A New Method for Protein Sequence Analysis Using Edman-degradation, Field-desorption Mass Spectrometry and Computer Calculation. Sequence Determination of the *N*-Terminal BrCN Fragment of *Streptomyces erythraeus* Lysozyme

Yasutsugu SHIMONISHI,* Yeong-Man HONG, Itsuo KATAKUSE,† and Saburo HARA††

Institute for Protein Research, Osaka University, Suita, Osaka 565

†*Department of Physics, Faculty of Science, Osaka University, Toyonaka, Osaka 560*

††*Department of Chemistry, Faculty of Science, Osaka University, Toyonaka, Osaka 560*

(Received April 22, 1981)

A new method is described for protein sequence analysis. The principle is realized as follows: 1) determination by field-desorption mass spectrometry of the molecular weights of the constituents in peptide mixtures prepared from a polypeptide or protein by specific cleavage methods, 2) quantitation of the 3-phenyl-2-thiohydantoin derivatives of amino acids liberated successively from the constituent peptides in 1) by Edman-degradation, and 3) determination of amino acid sequences with a computer from the data obtained in 1) and 2). The method was applied to sequence determination of a polypeptide of unknown structure, the *N*-terminal BrCN fragment of *Streptomyces erythraeus* lysozyme.

The Edman method[1] is widely used for sequencing of polypeptides and proteins, making it possible to determine their partial or even almost complete sequences by cleavage and identification of the *N*-terminal amino acid residue in each degradation reaction. Recently, we[2-4] reported a new method for sequencing of peptide mixtures that cannot be achieved by Edman-degradation alone, by a combination of Edman-degradation and field-desorption mass spectrometry.[5] The principle of the method is based on the determination of the molecular weights of peptides and peptide fragments degraded by the Edman method in a mixture by the field-desorption ionization technique and calculation of possible mass differences before and after degradation. We also reported that when preparations of peptide mixtures are obtained from a polypeptide by two or more kinds of specific cleavage methods, the sequence of the polypeptide can be examined using the computer program "PROSEQ,"[6] the data for which consist of the molecular weights and partial amino acid sequences, determined by the method[2-4] described above, of constituent peptides in mixtures.

More recently, some of the authors and others[7,8] developed another method for sequencing of polypeptides and proteins. The procedure consists of the following steps: 1) determination by the field-desorption ionization technique of the molecular weights of peptides in mixtures, which are prepared from a polypeptide by two or more kinds of specific cleavage methods, and their peptide fragments obtained after one-cycle of degradation, 2) estimation of the 3-phenyl-2-thiohydantoin derivatives of amino acids released successively from peptides in mixtures, prepared in 1), by Edman-degradation, and 3) deduction of amino acid sequences from the data obtained in 1) and 2) using the computer program "PROSEQ2."[8] The program "PROSEQ2" uses the data on the 3-phenyl-2-thiohydantoin amino acids released from peptide mixtures in each cycle of Edman-degradation in place of those on partial amino acid sequences of constituent peptides used in "PROSEQ1." We named the procedure the "IPR-sequencing method." To appreciate the utility of the procedure, the sequence of a polypeptide of unknown structure was examined.

In this paper we report in detail the sequence determination of the *N*-terminal BrCN fragment of *Streptomyces erythraeus* lysozyme[9] (SE lysozyme) of unknown sequence, which consists of 55 amino acid residues with an Ala residue at the *N*-terminus,[10] by the procedure. The amino acid sequence of this polypeptide was determined from data on peptides prepared by tryptic and chymotryptic digestion.

Materials and Methods

The *N*-terminal BrCN fragment of SE lysozyme was isolated as described elsewhere.[10] Trypsin treated with L-(1-tosylamino-2-phenylethyl) chloromethyl ketone was purchased from Worthington Biochemical Corp. (USA) and chymotrypsin treated with L-(1-tosylamino-5-aminopentyl) chloromethyl ketone from Sigma Chemical Co. (USA). [Gly[1]]-ACTH-(1—18)-NH_2[11] was obtained from Shionogi Research Laboratory, Shionogi and Co., Ltd. (Osaka), by courtesy of Dr. K. Inouye. H-Val-Tyr-Ile-His-Pro-Phe-OH was synthesized by a conventional method in this laboratory. Reagents for Edman-degradation were obtained from Wako Pure Chemical Ind. Ltd. (Osaka). The solvents used were redistilled. All other reagents were reagent grade and were used without further purification.

BrCN Cleavage. A peptide was dissolved at a concentration of 1% (w/v) and in a ratio of its Met residue to BrCN of 50/1 (mol/mol) in 70% formic acid.[12] The solution was stood for 20 h at 25 °C, diluted with distilled water and lyophilized repeatedly.

