Phytochemistry, 1977, Vol. 16, pp 363–365, Pergamon Press, Printed in England.

N-TERMINAL AMINO ACID SEQUENCES OF PLASTOCYANINS FROM VARIOUS MEMBERS OF THE COMPOSITAE

B. G. HASLETT, T. GLEAVES and D. BOULTER
Department of Botany, University of Durham, Durham DH1 3LE, England

(*Revised received* 22 *August* 1976)

Key Word Index—Compositae; plastocyanin; amino acid sequence; automatic sequencing.

Following the use of amino acid sequence data from animal proteins to construct phylogenies [1], higher plant phylogeny has also been studied by comparison of sequence data obtained from cytochrome *c* [2]. However, technical difficulties in extracting and purifying sufficient protein for sequencing purposes, coupled with the low rate of evolution of cytochrome *c* limits the usefulness of this protein for such studies.

Plastocyanin is a blue Type I copper protein first discovered in *Chlorella ellipsoidea* by Katoh [3] and subsequently found in chloroplasts of numerous higher plants, from which it can be obtained in relatively large yields [4].

Since the rate of evolution of plastocyanin is greater than that of cytochrome *c* [5], the differences between plastocyanin sequences is reflected in the differences found in the first third (*N*-terminal sequence) of the total sequence. Thus, the rank correlation between the number of differences in the first forty residues and the number of differences in residues 41–*C*-terminus for 11 species, whose complete sequence is known [5], was 0.744, (i.e. significant at the 0.1 % level). Since *N*-terminal sequence data can be collected extremely rapidly now that automated sequencing methods are available [6], our strategy has changed from one of comparing data from full sequences of plastocyanins of a limited number of plant species to comparing *N*-terminal sequences from a much larger number of species.

The *N*-terminal amino acid sequences of nine plastocyanins obtained from members of the Compositae are given in Table 2, together with that obtained for lettuce plastocyanin by Ramshaw *et al.* [7]. The sequences show a high degree of structural homology.

In some species two amino acid residues occur in some positions (nos. 1, 2, 11, 15, 18, 22), Table 1 shows the relative yields of each residue in these polymorphic positions. When considering the ratio of PTH amino acids found at a polymorphic position, account must be taken of the recoveries of different derivatives; generally, Leu, Ile, Phe, Ala and Val give fairly reproducible recoveries, whereas, Glu and Asp are often poorly extracted from the sequencer cup by butyl chloride because of their greater polarity, while the yield of Gly is often low because of its requirement for a longer conversion time. The yield of Pro is also non-reproducible, since cleavage at this residue is variable during Edman degradation.

It is not known whether polymorphism is due to the presence of more than one plastocyanin in a single plant, or to the occurrence of more than one plastocyanin type within a population. The former situation does occur, since polymorphism was found in the *N*-terminal sequence of plastocyanin from protein extracted from the leaves of a single *Malva* bush (Haslett, B. G., unpublished).

EXPERIMENTAL

Materials. Plastocyanin was extracted from nine species of the Compositae by the method of ref. [8]. After fractionation of the homogenate by Me_2CO-precipitation, the plastocyanin-containing fraction was redissolved in 25 mM Tris–HCl buffer

Table 1. Yields of PTH derivatives at sequence positions where polymorphism occurred

Species	Position in sequence	Yields of PTH (nmol)	Ratio
Helianthus annuus	1	Phe = 95.5; Ile = 57.0	1.68
	2	Asp = 67.5; Glu = 45.5	1.48
	22	Ala = 26.4; Pro = 18.3	1.44
Guizotia abyssinica	1	Phe = 315.0; Leu = 235.0	1.34
	2	Asp = 196.0; Gln = 42.0	4.67
Senecio vulgaris	11	Ala = 36.0; Gly = 26.0	1.39
	15	Val = 29.0; Gln = 16.2	1.79
		n.d.	
Cirsium vulgare	1	Val = 115.0; Ile = 74.0	1.56
		n.d.	

PTH yields quantified by GLC

Table 2. *N*-terminal amino acid sequences

	5	10	15
Helianthus annuus:	Ile/Phe–Glu/Asp–Val–Leu–Leu–Gly–Asp–Asn–Asp–Gly–Gly–Leu–Ala–Phe–Glu–Pro–Ser–		
Guizotia abyssinica:	Leu/Phe–Glu/Asp–Val–Leu–Leu–Gly–Asp–Asn–Asp–Gly–Ala–Leu–Ala–Phe–Glu–Pro–Ser–		
Senecio vulgaris:	Ile–Glu–Val–Leu–Leu–Gly–Asp–Asn–Asp–Gly–Gly/Ala–Leu–Ala–Phe–Val/Glu–Pro–Ser–		
Chrysanthemum vulgare:	Ile–Asp–Val–Leu–Leu–Gly–Ala–Asn–Asp–Gly–Gly–Leu–Ala–Phe–Glu–Pro–Ala–		
Cirsium vulgare:	Ile/Val–Glu–Val–Leu–Leu–Gly–Ala–Ser–Asp–Gly–Gly–Leu–Val–Phe–Glu–Pro–Ser–		
Centaurea niger:	Val–Asp–Val–Leu–Leu–Gly–Gly–Asp–Asp–Gly–Gly–Leu–Val–Phe–Glu–Pro–Ser–		
Tragopogon porrifolius:	Val–Glu–Val–Leu–Leu–Gly–Asp–Asn–Asp–Gly–Ser–Leu–Val–Phe–Glu Pro–Ser–		
*Lactuca sativa**:	Ala–Glu–Val–Leu–Leu–Gly–Ser–Ser–Asp–Gly–Gly–Leu–Val–Phe–Glu–Pro–Ser–		
Hieracium vulgatum:	Val–Glu–Val–Leu–Leu–Gly–Asp–Asn–Asp–Gly–Gly–Leu–Val–Phe–Glu–Pro–Ser–		
Taraxacum officinalis:	Val–Glu–Val–Leu–Leu–Gly–Asp–Asn–Asp–Gly–Gly–Leu–Val–Phe–Glu–Pro–Ser–		

