Biochem. J. (1976) **153**, 663–679
Printed in Great Britain

Mass-Spectrometric Identification and Sequence Location of the Ten Residues of the New Amino Acid (γ-Carboxyglutamic Acid) in the *N*-Terminal Region of Prothrombin

By HOWARD R. MORRIS* and ANNE DELL*
Medical Research Council Group, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

and TORBEN E. PETERSEN, LARS SOTTRUP-JENSEN and STAFFAN MAGNUSSON
Department of Molecular Biology, University of Aarhus, 8000 Aarhus C, Denmark

(*Received* 15 *August* 1975)

The detailed mass-spectrometric evidence for our original findings [Magnusson *et al.* (1974) *FEBS Lett.* **44**, 189–193] of ten γ-carboxyglutamic acid residues in the *N*-terminal calcium-binding polypeptide of prothrombin is presented. The identification and sequence location of γ-carboxyglutamic acid was made by electron-impact and field-desorption studies on acetyl permethyl peptide derivatives, and on the free amino acid. Details of the derivatives formed, and how this new amino acid may be easily recognized and sequenced from the mass spectrum, are given as a basis for future work.

Prothrombin consists of a single polypeptide chain of 582 amino acid residues with 12 disulphide bridges (Magnusson *et al.*, 1975*b*). It contains a total of approx. 12% carbohydrate consisting of glucosamine, mannose, galactose and sialic acid (Magnusson, 1965*b*), attached as three oligosaccharide units to asparagine residues 77, 101 and 376 (Magnusson *et al.*, 1975*b*).

In the activation of prothrombin, Factor X*a* (EC 3.4.21.6) catalyses the cleavage of the two peptide bonds connecting residues 274–275 and 323–324, leading to the formation of thrombin (residues 275–323, disulphide-bridged to residues 324–582) (EC 3.4.21.5) and the 'pro' fragment (residues 1–274), which can be cleaved by thrombin (either before or after the action of Factor X*a*) at peptide bond 156–157, producing the A-fragment (residues 1–156) and the S-fragment (residues 157–274) (Magnusson *et al.*, 1975*b*).

Vitamin K is required for the biosynthesis of prothrombin (Dam, 1935) in the hepatic cells of the liver (Barnhart, 1960). Rapid activation of prothrombin in the presence of Ca^{2+} ions involves the formation of a complex of prothrombin, phospholipids and Factor V (Barton & Hanahan, 1969; Bull *et al.*, 1972). In conditions of vitamin K deficiency and during treatment with dicoumarol (Campbell & Link, 1941) or other 'vitamin K antagonists' the prothrombin activity in the blood plasma is low, as measured in calcium/phospholipid-dependent assays. Using precipitating antibodies against normal prothrombin Josso *et al.* (1968) and Ganrot & Niléhn (1968*a*) have found that plasma from dicoumarol-treated patients contained a protein immunologically identical with prothrombin. The concentration was equivalent to that of prothrombin in normal plasma. This 'dicoumarol prothrombin' was not adsorbed on barium citrate and did not bind Ca^{2+} ions (Ganrot & Niléhn, 1968*b*). The quantitative amino acid and carbohydrate compositions of the isolated dicoumarol prothrombin (Stenflo, 1972; Nelsestuen & Suttie, 1972*a*) were found not to be significantly different from those of normal prothrombin (Magnusson, 1965*a*,*b*). The amino acid sequence of residues 1–5 was found to be identical in the two proteins (Stenflo, 1973). Ca^{2+}-binding studies indicate that, in contrast with dicoumarol prothrombin, normal prothrombin binds 3–4 Ca^{2+} ions tightly (Nelsestuen & Suttie, 1972*b*; Stenflo & Ganrot, 1973).

The first clue to the presence of an unusual structure in normal prothrombin was the isolation (Magnusson, 1972) of the tryptic peptides Gly-Phe-Leu-Glx-Glx-Val-Arg (residues 4–10) and Gly-Phe-Leu-Glx-Glx-Val-Arg-Lys (residues 4–11), which had electrophoretic mobilities at pH 6.5 that could only be explained by the presence of one or two negative charges more than can be accounted for by both Glx residues being Glu residues (Magnusson, 1973), probably as a result of a vitamin K-dependent substitution on either or both of Glx residues 7 and 8 (Magnusson *et al.*, 1974*a*). Both peptides were free of carbohydrate. Nelsestuen & Suttie (1973) adsorbed a different tryptic peptide on barium citrate. Stenflo (1974) confirmed the presence of extra negative

* Present address: Department of Biochemistry, Imperial College, London SW7 2AZ, U.K.

charges on peptides 4–10 and 4–11 from normal prothrombin and found that the corresponding peptides from dicoumarol prothrombin had electrophoretic mobilities that could be accounted for by residues 7 and 8 being normal Glu residues.

In establishing the primary structure of normal prothrombin by classical methods, at least six of the Glx residues (7, 8, 26, 27, 30 and 33) were found to carry extra negative charges. Mass spectrometry of an acetylated permethylated peptide showed the presence of one extra methyl group on each of the derivatives of Glx residues 26, 27 and 30 as well as patterns indicating that the methylated Glu had been formed by decarboxylation (Magnusson *et al.*, 1975*a*). Stenflo (1975) and Stenflo *et al.* (1974) proposed the structure γ-carboxyglutamic acid for residues 7 and 8, based on n.m.r. and mass-spectrometric evidence. In our mass-spectrometric work we have shown conclusively (Magnusson *et al.*, 1974*b*) that all ten Glx residues (7, 8, 15, 17, 20, 21, 26, 27, 30 and 33) in the region 1–42 are γ-carboxyglutamic acid residues. Nelsestuen *et al.* (1974) have confirmed the presence of one γ-carboxyglutamic acid in a dipeptide derived from prothrombin. The structure of γ-carboxyglutamic acid has been confirmed by synthesis (Morris *et al.*, 1975; S. Bajusz & A. Juhasz, unpublished work).

The present paper reports the first detailed mass-spectrometric evidence used to determine the structure of the new amino acid and its sequence location in ten positions in prothrombin. We believe that this work will form a simple basis for future direct recognition of this amino acid in structures, since amino acid analyses and classical sequencing procedures (manual or automated) have as yet (e.g. Enfield *et al.*, 1975) failed to give the definitive data required.

Materials and Methods

Enzymes and reagents

Pronase was a gift from Dr. M. Nomoto. Chymotrypsin (thrice crystallized, 6046-7), pepsin (twice crystallized from ethanol, 681) and soya-bean trypsin inhibitor (81 53M 457) were obtained from Worthington Biochemical Corp., Freehold, NJ, U.S.A. Thermolysin (declared activity 68 caseinolytic units/mg of protein, 92C-2500) was obtained from Sigma Chemical Co., St. Louis, MO, U.S.A. Diphenylcarbamoyl chloride was obtained from Eastman-Kodak, Rochester, NY, U.S.A. All other reagents were the best commercial grades. Sephadex G-25 (fine grade) was obtained from Pharmacia, Uppsala, Sweden.

Amino acid analyses

Peptide samples (5–10 nmol) hydrolysed in 6M-HCl (110°C, 16–20h, in sealed tubes at reduced pressure) were analysed essentially as described by Spackman *et al.* (1958) on Locarte amino acid analysers, by using either a one-column or a two-column system.

Amino acid sequences

These were determined manually by the Dns–Edman method (Gray, 1967). Dns-amino acid derivatives were identified by t.l.c. on polyamide (Cheng-Chin Trading Co., Taipeh, Taiwan) (Woods & Wang, 1967). Sequences of the peptides having unusual electrophoretic mobility were determined independently by the mass-spectrometric methods given below.

Electrophoresis

Electrophoretic separation of peptides was carried out on Whatman no. 3MM or no. 1 paper at 3kV (approx. 68V/cm) in buffers of pH6.5 and pH2.1 (Ambler, 1963). Peptides were detected by staining with Cd–ninhydrin (Heilmann *et al.*, 1957), and phenanthraquinone (Yamada & Itano, 1966) was used to detect arginine residues. Electrophoretic mobilities were correlated with charge and molecular weight as described by Offord (1966).

Prothrombin

Bovine prothrombin was purified as described by Magnusson (1965*b*, 1970). It had a specific activity of 1450 NIH (National Institutes of Health) units/mg dry wt. After reduction with mercaptoethanol it gave rise to a single band on disc-gel electrophoresis in the presence of sodium dodecyl sulphate by the method of Weber & Osborn (1969).

Peptides from prothrombin

The amino acid sequence of residues 1–35 of bovine prothrombin (Magnusson *et al.*, 1974*b*) is shown in Fig. 1. Over 30 different peptides were used to establish this sequence. However, sufficient evidence to prove both the structure of the new amino acid and its location in the sequence can now be obtained from one Pronase peptide (Prl) and one chymotryptic fragment (C) further digested with thermolysin and pepsin, and only these peptides will be described here.