Enzymic Digestion. The *N*-terminal BrCN fragment (0.7 mg (110 nmol) and 1.0 mg (160 nmol)) of SE lysozyme were digested with trypsin and chymotrypsin, respectively, at a substrate concentration of 1% (w/v) and in a ratio of substrate to enzyme at 50/1 (w/w) in 1% NH_4HCO_3 at pH 8.0 for 4 h at 37 °C, and the digests were lyophilized immediately after the reaction.

Field-desorption Mass Spectra. Field-desorption

mass spectra were measured with a second-order double focusing mass spectrometer[13] with a mono field-desorption ion source, equipped with a data processor (JEOL JMA-2000 mass data analysis system). Mass assignment was made using polypropylene glycol as a standard.[14] Silicon emitters[15] grown on a tungsten wire were used for measurements of field-desorption mass spectra. The lyophilized material of sample peptides was dissolved in a mixture of pyridine and water (1/1, v/v) and the solution (1—2 μl containing 1—2 μg of sample peptides) was loaded by the syringe technique[16] on the emitter. The conditions for measurement of field-desorption mass spectra were as described.[15]

Edman-degradation. Degradations were performed manually as described[1,4] using the following buffer solutions: a) a mixture of pyridine and water (1/1, v/v), pH 9.35; b) a mixture of pyridine and water (1/1, v/v), adjusted to pH 9.5 by adding *N*-methylmorpholine; c) a mixture of pyridine and water (1/1, v/v), adjusted to pH 9.5 by adding 30% aq trimethylamine and glacial acetic acid, and d) a mixture of pyridine and water (1/1, v/v), adjusted to pH 9.5 by adding *N,N*-dimethylallylamine and trifluoroacetic acid.[1] The resultant 3-phenyl-2-thiohydantoin derivatives of amino acids were subjected to high-performance liquid chromatography, and the remaining peptides in the water layer were subjected to field-desorption mass spectrometry.

High-performance Liquid Chromatography (HPLC). HPLC was performed on a Zorbax ODS column (4.6 mm×25 cm) (Dupont) in a high-performance liquid chromatograph (Shimadzu HPLC LC-3A equipped with a data processor chromatopac C-R1A), for the quantitation of the 3-phenyl-2-thiohydantoin derivatives of amino acids released from sample peptides by Edman-degradation. Samples were dissolved in methanol and introduced into the sample holder. The solvent employed for chromatography was a mixture of CH_3CN and 0.01 M NaOAc (pH 4.5) (42/58, v/v).[17] Chromatography was performed at 62 °C at a flow rate of 1.0 ml/min for samples from the organic phase and at 2.0 ml/min for those from the aqueous phase.

Computer Calculation. The computer programs "PROSEQ1"[6] and "PROSEQ2"[8] were used for deducing the major portions of the sequences or the complete sequences of the polypeptide from the data on the peptides obtained by digestions of the polypeptide with trypsin and chymotrypsin. Calculations were carried out in the ACOS 700 computer of the Crystallographic Research Center (Institute for Protein Research, Osaka University).

Results and Discussion

Mass Spectra of Peptides Containing Homoserine Residues at the C-*Terminus.* The cleavage of a polypeptide or protein containing a Met residue(s) by BrCN[12] gives a peptide fragment(s) containing a homoserine residue at its (their) *C*-terminus and the *C*-terminal fragment of the original polypeptide or protein, as *N*-terminal BrCN fragments. Therefore, in this experiment,

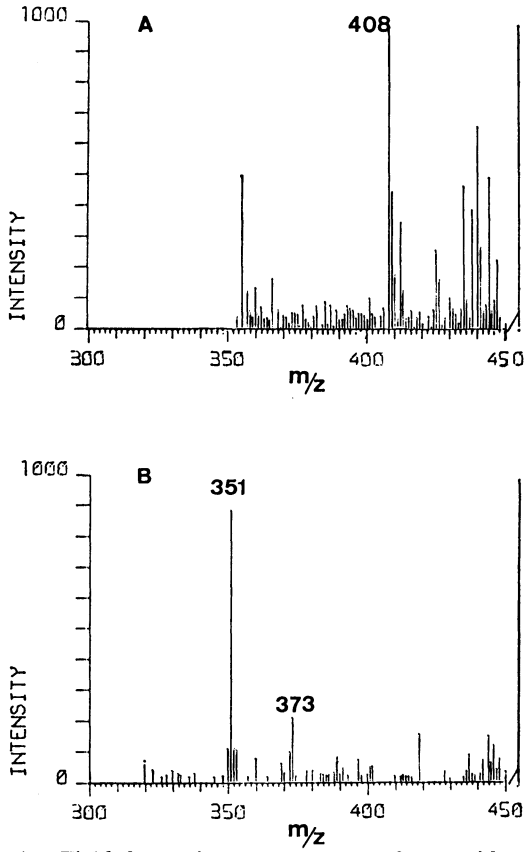

Fig. 1. Field-desorption mass spectra of a peptide mixture (A) prepared by cleavage of [Gly¹]-ACTH-(1—18)-NH_2 with BrCN and its one-step degraded peptide fragments (B). Spectra are shown in the range from 300 to 450 atomic mass units.

