"SpectraGraph" and "SpectraSort": Mass Spectral Display and Interpretation Software for the Macintosh

Paul R. West*,† and Andrew J. Mort‡

Departments of Chemistry and Biochemistry, Oklahoma State University, Stillwater, Oklahoma 74078

Received October 16, 1992

Two computer programs entitled "SpectraGraph" and "SpectraSort" have been written for the Apple Macintosh. SpectraGraph allows graphical display, manipulation, storage, and printing of an input mass spectrum list that has been imported from a mass spectrometer or entered manually. SpectraGraph gives the user the ability to display, normalize, and multiply different mass ranges, annotate peaks, and perform various other operations on the spectral display. Also the mass spectrum and other graphics may be copied to and from other Macintosh application documents. SpectraSort has been developed to aid in the interpretation of mass spectra, particularly those of biopolymers, by calculating the mass differences between peaks in a mass spectrum. The user then has the option of matching the mass differences with masses of fragments or residues stored in several user-definable look-up tables.

INTRODUCTION

As rapid communication in chemistry and mass spectrometry becomes increasingly important and the amount of information to be processed is multiplying at an alarming rate, a significant need has arisen for mass spectral interpretation and display software that operates on a personal computer. Data transfer from a mass spectrometer data system to a personal computer can be achieved quite easily by direct interface using a terminal emulator, modem, or networking software. All mass spectrometer vendors now supply data manipulation software with their instruments, and many are PC driven. This, software must be linked by direct means or by network to the mass spectrometer data system. Unfortunately, there are no inexpensive, standalone programs available which can display and interpret mass spectral data on a Macintosh (Mac). Independent display and printing of mass spectral data on a PC or Macintosh is usually performed with commercially available spreadsheet or graphing software packages. These types of software are generally very inflexible for mass spectral data manipulation (i.e. normalization and peak annotation among others). Additionally, if one desires to draw chemical structures, label peaks, or otherwise enhance the mass spectrum, a separate graphics program must be used. Much of the mass spectral interpretation software developed in individual laboratories is very specific for a certain type of computer and not readily adaptable to other data systems. Computer analysis of mass spectral data has primarily been performed on mini- and mainframe computers by library search and/or pattern recognition.[1-9] Some mass spectral data interpretation software including PC Gene,[10] PROCOMP,[11] GPMA,[12] and PEPTOP[13] are commercially available for IBM personal computers. MacProMass[14] is commercially available interpretation software for Apple Macintosh personal computers. All of the above programs are specialized for specific types of mass spectral data interpretation.

This paper describes two HyperCard programs entitled "SpectraGraph" and "SpectraSort", written for the Macintosh. These programs are designed to display mass spectra and assist in the interpretation of the data. Both programs take full advantage of the Macintosh user interface.

SpectraGraph allows graphical display, manipulation, storage, and printing of an input mass spectrum list imported from a mass spectrometer or entered manually. With use of this program, the data can be viewed and manipulated by the mass spectrometrist, sample provider, or other interested party on a Mac in their own office, home, or lab. SpectraGraph gives the user the ability to display, normalize, and multiply different mass ranges, annotate peaks, and perform various other operations on the spectral display. SpectraGraph also allows graphics (such as chemical structures) to be copied from other Mac application documents and pasted anywhere on the spectrum. The spectral display may also be copied and pasted into other documents. The enhanced mass spectral data can then be printed for presentation and publication.

The second program, "SpectraSort", has been developed to aid in the interpretation of mass spectra, particularly those of biopolymers. This program also processes imported or manually entered mass spectra. SpectraSort uses a simple "mass loss" algorithm to find the mass differences between peaks in an imported or manually input mass spectrum according to subtraction parameters defined by the user. Once these mass differences are calculated, the user has the option of matching them with masses of fragments or residues stored in several user-definable look-up tables. Mass spectra can be transferred directly to SpectraSort from SpectraGraph and vice versa.

