	1.2[1][e]	1.2	1.2	1.2
Val	2.0(2)	2.0	2.0	2.0	2.0	2.0	2.0
Met	n.f.(0)	n.f.	n.f.	n.f.	n.f.	n.f.	n.f.
Ile	2.8(3)	2.7	2.9	2.9	1.9[3]	1.9[3]	1.8[3]
Leu	4.2(4)	4.4	2.7[3][e]	3.2	3.0	3.0	3.1
Phe	1.1(1)	1.1	1.0	1.0	1.0	1.0	n.f.[1]
Lys	4.2(4)	3.9	4.2	4.4	4.7	4.5	4.4

[a] Values found from duplicate peptide hydrolysates. Calculations were made by normalization to a Val content of 2, and are given as mole residue per mole peptide.

[b] (H64) to (H59) are the predicted peptides in the case of a step-by-step release of single amino acid by CPY (1–64 to 1–59 forms).

[c] Cysteic acid.

[d] Numbers in parentheses indicate the expected values based on the amino acid sequence of native rHirudin, 1–65.

[e] Predicted losses and, in square brackets, the theoretical number of amino acid residues expected in the case of a step-by-step release of single amino acids (1–64 to 1–59 supposed forms), indicating that the experimental values did not fit the predicted ones.

[f] n.f. = not found.

analyses were inadequate for the unambigous identification of the composition of the truncated peptides.

LSIMS. A preliminary study conducted on intact rHirudin indicated that LSIMS was an excellent analytical tool for mass measurements of rHirudin-related proteins. A typical spectrum obtained from 5 μg of rHirudin is shown in Fig. 3. The analyses were performed at low resolution not only to optimize sensitivity but also to avoid misinterpretation of the monoisotopic peak within the molecular ion cluster. Similar assignment problems when applying high resolution to the mass determination of proteins have been described elsewhere [15–17].

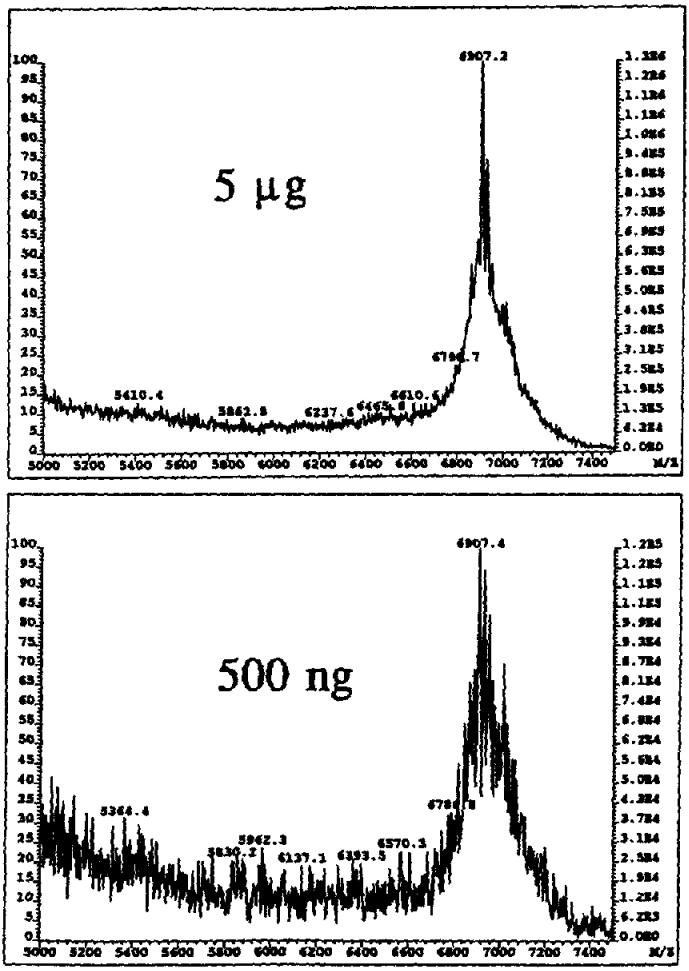

Fig. 3. LSIMS spectra obtained from 5 µg and 500 ng of intact rHirudin.

