A MASS SPECTROMETRIC STRATEGY FOR THE RAPID SCREENING OF HOMOLOGOUS PROTEINS: STUDIES ON A PSEUDOMONAS AZURIN

Anne Dell and Howard R. Morris

Department of Biochemistry, Imperial College, London, U.K.

Received August 23,1977

SUMMARY

We describe a mass spectrometric strategy for the rapid screening of homologous proteins. The method involves non-specific enzymic digestion followed by a single ion-exchange purification step. Mixtures of peptides are then sequenced by low resolution electron impact mass spectrometry. Peptides are aligned by homology with a type sequence thus obviating the necessity for obtaining overlap information. The procedure, which is rapid and reliable, is illustrated by work on a Pseudomonas azurin.

During the past few years there has been an accumulation of protein sequence data which has been made possible by the development of ever more sophisticated and efficient sequencing methods. The importance of these data in many fields of study e.g. structure-function relationships, enzyme catalysis, enzyme specificity, genetics and evolution is apparent and the continuing demand for rapid and reliable sequence analysis has stimulated the introduction of instrumentation and automation (1). Despite recent developments, the determination of the complete primary structure of a protein remains a long, tedious and expensive task and there is still scope for the improvement of methods. In particular, there is a need for sequence strategies which are tailor made for specific problems and which rapidly yield reliable data applicable to the overall programme under investigation. A pertinent example is the study of enzyme evolution where information on the nature and extent of

ISSN 0006-291X

variation in many homogous proteins is required rather than the total sequence of each. In this paper we describe a mass spectrometric sequencing strategy which is ideal for the rapid screening of homologous proteins and can be very usefully applied to evolution work.

Azurin, a respiratory protein from Pseudomonas fluoresens biotype G, was chosen for study because it appeared to be an ideal molecule with which to develop sequencing methods. It is suitably small (M.Wt. 13,000 dalton) and its sequence was unknown, although the primary structures of azurins from a number of different Pseudomonads had been established by Ambler and co-workers during their enzyme evolution studies (2,3). The object of our study was to develop a method which would yield a large number of random sequences. It was envisaged that these could then be aligned by homology with an already established type sequence (2) thus obviating the necessity for obtaining overlap information. A preliminary study (4) of a cyanogen bromide digest of azurin had yielded some useful sequence information but a number of difficulties encountered during the handling of large cyanogen bromide fragments made it apparent that a procedure utilising specific digestion could not be developed as a suitable general method for sequencing mutant proteins by mass spectrometry. In contrast, non-specific enzymic cleavage which had been used in other mass spectrometric sequencing work in this laboratory (5) proved to be highly successful. In this paper we report the results obtained from elastase and subtilisin digests of azurin and outline a general scheme for examining mutant proteins.

MATERIALS AND METHODS

Azurin (2.5 μmol) from P. fluorescens biotype G was

a gift from Dr. R.P. Ambler, University of Edinburgh. General methods of protein chemistry were as described by Shotton and Hartley (6). After removal of copper with trichloroacetic acid (2) the protein was digested with elastase (substrate to enzyme ratio 50:1) for 2 h. at 37°C and for a further 16 h. at 18°C. The digest was freeze dried, taken up in 0.1M NH_3(1 ml), applied to a Sephadex G25 column (80 x 1 cm) (flow rate 10 ml/h) and eluted with 0.1 M NH_3. The effluent was monitored at 280 nm. The void volume fraction was freeze dried and re-digested with subtilisin (substrate to enzyme ratio 60:1) for 5 h. at 37°C. The remaining fractions were pooled and freeze dried. Peptide analyser runs were carried out using a column (54 x 0.9 cm) containing Locarte cation exchange resin. The peptides were eluted using an eight chamber pyridine-acetic acid gradient. The gradient buffers (pH 3.1, 5.0 and 6.5) were prepared according to the published recipes (7) and the gradient was made up as follows: chambers 1 to 3:- 60 ml of pH 3.1; chamber 4:- 40 ml of pH 3.1 and 20 ml of pH 5.0; chamber 5:- 20 ml of pH 3.1 and 40 ml of pH 5.0; chamber 6:- 60 ml of pH 5.0; chamber 7:- 30 ml of pH 5.0 and 30 ml of pH 6.5; chamber 8:- 60 ml of pH 6.5. The pumping speed was 10 ml/h. and 180 x 2 ml fractions were collected. 1/20 of each fraction was examined by high voltage paper electrophoresis to determine the resolution of the column and allow a rational choice of fractions for pooling. Fractions were pooled, rotary evaporated to a small volume, diluted with water and freeze dried. 1/10 of each sample was taken for amino acid analysis and the remainder was derivatised according to the published procedures (8). All mass spectra were recorded on an AEI MS 902 mass spectometer. Sample handling procedures and instrument parameters have been described previously (9).

