APPENDIX

The Use of Mass-Spectrometry to Complement Conventional Techniques for Protein Sequence Determination

By J. D. PRIDDLE

Laboratory of Molecular Biophysics, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.

(Received 24 July 1973)

It has been established that the fragmentation pattern of an oligopeptide under electron impact provides sequence information and that suitable chemical modification, particularly permethylation of amide nitrogen, can enhance volatility and regulate fragmentation to the point where the sequence is directly interpretable by inspection of the low-resolution mass spectrum. Early studies were in the most part done on large quantities of synthetic peptides or readily obtained peptide antibiotics. The technique has only recently been applied to peptides derived from proteins in quantities commonly used in conventional techniques of sequence determination (Morris *et al.*, 1971), and with success to peptides containing residues of histidine, arginine and cysteine (see *inter alia* Morris, 1972; Morris *et al.*, 1973).

Small quantities of peptide left over from conventional work on chicken muscle triose phosphate isomerase (Furth *et al.*, 1974, accompanying paper) have been studied by mass spectrometry by the methods of Morris *et al.* (1971) and Morris (1972).

The spectra obtained (see Figs. 1–4) confirm and extend the results of Furth *et al.* (1974) and show the applicability of the technique to peptides containing histidine and carboxymethylcysteine and to mixtures produced by proteolytic digestion of small quantities of peptide. Mass-spectrometric information has been obtained for 35 residues of the sequence of the protein, that is about 15% of the total.

It has been noted that the assignment of amide groups by mass spectrometry need not always be unambiguous. Whereas the data of Fig. 1 illustrate the direct determination of the position of an amidated acidic residue that could not be assigned on the basis of the electrophoretic mobility of the whole peptide, those of Fig. 4(a) illustrate a case in which the accepted interpretation of the mass spectrum contradicts the rather clear indication given by the acidic mobility of the whole peptide. However, the residue in question is at the C-terminus of the peptide and anomalous fragmentations are known to be possible in this position (Brunfeldt *et al.*, 1972).

Fig. 1. Mass spectrum (above m/e 120) of the derivative of peptide T58P

Peptide T58P (250 nmol) was acetylated and permethylated by the method of Morris *et al.* (1971). Probe temperature was 280°C, source temperature 210°C. To simplify this and subsequent Figures, the normal amino acid abbreviations are used to represent the modified residues. The mass differences of sequence ions do not, of course, distinguish leucine from isoleucine but, in the present cases, prior amino acid analysis showed that no ambiguity would arise. The sequence ions show the order Val-Val-Phe-Gln-. The ion at m/e 729 could represent a subsequent residue of threonine or result by the loss of carbon monoxide from a subsequent residue of glutamic acid. The ambiguity is fortunately a rare case and could be excluded in other peptides by a knowledge of their amino acid composition. Furth *et al.* (1974), with conventional methods, found the sequence Val-Val-Phe-Glx-Glx-Thr-Lys. Thus a combination of the two techniques can establish the amidation of the acidic residues as follows: Val-Val-Phe-Gln-Glu-Thr-Lys.

Fig. 2. Mass spectrum (above m/e 140) of the derivative of peptide T58Q

Peptide T58Q (105 nmol) was digested with carboxypeptidase B as described by Furth *et al.* (1974) and the reaction mixture dried, acetylated and permethylated by the rapid method of Morris (1972). Probe temperature was 360°C, source temperature 210°C. The sequence ions show the order Ile-Gly-Val-Ala-Ala-Gln-, with the peak at m/e 192 characteristic of a C-terminal tyrosine residue. The conventional techniques used by Furth *et al.* (1974), established the sequence Ile-Gly-Val-Ala-Ala-Gln-(Asn,CmCys)Tyr-Lys.

