APPLICATION OF MASS SPECTROMETRY TO THE ANALYSIS OF PROTEINS CONTAINING A *N*-TERMINAL PYROGLUTAMIC ACID RESIDUE*

G.H.de HAAS
Biochemisch Laboratorium der Rijksuniversiteit, Utrecht, Holland

F.FRANĚK and B.KEIL
Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Prague, Czechoslovakia

and

D.W.THOMAS and E.LEDERER
Institut de Chimie des Substances Naturelles, C.N.R.S., 91-Gif-sur-Yvette, France

Received 5 June 1969

1. Introduction

Mass spectrometry of permethylated peptide derivatives allows the sequence determination of amino acids up to approximately ten *N*-terminal residues (for a review, see ref. 1). The preceding paper of this series [2] has described the sequence analysis of an octadecapeptide isolated from a λ-chain of an immunoglobulin, which, after permethylation by the technique of Coggins and Benoiton [3], gave a mass spectrum showing the sequence of the ten *N*-terminal amino acids. It was shown on that occasion that mass spectrometry has the advantage over the usual chemical degradation methods of distinguishing unequivocally glutaminyl (and asparaginyl) residues from glutamic (and aspartic) acid residues.

Another advantage of mass spectrometry, which is illustrated in the present paper, is the easy determination of *N*-terminal sequences of peptides having no free NH_2 group, which are difficult to determine by conventional degradation methods. The two examples described below concern peptides having each an *N*-terminal pyroglutamyl residue.

2. Results

2.1. *Heptapeptide of the zymogen of phospholipase A*

Studies in Utrecht [4] on phospholipase A from pig pancreas had indicated that this enzyme occurs mainly in the form of an enzymatically inactive zymogen molecule which can be converted by trypsin catalysis into the active enzyme. The conversion was accompanied by the splitting off of one peptide which had the following composition: Arg, Glu_2, Gly, Ile, Ser_2. This peptide contained an NH_2-protected glumatic acid residue, which was thought to be either *N*-acetylglutamic or pyroglutamic acid, but chemical studies could not provide an unambiguous answer to it's structure.

Mass spectrometry was considered as a solution to

* Part XVIII in the series "Determination of amino acid sequences in oligopeptides by mass spectrometry"; part XVII, reference 2.

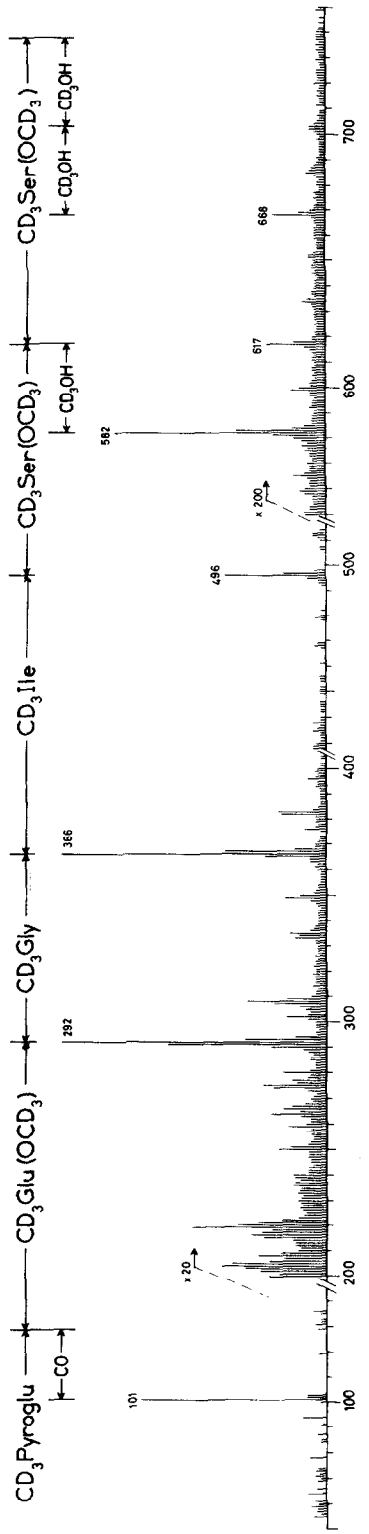

Fig. 1. Mass spectrum of *O,N*-perdeuteromethylated *N*-terminal heptapeptide from phospholipase A zymogen. Scale *m/e*.

this structural problem, but initial attempts with the methyl ester were not successful, due to insufficient volatility. However, after *O,N*-permethylation (with CD_3I and NaH in DMF [3]), an excellent spectrum was obtained (fig. 1) which demonstrated clearly the following sequence:

$$CD_3Pyroglu.CD_3\underset{\underset{OCD_3}{|}}{Glu}.CD_3Gly.CD_3Ile.CD_3\underset{\underset{OCD_3}{|}}{Ser}....$$

$$CD_3\underset{\underset{OCD_3}{|}}{Ser}....$$

In this spectrum, peaks are observed for losses of CD_3OH from the serine residues, which are characteristic of methylated serine and threonine peptides [5]. The arginine residue is not observed, but was already known to be present at the *C*-terminus due to preparation of the peptide by trypsin hydrolysis. Thus, the structure of this "activation" peptide is:

Pyroglu.Glu.Gly.Ile.Ser.Ser.Arg.OH

2.2. *Docosapeptide of pig immunoglobulin*

The *N*-terminal docosapeptide from the λ-chain of pig immunoglobulin was prepared in Prague [6]. By analogy with immunoglobulins of other animal species, there was a strong suspicion that the *N*-terminal residue was pyroglutamic acid, and several amino acid replacements were expected in the chain. However, because the *N*-terminal amino group was blocked, routine methods for sequential degradation were not possible. It was considered that even a short *N*-terminal sequence determined by mass spectrometry would be useful.

