Phytochemistry 1974 Vol 13 pp 2777 to 2782 Pergamon Press Printed in England

THE AMINO ACID SEQUENCE OF CYTOCHROME *c* FROM *ENTEROMORPHA INTESTINALIS*

BARRY T. MEATYARD and DONALD BOULTER*
Department of Botany, University of Durham, Durham DH1 3LE, England

(*Received* 14 *June* 1974)

Key Word Index—*Enteromorpha intestinalis*, Chlorophyta, ulvales, green alga, basic cytochrome *c*, amino acid sequence, ε-*N*-trimethyllysine

Abstract—Proposed amino acid sequence of cytochrome *c* from *Enteromorpha intestinalis* is presented. The cytochrome is a basic protein, homologous with higher plant and animal cytochromes *c* of mitochondrial origin Peptides from chymotryptic and tryptic digests were analysed by the dansyl-phenylisothiocyanate method and aligned by comparison with other homologous cytochromes *c* The proposed sequence consists of a single polypeptide chain of 111 residues which is acetylated at its *N*-terminus The sequence contains one residue of the unusual amino acid ε-*N*-trimethyllysine in position 80

INTRODUCTION

THE AMINO acid sequences of numerous higher plant cytochromes *c* have been determined and used to construct an affinity tree relating these sequences.[1,2] The aim of this investigation was to extend this work by studying the sequence of a cytochrome *c* from a lower plant.

RESULTS AND DISCUSSION

Species of *Enteromorpha* are difficult to identify except from details of their life cycles.[3,4] Due to the problem of collecting sufficient quantity of *Enteromorpha*, it was not possible, therefore, to unequivocally identify the source material as a single species and it may have consisted of a mixture of the closely related *E. linza*, *E. flexuosa* and *E. intestinalis*. However, there was no extensive heterogeneity found during the sequence analysis.

The yield of cytochrome *c* obtained, 0·13 mg/kg, was lower than the yields obtained with the majority of higher plants.[5,6] However, in this case it has been demonstrated that cytochrome *c* can be purified in reasonable amounts from green vegetative material and

* To whom reprint requests should be sent

[1] BOULTER, D., RAMSHAW, J. A. M., THOMPSON, E. W., RICHARDSON, M. and BROWN, R. H. (1972) *Proc. Roy. Soc. (London)* **B 181**, 441
[2] BOULTER, D. (1973) in *Chemistry in Botanical Classification*, Nobel Symp. 25 (BENDZ, G. and SANTESSON, J. eds.) Nobel Foundation-Stockholm, Academic Press, New York
[3] BLIDING, C. (1963) *Opera Botany* **8**, 1
[4] COOK, P. W. and HOFFMAN, L. R. (1971) in: *Selected Papers in Phycology* (ROSOWSKI, J. R. and PARKER, B. D. eds.), University of Nebraska Press
[5] RICHARDSON, M., LAYCOCK, M. V., RAMSHAW, J. A. M., THOMPSON, E. W. and BOULTER, D. (1970) *Phytochemistry* **9**, 2271
[6] RICHARDSON, M., RICHARDSON, D., RAMSHAW, J. A. M., THOMPSON, E. W. and BOULTER, D. (1971) *J. Biochem. (Tokyo)* **69**, 811

the yields obtained were about twice those obtained from spinach leaves, 0·07 mg/kg [7] The behaviour of the cytochrome during preparation and its mobility on gel electrophoresis showed that it was a basic protein very similar to other cytochromes *c* from plant, animal and fungal sources and is probably of mitochondrial origin. An acidic cytochrome c_{554} was also present in *Enteromorpha*, which may function in photosynthesis Previously, the algal cytochromes which have been examined, have all been acidic proteins which resemble cytochrome *f* and probably function in photosynthesis [8,9]

TABLE 1 THE AMINO ACID COMPOSITION OF *Enteromorpha* CYTOCHROME *c*

Amino acid	Analysis value*	Sequence value
Asp	13·0	15–16
Thr	8·6†	8
Ser	5·0†	4
Glu	6·4	3–4
Pro	6·2	7
Gly	11·6	11
Ala	17·0	19
Val	3·0‡	2
Cys	2·1	2
Met	0·9	1
Ile	3·9‡	3
Leu	7·6‡	7
Tyr	4·4†	4
Phe	5·6	6
TML }	13·0	1
Lys }		12
His	1·9	2
Arg	2·0	2
Trp	n d	1

* The average of several separate 24 and 72 hr hydrolysates, except that values of Val Ile, Leu were corrected for slow release and Thr, Ser, Tyr for amino acid breakdown

