Primary Structure of Alfalfa Ferredoxin*

S. KERESZTES-NAGY, F. PERINI, AND E. MARCOLLAISH

From the Department of Biochemistry and Biophysics, Stritch School of Medicine, Loyola University, Hines, Illinois 60141, and the Department of Molecular Biology, Abbott Laboratories, North Chicago, Illinois 60064

SUMMARY

The amino acid sequences of the six peptides obtained from tryptic digests of alfalfa ferredoxin were determined by chemical and enzymic degradations. The order of tryptic peptides was established from the properties of the two segments of the succinylated protein obtained by tryptic hydrolysis at the single arginyl residue. Alfalfa ferredoxin consists of a polypeptide chain of 97 residues, characterized by clusters of acidic and hydrophilic amino acids. It is suggested that this type of distribution of hydrophilic and strongly hydrophilic residues implies a low α-helical content of the native structure. Of the 5 cysteinyl residues only 2, separated by 2 serine residues, show a grouped arrangement similar to that observed in clostridial ferredoxins. Alfalfa ferredoxin differs from spinach ferredoxin at 19 residue positions. The over-all effect of these substitutions is largely conservative in terms of the total numbers of charged, hydroxy- and hydrophobic amino acids. Most of the variant positions are clustered in the vicinity of Ala-Ala sequences, leaving several long stretches unaltered.

Comparisons of peptide segments based on minimal mutation distances indicate the occurrence of internal repetitions of similar sequences 27 and 14 residues long in alfalfa ferredoxin. A rigorous statistical comparison of alfalfa and clostridial ferredoxins demonstrates a degree of amino acid sequence similarity clearly greater than would be expected on the basis of random occurrences. Although this raises the possibility of a common evolutionary origin for all ferredoxins, without knowledge of the minimal number of invariant residues required for function, such analyses cannot conclusively prove ancestral homology.

Ferredoxins are low molecular weight non-heme iron proteins having strongly negative oxidation-reduction potentials, which function as electron carriers on the hydrogen side of the pyridine nucleotides (see References 1 to 3). They have been obtained from many green plants, including algae, in which they are associated with chloroplasts and are part of the photosynthetic energy transducing system, as well as from some photosynthetic and non-photosynthetic bacteria (1–8). The molecular weight series represented by the proteins from these three different sources, approximately 12,900, 9,000, and 6,000, respectively (9–12), and the obvious amino acid sequence similarity of the two segments of clostridial ferredoxins (13, 14), have led to speculations concerning the possible development of present day ferredoxin genes through successive duplications (14), which have even been claimed to have started with gene strands no longer than four codons (15).

Plant ferredoxins, because of their ubiquitous distribution, relative ease of preparation, and small size, all of which make it practicable to determine the primary structures of the proteins from numerous species, may present a particularly useful object for studies of the relation of protein structure to organic evolution in the plant kingdom. Such investigations also afford numerous opportunities for the exploration of structure-function relations in a typical non-heme iron protein. To date, the amino acid sequences of the ferredoxins from spinach leaves (10, 16, 17) and from the green alga Stoeckelmannia have been determined. The present paper reports the primary structure of the ferredoxin from alfalfa (Medicago sativa var. Sonora) and the results of comparisons of the amino acid sequences of the alfalfa, spinach, and clostridial proteins.

EXPERIMENTAL PROCEDURE

Preparation of Alfalfa Ferredoxin Carboxymethyl and Succinyl Derivatives—Alfalfa ferredoxin was prepared as previously described (9) and, following deaeration under reduced pressure, was stored under N₂ at −15°C in 1 to 3% solutions in 0.02 M Tris-HCl buffer, pH 7.4, 0.8 M in NaCl. The material used had an absorbance ratio, A₄₄₂/₄₇₀, of 0.44 or higher.

Carboxymethylation was performed by a procedure similar to that of Crestfield, Moore, and Stein (18). The reaction mixture was deaerated and kept under N₂ in the dark. It contained 30 μmoles of ferredoxin, 1 ml of mercaptoethanol, 100 mg of di-sodium EDTA, 36 g of urea, and 0.05 M Tris-HCl buffer, pH 8.6, in a total volume of 75 ml. Incubation at 40°C was continued until complete decolorization had occurred (2 to 3 hours). After cooling to room temperature, 2.68 g of iodoacetamide (14.5 mmol) dissolved in 10 ml of 1 N NaOH were added and the reaction was allowed to proceed for 15 min. The mixture was

* This work was supported in part by United States Public Health Service Research Grant MG-14831-01 to one of the authors (S. K-N.).

passed through a column (45 x 4.5 cm) of Sephadex G-25, coarse bead form (Pharmacia), equilibrated with 0.02 M ammonium bicarbonate. The protein was detected in the effluent by its absorbance at 277 nm, and the excluded fraction (110 ml), containing the Cmc-ferredoxin was lyophilized three times to remove ammonium bicarbonate. The recovery was 95%.

Cmc-ferredoxin (4 μmoles in 2 ml of 0.02 M ammonium bicarbonate) was ceevinylated by slowly adding a 500 fold molar excess of solid succinic anhydride (200 mg) under continuous stirring over a period of 1 hour at room temperature. The pH was maintained near 8.0 by additions of finely powdered NaOH. The protein was recovered by gel filtration (2 x 45 cm column) as described above for Cmc-ferredoxin.

Enzymic Digests and Isolation of Peptides—Trypsin (Worthington, crystallized twice) was treated with t-1-tosylamido-2-phenylethyl chloromethyl ketone according to the method of Kostka and Carpenter (19) and recrystallized. Cmc-ferredoxin (15 μmoles in 15 ml of 0.05 M ammonium bicarbonate) was digested with 5 mg of enzyme at 38° for 3 hours. The digest was immediately applied to a column of Sephadex G-50, and 4 fractions were isolated from the eluate (see Fig. 1). The peptides were detected by their absorbance at 220 and 275 nm, and the pooled Fractions II, III, and IV were lyophilized three times to remove ammonium bicarbonate. Fraction I could not be lyophilized without becoming insoluble and was stored at -15°.

The components of Fraction IV were resolved by Dowex 50W-X2 (Bio-Rad) column chromatography (see Fig. 2), those of Fraction I by chromatography on a DEAE-cellulose (Eastman) column (see Fig. 3).

Peptides T-3, T-5, and T-6 were digested with chymotrypsin (Worthington, crystallized three times), 3 to 7 μmoles of peptide, 2% (mole per mole) enzyme, 3 to 15 ml of 0.05 M ammonium bicarbonate, 37°. Peptides T-3 and T-5 were hydrolyzed for 3 hours, Peptide T-6 for 8 hours. The reaction was stopped by immersing in boiling water for 3 min, the precipitate was centrifuged off, and the supernatant was lyophilized three times. The chymotryptic digests of Peptides T-3 and T-5 were fractionated by gel filtration (Figs. 4 and 5). The mixture of two dipeptides and one pentapeptide found in the more retarded fraction from the chymotryptic digest of Peptide T-5 (see Fig. 5) was resolved by preparative paper chromatography in 1-butanol-pyridine-acetic acid-water solvent (600:400:120:480 by volume) for 16 hours as previously described (20). The chymotryptic fragments from Peptide T-6 were separated by preparative paper peptide mapping (20), applying 0.5 μmole of digest on one sheet of Whatman No. 3MM or Schleicher and Schuell 593C paper (electrophoresis: pyridine-acetic acid buffer, pH 6.4, 20 volts per cm, 90 min; chromatography: 1-butanol-pyridine-acetic acid-water, 16 hours). In general, the latter brand of paper contained fewer peptide and amino acid impurities than the former and was preferred. As soon as the peptide positions were revealed with 0.02% ninhydrin in ethanol, the spots were cut out, excess ninhydrin was washed off with anhydrous acetone, the peptides were eluted with 2 M pyridine acetate buffer, pH 5.0, and the extracts were lyophilized.

The larger chymotryptic fragments of Peptides T-3 and T-5 were hydrolyzed with pepsin, 2 to 4 μmoles of peptide, 0.2 to 0.4 mg of pepsin (Worthington, crystallized twice) in 1.5 to 3.0
dissolved in a minimal volume of 10% NaCO₃ before addition to the reaction mixture. Leucine aminopeptidase (Worthington) and prolidase (28) digestions were performed as previously described (29). When prolidase was used it was added to a mixture which had been incubated for a suitable time with leucine aminopeptidase. Exopeptidase reactions were followed by thin layer peptide mapping and by analyzing aliquots taken at suitable intervals for their content of free amino acids.