Fig. 3. Mass spectrum (above m/e 100) of the derivatives of a mixture of peptides T58RTh1 and T58RTh2

Peptide T58R (80 nmol) was digested with carboxypeptidase B then with thermolysin as described by Furth *et al.* (1974) and the reaction mixture dried, acetylated and permethylated by the rapid method of Morris (1972). The sequence ions for the order Ile-Ile-Tyr-Gly- were obtained at probe temperature 180°C and source temperature 250°C. ---- denotes peaks whose relative intensity increased on further increasing probe temperature, reaching the values shown at 265°C. They result from an 'in-chain' fragmentation (Morris *et al.*, 1971), possibly pyrolytic, at the tyrosine residue to give the C-terminal sequence Tyr-Gly-Gly-Ser. The resulting sequence Ile-Ile-Tyr-Gly-Gly-Ser agrees with that of Furth *et al.* (1974) for the fragment T58R Th1. The only evidence of the presence of the carboxymethylcysteine-containing fragment Th2 [Val-Thr-Gly-Gly(Asn,CmCys)Lys] is the N-terminal valine fragment at m/e 156.

Priddle & Offord (1974) have also found that the mass-spectrometric technique is especially useful in cases where consecutive identical residues can lead to ambiguity in conventional studies.

Mass-spectrometric data, in being of an essentially different kind to those obtained by any sequential chemical degradation, provide a useful complement to such methods of increasing the accuracy of protein-sequence determination.

Dr. R. T. Aplin, Dyson Perrins Laboratory made available the mass spectrometer used in this work. J. D. P. was supported by the Science Research Council through the Oxford Enzyme Group.

References

- Brunfeldt, K., Christiansen, T. & Roepstorff, P. (1972) *FEBS Lett.* **25**, 184-188
- Furth, A. J., Milman, J. D., Priddle, J. D. & Offord, R. E. (1974) *Biochem. J.* **139**, 11-22
- Morris, H. R. (1972) *FEBS Lett.* **22**, 257-260
- Morris, H. R., Williams, D. H. & Ambler, R. P. (1971) *Biochem. J.* **125**, 189-201
- Morris, H. R., Dickinson, R. J. & Williams, D. H. (1973) *Biochem. Biophys. Res. Commun.* **51**, 247-255
- Priddle, J. D. & Offord, R. E. (1974) *FEBS Lett.* in the press

Fig. 4. Mass spectra of derivatives of peptides T45aTh1, T45aTh2 and T45aTh3

Peptide T45a (400nmol) was digested with thermolysin and the major peptide products separated by paper electrophoresis as described by Furth *et al.* (1974). The peptides (<400nmol) were acetylated, and peptides Th1 and Th2 permethylated by the method of Morris *et al.* (1971) and peptide Th3 (which contained histidine) by the rapid method of Morris (1972). (a) The mass spectrum (above m/e 100) of the derivative of peptide T45aTh1. Probe temperature was 280°C, source temperature 250°C. The sequence ions for Ala-Gln- are followed by a prominent peak at m/e 468. This would ordinarily be interpreted as a subsequent residue of glutamine, but the electrophoretic mobility and other data of Furth *et al.* (1974) would normally be taken to establish unequivocally the sequence Ala-Gln-Glu. The origin of the spurious ($M-18$) peak is as yet unestablished. (b) The mass spectrum (above m/e 140) of the derivative of peptide T45aTh2. Probe temperature was 280°C, source temperature 210°C. The sequence ions establish the order Thr-Ala-Thr-Pro-Gln-, whereas Furth *et al.* (1974) found Thr-Ala-Thr-Pro-Gln-Gln. (c) The mass spectrum (above m/e 100) of the derivative of peptide T45aTh3. Although the amount of this peptide taken was, presumably, approximately equimolar with those of peptides Th1 and Th2, the mass spectrum obtained after derivative formation was considerably weaker than in the other two cases. Probe temperature was 240°C, source temperature 210°C. The sequence ions establish the order Val-His-Glu-. This fragment has the sequence Val-His-Glu-Lys from the data of Furth *et al.* (1974).