*Phytochemistry*, 1978, Vol 17, pp 901–905. © Pergamon Press Ltd Printed in England. 0031-9422/78/0501-0901 $02.00/0

# THE AMINO ACID SEQUENCES OF PLASTOCYANIN FROM *MERCURIALIS PERENNIS* AND *CAPSELLA BURSA-PASTORIS*

MICHAEL D. SCAWEN*, JOHN A. M. RAMSHAW†, RICHARD H. BROWN‡ and DONALD BOULTER§
Department of Botany, University of Durham, Durham DH1 3LE, U.K.

(*Received* 30 *November* 1977)

**Key Word Index**—*Mercurialis perennis*; Euphorbiaceae; *Capsella bursa-pastoris*; Cruciferae; amino acid sequence; plastocyanin; CBNr fragments.

**Abstract**—The amino acid sequences of the plastocyanins from *Mercurialis perennis* and *Capsella bursa-pastoris* have been determined. The amides at positions 64 and 68 in the *Mercurialis* sequence were positioned by 'homology' Both proteins are single polypeptide chains of 99 residues and are closely related to other higher plant plastocyanins.

## INTRODUCTION

The purification of the plastocyanin from *Mercurialis perennis* and some of its properties have been reported previously [1]. The aim of the present investigation was to determine the complete amino acid sequences of two higher plant plastocyanins in order to generate more data for the construction of an affinity tree relating the evolution of higher plants [2]. In addition, comparative sequence data of plastocyanin is of value in integrating various spectroscopic studies aimed at elucidating the nature of the copper binding site [3].

* Present address: Microbiological Research Establishment, Porton, Salisbury; † Present address: Museum of Comparative Zoology, Harvard University, Cambridge, Mass, U.S.A.; ‡ Present address: Occupational Medicine and Hygiene Labs., Cricklewood, London; § To whom reprint requests should be sent.

## RESULTS

### Mercurialis *plastocyanin*

Fig. 1 shows the amino acid sequence of *Mercurialis*

Table 1. Amino acid composition of *Mercurialis* plastocyanin and its CNBr fragments. The results are expressed as residues/mol. For the total protein the values taken were the average of 24 and 72 hr hydrolysates except that values for threonine and serine were corrected for losses [13] and maximal values (72 hr) were taken for valine and isoleucine. ½-cysteine was determined as *S*-carboxymethyl cysteine

| | Total protein | | | | X-1 fragment | | X-2 fragment | | X-3 fragment | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 24 hr | 72 hr | Value taken | Sequence | 24 hr | Sequence | 24 hr | Sequence | 24 hr | Sequence |
| Asp | 15,8 | 15.3 | 15.8 | 8 | 11.3 | 6 | 3.1 | 2 | 0.9 | 0 |
| Asn | | | | 7 | | 5 | | 1 | | 1 |
| Thr | 3.6 | 2.6 | 4.2 | 4 | 0.2 | 0 | 1.7 | 2 | 0.6 | 1 |
| Ser | 7.1 | 6.3 | 7.6 | 8 | 4.7 | 5 | 3.1 | 3 | 0 | 0 |
| Glu | 7.0 | 7.5 | 7.3 | 6 | 3.1 | 3 | 4.2 | 3 | 0 | 0 |
| Gln | | | | 1 | | 0 | | 1 | | 0 |
| Pro | 5.1 | 4.8 | 4.9 | 5 | 3.1 | 3 | 2,2 | 2 | 0 | 0 |
| Gly | 9.9 | 10.4 | 10.1 | 10 | 6.0 | 6 | 3.8 | 4 | 1.0 | 1 |
| Ala | 9.1 | 8.9 | 9.0 | 9 | 5.1 | 5 | 3.9 | 4 | 0 | 0 |
| Cys | * | * | | 1 | 0 | 0 | 0.7 | 1 | 0 | 0 |
| Val | 8.7 | 9.9 | 9.9 | 10 | 5.5 | 6 | 1.0 | 1 | 2.7 | 3 |
| Met | 2.0 | 1.8 | 1.9 | 2 | | | | | | |
| Ile | 3.5 | 3.9 | 3.9 | 4 | 3.7 | 4 | 0.2 | 0 | 0 | 0 |
| Leu | 6.8 | 6.9 | 6.9 | 7 | 4.3 | 4 | 2.5 | 3 | 0 | 0 |
| Tyr | 2.5 | 2.4 | 2.5 | 3 | 0 | 0 | 2.5 | 3 | 0 | 0 |
| Phe | 5.3 | 5.8 | 5.6 | 6 | 4.8 | 5 | 1.1 | 1 | 0 | 0 |
| His | 2.1 | 2.2 | 2.2 | 2 | 1.2 | 1 | 1.1 | 1 | 0 | 0 |
| Lys | 5.1 | 5.3 | 5.2 | 5 | 2.7 | 3 | 0.9 | 1 | 1.1 | 1 |
| Arg | 0.2 | 0.3 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Hse | | | | | 0.7 | 1 | 0.6 | 1 | | |
| Totals | | | | 99 | | 57 | | 35 | | 7 |