† Extrapolated to zero time assuming first order rates for destruction [20]

‡ Maximal values were taken

§ The spectral ratios of the pure cytochrome suggest that one residue of tryptophan is present n d -Values were not determined

The amino acid compositions of *Enteromorpha* cytochrome *c* determined by analysis and calculated from the sequence, are given in Table 1. The analysis values were obtained from independent sets of analyses on two different instruments using several different preparations of cytochrome *c*, extracted from different collections of algal material The results show slight variations between the values obtained, this was due either to variations between the instruments and conditions used, or to the presence of differences in the samples However, when both sets of analyses are considered the values obtained are generally in reasonable agreement with those calculated from the complete sequence (Fig 1), except that the aspartic and alanine values were low and the serine, glutamic acid, valine and isoleucine values were high

Figure 1 gives the sequence of *Enteromorpha* cytochrome *c* and shows the chymotryptic and tryptic peptides which were purified and analysed Both enzymes cleaved the protein

[7] ASADA, K and TAKAHASHI, M (1971) *Plant and Cell Physiol* **12**, 361
[8] LAYCOCK, M V (1972) *Can J Biochem* **50**, 1311
[9] LIMBERG R and BARRETT, J (1973) *Cytochromes* Academic Press, London

in the positions expected from their known specificities The peptides obtained from the three chymotryptic digests account for almost the entire sequence. The peptides obtained from each digest were essentially similar However, certain peptides, e.g. C6, C9, were obtained in low yield or were difficult to purity and so were not analysed from each digest. Also, peptides C2, and C8 were not located in all three digests probably as a result of low yields The low yield of peptide C2 may have been due to poor cleavage at Phe-3

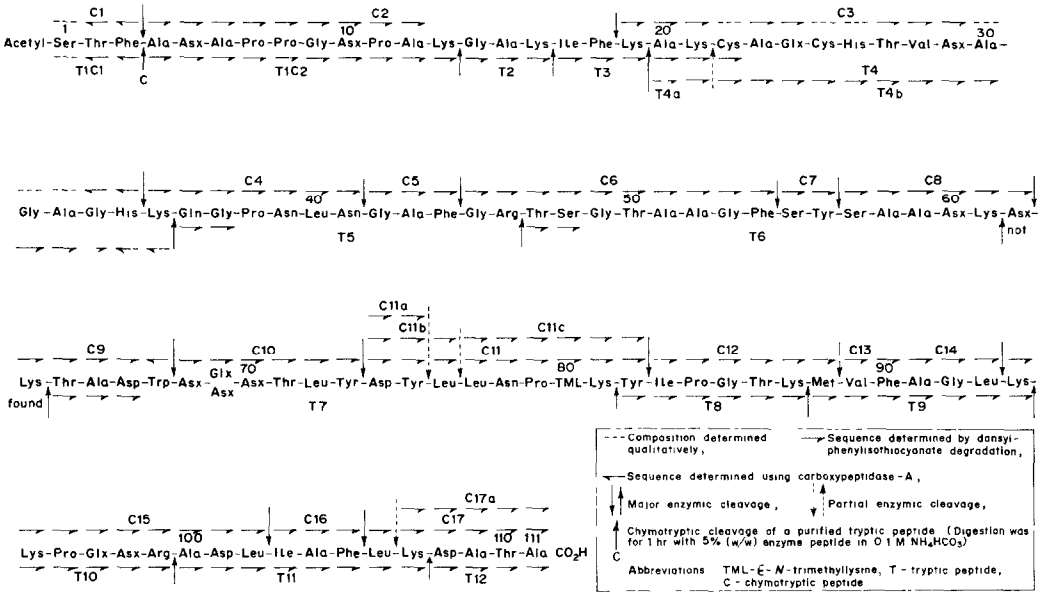

FIG 1 AMINO ACID SEQUENCE OF CYTOCHROME *c* FROM *Enteromorpha intestinalis*

The peptides obtained from the two tryptic digests were sufficient to identify the remaining unidentified residues in the complete sequence. Thus, every residue in the proposed sequence was identified in at least one digest. Many of the tryptic peptides were only recovered in low yield and were difficult to purify due to their poor electrophoretic properties. Similar peptides presented equivalent problems during the analysis of several higher plant cytochromes. No differences were found between equivalent peptides from the two tryptic digests The tryptic peptides enable the majority of the chymotryptic peptides to be aligned The remaining unaligned peptides, e.g C5, C6, C7 and C8 are either the same as or very similar to peptides from higher plant cytochromes *c* and so can be placed in the sequence by homology with this group of protein sequences. The compositions of tryptic peptides which were not completely sequenced because of insufficient material, support these placements based on homology; their compositions were determined semi-quantitatively

The purified peptides were normally analysed by using the dansyl-phenylisothiocyanate method,[10] in certain cases, however, carboxypeptidase analysis was needed. The acetylated *N*-terminal peptides, C1 and T1C1 were digested with carboxypeptidase A. After 2 hr only

[10] GRAY, W R and HARTLEY, B S (1963) *Biochem J* **89**, 379

TABLE 2: PEPTIDE MOBILITIES AT pH 6 5 USED FOR ASSIGNMENT OF AMIDE RESIDUES.

