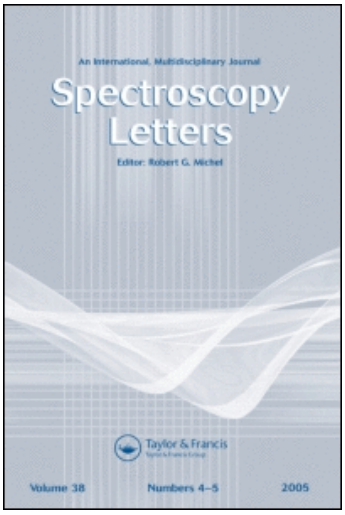


Spectroscopy Letters
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597299

A Computer Program for the Prediction of Fragmentation in the Fast Atom Bombardment Mass Spectra of Peptides

Federico M. Rubino[a]
[a] Istituto di Tecnologie Biomediche Avanzate, Milano, Italy

To cite this Article Rubino, Federico M.(1992) 'A Computer Program for the Prediction of Fragmentation in the Fast Atom Bombardment Mass Spectra of Peptides', Spectroscopy Letters, 25: 6, 811 — 820
To link to this Article: DOI: 10.1080/00387019208020713
URL: http://dx.doi.org/10.1080/00387019208020713

SPECTROSCOPY LETTERS, 25(6), 811-820 (1992)

A COMPUTER PROGRAM FOR THE PREDICTION OF FRAGMENTATION IN THE FAST ATOM BOMBARDMENT MASS SPECTRA OF PEPTIDES

Key-words: FAB-MS, Peptides, Computer Program

Federico M. Rubino

Istituto di Tecnologie Biomediche Avanzate, CNR

v.Ampere 56 I-20131 Milano (Italy)

SUMMARY

A computer program in BASIC language is described, which allows to calculate the masses of the sequence fragments in the positive-ion Fast Atom Bombardment spectra of peptides.

INTRODUCTION

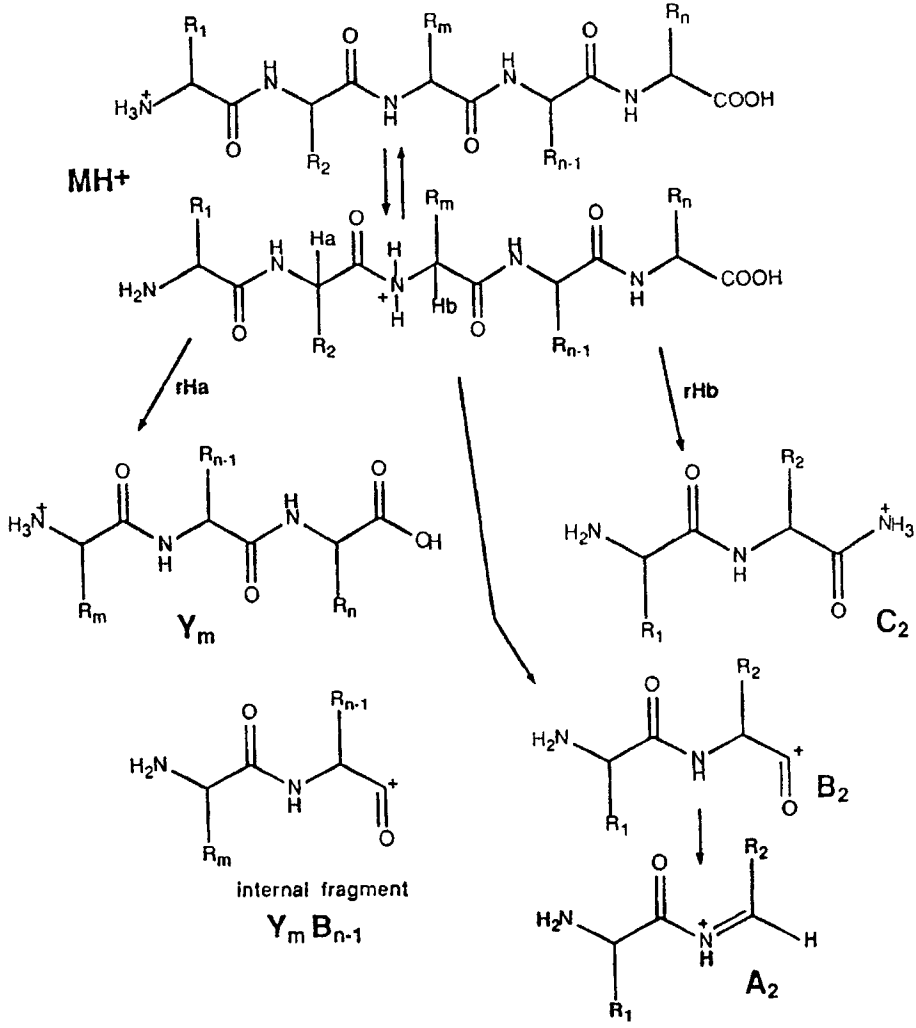
Since its introduction in the early 80's, Fast Atom Bombardment [1] has become the technique of choice for the analysis of underivatized peptides by mass spectrometry [2,3]. Sequence information is most commonly derived from analysis of the decomposition products of protonated peptide molecules (MH^+) occurring either in the source or in the field-free region of a mass spectrometer. In both cases, open and closed-shell fragment ions are formed through fission of the bonds in the peptide backbone. (SCHEME 1) and their names are attributed according to a convention proposed by Roepsdorff and Fohlmann [4].

