Phytochemistry, Vol. 23, No. 4, pp. 773–776, 1984.
Printed in Great Britain.

0031–9422/84 $3.00 + 0.00

AMINO ACID SEQUENCE OF A FERREDOXIN FROM *RHODYMENIA PALMATA*, A RED ALGA IN THE FLORIDEOPHYCEAE

KATSURA INOUE*, TOSHIHARU HASE, HIROSHI MATSUBARA, MICHAEL P. FITZGERALD†‡ and LYNDON J. ROGERS†

Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan; †Department of Biochemistry, University College of Wales, Aberystwyth, SY23 3DD, Wales, U.K.

(*Received* 12 *August* 1983)

Key Word Index—*Rhodymenia palmata*; Rhodophyta; Florideophyceae; red algae; ferredoxin; protein sequence; taxonomy.

Abstract—The amino acid sequence of a [2Fe-2S] ferredoxin from a red alga, *Rhodymenia palmata* in the family Florideophyceae, was determined by conventional methods. The ferredoxin is composed of 97 amino acid residues having five cysteines, but lacking methionine and tryptophan. It possesses a number of structural features of particular interest. The amino acid sequence is compared with those previously determined for ferredoxins from two red algae in the family Bangiophyceae. Conclusions from a comparison of the structures, by noting features such as the presence of gaps in the sequences and by constructing a phylogenetic tree, were consistent with the proposed taxonomic relationship among these algae.

INTRODUCTION

Comparison of the amino acid sequences of [2Fe-2S] ferredoxins from various photosynthetic organisms has enabled the evolution of this protein to be traced, and has greatly helped our understanding of structure–function relationships [1, 2]. In particular, taxonomic relationships amongst the algae have been clarified by comparison of the sequences of their ferredoxins [3, 4].

To date only two complete amino acid sequences [5, 6] and three partial amino (*N*)-terminal sequences [7, 8] are available for ferredoxins from red algae (Rhodophyta). One of the complete sequences is that of the ferredoxin isolated from *Porphyra umbilicalis*, a member of the primitive red algae, the Bangiophyceae. The other is that of the ferredoxin from *Cyanidium caldarium*, an alga which can be placed in this family on the basis of recent comparative studies of the structure of its ferredoxin [1, 3, 4, 6]. Two of the three partial sequences were those of ferredoxins from a marine and fresh water species of *Porphyridium*, which is also placed in the Bangiophyceae. The remaining partial sequence was that of a ferredoxin isolated from *Rhodymenia palmata* (L) Greville, alternavively named *Palmaria palmata* (L) Stackhouse [8], which is a member of the Florideophyceae, the other family in the Rhodophyta. The *N*-terminal sequence of 48 amino acid residues was reported, but with a number of ambiguous identifications [8]. The full amino acid sequence of this latter ferredoxin is of interest, both to compare its structure with those of other red algal ferredoxins, and to gain further insight into taxonomic and evolutionary relationships within this important group of algae.

*Present address: Department of Medical Chemistry, Osaka Medical College, Takatsuki, Osaka 569, Japan.

‡Present address: The Horticultural Centre, Loughgall, Co. Armagh, N. Ireland.

RESULTS AND DISCUSSION

A summary of the conclusions from the amino acid sequence studies is given in Fig. 1 and this sequence was based on the following studies.

Amino acid composition and terminal sequences of carboxymethyl (Cm)-ferredoxin

The amino acid composition of *R. palmata* Cm-ferredoxin based on a 24 hr hydrolysis and that from the established sequence is given in Table 1 together with that reported previously [8]. The discrepancies between the compositions deduced from the sequence and of each analysis might be the result of slight heterogeneity of the ferredoxin, which had proved more difficult to purify than those from other algae [8].