EXPERIMENTAL SECTION

SpectraGraph and SpectraSort are both HyperCard stacks, written in HyperTalk. They will run on a Macintosh Plus or any later model with 2 Mbyte of RAM and a hard disk. HyperCard version 2.0 or later and System 6.04 or later, including System 7, are also required. Both programs, including user manuals and help stacks, are available as shareware from Paul West at the above address.

LSIMS (liquid secondary ion mass spectrometry) mass spectra of a sapogenin (medicagenic acid) and a related saponin isolated from alfalfa (Medicago sativa) shown as examples were acquired on a VG ZAB2-SE reverse geometry, double-focusing mass spectrometer, operating at 8-kV acceleration voltage. Sample ionization was provided by a 35-kV Cs^+ primary ion beam. The sample matrix used was a 1:1 mixture of thioglycerol and glycerol. Data were collected and displayed

* To whom all correspondence should be addressed.
† Department of Chemistry.
‡ Department of Biochemistry.

0095-2338/93/1633-0234$04.00/0

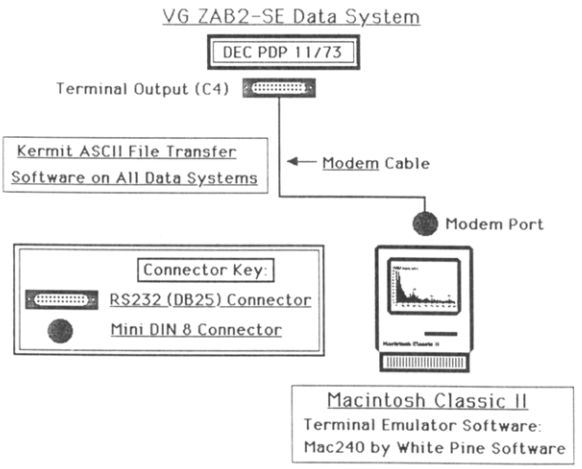


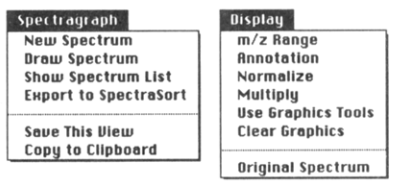
Figure 1. Hardware and software configuration for interfacing the VG 11-250J data system with a Macintosh Classic II personal computer.

by a DEC PDP-11/73 computer operating with VG 11-250J software. The VG software generated an ASCII mass spectrum list file which was then transferred directly to a Macintosh Classic II. The computers were interfaced by connecting the J4 serial output of the DEC PDP-11/73 computer by way of an RS-232 cable to the modem port of the Macintosh with an RS-232 to mini-DIN8 modem adapter cable (Figure 1). All of the hardware is readily available from personal computer supply companies.

Transfer of the ASCII mass spectrum list was accomplished by using the Macintosh program "Mac240" (White Pine Software, Nashua, NH), a DEC terminal emulator program. ASCII file transfer is explained in detail in the SpectraGraph and SpectraSort user manuals; however a brief explanation is provided here. The entire operation of file transfer was performed at the Macintosh terminal. First the Mac was used as a DEC terminal by launching the Mac240 application and then entering the login sequence to gain access to the VG software on the DEC PDP-11/73. A list of the mass spectrum of interest was then created using the VG LIST program by typing LIS SPE {Filename}/CU. The C option instructs the LIST program to create an ASCII file of the mass spectrum list while the U option makes the LIST update page available for the user to format the ASCII file properly for export to the Macintosh. Once the parameters such as mass range, intensity threshold, etc., were set in LIS, the list was created by typing 'G. This type of ASCII spectrum list is stored on the VG data system hard disk in the [71,33] directory as [71,33] {Filename}.Lis. This file was then transferred to the Macintosh using the Kermit (Columbia University, New York, NY) ASCII file transfer protocol (which is included in Mac240 and most other commercial modem software) in conjunction with Kermit-11 on the VG PDP-11/73 system. This transfer was accomplished by typing CONTROL\ to access RSX and then launching Kermit by typing KERMIT. The file was transferred by typing SEND [71,33] {Filename}.Lis from the KERMIT prompt and then pulling down the Mac240 "File" menu and selecting "Receive File (Kermit)...". The file was then transferred and stored in the Macintosh at a designated location on hard or floppy disk, where it was imported into SpectraGraph or SpectraSort.