* *S*-Carboxymethyl cysteine was not resolved from aspartic acid.

plastocyanin and summarizes the evidence for it. The amino acid composition calculated from the sequence is in good agreement with that found by analysis of the total protein (Table 1).

CNBr cleaved the *S*-carboxymethylcysteinyl protein into three fragments which were separated and ordered as described previously [4]. The amino acid composition of the purified fragments are shown in Table 1, and there is generally good agreement between the values obtained by analysis and those calculated from the sequence. The recoveries of homoserine were consistently low.

The automated sequence analysis of the total protein enabled the first 32 residues in the sequence to be identified (Fig. 1). This result provided an overlap with peptides X1H1 to X1H6 obtained from a thermolysin digest of fragment X-1. The single papain peptide isolated, X1P [5] provided an overlap with thermolysin peptides X1H7 and X1H8, leaving only the existence of a peptide bond between peptides X1H6 and X1H7 unconfirmed in the sequence of fragment X1.

The complete sequence of fragment X-2 was established from the dansyl-phenylisothiocyanate analysis of thermolytic and tryptic peptides, which gave complete overlap (Fig. 1). Evidence from the peptides for Met-92 was not strong, although the CNBr cleavage at this point confirmed the presence of a methionine residue. The smallest fragment, X3 was further purified by high voltage paper electrophoresis at pH 6.5 before its sequence was established by dansyl-phenylisothiocyanate analysis.