The amino (*N*)-terminal analyses were carried out by both manual and automated solid-phase Edman degradations to establish *N*-terminal sequences of 16 residues and 36 residues, respectively, but with a number of ambiguous identifications. The previous *N*-terminal analysis [8] identified position 23 as serine, but the present analysis showed it to be threonine, and this was confirmed by other sequence analyses as described below. The reason for this discrepancy is unclear, and the rest of this part of the sequence confirmed the earlier work [8]. Carboxypeptidase Y released tyrosine (1.0) and leucine (0.5) after 30 min digestion to suggest the carboxyl (*C*)-terminal sequence to be –Leu–Tyr.

Sequence studies of Cm-ferredoxin

A tryptic digest fractionated by chromatography on Bio-Gel P-6 gave five tryptic peptides, T-1 to T-5; as necessary (T-2, T-3) these were further purified on a DE-32 column. The amino acid compositions of T-1 to T-5 are given in Table 1.

Each peptide was subjected to manual Edman degra-

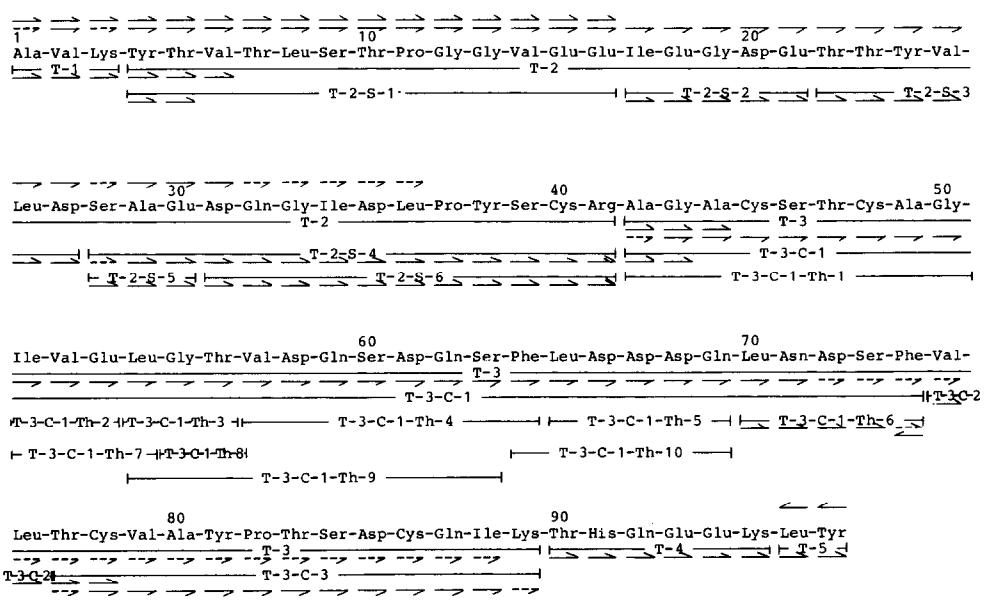

Fig. 1. Summary of sequence studies of *R. palmata* Cm-ferredoxin. T, S, C and Th refer to tryptic, staphylococcal protease, chymotryptic and thermolysin peptides, respectively. Cys was identified as Cm-cysteine. The amino acid sequence was determined by a combination of manual Edman degradation (→), automated Edman degradation (⇀), and carboxypeptidase Y digestion (←), respectively. Arrows above the sequence represent analyses on Cm-ferredoxin while those below the sequence are analyses on derived peptides; the latter include identification of the free amino acid (↠) after Edman degradations to the final residue. Dotted arrows represent ambiguous identifications.

Table 1. Amino acid composition of Cm-ferredoxin and of the peptides derived by tryptic digestion

