Phytochemistry, 1976, Vol. 15, pp. 1199–1202. Pergamon Press Printed in England.

THE AMINO ACID SEQUENCE OF PLASTOCYANIN FROM *LACTUCA SATIVA* (LETTUCE)

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

(*Received* 23 *January* 1976)

Key Word Index—*Lactuca sativa*; Compositae; plastocyanin; amino acid sequence; lettuce.

Abstract—The amino acid sequence of plastocyanin from lettuce (*Lactuca sativa* L.) was determined by using a Beckman 890C automatic sequencer and by dansyl-phenylisothiocyanate analysis of peptides obtained by enzymic digestion of purified CNBr fragments. The protein consists of a single polypeptide chain of 99 residues, and shows close homology with other higher plant plastocyanins. The data are discussed in relation to the possible residues involved in the binding of copper in plastocyanin.

INTRODUCTION

Plastocyanin is a 'blue', Type 1, copper protein which is involved in photosynthetic electron transport [1]. Recently, studies of its amino acid sequence have been initiated in order to deduce a plant phylogeny based on amino acid sequence comparisons [2]. In the absence of a crystal structure analysis for plastocyanin, the sequence data may give an insight into the nature of the copper binding in the protein, since, when the lettuce and other plastocyanin sequences are compared, the number of invariant residues which could act as ligands is now fairly limited.

RESULTS AND DISCUSSION

The amino acid sequence of plastocyanin from lettuce is shown in Fig. 1, together with the details of the fragments and peptides which were analysed. The protein consists of a single polypeptide chain of 99 residues, and shows great homology with other higher plant plastocyanins [4, 9, 11–15].

The automatic sequencer unambiguously identified 34 of the *N*-terminal 43 residues and allowed probable identification of a further 8 of these residues; this sequence corresponds to the largest CNBr fragment. The sequence data were confirmed, and the remainder of the sequence of the largest CNBr fragment established, by analysis of tryptic and thermolytic peptides. The two large tryptic peptides X1T1 and X1T3 could not be readily separated and so were not further examined. The evidence obtained by degradation for residues His-37, Ala-52, Ser-53 and Lys-54 was weaker than for other residues. However, peptide compositions support the assignments of His-37 and Lys-54 and the tryptic cleavage before Ile-55 also suggests a lysine residue is present, as arginine is absent in the protein. In position 52 both alanine and serine were observed, but with alanine being the stronger. At position 53 both these residues were again observed but with serine being the stronger. It is possible that some heterogeneity may exist in these positions. The automatic sequencer analysis established overlaps between all the residues in the fragment except for that between residues 54 and 55 which has been assumed by homology [4, 11].

Automatic sequence analysis of the second CNBr fragment identified 12 of the *N*-terminal 14 residues. These were confirmed and the remainder of the sequence established by analysis of tryptic and thermolytic peptides. In this fragment, the evidence obtained for Pro-66 was weaker than for other residues in the sequence.

The sequence of the smallest CNBr fragment was established directly and together with the sequences of the other two fragments completes the sequence of lettuce plastocyanin. Overlaps between the CNBr fragments have not been shown but have been assumed by homology with plastocyanins where this has been shown [13, 14]. Where amide residues could not be directly identified, they were established from the peptide mobilities at pH 6.5 [16].

The amino acid compositions of the protein (Table 1) determined by analysis and from the sequence are generally in agreement.

When the lettuce plastocyanin sequence is compared with the 9 other complete higher plant plastocyanin sequences [4, 9, 11–15; Boulter *et al.*, unpublished], and also with the plastocyanin sequences for the algae, *Chlorella fusca* [17] and *Anabaena variabilis* [18] only 30 residues remain invariant (see Figs. 2 and 3). Of these there are 8 glycine and 4 proline residues, many of which may be important for proper folding of the polypeptide chain. In addition the invariant hydrophobic residues may in many cases be important internal residues in the protein.

The number of possible invariant residues which could act as copper ligands is now very limited. Of these, the invariant Cys residue is of interest since X-ray photoelectron spectroscopy studies have demonstrated that a single sulphur atom is involved in the copper binding site

* Present address: Department of Inorganic Chemistry, University of Sydney, N.S.W. 2006, Australia.

† Present address: Microbiological Research Establishment, Porton, Salisbury, Wiltshire, England.

‡ Present address: Bradford University, Postgraduate School of Studies in Biological Sciences, Bradford, BD7 1DP, England.