We now describe a simple computer program which allows to calculate the mass of the backbone fragments expected for a peptide of given amino acid sequence.

DESCRIPTION OF THE PROGRAM

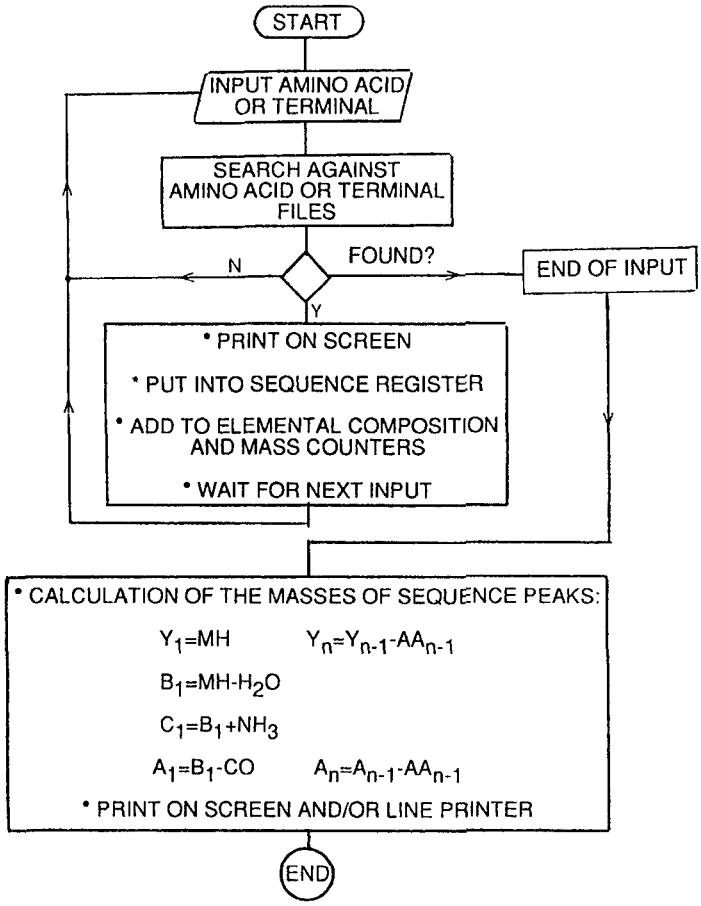
The PEPTIFAB program runs on a Commodore CBM64 home computer with minimal hardware requirement (tape or single disk drive and line printer) under the extended instruction set Simons Basic. The algorithm can be of course coded to run on any other computer, once the modifications relevant to the different BASIC dialects are performed. A flowchart of the main program routines is displayed in SCHEME 2 [5].