RESULTS AND DISCUSSION

SpectraGraph. This program provides display and enhancement of an input centroided mass spectrum list. Spec-




```
P2206110R#1    x1   F:D FFR:1 22-JUN-91 15:40+0:02:37  ZAB-SE    FB+      1.1
BpM=0   I=9.9v  Hm=0   TIC=0                  Acnt:PW      Sys:DAUGHTERS
B/E CAD OF MEDICAGENIC ACID [M+H]+                        PT= 0    FM1:501.100
     Mass    % Base
    501.1    100.00
    499.0      0.08
    497.8      0.02
    496.9      0.02
    496.1      0.04
    495.0      0.03
    494.3      0.05
    492.9      0.07
    492.0      0.04
    490.9      0.05
    488.8      0.02
    487.9      0.03
    486.0      0.40
```

Figure 2. (a, top) SpectraGraph's custom menus. Most routines are initiated using these menus. (b, bottom) A partial ASCII text spectrum list as generated by the List subroutine in the VG 11-250J software on the DEC PDP-11/73 computer. This is how the list appears on the Macintosh after being transferred using Kermit and then being opened with a word processing application.

traGraph takes full advantage of the Macintosh environment by providing the user with custom menus (Figure 2a) to initiate most of the routines. A mass spectrum list (in x,y format with $x = m/z$ and y = intensity) may be entered manually or a mass spectrum ASCII file (Figure 2b) that has been transferred from a mass spectrometer data system may be input automatically. This ASCII mass spectrum list can be imported into SpectraGraph from any mass spectrometer capable of generating such a list; however there are currently three versions of the SpectraGraph available, designed to import and automatically modify to usable form (remove headers, extract sample ID information, etc.) an ASCII mass spectrum list from VG 11-250 software (Figure 3), a JEOL DA5000, and a Finnigan TSQ-70.

When SpectraGraph imports a VG, JEOL, or Finnigan ASCII file, it places the filename and text for the file into the proper fields on the SpectraGraph display and then removes unnecessary characters (i.e. headings and extra spaces) before placing the spectrum list into the Spectrum List field. ASCII files from other data systems may also be imported, but since the format of these files is different, the following special formatting rules must be followed when editing the list in order for SpectraGraph to recognize and plot the spectrum list:

1. The list must be two columns, the left column being the m/z value, the right column its corresponding intensity.
2. The masses may be entered in any order. SpectraGraph will sort the list prior to plotting the spectrum.
3. There must be no other characters in the list (including headings) other than m/z and intensity.
4. The spectrum should be normalized to the most intense peak in the mass range. In other words, one of the peaks should be at 100% and all others normalized to it. This is not absolutely necessary, but it helps to turn out a better drawing.
5. The spectrum list may be either Tab or Space delimited.

Once the spectrum list has been entered, or imported, the spectrum is plotted (Figure 4). Some of the mass spectral display features of SpectraGraph (as demonstrated in Figure 5) include variable display mass range, normalization of the mass spectrum, multiplication of a defined mass range, and the ability to save up to 10 displays for later recall. Peak annotation with 0–4 decimal places can be performed