Amino acid	Cm-ferredoxin*	(9)†	T-1	T-2	T-3	T-4	T-5
Cmc	4.1 (5)	5		1.3 (1)	3.6 (4)		
Asp	12.4 (12)	12		4.9 (4)	8.4 (8)		
Thr	8.4 (10)	9		4.4 (5)	4.4 (4)	0.8 (1)	
Ser	7.1 (8)	8	0.3	2.9 (3)	4.4 (5)		
Glu	14.8 (14)	13	0.3	6.1 (6)	6.1 (5)	2.8 (3)	
Pro	2.8 (3)	4		1.6 (2)	1.1 (1)		
Gly	9.6 (7)	8	0.3	4.0 (4)	3.6 (3)		
Ala	6.9 (6)	7	1.0 (1)	1.6 (1)	3.5 (4)		
Val	7.4 (8)	8	1.0 (1)	3.0 (3)	3.7 (4)		
Met	0.5 (0)	0					
Ile	3.3 (4)	4		1.7 (2)	1.6 (2)		
Leu	7.1 (8)	9		3.2 (3)	4.2 (4)		1.0 (1)
Try	3.7 (5)	4		2.1 (3)	0.9 (1)		0.8 (1)
Phe	2.1 (2)	2		0.7 (0)	1.8 (2)		
Lys	2.7 (3)	3	1.0 (1)	0.2	0.9 (1)	1.0 (1)	
His	1.1 (1)	1				1.1 (1)	
Arg	1.4 (1)	1		0.9 (1)			
Total residues	97		3	38	48	6	2
Yield (%)			55	32	44	44	36
Purification				DE-32	DE-32		

*Based on a 24 hr hydrolysis; the numbers in parentheses are those calculated from the final sequence.

†Composition from ref. [9].

dations (peptide T-1: three steps, peptide T-2: three steps, peptide T-3: three steps, peptide T-4: six steps, peptide T-5: two steps). Sequences of peptides T-1, T-4, and T-5 were completely determined by this method. Automated Edman degradation resolved peptide T-3 to 30 residues, with a further 17 residues subject to some uncertainty. Residue 6 of T-3 (47 in the full sequence) was identified as threonine; the earlier *N*-terminal analysis [8] suggested this was aspartate.

Peptide T-2 was further digested with staphylococcal protease and the digest was separated by paper electrophoresis at pH 3.6. Six peptides, T-2-S-1 to T-2-S-6, were isolated and completely sequenced by manual Edman degradations except for peptide T-2-S-1 which was subjected to only two steps of the degradation. Thus, with the information from the automated Edman degradation of the original protein, peptide T-2 was completely sequenced.

Peptide T-3 was digested with chymotrypsin and the digest was purified by paper electrophoresis at pH 3.6. Three peptides, T-3-C-1 to T-3-C-3 were isolated. After two steps of manual Edman degradation peptide T-3-C-1 was further digested with thermolysin and the digest was fractionated by high performance liquid chromatography. Ten peptides were isolated and they were all unequivocally assigned to the appropriate positions in peptide T-3-C-1 on the basis of their amino acid compositions. Only peptide T-3-C-1-Th-6 needed to be subjected to manual Edman degradation and carboxypeptidase Y digestion to complete this part of the sequence. The sequence of peptide T-3-C-2 was determined by manual Edman degradation. Peptide T-3-C-3 was subjected to manual (two steps) and solid-phase (13 steps) Edman degradations. The *C*-terminal residue of this peptide was not unequivocally identified, but can be assigned as lysine in accord with trypsin specificity. Moreover, from the amino acid composition of the ferredoxin (Table 1) the single arginine had already been placed and one lysine was unaccounted for. Thus, peptide T-3 was completely sequenced. There were no overlaps involving two peptides, T-3 and T-4, but so many homologous amino acid sequences of [2Fe-2S] ferredoxins isolated from various organisms [3] are available that these fragments could be assigned in order on that basis.

The total number of amino acid residues was 97; and the amino acid composition from the sequence was Asp_{11}, Asn_1, Thr_{10}, Ser_8, Glu_8, Gln_6, Pro_3, Gly_7, Ala_6, Cys_5, Val_8, Ile_4, Leu_8, Tyr_5, Phe_2, Lys_3, His_1, Arg_1, with methionine and tryptophan absent. The MW of *Rhodymenia palmata* ferredoxin, including the iron and sulfur atoms, is thus 10740.

When the aligned sequences [3] for the 38 ferredoxins whose structures have been determined are compared, several amino acids in *R. palmata* ferredoxin occur for the first time at specific positions in the sequences. These are valine-2 and -14, lysine-3, serine-9, -28, and -73, glycine-12 and -19, threonine-22, leucine-54, asparagine-71, aspartic acid-72, serine-73 and glutamine-87. At position 73 all the other ferredoxins so far sequenced have glycine.