Symbols and Calculation of Analytical Results—Peptides isolated from the initial tryptic digest of Cmc-ferredoxin are denoted by T. In cases in which these peptides were the result of both tryptic- and chymotryptic-like splits they are denoted by CT. The chymotryptic and peptic fragments derived from the initial tryptic peptides are denoted C and P, respectively. When tryptic peptides were further hydrolyzed by trypsin, and chymotryptic peptides by chymotrypsin, the fragments isolated from such digestes were distinguished by lower case letters. For example T-6a and T-6b are the tryptic fragments of Peptide T-6. In the tables and in the text the numbers given in parentheses following the analytical values are the assumed stoichiometric numbers of residues per molecule of pure peptide. Except when necessary to show the disappearance of a residue following Edman degradation, analytical values for amino acids lower than 0.1 residue are not reported. The average residue value for each analysis was calculated on the basis of all amino acids not significantly destroyed during acid hydrolysis or removed by preceding Edman degradation steps. For partial exopeptidase digests the residues released are reported as molar ratios, calculated by taking the residues released in largest amount, and known to occur as a single residue in the peptide, as unity. Since alfalfa ferredoxin carries a single tryptophan, peptides which gave a positive Ehrlich reaction were considered to contain 1 residue of tryptophan. In the tables, following the analytical values for each peptide, the percentage yield, the electrophoretic and chromatographic mobilities, and the color reactions on paper are listed in this order within brackets. The yields of all peptides were calculated without taking into account losses incurred during purification and as a percentage of the initial amount of protein subjected to tryptic digestion, even in the cases of fragments derived from larger peptides by enzymic hydrolysis.

The electrophoretic and chromatographic mobilities, in centimeters under the standard conditions employed, are designated by el and ch, respectively, a minus sign (−) denoting cathodic migration, a plus sign (+), anodic mobility, and a 0 indicating no electrophoretic movement. Peptides which gave a yellow color with the ninhydrin reagent are indicated as "yellow"; no notation is made if the ninhydrin color was the usual purple-blue; positive reactions to residue specific reagents derived from larger peptides by enzymic hydrolysis.

RESULTS

**Amino- and Carboxyl-terminal Amino Acid Sequences of Alfalfa Ferredoxin**

Direct identification of the PTH-amino acids for 13 steps of Edman degradation yielded the following sequence: NH₃Ala–Ser–Tyr–Lys–Val–Lys–Leu–Val–Thr–Pro–Glu–Gly–Thr. The results for subsequent cycles were not entirely satisfactory.

Digestion of 0.2 μmole of Cmc-ferredoxin with carboxypeptidase A for 1 hour released only the following residues: alanine, 0.108 μmole; threonine, 0.066 μmole; leucine, 0.026 μmole; suggesting the carboxyl-terminal sequence Leu–Thr–AlaCOOH.

**Isolation of Tryptic Peptides**

The gel filtration elution diagram of the tryptic digest of Cmc-ferredoxin is presented in Fig. 1. Fractions II and III contained single peptides of 15 residues (T-6) and 10 residues (T-4), respectively, and required no further purification. Fraction IV contained two authentic tryptic peptides, a tetrapeptide (T-1) and a dipeptide (T-2), as well as a tripeptide (CT-1) and a dipeptide (CT-2) resulting from both tryptic- and chymotryptic-
like hydrolysates. These were separated by cation exchange chromatography as shown in Fig. 2. Fraction I of Fig. 1 contained large peptides which were resolved by DEAE-cellulose column chromatography (Fig. 3). The first peak consisted of a 32-residue tryptophan-containing peptide (T-5), the second peak an arginine-containing 34-residue peptide (T-3), while the third peak was a mixture of these two peptides from which had been removed the carboxy-terminal di- and tripeptide sequences, respectively, by chymotryptic like action (CT-3 and CT-4). This tripeptide and this dipeptide were components of Fraction IV (Fig. 4). A single step of Edman degradation established its structure.

Edman 1: Val, 0.02; Lys, 1.00

Residues 7 through 40: Leu-Val-Thr-Pro-Glu-Gly-Thr-Gln-Glu-Phe-Glu-Cys(Cm)-Pro-Asp-Asp-Val-Tyr-Ile-Leu-Asp-His-Ala-Glu-Glu-Gly-Ile-Val-Leu-Pro-Tyr-Ser-Cys(Cm)-Arg (Peptide T-3, Table I)—This 34-amino acid peptide contained the single arginyln residue in the protein. Edman degradation indicated the amino-terminal sequence to be Leu-Val-Thr-Pro, and this structure was confirmed by digestion with leucine aminopeptidase with and without prolidase. Digestion with carboxypeptidases A and B indicated that arginine was carboxyl-terminal, as expected from the specificity of trypsin, and that it was preceded by carboxymethylcysteine, serine, and tyrosine.

Three fragments were isolated from a chymotryptic digest of T-3 (Fig. 4). C-1, a 17-residue fragment, had the amino-terminal sequence of T-3, while C-3, a tripeptide, contained the arginyln residue as well as serine and carboxymethylcysteine, indicating it represented the carboxyl-terminal fragment of T-3. Thus, from amino-to-carboxyl-terminal the order of the chymotryptic fragments of T-3 must be C-1 followed by C-2 and C-3.

Fourteen steps of Edman degradation indicated the amino-terminal sequence of C-1 to be, Leu-Val-Thr-Pro-Glu-Gly-Thr-Glu-Phe-Glu-Cys(Cm)-Pro-Asp, while leucine aminopeptidase digestion liberated only leucine and valine, as it had in the case of T-3, confirming that proline was the 4th residue in the sequence. Carboxypeptidase A digestion showed the carboxyl-terminal sequence to be Val Tyr. Five fragments of C-1 were isolated from a peptic digest. They were employed to confirm partially the structure derived by Edman degradation and to establish the carboxyl-terminal sequence of C-1. P-1, from its composition, clearly represented the amino-terminal 5 residues. P-2, containing only the glycine in C-1, represented the next 3 residues. Edman degradation and carboxypeptidase A digestion confirmed that its structure was Gly-Thr-Gln. P-3, an 11-residue fragment, had the amino-terminal sequence of P-2, and a carboxyl-terminal sequence of Asp-Val, indicating that this peptide extended to the penultimate residue of the over-all chymotryptic peptide C-1. P-4, a dipeptide, Glu-Phe, represented a portion of P-3, while P-5 was the carboxyl-terminal dipeptide sequence of C-1, Val-Tyr. Both C-1 and P-3 contained 2 aspartyl residues, and P-3 had a single valine. The 1st aspartyl residue was detected at the 14th step of Edman degradation of C-1, the second was contained in the carboxyl-terminal sequence Asp-Val of P-3, while the dipeptide P-5, Val-Tyr, contained the carboxyl-terminal sequence of C-1. Thus, the carboxyl terminus of C-1 must carry the sequence Asp-Asp-Val-Tyr. Together with the results of Edman degradation, this establishes the proposed structure of C-1.

Even though the direct identifications of the PTH-amino acids in the Edman degradation of C-1 were unequivocal, the analyses of the residual peptides at the last few steps were not entirely satisfactory. Therefore, to confirm the indicated structure, Peptide C-1 was subjected to prolonged chymotryptic digestion (2 moles of peptide, 5% (moles per mole) enzyme, 2 ml of 0.05 M ammonium bicarbonate, 37°, 16 hours), and two fragments, C-1a and C-1b, were isolated from the digest by preparative paper chromatography (Table I). C-1a contained Residues 7 to 16, C-1b, Residues 17 to 23. Edman degradat-
### Table I

