Amino acid sequence of a protease inhibitor isolated from *Sarcophaga bullata* determined by mass spectrometry

I.A. PAPAYANNOPOULOS AND K. BIEMANN

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(RECEIVED July 30, 1991; ACCEPTED September 11, 1991)

Abstract

The amino acid sequence of a protease inhibitor isolated from the hemolymph of *Sarcophaga bullata* larvae was determined by tandem mass spectrometry. Homology considerations with respect to other protease inhibitors with known primary structures assisted in the choice of the procedure followed in the sequence determination and in the alignment of the various peptides obtained from specific chemical cleavage at cysteines and enzyme digests of the *S. bullata* protease inhibitor. The resulting sequence of 57 residues is as follows: Val Asp Lys Ser Ala Cys Leu Gln Pro Lys Glu Val Gly Pro Cys Arg Lys Ser Asp Phe Val Phe Phe Tyr Asn Ala Asp Thr Lys Ala Cys Glu Glu Phe Leu Tyr Gly Gly Cys Arg Gly Asn Asp Asn Arg Phe Asn Thr Lys Glu Glu Cys Glu Lys Leu Cys Leu.

Keywords: amino acid sequence; cysteine cleavage; cysteine cyanylation; laser desorption mass spectrometry; protease inhibitor; *Sarcophaga bullata*; tandem mass spectrometry

The immune system of insects and other arthropods, even though not highly developed, is capable of efficient identification of foreign materials and microorganisms and of appropriate response to their intrusion. Such immune system response often involves engulfing the foreign object in melanin, which is produced biosynthetically by the action of the enzyme phenoloxidase. In most insects phenoloxidase is present as the inactive proenzyme (prophenoloxidase) (Soderhall, 1982; Ratcliffe et al., 1984), and it is activated by a protein found in the hemolymph and the cuticle (Ashida, 1971; Ashida & Dohke, 1980). The activating protein has been found to be a serine protease (Dohke, 1973; Ashida et al., 1974) and appears to activate the prophenoloxidase by limited proteolysis, which is, therefore, irreversible (Ashida & Dohke, 1980). Prophenoloxidase is present in the hemolymph along with its phenolic substrates as well as the activating protease (Brunet, 1980). Accidental phenoloxidase activation and melanization is avoided under normal physiological conditions by the presence of large amounts of serine protease inhibitors in the hemolymph, which, by inhibiting the serine protease, prevent prophenoloxidase activation.

One such protease inhibitor, which was isolated from

the hemolymph of *Sarcophaga bullata* larvae, appears to inhibit chymotrypsin-mediated prophenoloxidase activation (Sugumaran et al., 1985). A similar inhibitor from the tobacco hornworm (*Manduca sexta*) inhibits the prophenoloxidase activation caused by an endogenous protease from the same organism (Saul & Sugumaran, 1986). There is a certain degree of homology among the known amino acid sequences of various protease inhibitors that have been isolated from different species, such as bovine pancreatic trypsin inhibitor (BPTI, also known as aprotinin), and protease inhibitors from the tobacco hornworm (*M. sexta*), bovine colostrum, turtle egg white, Russel viper venom, sea anemone, and snail isoinhibitor K (Ramesh et al., 1988). All these protease inhibitors have identically spaced cysteines, usually six, along a peptide chain of varying length, as generally indicated below (with the number of amino acids between consecutive cysteines in parentheses):

It was expected, therefore, that the *S. bullata* protease inhibitor (SBPI) would also exhibit a similar degree of homology. It was decided that cleavage at the N-terminus of cysteines after reduction of disulfides and cyanylation

Reprint requests to: K. Biemann, Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

Scheme 1. Reduction (a), cyanylation (b), and cleavage (c) of a disulfide-containing peptide (Stark, 1977).

(Stark, 1977), as shown in Scheme 1, would provide information on the structure and show whether SBPI were homologous with the other protease inhibitors as anticipated. It should be noted that this method of specific protein cleavage has not been extensively utilized because it produces peptides with blocked N-termini that cannot, therefore, be sequenced by the Edman method. However, tandem mass spectrometry readily provides sequence information for peptides with modifications such as a blocked N-terminus (Biemann, 1990). The primary structure of purified (Sugumaran et al., 1985) SBPI was determined by mass spectrometry and is reported herein.

