The Amino Acid Sequence of Ferredoxin from *Triticum aestivum* (Wheat)

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

(Received 10 October 1978)

The amino acid sequence of the ferredoxin of *Triticum aestivum* (wheat) was determined by using a Beckman 890C sequencer in combination with the dansyl–phenylisothiocyanate method to characterize peptides obtained by tryptic, chymotryptic and thermolytic digestion of CNBr-cleavage fragments. The molecule consists of a single polypeptide chain of 97 residues and has an unblocked *N*-terminus. It shows considerable similarity to other plant-type ferredoxins.

Ferredoxins have been isolated and characterized from various bacteria, algae and higher plants. They are small proteins with a mol.wt. of approx. 6000 or 11000 depending on the source. Those from higher plants, eukaryotic algae and blue-green algae (blue-green bacteria) possess the 2Fe–2S active centre and have a mol.wt. of approx. 11000, whereas bacterial ferredoxins contain two 4Fe–4S active centres and have a mol.wt. of approx. 6000. These non-haem iron-containing proteins exhibit characteristic u.v. and visible absorption spectra, have negative redox potentials and show a characteristic e.p.r. signal in the reduced state. They act as electron carriers in a number of different biochemical processes, including photosynthetic electron transport and nitrogen fixation.

The amino acid sequence of *Triticum aestivum* (wheat) ferredoxin shows great similarity to those of other ferredoxins that have been isolated from higher plants and algae and sequenced (Takruri *et al.*, 1978; Wakabayashi *et al.*, 1978).

Experimental

Materials

Triticum aestivum (wheat) leaves were obtained from a local farmer's field during Spring 1977 and used immediately. Trypsin (EC 3.4.21.4) (1-chloro-4-phenyl-3-L-toluene-*p*-sulphonamidobutan-2-one-treated), and α -chymotrypsin (EC 3.4.21.1) were from Worthington Biochemical Corp., Freehold, NJ, U.S.A. Thermolysin (EC 3.4.24.4) (crystalline) was from Daiwa Kasei KK, Osaka, Japan. Carboxypeptidase A (di-isopropyl phosphorofluoridate-treated) (EC 3.4.12.2) and carboxypeptidase B, from Sigma (London) Chemical Co., London, S.W.6, U.K. Acrylamide and *NN'*-methylenebisacrylamide were obtained from BDH Chemicals Ltd., Poole, Dorset, U.K.; 40% Ampholyte carrier (pH 3–5) was obtained from Serva, Heidelberg, Germany.

Methods

Ferredoxin was isolated and purified as described by Petering & Palmer (1970) and carboxymethylated as described by Milne & Wells (1970) and Takruri *et al.* (1978). The amino acid analyses and peptide purification were as described in Takruri *et al.* (1978).

Isoelectric focusing. Isoelectric focusing was carried out on native and carboxymethylated ferredoxins as described by Wrigley (1968) in gels containing 5% (w/v) polyacrylamide, 2% (w/v) Ampholine (pH 3–5) and 6M-urea. After focusing, the gels were fixed and the protein detected as a white precipitation band(s) by immersion in 25% (w/v) trichloroacetic acid. The isoelectric point (pI) was determined by measuring the pH gradient in similar unfixed gels.

Cleavage methods. CNBr cleavage was carried out by the method of Steers *et al.* (1965). Enzymic digestions were as described in Takruri *et al.* (1978).

Amino acid-sequence methods. Peptide sequences were determined by using the method of Gray & Hartley (1963) as described by Thompson *et al.* (1970). Separation of dansyl-amino acids on polyamide sheets (Woods & Wang, 1967) was performed by using the solvent systems described by Ramshaw *et al.* (1970). After digestion with carboxypeptidase A or B, released C-terminal amino acids were identified as their dansyl derivatives.

Nomenclature. The CNBr fragments are numbered on the basis of their position in the complete sequence. Peptides derived from digestion of the CNBr fragments, or a large peptide, are numbered on the basis of their order within the parent fragment. The following abbreviations are used below: X, CNBr fragments; H, thermolytic peptides; C, chymotryptic peptides; T, tryptic peptides.

Results

T. aestivum ferredoxin was isolated in a high degree of purity. The homogeneity of the protein was shown

by isoelectric focusing, when the protein migrated as a single band to a position corresponding to a pI value of 3.9. Carboxymethylated ferredoxin gave two bands (pI values, 4.1 and 3.9) in isoelectric-focusing gels.

