A Mass-Spectrometric Method for the Estimation of the Ratio of γ -Carboxyglutamic Acid to Glutamic Acid at Specific Sites in Proteins

APPLICATION TO THE *N*-TERMINAL REGION OF BOVINE PROTHROMBIN

Keith ROSE,* John D. PRIDDLE,* Robin E. OFFORD* and M. Peter ESNOUF†

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

†Nuffield Department of Clinical Biochemistry, Radcliffe Infirmary, Oxford OX2 6HE, U.K.

(Received 3 December 1979)

When a polypeptide containing γ -carboxyglutamic acid is decarboxylated in $^2\text{H}_2\text{O}$, residue of $(\gamma\gamma\text{-}^2\text{H}_2)\text{glutamic acid}$ are formed. Subsequent proteolytic digestion produces peptides which contain at each site $^2\text{H}_2$ -substituted and unsubstituted glutamic acid in the same ratio as existed for γ -carboxy-substitution. The peptides may be identified and this ratio determined by combined gas chromatography–mass spectrometry. We also discuss decarboxylation in $^3\text{H}_2\text{O}$ followed by amino-acid analysis and Edman degradation.

Several proteins, particularly some of those involved in the blood-clotting mechanism, undergo a post-translational modification, dependent on vitamin K, whereby a second carboxyl group is introduced at the γ -carbon of the side chain of certain of their glutamic residues (e.g. Esmon *et al.*, 1975). The product of the modification, γ -carboxyglutamic acid, is a key structural and functional element, so when studying such proteins it is important to be able to determine the extent of γ -carboxylation at each and every site. Anti-coagulants such as Warfarin exert their effect by suppressing, either generally or selectively, this γ -carboxylation (Esnouf & Prowse, 1977). Amino-acid-sequence determination is hampered by the ease with which γ -carboxyglutamic residues decarboxylate back to glutamic acid.

Mass spectrometry has been applied to such proteins: fractional evaporation of the products obtained after acetylation and permethylation served to establish the presence of γ -carboxyglutamic acid in ten positions in the *N*-terminal region of bovine prothrombin (Morris *et al.*, 1976) and in twelve positions in clotting-factors X_1 and X_2 (Thøgersen *et al.*, 1978). As a result of the chemical nature of γ -carboxyglutamic acid, a multiplicity of derivatives is formed during this procedure, so use of the method for quantitative estimates of the proportion of modified and unmodified residues would be difficult. The production of a more stable derivative by reduction of γ -carboxyglutamic acid to 5,5'-

dihydroxyleucine has been proposed (Zytkovicz & Nelsestuen, 1975) but so far the method seems only to have been used prior to total hydrolysis and amino acid analysis (Howard & Nelsestuen, 1975).

As a general sequencing method, g.l.c.–m.s. offers advantages of speed and redundancy of information (Priddle *et al.*, 1976a,b, 1978; Rose & Priddle, 1978). Since identification is by mass spectrometry, the method is particularly well suited to the study of material containing stable isotope substitutions (Rose & Priddle, 1978), and it seemed to us (Rose, 1979) to be applicable to the estimation of the degree of γ -carboxylation of glutamic residues while at the same time determining the sequence around the site in question.

Rather than being a hindrance, the tendency of γ -carboxyglutamic acid to decarboxylate is turned to advantage. The stages of the procedure are as follows: (1) Treatment of the protein, as soon as it is isolated, with $^2\text{H}_2\text{O}$ in which isotopic exchange of the γ -H is very rapid (N. D. Pluck & R. J. P. Williams, personal communication). (2) The protein is then subjected to γ -decarboxylation in the same medium. The lost carboxyl group is thereby replaced by a further ^2H atom, with the result that a residue of $(\gamma\gamma\text{-}^2\text{H}_2)\text{glutamic acid}$ is formed. Now that there is only one carboxyl group on the γ -carbon atom, the ^2H atoms become effectively non-exchangeable under all the conditions that the peptide subsequently experiences. It is for this reason that residues of normal glutamic acid are unaffected by the procedure in so far as hydrogen bound to carbon is concerned. ^2H atoms picked up elsewhere on this

Abbreviation used: g.l.c.–m.s., gas-liquid chromatography coupled to mass-spectrometry.

or other residues (by exchange with protons attached to oxygen or nitrogen atoms) either exchange back during later handling in H_2O or are replaced by $-CH_3$ when the peptide is permethylated (see below). (3) The ratio of 2H_2 -substituted glutamic acid to the unsubstituted form is then determined at each site in the course of the g.l.c.-m.s. analysis of the sequence. The peptide is trifluoroacetylated and permethylated for this purpose. The ratio of intensities in the doublets of ions that are produced should reflect the original extent of γ -carboxylation. Since two 2H atoms are incorporated, the masses of the ions with deuterium would differ by two mass units from those with hydrogen; this probably renders unnecessary a correction for the natural abundance of ^{13}C .