RESULTS AND DISCUSSION

Azurin (2.5 μmol) was digested with elastase and the digest passed down a Sephadex G25 column. The first fraction, which eluted in the void volume, was further digested with subtilisin. The remainder of the effluent was pooled and chromatographed on a peptide analyser (see experimental). Mixtures of derivatised peptides were then examined by mass spectrometry using the "mixture analysis" technique first described by Morris *et al* (9) and subsequently used in a number of protein studies (4, 5, 10, 11, 12). The sequences of peptides obtained from the elastase digest are given in Table 1. The subtilisin digest was treated in a similar manner, i.e. after partial separation of the complex digest

Table 1. Sequences of peptides obtained from an elastase digest of azurin.

Asp-Gly-Met-Ala-Ala	Asn-Leu-Pro-Lys-Asn
Thr-Val-Asp-Ser-Thr	Glu-Leu
Phe-Thr-Val-Glu-Leu-Thr	Lys-Leu$^{\downarrow}$-Leu$^{\downarrow}$-Gly-Glu
Asn-Met-Gln-Pro	Tyr-Leu-Lys-Pro
Val-Glu-Leu-Thr	Lys-Ala-Ile
Asp-Gly-Met-Ala	Leu-Lys-Glu-Ala*
Phe-Thr-Val-Glu-Leu	Lys-Glu-Leu$^{\downarrow}$
Thr-Phe-Asp-Val-Ser	His-Ser-Gly-Asn-Leu-Pro-Lys-As
Glu-Leu-Thr	Lys-Gly-Thr-Val-Thr
Ser-PheAsn-Thr	Lys-Ala
Leu-Glu-Ala*	Leu$^{\downarrow}$-Arg
Ala-Thr	Val-Lys
Asp-Tyr-Ala*	Lys-Ala-Ile
Phe-Phe	Cys-Lys-Thr
Ala-Ala	Ala-Glu-Cys-Lys
Gly-Asn-Leu-Pro-Lys-Asn	Val-Lys-Arg*
Leu$^{\downarrow}$-Glu-Leu$^{\downarrow}$*	Phe-Asn-Thr
PCA-Val-Leu$^{\downarrow}$-Glu*	Ser-Phe-Asn
Met-Gly-His-Asn-Trp-Val	Phe-Thr-Val

$^{\downarrow}$ indicates Leu and Ile not distinguished.
* indicates peptides not aligned with the type sequence.

on a peptide analyser the resulting peptide mixtures were examined by mass spectrometry. Data obtained from the subtilisin digest are presented in Table 2.

Most of the peptides in Tables 1 and 2 were easily aligned with the type sequence (azurin from P. aeruginosa) and Figure 1 shows the type sequence together with the peptides

Table 2. Sequences of peptides obtained from a subtilisin digest of high molecular weight material recovered from the elastase digest.

Thr-Val-Asp-Ser	Asn-Leu-Pro
Val-Asp-Ser-Thr	Gly-Ile-Asp-Lys-Asp-Tyr
Asp-Val-Ser	Leu-Lys-Pro
Leu$^{\downarrow}$-Ala*	Trp-Val-Leu
Leu$^{\downarrow}$-Glu-Leu$^{\downarrow}$*	Leu-Pro-Val-Val-Gly*
Ser-Phe-Asn-Thr	Val-Lys
Asp-Tyr	Lys-Ala
Ser-Val	Asp-Asp-Ile-Arg
Met-Ser-Phe	

$^{\downarrow}$ indicates Leu and Ile not distinguished.
* indicates peptides not aligned with the type sequence.

of the P. fluorescens azurin which have been aligned by homology. The few peptides whose positions in the sequence were not obvious, either because of total lack of homology of because they were too short to place with confidence, are indicated by an asterix in Tables 1 and 2. Approximately 85% of the sequence of azurin was obtained from the two digests of which 76% has been aligned with the type protein. This work was completed in the space of four weeks. Non-specific digestion coupled with mixture analysis mass spectrometry is clearly a suitable strategy for the rapid acquisition of a considerable quantity of sequence data.