The amino acid composition of *T. aestivum* together with that of the CNBr fragments is shown in Table 1. Composition of the fragments shows good agreement with that of the protein. The C-terminal sequence of the protein was found to be -Leu-Thr-Ala.

Cleavage with CNBr produced two fragments, which were purified by gel chromatography. The larger fragment, X1, had the same N-terminal residue as the protein, and had a homoserine residue at the C-terminus. Fragment X2 contained no homoserine and had a C-terminal residue identical with that of the intact protein. The order of the fragments in the intact protein could therefore be unequivocally established.

The N-terminal 48 residues of the protein have been determined with an automated Beckman 890C protein sequencer by using the 'fast protein' program supplied with the instrument (Beckman program no. 072172C). Phenylthiohydantoin derivatives were identified as described by Haslett & Boulter (1976). Digestion of the X1 fragment with trypsin produced five peptides (see Fig. 1 and Table 2) of which the largest, T3, was purified by gel chromatography of the mixture,

and others by paper electrophoresis. The order of tryptic peptides in the parent fragment X1 could be established from these data, since peptides T1, T2, T3 and T4 were overlapped by that N-terminal sequence. The remaining tryptic peptide, T5, which has homoserine as the C-terminal residue, could also be positioned in the X1 fragment. Furthermore, all these tryptic peptides were overlapped by chymotryptic peptides (see below). The sequence of peptides T1, T2, T4 and T5 were established by direct Edman degradation. The sequence of peptide T3 was determined after digestion of the peptide by thermolysin, which produced nine peptides that were separated by paper electrophoresis and their sequences established by Edman degradation. Their order was established from overlap by chymotryptic peptides. All acidic residues of fragment X1, except those of T5 tryptic peptides, were established as such by direct identification of phenylthiohydantoin derivatives. These assignments agreed with the electrophoretic mobility of the constituent peptides (see Table 2). The amide residues in peptide T5 were determined by the electrophoretic mobility of the two peptides C6 and C7. The positions of amides within peptide X1C6 were established from changes in its electrophoretic mobility after successive cycles of Edman degradation (see Fig. 1). The position of the amide group in

Table 1. Amino acid composition of ferredoxin of *Triticum aestivum* and its CNBr fragments

The results are expressed as residues/molecule. Cysteine was measured as cysteic acid (Hirs, 1956). Tryptophan was determined after alkaline hydrolysis for 72h (Noltmann *et al.*, 1962). Average values ($\pm 5\%$) are given, except for serine and threonine, which are extrapolated to zero time where possible, assuming first-order rate of destruction (Moore & Stein, 1963), and for valine and isoleucine, where maximal values (72h hydrolysis) are given. Abbreviations used: n.d., not determined; Hse, homoserine; Seq, values from sequence determination.

Amino acid	Protein					X1 fragment				X2 fragment				
	24h	48h	72h	Average	Seq	24h	72h	Average	Seq	24h	72h	Average	Seq	
Asp	10.35	11.13	9.96	10.48	10	9.39	8.95	9.17	9	1.31	1.39	1.35	1	
Thr	5.13	5.02	4.82	5.30	5	2.20	2.13	2.24	2	3.01	2.91	3.06	3	
Ser	6.89	6.35	5.68	7.61	8	5.98	5.20	6.42	7	0.92	0.91	0.93	1	
Hse	0	0	0	0	0	0.55	0.50	0.52	1	0	0	0	0	
Glu	17.32	17.41	17.22	17.32	17	12.29	12.41	12.35	12	4.72	4.95	4.84	5	
Pro	3.66	4.21	4.00	3.96	4	3.02	2.81	2.90	3	0.89	0.90	0.89	1	
Gly	6.26	6.60	6.20	6.35	6	5.31	5.32	5.31	5	0.82	0.92	0.87	1	
Ala	6.88	6.98	6.67	6.84	7	4.20	4.24	4.22	4	2.81	2.98	2.89	3	
Cys	4.33	n.d.	n.d.	4.33	4	n.d.	n.d.	n.d.	3	1.33	1.23	1.28	1	
Val	7.41	7.64	8.11	8.11	8	5.50	5.71	5.71	6	1.79	1.98	1.98	2	
Met	0.94	0.77	0.76	0.82	1	0	0	0	0	0	0	0	0	
Ile	4.51	4.66	4.71	4.71	5	2.91	3.20	3.06	3	1.69	1.87	1.87	2	
Leu	7.57	8.14	8.47	8.47	8	5.88	6.31	6.10	6	2.06	2.12	2.09	2	
Tyr	3.92	3.95	3.89	3.92	4	2.66	2.52	2.59	3	0.63	0.51	0.57	1	
Phe	0.95	0.91	0.88	0.91	1	0.98	0.92	0.95	1	0	0	0	0	
His	1.85	1.81	1.73	1.80	2	0	0	0	0	1.68	1.79	1.74	2	
Lys	4.81	4.82	4.53	4.72	5	3.33	3.16	3.16	3	2.00	2.06	2.03	2	
Arg	0.93	0.91	0.87	0.91	1	0.95	0.95	0.95	1	0	0	0	0	
Trp	n.d.	n.d.	0.93	0.93	1	n.d.	n.d.	n.d.	0	n.d.	n.d.	n.d.	1	
Total residues					97					69				28