The program allows to calculate the nominal (monoisotopic), precise and average mass of a protonated peptide molecule with up to 25 amino acid residues. The sequence may contain any of the standard protein amino acids, and many of the modified or non-natural ones, and this set can be enlarged at the need of the user, by just entering a new line into the appropriate program routine.



SCHEME 1

Fragmentation Pattern of Peptides under Positive-ion FAB-MS



SCHEME 2

Flowchart of the PEPTIFAB Program

The relevant data on the amino acid residues (-HN-CHR-CO-) are contained in a built-in database, and a similar, but separate database contains those of the N- and C-terminal group. The record of each amino acid contains the three-letter IUPAC-IUB code, the elemental formula and the precise residue mass. Additional fields contain the Bull&Breese

hydrophobicity index [6] and the Rekker fragmental constant [7], which both can be employed to calculate parameters related to discrimination in the FAB-MS examination of peptide mixtures [8] and to prediction of the RP-HPLC elution time.

All operations in the program are menu-driven and a full explanation of the purpose and the directions for use are given in a brief HELP page.

The user is at first asked to type a name for the peptide sequence to be elaborated (a 256-character string allows full description), then the program prompt asks for the first amino acid or terminal code. Each input is searched against the database and a new one is asked for, if the entry is mistyped or not accepted. Once a new input is accepted, the program shows the sequence hitherto, under its running title. As soon as the input of the sequence is terminated, the program calculates and displays the nominal, precise and average molecular weight of the peptide and the mass of the $(M+H)^+$ species. A menu then asks if sequence peak calculation is required, and prints it as four columns of precise mass data.

Calculation of the mass of the sequence peaks is performed by stepwise subtraction, from the i-th fragment of each type, of the mass of the amino acid residue at the i-th place, to yield the corresponding (i+1)-th fragment. The first Y-type fragment (Y_1) is the molecular peak itself, while A_1, B_1 and C_1 are calculated by subtracting 18u ($-H_2O$) or 46u ($-H_2O-CO$) to the mass of MH^+, or respectively adding 17u ($+NH_3$) to the mass of B_1. It is straightforward that C_1 is only present (and has the same mass of Y_1) when the C-terminal group of the peptide is an amide.

DISCUSSION

Several computer programs have been forwarded in the past [9-15] to aid the sequence interpretation of the source and CID MS/MS spectra of

