Biotecnología Aplicada, Vol. 9, No. 1, 1-15 (1992)

Trabajos de revisión

Mass spectrometry in protein analysis

V. BESADA, W. ANTUCH, R. RODRÍGUEZ, L. GONZÁLEZ, G. CHINEA and G. PADRÓN

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

Received in September 1991
Approved in October 1991

SUMMARY

Fast Atom Bombardment Mass Spectrometry (FAB-MS) is an established method for peptide and protein analysis. In particular, the amino acid sequence of recombinant proteins is verified by this method, in order to detect mutations or posttranslational modifications.

In this paper, we discuss the general method and its application to the determination of different posttranslational modifications such as: acylation, glycosylation, scrambled disulfide bridges and protein degradation. The use of FAB-MS in peptide sequencing is also presented.

RESUMEN

La espectrometría de masas con ionización por bombardeo con átomos acelerados (FAB-MS) es un método establecido para el análisis de péptidos y proteínas. La secuencia aminoacídica de las proteínas recombinantes se verifica utilizando este método, para detectar posibles mutaciones o modificaciones post-traduccionales.

En este artículo se presenta el método general y su aplicación en la determinación de modificaciones post-traduccionales como: acilación, glicosidación, formación incorrecta de puentes de disulfuro y degradación de las proteínas, entre otras.

For more than 30 years, mass spectrometry and, particularly, the combination of gas chromatography and mass spectrometry (GC-MS) has been one of the most powerful tools in the analysis of organic compounds. However, its application to biomolecules such as carbohydrates, proteins and nucleic acids was extremely limited up to 1981, because the ionization methods available up to that time needed evaporation of the sample before ionization and most of the bioorganic compounds are non-volatile and very sensitive to temperature.

In 1981, Michael Barber at the Institute of Science and Technology of the University of Manchester developed the Fast Atom Bombardment (FAB) ionization method using accelerated argon atoms as ionizing agent (Barber *et al.*, 1981a). The name of this method hides its most outstanding feature. Several similar systems were known in 1981 by the generic name of Secondary Ion Mass Spectrometry (SIMS) that use accelerated ions as ionizing agents. In these methods, the sample is deposited on a solid surface and, after a few seconds

of bombardment, the solid surface is completely destroyed, avoiding recording the whole spectrum.

The major advantage of the Barber method is the use of a liquid matrix (glycerol) instead of a solid surface. In spite of the fact that the surface of the liquid matrix is also damaged, new molecules of the liquid rise to the surface, restoring it. This is even more evident in the new instruments on the market that can use both xenon atoms or cesium ions as the ionizing agents.

Since Barber introduced the FAB ionization method, it has been readily applied to the analysis of peptides and proteins (Barber *et al.*, 1981b; Morris *et al.*, 1981). This method has been focused on the characterization of recombinant proteins (Fukuhara *et al.*, 1985; Takao *et al.*, 1987), because a knowledge of the gene sequence is not sufficient to ensure the correct amino acid sequence of the protein.

Analysis by FAB mass spectrometry (FAB-MS) of enzymatic digests of proteins is now an established method for the confirmation of amino acid sequences and the determination of post-translational modifications (Morris *et al.*, 1983; Takao *et al.*, 1984; Gibson and Biemann, 1984).

This method has been applied to the characterization of several recombinant proteins produced in *E. coli* or in yeast at the Center for Genetic Engineering and Biotechnology in Havana. Further, a combination of reverse phase liquid chromatography (RP-HPLC) and FAB-MS was found to be a powerful tool for examining the fermentation, renaturation and purification processes of recombinant proteins (Padrón *et al.*, 1989; Besada *et al.*, 1990).

We have already studied human alpha-2b interferon (α-IFN), human gamma interferon (γ-IFN), human interleukin-2 (IL-2), streptokinase, α-amylase, human epidermal growth factor (EGF), hepatitis B surface antigen, murcorpepsin and sucrose invertase. We have also sequenced several peptides including a protein proteinase inhibitor isolated from a sea anemone.

By mass spectrometry we have been able to verify the amino acid sequence, including the location of S-S bonds, to detect such post-translational modifications as acetylation, formylation or glycosylation, artifacts produced during the purification process, ragged C-terminal end, etc. The sequence of a 55-amino acid polypeptide was determined combining mass spectrometry, exopeptidases, endopeptidases and manual Edman degradation. In this article, a summary of these results is presented.

GENERAL STRATEGY

In Fig. 1 is shown the general strategy for characterization of recombinant proteins by FAB-MS. The pure protein is digested with different proteinases, separately or in tandem, and the mass spectra of the digests are recorded. The remainders of the peptide mixtures are purified by reverse phase liquid chromatography (RP-HPLC) and each isolated peptide is analyzed by FAB-MS.

FAB-MS is a "soft" ionization method and the spectrum of a peptide consists of practically only one peak due to the pseudomolecular ion: $(M+H)^+$ or $(M-H)^-$, depending on the detection mode (positive ion mode is currently used) without any important fragmentation. Therefore, a mass spectrum of a peptide mixture is a collection of pseudomolecular ions

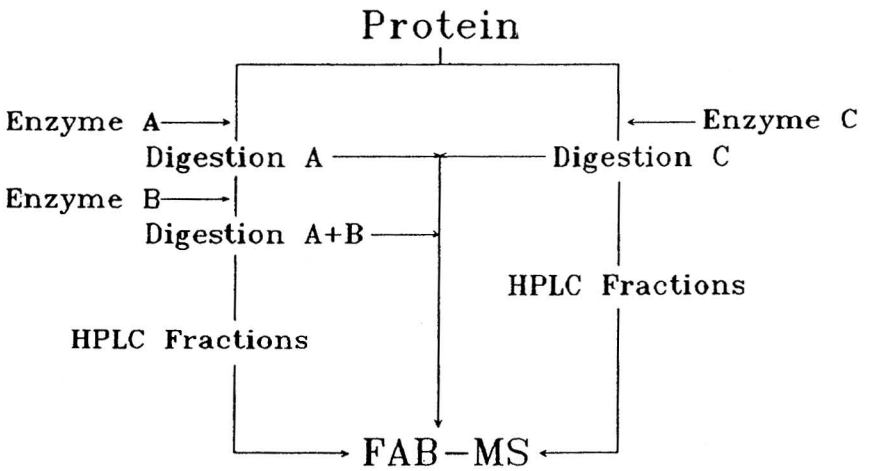

FIG. 1. General strategy for verification of amino acid sequences of recombinant proteins by FAB-MS.

corresponding to the peptides in the mixture and, in principle, it is not necessary to purify it. However, hydrophobic peptides are ionized more easily than hydrophilic ones and often, hydrophilic peptides are not observed in the mass spectra. A partial purification of peptide mixtures by reverse phase liquid chromatography avoids this inconvenience.