Phylogenetic relationship of R. palmata *ferredoxin to other red algal ferredoxins*

Comparisons of the sequences of all the eukaryotic and prokaryotic algal ferredoxins so far reported have been presented previously [1, 3, 4] and discussion here will be confined to the relationship between ferredoxins from red algae.

Only three full amino acid sequences for ferredoxins from red algae are available; these being the proteins from *P. umbilicalis* [5], *C. caldarium* [6] and *R. palmata* (present work). The partial sequences for the ferredoxins from two species of *Porphyridium* have also been reported [7]. These sequences are compared in Fig. 2, the alignment being made to obtain the highest homology. On the basis of this alignment the amino acid differences for the three complete sequences were calculated; as in previous comparisons [3, 4, 9] one gap was counted as five amino acid differences, because the events producing gaps or additions must be rarer than those of point mutations. A matrix for the differences found is given in Fig. 3. Although only partial sequences are known for the two *Porphyridium* ferredoxins, they are on this basis more similar to the ferredoxins from *P. umbilicalis* and *C. caldarium* than to that from *R. palmata*. *P. umbilicalis*, *Po. aerugineum*, *Po. cruentum* and *C. caldarium* are all placed in the Bangiophyceae and *R. palmata* is in the Florideophyceae. The distinction is clearly reflected by characteristics in the sequences; in particular the presence and absence of gaps at positions 11 and 15. The presence of two gaps at positions 11 and 15 in *R. palmata* ferredoxin suggest this organism may be closer to green algae than to other red and yellow–green algae.

EXPERIMENTAL

Algal cells and isolation of ferredoxin. Rhodymenia palmata cells were harvested from the *Laminaria* Zone of the seashore at

```
(1) Porphyridium aerugineum    AKKYKVRLLSEAEGIDVTIDSADDVYILDA
(2) Porphyridium cruentum      AT-YKVRLLSEAEGIDVTIDCADDVYILDA
(3) Porphyra umbilicalis       AD-YKIHLVSKEEGIDVTFDCSEDTYILDAAEEEGIELPYSCRAGACSTCAGKVTEGTVD
(4) Cyanidium caldarium        AS-YKIHLVNKDQGIDETIECPDDQYILDAAEEQGLDLPYSCRAGACSTCAGKLLEGEVD
(5) Rhodymenia palmata         AVKYTVTLST-PGG-VEEIEGDETTYVLDSAEDQGIDLPYSCRAGACSTCAGIVELGTVD
                               1         1         2         3         4         5         6
                                         0         0         0         0         0         0

(3) QSDQSFLDDEQMLKGYVLTCIAYPESDCTILTHVEQELY
(4) QSDQSFLDDDQVKAGFVLTCVAYPTSNATILTHQEESLY
(5) QSDQSFLDDDQLNDSFVLTCVAYPTSDCQIKTHQEEKLY
                   7         8         9         9
                   0         0         0         9
```

Fig. 2. Comparison of amino acid sequences of red algal ferredoxins. Gaps are inserted to produce the highest homology among the sequences. (1) and (2) are the partial sequences of ferredoxins from two *Porphyridium* species. Residues underlined represent ambiguous identifications. The sequences are taken from refs [5–7], except *Rhodymenia palmata* (present work).

	P. u.	C. c.	R. p.
P. umbilicalis	0	28	52
C. caldarium	28	0	48
R. palmata	52	48	0

Fig. 3. Matrix of amino acid differences. The matrix was derived on the basis of the alignment given in Fig. 2 with each gap counted as five amino acid differences.

Aberystwyth and the ferredoxin was isolated and purified as described previously [8].

Amino acid sequence determination. Cm-ferredoxin was prepared according to the method described in ref. [10]. Cm-ferredoxin (50 and 100 nmol, respectively) was separately subjected to manual and automated solid-phase Edman degradation to determine the *N*-terminal sequence. A further sample (17 nmol) was digested with carboxypeptidase Y (3 μg) in 0.1 M pyridine–NaOAc buffer, pH 5.5, at 40° for 5–30 min to determine the *C*-terminal sequence.