Peptide Prl. Prothrombin (500mg, approx. 6.7 μmol) was dissolved in 58ml of 0.15M-NH_4HCO_3, pH7.8, and incubated at 37°C with 5mg of Pronase for 2h. Another 5mg of Pronase was then added, and incubation continued for a further 20h. The digest was freeze-dried and dissolved in 38ml of 20mM-sodium citrate, and 38ml of 1M-$BaCl_2$ was then added, all at 4°C, and the mixture left stirring for 1h. The resulting barium citrate precipitate carrying the

```
 1   2   3   4   5   6   7   8   9  10  11  12  13
Ala-Asn-Lys-Gly-Phe-Leu-Gla-Gla-Val-Arg-Lys-Gly-Asn-

 14  15  16  17  18  19  20  21  22  23  24  25
-Leu-Gla-Arg-Gla-Cys-Leu-Gla-Gla-Pro-Cys-Ser-Arg-

 26  27  28  29  30  31  32  33  34  35
-Gla-Gla-Ala-Phe-Gla-Ala-Leu-Gla-Ser-Leu
```

Fig. 1. *Amino acid sequence of residues* 1–35 *of bovine prothrombin* (*Magnusson et al.*, 1974*b*)

The γ-carboxyglutamic acid residues in positions 7, 8, 15, 17, 20, 21, 26, 27, 30 and 33 were determined by mass spectrometry and are referred to as Gla residues in this paper. For details see the text.

adsorbed peptide was separated by centrifugation (4°C, 1000*g*, 15min) and then washed twice, each time by stirring for 5min with 19ml of 10mM-sodium citrate, then adding 1ml of 1M-$BaCl_2$ and separating the precipitate by centrifugation. The washed precipitate was stirred with 9.6ml of 1M-Na_2SO_4. The resulting precipitate of $BaSO_4$ is in a sufficiently high concentration of sodium citrate to prevent adsorption of prothrombin (Lewis & Ware, 1953; Goldstein *et al.*, 1959) and therefore presumably of peptides from the vitamin K-dependent region. It was separated by centrifugation (as described above). The supernatant was applied to a column (2cm×125cm) of Sephadex G-25 (fine grade), equilibrated and eluted with 50mM-NH_3 (flow rate 15ml/h). The E_{280} of the effluent was monitored. The peptide fraction was freeze-dried without further purification. The yield was 5.2mol (78% of theoretical). The amino acid composition is given in Table 1. This composition and the results of *N*-terminal analysis are consistent with peptide Prl being mainly an approximately equimolar mixture of residues 14–30 and 14–31, with a small fraction being residues 9–30/31 and 12–30/31 (see Fig. 1).

Peptide C. Prothrombin (2g, approx. 27μmol) was dissolved in 400ml of 0.15M-NH_4HCO_3, pH7.8, containing 10mg of soya-bean trypsin inhibitor and digested with 45mg of chymotrypsin at 37°C for 4–5h. Then 2mg of diphenylcarbamoyl chloride in 1ml of methanol was added. The digest was then freeze-dried and dissolved in 400ml of 20mM-sodium citrate at 4°C. Then 18ml of 1M-$BaCl_2$ was added, and the mixture was stirred for 1h. The precipitate was collected by centrifugation (22000*g*, 15min, 4°C) and washed with 80ml of 10mM-sodium citrate; another 2ml of 1M-$BaCl_2$ was added. After a second centrifugation the precipitate was stirred with 1M-Na_2SO_4 for 10min. The resulting supernatant was desalted as described for peptide Prl. The yield of freeze-dried material was 100mg (84% yield).

Subdigestion of peptide C. Peptide C (50mg, approx. 11μmol) was dissolved in 10ml of 2M-pyridine/acetic acid, pH6.5, containing 20mM-$CaCl_2$, and digested with 1.5mg of thermolysin at 50°C. During the first hour a heavy precipitate formed in the initially yellowish milky suspension. Then 1ml of 0.6M-Tris/HCl, pH8.5, was added to give pH8.1, but the precipitate did not dissolve. The incubation was continued for a further 4h with occasional shaking. Then 2ml of 0.1M-EDTA, adjusted to pH8.5 with NaOH, was added and the precipitate dissolved. The resulting solution was applied to a column (2cm× 120cm) of Sephadex G-25 (fine), equilibrated and eluted with 50mM-NH_4HCO_3, pH8.5 (flow rate 15ml/h, 20min fractions). The effluent was monitored by high-voltage paper electrophoresis at pH6.5. The peptide 'maps' resulting from staining with Cd–ninhydrin and phenanthraquinone were used to decide on pooling. Three cuts were made, and the pooled material was freeze-dried.

Pool III (fractions 57–63) contained the peptides Ala-Asn-Lys-Gly (residues 1–4) and Gly-Asn (12–13), as judged from mobility and staining characteristics. These peptides were not isolated from this particular digest.

Pool II (fractions 46–51) contained the peptide CTL 1 (residues 5–11; yield 1.8μmol). For the mass-spectrometric analysis a much smaller amount of the corresponding peptide from a twin experiment was used.

Pool I (fractions 33–39), according to the amino acid and *N*-terminal analysis, contained mostly residues 5–35 with small amounts of 12–35 and 14–35. It was reduced and carboxymethylated as follows. Approx. 10μmol was dissolved in 5.2ml of 0.6M-Tris/HCl, pH8.5, containing 4mM-EDTA and 5mM-dithiothreitol, for 20min and then sodium iodoacetate was added to a final concentration of 12mM. After 30min the solution was applied to the Sephadex G-25 column as described above. The void-volume

Table 1. *Amino acid compositions and electrophoretic mobilities of isolated peptides*

Results of amino acid analyses are expressed in molar amounts relative to the underlined values.

	Prl	CTLP1	CTLP2	CTLP3	CTLP4	CTLP5	CTLP6	CTL1
CmCys		1.14	0.96	0.85			1.84	
Asx	0.74						0.28	0.19
Ser	1.03			1.03	0.93	0.19	0.64	
Glx	5.71	3.83	4.77	3.81	1.20	2.55	5.88	1.89
Pro	0.86			0.89			0.78	
Gly	0.69							0.25
Ala	1.69			1.00	1.00	0.40	0.72	
½Cys	1.12							
Val	0.42							1.06
Leu	2.08	1.00	1.60		2.01	0.92	2.00	0.87
Phe	1.00			1.14		1.00	0.60	1.03
Lys	0.38			0.24				1.03
Arg	1.77	1.02	1.00	1.16		0.36	1.66	1.00
*m_{Asp}	—	1.15	1.00	0.86	0.69	0.94	1.01	0.42
†Residues	14–30	15–21	14–21	22–30	31–35	5–8	14–30	5–11
‡M.W.	2491	1140	1253	1256	575	624	2491	1007
Charge found		−5	−5	−4	−2	−3		−2
§Theoret. charge		−8	−8	−6	−2	−4		−2

* m_{Asp}, Electrophoretic mobility at pH6.5 relative to that of aspartic acid (1.00).
† Residues, Position of peptide in prothrombin sequence (see Fig. 1).
‡ M.W., Molecular weight calculated from sequence.
§ Theoret. charge, Net charge at pH6.5 derived by attributing two negative charges to each γ-carboxyglutamic acid residue.

material was freeze-dried (weight 35mg, yield approx. 7μmol), and then dissolved in 4ml of 0.1M-HCl, mixed with 5mg of pepsin in 1ml of 0.1M-HCl and incubated at 37°C for 16h. Then 5M-NH_3 was added to give pH8.5, and the digest was applied to the Sephadex G-25 column. The peptide separation was monitored by paper electrophoresis at pH6.5 of samples from every second column fraction, by staining two sets of 'fingerprints' with Cd–ninhydrin and phenanthraquinone respectively. The following peptides were isolated by preparative paper electrophoresis at pH6.5. From pool I (fractions 33–38), peptide CTLP1 (residues 15–21; yield 1.4μmol); from pool II (fractions 39–42), peptides CTLP2 (residues 14–21; yield 0.6μmol) and CTLP3 (residues 22–30; yield 1.5μmol); from pool III (fractions 43–52), peptides CTLP4 (residues 31–35; yield 2.7μmol) and CTLP5 (residues 5–8; yield 2.5μmol). Peptide material corresponding to residues 9–14 was seen in the 'fingerprints' of fractions 53–59, but was not isolated. The amino acid compositions, electrophoretic mobilities and molecular weights of these peptides as well as their position in sequence and approximate charge at pH6.5 are listed in Table 1.

A second digest was prepared from 25mg (approx. 5.6μmol) of peptide C by using thermolysin and pepsin as described above, except that $CaCl_2$ was not added. From this digest, where no precipitate formed, peptide CTLP6 (residues 14–30; yield 0.7μmol) was isolated by paper electrophoresis at pH6.5 and pH2.1.

Isolation of the new amino acid by partial acid hydrolysis

Peptide C (20mg; approx. 4.5μmol) was dissolved in 8ml of 6M-HCl and incubated at 45°C for 24h. The sample was then dried *in vacuo*, dissolved in 2M-NH_3, applied as a 35cm line (20cm from the top of the paper) to a sheet of Whatman 3MM paper and subjected to electrophoresis at pH6.5, 3kV, 40min. A guide strip stained with Cd–ninhydrin showed several positive bands, one of which had $m_{Asp} = 1.35$. The rest of this band was cut out, stitched to a new sheet of Whatman 3MM paper and subjected to electrophoresis at pH2.1, 3kV, 45min. The two major ninhydrin-positive bands with mobilities $m_{Ser} = 0.41$ (corresponding to free γ-carboxyglutamic acid) and $m_{Ser} = 0.30$ (corresponding to the dipeptide Gla-Gla; see under 'Nomenclature') were eluted with 50mM-NH_3, and the samples were dried *in vacuo*. When subjected to hydrolysis under the standard conditions used for amino acid analysis (6M-HCl, 110°C, 16–20h) both samples gave rise to only one peak on the amino acid-analysis chart, apparently identical with glutamic acid. Assuming that each γ-carboxyglutamic acid residue had been quantitatively decarboxylated to

glutamic acid the yield of γ-carboxyglutamic acid plus Gla-Gla was no more than 10% of theoretical.

