Phytochemistry, 1978, Vol. 17, pp. 735–739. Pergamon Press. Printed in England.

AMINO ACID SEQUENCE OF PLASTOCYANIN FROM *SOLANUM CRISPUM* USING AUTOMATIC METHODS

BARRY G. HASLETT, I. MARTA EVANS and DONALD BOULTER
Department of Botany, University of Durham, Durham DH1 3LE, U.K.

(*Received* 4 *November* 1977)

Key Word Index—*Solanum crispum*; Solanaceae; amino acid sequence; plastocyanin; automatic sequencing; solid-phase sequencer; solution sequencer.

Abstract—The complete amino acid sequence of plastocyanin of *Solanum crispum* L. was determined on 6 micromole protein using automatic solid-phase and spinning cup sequencer methods. Overlap between residues Lys 54 and Ile 55 was however not demonstrated. The protein consists of 99 residues.

INTRODUCTION

The development of automatic methods for Edman degradation of peptides and proteins has been the most important methodological advance in the field of amino acid sequence determination in the last ten years. Bridgen [1, 2] has published a comparison of the two methods available, using a spinning cup (solution) sequencer and a solid phase sequencer.

In recent years a number of complete amino acid sequences of plastocyanin from higher plants have been determined from data obtained by using a Beckman spinning cup sequencer in conjunction with that from manual subtractive Edman degradation [3]. This paper presents the amino acid sequence of plastocyanin from a higher plant *Solanum crispum* L., determined solely by automatic methods; the use of the solid-phase method is briefly discussed.

RESULTS

Cleavage of *S*-carboxymethyl plastocyanin by CNBr resulted in the production of three peptides which were purified by gel filtration. These peptides were designated X-1, X-2 and X-3 in accordance with their order of elution from the column. *N*- and *C*-terminal analyses of these peptides and of the uncleaved protein are given in Table 1. The fragments can be unequivocally positioned in the order X-1, X-2, X-3 in the intact protein from their amino-acid compositions, and their *N*- and *C*-termini as compared to those of the uncleaved protein (Tables 1 and 2). Similar fragments have been obtained previously by CNBr cleavage of potato plastocyanin [4] and French bean plastocyanin [5].

Amino acid sequence of X-1

The first 40 residues of the X-1 peptide (see Fig. 1) were established by automatic Edman degradation using 6 mg of the native protein in a Beckman 890C spinning cup sequencer. X-1 (500 nmol) was digested with trypsin and the resulting peptide mixture was subjected to the homoserine lactone coupling procedure in the presence of 40 mg of TETA resin. The resin was dried, mixed with 450 mg glass beads and packed into the reaction column of the solid-phase sequencer.

Identification of PTH derivatives revealed the single amino acid sequence, Ile–Ser. This showed that only one peptide (X-1A in Fig. 1) had attached to the support and it was assumed, on the basis of the specificity of the method used, that this peptide terminated in homoserine. The experiment was repeated using AEAPG as the solid support but no amino acid sequence data were obtained.

In a similar experiment, 500 nmol of X-1 peptide was digested with chymotrypsin and the resulting peptide mixture subjected to the homoserine lactone coupling procedure in the presence of 400 mg AEAPG as a solid support. Identification of PTH derivatives from the coupled material gave a single sequence of 20 residues in length (peptide X-1B of Fig. 1). The first 11 of these overlapped with residues 30 to 40 of the intact protein, as determined in the spinning cup sequencer. Sequencing through the acidic region of peptide X-1B (residue 42 to 45) led to a marked reduction in repetitive yield, indicating possible blocking of α-amino groups in this region (see Discussion). The proximity of this acidic region may also be the reason why chymotrypsin failed to cleave X-1 after the Phe^{41} as expected from the specificity of this enzyme.

Another 700 nmol of X-1 was digested with trypsin and peptide X-1C (see Fig. 1) was separated from the

Table 1. *N*- and *C*-terminal analyses of plastocyanin from *Solanum crispum* and CNBr peptides X-1, X-2 and X-3

Material	*N*-terminus	*C*-terminal residues
Plastocyanin	Ile	Asn*, Val, Thr, Gly
X-1	Ile	Hse, Ser, Ile
X-2	Pro	Hse, Gly, Ala
X-3	Val	Asn, Val, Thr, Gly

* Underlined residue = *C* terminus.