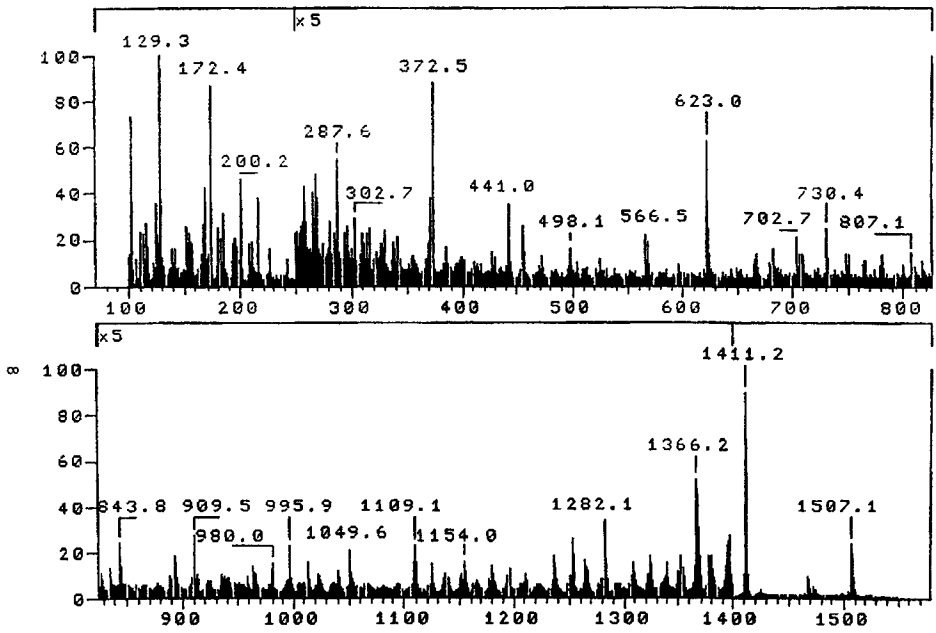


FIGURE 1
FAB spectrum of tridecapeptide SLSQSKVLPVPQK. Operating conditions: Finnigan MAT90 mass spectrometer with IonTech atom gun (Xe, 8 KeV, R>2000, scan 100-1550u at 20s/dec) sample (about 100ug) in 5uL of glycerol(1M trifluoroacetic acid).

peptides, while little information has been published on programs performing the opposite task, i.e. to predict the mass spectral fragmentation pattern for a given peptide sequence [15,16].

The one here reported has been successfully employed by us for several years, for sequence checking of peptides of synthetic origin or derived from proteolytic digest of larger molecules [16-21] through analysis of their source spectra, and in some cases of the CID MS/MS ones.

TABLE 1 Output of the PEPTIFAB Program

SYNTHETIC TRIDECAPEPETIDE				
H-Ser-Leu-Ser-Gln-Ser-Lys-Val-Leu-Pro-Val-Pro-Gln-Lys-OH				
C 63 H 111 N 17 O 19 S 0				
Mass: Integer 1409 Precise 1409.82414 Average 1410.6835 Molecular peak: 1410.83196				
	Y	C	B	A
1	1410.832	-	1392.821	1364.821
2	1323.80	1281.753	1264.726	1236.726
3	1210.716	1153.694	1136.668	1108.668
4	1123.684	1056.642	1039.615	1011.615
5	995.625	957.573	940.547	912.547
6	908.593	860.521	843.494	815.494
7	780.494	747.436	730.410	702.410
8	681.430	648.368	631.342	603.342
9	568.346	520.273	503.246	475.247
10	471.293	433.241	416.215	388.214
11	372.225	305.182	288.156	260.156
12	275.172	218.151	201.124	173.124
13	147.113	105.066	88.040	60.039

A typical application of the PEPTIFAB program is exemplified in the structural characterization of the HPLC-purified synthetic tridecapeptide SLSQSKVLPVPQK. The source spectrum (FIGURE 1) was run from a glycerol matrix acidified with 1M trifluoroacetic acid and afforded the

protonated molecule and several fragment ions. Precise mass measurement of the MH^+ peak was also performed at a resolution of about 6000 and was in excellent agreement to the calculated value (M_r found: 1410.887).

Under the experimental conditions employed by us, source and CID MS/MS spectra essentially yield the same type of fragments, so that the program can be applied to predict the fragments expected from both experiments. In particular, fragmentations leading to D- and W- type ions (which allow to discriminate isoleucine from leucine) are not detected, since these only appear in the MS/MS spectra obtained by high-energy collision induced decomposition on tandem high-resolution magnetic mass spectrometers [2,21].

TABLE 1 shows the printer output of the PEPTIFAB program once the sequence of the analyzed peptide is input. Checking of the fragments in the spectrum against the calculated ones allows to quickly confirm the proposed sequence.

REFERENCES

1) BARBER M., BORDOLI RC., ELLIOTT GJ., SEDGWICK RD., TAYLOR AN.: Fast Atom Bombardment Mass Spectrometry. Anal. Chem. 1982; **54**, 645A.

2) BIEMANN K.: Sequencing by Mass Spectrometry in: Finlay JBC. and Geisow MJ. (Eds.) Protein Sequencing a Practical Approach. JRL at Oxford University Press (1989) pp. 91-118

3) HUNT DF., SHABANOWITZ J., YATES JR., GRIFFIN PR., ZHU NZ.: Protein Sequence Analysis by Tandem Mass Spectrometry in: McEwen CN. and Larsen BS. (Eds.) Mass Spectrometry of Biological Materials. Marcel Dekker (1990) pp 169-96

4) ROEPSTORFF P. and FOHLMANN J.: Proposal of a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides, Biomed. Environm. Mass Spectrom. 1987; **11**, 601.