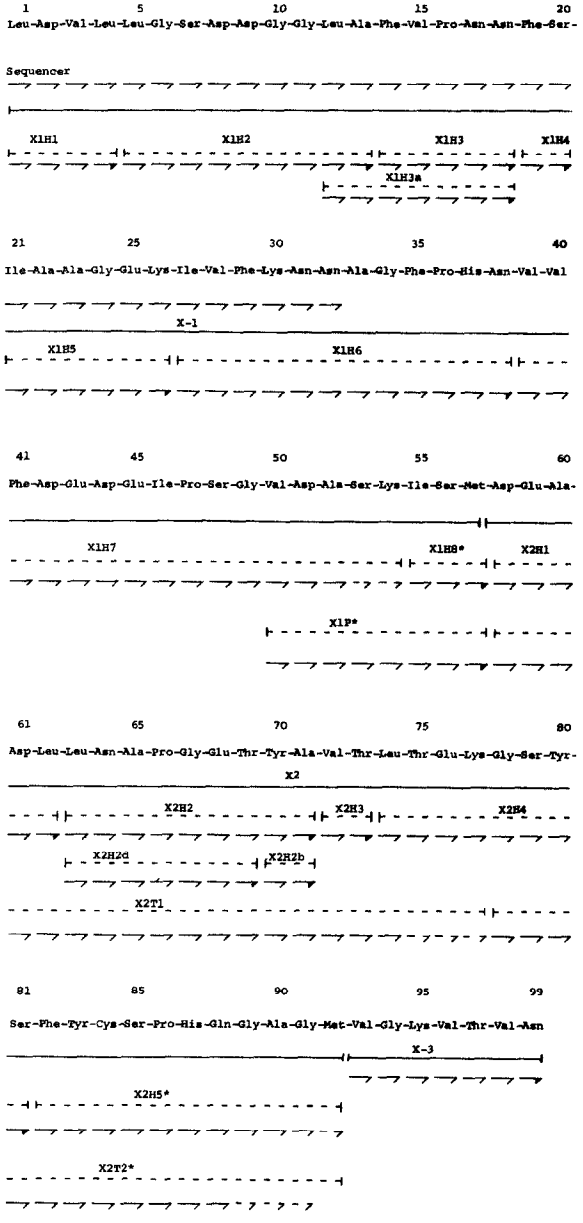


Fig. 1. The amino aicd sequence of *Mercurialis perennis* plastocyanin. Full lines under peptides indicate quantitative amino acid analyses and broken lines qualitative analyses. Solid arrows (——→) indicate that the residue was identified unambiguously; dashed arrows (— — →) indicate that the residue could not be unambiguously identified. Arrows (——→) indicate that the *C*-terminal residue of a peptide was identified as the free amino acid. In peptides marked * methionine was identified as homoserine.

Table 2. Properties of peptides isolated from *Mercurialis* plastocyanin. Electrophoretic mobilites of peptides at pH 6.5 were measured from the neutral amino acids relative to aspartic acid and at pH 1.9 were measured from 1-dimethylaminonaphthalene-5-sulphonic acid relative to Dns-arginine.

| | Mobility at | | |
|---|---|---|---|
| Peptide | pH 6.5 | pH 1.9 | BAWP |
| X1H1 | −0.44 | 0.41 | - |
| X1H2 | −0.62 | 0.31 | — |
| X1H3 | 0 | 0.35 | 0.31 |
| X1H3a | 0 | 0.30 | — |
| X1H4 | 0 | 0.60 | 0.47 |
| X1H5 | 0 | 0.75 | — |
| X1H6 | 0.40 | 0.72 | — |
| X1H7 | −0.62 | 0.47 | — |
| X1H8 | 0 | 0.50 | 0.34 |
| X1P | 0/0.34 | — | — |
| X2H1 | −0.99 | — | — |
| X2H2 | −0.27 | 0.27 | — |
| X2H2a | −0.33 | 0.33 | — |
| X2H2b | 0 | 0.60 | 0.58 |
| X2H3 | 0 | 0.60 | 0.48 |
| X2H4 | 0 | 0.60 | 0.29 |
| X2H5 | 0.08 | — | — |
| X2T1 | −0.65 | 0.34 | — |
| X2T2 | 0.11 | — | — |
| X3 | 0.32 | 0.74 | — |

BAWP, $R_f$ of peptide after descending PC in *n*-BuOH–HOAc–$H_2O$–Py (15:3:12:10). — = n.d.

The electrophoretic mobilities and $R_f$ values on paper chromatography of the peptides isolated are given in Table 2. All of the acid/amide residues in the sequence except for Asn 64 and Glu 68, have been assigned, from a consideration of the electrophoretic mobilities of the peptides at pH 6.5 [6] or from the automated sequencer results. Asn 64 and Glu 68 were assigned by homology with the other higher plant plastocyanins from which amino aicd sequence data are available [3].