At present, as the isotope clusters are not resolved, only the average chemical mass can be assigned to each peak [18–20].

In this case, the average molecular mass was found to be 6907.2 (MH^+, expected value 6907.5). In addition, a prominent peak associated with the doubly charged protonated molecular ion (MH_2^{2+}) was also observed.

The reproducibility was checked with a 5 µg/µl solution of rHirudin. Eight consecutive measurements were done, and the mean value observed was 6907.64 (MH^+) with a standard deviation (σ) of 0.42. The precision was better than that achieved with plasma desorption mass spectrometry (PDMS) [21], and was equivalent to the results of LSIMS analysis of small proteins [15,22].

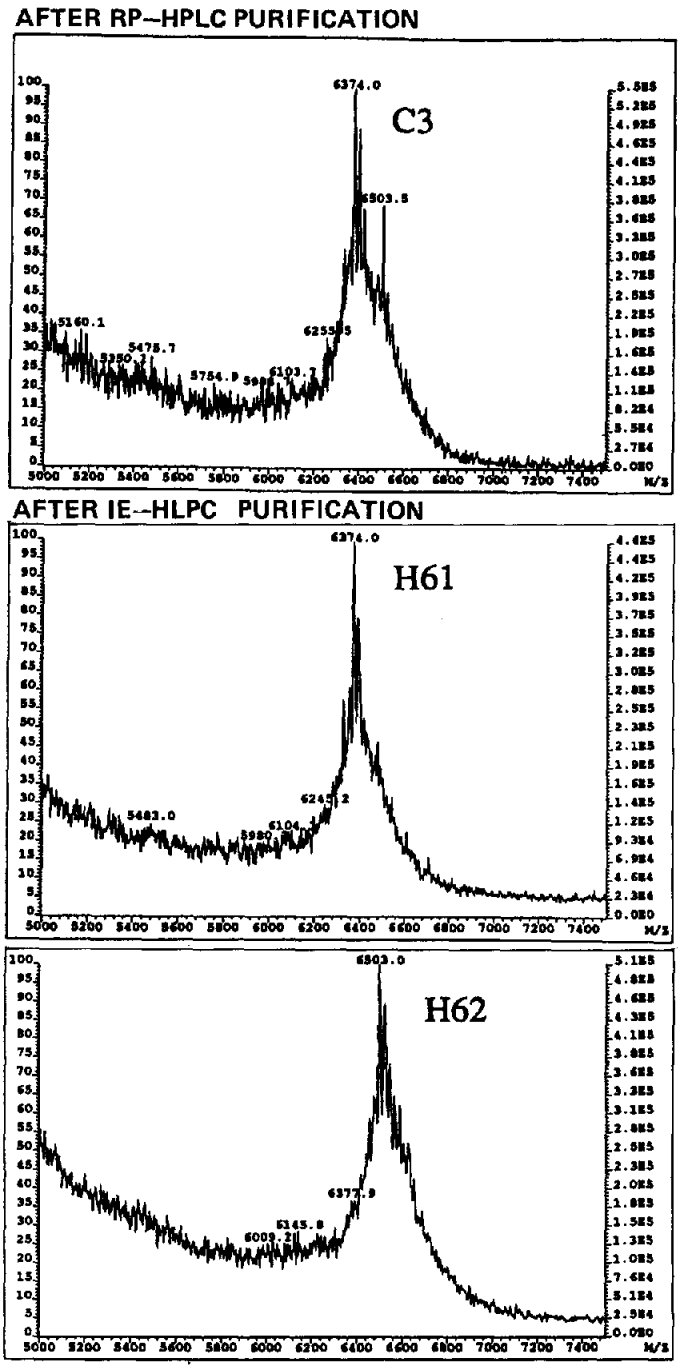
a AFTER RP--HPLC PURIFICATION
C3
6374.0
6803.8
5160.1
5475.7
5784.9
6103.7
AFTER IE--HLPC PURIFICATION
H61
6374.0
5483.0
6104
H62
6803.0
6377.9
5145.8
6009.2
5000 5200 5400 5600 5800 6000 6200 6400 6600 6800 7000 7200 7400
M/z

Fig. 4.

