Durham University

Durham E-Theses

Molecular evolution of chloroplast type Ferredoxins

Ishaq Abdel - Hamid Takruri

How to cite:

Takruri, Ishaq Abdel - Hamid (1979) Molecular evolution of chloroplast type Ferredoxins. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/8372/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

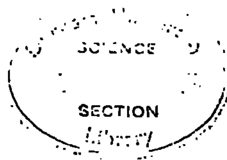
MOLECULAR EVOLUTION OF CHLOROPLAST-TYPE FERREDOXINS

A Thesis
submitted in accordance with
the requirements of the University
of Durham for the Degree of
Doctor of Philosophy

By

ISHAQ ABDEL-HAMID TAKRURI

January 1979



Department of Botany

The copyright of this thesis rests with the author.
No quotation from it should be published without
his prior written consent and information derived
from it should be acknowledged.

TO

My Parents and

My brother Abdel-Aziz A. Takruri

CONTENTS

Page No.

ACKNOWLEDGEMENTS I

SUMMARY II

INTRODUCTION 1

MATERIALS AND METHODS

- I. General 12
- II. Purification methods 18
- III. Protein sequence determination 23
- IV. Calculations based on amino acid sequence data 32

RESULTS

- I. Isolation and purification of ferredoxins 33
- II. The amino acid sequence of *Porphyra* ferredoxin 35
- III. The amino acid sequence of *Triticum* ferredoxin 44
- IV. The amino acid sequence of *Sambucus* ferredoxin 53
- V. The amino acid sequence of *Brassica* ferredoxin 60
- VI. Partial amino acid sequence of *Hordeum* ferredoxin 67

DISCUSSION

- I. Technical problems in protein sequencing 68
- II. Comparison of ferredoxin sequences 71
- III. Evolutionary aspects of the sequence data obtained 77

BIBLIOGRAPHY

APPENDIX

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Professor. D. Boulter for his continued help and guidance at all stages in this work and for the use of the facilities of the Botany Department in the University of Durham.

I also wish to thank Dr. B.G. Haslett for his guidance and for much helpful advice during the first year of my work and Dr. T. Gleaves for writing the ancestral sequence method tree construction program.

I am grateful to Mr. J. Gilroy for assistance with the extraction and purification of ferredoxin, for Mrs. R. McIntosh for assistance with the dansyl-Edman analysis and Mr. T.G. Krishna and Mrs. C. Webb for typing this thesis.

I should like to acknowledge the financial support received from the British Council.

SUMMARY

Ferredoxins have been extracted and purified from four species of higher plants, *Triticum aestivum* (Wheat), *Brassica napus* (Rape), *Sambucus nigra* (Elder) and *Hordeum vulgare* (Barley). The complete amino acid sequence of the first three ferredoxins and that of *Porphyra umbilicalis* (Red algae) have been determined; a partial sequence of the ferredoxin of *Hordeum vulgare* was also obtained. In case of Elder, owing to limitation on the amount of pure protein obtainable, five residues were placed by homology.

The sequence data obtained in this investigation show that ferredoxin from the red algae are homologous with other eukaryotic chloroplast ferredoxins and also with prokaryotic blue green algae ferredoxins. However, the red algae are a divergent group. Sequence comparisons were used to examine the divergence of various taxonomic groups and to construct phylogenetic trees using a modified ancestral sequence method. However, this method did not give a unique phylogenetic tree, but the alternative trees obtained were all similar.

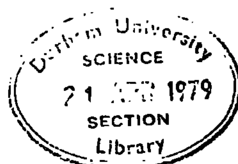
In general, the ferredoxin data support the outline of the existing evolutionary tree. Thus, the blue green algae and red algae are on a separate branch to the green algae and higher plants. *Equisetum* (fern allies) are seen as evolutionary offshoots and the higher plants group together. It is in considering the interrelation-

ships of major higher plant groups that disagreements appear. The present data set, even taken in conjunction with existing plant cytochrome c and plastocyanin data, is insufficient to establish a new phylogeny. However, the molecular data show sufficient consistency to indicate that the existing schemes e.g., Cronquist are suspect.

INTRODUCTION

Taxonomists can use either phenetic or phyletic classificatory methods to group organisms (Sokal & Sneath, 1963; Cronquist, 1968). Phenetic classifications group organisms by using a similarity or dissimilarity index of characters whereas phyletic classifications aim to reflect the evolutionary relationships of the different organisms. However, the only certain way to establish phylogeny is from an adequate fossil record. This approach has been applied to the vertebrates of the animal kingdom (Romer, 1966; Collert, 1969) with good success. Unfortunately, the fossil record of plants including algae is inadequate and often ambiguous (Briggs and Walters, 1969) for this purpose. As a result, plant taxonomists have resorted to the use of comparisions of characters of living organisms. In this connection, it is generally believed, although not often stated, morphological characters are much less reliable markers of evolutionary history of the genes than are the amino acid sequences of specific homologous proteins because the primary structure of proteins reflect the sequence bases of the structural DNA (Swain, 1974).

Initially, it was hoped that the use of biochemical data would avoid some of the limitations inherent in the use of morphological data (Zuckerkandl and Pauling, 1965). However, as information has accumulated it is


SCIENCE
SECTION
Library

apparent that biochemical data suffer from the same limitation as do other chemical data. At first, the presence or absence of particular micromolecules were used to suggest polygenetic evolutionary schemes. More recently, more detailed characteristics of particular micromolecules have been compared. More or less from the outset, Zuckerkandl & Pauling (1965) hypothesised that the total chemical content of any species was a document of its evolutionary history. They classified molecules on the relative amount of information they contained as follows:

First Class: the semantides, consist of molecules which carry information or transcripts of that information either in the form of nucleic acids or proteins.

Second Class: the episemantic molecules which are synthesised under the control of semantides.

Third Class: asemantides molecules which are not produced by the organism and do not express any information about the organism.

They suggested that the amount of evolutionary history preserved in the molecules of greater complexity, **namely,** the semantides have the largest information content. Semantides were divided into three sub-categories: the primary semantides which are the genes

of the organism ; secondary semantides comprising messenger and transfer RNAs and tertiary semantides comprising the proteins.

Clearly macromolecules such as nucleic acids and proteins would appear to be more powerful characters than micromolecules since they are a direct transcript of the genes and less liable to environmental influences. It has been established that differences between species, and hence the main underlying processess of evolution giving rise to them are due to dissimilarity of their genes, that is, the variation of the base sequences of their informational chromosomal DNA. Therefore, the pathway of evolution should be fully understood when the base sequences of DNA of selected taxa are known. In order to obtain information about the sequence or degree of homology of DNAs of different species, three main approaches can be followed. First, hybridisation methods (see Southern, 1970). Second, comparing the base sequences of RNAs or transfer RNAs which are transcribed from DNA. Thirdly, comparing the amino acid sequences of specified protein which are translated copies of the base sequence in the corresponding gene. The detailed chemical analysis, characterisation and sequencing of primary and secondary semantides (DNA and RNA) has been rarely used specifically to examine relationships between groups of organisms. The base sequences of a small number of tRNA molecules (see

Dayhoff, 1972 & 1976) have been determined for this purpose and DNA hybridisation techniques (Kohne, 1968) have been used also to some extent to examine relationships between organisms (Lea & Norris, 1972), but none of these techniques have so far been applied to a wide range of plants because of technical difficulties. The results obtained from hybridisation techniques are particularly difficult to interpret. More recently, the use of restriction endonucleases and recombinant DNA has allowed the possibility to clone specific pieces of DNA and produce DNA in relatively large amounts. This coupled with the development of fast methods to determine the base sequences of DNA gives the possibility to sequence comparable genes from different organisms relatively rapidly once the body of techniques has been mastered. In the future therefore the use of this method is likely to complement protein sequencing and may even to some extent supplant it (Malcolm, 1978).

Comparisons of homologous tertiary semantides, proteins, from different species with a view to deducing taxonomic and phylogenetic relationships have been receiving increasing attention (Wilson, et al., 1977). Proteins have an added advantage over the nucleic acids because of their greater structural variability and the fact that most of them possess enzymic activity (see Swain, 1974) so making detection of homology simpler. The implication of proteins in taxonomic studies has

fallen into three areas of research. Protein immunology (Serology) (Fairbrothers, 1968; Wilson et al., 1977), protein electrophoresis (comparative biochemistry) (Boulter et al., 1966; Vaughn, 1968a) and protein sequences (Dayhoff, 1969; 1972; 1973; 1976; Peacock & Boulter, 1975; Boulter, 1972).

Serological techniques were the first methods to be used to compare homologous proteins from different species in an attempt to establish possible phylogenetic relationships. Serology in plant taxonomy was pioneered by Metz in the 1920s (see Chester, 1937). Metz and his collaborators established phylogenetic trees of the plant kingdom, but later a rival school produced a completely different set of relationships (see Davis & Heywood, 1963). The application of modern serological techniques has been applied successfully however for certain taxonomic problems (Vaughn, 1968a and b). Phylogenetic trees have recently been established for a wide variety of organisms on the basis of immunological comparisons of their proteins (Ho, et al., 1976) and with plants, Lee & Fairbrothers (1978) have constructed phylogenetic relationships among the Rubiaceae and other related families based on serological evidence.

Protein electrophoretic separations and comparative biochemical studies provide a further approach to taxonomic problem. Boulter et al., (1967) have

discussed the separation and analysis of plant proteins by zonal electrophoresis on gels and the use of band data on gels in order to compare proteins of different species so as to establish phylogenetic relationships. This method of analysis is of particular value to the chemotaxonomist because a rapid separation, high resolution and simultaneous analysis of many samples can be carried out.

During the last two decades much attention has been given to the usefulness of protein sequences in deducing phylogenetic relationships (Swain, 1974; Matsubara and Yamanaka, 1978). Thus it has become quite possible to determine the primary structure of many different proteins. The "Atlas of protein sequence and structure" for 1966 (Dayhoff & Eck, 1966) reported 184 complete primary structures of proteins and related macromolecules. By 1972 the same publication (Dayhoff, 1972) has listed more than 350 sequences and about 150 more were added in a supplement for 1973 (Dayhoff, 1973). In addition amino acid sequences of several proteins have been extensively used to study evolution and relationships between different species (see Dayhoff, 1972, 1973) but in most cases animal proteins have been used. However, a limited number of bacterial, algae and higher plant proteins have also been studied. For example, cytochrome C and plastocyanin proteins of

many higher plant species have been well studied by Boulter and his colleagues (1972, 1977).

Possibly the most important group of proteins from an evolutionary point of view is the iron-sulphur proteins. Knowledge of their biological function, mainly as electron carriers in many biological reactions such as photosynthesis and nitrogen fixation and their molecular properties and primary structures has grown rapidly during the last two decades (see Lovenberg, 1973 Vol. I & II). So far as the primary structures are concerned, a limited number of amino acid sequences of ferredoxin, rubredoxin and adrenodoxins have been determined (Dayhoff, 1972, 1973; Lovenberg, 1973). However, ferredoxins have received most attention probably because it is the most widespread protein of this group and also it was the first iron-sulphur protein whose sequence was determined, that is that of C. pasteurianum (Tanaka et al., 1966). From the present standpoint, ferredoxins offer the distinct advantage that they are readily purified and are of relatively small molecular weight so simplifying the sequencing process. They are also found in many different organisms. Thus they have been found in all organisms studied from the obligate anaerobic bacteria, the algae and higher plants (Wada et al,, 1975; Hall & Rao, 1977; Matsubara et al., 1978). They may well have been one of the first proteins

to be formed during evolution and ferredoxins therefore allow us to trace evolution much farther back than many other proteins (Hall et al., 1975; 1973). More than one ferredoxin occurs in some organisms (see Wakabayashi et al., 1978; Yoch et al., 1977) and it will be interesting to compare these proteins.

Ferredoxin was initially isolated from different sources under different names. Davenport et al., (1952) isolated the methemoglobin-reducing factor; Arnon et al., (1957) isolated TPN^+-reducing factor from spinach chloroplast; San Pietro and Lang (1958) isolated a photosynthetic pyridine nucleotide reductase. By 1961, it became clear that these were different names for the same protein which Mortenson et al., (1962) called ferredoxin.

Two distinct types of ferredoxins were recognised, the bacterial-type ferredoxins which are split into two sub-types, photosynthetic and non-photosynthetic ferredoxins, and the chloroplast type ferredoxins. These two types of ferredoxins share several properties, (a) possessing iron and labile sulphur in equimolar amounts (b) being strongly acidic (c) possessing low redox potential (about -400 mV at pH 7.0). They differ, however, in spectral characteristics, in the amino acid composition, in molecular weight and in the number of iron-sulphur groups. In each case the iron atoms are linked with the protein through the sulphur of cysteine

residues, and the labile sulphur forms a link between the iron atoms (Rao et al., 1971; Hall et al., 1973) Ferredoxins containing eight iron, eight sulphur atoms arranged as two 4Fe : 4S active centres are found probably only in anaerobic and photosynthetic bacteria (Adman et al., 1973). Such ferredoxins accept two electrons on reduction. However, Desulphovibrio and Bacillus ferredoxins have only a single 4Fe : 4S active centre and transfer one electron on reduction (Le Gall & Dragoni, 1966; Travis et al., 1971; Orme Johnson et al., 1972). These ferredoxins may represent intermediate stages in evolution between the anaerobic bacteria and higher plants which have 2Fe : 2S active centre and accept one electron on reduction (see De Lay and Kersters 1974) (see Appendix 1).

Ferredoxins are located in the chloroplast of plants and also green and red algae and in the photosynthetic lamellae of the blue green algae.Their major role is to carry electrons in the photosynthetic reactions. Thus they act as electron carriers in cyclic and non-cyclic photophosphorylation and also in photoreduction of NADP (Arnon, 1977; Buchanan & Arnon, 1970; 1971).

Prior to this study, several ferredoxin sequences were known (see Wakabayashi et al., 1978 and discussion of this thesis). In addition to this sequence data the immunochemical characteristics of ferredoxins have also been investigated and this method provides a comparatively powerful biochemical method for the understanding of taxonomy and evolution of this

important protein (Matson et al., 1975; Huisman et al., (1977). Tel-Or et al., (1977) studied the ferredoxin-antiferredoxin interrelations using antibodies of ferredoxins from five species of higher plants, algae and bacteria. The antibodies were cross reacted with ferredoxins from 30 organisms to obtain possible phylogenetic affinities of the organisms from which the ferredoxins were obtained. They found that there was a correlation between the degree of immunoprecipitation and the similarity of their amino acid sequences. Ferredoxins with identical primary structure isolated from different species could not be distinguished. Huisman and his coworkers (1977) in comprehensive studies of antigenic cross-reactivity show that a complete cross-reaction was observed between antibodies of Nicotiana tabacum with ferredoxins from five other Nicotiana species, thus it was impossible to distinguish between these ferredoxins in this manner. Furthermore, Tel-Or et al. (1977) showed that four substitutions in the amino acid sequences between Spirulina maxima and Spirulina platencis had no significant effect on the antigenic cross-reaction suggesting that these techniques are inadequate for the construction of close phylogenetic relationship.

The aim of this investigation, whose results are presented here, was to purify and sequence the chloroplast type ferredoxins from various species and

to use these sequence data together with those already published (Dayhoff, 1972; 1973; 1976) to construct a phylogenic tree.

A large scale approach to the purification was adopted using similar strategies to those developed for ferredoxins from spinach by Petering & Palmer (1970). The sequence methodology was similar to that devised by Gray & Hartley (1963a,b) and used cyanogen bromide clevage followed by proteolytic digestion of the separated fragments as well as enzymic digestion directly on the intact protein. The complete sequence data together with that available in the literature were then subjected to ancestral amino acid sequence analysis to generate phylogenies (Dayhoff, 1972).

As in other investigations, difficulty in obtaining material either due to its relative inaccesibility in nature or because of extraction and purification problems has meant that the final amino acid sequence data set has not been phylogenetically ideal.

MATERIALS AND METHODS

I. GENERAL

(1) Biological Materials

Purified *Porphyra umbilicalis* ferredoxin was obtained from Lyndon J. Rogers, Department of Biochemistry and Agricultural Biochemistry, University College of Wales, Aberystwyth.

Leaves of *Triticum aestivum* (wheat) and *Hordeum vulgare* (barley) were collected locally from Houghall School of Agriculture.

Sambucus nigra L. (elder) was collected locally in Durham. Seeds of *Brassica napus* L. (rape) were obtained from the Tyneside Seed Stores, Gateshead.

(2) Chemicals and Reagents

All chemicals, except the ones listed below, were obtained from British Drug Houses (BDH) Limited, Poole, Dorset, and were of analytical grade when available.

Trypsin	E.C.3.4.4.4.	(twice recrystallized)
Chymotrypsin	E.C.3.4.4.5.	(thrice recrystallized)
Papain	E.C.3.4.4.10.	(twice recrystallized)

were obtained from the Worthington Biochemical Corporation, Freehold, New Jersey, U.S.A.

Thermolysin E.C.3.4.24.4. (crystalline) was obtained from Daiwa Kasei K.K., Osaka, Japan.

Carboxypeptidase A E.C.3.4.2.1. (di-isopropylphosphorofluridate, crystalline suspension in water)

Carboxypeptidase B E.C.3.4.2.2.

Carboxypeptidase C

4-Dimethylamino benzaldehyde

were obtained from Sigma (London) Chemical Co., London SW6, U.K.

Sephadex DEAE A-50

Sephadex G-10

Sephadex G-50

Sephadex G-75

were obtained from Pharmacia Ltd., Uppsala, Sweden.

Biogel P-10

Biogel P-4

were obtained from BioRad Laboratories Ltd., London.

Guanidine hydrochloride, sequanal grade

was obtained from Pierce, Rockford, Illinois 61105, U.S.A.

Ninhydrin (Indantrione hydrate)

was obtained through Koch-Light Laboratories Limited, Colnbrook, Bucks., England.

Pyridine

was obtained through Rathburn Chemicals, Walkerburn, Peebleshire.

Arginylarginine

was obtained from Cyclochemical Corporation, Los Angeles, California, U.S.A.

(3) Other Materials

Polyamide sheets were obtained from

(a) Cheng Chin Trading Co.,Ltd., Tapei, Taiwan, through BDH Chemicals Ltd.

or (b) Pierce, Rockford, Illinois 61105, U.S.A.

(4) Preparation of Solutions

a. Buffer solutions:

(i) 10 mM Tris-HCl, Tris-(hydroxymethyl)-aminoethane-HCl buffer, pH 7.2 for protein extraction.

1.21 g Tris was dissolved in 500 ml distilled H_2O

0.1 M HCl was added slowly to make the pH 7.2

2 ml EDTA were added

the vol. was made up to 1000 ml with distilled H_2O

(ii) 150 mM Tris-HCl, 1 M NaCl buffer, pH 7.2, for protein elution.

18.15 g Tris was dissolved in 500 ml H_2O

0.1 M HCl was added slowly to make the pH 7.2

58 g NaCl were dissolved in the buffer

the vol. was made up to 1000 ml with distilled H_2O.

b. Electrophoretic and Chromatographic location reagents:

(i) Cadmium ninhydrin reagent (Heilman, *et al.*, 1957)

Solution A. 100 mg Cadmium acetate
10 ml H_2O
5 ml acetic acid
100 ml acetone

Solution A was used to prepare a fresh 1% (w/v) ninhydrin solution in which the electrophoretic and chromatographic papers were dipped and allowed to dry at room temperature. Electrophoretic paper was heated at 105°C for 2-3 min. when colored spots on a white background showed a positive reaction. These spots were marked before the paper was heated again to locate slowly developing spots.

(ii) Ehrlich reagent:

2% (w/v) p-dimethylaminobenzaldehyde in 20%(v/v) HCl in acetone was prepared just before required. The paper was dipped in the above solution and allowed to dry at room temperature. A purple color showed a positive reaction with tryptophan. Greater sensitivity was obtained when this test was followed by ninhydrin staining (Easley, 1965). The pink ninhydrin spots become colorless and a positive purple color appeared.

c. Electrophoretic separation

(i) Standard marker solution

Arginylarginine was dissolved in 1 M $NaHCO_3$ to give a 0.1 M solution and this was treated with an equal vol. of 0.2 M dansyl chloride in acetone. After 1 hr. at 37° C, the mixture was diluted 1000-fold and ethanolic dansyl-arginine was added to give a concentration of 0.1 mM.

(ii) Electrophoresis buffers

pH 6.5	Pyridine	25 ml
	acetic acid	1 ml
	water	225 ml

pH 1.9	acetic acid	20 ml
	formic acid	5 ml
	water	225 ml

d. Chromatographic separation

(i) Marker solution

0.1 mM solution of dansyl-arginine in 95% (v/v) ethanol.

(ii) Solution used, BAWP solvent

Butan-1-ol	75 ml
acetic acid	15 ml
pyridine	50 ml
water	60 ml

e. Chromatography of dansyl derivatives

(i) Standard marker solution

0.1 mg/ml of dansyl-proline, dansyl isol-leucine, dansyl-phenylalanine, dansyl-glycine, dansyl glutamic acid and dansyl-serine in 95% (v/v) ethanol.

(ii) Chromatographic solvents (Ramshaw *et al.*, 1970)

Solvent A	1.5% (v/v) formic acid
Solvent B	Toluene: Acetic acid (9:1, (v/v))
Solvent C	Butylacetate: Methanol: Acetic acid (30: 20: 1, (v/v))

f. Amino acid Autoanalyser solutions

(i) Citric acid solution buffer, pH 3.25

2 M citric acid solution	100 ml
pellets NaOH (Aristar)	16 g
thiodiglycol	10 ml
Brij 35 solution	5.4 ml
sodium n-octanoate	0.2 g

Adjusted to pH 3.25 with HCl, made upto 2 liters with distilled H_2O, and stored at 2° C.

(ii) Citric acid solution, pH 4.25 buffer

2 M citric acid solution	100	ml
pellets NaOH (Aristar)	16	g
thiodiglycol	10	ml
Brij 35 solution	5.4	ml
sodium n-octanoate	0.2	g

Adjusted to pH 4.25 with HCl, make upto 2 liters with distilled water and stored at 2°C.

(iii) Citric acid solution, pH 8.0 buffer

2 M citric acid solution	210	ml
pellets NaOH (Aristar)	50.4	g
thiodiglycol	10	ml
Brij 35 solution	5.4	ml
sodium n-octanoate	0.2	g
sodium chloride	43.2	g

Adjusted to pH 8.0 with NaOH, make upto 2 liters with distilled H_2O, and stored at 2°C.

II. PURIFICATION METHODS

A generalized purification scheme is shown in Fig. 1.

(1) Ferredoxin assay

Ferredoxin was assayed qualitatively in solution by its U.V absorption spectrum since it exhibits absorption maxima at about 278, 330, 420 and 465 nm.

Ferredoxin was quantitatively estimated spectrophotometrically by measuring the absorbance of the oxidised and reduced protein at 420 nm. Assuming the molecular weight of thr protein to be 11,000 and using the extinction coefficient of 0.835 $(mg/ml)^{-1}cm^{-1}$ at 420 nm Tagawa & Arnon (1962).

The purity of the protein was estimated spectrophotometrically from the ratio (R) of the absorbance at 420 nm to that at 278 nm (R= A(420)/A(278)). This ratio was compared with those of ferredoxins of known structure and purity. The purity of carboxymethylated ferredoxin was ascertained by $-NH_2$ and -COOH terminal analysis.

(2) Ferredoxin extraction

All preparations were carried out in the cold room at 0-4 C. Fresh plant leaves were collected and used immediately. Each 500 g leaves were placed in a 5 litre blender and homogenised with 1 litre of 10 mM Tris-HCl buffer at pH 7.2 with 2 ml of (10 mM) EDTA added. Homogenisation was continued for 2-3 min. The homogenate

FIGURE 1.

Generalised scheme for the purification of ferredoxin from plants

Fresh material was blended with 10 mM Tris-HCl pH 7.2

Filtered through muslin and centrifuged; supernatant taken

Adsorbed on DEAE-cellulose.

Eluted with 150 mM Tris-HCl, 1 M NaCl, pH 7.2

Ammonium sulphate fractionation.

Dialysed.

Concentrated on DEAE-cellulose resin and diaflo.

Gel filtration on sephadex G-75 or Biogel P-10.

Chromatography on sephadex DEAE A-50.

Desalted on Amberlite MB-1 and lyophilised.

was filtered through two layers of muslin and centrifuged at 12,000 x g for 20 min (MSE, 4L).

(3) DEAE- Cellulose adsorbtion

7.5 ml of well settled DEAE- 23 cellulose resin which was equilibrated in 10 mM Tris-HCl buffer at pH 7.2 was added to each one liter of the dark green ferredoxin-containing supernatant and stirred gently for about 1 h. The dark green cellulose resin was collected by centrifugation at 2,000 rpm for about 2 min. (MSE, 4L). The resin containing the protein was repeatedly washed with 10 mM Tris-HCl buffer, pH 7.2 and collected after each wash by centrifugation at 120 x g for about 30 sec. (MSE, 4L).

(4) Elution

The resin containing the ferredoxin was packed into a column and washed with 1-2 l of 10 mM Tris-HCl buffer at pH 7.2. The protein was eluted with 150 mM Tris-HCl buffer, pH 7.2, containing 1 M NaCl. Fractions with distinct brown color were collected and pooled together.

(5) Ammonium sulphate fractionation

Solid ammonium sulphate was added slowly with stirring to the brown ferredoxin solution to give 85% saturation, (559 g/ litre). The degree of saturation was determined by using the nomogram of Dixon (1953). After 2-3 h stirring the precipitate was removed by centrifuging at 10,000 rpm for 20 min. (MSE. High-speed 18 centrifuge, 6 x 250 ml angle rotor). The supernatant was then dialysed.

(6) Dialysis

The combined supernatant was dialysed twice against ten vol. 10 mM Tris-HCl buffer, pH 7.2, for total of at least 8 h. The solution was diluted 2-fold with distilled H_2O.

(7) Concentration

The dialysed solution was loaded onto a column of DEAE- 23 cellulose equilibrated with 10 mM Tris-HCl buffer at pH 7.2. The column was then washed with 0.5-1 litre of 10 mM Tris-HCl buffer. Ferredoxin was eluted with 150 mM Tris-HCl buffer, pH 7.2, containing 160 mM NaCl and 10 mM mercaptoethanol. The eluted ferredoxin solution was concentrated on the Amicon Diaflow apparatus to about 5 ml using a UM- 2 membrane.

(8) Gel filtration

The column (3 cm X 90 cm) was prepared by pouring a slurry of either Biogel P-10 or Sephadex G-50 equilibrated in 50 mM Tris-HCl buffer, pH 7.2 made 50 mM with KCl. The concentrated sample (volume of 3-6 ml) was added and the column developed at a flow rate of 40-60 ml/h. The protein fractions collected from the column were analysed for spectral ratios ($R = \frac{A\ (420)}{A\ (278)}$) as an index of purity. At this stage fractions containing ferredoxin exhibited a low value of (R) (less than 0.35), fractions were pooled and concentrated using a diaflo (UM-2) membrane.

(9) DEAE A-50 Sephadex chromatography

The concentrated solution containing the protein was adsorbed onto a column (1.5 cm X 15 cm) of DEAE sephadex equilibrated with 50 mM Tris-HCl buffer, pH 7.2 containing 50 mM KCl. The protein was eluted under a linear ionic gradient (from 50 mM Tris-HCl buffer and 50 mM KCl (300 ml) to 50 mM Tris-HCl buffer and 450 mM KCl (300 ml)). The gradient was constructed using an apparatus as described by Bock & Ling (1954). The ferredoxin collected was pure as indicated by its (R) ratio and other criteria (see step 11). Fractions containing the pure ferredoxin were pooled and concentrated using a diaflo (UM-2) membrane.

(10) Desalting by an ion-exchange method

Pure concentrated ferredoxin was desalted in 2 ml batches on a 1 cm X 15 cm column of Amberlite MB-1 resin equilibrated in distilled water. Salt free ferredoxin was lyophilised and stored at -20°C.

(11) Homogeneity of the protein

Isoelectric focusing on gels was carried out to determine the homogeneity of the ferredoxin as described by (Wrigley (1968)). The gels were 7 mm in diameter and 100 mm in length and contained 5% (w/v) polyacrylamide, 2% (w/v) Ampholine (pH 3-10; pH 3-5). Samples of the protein (400-600 ug) were applied to the upper (cathodic) end of polymerised gels in 100 ul of 5% (w/v) Ampholine, 5% (w/v) sucrose in 6M urea. A protective layer of 5%

(w/v) ampholine, 5% (w/v) sucrose in 6M urea was applied over the samples. The upper electrode compartment (cathode) was filled with 0.4% (v/v) ethanolamine; the lower compartment (anode) contained 0.2% (v/v) H_2SO_4. A current of 2 mA/gel was applied until the voltage reached 300 V, then voltage was maintained at 300 V for further 5 h. After focusing, the protein zones were detected as white precipitation bands by immersion of gels in 25% (w/v) trichloroacetic acid. The isoelectric points (pI) of the bands were determined by measuring the pH gradient in unfixed gel electrofocused simultaneously with fixed gels. The pH gradient in the unfixed gel was measured by cutting the gel (into 10 mm slices) and soaking each slice in 2 ml water for 12 h and then measuring the pH.

III. PROTEIN SEQUENCE DETERMINATION

(1) Carboxymethylation of ferredoxin

Ferredoxin (15-20 mg) was denatured in 2 ml of 6M guanidinehydrochloride-1 M Tris-HCl, pH 8.5 and carboxymethylated as described by Milne & Wells (1970).

Carboxymethylated protein was desalted by passage through a 1.5 cm X 30 cm column of sephadex G-10 equilibrated with distilled water. Desalting was carried out in a dark room. Eluted protein was subsequently freeze-dried.

