Phytochemistry, 1979, Vol. 18, pp. 1481–1484. © Pergamon Press Ltd. Printed in England. 0031-9422/79/0901-1481 $02.00/0

THE AMINO ACID SEQUENCE OF FERREDOXIN FROM *SAMBUCUS NIGRA*

ISHAQ A. H. TAKRURI and DONALD BOULTER
Department of Botany, University of Durham, Durham, DH1 3LE, U.K.

(*Received* 9 *February* 1979)

Key Word Index—*Sambucus nigra*: Caprifoliaceae; amino acid sequence; ferredoxin; matrix of differences.

Abstract—The amino acid sequence of the ferredoxin from *Sambucus nigra* consists of a single polypeptide chain of 97 amino acid residues, 5 of which are cysteine. The positions of the 4 cysteine residues which bind the iron atoms of the active centre are identical to those of other ferredoxins. Due to difficulties of obtaining pure protein, residues 87–90 have only been identified from the amino acid analysis of peptide C 10 and by homology with other higher plant ferredoxins.

INTRODUCTION

Ferredoxins have been isolated and characterized from various bacteria, algae and higher plants [1]. The primary structures of the (2Fe-2S) ferredoxins from one green alga, one red alga, some blue-green algae and higher plants have also been determined [2, 3]. These ferredoxins are small proteins with a MW of approximately 11500, and they act as electron carriers in a number of different biochemical processes including photosynthetic electron transport and nitrogen fixation.

The amino acid sequence of *Sambucus* ferredoxin shows great similarity to those of other ferredoxins which have been isolated and sequenced from algae and higher plants [2, 3]. It was determined in order to generate more data for the construction of an affinity tree relating the evolution of higher plants [4].

RESULTS

Sambucus ferredoxin was isolated to a high degree of purity with an absorption index $A(420)/A(280) = 0.43$. The homogeneity of the protein was shown by isoelectric focusing since a single band corresponding in position to a pI value of 3.8 was found.

The complete amino acid sequence is shown in Fig. 1; the amino acid composition calculated from the sequence is in good agreement with that found by composition analysis of the total protein (Table 1). The *N*-terminal 40 residues of the protein have been determined by using an automated Beckman 890C protein sequencer; the *C*-terminal sequence of the protein was found to be Leu-Thr-Ala after digestion with carboxypeptidase A for different periods of time.

Ten peptides, C-1 to C-10, were isolated from a chymo-

Table 1. The amino acid composition of *Sambucus nigra* ferredoxin

	Mean values 24 hr hydrolysis	Mean values 48 hr hydrolysis	Mean values 72 hr hydrolysis	Average values	Sequence values
Asx	9.93	9.02	9.26	9.40	9
Thr	5.61	5.27	5.01	5.11*	5
Ser	8.74	7.90	7.35	9.14*	9
Glx	16.91	15.81	15.78	16.17	16
Pro	5.65	5.20	4.92	5.26	5
Gly	6.00	6.02	5.87	5.96	6
Ala	7.03	7.16	7.09	7.09	7
Val	6.42	6.61	7.01	7.01†	7
Ile	5.37	5.71	6.49	6.49†	6
Leu	6.81	6.76	6.82	6.82†	7
Tyr	3.29	4.08	4.06	3.81	4
Phe	2.07	2.08	2.02	2.06	2
His	2.24	2.05	2.09	2.13	2
Lys	4.58	4.58	5.00	4.72	5
Arg	1.00	1.03	1.00	1.01	1
Cm-Cys	5.11	5.02	5.14	5.09	5
Trp	—	—	—		1

* Calculated from 24 and 72 hr values assuming first order kinetics for destruction [5].
† For valine, isoleucine and leucine, maximal values (72 hr hydrolysis) were taken.