**Amino acid sequence of Residues 7 through 40 (Peptide T-2)**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence: Leu-Val-Thr-Pro-Glu</th>
<th>Gly-Thr-Gln-Glu-Cys(Cm)-Pro-Asp-Val-Tyr</th>
</tr>
</thead>
</table>
| T-3     | Leu, 3.06(3); Val, 3.01(3); Thr, 1.80(2); Pro, 2.05(3); Gly, 7.21(7); Glu, 2.13(2); Phe, 1.00(1); Cys(Cm), 1.81(2); Asp, 3.28(3); Tyr, 1.75(2); His, 1.11(1); Ala, 0.97(1); Ser, 0.91(1); Arg, 0.88(1) [45%; el, +6.6; ch, 19.6; Sakaguchi, Pauly].
|         | Edman 1: PTH-Leu. Edman 2: PTH-Val. Edman 3: PTH-Thr. Edman 4: PTH-Pro. Leucine aminopeptidase (1 hr): Leu, 1.00; Val, 0.88.
|         | Leucine aminopeptidase (1 hr) + prolidase (2 hrs): Leu, 1.00; Val, 0.04; Thr, 0.35.
|         | Carboxypeptidases A and B (30 min): Arg, 1.00; Cys(Cm), 0.93; Ser, 0.87; Tyr, 0.81.
|         | Leu, 1.11(1); Val, 2.14(2); Thr, 1.81(1); Pro, 1.83(2); Gly, 4.07(4); Gly, 1.11(1); Phe, 0.96(1); Cys(Cm), 0.95(1); Asp, 2.00(2); Tyr, 0.97(1) [42%; el, +7.0; ch, 20.0; Pauly].
|         | Edman 1: PTH-Leu. Residue: Leu, 0.16; Val, 1.98; Thr, 1.81; Pro, 1.94; Glu, 4.10; Gly, 0.99; Phe, 0.99; Cys(Cm), 0.79; Asp, 2.02; Tyr, 0.90.
|         | Edman 2: PTH-Val. Residue: Leu, trace; Val, 0.95; Thr, 1.87; Pro, 1.89; Glu, 4.11; Gly, 1.00; Phe, 0.98; Cys(Cm), 0.89; Asp, 2.01; Tyr, 0.80.
|         | Edman 3: PTH Thr. Residue: Val, 0.96; Thr, 1.06; Pro, 1.96; Gly, 4.07; Gly, 1.06; Phe, 0.97; Cys(Cm), 0.93; Asp, 1.95; Tyr, 0.79.
|         | Edman 4: PTH-Pro. Edman 5: PTH-Glu. Residue: Val, 0.93; Thr, 0.99; Pro, 1.37; Glu, 3.45; Gly, 1.12; Phe, 0.87; Cys(Cm), 1.02; Asp, 2.08; Tyr, 0.85.
|         | Edman 6: PTH-Gly. Residue: Val, 0.98; Thr, 0.98; Pro, 1.14; Glu, 3.08; Gly, 0.47; Phe, 1.01; Cys(Cm), 0.95; Asp, 2.03; Tyr, 0.80.
|         | Edman 7: PTH-Thr. Residue: Val, 0.96; Thr, 0.57; Pro, 1.13; Glu, 3.15; Gly, 0.41; Phe, 0.95; Cys(Cm), 0.83; Asp, 2.07; Tyr, 0.85.
|         | Edman 8: PTH-Gln. Residue: Val, 0.88; Thr, 0.43; Pro, 0.93; Glu, 2.51; Gly, 0.48; Phe, 0.93; Cys(Cm), 0.88; Asp, 2.19; Tyr, 0.89.
|         | Edman 9: PTH-Glu. Edman 10: PTH-Phe. Residue: Val, 1.01; Thr, 0.39; Pro, 1.01; Glu, 1.97; Gly, 0.41; Phe, 0.41; Cys(Cm), 0.99; Asp, 2.07; Tyr, 0.92.
|         | Edman 11: PTH-Glu. Edman 12: PTH-Cys(Cm). Edman 13: PTH-Pro. Residue: Val, 1.04; Thr, 0.32; Pro, 0.50; Glu, 1.36; Gly, 0.37; Phe, 0.39; Cys(Cm), 0.31; Asp, 1.96; Tyr, 0.98.
|         | Edman 14: PTH-Asp. Leucine aminopeptidase (30 min): Leu, 1.00; Val, 0.76.
|         | Carboxypeptidase A (15 min): Tyr, 1.00; Val, 0.72; Asp, trace.
|         | Carboxypeptidase A (1 hr): Gly, 1.00; Thr, 0.27; Gly, 0.26.
| P-1     | Gly, 1.09(1); Thr, 0.91(1); Glu, 1.01(1) [9%; el, +2.9; ch, 28.5].
|         | Edman 1: PTH-Gly. Carboxypeptidase A (1 hr): Gly, 1.00; Thr, 0.27; Gly, 0.26.
| P-2     | Gly, 0.99(1); Thr, 0.87(1); Gly, 3.05(3); Phe, 0.95(1); Cys(Cm), 0.97(1); Pro, 1.07(1); Asp, 2.11(2); Val, 1.05(1) [9%; el, +6.4; ch, 14.8; yellow].
|         | Edman 1: PTH-Gly. Edman 2: PTH-Thr. Carboxypeptidase A (1 hr): Val, 1.00; Asp, 0.25.
| P-3     | Glu, 1.05(1); Phe, 0.95(1) [4%; el, +3.6; ch, 16.8].
| P-4     | Val, 1.11(1); Tyr, 0.80(1) [3%; el, 0; ch, 31.9; Pauly].
| P-5     | Leu, 0.98(1); Val, 1.11(1); Thr, 1.86(2); Pro, 0.91(1); Glu, 3.05(3); Gly, 1.15(1); Phe, 0.96(1) [20%; el, +5.6; ch, 28.5].
| C-1a    | Edman 1: PTH-Leu. Edman 2: PTH-Val. Carboxypeptidase A (1 hr): Phe, 1.00; Glu, 0.18.
TABLE I—Continued

| Peptide | Glu, 1.05(1); Cys(Cm), 0.92(1); Pro, 0.96(1); Asp, 2.10(2); Val, 1.07(1); Tyr, 0.90(1) [30%; el, +8.7; ch, 15.0; Pauly]. Edman 1: PTH-Glu. Residue: Glu, 0.12; Cys(Cm), 0.88; Pro, 0.95; Asp, 2.04; Val, 1.02; Tyr, 0.93. Edman 2: PTH-Cys(Cm). Residue: Glu, 0.14; Cys(Cm), 0.17; Pro, 0.93; Asp, 2.02; Val, 1.05; Tyr, 0.94. Edman 3: PTH-Pro. Residue: Glu, trace; Cys(Cm), 0.15; Pro, 0.21; Asp, 2.00; Val, 1.00; Tyr, 0.01. Edman 4: PTH-Asp. Residue: Cys(Cm), 0.16; Pro, 0.20; Asp, 1.37; Val, 1.02; Tyr, 0.98. Edman 5: PTH-Asp. Residue: Cys(Cm), trace; Pro, 0.13; Asp, 0.56; Val, 1.10; Tyr, 0.89. Carboxypeptidase A (1 hr): Tyr, 1.00; Val, 0.96; Asp, 0.18. Edman 1: PTH-Ile. Residue: Ile, 0.92; Leu, 1.01; Asp, 1.11; His, N.D.; Ala, 0.97; Glu, 3.07; Gly, 1.05; Val, 1.07; Pro, 0.96; Tyr, N.D. Edman 2: PTH-Leu. Residue: Ile, 0.96; Leu, 1.13; Asp, 1.11; His, 0.95; Ala, 1.02; Glu, 3.15; Gly, 1.07; Val, 1.09; Pro, 0.85; Tyr, 0.83. Edman 3: PTH-Asp. Residue: Ile, 0.97; Leu, 1.01; Asp, 0.39; His, 1.00; Ala, 1.02; Glu, 3.02; Gly, 1.09; Val, 0.99; Pro, 0.91; Tyr, 0.85. Edman 4: PTH-His. Residue: Ile, 0.89; Leu, 0.99; Asp, 1.37; His, 0.16; Ala, 0.94; Glu, 2.35; Gly, 1.11; Val, 1.00; Pro, 1.10; Tyr, 0.88. Edman 5: PTH-Ala. Residue: Ile, 0.85; Leu, 1.05; Asp, 0.44; His, 0.40; Ala, 0.50; Glu, 3.15; Gly, 1.05; Val, 1.02; Pro, 0.87; Tyr, 0.92. Edman 6: PTH-Glu. Edman 7: PTH-Glu. Residue: Ile, 0.98; Leu, 0.98; Asp, 0.41; His, N.D.; Ala, 0.37; Glu, 1.35; Gly, 1.25; Val, 1.07; Pro, 0.90; Tyr, 0.81. Edman 8: PTH-Glu. Leucine aminopeptidase (15 min): Ile, 1.00; Leu, 0.85; Asp, 0.28. Carboxypeptidase A (1 hr): Tyr only. Ile, 0.98(1); Leu, 1.02(1) [7%; el, 0; ch, 81.6]. P-1 Asp, 1.07(1); His, 1.01(1); Ala, 1.02(1); Glu, 3.02(3); Gly, 0.90(1) [9%; el, +7.5; ch, 12.3]. Carboxypeptidase A (8 hrs): Gly, 1.00; Glu, 0.35. P-3 Gly, 1.11(1); Ile, 1.02(1); Val, 0.88(1) [8%; el, 0; ch, 15.3; yellow]. Edman 1: PTH-Gly. Carboxypeptidase A (5 min): Val, 1.00; Ile, 0.57; Gly, 0.52. P-4 Ile, 0.98(1); Val, 1.11(1); Leu, 1.10(1); Pro, 0.96(1); Tyr, 0.93(1) [6%; el, 0; ch, 23.2; Pauly]. Edman 1: PTH-Ile. Leucine aminopeptidase (2 hrs): Ile, 1.00; Val, 0.95. Leucine aminopeptidase (2 hrs) + prolidase (2 hrs): Ile, 1.00; Val, 0.98; Leu, 0.58. Carboxypeptidase A (3 hrs): Tyr only. Residues 41 through 50: Ala-Gly-Ser-Cys(Cm)-Ser-Ser-Cys(Cm)-Ala-Gly-Lys (Peptide T-4, Table II)—The structure and carboxypeptidase digestions yielded the expected structures. The order of the amino-terminal 6 residues of Peptide C-2 was established by Edman degradation. The next two degradation cycles, like the sixth cycle, yielded only PTH-Glu, presumably indicating a Glu-Glu-Glu sequence. Leucine aminopeptidase digestion served to confirm the order of the first 3 residues, while carboxypeptidase A released only the carboxyl-terminal tyrosine. Four fragments were purified from a peptic digest of C-2. P-1 contained the amino-terminal dipeptide sequence of the parent peptide. From its unique composition, P-2 covered the succeeding 7 residues, as indicated by the amino-terminal sequence of C-2 established above. P-2 contained the 3 glutamyl residues in C-2, and extended the sequence determined by Edman degradation of C-2 by 1 residue. The single glycine in C-2 was the carboxyl-terminal residue of P-2 and the amino-terminal residue of P-3, thus showing that P-3 extends the sequence of P-2 at its carboxyl terminus. P-3 was the tripeptide Gly-Ile-Val, as established by Edman degradation and carboxypeptidase A digestion. The pentapeptide P-4, in addition to isoleucine and valine, contained the 3 residues of C-2 not accounted for by the other peptic fragments, including the carboxyl-terminal tyrosine. Edman degradation and leucine aminopeptidase digestion, with and without prolidase, indicated that the amino-terminal sequence of P-4 was Ile-Val-Leu, and that proline was placed in the fourth position. Carboxypeptidase A digestion showed tyrosine to be carboxyl-terminal, completing the structure of P-4. The tripeptide C-3 contained the single arginyl residue in the protein, and its structure was determined on the tripeptide CT-1 (Fig. 2), originally isolated directly from the trypsin digest of Cmbo-ferredoxin. CT-1 was present in the trypsin digest presumably as a result of chymotryptic-like hydrolysis at tyrosyl Residue 37 (see Fig. 6) and tryptic cleavage at arginine 40.
of this decapeptide was determined by nine steps of Edman degradation. It was partially confirmed by leucine aminopeptidase and carboxypeptidase B and A digests.

