Phytochemistry, Vol. 21, No. 2, pp. 325–327, 1982.
Printed in Great Britain.

0031-9422/82/020325-03$03.00/0

AMINO ACID SEQUENCE OF FERREDOXIN FROM *ARCTIUM LAPPA*

ISHAQ A. H. TAKRURI, JOHN GILROY and DONALD BOULTER

Department of Botany, University of Durham, Durham DH1 3LE, U.K.

(*Received* 5 *May* 1981)

Key Word Index—*Arctium lappa*; Compositae; burdock; amino acid sequence; ferredoxin.

Abstract—Ferredoxin from *Arctium lappa* consists of a single polypeptide chain of 97 residues, four of which are cysteine. These residues, which are in the active centre, are in identical positions to those of other ferredoxins. Overlap between residue positions 50/51 was not obtained, but amino acid composition of the two cyanogen bromide fragments which were overlapped corresponded with the amino acid composition of the total protein.

INTRODUCTION

The amino acid sequence of *Arctium* ferredoxin presented in this paper is very similar to those of other 2Fe–2S ferredoxins which have been isolated and sequenced from algae and higher plants [1, 2]. It was determined as part of a programme to construct an affinity tree relating the evolution of higher plants [3].

RESULTS

Arctium (burdock) ferredoxin was isolated and purified to give an absorption index A_{420}/A_{280} of 0.50. The complete amino acid sequence is shown in Fig. 1 and the amino acid composition calculated from this is in general agreement with that found by compositional analysis of the total protein (Table 1). The *N* and *C*-terminal residues of the protein were shown to be Ala, which was preceded at the *C*-terminus region by Thr and Leu. Cyanogen bromide cleavage of the protein gave two fragments called X1 and X2, although a cleavage in low yield occurred at Try, position 73. The X1 fragment had an *N*-terminus of Ala and a *C*-terminus of homoserine, and the X2 fragment had Glu at its *N*-terminus and Ala at the *C*-terminus.

A tryptic digest of the total protein gave six peptides in accordance with the number expected from the four Lys and one Arg residues determined by amino acid analysis. However, a peptide from residue 14 to residue 40, i.e. expected T3, was not isolated, but an additional peptide from residue 24 to residue 40 was found. This peptide was subsequently shown, by evidence from the chymotryptic digest, to have been the result of chymotryptic activity in the trypsin preparation used, i.e. there was cleavage at Tyr 23. This does not explain the absence of T3 in the tryptic digest, but the reason for that finding became clear when subsequent sequence analysis using the equivalent chymotryptic peptide showed that position 14 was Gln and probably cyclised, so blocking the *N*-terminus of the T3 peptide. A large peptide separated by molecular sieving of the tryptic digest was sub-digested with thermolysin and this yielded four useful peptides designated T6 H1–T6 H4, which when sequenced established overlap at the X1, X2 junction.

A tryptic digest was also carried out on the XI CNBr fragment; this yielded seven peptides, instead of the expected five, due to some chymotryptic cleavages. Peptide X1 T3 was sequenced to 14 resi-

Table 1. Amino acid composition of burdock ferredoxin*

	Amino acid composition values	Sequence values
CM-Cys	5	4
Asx	9	10
Thr	9	9
Ser	8	8
Glx	14	13
Pro	4	4
Gly	6	7
Ala	9	9
Val	9	9
Met	1	1
Ile	3	3
Leu	6	6
Phe	1	1
Tyr	5	5
His	2	2
Lys	4	4
Arg	1	1
Trp	—	1

*Calculated to nearest whole number from values obtained in 24 and 72 hr hydrolysates [4], Thr and Ser assuming first order kinetics for destruction; Ile, Val and Leu (72 hr value), others from the average values.