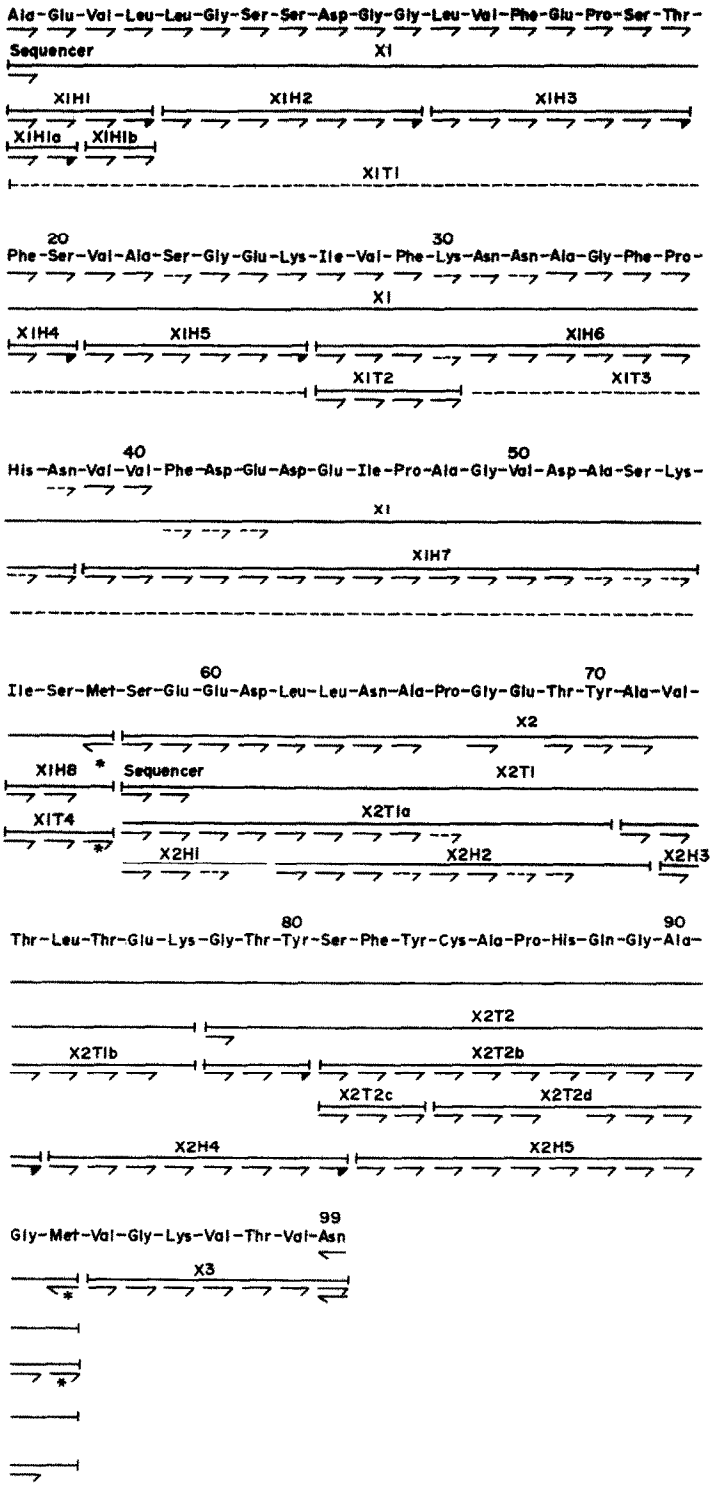

Fig. 1. The amino acid sequence of plastocyanin from lettuce (*Lactuca sativa*). Peptides which were purified for analysis are shown by a solid line. Cyanogen bromide fragments are labelled X; tryptic peptides T; thermolytic peptides, H: Arrows —→ indicate residues which were identified by either the automated or manual sequencing. Where some ambiguity existed in identification an arrow --→ is used. An arrow —→ indicates that the free amino acid was also observed and an arrow ←— indicates identification by carboxypeptidase analysis. An * indicates that methionine was identified as homoserine after CNBr cleavage.

Table 1. Amino acid analysis of lettuce plastocyanin

	By analysis (20 hr)	Sequence
Asp	9.9	5
Asn		5
Thr*	6.3	6
Ser*	8.8	9
Glu	10.6	9
Gln		1
Pro	4.7	5
Gly	10.5	10
Ala	8.8	9
Val	11.6	11
Cys	0.9	1
Met	1.9	2
Ile	2.5	3
Leu	6.5	6
Tyr	3.0	3
Phe	6.4	6
His	1.9	2
Lys	5.3	5
Arg	0	0
Trp†	0	0

* Corrected to zero time of hydrolysis. † Determined spectrophotometrically.