Table 2. Amino acid compositions of *Solanum plastocyanin* and its CNBr fragments

	Total protein					X-1 Fragment				X-2 Fragment				X-3 Fragment			
	24 hr	48 hr	72 hr	Av.*	Seq.	24 hr	72 hr	Av.*	Seq.	24 hr	72 hr	Av *	Seq.	24 hr	72 hr	Av.*	Seq.
Asp	12.84	12.39	12.39	12.54	12	8.71	8.82	8.76	9	2.59	2.57	2.58	2	1.12	1.45	1.29	1
Thr	4.58	4.64	4.27	4.66	5	1.25	1.36	1.31	1	2.52	2.22	2.68	3	1.18	1.00	1.28	1
Ser	8.3	7.2	7.34	9.07	9	4.63	3.92	5.03	5	3.54	3.17	3.74	4*	0	0	0	0
Hse	0	0	0	0	0	0.78	0.81	0.79	1	0.70	0.85	0.78	1	0	0	0	0
Glu	9.63	9.29	9.29	9.40	9	3.85	4.15	4.00	4	5.06	4.94	5.00	5	0	0	0	0
Pro	5.74	5.94	6.26	5.98	6	3.94	3.56	3.75	4	2.53	2.63	2.58	3	0	0	0	0
Gly	11.93	12.39	12.39	12.24	12	7.19	7.13	7.16	7	4.22	4.30	4.26	4	1.35	1.42	1.39	1
Ala	7.50	7.15	7.40	7.35	7	4.77	5.15	4.96	5	2.41	2.48	2.45	2	0	0	0	0
CySH	1.00			1.00	1	0		0	0	0.9		0.9	1	0		0	0
Val	8.98	9.57	9.29	9.28	9	5.09	5.08	5.09	5	1.41	1.45	1.45	1	2.98	3.38	3.38	3
Met	1.93	2.01	2.14	2.03	2	0	0	0	0	0	0	0	0	0	0	0	0
Ile	4.81	4.95	4.89	4.88	5	4.40	4.64	4.64	5	0	0	0	0	0	0	0	0
Leu	6.03	6.16	6.10	6.10	6	3.34	3.46	3.46	3	2.72	2.79	2.79	3	0	0	0	0
Tyr	2.86	3.25	3.25	3.12	3	0	0	0	0	2.88	3.01	2.95	3	0	0	0	0
Phe	5.81	6.04	6.04	5.96	6	4.74	4.83	4.83	5	1.25	1.33	1.29	1	0	0	0	0
His	2.12	2.51	1.83	2.17	2	0.95	0.90	0.93	1	0.79	0.85	0.82	1	0	0	0	0
Lys	4.62	4.89	5.11	4.79	5	3.04	3.19	3.12	3	1.36	1.41	1.39	1	1.00	1.17	1.09	1
Arg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total					99				57				35				7

The results are expressed as residues/mol and are based on 24, 48 and 72 hr hydrolyses of the total protein and 24 and 72 hr hydrolyses for the CNBr fragments.

* Values were taken as means of the different hydrolyses values except that value for serine and threonine were corrected for losses [6] and 72 hr hydrolysis values were taken for valine, leucine, isoleucine and phenylalanine.

Cyst(e)ine was determined as cysteic acid after performic acid oxidation prior to hydrolysis

resulting peptide mixture by ion-exchange chromatography on Aminex A5 resin. This peptide was then attached to 300 mg of APG using the DITC procedure. The *N*-terminal 23 residues were identified but since no PTH derivatives were detected by TLC subsequently, the 24th residue was assumed to be *C*-terminal lysine which remained attached to the solid support during Edman degradation.

A significant lowering of repetitive yield occurred when sequencing the acidic region between cycles 12–15 as reported above with peptide X-1B.

The amino-acid sequence of the *C*-terminal region of peptide X-1C was verified by digesting 500 nmol of X-1 with thermolysin. The mixture of peptides was fractionated using high voltage paper electrophoresis at pH 1.9. The peptide X-1D (see Fig. 1) was isolated and attached to 300 mg of APG by the DITC method, and 15 *N*-terminal amino acid residues identified. The *C*-terminal residues of X-1D was assumed to be lysine which remained attached to the support during sequencing.