(2) Cyanogen bromide (CNBr) cleavage

CNBr cleavage was carried out by the method of Steers *et al* (1965). Lyophilised carboxymethylated ferredoxin (10-20 mg) was dissolved in 500 ul of 85% (v/v) formic acid (*Porphyra* ferredoxin) or 70% (v/v) formic acid (Wheat ferredoxin). An 8-fold excess (w/w) of CNBr was added and kept for 24 h in a dark place at room temperature. The reaction was terminated by lyophilisation and the fragments were separated by gel chromatography.

i. A col. (1.5 cm X 130 cm) of sephadex G-50 superfine grade equilibrated with 85% (v/v) formic acid was used to separate the CNBr fragments (Porphyra).

ii. A col. (1 cm X 190 cm) of sephadex G-75 equilibrated with 70% (v/v) formic acid was used to separate CNBr fragments (Wheat).

Elution profiles were followed by measurement of the

A_{280} and A_{206} with an LKB Uvicord III instrument and by N-terminal analysis by the dansyl technique of Gray & Hartley (1963a).

(3) Proteolytic digestion

(a) Chymotryptic or Tryptic digestion

The carboxymethylated protein or CNBr fragment (10-20 mg) was dissolved in 250 ul of 0.2 M N-ethylmorpholine acetate buffer at pH 8.5. 1 mg of enzyme was dissolved in 200 ul buffer and 40-80 ul of the enzyme solution was added to the protein to give 2% (w/w) enzyme/protein and incubated for 1-2 h at 37° C.

The enzymic digestion was terminated by the addition of an equal vol. of acetic acid and freeze-dried.

(b) Thermolysin digestion

The digest conditions and enzyme concentration were as for the chymotryptic digestion except that the buffer contained 5 mM $CaCl_2$.

(c) Papain digestion

Peptides (1-2 uM) were dissolved in 500 ul of 0.1% (v/v) mercaptoethanol in pyridine acetate buffer, pH 6.5. The enzyme (30 mg/ml saline suspension) was added to give 2% (w/w) enzyme/peptide. Incubated for 1-2 h at 37° C. The digestion was terminated by adding an equal vol. acetic acid and then freeze-dried.

(4) Peptides purification

(a) Column chromatography

Peptides resulting from tryptic or chymotryptic

digestion were subjected to gel chromatography on a 1 cm X 190 cm col. of Bio-Gel P-4 equilibrated with 70% (v/v) formic acid except for peptides of Porphyra which were separated by using a col. (1.5 cm X 130 cm) of sephadex G-50 superfine equilibrated with 85% (v/v) formic acid. Eluted fractions were followed by measuring the absorbance at 280 nm and 206 nm with LKB Uvicord III instrument and by the N-terminal analysis with the technique of Gray & Hartley, (1963a).

(b) Electrophoresis

Peptides, which were not pure after column chromatography, were separated as described by Thompson et al (1970) by high-voltage electrophoresis at pH 6.5 (pyridine acetic acid - water, 25:1:225, by vol.) on a flat-plate apparatus (107 cm X 15 cm; The Locarte Co., London S.W.3, U.K.). Separation was achieved on Whatman 3MM paper (15 cm) using a voltage of 9 KV to give a current of 40-50 mA at 7 p.s.i. pressure for 60-90 min.

(c) Paper chromatography

Peptides were separated by descending chromatography using BAWP solvent system (butanol - acetic acid - water pyridine, 75:15:60:50 by vol.). Samples were chromatographed on Whatman 3MM chromatography paper for 20 h. at room temperature in a Gallenkamp chromatography frames and bags.

(d) Peptides location

i. Cadmium ninhydrin reagent (Heilman et al, 1957). Peptides were located using 10% strips of the electro-

phoretic or chromatographic paper. The paper- guide strip was dipped in the reagent and allowed to dry at room temperature. Electrophoresis strips were dried in 105° C for 2-3 min.

ii. Ehrlich reagent

Electrophoretic and chromatographic guide strips were passed through freshly prepared Ehrlich reagent (2% 4-dimethylaminobenzaldehyde in 20% (v/v) HCl in acetone). The pink ninhydrin spots become colorless and purple color appears indicating a positive reaction with tryptophan.

(e) Peptide mobilities

At pH 1.9 peptide mobilities were measured from 1-dimethylaminonaphthalene-5-sulfonic acid relative to dansyl arginine and from dansyl arginine relative to dansyl arginylarginine at pH 6.5

(f) Elution of peptides

Peptides were eluted with 20% (v/v) pyridine into (1 cm X 5 cm) screw-cap tubes. The samples were freeze dried *in vacuo* over NaOH.

(5) Quantitative amino acid composition of proteins, CNBr fragments and peptides.

The amino acid composition of protein, CNBr fragment and peptide were carried out with a Locarte amino acid analyser.

Protein (0.03-0.05 μM), CNBr fragment (0.05-0.15 μM) or peptide (0.05-0.15 uM) were hydrolysed with 0.5 ml of 6M HCl at 105° C in evacuated, sealed pyrex tubes.

Protein samples were hydrolysed for 24, 48, 72 h and CNBr, fragments for 24, 72 h to calculate the zero time values for threonine and serine and the maximum values for valine and isoleucine (Moore & Stein, 1963). Peptides were hydrolysed for 24 h only. After hydrolysis the samples were dried in vacuo and kept at -20° C until analysed.

Cysteine was calculated as cysteic acid with 6M HCl at 105° C for 24 h (Hirs, 1956).

Tryptophan was estimated after alkaline hydrolysis as described by Noltmann et al (1962).

(6) Semi-quantitative amino acid composition of peptides

A purified peptides (30-50 nM) were dried in vacuo over NaOH in Durham tubes (6 mm X 30 mm, Gallenkamp Limited, London), 50 ul of 6M HCl added and hydrolysed for 18 h at 105° C. The acid was removed by drying in vacuo over NaOH and the free amino acids were labelled by the dansyl method of Gray & Hartley (1963a) and the dansyl amino acids identified by polyamide sheet chromatography (see materials and methods, 7 (iv)).

(7) Methods of peptide sequencing

a. Manual sequencing methods

Peptide sequences were determined by using the method of Gray & Hartley (1963b) as described by Thompson et al (1970). The dansyl amino acids were identified by polyamide sheets Woods & Wang (1967), using the solvent

system described by Ramshaw *et al* (1970).

(i) N-Terminal method

The dansyl-Edman procedure of Gray & Hartley (1963b) was used. 10% of the peptide material was used to determine the N-terminal amino acid after each cycle of Edman degradation.

(ii) C-Terminal method

C-terminal sequences were determined with Carboxypeptidase A, B, C.

1-1.5 mg enzyme was suspended in 100 ul of 0.2 M $NaHCO_3$ after washing it twice with distilled water. Carboxypeptidase C was used without washing. 100-150 ul of 0.1 M NaOH were added to dissolve the enzyme and then 100-150 ul of 0.1 M HCl followed by N-ethylmorpholine acetate buffer, pH 8.5 to 1.5 ml. The enzyme solution was kept in ice and used as soon as possible.

Peptides were dried *in vacuo* over NaOH. 20 ul of the enzyme solution were added to each peptide and the tubes were covered with parafilm and incubated at 37° C for various times (1 min. to 3 h). The digestion was terminated by drying *in vacuo* over NaOH. The liberated amino acid was identified by dansylation without hydrolysis.

(iii) Dansyl method (Gray & Hartley, 1963b)

10 ul of 0.1 M $NaHCO_3$ were added to a dried peptide (1-10 nM) in a Durham tube and dried again *in vacuo* over NaOH. 5 ul of distilled water and 5 ul of dansyl chloride

(5 mg/ml in acetone) were premixed and added to the peptide. The tube was sealed with parafilm and incubated at 45° C for 45 min. The reaction was terminated by drying *in vacuo* over NaOH. 50 ul of 6 M HCl was added and the dansyl peptide was hydrolysed in a partially vaccum sealed tube at 105° C for 5-18 h. The hydrolysate was dried *in vacuo* over NaOH.

(iv) Chromatography of dansyl derivatives

Dansyl amino acids were identified by polyamide sheet chromatography (Woods & Wang, 1967). The sample was dissolved in 10 ul of 50% (v/v) pyridine and spotted on both sides of the sheet in a 4:1 ratio and dried under a hot air draught. 1 ul of a standard marker solution (see materials and methods, (4)e(i)) was spotted with the 20% of the sample. When dry, the chromatograms were developed in two dimensions using the solvent system described by Ramshaw *et al* (1970). The chromatograms were developed in solvent A (see materials and methods) for 45 min. After drying in a hot air draught, they were run in solvent B for 45 min. The results of the two dimensional separation of the dansyl amino acids were recorded by illuminating the sheets under a UV lamp (350 nm), (Fig. 2). The sheets were then run in solvent C in the second direction for 45 min. After drying, the separation was recorded (Fig. 3).

Polyamide sheets were reused again after they had been washed in an acetone - 1M ammonia solution (1:1 by vol.) for 0.5-1 h.

FIGURE 2.

Chromatography of dansyl-amino acids on polyamide thin layers.

Development was by Solvent A in the first dimension and Solvent B in the second dimension.

FIG. 2

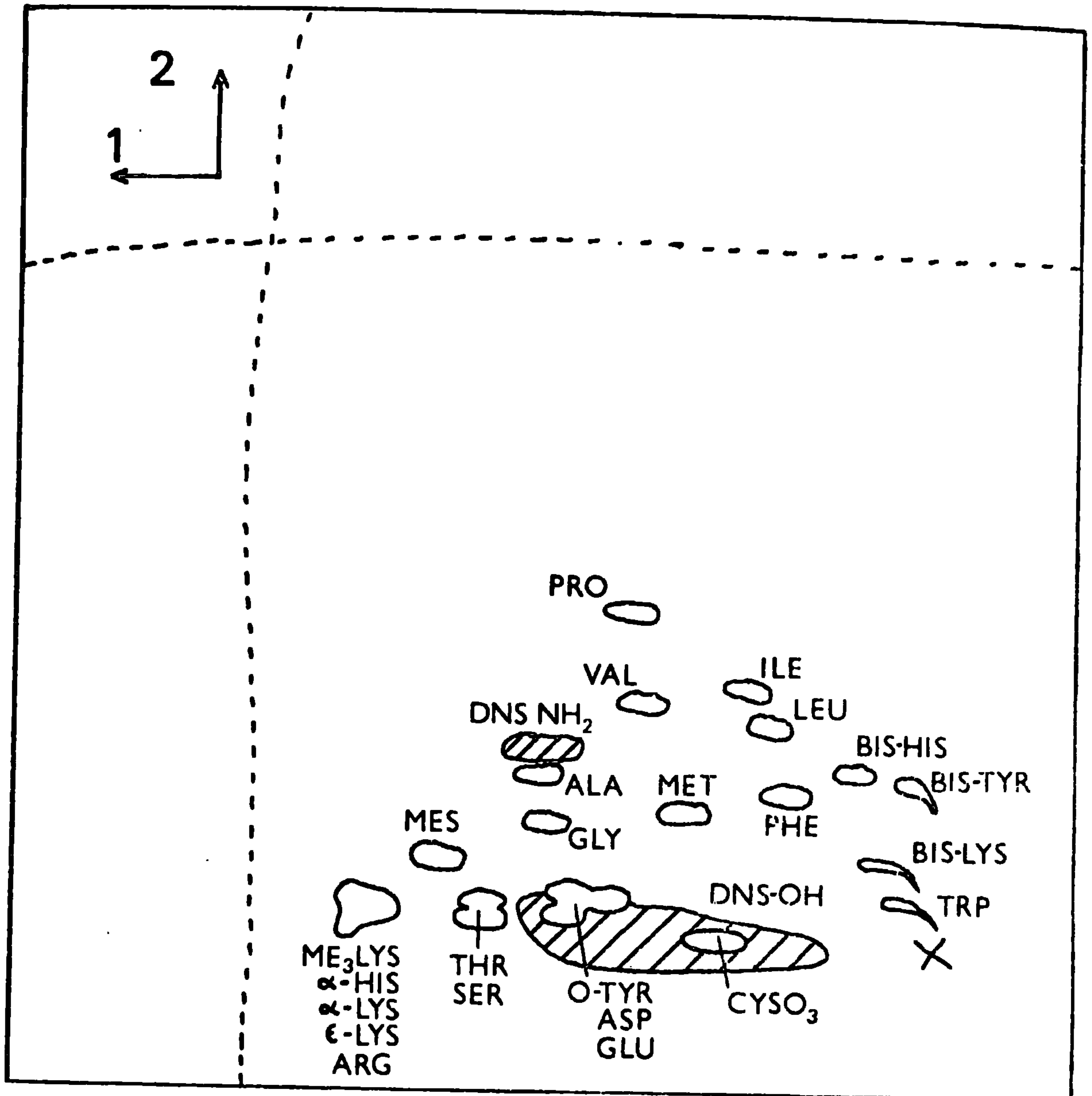
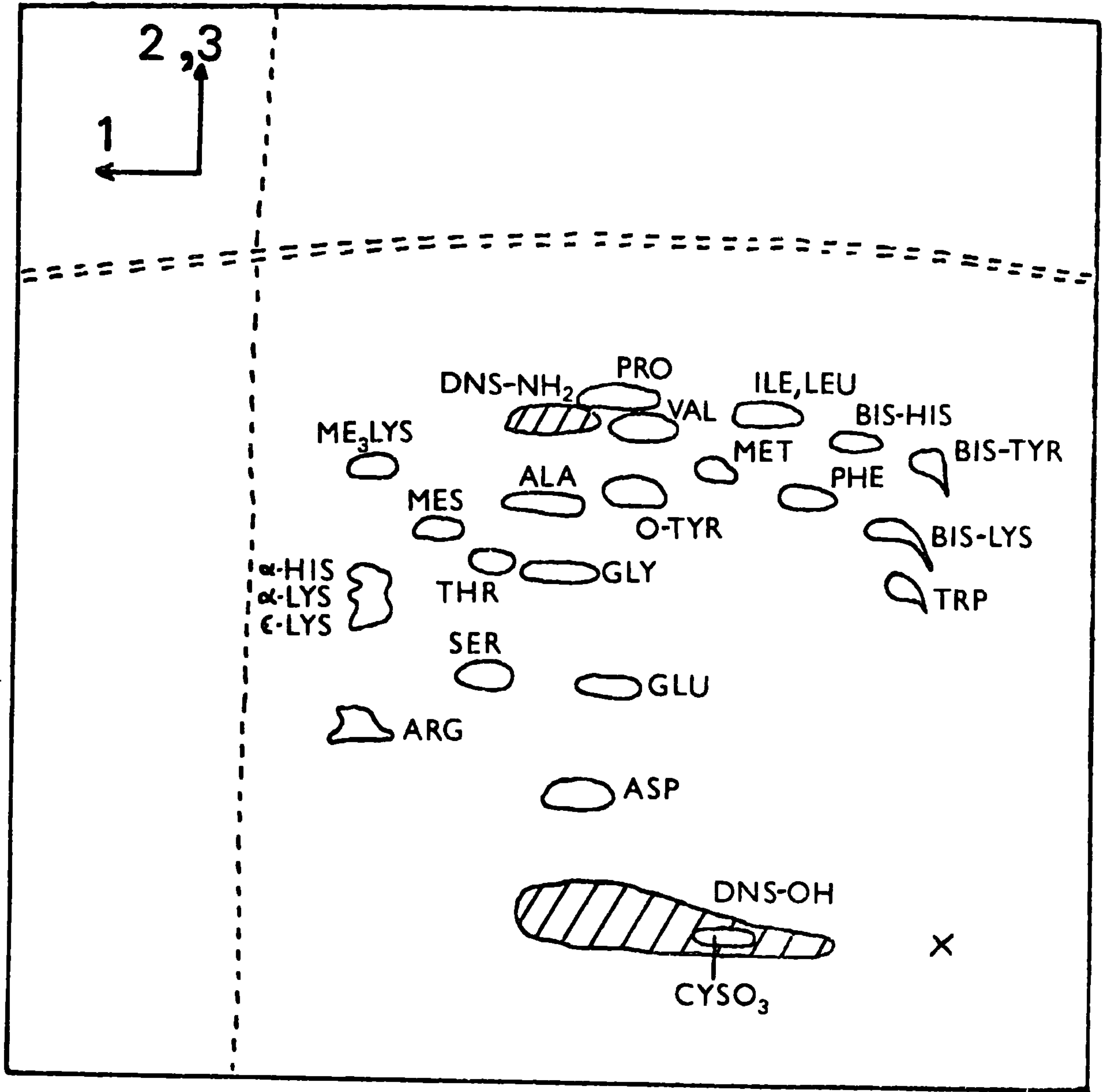


FIGURE 3.

Chromatography of dansyl-amino acids on polyamide thin layers.

Development was by Solvent A in the first dimension and Solvent B followed by Solvent C in the second dimension.

FIG. 3



(v) Edman degradation method. Edman, (1956) and Blomback *et al* (1966)

150 ul of 5% (v/v) redistilled Phenylisothiocyanate (PITC) in pyridine were added to peptide dissolved in 100-200 ul of 20% (v/v) pyridine. The sample was immediately flushed with nitrogen for 15 sec. then mixed briefly(Ilse & Edman, 1963), tubes capped and incubated at 45° C for 1 h. The excess reagents were removed by drying *in vacuo* over NaOH and P_2O_5 at 60° C. When dry, 200 ul of anhydrous trifluroacetic acid (TFA) was added (Elmore & Toseland, 1956) and the tube was covered with parafilm and incubated at 45° C for 30 min. Excess TFA was removed by drying *in vacuo* over NaOH at 60° C. The peptide was dissolved in 200 ul of distilled water and extracted twice with 2 ml of butyl acetate, Gray, (1967) then dried *in vacuo* over NaOH and 4 ml concentrated H_2SO_4.

(vi) Determination of amide residues

Amide residues were determined from the peptide mobilities after electrophoresis at pH 6.5 Offord, (1966) (Fig. 3, 4).

b. Automatic sequencing method

Automated sequence analysis of reduced and carboxymethylated protein was carried out on a Beckman 890 C Protein Sequencer by the method recommended in the Beckman (1972) Operation Manual. The phenylthioazolinone derivatives of amino acids were collected in a fraction

FIGURE 4.

The mobility of peptides on pH 6.5 electrophoresis.

The electrophoretic mobility of peptides relative to dansyl-arginylarginine at pH 6.5 is plotted against their molecular weight for charges (E) of $\pm$ 1 to $\pm$ 3.

Peptides containing histidine or cysteic acid do not conform directly to this diagram.

FIG. 4

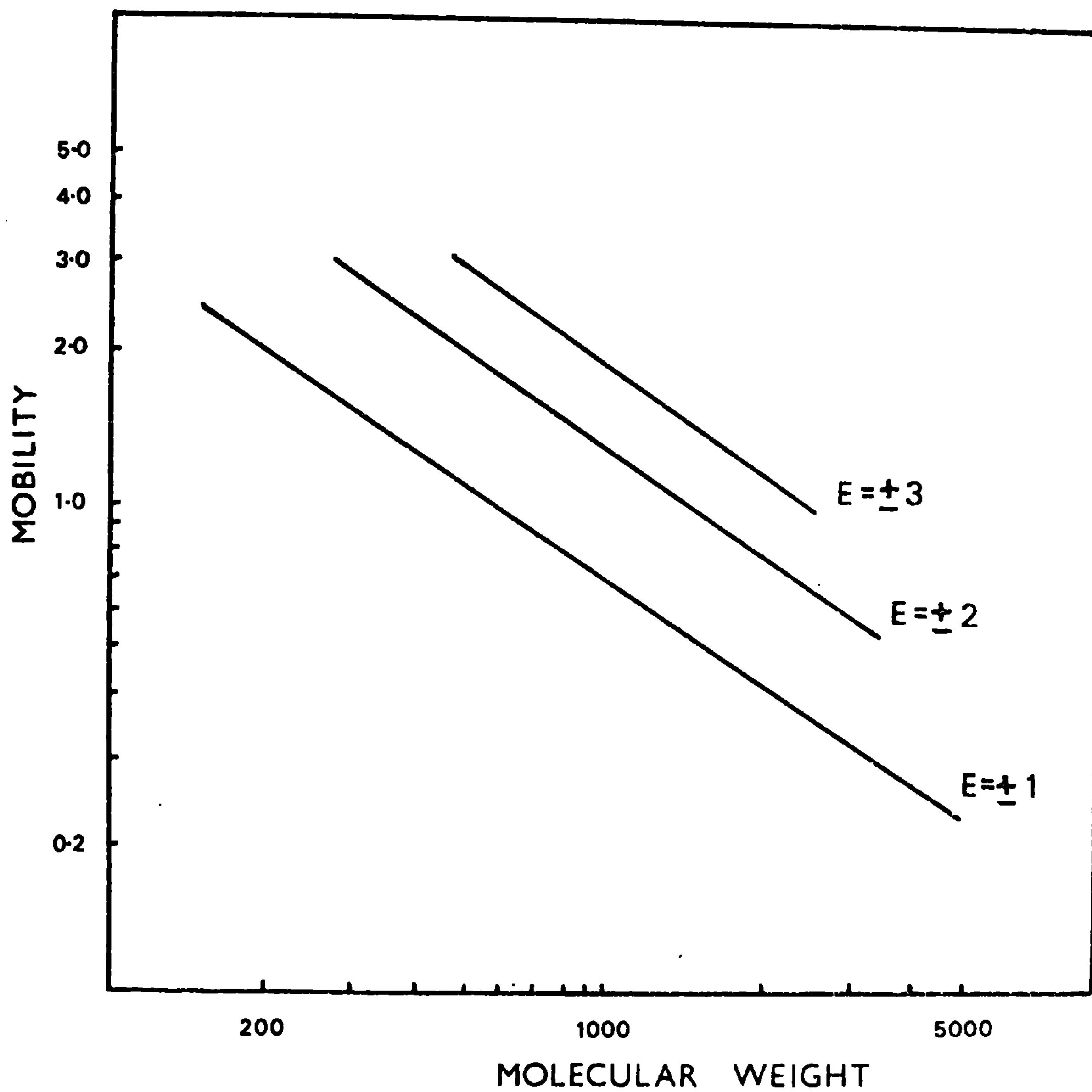
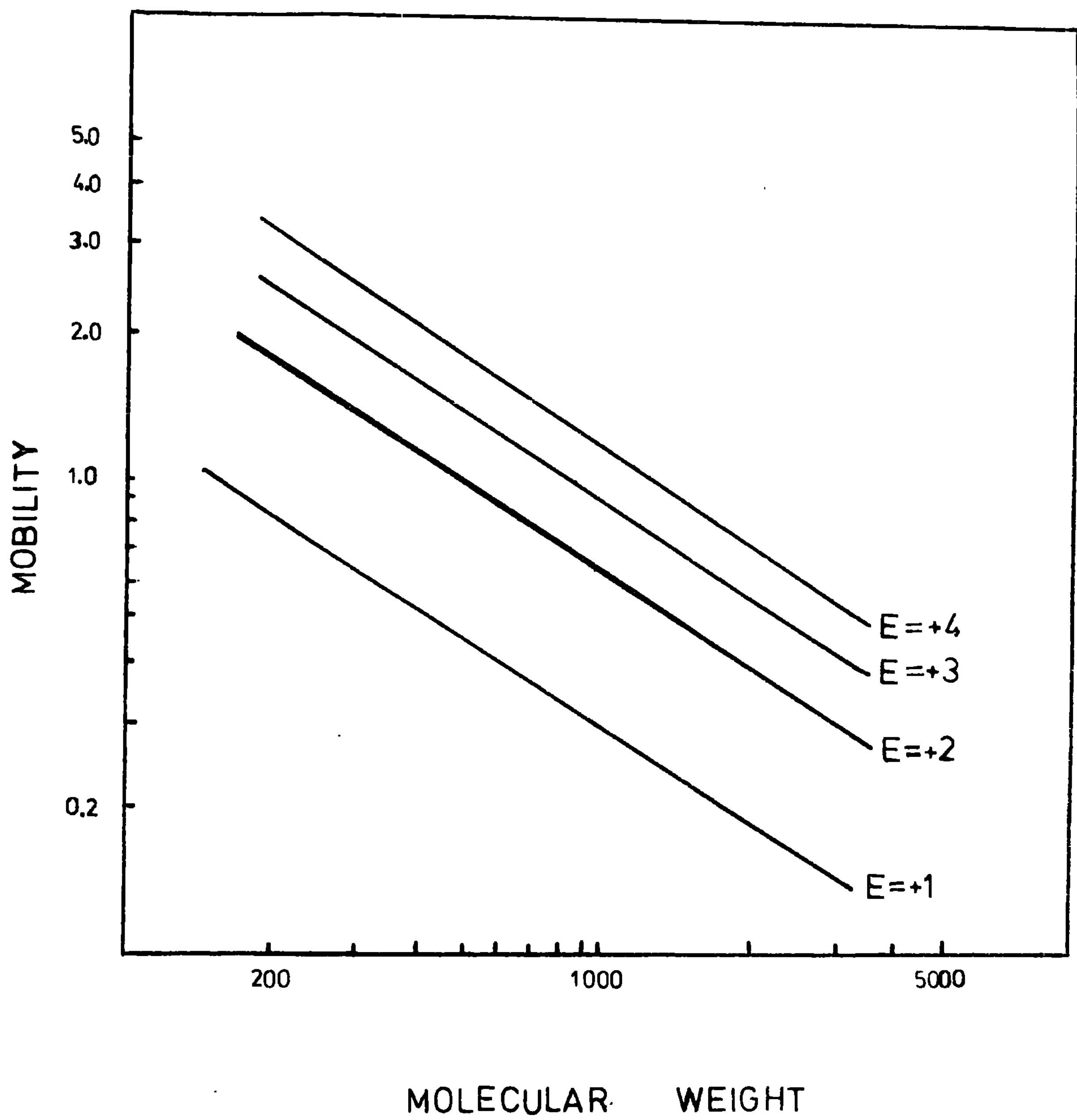


FIGURE 5.

The mobility of peptides on pH 1.9 electrophoresis.

The electrophoretic mobility of peptides at pH 1.9 is plotted against their molecular weight for charges (E) of +1 to +4. Peptides containing histidine or cysteic acid do not conform to this diagram.


5.0
4.0
3.0
2.0
1.0
0.2
MOBILITY
E=+4
E=+3
E=+2
E=+1
200
1000
5000
MOLECULAR WEIGHT

collector and flushed under a stream of nitrogen until dryness. 0.2 ml of 1M HCl was added to each fraction to obtain a more stable phenylthiohydantoin (PTH) derivatives and kept for 10 min. at 80° C under nitrogen. The PTH derivatives were extracted twice with 0.7 ml of ethyl acetate. Both phases, the organic phase and the aqueous phase, were dried down and retained. The organic phase contains most of the PTH amino acids, but the aqueous phase contains the PTH derivatives of arginine, histidine and cysteic acid.

PTH derivatives were identified as described by Haslett & Boulter, (1976) by gas chromatography, thin layer chromatography on silica plates and the hydriodic acid procedure of Inglis et al (1971).

(8) Nomenclature used to describe sequence analysis data

CNBr fragments and peptides derived from digestion of the whole protein are numbered on the basis of their order within the complete sequence starting at the N-terminus of the protein. Peptides derived from digestion of a larger peptide are numbered on the basis of their order within the parent fragment. Peptides resulting from partial clevage within a major peptide are given a letter subscript to the major peptide.

IV CALCULATIONS BASED ON AMINO ACID SEQUENCE DATA

1) Construction of the amino acid difference matrix

Sequence alignments of ferredoxins were made relative to their invariant cysteine residues to which the iron and the labile sulphur are attached. Differences or gaps due to deletions were considered as single changes.

2) Phylogenetic tree construction using an ancestral sequence method (J.T. Gleaves, unpublished)

This method contains two basic parts:

(a) A method of counting the minimum number of substitutions required to account for a given tree. This is basically the preliminary phase of Fitch's method (Fitch, 1971).

(b) A tree building procedure similar to that of Dayhoff (1972) in which species are added to an existing tree, starting with a tree of three species, in each possible position. In Dayhoff's method, the first of the most parsimonious alternatives is kept for the addition of subsequent species, whereas in this method all of the most parsimonious alternatives are kept.

RESULTS

1) Isolation and purification of ferredoxins

A large scale preparation of ferredoxins using the method of Petering and Palmer (1970) was chosen for the purification of ferredoxins from *Triticum aestivum*, *Brassica napus*, *Sambucus nigra* and *Hordeum vulgare*.

The fresh leaves were homogenised with 10 mM Tris-HCl buffer, pH 7.2, filtered through muslin and centrifuged to eliminate unwanted precipitated materials. The protein was adsorbed by adding DEAE-23 cellulose to the dark brown supernatant and stirring gently for about 1 hr. The dark brown colored resin was collected by centrifugation and washed repeatedly with 10 mM Tris-HCl buffer, pH 7.2 until the washing buffer became clear. The ferredoxin was eluted as a brown liquid with 150 mM Tris-HCl buffer at pH 7.2 containing 1 M NaCl. Ammonium sulphate fractionation was carried out at pH 7.2 and 4° C. No precipitation of ferredoxin was detected at saturations upto 85% at which point the fractionation was terminated. Large quantities of brown precipitate were discarded following centrifugation. The 85% ammonium sulphate supernatant (ferredoxin solution) was exhaustively dialysed against several changes of 10 mM Tris-HCl buffer, pH 7.2, before concentration on DEAE-23. The dialysate was diluted with 1:1 distilled water, added to the column and was then eluted with 150 mM Tris-HCl containing 160 mM NaCl, concentrated again to a 5 ml volume by using a diaflo apparatus. The eluted ferredoxin solution

was so dark brown that spectrophotometric measurements could not be made. Elder ferredoxin was subjected to six successive DEAE cellulose columns before it was chromatographed on a 3 cm X 90 cm column of Biogel P-10. Wheat and barley ferredoxins were chromatographed on a 3 cm X 90 cm column of sephadex G-75 while rape ferredoxin was chromatographed on a 3 cm X 90 cm column of Biogel P-10, the eluted ferredoxins had less than 0.35 ratio values (R= A(420)/A(278)). The concentrated fractions of ferredoxins were chromatographed on a 1.5 cm X 15 cm of DEAE sephadex A-50. The ferredoxin adsorbed to the top of the column and was eluted using a linear ionic gradient. The eluted ferredoxins exhibited the typical purified absorbance ratios (see table 1 and Fig. 6, 7 and 8) of pure plant ferredoxins.