Peptide	Mobility at pH 6 5*	Assignments and comments
C2	0	Asp-5, Asp-10, since the length of the peptide was not definitely established during analysis, these residues are left as Asx in Fig 1
C3	0	Glx-24, Asx-29, the mobility suggests that one amide residue is present when peptide mobility is compared with other cytochrome haem peptides of known charge
T4	0	
C4	+ve	Gln-36, Asn-39, Asn-41, Mobility not accurately determined Since it had a net positive charge at pH 6 5, however, residues 36 39 and 41 must be present as amide residues
C8	0	Asx-60, Asx-62, 1 amide residue present The neutral mobility of peptide C8 suggests that either residue 60 or 62 is present as aspartic acid No peptide due to tryptic cleavage between lysine residues 61 and 63 was found to resolve this problem In other plant cytochromes these two positions have always been amide residues Margoliash[19] has shown however that in horse heart cytochrome c the two amide residues in equivalent positions in the sequence are the most likely to be deamidated, and it is possible that deamidation could have occurred therefore in this preparation
T6	0	Asp-60
C9	0	Asp-66
C10	−1 1	Asx-68, (Glx/Asx)-69 Asx-70, the expected mobility of this peptide if two amides were present was −0 9, if one amide were present the expected mobility was −1 7 Thus although the mobility is slightly high it suggested that two amides were present whose position remains undecided
C11a	−1 63	Asp-74
C11b	−1 30	Asp-74
C11c	1 65	Asn-78
C15	0 68	Glx-97, Asx-98, Asp-101 (see T11), 1 amide present
T11	0	Asp-101
C17	0	Asp-108
C17a	0	Asp-108

* Mobility at pH 6 5 was measured from the position of the neutral amino acids relative to Dansyl-Arg-Arg Amide content was determined using the method of Offord [21]

phenylalanine had been released but after 24 hr digestion threonine had also been released. Since the composition of the peptide, determined semi-quantitatively, indicated the presence of a serine residue in addition to these two residues, this was placed at position 1 by difference, giving the sequence Acetyl-Ser-Thr-Phe The acetyl group was determined using peptide TICl. The haem peptides C3 and T4b were always recovered in low yield after removal of the haem moiety Carboxypeptidase A digestion of peptide C3 released only histidine after 15 min and both histidine and glycine after 1 hr After 1 hr digestion of peptide T4b by carboxypeptidase A both lysine and histidine had been released. Peptides C9 and T7 both gave positive reactions to the Ehrlich test, indicating the presence of tryptophan This amino acid was confirmed to be present in the peptide on the first

step of the dansyl-phenylisothiocyanate analysis when characteristic pink colours were observed in the trifluoroacetic acid.[11] Digestion of peptide C9 with carboxypeptidase A for 1 hr demonstrated that tryptophan was the C-terminal residue of this peptide The spectral ratios of the native protein suggested that one tryptophan was present in the protein.

Semi-quantitative analysis of the composition of peptide T7 showed that it contained ε-*N*-trimethyllysine. That it occurred at position 80 in the sequence was confirmed by analysis of peptides C11 and C11b In higher plants a second residue of ε-*N*-trimethyllysine occurs at position 94, whereas in *Enteromorpha* this position is occupied by a lysine residue, and this allows tryptic cleavage of the polypeptide chain at this point, cleavage does not occur at this position when ε-*N*-trimethyllysine is present.

The C-terminal residue of the protein was shown to be alanine from peptides C17, C17a and T12. In each peptide, after the appropriate numbers of steps of degradation, alanine was shown to be present as the free amino acid.

In peptide C10, which was obtained in good yield from all three chymotryptic digests, both glutamyl and aspartyl residues were found in the second position (residue 69) on each occasion. This was not caused by carry-over of the *N*-terminal aspartyl residue as a result of incomplete degradation, since further degradations gave clean results It is presumed that this result is due to heterogeneity in the protein preparation. This may have been caused by more than one *Enteromorpha* species being present or by polymorphism of the protein in a single species Any other heterogeneity in the sequence, if present, was at a low level as it was not detected during any of the dansyl–phenylisothiocyanate analyses.