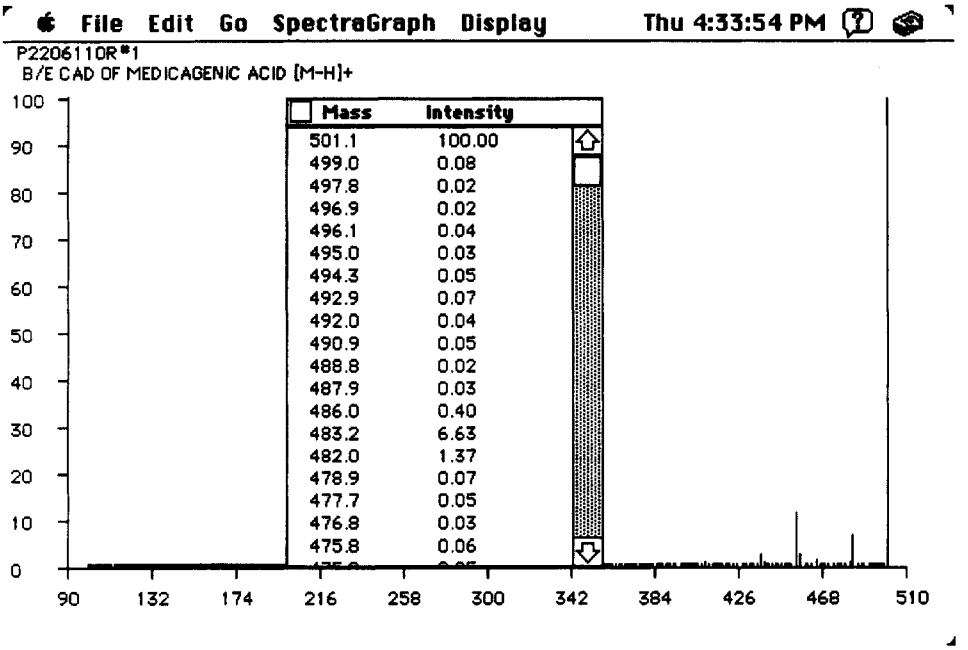


Figure 3. ASCII test spectrum list from Figure 2b as imported and formatted by SpectraGraph. The filename and sample ID are automatically placed in the upper left corner of the spectrum.

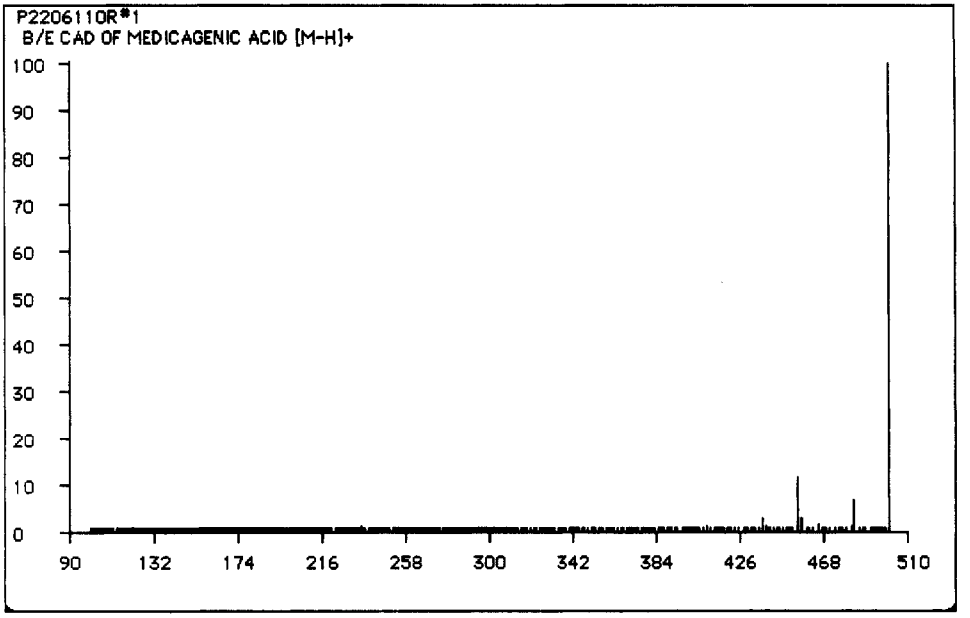


Figure 4. Original spectrum as plotted by SpectraGraph from the list in Figure 3.

automatically for the entire display range, or individual peaks can be annotated by simply clicking the mouse on the peak.