In addition to the likely advantages of the mass spectrometric analysis itself, we believe the fact that the decarboxylation comes first in the procedure to be a substantial benefit. Without this step the pattern of proteolytic cleavage of a given sequence into fragments would depend on whether the γ -carboxyl group was present or not, as would the separation behaviour and recoveries of the individual peptides formed. By contrast, the effect of 2H_2 -substitution should be much less marked.

The above scheme has been evaluated by work with peptide C from bovine prothrombin. This peptide comprises residues 1–35 of the protein, and all the glutamic residues are thought to be fully γ -carboxylated (Morris *et al.*, 1976).

Decarboxylation in the presence of 3H_2O has been used, followed by total hydrolysis and scintillation counting of the effluent from an amino acid analyser, as a measure of content of γ -carboxyglutamic acid (Hauschka, 1979). The option is available for the Edman degradation of peptides tritiated in this way.

Materials and Methods

2H_2O (99.8% isotopic abundance) was obtained from Ryvan Chemical Co., Southampton, U.K. The sources of the other reagents, which were generally of analytical or equivalent grade, are given by Rose (1979), as are the procedures for the drying of dimethylsulphoxide and the preparation of the 2M-solution of dimethylsulphinyll sodium.

Prothrombin was prepared as described by Prowse *et al.* (1976) and peptide C was obtained from it by digestion with chymotrypsin and adsorption onto barium citrate (Morris *et al.*, 1976). Peptide C (21.8 mg) was dissolved in 1.5 ml of 2H_2O and acidified to an apparent pH of 2 (indicator paper pH 1–4 range, Whatman, Maidstone, Kent, U.K.) with 5 μ l of concentrated HCl. The protein was precipitated, but redissolved during the decarboxylation step, which was carried out by heating at 95°C for 6 h. After cooling, the tube was opened and the solvent was evaporated under reduced pressure.

The sample was dissolved in 1 ml of a buffer made from 155 mg of boric acid in 10 ml of 8M-urea (freshly de-ionized on an Amberlite MB-3 column) containing dithiothreitol (29 mg/ml) and adjusted to pH 8.6 with 5M-NaOH. Aziridine (20 μ l) was added and the pH was maintained at about 8.6 by manual addition of 1M-HCl. After 15 min, a further 20 μ l of aziridine was added and the pH was maintained at 8.6 for 1 h. The sample was then diluted with 1 ml of aq. 50% (v/v) acetic acid and applied to a column of Sephadex G-50 (45 cm $\times$ 1.9 cm dia.) eluted with 1% acetic acid. Peptide-containing fractions were detected by their A_{260} , pooled, and freeze-dried. The decarboxylated, S -aminoethylated peptide C was dissolved in 0.2 ml of H_2O , and 0.2 ml of triethylamine, 0.4 ml of ethanol and 0.4 ml of pentane-2,4-dione were added, in that order, with mixing. The tube was sealed and heated at 95°C for 4 h. After cooling, 2 ml of 98% formic acid was added to the reaction mixture and the solution was applied to a Sephadex G-50 column (92 cm $\times$ 2.65 cm dia.) eluted with aq. 50% (v/v) acetic acid. The elution of the peptide was conveniently monitored at 310 nm, the approximate absorption maximum of dimethylpyrimidylornithine residues (Vetter-Diechtel *et al.*, 1968). The corresponding fractions were pooled, rotary-evaporated to below 5 ml, diluted with water, and freeze-dried. The recovery of peptide was 14.5 mg. Separate samples (1.5–2 mg) of this material were digested in 0.5 ml of 1% NH_4HCO_3 solution with subtilisin [enzyme/substrate ratio 1:100 (w/w) for 4 h at 37°C], elastase [enzyme/substrate ratio 1:40 (w/w) for 4 h at 37°C], or Pronase [enzyme/substrate ratio 1:100 (w/w) for 2.5 h at 37°C], then freeze-dried.