There is no overlap between peptide X-1A and peptides X-1 C and D, and therefore it is possible that amino acid residues occur between positions 54 and 55 in the sequence. However, the amino acid composition of X-1 agrees with the sequence determination of the peptide and although considerations of homology should be treated with caution, the sequence of Solanum X-1 shows a high degree of similarity to the *N*-terminal amino-acid sequences of all the other higher plant plastocyanins determined so far [3].

Amino acid sequence of X-2

X-2 (400 nmol) was subjected to homoserine lactone coupling in the presence of 300 mg of APG. Subsequent solid-phase sequencing of the attached peptide enabled the determination of only 10 amino acid residues, and from the strength of the fluorescence-quenching spots on TLC relative to that of standard PTH derivatives, it was estimated that less than 20 nmol of peptides used for coupling had undergone Edman degradation. However, results from an amino acid analysis after hydrolysis of the peptide attached to the support showed that 130 nmol of X-2 had coupled to APG.

400 nmol of X-2 were reacted in the presence of 300 mg of APG using the DITC method, and a sequence of 19 amino acid residues was established from subsequent solid-phase sequencing of the attached peptide. No PTH derivatives were obtained after Glu^{76} (see Fig. 1) suggesting that the peptide had become uncoupled from the support after this sequence position. Amino acid analysis of X-2 (see Table 2) showed the presence of a single lysine residue assumed to be a sequence position 77 and which remain covalently-bound to the support after cleavage of the peptide bond between residues 77 and 78 on Edman degradation. The residual X-2, after Lys^{77}, was assumed to have been washed out of the reaction column since the DITC coupling method attaches peptides to the support via the ε-amino side-chains of lysine. The experiment was repeated and the sample from the solid-phase sequencer at cycle 20 was lyophilized. *N*-terminal analysis of this sample demonstrated *N*-terminal glycine, supporting the above suggestion.

500 nmol of X-2 was digested with trypsin and the resulting peptide mixture subjected to the homoserine lactone coupling method in the presence of 400 mg of AEAPG. Solid-phase sequencing of the attached peptide material established the amino acid sequence of