Table 2. *Mobilities and amino acid compositions of peptides from T. aestivum ferredoxin*

The mobilities of peptides during electrophoresis at pH 6.5 were measured from dansyl(5-dimethylaminonaphthalene-1-sulphonyl)-arginine relative to aspartic acid and at pH 1.9 from 5-dimethylaminonaphthalene-1-sulphonic acid relative to dansyl-arginine. Amino acids present in less than 0.3 residue were disregarded; this indicates that some peptides were not completely pure. Abbreviations used: Hse, homoserine; CmCys, carboxymethylcysteine; —, values not determined. Peptides not analysed for amino acid composition were characterized by direct dansyl-Edman analysis.

Peptides	Mobilities		Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	His	Arg	Lys	Hse	CmCys	Total residues
	pH6.5	pH1.9																		
X1T1	0.46	0.93	0.85						1.09				0.62				0.82			4
X1T2	0.67	1.44								0.61							0.83			2
X1T3	-0.70	0.26	4.45	0.90	1.30	8.03	2.75	2.13	1.09	3.65	1.77	3.66	1.66			0.75			0.91	34
X1T3H1	-0.48	0.28																		
X1T3H2	-0.63	0.61																		
X1T3H3	-0.80	0.32																		
X1T3H3a	-0.63	0.61																		
X1T3H3b	-0.72	0.41																		
X1T3H4	0	0.61																		
X1T3H5	-0.80	0.24																		
X1T3H6	-0.62	0.60																		
X1T3H7	0.30	0.58																		
X1T4	-0.27	0.50						2.11	2.02								1.07	0.43	(2)	10
X1T5	-0.61	0.16	4.75		2.63	4.41		1.24		1.03	0.91	1.76	1.03	0.72						19
X1C1	0	0.55		1.10				1.04		0.72		0.56					1.70			3
X1C2	0.94	1.40																		4
X1C3	-0.75	0.20																		
X1C3a	-0.66	0.25																		
X1C4	-0.73	0.21																		
X1C5	0.22	0.38						2.00	2.00		2.05	2.01	0.90			0.89	0.88		2.60	14
X1C6	-0.57	0.20			3.88															
X1C7	-0.73	0.32																		
X2T1	0	0.40																		
X2T2	0	0.51	1.17	0.79	0.62	0.95				0.99	2.15	1.19	1.09		0.77	1.04	0.55		0.66	13 (Trp)
X2T3	-0.81	0.30		0.78		3.11			0.99			0.95					1.33			9
X2C1	-0.45	0.45				1.05		0.98	0.88											6
X2C2	0	0.70								1.05		0.78			1.00				—	4 (Trp)
X2C2a	0	0.57		1.22						0.93		0.99							—	2
X2C2b	0	0.80		0.58										1.00					—	2
X2C3	-0.24	0.48																		
X2C4	0	0.85		0.95					1.02											2

Peptides not analysed for amino acid composition were characterised by direct change and Edman analysis.

Table 3. Matrix of amino acid-sequence differences for representative plant-type ferredoxins I and II refer to two different ferredoxin types.

Organism																
Type	Common name	Latin name														
Higher plants	Wheat	<i>Triticum aestivum</i>	0	21	0											
	Spinach	<i>Spinacea oleracea</i>	19	19	0											
	Alfalfa	<i>Medicago sativa</i>	20	18	16	0										
	Taro	<i>Colocasia esculenta</i>	20	20	23	20	0									
	Koa	<i>Leucaena glauca</i>	20	34	40	40	37	36	1	0						
Horsetails	—	<i>Equisetum arvense</i> I	32	39	39	36	36	36	1	0						
	—	<i>Equisetum telmateia</i> I	44	46	46	44	40	31	30	0						
	—	<i>E. arvense</i> II	42	45	45	43	39	30	29	1	0					
	—	<i>E. telmateia</i> II	30	30	29	26	33	31	30	43	42	0				
	—	<i>Scenedesmus</i>	30	33	34	37	34	41	41	41	40	29	0			
Blue-green algae	—	<i>Aphanothece</i>	33	34	37	36	39	41	40	45	44	25	30	0		
	—	<i>Spirulina maxima</i>	27	34	34	36	32	39	38	41	40	28	26	21	0	
Red alga	—	<i>Nostoc muscorum</i>	37	37	37	37	40	39	38	41	43	31	41	28	37	0
	—	<i>Porphyra umbilicalis</i>														

