STUDY OF 'DIFFICULT PEPTIDES' FROM *PARACOCCUS* CYTOCHROME *c*-550 AND A DOLPHIN CYTOCHROME *c*

Fast atom bombardment: a new method for molecular weight and sequence determination of peptides

Dudley H. WILLIAMS, Gustav BOJESEN, Anthony D. AUFFRET and Lester C. E. TAYLOR+

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW and +*Kratos Ltd., Barton Dock Road, Urmston, Manchester, England*

Received 25 March 1981

1. Introduction

In sequence investigations of peptides by the dansyl-Edman method, 'difficult peptides' are occasionally encountered. Among such peptides are included those with acylated N-terminal amino acids, and those with N-terminal pyroglutamic acid. In these cases, the sequences can frequently be determined by mass spectrometry [1,2]. Other 'difficult peptides' are those which are relatively long, and for which enzymic subdigestion is difficult or impossible; the interpretation of data obtained by the dansyl-Edman method may then be ambiguous after a large number of cycles. We now report a mass spectrometric sequence study of three peptides which fall into the above categories.

Two of the peptides are derived from cytochrome *c*-550 from the denitrifying aerobic bacterium *Paracoccus denitrificans*. The amino acid sequence originally reported [3] for the N-terminal portion of the protein was acetyl–Asn–Glu–Gly–Asp–Ala–Ala–Lys, but a subsequent study [4] has shown that the terminal residue can be removed from a tryptric heptapeptide with pyroglutamate aminopeptidase, and the residual sequence determined as Asp–Gly–Asp–Ala–Ala–Lys. We have therefore sought to confirm the presence of N-terminal pyroglutamic acid in the heptapeptide. The second peptide from cytochrome *c*-550 to be studied is derived from the C-terminus of the protein, which has proved difficult to sequence by both automated sequencing [3] and dansyl-Edman [4] methods. The tentative sequence derived from the latter method [4] is Leu–Ala–Gln–Asn–Ser–Pro–Asp–Ala–Gly–Gly–Asp–Gly–Glu–Ala–(Ala); with evidence that about half the molecules in their protein contain the extra alanine shown in parentheses. This sequence differs in several respects from that originally reported [3].

The third peptide we have studied is from the N-terminal portion of a dolphin (*Coryphaena hippurus*) cytochrome *c*. Its probably sequence is acyl–Gly–Asp–Val–Ala–Lys (Ambler, R. P., personal communication); our aim was to determine independently the sequence, and to uncover the nature of the N-terminal blocking group.

2. Materials and methods

The peptides, provided by Dr R. P. Ambler, had been obtained and purified by the following methods,

(i) N-Terminal heptapeptide and C-terminal peptide from *Paracoccus denitrificans* cytochrome *c*-550: obtained by tryptic digest, and purified by gel filtration on Sephadex G-25 (developed with 5% formic acid), followed by high-voltage paper electrophoresis at pH 6.5 (pyridine acetate buffer). The peptide was eluted from paper with 0.1 M NH_3.

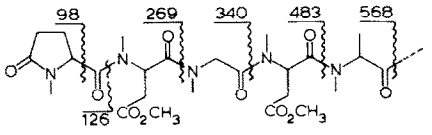
(ii) N-Terminal pentapeptide from dolphin (*Coryphaena hippurus*) cytochrome *c*: obtained as a tryptic fragment from a large staphylococcal protease peptide. Purified by high-voltage paper electrophoresis as described under (i).