The molecular weight of each peptide is then matched with the expected sequence of the protein. Mistakes in the DNA sequence in the synthesis of the protein by the host cell or modified amino acids can be detected.

Figure 2 shows the verification of the sequence of α-IFN. The protein was digested with trypsin and endoproteinase

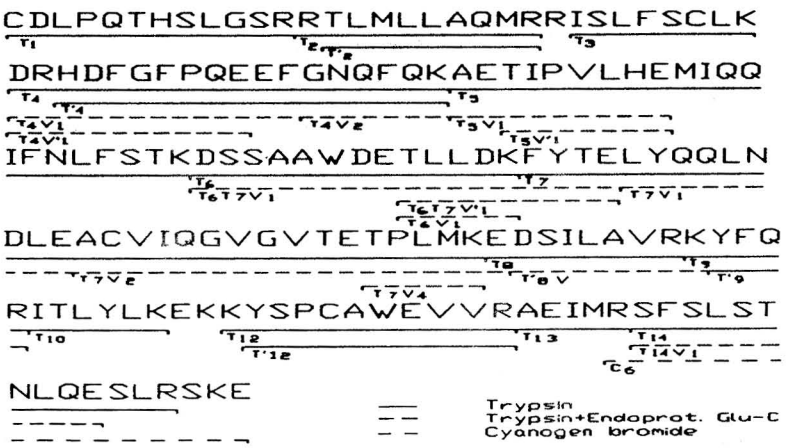

FIG. 2. Verification of amino acid sequence of α-IFN as predicted from gene sequence. Peptides obtained by digestion with trypsin, trypsin and endoproteinase Glu-C in tandem and cyanogen bromide are labeled with (T), (TV) and (C) respectively (98% of the sequence was verified).

Glu-C in tandem. The C-terminal end was verified by cyanogen bromide cleavage. Ninety-eight percent (98%) of the sequence was confirmed (Padrón *et al.*, 1989).

Location of S-S linkages

FAB-MS is a powerful tool to determine the location of the S-S bonds. The general strategy outlined in Fig. 1 is also applied. After enzymatic cleavage of a protein, peptides bonded by S-S bridges can be observed together with the signals of the corresponding reduced peptides (the use of reducing matrix as 2-mercaptoethanol or dithiotreitol could help in detecting the reduced peptides, but the atom bombardment is a reducing agent itself).

From this information it is possible to assign the cysteines linked by S-S bridges (Morris and Pucci, 1985; Yazdanparast *et al.*, 1986). Natural α-IFN has four cysteines with Cys1-Cys98 and Cys29-Cys138 bonded by disulphide bridges. The tryptic digest of α-IFN was purified by RP-HPLC and each peptide was analyzed by FAB-MS. Tryptic peptides at 2246.0, 2118.0, 4617.0 and 6049.4 confirmed the correct S-S bonds in the recombinant protein (Table 1, see also. Fig. 2) (Padrón *et al.*, 1989).

In a similar way, the three disulphide bridges of EGF were confirmed. EGF was mildly hydrolyzed with 2% acetic acid and the correct S-S bonds were detected (Fig. 3) (Besada *et al.*, 1990).

Incorrect S-S linkages

Alpha-IFN is obtained in *E. coli* as an insoluble product. The renaturation process was followed by RP-HPLC and FAB-MS. Two peaks were always observed in the RP-HPLC profile (peaks 1 and 2 in Fig. 4), but in different experiments, third peaks were also obtained with small changes in retention time, but with variable yields (Fig. 4B). These components were digested with trypsin and the purified peptides were analyzed by MS, showing scrambled S-S bonds. Mass spectra of cysteine containing peptides gave signals at m/z 2221.9, 4112.7 and 4510.8. These values were found to correspond to the tryptic fragments T1-T7, T'3-T7 and T7-T'12 linked by disulphide bonds between Cys1-Cys29, Cys29-Cys98, and Cys98-Cys138, respectively (Table 2). Finally, a process without S-S scrambling was established (Padrón *et al.*, 1989).

Table 1
OBSERVED AND EXPECTED MASS VALUES OF TRYPTIC PEPTIDES OF α-IFN BONDED BY DISULPHIDE BRIDGES.

Peptides bonded by disulphide bridges		Observed mass values	Expected mass values
Ile24-Lys31	Lys134-Arg144	2246.0	2246.1
Ile24-Lys31	Tyr135-Arg144	2118.0	2118.1
Cys1-Arg12	Phe84-Lys112	4617.0	4616.4
Cys1-Arg12	Asp71-Lys112	6049.4	6048.8