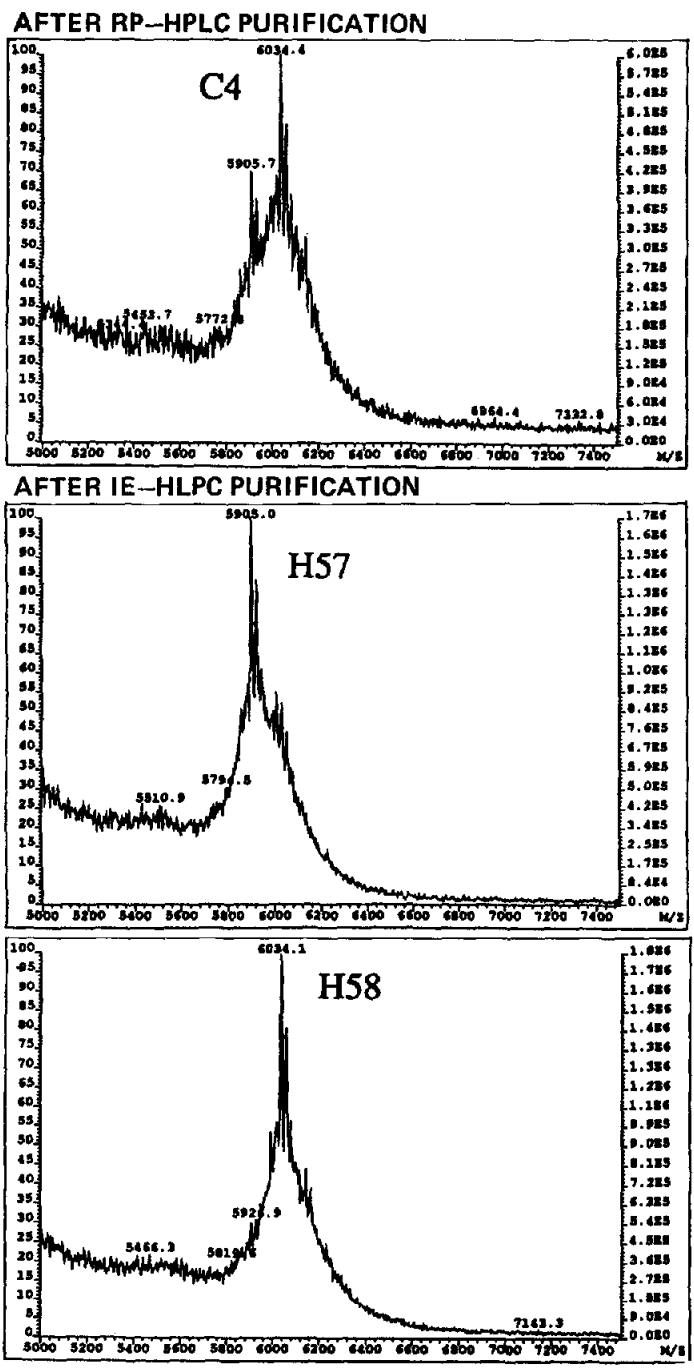

Fig. 4.

(Continued on p. 430)

