```
EQU2  SADVAGAVIDGASLSFDILKTVLEALGNVKRKIAVGVDNESGKTWTALNTYFRSGTSDIV
CYT3  A--LAGTIIAGASLTFQVLDKVLEELGKVSR----------------------SGTTDVI
STIC  A--LAGTIIAGASLTFQVLDKVLEELGKVSRKIAVGTDNESGGTWTALNAYFRSGTTDVI
      .--.**..*-****.*..*--***-**.*-*----------------------***.*..
EQU2  LPHKVPHGKALLYNGQKDRGPVATGAVGVLAYLMSDGNTLAVLFSVPYDYNWYSNWWNVR
CYT3  LPEFVPNTKALLYSGRKDTGPVATGAVAAFQYYMSSGNTLGVMFSVPFDYNWYSNWWDVK
STIC  LPEFVPNTKALLYSGRKDTGPVATGAVAAFQYYMSSGNTLGVMFSVPFDYNWYSNWWDVK
      **.-**.-*****.*.**-********.-.-*-**-****.*.****.*********.*.
EQU2  IYKGKRRADQRMYEELYYNLSPFRGDNG-WHTRNLGYGLKSRGFMNSSGHAILEIHVSKA
CYT3  IYSGKRRADQGMYEDLYYG-NPYRGDNGHWE-KNLGYGLRMKGIMTSAGEAKMQIKISR-
STIC  IYSGKRRADQGMYEDLYYG-NPYRGDNGHWE-KNLGYGLRMKGIMTSAGEAKMQIKISR-
      *********-***.***-.*.*****-*.-.******.-.*.*-*.*.*-..*-.*.
```

has also been isolated from *St. helianthus* (1) and with similar biological function. Unexpectedly, from cycle 30 to 51 amino acids detected did not match at all with Cytolysin-III, but coincide again after cycle 51 to 67.

In order to verify the primary structure the protein was digested with trypsin. Isolated peptides were sequenced and their mass were confirmed by FAB mass spectrometry. 95% of the protein sequence was verified and 87% coincides with Cytolysin-III (CYT3). On the other hand, sequence alignment with protein database also revealed more than 65% homology with Equinatoxin-II (EQU2) isolated from the australian sea anemone *Actinia tenebrosa* and the european *Actinia equina,* which also contains the 22 amino acids insertion fragment.

The amino acid composition match very well with the theoretical composition considering also complete similarity with Cytolysin-III in the still non-verified 5% of the molecule sequence, however the experimental molecular mass differ from that calculated (19341 Da) thus we suppose the existance of minor differences within the non-verified amino acid sequences.

REFERENCES

BLUMENTHAL, K. M. and W. R. KEM (1983). *J. Biol. Chem.* **258** (9): 5574-5580.

A PROGRAM TO AID THE INTERPRETATION OF THE MS/MS SPECTRA OF PEPTIDES

Jorge F. Cossío, Javier González and Vladimir Besada.

Center for Genetic Engineering and Biotechnology. P.O. Box 6162. La Habana 6, Cuba.

INTRODUCTION

The combination of Collision Activated Decomposition (CAD) and linked-scan measurement has demonstrated its utility for sequencing unknown and chemically modified peptides. When a peptide pass through the collision cell yield many daughter ions that can be pooled in two groups, the N- and C-terminal ions.

Both ion series contain complementary information on the peptide sequence, the N-terminal ions (**a,b,c"**) provide us the sequence from the N- to the C-terminus and the C-terminal ions (**x, y", z**) from the C- to the N-terminus. In the manual interpretation of the CAD spectra it is not obvious the assignment of the daughter ions as an **N-** or **C-**terminal ion unless the peptide be partially labeled with ^{18}O at their C-terminus (1,2). Even so, for an skillful specialist the interpretation of a complicated CAD spectra often takes long time and the results can be ambiguosly assigned to several sequences.

Various computer algorithms has been described in order to aid the peptide sequencing but few of them behave well for peptides with molecular weight over 1 000. Our algorithm based on the graphos theory (3)

was modified on to obtain faster and reliable results of the "N" sequences with the best scores. To compare it we used a computer with lower performances than the previous papers.

MATERIALS AND METHODS

The CAD spectra were obtained in a JEOL JMS HX-110HF two sector mass spectrometer, operated with JEOL JMA DA-5000 data system. The computer used to evaluated the performances of this program was an IBM XT 8086/8 Mhz without mathematical coprocesor.

RESULTS AND DISCUSSIONS

Our program only need as input data: the molecular weight of the peptide, the mass of the N- and C-terminal groups and the mass of the daughter ions observed in the CAD spectrum, neither the amino acid composition nor the relative intensities of the daughter ions are necessaries although both informations can also be introduced to the computer in order to increase the speed and use another scoring method. Partial sequences obtained by exopeptidase analysis or Edman degradation can be used by the program if this information is available. In the following table we compare the results obtained by the program reported by Bartels et. al. (3) and our program with the modification mentioned above.

We searched for the original spectrum of sample No.1 and the double in number of peaks were input to our program and the execution time was lower in comparison with the time needed by the Bartels's program executed in an IBM-AT compatible UNISYS (80386) computer superior in performances than the used in this paper.

It is very important to point out that the modification implemented in our program is the main cause in the reduction of the execution time. For example in sample No. 1 the non-modified program needed 1 min 38 sec to find 15378 possible sequences with the same M.W. than the analyzed peptide but the modified one only analyzed 20 sequences in order to find the top ten candidates. The execution time is increased linearly with number of peaks and asymptotically (to the time needed for the unmodified program) while increase the "N" top sequences selected by the users.

The program also search for pair of amino acids to explain the gaps if peaks are missing. Actually this program has been used to sequence several peptides obtained from natural and recombinant proteins. It can also be helpful to process CAD spectra of peptides partially laveled with ^{18}O at their C-terminus and peptides with a fixed positive charge at the N-terminus. The program needed less than 1 minute to analyze very complicated spectrum it can also differentiate two isobaric amino acid such as Leu and Ile when the side chain fragmentation were observed in the spectrum.

REFERENCES

1. TAKAO, T. *et al.* (1990). Rapid. Commun. *Mass Spectrom,* **5**:312-318.
2. TAKAO, T. *et al.* (1993). *Anal. Chem.* **65**:2394-2399.
3. BARTELS, C. *et al.* (1990). Biomed. Environ. *Mass Spectrom.* **19**:363-368.

Bartels *et al.*					Cossío *et al.*		
Sample No.	No. peaks	Time (s)	Top score	Score. right structure	No. peaks	Time (s)	Ranks
1	21	14	475	464	43	10.89	2
2	8	4	352	352	8	0.71	1
3	31	13	418	386	31	5.44	1
4	13	14	372	360	14	3.24	10
5	30	15	405	365	30	4.01	3
6	26	21	386	364	33	5.10	1
7	34	10	390	390	34	4.72	1