Prior to obtaining electron impact spectra, the

peptides (approx. 100 nmol) were subjected to acetylation conditions, and then permethylated. Acetylation was carried out as in [5] using 0.5 ml of acetic anhydride (or [U-^{2}H]acetic anhydride) in methanol (1:4, v/v) for 3 h at room temperature. The short permethylation conditions of [6] were adopted: the peptide was dissolved in 25 μl dimethyl sulphoxide and reacted with an excess of methyl sulphinyl carbanion base in dimethyl sulphoxide, followed by addition of a large excess (40 μl) of methyl iodide (or [U-^{2}H]methyl iodide). Low resolution electron impact mass spectra were recorded on AEI MS 902 or Kratos MS 50 instruments operating at 8 kV and an electron beam energy of 70 eV. Fast atom bombardment spectra [7] were obtained by bombardment of the unmodified peptide, in a matrix of glycerol on a copper probe tip, with 4-kV argon atoms; these spectra were recorded on a Kratos MS 50 instrument at accelerating voltages between 4–8 kV.

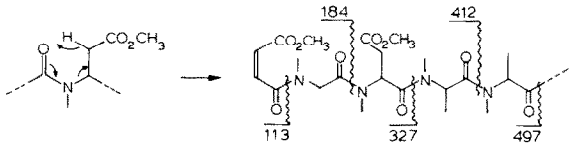
3. Results and discussion

Since the originally reported sequence [3] of the N-terminal portion of the *Paracoccus* cytochrome *c*-550 contained an *N*-acetyl group, acetylation of the peptide was carried out with [U-^{2}H]acetic anhydride. After permethylation, the electron impact mass spectrum of the product contained a base peak at m/z 98; this peak is characteristic of N-terminal Gln when the sequence ion m/z 126, from which it is formed by CO loss, is of low abundance (as in the present case). Abundant acylium ion sequence ions are observed at m/z 269, 340, 483 and 568, establishing the sequence <Glu–Asp–Gly–Asp–Ala . . . (see *1*).

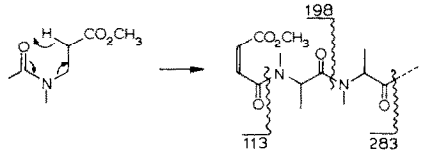


1

In-chain cleavage [2] occurs at both aspartic acid residues (*2* → *3* and *4* → *5*). Hence relatively abundant sequence ions are observed at m/z 184, 327, 412 and 497, indicating the sequence . . . Asp–Gly–Asp–Ala–Ala . . . (see *3*); and at m/z 198 and 283, indicating the sequence . . . Asp–Ala–Ala . . . (see *5*). Thus, the sequence <Glu–Asp–Gly–Asp–Ala–Ala . . . is established.



2 3



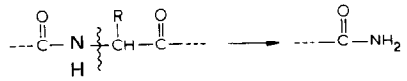
4 5

The sequence determination was completed by running both positive and negative ion FAB mass spectra of the underivatised peptide; the spectra showed abundant ions at m/z 687 (MH^+) and m/z 685 ($M\text{-}H^-$), respectively. The complete sequence is therefore <Glu–Asp–Gly–Asp–Ala–Ala–Lys.

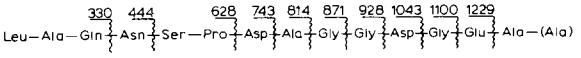
Prior studies (Ambler, R. P., personal communication) had shown that no significant digestion of the large C-terminal *Paracoccus* cytochrome *c*-550 peptide was achieved with trypsin, chymotrypsin, staphylococcus protease, or thermolysin. In our hands, attempted digestions with papain, pronase, and pyroglutamate aminopeptidase (hopefully utilising an impurity protease in the last of these enzymes) also failed. Following acetylation and permethylation, an N-terminal sequence Leu–Ala–Gln–Asn–Ser . . . was deduced from the electron impact spectrum from sequence ions at m/z 170, 255, 425, 581 and 696. This sequence was confirmed by the appropriate mass shifts of these ions when the permethylation was carried out with [U-^{2}H]methyl iodide. Limited digestion of the large C-terminal peptide was apparently achieved with subtilisin. When this digest was acetylated and then permethylated with [U-^{2}H]methyl iodide, the electron impact spectrum contained sequence ions at m/z 117, 280 and 368, indicating the sequence Gly–Glu–Ala . . .