In view of the relative lack of "redundant" information and the fact that differences exist between the analytical and sequence amino acid compositions, the proposed sequence is tentative. The presence of amide groups cannot be directly established using the dansyl–phenylisothiocyanate analysis Their presence was determined therefore, when possible, from the mobilities of the peptides at pH 6·5 (Table 2). This sets out those positions to which amides have been assigned together with the reasoning involved.

The sequence determination clearly indicates homology of this sequence with the higher plant cytochromes *c*.[1,2] for example, the tail of the protein is acetylated and is 8 residues longer than the animal cytochromes *c* Certain residues, for example, Ala-32, Gly-33, Gln-36, Thr-50, Tyr-73 and Leu-77, only occur in these positions elsewhere in cytochromes *c* of higher plants. The sequence, however, contains only one residue of the unusual amino acid ε-*N*-trimethyllysine, in position 80, whereas all the higher plant cytochromes *c* contain two residues of this amino acid, in positions 80 and 94.

EXPERIMENTAL

Materials Approx 300 kg of *Enteromorpha* (fr wt) were collected from Eyemouth Harbour, Berwickshire, on various occasions Other materials were as described by Thompson *et al* [12,13]

Methods Enteromorpha cytochrome *c* was purified using the methods described by Richardson *et al* [5,6] for higher plant cytochrome *c*, except that the homogenate was adjusted to pH 8 with 1 M-tris prior to filtration Cytochrome *c* [E_{550} (red)/E_{280} (ox) = 1 24] was obtained with a yield of approx 0 1 mg/kg and was shown to be a single component when examined by gel electrophoresis at pH 4 5 and at pH 8 3 [14–16] The cytochrome

[11] Uphaus, R A, Grossweiner, L I, Katz, J J and Kopple, K D (1959) *Science* **129**, 641
[12] Thompson, E W, Laycock, M V, Ramshaw, J A M and Boulter, D (1970) *Biochem J* **117**, 183
[13] Thompson, E W, Richardson, M and Boulter, D (1970) *Biochem J* **121**, 439
[14] Reisfeld, R A, Lewis, O J and Williams, D E (1962) *Nature* **195**, 281
[15] Davis, B J (1964) *Ann N Y Acad Sci* **121**, 305
[16] Ornstein, L (1964) *Ann N Y Acad Sci* **121**, 321

was oxidized with $K_3Fe(CN)_6$ and denatured with EtOH prior to digestion with proteolytic enzymes. The sequence was determined by the analysis of chymotryptic and tryptic peptides prepared and purified as described by Thompson *et al*.[12] Peptides from three chymotryptic digests and two tryptic digests were analysed, approx. 6 mg (0.5 μmol) cytochrome *c* was used for each digest. The peptides were purified by high-voltage paper electrophoresis at pH 6.5 and 1.9 as described previously,[12] and their amino acid sequence determined using the dansyl-phenylisothiocyanate method[10] as described previously.[12] dansyl-amino acids were identified by chromatography on polyamide sheets[17] using the solvent systems described by Ramshaw *et al*.[18]. Digestion of peptides with carboxypeptidase A and identification of the released amino acids as their dansyl derivatives and further digestion of peptides with chymotrypsin, were both as described previously.[12] Quantitative amino acid analyses were performed on Technicon and Locarte amino acid analysers using the methods previously described.[12]

Acknowledgements. We thank Dr. I. A. M. Ramshaw for help with the manuscript, Mr. G. Bainbridge, Mrs. E. Innes, Mrs. L. Mayo and Mrs. L. Cooper, for invaluable technical assistance, and the Nuffield Foundation and the Science Research Council for financial support.

[17] Woods, W. R. and Wang, K. T. (1967). *Biochim. Biophys. Acta* **133**, 369
[18] Ramshaw, I. A. M., Thompson, E. W. and Boulter, D. (1970). *Biochem. J.* **119**, 535
[19] Margoliash, E. (1964). In unpublished experiments cited by: Margoliash, E. and Schecter, A. (1966). *Adv. Prot. Chem.* **21**, 136
[20] Moore, S. and Stein, W. H. (1963). In: *Methods in Enzymology*, Vol. 6, p. 819. (Colowick, S. P. and Kaplan, N. O., eds.) Academic Press, New York
[21] Offord, R. E. (1966). *Nature* **211**, 591