Results and discussion

The sequence of SBPI was determined by tandem mass spectrometry of peptides produced by specific cleavage of the protein with chemical and enzymatic methods. The high cysteine content and the potential homology of SBPI with other protease inhibitors suggested the use of 2-nitro-5-thiocyanobenzoic acid (NTCB), which cleaves at the N-terminus of the cysteine residues (Stark, 1977), as illustrated in Scheme 1. This method had been used successfully with *Escherichia coli* thioredoxin to cleave at its active site (...-Cys³²-Gly-Pro-Cys³⁵-...) and de-

Fig. 1. High performance liquid chromatography of peptide mixture obtained from SBPI by S-cyanylation with NTCB and cleavage in 0.1 M Na₃BO₃. The markings indicate the fractions collected for analysis by mass spectrometry.

termine by fast atom bombardment mass spectrometry (FABMS) the molecular weights of the peptides thus obtained (Biemann et al., 1986).

Therefore, SBPI was subjected to NTCB cleavage, the products were partially fractionated by high performance liquid chromatography (HPLC) (Fig. 1), and the molecular weights of the peptides in each fraction were determined by FABMS (Table 1, Fig. 2). The amino acid sequences for three of these peptides were subsequently determined from their collision-induced dissociation (CID) mass spectra (Table 1; Figs. 3, 4). It is of interest to note that the CID mass spectra of some of the protonated

Table 1. (M+H)⁺ ions of peptides obtained from cleavage of SBPI at cysteines after cyanylation with NTCB^a

No.	(M+H) ⁺ m/z	HPLC fraction	Sequence	Position
1	1,607.9	2		39-51
2	519.2	3		1-5
3	995.5	3	CLQPKEVGP	6-14
4	1,495.8	3		1-14
5	758.4	4		52-57
6	942.5	4	CEEFLYGG	31-38
7	2,533.2	4		31-51
8	733.4	5	CEKLCL	52-57
9	1,937.0	5		15-30
10	2,838.3	6		15-38

^a High performance liquid chromatographic fractions indicated are those marked on the chromatogram shown in Figure 1. The positions correspond to the final structure (Fig. 10), and where the amino acid sequence is shown it was determined by CIDMS. The second cysteine in the peptide of (M+H)⁺ m/z 733.4 is free whereas it is presumed S-cyanylated in the peptide of (M+H)⁺ m/z 758.4.

Fig. 2. Partial FAB mass spectrum of HPLC fraction 4 of the NTCB digest of SBPI indicating the presence of two peptides with $(M+H)^+$ m/z 758.4 and 942.5 (see Table 1). Ions marked **m** are $(\text{Iglycerol})_n + H)^+$ matrix clusters.

peptides bearing the 2-iminothiazolidine-5-carboxylic acid moiety are relatively simple and indicate fragmentation predicted for a peptide with a weakly basic N-terminus (Biemann, 1988; Johnson et al., 1988). The abundant

ions are of type a_n and b_n (see Scheme 2) encompassing the peptide N-terminus (Fig. 3), whereas C-terminal ions are only formed when a basic amino acid is retained in the fragment: Figure 4 shows y_4 and w_5 , but no y_{1-3} , in addition to a complete b_{1-6} series because of the lysine at position 3. It is noteworthy that these spectra exhibit significant b_1 ions that are generally observed only for N_α -acylated peptides (Biemann, 1990). The CID mass spectra of some of the other protonated peptide molecules were difficult to interpret for reasons that will be discussed later.