it was necessary to obtain information on whether peptides containing a homoserine residue at their *C*-termini give mass peaks with mass values of homoserine-peptides or those of homoserine-lactone peptides in mass spectra. As examples, we measured the field-desorption mass spectra of a peptide mixture obtained by cleavage with BrCN of [Gly¹]-ACTH-(1—18)-NH_2[11] containing a Met residue at the 5th position from the *N*-terminus and its peptide fragments prepared after one-cycle of Edman-degradation, as shown in Figs. 1A and 1B. The mass peaks of the peptide containing a homoserine residue in the mixture and its peptide fragment after Edman-degradation were observed at m/z=408 and 351, respectively, which corresponded to mass values of peptides containing homoserine lactones. However, when the peptides were measured after being kept for a long time, they gave mass values of homoserines opened at the *C*-terminal lactone rings. Generally, free "underivatized" peptides give mass peaks as quasi-molecular ions ($[M+H]^+$).[18,19] It is unknown whether peptides at their *C*-termini containing homoserine-lactones are observed as molecular ions $[M]^+$ or as quasi-molecular ions ($[M+H]^+$). The mass values (m/z=408 and 351) of the peptides containing homoserine residues observed above are responsible for the molecular ions. However, these mass values may be observed with a discrepancy of 1 atomic mass unit, because the mass peak at m/z=373 shown in Fig. 1B

is considered to be ([M+Na]$^+$) derived probably from 351 by addition of Na. This problem remains to be investigated further.

Edman-degradation and Field-desorption Mass Spectra. In our procedure of sequence determination, it was desirable to measure the mass spectra of peptides kept in the buffer solution for Edman-degradation. Therefore, we attempted to measure the mass spectra of several peptides in various buffer solutions. First we measured the mass spectra of peptides dissolved in a buffer solution containing *N,N*-dimethylallylamine as a base,[1] which is generally used in the coupling reaction of peptides with phenyl isothiocyanate in the Edman method. However, in this buffer the mass peaks could not be observed with high intensities in the spectra. The same phenomena were observed when some peptides were treated by the Edman method in the buffer solution and applied to the emitter. However, intense mass peaks could be obtained when the peptides were degraded in a mixture of pyridine and water without *N,N*-dimethylallylamine.[2,4]

Then we measured the mass spectra of a model peptide (H–Val–Tyr–Ile–His–Pro–Phe–OH) and its peptide fragments degraded successively under various conditions, as described in the Experimental section. Examples are illustrated in Figs. 2A and 2B. The model peptide and its peptide fragment gave mass peaks at m/z=775 and 676, respectively, with high intensities in all the buffer solutions examined except that containing *N,N*-dimethylallylamine. It was made not clear why the spectrum could not be obtained with high intensity in buffer solution with *N,N*-dimethylallylamine as base.

On the other hand, the 3-phenyl-2-thiohydantoin

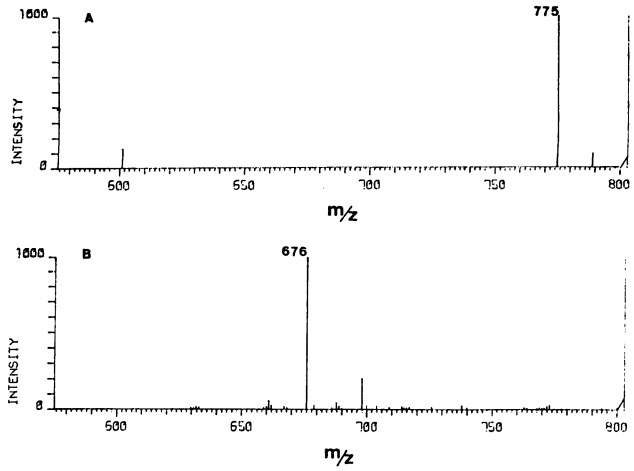

Fig. 2. Field-desorption mass spectra of H–Val–Tyr–Ile–His–Pro–Phe–OH (A) dissolved in a buffer solution containing *N*-methylmorpholine as a base (see text) and its peptide fragment (B) after one-cycle of degradation in the buffer solution.

derivatives of amino acids released from the model peptide in the degradation were quantitated by HPLC. The recoveries were similar in the all buffer solutions, but results seemed best in buffer solution containing *N*-methylmorpholine (data not shown), so, we used this buffer solution in Edman-degradation of the unknown polypeptide, as described below.