A set of drawing tools (including line, polygon, rectangle, freehand pencil, eraser, lasso, and selection) is available, enabling the user to draw graphics such as molecular structures or labels directly on the spectral display (Figure 5). Graphics and/or text may be copied from other application documents and then pasted onto the SpectraGraph mass spectrum display. The spectrum is originally drawn in the background (lower) layer; therefore the actual stick spectrum is unaffected by any cutting or erasing of graphics in the card (upper) layer. The entire display may be copied and pasted (via the Macintosh clipboard) into other application documents. The mass spectrum list may also be output automatically to SpectraSort for interpretation.

Another advantage of SpectraGraph is that once a mass spectrum ASCII file has been transferred to a Macintosh computer, it may then be stored on floppy disk and transported to other locations, such as the sample provider's laboratory. If the sample provider has SpectraGraph on his or her Macintosh, the spectrum can then be viewed, manipulated, labeled, and printed at that location at any time. This is also very conducive for data archiving.

SpectraSort. SpectraSort has been developed to aid in the interpretation of mass spectra. SpectraSort uses a simple "mass loss" algorithm to find mass differences between peaks in the spectrum. Determining these mass differences can be very helpful for chemical structure elucidation.

A new mass spectrum list may be entered manually or imported automatically from an ASCII text file (similar to SpectraGraph file importation) or directly from SpectraGraph. When importing a spectrum list, an intensity threshold is provided in order to filter out background noise and insignificant peaks. The mass list is then sorted according to descending mass order. This enables the user to examine the list for duplication errors due to manual entry and helps to

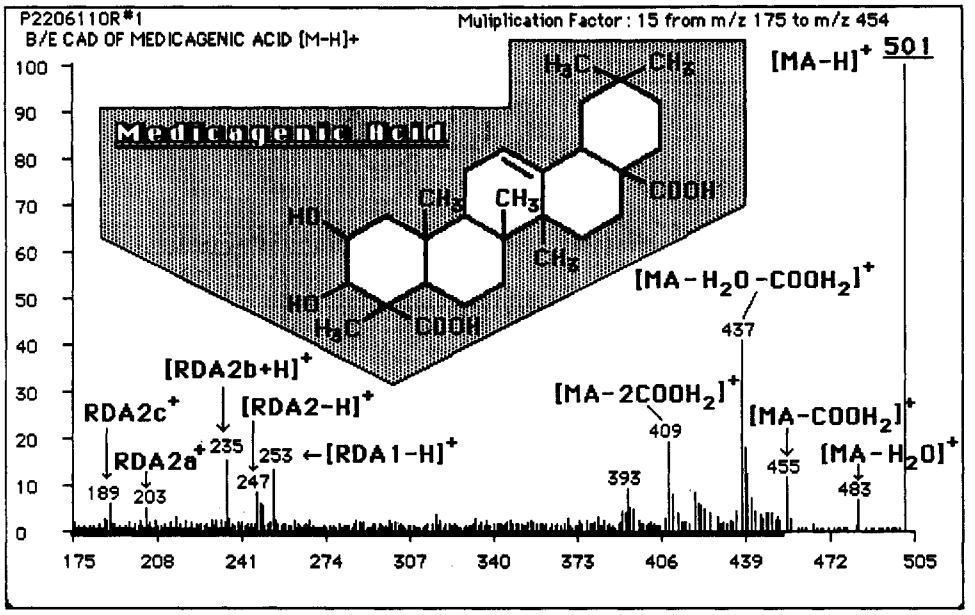


Figure 5. Mass spectrum from Figure 4. The spectrum has been multiplied by a factor of 15 from m/z 175 to m/z 454, and the major peaks have been annotated and labeled. The structure of medicagenic acid was first drawn in ChemDraw and then pasted onto the SpectraGraph display, where it was graphically enhanced using the graphics tools. The enhanced mass spectrum is now ready for printing.

compensate for some differences between mass spectrometer data systems. SpectraSort then calculates mass differences between peaks in the spectrum according to upper and lower subtraction limits set by the user. Mass differences are calculated in the following manner.