Preparation of derivatives for mass spectrometry

Peptides (100–200nmol) were acetylated in acetic anhydride/methanol (1:4, v/v) (Thomas *et al.*, 1968) after being dissolved, where necessary, in 10–20μl of water. The acetyl derivatives were permethylated by the Hakamori (1964) reaction with a very short reaction time of 75s (Morris, 1972) to avoid salt formation with some amino acid residues. Both of the above procedures were modified by using the isotopic reagents C^2H_3I, $C^2H_3O^2H$, $(C^2H_3CO)_2O$ and $C^2H_3SOC^2H_3$ at the appropriate place in the reaction sequence.

Mass spectrometry

Electron impact. Spectra were recorded on a GEC–AEI MS902 mass spectrometer operating with an accelerator voltage of 8kV and an electron beam energy of 70eV.

Sample handling, temperature gradient and partial-fractionation techniques for mixtures have been described elsewhere (Morris *et al.*, 1971).

Interpretation of spectra was based on the most recent data on the fragmentation properties of acetylated permethylated peptides (Morris *et al.*, 1974; Dell & Morris, 1974).

Field desorption. Field-desorption spectra of the naturally occurring γ-carboxyglutamic acid were obtained on an AEI field-ionization source, modified in this laboratory (Cambridge) by using drawings provided by Professor H. Beckey, University of Bonn, to convert it into a field-desorption system, and also on a Varian CH5 mass spectrometer (courtesy of Dr. D. E. Games, University of Cardiff). Our own source was operated with an anode (wire) voltage of 8kV and a cathode voltage of −4kV.

Nomenclature

The nomenclature for describing sequences deduced from spectra and the derivatives of the common amino acids is as previously described (Morris *et al.*, 1974). We have introduced the tentative three-letter code Gla for the new amino acid γ-carboxyglutamic acid (Magnusson *et al.*, 1974*b*), and the trivial codes X, X_c, Y and Y_c are used for derivatives of the new amino acid.

Results and Discussion

To facilitate understanding of the arguments for sequence assignment and structure assignment of the new amino acid and its derivatives these will be discussed separately, dealing with sequence assignment first.

In the discussion of mass spectra which follows, two new *N*-terminal mass numbers *m/e* 140 and *m/e* 198 and two new mass differences of 171 and 229 mass units will be apparent. These will be referred to in spectra and in the text as X, X_c, Y and Y_c respectively (Table 2). [For a list of the mass numbers and mass differences of the common amino acids see Morris *et al.* (1971).] We shall prove in the latter part of the discussion that Y_c is a derivative of the new amino acid γ-carboxyglutamic acid, giving rise to Y by 'decarboxylation', X_c by cyclization and X by 'decarboxylation' and cyclization. For the purpose of the sequence discussion therefore, whenever we can assign to the spectrum X, X_c, Y or Y_c, these signals may be thought of as originating from one amino acid.

A distinct advantage of mass spectrometry over other sequencing methods is that the technique allows the examination of peptide mixtures. Some of the samples investigated were seen to contain more than one peptide after paper purification. Further, in many cases the derivative-formation procedure caused chain cleavage at γ-carboxyglutamic acid residues, giving rise to mixed spectra (see, e.g., Fig. 2). Although this fragmentation complicated interpretation to some extent, it also helped decisively in solving the sequence of two big peptides of 17–18 residues, which would not normally be accessible to sequencing by mass spectrometry. Of the data actually collected only the minimum necessary to prove the sequence assignment is presented below.

Sequence location of the ten γ-carboxyglutamic acid residues

(*a*) *Positions* 26, 27 *and* 30. One of the first peptides examined by mass spectrometry was Prl (see Table 1). This peptide contains arginine, but, although we normally use hydrazine treatment (Morris *et al.*, 1973) to convert arginine residues into ornithine residues, it was decided not to treat peptide Prl with hydrazine in order to minimize possible chemical modification of the unknown amino acid. Fig. 2 shows the mass spectrum of an acetylated permethylated sample (200nmol) of peptide Prl at a source temperature of 220°C. The spectrum is clearly complex,

Table 2. *New mass numbers and mass differences observed*

X, X_c, Y and Y_c are codes for amino acid derivatives. For details see the text.

Mass number (*N*-terminus)	Code for derivative
m/e 140	X
m/e 198	X_c
Mass difference (in chain)	
171 mass units	Y
229 mass units	Y_c

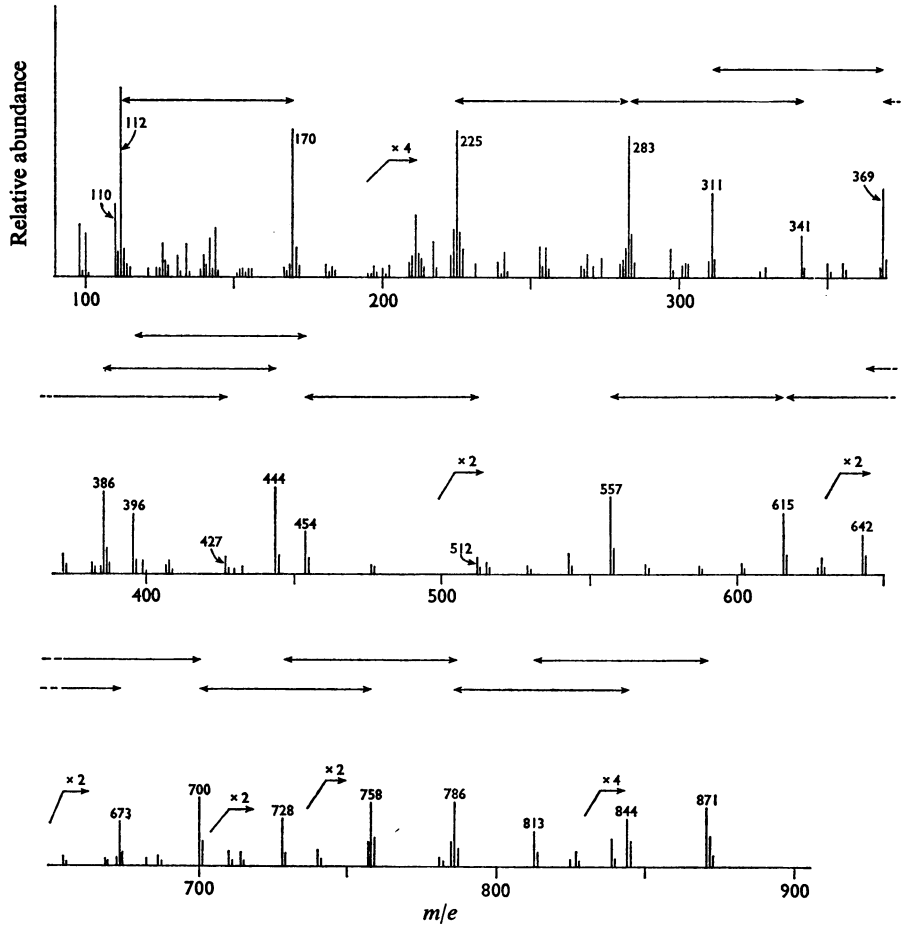


Fig. 2. *Partial mass spectrum, above m/e* 90, *of the acetyl permethyl derivative of peptide Prl*

Source temperature 220°C. Double-headed arrows indicate 58 mass unit differences; for explanation see the text. For clarity, less abundant signals are increased by a suitable factor, e.g. ×4.

having the appearance of a mixture. Fig. 2 is a scan representative of the ten spectra which were taken at different source temperatures to aid interpretation. Table 3 shows the interpretation arrived at by examination of all the scans by a normal mixture-analysis approach (Morris *et al.*, 1971). This shows that, although complex, the major signals present in Fig. 2 denote variations on only one sequence, namely:

$$X_c\text{-}Y_c\text{-Ala-Phe-}Y_c\text{-Ala}$$

All other variations are either cyclized, decarboxylated or both, as seen in Table 3. Interpretation of the spectra was made possible by observing certain patterns in the signals present, particularly the numerous 58 mass-unit differences suggestive of some species with and without carbomethoxyl groups (carboxyl groups before derivative formation). The above interpretation is substantiated by the isotope-labelling and high-resolution data described below.

At this stage in the study, the determination of the sequence of prothrombin by classical methods was completed (Magnusson *et al.*, 1975*a*), and comparison of the two sets of data showed the clear relationship seen in the scheme below:

26	27	28	29	30	31	
Glx-	Glx-	Ala-	Phe-	Glx-	Ala	Dns–Edman
X_c -	Y_c -	Ala-	Phe-	Y_c -	Ala	Mass spectrometric

The mass spectra show that residues 26, 27 and 30 are not in fact glutamic acid residues, since these

would have been apparent with a mass difference of 157 mass units for the glutamic acid derivative, or *m/e* 126 (losing CO to give *m/e* 98) for its cyclized counterpart, pyroglutamic acid. The new amino acid is therefore assigned to residues 26, 27 and 30, which have a mass of 229 mass units (Y_c) after derivative formation, and can give rise to X_c, X and Y by processes described in detail below.