peptide X1C7 was determined by carboxypeptidase A digestion, since the mobility of this peptide indicated the presence of only one amide group.

The X2 fragment was cleaved into three peptides by trypsin. The order of peptides X2T1, X2T2 and X2T3 in the parent fragment X2 could be partially established from these data, since peptide T1 had the same three *N*-terminal residues as those of fragment X2, and the sequence of the *C*-terminal two residues of peptide T3 was identical with those of fragment X2. Furthermore, all these tryptic peptides were overlapped by chymotryptic peptides. The sequences of these peptides were established by direct Edman degradation.

Discussion

The ferredoxin of *T. aestivum* contains 97 amino acids as compared with other plant/algal ferredoxins, which contain 96–98 amino acids. Its pI is very similar to those of barley (*Hordeum vulgare*), 3.75 (I. Takruri & D. Boulter, unpublished work), the blue-green alga *Anabaena flosaquae*, 3.7 (Andrew *et al.*, 1975), the alga *Porphyra umbilicalis*, 3.7 (Andrew *et al.*, 1976) and the bacterium *Clostridium pasteurianum* 3.7 (Lovenberg *et al.*, 1963). The molecular weight deduced from the sequence is 10801, including the active centre. *T. aestivum* ferredoxin contained only four cysteine residues, which are the invariant residues required for the binding of the two iron atoms involved in the catalytic site of the protein. They are located at positions 39, 44, 47 and 77. The molecule also contained one tryptophan and one methionine residue and it was found to be lacking in asparagine.

Table 3 gives the numbers of differences that exist

between various plant-type ferredoxins. *Medicago sativa* (alfalfa) ferredoxin shows the greatest sequence similarity to the *T. aestivum* ferredoxin, followed by *Colocasia esculenta* (taro) and *Leucaena glauca* (koa).

References

- Andrew, P. W., Delaney, M. E., Rogers, L. J. & Smith, A. J. (1975) *Phytochemistry* **14**, 931–935
- Andrew, P. W., Rogers, L. J., Boulter, D. & Haslett, B. G. (1976) *Eur. J. Biochem.* **69**, 243–248
- Gray, W. R. & Hartley, B. S. (1963) *Biochem. J.* **89**, 378–380
- Haslett, B. G. & Boulter, D. (1976) *Biochem. J.* **153**, 33–38
- Hirs, C. H. W. (1956) *J. Biol. Chem.* **219**, 611–621
- Lovenberg, W., Buchanan, B. B. & Rabinowitz, J. C. (1963) *J. Biol. Chem.* **238**, 3899–3913
- Milne, P. R. & Wells, J. R. E. (1970) *J. Biol. Chem.* **245**, 1566–1574
- Moore, S. & Stein, H. W. (1963) *Methods Enzymol.* **6**, 817–831
- Noltmann, E. A., Mahowald, T. A. & Kirby, S. A. (1962) *J. Biol. Chem.* **237**, 1146–1154
- Petering, D. H. & Palmer, G. (1970) *Arch. Biochem. Biophys.* **141**, 456–464
- Ramshaw, J. A. M., Thompson, E. W. & Boulter, D. (1970) *Biochem. J.* **119**, 535–539
- Steers, E., Craven, G. R. & Anfinsen, C. B. (1965) *J. Biol. Chem.* **240**, 2478–2484
- Takruri, I., Haslett, B. G., Boulter, D., Andrew, P. W. & Rogers, L. J. (1978) *Biochem. J.* **173**, 459–466
- Thompson, E. W., Laycock, M. W., Ramshaw, J. A. M. & Boulter, D. (1970) *Biochem. J.* **117**, 183–192
- Wakabayashi, S., Hase, T., Wada, K., Matsubara, H., Suzuki, K. & Takaichi, S. (1978) *J. Biochem. (Tokyo)* **83**, 1305–1319
- Woods, W. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta* **133**, 369–370
- Wrigley, C. (1968) *Sci. Tools* **15**, 17–23