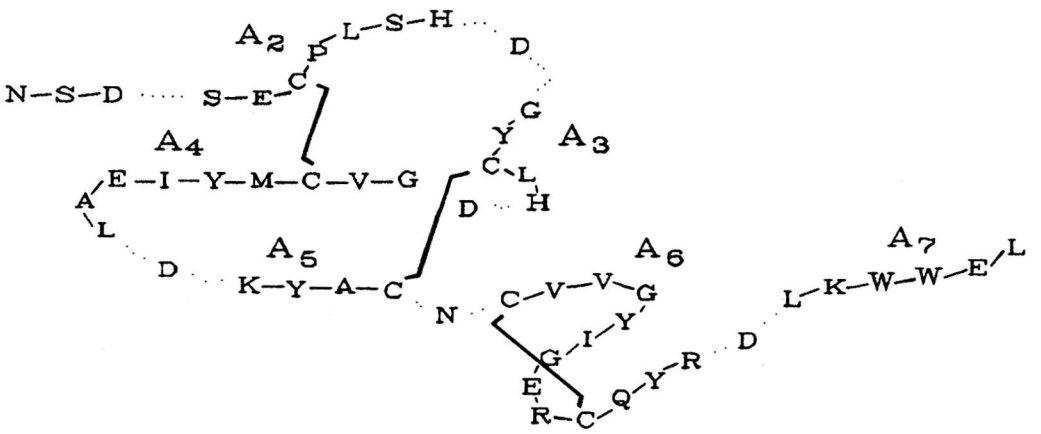

FIG. 3. Location of S-S bonds in EGF. Peptides are obtained by mild acid hydrolisis of EGF. Aspartic acids (or asparagines) were selectively released. Observed and theoretical (in parenthesis) mass values of peptides involved in disulphide bridges are shown.

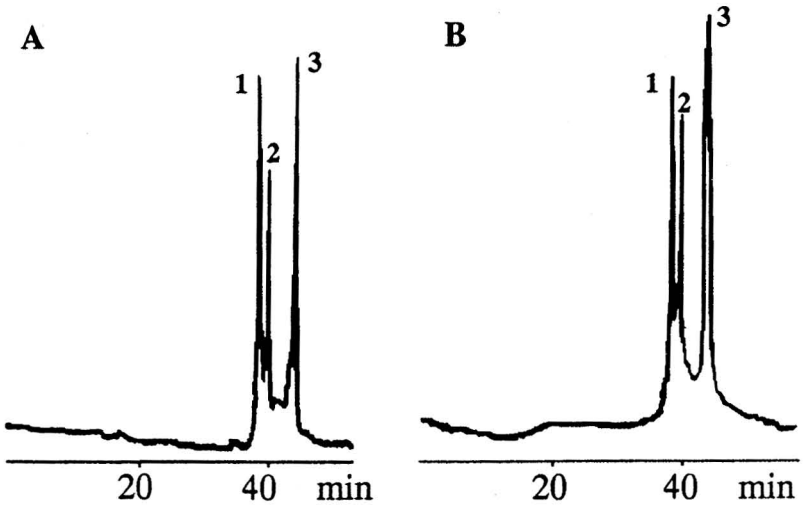

FIG. 4. RP-HPLC profile for α-IFN obtained from *E: coli* and renaturated by different processes. Peaks 1 and 2 correspond to normal and acetylated α-IFN, respectively. At peak 3 elutes an artifact with 2-mercaptoethanol bonded to Cys1 and Cys98 (Fig. A) or α-IFN with scrambled disulphide bonds (Fig. B).

Table 2
OBSERVED AND EXPECTED MASS VALUES OF TRYPTIC PEPTIDES OF α-IFN BONDED BY INCORRECT DISULPHIDE BRIDGES

Peptides bonded by incorrect disulphide bridges		Observed mass values	Expected mass values
Cys1-Arg12	Ile24-Lys31	2221.9	2222.1
Cys1-Arg12	Lys134-Arg144	2521.3	2521.2
Phe84-Lys112	Tyr135-Arg144	4510.8	4511.2
Ile24-Lys31	Phe84-Lys112	4212.7	4212.1

MODIFICATIONS AT THE N-TERMINAL END

The addition of a methionine at the N-terminal end is the most common modification of a protein obtained in *E. coli* because the methionine coded by the initiation codon is not efficiently removed. α-IFN contained only a very low percent of Met at the N-terminus (the tryptic peptide T1 was shifted from 1313.5 to 1444.5). However, in γ-IFN (Pérez *et al.*, 1990) and IL-2, practically 100% of the protein has a Met at the N-terminal end.

Mass Spectrometry becomes a very valuable tool for detection and identification of blocked N-terminal end, overcoming the inconveniences of techniques based on Edman degradation.

The second peak observed in the RP-HPLC profile of α-IFN (Figs. 4a-b) corresponded to the acetylated α-IFN. Its sequence was identical to the sequence of natural α-IFN, but peptide T1 (see Fig. 2) and peptides containing the S-S bond Cys1-Cys98 were shifted 42 Da (Fig. 5), due to the acetyl group at the N-terminal end.

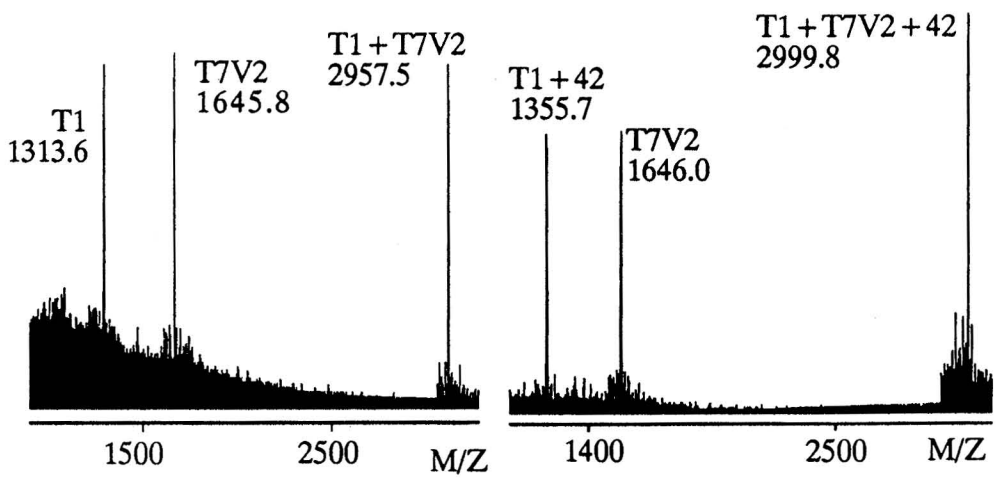

FIG. 5. Alpha-IFN obtained from yeast was predominately N-terminal acetylated (about 70%). FAB mass spectra of normal (left) and acetylated (right) peptides $T_1+T_7V_2$ (disulphide linked, see fig. 2). Reduced peptides T_1 and T_7V_2 are also observed. Peptides T_1 and $T_1+T_7V_2$ containing the N-terminal Cys_1 are shifted 42 Da.