After permethylation with CH_3I and NaH in DMF, a mass spectrum was obtained (fig. 2) which permitted the identification of the six *N*-terminal residues. One clear sequence accounted for about 75% of the total material:

Pyroglu.Thr.Val.Leu*.GlN.Glu....

The expected *N*-terminal pyroglutamic acid was indicated by an intense *m/e* 98 peak (not shown in fig. 2, but corresponding to the trideuterated fragment at

* Leu or Ile.

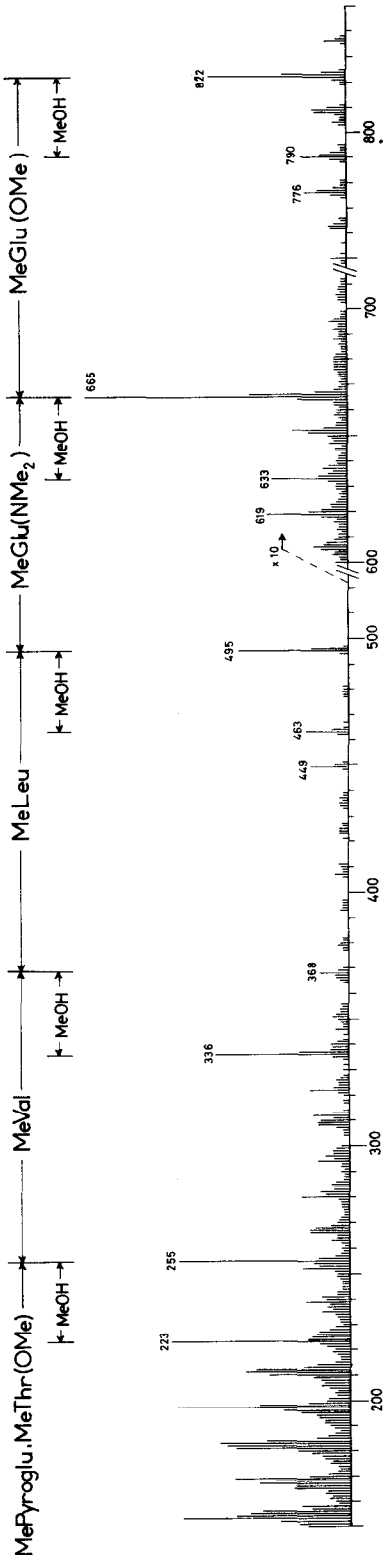

Fig. 2. Mass spectrum of *O,N*-permethylated *N*-terminal docosapeptide from pig immunogl ulin. Scale *m/e*.

m/e 101 in fig. 1) due to the fragment

$$CH_3\overset{+}{N}=CHCH_2CH_2CO$$

which is characteristic of this residue.

Peaks at *m/e* 449, 619, and 776 indicate a second component which contains amino acid replacements at one or more of the positions 2, 3 or 4. The following sequences may account for these peaks:

MePyroglu.MeSer.MeVal.MeLeu.MeGlu.MeGlu...
OME NMe$_2$ OMe

-MeOH 209 322 449 619 776

MePyroglu.Pro.MeVal.MeVal.MeGlu.MeGlu...
NMe$_2$ OMe

223 336 449 619 776

These two possibilities are still under investigation.

3. Discussion

From the results presented here on these two peptides, it is clear that the presence of pyroglutamic acid presents no obstacle to amino acid sequence determination by mass spectrometry, and more remarkable is the success (even partial) of this method with a peptide of 22 amino acid residues. This represents the largest peptide yet analysed by mass spectrometry.*

References

[1] E.Lederer, 5th IUPAC Symposium on the Chemistry of Natural Compounds, London, July 1968. Pure Appl. Chem. 17 (1968) 489.
[2] F.Franěk, B.Keil, D.W.Thomas and E.Lederer, FEBS Letters 2 (1969) 309.
[3] J.Coggins and L.Benoiton, Abstracts of Papers, BIOL 18, 156th ACS National Meeting, Atlantic City, 1968.
[4] G.H.de Haas, N.M.Postema, W.Nieuwenhuizen and L.L.M. van Deenen, Biochem. Biophys. Acta 159 (1968) 118.
[5] D.W.Thomas, B.C.Das, S.D.Géro and E.Lederer, Biochem. Biophys. Res. Commun. 32 (1968) 199.
[6] F.Franěk and J.Novotny, European J. Biochem., in press.

* The insulin A chain of 21 residues (with the cysteines transformed to cysteic acids) has also given a mass spectrum after permethylation, which showed the first 5 *N*-terminal residues.