Mass Spectra of Tryptic and Chymotryptic Peptides of the N-Terminal BrCN Fragment of SE Lysozyme. We used trypsin and chymotrypsin, which specifically cleave different peptide bonds to each other, as methods for cleavage of the *N*-terminal BrCN fragment of SE

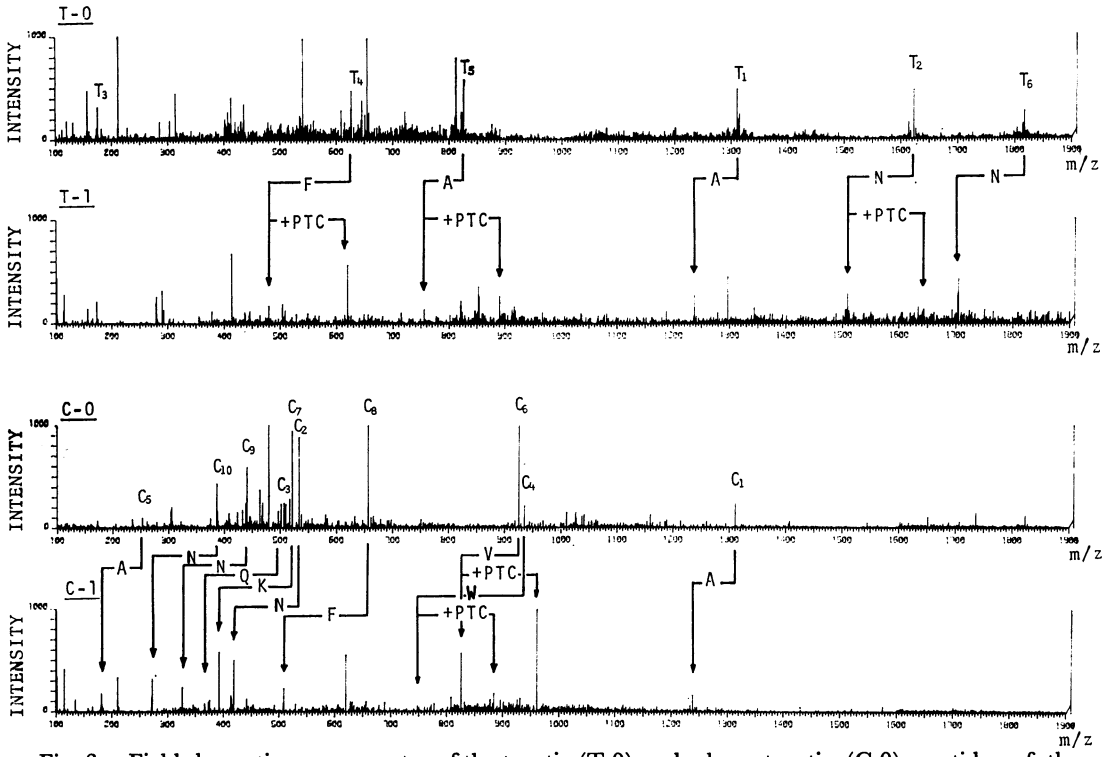

Fig. 3. Field-desorption mass spectra of the tryptic (T-0) and chymotryptic (C-0) peptides of the *N*-terminal BrCN fragment of SE lysozyme and their peptide fragments (T-1 and C-1) after one-step of degradation.

lysozyme. The tryptic and chymotryptic peptides and their peptide fragments obtained after the first-cycle of Edman-degradation gave the field-desorption mass spectra illustrated in Fig. 3. The spectra of the tryptic peptides and their degraded peptide fragments were simple, suggesting that the original polypeptide was specifically cleaved by this enzyme since it has only 3 Lys and 2 Arg residues. On the other hand, the spectra of the chymotryptic peptides and their degraded peptide fragments should be rather complicated, because the polypeptide contains many aromatic amino acid residues (6 Tyr, 3 Trp, and 2 Phe residues), which are susceptible to chymotryptic digestion. In fact, chymotryptic peptides and their degraded peptide fragments gave rather complicated spectra in the mass region up to 600 atomic mass units, as shown in Fig. 3. All the possible mass differences before and after the first-cycle of Edman-degradation of the tryptic and chymotryptic peptides were examined by the collation to the 3-phenyl-2-thiohydantoin derivatives of amino acids released in the degradation (Tables 1 and 2), as described previously.[2,4] In this case, the mass spectra were examined for mass peaks of $[M+Na]^+$, $[M+H-H_2O]^+$ and multiply charged ions $[M+2H]^{2+}$, *etc.*, which are occasionally observed with quasi-molecular ions $[M+H]^+$ in free "underivatized" peptides. Furthermore, the presence of peptides containing a Lys residue, giving mass peaks in the spectra with mass values that increased by addition of the phenylthiocarbamoyl group (135) after Edman-degradation,[4] was examined. The results revealed possible mass peaks of peptides obtained from the original polypeptide by enzymic digestion; namely, six tryptic peptides with R, F, A, and N[20] as *N*-termini and ten chymotryptic peptides with A, N, K, Q, F, W, and V as *N*-termini, as illustrated in Fig. 3. It was also found that the mass peaks at m/z=627, 826, and 1623 in tryptic peptides and at m/z=925 and 935 in chymotryptic peptides each contain a Lys residue. These findings are compatible with the presence of 3 Lys residues in the original polypeptide.