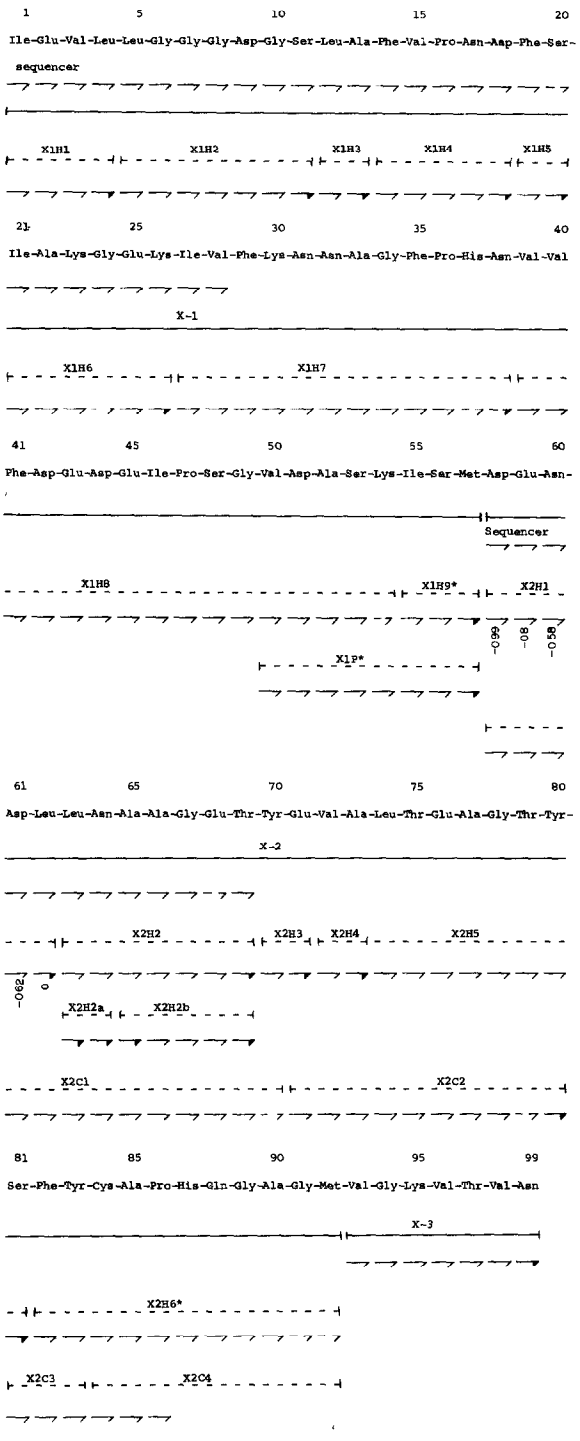


Fig. 2. Amino acid sequence of *Capsella bursa-pastoris* plastocyanin. Full lines under peptides indicate quantitative amino acid analyses and broken lines qualitative analyses. Solid arrows (——→) indicate that the residue was identified unambiguously; dashed arrows (– – →) indicate that the residue could not be unambiguously identified. Arrows (——▸) indicate that the *C*-terminal residue of a peptide was identified as the free amino acid. In peptides marked * methionine was identified as homoserine. The figures below peptide X2H1 show the electrophoretic mobility of the peptide at pH 6.5 after successive steps of phenylisothiocyanate degradation [8].

Capsella *plastocyanin*

Fig 2 shows the amino acid sequence of *Capsella* plastocyanin and summarizes the evidence for it. The amino acid composition calculated from the sequence is in good agreement with that found by analysis of the total protein (Table 3), except that the values found by analysis for glutamic acid and lysine were consistently high.

CNBr cleaved the *S*-carboxymethylcysteinyl protein into three fragments (Fig. 2) which were separated and ordered as described previously [4]. The amino acid compositions of the purified fragments are shown in Table 3. The values determined by the analysis are generally in good agreement with those calculated from the sequences of the three fragments. The recoveries of homoserine and tyrosine were consistently low.