The sequences of the peptides of $(M+H)^+$ m/z 733.3, 942.5, and 995.5 (8, 6, and 3 in Table 1) indicated that there are three pairs of cysteines that are separated by 3, 7, and 8 amino acids, respectively. This spacing matches that of all other known protein inhibitors of this class and suggests that the spacing of all cysteines is conserved in SBPI. With this preliminary hypothesis and considering the molecular weights of the peptides listed in Table 1, it is possible to assemble the following general structure:

X_n , X_{15} , and X_{12} indicate sequences of n (as yet undetermined), 15, and 12 amino acids, respectively. Both X_{15}

Fig. 3. The CID mass spectrum (utilizing all four sectors of a high performance tandem mass spectrometer) of $(M+H)^+$ m/z 942.5 (see Fig. 2). Ions C^* , L, E, F, and Y are the immonium ions (Scheme 2) of cyclized cysteine (2-iminothiazolidine-4-carboxylic acid), leucine, glutamic acid, phenylalanine, and tyrosine, respectively. Ions $-\text{Y}$, $-\text{F}$, and $-\text{L}$ are due to loss of the side chain of the respective ion from the protonated peptide molecule. The N-terminal cysteine, which has been converted to 2-iminothiazolidine-4-carboxylic acid, is indicated schematically by the three corner symbol. The fragment ion nomenclature (Biemann, 1988) is shown in Scheme 2. The collision cell potential was 0 kV.

Theoretically, the CID mass spectra of all peptides

Fig. 5. Mass spectra of the intact SBPI molecule (a) by matrix-assisted LD-TOF, sum of 40 laser shots, at a resolution of 1:300; and (b) by FAB, sum of nine 8-s scans over 250 u (100 Hz filter), at a resolution of 1:1,000.

generated by two separate complete digests of different specificities should produce the entire amino acid sequence of a protein, particularly one as small as that under investigation here. However, rather than optimizing the conditions of the cysteine cleavage of this particular protein, we decided to use other enzymatic digestions instead. Although we are presently investigating the best conditions for complete cleavage of cysteine-rich proteins, it is already clear that this reaction is at least as useful as the cyanogen bromide (CNBr) cleavage of methionine to produce relatively large fragments from proteins (Zaia et al., 1991) for further enzymatic digestion because cysteine is generally a rare amino acid.

Having obtained some sequence data and valuable mo-

lecular weight information (see partial structure [I] above), the protein was cleaved with trypsin. The m/z values of 13 $(M+H)^+$ ions (11–23, Table 2) were measured. The CID mass spectra of the major components (11, 13, 16, 21, and 22) revealed the sequences shown in Table 2. Tryptic peptide 21 (ACEEFLYGGCR, Fig. 6) completely overlaps the entire sequence of the NTCB-produced peptide 6 (CEEFLYGG) and extends it one amino acid before its N-terminus and two amino acids past its C-terminus; the position of the second cysteine is in agreement with the original assumption as to the spacing of the cysteines and is consistent with the specificity with the NTCB cleavage reaction.

The sequences of these tryptic peptides combined with

Fig. 6. The CID mass spectrum of $(M+H)^+$ m/z 1,361.6. The peptide was obtained from a tryptic digest of SBPI after partial fractionation with HPLC. The cysteines have been carbamidomethylated. The abundance of C-terminal fragment ions (v_n , w_n , x_n , y_n , and z_n) is consistent with the presence of the basic arginine at the C-terminus of this peptide. The collision cell potential was 3 kV.

Table 2. $(M+H)^+$ ions of peptides obtained from digestion of SBPI with trypsin^a

No.	$(M+H)^+$ m/z	Sequence	Position
11	405.2	LCL	55-57
12	711.3		24-29
13	717.4	EVGPCR	11-16
14	761.4	SDFVFF	18-23
15	803.4	SACLQPK	4-10
16	889.5	KSDFVFF	17-23
17	924.4		18-24
18	1,080.5		50-57
19	1,145.6		1-10
20	1,184.6		46-54
21	1,361.6	ACEEFLYGGR	30-40
22	1,581.8	---FVFFYNADTK	17-29
23	2,128.0		41-57

^a Cysteines were reduced and alkylated with iodoacetamide. Positions are those corresponding to the final structure (Fig. 10), and where the amino acid sequence is shown it was determined by CIDMS.

the molecular weight data obtained from the NTCB cleavage (Table 1) permitted extension of the partial structure [I]: X_{15} must be RKSDFFVFFYNADTKA. Furthermore, the C-terminal LCL was confirmed, the first amino acid of X_{12} was determined to be R, and the two C-terminal amino acids of X_n were determined to be SA

(compare the overlapping sequences of peptides 8 and 11, peptides 6 and 21, and peptides 3 and 15, respectively).