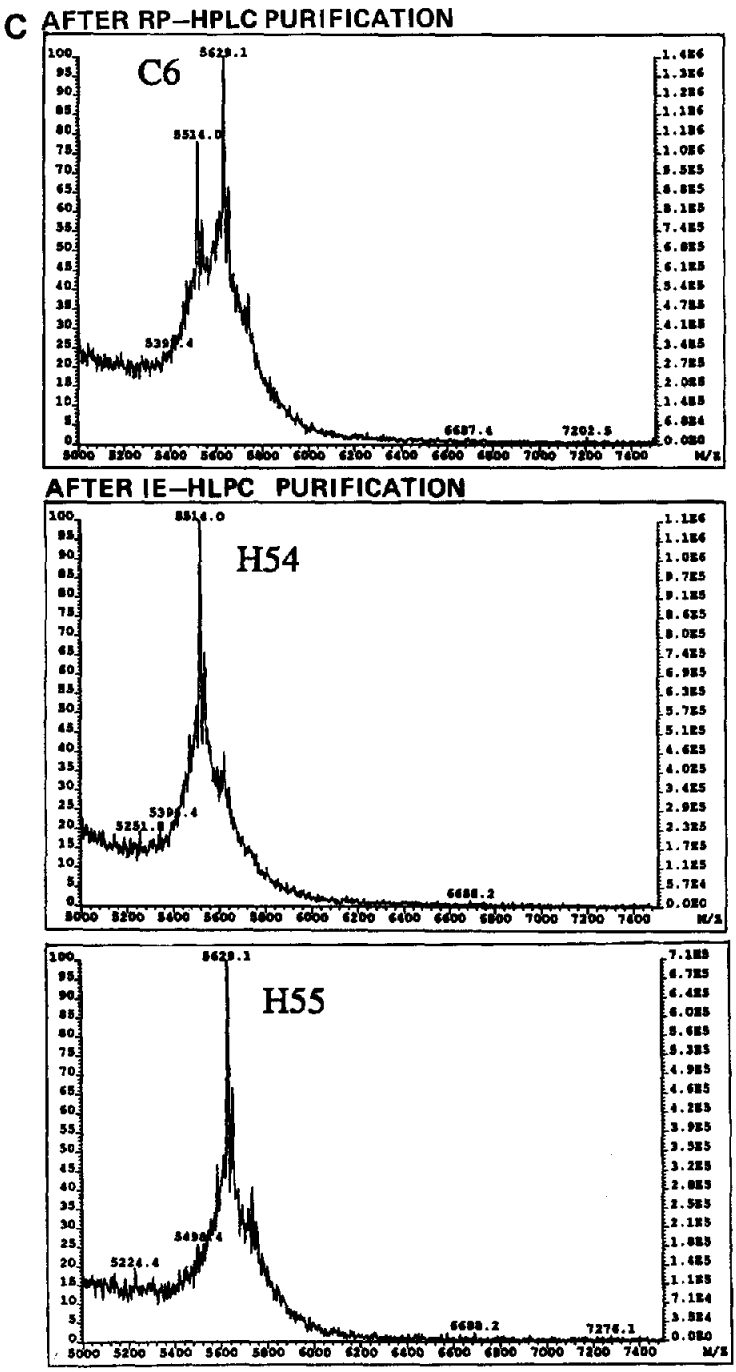

Fig. 4. LSIMS spectra of the RP-HPLC collected peptides [upper spectra, C3 on (a), C4 on (b) and C6 on (c)] and the corresponding pure peptides after IE-HPLC purification (lower spectra).

A common feature of all the spectra was a broad asymmetric response, several hundred daltons wide, extending above and below the singly charged molecular ion. This broad signal was dominated by a narrow peak corresponding to the average molecular mass of the single charged molecule. The reason for this broadness is uncertain [23], particularly concerning the ions occurring at higher masses than the molecular ion. Adducts with the liquid matrix, or the presence of impurities in the sample, may contribute.

We also checked the sensitivity of mass determination by successive dilution of the 5 $\mu g/\mu l$ solution of rHirudin. A reliable measurement was still possible down to 500 ng (72 pmol) introduced into the mass spectrometer (Fig. 3). In this case, a signal-to-noise ratio of 6 was estimated on the MH^+ ion, which was found at 6907.4 a.m.u. The measurement was rapid, with a typical total time to acquire and process the data of 5–10 min per sample.

This preliminary study showed the potential of LSIMS as a tool for precise mass measurements of the six rHirudin-truncated forms prepared by CPY digestion (C1 to C6). The results obtained are presented in Table II and Fig. 4 (upper panels). Indeed, we found that C1, C2 and C5 each corresponded to a unique polypeptide, whereas the mass spectra of C3 (Fig. 4a), C4 (Fig. 4b), and C6 (Fig. 4c) clearly demonstrated that each was a mixture of two compounds. Consequently, the separation of these polypeptides was not achieved by RP-HPLC (Fig. 5. upper panel).