5) The fully operative program on a 5 1/4" floppy disk is freely available from the author.

6) BULL JB. and BREESE K.: Surface Tension of Amino Acid Solutions: a Hydrophobicity Scale of the Amino Acid Residues, Arch. Biochem Biophys. 1974; **161**, 665.

7) MOLNAR A. and HORVATH C.: Separation of Amino Acids and Peptides on non-Polar Stationary Phases by High Performance Liquid Chromatography. Jour. Chrom. 1977; **142**, 623.

8) NAYLOR S., FINDEIS AF., GIBSON BW., WILLIAMS DH.: An Approach Towards the Complete FAB Analysis of Enzymic Digests of Peptides and Proteins. J. Am. Chem. Soc. 1986; **108**, 6359.

9) MATSUO T., MATSUDA H., KATAKUSE I.: Computer Program PAAS for the Extimation of Possible Amino Acid Sequence of Peptides. Biomed. Mass Spectrom. 1981; **8**, 137.

10) MATSUO T., SAKURAI T., MATSUDA H., WOLLWICK H., KATAKUSE I.: Improved PAAS, a Computer Program to Determine Possible Amino Acid Sequences of Peptides. Biomed. Mass Spectrom. 1983; **10**, 57.

11) ISHIKAWA K., NIWA Y.: Computer-Aided Peptide Sequencing by Fast Atom Bombardment Mass Spectrometry., Biomed. Environm. Mass Spectrom. 1986; **13**, 373.

12) HAMM C.W., WILSON WE., HARVAN DJ.: Peptide Sequencing Program. CABIOS, 1986; **2**, 115.

13) JOHNSON RS. and BIEMANN K.: Computer Program (SEQPEP) to Aid in the Interpretation of High-energy Collision Tandem Mass Spectra of Peptides. Biomed. Environm. Mass Spectrom. 1989; **18**, 945.

14) ZHIDAROV D., THIBAULT P., EVANS MJ., BERTRAND MJ.: Determination of the Primary Structure of Peptides Using Fast Atom Bombardment Mass Spectrometry. Biomed. Environm. Mass Spectrom. 1990; **19**, 13.

15) C.BARTHELS: Fast Algorithm for Peptide Sequencing by Mass Spectrometry. Biomed. Environm. Mass Spectrom. 1990; **19**, 363.

16) LEE TD. and VEMURI S.: MacProMass: a Computer Program to Correlate Mass Spectral Data to Peptide and Protein Structures., Biomed. Environm. Mass Spectrom. 1990; **19**, 639 (1990)

17) PAPAYANNOPOULOS IA. and BIEMANN K.: A Computer Program (COMPOST) for Predicting Mass Spectrometric Information from Known Amino Acid Sequences. Jour. Amer. Soc. Mass Spectrom. 1991; **2**, 174.

18) ARLANDINI E., GIOIA B., PERSEO G., DANIELI B., RUBINO FM.: FD and FAB Mass Spectra of Some Oligopeptides of the Tryptophyllin Family. Biomed. Environm. Mass Spectrom. 1987; **14**, 487.

19) BANFI S., RUBINO FM., SOMMARUGA M., RESTANI P.: Synthesis and Fast Atom Bombardment Mass Spectrometry of B-Casomorphin-5. Gazz. Chim. It. 1988; **118**, 41.

20) PASQUALUCCI C., VISCONTI M., PINZA M., DANIELI B., RUBINO FM.: FAB Mapping of Elcatonin, a Novel Synthetic Analogue of Eel Calcitonin. Adv. Mass Spectrom. 1989; **11B**, 1414.

21) RUBINO F.M., DANIELI B., CHILLEMI F., CREMONESI A.: Fast Atom Bombardment and Tandem Mass Spectrometry at High and Low Collision Energy for the Sequence Analysis of Low-to-Middle Mass Peptides. Biol. Mass Spectrom. 1991; accepted for publication.

Date Received: 02/07/92
Date Accepted: 03/10/92