None of the purified ferredoxins from these different plants differed significantly from the absorbance spectrum given by Spinach ferredoxin (see Hall et al., 1973). The purified ferredoxins showed absorption maxima in the U.V region at 278 and 330 nm and in the visible region at 420 and 465 nm that is typical of purified ferredoxins (Buchanan & Arnon, 1971). The absorbance at 420 nm decreases by about 50% when the protein is reduced with sodium dithionite (see Fig. 6).

The homogeneity of the ferredoxins was demonstrated by isoelectric focusing studies. Ferredoxins migrated as single bands corresponding to their pIs (see table 1 and Fig. 9) except carboxymethylated wheat ferredoxin which gave two closely migrating bands.

TABLE 1.

Yield of ferredoxin, purity ratio (R) values and isoelectric focusing points for the different plants.

Source	Yield mg/Kg (F.W.)	R A(420)/A(278)	pI
*Porphyra umbilicalis**	12.5	0.45	3.8
Triticum aestivum	25	0.48	3.9
Hordeum vulgare	20	0.48	3.75
Sambucus nigra	3.5	0.42	3.8
Brassica napus	10	0.79	3.8

* Andrew *et al.*, 1976

FIGURE 6 .

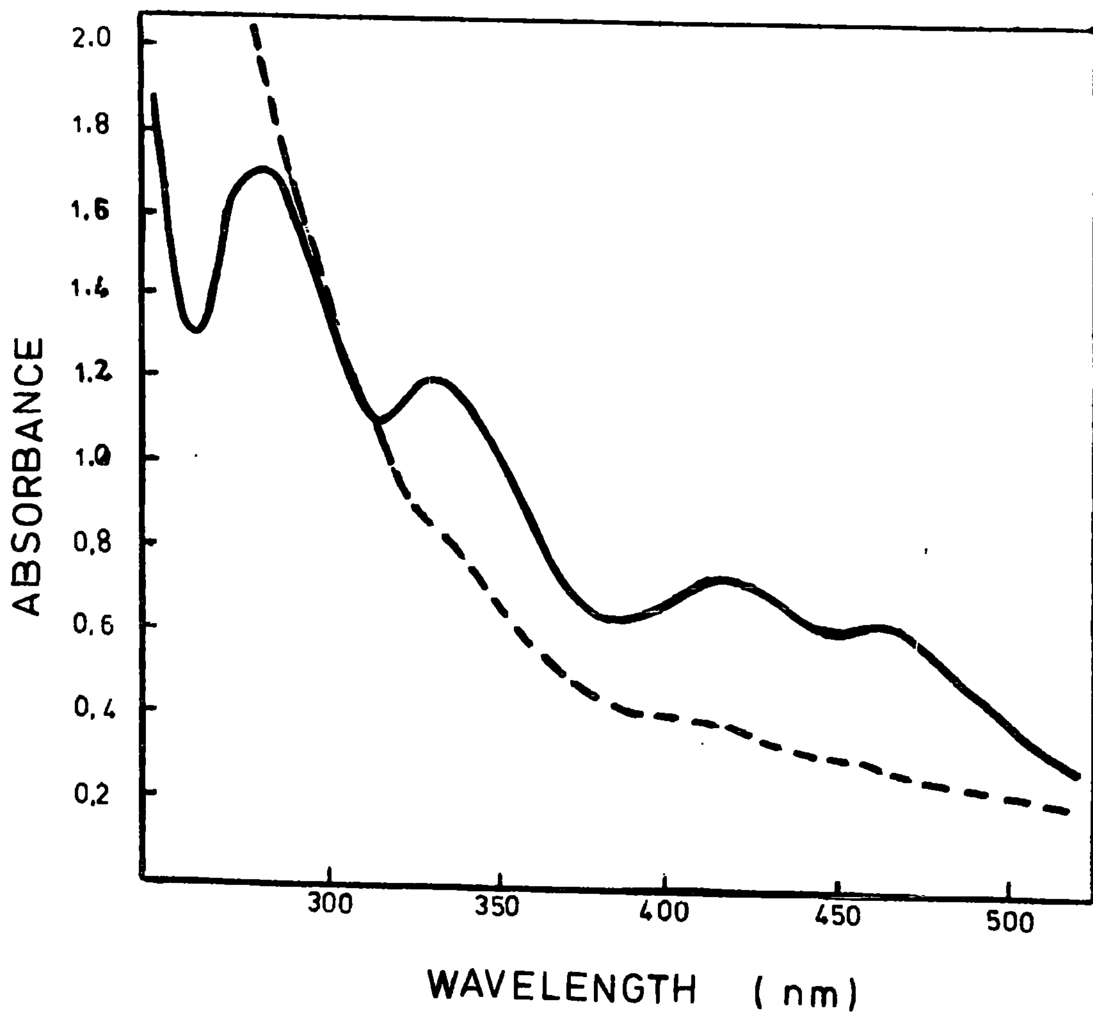
Absorption spectra of the oxidised and reduced Wheat ferredoxin. The protein was dissolved in 50 mM Tris-HCl buffer (pH 7.2) containing 50 mM KCl.

——— Oxidised spectra

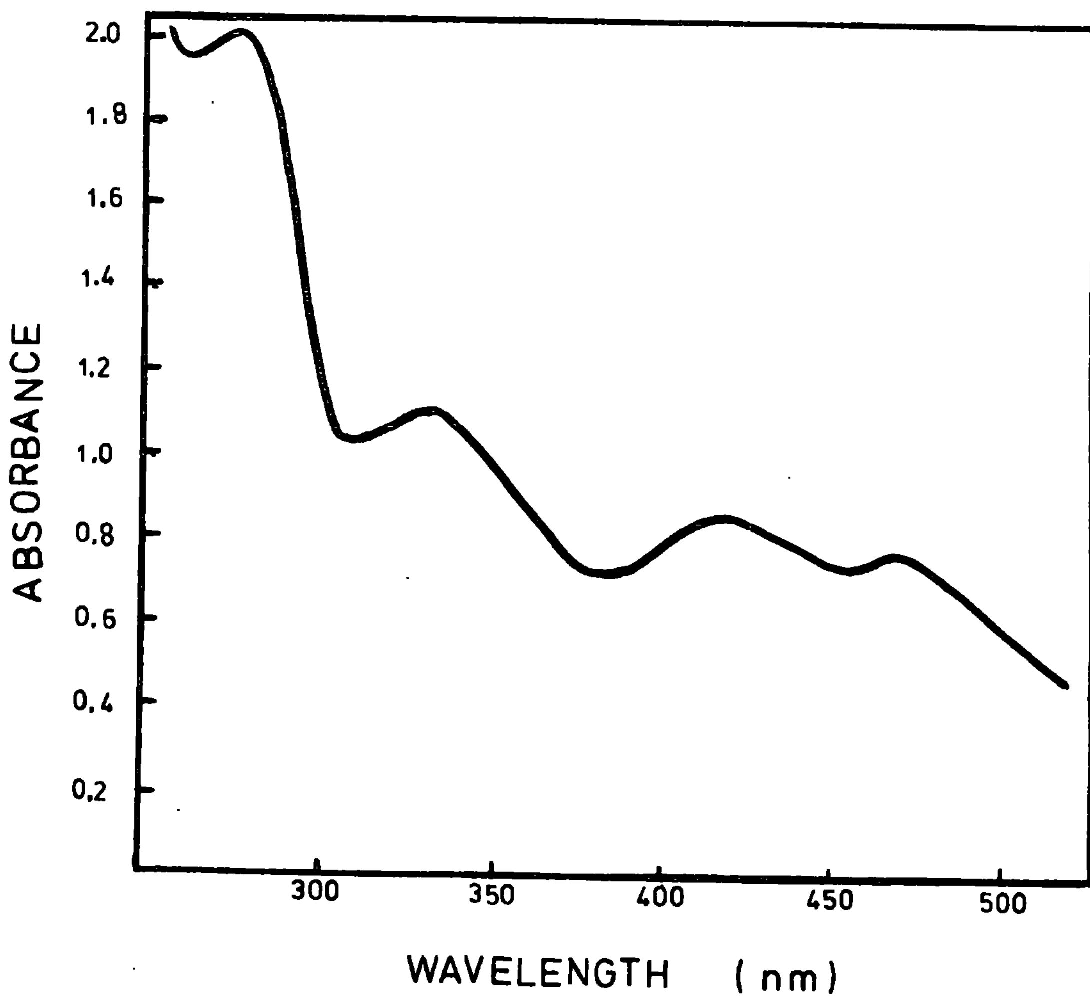
- - - - Spectra after reduction by Sodium dithionite

FIGURE 7 .

Absorption spectra of Elder ferredoxin. The protein was dissolved in 50 mM Tris-HCl buffer (pH 7.2) containing 50 mM KCl.



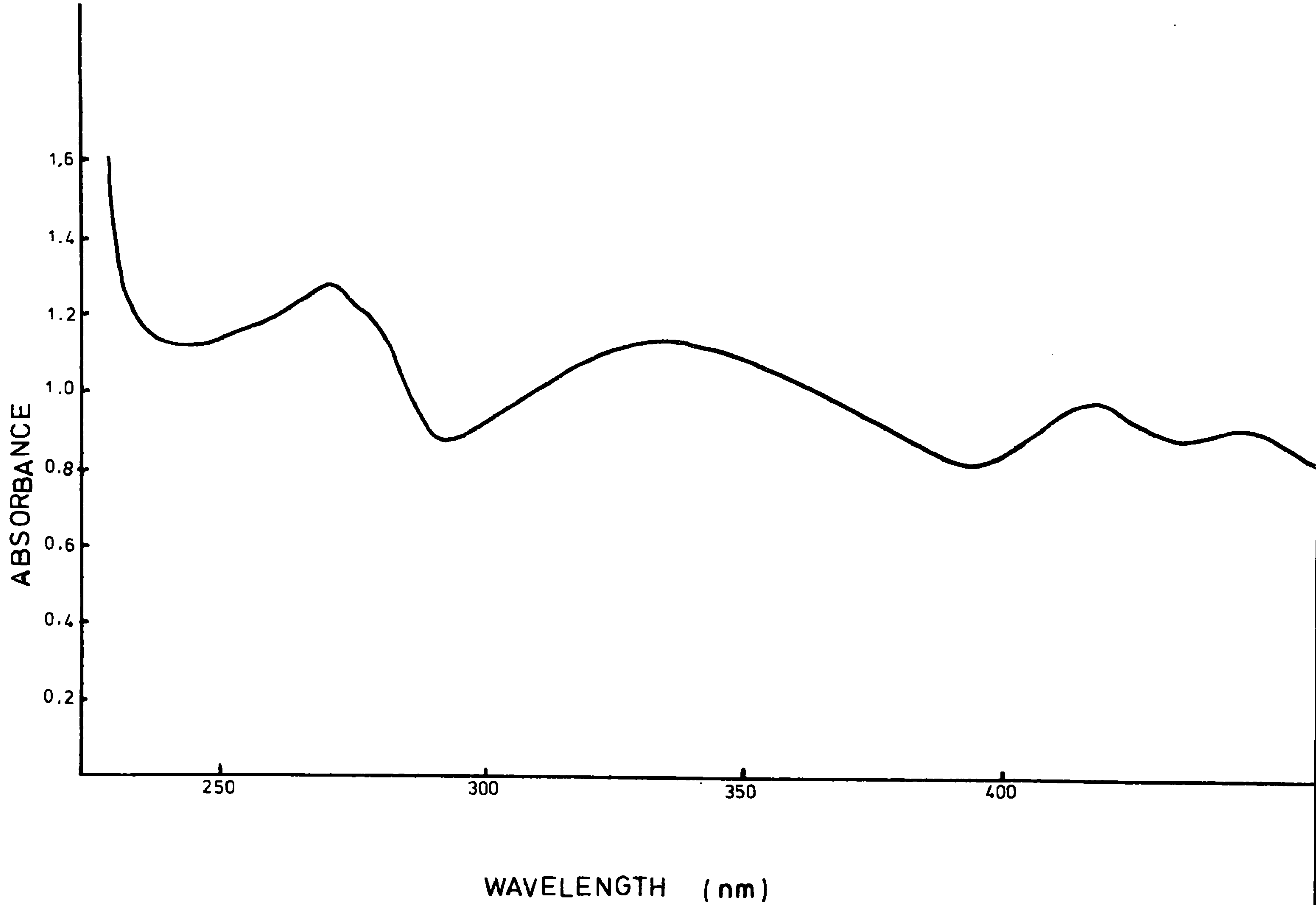
ABSORBANCE
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
300
350
400
450
500
WAVELENGTH (nm)



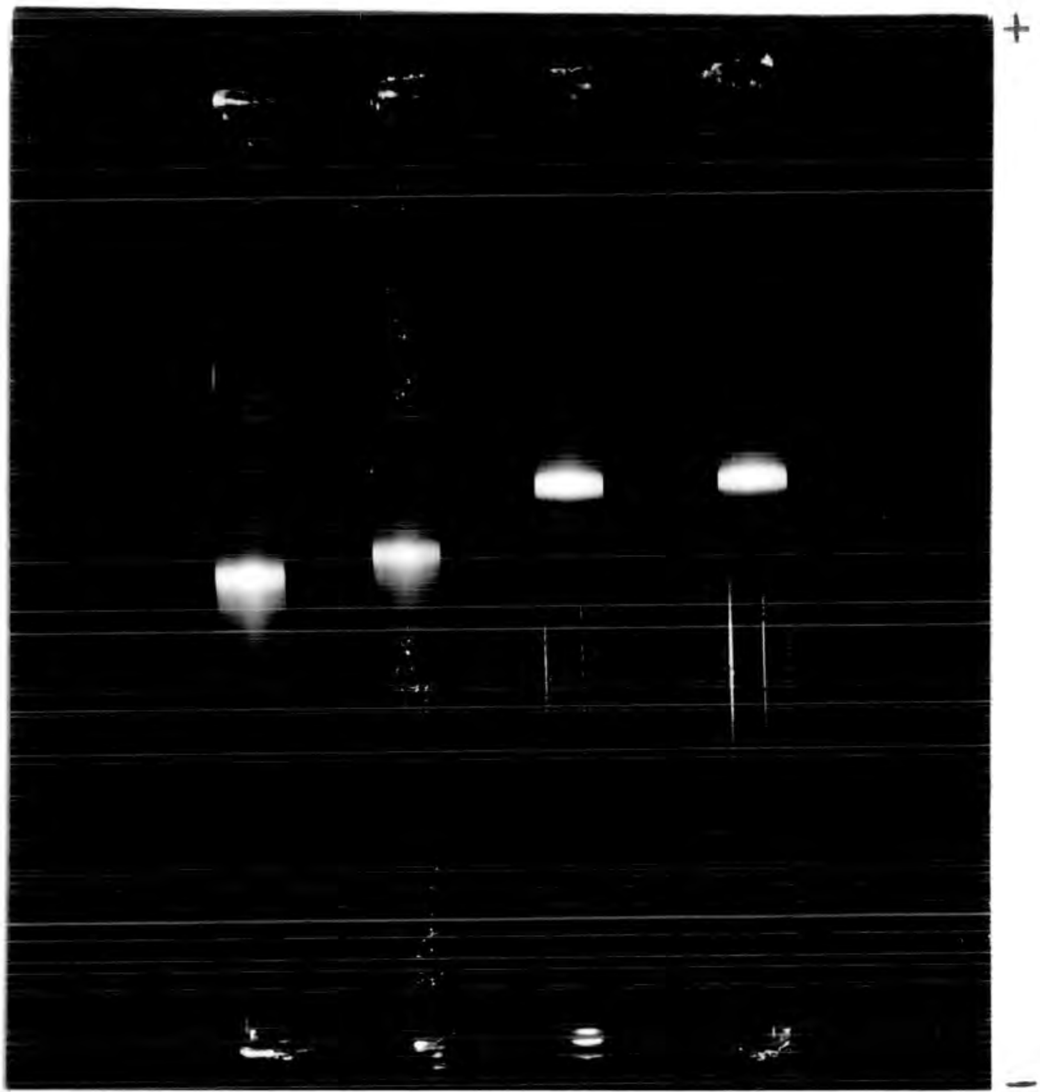
ABSORBANCE
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
300
350
400
450
500
WAVELENGTH (nm)

FIGURE 8.

Absorption spectra of *Rape* ferredoxin. The protein was dissolved in 50 mM Tris-HCl buffer (pH 7.2) containing 50 mM KCl.



ABSORBANCE
1,6
1.4
1.2
1.0
0.8
0.6
0.4
0,2
250
300
350
400
WAVELENGTH (nm)



B A

Isoelectric Focusing of Purified Ferredoxin

From A) Wheat B) Barley

Focusing was done in pH 3−5. + and −

indicate anode and cathode respectively.

FIGURE 9

II) The amino acid sequence of *Porphyra* ferredoxin

The amino acid sequence of *Porphyra* ferredoxin was determined from the evidence of peptides obtained by tryptic, chymotryptic and thermolytic digestion of the protein or its CNBr fragments. The sequence is given in Fig.10 showing the points of enzyme and CNBr cleavage together together with the overlapping peptides from which it was deduced. The amino acid composition of *Porphyra* ferredoxin together with that of the CNBr fragments is shown in Table 1.

Cyanogen Bromide cleavage

15 mg of carboxymethylated protein was used for clevage. The CNBr fragments were chromatographed on a 1.5 cm X 130 cm Sephadex G-50 superfine column. The flow rate was 9 ml/h and fractions of volume 1 ml were collected. 50 ul aliquots were taken for N-terminal amino acid analysis. Two fragments (X1 & X2) were identified (Fig.11)one with N-terminus alanine and the other with N-terminus leucine. The first peak eluted from the column (see Fig.11) was the uncleavaged protein.

Digestion with carboxypeptidase A showed that the C-terminus of X1 fragment was -Glu-Gln-Hse., and the C-terminus of X2 fragment was -Leu-Tyr. The first fragment (X1) was digested with 2% (w/w) trypsin. The second fragment (X2), 50% was subjected directly to dansyl-Edman analysis (Table 2) and 50% was digested with 2% (w/w) thermolysin.

For chymotryptic digest, 12 mg of carboxymethylated protein was digested with 2% (w/w) chymotrypsin.

Separation of peptides

Peptides resulting from trypsin and chymotrypsin digestion were subjected to gel chromatography on a column (1.5 cm X 130 cm) of Sephadex G-50 superfine equilibrated with 85% (v/v) formic acid. Elution profiles were followed by measurement of A_{280} and A_{206} with LKB Uvicord instrument and by N-terminal analysis by the dansyl technique. Peptides which remained impure after this purification step and peptides which were produced by thermolysin digestion were subjected to high voltage paper electrophoresis at pH 6.5 and pH 1.9.

A list of chymotryptic peptides together with mobility and sequence data is given in Table 4 and a similar list of tryptic and thermolytic peptides deduced from CNBr fragments are given in Table 5 and Table 6. The amino acid composition of peptides is given in Table 3.

Peptide X1T1 (1-4) (Ala-Asx-Tyr-Lys)

The mobility of this peptide at pH 6.5 indicated that residue 2 was aspartic acid. After three steps of Edman degradation, lysine was identified by dansylation without hydrolysis as bis-dansyl-lysine.

Peptide X1T2 (5-10) (Ile-His-Leu-Val-Ser-Lys)

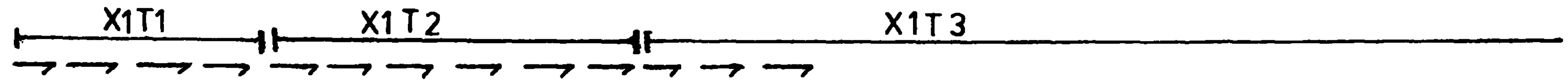
The histidine was identified as the bis-dansyl-histidine derivative. Dansylation without hydrolysis after five steps of Edman degradation yielded bis-dansyl-lysine.

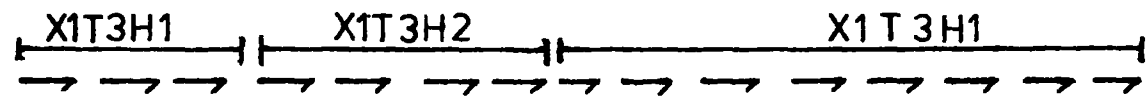
FIGURE 10.

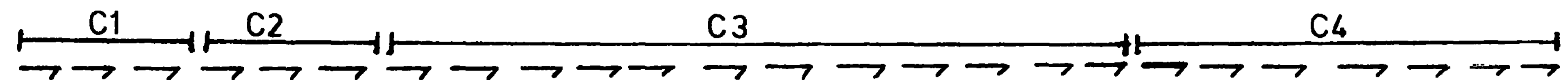
The amino acid sequence of Porphyra ferredoxin. Residues which were identified by automatic sequencer, dansyl-Edman analysis and carboxypeptidase digestion are indicated by ⟶ , ⇀ , ↽ respectively. The arrows --⇀ indicate that residues was ambiguously identified. The arrow ↓ indicates CNBr cleavage and * indicates homoserine was identified as the derivative from CNBr cleavage. For peptides X1T5H2 and X2H5 the mobilities at pH 6.5 after some degradation cycles are indicated. T, C and H represent peptides obtained by tryptic, chymotryptic and thermolysin digestions respectively.

1 5 10 15 20 25

Ala-Asp-Tyr-Lys-Ile-His-Leu-Val-Ser-Lys-Glu-Glu-Gly-Ile-Asp-Val-Thr-Phe-Asp-Cys-Ser-Glu-Asp-Thr-Tyr

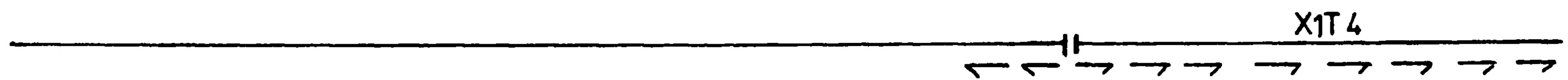


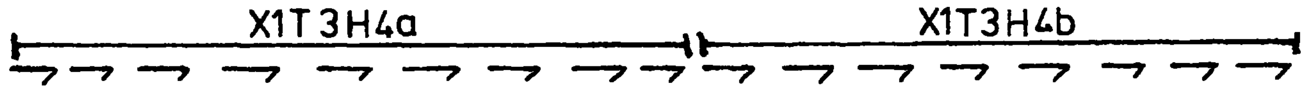


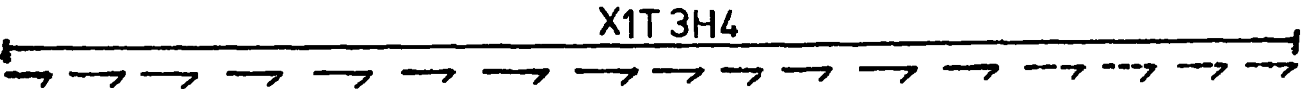


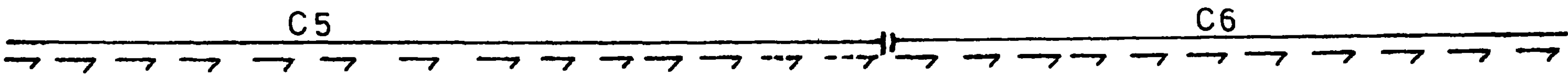
26 30 35 40 45 50

Ile-Leu-Asp-Ala-Ala-Glu-Glu-Glu-Gly-Ile-Glu-Leu-Pro-Tyr-Ser-Cys-Arg-Ala-Gly-Ala-Cys-Ser-Thr-Cys-Ala

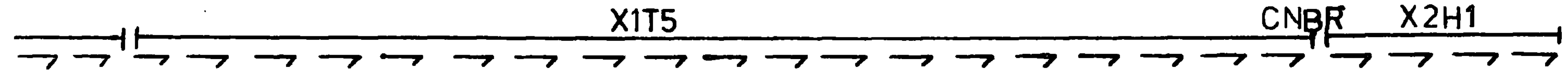


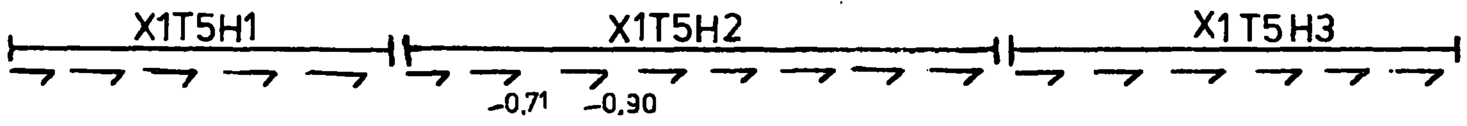


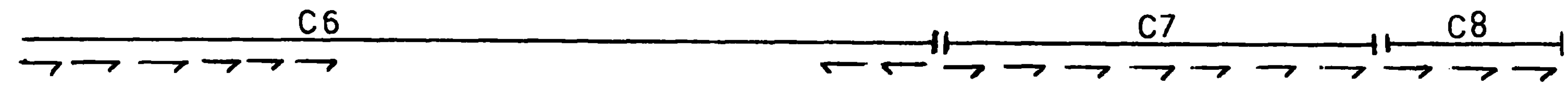




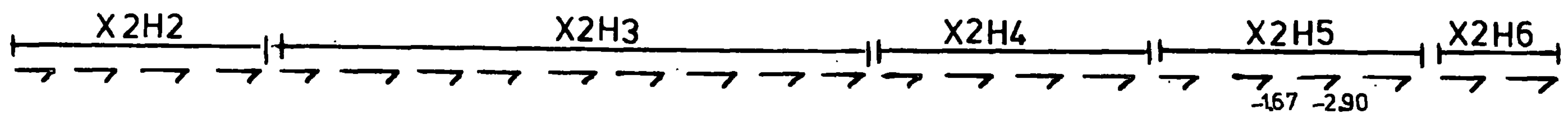
51 55 60 65 70 75
Gly-Lys-Val-Thr-Glu-Gly-Thr-Val-Asp-Gln-Ser-Asp-Gln-Ser-Phe-Leu-Asp-Asp-Glu-Gln-Met-Leu-Lys-Gly-Tyr







76 80 85 90 95
Val-Leu-Thr-Cys-Ile-Ala-Tyr-Pro-Glu-Ser-Asp-Cys-Thr-Ile-Leu-Thr-His-Val-Glu-Gln-Glu-Leu-Tyr



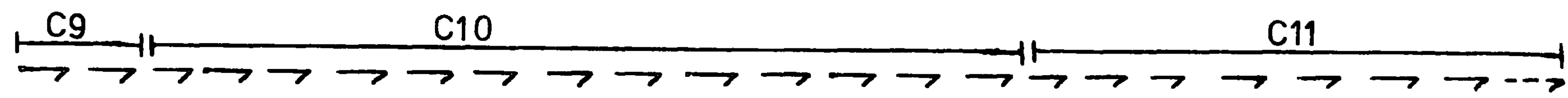


Figure 11.

Chromatography of X1 and X2 fragments from *Porphyra* ferredoxin on Sephadex G-50 superfine.

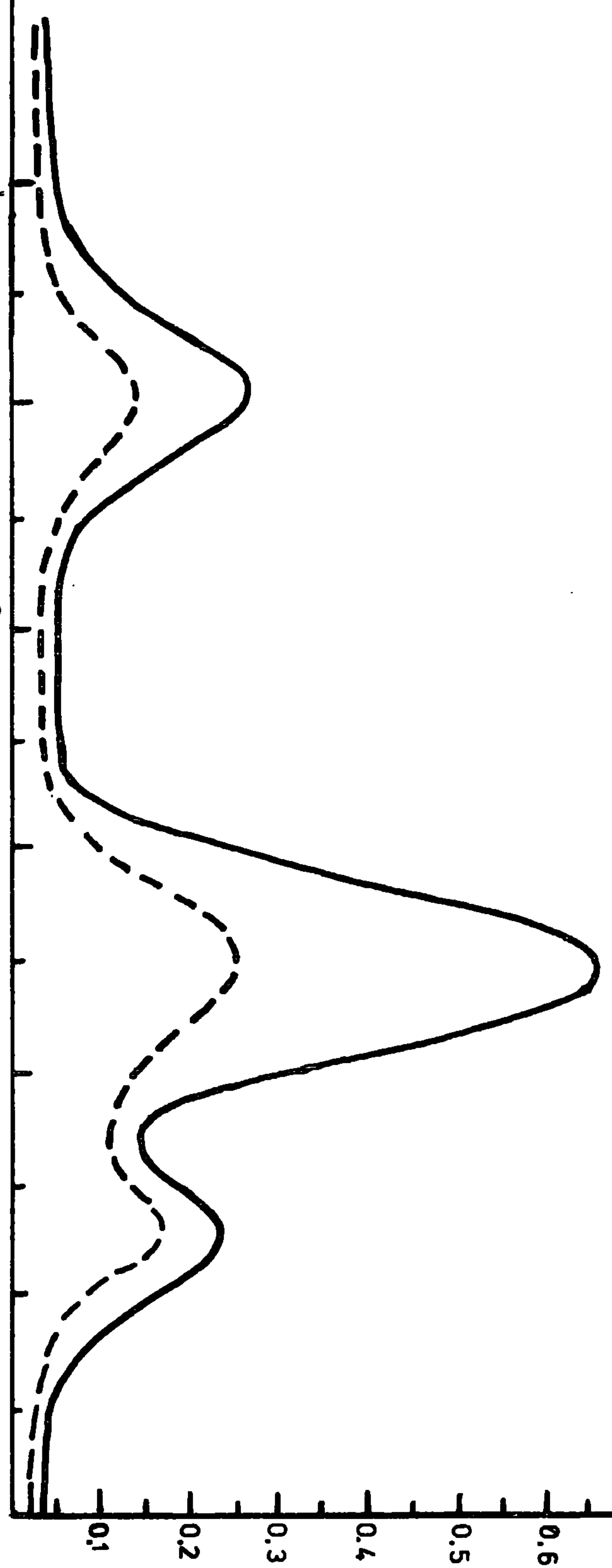
A 1.5 cm X 130 cm column of resin was used; it was equilibrated and eluted by 85% formic acid.