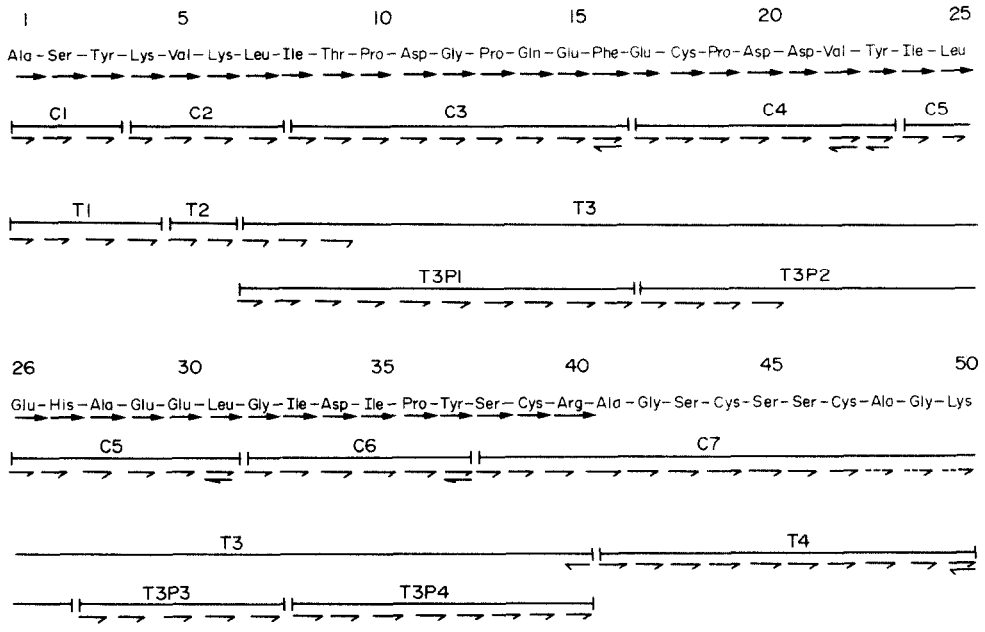


Fig. 1. Amino acid sequence of *Sambucus nigra* ferredoxin. Arrows (→), (⇀), (←) represent automatic sequencing, manual Edman degradation and carboxypeptidase digestion, respectively. Dashed arrows (⇢) indicate that the residue could not be unambiguously identified. T, C and P represent peptides obtained from tryptic, chymotryptic and papain digestion, respectively.

tryptic digest of carboxymethylated ferredoxin. All these peptides were completely sequenced by Edman degradation except C-7 and C-10 whose sequences were only partially determined. C-10 must be the *C*-terminal peptide of the protein, since its *C*-terminal sequence, Leu-Thr-Ala, was identical with that of the intact protein. The order of chymotryptic peptides, C 1–C 7, was established from overlap by the *N*-terminal sequence determined automatically and by tryptic peptides: that of C 8–C 10 by overlapping with tryptic peptides.

Digestion of the protein with trypsin produced six peptides (see Fig. 1) of which T-3 and T-5 were purified by gel chromatography of the mixture and the others by paper electrophoresis. The sequence of T-3 was determined after digestion of the peptide by papain; this enzyme produced four peptides which were separated by paper electrophoresis and their sequences were established by Edman degradation. The order of papain peptides in the parent peptide T-3 could be established since they were overlapped by the *N*-terminal sequence. The remaining tryptic peptides were completely sequenced by Edman degradation except peptide T-5 which was partially sequenced. Their order was established from overlap by chymotryptic peptides except in position 50–51.

All the acidic and amide residues indicated in the sequence were placed from the electrophoretic mobilities at pH 6.5 (Table 2) of the intact peptides except residues 57, 58, 60, 61 and 68 which were placed as Asp, Gln, Asp, Gln, Gln, respectively by homology with the other chloroplast-type ferredoxins from which amino acid sequence data are available. The assignments of amides and acid residues in the *N*-terminal 40 residues agreed with direct identification by their phenylthiohydantion derivatives.

The amino acid compositions of tryptic and chymotryptic peptides are given in Table 3 and the sequence data of these peptides are shown in Fig. 1.

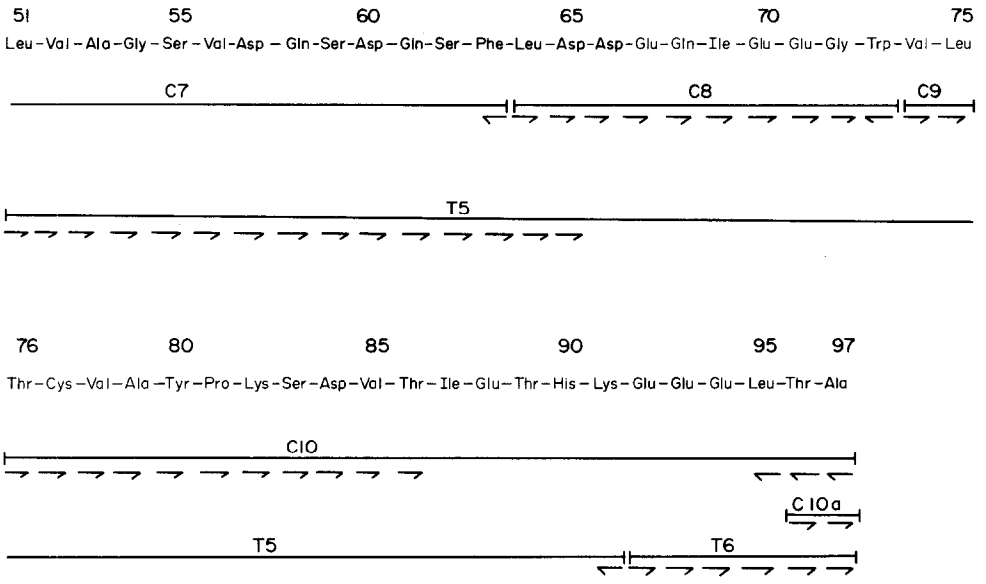
Table 2. Electrophoretic mobilities of peptides at pH 6.5 were measured from the neutral amino acids relative to aspartic acid and at pH 1.9 measured from 1-dimethylaminonaphthalene-5-sulphonic acid relative to dansyl arginine