Further purification was undertaken by IE-HPLC, which allowed good separation of the nine polypeptides with the exception of rHirudin and the 1–64 truncated form (Fig. 5, lower panel). In addition, mass spectra confirmed the

TABLE II

LSIMS MOLECULAR MASSES OF THE NINE C-TERMINAL-TRUNCATED FORMS OF rHIRUDIN

RP-HPLC-purified peptides	Measured MH^+	Corresponding truncated form	Theoretical value (MH^+)	Error (a.m.u.)	Amino acid loss[a]
H65	6907.2	1–65	6907.5	−0.3	
C1	6779.4	1–64	6779.4	+0.0	Gln
C2	6666.9	1–63	6666.3	+0.6	Leu
C3	6503.5	1–62	6503.1	+0.4	Tyr
	6374.0	1–61	6373.9	+0.1	Glu
C4	6034.1	1–58	6034.3	−0.2	Glu, Pro, Ile
	5905.0	1–57	5905.2	−0.2	Glu
C5	5776.5	1–56	5776.1	+0.4	Glu
C6	5629.1	1–55	5628.9	+0.2	Phe
	5514.0	1–54	5513.8	+0.2	Asp

[a] Amino acid released from the previous peptide.

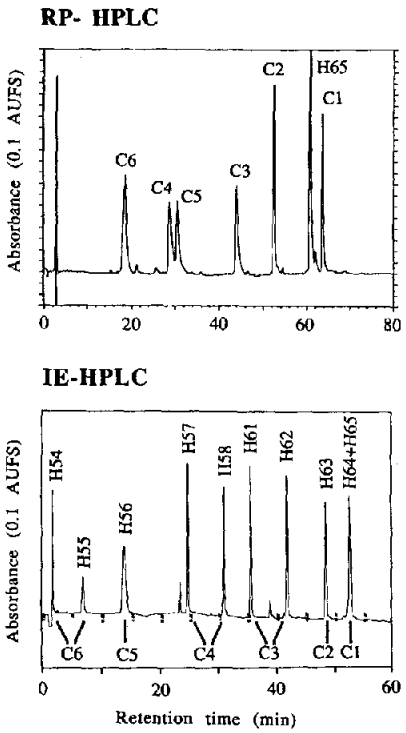

Fig. 5. RP- and IE-HPLC separations of the nine C-terminal-truncated reference forms of rHirudin produced by CPY cleavage. For chromatographic conditions, see Experimental.

purity of the new samples obtained (Fig. 4, lower panels). Thus, their exact structures were readily determined.

The combined use of two different chromatographic purification steps allowed the preparation of nine C-terminal-truncated derivatives of rHirudin. LSIMS characterized these forms as H1–64 to H1–61 and H–58 to H1–54 derivatives. This identification of the truncated peptides by LSIMS analyses offered several advantages over the traditional methods based on analysis of the released amino acid residues. In particular, the time and sample consumption were considerably reduced: the sensitivity was in the picomole range, and it was possible to analyse six to eight samples within 1 h. The accuracy of the molecular mass determination was excellent ($\leq$0.6 a.m.u.) leading to the unambigous identification of the analysed peptide. LSIMS analysis allowed differentiation and mass characterization of peptides not easily resolved by RP-HPLC and differing only by one amino acid residue.

Two C-terminal-truncated forms, H1–60 and H1–59, were not recovered during CPY processing of rHirudin. This is probably due to a very rapid loss of both Pro and Ile residues, leading to the rapid appearance of the consequent H1–58 peptide. This observation is consistent with reported data describing Pro and Ile

as rapidly released amino acid residues under CPY action [24], and was also reported in the recent work of Klarskov *et al.* [21].

CONCLUSIONS

Analysis by LSIMS was essential to identify and to test the purity of chromatographically isolated C-terminal-truncated standards of rHirudin produced by enzymic hydrolysis with CPY. LSIMS was rapid, precise, and sensitive, and it allowed good characterization of each compound. Certainly, it must be considered a valuable tool in the characterization of isolated recombinant hirudin derivatives of various origins (*i.e.* degradation by-products in biotechnology processes, biotransformation derivatives in drug metabolism studies) and may also give useful complementary information concerning the homogeneity of proteins purified by chromatographic means.