The following notations has been used:-

——— absorbance at 280 nm

- - - - absorbance at 206 nm

The absorbance at 206 nm was very low due to the lack of sensitivity of the Uvicord instrument.



ELUTION VOLUME (ml)
20
40
60
80
100
120
ABSORBANCE
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8

TABLE 2.

Cyanogen Bromide fragments of *Porphyra* ferredoxin

Fragment	Dansyl-Edman and carboxypeptidase A results
X1 (1-71)	Ala-(digested with trypsin)-Glu-Gln-Hse
X2 (72-98)	Leu-Lys-Gly-Tyr-Val-Leu-Thr-Cys-Ile-Ala-Tyr-Pro-Glu-Ser-Asp-Cys-Thr-(- - - - -)-Leu-Tyr

X1 was digested with trypsin.

X2 was digested with thermolysin.

TABLE 3

The Amino Acid Composition of *Porphyra* ferredoxin and its CNBr fragments

The results are expressed as residues/molecule and cysteine was measured as cysteic acid. Values for serine and threonine are extrapolated to zero time where possible, assuming first-order rate of destruction (Moore and Stein 1963). For valine and isoleucine maximal values (72h hydrolysis) were taken. n.d. = not determined; Hse = homoserine; Seq. = values from sequence determination.

	Total protein					X1 fragment				X2 fragment			
	24h	48h	72h	Average	Seq.	24h	72h	Average	Seq.	24h	72h	Average	Seq.
Asp	9.71	10.20	10.63	10.18	10	9.31	9.46	9.39	9	1.38	0.70	1.05	1
Thr	8.38	8.32	8.31	8.41	8	4.64	5.13	4.89	5	3.22	2.77	3.47	3
Ser	6.32	6.87	6.90	6.70	7	5.70	5.59	5.76	6	1.24	1.14	1.29	1
Hse	0	0	0	0	0	0.52	0.7	0.61	1	0	0	0	0
Glu	16.51	16.43	15.38	16.11	16	11.95	12.00	11.98	12	4.08	3.89	3.99	4
Pro	1.67	1.94	1.70	1.77	2	0.56	0.92	0.74	1	0.84	1.08	0.96	1
Gly	6.32	6.19	6.24	6.25	6	5.38	5.42	5.40	5	1.38	1.26	1.32	1
Ala	6.93	6.85	6.81	6.86	7	5.64	6.34	5.99	6	1.21	1.27	1.24	1
Cys	6.11	n.d.	n.d.	6.11	6	n.d.	n.d.	n.d.	(4)	n.d.	n.d.	n.d.	(2)

Val	5.27	5.57	5.72	5.72	6	3.74	3.96	3.96	4	1.99	2.00	2.00	2
Met	0.64	0.51	0.63	0.59	1	0	0	0	0	0	0	0	0
Ile	5.65	5.68	5.72	5.72	6	3.51	4.14	4.14	4	1.90	1.91	1.91	2
Leu	7.67	7.82	7.81	7.77	8	4.22	4.05	4.14	4	3.85	3.90	3.88	4
Tyr	5.56	5.65	5.72	5.64	6	3.39	2.44	2.92	3	2.54	2.63	2.59	3
Phe	1.85	2.03	1.95	1.94	2	1.91	1.77	1.84	2	0	0	0	0
His	1.64	1.81	1.80	1.75	2	1.38	0.78	1.08	1	0.82	0.89	0.86	1
Lys	3.61	3.84	3.76	3.74	4	3.38	2.84	3.12	3	0.95	1.17	1.06	1
Arg	1.02	1.02	1.14	1.06	1	1.37	1.33	1.35	1	0	0	0	0
Total residues					98				71				27

TABLE 4.

Chymotryptic peptides from *Porphyra* ferredoxin

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
C1 (1-3)	-1.15	0.45	Ala-Asx-Tyr
C2 (4-6)	2.60	—	Lys-Ile-His
C3 (7-18)	-1.01	0.45	Leu-Val-Ser-Lys-Glx-Glx-Gly-Ile-Asx-Val-Thr-Phe
C4 (19-25)	-2.90	—	Asx-Cys-Ser-Glx-Asx-Thr-Tyr
C5 (26-39)	-3.37	—	Ile-Leu-Asx-Ala-Ala-Glx-Glx-Glx-Ile-Glx-Leu-Pro-Tyr
C6 (40-65)	—	—	Ser-Cys-Arg-Ala-Gly-Ala-Cys-Ser-Thr-Cys-Ala-Gly-Lys-Val-Thr-Glx-Gly-Thr-Val-Asx-Glx-Ser-Asp-Glx-Ser-Phe
C7 (66-72)	-2.25	0.34	Leu-Asx-Asx-Glx-Glx-Met-Leu
C8 (73-75)	1.42	—	Lys-Gly-Tyr
C9 (76-77)	0	0.68	Val-Leu
C10(78-90)	-1.97	—	Thr-Cys-Ile-Ala-Tyr-Pro-Glx-Ser-Asx-Cys-Thr-Ile-Leu
C11(91-98)	-0.77	0.59	Thr-His-Val-Glx-Glx-Glx-Leu-Tyr

TABLE 5.

Tryptic and thermolytic peptides of X1 fragments from *Porphyra* ferredoxin

Peptide/ position	Mobility (pH 6.5)	Mobility (pH 1.9)	Dansyl-Edman results
X1T1 (1-4)	0	0.9	Ala-Asx-Tyr-Lys
X1T2 (5-10)	1.67	—	Ile-His-Leu-Val-Ser-Lys
X1T3 (11-43)	-1.73	—	Glx-Glx-Gly-(- - -)-Cys-Arg This peptide was digested with thermolysin
X1T3 H1(11-13)	—	0.5	Glx-Glx-Gly
X1T3 H2(14-17)	—	0.39	Ile-Asx-Val-Thr
X1T3 H3(18-25)	—	0.25	Phe-Asx-Cys-Ser-Glx-Asx-Thr-Tyr
X1T3 H4(26-42)	—	0.29	Ile-Leu-Asx-Ala-Ala-Glx-Glx-Glx-Gly-Ile-Glx-Leu-Pro-Tyr-Ser-Cys-Arg
X1T3 H4a(26-34)	—	0.34	Ile-Leu-Asx-Ala-Ala-Glx-Glx-Glx-Gly
X1T3 H4b(35-42)	—	0.5	Ile-Glx-Leu-Pro-Tyr-Ser-Cys-Arg
X1T4 (43-52)	-1.15	—	Ala-Gly-Ala-Cys-Ser-Thr-Cys-Ala-Gly-Lys
X1T5 (53-71)	-3.07	—	Val-Thr-Glx-Gly-Thr-Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe-Leu-Asx-Asx-Glx-Glx-Hse

(continued....)

TABLE 5.

(continued.....)

X1T5 H1(53-57)	-1.01	0.4	Val-Thr-Glx-Gly-Thr-
X1T5 H2(58-65)	-1.21	—	Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe
X1T5 H3(66-71)	-2.36	0.34	Leu-Asx-Asx-Glx-Glx-Hse

TABLE 6.

Thermolytic peptides of X2 fragment from *Porphyra* ferredoxin

Peptide/ position	Mobility (pH 6.5)	Mobility (pH 1.9)	Dansyl-Edman results
X2H1 (72-75)	1.21	1.00	Leu-Lys-Gly-Tyr
X2H2 (76-80)	-1.34	0.45	Val-Leu-Thr-Cys
X2H3 (81-88)	-2.16	0.35	Ile-Ala-Tyr-Pro-Glx-Ser-Asx-Thr
X2H4 (89-92)	1.20	0.80	Ile-Leu-Thr-His
X2H5 (93-96)	-2.25	0.45	Val-Glx-Glx-Glx
X2H6 (97-98)	0	0.50	Leu-Tyr

<u>Peptide X1T3</u> (11-42) (Glx-Glx-Gly-(was digested with thermolysin)-Cys-Arg)

This large peptide was separated and purified by gel chromatography. The amino acid analysis showed it consisted of 32 residues. Two steps of Edman degradation and dansyl analysis suggested that the N-terminal of this peptide corresponded to that of peptide T3. Digestion with carboxypeptidase C released arginine on short digestion and arginine with trace of cysteine on prolonged (2 h) digestion. Digestion with thermolysin for 1 h produced six peptides which were isolated by paper electrophoresis at pH 1.9. Edman degradation and dansyl analysis of each thermolytic peptide was carried out.

<u>Peptide X1T3H1</u> (11-13) (Glx-Glx-Gly)

PTH-derivatives of residues 11 and 12 indicated they are glutamic acids. This was confirmed from peptide C3.

<u>Peptide X1T3H2</u> (14-17) (Ile-Asx-Val-Thr)

Residue 15 was identified as aspartic acid from its PTH-derivative and from peptide C3.

<u>Peptide X1T3H3</u> (18-25) (Phe-Asx-Cys-Ser-Glx-Asx-Thr-Tyr)

PTH-derivatives of residues 19, 22 and 23 showed they are aspartic acid, glutamic acid and aspartic acid respectively. Dansylation without hydrolysis after seven steps of Edman degradation yielded dansyl-bis-tyrosine.

<u>Peptide X1T3H4</u> (26-42) (Ile-Leu-Asx-Ala-Ala-Glx-Glx-Glx-Gly-Ile-Glx-Leu-Pro-Tyr-Ser-Cys-Arg)

Dansyl Edman analysis become inconclusive after twelve steps of degradation. Residues 39-42 were placed

by examination of the semi quantitative amino acid analysis and from peptide X1T3H4b.

Peptide X1T3H4a (26-34) (Ile-Leu-Asx-Ala-Ala-Glx-Glx-Glx-Gly)

Residues 28, 31, 32 and 33 were identified as acids from their PTH-derivatives. Dansylation without hydrolysis after eight steps of Edman degradation yielded dansyl-glycine.

Peptide X1T3H4b (35-42) (Ile-Glx-Leu-Pro-Tyr-Ser-Cys-Arg)

Arginine was identified as the C-terminal of the peptide by dansylation without hydrolysis after seven steps of Edman degradation. Residue 36 was identified as PTH-glutamic acid.

Peptide X1T4 (43-52) (Ala-Gly-Ala-Cys-Ser-Thr-Cys-Ala-Gly-Lys)

Amino acid analysis and mobility at pH 1.9 showed the presence of one lysine. Dansylation without hydrolysis after nine Edman degradation steps confirmed lysine as the C-terminal.

Peptide X1T5 (53-71) (Val-Thr-Glx-Gly-Thr-Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe-Leu-Asx-Asx-Glx-Glx-Hse)

Eighteen steps of dansyl-Edman degradation were performed succesfully of this peptide. Digestion with carboxypeptidase A released homoserine after ½ h. Prolonged digestion released glutamine and trace of glutamic acid. This peptide must have been from C-terminal region of X1 fragment because its C-terminal sequence was identical with that of X1 fragment. Amides and acids of this peptide were determined after it was digested

with thermolysin. Three peptides X1T5H1, X1T5H2 and X1T5H3 were produced from the parent peptide. They were purified by paper electrophoresis at pH 6.5.

Peptide X1T5H1 (53-57) (Val-Thr-Glx-Gly-Thr)

Mobility at pH 6.5 suggested that residue 55 was glutamic acid.

Peptide X1T5H2 (58-65) (Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe)

Mobility at pH 6.5 indicated the presence of two amides and two acidic residues. The pH 6.5 mobilities after two, three and five steps of Edman degradation suggested that residue 59 was aspartic acid, residue 60 was glutamine, residue 62 was aspartic acid and residue 63 was glutamine.

Peptide X1T5H3 (66-71) (Leu-Asx-Asx-Glx-Glx-Hse)

The mobility at pH 6.5 suggested that this peptide had a charge of -3. From the C-terminal analysis of X1 fragment and X1T5, residue 70 was glutamine, then residues 67, 68 and 69 are acidic.

Peptide X2H1 (72-75) (Leu-Lys-Gly-Tyr)

Amino acid analysis and mobility at pH 1.9 of this peptide indicated the presence of one lysine. Dansylation without hydrolysis after 3 steps of Edman degradation yielded dansyl-bis-tyrosine.

Peptide X2H2 (76-79) (Val-Leu-Thr-Cys)

Cysteine was identified as C-terminal by dansylation without hydrolysis after 3 steps of Edman degradation.

Peptide X2H3 (80-88) (Ile-Ala-Tyr-Pro-Glu-Ser-Asp-Cys-Thr)

Mobility at pH 6.5 indicated that residues 84 and 86 were glutamic acid and aspartic acid respectively.

Peptide X2H4 (89-92) (Ile-Leu-Thr-His)

Dansylation without hydrolysis after 3 steps of Edman degradation yielded dansyl-bis-histidine.

Peptide X2H5 (93-96) (Val-Glx-Glx-Glx)

Mobility at pH 6.5 of this peptide suggested that it had a charge of -2. Mobilities at pH 6.5 after 2 and 3 steps of Edman degradation indicated that residue 94 was glutamic acid, residue 95 was glutamine and residue 96 was glutamic acid.

Peptide X2H6 (97-98) (Leu-Tyr)

After a single Edman degradation, dansylation without hydrolysis gave dansyl-bis-tyrosine.

Eleven peptides were produced after the whole protein was digested with chymotrypsin. Peptide C6 was obtained pure after gel chromatography, the remaining peptides were purified by paper electrophoresis.

Peptide C1 (1-3) (Ala-Asx-Tyr)

Residue 2 was identified as aspartic acid from the peptide mobility at pH 6.5.

Peptide C2 (4-6) (Lys-Ile-His)

The peptide mobility at pH 1.9 and pH 6.5 suggested the charge of +2. Dansylation without hydrolysis after two steps of Edman degradation yielded dansyl-bis-Histidine.

Peptide C3 (7-18) (Leu-Val-Ser-Lys-Glx-Glx-Gly-Ile-Asx-Val-Thr-Phe)

Mobility data confirmed that residues 11, 12 and 15 were acidic. Phenylalanine was determined as the C-terminal of the peptide by dansylation without hydrolysis after 11 steps of Edman degradation.

Peptide C4 (19-25) (Asx-Cys-Ser-Glx-Asx-Thr-Tyr)

Digestion with carboxypeptidase A for 30 min. followed by dansylation yielded an excess of bis-dansyl-tyrosine and a trace of dansyl-threonine. The pH 6.5 electrophoretic mobility indicated that residues 19, 22 and 23 were acidic.

Peptide C5 (26-39) (Ile-Leu-Asx-Ala-Ala-Glx-Glx-Glx-Gly-Ile-Glx-Leu-Pro-Tyr)

This peptide was purified in relatively low yield, as estimated by the intensity of the dansyl derivatives. Dansyl-Edman analysis became inconclusive after eleven steps of degradation. Residues 28, 31, 32, 33 and 36 were placed as acids from the peptides X1T3H4a and X1T3H4b and from the evidence obtained from the automatic sequencer.

Peptide C6 (40-65) (Ser-Cys-Arg-Ala-Gly-Ala-Cys-Ser-Thr-Cys-Ala-Gly-Lys-Val-Thr-Glx-Gly-Thr-Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe)

This peptide was purified by gel chromatography. The mobility of such a long peptide could not be determined because it tends to smear on paper electrophoresis. Digestion with carboxypeptidase A for 30 min. followed by dansylation yielded dansyl-phenylalanine and trace

of dansyl-serine. Dansyl-Edman analysis become inconclusive after sixteen steps of degradation. Residues 57-63 were placed from composition data and the sequence of X1T5. Position 55, 59 and 62 were placed as acidic residues and positions 60 and 63 as amides from the mobility evidence of peptides X1T5H1 and X1T5H2.

Peptide C7 (66-72) (Leu-Asx-Asx-Glx-Glx-Met-Leu)

The pH 6.5 electrophoretic mobility indicated the presence of three acidic and one amide residues. Positions 67, 68 and 69 were placed as acidic residues and position 70 as amide from the mobility evidence of peptide X1T5H3.

Peptide C8 (73-75) (Lys-Gly-Tyr)

Dansylation without hydrolysis after two Edman degradation steps yielded bis-dansyl-tyrosine.

Peptide C9 (76-77) (Val-Leu)

Dansylation without hydrolysis after one Edman degradation yielded dansyl-leucine.

Peptide C10 (78-90) (Thr-Cys-Ile-Ala-Tyr-Pro-Glx-Ser-Asx-Cys-Thr-Ile-Leu)

The pH 6.5 electrophoretic mobility indicated that residue 84 was glutamic acid and residue 86 was aspartic acid. Dansylation without hydrolysis after twelve steps of degradation yielded dansyl-leucine.

Peptide C11 (91-98) (Thr-His-Val-Glx-Glx-Glx-Leu-Tyr)

Dansylation without hydrolysis after seven steps of Edman degradation yielded bis-dansyl-tyrosine. The pH 6.5 mobility indicated the presence of two acid residues and these were placed at positions 94 and 96 and position 95 was placed as amide from the mobility evidence of peptide X2H5.

The overlapping of chymotryptic, tryptic and thermolytic peptides of ferredoxin, of its CNBr fragment gave the sequence of Porphyra ferredoxin as shown in Fig. 10. The sequence agreed with the amino acid composition of the protein together with that of the CNBr fragments as shown in Table 3. The amino acid composition of the peptides of Porphyra ferredoxin is given in Table 7. The N-terminal 40 residues of the protein have been determined by using an Automatic Sequencer.

All the acidic and amide residues indicated in the sequence were placed from the pH 6.5 electrophoretic mobilities of intact or partially degraded peptides.

The observed enzyme specificities were consistent with those expected (Smyth, 1967), except that full chymotryptic cleavage was observed at histidine 6.

TABLE 7

Amino acid compositions of peptides from *Porphyra* ferredoxin

Amino acids present in less than 0.3 residue were disregarded. This indicates that some peptides were not completely pure. Abbreviations: Hse, homoserine; CmCys, carboxymethylated cysteine; ——, value not determined.

Peptides	Asp	Thr	Ser	Glu	Pro	Gly	Ala	CmCys	Val	Met	Ile	Leu	Tyr	Phe	His	Lys	Arg	Hse
X1T1	0.61						0.60		0.65		0.85	0.87	0.75			0.70		
X1T2			0.72						1.36		2.37	2.25			0.55	0.67		
X1T3	3.83	2.40	1.52	7.00	0.36	2.27	2.30	——					1.54	0.78			0.80	
X1T4		0.95	0.61	0.62		1.90	2.30	——								0.98		
X1T5	4.10	1.70	1.63	5.40		1.20			2.10			0.86		0.83				0.52
X2H1						0.71						0.88	0.72			1.17		
X2H2		0.64						0.53	0.83			0.83						
X2H3	0.87	0.97	0.78	0.99	0.70		0.79	0.65			0.62		0.78					
X2H4		0.97									0.83	0.88			1.03			
X2H5				2.94					0.76									
X2H6												0.89	0.85					

C1	0.85						0.72						0.69				
C2											0.76				0.56	0.76	
C3	0.51	1.34	1.20	2.00		1.5			2.10		0.91	0.52		0.72		0.72	
C4	2.22	1.04	1.04	1.32				—					0.60				
C5	1.47			4.21	0.68	1.09	1.56				1.86	1.78	0.64				
C6	1.75	2.65	3.83	3.03		2.93	2.41	—	1.65					0.63		0.70	0.93
C7	1.58			1.87						0.51		1.64					
C8						1.21							0.70			0.50	
C9									1.23			1.30					
C10	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
C11		0.79		2.46					1.46			0.48	0.40		0.44		

III) The amino acid sequence of Triticum ferredoxin

The amino acid sequence of Triticum ferredoxin was determined by characterisation of the tryptic, chymotryptic peptides of CNBr fragments and by further digestion of one large tryptic peptide with thermolysin followed by sequence determination. The N-terminal 48 residues were determined by using an automatic sequencer. The sequence is shown in Fig.12 giving the points of CNBr and enzyme cleavage together with the overlapping peptides from which the sequence was deduced. The amino acid composition of Triticum ferredoxin together with that of CNBr fragments is shown in Table 8. A list of chymotryptic peptides of both fragments, together with electrophoretic mobility and sequence is given in Table 9 , and a similar table of tryptic peptides is given in Table 10. X1T3 peptide was purified by gel chromatography and subdigested with thermolysin. A list of the resulting peptides together with their mobility and sequence data is given in Table 11.

Cyanogen Bromide cleavage

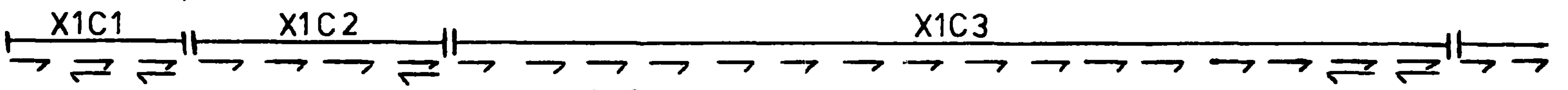
15-20 mg of carboxymethylated ferredoxin was used for CNBr cleavage. More than one cleavage was necessary because of different enzyme digestion for each fragment. As expected from the presence of a single methionine in the protein, CNBr cleavage produced two fragments which were purified by gel chromatography (Fig.13). The first peak eluting from the column was the uncleaved protein. The larger fragment (X1) had alanine at the

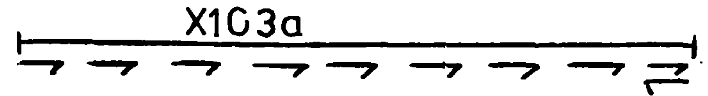
FIGURE 12.

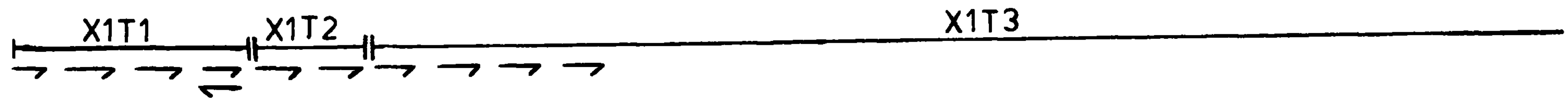
The amino acid sequence of Wheat ferredoxin. Residues which were identified by automatic sequencer, dansyl-Edman analysis and carboxypeptidase digestion are indicated by ⟶, ⟶, ⟵ respectively. The arrows - - -› indicate that residues were ambiguously identified. The arrow indicates CNBr clevage and * indicates homoserine was identified as the derivative from CNBr cleavage. For the peptide X1C6 the mobilities at pH 6.5 after some degradation cycles are indicated. T, C and H represent peptides obtained by tryptic, chymotryptic and thermolysin digestions respectively.

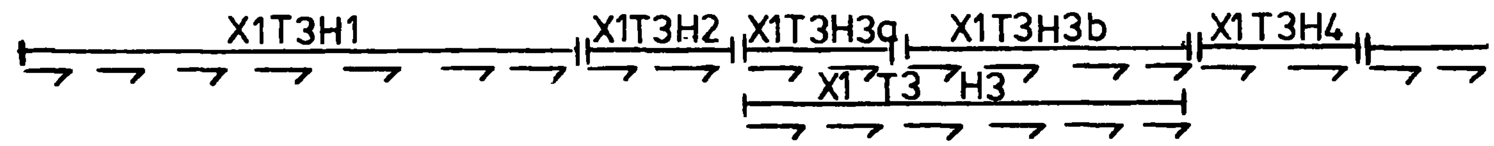
1 5 10 15 20 25

Ala-Thr-Tyr-Lys-Val-Lys-Leu-Val-Thr-Pro-Glu-Gly-Glu-Val-Glu-Leu-Glu-Val-Pro-Asp-Asp-Val-Tyr-Ile-Leu



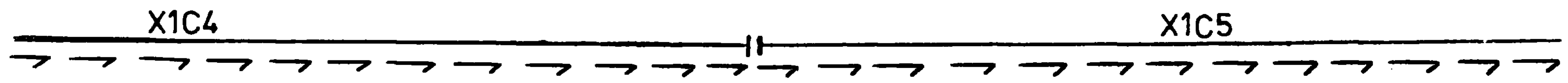


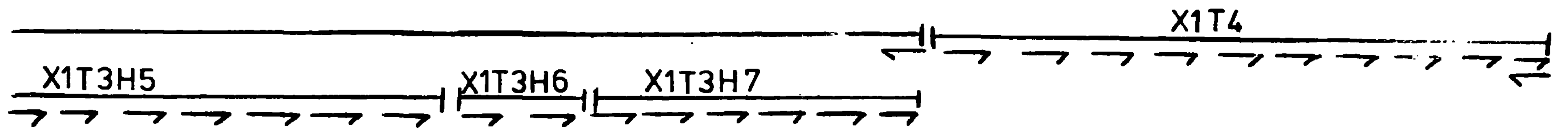




26 30 35 40 45 50

Asp-Gln-Ala-Glu-Glu-Glu-Gly-Ile-Asp-Leu-Pro-Tyr-Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys





51 55 60 65 70 75

Leu-Val-Ser-Gly-Glu-Ile-Asp-Gln-Ser-Asp-Gln-Ser-Phe-Leu-Asp-Asp-Asp-Gln-Met-Glu-Ala-Gly-Trp-Val-Leu

0,51 0,34 0,36 0,0

X1C6 X1C7 * X2C1 X2C2a

X2C2

X1T5 * X2T1

76 80 85 90 95

Thr-Cys-His-Ala-Tyr-Pro-Lys-Ser-Asp-Ile-Val-Ile-Glu-Thr-His-Lys-Glu-Glu-Glu-Leu-Thr-Ala

X2C2b X2C3 X2C4

X2T2 X2T3

FIGURE 13.

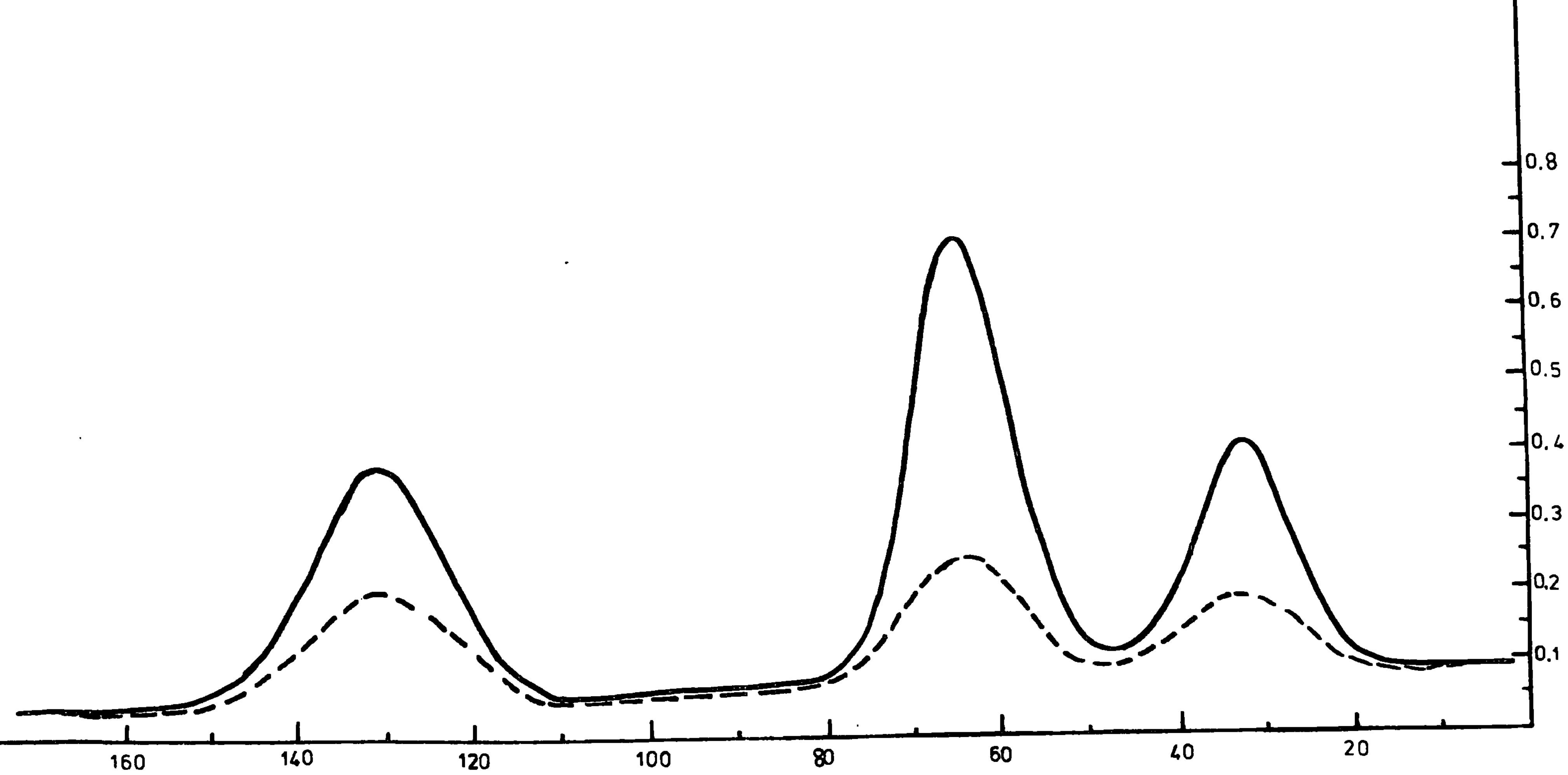
Chromatography of X1 and X2 fragments from *Triticum aestivum* ferredoxin on Sephadex G-75.