[19]. The Cys residue is the only invariant sulphur-containing side chain; chemical [20] and spectroscopic [21, 22] studies also support this residue as being a copper ligand.

Although several possible invariant N ligands still exist, as well as the polypeptide backbone, the two invariant histidine residues are of note, since both variation of redox potential with pH [20] and NMR studies [23] suggest a close association of histidine to the copper binding site. Also the homology noted between the *C*-terminal regions of plastocyanin and azurin [4, 17, 24] includes homology with one of the invariant histidine residues as well as with the invariant cysteine. A histidine residue homologous to the other invariant histidine residue in plastocyanins may also be present in azurin [4, 17, 24]. The *C*-terminal region of plastocyanin also contains the single invariant tyrosine residue, for which a homologous residue is also present in azurin. While studies with model copper complexes suggest that strong oxygen binding to the copper is unlikely [25], Miskowski *et al.* [22] have suggested that oxygen ligands could occupy weak axial positions in a trigonal bipyrimadal geometry.

−2	−1	1				5					10									20					
Glu	*Asp* *Thr*	*Ala* Ile Leu *Tyr* Val	Asp Glu Thr	Ile Val	Leu Lys	Leu	Gly	Ala Gly Ser	Asp *Glu* Gly Ser	Asp *Lys* *Ser*	Gly	Ala *Glu* Gly *Leu* Ser	Leu	Ala Val	Phe	Glu Ile *Leu* Val	Pro	*Ala* Asn Gly Ser	Asn Asp *Glu* *Lys* Ser *Thr*	*Leu* *Val* Phe	*Glu* Ser Thr	Ile Val	Ala Lys Pro Ser	Ala *Lys* *Pro* Ser	Gly	Asp Glu
Glu Lys Thr	Ile Val	*Glu* Thr Val	Phe *Trp*	*Leu* Lys *Val*	Asn	Asn	Ala *Lys*	Gly *Val*	Phe *Pro*	Pro	His	Asn	Ile Val	40 Val	Phe	Asp	*Ala* Glu	*Ala* Asp	Glu *Leu*	*Asn* Ile Val	Pro	Ala Ser	Gly *Lys*	50 *Ala* *Ser* Val	*Ala* *Asn* Asp	Ala *Asp* *Ser*
Ala *Glu* *Gly* *Leu* Ser *Val*	Ala Lys	Ile *Leu* *Lys*	Ser	* *Leu* Met	* *Ala* *Asn* Asp Pro Ser	Glu His	60 *Ala* *Asn* *Asp* Glu *Lys*	Asp *Gln* *Glu*	Leu *Tyr*	Leu	Asn *Met*	Ala *Ser*	Ala Pro	Gly	*Gln* Glu	Ser Thr Val	70 *Thr* Tyr	Ala *Glu* Lys Ser *Val*	*Ala* *Thr* Val	*Ala* *Asn* Lys Thr	Leu Phe	Asp Ser Thr	Ala Glu Thr	Ala Lys *Ser*	Gly	*Glu* Ser Thr
80 Tyr	*Gly* Lys Ser Thr	Phe *Tyr*	*Phe* Tyr	Cys	Ala Glu Ser	Pro	His	*Arg* Gln	Gly	90 Ala	Gly	*Leu* Met	*Lys* Val	Gly	*Gln* Lys *Thr*	Ile Val	Thr	Val	*Ala* Asn Gln						

Fig. 2. Invariant amino acid residues of the plastocyanin data set. * = Deletion in *Chlorella*.

	Anabaena	*Chlorella*	*Rumex*	*Sambucus*	*Spinacia*	*Cucurbita*	*Capsella*	*Solanum*	*Vicia*	*Phaseolus*	*Mercurialis*	*Lactuca*
Anabaena variabilis	0											
Chlorella fusca	50	0										
Rumex obtusifolius	59	43	0									
Sambucus nigra	60	47	20	0								
Spinacia oleracea	60	48	19	12	0							
Cucurbita pepo	61	48	19	19	12	0						
Capsella bursa-pastoris	61	49	22	20	15	14	0					
Solanum tuberosum	58	47	19	17	17	17	18	0				
Vicia faba	58	42	24	21	17	20	23	22	0			
Phaseolus vulgaris	57	47	22	19	16	19	20	19	20	0		
Mercurialis perennis	59	49	21	16	17	15	17	15	22	16	0	
Lactuca sativa	57	47	18	20	14	17	17	17	21	13	15	0

Fig. 3. Matrix of amino acid differences among plastocyanins. See text for source of sequence data.