Peptide	Mobility at pH 6.5	Mobility at pH 1.9
C1	0	0.58
C2	0.87	1.39
C3	−0.49	0.20
C4	−0.60	—
C5	−0.50	—
C6	−0.24	0.28
C7	—	—
C8	−0.81	—
C9	0	0.63
C10	−0.10	—
T1	0.40	0.85
T2	0.70	1.33
T3	—	—
T3P1	−0.54	—
T3P2	−0.46	—
T3P3	−0.55	—
T3P4	−0.20	—
T4	0	—
T5	—	—
T6	−0.73	0.26

— = n.d.

DISCUSSION

The ferredoxin of *Sambucus nigra* consists of 97 amino acids as compared to other plant/algal ferredoxins which vary in length from 96 to 98 residues; all higher plants,



however, have 97 residues. Its pI is very similar to those of *Hordeum vulgare*, 3.7 (Takruri and Boulter, unpublished); *Porphyra umbilicalis*, 3.8 [6] and *Clostridium pasteurianum*, 3.7 [7]. The MW deduced from the sequence is 11033 including the active centre. The sequence contains five cysteine residues, four of which bind the iron atoms of the active centre and are identical in position to those of other ferredoxins. All residues were positively identified except residues 87–90 which were placed from the amino acid composition of peptide C 10 and by homology of other higher plant ferredoxins. Acids and amides in positions 57, 58, 60, 61 and 68 were also placed by homology of other ferredoxin sequences. The molecule was found to be lacking in methionine.

Table 4 gives the number of differences which exist between various plant ferredoxins. *Medicago sativa* ferredoxin shows the greatest sequence similarity to the *Sambucus* ferredoxin, followed by *Leucaena glauca*, then *Colocasia esculenta*; similarity between the amino acid composition of these species and that of *Sambucus racemosa* has also been reported [8].

EXPERIMENTAL

Materials. Sambucus nigra leaves were collected locally. Other materials were as described in ref. [3]. Ferredoxin was isolated and purified as described in ref. [9].

Methods. Isoelectric focusing was carried out on native ferredoxins as described in ref. [10] in gels containing 5% (w/v) polyacrylamide, 2% (w/v) Ampholine (pH 3–5) and 6 M urea. After focusing, the gels were fixed and the protein detected as a white precipitation band by immersion in 25% (w/v) trichloroacetic acid. The isoelectric point (pI) was determined by measuring the pH gradient in similar unfixed gels.

Table 3. Amino acid composition of peptides from *Sambucus nigra* ferredoxin

Peptide	Asx	Thr	Ser	Glx	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	His	Arg	Lys
C1			0.90				1.12				0.85				
C2								0.82		1.12					1.81
C3	0.79	0.59		1.70	2.22	0.72	0.60		0.64			1.39			
C4	2.07			0.73	1.36			0.99			0.40				
C5				3.07			0.84		0.83	1.84			0.92		
C6	1.40				1.01	0.73			1.90		0.39				
C8	3.13			2.97		1.06			0.90	0.74					
C9								0.99		0.98					
C10	0.55	3.95	0.48	4.23	1.00		1.89	1.72	0.85	1.21	0.70		1.06		1.86
C10a		0.66					1.00								
T1			1.01				0.87				1.13				0.82
T2								1.13							0.87
T4			2.67			2.39	1.60								0.92
T6		0.96		2.98			1.60								