Residues 51 through 82: Val-Ala-Ala-Gly-Glu-Val-Asn-Gln-Ser-Gly-Ser-Phe-Leu-Asp-Asp-Glu-Val-Leu-Thr-Cys(Cm)-Val-Ala-Tyr-Ala-Lys (Peptide T-5, Table III) This 22-residue peptide contained the only tryptophan in alfalfa ferredoxin. Edman degradation showed that the amino-terminal sequence was Val-Ala-Ala-Gly. Digestion with carboxypeptidases A and B indicated that lysine was carboxyl-terminal and was preceded by alanine and tyrosine. From the kinetics of release of free amino acids by these enzymes, employed at one-fourth their usual concentrations, the carboxy-terminal sequence was shown to be Ala Tyr Ala-Lys. These results are not listed in Table III.

As described under "Experimental Procedure", four fragments were isolated from a chymotryptic digest of Peptide T-5 by gel filtration (Fig. 5), followed by paper chromatography to separate the three peptides in the retarded fraction. Peptide C-1, a 23-residue segment, contained the single tryptophan residue at its carboxyl terminus, preceded by glycine, and had the amino-terminal sequence of the parent tryptic peptide, T-5. Fourteen steps of Edman degradation yielded the sequence: Val-Ala-Ala-Gly-Glu-Val-Asn-Gln-Ser-Asp-Gly-Ser-Phe-Leu-Asp-Asp-Asp-Gln-Ile-Glu-Glu-Tyr. Edman 8: PTH-Ala. Residue: Ala, 0.38; Gly, 1.00; Ser, 0.21; Cys(Cm), 1.09; Lys, 0.99; (1) [87%; el, 0.3; ch, 9.9].

Edman 1: PTH-Ala. Residue: Ala, 1.18; Gly, 2.12; Ser, 2.94; Cys(Cm), 1.87; Lys, 1.05.

Edman 2: PTH-Gly. Residue: Ala, 1.04; Gly, 1.30; Ser, 2.96; Cys(Cm), 2.00; Lys, N.D.

Edman 3: PTH-Ser. Residue: Ala, 1.09; Gly, 1.21; Ser, 1.98; Cys(Cm), 1.91; Lys, N.D.

Edman 4: PTH-Cys(Cm). Residue: Ala, 1.00; Gly, 1.16; Ser, 1.94; Cys(Cm), 1.20; Lys, N.D.

Edman 5: PTH-Ser. Residue: Ala, 1.00; Gly, 1.14; Ser, 1.50; Cys(Cm), 0.93; Lys, N.D.

Edman 6: PTH-Ser. Residue: Ala, 1.09; Gly, 1.11; Ser, 0.55; Cys(Cm), 0.91; Lys, N.D.

Edman 7: PTH-Uys(Cm). Ser, 0.19; Cys(Cm), 0.15.

Edman 8: PTH-Ala. Residue: Ala, 0.38; Gly, 1.00; Ser, 0.49; Cys(Cm), 0.34; Lys, N.D.

Edman 9: PTH-Gly. Leucine aminopeptidase (2 hrs): Ala, 1.00; Gly, 0.21; Ser, 0.19; Cys(Cm), 0.15.

Carboxypeptidase B (15 min): Lys only.

Carboxypeptidases A and B (1 hr): Lys, 1.00; Gly, 0.21; Ala, 0.18; Cys(Cm), 0.11; Ser, 0.14.

(5 hrs): Lys, 1.00; Gly, 0.39; Ala, 0.32; Cys(Cm), 0.21; Ser, 0.31.

Table II

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence:</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-4</td>
<td>Ala, 2.14(2); Gly, 2.10(2); Ser, 2.80(3); Cys(Cm), 1.91(2); Lys, 0.99(1) [87%; el, 0.3; ch, 9.9].</td>
</tr>
</tbody>
</table>

To confirm the above proposed structure of C-1, this peptide was subjected to prolonged chymotryptic digestion (2 amoles of peptide, 5% enzyme (moles per mole), 16 hours). Three peptide fragments were isolated by preparative paper chromatography and their amino- and carboxy-terminal residues identified by Edman degradation and carboxypeptidase digestion (Table III). Fragment C-la consisted of the 13-residue sequence at the amino end of C-1, C-1b the carboxy-terminal 10 residues, and C-1c was the same as C-1b except that it lacked the amino-terminal leucine of the latter fragment. The compositions of these peptides confirmed the over-all position assignments of the peptic fragments of C-1 and the hydrolyses at the phenylalanyl and leucyl residues (Residues 63 and 64) are according to normal chymotryptic specificity.
**Table III**

Amino acid sequence of Residues 51 through 82 (Peptide T-5)

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence:</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T-5</strong></td>
<td>Val-Ala-51</td>
<td>Val-Ala-Thr-Cys(Cm) Val-Ala-Glu-Glu-Glu-Glu-Glu-Trp</td>
</tr>
<tr>
<td>C-1</td>
<td>Val-Leu-51</td>
<td>Val-Leu-Thr-Cys(Cm) Val-Ala-Glu-Glu-Glu-Glu-Glu-Trp</td>
</tr>
</tbody>
</table>

**Carboxypeptidase A (3 hrs):**
- Val, 1.09(2); Ala, 1.08(2); Gly, 1.98(3); Glu, 5.02(5); Asp, 4.08(5); Ser, 1.94(2); Phe, 1.03(1); Leu, 1.13(1); Ile, 0.97(1); Trp, (1) 142%; el, +8.7; ch, 18.2; Ehrlich, Pauly.