EXPERIMENTAL

Plastocyanin was prepared from lettuce leaves as previously described [3]. *S*-carboxymethylated lettuce plastocyanin (3 μmol) was prepared and cleaved with CNBr using the methods previously described [4]. The resulting CNBr fragments were separated by gel-filtration [4] and samples of the two largest fragments were digested with trypsin and with thermolysin using the conditions previously described [4]. Peptides were purified by high voltage paper electrophoresis and by PC [4]. Their amino acid sequences were determined by the dansylphenylisothiocyanate method [5] and by carboxypeptidase A digestion, using the procedures described previously [4, 6, 7]. Automated sequence analysis of the native protein (400 nmol) and of the second CNBr fragment (100 nmol) were carried out on a Beckman 890 C sequenator using 'fast-protein' and 'fast-peptide' programs respectively [8, 9] Identification of the phenylthiohydantoin amino acids by TLC GLC and after regeneration of the amino acid by HI hydrolysis [10] followed the methods previously described [4, 9]. Amino acid analyses were performed on a Locarte amino acid analyser.

REFERENCES

1. Malkin, R. and Malmstrom, B. G. (1970) *Adv. Enzymol.* **33**, 177.
2. Boulter, D. (1973) *Nobel Symp.* **25**, 211.
3. Ramshaw, J. A. M., Brown, R. H., Scawen, M. D. and Boulter, D. (1973) *Biochim. Biophys Acta* **303**, 269.
4. Ramshaw, J. A. M., Scawen, M. D., Bailey, C. J. and Boulter, D. (1974) *Biochem. J.* **139**, 583.
5. Gray, W. R. and Hartley, B. S. (1963) *Biochem. J.* **89**, 378.
6. Thompson, E. W., Laycock, M. V., Ramshaw, J. A. M. and Boulter, D. (1970) *Biochem. J.* **117**, 183.
7. Ramshaw, J. A. M., Thompson, E. W. and Boulter, D. (1970) *Biochem. J.* **119**, 535.
8. Beckman (1972) *Model 890 C Sequencer Instruction Manual.* Spinco Division, Beckman Instruments Inc., Palo Alto.
9. Scawen, M. D. and Boulter, D. (1974) *Biochem. J.* **143**, 257.
10. Inglis, A. S., Nicholls, P. W. and Roxburgh, C. M. (1971) *Australian J. Biol. Sci.* **24**, 1247.
11. Ramshaw, J. A. M., Scawen, M. D. and Boulter, D. (1974) *Biochem. J.* **141**, 835.
12. Scawen, M. D., Ramshaw, J. A. M., Brown, R. H. and Boulter, D. (1974) *European J. Biochem.* **44**, 299.
13. Milne, R. P., Wells, J. R. E. and Ambler, R. P. (1974) *Biochem. J.* **143**, 691.
14. Haslett, B., Bailey, C. J., Ramshaw, J. A. M., Scawen, M. D. and Boulter, D. (1974) *Biochem. Soc. Trans.* **2**, 1329.
15. Scawen, M. D., Ramshaw, J. A. M. and Boulter, D. (1975) *Biochem. J.* **147**, 343.
16. Offord, R. E. (1966) *Nature (London)* **211**, 591.
17. Kelly, J. and Ambler, R. P. (1974) *Biochem. J.* **143**, 681.
18. Aitken, A. (1975) *Biochem. J.* **149**, 657.
19. Solomon, E. I., Clendening, P. J., Gray, H. B. and Grunthaner, F. J. (1975) *J. Amer. Chem. Soc.* **97**, 3878.
20. Katoh, S., Shiratori, I. and Takamiya, A. (1962) *J. Biochem. (Tokyo)* **51**, 32.
21. McMillin, D. R., Rosenberg, R. C. and Gray, H. B. (1974) *Proc. Natl. Acad. Sci. U.S.* **71**, 4760.
22. Miskowski, V., Tang, S. P. W., Spiro, T. G., Shapiro, E. and Moss, T. H. (1975) *Biochemistry* **14**, 1244.
23. Markley, J. L., Ulrich, E. L., Berg, S. P. and Krogmann, D. W. (1975) *Biochemistry* **14**, 4428.
24. Ambler, R. P. (1971) In *Recent Developments in the Chemical Study of Protein Structure*, pp. 289–305. INSERM, Paris.
25. Freeman, H. C. (1973) Ch. 4. Metal Complexes of Amino acids and Peptides. In *Inorganic Biochemistry* Vol. 1, (Ed. Eichhorn, G. L.), pp. 121–166. Elsevier, Amsterdam.