(*b*) *Position* 33. Peptide CTLP4 (see Table 1) (200nmol) was examined after acetylation and permethylation. A representative spectrum is shown in Fig. 3. Again, until a pattern is recognized, the spectrum appears to be complex, but note the 58 mass unit differences between *m/e* 426 and 484 and between *m/e* 541 and 599, showing evidence of decarboxylation. Two major sequences can be assigned via signals at *m/e* 128 (acetyl-*N*-methylalanine), 255, 426, 484, 541 and 599 as Ala-Leu-Y_c-Ser and Ala-Leu-Y-Ser. In addition, signals at *m/e* 170 (X_c-CO), 313, 440 and 471 can be assigned to X_c-Ser-Leu-OMe (a *C*-terminal sequence). X_c is formed by cyclization of Y_c (see below), and the complete sequence of the peptide is thus determined as Ala-Leu-Y_c-Ser-Leu. Comparison with the sequence obtained by classical methods (Fig. 1) thus allows assignment of residue 33 as the new amino acid.

(*c*) *Positions* 7 *and* 8. Peptide CTL1, not treated with hydrazine, was acetylated and permethylated and examined in the mass spectrometer. The mass spectrum shown in Fig. 4 was obtained. Three clear *N*-terminal sequences are present, Phe-Leu-Y-Y, Phe-Leu-Y-Y_c and Phe-Leu-Y_c-Y, as shown in Fig. 4, together with a sequence X-Val, via signals at *m/e* 112

Table 3. *Sequences assigned from the mass spectra of acetyl permethyl Prl*

X, $\dot{X}_c$, Y and Y_c are codes for amino acid derivatives. For details see the text and Fig. 2.

Sequences	*m/e* values
X-Y-Ala-Phe-Y-Ala	112, 311, 396, 557, 728, 813
X-Ala-Phe-Y-Ala	112, 225, 386, 557, 642
X_c-Y-Ala-Phe-Y-Ala	170, 369, 454, 615, 786, 871
X_c-Ala-Phe-Y-Ala	170, 283, 444, 615, 700
X-Y_c-Ala-Phe-Y-Ala	112, 369, 454, 615, 786, 871
X-Y-Ala-Phe-Y_c-Ala	112, 311, 396, 557, 786, 871
X-Y_c-Ala-Phe-Y_c	112, 369, 454, 615, 844
X_c-Y_c-Ala-Phe-Y	170, 427, 512, 673, 844
X_c-Y-Ala-Phe-Y_c	170, 369, 454, 615, 844
X-Ala-Phe-Y_c-Ala	112, 225, 386, 615, 700
X_c-Ala-Phe-Y_c-Ala	170, 283, 444, 673, 758

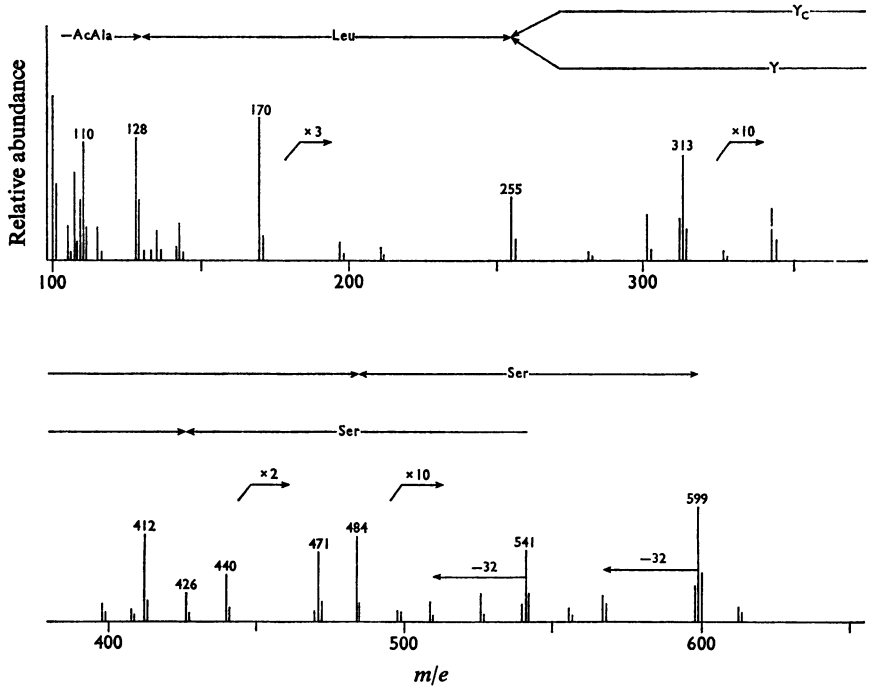


Fig. 3. *Partial mass spectrum, above* m/e 100, *of the acetyl permethyl derivative of peptide CTLP*4

Source temperature 140°C. For detailed interpretation see the text.

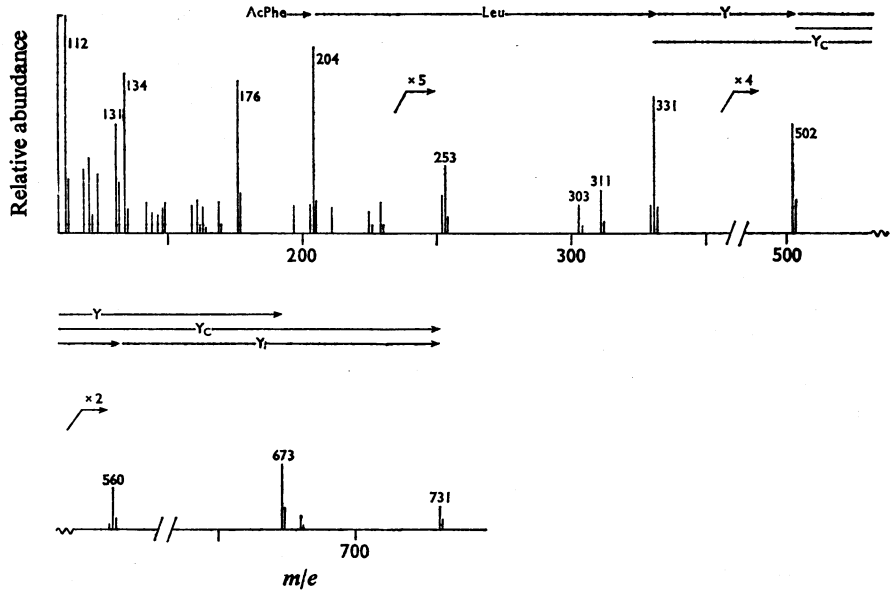


Fig. 4. *Partial mass spectrum, above m/e* 110, *of the acetyl permethyl derivative of peptide CTL*1

Source temperature 230°C. For detailed interpretation see the text.

(X−CO) and 253. Comparison with the classically obtained sequence of residues 1–35 of prothrombin (Fig. 1) allows assignment of the new amino acid to residues 7 and 8. A smaller peptide Phe-Leu-Glx-Glx (residues 5–8) from this region has also been studied by Stenflo *et al.* (1974) by n.m.r. and mass spectrometry. Although these authors misinterpreted the data from their mass-spectrometry experiment (see below), our analysis agrees with their assignment of two γ-carboxyglutamic acid residues to positions 7 and 8 in prothrombin. We have used the same peptide (CTLP5) for an isotope study, which will be discussed below.

(*d*) *Positions* 15, 17, 20 *and* 21. This part of the prothrombin sequence was the most difficult to study both by classical methods and by mass spectrometry. From a mass-spectrometric viewpoint the presence of arginine and cysteine (carboxymethylated) residues posed an additional problem in the peptide studied (CTLP6). Comparison with the classically obtained sequence (Fig. 1) suggests that this peptide corresponds to residues 14–30 in the sequence. Because of its size (17 residues) this peptide was expected to be particularly difficult to sequence by mass spectrometry.

Examination of peptide CTLP6 and other similar peptides, from which derivatives were made without prior hydrazine treatment, enabled us to assign the Glx residues 17, 20 and 21 of the classically derived sequence to the new amino acid. However, it was not possible to obtain unambiguous information on Glx-15 by this method. Therefore a sample of peptide CTLP6 was treated with hydrazine before acetylation and permethylation to convert arginine into ornithine (Morris *et al.*, 1973). A control experiment had been carried out to see if hydrazine treatment affected the new amino acid in a peptide that contained no arginine. No difference was observed between spectra obtained with and without hydrazine treatment.

The spectrum of the derivative of peptide CTLP6 proved complex, and in addition an ambiguity arose owing to the fact that *m/e* 170 can correspond both to X_c-CO and AcLeu. A further isotopic derivative was therefore prepared as follows: (1) treatment with hydrazine as before; (2) acetylation in $(C^2H_3CO)_2O/CH_3OH$; (3) permethylation as before. The mass spectrum produced from this derivative is shown in Fig. 5. Before proceeding with the interpretation two points need to be noted.

(1) Treatment of the peptide with hydrazine converts arginine into ornithine. This amino acid would give a mass difference of 184 mass units for its normal acetylated derivative, but, since $(C^2H_3CO)_2O$ was used, the mass difference will shift to 187 mass units.

(2) When CmCys*-containing peptides are permethylated, β-elimination is almost quantitative, giving rise to dehydroalanine (83 mass units) in the mass spectrum (Morris *et al.*, 1973).

* Abbreviation: CmCys, carboxymethylcysteine.