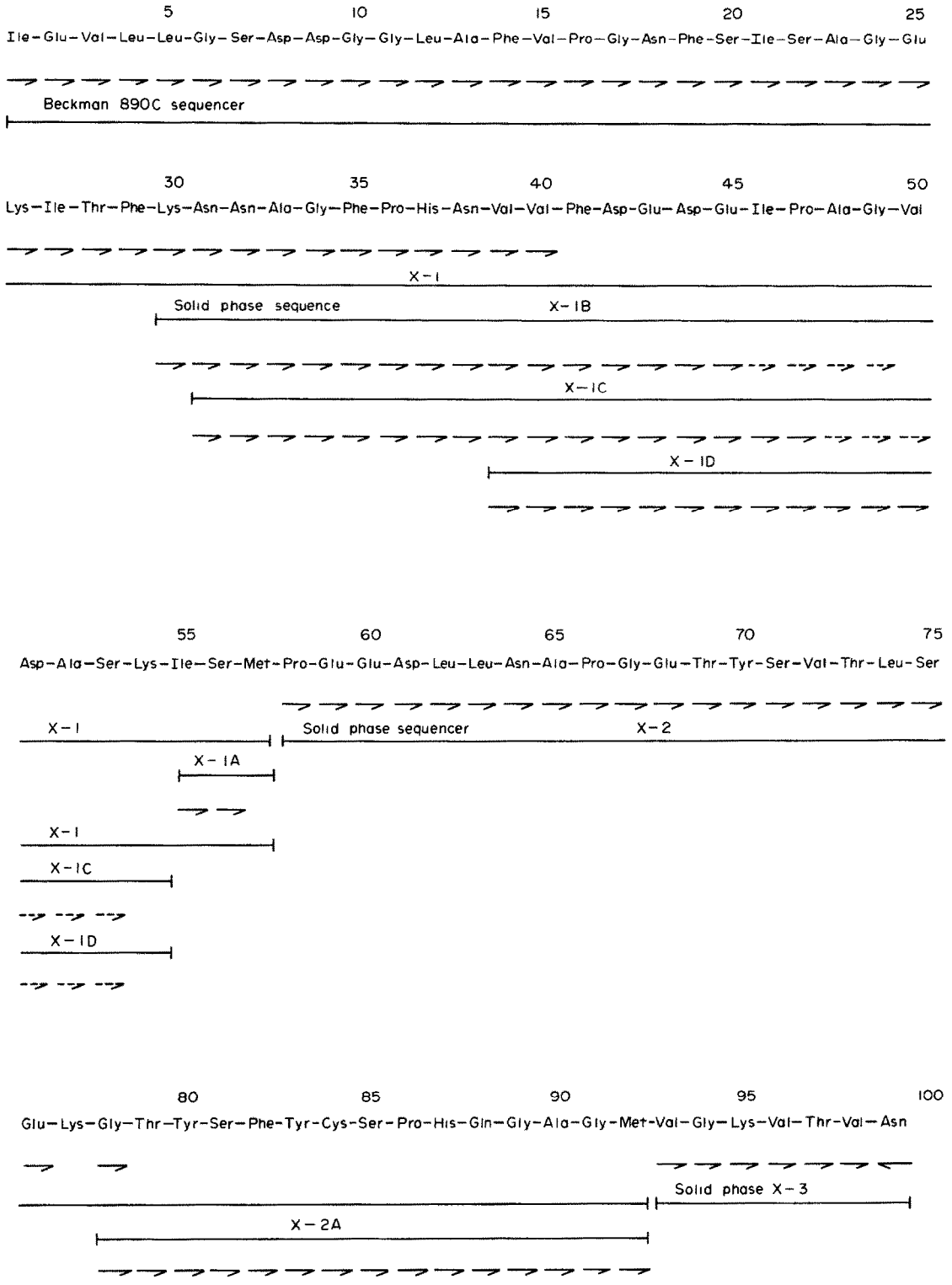

Fig. 1. Amino acid sequence of plastocyanin of *Solanum crispum*. X-1, X-2 and X-3 represent the CNBr fragments. A B C D subpeptides obtained from parent fragment by enzymic digestion. Arrows ( ) and ( ) represent automatic sequencing and carboxypeptidase digestion, respectively. Residues which were unambiguously identified are indicated with solid arrows others with dotted arrows.

residues 78 to 92 of the plastocyanin (peptide X-2A of Fig. 1). The amino acid analysis of peptide X-2 agrees with the composition obtained from the sequence determination.

Amino acid sequence X-3

100 nmol of X-3 was reacted with *t*-butyloxycarbonyl azide and subjected to the carbodiimide coupling technique in the presence of 40 mg of TETA resin. Analysis of samples from the solid-phase sequencer showed that *ca* 10 nmol of the X-3 peptide had coupled to the resin, as estimated qualitatively from the strength of the fluorescence-quenching spots relative to that of known amounts of PTH derivatives. Nevertheless it was possible to establish the amino-acid sequence of X-3 apart from the *C*-terminal asparagine (see Fig. 1), which remained covalently bound to the solid support. However, *C*-terminal analysis of X-3 (see Table 1) showed the presence of asparagine.

DISCUSSION

The strategy used to determine automatically the amino acid sequence of a plastocyanin differs markedly from that using manual Edman degradations [3]. In both cases, the protein is treated with CNBr to yield three peptides which can be purified by gel filtration, and the amino acid sequence of the short *C*-terminal peptide can be determined by manual Edman degradation without further cleavage. However, the low efficiency of the manual Edman technique requires digestion of the larger CNBr fragments with enzymes which yield numerous, relatively short peptides which must then be purified and sequenced. It is also necessary to establish the order of these short peptides in the original fragment, and a large amount of redundant sequence information must be obtained to provide the necessary overlaps to do so.