It is apparent that α-IFN shows a special susceptibility to acylation because acetylation was also observed when it is obtained from yeast (about 70%), and formylation (about 35%) and acetylation (about 15%) occurred when it is produced from *Pseudomonas* (Fig. 6).

N-acetyl methionine has been observed blocking the N-terminal residue of some proteins. This residue can be released after CNBr treatment, permethylated and quantitatively identified by gas chromatography-mass spectrometry (Hemling *et al.*, 1988).

Ragged C-terminal end

Automatic sequencing is useful to detect degradation at the N-terminal end of proteins, but degradations at C-terminus are very difficult to verify. The use of MS makes any degradation at N or C-terminal ends relatively easy to detect.

The RP-HPLC profile of EGF produced from yeast always showed two main peaks, but the relative intensities were different from batch to batch (Fig. 7). Each component (peaks 1 and 2 in Fig. 7) was purified by RP-LC and analyzed by FAB-MS (Fig. 8). Observed molecular ion cluster peaks (centroid mode) at 5948.4 and 6061.3, respectively, suggested that the only difference between peaks 1 and 2 is at the C-terminal end. Peak 2 lacks the C-terminal arginine and peak 1 lacks the two C-terminal residues leucine- arginine. This result was confirmed digesting each component (reduced and carboxymethylated) with endoproteinases Glu-C and Lys-C separately and studying the peptides containing the C-terminal end by FAB-MS (Besada *et al.*, 1990).

In some experiments, higher degradation of the C-terminus of EGF was observed. Peaks with molecular weights of 5318.4, 5446.3 and 5632.8 were detected,

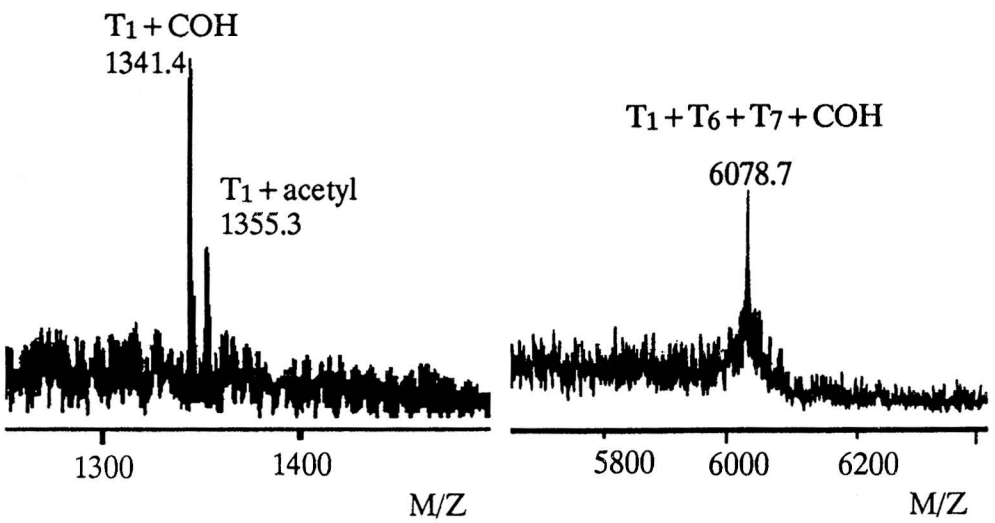

FIG. 6. Formylation at the N-terminus was detected in α-IFN expressed in pseudomonas. Tryptic peptides T_1 and $T_1+T_6T_7$ are observed at 1341.4 and 6078.7, shifted 28 Da (see Fig. 2 and Table 1). A minor population was N-acetylated (peptide T_1 at 1355.3).

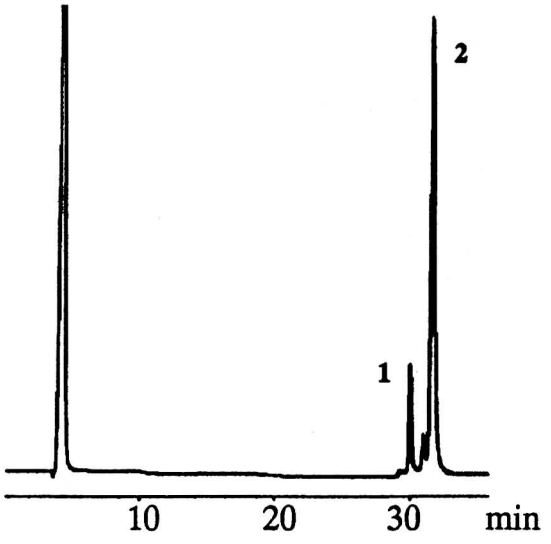

FIG. 7. RP-HPLC profile of EGF produced from yeast.

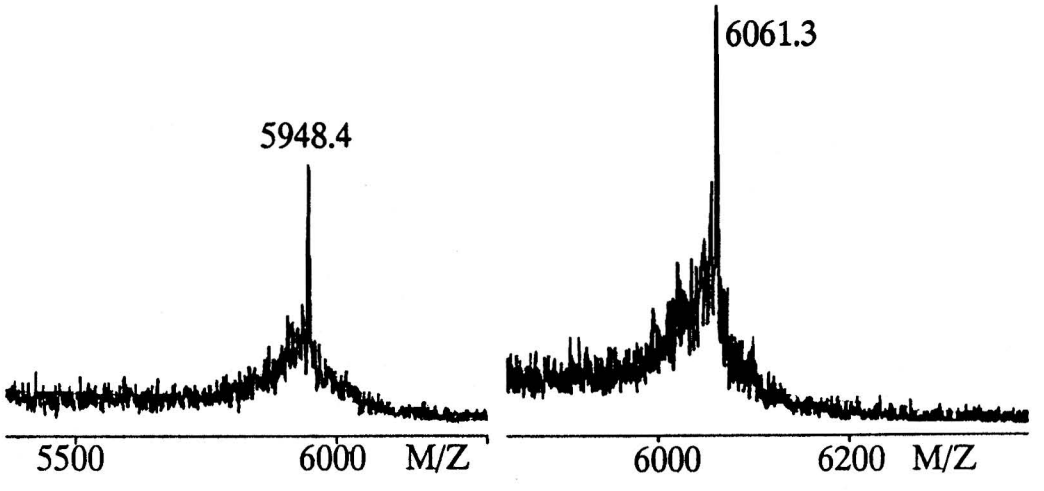

FIG. 8. FAB mass spectra of peaks 1 and 2 (in fig. 7) that correspond to the pseudomolecular ions of EGF containing 51 (left) and 52 (right) amino acids, respectively.

corresponding to the loss of 6, 5 and 4 amino acids from C-terminal end respectively (Besada *et al.*, 1990).