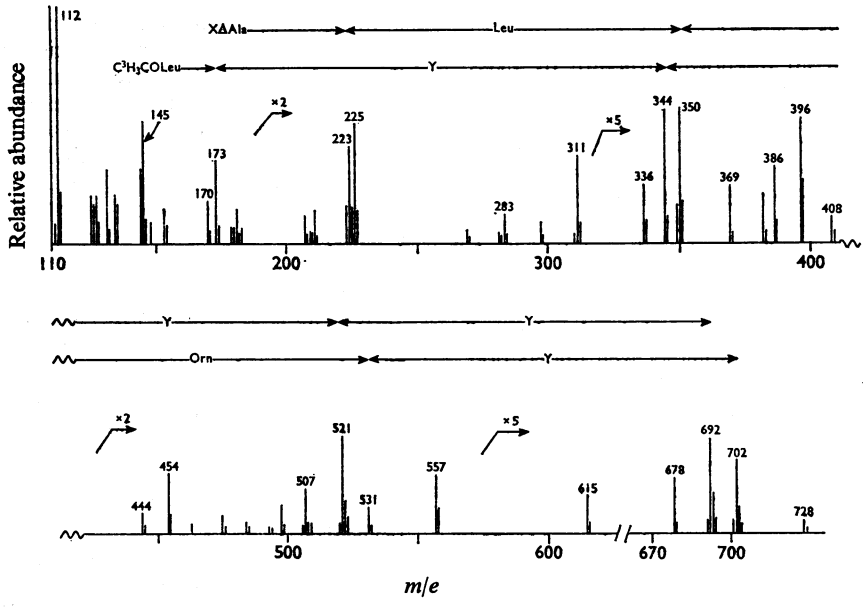


Fig. 5. *Partial mass spectrum, above m/e* 110, *of the deuteroacetyl permethyl derivative of hydrazinolysed peptide CTLP6*

Source temperature 270°C. For detailed interpretation see the text.

The use of $(C^2H_3CO)_2O$ in the acetylation reaction allows an *N*-terminal leucine to be assigned to *m/e* 173 (−CO at *m/e* 145) in Fig. 5. A signal for X_c-CO remains at *m/e* 170. The only other signals to shift when $(C^2H_3CO)_2O$ was used instead of $(CH_3CO)_2O$ were *m/e* 344 (shifted from *m/e* 341), *m/e* 531 (from *m/e* 525) and *m/e* 702 (from *m/e* 696). A clear connexion is thus formed between these signals, and examination of the mass differences present reveals the sequence Leu-Y-Orn-Y (derived from Leu-Y-Arg-Y). Comparison with the classically deduced sequence shows that this sequence corresponds to Leu-Glx-Arg-Glx-(14–17), and therefore residues 15 and 17 are also the new amino acid.

Since none of the remaining signals in Fig. 5 shift when $(CH_3CO)_2O$ is exchanged for $(C^2H_3CO)_2O$ in the acetylation reaction, it follows that they must all originate from cyclic *N*-termini (or from N–C cleavage, which is not relevant here). These are clearly present at *m/e* 112 (X−CO) and *m/e* 170 (X_c-CO) in Fig. 5. Comparison of the spectrum with Fig. 2 and Table 3, discussed above, shows that *m/e* 225, 311, 369, 386, 396, 454, 557, 615 and 728 originate from residues 26–30 in the prothrombin sequence and can be assigned to X-Ala-Phe, X-Y-Ala-Phe-Y, X_c-Y-Ala-Phe and X-Y_c-Ala-Phe. These signals are expected, since residues 26–30 constitute part of the sequence of peptide CTLP6, and ready cyclization leads to the smaller fragments observed above. Even if losses of CO from the above sequence ions and some expected degree of undermethylation (Morris, 1972) are taken into account, there are still major signals left unassigned, namely *m/e* 223, 350, 521 and 692 (see Fig. 5).

The signal at *m/e* 223 can be assigned only to X-ΔAla (X-dehydroalanine derived from X-CmCys as described above); *m/e* 350 minus *m/e* 223 gives 127 mass units, a leucine fragment. For the signals at *m/e* 521 and 692 we can assign the sequence:

X-ΔAla-Leu-Y-Y

This assignment was also verified by studies of a number of other peptide samples where this sequence was the major component. Comparison with the classically derived sequence (Fig. 1) thus shows that the new amino acid occurs in positions 17, 20 and 21 in prothrombin.

In summary, the mass-spectrometric evidence shows that the new amino acid occurs as residues 7, 8, 15, 17, 20, 21, 26, 27, 30 and 33 in prothrombin. In the sequencing by classical methods (Dns–Edman procedure using hydrolysis with 6M-HCl at 110°C for 10h to release Dns-amino acids from Dns-peptides) these residues were identified as Dns-Glu, although the electrophoretic mobilities of the peptides had shown the presence of extra negative charges (Magnusson, 1973; Magnusson *et al.*, 1974*a*,*b*).

Structure of X, X_c, Y and Y_c

When studying an unknown structure by mass spectrometry it is important to choose a procedure for forming volatile derivatives, which allows confident deduction of the unknown structure from the resulting spectra. Ordinary acyl esters are among the simplest peptide derivatives, but this derivative is suitable only for small non-polar peptides, and is of no value in the study of more complex peptides such as those examined here. For this reason we chose the acetyl permethyl derivative. This has been in use in peptide studies for some years, and the chemistry of derivative formation is reasonably well understood (see, e.g., Morris *et al.*, 1973). Despite this, a new feature of one of these reactions was uncovered during the present study.

The act of forming a derivative can be put to advantage when interpreting the spectrum, by using stable isotope analogues, and these were used extensively in the structure determination of X, X_c, Y and Y_c.

(1) *Use of $(C^2H_3CO)_2O$ in the acetylation reaction.* The use of deuterated acetic anhydride in the acetylation procedure instead of $(CH_3CO)_2O$ followed by permethylation in the usual manner gave no change in the respective mass numbers of X, X_c, Y or Y_c, thus showing that no acetyl group is incorporated during their formation.

(2) *Use of C^2H_3I in the permethylation reaction.* The use of the normal acetylation procedure followed by permethylation with C^2H_3I instead of CH_3I produced the shifts given in Table 4 for X, X_c, Y and Y_c, showing that each of the species X, X_c, Y and Y_c is formed by addition of two methyl groups during permethylation.

At the time this was a puzzling result, bearing in mind our observation of the 58 mass unit differences in the spectra, and the similarity in behaviour (ready cyclization) to that of glutamic acid. On the basis of these early results, a tentative assignment of some type of carboxylated glutamic acid was made. However, such a precursor would be expected to add at least three methyl groups on permethylation, namely the two normally added on to glutamic acid residues (N-Me, -OMe) and one on the extra carboxyl group (-OMe).

(3) *High-resolution mass measurements.* The mass measurements shown in Table 5 were made on various samples at times when sufficient peptide was available for both low- and high-resolution analysis. The data can be analysed to give the following atomic compositions: X = $C_7H_{10}NO_2$; X_c = $C_9H_{12}NO_4$; Y = $C_8H_{13}NO_3$; Y_c = $C_{10}H_{15}NO_5$. The 58 mass unit difference observed on the low-resolution scans is also confirmed as an exchange of H for $C_2H_3O_2$ on going from X and Y to X_c and Y_c respectively.

The above data again point to the precursor of these signals being a carboxylated glutamic acid. X corresponds in mass and atomic composition to a methylated pyroglutamic acid residue of the type

Table 4. *Mass shifts of X, X_c, Y and Y_c on deutero-permethylation of acetylated peptides*

X, X_c, Y and Y_c are codes for amino acid derivatives. For details see the text.

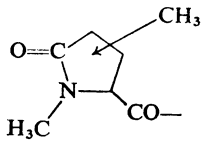
Code	Normal mass value (CH_3I)	Shift observed (C^2H_3I)	No. of methyl groups added
X–CO	112	118	2
X_c–CO	170	176	2
Y	171	177	2
Y_c	229	235	2

Table 5. *High-resolution mass measurements on peptide fragments from prothrombin*

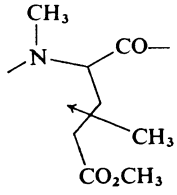
Other compositions (other than the elemental composition) whose calculated masses are within 10 p.p.m. of the measured value are possible, but none are sensible for peptide derived fragments. An example is $C_4H_8N_4$ for *m/e* 112. The error is the difference between measured mass and calculated mass of the proposed formula. X, X_c, Y and Y_c are codes for amino acid derivatives. For details see the text.

Sample	Assignment	Measured mass	Elemental composition	Error (p.p.m.)
CTLP4	X_c minus $(CO+C_2H_4O_2)$	110.0608	C_6H_8NO	1.7
Prl	X minus CO	112.0759	$C_6H_{10}NO$	2.7
Prl	X_c minus CO	170.0818	$C_8H_{12}NO_3$	<1
Prl	X-Ala	225.1234	$C_{11}H_{17}N_2O_3$	2.0
Prl	X-Y	311.1601	$C_{15}H_{23}N_2O_5$	4.5
CTLP4	X_c-Ser	313.1404	$C_{14}H_{21}N_2O_6$	1.3
Prl	X_c-Y_c minus CO	399.1772	$C_{18}H_{27}N_2O_8$	1.0
CTLP4	Ala-Leu-Y	426.2611	$C_{21}H_{36}N_3O_6$	1.5
Prl	X_c-Y_c	427.1716	$C_{19}H_{27}N_2O_9$	5.0
CTLP4	Ala-Leu-Y_c	484.2653	$C_{23}H_{38}N_3O_8$	1.2

shown below (arrow indicates that the position of substitution is not assigned):



The C^2H_3I experiment described above showed that both methyl groups originate from the permethylation, and are therefore not present in the original structure. Analogously, Y corresponds in mass and atomic composition to a methylated glutamic acid residue of the type shown below:



and X_c and Y_c correspond in mass and atomic composition to carbomethoxylated versions of X and Y respectively. Further, Y_c and Y would be expected to yield X_c and X on cyclization, in the same manner in which a glutamic acid derivative yields a pyroglutamic acid derivative in normal peptides (see, e.g., Morris *et al.*, 1971).