A 1 cm X 190 cm column of resin was used. The column was equilibrated and eluted by 70% formic acid.

——— absorbance at 280 nm

- - - - absorbance at 206 nm

The absorbance at 206 nm was very low due to the lack of sensitivity of the Uvicord instrument.



0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
ABSORBANCE
160
140
120
100
80
60
40
20
ELUTION VOLUME (ml)

TABLE 8.

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

The results are expressed as residue/molecule. Cysteine was measured as cysteic acid . Tryptophan was determined after alkaline hydrolysis for 72 h (Noltmann et al, 1962). Average values are given except for serine and threonine, which are extrapolated to zero time where possible, assuming first-order rate of destruction (Moore and Stein, 1963) and for valine and isoleucine where maximal values (72 h hydrolysis) are given. Sequence values from sequence determination.

Abbreviations: n.d. not determined; Hse, homoserine.

TABLE 8

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

	Total protein					X1 fragment				X2 fragment			
	24h	48h	72h	(1) Average	(2) Seq	24h	72h	(1) Average	(2) Seq	24h	72h	(1) Average	(2) Seq
Asx	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
Glx	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	0	0	0	0	n.d.	n.d.	n.d.	1
total residues					97				69				28

TABLE 9.

Chymotryptic peptides of CNBr fragments from *Triticum* ferredoxin

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
X1C1 (1-3)	0	0.55	Ala-Thr-Tyr
X1C2 (4-7)	2.58	1.40	Lys-Val-Lys-Leu
X1C3 (8-23)	-2.06	0.20	Val-Thr-Pro-Glx-Gly-Glx-Val-Glx-Leu-Glx-Val-Pro-Asx-Asx-Val-Tyr
X1C3a (8-16)	-1.81	0.25	Val-Thr-Pro-Glx-Gly-Glx-Val-Glx-Leu
X1C4 (24-37)	-2.00	0.21	Ile-Leu-Asx-Glx-Ala-Glx-Glx-Glx-Gly-Ile-Asx-Leu-Pro-Tyr
X1C5 (38-51)	0.60	0.38	Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys-Leu
X1C6 (52-63)	-1.56	0.20	Val-Ser-Gly-Glx-Ile-Asx-Glx-Ser-Asx-Glx-Ser-Phe
X1C7 (64-69)	-2.00	0.32	Leu-Asx-Asx-Asx-Glx-Hse
X2C1 (70-73)	-1.23	0.45	Glx-Ala-Gly-Trp
X2C2 (74-78)	0	0.70	Val-Leu-Thr-Cys-His
X2C2a (74-75)	0	0.57	Val-Leu
X2C2b (76-78)	0	0.80	Thr-Cys-His

Continued. . . .

Continued. . . . .

X2C3 (79-95)	-0.66	0.48	Ala-Tyr-Pro-Lys-Ser-Asx-Ile-Val-Ile-Glx-Thr-His-Lys-(Glx-Glx-Glx)-Leu
X2C4 (96-97)	0	0.85	Thr-Ala

TABLE 10.

Tryptic peptides of CNBr fragment from *Triticum* ferredoxin

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
X1T1 (1-4)	1.26	0.93	Ala-Thr-Tyr-Lys
X1T2 (5-6)	1.84	1.44	Val-Lys
X1T3 (7-40)	-1.92	0.26	Leu-Val-Thr-Pro-(was digested with thermolysin)-Arg
X1T4 (41-50)	-0.74	0.50	Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys
X1T5 (51-69)	-1.67	0.16	Leu-Val-Ser-Gly-Glx-Ile-Asx-Glx-Ser-Asx-Glx-Ser-Phe-Leu-(Asx-Asx-Asx)-Glx-Hse
X2T1	0	0.40	Glx-Ala-Gly-(Trp)-Val-Leu-Thr-Cys-His-Ala-Tyr-Pro-Lys
X2T2	0	0.51	Ser-Asx-Ile-Val-Ile-Glx-Thr-His-Lys
X2T3	-2.22	0.30	Glx-Glx-Glx-Leu-Thr-Ala

TABLE 11.

Thermolytic peptides from X1T3

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman Results
X1T3H1	-1.32	0.28	Leu-Val-Thr-Pro-Glx-Gly-Glx
X1T3H2	-1.73	0.61	Val-Glx
X1T3H3	-2.20	0.32	Leu-Glx-Val-Pro-Asx-Asx
X1T3H3a	-1.73	0.61	Leu-Glx
X1T3H3b	-1.97	0.41	Val-Pro-Asx-Asx
X1T3H4	0	0.61	Val-Tyr
X1T3H5	-2.20	0.24	Ile-Leu-Asx-Glx-Ala-Glx-Glx-Glx-Gly
X1T3H6	-1.70	0.60	Ile-Asx
X1T3H7	0.82	0.58	Leu-Pro-Tyr-Ser-Cys-Arg

N-terminus identical with that of the protein and homoserine as the C-terminus. Fragment X2 had glutamic acid as the N-terminus and a C-terminus identical with that of the intact protein. The order of the fragments in the protein could therefore be unequivocally established.

Subdigestions

Five tryptic peptides were produced when fragment X1 was digested with trypsin. The order of the five resulting peptides in the parent fragment X1 could be established from the results, since peptides T1, T2, T3 and T4 were overlappedby that of the 48 N-terminal residues. The remaining peptide, T5, which has homoserine as its C-terminal could also be positioned from the information in the X1 fragment. Furthermore, all the tryptic peptides were overlapped by the seven chymotryptic peptides produced when X1 fragment was digested with chymotrypsin. The sequence of tryptic and chymotryptic peptides were established by direct Edman degradation except T3 which was further digested with thermolysin and the resulting peptides purified by paper electrophoresis at pH 1.9, pH 6.5 and by paper chromatography.

The X2 fragment was cleaved into three peptides by trypsin. The order of the peptides X2T1, X2T2 and X2T3 in the parent fragment X2 could be established from these data, since peptide T1 had the same N-terminal as those 3 residues of fragment X2 and the sequence of the two C-terminal residues of the peptide T3 was identical with

those of fragment X2. Furthermore, all these tryptic peptides were overlapped by chymotryptic peptides. The sequence of these peptides were established by direct Edman degradation.

The amino acid composition of peptides was obtained from one sample hydrolysed for 24 h and this is shown in Table 12.

Chymotryptic peptides

Peptide X1C1 (1-3) (Ala-Thr-Tyr)

Digestion with carboxypeptidase A for 45 min followed by dansylation yielded an excess of bis-dansyl-tyrosine and a trace of dansyl-threonine.

Peptide X1C2 (4-7) (Lys-Val-Lys-Leu)

Digestion with carboxypeptidase A for 45 min followed by dansylation yielded dansyl-leucine.

Peptide X1C3 (8-23) (Val-Thr-Pro-Glx-Gly-Glx-Val-Glx-Leu-Glx-Val-Pro-Asx-Asx-Val-Tyr)

PTH-derivatives of residues 11, 13, 15 and 17 were identified as glutamic acids and residues 20 and 21 as aspartic acids. This was confirmed from thermolytic peptides followed by dansylation yielded bis-dansyl-tyrosine and a trace of dansyl-valine.

Peptide X1C3a (8-16) (Val-Thr-Pro-Glx-Gly-Glx-Val-Glx-Leu)

Peptide X1C4 (24-37) (Ile-Leu-Asx-Glx-Ala-Glx-Glx-Glx-Gly-Ile-Asx-Leu-Pro-Tyr)

Dansylation without hydrolysis after thirteen steps of Edman degradation yielded bis-dansyl-tyrosine. PTH-derivative of residue 27 was identified as glutamine. Residues 29, 30 and 31 were identified as glutamic acids

from their PTH-derivatives and residues 26 and 34 were identified as aspartic acids.

Peptide X1C5 (38-51) (Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys-Leu)

Digestion with carboxypeptidase A for 45 min gave dansyl-leucine after dansyl analysis.

Peptide X1C6 (52-63) (Val-Ser-Gly-Glx-Ile-Asx-Glx-Ser-Asx-Glx-Ser-Phe)

Amides and acids were established from changes in the electrophoretic mobility at pH 6.5 after successive cycles of Edman degradation (see Fig.12). Residues 58 and 61 were identified as glutamines, residue 55 as glutamic acid and residues 57 and 60 as aspartic acids. Digestion with carboxypeptidase A for 30 min gave dansyl-phenylalanine after dansyl analysis.

Peptide X1C7 (64-69) (Leu-Asx-Asx-Asx-Glx-Hse)

Digestion with carboxypeptidase A for 2 h followed by dansylation gave dansyl-homoserine together with some dansyl-glutamine. The pH 6.5 mobility suggested that this peptide had a charge of -3, therefore, residues 65, 66 and 67 were aspartic acids.

Peptide X2C1 (70-73) (Glx-Ala-Gly-Trp)

This peptide a positive reaction with the Ehrlich reagent indicating the presence of tryptophan. Digestion with carboxypeptidase A for 1 h followed by dansylation gave dansyl-tryptophan together with some dansyl-glycine. The mobility at pH 6.5 indicated that residue 70 was glutamic acid.

Peptide X2C2 (74-78) (Val-Leu-Thr-Cys-His)

Dansylation without hydrolysis after four steps of Edman degradation yielded a weak spot of bis-dansyl-histidine although the semi quantitative analysis gave an excess of bis-dansyl-histidine after dansylation.

Peptide X2C2a (74-75) (Val-Leu)

Peptide X2C2b (76-78) (Thr-Cys-His)

Peptide X2C3 (79-95) (Ala-Tyr-Pro-Lys-Ser-Asx-Ile-Val-Ile-Glx-Thr-His-Lys-Glx-Glx-Glx-Leu)

Dansyl Edman analysis of this peptide became inconclusive after twelve steps of degradation. Digestion with carboxypeptidase A for 1 h followed by dansylation yielded dansyl-leucine. A consideration of the pH 6.5 electrophoretic mobility indicated that residue 84 was aspartic acid and residue 88 as glutamic acid. Residues 91, 92 and 93 were placed as glutamic acids from the electrophoretic mobility and the sequence evidence of peptide X2T3.

Peptide X2C4 (96-97) (Thr-Ala)

Dansylation without hydrolysis after one step of Edman degradation gave dansyl-alanine.

Tryptic peptides

Peptide X1T1 (1-4) (Ala-Thr-Tyr-Lys)

Digestion with carboxypeptidase B for 1 h followed by dansylation yielded dansyl-lysine.

Peptide X1T2 (5-6) (Val-Lys)

Dansylation without hydrolysis after one step of Edman degradation yielded bis-dansyl-lysine.

Peptide X1T3 (7-40) (Leu-Val-Thr-Pro-was digested with thermolysin-Arg)

The amino acid analysis showed that this peptide consisted of 34 residues. Three steps of Edman degradation and dansyl analysis suggested that the N-terminal of this peptide corresponded to that of peptide X1T3. Digestion with carboxypeptidase B released arginine. Digestion with thermolysin for 2 h produced nine peptides which were isolated and purified by paper electrophoresis at pH 6.5, pH 1.9 and paper chromatography. Edman degradation and dansyl analysis of each thermolytic peptide was carried out.

Peptide X1T3H1 (7-13) (Leu-Val-Thr-Pro-Glx-Gly-Glx)

Residues 11 and 13 were identified as glutamic acids from their PTH-derivatives and the peptide electrophoretic mobility at pH 6.5.

Peptide X1T3H2 (14-15) (Val-Glx)

The PTH-derivative of residue suggested that it was glutamic acid and it was confirmed from the peptide electrophoretic mobility at pH 6.5.

Peptide X1T3H3 (16-21) (Leu-Glx-Val-Pro-Asx-Asx)

Dansylation without hydrolysis after five steps of Edman degradation yielded dansyl-aspartic acid. Residues 17, 20 and 21 were identified as acids from their PTH-derivatives and the electrophoretic mobilities at pH 6.5 of this peptide, X1T3H3a and X1T3H3b peptides.

Peptide X1T3H3a (16-17) (Leu-Glx)

Peptide X1T3H3b (18-21) (Val-Pro-Asx-Asx)

Peptide X1T3H4 (22-23) (Val-Tyr)

Dansylation without hydrolysis after one step of Edman degradation yielded bis-dansyl-tyrosine.

Peptide X1T3H5 (24-32) (Ile-Leu-Asx-Glx-Ala-Glx-Glx-Glx-Gly)

The acidic and the amide residues of this peptide were identified from their PTH-derivatives. Residue 27 was identified as glutamine, residue 26 as aspartic acid and residues 29, 30 and 31 as glutamic acids. Dansylation without hydrolysis after eight steps of Edman degradation yielded dansyl-glycine.

Peptide X1T3H6 (33-34) (Ile-Asx)

The electrophoretic mobility at pH 6.5 suggested that residue 34 was aspartic acid. This was confirmed from the PTH-derivative.

Peptide X1T3H7 (35-40) (Leu-Pro-Tyr-Ser-Cys-Arg)

Dansylation without hydrolysis after five steps of Edman degradation yielded dansyl-arginine which is the C-terminal of the whole peptide X1T3.

Peptide X1T4 (41-50) (Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys)

Digestion with carboxypeptidase B for 1 h released lysine which was identified as bis-dansyl-lysine after dansylation.

Peptide X1T5 (51-69) (Leu-Val-Ser-Gly-Glx-Ile-Asx-Glx-Ser-Asx-Glx-Ser-Phe-Leu-Asx-Asx-Asx-Glx-Hse)

Digestion with carboxypeptidase A for 2 h followed by dansylation yielded dansyl-homoserine and hint of dansyl-glutamine. Amides and acids were identified from

the mobility evidence of peptides X1C6 and X1C7. Residue 55 was identified as glutamic acid, residues 58 and 61 as glutamines and residues 57, 60, 65, 66 and 67 as aspartic acids. Residue 68 was identified as glutamine from the C-terminal analysis.

Peptide X2T1 (70-82) (Glx-Ala-Gly-Trp-Val-Leu-Thr-Cys-His-Ala-Tyr-Pro-Lys)

Digestion with carboxypeptidase B followed by dansylation yielded a trace of bis-dansyl-lysine. The peptide gave a positive reaction with Ehrlich reagent indicating the presence of tryptophan, which was not identified on N-terminal analysis after three degradation steps because of the acid destruction of dansyl-trypto phan. Residue 70 was identified as glutamic acid from the evidence of peptide X2C1.

Peptide X2T2 (83-91) (Ser-Asx-Ile-Val-Ile-Glx-Thr-His-Lys)

The electrophoretic mobility at pH 6.5 suggested that residue 84 was aspartic acid and residue 88 was glutamic acid. Digestion with carboxypeptidase B for 1 h yielded bis-dansyl-lysine. The peptide was purified in low yield, as estimated by the intensity of the dansyl derivatives.

Peptide X2T3 (92-97) (Glx-Glx-Glx-Leu-Thr-Ala)

The electrophoretic mobility at pH 6.5 suggested that residues 92, 93 and 94 were glutamic acids. Digestion with carboxypeptidase A for 1 h followed by dansylation yielded a strong spot of dansyl-alanine and trace of dansyl-threonine together with a hint of dansyl-leucine.

The overlapping chymotryptic and tryptic peptides of both the CNBr fragments gave the sequence of wheat ferredoxin as shown in Fig. 12. The sequence agreed with the amino acid composition of the protein and also with those of the CNBr fragments as shown in Table 8 . The N-terminal 48 residues of the protein were determined by using an automatic sequencer. Amides and acids of residues included in these 48 residues were determined from their PTH-derivatives. The rest of the amides and acids indicated in the sequence were placed from the pH 6.5 electrophoretic mobilities of intact or degraded peptides. Glutamine 68 was placed from the C-terminal analysis of both X1T5 and X1C7 peptides.

The observed enzyme specificities were consistent with those expected (Smyth, 1967) except that full chymotryptic cleavage occurred at histidine 78 and tryptic cleavage at lysine 82. Partial thermolytic cleavage was observed at valine 18, leucine 25.

TABLE 12

Amino acid compositions of peptides from *Triticum aestivum* ferredoxin

Peptide	Asx	Thr	Ser	Glx	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	His	Arg	Lys	Hse	Cm-Cys
X1T1		0.85					1.09				0.62				0.82		
X1T2								0.61							0.83		
X1T3	4.45	0.90	1.30	8.03	2.75	2.13	1.09	3.65	1.77	3.66	1.66			0.75			0.9
X1T4			2.85			2.11	2.02								1.07		(2)
X1T5	4.75		2.63	4.41		1.24		1.03	0.91	1.76		0.72				0.43	
X1C1		1.10					1.04				1.03						
X1C2								0.72		0.56					1.70		
X1C4	2.11			4.21	0.91	0.69	0.78		2.05	2.01	0.90						
X1C5			3.88			2.00	2.00			0.66				0.89	0.88		2.6
X2T1		0.77		1.29	0.66	0.95	2.05	1.00		1.19	1.09		0.77		0.55		0.6
X2T2	1.17	0.79	0.62	0.95				0.99	2.15				1.04		1.33		
X2T3		0.78		3.11			0.99			0.95							
X2C1				1.05		0.98	0.88										
X2C2		1.22						1.05		0.78		1.00					

X2C2a			0.93	0.99	
X2C2b	0.58				1.00
X2C4	0.95	1.02			

IV) The amino acid sequence of *Sambucus* ferredoxin

The amino acid sequence of *Sambucus* ferredoxin was determined from the evidence obtained from one chymotryptic and one tryptic digestions of the protein, from data obtained with a papain digest of T3 peptide together with the data obtained by using a Beckman Automatic Sequencer. The sequence is shown in Fig.14 giving points of enzymic cleavage together with the overlapping peptides from which the sequence was deduced. A list of chymotryptic peptides together with their electrophoretic mobilities and sequences is given in Table13 and a similar list of tryptic peptides is given in Table14. The amino acid composition of the protein was obtained from duplicate samples hydrolysed for 24, 48 and 72 h respectively and this information is shown in Table15. The amino acid composition of the peptides was obtained from a single sample hydrolysed for 24 h and this is shown in Table16.

The tryptic and chymotryptic digestions were performed on about 15 mg of carboxymethylated ferredoxin for each digest. The automatic sequencer data was obtained using 6 mg of carboxymethylated ferredoxin.

Chymotryptic peptides

Peptide C1 (1-3) (Ala-Ser-Tyr)

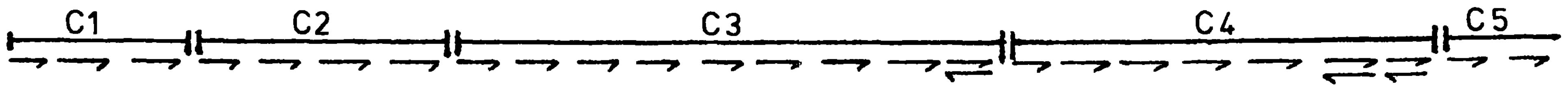
Dansylation without hydrolysis after two steps of Edman degradation yielded bis-dansyl-tyrosine.

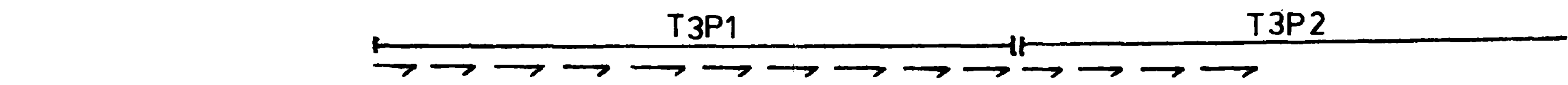
FIGURE 14

The amino acid sequence of Sambucus ferredoxin. Residues which were identified by automatic sequencer, dansyl-Edman analysis and carboxypeptidase digestion are indicated by ⟶ , ⟶ , ⟵ respectively. T, C and P represent peptides obtained by tryptic, chymotryptic and papain digestion respectively.

1 5 10 15 20 25

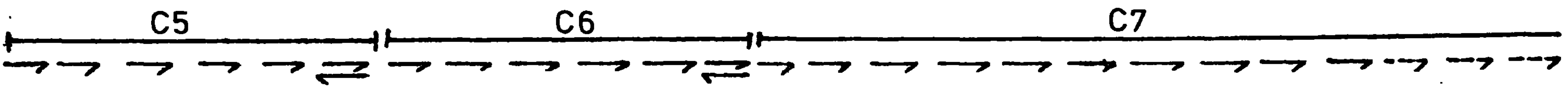
Ala-Ser-Tyr-Lys-Val-Lys-Leu-Ile-Thr-Pro-Asp-Gly-Pro-Gln-Glu-Phe-Glu-Cys-Pro-Asp-Asp-Val-Tyr-Ile-Leu

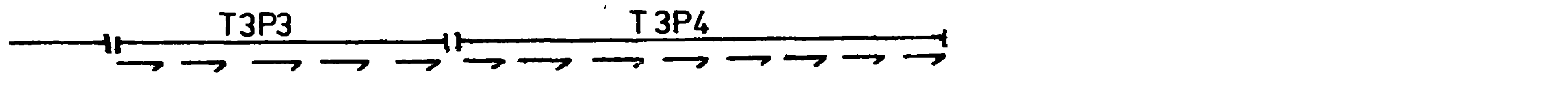




26 30 35 40 45 50

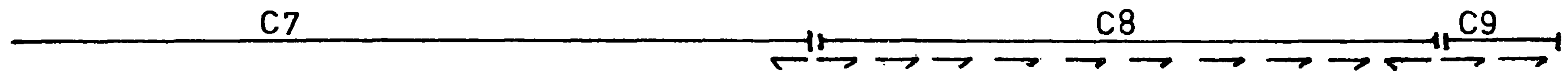
Glu-His-Ala-Glu-Glu-Leu-Gly-Ile-Asp-Ile-Pro-Tyr-Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys

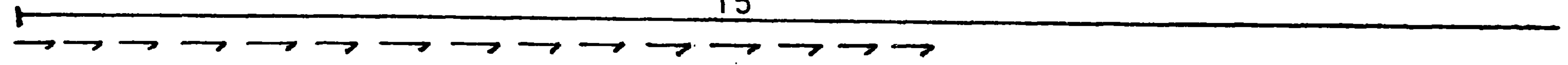




51 55 60 65 70 75

Leu-Val-Ala-Gly-Ser-Val-Asp-Gln-Ser-Asp-Gln-Ser-Phe-Leu-Asp-Asp-Glu-Gln-Ile-Glu-Glu-Gly-Trp-Val-Leu





76 80 85 90 95 97

Thr-Cys-Val-Ala-Tyr-Pro-Lys-Ser-Asp-Val-Thr-Ile-Glu-Thr-His-Lys-Glu-Glu-Glu-Leu-Thr-Ala

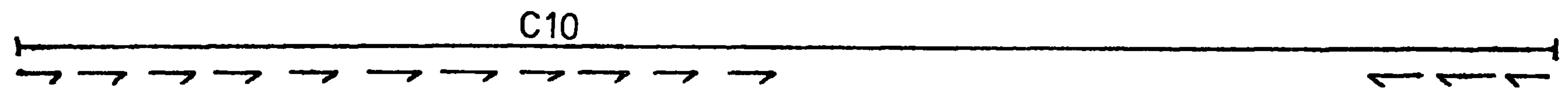


TABLE 13.

Chymotryptic peptides of *Sambucus* ferredoxin

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
C1 (1-3)	0	0.58	Ala-Ser-Tyr
C2 (4-7)	2.39	1.39	Lys-Val-Lys-Leu
C3 (8-16)	-1.34	0.20	Ile-Thr-Pro-Asx- Gly-Pro-Glx-Glx- Phe
C4 (17-23)	-1.65	—	Glx-Cys-Pro-Asx- Asx-Val-Tyr
C5 (24-31)	-1.38	—	Ile-Leu-Glx-His- Ala-Glx-Glx-Leu
C6 (32-37)	-0.65	0.28	Gly-Ile-Asx-Ile- Pro-Tyr
C7 (38-63)	—	—	Ser-Cys-Arg-Ala- Gly-Ser-Cys-Ser- Ser-Cys-(Ala-Gly- Lys-Leu-Val-Ala- Gly-Ser-Val-Asx- Glx-Ser-Asx-Glx-) Ser-Phe
C8 (64-73)	-2.23	—	Leu-Asx-Asx-Glx- Glx-Ile-Glx-Glx- Gly-Trp
C9 (74-75)	0	0.63	Val-Leu
C10 (76-97)	-0.23	—	Thr-Cys-Val-Ala- Tyr-Pro-Lys-Ser- Asx-Val-(Thr-Ile- Glx-Thr-His-Lys- Glx-Glx-Glx-)Leu- Thr-Ala

TABLE 14.

Tryptic peptides from *Sambucus* ferredoxin

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
T1 (1-4)	1.09	0.85	Ala-Ser-Tyr-Lys
T2 (5-6)	1.92	1.33	Val-Lys
T3 (7-40)	——	——	This peptide was digested with Papain
T3P1 (7-16)	-1.47	——	Leu-Ile-Thr-Pro-Asx-Gly-Pro-Glx-Gly-Phe
T3P2 (17-27)	-1.25	——	Glx-Cys-Pro-Asx-(Asx-Val-Tyr-Ile-Leu-Glx-His)
T3P3 (28-32)	-1.52	——	Ala-Glx-Glx-Leu-Gly
T3P4 (33-40)	-0.55	——	Ile-Asx-Ile-Pro-Tyr-Ser-Cys-Arg
T4 (41-50)	0	——	Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys
T5 (51-91)	——	——	Leu-Val-Ala-Gly-Ser-Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe-Leu-Asx-(Asx-Glx-Glx-Ile-Glx-Glx-Gly-Trp-Val-Leu-Thr-Cys-Val-Ala-Tyr-Pro-Lys-Ser-Asx-Val-Thr-Ile-Glx-Thr-His-) Lys
T6 (92-97)	-2.00	0.26	Glx-Glx-Glx-Leu-Thr-Ala

TABLE 15.

The amino acid composition of *Sambucus nigra* ferredoxin

	Mean values 24h hydrolysis	Mean values 48h hydrolysis	Mean values 72h 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	1.00	1.00	1.00	1
Cm-Cys	5.11	5.02	5.14	5.09	5
Trp	---	---	---		1

*Calculated from 24h and 72h values assuming first order kinetics for destructions (Moore and Stein 1963).

+For valine, isoleucine and leucine maximal values (72h hydrolysis) were taken.

Peptide C2 (4-7) (Lys-Val-Lys-Leu)

Dansylation without hydrolysis after three steps of Edman degradation gave dansyl-leucine.

Peptide C3 (8-16) (Ile-Thr-Pro-Asx-Gly-Pro-Glx-Glx-Phe)

Digestion with carboxypeptidase A for 1 h yielded dansyl-phenylalanine after dansyl analysis. The pH 6.5 electrophoretic mobility indicated the presence of two acidic residues and these were placed at positions 11 and 15 as aspartic and glutamic acid respectively from evidence obtained from the automatic sequencer. The PTH-derivative of residue 14 was identified as glutamine.

Peptide C4 (17-23) (Glx-Cys-Pro-Asx-Asx-Val-Tyr)

Digestion with carboxypeptidase A for 30 min yielded bis-dansyl-tyrosine after dansyl analysis as did dansylation without hydrolysis after six Edman degradation steps. The pH 6.5 electrophoretic mobility showed that residues 17, 20 and 21 were acidic and this was confirmed from evidence obtained from automatic sequencer.

Peptide C5 (24-31) (Ile-Leu-Glx-His-Ala-Glx-Glx-Leu)

Digestion with carboxypeptidase A for 1 h yielded dansyl-leucine after dansyl analysis. The pH 6.5 electrophoretic mobility of the peptide and the PTH-derivatives showed that residues 26, 29 and 30 were acidic.

Peptide C6 (32-37) (Gly-Ile-Asx-Ile-Pro-Tyr)

Dansylation without hydrolysis after five steps of Edman degradation gave bis-dansyl-tyrosine. The pH 6.5

electrophoretic mobility indicated that residue 34 was aspartic acid. This was supported by the evidence of the automatic sequencer.

Peptide C7 (38-63) (Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys-Leu-Val-Ala-Gly-Ser-Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe)

The dansyl-Edman analysis was inconclusive beyond nine degradation steps. Digestion with carboxypeptidase A for 30 min followed by dansyl analysis yielded dansyl-phenylalanine, whilst digestion for 1.5 h gave dansyl-phenylalanine and dansyl-serine. Attempts to find the electrophoretic mobility were not successful because the peptide tended to smear along the electrophoresis paper. Residues 57, 58 and 61 and 62 were placed according to homology of other ferredoxin sequences as aspartic acid, glutamine, aspartic acid and glutamine respectively.

Peptide C8 (64-73) (Leu-Asx-Asx-Glx-Glx-Ile-Glx-Glx-Gly-Trp)

The peptide gave a positive result with Ehrlich reagent and digestion with carboxypeptidase A for 30 min followed by dansyl analysis yielded dansyl-tryptophan, whilst a similar analysis after 1.5 h of digestion yielded dansyl-tryptophan and dansyl-glycine. The electrophoretic mobility at pH 6.5 showed that this peptide is highly acidic. Residue 68 was located as glutamine from the homology of other ferredoxin sequences.