Again, the sum of molecular weights of all tryptic peptides exceed the experimentally determined molecular weight of SBPI, indicating that some of the minor components, which were too low in abundance to produce good CID mass spectra, must be due to incomplete cleavage at Arg or Lys. In such a case one could repeat the tryptic digest for a longer time, use a higher enzyme-to-substrate ratio, or add fresh enzyme after a certain period of time. Alternatively, an enzyme of different specificity could be used. Although the latter option includes the risk of another unoptimized digest we prefer it, because the resulting peptides will provide different rather than redundant sequence information. In addition, the CID mass spectra of peptides from a different digest, which lack the C-terminal basic amino acids (Lys or Arg) of tryptic peptides, exhibit different types of fragment ions (Biemann, 1988).

For these reasons, chymotrypsin was chosen to digest another aliquot of reduced and carbamidomethylated SBPI. From the molecular weight information (Table 3) it is possible to match, using the computer program COMPOST (Papayannopoulos & Biemann, 1991) peptides 24, 25, 26, 29, and 30 to already known parts of the sequence, leaving peptides 27 and 28 $[M+H]^+$ m/z 792.4 and 1,150.6) as the best candidates for providing additional useful sequence information. The CID mass spectra of these, along with a few others, were measured and

Fig. 7. The CID mass spectrum of $(M+H)^+$ m/z 792.4. The peptide was obtained from a chymotryptic digest of reduced and carbamidomethylated SBPI after partial fractionation by HPLC. The collision cell potential was 3 kV.

Table 3. $(M+H)^+$ ions of peptides obtained from digestion of SBPI with chymotrypsin^a

No.	$(M+H)^+$ m/z	Sequence	Position
24	412.2	VFF	21–23
25	496.2		17–20
26	575.3		21–24
27	792.4	VDKSACL	1–7
28	1,150.6	NTKEECEKL	47–55
29	1,297.7		25–35
30	1,347.7	YNADTKACEEF	24–34

^a Cysteines were reduced and alkylated with iodoacetamide. Positions are those corresponding to the final structure (Fig. 10), and where the amino acid sequence is shown it was determined by CIDMS.

yielded the sequences shown in Table 3. These data revealed the N-terminus X_n (Fig. 7) and the five C-terminal amino acids of X_{12} of the partial structure [I], but still left a gap of probably seven amino acids, representing the N-terminal part of X_{12} . Another aliquot of reduced but not alkylated SBPI was digested with chymotrypsin resulting in peptides 31, 32, 33, 34, and 35 (Table 4).¹

¹ It is of interest to note that SBPI, which in its native form inhibits chymotrypsin (Sugumaran et al., 1985), loses activity upon reduction and may be cleaved with this enzyme.

Table 4. $(M+H)^+$ ions of peptides obtained from digestion of SBPI with chymotrypsin, after reduction of the cysteines but without any further modification^a

No.	$(M+H)^+$ m/z	Sequence	Position
31	735.4	VDKSACL	1–7
32	1,093.5	NTKEECEKL	47–55
33	1,095.5	GGCRGNDNRF	37–46
34	1,290.6		24–34
35	1,589.8		8–21

^a Positions are those corresponding to the final structure (Fig. 10), and where the amino acid sequence is shown it was determined by CIDMS.

Of these, all but peptide 33 could be assigned to already known portions of the sequence. The CID mass spectrum of peptide 33 ($[M+H]^+$ m/z 1,095.5) was, therefore, measured, along with others in the same fraction. From this spectrum (Fig. 8) the sequence GGCRGNDNRF was deduced. Its N-terminus overlaps the C-terminal amino acids of peptides 6 and 21; however, the C-terminus of this peptide did not overlap with any of the hitherto determined sequence. Thus the question arises whether it completes the SBPI sequence, immediately preceding peptide 32 (and its carbamidomethylated analogue 28), or whether there is a section missing. There is abundant

Fig. 8. The CID mass spectrum of $(M+H)^+$ m/z 1,095.5. The peptide was obtained from a chymotryptic digest of reduced but not alkylated SBPI after partial fractionation by HPLC. The spectrum was recorded with a microchannel array detector, and the collision cell potential was 6 kV.

evidence for the former: not only does the molecular weight of SBPI not allow additional amino acids, but the molecular weights of peptides 1, 7, 20, and 23 now turn out to overlap in this region.