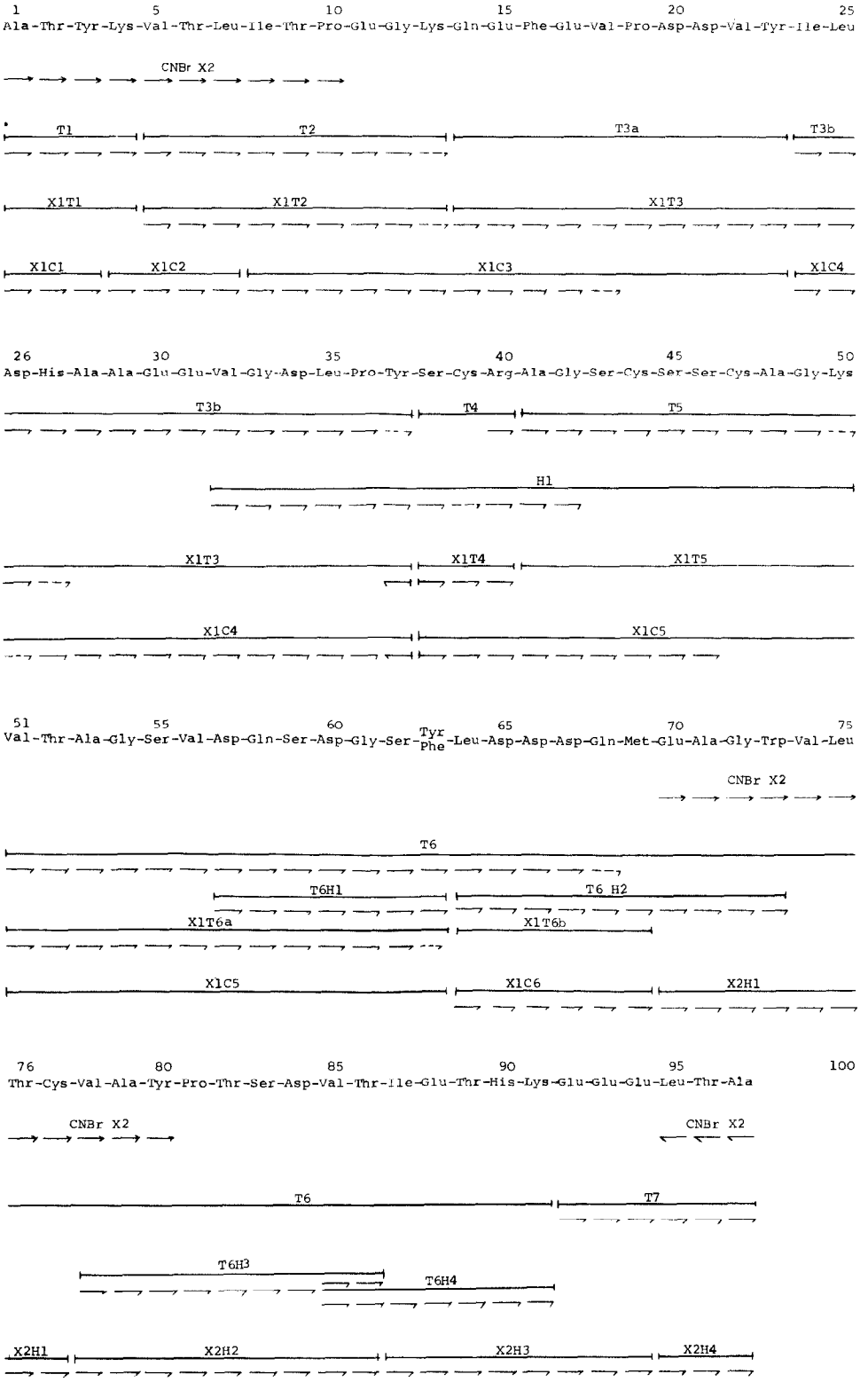

Fig. 1. Amino acid sequence of *Arctium lappa* ferredoxin. Arrows (→), (←) represent residues determined by the DABITC method and carboxypeptidase digestion respectively. Dashed arrows (– – →) indicate that the residue could not be unambiguously identified. T, C and H represent peptides obtained from tryptic chymotryptic and thermolytic digestion, respectively. X1 and X2 are CNBr fragments.

dues; X1 T1, X1 T5 and X1 T7 were not further analysed as their sequences had already been determined from the previous tryptic digest. The sequence of X1 T4 was determined and this explains the original failure to sequence T4 to its expected conclusion, i.e. due to an anomalous chymotryptic cleavage at Tyr 37. However, the thermolysin digest which yielded peptides T6 H1–T6 H4 also contained several other peptides of which H1 was sequenced so overlapping residue 37/38.