**Carboxypeptidase A and B (15 min, room temperature):**
- Lys, 1.00; Ala, 1.76; Tyr, 0.87; Val, 0.17.

Val, 1.08(2); Ala, 1.98(2); Gly, 2.98(3); Glu, 5.02(5); Asp, 4.08(5); Ser, 1.94(2); Phe, 1.03(1); Leu, 1.13(1); Ile, 0.97(1); Trp, (1) 142%; el, +8.7; ch, 18.2; Ehrlich, Pauly.


**Edman 1:** PTH-Val. Residue: Ser, Glu, Asp, Trp, N.D.; Val, 1.21; Ala, 1.95; Gly, 3.06; Phe, 0.97; Leu, 1.14; Ile, 0.91.

Edman 2: PTH-Ala. Residue: Val, 0.96; Ala, 0.95; Gly, 3.05; Glu, 5.00; Asp, 4.92; Ser, 1.96; Phe, 1.01; Leu, 1.10; Ile, 0.97; Trp, N.D.

Edman 3: PTH-Ala. Residue: Val, 1.03; Ala, 0.37; Gly, 3.00; Glu, 5.06; Asp, 4.84; Ser, 1.92; Phe, 1.05; Leu, 1.11; Ile, 0.97; Trp, N.D.

Edman 4: PTH-Gly. Residue: Val, 1.14; Ala, 0.33; Gly, 2.28; Glu, 5.01; Asp, 4.80; Ser, 1.90; Phe, 1.05; Leu, 1.07; Ile, 0.98; Trp, N.D.

Edman 5: PTH-Glu. Residue: Val, 1.11; Ala, 0.22; Gly, 2.41; Glu, 4.24; Asp, 4.80; Ser, 2.12; Phe, 0.87; Leu, 1.07; Ile, 0.96; Trp, N.D.

Edman 6: PTH-Val. Residue: Val, 0.35; Ala, 0.21; Gly, 2.25; Glu, 4.17; Asp, 4.95; Ser, 1.98; Phe, 1.02; Leu, 1.09; Ile, 0.96; Trp, N.D.

Edman 7: PTH-Glu. Residue: Val, 0.27; Gly, 2.13; Glu, 4.15; Asp, 4.07; Ser, 2.09; Phe, 1.01; Leu, 1.05; Ile, 0.91; Trp, N.D.

Edman 8: PTH-Val. Residue: Val, 0.25; Gly, 2.08; Glu, 3.47; Asp, 4.07; Ser, 1.99; Phe, 1.02; Leu, 1.06; Ile, 0.93; Trp, N.D.

Edman 9: PTH-Ser. Edman 10: PTH-Asp. Residue: Val, 0.15; Gly, 2.05; Glu, 3.33; Asp, 3.41; Ser, 1.27; Phe, 1.01; Leu, 1.02; Ile, 0.98; Trp, N.D.

Edman 11: PTH-Glu. Edman 12: PTH-Ser. Edman 13: PTH-Phe. Residue: Gly, 1.47; Glu, 3.44; Asp, 3.33; Ser, 0.58; Phe, 0.46; Leu, 1.03; Ile, 0.97; Trp, N.D.

Edman 14: PTH-Leu. Carboxypeptidase A (3 hrs): Trp, 1.00; Gly, 0.15.

**P-1** Val, 1.02(1); Ala, 0.98(1) [8%; el, 0; ch, 21.5].

**P-2** Ala, 1.08(1); Gly, 0.91(1); Glu, 1.01(1) [8%; el, +4.5; ch, 16.8].

**P-3** Val, 1.05(1); Asp, 2.01(2); Glu, 0.97(1); Ser, 1.87(2); Gly, 1.12(1) [8%; el, +3.6; ch, 14.4].

Edman 1: PTH-Val. Carboxypeptidase A (3 hrs): Ser, 1.60; Gly, 0.75.

**P-4** Val, 1.09(1); Asp, 1.04(2); Glu, 1.06(1); Ser, 1.88(2); Gly, 1.07(1); Phe, 0.97(1) [10%; el, +3.9; ch, 15.8].

Edman 1: PTH-Val. Residue: Val, 0.25; Asp, 1.96; Glu, 1.06; Ser, 1.97; Gly, 1.07; Phe, 0.96.

Edman 2: PTH-Asp. Edman 3: PTH-Glu.

**Edman 4:** PTH-Ser. Residue: Val, 0.11; Asp, 1.24; Glu, 0.58; Ser, 1.29; Gly, 1.11; Phe, 0.89.

**Edman 5:** PTH-Asp. Carboxypeptidase A (45 min): Phe, 1.00; Ser, 0.35; Gly, 0.21.

**P-5** Val, 1.00(1); Asp, 2.15(2); Glu, 1.11(1); Ser, 1.92(2); Gly, 1.02(1); Phe, 0.91(1); Leu, 0.92(1) [8%; el, +3.7; ch, 18.7].

Edman 1: PTH-Val. Carboxypeptidase A (45 min): Leu, 1.00; Phe, 0.89; Ser, 0.35; Gly, 0.19.
### TABLE III—Continued

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Leu, 1.00; Asp, 3.05; Glu, 0.97</th>
<th>Edman 1: PTH-Leu. Residue: Leu, 0.20; Asp, 2.97; Glu, 1.02.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-2</td>
<td>Val, 2.00; Ala, 1.97; Gly, 2.05</td>
<td>Edman 1: PTH-Val. Residue: Val, 1.10; Tyr, 0.80.</td>
</tr>
<tr>
<td>C-3</td>
<td>Thr, 0.94; Cys(Cm), 1.08; Val, 1.10; Tyr, 0.80</td>
<td>Edman 1: PTH-Thr. Edman 2: PTH-Cys(Cm).</td>
</tr>
<tr>
<td>C-4</td>
<td>Ala, 1.02; Lys, 0.98</td>
<td>Edman 1: PTH-Ala.</td>
</tr>
<tr>
<td>CT-2</td>
<td>Ala, 1.01; Lys, 0.99</td>
<td>Edman 1: PTH-Ala.</td>
</tr>
</tbody>
</table>

C-2 was the dipeptide Val-Leu, C-3 the pentapeptide Thr-Cys(Cm)-Val-Ala-Tyr, and C-4, the dipeptide Ala-Lys. These three fragments must be placed at the carboxyl terminus of T-5 since C-1 represents the amino-terminal 23 residues of T-5. Carboxypeptidase A and carboxypeptidases A and B digestions of T-5 (Table III) showed lysine to be carboxyl-terminal and also released alanine and the single tyrosine in the peptide, but no leucine. Hence, C-3, carrying a carboxyl-terminal tyrosine, must precede the dipeptide C-4, Ala-Lys, and C-2, Val-Leu, can be placed by difference between C-1 and C-3. This explains the kinetics of release of residues by carboxypeptidase digestion of the parent tryptic peptide T-5. CT-2 was the same dipeptide as C-4. It had been obtained from the original tryptic digest of the protein (Figs. 1 and 2), and represented a product of both tryptic- and chymotryptic-like hydrolys.

### Residues 83 through 97: Ser-Asp-Val-Thr-Ile-Glu-Thr-His-Lys-Glu-Glu-Lys-Leu (Peptide T-6, Table IV)—The structure of T-6 was established by 12 steps of Edman degradation and carboxypeptidase digestion. Because the evidence, from Edman degradation, on the location of the 3 glutamyl residues could not be considered definitive, and because it was not certain that no glutamine was present, the chymotryptic fragments of T-6 were isolated. Their compositions and partial sequences served to confirm the structure given above. C-1 contained the amino-terminal 8 residues of T-6. Fragment C-2

### Fig. 5. Elution pattern of peptides from a chymotryptic digest of Peptide T-6. Sephadex G-50 column chromatography under the conditions given for Fig. 1. The fractions pooled (solid bars) were lyophilized.
TABLE IV

Amino acid sequence of Residues 88 through 97 (Peptide T-6)