Table 4. Matrix of amino acid sequence difference of plant ferredoxins

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. *Colocasia esculenta*	0												
2. *Leucaena glauca*	20	0											
3. *Spinacia oleracea*	18	20	0										
4. *Medicago sativa*	16	23	19	0									
5. *Triticum aestivum*	20	20	21	19	0								
6. *Sambucus nigra*	17	16	21	13	19	0							
7. *Phytolacca americana* (I)	28	35	30	27	32	29	24	0					
8. *Phytolacca americana* (II)	29	37	31	27	34	31	31	23	0				
9. *Equisetum telmateia* (I)	36	36	39	39	32	38	34	35	39	0			
10. *Equisetum telmateia* (II)	43	39	45	45	42	40	44	43	47	29	0		
11. *Equisetum arvense* (I)	37	37	40	40	34	37	35	36	40	1	30	0	
12. *Equisetum arvense* (II)	44	40	46	46	44	46	44	44	48	30	1	31	0

Amino acid composition and sequence. Ferredoxin was denatured in 6 M guanidine chloride, reduced and carboxymethylated [11]. Amino acid composition of CM-ferredoxin was obtained after hydrolysis for 24, 48, 72 hr with 6 N HCl at 105° in sealed, evacuated tubes. Duplicate hydrolyses were carried out for each hydrolysis time. Digestion of carboxymethylated protein by trypsin or chymotrypsin was carried out as in ref. [3] and T-3 was digested with papain, (see [12]). Carboxypeptidase A or B was made soluble by method 1 of ref. [13], and digestion was with a 2% (w/w) ratio of enzyme to substrate for periods of up to 2 hr in 0.2 M *N*-ethylmorpholine acetate buffer, pH 8.5. Peptides resulting from tryptic and chymotryptic digestions were applied to a Bio-Gel P-4 column (1.5 × 190 cm). Elution profiles were followed by measurement of A_{280} and A_{206} with an LKB Uvicord III instrument and by *N*-terminal analysis using the dansyl technique [14]. Tryptic and chymotryptic peptides, which remained impure after gel chromatography, and papain peptides were subjected to high voltage paper electrophoresis at pH 6.5 and 1.9 [15]. Peptides were detected on paper by using cadmium ninhydrin reagent [16]. Peptide sequences and electrophoretic mobilities were determined as in ref. [15]. Automated sequencer analysis was carried out on a Beckman model 890C automatic protein sequencer and phenylthiohydantoin (PTH) derivatives were identified as described by ref. [17].

REFERENCES

1. Hall, D. O., Rao, K. K. and Cammack, R. (1975) *Sci. Prog. Oxf.* **62**, 285.
2. Hase, T., Wakabayashi, S., Matsubara, H., Rao, K. K., Hall, D. O., Widmer, H., Gysi, J. and Zuber, H. (1978) *Phytochemistry* **17**, 1863.
3. Takruri, I., Haslett, B. G., Boulter, D., Andrew, P. W. and Rogers, L. J. (1978) *Biochem. J.* **173**, 459.
4. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure*, Vol. 5. National Biomedical Research Foundation, Washington, D.C.
5. Moore, S. and Stein, H. W. (1963) *Methods Enzymol.* **6**, 817.
6. Andrew, P. W., Rogers, L. J., Boulter, D. and Haslett, B. G. (1976) *Eur. J. Biochem.* **69**, 243.
7. Lovenberg, W., Buchanan, B. B. and Rabinowitz, J. C. (1963) *J. Biol. Chem.* **238**, 3899.
8. Altosaar, I., Bohm, B. A. and Taylor, I. E. P. (1977) *Can. J. Biochem.* **55**, 159.
9. Petering, D. H. and Palmer, G. (1970) *Arch. Biochem. Biophys.* **141**, 456.
10. Wrigley, C. (1968) *Sci. Tools.* **15**, 17.
11. Milne, P. R. and Wells, J. R. E. (1970) *J. Biol. Chem.* **245**, 1566.
12. Scawen, M. D., Ramshaw, J. A. M. and Boulter, D. (1975) *Biochem. J.* **147**, 343.
13. Ambler, R. P. (1972) *Methods Enzymol.* **25**, 143.
14. Gray, W. R. and Hartley, B. S. (1963) *Biochem. J.* **89**, 378.
15. Thompson, E. W., Laycock, M. W., Ramshaw, J. A. M. and Boulter, D. *Biochem. J.* **117**, 183.
16. Heilman, J., Barrollier, J. and Watzke, E. (1957) *Hoppe-Seyler's Z. Physiol. Chem.* **309**, 219.
17. Haslett, B. G. and Boulter, D. (1976) *Biochem. J.* **153**, 33.