In summary, a general procedure for sequencing mutant proteins by mass spectrometry is as follows:

1. digestion of the mutant with a non-specific enzyme e.g. elastase

P.aer. Ala-Glu-Cys-Ser-Val-Asp-Ile-Gln-Gly-Asn-Asp-Gln-Met-Gln-Phe-Asn-Thr-
P.fl. Ala-Glu-Cys-Lys- Met-Ser-Phe-Asn-Thr-

P.aer. Asn-Ala-Ile-Thr-Val-Asp-Lys-Ser-Cys-Lys-Gln-Phe-Thr-Val-Asn-Leu-Ser-
P.fl. Lys-Ala-Ile-Thr-Val-Asp-Ser-Thr-Cys-Lys-Thr-Phe-Thr-Val-Glu-Leu-Thr-

P.aer. His-Pro-Gly-Asn-Leu-Pro-Lys-Asn-Val-Met-Gly-His-Asn-Trp-Val-Leu-Ser-
P.fl. His-Ser-Gly-Asn-Leu-Pro-Lys-Asn- Met-Gly-His-Asn-Trp-Val-Leu-

P.aer. Thr-Ala-Ala-Asp-Met-Gln-Gly-Val-Val-Thr-Asp-Gly-Met-Ala-Ser-Gly-Leu-
P.fl. Ala-Ala-Asn-Met-Gln-Pro- Ala-Thr-Asp-Gly-Met-Ala-Ala-Gly-Ile-

P.aer. Asp-Lys-Asp-Tyr-Leu-Lys-Pro-Asp-Asp-Ser-Arg-Val-Ile-Ala-His-Thr-Lys-
P.fl. Asp-Lys-Asp-Tyr-Leu-Lys-Pro-Asp-Asp-Ile-Arg- Lys-

P.aer. Leu-Ile-Gly-Ser-Gly-Glu-Lys-Asp-Ser-Val-Thr-Phe-Asp-Val-Ser-Lys-Leu-
P.fl. Leu-Leu-Gly-Glu- Thr-Phe-Asp-Val-Ser-

P.aer. Lys-Glu-Gly-Glu-Gln-Tyr-Met-Phe-Phe-Cys-Thr-Phe-Pro-Gly-His-Ser-Ala-
P.fl. Phe-Phe-

P.aer. Leu-Met-Lys-Gly-Thr-Leu-Thr-Leu-Lys
P.fl. Lys-Gly-Thr-Val-Thr-Val-Lys

Figure 1. Sequence of azurin from P. aeruginosa together with regions of P. fluorescens azurin which have been aligned by homology. Underlined residues differ from the type.

2. gel filtration to separate undigested or partially digested protein from small peptides
3. partial separation of the peptide pool on a peptide analyser followed by electrophoretic monitoring of the effluent
4. mixture analysis mass spectrometry of suitable pools of peptides
5. alignment of peptides with an already established type sequence
6. redigestion of high molecular weight material recovered in step 2 and a repeat of steps 3 to 6

Acknowledgements

We thank Dr. R.P. Ambler for a gift of Azurin and The Medical Research Council for financial support.

References

1. Perham, R.N. ed. (1975) Instrumentation in Amino Acid Sequence Analysis, Academic Press, London.
2. Ambler, R.P. and Brown, L.H. (1967) Biochem. J., 104, 784-825.
3. Ambler, R.P. (1971) in Recent Developments in the Chemical Study of Protein Structures, Previero, A., Pechere, J.F. and Coletti-Previero, M.A. eds., pp 289-305, Inserm, Paris.
4. Dell, A., Morris, H.R., Williams, D.H., and Ambler, R.P. (1974) Biomed. Mass Spectrom., 1, 269-273.
5. Morris, H.R., Batley, K.E. Harding, N.G.L., Bjur, R.A., Dann, J.G. and King R.W. (1974) Biochem J., 137, 409-411.
6. Shotton, D.M. and Hartley, B.S. (1973) Biochem. J., 131, 643-675.
7. Schroeder, W.A. (1972) Methods Enzymol., 25, 203-221.
8. Morris, H.R., Dickinson, R.J. and Williams, D.H. (1973) Biochem. Biophys. Res. Commun. 51, 247-255.
9. Morris, H.R., Williams, D.H. and Ambler, R.P. (1971) Biochem. J., 125, 189-201.
10. Morris, H.R., Dell, A., Petersen, T.E., Sottrup-Jensen, L. and Magnusson, S.(1976) Biochem. J., 153, 663-679.
11. Morris, H.R. , Williams, D.H. Midwinter, G.G. and Hartley, B.S. (1974) Biochem. J., 141, 701-713.
12. Batley, K.E. and Morris, H.R. (1977) Biochem. Biophys. Res. Commun., 75, 1010-1014.