Example: For the following observed m/z values

peak 1: 500

peak 2: 492

peak 3: 442

peak 4: 230

SpectraSort calculates the difference between

peak 1 and peak 2: 500 – 492 = 8 (A)

peak 1 and peak 3: 500 – 442 = 58 (B)

peak 1 and peak 4: 500 – 230 = 270 (C)

SpectraSort then calculates the difference between

peak 2 and peak 3: 492 – 442 = 50 (D)

peak 2 and peak 4: 492 – 230 = 262 (E)

Finally, SpectraSort calculates the difference between

peak 3 and peak 4: 442 – 230 = 212 (F)

If the upper subtraction limit is set to 260, then the following mass difference results will not be displayed:

(C) because 270 > 260

(E) because 262 > 260

Similarly, if the lower subtraction limit is set to 17 then (A) will not be displayed because 8 < 17.

The other results, (B), (D), and (F) will be displayed since they fall within the upper and lower subtraction limits. A hard copy of the results of the mass difference calculation results may be printed.

Residues ± 1 amu	Mass Difference
	345 - 321 = 24.0
acid ;	345 - 300 ▪ 45.0
	345 - 290 = 55.0
	345 - 123 = 222.0
	345 - 102 = 243.0
Na ;	321 - 300 = 21.0
	321 - 290 = 31.0
	321 - 123 = 198.0
	321 - 102 = 219.0
hexa ;	300 - 123 = 177.0
	300 - 102 = 198.0
	290 - 123 = 167.0
	290 - 102 ▪ 188.0
Na ;	123 - 102 = 21.0

Figure 6. Results displayed after a SpectraSort mass difference calculation and residue mass search. "Hexa" and "acid" are user-defined residue abbreviations corresponding to a hexuronic acid residue (MW 176) and COOH (MW 45), respectively.

Once the calculations are complete, a user-defined "Residue File" can be searched to find residues with masses which match the mass differences. The Residue File to be searched is chosen by pulling down the "Search Files" menu and selecting the appropriate file. SpectraSort then searches through the chosen Residue File to find release masses which match the calculated mass difference to within a specified mass window (from 0 to ±2 amu) which is specified by the user. When the search is completed, the matching residues are displayed in the "Residues" field on the left, corresponding to its mass difference on the right, as shown in Figure 6. A report can also be made which shows only the mass differences to which corresponding matching residues have been found. A very abbreviated SpectraSort report is shown in Figure 7. This figure also illustrates the ability of SpectraGraph and SpectraSort to work together.

Residue Files. The most important feature of SpectraSort is that the user has the ability to create and modify custom Residue files. These are the files that SpectraSort searches to find matches to the calculated mass differences. Each Residue file consists of a list of Residues (i.e. fragments). Each line in the file contains three pieces of information:

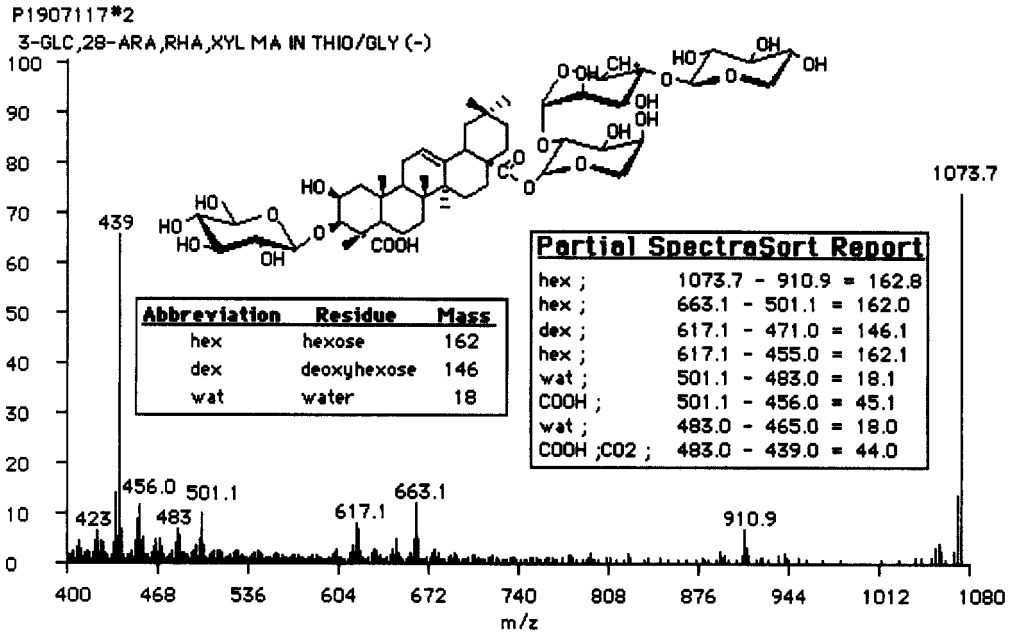


Figure 7. LSIMS spectrum of 3-Glc, 28-Ara Rha Xyl Medicagenic Acid as plotted and graphically enhanced in SpectraGraph. The structure of the compound was first drawn in ChemDraw and then pasted onto the display. A partial SpectraSort Report for this spectrum was also placed on the spectrum. The entire figure as shown was created on a Macintosh SE with no manual cutting and pasting needed.

Residue File:		'Bio Residues'
57.02146	gly	glycine
71.03711	ala	alanine
87.03203	ser	serine
97.05276	pro	proline
101.04768	thr	threonine
103.00919	cys	cysteine
113.04768	hyp	hydroxyproline
113.08406	ile	isoleucine
113.08406	leu	leucine
115.02694	asp	aspartic acid
128.09496	lys	lysine
129.04259	glu	glutamic acid
131.04049	met	methionine
137.05891	his	histidine
147.06841	phe	phenylalanine

(a)

Exact Mass File:	'Biochem'
Ala, 71.03711	
Arg, 156.10111	
Asn, 114.04293	
Asp, 115.02694	
Cys, 103.00919	
Glu, 129.04259	
Gln, 128.05858	
Gly, 57.02146	
His, 137.05891	
Ile, 113.08406	
Leu, 113.08406	
Lys, 128.09496	
Met, 131.04049	
Orn, 114.07931	

(b)

Figure 8. (a, top) One of the SpectraSort residue files available for searching. (b, bottom) SpectraSort exact mass file used for mass calculation.

1. the name of the residue
2. an abbreviation for the name
3. the mass of the residue

A new residue file may be created by choosing "Create New File" under the "Residue Files" menu. A typical residue file, "Bio Residues" is shown in Figure 8a. A hard copy of a residue file can be printed, and residues may be added to or deleted from a residue file by simply clicking on the appropriate buttons.

Exact Mass Calculator. An exact mass calculator has been provided which uses "Exact Mass Files" containing user defined residues and their masses. This feature has been implemented to assist in calculating the masses of large molecules such as peptides, polysaccharides, etc.; however any type of mass calculation can be performed. The Exact Mass Calculator may be accessed directly from the "Mass Entry" page by choosing the "Exact Mass Calculator" menu item under the "SpectraSort" menu.

The "Exact Mass Files" provided with SpectraSort contain masses for the 106 elements (most abundant isotopes) and residue masses for the 20 most common amino acids and some common sugars. Each of these biochemical residue masses is the molecular weight of the compound minus the mass of water in order to facilitate calculation of biopolymers. When calculating the mass of a peptide for example, the mass of water must be added once per peptide. All parts of the formula are entered separated by commas as shown in these examples:

$C_6H_{12}N$ should be entered as: C,6,H,12,N,1

For example, to calculate the mass of maltoheptaose (an oligosaccharide consisting of seven glucose (hexose) residues plus water) the formula would be entered as hex,7,wat,1.