The automated sequence analysis of the total protein enabled the first 28 residues in the sequence to be identified (Fig. 2). This result provided an overlap with peptides X1H1 to X1H7 obtained from a thermolysin digest of fragment X1. The single papain peptide isolated, X1P [5] provided an overlap with thermolysin peptides X1H8 and X1H9 leaving only the peptide bond between X1H7 and X1H8 unconfirmed.

Automated sequence analysis of fragment X2 using a peptide programme [7] enabled the first 12 residues in the fragment to be identified. The complete sequence of the fragment was established by the dansyl-phenylisothiocyanate analysis of thermolytic and chymotryptic peptides (Fig. 2). The evidence from the peptides for Gly-91 and Met-92 was weak, although the CNBr cleavage confirms the presence of Met-92.

The smallest fragment, X3 was further purified by high-voltage paper electrophoresis at pH 6.5 before its sequence was established by dansyl-phenylisothiocyanate analysis.

The electrophoretic mobilities and where applicable $R_f$ values on paper chromatography of the isolated peptides are given in Table 4. All of the amide residues in the sequence have been assigned, either from a consideration of the electrophoretic mobilities of the peptides at pH 6.5 [6], or from the automatic sequencer results. The identification of Asn-60 was confirmed by determining the electrophoretic mobility of peptide X2H1 after each degradation step [8]. This residue was also clearly identified as asparagine during the automated sequence analysis of CNBr fragment X-2.

## DISCUSSION

In both plastocyanin molecules the amino acid sequences of the three CNBr fragments isolated account completely for the composition of the total protein. Since no other CNBr fragments were isolated, it is considered that the sequences of the three fragments provide sufficient evidence to establish the complete sequences of the two molecules, even though peptides which overlapped the two methionine residues were not isolated. Amino acid analyses of individual peptides were not obtained, but the generally good agreement obtained between sequence and composition for the CNBr fragments means that there is still reasonable evidence for the proposed amino acid sequences.

In both sequences the evidence from the peptides for Gly-91 and Met-92 is not strong, as these residues were

Table 3. Amino acid composition of *Capsella* plastocyanin and its CNBr fragments. The results are expresses as residues/mol. For the total protein the values taken were the averages of 24 and 72 hr hydrolysates except that values for threonine and serine were corrected for losses [13] and maximal values (72 hr) were taken for valine and isoleucine

| | Total protein | | | | X-1 fragment | | X-2 fragment | | X-3 fragment | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 24 hr | 72 hr | Value taken | Sequence | 24 hr | Sequence | 24 hr | Sequence | 24 hr | Sequence |
| Asp | 12.8 | 12.8 | 12.8 | 6 | 9.4 | 5 | 3.7 | 1 | 1.0 | 0 |
| Asn | | | | 7 | | 4 | | 2 | | 1 |
| Thr | 3.6 | 3.1 | 3.9 | 4 | 0.4 | 0 | 3.0 | 3 | 1.0 | 1 |
| Ser | 5.7 | 4.5 | 6.4 | 6 | 3.8 | 5 | 1.1 | 1 | 0 | 0 |
| Glu | 11.2 | 11.0 | 11 1 | 9 | 4.5 | 4 | 6.0 | 5 | 0 | 0 |
| Gln | | | | 1 | | | | 1 | | |
| Pro | 3.7 | 3.7 | 3.7 | 4 | 2.9 | 3 | 1.1 | 1 | 0 | 0 |
| Gly | 11.8 | 11.9 | 11.9 | 12 | 7.0 | 7 | 4.3 | 4 | 1.0 | 1 |
| Ala | 9.8 | 9.9 | 9.9 | 10 | 3.8 | 4 | 5.6 | 6 | 0 | 0 |
| Cys | 0.9 | | 0.9 | 1 | 0 | 0 | * | 1 | 0 | 0 |
| Val | 9.0 | 9.9 | 9.9 | 10 | 5.5 | 6 | 1.1 | 1 | 2.8 | 3 |
| Met | 1.8 | 1.9 | 1.9 | 2 | | | | | | |
| Ile | 4.8 | 4.9 | 4.9 | 5 | 4.9 | 5 | 0.3 | 0 | 0 | 0 |
| Leu | 6.4 | 6.4 | 6.4 | 6 | 3.2 | 3 | 2.9 | 3 | 0 | 0 |
| Tyr | 2.8 | 2.6 | 2.7 | 3 | 0 | 0 | 2.3 | 3 | 0 | 0 |
| Phe | 6.0 | 6.0 | 6.0 | 6 | 5 1 | 5 | 1.0 | 1 | 0 | 0 |
| His | 1.8 | 1.8 | 1.8 | 2 | 1.0 | 1 | 0.9 | 1 | 0 | 0 |
| Lys | 5.4 | 5.6 | 5.5 | 5 | 4.5 | 4 | 0.4 | 0 | 1.3 | 1 |
| Arg | 0.3 | 0.3 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Hse | | | | | 0.3 | 1 | 0.4 | 1 | | |
| Total | | | | 99 | | 57 | | 35 | | 7 |