The molecular weight of the polypeptide was assumed to be within the range of 6280 and 6293 from the amino acid composition,[10] as described in Fig. 6, because Asp and Asn and Glu and Gln are quantitated together on amino acid analysis of acid hydrolysates of peptides and proteins. This molecular weight was compared with the sum of the molecular weights of the tryptic and chymotryptic peptides, which were obtained from the mass peaks in Fig. 3. The molecular weight of each peptide was obtained by subtracting one proton from the mass value of each peptide, because the mass peaks of free "underivatized" peptides are generally observed as quasi-molecular ions ($[M+H]^+$). Consequently, the tryptic peptides were assumed to be composed of peptides with mass values at m/z=175(R) : T_3, 627(F) : T_4, 826(A) : T_5, 1310(A) : T_1, 1623(N) : T_2 and 1818(N) : T_6 (amino acid residues in parentheses indicate the *N*-terminal amino acid residue of each peptide and T stands for tryptic peptide). Thus, the sum of the molecular weights of these six tryptic peptides was calculated as 6283. On the other hand, the chymotryptic peptides were suggested to be 253(A) : C_5, 386(N) : C_{10}, 440(N) : C_9, 496(Q) : C_3, 521(K) : C_7, 533(N) : C_2, 656(F) : C_8, 925(V) : C_6, 935(W) : C_4 and 1310(A) : C_1 (C stands for chymotryptic peptide). The peptide with a mass value of m/z=496 was assumed to have Q as the *N*-terminal residue, because the intense mass peaks at 479 and 501 were considered to be derived from a peptide with a mass value of m/z=496 by elimination of NH_3 during Edman-degradation or measurements of mass spectra and to correspond to $[M+H-NH_3]^+$ and $[M+Na-NH_3]^+$, respectively. Therefore, another peptide with the same mass difference (128) before and after the first-cycle of Edman-degradation as that of the peptide (m/z=496) and with a mass value at m/z=521 was expected to have *K* as the *N*-terminal residue, because one Lys residue, having the same residual weight (128) as that of Q, was observed in the 3-phenyl-2-thiohydantoin derivatives of amino acids released in the first-cycle of degradation, as shown in Table 2. The sum of the molecular weights of the chymotryptic peptides was calculated as 6283, and thus, completely coincided with that of the tryptic peptides and was definitely within the range (6280—6293) of the expected molecular weight of the polypeptide. Thus, the mass peaks of the tryptic and chymotryptic peptides

TABLE 1. RECOVERIES OF PHENYLTHIOHYDANTOIN DERIVATIVES RELEASED FROM TRYPTIC PEPTIDES OF THE *N*-TERMINAL BrCN FRAGMENT OF *Streptomyces erythraeus* LYSOZYME BY EDMAN-DEGRADATION

	Cycle of Edman-degradation										
	1	2	3	4	5	6	7	8	9	10	11
Asp	24	3	26	4	2	3	20	5	6	2	2
Asn	96	3*			2	2		—*	5	2	1
Glu			62	11	6	7	1	1	2	1	
Gln		—*	11*	8*	5	6	5	1*		t*	1
Ser		4*	—*	—*				—*	5	2*	3
Thr	1	80	8	1	12	3	1	1	2		2
Gly	6	12	3	25	28	16	7	3	5	14	6
Tyr			22*	—*	6	4	9	7	4	2	4
Ala	121	55	—*	32*	17	2					
Met											
Hse											
Val	3	58	29	8	3	1	1	14	6	1	3
Trp	2			2							
Pro		40	2						1	3	1
Lys	1	1			1			6	2		
Phe	44	2		9	1						1
Ile	1	4	1		1	20	4	1	1		1
Leu											
His											6
Arg	26	4									

Numerals are recoveries (nmol) from the *N*-terminal BrCN fragment (110 nmol). Ten % of the peptide fractions were removed for mass measurements before the 1st, 2nd and 3rd cycles of degradation. Thr was estimated in addition to *Δ*-Thr. The separations of Asn and Gln and Ser-derivatives and of Ala and Tyr-derivatives were not satisfactory in some cases. In these cases (*), recoveries were estimated as those of the main component. t means the presence of a trace of material.

Table 2. Recoveries of phenylthiohydantoin derivatives released from chymotryptic peptides of the *N*-terminal BrCN fragment of *Streptomyces erythraeus* lysozyme by Edman-degradation

	Cycle of Edman-degradation										
	1	2	3	4	5	6	7	8	9	10	11
Asp	62	30	46	8	4	3	33	8	3		1
Asn	277*	113*	11	2	1	1					
Glu	59	16*	75	37	45	9	12	4			
Gln	—*	11*	65*	31	5	1	1	2			2
Ser			—*				1		5	2	
Thr		64	13	25	7	7	10	3	3	1	1
Gly	28	51	33	39	43	34	10	19	8	19	6
Tyr		23	—*	25	17*	4	3	2	5	2	2
Ala	211	102	48*	64	22*	5	1	1			
Met											
Hse											
Val	67	47	75	15	3	2	2	24	5	1	1
Trp	69		—*	1							
Pro		2	50*	3							
Lys	113	71	5	29	2	1		1			
Phe	97	3	1				t				
Ile	3	15	2	1	1	39	5	1	1		
Leu											
His											10
Arg	8	8			3	5					