Peptide C9 (74-75) (Val-Leu)

Dansylation without hydrolysis after one step of Edman degradation yielded dansyl-leucine.

Peptide C10 (76-97) (Thr-Cys-Val-Ala-Tyr-Pro-Lys-Ser-
Asx-Val-Thr-Ile-Glx-Thr-His-Lys-
Glx-Glx-Glx-Leu-Thr-Ala)

The dansyl-Edman analysis became inconclusive after ten degradation steps. Residues 87-94 were placed from the amino acid composition data and from the evidence obtained with T6. Digestion with carboxypeptidase A for one hour yielded dansyl-alanine and trace of dansyl-threonine after dansyl analysis, whilst 3 h of digestion followed by dansyl analysis yielded dansyl-alanine and dansyl-threonine together with a significant amount of dansyl-leucine. The electrophoretic mobility at pH 6.5 indicated that residues 88, 92, 93 and 94 were glutamic acids and residue 84 was aspartic acid.

Peptide C10a (96-97) (Thr-Ala)

Tryptic peptides

Six tryptic peptides were produced when the protein was digested for 2 h with trypsin. The peptides were separated by gel chromatography. Two large peptides T3 and T5 were subdigested with papain and thermolysin respectively. The resulting peptides were purified by paper electrophoresis at pH 1.9 and 6.5.

Peptide T1 (1-4) (Ala-Ser-Tyr-Lys)

Dansylation without hydrolysis after three steps of Edman degradation yielded bis-dansyl-lysine.

Peptide T2 (5-6) (Val-Lys)

Dansylation without hydrolysis after one Edman degradation step gave excess of bis-dansyl-lysine.

Peptide T3 (7-40)

Digestion with carboxypeptidase B for 1 h gave dansyl-arginine after dansyl analysis. The electrophoretic mobility was not determined because such long peptide tends to smear on the electrophoresis paper. Four peptides were obtained after this peptide was digested with papain for 2 h. The resulting peptides were separated by paper electrophoresis.

Peptide T3P1 (7-16) (Leu-Ile-Thr-Pro-Asx-Gly-Pro-Glx-Glx-Phe)

Residues 11 and 15 were identified as acids and residue 14 as amide from their PTH-derivatives and the C3 electrophoretic mobility at pH 6.5.

Peptide T3P2 (17-27) (Glx-Cys-Pro-Asx-Asx-Val-Tyr-Ile-Leu-Glx-His)

The dansyl-Edman analysis was inconclusive beyond four degradation steps. Residues 21 to 27 were placed from the semiquantitative amino acid composition data of T3P2 and the sequence data of C4 and C5. The electrophoretic mobility at pH 6.5 show that no amide is present. This conclusion was supported by the evidence from the Automatic Sequencer.

Peptide T3P3 (28-32) (Ala-Glx-Glx-Leu-Gly)

Residues 29 and 30 were identified as acidics from the electrophoretic mobility at pH 6.5. This was confirmed by the evidence from the Automatic Sequencer.

Peptide T3P4 (33-40) (Ile-Asx-Ile-Pro-Tyr-Ser-Cys-Arg)

The electrophoretic mobility at pH 6.5 indicated that residue 34 was aspartic acid and this was confirmed from the mobility of the peptide C6.

Peptide T4 (41-50) (Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys)

Digestion with carboxypeptidase B for 1 h followed by dansylation analysis yielded bis-dansyl-lysine and trace of dansyl-glycine.

Peptide T5 (51-91) (Leu-Val-Ala-Gly-Ser-Val-Asx-Glx-
Ser-Asx-Glx-Ser-Phe-Leu-Asx- - -
- - - - - - - Lys)

This large peptide was purified by gel chromatography. The first fifteen residues were identified by direct dansyl-Edman analysis and the C-terminal was proved as lysine after digestion with carboxypeptidase B for 1 h followed by dansylation analysis. Residues from 66-90 were placed from sequence data of C8, C9 and C10.

Peptide T6 (92-97) (Glx-Glx-Glx-Leu-Thr-Ala)

The electrophoretic mobility at pH 6.5 indicated that residues 92, 93 and 94 were acids. Dansylation without hydrolysis after five steps of Edman degradation yielded dansyl-alanine.

Amino acid sequence

The determination of the sequence of the overlapping chymotryptic and tryptic peptides allowed the complete sequence of Sambucus ferredoxin to be deduced as shown in Fig. 14. Peptide T3 was subdigested with papain. The sequence data agreed with the amino acid composition data shown in Table 15 although tryptophan was not determined as part of the amino acid composition.

From Fig. 14 it can be seen that all the residues

were positively identified by chymotryptic, tryptic and papain digestions except residues 87-90 which were located according to amino acid composition of C10 and homology of other ferredoxin sequences. The first 40 residues were identified by the Automatic Sequencer. The tryptophan residue present was not identified in the appropriate tryptic peptide because of the destruction of the dansyl derivative during the acid hydrolysis stage of dansyl analysis, but it was positively identified in peptide C8 by Ehrlich test and C-terminal analysis with carboxypeptidase A.

All the acidic and amide residues indicated in the sequence were placed from the electrophoretic mobilities at pH 6.5 of intact peptides except residues 57, 58, 60, 61 and 68 which were located according to homology of other ferredoxin sequences as aspartic acid, glutamine, aspartic acid, glutamine and glutamine respectively. Amides and acids in the first 40 residues were placed from a combination of the above, together with analysis associated with the Automatic Sequencer.

The peptides obtained by digestion with chymotrypsin and trypsin had terminal amino acid that were consistent with the specificity of these enzymes.

TABLE 16

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		1.06			

V) The amino acid sequence of *Brassica napus* ferredoxin

The amino acid sequence of *Brassica napus* was determined from the evidence of one chymotryptic digestion and two tryptic digestions using a total of 5 umol. (60 mg) of protein. Thermolysin was used to subdigest two large tryptic peptides T3 and T5. The complete sequence is given in Fig.15 showing the points of enzyme cleavage together with the overlapping peptides from which the sequence was deduced. A list of chymotryptic peptides, together with mobility and sequence data is given in Table 17, and similar lists of tryptic and thermolytic peptides are given in Table 18 and Table 19. The amino acid composition of *Brassica napus* was obtained from two duplicate samples at each hydrolysis time i.e., 24, 48 and 72 h and this is given in Table 20 and similar list for peptide amino acid composition obtained from one sample of each peptide hydrolysed for 24 h is given in Table 21.

Chymotryptic peptides

20 mg of protein were digested for 2 h with 2% (w/w) enzyme. The resulting peptides (nine) were separated by gel chromatography using Biogel P-4 in 70% formic acid. The peptides which remained contaminated were further purified by paper electrophoresis at pH 6.5 and 1.9.

Peptide C1 (1-3) (Ala-Thr-Tyr)

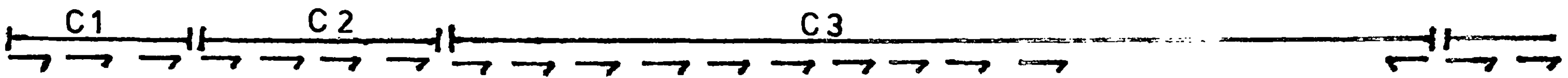
Dansylation without hydrolysis after two dansyl Edman degradation yielded bis-dansyl-tyrosine.

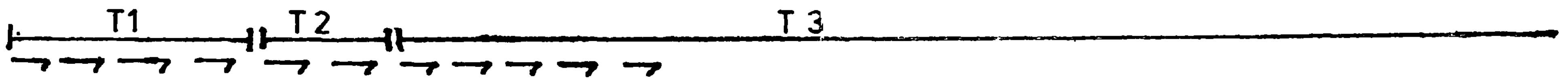
FIGURE 15.

The amino acid sequence of *Brassica napus* ferredoxin. Residues which were identified by automatic sequencer, dansyl-Edman analysis and carboxypeptidase digestion are indicated by ⟶, ⇁, ↽ respectively. T, C, and H represent peptides obtained by tryptic, chymotryptic and thermolytic digestions respectively.

1 5 10 15 20 25

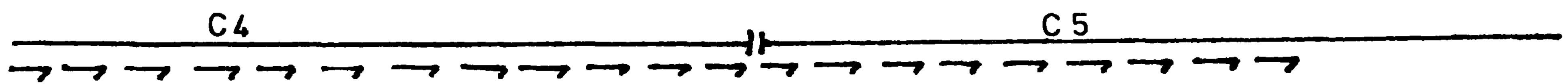
Ala-Thr-Tyr-Lys-Val-Lys-Phe-Ile-Thr-Pro-Glu-Gly-Glu-Gln-Glu-Val-Glu-Cys-Asp-Asp-Asp-Val-Tyr-Val-Leu

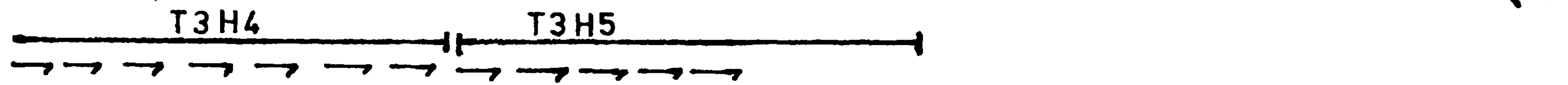




26 30 35 40 45 50

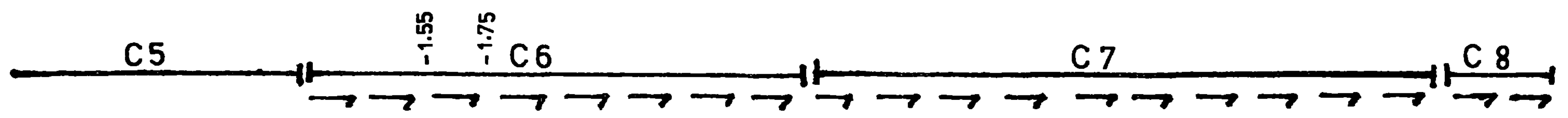
Asp-Ala-Ala-Glu-Glu-Ala-Gly-Ile-Asp-Leu-Pro-Tyr-Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys

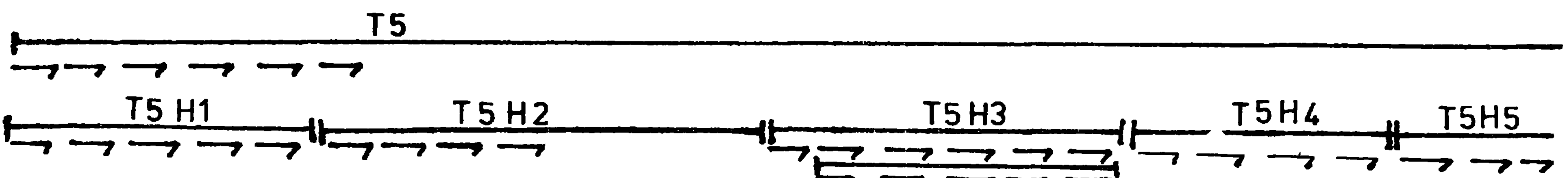




51 55 60 65 70 75

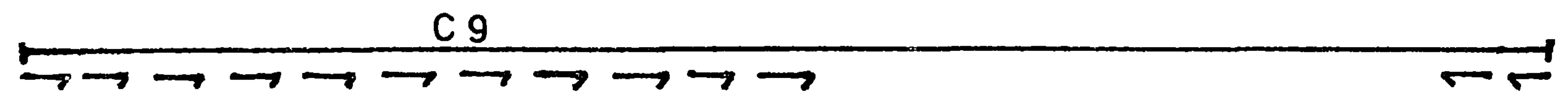
Val-Val-Ser-Gly-Ser-Val-Asp-Gln-Ser-Asp-Glu-Ser-Phe-Leu-Asp-Asp-Asp-Gln-Ile-Ala-Glu-Gly-Phe-Val-Leu





76 80 85 90 95

Thr-Cys-Ala-Ala-Tyr-Pro-Thr-Ser-Asp-Val-Thr-Ile-Glu-Thr-His-Lys-Glu-Glu-Glu-Leu-Val



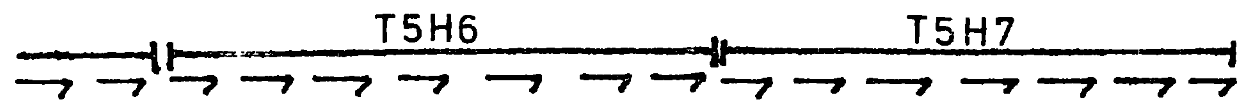


TABLE 17.

Chymotryptic peptides of Brassica napus

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
C1 (1-3)	0	0.60	Ala-Thr-Tyr
C2 (4-7)	2.50	1.38	Lys-Val-Lys-Phe
C3 (8-23)	-2.13	0.19	Ile-Thr-Pro-Glx-Gly-Glx-Glx-Glx-Val-Glu-(Cys-Asx-Asx-Asx-Val)-Tyr
C4 (24-37)	-2.22	0.19	Val-Leu-Asx-Ala-Ala-Glx-Glx-Ala-Gly-Ile-Asx-Leu-Pro-Tyr
C5 (38-55)	—	0.35	Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-(Cys-Ala-Gly-Lys-Val-Val-Ser-Gly-Ser)
C6 (56-63)	-2.04	—	Val-Asx-Glx-Ser-Asx-Glx-Ser-Phe
C7 (64-73)	-2.18	0.25	Leu-Asx-Asx-Asx-Glx-Ile-Ala-Glx-Gly-Phe
C8 (74-75)	0	0.79	Val-Leu
C9 (76-96)	-0.82	—	Thr-Cys-Ala-Ala-Tyr-Pro-Thr-Ser-Asx-Val-Thr-(Ile-Glx-Thr-His-Lys-Glx-Glx-Glx)-Leu-Val

TABLE 18.

Tryptic peptides of *Brassica napus*

Peptide/ position		Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
T1	(1-4)	1.15	0.88	Ala-Thr-Tyr-Lys
T2	(5-6)	1.92	1.34	Val-Lys
T3	(7-40)	----	----	Phe-Ile-Thr-Pro-Glx-(-------)-Arg This peptide was digested with thermolysin. (see Table 19)
T4	(41-50)	-0.80	0.51	Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys
T5	(51-91)	----	----	Val-Val-Ser-Gly-Ser-Val- (------------)-Lys This peptide was digested with thermolysin (see Table 19)
T6	(92-96)	-2.05	0.29	Glx-Glx-Glx-Leu-Val

TABLE 19.

Thermolytic peptides of T3 and T5

Peptide/ position	Mobility (pH 6.5)	(pH 1.9)	Dansyl-Edman results
T3H1 (7-15)	-2.11	—	Phe-Ile-Thr-Pro-(Glx-Gly-Glx-Glx-Glx)
T3H2 (16-17)	-1.50	0.58	Val-Glx
T3H3 (18-24)	—	—	Was not isolated.
T3H4 (25-32)	-1.87	—	Leu-Asx-Ala-Ala-Glx-Glx-Ala-Gly
T3H5 (33-40)	-0.45	—	Ile-Asx-Leu-Pro-Tyr- (Ser-Lys-Arg)
T5H1 (51-55)	0	0.40	Val-Val-Ser-Gly-Ser
T5H2 (56-62)	-1.93	0.28	Val-Asx-Glx-Ser-(Asx-Glx-Ser)
T5H3 (63-68)	-2.25	0.35	Phe-Leu-Asx-Asx-Asx-Glx
T5H3a (64-68)	-2.26	0.37	Leu-Asx-Asx-Asx-Glx
T5H4 (69-72)	-1.18	0.24	Ile-Ala-Glx-Gly
T5H5 (73-77)	-0.75	0.37	Phe-Val-Leu-Thr-Cys
T5H6 (78-84)	-0.59	0.30	Ala-Ala-Tyr-Pro-Thr-Ser-Asx
T5H7 (85-91)	0.73	0.60	Val-Thr-Ile-Glx-Thr-His-Lys

TABLE 20

The amino acid composition of *Brassica napus* ferredoxin

	Mean values 24h hydrolysis	Mean values 48h hydrolysis	Mean values 72h hydrolysis	Average values	Sequence values
Asx	11.42	10.59	11.43	11.15	11
Thr	5.64	5.61	6.44	5.75*	6
Ser	7.89	7.55	6.87	8.70*	9
Glx	13.12	15.78	15.03	14.64	16
Pro	3.58	3.21	3.11	3.30	3
Gly	6.52	6.22	6.26	6.33	6
Ala	10.00	9.00	9.5	9.5	9
Val	9.66	9.78	9.85	9.85+	10
Ile	3.23	3.50	4.02	4.02+	4
Leu	4.98	4.83	4.21	4.67	5
Tyr	3.64	3.66	3.22	3.51	4
Phe	2.98	3.09	3.69	3.25	3
His	1.20	1.22	1.23	1.22	1
Lys	3.39	3.70	3.69	3.59	4
Arg	1.78	1.68	2.00	1.82	1
Cm-Cys	--	--	--		4

*Calculated from 24h and 72h values assuming first order kinetics for destructions (Moore and Stein, 1963).

+For Valine and isoleucine maximal values (72h hydrolysis) were taken.

Cm-Cys was not determined after hydrolysis

TABLE 21

Amino acid composition of peptides from *Brassica napus* ferredoxin

Peptides	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	His	Lys	Arg	Cm-Cys
C1		0.81					1.20				0.99					
C2								0.92				0.77		2.15		
C3	2.90	1.02		5.46	0.71	1.06	0.29	2.24	0.95		0.92					
C4	2.17			2.31	0.83	0.71	2.66	1.00	0.68	1.88	0.93					
C5			5.54			2.75	1.84	2.30						1.00	1.05	2.25
C6	2.19		1.55	2.01				0.87				1.00				
C7	2.75			2.11		0.77	0.85		0.77	0.89		0.89				
C8								0.93		0.89						
C9	0.83	4.08	0.50	4.15	0.65		2.23	1.88	1.11	0.98	0.45		1.28	0.39		1.22
T1		1.01					1.15				1.01			0.79		
T2								0.99						1.05		
T4			2.59			1.77	1.76							0.93		(2)
T6				3.13				0.92		0.78						

Peptide C2 (4-7) (Lys-Val-Lys-Phe)

The electrophoretic mobility at pH 1.9 and 6.5 indicated the presence of two basic residues and the amino acid composition suggested the presence of two lysine residues. Dansylation without hydrolysis after three steps of Edman degradation yielded dansyl-phenylalanine.

Peptide C3 (8-23) (Ile-Thr-Pro-Glu-Gly-Glu-Gln-Glu-Val-Glu-Cys-Asp-Asp-Asp-Val-Tyr)

Digestion with carboxypeptidase A for 1 h yielded bis-dansyl-tyrosine after dansyl analysis. The dansyl Edman analysis was inconclusive beyond nine degradations. Residues from 17-23 were placed from the peptide amino acid composition together with the sequence evidence of the protein using the Automatic Sequencer.

Peptide C4 (24-37) (Val-Leu-Asx-Ala-Ala-Glx-Glx-Ala-Gly-Ile-Asx-Leu-Pro-Tyr)

Digestion with carboxypeptidase A for 1 h yielded a trace of bis-dansyl-tyrosine after dansyl analysis. Tyrosine was confirmed as the C-terminal residue by dansylation without hydrolysis after thirteen steps of Edman degradation. The electrophoretic mobility at pH 6.5 indicated that residues 26 and 34 were aspartic acids and residues 30 and 31 were glutamic acids.

Peptide C5 (38-55) (Ser-Cys-Arg-Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys-Val-Val-Ser-Gly-Ser)

The dansyl Edman analysis was inconclusive beyond eight degradation steps and residues 47-55 were placed from the peptide amino acid composition together with the sequence evidence for peptides T4 and T5.

Peptide C6 (56-63) (Val-Asp-Gln-Ser-Asp-Glu-Ser-Phe)

A consideration of the pH 6.5 electrophoretic mobility indicated the presence of one amide residue and this was placed at position 58 from the pH 6.5 mobility after two and three Edman degradation steps. Dansylation without hydrolysis after seven steps of Edman degradation yielded dansyl-phenylalanine.

Peptide C7 (64-73) (Leu-Asp-Asp-Asp-Gln-Ile-Ala-Glu-Gly-Phe)

Digestion with carboxypeptidase A for 1 h followed by dansyl analysis yielded dansyl-phenylalanine and dansyl glycine and the C-terminal was confirmed as phenylalanine by dansylation without hydrolysis after nine steps of Edman degradation. The electrophoretic mobility at pH 6.5 suggested the presence of one amide and it was placed at position 68 from the sequence evidence for peptide T5H3, and residues 65, 66 and 67 were identified as aspartic acids and residue 71 as glutamic acid from the electrophoretic mobilities of peptides T5H3 and T5H4.

Peptide C8 (74-75) (Val-Leu)

Dansylation without hydrolysis after one step of Edman degradation yielded dansyl-leucine.

Peptide C9 (76-96) (Thr-Cys-Ala-Ala-Tyr-Pro-Thr-Ser-Asp-Val-Thr-Ile-Glu-Thr-His-Lys-Glu-Glu-Glu-Leu-Val)

Digestion with carboxypeptidase A followed by dansyl analysis yielded dansyl-valine after 30 min. and dansyl valine together with dansyl leucine after 1 h. The dansyl Edman analysis become inconclusive beyond ten degradation

steps. Residues from 86 - 94 were placed from the peptide amino acid composition together with the sequence evidence for peptides T5H7 and T6.

Tryptic peptides

Six peptides were obtained when the protein was digested with 2% (w/w) enzyme for 2 h. The resulting peptides were separated by gel chromatography using Bio-gel P-4 in 70% formic acid. Further purifications were achieved by paper electrophoresis at pH 1.9 and 6.5.

Peptide T1 (1-4) (Ala-Thr-Tyr-Lys)

Dansylation without hydrolysis after three Edman degradation steps yielded bis-dansyl-lysine.

Peptide T2 (5-6) (Val-Lys)

Dansylation without hydrolysis after one step of Edman degradation yielded bis-dansyl-lysine.

Peptide T3 (7-40) (Phe-Ile-Thr-Pro-Glx- - - - - - -Arg)

This peptide was digested with thermolysin for 2 h. The first five residues were identified after direct dansyl Edman analysis of the whole peptide. Digestion with carboxypeptidase B for 1 h followed by dansylation yielded dansyl-arginine.

Peptide T3H1 (7-15) (Phe-Ile-Thr-Pro-Glx-Gly-Glx-Glx-Glx)

Dansyl-Edman analysis became inconclusive beyond three steps of Edman degradation. Residues 11-15 were placed from the semiquantitative amino acid analysis of the peptide together with the evidence sequence of peptide C3.

Peptide T3H2 (16-17) (Val-Glx)

The electrophoretic mobility at pH 6.5 indicated that residue 17 was glutamic acid.

Peptide T3H3 (18-24)

This peptide was not isolated.

Peptide T3H4 (25-32) (Leu-Asx-Ala-Ala-Glx-Glx-Ala-Gly)

The electrophoretic mobility at pH 6.5 indicated that residues 26, 29 and 30 were acidic.

Peptide T3H5 (33-40) (Ile-Asx-Leu-Pro-Tyr-Ser-Cys-Arg)

Dansyl-Edman analysis became inconclusive beyond five steps of Edman degradation. Residues 37-40 were placed from the semiquantitative amino acid analysis of the peptide and the evidence sequence of peptides C4 and C5.

Peptide T4 (41-50) (Ala-Gly-Ser-Cys-Ser-Ser-Cys-Ala-Gly-Lys)

Digestion with carboxypeptidase B for 1 h followed by dansyl analysis yielded bis-dansyl-lysine.

Peptide T5 (51-91) (Val-Val-Ser-Gly-Ser-Val- - - -Lys)

This peptide was purified by the gel chromatography. The electrophoretic mobilities at pH 6.5 and 1.9 could not be determined due to smearing.Digestion with carboxypeptidase B followed by dansylation yielded bis-dansyl-lysine. The peptide was digested with thermolysin for 2 h and the resulting peptides were isolated by paper electrophoresis at pH 6.5 and 1.9.

Peptide T5H1 (51-55) (Val-Val-Ser-Gly-Ser)

Peptide T5H2 (56-62) (Val-Asp-Gln-Ser-Asp-Glu-Ser)

Dansyl-Edman analysis became inconclusive beyond three steps of Edman degradation. Residues from 60-62 were placed from the semiquantitative amino acid analysis of the peptide together with the evidence sequence of peptide C6. The electrophoretic mobility at pH 6.5 indicated the presence of one amide which was placed at position 58 from the mobility index of peptide C6.

Peptide T5H3 (63-68) (Phe-Leu-Asp-Asp-Asp-Gln)

The electrophoretic mobility indicated the presence of one amide residue and dansylation without hydrolysis after five steps of Edman degradation yielded dansyl glutamine, therefore, residue 65, 66 and 67 were aspartic acids.

Peptide T5H3a (64-68) (Leu-Asp-Asp-Asp-Gln)

Dansylation without hydrolysis after four degradations yielded dansyl-glutamine.

Peptide T5H4 (69-72) (Ile-Ala-Glu-Gly)

Dansylation without hydrolysis after three degradation steps yielded dansyl-glycine. The electrophoretic mobility at pH 6.5 indicated that residue 71 was glutamic acid.

Peptide T5H5 (73-77) (Phe-Val-Leu-Thr-Cys)

Peptide T5H5a (75-77) (Leu-Thr-Cys)

Semiquantitative amino acid analysis showed that this peptide consisted of three amino acids and the sequence was constructed from direct dansyl-Edman analysis.

Peptide T5H6 (78-84) (Ala-Ala-Tyr-Pro-Thr-Ser-Asp)

The electrophoretic mobility at pH 6.5 indicated that residue 84 was aspartic acid and this was confirmed by dansylation without hydrolysis after six steps of Edman degradation.

Peptide T5H7 (85-91) (Val-Thr-Ile-Glu-Thr-His-Lys)

Electrophoretic mobility at pH 6.5 indicated that residue 88 was glutamic acid. Dansylation without hydrolysis after six Edman degradation steps yielded bis dansyl-lysine.

Peptide T6 (92-96) (Glu-Glu-Glu-Leu-Val)

Electrophoretic mobility at pH 6.5 indicated that residues 92, 93 and 94 were acidic.

The overlapping chymotryptic and tryptic peptides give the sequence of *Brassica napus* ferredoxin as in Fig.15 . This sequence is in accord with the determined amino acid composition of the protein except for arginine and glutamic acid. Every residue was positively identified during the sequence analysis except for residues 18-23 which were identified by data obtained from the Beckman Automatic Sequencer and was confirmed from the peptide C3 composition. Amides and acid residues included in the first 40 residues were determined from their PTH derivatives. The rest of the amides and acids indicated in the sequence were placed from the pH 6.5 electrophoretic mobilities of intact or degraded peptides.

VI) Partial amino acid sequence of Barley ferredoxin

```
 1              5                   10
Ala-Thr-Tyr-Lys-Val-Lys-Leu-Val-Thr-Pro-
 11             15                  20
Glu-Gly-Glu-Val-Glu-Leu-Glu-Val-Pro-Asp-
 21             25                  30
Asp-Val-Tyr-Ile-Leu-Asp-Gln-Ala-Glu-Glu-
 31             35                  40
Glu-Gly-Ile-Asp-Leu-Pro-Tyr-Ser-Cys-Arg-
```

The N-terminal sequence (1-40 residues) were determined using a Beckman 890 C Sequencer using 700 n mol. of carboxymethylated protein. The acidic and amide residues were identified as their PTH-derivatives by thin layer chromatography (TLC) and by gas chromatography (GLC). The lysine residues in positions 4 and 6 were identified as their PTH-derivatives on thin layer chromatography and by dansyl analysis of the regenerated parent amino acid. Arginine was identified in position 40 after regeneration of the aqueous phase followed by dansyl analysis. The remaining residues were identified as their PTH-derivatives on TLC and GLC. The notable feature of the N-terminal sequence compared to that of Wheat is that it is identical.

The amino acid composition of the protein and its CNBr fragments are given in Table 22. Tryptophan was not determined as to be expected after acid hydrolysis, but the protein gave a positive reaction with Ehrlich reagent.

TABLE 22

The Amino acid composition of ferredoxin of *Hordeum vulgare* and its CNBr fragments

The results are expressed as residual molecule. Average values are given except for serine and threonine which are extrapolated to zero time where possible assuming first-order rate of destruction (Moore and Stein, 1963) and for valine, isoleucine and leucine where maximal values (72h hydrolysis) are given. Tryptophan was not determined but Ehrlich reagent with the protein gave a positive result. n.d. = not determined; Hse = homoserine.