(pH 7.4) containing 0.5% Tween 80; the detergent was added to lessen the formation of H_2O-insoluble protein–tannin complexes. After removal of insoluble material by centrifugation, plastocyanin was purified by chromatography on DEAE cellulose, DEAE-Sephadex and Sephadex G75, as described in ref [4]. Preparations with an *A* ratio A_{280}/A_{597} of 3.0 were used for *N*-terminal sequence determinations. Yields of pure plastocyanin together with the total amounts extracted are given in Table 3.

Automated sequencing. Prior to degradation, proteins were desalted by passage through a column (1 × 10 cm) of Amberlite MB-1 ion-exchange resin and freeze-dried. Edman degradation was carried out with a Beckman 890C automatic sequencer: 5–7 mg protein (450–650 nmol) was applied to the sequencer in 0.4 ml of 75% HCO_2H. Before degradation, the dried film was treated with HFBA followed by BuCl washing in order to improve stability of the film [9]. Single degradations were performed on each protein using the 'fast protein' programme as supplied with the instrument (Beckman prog. no. 072172C). Resultant thiazalinone derivatives were converted to the more stable phenylthiohydantoins by heating in N HCl at 80° for 10 min.

Identification of phenylthiohydantoin derivatives of amino acids. Phenylthiohydantoins were identified as described previously [10]. After preliminary TLC, identification was confirmed either by GLC or by polyamide-TLC of the dansylated parent amino

Table 3. Yields of purified plastocyanin from species of the Compositae

Species	Fresh weight of leaf material extracted (kg)	Yield (mg pure plastocyanin per kg leaves)
Hieracium vulgatum	5	1.2
Tragopogon porrifolius	4	8.5
Cirsium vulgare	8	5.9
Helianthus annuus	5	6.9
Taraxacum officinalis	6	3.5
Senecio vulgaris	6.5	5.7
Guizotia abyssinica	5	7.7
Chrysanthemum vulgare	8	2.6
Centaurea niger	10	2.4

For details of extraction and purification methods see text.

of plastocyanins from the Compositae

20	25	30	35	40
Thr–Phe–Ser–Val–$\frac{\text{Ala}}{\text{Pro}}$–Ala–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				
Thr–Phe–Ser–Val–Pro–Ser–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				
$\frac{\text{Asn}}{\text{Thr}}$–Phe–Ser–Val–Ala–Ala–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				
Thr–Phe–Ser–Val–Pro–Ala–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				
$\frac{\text{Asn}}{\text{Thr}}$–Phe–Thr–Val–Ala–Ser–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ala–Gly–Phe–Pro–His–Asn–Val–Val				
Thr–Phe–Ser–Val–Ala–Ser–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ala–Gly–Phe–Pro–His– † – † – †				
Thr–Phe–Ser–Val–Ala–Ser–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				
Thr–Phe–Ser–Val–Ala–Ser–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ala–Gly–Phe–Pro–His–Asn–Val–Val				
Thr–Leu–Ser–Val–Ala–Ser–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				
Thr–Phe–Ser–Val–Pro–Ala–Gly–Glu–Lys–Ile–Val–Phe–Lys–Asn–Asn–Ser–Gly–Phe–Pro–His–Asn–Val–Val				

* From ref. [7] † Not identified.

acid after hydrolysis with HI. Criteria used for identification of the derivatives were those described previously [10].

REFERENCES

1. Dayhoff, M. O. (1969) *Atlas of Protein Sequence and Structure* 4. National Biomedical Research Foundation, Maryland.
2. Boulter, D., Ramshaw, J. A. M., Thompson, E. W., Richardson, M. and Brown, R. H. (1972) *Proc. R. Soc. ondon Ser. B*, **181**, 441.
3. Katoh, S. (1960) *Nature* **186**, 533.
4. Ramshaw, J. A. M., Brown, R. H., Scawen, M. D. and Boulter, D. (1973) *Biochim. Biophys. Acta* **303**, 269.
5. Boulter, D., Haslett, B. G., Peacock, D., Ramshaw, J. A. M. and Scawen, M. D. (1976) *International Review of Biochemistry—Series II. Plant Biochemistry* (ed. D. H. Northcote). Medical and Technical Publishing, Lancaster. In Press.
6. Edman, P. and Begg, G. (1967) *European J. Biochem.* **1**, 80.
7. Ramshaw, J. A. M., Scawen, M. D., Jones, E. A., Brown, R. H. and Boulter, D. (1976) *Phytochemistry*, **15**, 1199.
8. Plesnicar, M. and Bendall, D..S. (1970) *Biochim. Biophys. Acta* **216**, 192.
9. Bridgen, J. and Secher, D. S. (1973) *FEBS Letters* **29**, 55.
10. Haslett, B. G. and Boulter, D. (1976) *Biochem. J.* **153**, 33.