REFERENCES

1 G. Loison, A. Findeli, S. Bernard, M. Nguyen-Juilleret, M. Marquet, N. Riehl-Bellon, D. Carvallo, L. Guerra-Santos, S. W. Brown, M. Courtney, C. Roitsch and Y. Lemoine, *Biotechnology*, 6 (1988) 72.
2 F. Markwardt, *Methods in Enzymology*, Academic Press, Orlando, FL, 1970, p. 924.
3 J. W. Fenton, *Ann. N. Y. Acad. Sci.*, 370 (1981) 468.
4 S. R. Stone and J. Hofsteenge, *Biochemistry*, 25 (1986) 4622.
5 M. Courtney, G. Loison, Y. Lemoine, N. Riehl-Bellon, E. Degryse, S. W. Brown, J. P. Cazenave, G. Delreyn, D. Delebassee, A. Bernat, J. P. Maffrand and C. Roitsch, *Semin. Thromb. Hemostasis*, 15 (1989) 288.
6 R. P. Harvey, E. Degryse, L. Stefani, J. P. Cazenave, M. Courtney, P. Tolstochev and J. P. Lecocq, *Proc. Natl. Acad. Sci. U.S.A.*, 83 (1986) 1084.
7 J. Dodt, W. Machleidt, U. Seemüller, R. Maschler and H. Fritz, *Biol. Chem. Hoppe-Seyler*, 367 (1986) 803.
8 S. J. T. Mao, M. T. Yates, D. T. Blankenchip, A. D. Cardin, J. L. Krtenansky, W. Lovenberg and R. L. Jackson, *Anal. Biochem.*, 161 (1987) 514.
9 D. Tripier, *Thromb. Res.*, Suppl. 7 (1987) 514.
10 S. J. T. Mao, M. T. Yates, T. J. Owen and J. L. Krtenansky, *Biolchemistry*, 27 (1988) 8170.
11 J. Y. Chang, *FEBS Lett.*, 164 (1983) 307.
12 A. Henschen, F. Markwardt and P. Walsmann, *Folia Haematol.*, 115 (1988) 59.
13 B. A. Bidlingmeyer, S. A. Cohen and T. L. Tarvin, *J. Chromatogr.*, 336 (1984) 93.
14 G. E. Tarr, in J. E. Shively (Editor), *Methods of Protein Microcharacterization, A Practical Handbook*, Humana Press, Clifton, NJ, 1986, p. 155.
15 A. Van Dorssclaer, P. Lepage, F. Bitsch, O. Whitechurch, N. Riehl-Bellon, D. Fraisse, B. Green and C. Roitsch, *Biochemistry*, 28 (1989) 2949.
16 J. A. Yergey, R. J. Cotter, D. Heller and C. Fenselau, *Anal. Chem.*, 56 (1984) 2236.
17 B. N. Green and R. S. Bordoli, in S. J. Gaskell (Editor), *Mass Spectrometry in Biomedical Research*, Wiley, Chichester, 1986, p. 235.
18 R. J. Cotter, B. S. Larsen, D. N. Heller, J. E. Campana and C. Fenselau, *Anal. Chem.*, 57 (1985) 1479.
19 B. Larsen, in C. N. McEwen and B. S. Larsen (Editors), *Mass Spectrometry of Biological Materials*, Marcel Dekker, New York, 1990, p. 197.
20 K. Biemann, in C. N. McEwan and B. S. Larsen (Editors), *Mass Spectrometry of Biological Materials*, Marcel Dekker, New York, 1990, p. 3.

21 K. Klarskov, K. Breddam and P. Roepstorff, *Anal. Biochem.*, 180 (1989) 28.
22 W. J. Richter, F. Raschdorf and W. Maerki, *Mass Spectrom. Health Life Sci.*, 24 (1985) 193.
23 M. Barber and B. N. Green, *Rapid. Commun. Mass Spectrom.*, 1 (1987) 80.
24 B. N. Jones, in J. E. Shively (Editor), *Methods of Protein Microcharacterization, A Practical Handbook*, Humana Press, Clifton, NJ, 1986, p. 337.