A significant reduction of the enzymatic degradation was obtained by controlling the fermentation process and increasing the speed of the purification process.

We have obtained human γ-IFN from *E. coli* and have verified the integrity of its C-terminus (Pérez *et al.*, 1990). However, γ-IFN has also been reported with ragged C-terminal end (Morris and Greer, 1988).

In figure 9 is reported the mass spectrum of the C-terminal peptide of a bacterial membrane protein. Four peaks are observed with succesive loss of lysine. This suggests that the protein has been proteolytically clipped at the C-terminus.

ARTIFACTS

During renaturation or purification processes, formation of different artifacts could take place. Oxidation of methionines or deamidation of asparagine or glutamine have been reported (Besada *et al.*, 1990; Carr *et al.*, 1990), but other less common artifacts could also occur.

In the first step of the renaturation process, α-IFN was dissolved in guanidinium chloride. In some experiments, 2-mercapto-ethanol was also used and the corresponding chromatograms showed a third component (Fig. 4A), that was found to be a protein with a sequence equal to that of α-IFN. However, the mass spectra of tryptic peptides isolated by RP-HPLC gave no signals related to the peptides with the disulphide bond between Cys1 and Cys98. Instead, the peptides containing Cys1 or Cys98 obtained in its tryptic digest showed signals at 1313.3 and 1389.3 or 3304.0 and 3380.5, respectively. The values 1313.3 and 3304.0 were identical with the theoretical mass values of peptides T1 and T7, and the values 1389.3 and 3380.5 corresponding to peptides T1 and T7 shifted 76 Da. This means that, instead

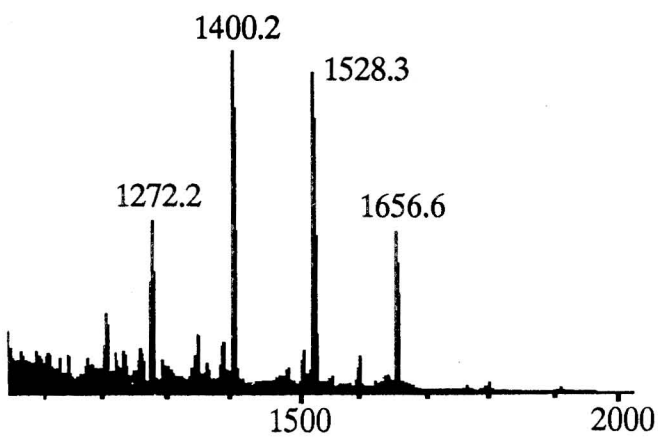

FIG. 9. Chymotryptic peptides corresponding to the C-terminal end of a bacterial membrane protein. The spectrum shows succesive loss of lysine, probably due to the action of carboxypeptidases during the isolation process.

of forming a disulphide bridge between Cys1 and Cys98, each cycteine was linked to a 2-mercaptoethanol molecule by a disulphide bond (Padrón *et al.*, 1989).

GLYCOSYLATION

Glycosylation is a common post-translational modification of proteins produced in eukaryotic cells. There are two types of glycosylation: N and O-glycosylation, and both can be detected by FAB-MS. The strategy for the detection of N-glycosylation is similar to the general strategy in Fig. 1, but including deglycosylation steps with endoglycosidase H (Endo H) and/or Peptide N-glycosidase F (PNGase F) (Carr and Roberts, 1986; Carr *et al.*, 1988).

The enzyme Endo H cleaves the β-glycosydic bond between the two N-acetylglucosamines attached to asparagine (oligosaccharide chain, high mannose or hybrid type, must contain at least three mannose units), leaving an N-acetyl glucosamine molecule bonded to the asparagine. Therefore, a peptide, obtained by enzymatic proteolysis of an Endo H deglycosylated glycoprotein, containing a glycosylation site will be shifted 203 Da.

The enzyme PNGase F cleaves the protein-sugar bond, converting the asparagine into asparticacid, which weighs 1 Da more. Its use in tandem with Endo H permits the detection of shorter sugar chains that are resistant to Endo H. Furthermore, comparing the treatment of a glycoprotein with endoproteinase Asp-N (cleaves at N-terminal of aspartic acid) with and without previous deglycosylation with PNGase F simplifies the detection of the glycosylation sites because of the appearance of new cleavage sites.

These procedures not only allow the location of the glycosylation sites, but also the determination of the glycosylation heterogeneity; that is, the site could be wholly, partially or non-glycosylated. The example of α-amylase produced in yeast is shown in Fig. 10. Asn 320