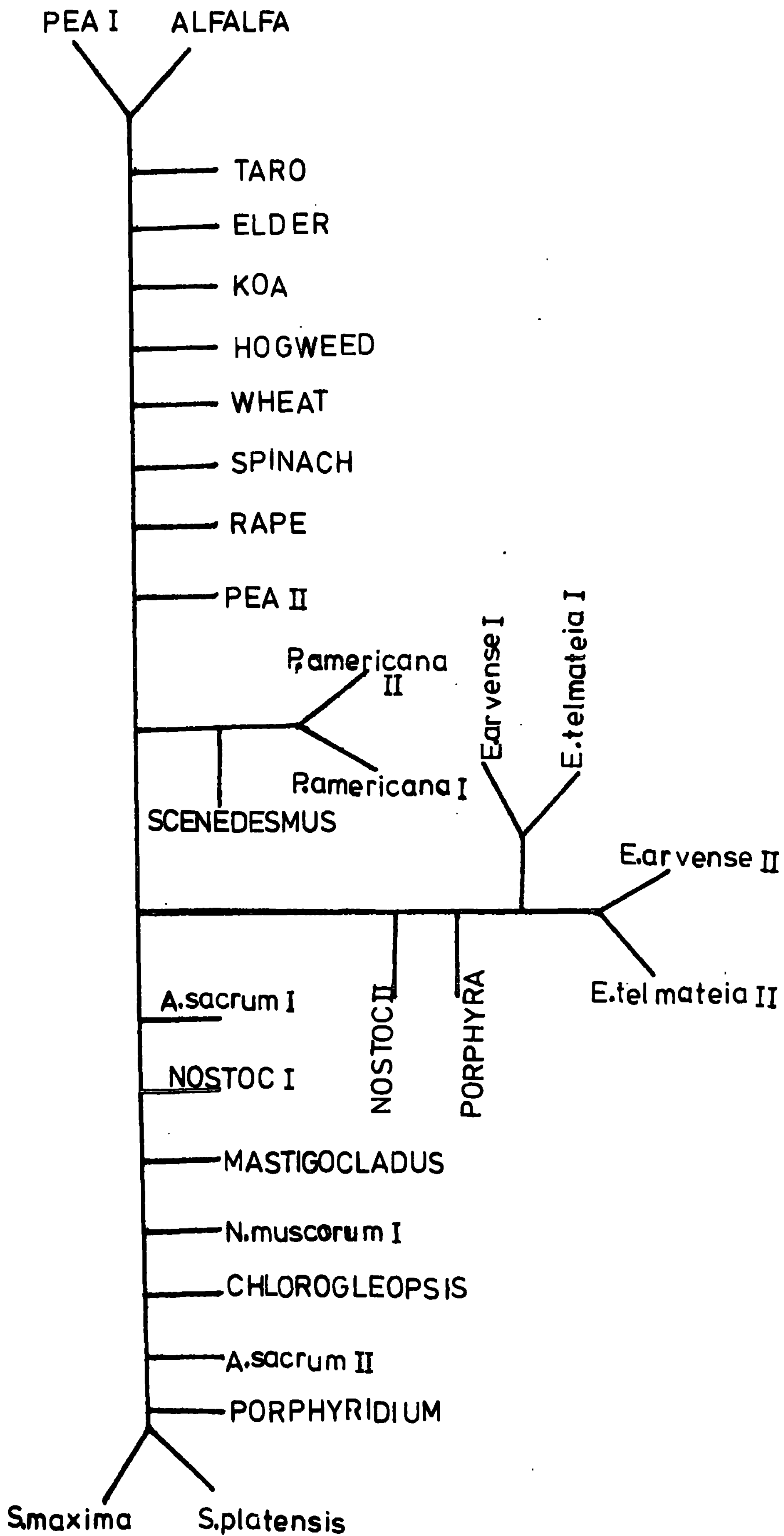
	Total Protein				X1 fragment			X2 fragment		
	24h	48h	72h	Av.	24h	72h	Av.	24h	72h	Av.
Asx	10.85	11.20	11.66	11.24	9.30	9.20	9.25	1.89	2.16	2.02
Thr	4.90	4.64	4.54	5.11 *	2.16	2.14	2.25 *	2.58	2.25	2.63 *
Ser	7.57	6.84	5.96	8.00 *	6.45	5.56	6.83 *	1.36	1.24	1.35 *
Hse	0	0	0	0	0.34	0.38	0.36	0	0	0
Glx	17.51	18.21	18.65	18.12	13.49	13.50	13.50	5.78	5.18	5.48
Pro	n.d.	3.71	3.99	3.85	n.d.	2.35	2.35	0.70	0.67	0.69
Gly	6.09	6.71	6.39	6.39	5.70	5.17	5.43	1.05	1.26	1.15

Ala	6.53	6.75	6.93	6.74	4.33	4.46	4.39	2.83	2.84	2.84
Cys	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0.31	0.40	0.35
Val	6.91	7.19	7.50	7.50	5.77	6.50	6.50	1.86	2.20	2.20
Met	0.40	0.52	0.34	0.42	0	0	0	0	0	0
Ile	4.08	4.14	4.29	4.29	2.92	3.22	3.22	1.50	1.84	1.84
Leu	7.65	7.28	7.80	7.80	6.07	6.29	6.29	2.20	2.29	2.29
Tyr	3.05	2.63	2.44	2.71	1.91	2.51	1.71	0.36	0.31	0.34
Phe	0.80	0.79	0.78	0.79	0.76	0.83	0.79	0	0	0
His	1.63	1.58	1.76	1.66	0	0	0	1.60	1.48	1.54
Lys	4.53	4.56	4.68	4.59	3.30	3.55	3.42	2.00	1.98	1.99
Arg	0.82	0.88	0.83	0.84	0.72	0.83	0.77	0	0	0
Trp	0	0	0	(1)	0	0	0	0	0	(1)

FIGURE 16.

The phylogenetic tree relating sixteen green plants and thirteen algal ferredoxins. The tree was constructed using the ancestral sequence method.

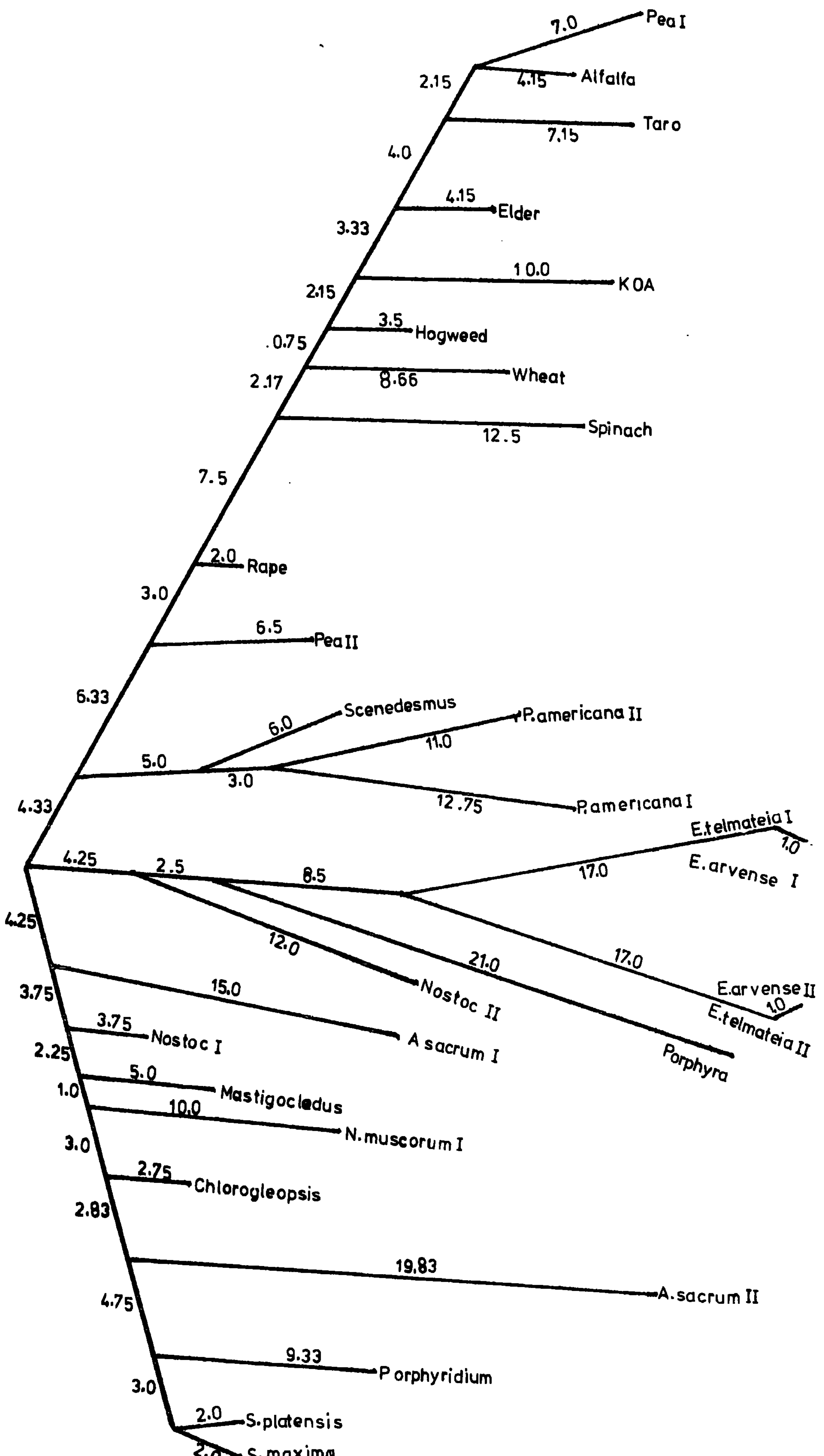
Barley is in the same position as wheat. Rhodymenia has been ommitted since its position does not fit with the algae group.



PEA I
ALFALFA
TARO
ELDER
KOA
HOGWEED
WHEAT
SPINACH
RAPE
PEA II
P.americana II
P.americana I
SCENEDESMUS
E.arvense I
E.telmateia I
E.arvense II
E.telmateia II
NOSTOC II
PORPHYRA
A.sacrum I
NOSTOC I
MASTIGOCLADUS
N.muscorum I
CHLOROGLEOPSIS
A.sacrum II
PORPHYRIDIUM
S.maxima
S.platensis

FIGURE 17.

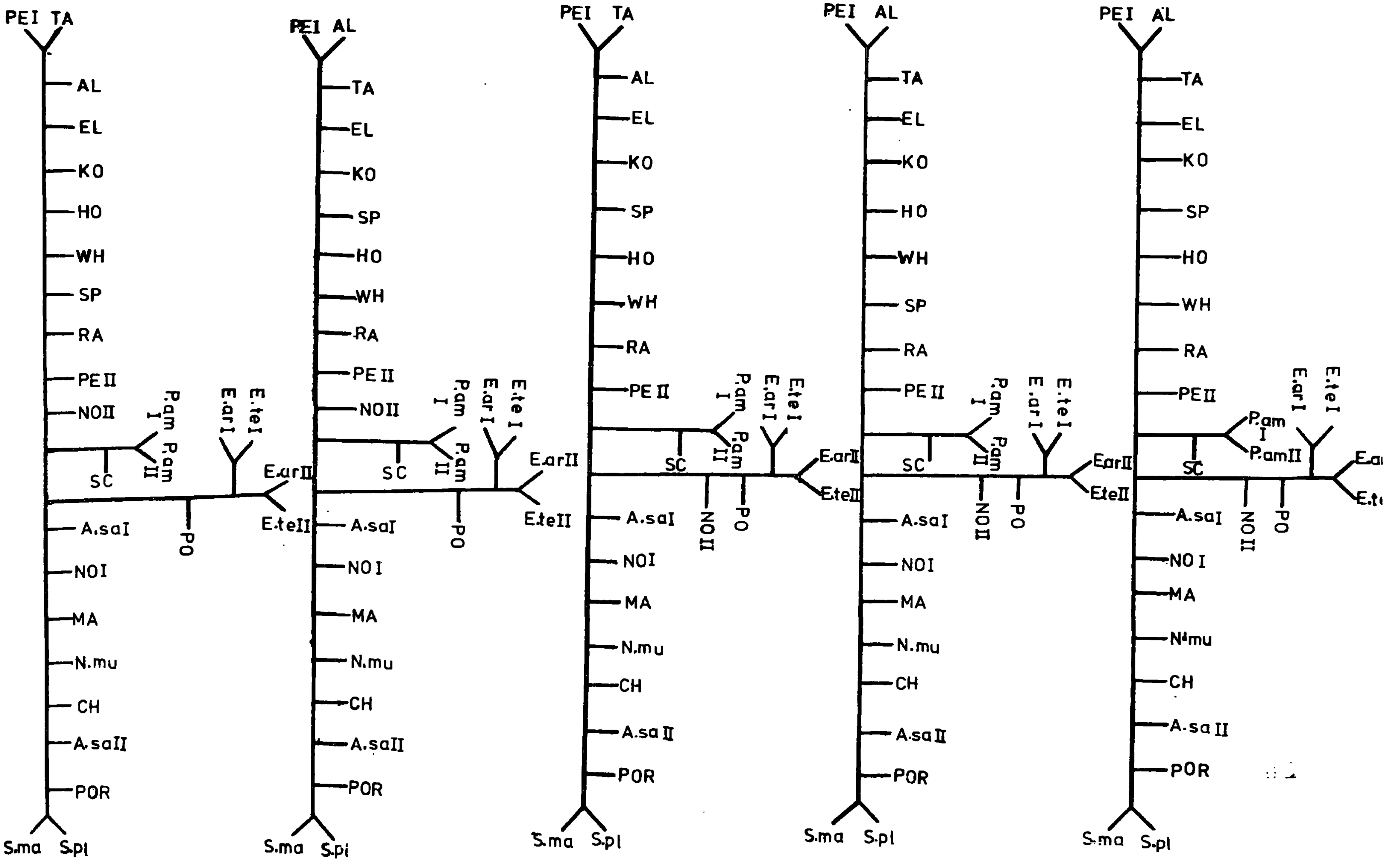
The overall phylogenetic tree relating twentynine ferredoxin sequences. The branch lengths are expressed in "minimum number of amino acid substitutions" ("MNS"). Figures in brackets refer to MNS values calculated.



7.0
Pea I
2.15
4.15
Alfalfa
Taro
7.15
4.0
4.15
Elder
3.33
10.0
KOA
2.15
3.5
Hogweed
0.75
Wheat
8.66
2.17
Spinach
12.5
7.5
2.0
Rape
3.0
6.5
Pea II
6.33
Scenedesmus
6.0
P.americana II
11.0
5.0
3.0
12.75
P.americana I
4.33
E.telmateia I
1.0
E.arvense I
17.0
4.25
2.5
8.5
12.0
Nostoc II
21.0
17.0
E.arvense II
1.0
E.telmateia II
Porphyra
4.25
15.0
A sacrum I
3.75
3.75
Nostoc I
2.25
5.0
Mastigocladus
1.0
10.0
N.muscorum I
3.0
2.75
Chlorogleopsis
2.83
19.83
A.sacrum II
4.75
9.33
Porphyridium
3.0
2.0
S.platensis
S.maxima

FIGURE 18.

The alternative trees obtained to the best tree which have the same number of amino acid substitutions.



PEI TA
AL
EL
KO
HO
WH
SP
RA
PE II
NO II
P.am I
P.am II
SC
E.ar I
E.te I
E.ar II
E.te II
PO
A.sa I
NO I
MA
N.mu
CH
A.sa II
POR
S.ma
S.pl
PEI AL
TA
EL
KO
SP
HO
WH
RA
PE II
NO II
P.am I
P.am II
SC
E.ar I
E.te I
E.ar II
E.te II
PO
A.sa I
NO I
MA
N.mu
CH
A.sa II
POR
S.ma
S.pl
PEI TA
AL
EL
KO
SP
HO
WH
RA
PE II
P.am I
P.am II
SC
E.ar I
E.te I
E.ar II
E.te II
NO II
PO
A.sa I
NO I
MA
N.mu
CH
A.sa II
POR
S.ma
S.pl
PEI AL
TA
EL
KO
HO
WH
SP
RA
PE II
P.am I
P.am II
SC
E.ar I
E.te I
E.ar II
E.te II
NO II
PO
A.sa I
NO I
MA
N.mu
CH
A.sa II
POR
S.ma
S.pl
PEI AL
TA
EL
KO
SP
HO
WH
RA
PE II
P.am I
P.am II
SC
E.ar I
E.te I
NO II
PO
A.sa I
NO I
MA
N.mu
CH
A.sa II
POR
S.ma
S.pl

DISCUSSION

1) Technical problems in the protein sequencing

The method of peptide sequence determination was based on that derived by Gray & Hartly (1967) which involved the proteolytic and chemical digestions of the protein, the isolation of the resulting peptides and analysis of their sequences by the dansyl-Edman method. It is an extremely sensitive technique as compared with some other methods for example, the subtractive Edman method, it has been used in a microform to determine the sequence of as little as 10 p mol. of peptide (Bruton & Hartly, 1970). Using Woods and Wang's method (1967) to separate the dansyl derivatives of amino acids has the advantage of providing an excellent resolution of all naturally occurring protein amino acid derivatives. The main difficulties associated with this method were as follows:

a) The identification of amino acids whose dansyl derivatives were labile during acid hydrolysis of the dansylated peptide. Thus, both asparagine and glutamine were deaminated to their corresponding acids. Also bis-dansyl-histidine was degraded to α-N-dansyl-histidine and lysine can also be difficult to identify after several steps of Edman degradation. Tryptophan and its dansyl derivative were totally destroyed after acid hydrolysis while dansyl-proline was usually degraded on prolonged acid hydrolysis. When proline was expected in

the sequence, therefore the labelled peptide was hydrolysed for a short time (5 h).

The electrophoretic mobility of the peptide at pH 6.5 was often sufficient to assign most of the amide residues (Offord, 1966). If a peptide contained a mixture of acid and amide residues, or had a charge more than $\pm$ 4, ambiguity may still exist. This problem can be resolved either by following the changes in the electrophoretic mobility of the peptide resulting from a series of Edman degradation steps (see T5 of Porphyra sequence) or alternatively amides may be tentatively assigned by homology with other ferredoxins in which the amide residues have been exprimentally determined (see T6 of Elder sequence).

b) Various undesirable side reactions tend to block the amino group of peptides and limit the number of dansyl-Edman cycles that can be performed. Oxygen which reacts with the phenylthiocarbonyl group and acts as blockage to the degradation (Ilse & Edman, 1963) must be excluded at the PITC and TFA stages by flushing the reactants with nitrogen.

c) Lastly, the manual dansyl-Edman method is a slow procedure and therefore time consuming technique.

Recently, the Automatic Sequencer has proved to be an ideal tool for the determination of the sequences of proteins (Edman & Begg, 1967) especially when used in conjunction with manual methods, or with radioactive Edman reagents (Silver & Hood, 1975; Bridgen, 1976). These methods, in spite of their great sensitivity, require

either expensive instruments or hazardous radioactive materials which make them generally unacceptable. More recently, a micro-sequence analysis of peptides and proteins using a colored Edman reagent 4 N,N-dimethlyaminoazobenzene 4'-isothiocyanate (DABITC) in combination with Phenyliso-thiocyanate has been reported (Chang, et al., 1978). This new manual method could provide a fast, simple and inexpensive technique. One degradation cycle takes about 140 min. and needs no acid hydrolysis. Thus it should provide a generally applicable method for micro-sequence analysis.

The performic acid oxidation method (Hirs,1956) used to measure the cysteine content and alkaline hydrolysis of the ferredoxin to estimate the tryptophan content (Noltmann, et al., 1962) were found to be satisfactory. However, use of Ehrlich reagent provided valuable information for the sequencing of tryptophan containing peptides.

The technical problems involved in the protein sequence analysis and the probable sources of inaccuracy have been discussed in detail by Ambler and Wynn (1973), Tarr (1977), and Niall (1977).

2) Comparison of ferredoxin sequences

Ferredoxin is present in both prokaryotic and eukaryotic organisms, and the enzyme from both groups can effectively substitute for one another *in vitro* (Tagawa & Arnon, 1962). This remarkable functional similarity is accompanied by great similarities in their primary structures.

The amino acid sequence of the various plant and algal ferredoxins including blue green algae, which are often assumed to be the precursor of the other algal and the higher plant ferredoxins, show many similar characteristics which can be considered as being typical of ferredoxin group (Orme-Johnson, 1973). All consist of a single polypeptide chain of 95-98 residues and are generally longer than bacterial ferredoxins (55-81 residues). All those whose amino acid sequences which have been determined have an alanine residue as their amino terminus and the data sequences sets are considerably homologous.

The primary structure of fifteen plant, eight algae and nine bacterial ferredoxins have been determined (see Wakabayashi *et al*., 1978). Among them is only one red algae *Porphyra umbilicalus* the sequence of which has been determined during the present investigation.

It is possible to align the chloroplast type ferredoxin sequences with reference to the four invariant cysteine residues which are involved in the linkage of the iron-sulphur active group. It is common practice to number

the residues in the sequence from the invariant NH_2-terminal alanine residue so that the invariant cysteine residues occupy positions of 41, 46, 49 and 79. However, it is necessary to shift the reading frame of Equisetum and Koa ferredoxins by one residue from the NH_2 terminus of the protein to get good alignment with other chloroplast type ferredoxins. The alignment also demonstrates that the difference in the length of various chloroplast ferredoxins are due to an additional one or two residues at the C-terminus, whilst, differences in the length occasioned by variation at the N-terminal occurs only in ferredoxin II of Phytolacca america (see Wakabayashi et al., 1978) which has one extra alanine residue. Gaps are introduced in positions 3, 11, 15 and 61 to provide the best match of residues among the various chloroplast ferredoxins.

There is little doubt that the sequences of the various plant and algal ferredoxins show homologous sequences, since there is a greater similarity between them than that needed for a common function. Thus, about 35 residues including the four cysteines are invariant. Segments containing residues 64-68, 76-79 and 91-92 are identical in all sequences. The most distinguished of the invariant segment detected is the constancy of ten amino acid residues located in the region between 37-51 (see Appendix 4) which include three cysteine residues separated by four and two residues which may be essential to the proper covalent binding of the iron-sulphur (see Appendix 1).

Chloroplast ferredoxins contain generally 10-17 glutamic acid and 8-13 aspartic acid residues. The acidic residue exceed the basic ones which accounts for the acidic isoelectric point (pI) of the protein. The isoelectric points vary slightly from species to species, for example pI for Wheat ferredoxin was determined as 3.9 and that of Barley as 3.75. Two closely associated bands appeared on gels when carboxymethylated Wheat ferredoxin was electrofocused while only one band was present when the native protein was examined. This presumably is due to incomplete reduction and carboxymethylation.

Some protein amino acids are absent from some sequences. Thus tryptophan is present e.g in Wheat, Koa, Spinach and Taro at position 73 while it has not been detected in the algal ferredoxins except A sacrum II. Methionine is lacking in a number of ferredoxins and does not appear essential for activity. It has been detected at position 71 only in Wheat, Porphyra Scenedesmus and Equisetum, and there are two residues in A. sacrum II.

Iron-sulphur proteins, namely, ferredoxin, rubredoxin and adrenodoxin and perhaps some other protein may have diverged from a common ancestor i.e., are paralogous proteins. Homology has been suggested between ferredoxin bacterial type and rubredoxin since when both sequences are aligned there is considerable similarity (Bruschi & Le Gall, 1978). Even so they were shown to give different topologies when a phylogenetic tree of both sequences from the same set of species was completed. Ycas (1976)

using the ancestral sequences of ferredoxins and rubredoxins suggested that both can be aligned as homologous proteins. Similar proposals have been suggested by Baltscheffsky (1974). He proposed a hypothesis on the evolution of electron transport system and suggested that flavoproteins and cytochrome c evolved from a ferredoxin-like precursor. De Lay and Kersters (1974) and Hall et al., (1971) also suggested that anaerobic bacterial ferredoxins may have been one of the first proteins to appear on the earth. In attempting to compare the similarity of more distantly related iron-sulphur proteins namely, adrenodoxin, Hall et al., (1973) have identified a number of probable homologous segments in the primary structures of adrenodoxin and ferredoxin. They exhibit a fair degree of similarity in their sequences and the active centre of these two proteins is almost identical (Cammack et al., 1971) which confirms the relationship between adrenodoxin (animal protein) and ferredoxin (plant protein). On the other hand, little evidence of homology has been observed between ferredoxins and flavodoxins even though they possess similar or even interchangable functions. The amino acid sequences of flavodoxins and ferredoxins of various organisms (see Kobayashi & Fox, 1978) show little evidence that they have diverged from a common ancestor. From a consideration of their size and nature, it has been suggested that ferredoxins may have originated from a primitive tetra-peptide (Ala-Asp-Ser-Gly) (Eck & Dayhoff, 1966) followed

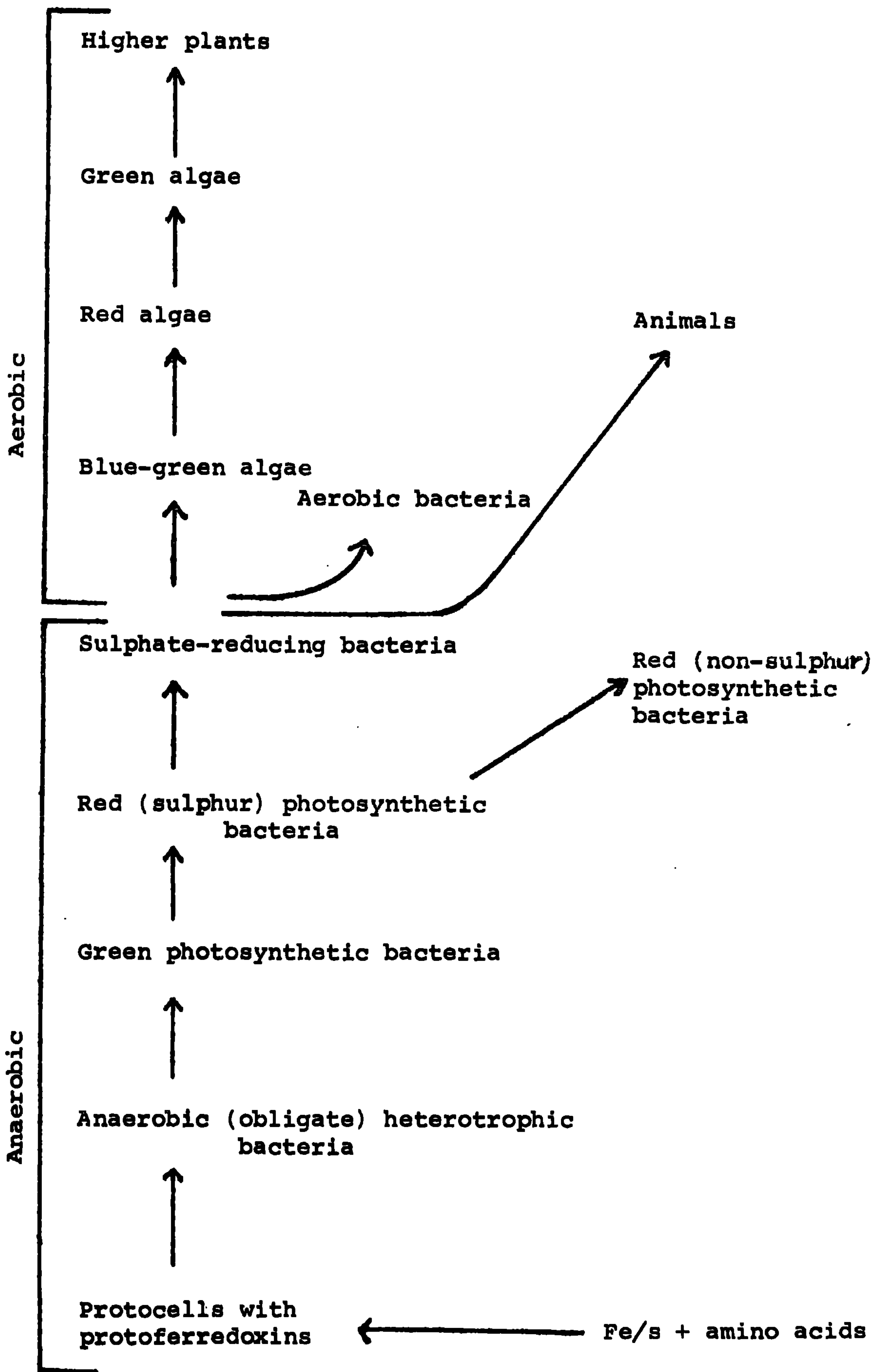
by elongation of the protein by internal gene duplication of the bacterial ferredoxin while a pentapeptide (Lys-Gly-Ala-Asp-Val) repeating structure appears in bacterial flavodoxin. In some cases two ferredoxins were isolated from one organism. The first photosynthetic organism in which two ferredoxins were found was Rhodospirillum rubrum (Shanmugam et al., 1972). Recently, two ferredoxins have been reported in several other plants (Hase et al., 1975; Hase et al., 1977a and 1977b; Wakabayashi et al., 1978). These show very large amino acid differences suggesting that gene duplication had occurred quite early in evolution (Hase et al., 1978 a). There are differences of 37, 29, 31 and 23 between Aphanothece sacrum ferredoxins I and II, Equisetum telmateia ferredoxins I and II, Equisetum arvense ferredoxins I and II and Phytolacca americana ferredoxins I and II respectively (see Appendix 3). Similarly, large differences are to be expected in the total sequences for ferredoxins of Nostoc I and II, Pea I and II ferredoxins as judged by their partial sequences. From their sequence comparison, ferredoxin I of Aphanothece sacrum is orthologous with those of higher plant ferredoxins with gaps at positions 11 and 15 and ferredoxin II is orthologous with those of blue-green algal ferredoxins without those gaps. This can be clearly seen on the phylogenetic tree (see Fig. 16, 17) where Aphanothece sacrum ferredoxin II

is located with the blue-green algae while ferredoxin I of *Aphanothece sacrum* is located near the plant ferredoxin. The ferredoxin gene corresponding to *Aphanothece sacrum* ferredoxin I might be orthologous with *Nostoc* II ferredoxin gene whilst *Nostoc* I corresponds to *Aphanothece sacrum* II. Matsubara *et al* (1978) had suggested that gene duplication occurred before the divergence of blue-green algae and eukaryotes and that gene duplication of ferredoxin genes must have occured often.