Amino acid analysis of the intact SBPI had indicated the presence of only leucine and no isoleucine. This was corroborated by some of the CID mass spectra (e.g., Figs. 6, 9), which exhibited amino acid side chain-related fragmentations indicative of leucine (Johnson et al., 1987, 1988) as well as the results from the digestions with chymotrypsin, which cleaves at the C-terminus of leucine but not isoleucine.

The primary structure of SBPI deduced from the data discussed above is shown in Figure 10. The position of each amino acid is defined by at least one CID mass spectrum (heavy underlining) and is further supported by the molecular weights of the larger overlapping peptides (light underlining). Furthermore, the SBPI amino acid sequence thus determined corresponds to an average molecular weight of 6,512.3 Da; this is in excellent agreement with that determined experimentally by FAB (two different measurements: 6,512.3 and 6,512.5 Da) and matrix-assisted laser desorption time-of-flight (LD-TOF) mass spectrometry (6,512.6 Da). The latter method was also used to measure the molecular weight of intact SBPI after reduction and alkylation of the cysteines with iodoacetamide. The result obtained (6,860.3 Da) corresponds

to addition of six carbamidomethyl groups to SBPI and represents entirely independent evidence for the presence of six cysteines.

A few comments concerning certain features of the CID mass spectra shown in Figures 3, 4, 7, and 8 may be in order. By coincidence the residue mass ([molecular weight] - 18) of the cyclized cysteine derivative is 128 Da, the same as that of lysine and glutamine. This potential ambiguity is easily resolved by the fact that peptides produced by cyanylation of cysteine must have a 2-iminothiazolidine-4-carboxyl group at the N-terminus. The exception would be the N-terminal peptide of the protein. N-terminal glutamine would easily cyclize to pyroglutamic acid (residue mass 111), but if that were not the case, in the CID mass spectrum, N-terminal fragment ions would be accompanied by peaks 17 u lower due to loss of NH_3 (Baldwin et al., 1990). The presence of N-terminal lysine in the peptide whose CID mass spectrum is shown in Figure 3 can be eliminated by the absence of an abundant ion at m/z 84. This argument could not be used for Figure 4, because this peptide contains an internal lysine. However, the abundant molecular weight information of peptides containing that cysteine and the CID mass spectra of peptides 28 and 31 clearly demonstrate the presence of cysteine at position 52. If there were any remaining doubt, acetylation of the peptide in question could remove it. The presence of N-terminal ly-

Fig. 9. The CID mass spectrum of $(M+H)^+$ m/z 803.4. The peptide was obtained from a tryptic digest of SBPI after partial fractionation by HPLC. Ions labeled with an asterisk (*) are due to the glycerol matrix. The cysteine present has been carbamidomethylated. The w_4 ion shows that the fourth amino acid is leucine rather than isoleucine (Johnson et al., 1987). The collision cell potential was 3 kV.

Fig. 10. The sequence of SBPI as determined by tandem mass spectrometry. Heavy underlining (—) indicates peptides sequenced by CIDMS, and light underlining (—) indicates peptides of which the molecular weight was determined by FABMS. The type of digest that resulted in the production of the underlined peptides is indicated as follows: A, NTCB; B, trypsin; C, chymotrypsin (cysteines reduced and carbamidomethylated); D, chymotrypsin (cysteines reduced but not otherwise modified). The six cysteines and all amino acids present in the sequence of other protease inhibitors (Ramesh et al., 1988) at the same position with respect to the cysteines appear in boxes; Asn-25 and Gly-41 occur in these positions relative to the cysteines in more than half of the sequences of the protease inhibitors.

sine would add 84 Da to the molecular weight of the peptide (addition of two acetyl groups), whereas the cyclized cysteine would be unreactive.

For the chymotryptic peptide that covers the SBPI N-terminus (Table 3, No. 27), the only ion in its CID mass spectrum (Fig. 7) that corresponds to simple backbone cleavage between the first and second amino acids is z_6 , indicating N-terminal Val-Asp rather than Asp-Val. The former arrangement is confirmed by the abundant internal fragment ions (see Scheme 2) DK and DKS.