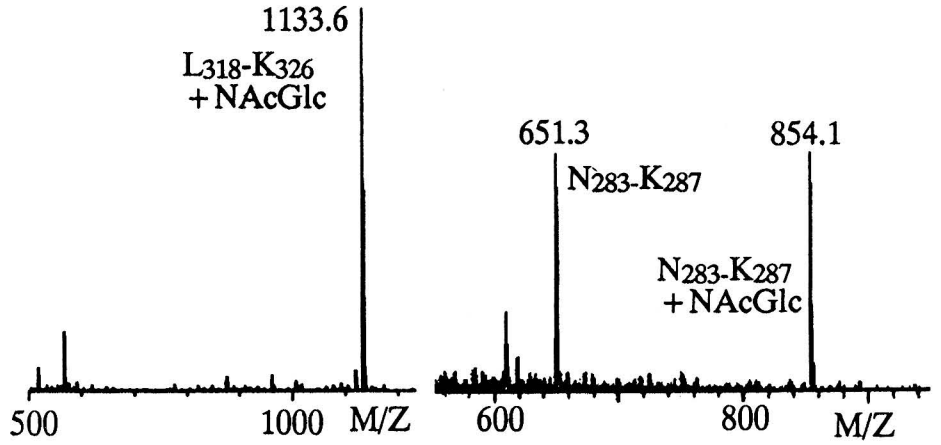

FIG. 10. FAB mass spectra of peptides obtained by digestion of α-Amylase with endoproteinases Glu-C and Lys-C in tandem. The protein was previously deglycosilated with Endo H. Asn 320 was always found glycosilated (peptides containing Asn 320 were shifted 203 Da), but Asn 286 was only partially glycosylated as indicated by the appearance of both peaks, at expected mass values and shifted 203 Da.

and Asn 286 are two of the glycosylation sites. The first one is practically fully glycosylated, but Asn 286 is only partially glycosylated (González, 1990).

We have also determined the glycosylation sites of mucorpepsin (3 potential glycosylation sites: Asn-X-Thr/Ser,X≠Pro). Only two of these sites were actually glycosylated and the third one was always found without any sugar bonded (Fig. 11) (Chinea, 1990).

By using this technique, it was found that sucrose invertase secreted in *Hansenula polymorpha* has a different glycosylation pattern than the one produced in *Saccharomyces cerevisiae* (result will be published elsewhere).

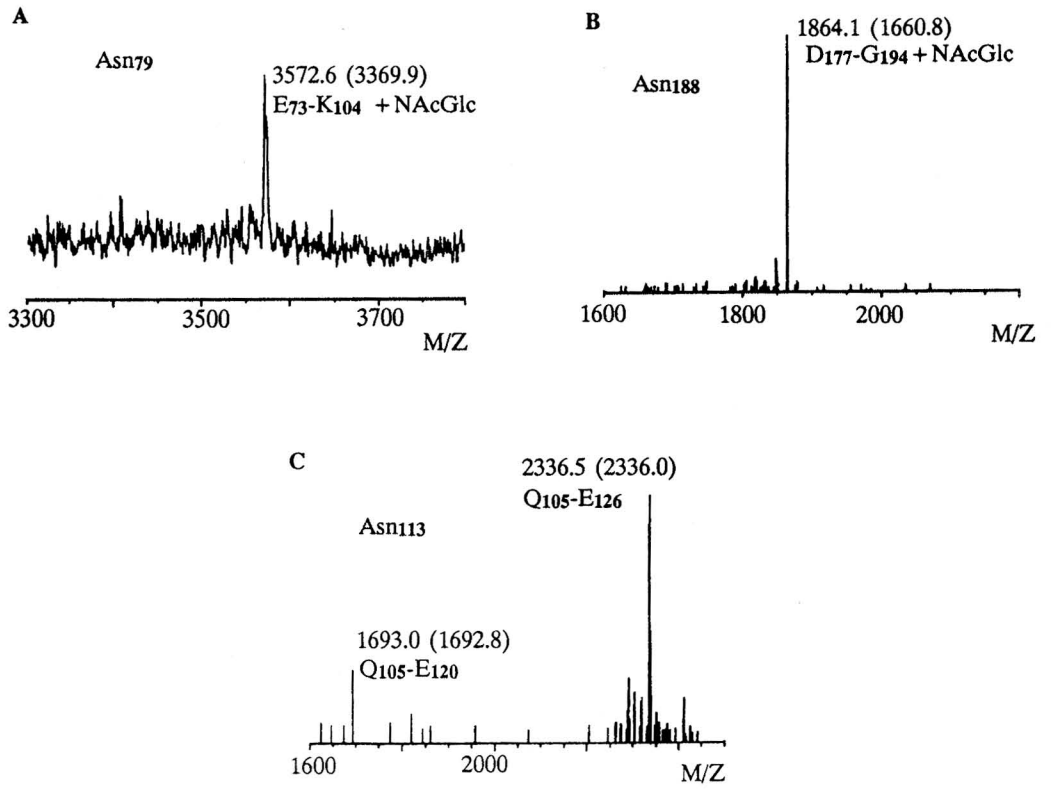

FIG. 11. Mass spectra of peptides containing the potential glycosylation sites of mucorpepsin. Peptides, in Figs. (A) and (C), were obtained by treatment with endoproteinases Lys-C and Glu-C in tandem, and peptide, in Fig. (B), was isolated from a digestion with endoproteinase Asp-N. Previously to enzymatic treatments, the protein was deglycosylated with Endo H. Theoretical mass values are shown in parenthesis. Observed mass values of peptides, in Figs. (A) and (B), are shifted 203 Da indicating that asparagines 79 and 108 are glycosylated.

Peptide sequencing by FAB-MS

The peptide sequencing by FAB-MS-MS using collisionally activated decomposition (CAD) is a very well known and efficient method (Johnson and Biemann, 1987; Johnson *et al.*, 1988). However, it is also possible to sequence a peptide using a single FAB-MS system with the help of manual Edman degradation and exo end endo proteinases (Bradley and Williams, 1982).

We have sequenced a 16 amino acid peptide combining FAB-MS and digestions with carboxypeptidase P, aminopeptidase M and endoproteinase Glu-C.

The molecular weight of the peptide was determined by FAB-MS (1881.2). Seven amino acids at C-terminal were determined by digestion with carboxypeptidase P: -VNDDAIK (Figs. 12 A and B). Mass spectra are measured at different times after the addition of an exopeptidase and the sequential loss of amino acids can be deduced.

Digesting the peptide with amino-peptidase M made it possible to sequence four amino acids at N-terminal: SVVH (Fig. 12 C).