Numerals are recoveries (nmol) from the *N*-terminal BrCN fragment (160 nmol). Ten % of the peptide fractions were removed for mass measurements before the 1st, 2nd, and 3rd cycles of degradation. Thr was estimated in addition to *Δ*-Thr. The separations of Asn and Gln and Ser-derivatives and of Ala and Tyr-derivatives were not satisfactory in some cases. In these cases (*), recoveries were estimated as those of the main component. t means the presence of a trace of material.

seemed to be selected correctly.

High-performance Liquid Chromatography of the 3-Phenyl-2-thiohydantoin Derivatives of Amino Acids Released from Tryptic and Chymotryptic Peptides. The 3-phenyl-2-thiohydantoin derivatives of amino acids released from peptide mixtures by Edman-degradation were quantitated using HPLC.[17] The chromatograms of the 3-phenyl-2-thiohydantoin derivatives of amino acids released from tryptic and chymotryptic peptides are illustrated in Figs. 4 and 5, respectively. The recoveries in each cycle of Edman-degradation are summarized in Tables 1 and 2. From these values, the kind and ratio of the 3-phenyl-2-thiohydantoin derivatives of amino acids released from the tryptic and chymotryptic peptides in each cycle of degradation were selected as interger values within the limits between possible maxima and minima, as described in Fig. 6, although the composition of the 3-phenyl-2-thiohydantoin derivatives of amino acids released were rather complicated. In these cases, the specificity of the enzymes used and the expected number of peptides in mixtures were considered. Furthermore, the fact that the recoveries of 3-phenyl-2-thiohydantoin derivatives of amino acids from larger peptides or Arg-containing peptides are generally better than those from smaller peptides was

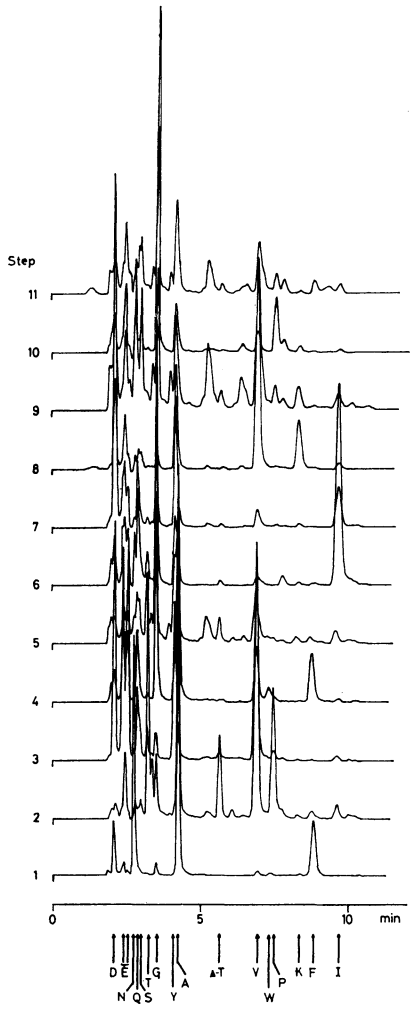

Fig. 4. High-performance liquid chromatograms of the 3-phenyl-2-thiohydantoin derivatives of amino acids released successively in the organic phase from the tryptic peptides of the *N*-terminal BrCN fragment of SE lysozyme. Numerals indicate the cycle of Edman-degradation. For conditions see text.

taken into consideration.

In this experiment, the 3-phenyl-2-thiohydantoin derivative of tryptophan was recovered with difficulty, because Edman-degradation was performed manually. Moreover, the 3-phenyl-2-thiohydantoin derivatives of homoserine and its lactone were not detected on the chromatograms.

Sequence Deduction with a Computer Program. The amino acid sequence of the *N*-terminal BrCN fragment of SE lysozyme was examined using the computer program "PROSEQ2,"[8] which was designed for deducing sequences of polypeptides or proteins from the molecular weights of the constituent peptides prepared from the polypeptides or proteins by two or more kinds of specific cleavage methods, the *N*-terminal amino acid residues or partial sequences of the constituent peptides and the 3-phenyl-2-thiohydantoin derivatives of amino acids released from the constituent peptides by successive degradation. The input data were as described in Fig. 6 and A and U (U denotes an Hse residue in this paper) as *N*- and *C*-terminal residues of the original

polypeptide,[10] respectively. There was only one output sequence, and it was complete as seen in Fig. 6. The results show that the procedure described above is applicable as a method for peptide or protein sequence analysis, and compares favorably with other methods,[1,21] although there may be no problem in sequencing a peptide as long as the *N*-terminal BrCN fragment tested with an automated sequencer.[22]