Automatic sequence analysis, on the other hand, generally allows a larger number of amino acid residues to be determined during a single degradation as compared with the manual technique. Thus the amino-acid sequence of CNBr fragments can be established from a relatively small number of large constituent peptides obtained by enzymic digestion, so reducing the amount of redundant information needed to provide overlaps. The need for purification of peptides produced by enzymic digestion of CNBr fragments is reduced by the specificity of solid-phase coupling methods such as the DITC technique (for coupling ε-amino groups of lysine) and homoserine lactone coupling (for *C*-terminal homoserine) since, in many cases, only one of the peptides in a mixture will couple, i.e. that containing the specific group.

The DITC method gave the best coupling yields of the three solid-phase attachment methods used, on average 40%. Homoserine lactone coupling gave variable attachment yields of the order of 20%. The carbodiimide method gave poor yields of about 5%. The true attachment yield is determined from amino acid analysis after an aliquot of support material with peptides attached has been subjected to hydrolytic conditions. However, as in the experiment in which X-2 was coupled to APG by the homoserine lactone method, there may be a large discrepancy between the true attachment yield and the attachment yield estimated from the yield of PTH derivatives after the first few cycles of sequencing. This 'apparent' attachment yield is a more useful quantity, since from this, the number of amino acid residues potentially identifiable during sequencing can be estimated. The discrepancy between the true and apparent attachment yields could be due to partial blocking of the peptide α-amino groups during coupling to the solid support, but might also be explained by adverse conditions for the Edman reaction in this situation. The fact that peptides were attached to APG by the homoserine lactone coupling method suggested that an intramolecular general base-catalysis is not responsible for facilitating proton removal from the attacking nitrogen atom of the derivatized glass, as has been suggested [7]. Short peptides such as X-1A could be coupled to TETA resin but not to derivatized glass.

A comparison between this sequence and that of *Solanum tuberosum* L. (potato) [4] shows ten substitutions. It is of interest to compare the *N*-terminal 36 residues of these two sequences with those of *Lycopersicon esculentum* (unpublished, B. G. Haslett). There are six differences between *Solanum crispum* and potato whereas there are only two between potato and tomato, indicating a closer similarity between the latter sequences. *Solanum crispum* is a woody member of the genus, whereas potato is herbaceous, and could represent a 'primitive' member. Boulter [8] has previously suggested that many present day families of flowering plants have had a long separate evolution history and that considerable evolutionary change has taken place within families relative to differences developed between divergence times of families.

EXPERIMENTAL

Materials. Solanum crispum leaves were collected from a single bush growing in the Durham Botanical Gardens. Chemicals used for sequencing were obtained from the Pierce Chemical Co. (Rockford, Ill. USA) or Rathburn Chemical Co. (Peebleshire, Scotland).

Preparation of plastocyanin. Plastocyanin was extracted and purified from *Solanum crispum* L. as described in [9]. The yield of purified plastocyanin was 25 mg/kg fr. wt of leaves.

Preparation of solid supports used in solid-phase sequencing. Aminopolystyrene was prepared from Bio-Beads S-X1 polystyrene, 200–400 mesh, (Bio-Rad Labs., Bromley, Kent) by the method of [10]. Triethylenetetramine (TETA) resin was prepared by the method of [11], using chloromethyl divinylbenzene–polystyrene, 1% cross-linked, 200–400 mesh (Sigma) as starting material. The chloride content of the latter was 0.9 mequiv./g resin as measured, using the Volhard method (modified by [12]), by potentiometric titration of chloride ions released after heating 150 mg of the resin in 2 ml Py at 100° for 15 min, dilution with 50 ml H_2O and acidification with conc HNO_3. Final preps contained 1–1.5 μmol NH_2/g resin. 3-Aminopropyl glass (APG) and *N*-(2-aminoethyl) 3-aminopropyl glass (AEAPG) were prepared from controlled pore glass beads (Corning CPG-10, 200–400 mesh, 75 Å mean diameter) by incubation of the beads in a 4% (v/v) soln of either 2-aminopropyl-triethoxysilane or *N*-(2-aminoethyl) 3-aminopropyltriethoxysilane at 45° for 24 hr [13]. The derivatized glass was subsequently washed in Me_2CO, then MeOH and dried *in vacuo*. Batches of APG contained 80–120 μmol NH_2/g glass, while the figure for AEAPG was 150–200 μmol.