The peptide was digested with endoproteinase Glu-C. Two new peptides were obtained (Fig. 12 D). One of them (MW 1343.4) was digested with aminopeptidase M and seven amino acids were sequenced (Fig. 12 E). The last two amino acids overlapped with those sequenced by carboxypeptidase P (Fig. 12 A and B). The mass difference between the molecular weight of the peptide and the sum of the amino acids already sequenced clearly indicates the existence of another aspartic acid and that also agrees with the specificity of the endoproteinase-Glu-C. Therefore, the sequence of the peptide is: SVVHDVFWHVNDDAIK.

By a similar strategy, a 55-amino-acid protein proteinase inhibitor was completely sequenced (unpublished work).

Perspectives in Mass Spectrometry

Instruments with new ionization systems are now on the market.

A laser desorption ionization system and a time-of-flight MS have been used for protein analysis, and molecular weights higher than 200,000 have been measured (Karas *et al.*, 1989).

Electrospray ionization is also very useful for protein analysis (Fenn *et al.*, 1989). Nebulization of liquid in vacuum under the ifnluence of an electric field liberates intact molecules and produces a coherent series of multiply protonated molecular ions. The charge state of each of these peaks can be easily determined, and from this information the molecular weight is obtained. As the m/z ratio is recorded, standard quadrupole mass spectrometers can be used (up to 2000 mass range) for analysis of high molecular weight proteins. A quadrupole MS-MS system with electrospray ionization is able to measure the molecular weight of a protein, verify the sequence, detect post-translational modifications or sequence the protein, using pmoles or possibly less. Combination with microbore HPLC or with capillary electrophoresis increases its potentiality.

Mass spectrometry promises to become one of the most important, essential and powerful tools in the analysis of proteins and other biomolecules.

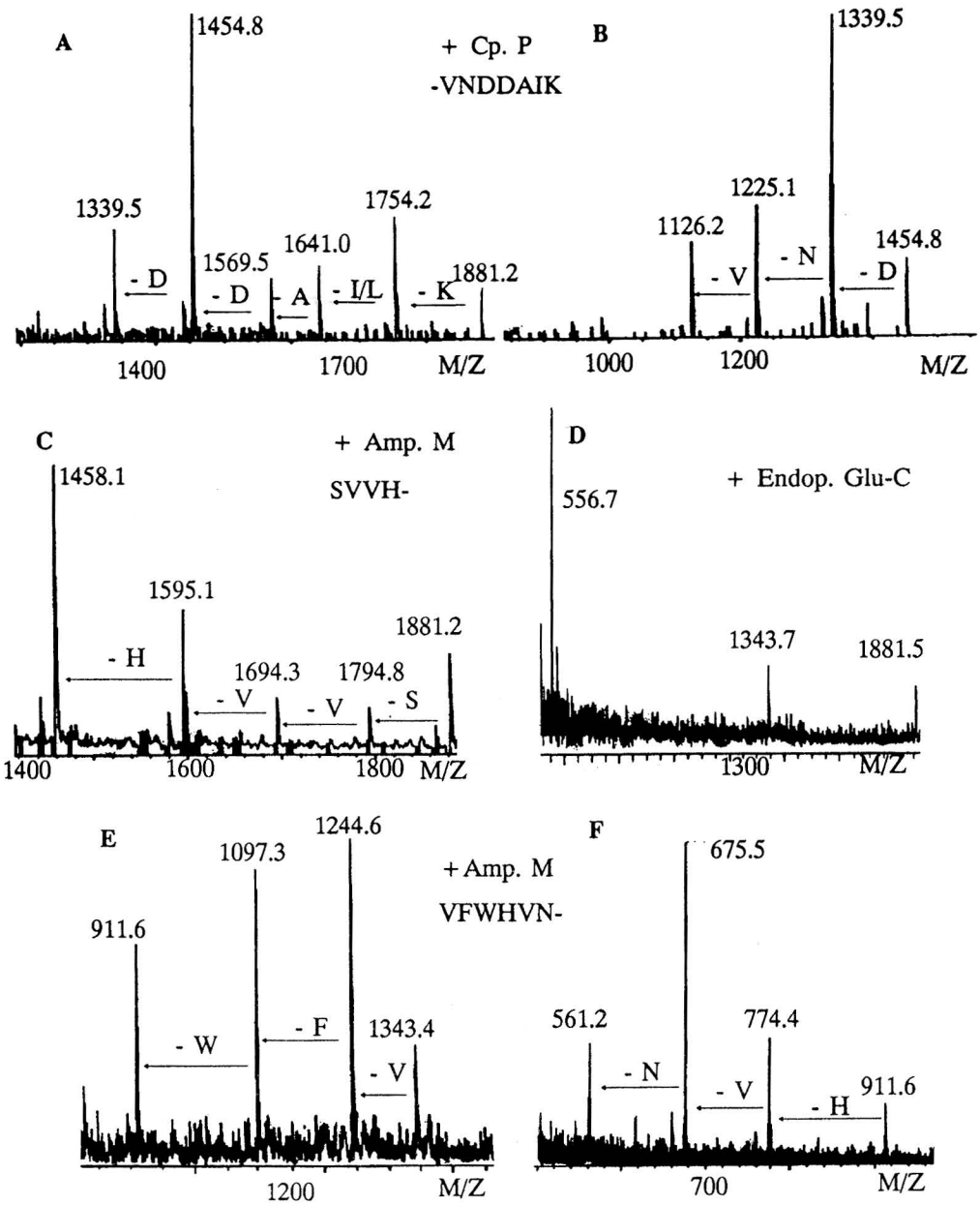

FIG. 12. Sequencing of a 16 amino acid peptide by FAB-MS combined with endo and exopeptidases. (A) and (B) show the digestion with carboxypeptidase P and (C) the digestion with aminopeptidase M. The peptide was also digested with endoproteinase Glu-C (D) and the resultant peptide was further digested with aminopeptidase M (E and F). Spectra (B) and (F) were obtained using greater amounts of the enzyme and longer digestion times than spectra (A) and (E), respectively. The sequence of the peptide can be deduced from these spectra (A to F) as SVVHDVFWHVNDDAIK (Ile was confirmed by amino acid analysis).