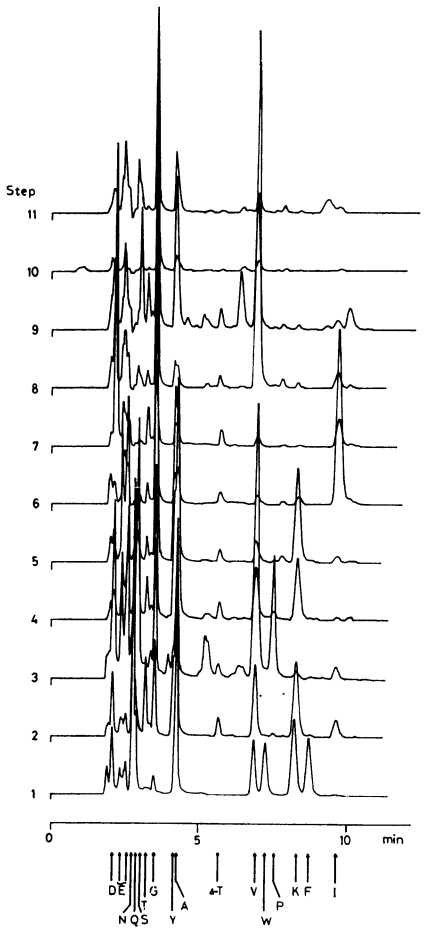

Fig. 5. High-performance liquid chromatograms of the 3-phenyl-2-thiohydantoin derivatives of amino acids released successively in the organic phase from the chymotryptic peptides of the *N*-terminal BrCN fragment of SE lysozyme. Numerals indicate the cycle of Edman-degradation. For conditions see text.

We express our sincere thanks to Professors Y. Izumi (Institute for Protein Research, Osaka University), H. Matsuda (College of General Education, Osaka University) and T. Ikenaka (Faculty of Science, Osaka University) for much helpful discussion and encouragement throughout this work.

References

1) P. Edman and A. Henschen, "Protein Sequence Determination," ed by S. B. Needleman, Springer-Verlag, Berlin-Heidelberg-New York (1975), p. 232.

2) Y. Shimonishi, Y.-M. Hong, T. Matsuo, I. Katakuse, and H. Matsuda, *Chem. Lett.*, **1979**, 1369.

3) T. Matsuo, I. Katakuse, H. Matsuda, Y. Shimonishi, Y.-M. Hong, and Y. Izumi, *Mass Spectroscopy (Japan)*, **28**, 169 (1980).

4) Y. Shimonishi, Y.-M. Hong, T. Kitagishi, T. Matsuo, H. Matsuda, and I. Katakuse, *Eur. J. Biochem.*, **112**, 251 (1980).

5) H. D. Beckey, *Int. J. Mass Spectrom. Ion Phys.*, **2**, 500 (1969).

6) T. Kitagishi, Y.-M. Hong, and Y. Shimonishi, *Int. J. Pept. Protein Res.*, **17**, 436 (1981).

7) Y. Shimonishi, Y.-M. Hong, T. Kitagishi, I. Katakuse, T. Matsuo, H. Matsuda, S. Hara, T. Ikenaka, and Y. Izumi, *Chem. Lett.*, **1981**, 499.

8) T. Kitagishi, Y.-M. Hong, and Y. Shimonishi, "Peptide Chemistry 1980," Proceedings of the 18th Symposium on Peptide Chemistry, ed by K. Okawa, Protein Research Foundation, Minoh, Osaka (1981), p. 135.

9) T. Morita, S. Hara, and Y. Matsushima, *J. Biochem. (Tokyo)*, **83**, 893 (1978).

10) S. Hara, M. Akira, E. Taketani, and T. Ikenaka, The 53rd Annual Meeting of the Biochemical Society of Japan (Tokyo), October 13—16 (1980).

11) H. Otsuka, M. Shin, Y. Kinomura, and K. Inouye,

INPUT DATA

AMINO ACID COMPOSITION

		N-TERM	C-TERM
G	7		
A	5	A	
U	1		U
S	2		
P	1		
V	4		
T	3		
L	0		
I	2		
N+D	7		
Q+E	6		
K	3		
M	0		
H	1		
C	0		
F	2		
R	2		
Y	6		
W	3		
TOTAL	55		