The resulting mass is displayed as:

hex7wat1 = 1152.38031.

The most significant feature of the exact mass calculator is that the user has the ability to modify or delete the existing files and create new files. A new Exact Mass File is created by choosing the appropriate menu and then entering the names of the elements or residues and their corresponding masses to the file by clicking on the proper buttons. A hard copy of the current file on view can be obtained by clicking on the "Print Exact Mass File" button. A typical Exact Mass Residue File ("Biochem") is shown in Figure 8b.

CONCLUSIONS

SpectraGraph and SpectraSort are very helpful and easily accessible utility programs for mass spectral display and interpretation, taking full advantage of the Macintosh user interface. Mass spectra can be transferred from a mass spectrometer data system to a Macintosh and then in turn to any other Macintosh via floppy disk. SpectraGraph then allows display, enhancement, and printing of the mass spectra, while SpectraSort aids in the interpretation of the data.

ACKNOWLEDGMENT

The authors wish to thank the National Science Foundation for Research Grant DMB 9006052 and Instrument Grant BBS 8702089 which contributed toward the purchase of the mass spectrometer and data system. P.R.W. would also like to thank Bob West for his invaluable assistance.

REFERENCES AND NOTES

(1) Martinsen, D. P.; Song, B.-H. Computer applications in mass spectral interpretation: A recent review. *Mass Spectrom. Rev.* **1985**, *4*, 461–490.

(2) Enke, C. G.; Wade, A. P.; Palmer, P. T.; Hart, K. J. Solving the MS/MS Puzzle: Strategies for Automated Structure Elucidation. *Anal. Chem.* **1987**, *59*, 1363A–1371A.

(3) Wade, A. P.; Palmer, P. T.; Hart, K. J.; Enke, C. G. Development of Algorithms for Automated Elucidation of Spectral Feature/Substructure Relationships in Tandem Mass Spectrometry. *Anal. Chim. Acta* **1988**, *215*, 169–186.

(4) Loninger, H.; Varmuza, K. Selective Detection of Classes of Chemical Compounds by Gas Chromatography/Mass Spectrometry/Pattern Recognition: Polycyclic Aromatic Hydrocarbons and Alkanes. *Anal. Chem.* **1987**, *59*, 236–244.

(5) Zhudamo; She, J.; Hong, Q.; Liu, R.; Lu, P.; Wang, L. ASES/MS: An Automatic Structure Elucidation System for Organic Compounds Using Mass Spectrometric Data. *Analyst* **1988**, *113*, 1261–1265.

(6) Luinge, H. J. A knowledge-based system for structure analysis from infrared and mass spectral data. *Trends Anal. Chem.* **1990**, *9*, 66–69.

(7) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.

(8) Munk, M. E.; Christie, B. D. The Characterization of Structure by Computer. *Anal. Chim. Acta* **1989**, *216*, 57–68.

(9) Siegel, M. M.; Gill, G. MASSPEC: a graphics-based data system for correlating a mass spectrum with a proposed structure. *Anal. Chim. Acta* **1990**, *237*, 459–472.

(10) *PC Gene*, *5.17*, for the IBM PC; Intelligenetics: Mountain View, CA, 1988.

(11) Andrews, P. C. *PROCOMP*, 1.0, for the IBM PC; Biochemistry Department, Purdue University: West Lafayette, IN, 1989.

(12) Hojrup, P. *GPMA*, 4.2A, for the IBM PC; Department of Molecular Biology, Odense University: Odense, Denmark, 1988.

(13) Fraser, B. *PEPTOP*, 3.0, for the IBM PC; Center for Biologics Evaluation and Research, FDA: Bethesda, MD, 1988.

(14) Lee, T. D.; Vemuri, S. MacProMass: A Computer Program to Correlate Mass Spectral Data to Peptide and Protein Structures. *Biomed. Environ. Mass Spectrom.* **1990**, *19*, 639–645.