The freeze-dried samples were N -trifluoroacetylated in dimethylsulphoxide solution (Priddle *et al.*, 1978; Rose, 1979) using for each sample 0.5 ml of dimethylsulphoxide/triethylamine/ S -ethyltrifluoroacetate (50:1:1, by vol.). After 1 h at room temperature, the reagents, but not the solvent, were pumped away in a vacuum desiccator containing NaOH and H_2SO_4 . Permethylation was carried out by addition of 100 μ l of 2M-dimethylsulphinyll sodium immediately followed by 30 μ l of methyl iodide. After 60 s, the reaction mixture was quenched with 1 ml of H_2O and the product was extracted into 1 ml of $CHCl_3$. After washing twice with 1 ml portions of H_2O the organic solvent was evaporated under reduced pressure. The product, a pale yellow oil, was taken up in 10 μ l of $CHCl_3$ and 2 μ l was analysed by g.l.c.-m.s. in the manner previously described (Priddle *et al.*, 1976b) except that the MM 30F mass spectrometer is now equipped with a 2035 F/B data system (VG Data Systems, Altrincham, Cheshire, U.K.). The instrument scans the range from approximately m/e 1500 down to below m/e 28 at about 3 s/decade every 10 s. Data from every scan

Fig. 1. Peptides identified by g.l.c.-m.s. shown aligned under the known sequences of peptide C. The symbol — indicates that the N-terminus was found; the symbol () indicates that the molecular ion was found. El indicates those peptides found in the elastase digest (Sigma, type III), S those found in the subtilisin digest (Sigma, type VIII), and Pr those found in the Pronase digest (Kaken, type P). Those sequences shown in parentheses are not of sufficient length to allow them to be placed uniquely. Gla, γ-carboxyglutamic acid.

are stored, together with the time from the start of the g.l.c. run, on a magnetic disc. The instrument parameters were: ionizing voltage 70 V, accelerating voltage 4 kV, source temperature 185°C. The g.l.c. parameters were: 30 cm $\times$ 2 mm i.d. column, 1% OV1 on WHP 100–120 mesh (Chrompack U.K., London SW16, U.K.), 15 ml/min He (5 ml/min eluent to mass spectrometer source, via a T splitter), initial column temperature 150°C, raised at 4°C/min from the appearance of the solvent front in the spectrometer until termination of the run; the injection zone was held at 300°C.

Result and Discussion

Fig. 1 shows the peptides identified by g.l.c.–m.s., aligned under the known sequence (Morris *et al.*, 1976) of peptide C. Taking the three digests together, each of the 10 γ -carboxyglutamic residues is found [as ($^2\text{H}_2$)glutamic acid] in at least one peptide of unique sequence. We can thus screen rapidly and unequivocally samples of peptide C that may be partly carboxylated at certain sites.

In order to obtain a sample of a polypeptide with a known degree of substitution at a series of sites, 1.85 mg of native peptide C was decarboxylated in $^2\text{H}_2\text{O}$ as above, and mixed with 0.925 mg of peptide C that has been decarboxylated in $^1\text{H}_2\text{O}$, to provide decarboxylated peptide C with (γ - $^2\text{H}_2$)glutamic acid and glutamic acid in the ratio 2:1 at each of the 10 positions. The mixture was digested with subtilisin [enzyme/substrate ratio 1:25 (w/w) for 2 h at

37°C], *N*-trifluoroacetylated, permethylated and analysed by g.l.c.–m.s. as before. As hoped for, spectra of peptides that contained substituted glutamic acid showed a characteristic pattern, with an isotope ratio of very nearly 2:1. Fig. 2 shows the mass spectrum of one such peptide. The mean ratio for the doublets present as sequence ions or as related fragments is 1.95:1 [s.d. = 0.12 ($n = 8$)]. No corrections for background ions or for ^{13}C contributions have been made. [Hauschka (1979) was only able to substitute the γ -carboxyglutamic residues in polypeptides to the extent of 0.95 ± 0.14 mol/mol, rather than to the predicted extent of 2 mol/mol. Our results imply that our conditions may be more satisfactory for quantitative work.]