However, the anomaly remains that each derivative X, X_c, Y and Y_c adds on only two methyl groups during permethylation. This anomaly is resolved in the following experiment, which is crucial to the understanding of the formation of X, X_c, Y and Y_c and therefore to the assignment of structure to the new amino acid. It is worth noting here, with respect to the discrepancy in mass difference of the acetylated permethylated derivative between our data, 229 mass units (Magnusson *et al.*, 1974*b*), and those of Stenflo *et al.* (1974), 201 mass units, that, had these authors used C^2H_3I to confirm the interpretation of their mass spectrum, it would have been apparent that their conclusions about derivative formation were incorrect. Fortuitously, however, this misinterpretation, together with n.m.r. evidence, led the latter authors to a structure for the precursor amino acid that our data show is correct.

(4) *Use of $C^2H_3O^2H$ as solvent in the acetylation reaction.* Peptide CTLP5 was acetylated in acetic anhydride/$C^2H_3O^2H$ (1:4, v/v) and then permethylated in the normal manner. Fig. 6 shows the mass spectrum at different source temperatures. Comparison with Fig. 4 shows immediately that those signals containing X_c, Y and Y_c, but not X, have shifted in mass. This indicates that one or two methoxyl groups from the methanol are incorporated during the acetylation step in the formation of X_c, Y and Y_c.

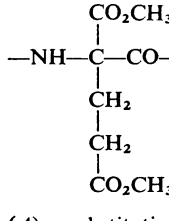
The weaker signals 3 or 6 mass units less than the numbered signals in Fig. 6 may be ignored for the purpose of our discussion, since they are mainly due to transmethylation (and therefore loss of deuterium label) during the permethylation reaction.

Fig. 6(*a*) shows that X_c incorporates one methoxyl group during the acetylation step, via the signal from ('X_c+$^2H_3$')−CO at *m/e* 173. Figs. 6(*a*), (*b*) and (*c*) all show evidence for the incorporation of two methoxyl groups in Y_c from the numerous signals containing a 'Y_c+$^2H_6$' moiety. Fig. 6(*c*) is the closest in appearance to Fig. 4 (no deuterium label), and the mass shifts can be clearly seen. Note, for example, the shift in the Phe-Leu-Y signal from *m/e* 502 to *m/e* 505, showing the incorporation of one methoxyl group in the formation of Y.

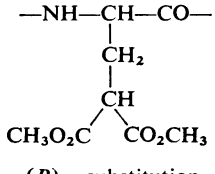
These results are summarized in Table 6 and provide the key to solving the structures of X, X_c, Y and Y_c, and therefore also of the new amino acid.

Combining the results of experiments (1)–(4) above, Y_c is shown to be a carbomethoxylated *C*-methylated *N*-, *O*-permethylated glutamic acid residue, in which both of the ester methyl groups come from the methanol used in the acetylation reaction, and the two remaining methyl groups come from the methyl iodide in the permethylation reaction. X_c can then be formed from Y_c simply by cyclization, retaining one methyl (methoxyl) group (from the methanol) and both methyl groups from the methyl iodide. The origin of X and Y will be discussed below.

Since we have proved that all of the methyl groups present in Y_c come from the reagents used in derivative formation, it follows that the new amino acid is a carboxylated glutamic acid residue. The mass of this amino acid is confirmed by the field-desorption spectrum of a sample of the natural product, shown in Fig. 7. The base peak at *m/e* 192 (*M*+H) and the associated losses of water and CO_2 are fully consistent with the carboxyglutamic acid structure. The ease with which this carboxyglutamic acid is decarboxylated suggests that the extra carboxyl group is substituted on either the α-carbon or the γ-carbon. Both these potential precursors of Y_c could form a diester on treatment with acetic anhydride in methanol, as shown below:



(*A*) α-substitution



(*B*) γ-substitution

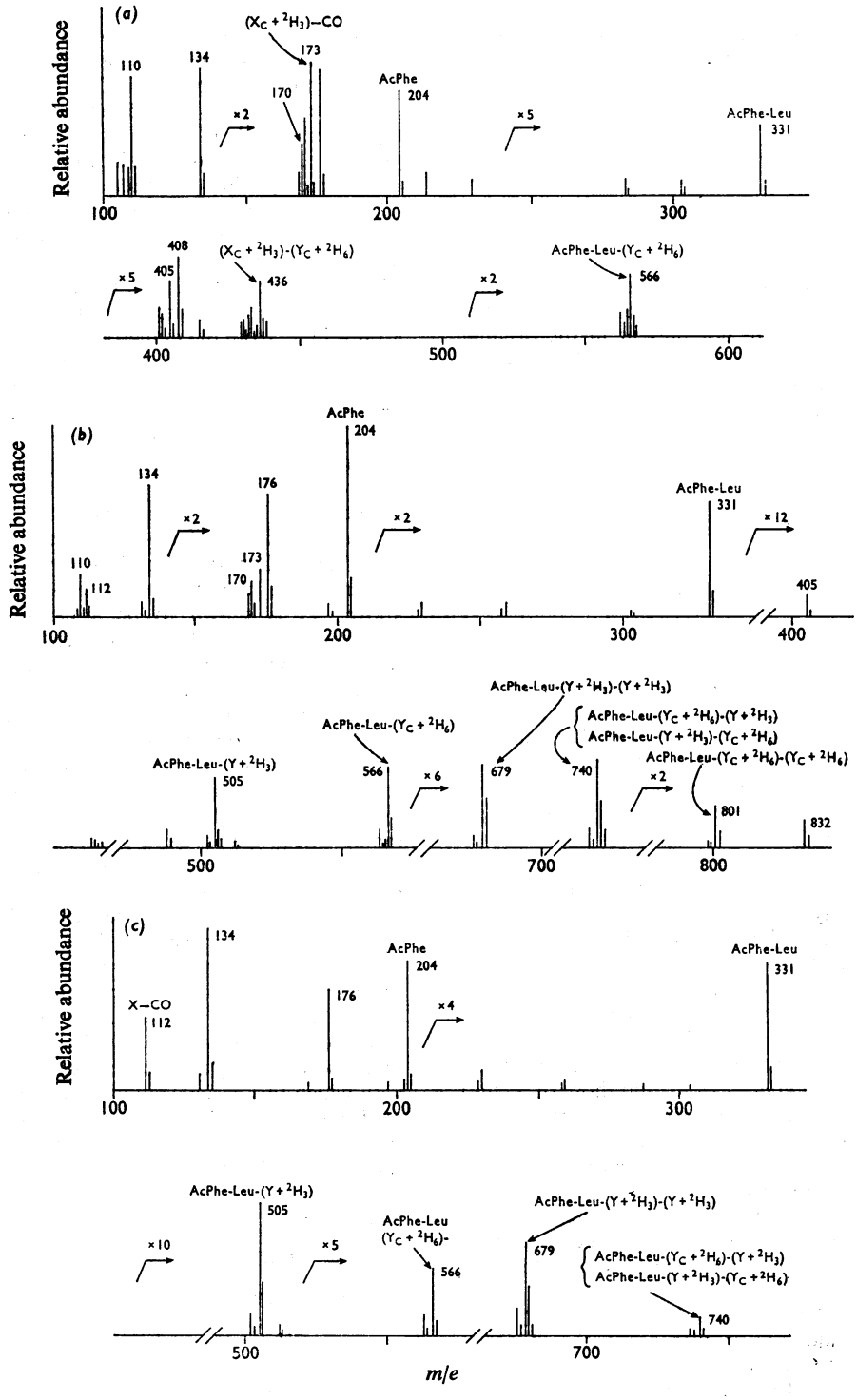


Fig. 6. *Partial mass spectrum, above m/e* 100, *of a derivative of peptide CTLP5, acetylated in deuteromethanol, and permethylated* (*a*) Source temperature 160°C; (*b*) source temperature 200°C; (*c*) source temperature 250°C. Where two sequence possibilities are written, either or both may contribute to the signal. For detailed interpretation see the text.

Table 6. *Mass shifts of X, X_c, Y and Y_c on acetylation with $(CH_3CO)_2O$ in $C^2H_3O^2H$ and permethylation*

X, X_c, Y and Y_c are codes for amino acid derivatives. For details see the text and Fig. 6.

Code	Normal mass value	Shift observed	Number of methyl groups added from $C^2H_3O^2H$
X–CO	112	No change	0
X_c–CO	170	173	1
Y	171	174	1
Y_c	229	235	2

The glutamic acid side chain in structure (*A*) is preserved, and we can predict the chemistry of this compound during the permethylation reaction from a knowledge of the behaviour of γ-glutamate esters under the same reaction conditions. We have shown (K. E. Batley, A. Dell, D. Mak & H. R. Morris, unpublished work) that permethylation of peptides containing γ-methylglutamate yields a complex mixture of products, none of which show the expected sequence fragment for *N,O*-permethylglutamic acid. On this basis it would be impossible for structure (*A*) above (the α-carbomethoxy linkage) to yield the clean signal (229 mass units) actually observed for Y_c. As a final proof of the γ-linkage, we have synthesized γ-carboxyglutamic acid (Gla) (Morris *et al.*, 1975), and its properties are identical with those of the naturally occurring amino acid isolated from a partial acid hydrolysis of peptide C. Both the synthetic and natural γ-carboxyglutamic acids give the mass spectrum shown in Fig. 8 (for the synthetic) as the acetyl ester derivatives. It should be pointed out here that, although we have no convincing evidence for a substitution on any of the 20 γ-carboxyl groups in the ten γ-carboxyglutamic acid residues of native prothrombin, our discovery of the easy esterification of these functions opens up the possibility that a substituted γ-carboxyglutamic acid residue could give the same derivatives as an unsubstituted one, via an initial transesterification step.