The digestion of the Cm-ferredoxin with trypsin, separation of the resulting peptides and their sequence determination were essentially as described in detail in previous papers [11, 12]. Cm-ferredoxin (1 μmol) was digested with trypsin (0.21 mg) in 0.1 M Tris–HCl buffer, pH 8, for 4 hr at 40°. The digest was separated on a Bio-Gel P-6 column (2 × 180 cm) developed with 0.2 M NH_4HCO_3–NH_4OH buffer, pH 9. Peptides eluted were monitored by measuring the *A* at 230 nm. Peptides T-2 and T-3 were further chromatographed on a DE-32 column (1.5 × 30 cm) with a linear gradient of NH_4HCO_3 from 20 to 800 mM (300 ml in each reservoir) at a flow rate of 18 ml/hr. Some peptides were subjected to further digestion with staphylococcal protease, chymotrypsin or thermolysin and the resulting peptides were purified by paper electrophoresis at pH 3.6 or by HPLC as described previously [13]. The amino acid compositions of proteins and peptides were analysed as described previously [14].

Nomenclature of peptides. T- refers to the tryptic peptides of Cm-ferredoxin, and S- to staphylococcal protease peptides derived by secondary fragmentation. C- refers to chymotryptic peptides derived by secondary fragmentation of tryptic peptides and Th- to peptides then obtained by further digestion with

Acknowledgements—This work was supported in part by Grants-in-Aid for Special Project Research (NO. 57120005 and 58470131) from the Ministry of Education, Science and Culture of Japan to H.M. and by a U.K. Science Research Council Grant (GR/A 1427-8) to L.J.R.

REFERENCES

1. Matsubara, H., Hase, T., Wakabayashi, S. and Wada, K. (1980) in *The Evolution of Protein Structure and Function* (Sigman, D. S. and Brazier, M. A. B., eds) pp. 245–266. Academic Press, New York.
2. Tsukihara, T., Kobayashi, M., Nakamura, M., Katsube, Y., Fukuyama, K., Hase, T., Wada, K. and Matsubara, H. (1982) *BioSystems* **15**, 243.
3. Matsubara, H. and Hase, T. (1983) in *Proteins and Nucleic Acids in Plant Systematics* (Jensen, U. and Fairbrothers, D., eds) Springer, Berlin (in press).
4. Hase, T., Inoue, K., Hagihara, N., Matsubara, H., Williams, M. M. and Rogers, L. J. (1983) *J. Biochem.* **94**, 1457.
5. Takruri, I., Haslett, B. G., Boulter, D., Andrew, P. W. and Rogers, L. J. (1978) *Biochem. J.* **173**, 459.
6. Hase, T., Wakabayashi, S., Wada, K., Matsubara, H., Jüttner, F., Rao, K. K., Fry, I. and Hall, D. O. (1978) *FEBS Letters* **96**, 41.
7. Andrew, P. W., Rogers, L. J., Haslett, B. G. and Boulter, D. (1981) *Phytochemistry* **20**, 1293.
8. Andrew, P. W., Rogers, L. J., Haslett, B. G. and Boulter, D. (1981) *Phytochemistry* **20**, 579.
9. Fitch, W. M. and Smith, T. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1382.
10. Crestfield, A. M., Moore, S. and Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622.
11. Hase, T., Matsubara, H., Ben-Amotz, A., Rao, K. K., and Hall, D. O. (1980) *Phytochemistry* **19**, 2065.
12. Hase, T., Matsubara, H., Hutber, G. N., and Rogers, L. J. (1982) *J. Biochem.* **92**, 1347.
13. Inoue, K., Hase, T., Böger, P., and Matsubara, H. (1983) *J. Biochem.* **94**, 1451.
14. Spackman, D. H., Stein, W. H., and Moore, S. (1958) *Analyt. Biochem.* **30**, 1190.