This approach seems to provide a general method for the detection of γ -carboxyglutamic residues in proteins and for the quantification of the degree of γ -carboxylation at particular positions. Its advantage lies in that the first step, decarboxylation in $^2\text{H}_2\text{O}$, displaces a substituent group which could, because of its lability, give trouble later on in the analysis. The decarboxylation produces a stable, covalently labelled product. Although unambiguously labelled, this product is nonetheless chemically identical with the unsubstituted residue which might be present in a proportion of the sample. The absolute yields of material from the subsequent derivative formation and proteolytic digestion do not affect the isotope ratio in any peptide.

Fig. 2. Mass spectrum showing the sequences ($^2\text{H}_2$)Glu-Ala-Leu and Glu-Ala-Leu in a ratio of about 2:1. In this Figure, the usual three-letter code is used to denote the *N,O*-permethylated residues. This spectrum was stored as scan no. 198 of the g.l.c.–m.s. run (see text). The sequence deduced is a unique segment of the sequence of peptide C (Fig. 1), being residues 30–32. The presence of a strong signal at m/e 58, which corresponds to $[\text{CH}_3\text{—NH=CH—CH}_3]$, indicates that alanine is present in a position other than at the *N*-terminus (Rose, 1979). Signals at m/e 166 and 194, seen here as doublets, are considered to be due to loss of HCOOCH_3 (60 a.m.u.) and of $\text{HCOOCH}_3 + \text{CO}$ (88 a.m.u.), respectively, from the *N*-terminal sequence ion (m/e 256 for the $^2\text{H}_2$ -containing material). Corresponding signals are present in mass spectra of all the *N*-terminal glutamyl peptides examined so far (K. Rose, unpublished results); they are not prominent in the case of *N*-acetyl permethyl derivatives, where loss of $\text{CH}_2\text{CO} + \text{CO}$ (70 a.m.u.), to give the ion $[\text{CH}_3\text{—NH=CH—CH}_2\text{—CH}_2\text{—OC}_2\text{CH}_3]$ is strongly favoured.

We thank the Medical Research Council and the Science Research Council for providing the equipment. K. R. and J. D. P. thank the Medical Research Council for financial support. K. R. also thanks the Wellcome Trust for financial support. This is a publication of the Oxford Enzyme Group.

References

- Esmon, C. T., Sadowski, J. A. & Suttie, J. W. (1975) *J. Biol. Chem.* **250**, 4744–4748
- Esnouf, M. P. & Prowse, C. V. (1977) *Biochim. Biophys. Acta* **490**, 471–476
- Hauschka, P. V. (1979) *Biochemistry* **18**, 4992–4999
- Howard, J. B. & Nelsestuen, G. L. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1281–1285
- Morris, H. R., Dell, A., Petersen, T. E., Sottrup-Jensen, L. & Magnusson, S. (1976) *Biochem. J.* **153**, 663–679
- Priddle, J. D., Rose, K. & Offord, R. E. (1976a) in *Advances in Mass Spectrometry in Biochemistry and Medicine* (Frigerio, A., ed.), vol. 2, pp. 477–485, Spectrum Publications, New York
- Priddle, J. D., Rose, K. & Offord, R. E. (1976b) *Biochem. J.* **157**, 777–780
- Priddle, J. D., Rose, K. & Offord, R. E. (1978) in *Advances in Mass Spectrometry* (Daly, N., ed.), vol. 7B, pp. 1502–1505, Heyden and Son, London
- Prowse, C. V., Mattock, P., Esnouf, M. P. & Russel, A. M. (1976) *Biochim. Biophys. Acta* **434**, 265–279
- Rose, K. (1979) *D. Phil. Thesis*, University of Oxford
- Rose, K. & Priddle, J. D. (1978) in *Semisynthetic Peptides and Proteins* (Offord, R. E. & DiBello, C., eds.), pp. 387–395, Academic Press, London, New York and San Francisco
- Thøgersen, H. C., Petersen, T. E., Sottrup-Jensen, L., Magnusson, S. & Morris, H. R. (1978) *Biochem. J.* **175**, 613–627
- Vetter-Diechtl, H., Vetter, W., Richter, W. & Biemann, K. (1968) *Experientia* **24**, 340–341
- Zytkovicz, T. H. & Nelsestuen, G. L. (1975) *J. Biol. Chem.* **250**, 2968–2972