REFERENCES

BARBER, M.; R.S. BORDOLI; R.D. SEDGWICK and A.N. TYLER (1981). Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry. *J. Chem. Soc. Chem. Commun.* pp. 325-327.

BARBER, M.; R.S. BORDOLI; R.D. SEDGWICK and A.N. TYLER (1981). Fast atom bombardment of bradikynin and related oligopeptides. *Biomed. Mass Spectrom.* **8**: 337-342.

BESADA, V.; W. ANTUCH; A. CINZA; I. ROJAS; M. QUINTANA; G. PADRON; T. TAKAO and Y. SHIMONISHI (1990). Chemical characterization of recombinant human epidermal growth factor. *Anal. Chim. Acta* **239**: 301-305.

BRADLEY, C.V. and D.H. WILLIAMS (1982). Peptide sequencing using the combination of Edman degradation, carboxypeptidase digestion and fast atom bombardment mass spectrometry. *Biochem. Biophys. Res. Commun.* **104**: 1223-1230.

CARR, S.A. and G.D. ROBERTS (1986). A novel method for identifying attachment sites of Asn-linked sugars in glycoproteins. *Anal. Biochem.* **157**: 396-406.

CARR, S.A.; G.D. ROBERTS; A. JUREWICZ and B. FREDERICK (1988). Structural fingerprinting of Asn-linked carbohydrates from specific attachment sites in glycoproteins by mass spectrometry: application to tissue plasminogen activator. *Biochemie* **70**: 1445-1454.

CARR, S.A.; G.D. ROBERTS and M.E. HEMLING (1990). "Structural analysis of posttranslationally modified proteins by mass spectrometry". In: *Mass Spectrometry of Biological Materials*, pp. 87-136. C.N. McEwen and B.S. Larsen, Eds., Marcel Dekker, Inc.

CHINEA, G. (1990). *Characterization of recombinant mucorpepsin by mass spectrometry.* M. Sc. Dissertation, Inst. of Nuclear Science and Technology, La Habana.

FENN, J.B.; M. MANN; C.K. MENG; S.F. WONG and C.M. WHITEHOUSE (1989). Electrospray ionization for mass spectrometry of large molecules. *Science* **246**: 64-71.

FUKUHARA, K.; T. TSUJI; K. TOI; T. TAKAO and Y. SHIMONISHI (1985). Verification by mass spectrometry of the primary structure of human interleukine-2. *J. Biol. Chem.* **260**: 10487-10494.

GIBSON, B. and K. BIEMANN (1984). Strategy for the mass spectrometric verification and correction of the primary structures of proteins deduced from their DNA sequences. *Proc. Natl. Acad Sci. U.S.A.* **81**: 1956-1960.

GONZALEZ, L. (1990). *Study of recombinant proteins by mass spectrometry.* M. Sc. Dissertation, Fac. of Chemistry, Havana University.

HEMLING, M.E.; S.A. CARR; C. CAPIAU and J. PETRE (1988). Structural characterization of recombinant hepatitis B surface antigen protein by mass spectrometry. *Biochemistry* **27**: 699-705.

JOHNSON, R.S. and K. BIEMANN (1987). The primary structure of thioredoxin from *Chromatium vinosum* determined by high-performance tandem mass spectrometry. *Biochemistry* **26**: 1209-1214.

JOHNSON, R.S.; W.R. MATHEWS; K. BIEMANN and S. HOPPER (1988). Amino acid sequence of thioredoxin isolated from rabbit bone marrow determined by tandem mass spectrometry. *J. Biol. Chem.* **263**: 9589-9597.

KARAS, M.; U. BAHR; A. INGENDOH and F. HILLENKAMP (1989). Laser desorption/ionization mass spectrometry of mass 100 000 to 250 000 dalton. *Angew. Chem., Int.* Ed. Engl., **28**: 760-761.

MORRIS, H.R. and F.M. GREER (1988). Mass spectrometry of natural and recombinant proteins and glycoproteins. *Trends in Biotechnology* **6**: 140-147.

MORRIS, H.R. and P. PUCCI (1985). A new method for rapid assignment of S-S bridges in proteins. *Biochem. Biophys. Res. Commun.* **126**: 1122-1128.

MORRIS, H.R.; M. PANICO; M. BARBER; R.S. BORDOLI; R.D. SEDGWICK and A.N. TYLER (1981). Fast atom bombardment a new mass spectrometric method for peptide sequence analysis. *Biochem. Biophys. Res. Commun.* **101**: 623-631.

MORRIS, N.R.; M. PANICO and G.W. TAYLOR (1983). FAB-mapping of recombinant-DNA protein products. *Biochem. Biophys. Res. Commun.* **117**: 299-305.

PADRON, G.; V. BESADA; A. AGRAZ; Y. QUIÑONES; L. HERRERA; Y. SHIMONISHI and T. TAKAO (1989). Mass spectrometric analysis of recombinant human alpha-2 interferon. *Anal. Chim. Acta* **223**: 361-369.

PEREZ, L.; J. VEGA; C. CHUAY; A. MENENDEZ; R. UBIETA; M. MONTERO; G. PADRON; A. SILVA; C. SANTIZO; V. BESADA and L. HERRERA (1990). Production and characterization of human gamma interferon from *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **33**: 429-434.

TAKAO, T.; T. ITOUYE; Y. SHIMONISHI; T. TANABE; S. INOUYE and M. INOUYE (1984). Verification of protein sequence by fast atom bombardment mass spectrometry. Amino acid sequence of protein S. A development specific protein of *Myxococcus xanthus*. *J. Biol. Chem.* **259**: 6105-6109.

TAKAO, T.; M. KOBAYASHI; O. NISHIMURA and Y. SHIMONISHI (1987). Chemical characterization of recombinant human leukocyte interferon A using fast atom bombardment mass spectrometry. *J. Biol. Chem.* **262**: 3541-3547.

YAZDANPARAST, R.; P.C. ANDREWS; D.L. SMITH and J.E. DIXON (1986). A new approach for detection and assignment of disulfide bonds in peptides. *Anal. Biochem.* **153**: 348-353.