A similar situation is represented by the spectrum shown in Figure 8. Because the residue mass of Gly-Gly equals that of Asn, the latter could be the N-terminal amino acid, followed by Cys. If that were the case, the ion at m/z 115, labeled b_2 , would have to be b_1 ; this is unlikely because in our experience the N-terminal amino acyl ion is formed only when the N-terminus is acylated or a cyclized cysteine (Figs. 3, 4), but not for Asn.

Conclusion

At the outset we had based partial structure [I] on the hypothesis that SBPI would have the same distribution of cysteines as the other members of this family of protease inhibitors. However, the final sequence shown in Figure 10 is entirely independent of this original assumption and based exclusively on the molecular weights of the products of chemical and enzymatic cleavage of SBPI and the CID mass spectra of a selected set of these peptides.

Materials and methods

Fast atom bombardment mass spectrometry

Fast atom bombardment mass spectrometry of the peptides (Barber et al., 1981; Martin et al., 1982) from the enzymatic and chemical cleavage reactions was carried out in the first (MS-1) of two mass spectrometers of the JEOL HX110/HX110 tandem high resolution mass spectrometer (JEOL Ltd., Akishima, Japan) at an accelerating voltage of 10 kV and a resolution of 1:2,500. Using about 20% of each HPLC fraction (see below), single scans were acquired at a rate to scan from 300 to 4,000 in about 1.5 min and with 100–300 Hz filtering. For calibration, $(CsI)_nCs^+$ cluster ions were used. A JEOL FAB gun, operated at 6 kV with xenon as the FAB gas, or an ANTEK Cs^+ gun, operated at 7 kV, was used.

Tandem mass spectrometry

The remainder (approximately 70%) of each HPLC fraction was used for tandem mass spectrometry, carried out with all four sectors of the JEOL HX110/HX110, an instrument of $E_1B_1E_2B_2$ configuration. Collision-induced dissociation of protonated peptide molecules took place

in the field-free region after B_1 , thus operating both MS-1 (E_1B_1) and MS-2 (E_2B_2) as double-focusing instruments. The collision gas, helium, was introduced into the collision chamber at a pressure sufficient to reduce the precursor ion signal by 65–70%. The CID mass spectra were recorded either with a conventional electron multiplier with 100 Hz filtering and a scan rate that was the same as that used for FABMS, or with a microchannel array detector (Hill et al., 1989). The resolution of MS-1 was adjusted so that only the ^{12}C species of the protonated molecule to be analyzed was transmitted into the collision cell. The MS-2 was operated at a resolution of 1:1,000, and was calibrated with a mixture of CsI, NaI, KI, RbI, and LiCl (Sato et al., 1987).

Matrix-assisted LD-TOF mass spectrometry

Matrix-assisted LD-TOF mass spectra (Karas & Hillenkamp, 1988) of 1–2 pmol of protein were acquired with a VT2000 TOF mass spectrometer (VESTEC Corp., Houston, Texas), operating at 30 kV accelerating voltage and equipped with a Lumonics HY400 Nd:YAG laser (Biemann et al., 1991; Juhasz et al., 1991). Sample ionization was achieved by 355 nm irradiation, with 8-ns laser pulses at a repetition rate of 5 Hz. Sinapinic acid (3,5-dimethoxy-4-hydroxy cinnamic acid) was used as the matrix (Beavis & Chait, 1989).

Chemical cleavage of SBPI

2-Nitro-5-thiocyanobenzoic acid (Sigma Chemical Co.) was used to cyanylate the cysteines generated after reduction of cystines with 1,4-dithio-D,L-threitol (Pierce Chemical Co.). Both reduction and cyanylation of 4 nmol SBPI were carried out in 0.3 mL denaturing buffer (6 M guanidinium chloride, 1 mM EDTA, 0.2 M Tris-acetate, pH 8) (Stark, 1977; Allen, 1981). The excess reagents were removed by dialysis against 4 L 25% acetic acid using a 3,000-molecular weight cutoff dialysis membrane. The sample was dried in vacuum and subsequently dissolved in 0.3 mL of 0.1 M sodium borate (pH 9). Cleavage of the cyanylated SBPI, with cyclization of cysteine to 2-iminothiazolidine-4-carboxylic acid, was accomplished after incubation at 37 °C for 16 h.