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-6</td>
<td>Ser-Asp-Val-Thr-Ile-Glu-Thr-His-Lys-90-Glu-Glu-Glu-Leu-Thr-Ala-97</td>
</tr>
<tr>
<td>C-1</td>
<td>Ser, 0.88(1); Asp, 1.14(1); Val, 1.05(1); Thr, 2.85(3); Ile, 0.98(1); Glu, 4.09(4); His, 1.03(1); Lys, 0.97(1); Leu, 1.05(1); Ala, 0.97(1) [88%; el, +6.3; ch, 18.5 (trailing); Pauly].</td>
</tr>
<tr>
<td>Edman 1: PTH-Ser. Residence: Ser, 0.14; Asp, 1.11; Val, 1.04; Thr, 2.81; Ile, 0.99; Glu, 4.02; His, 1.05; Lys, 0.95; Leu, 1.07; Ala, 0.98.</td>
<td></td>
</tr>
<tr>
<td>Edman 2: PTH-Asp.</td>
<td></td>
</tr>
<tr>
<td>Edman 3: PTH-Val. Residence: Asp, 0.37; Val, 0.43; Thr, 2.38; Ile, 0.90; Glu, 4.07; His, 1.05; Lys, N.D.; Leu, 1.04; Ala, 1.02.</td>
<td></td>
</tr>
<tr>
<td>Edman 4: PTH-Thr. Residence: Asp, 0.35; Val, 0.41; Thr, 2.12; Ile, 0.96; Glu, 4.02; His, 1.03; Lys, 0.97; Leu, 1.02; Ala, 1.01.</td>
<td></td>
</tr>
<tr>
<td>Edman 5: PTH-Ile. Residence: Asp, 0.33; Val, 0.39; Thr, 2.05; Ile, 0.41; Glu, 4.05; His, 1.11; Lys, 0.89; Leu, 1.01; Ala, 0.94.</td>
<td></td>
</tr>
<tr>
<td>Edman 6: PTH-Glu. Residence: Asp, 0.25; Val, 0.31; Thr, 2.00; Ile, 0.35; Glu, 3.46; His, 1.07; Lys, 0.93; Leu, 1.05; Ala, 0.95.</td>
<td></td>
</tr>
<tr>
<td>Edman 7: PTH-Thr.</td>
<td></td>
</tr>
<tr>
<td>Edman 8: PTH-His. Residence: Asp, 0.20; Val, 0.21; Thr, 1.02; Ile, 0.25; Glu, 3.28; His, 0.31; Lys, 0.83; Leu, 1.10; Ala, 1.07.</td>
<td></td>
</tr>
<tr>
<td>Edman 9: PTH-Lys. Residence: Asp, 0.15; Val, trace; Thr, 1.11; Ile, trace; Glu, 3.11; His, 0.31; Lys, 0.23; Leu, 1.03; Ala, 0.97.</td>
<td></td>
</tr>
<tr>
<td>Edman 10, 11, and 12: PTH-Glu.</td>
<td></td>
</tr>
<tr>
<td>Carboxypeptidase A (1 hr): Ala, 1.00; Thr, 0.86; Leu, 0.49; Glu, 0.18.</td>
<td></td>
</tr>
<tr>
<td>Edman 1: PTH-Ser. Edman 2: PTH-Asp.</td>
<td></td>
</tr>
<tr>
<td>Carboxypeptidases A and B (1 hr): His, 1.00; Thr, 0.55; Glu, 0.12.</td>
<td></td>
</tr>
<tr>
<td>Edman 4: PTH-Glu. Residue: Lys, 0.11; Glu, 1.51; Leu, 1.00.</td>
<td></td>
</tr>
<tr>
<td>Edman 5: PTH-Glu.</td>
<td></td>
</tr>
<tr>
<td>Leucine aminopeptidase (16 hrs): Lys, 1.00; Glu, 1.37; Leu, 0.13.</td>
<td></td>
</tr>
<tr>
<td>Carboxypeptidase A (2 hrs): Leu, 1.00; Glu, 0.53.</td>
<td></td>
</tr>
<tr>
<td>Edman 1: PTH-Thr.</td>
<td></td>
</tr>
<tr>
<td>Edman 2: PTH-Asp.</td>
<td></td>
</tr>
<tr>
<td>Carboxypeptidase A (1 hr): Ala, 1.00; Thr, 0.66; Leu, 0.49; Glu, 0.18.</td>
<td></td>
</tr>
<tr>
<td>Edman 1: PTH-Ser. Edman 2: PTH-Asp.</td>
<td></td>
</tr>
<tr>
<td>Carboxypeptidase A and B (1 hr): His, 1.00; Thr, 0.55; Glu, 0.12.</td>
<td></td>
</tr>
<tr>
<td>Edman 3: PTH-Glu.</td>
<td></td>
</tr>
<tr>
<td>Prolonged digestion (16 hours) of T-6 with trypsin, using 5% enzyme (moles per mole), yielded two fragments which were separated by preparative paper chromatography. T-6a contained the amino-terminal 9 residues of T-6, and T-6b the carboxyl-terminal 6 residues. Their compositions and partial sequences confirmed the structure of T-6 given above.</td>
<td></td>
</tr>
</tbody>
</table>

Order of Tryptic Peptides

The amino-terminal amino group and the e-amino groups of the 5 lysyl residues of Cmc-ferredoxin were blocked by succinyl-
of the intact protein. Hence the tryptic peptides in Peptide B were in the following order: T-4, T-5, T-6; completing the over-all amino acid sequence of alfalfa ferredoxin shown in Fig. 6.

**Amino Acid Composition of Alfalfa Ferredoxin**

The sums of the amino acid compositions of the tryptic peptides obtained from Cmc-ferredoxin and from succinyl-Cmc-ferredoxin are listed in Table VI and compared to the composition of the protein derived from its amino acid sequence. The agreement is satisfactory. However, this over-all composition differs from that previously reported for native alfalfa ferredoxin.

The sums of the amino acid compositions of the tryptic peptides from succinyl-Cmc-ferredoxin and from succinyl-Cmc-ferredoxin are listed in Table VI and compared to the composition of the protein derived from its amino acid sequence. The agreement is satisfactory. However, this over-all composition differs from that previously reported for native alfalfa ferredoxin.

### Table V

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Ala, 1.99(2); Ser, 1.83(2); Tyr, 2.77(3); Lys, 2.16(2); Val, 4.08(4); Leu, 2.96(3); Thr, 1.95(2); Pro, 2.85(3); Gly, 2.07(2); Glu, 7.27(2); Phe, 0.96(1); Cys(Cm), 1.79(2); Asp, 3.25(3); Ile, 2.07(2); His, 1.11(1), Arg, 0.97(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Succinyl-A</td>
<td>Edman 1: No residue released. Carboxypeptidase A (1 hr): No residue released. Carboxypeptidase B (15 min): Arg, 1.00; Cys(Cm), 0.27; Ser, 0.24; Tyr, 0.21.</td>
</tr>
<tr>
<td>Succinyl-B</td>
<td>Edman 1: PTH-Ala. Edman 2: PTH-Gly. Edman 3: PTH-Ser. Edman 4: PTH-Cys(Cm). Edman 5: PTH-Ser. Carboxypeptidase A (1 hr): Ala, 1.00; Thr, 0.89; Leu, 0.88.</td>
</tr>
</tbody>
</table>

### Table VI

**Amino acid composition of alfalfa ferredoxin**

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Cmc-ferredoxin</th>
<th>Cmc-ferredoxin tryptic peptides</th>
<th>Succinyl-Cmc-ferredoxin tryptic peptides</th>
<th>From amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>5.25</td>
<td>5.19</td>
<td>5.13</td>
<td>5</td>
</tr>
<tr>
<td>Histidine</td>
<td>1.89</td>
<td>2.14</td>
<td>2.23</td>
<td>2</td>
</tr>
<tr>
<td>Arginine</td>
<td>0.88</td>
<td>0.88</td>
<td>0.97</td>
<td>1</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>9.38</td>
<td>9.63</td>
<td>9.20</td>
<td>9</td>
</tr>
<tr>
<td>Threonine</td>
<td>6.05</td>
<td>5.77</td>
<td>5.81</td>
<td>6</td>
</tr>
<tr>
<td>Serine</td>
<td>7.75</td>
<td>7.55</td>
<td>7.55</td>
<td>8</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>16.39</td>
<td>16.43</td>
<td>16.54</td>
<td>16</td>
</tr>
<tr>
<td>Proline</td>
<td>3.97</td>
<td>2.95</td>
<td>2.95</td>
<td>3</td>
</tr>
<tr>
<td>Glycine</td>
<td>7.32</td>
<td>7.43</td>
<td>7.09</td>
<td>7</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.99</td>
<td>8.93</td>
<td>8.86</td>
<td>9</td>
</tr>
<tr>
<td>Half-cystine</td>
<td>4.95</td>
<td>4.78</td>
<td>4.63</td>
<td>5</td>
</tr>
<tr>
<td>Valine</td>
<td>8.99</td>
<td>8.99</td>
<td>9.24</td>
<td>9</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>3.92</td>
<td>3.84</td>
<td>4.21</td>
<td>4</td>
</tr>
<tr>
<td>Leucine</td>
<td>6.05</td>
<td>5.98</td>
<td>6.11</td>
<td>6</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>4.08</td>
<td>3.57</td>
<td>3.61</td>
<td>4</td>
</tr>
<tr>
<td>Phenylationamine</td>
<td>1.84</td>
<td>1.80</td>
<td>1.94</td>
<td>2</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>1.85</td>
<td>1.85</td>
<td>1.85</td>
<td>1</td>
</tr>
</tbody>
</table>

* Values from analyses of 24-, 48-, and 72-hour acid hydrolysates of Cmc-ferredoxin. The half-cystine was determined as carboxymethylcysteine and the tryptophan content was assumed to be 1 residue per molecule of protein. The values listed for threonine, serine, proline, half-cystine, and tyrosine were obtained by extrapolation to zero time, those for all other residues by averaging the values determined for the different hydrolysates.