The X1 CNBr fragment was also subjected to chymotryptic digestion and yielded six peptides. Subsequent sequence analysis provided the evidence mentioned above to explain the results obtained with the tryptic digest. Two of the chymotryptic peptides, i.e. X1 C3 and X1 C5, were not sequenced to completion, in spite of several attempts to do so. This was not of great consequence in the case of X1 C3 since although there was no overlap between 37 and 38 this was established from the thermolytic digest. The failure to determine the complete sequence of X1 C5, however, means that there is no overlap of the T5 and T6 peptides, i.e. at residues 50, 51. The X2 CNBr fragment was subjected to thermolytic digestion which yielded four peptides; these were all completely sequenced providing overlaps of peptides from residue 70 to the *C*-terminus.

Both Tyr and Phe were found at position 63. A conservative difference of this nature in a molecule would not however normally lead to the separation of the two parent molecules by the usual purification techniques. Although peptides T5 and T6 were not overlapped, we are confident that the complete sequence has been determined. Firstly, it has the same number of residues as the other higher plant ferredoxins, where complete information is available, i.e. 97 residues. Secondly, we have made total hydrolysates of the peptides and dansylated them, and their amino acid compositions agree with that expected from the 97 established residues (data not shown). Lastly, the amino acid composition of the CNBr fragments, X1 and X2, which were overlapped, corresponded with the amino acid composition of the total protein.

When wheat, rape, spinach, taro, alfalfa, elder, koa and burdock ferredoxin sequences were analysed by a total search ancestral method (T. Gleaves, unpublished), the burdock sequence was most closely associated with wheat in a unique tree. The affinity tree has not been presented, however, as its significance cannot be assessed as so few ferredoxin sequences are available, the trees would not be stable on the addition of new sequences; 20 out of a total of 67 substitutes are convergent.

EXPERIMENTAL

Materials. *Arctium lappa* L. leaves were collected locally. Other materials were as described in ref. [2] and ferredoxin was isolated and purified as in ref. [5] except that diethylammonium diethyl-dithiocarbamate was added to the extraction buffer in order to reduce oxidation reactions and after $(NH_4)_2SO_4$ fractionation the dialysis step was omitted and the DE-23 cellulose column was equilibrated in 90% $(NH_4)_2SO_4$.

Amino acid composition and sequence. Ferredoxin was denatured in 6 M guanidine chloride, reduced and carboxymethylated [6]. Amino acid composition of CM-ferredoxin was obtained after hydrolysis for 24 or 72 hr with 6 M 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. [2] and T6 was digested with thermolysin. Carboxypeptidase A or B was made soluble by method 1 of ref. [7] 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} and by *N*-terminal analysis using the dansyl technique [8]. Tryptic and chymotryptic peptides, which remained impure after gel chromatography and thermolytic peptides were subjected to high voltage paper electrophoresis at pH 6.5 and 1.9 [9]. Peptides were detected on paper using cadmium ninhydrin reagent [10].

CNBr cleavage was as in ref. [11] and thermolysin digestion as in ref. [12]. Peptide sequences were determined by the DABITC method [13].

REFERENCES

1. Hase, T., Wakabayashi, S., Matsubara, H., Rao, K. K., Hall, D. O., Widmer, H., Gysi, J. and Zuber, H. (1978) *Phytochemistry* **17**, 1863.
2. Takruri, I., Haslett, B. G., Boulter, D., Andrew, P. W. and Rogers, L. J. (1978) *Biochem. J.* **173**, 459.
3. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure*, Vol. 5. National Biomedical Research Foundation, Washington, D.C.
4. Moore, S. and Stein, H. W. (1963) *Methods Enzymol.* **6**, 817.
5. Petering, D. H. and Palmer, G. (1970) *Arch. Biochem. Biophys.* **141**, 456.
6. Milne, P. R. and Wells, J. R. E. (1970) *J. Biol. Chem.* **245**, 1566.
7. Ambler, R. P. (1972) *Methods Enzymol.* **25**, 143.
8. Gray, W. R. and Hartley, B. S. (1963) *Biochem. J.* **89**, 378.
9. Thompson, E. W., Laycock, M. W., Ramshaw, J. A. M. and Boulter, D. (1970) *Biochem. J.* **117**, 183.
10. Heilman, J., Barrollier, J. and Watzke, E. (1957) *Hoppe-Seyler's Z. Physiol. Chem.* **309**, 219.
11. Steers, Jr. E., Craven, G. R., Anfinsen, C. B. and Bethune, J. L. (1965) *J. Biol. Chem.* **240**, 2478.
12. Heinrikson, R. L. *Methods Enzymol.* **47**, 175.
13. Chang, J. Y., Brauer, D. and Wittman-Liebold, B. (1978) *FEBS Letters* **93**, 205.