Solid-phase coupling methods. All attachment procedures were carried out in 1.5 × 4 cm tubes with Teflon-lined caps. After attachment, peptides were either analysed immediately or stored at −20°.

Coupling with diisothiocyanate (DITC). Prior to attachment, samples were lyophilized twice in 7% aq. NEt_3 to remove any traces of ammonia. Lysine-containing peptides were attached using either APG or AEAPG by the method of [14]. Experiments with oxidized B-chain of insulin showed that best coupling was obtained if a 10- to 25-fold molar excess of DITC over peptide amino groups was used, plus a 4- to 6-fold molar excess of glass amino groups over DITC. The best results occurred when the vol. of reaction mixture was the minimum required to form a slurry with the added glass beads.

Homoserine lactone coupling. Attachment via homoserine was carried out as described in [11] using either derivatized glass or TETA-polystyrene resin.

Coupling by the carbodiimide method. *N*-terminal amino groups of peptides were first blocked using *t*-butyloxycarbonyl azide as described in [10]. Carbodiimide coupling method was performed according to [15] or [16].

Automatic sequencing. The *N*-terminal sequence of the protein was determined by using a Beckman 890 soln sequencer as described in [17]; 6 mg of protein were used after desalting on a 1 × 10 cm column of Amberlite MB-1 ion-exchange resin. Analysis by the solid-phase method was carried out with an Anachem APS 2400 sequencer using the standard single and double column programmes as supplied with the instrument. After coupling peptides to various support materials these were packed into the reaction column and subjected to automatic Edman degradation. PTH derivatives were identified mainly by TLC on pre-coated Si gel sheets impregnated with a fluorescent indicator (Merck 5554); the solvent systems used were those of [18]. When necessary, PTH derivatives were identified by GLC [17] or the parent amino acid was regenerated in HI [19], dansylated and identified by polyamide-layer chromatography [20].

Cleavage with CNBr and separation of fragments. 20 mg aliquots of plastocyanin were denatured with 6M guanidine HCl–1M Tris (pH 8.6), reduced and carboxymethylated as described in [5]. The *S*-carboxymethyl protein was cleaved by a 5-fold excess (w/w) of CNBr using the method of [21] and the resulting fragments were separated by gel-filtration on a 1.5 × 150 cm column of Biogel P-10 in 70% formic acid with a flow rate of *ca* 10 ml/hr. Elution profiles were followed by measuring *A* at 206 and 280 nm with an LKB Uvicord III.

Enzymic digestions and peptide separation. Digestion of CNBr fragments with trypsin, chymotrypsin, thermolysin and papain were carried out as described in [4]. Tryptic peptides were fractionated on a 1 × 30 cm column of Aminex A5 cation-exchange resin at 50° with a flow rate of 25 ml/hr [22]. The peptide mixture was applied in 1 ml of starting buffer and eluted with a linear gradient using 200 ml 0.1 M ammonium formate pH 3.5 and 200 ml 2 M ammonium formate pH 6.5. Peptide elution was monitored with UV *A* at 206 and 280 nm, and also by *N*-terminal residue identification. Peptides resulting from digestion with other enzymes were separated by high-voltage paper electrophoresis at pH 1.9, as described in [23]. Carboxypeptidase A was solubilized by method (i) in [24] and used for *C*-terminal analysis by digesting peptides for 3 hr at 37° in 0.2 M *N*-ethylmorpholine–acetate buffer, pH 8.5. The ratio of enzyme to substrate was 5% (w/w) and the released amino acids were identified as their dansyl derivatives [23].

N-terminal identification. The *N*-terminal amino acids of peptides were determined after dansylation and hydrolysis in 6N HCl as described in [23]. Dansyl amino acids were identified by chromatography on polyamide thin-layer sheets using the solvent systems described in [20].