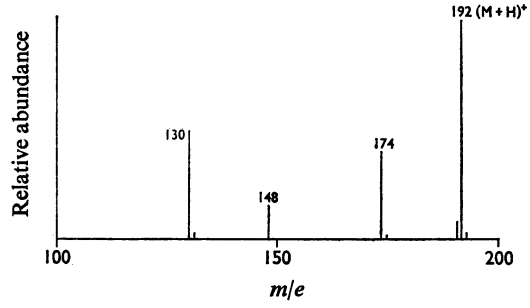
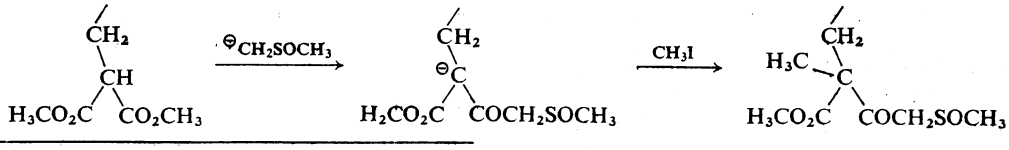


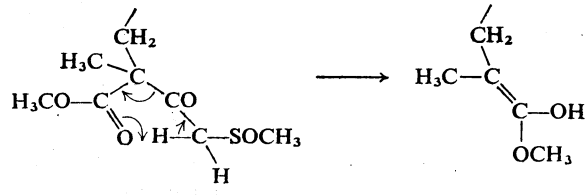
Fig. 7. *Field-desorption spectrum of the new amino acid γ-carboxyglutamic acid isolated from prothrombin*

The sample desorbed from the wire anode at a current of 15mA.

We are now in a position to return to the structures of X and Y. It may be noticed in Fig. 6 that, at the lower temperatures, species containing X_c and Y_c predominate, whereas at higher temperatures the decarboxylated analogues are more abundant. It follows that these species are either products of thermal decomposition of the X_c and Y_c analogues (taking place in the mass spectrometer) or alternatively that they are decomposition products of some moiety of lower volatility than X_c and Y_c. It is possible that both mechanisms are operative, although no metastable transitions are observed in any of the spectra. We have evidence for the operation of the second mechanism, via the formation of a β-keto sulphoxide on Y_c and X_c and subsequent decomposition. This is a chemically reasonable process to occur during permethylation of a compound containing ester groups, as we show in the scheme below, which depicts a possible reaction product of the γ-carboxyglutamic acid side chain in the permethylation reaction, after acetylation has produced the γ-carboxyglutamic acid dimethyl ester:



In the mass spectrometer the sulphoxide derivative may be degraded as shown below, giving rise to Y (shown as the enol form):



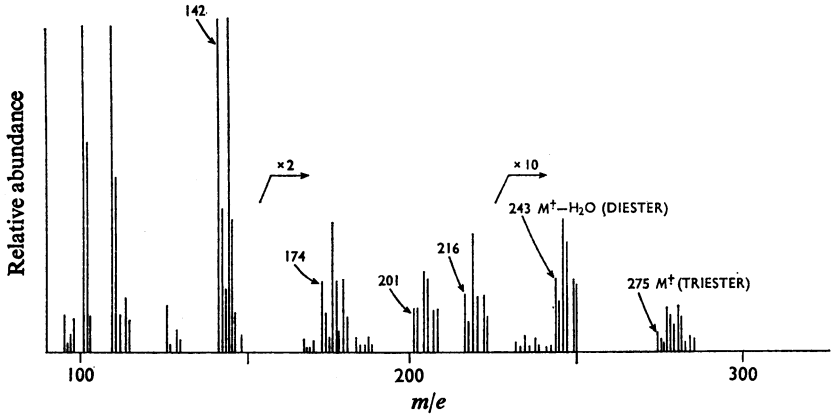
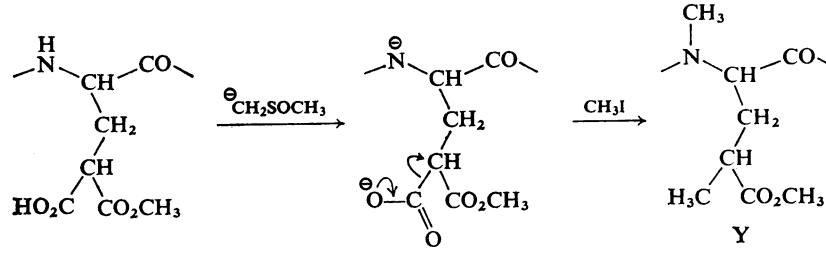


Fig. 8. *Partial mass spectrum, above m/e* 90, *of the acetyl derivative of γ-carboxyglutamic acid*

Acetylation was accomplished by using acetic anhydride in a 1:1 mixture of CH_3OH and $C^2H_3O^2H$. The triester (*m/e* 275) was a minor component, not present in all scans. Source temperature 110°C.

This mechanism is substantiated by the use of $(C^2H_3)_2SO$ in the permethylation reaction. This causes a shift of a substantial part of the X and Y signals from *m/e* 112 to *m/e* 113 (X−CO) and from *m/e* 171 to *m/e* 172 (Y) respectively. The occurrence of extensive ester formation during the mild acetylation step was unexpected, although again not chemically unreasonable. Under the mild conditions used, the γ-carboxyl group of glutamic acid itself is virtually unaffected after 3 h, although some modification of the α-carboxyl group in peptides is beginning to take place (K. E. Batley, A. Dell, D. Mak & H. R. Morris, unpublished work). It has been reported previously that acetic anhydride in methanol can slowly esterify free carboxyl groups in proteins (Blackburn & Phillips, 1944), although the reaction time for approx. 1% ester formation (20 h) was much longer than the procedure that we have followed in making peptide derivatives (3 h at room temperature). The pK_a of the γ-carboxyl group will, of course, be affected noticeably by the substitution of a second carboxyl group on the same carbon atom. Ester formation proceeds, probably via the cyclic and/or mixed anhydride, with some catalysis from the peptide backbone.

Of course Y (and therefore X) could be formed from the mono ester of γ-carboxyglutamic acid. During permethylation, this may be expected to happen as follows:



This reaction can occur only to a small extent, because the evidence of the mass spectra shows that the majority of X- and Y-containing fragments volatilize, or are observed, at higher temperatures than their X_c- and Y_c-containing counterparts.

Conclusion

The above data give clear mass-spectrometric evidence proving that the structure of the new amino acid in prothrombin is γ-carboxyglutamic acid, and that it is present in ten positions in the first 33 residues of the prothrombin sequence (Fig. 1).

The products of derivative formation, an understanding of which will be important to future studies of this type, are summarized (for the acetylation and permethylation reactions) in Fig. 9. The use of acetic anhydride in methanol for the acetylation reaction, although complicating the interpretation, was in one way fortunate. We find that if a reagent is chosen which does not give ester formation, permethylation

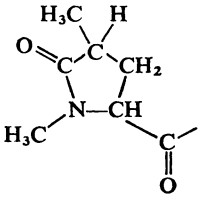
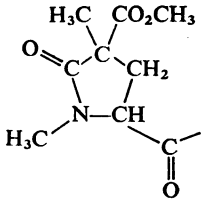
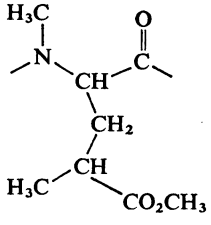
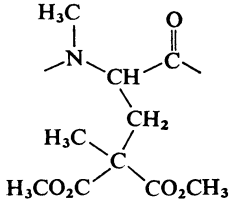
Code	Mass	Plausible structure	Common fragment ion
X	140		minus CO
X_c	198		minus CO
Y	171		–
Y_c	229		–

Fig. 9. *Plausible structures for the products of derivative formation of γ-carboxyglutamic acid*

Derivatives are formed by treatment of a peptide (or amino acid) with acetic anhydride in methanol, followed by permethylation. For details of these reactions see the text.

of the di-acid leads to extensive by-product formation, and must therefore be avoided. This study also emphasizes the relative ease with which unknown mixtures of peptides can be analysed by mass spectrometry. In several cases it turned out that cleavage of the peptide chain at γ-carboxyglutamic acid residues (e.g. at $Arg_{(16)}$-$Gla_{(17)}$, $Gla_{(17)}$-$Cys_{(18)}$ and $Arg_{(25)}$-$Gla_{(26)}$) occurred during the derivative-formation procedure. This type of cleavage was useful in obtaining small fragments from some of the large peptides.