### Fig. 6

**Fig. 6. Amino acid sequence of alfalfa ferredoxin.** The symbols T-1 through T-6 delimit the tryptic peptides. The arrows above the line indicate points of tryptic hydrolysis, those below the line, points of chymotryptic hydrolysis observed in the digestions of tryptic Peptides T-3, T-5, and T-6. The large arrow after Residue 40 marks the point of tryptic hydrolysis of succinyl-Cmc-ferredoxin, and the brackets on the left denote the two resultant succinyl peptides.
It would appear that native plant ferredoxins are only to a very minor extent in the α-helical conformation, as estimated from optical rotatory dispersion and circular dichroism studies. It may be suggested that the coincidence of these two phenomena, which occur in an α-helical region, in a protein too small to contain the α-helix entirely in its interior, some of the side chains would necessarily be facing the inside and others the surface of the molecule.

The two proteins carry identical residues in 78 positions. Such an extensive degree of identity is probably well beyond the minimal requirements for function. It is, therefore, unlikely that these similarities result from evolutionary functional clusters, but at positions so spaced that along an α-helix the side chains are presented to the “inside” of the protein out of contact with the solvent (32). If clusters of a particular type of residue occur in an α-helical region, in a protein too small to contain the α-helix entirely in its interior, some of the side chains would necessarily be facing the inside and others the surface of the molecule.

Of the 5 cysteine residues, 2 are in a Cys–Ser–Ser–Cys sequence reminiscent of the spacing of 6 of the 8 cysteine residues in clostridial ferredoxins (13, 14). The latter proteins have 7 iron atoms and 7 “labile sulfur” per molecule, as compared to 2 each for alfalfa ferredoxin (9). Cysteine sulfur has been implicated in the iron-binding sites of ferredoxins and other non-heme iron proteins (see References 3 and 33 to 36). Thus, the similarity of spacing of cysteiny! residues 44 and 47 in alfalfa ferredoxin to that in the bacterial proteins, and the corresponding iron and “labile sulfur” contents, suggest that these particular cysteines may be part of the iron-binding locus in the plant protein.

The Lys–Glu bond between Residues 91 and 92 was relatively resistant to tryptic hydrolysis, possibly because of the adjacent Glu–Glu–Glu sequence. No split at this bond was observed in the initial tryptic hydrolysis of the protein, the peptide T-6 containing Residues 91 and 92 being recovered in 88% yield. When Peptide T-6 was digested with 5% (moles per mole) trypsin for 16 hours at 37°, approximately 35% of the peptide was split at this bond. The peptide bonds at tyrosyl Residues 37 and 80 were particularly susceptible to the chymotryptic-like activity of the preparation of trypsin employed, even though it had been treated with 1-tosylamido-2-phenylethyl chloromethyl ketone (19) and recrystallized. Indeed, the peptides Ser–Cys–Arg (Residues 38 to 40) and Ala–Lys (Residues 81 and 82) were recovered in yields of 18% and 42%, respectively, indicating substantial digestion at the adjacent tyrosyl residues. No other unexpected points of hydrolysis were observed.

In the chymotryptic digestions of the larger tryptic peptides, T-3 and T-5, the bonds at phenylalanines 16 and 63 were hydrolyzed distinctly less readily than those at tyrosyl Residues 23, 37, and 80, and tryptophanyl Residue 73. Hydrolyses amounting to 70 and 65% at Residues 16 and 63, respectively, as judged from the yields of the corresponding fragments, were obtained only when 5% chymotrypsin (moles per mole of peptide) and digestion times of 16 hours at 37° were employed, as compared to the 2% (moles per mole) enzyme and digestion times of 3 hours, routinely used. Similarly, chymotryptic hydrolysis at histidyl Residue 90 was obtained in a yield of 23% only at the higher enzyme concentration. Of the 5 leucines in Peptides T-3 and T-5 only that at Residue 75 provided a point of substantial chymotryptic hydrolysis (63%). The leucine at position 7 was amino-terminal, that at 35 was followed by proline, and those at positions 25 and 64 were followed by aspartic acid and the Asp–Asp–Asp sequence, respectively. At the latter 2 leucines some chymotryptic hydrolysis was observed. It was not quantitated and probably amounted to no more than 15 to 25%.

Comparison of Primary Structures of Alfalfa and Spinach Ferredoxins—These ferredoxins consist of a single polypeptide chain of 97 residues (Fig. 7), and are functionally interchangeable in the photoreduction of TPN+ mediated by chloroplasts (9). The two proteins carry identical residues in 78 positions. Such an extensive degree of identity is probably well beyond the minimal requirements for function. It is, therefore, unlikely that these similarities result from evolutionary functional

---

**DISCUSSION**

Amino Acid Sequence of Alfalfa Ferredoxin—The primary structure of alfalfa ferredoxin (Fig. 6) is characterized by groupings of acidic and amided residues. Thus, there is one Asp–Asp, one Asp–Asp–Asp–Glu, one Gln–Glu, one Asn–Glu, one Thr–Asp, and two Glu–Glu–Glu sequences. Outside of these clusters there are only 3 aspartyl and 4 glutamyl residues. There is also distinct clustering of hydrophobic residues, such as in the sequences Tyr–Lys–Val–Val–Leu–Val (Residues 3 to 8), Val–Tyr–Ile–Leu–Val–Val–Tyr–Ile–Leu–Val (Residues 22 to 25), Ile–Val–Leu–Pro–Tyr–Leu (Residues 32 to 37), Phe–Leu–Val (Residues 65 and 64), Trp–Val–Leu–Val (Residues 73 to 75), Val–Asp–Tyr (Residues 75 to 80), and Val–Thr–Ile (Residues 85 to 87). Only 2 valyl residues and 1 residue each of phenylalanine, isoleucine, and leucine are excluded from these hydrophobic clusters. The few basic residues in the protein are dispersed. This type of distribution of hydrophilic and acidic residues bears a distinct resemblance to the situation in eukaroytic cytochromes c, in which hydrophilic and basic residues are clustered (see Reference 31). Like cytochromes c it would appear that native plant ferredoxins are only to a very minor extent in the α-helical conformation, as estimated from optical rotatory dispersion and circular dichroism studies.

It may be suggested that the coincidence of these two phenomena, namely overwhelmingly clustered distributions of hydrophilic and of strongly hydrophobic residues and a low helical content, is likely to be the general rule in proteins. Indeed, in those proteins which have high helical contents, as the myoglobin and hemoglobin, nonpolar residues are likely to occur not in
convergence, and probably indicate that the two proteins are homologous, namely, that they both derive from a common ancestral form (36).

Although spinach and alfalfa are dicots, they belong to orders, the Chenopodiaceae and Leguminosae, respectively, which appear to have diverged very early in the evolution of angiosperms, possibly as far back as 150 million years ago in the Jurassic period (37). Whether the 19 residue differences between alfalfa and spinach ferredoxins, corresponding to about 20% of the entire chain, represent an unusual or the common rate of evolutionary change for plant ferredoxins cannot be judged without primary structure information from many more such proteins. Interestingly, although only 3 of the 19 variations involve apparently conservative residue interchanges (Residues 51 and 56: leucine, valine; Residue 70: aspartic and glutamic acids), the net effect of all the substitutions on the over-all amino acid compositions of the two proteins is remarkably conservative. Thus, although valine is involved in four changes in addition to the conservative variations noted, and alanine in six, this results in only two fewer valines in the spinach protein and in no net change in the number of alanines. Similarly, serine and threonine occur 7 times in the residue changes but the total number of hydroxy-amino acids differs by only 1, being 15 in the spinach protein and 14 in alfalfa ferredoxin. Glutamyl and aspartyl residues are involved seven times in the variations between the two proteins, and the resulting change in the total number of acidic residues is only 1 extra such residue in the alfalfa as compared to the spinach protein (21 and 20 residues, respectively). This added negative charge is balanced by 1 extra residue each of lysine and histidine in alfalfa ferredoxin.