Amino acid analysis. Protein and peptides were hydrolyzed in evacuated (0.05 Torr), sealed tubes. Samples of the protein were hydrolysed for 24, 48 and 72 hr and peptides for 24 and 72 hr in 6 N HCl at 105°. Cysteine was determined separately as cysteic acid after oxidation with performic acid [25]. All hydrolyses were performed in duplicate and the resultant amino acid mixtures analysed on a Locarte amino-acid analyser. To estimate the amount of peptide attached to solid supports, the support material was washed in TFA and MeOH, dried and subjected to hydrolytic procedures for 24 hr in evacuated, sealed tubes either as above or in a mixture of propionic acid–12N HCl (1:1) for 15 min at 160° [26]. The resulting amino acid mixture was analysed on a Locarte amino acid analyser, and the amount of peptide calculated from its known amino acid composition. Both methods gave similar results.

Acknowledgements—We thank Mr. J. Gilroy, Mrs R. McIntosh and Mr B. Joicey for their technical assistance. IME thanks Mr C. Brock of M.R.C. Cambridge, for demonstrating the details of solid-phase sequencing using the diisothiocyanate coupling method.

REFERENCES

1. Bridgen. J. (1974) *Trans. Biochem. Soc.* **2**, 811.
2. Bridgen, J. (1977) *Science Tools* **24**, 1.
3. Boulter, D., Haslett, B. G., Peacock, D., Ramshaw, J. A. M. and Scawen, M. D. (1977) in *International Rev. of Biochem. Plant Biochem. II* (Northcote, D. H. ed.) Vol. 13, p. 1. University Park Press, Baltimore.
4. Ramshaw, J. A. M., Scawen, M. D., Bailey, C. J. and Boulter, D. (1974) *Biochem. J.* **139**, 583.
5. Milne, P. R. and Wells, J. R. E. (1970) *J. Biol. Chem.* **245**, 1566.
6. Moore, S. and Stein, W. H. (1963) *Meth. Enzymol.* **6**, 817.
7. Bridgen, J. (1975) *FEBS Letters* **50**, 159.
8. Boulter, D. (1974) in *Chemistry and Botanical Classification* (Bendz, G. and Santesson, J. eds.) p. 211. Academic Press, New York.
9. Plesnicar, M. and Bendall, D. (1970) *Biochim. Biophys. Acta* **216**, 192.
10. Laursen, R. A. (1971) *European J. Biochem.* **20**, 89.
11. Horn, M. H. and Laursen, R. A. (1973) *FEBS Letters* **36**, 285.
12. Stewart, J. M. and Young, J. D. (1969) in *Solid-Phase Peptide Synthesis* p. 55. W. H. Freeman, San Francisco.
13. Robinson, P. J., Dunnhill, P. and Lilly, M. D. (1971) *Biochim. Biophys. Acta* **242**, 659.
14. Laursen, R. A., Horn, M. J. and Bonner, A. G. (1972) *FEBS Letters* **21**, 67.
15. Previero, A., Derancourt, J., Colletti-Previero, M. A. and Laursen, R. A. (1973) *FEBS Letters* **33**, 135.
16. Wittman-Liebold, B. and Lehmann, A. (1975) in *Solid-Phase Methods in Protein Sequence Analysis* (Laursen, R. A. ed.) p. 81. Pierce Chemical Co., Rockford.
17. Haslett, B. G. and Boulter, D. (1976) *Biochem. J.* **153**, 33.
18. Jeppson, J. O. and Sjoquist, J. (1967) *Anal. Biochem.* **18**, 264.
19. Inglis, A. S., Nicholls, P. W. and Rosburgh, C. M. (1971) *Australian J. Biol. Sci.* **24**, 1247.
20. Ramshaw, J. A. M., Thompson, E. W. and Boulter, D. (1970) *Biochem. J.* **119**, 535.
21. Steers, E., Craven, G. R., Anfinson, C. B. and Bethune, J. L. (1965) *J. Biol. Chem.* **240**, 2478.
22. Hartley, B., Wheeler, A. M. and Ellis, R. J. (1975) *J. Mol. Biol.* **91**, 67.
23. Thompson, E. W., Laycock, M. V., Ramshaw, J. A. M. and Boulter, D. (1970) *Biochem. J.* **117**, 183.
24. Ambler, R. P. (1972) *Meth. Enzymol.* **25**, 143.
25. Schram, E., Moore, S. and Bigwood, E. J. (1954) *Biochem. J.* **57**, 33.
26. Westall, F. and Hesser, H. (1974) *Anal. Biochem.* **61**, 610.