Function of vitamin K

Available evidence suggests that dicoumarol prothrombin has normal glutamic acid residues in those ten positions where normal prothrombin has γ-carboxyglutamic acid residues, although so far only four of those residues (7, 8, 15 and 17) have actually been sequenced in dicoumarol prothrombin (Stenflo, 1974). Since no other structural differences between the two prothrombins have yet been detected it seems reasonable to conclude that the role of vitamin K in the biosynthesis of prothrombin is to support a post-synthetic γ-carboxylation of glutamic acid (Girardot *et al.*, 1974). Our data (Magnusson *et al.*, 1974*b*), presented in detail here, show that all ten glutamic acids in the *N*-terminal region of prothrombin are affected. The first glutamic acid that is not γ-carboxylated occurs in position 49 of normal bovine prothrombin.

A recent paper on bovine factor X (Howard & Nelsestuen, 1975) has made reference to a sequence for the *N*-terminus of bovine prothrombin which differs from that deduced from the data presented here (Fig. 1, Magnusson *et al.*, 1974*b*). However, no evidence for the sequence has yet been given, and only indirect methods were used to estimate the number of γ-carboxyglutamic acids present (Howard & Nelsestuen, 1975).

The homology between the A-fragment (residues 1–156) of prothrombin (Magnusson *et al.*, 1974*b*, 1975*a*,*b*) and the light chain of factor X_1 (Enfield *et al.*, 1975) is quite striking (75% in the first 32–33 residues, 62% in the first 44–45) in the *N*-terminal region, but the rest of the two structures show little or no sequence homology.

Function of γ-carboxyglutamic acid

The ability of prothrombin to be adsorbed on insoluble barium salts (sulphate, citrate) was used in the purification of the chymotryptic fragment 1–35, which contains all ten γ-carboxyglutamic acid residues (Magnusson *et al.*, 1974*b*), and of a tryptic fragment (Nelsestuen & Suttie, 1973), probably residues 12–35, containing the last eight of these residues. No other peptides were adsorbed in this way from chymotryptic digests of native prothrombin. Therefore it seems logical to conclude that the adsorption of prothrombin on barium salts is a consequence of its containing γ-carboxyglutamic acids.

The strong Ca^{2+}-binding of prothrombin, which is not shared by dicoumarol prothrombin, can also be attributed to the γ-carboxyglutamic acids. Not only prothrombin itself (Ganrot & Nilehn, 1968*b*) but also peptides containing γ-carboxyglutamic acid change their electrophoretic mobility if Ca^{2+} ions are added to the buffer (Stenflo, 1974).

The rapid Factor X_a-catalysed activation of prothrombin to thrombin in the presence of Factor V and phospholipids involves the formation of a Ca^{2+}-dependent complex of prothrombin and phospholipids (Bull *et al.*, 1972). The importance of the A-fragment, and more specifically the γ-carboxyglutamic acid residues, for the formation of this complex is demonstrated by the fact that neither neoprothrombin-S, which lacks the entire A-fragment (Magnusson, 1962, 1965*a*; Magnusson *et al.*, 1975*a*) because of thrombin-catalysed cleavage (Seegers *et al.*, 1967; Esmon *et al.*, 1974) at position 156 (Magnusson *et al.*, 1975*b*), nor dicoumarol prothrombin, which contains the A-fragment but lacks the extra carboxyl groups (Stenflo, 1974; Stenflo *et al.*, 1974), can be activated (Magnusson, 1962; Josso *et al.*, 1968; Ganrot & Nilehn, 1968*a*) in phospholipid-dependent systems.

The extensive sequence homology of the vitamin K-dependent Ca^{2+}-binding domains (residues 1–45) of prothrombin and Factor X shows that this region has been strongly conserved during evolution and almost certainly means that it has the same tertiary structure and function in the two proteins.

We thank Miss Laila Brøns, Miss Lene Kristensen, Miss Lone Christensen and Miss Margit Skriver for technical assistance. We are grateful to the Medical Research Council, Great Britain (Grant G972/807/B), and to the U.S. National Heart and Lung Institute (Grant HL16238) for financial support. T. E. P. and L. S. received fellowships from the Danish Science Research Council, and A. D. was recipient of an 1851 Exhibition fellowship.

References

Ambler, R. P. (1963) *Biochem. J.* **89**, 349–378
Barnhart, M. I. (1960) *Am. J. Physiol.* **119**, 360–366
Barton, P. G. & Hanahan, D. J. (1969) *Biochim. Biophys. Acta* **187**, 319–327
Blackburn, S. & Phillips, H. (1944) *Biochem. J.* **38**, 171–178
Bull, R. K., Jevons, S. & Barton, P. G. (1972) *J. Biol. Chem.* **247**, 2747–2754
Campbell, H. A. & Link, K. P. (1941) *J. Biol. Chem.* **138**, 21–28
Dam, H. (1935) *Nature (London)* **135**, 652–653
Dell, A. & Morris, H. R. (1974) *Biochem. Biophys. Res. Commun.* **61**, 1125–1132
Enfield, D. L., Ericsson, L. H., Walsh, K. A., Neurath, H. & Titani, K. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 16–19
Esmon, C. T., Owen, W. G. & Jackson, C. M. (1974) *J. Biol. Chem.* **249**, 606–611
Ganrot, P. O. & Nilehn, J. E. (1968*a*) *Scand. J. Clin. Lab. Invest.* **21**, 238–244
Ganrot, P. O. & Nilehn, J. E. (1968*b*) *Scand. J. Clin. Lab. Invest.* **22**, 23–28
Girardot, J.-M., Delaney, R. & Johnson, B. C. (1974) *Biochem. Biophys. Res. Commun.* **59**, 1197–1203
Goldstein, R., LeBolloc'h, A., Alexander, B. & Zondermann, E. (1959) *J. Biol. Chem.* **234**, 2857–2866
Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475
Hakamori, S. I. (1964) *J. Biochem. (Tokyo)* **55**, 205–213
Heilmann, J., Barrollier, J. & Watzke, E. (1957) *Hoppe-Seyler's Z. Physiol. Chem.* **309**, 219–220
Howard, J. B. & Nelsestuen, G. L. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1281–1285
Josso, F., Lavergne, J. M., Gouault, M., Prou-Wartelle, O. & Soulier, J.-P. (1968) *Thromb. Diath. Haemorrh.* **20**, 88–98
Lewis, M. L. & Ware, A. G. (1953) *Proc. Soc. Exp. Biol. Med.* **84**, 636–643
Magnusson, S. (1962) *Thromb. Diath. Haemorrh.* **7**, *Suppl.* **1**, 229–236
Magnusson, S. (1965*a*) *Ark. Kemi* **23**, 271–283
Magnusson, S. (1965*b*) *Ark. Kemi* **23**, 285–298
Magnusson, S. (1970) *Methods Enzymol.* **19**, 157–184
Magnusson, S. (1972) *Folia Haematol. (Leipzig)* **98**, 385–390
Magnusson, S. (1973) *Thromb. Diath. Haemorrh. Suppl.* **54**, 31–35

Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Klemmensen, P. & Kouba, E. (1974*a*) *Thromb. Diath. Haemorrh. Suppl.* **57**, 153–157

Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Morris, H. R. & Dell, A. (1974*b*) *FEBS Lett.* **44**, 189–193

Magnusson, S., Sottrup-Jensen, L., Petersen, T. E. & Claeys, H. (1975*a*) in *Boerhaave Symposium on Prothrombin and Related Coagulation Factors*, pp. 25–46, University Press, Leiden

Magnusson, S., Petersen, T. E., Sottrup-Jensen, L. & Claeys, H. (1975*b*) *Cold Spring Harbor Symp. Quant. Biol.* in the press

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

Morris, H. R., Williams, D. H., Midwinter, G. G. & Hartley, B. S. (1974) *Biochem. J.* **141**, 701–713

Morris, H. R., Thompson, M. R. & Dell, A. (1975) *Biochem. Biophys. Res. Commun.* **62**, 856–861

Nelsestuen, G. L. & Suttie, J. W. (1972*a*) *J. Biol. Chem.* **247**, 8176–8182

Nelsestuen, G. L. & Suttie, J. W. (1972*b*) *Biochemistry* **11**, 4961–4964

Nelsestuen, G. L. & Suttie, J. W. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3366–3370

Nelsestuen, G. L., Zytkovicz, T. H. & Howard, J. B. (1974) *J. Biol. Chem.* **249**, 6347–6350

Offord, R. E. (1966) *Nature (London)* **211**, 591–593

Seegers, W. H., Marciniak, E., Kipfer, R. K. & Yasunaga, K. (1967) *Arch. Biochem. Biophys.* **121**, 372–383

Spackman, D. H., Stein, W. H. & Moore, S. (1958) *Anal. Chem.* **30**, 1190–1206

Stenflo, J. (1972) *J. Biol. Chem.* **247**, 8167–8175

Stenflo, J. (1973) *J. Biol. Chem.* **248**, 6325–6332

Stenflo, J. (1974) *J. Biol. Chem.* **249**, 5527–5535

Stenflo, J. (1975) in *Boerhaave Symposium on Prothrombin and Related Coagulation Factors*, pp. 152–158, University Press, Leiden

Stenflo, J. & Ganrot, P. O. (1973) *Biochem. Biophys. Res. Commun.* **50**, 98–104

Stenflo, J., Fernlund, P., Egan, W. & Roepstorff, P. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2730–2733

Thomas, D. W., Das, B. C., Gero, S. D. & Lederer, E. (1968) *Biochem. Biophys. Res. Commun.* **33**, 519–527

Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412

Woods, K. R. & Wang, K.-T. (1967) *Biochim. Biophys. Acta* **133**, 369–370

Yamada, S. & Itano, H. A. (1966) *Biochim. Biophys. Acta* **130**, 538–540