In view of the difficulty encountered with this peptide, it was also subjected to the new technique of fast atom bombardment (FAB) [7]. The positive ion FAB spectrum gave two intense $[MH]^+$ ions at m/z 1301 and 1372 without prior derivatisation of the peptide, thus establishing it to be a mixture of two

components differing by the mass of one in-chain alanine. The spectrum contained less abundant sequence ions at m/z 330, 444, 628, 743, 814, 871, 928, 1043, 1100 and 1229. These sequence ions arise via cleavage of N–C_α bonds with the N atom having acquired, in the fragment, an additional hydrogen atom, viz:

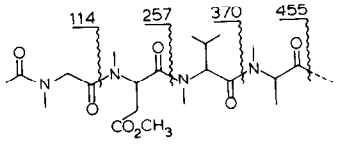


The positive charge of these fragments seems likely to arise due to protonation of the terminal amino group they contain. The sequence ions are due to the formation of fragments as indicated below; and establish, in conjunction with the electron impact data and the FAB molecular weight determinations of two components, the complete sequence which is given.



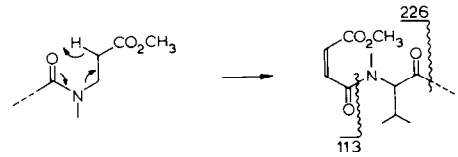
The sequence of the large C-terminal peptide, and its occurrence as a mixture of two components, is in accord with the tentative conclusions of [4].

The N-terminal peptide from dolphin cytochrome *c* was subjected to the acetylation reaction with [U-^{2}H]acetic anhydride, and permethylated. The electron impact mass spectrum of the product contained abundant sequence ions at m/z 114, 257, 370 and 455. The first sequence ion at m/z 114 corresponds to acetyl–Gly . . . (*6*). Thus, the peptide carried an *N*-acetyl group prior to derivatisation, and the absence of any peak at m/z 117 indicates the absence of any free N-terminal glycine in the original peptide. The sequence *N*-acetyl–Gly–Asp–Val–Ala . . . is determined (*6*), although the sequence ions did not extend to the C-terminal lysine which was believed to be present.



6

In-chain cleavage at the Asp residue gives rise to ions at m/z 113 and 226, independently confirming the Asp–Val . . . sequence (*7* → *8*).



7 8

The sequence determination of the peptide was completed by obtaining the negative ion FAB mass spectrum of the underivatised peptide. An abundant ($M-H^-$) ion was observed at m/z 529. Thus, the complete sequence is *N*-acetyl–Gly–Asp–Val–Ala–Lys.

The FAB technique clearly represents an extremely important development in the molecular weight and sequence determination of peptides. Ions are produced over a long period (minutes to approx. 1 h) and molecular weight determinations have been achieved down to the 1 nmol level (Williams, D. H., et al., unpublished results). The signal/noise ratio achieved with such sample sizes suggests that the limit for molecular weight determination can probably be lowered to 0.1 nmol of sample.

Acknowledgements

The Science Research Council and the Medical Research Council are thanked for financial support. We thank the authors of [4] for drawing these problems to our attention.

References

[1] Auffret, A. D., Williams, D. H. and Thatcher, D. R. (1978) FEBS Lett. 90, 324–326.
[2] Morris, H. R., Williams, D. H., Midwinter, G. G. and Hartley, B. S. (1974) Biochem. J. 141, 701–713.
[3] Timkovich, R., Dickerson, R. E. and Margoliash, E. (1976) J. Biol. Chem. 251, 2197–2206.
[4] Ambler, R. P., Meyer, T. E., Kaman, M. D., Schichman, S. A. and Sawyer, L. (1981) J. Mol. Biol., in the press.
[5] Thomas, D. W., Das, B. C., Gero, S. D. and Lederer, E. (1968) Biochem. Biophys. Res. Commun. 33, 519–521.
[6] Morris, H. R. (1972) FEBS Lett. 2, 257–260.
[7] Barber, M. et al. (1981) J. Chem. Soc. Chem. Commun., in the press.