* *S*-Carboxymethyl cysteine was not resolved from aspartic acid.

Table 4. Properties of peptides isolated from *Capsella* plastocyanin. Electrophoretic mobilities of peptides at pH 6.5 were measured from the neutral amino acids relative to aspartic acid and at pH 1.9 from 1-dimethylaminonaphthalene-5-sulphonic acid relative to Dns-arginine

| | Mobility at | | |
|---|---|---|---|
| Peptide | pH 6 5 | pH 1.9 | BAWP |
| X1H1 | −0.40 | 0.44 | 0.86 |
| X1H2 | −0.32 | 0.38 | 0.24 |
| X1H3 | 0 | 0.78 | 0.75 |
| X1H4 | 0.32 | 0.38 | 0.40 |
| X1H5 | 0 | 0 69 | 0.51 |
| X1H6 | 0.24 | 1.15 | — |
| X1H7 | 0.24 | 0.68 | 0.29 |
| X1H8 | −0.63 | 0.43 | — |
| X1H9 | 0/0.54 | 0.57 | 0.37 |
| X1P | 0/0.32 | | — |
| X2H1 | −0.99 | 0.36 | — |
| X2H2 | 0 33 | 0.33 | — |
| X2H2a | 0 | 0.61 | 0.38 |
| X2H2b | −0.35 | 0.46 | 0.20 |
| X2H3 | −0.35 | 0.51 | 0.45 |
| X2H4 | 0 | 0.78 | 0.62 |
| X2H5 | −0.25 | 0.28 | 0.39 |
| X2H6 | 0.1 | 0.52 | 0.18 |
| X2C1 | −0.8 | — | — |
| X2C2 | −0.51 | — | — |
| X2C3 | 0 | 0.47 | — |
| X2C4 | 0.17 | 0.62 | — |
| X3 | 0.36 | — | — |

BAWP, $R_f$ after descending PC in *n*-BuOH–HOAc–$H_2O$–Py (15:3:12:10) — = n.d.

only seen in one peptide in each case and the identifications could not be completely unambiguous because of carry-over. However the amino acid composition of the fragments X2 and the CNBr cleavage provide additional evidence for the presence of methionine.

The amino acid sequences of these plastocyanins are similar to the other higher plant plastocyanin sequences which have been published [see 3]. Of the two reported here the *Capsella* sequence is the most unusual in that residue 21, which had previously been considered an invariant valine is isoleucine. In both the *Mercurialis* and *Capsella* sequences, residue 60 is unusual in that it is alanine and asparagine respectively, instead of an acidic residue as in all the other higher plant plastocyanin sequences studied so far [3]. However, in all cases the net charge in this region of the sequence, residues 58–62 is maintained at −3.