MOL WEIGHTS, N-TERMINAL SEQUENCES AND PTH-AMINO ACIDS IN EACH CYCLE

CHYMOTRYPTIC PEPTIDES

		2	3	4	5	6	7	8	9	10	11
252	A	G 1,0	1,0	1,0	1,0	1,0	0	1,0	0	1,0	0
385	N	A 2,0	1,0	2,0	1,0	0	0	0	0	0	0
439	N	U 1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
495	Q	S 0	1,0	0	0	0	1,0	0	1,0	1,0	0
520	K	P 0	1,0	0	0	0	0	0	0	0	0
532	N	V 1,0	2,0	0	0	0	0	1,0	1,0	0	0
655	F	T 2,0	0	1,0	0	1,0	1,0	0	0	0	0
924	Y	L 0	0	0	0	0	0	0	0	0	0
934	W	I 1,0	0	0	0	1,0	0	0	0	0	0
1309	A	N 2,0	0	0	0	0	0	0	0	0	0
		D 0	1,0	0	0	0	1,0	0	0	0	0
		Q 1,0	2,0	1,0	0	0	0	0	0	0	0
		K 1,0	0	1,0	1,0	0	0	0	0	0	0
		E 1,0	1,0	1,0	1,0	1,0	1,0	0	0	0	0
		M 0	0	0	0	0	0	0	0	0	0
		H 0	0	0	0	0	0	0	0	0	1,0
		C 0	0	0	0	0	0	0	0	0	0
		F 0	0	0	0	0	1,0	0	0	0	0
		R 0	0	0	1,0	1,0	0	0	0	0	0
		Y 2,0	1,0	2,0	2,0	0	1,0	1,0	1,0	1,0	1,0
		W 0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

TRYPTIC PEPTIDES

		2	3	4	5	6	7	8	9	10	11
174	R	G 0	0	1,0	2,0	2,0	1,0	0	1,0	2,0	2,0
626	F	A 1,0	0	1,0	1,0	0	0	0	0	0	0
825	A	U 0	0	0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
1309	A	S 1,0	0	0	0	0	0	0	2,0	1,0	2,0
1622	N	P 1,0	0	0	0	0	0	0	0	1,0	0
1817	N	V 1,0	1,0	1,0	0	0	0	1,0	1,0	0	1,0
		T 2,0	0	0	1,0	1,0	0	0	0	0	1,0
		L 0	0	0	0	0	0	0	0	0	0
		I 0	0	0	0	2,0	0	0	0	0	0
		N 0	0	0	0	0	0	0	2,0	1,0	1,0
		D 0	1,0	0	0	0	2,0	0	0	0	1,0
		Q 0	1,0	1,0	1,0	1,0	1,0	0	0	1,0	0
		K 0	0	0	1,0	0	0	1,0	1,0	0	0
		E 0	2,0	1,0	1,0	1,0	0	0	0	0	0
		M 0	0	0	0	0	0	0	0	0	0
		H 0	0	0	0	0	0	0	0	0	1,0
		C 0	0	0	0	0	0	0	0	0	0
		F 0	0	1,0	0	0	0	0	0	0	0
		R 0	0	0	0	0	0	0	0	0	0
		Y 0	2,0	0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
		W 0	0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

CANDIDATE SEQUENCE

1 ATVAGIDVSGHQRNVDWQYWWHQGKRFAYVKATEGTGYKNPYFAQQYNGSYNIGU

Fig. 6. Computer output sequence of the *N*-terminal BrCN fragment of SE lysozyme using "PROSEQ2."[8] U and W were put into as many data as possible, because the former derivative was not identified and the latter residue was possibly destroyed under our experimental conditions, as described in the text.

Bull. Chem. Soc. Jpn., **43**, 196 (1970).
12) E. Gross and B.Witkop, *J. Biol. Chem.*, **237**, 1856 (1962).
13) H. Matsuda, *Atomic Masses Fundam. Constants* **5**, 185 (1976).
14) T. Matsuo, H. Matsuda, and I. Katakuse, *Anal. Chem.*, **51**, 1329 (1979).
15) T. Matsuo, H. Matsuda, and I. Katakuse, *Anal. Chem.*, **51**, 69 (1979).
16) H. D. Beckey, A. Hendricks, and H. U. Winkler, *Int. J. Mass Spectrom. Ion Phys.*, **3**, app. 9 (1970).
17) C. L. Zimmerman, E. Appella, and J. J. Pisano, *Anal. Biochem.*, **77**, 569 (1977).
18) H. U. Winkler and H. D. Beckey, *Biochem. Biophys. Res. Commun.*, **46**, 391 (1972).
19) S. Asante-Poku, G. W. Wood, and D. E. Schmidt, Jr., *Biomed. Mass Spectrom.*, **2**, 121 (1975).
20) R. V. Eck and M. O. Dayhoff, "Atlas of Protein Sequence and Structure," ed by National Biomedical Research Foundation Publishers, Silver Spring, Md. (1966); A : Ala, D: Asp, E: Glu, F: Phe, G: Gly, H: His, I: Ile, K: Lys, L: Leu, M: Met, N: Asn, P: Pro, Q: Gln, R: Arg, S: Ser, T: Thr, V: Val, W: Trp, Y: Tyr.
21) J. A. Kelley, H. Nau, H.-J. Forster, and K. Biemann, *Biomed. Mass Spectrom.*, **2**, 315 (1975).
22) P. Edman and G. Begg, *Eur. J. Biochem.*, **1**, 80 (1967).