Enzymatic digestion of SBPI

Sarcophaga bullata protease inhibitor was digested with trypsin and chymotrypsin (Allen, 1981). Three to four nanomoles of SBPI were used for each digest; the cysteines were reduced with dithiothreitol and alkylated with iodoacetamide (Fluka Chemie AG); a second chymotryptic digest was carried out with the SBPI cysteines reduced but not alkylated.

Digestions with enzymes (2% by weight of SBPI) took place in 0.3 mL of 0.1 M ammonium bicarbonate (pH

7.5), at 37 °C, for 2 h with trypsin or for 6 h with chymotrypsin. The enzymes were obtained from Boehringer Mannheim GmbH.

High performance liquid chromatography fractionation

Partial fractionation of the peptides obtained from the chemical or enzymatic cleavage of SBPI, was carried out by reversed-phase analytical HPLC (C_{18} or C_4 column). The gradient was 100% solvent A (water containing 0.05% trifluoroacetic acid)/0% solvent B (acetonitrile containing 0.035% trifluoroacetic acid) to 35% solvent A/65% solvent B, in 36 min at a flow rate of 1.5 mL/min, with detection of peptides at 214 nm. Each fraction was prepared for FAB mass spectrometric analysis by addition of 5–6 μ L of 50% glycerol in water and removal of the HPLC solvents by lyophilization; care was taken to avoid evaporation of the glycerol.

Acknowledgments

The authors thank Professor M. Sugumaran and Dr. S.J. Saul (Department of Biology, University of Massachusetts/Harbor Campus, Boston, Massachusetts) for providing the purified SBPI sample. This work was supported in part by grants from the National Institutes of Health (RR00317 and GM05472).