The 19 positions which carry different residues in alfalfa and spinach ferredoxins are not uniformly distributed throughout the chain (Fig. 7). There are three variable segments (Residues 1 to 17, 52 to 61, and 78 to 82) containing 15 amino acid changes, leaving several long stretches unchanged. Among these is the 16-residue segment (Residues 56 to 60) containing the Cys-Ser-Ser-Cys sequence discussed above. Two of the variable segments start in the Ala-Ala sequence in spinach ferredoxin (Residues 1 and 2 and 78 and 79) and the third begins with the same dipeptide sequence in the alfalfa protein (Residues 52 and 53). There is a fourth Ala-Ala sequence in spinach ferredoxin at positions 27 and 28, starting an 8-residue segment containing two more of the variant positions. The three Ala-Ala sequences mark the beginning of segments 25, 26, and 20 residues long, with the variable regions located in roughly the amino-terminal halves of these segments (Fig. 7). Such periodicities might reflect an evolutionary process of chain lengthening by gene duplication. A repeating segment about 26 residues long is suggested by the fact that the ferredoxin of the photosynthetic bacterium, Chlorobium (38), is 26 residues longer than the clostridial proteins (13, 14), and also that the latter consist of two remarkably similar segments of 29 and 26 residues, respectively (13, 14). All ferredoxins are only 1 extra such residue in the alfalfa as compared to the spinach protein (21 and 20 residues, respectively). This added negative charge is balanced by 1 extra residue each of lysine and histidine in alfalfa ferredoxin.

### Table VII

Amino acid sequence comparisons of segments of alfalfa and Clostridium pasteurianum ferredoxins

#### A. Internal similarities in alfalfa ferredoxin

**Comparison I**

<table>
<thead>
<tr>
<th>Residues</th>
<th>Residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu-Gly-Thr-Glu-Glu-Phe-Glu-Cys-Pro-Asp-Val-Tyr-Ile-</td>
<td>Val-Ala-Ala-Gly-Glu-Val-Asn-Gln-Ser-Gly-Leu-</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
</tr>
</tbody>
</table>

Nucleotide difference: 1 1 1 2 0 1 1 1 2 1 1

Average minimal nucleotide difference per codon: 26:27 = 0.96

**Comparison II**

<table>
<thead>
<tr>
<th>Residues</th>
<th>Residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp-Val-Tyr-Ile-Leu-Asp-His-Ala-Glu-Glu-Gly-Ile-Val</td>
<td>56</td>
</tr>
<tr>
<td>21</td>
<td>84</td>
</tr>
</tbody>
</table>

Nucleotide difference: 0 0 2 0 2 2 0 2 0 0 0 2 1 1

Average minimal nucleotide difference per codon: 12:14 = 0.86

#### B. Similar sequence regions in alfalfa and C. pasteurianum ferredoxins

**Alfalfa:**

<table>
<thead>
<tr>
<th>Residues</th>
<th>Residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly-Ser-Cys-Ser-Cys-Ala-Gly-Val-Ala-Ala-Gly-</td>
<td>42</td>
</tr>
</tbody>
</table>

Nucleotide difference: 1 0 0 1 1 1 0 1 2 1 1 2 1

Average minimal nucleotide difference per codon: 23:28 = 0.82

**Clostridial:**

<table>
<thead>
<tr>
<th>Residues</th>
<th>Residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val-Ser-Cys-Gly-Ala-Cys-Ala-Ser-Glu-Cys-Pro-Val-Ala-</td>
<td>56</td>
</tr>
</tbody>
</table>

Nucleotide difference: 1 1 0 1 0 1 1 0 1 2 0 1 1 1

Average minimal nucleotide difference per codon: 23:28 = 0.82
of these two apparently repeated sequences corresponds to one of
ferences per codon of 0.96 and 0.86, respectively. The shorter
to 77, as well as Residues 21 to 34 and 84 to 97. The minimal
ments of present day ferredoxins. Indeed, Matsubara,
possible to observe some degree of similarity between segments of
residue.
partial gene duplications conserving the original amino-terminal
ich functions of these two sets of proteins (36).

If evolutionary duplications did in fact occur, it might be
possible to observe some degree of similarity between segments of
the structures of present day ferredoxins. Indeed, Matsubara,
Sasaki, and Chain (10) found two similar segments in spinach
ferredoxin, Residues 1 to 9 and 78 to 86 (see Fig. 7) matching the
murn of their ancestral amino acid sequences may better
reveal the occurrence and exact location of duplicated segments
(39).

Possible Homology between Clostridial and Plant Ferredoxins
—In comparing spinach and C. butyricum ferredoxins, Matsu-
bara and Sasaki (17) found a pair of 19-residue segments with a
minimal mutation distance of 15, yielding an average minimal
nucleotide difference per codon of 0.79. In the comparison of
alfa and C. pasteurianum ferredoxins (Table VII) the simi-
arity in the same regions of the primary structures can be ex-
tended to segments of 28 residues. Thus, Residues 42 to 69 of
the alfalfa protein are at a minimal mutation distance of 23
from Residues 9 to 36 of the C. pasteurianum protein, corre-
sponding to an average minimal nucleotide difference per codon of
0.82. This result leads to an interesting conclusion. As noted
above, all clostridial ferredoxins show strong evidence of having
resulted from a duplication which left the amino-terminal 29-
residue sequence similar to the carboxyl-terminal 26-residue se-
quence (13, 14). Thus, if one accepts that plant and clostridial
ferredoxins have descended from a common ancestral form, the
fact that the similarity of the alfalfa and C. pasteurianum ferre-
doxins extends to both similar portions of the clostridial protein
in a continuous sequence, must mean that the divergence of the
lines of descent leading to plant and clostridial ferredoxins oc-
curred later than the duplication which yielded the structure of
the clostridial proteins.

That plant and clostridial ferredoxins may indeed be homolo-
gous (10) can be derived from a systematic comparison of the
minimal mutation distances of all possible 25-residue segments of
alfa ferredoxin with 25-residue segments of C. pasteurianum
ferredoxin, plotted according to method of Fitch (40) (Fig. 8).
This demonstrates that the degree of similarity between the two
proteins is clearly greater than would be expected on the basis of
random occurrences, as shown by the deviation of the cumulative
frequency of minimal mutation distances, at the lower left of
Fig. 8, from the straight line which represents the random dis-
tribution (40). It must be emphasized that this result does not
necessarily mean that plant and clostridial ferredoxins are
ancestrally homologous, since, in addition to demonstrating the
non-randomness of their degree of similarity, one would also have
to show that this similarity is more extensive than required by the
similar functions of these two sets of proteins (36).

Acknowledgments—We wish to thank Dr. Walter M. Fitch for
the analysis of the amino acid sequence of alfalfa ferredoxin for
internal duplications and for its statistical comparison with the
structure of C. pasteurianum ferredoxin. We are grateful to Dr.
R. M. Love of the University of California at Davis, and
to Dr. W. F. Lehman of the Imperial Valley Field Station of
the University of California, for their generous help in supplying
the Sonora variety of alfalfa. The skillful technical assistance of
Mr. T. Tanner is gratefully acknowledged.

FIG. 8. Comparison of minimal mutation distances (MUTA-
tions Required) for all possible 25-residue segments of
alfa and C. pasteurianum ferredoxins, according to the pro-
cedure of Fitch (40). The cumulative frequencies of the minimal
mutation distances occurring in the 2263 comparisons are plotted
on a probit scale on the ordinate (CUMULATIVE FREQUENCY
OF MUTATIONS REQUIRED). The random portion of
the distribution is represented by the straight line and where a given
minimal mutation distance occurs more frequently than ex-
pected for a random distribution it produces a departure from
linearity to the left. The probability that the observed degree of
non-randomness would occur by chance is less than 0.02 as
estimated according to an unpublished method of Dr. W. M.
Fitch.

1 Tsunoda, J. S., Whiteley, H. R., and Yasunobu, K. T., Ab-
2 F. Perini and E. Margoliash, unpublished results.
3 Keresztes-Nagy, F. Perini, and E. Margoliash, unpublished results.
REFERENCES

Primary Structure of Alfalfa Ferredoxin
S. Keresztes-Nagy, F. Perini and E. Margoliash

J. Biol. Chem. 1969, 244:981-995.

Access the most updated version of this article at http://www.jbc.org/content/244/4/981

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/244/4/981.full.html#ref-list-1