## EXPERIMENTAL

*Materials.* Leaves of Dog's Mercury (*Mercurialis perennis* and Shepherd's Purse (*Capsella bursa-pastoris*) were collected locally. Other materials were as described previously [5, 9]. Plastocyanin was prepared from both species as described in ref. [1]

*Sequence analysis.* Plastocyanin from both species was denatured in 6 M guanidine chloride, reduced and *S*-carboxymethylated [4, 10]. The *S*-carboxymethylcysteinyl plastocyanins were cleaved by CNBr in 70% formic acid [11] and the fragments separated by gel filtration on Sephadex G50 equilibrated in 70% formic acid [4] The purified CNBr fragments were digested with trypsin or chymotrypsin, thermolysin and papain, [4, 7]. The resulting peptides were purified by high voltage paper electrophoresis (HVE) and PC [5] The amino acid sequences of the purified peptides were determined by using the dansyl-phenylisothiocyanate method [12] as described

in ref. [9]. Automated sequencer analysis was carried out on a Beckman model 890C automatic protein sequencer [7]. Samples of *S*-carboxymethylcysteinyl plastocyanin for amino acid analysis were hydrolysed in constant boiling HCl at 105° for 24 and 72 hr. Samples of the purified CNBr fragments were hydrolysed for 24 hr only. The analyses were carried out on a Locarte amino acid analyser. Quantitative amino acid analyses of the purified peptides were not obtained, but every peptide examined was analysed qualitatively by dansylation following hydrolysis [9].

*Quantities of material used.* For CNBr cleavage, 25 mg of *S*-carboxymethyl cysteinyl protein was used. 0.75–1 μmol of the purified fragments was used for each enzymic digest 0.3 μmol of *S*-carboxymethylcysteinyl plastocyanins and of fragment X-2 in the case of *Capsella* plastocyanin, were used for the automated sequencer analysis.

*Nomenclature.* The CNBr fragments are labelled X and are numbered in Figs 1 and 2 on the basis of their position in the complete sequence. Peptides obtained by enzymic digestion of the purified fragments are numbered according to their position in the parent fragment. The following abbreviations are employed: T. trypsin; C, chymotrypsin; H, thermolysin; P, papain.

*Acknowledgements*—We thank John Gilroy and Elizabeth Jones for technical assistance and the Nuffield Foundation for financial support.

## REFERENCES

1. Ramshaw, J. A. M., Brown, R. H., Scawen, M. D. and Boulter, D. (1973) *Biochim. Biophys. Acta* **303**, 269.
2. Boulter, D. (1973) *Nobel Symp.* **25**, 211.
3. Boulter, D., Haslett, B. G., Peacock, D., Ramshaw, J. A. M. and Scawen, M. D. (1977) *International Review of Biochemistry, Plant Biochemistry II* (Northcote, D. H. ed.) Vol. 13, pp. 1–40. University Park Press Baltimore.
4. Scawen, M. D., Ramshaw, J. A. M. and Boulter, D. (1975) *Biochem. J.* **147**, 343.
5. Scawen, M. D., Ramshaw, J. A. M., Brown, R. H. and Boulter, D. (1974) *European J. Biochem.* **44**, 299.
6. Offord, R. E. (1966) *Nature* **211**, 591.
7. Scawen, M. D. and Boulter, D. (1974) *Biochem. J.* **143**, 257.
8. Gray, W. R. (1967) *Meth. Enzymol.* **11**, 469.
9. Thompson, E. W., Laycock, M. V., Ramshaw, J. A. M. and Boulter, D. (1970) *Biochem. J.* **177**, 183.
10. Milne, P. R. and Wells, J. R. E. (1970) *J. Biol. Chem.* **245**, 1566.
11. Steers, J. E., Craven, G. R., Anfinisen, C. B. and Bethune J. L. (1965) *J. Biol. Chem.* **240**, 2478.
12. Gray, W. R. and Hartley, B. S. (1963) *Biochem. J.* **89**, 378.
13. Moore, S. and Stein, W. H. (1963) *Meth. Enzymol.* **6**, 817.