References

- Allen, G. (1981). Sequencing of proteins and peptides. In *Laboratory Techniques in Biochemistry and Molecular Biology* (Work, T.S. & Burdon, R.H., Eds.). Elsevier/North Holland Biomedical Press, The Netherlands.
- Ashida, M. (1971). Purification and characterization of pre-phenoloxidase from the hemolymph of the silkworm, *Bombyx mori*. *Arch. Biochem. Biophys.* 144, 749–762.
- Ashida, M. & Dohke, K. (1980). Activation of pre-phenol oxidase by the activating enzyme of the silkworm, *Bombyx mori*. *Insect Biochem.* 10, 37–47.
- Ashida, M., Dohke, K., & Ohnishi, E. (1974). Activation of prophenol oxidase. Release of a peptide from prophenol oxidase by the activating enzyme. *Biochem. Biophys. Res. Commun.* 57, 1089–1095.
- Baldwin, M.A., Falick, A.M., Gibson, B.W., Prusiner, S.B., Stahl, N., & Burlingame, A.N. (1990). Tandem mass spectrometry of peptides with N-terminal glutamine. Studies on a prion protein peptide. *J. Am. Soc. Mass Spectrom.* 1, 258–264.
- Barber, M., Bordoli, R.S., Sedgwick, R.D., & Tyler, A.N. (1981). Fast atom bombardment of solids (F.A.B.): A new ion source for mass spectrometry. *J. Chem. Soc. Chem. Commun.* 325–327.
- Beavis, R.C. & Chait, B.T. (1989). Matrix-assisted laser-desorption mass spectrometry using 355 nm radiation. *Rapid Commun. Mass Spectrom.* 3, 436–439.
- Biemann, K. (1988). Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* 16, 99–111.
- Biemann, K. (1990). Sequencing of peptides by tandem mass spectrometry upon high energy collision-induced dissociation. *Methods Enzymol. Mass Spectrom.* 193, 455–479.
- Biemann, K., Costello, C.E., Papayannopoulos, I.A., & Juhasz, P. (1991). Applications of matrix-assisted laser desorption time-of-flight mass spectrometry. In *Proceedings of the 39th American Society for Mass Spectrometry Conference*, pp. 939–940.
- Biemann, K., Martin, S.A., Scoble, H.A., Johnson, R.S., Papayannopoulos, I.A., Biller, J.E., & Costello, C.E. (1986). How to obtain and how to use mass spectral data at high mass. In *Mass Spectrometry in the Analysis of Large Molecules: Proceedings of the Texas Symposium on Mass Spectrometry* (McNeal, C., Ed.), pp. 131–149. John Wiley & Sons, Sussex, England.
- Brunet, P.C.J. (1980). The metabolism of aromatic amino acids concerned in the crosslinking of insect cuticle. *Insect Biochem.* 10, 467–500.
- Dohke, K. (1973). Prophenoloxidase-activating enzyme from cuticle of the silkworm *Bombyx mori*. Purification and characterization of the enzyme. *Arch. Biochem. Biophys.* 157, 203–221.
- Hill, J.A., Biller, J.E., Martin, S.A., Biemann, K., Sato, K., & Yoshidome, K. (1989). Design considerations, calibration, and applications of an array detector for a four sector tandem mass spectrometer. *Int. J. Mass Spectrom. Ion Processes* 92, 211–230.
- Johnson, R.S., Martin, S.A., & Biemann, K. (1988). Collision induced fragmentation of $(M+H)^+$ ions of peptides. Side chain specific sequence ions. *Int. J. Mass Spectrom. Ion Processes* 86, 137–154.
- Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., & Watson, J.T. (1987). Novel fragmentation processes of peptides by collision induced decomposition in a tandem mass spectrometer: Differentiation of leucine and isoleucine. *Anal. Chem.* 59, 2621–2625.
- Juhasz, P., Biller, J.E., Costello, C.E., Papayannopoulos, I.A., & Biemann, K. (1991). Evaluation of factors contributing to the accuracy and reliability of the mass assignments in matrix-assisted laser desorption mass spectrometry. In *Proceedings of the 39th Annual American Society for Mass Spectrometry Conference*, pp. 344–345.
- Karas, M. & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. *Anal. Chem.* 60, 2299–2301.
- Martin, S.A., Costello, C.E., & Biemann, K. (1982). Optimization of experimental procedures for fast atom bombardment mass spectrometry. *Anal. Chem.* 54, 2366–2368.
- Papayannopoulos, I.A. & Biemann, K. (1991). A computer program (COMPOST) for predicting mass spectrometric information from known amino acid sequences. *J. Am. Soc. Mass Spectrom.* 2, 174–177.
- Ramesh, N., Sugumaran, M., & Mole, J.E. (1988). Purification and characterization of two trypsin inhibitors from the hemolymph of *Manduca sexta* larvae. *J. Biol. Chem.* 263, 11523–11527.
- Ratcliffe, N.A., Leonard, C., & Rowley, A.F. (1984). Prophenol oxidase activation: Nonself recognition and cell cooperation in insect immunity. *Science* 226, 557–559.
- Sato, K., Asada, T., Ishihara, M., Kunihiro, F., Kammei, Y., Kubota, E., Costello, C.E., Martin, S.A., Scoble, H.A., & Biemann, K. (1987). High performance tandem mass spectrometry: Calibration and performance of linked scans of a four-sector instrument. *Anal. Chem.* 59, 1652–1659.
- Soderhall, K. (1982). Prophenoloxidase activating system and melanization—A recognition mechanism of arthropods? A review. *Dev. Comp. Immunol.* 6, 601–611.
- Stark, G.R. (1977). Cleavage at cysteine after cyanilation. *Methods Enzymol. Enzyme Struct.* 47, 129–132.
- Sugumaran, M., Saul, S.J., & Ramesh, N. (1985). Endogenous protease inhibitors prevent undesired activation of prophenolase in insect hemolymph. *Biochem. Biophys. Res. Commun.* 132, 1124–1129.
- Zaia, J., Papayannopoulos, I.A., & Biemann, K. (1991). Utility of a cysteine-specific cleavage reaction in the structure determination of proteins by laser desorption time-of-flight and tandem mass spectrometry. In *Proceedings of the 39th Annual American Society for Mass Spectrometry Conference*, pp. 1414–1415.