3) Evolutionary aspects of the sequence data obtained

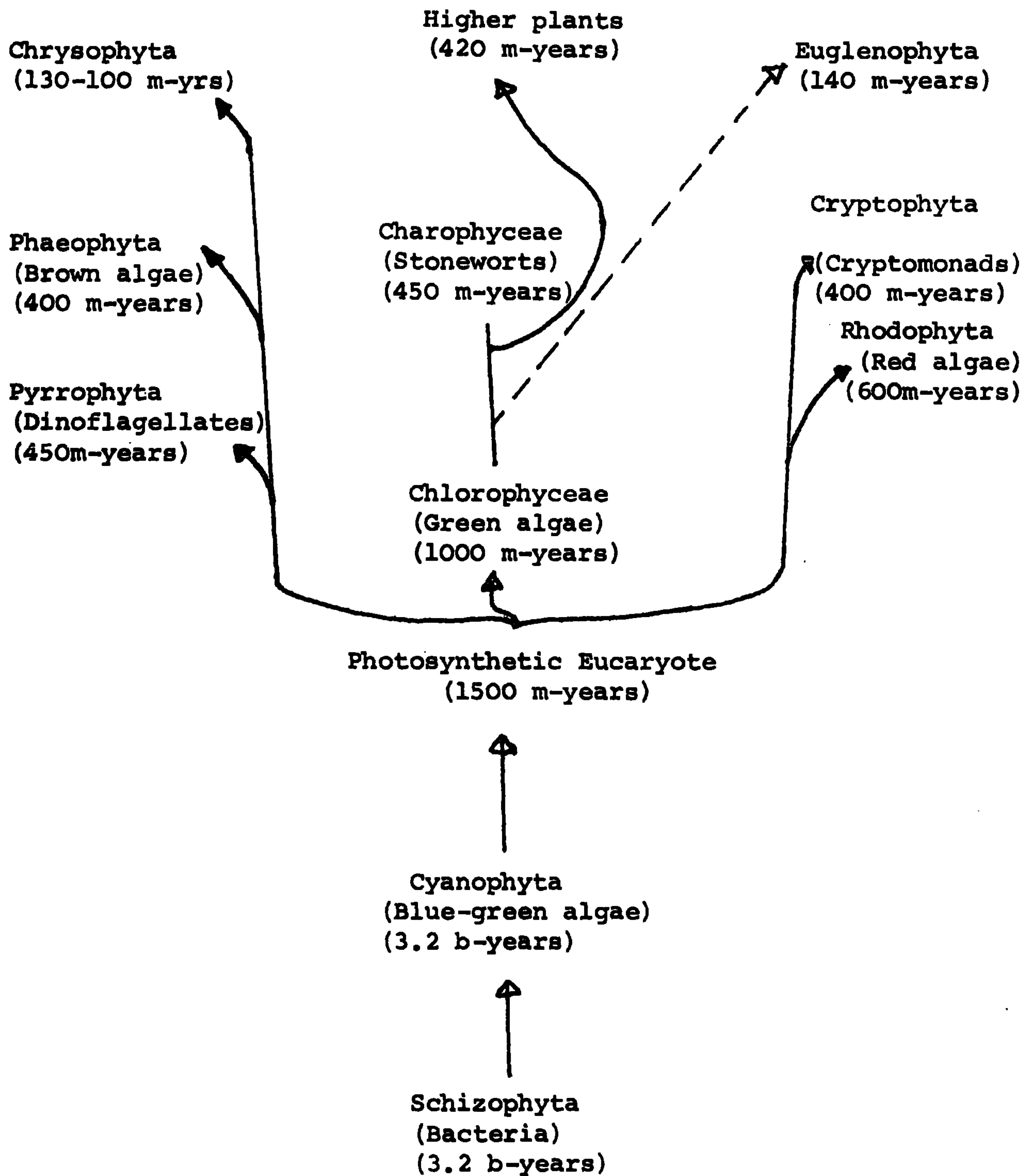
Although it is not intended to consider the bacterial ferredoxins in detail, a brief mention of their evolution is required since these are the primitive ferredoxins from which the others have derived. Thus Hall et al., (1977) have proposed a phylogenetic scheme as shown in (Fig. 19). Several other proposals for the possible evolution within the bacterial ferredoxins have been made. Eck and Dayhoff (1966) postulated that ferredoxin was the first protein to evolve from a tetrapeptide precursor which underwent a polymerization process to give a 28 amino acid residue containing ferredoxin. Fitch (1966) and Yasunobu & Tanaka (1973) have also suggested that present day bacterial ferredoxins have evolved by gene duplication with fusion.

The anaerobic bacterial ferredoxins and bacterial photosynthetic ferredoxins were claimed to have homologous sequences, based on the similarity of residues 1-41 (Matsubara et al., 1968). Furthermore the photosynthetic bacterial ferredoxins show quite strong similarity for residues 9-50 with those of blue-green algae, in the region 41-83 and also with the same region of chloroplast ferredoxin (Yasunobu and Tanaka 1973). They also suggested that a portion of the gene coding for residues 83-97 in the chloroplast ferredoxins has arisen from the portion of the gene coded for residues 1-9 in the bacterial ferredoxin gene, implying that translocations have occured. Gene duplication with fusion leading to a ferredoxin



Evolutionary development of ferredoxins (Hall <u>et</u> <u>al</u>., 1977)

<u>Fig. 19</u>



Possible evolution of lower plants. Figures in brackets give age of oldest fossil record (Swain 1974).
m = million (10^6) b = billion (10^9)

Fig. 20

molecule which is doubled in size is not the only phenomenon of gene duplication since separate ferredoxins called Fd1 and Fd2 occur in some organisms e.g. blue-green algae, Equisetum and some higher plants. Matsubara et al., (1978) using an assumed unit evolutionary period (see later) have placed the gene duplication event in the Equisetum line at 250 million years ago. However, there are many assumptions inherent in this calculation.

The unit evolutionary period is the time required for a single amino acid change to have been fixed in two sequences since they diverged from a common ancester. If the rate of change were constant therefore, it should be possible to calculate from the amino acid differences between two groups their time of divergence from a common ancester. However, there is much argument about whether the clock is steady or not (Wilson et al., 1977).

When the differences between members of any one group and the higher plants are compared, they follow a pattern much as to what is expected from the fossil record. Thus angiosperms evolved about 135 million years ago and show 14-24 differences between themselves. When the green algae and higher plants are compared the values are 25-33 and when the latter are compared with red algae the values rise to 35-42 and for higher plants and blue-green algae 27-46.

The Equisetum (fernallies) value does not fit very well being 32-40 and 39-46 respectively when Equisetum Ferredoxin 1 or II are used. A fact also noted by Matsubara et al (1978) who used corrected substitutions rather than the raw data as used in the above calculations. In view of the uncertainty that the molecular clock is steady (constant rate) plus the very few data available for each group makes it likely that a useful unit evolutionary period cannot be calculated. It is for this reason that it is difficult to decide the authenticity between schemes such as Hall et al., (19 77)(see Fig. 19), and Swain (1974) (Fig. 20). Thus according to the latter unicellular members of the green algae (chlorophyta) are believed to be the most primitive eukaryotes of the plant kingdom even though the red algae (Rhodophyta) have some biochemical features similar to the blue green algae. According to this view the eukaryotic higher plants and algae, except for the Rhodophyta, have evolved along one branch which split off from fungi (Swain, 1971). The Rhodophyta possess some major differences from all other classes of eukaryotic algae and contain for example carbohydrates and phycobilins and auxilary photosynthetic pigments similar to the blue green algae (procaryotes). Therefore, it has been suggested that Rhodophyta may have been derived from some Cyanophytic ancester (see Fig. 19). Other suggestions for the origin of this and other classes

of algae have been discussed by Swain (1974). The fossil record is also ambiguous.

In this investigation, the primary structure of ferredoxins which have been determined from algae and plants have been utilised together with published sequence data excluding bacteria, to study the evolutionary relationship of the organisms from which they came. However, difficulties of interpretation arise here also.

To date, the primary structure of twenty three ferredoxins excluding bacterial ferredoxins have been determined. The determination of the amino acid sequence of *Porphyra* ferredoxin, the only red algal one was part of the present investigation. It has some features similar to the prokaryotic blue green algae and other characteristics like the green algae and plant ferredoxins. Thus its sequence, like most of the blue green algae sequences, contain two extra cysteine residues in addition to the four invariant cysteine residues. These two extra cysteines are located at position 20 and 87 and they are not involved in the active group. *Porphyra* ferredoxin, however, contains one methionine residue at position 73 like some plant ferredoxins such as Wheat, whereas methionine has not usually been detected in any of the blue green algae ferredoxins whose sequences have been determined so far. In general, *Porphyra* ferredoxin differs from those of blue green algae as much as it does from those of the chloroplast ferredoxins (Appendix 3). Furthermore, comparison of the *Porphyra*

sequence with preliminary sequence data for ferredoxins from other red algae Porphyridium and Rhodymenia (Boulter, unpublished work) suggests that big differences exist among Rhodophyta.

When plants and algal ferredoxins were aligned, two distinct representative types of sequences were obtained as a result of the gaps which had to be introduced;one type had gaps at positions 11 and 15 and the other no gaps in these positions. It is interesting that the first group includes all eukaryotic plants except for Porphyra and Porphyridium ferredoxins. The other type is represented by the blue green algae and red algae except Rhodymenia and Aphonatheca sacrum I ferredoxins, these have no gaps at these positions. Therefore, red algae are represented in both groups which suggests that there is a large diversity among Rhodophyta species. This is in agreement with the information obtained from ferredoxin-antiferredoxin serological reactions between algal ferredoxins. Thus Tel-or et al., (1977) showed that the three red algal ferredoxins, Porphyra, Porphyridium and Cyanidium were clearly differentiated from each other and they showed different cross reactions with the green algal anti-ferredoxin.

When establishing species phylogenies, it is important that distinction is made between "paralogous" and "orthologous"

genes (Fitch and Margoliash, 1971). Genes which are originally identical immediately after duplication may diverge subsequently to an extent that they code for different proteins. Such genes are homologous but have been termed paralogous to distinguish them from the other class of homologous genes "orthologous genes" those which remain identical in function throughout. From their location in the tree, Pea II and Nostoc II ferredoxins seems to be paralogous and not orthologous proteins as compared with the rest of ferredoxins.

When constructing a phylogenetic tree from all plants and algal ferredoxin data, the ancestral sequence method of Gleaves (unpublished) was followed. In this method, the ancestral sequences are inferred. The computing strategy employed was based on that described by Dayhoff (1972) and Fitch (1969). The tree consists of branches and nodes (junctions). Each node has three branches which lead either to an adjacent node or to a determined sequence. The ancestral sequence method involves two assumptions: (1) the evolution has occurred by the minimum number of amino acid substitutions, (2) the final tree obtained is that which has the minimum number of amino acid substitutions of all possible trees. Initially, three species are arbitarily chosen and the topologies calculated when a fourth species

is added, all topologies, which have minimum number of substitutions, are used for subsequent construction and in this sense is a more powerful method than the original of Dayhoff (1972). Additional species are added in turn until a complete tree is obtained. Using the ferredoxin data, several alternative best trees have been formed all having the same minimum number of amino acid substitutions. However, it has been found that with this method a single change in the sequence may be sufficient in certain cases to cause a change in topology and the conclusion drawn therefore may be susceptible to sequencing errors. The final best tree constructed is given in Fig. 17 and the lengths of the branches are proportional to the inferred amount of change in the sequences. It was not unique, but the alternatives (see Fig. 18) are minor variations.

The species (sequences) used in the computation of the tree fall into two groups:- higher plants and green algae (group one) and red algae and blue-green algae (group two). The red algae subset comprised only three sequences _Porphyra_, _Rhodymenia_ and _Porphyridium_ and they appear far apart on the tree. This is almost certainly due to 1) lack of sequence data availabe, 2) the incomplete sequence data for _Rhodymenia_ and _Porphyridium_ limiting the computing strategy employed to that of comparing the known

sequence fragments with the homologous sections of other sequences, 3) to their phylogenetic diversity.

In general, the ferredoxin sequences (see Appendix 4) follow the evolutionary tree already accepted on the basis of morphological and fossil data. Unfortunately, the latter is very sparse. Thus, the sequence data derive the red and blue green algae on one branch and the higher plants on another. Equisetum (fern allies) have long been accepted as an evolutionary offshoot difficult to place exactly and this is also suggested by their isolated position away from the green plants sited at the end of the blue green and red algal branch. Phytolacca is a member of Phytolaccaceae which is placed in the Caryophyllales by Takhtajan (1969) and thought to be derived from the Ranunculales and to be quite primitive in the dicotyledons. This idea is supported by the present sequence data which derive it at the earliest point in time for all the dicotyledon species investigated. The next most notable feature of the sequence tree is the fact that some dicotyledon sequences, that is Spinach and Rape are derived earlier than monocotyledons such as Wheat. This again is the accepted classical view that is, that some dicotyledons are more primitive than the monocotyledons. From its position Pea II ferredoxin (as opposed to Pea I ferredoxin) would appear to be paralogous not orthologous with the other ferredoxins studied that is, they are all

ferredoxin I sequences. This is based on the fact that Pea and Alfalfa are members of the same family. In this respect the present data support the suggestions made by Boulter and his colleagues (1976) that the leguminosae are not a primitive family as suggested by Corner (1949), Takhtajan (1969) and Cronquist (1968). Also of interest is the early divergence of Spinach a member of the Caryophyllaceae prior to Wheat a suggestion also made by the cytochrome c data of Boulter (1973).

Although existing schemes for the evolution of higher plants such as those of Takhtajan (1969) and Cronquist (1968) are unsatisfactory, the sequence data set is insufficiently large for the trees here presented to firmly replace them. However, the ferredoxin sequences do support the contention of Boulter *et al.*, (1972) using cytochrome c sequence data that the existing schemes are suspect.

It is quite clear that the serious weakness of the present morphological schemes, of grouping of similarities due to retention of primitive characters, convergent similarity as well as diverged character similarity, cannot be accepted for too much longer as molecular data accumulate, since the latter, when in sufficient amount, will to a large extent overcome those difficulties.

BIBLIOGRAPHY

ADMAN, E.T., SIEKER, L.C. and JENSEN, L.H. (1973). J. Biol. Chem. 248, 3987-3996.

AMBLER, R.P. and WYNN, M. (1973). Biochem. J. 131, 485.

ANDREW, P.W., ROGERS, L.J., BOULTER, D., and HASLETT, B.G. (1976). Eur. J. Biochem. 69, 243-248.

ARNON, D.I. (1977). Photosynthesis 1950-75 : Changing Concepts and Perspectives. In: Encyclopedia of Plant Physiology, New Series, Vol. 5 (Trebst A. and Arnon M. Eds). pp 7-56. Springer-Verlag, Berlin, Heidelberg, New York.

ARNON, D.I., WHATLEY, F.R. and ALLEN, M.B. (1957). Nature 180, 182-185.

BALTSCHEFFSKY, H. (1974). Origins of Life 5, 387.

BECKMAN (1972). Instruction Manual for Model 890C Sequencer. Spinco Div. of Beckman Instrument Inc., Palo Alto, California, U.S.A.

BLOMBACK, B., BLOMBACK, M., EDMAN, P. and HESSEL, B. (1966). Biochim. Biophys. Acta 115, 371.

BOCK, R.M. and LING, N.S. (1954). Analyt. Chem. 26, 1543-1546.

BOULTER, D. (1972). Sci. Prog. Oxford. 60, 217.

BOULTER, D. (1973). The Use of Amino Acid Sequence Data in the Classification of Higher Plants. In: Nobel 25,

Chemistry in Botanical Classification (Bendz, G. and Santesson, J. Eds.) pp 211-216. Academic Press, New York and London.

BOULTER, D. (1976). The Evolution of Plant Proteins with special reference to higher plant cytochrome c. In: Commentaries in Plant Science (Smith, H. Ed) Pergamon Press, Oxford, New York, Toronto, Paris and Frankfurt.

✓ BOULTER, D., HASLETT, B.G., PEACOCK, D., RAMSHAW, J.A.M., and SCAWEN, M.D. (1977). Chemistry, Function and Evolution of Plastocyanin. In: International Review of Biochemistry, Plant Biochemistry II. Vol. 13 (Northcote, D.H. Ed). University Park Press, Baltimore.

✓BOULTER, D., RAMSHAW, J.A.M., THOMPSON, E.W., RICHARDSON, M. and BROWN, R.H. (1972). Proc. R. Soc. Lond. B. 181, 441-455.

✓BOULTER, D., THURMAN, D.A. and DERBYSHIRE, E. (1967). New Phytol. 66, 27.

✓BOULTER, D., THURMAN, D.A. and TURNER, B.L. (1966). Taxon 15, 135.

✓BRIDGEN, J. (1976). Biochemistry 15, 3600-3604.

✓BRIGGS, D. and WALTERS, S.M. (1969). Plant Variation and Evolution, McGraw-Hill, New York.

BRUSCHI, M., and LeGALL, J. (1978). Phylogenetic studies of three families of electron transfer proteins from sulfate reducing bacteria : Cytochromes C_3, ferredoxins, and rubredoxins. In: Evolution of Protein Molecules. (Matsubara, H. and Yamanaka, T. Eds). Japan Scientific Societies Press, Tokyo, Japan.

BRUTON, C.J. and HARTLEY, B.S. (1970). J. Mol. Biol. 52, 165-178.

BUCHANAN, B.B. and ARNON, D.I. (1970). Adv. Enzymol. 33, 119-176.

BUCHANAN, B.B. and ARNON, D.I. (1971). Ferredoxins from photosynthetic bacteria, algae, and higher plants. In: Methods in Enzymol. Vol. XXIII, pp 413-425. (San Pietro A. Ed) New York, Academic Press.

CAMMACK, R., HALL, D.O., RAO, K.K. and JOHNSON, C.E. (1971). Biochem. J. 125, 849.

CHANG, J.Y., BRAUER, D. and WITTMANN-LIEBOLD, B. (1978). FEBS Letts. 93, 205-214.

CHESTER, K.S. (1937). Quart. Rev. Biol. 12, pp 19 and 165.

CORNER, E.J.H. (1949). Ann. Bot. 52, 367-414.

COLLERT, E.H. (1969). The Evolution of the Vertebrates. 2nd Edition, New York Press.

CRONQUIST, A. (1968). The Evolution and Classification of Flowering Plants. London and Edinburgh, Nelson.

DAVENPORT, H.E., HILL, R. and WHATLEY, F.R. (1952). Proc. Roy. Soc., Ser. B. 139, 356-358.

DAVIS, P.H. and HEYWOOD, V.H. (1963). Principles of Angiosperm Taxonomy. Oliver and Boyd, Edinburgh and London.

DAYHOFF, M.O. and ECK, R.V. (1966). Atlas of Protein Sequence and Structure. Vol. 2. National Biomedical Research Foundation, Silver Spring, Maryland, U.S.A.

DAYHOFF, M.O. (1969). Atlas of Protein Protein Sequence and Structure. Vol. 4. National Biomedical Research Foundation, Silver Spring, Maryland, U.S.A.

DAYHOFF, D.O. (1972). Atlas of Protein Sequence and Structure. Vol. 5. National Biomedical Research Foundation, Silver Spring, Maryland, U.S.A.

DAYHOFF, D.O. (1973). Atlas of Protein Sequence and Structure. Vol. 5. Supplement. National Biomedical Research Foundation, Silver Spring, Maryland, U.S.A.

DAYHOFF, D.O. (1976). Atlas of Protein Sequence and Structure. Vol. 5 Supplement 2. National Biomedical Research Foundation, Silver Spring, Maryland, U.S.A.

DeLAY, J. and KERSTERS, K. (1974). In: Comprehensive Biochemistry. Vol. 29B (Florkin, M. and Stotz, E.H. Eds.) pp 1-72.

DIXON, M. (1953). Biochem. J. 54, 457.

EASLEY, C. (1965). Biochim. Biophys. Acta. 107, 386-388

ECK, R.V. and DAYHOFF, D.O. (1966). Science, 152, 363-366

EDMAN, P. (1956). Acta. Chem. Scand. 10, 761

EDMAN, P. and BEGG, G. (1967). Eur. J. Biochem. 1, 80-91

ELMORE, D.T. and TOSELAND, P.A. (1956). J. Chem. Soc. 1956, 188-196.

FAIRBROTHERS, D.E. (1968). In: Modern Methods in Plant Taxonomy (Heywood, V.H. Ed.) Academic Press, London p. 141

FITCH, W.M. (1971). Systematic Zoology 20, 406-416

FITCH, W.M. (1966). J. Mol. Biol. 16, 17-27

GRAY, W.R. (1967). Sequential Degradation Plus Dansylation. In: Methods in Enzymol. (Hirs, C.H.W. Ed.) Vol.11 Academic Press, N.Y. and London. P. 469.

GRAY, W.R. and HARTLEY, B.S. (1963a). Biochem. J. 89, 378-380

GRAY, W.R. and HARTLEY, B.S. (1963b). Biochem. J. 89, 59

HALL, D.O. and RAO, K.K. (1977). Ferredoxin. In: Encyclopedia of Plant Physiology. New Series. Vol. 5. (Trebst, A. and Arnon, M. Eds.) pp 206-216. Springer-Verlag, Berlin, Heidelberg, New York.

HALL, D.O., CAMMACK, R. and RAO, K.K. (1971). Nature, 233, 136.

HALL, D.O., CAMMACK, R. and RAO, K.K. (1973). Pure Appl.

Chem. 34, 553-577.

✓ HALL, D.O., RAO, K.K. and CAMMACK, R. (1975). Sci. Prog. Oxford. 62, 285-317.

✓HALL, D.O., RAO, K.K. and MULLINGER, R.N. (1977). A Role for Fe-S Proteins in the Origin in The Origin and Evolution of Life. In: Protein Structure and Evolution (Fox, J.L., Deyl, Z. and Blazej, A. Eds) p. 233. Marcel Dekker, Inc. N.Y. and Basel.

✓HASE, T., WADA, K., and MATSUBARA, H. (1975). J. Biochem. 78, 605-610.

✓HASE, T., WADA, K., OHMIYA, M. and MATSUBARA, H. (1976a). J. Biochem. 80, 993-999.

✓HASE, T., WADA, K. and MATSUBARA, H. (1976b). J. Biochem. 79, 329-343.

✓HASE, T., WADA, K. and MATSUBARA, H. (1977a). J. Biochem. 82, 277-286.

✓HASE, T., WADA, K., and MATSUBARA, H. (1977b). J. Biochem. 82, 267-276.

✓HASE, T., WAKABAYASHI, S., WADA, K. and MATSUBARA, H. (1978a). J. Biochem. 83, 761-770.

✓HASE, T., WAKABAYASHI, S., MATSUBARA, H., RAO, K.K., HALL, D.O., WIDMER, H., GYSI, J. and ZUBER, H. (1978b). Phytochemistry, 17, 1863-1867.

HASLETT, B.G. and BOULTER, D. (1976). Biochem. J. 153, 33-38.

HEILMAN, J., BAROLLIER, J. and WATZKE, E. (1957). Hoppe-Seyler's Z. Physiol. Chem. 309, 219.

HIRS, G.H.W. (1956). J. Biol. Chem. 219, 611-621.

HO, C.Y.K., PRAGER, E.M., WILSON, A.C., OSUGA, D.T. and FEENEY, R.E. (1976). J. Mol. Evol. 8, 59-78.

HUISMAN, J.G., BERNARDS, A., LIEBREGTS, P., GEBBINK, M.G. and STEGWEE, D. (1977). Planta, 137, 279-286.

ILSE, D. and EDMAN, D.E. (1963). Aust. J. Chem. 16, 411-416.

INGLIS, A.S., NICHOLLS, P.W. and ROXBURGH, C.M. (1971). Aust. J. Biol. Sci. 24, 1247.

KOBAYASHI, K., and FOX, J.L. (1978). Evolution of Flavodoxin: Repetitions Gene Duplication. In: Evolution of Protein Molecules (Matsubara, H. and Yamanaka, T. Eds.) Japan Scientific Socieities Press, Tokyo, Japan.

KOHNE, D.E. (1968). Taxonomic Applications of DNA Hybridization Techniques In: Chemotaxonomy and Serotaxonomy (Hawkes, J.G. Ed.) p. 117. Academic Press, London and N.Y.

LEA., P.J. and NORRIS, R.D. (1972). Photochem. 11, 2897.

LEE., Y.S. and FAIRBROTHERS, D.E. (1978). Taxon 27, 159-185.

LeGALL, J., and DRAGONI, N. (1966). Biochem. Biophys. Res. Commun. 23, 145.

LOVENBERG, W.(Ed.)(1973). Iron Sulfur Proteins. New York, London. Academic Press. Vols. I and II.

MALCOLM, D.B.A. (1978). Nature, 275, 90-91.

MATSON, R.S. and KIMURA, T. (1975). Biochim. Biophys. Acta. 396, 293-300.

MATSUBARA, H. and YAMANAKA, T.(Eds.) (1978). Evolution of Protein Molecules. Japan Scientific Societies Press, Center for Academic Pub. Tokyo ,Japan.

MATSUBARA, H., HASE, T., WAKABAYASHI, S., and WADA, K. (1978). Gene Duplications during Evolution of Chloroplast-type Ferredoxin. In: Evolution of Protein Molecules (Matsubara, H. and Yamanaka, T. Eds.) Japan Scientific Societies Press. Center for Academic Publications, Tokyo, Japan.

MILNE, P.R., and WELLS, J.R.E. (1970). J. Biol. Chem. 245, 1566-1574.

MOORE, S. and STEIN, W.H. (1963). In: Methods in Enzymol. Vol. 6, (Colowick, S.P. and Kaplan, N.O. Eds.) p. 817-831. Academic Press, N.Y.

MORTENSON, L.E., VALENTINE, R.C. and CARNAHAN, J.E. (1962). Biochem. Biophys. Res. Commun. 7, 448-452.

NIALL, H.D. (1977). Modern Micromethods for Sequence Analysis. In: Protein Structure and Evolution. Fox, J.L., Deyl, Z. and Blazej, A. Eds.) pp 5-22, Marcel Dekker, Inc. N.Y. and Basel.

NOLTMANN, E.A., MAHOWALD, T.A. and KIRBY, S.A. (1962). J. Biol. Chem. 237, 1146-1154.

✓OFFORD, R.E. (1966). Nature, 211, 591-593.

ORME-JOHNSON, W.H. (1973). Ann. Rev. Biochem. 42, 159-204.

ORME-JOHNSON, W.H., STOMBAUGH, N.A. and BURRIS, R.H. (1972). Fed. Proc. 31, 448.

PEACOCK, D. and BOULTER, D. (1975). J. Mol. Biol. 95, 513-527.

PETERING, D.H. and PALMER, G. (1970). Arch. of Biochem. and Biophys. 141, 456-464.

RAMSHAW, J.A.M., THOMPSON, E.W. and BOULTER, D. (1970). Biochem. J. 119, 535-539.

RAO, K.K., CAMMACK, R., HALL, D.O. and JOHNSON, C.E (1971). Biochem. J. 122, 257-265.

ROMER, A.S. (1966). Vertebrate Paleontology, 3rd edition. Chicago, The University of Chicago Press.

SAN PIETRO, A. and LANG, H.M. (1958). J. Biol. Chem. 231, 211-229.

SHANMUGAM, K.T., BUCHANAN, B.B. and ARNON, D.I. (1972).

Biochim. Biophys. Acta 256, 477.

SILVER, J. and HOOD, L. (1975). Anal. Biochem. 67, 392-396.

SMYTH, D.G. (1967). Techniques in Enzymic Hydrolysis. In: Methods in Enzymol. Vol. 11 (Colowick, S.P. and Kaplan, N.O. Eds.), p 214. New York, Academic Press.

SOKAL, R.R. and SNEATH, P.H.A. (1963). Principles of Numerical Taxonomy, San Francisco, Freeman and Co.

SOUTHERN, E.M. (1970). Nature, 227, 794.

STEERS, E., CRAVEN, G.R. and ANFINSEN, C.B. (1965). J. Biol. Chem. 240, 2478-2484.

SWAIN, T. (1971). Biochemical Evolution in Plants. In: Biochemical evolution and the origin of life (Schoffeniels, E. Ed.) pp 274-296. North-Holland Co., Amsterdam, London.

SWAIN, T. (1974). Biochemical Evolution of Plants. In: Comprehensive Biochem. Mol. Evolution (Florkin, M. and Stotz, E.H. Eds.) Vol. 29A. Elsevier Scientific Pub. Co. Amsterdam, London, New York.

TAGAWA, K. and ARNON, D.I. (1962). Nature, 195, 537.

TAKHTAJAN, A. (1969). Flowering plants - Origin and Dispersal. Edinburgh, Oliver and Boyd.

TAKRURI, I., HASLETT, G.B., BOULTER, D., ANDREW, P.W. and ROGERS, J.L. (1978). Biochem. J. 173, 459-466.

TANAKA, M., HANIU, M., ZEITLIN, S., YASUNOBU, K., EVANS, M.C.W., RAO, K.K. and HALL, D.O. (1975). Biochem. Biophys. Res. Commun. 64, 399-407.

TANAKA, M., NAKASHIMA, T., BENSON, A.M., MOWER, H.F. and YASUNOBU, K.T. (1966). Biochemistry 5, 1666.

TARR, G.E. (1977). In Methods in Enzymol. Vol XLVII (Colowick S.P. and Kaplon, N.O.Eds. in Chief pp 335-357). Academic Press, New York and London.

TEL-OR, E., CAMMACK, R., RAO, K.K., ROGERS, J.L., STEWART, D.W., and HALL, D.O. (1977). Biochim. Biophys. Acta. 490, 120-131.

THOMPSON, E.W., LAYCOCK, M.W., RAMSHAW, J.A.M. and BOULTER, D. (1970). Biochem. J. 117, 183-192.

TRAVIS, J., NEWMAN, D.J., LeGALL, J. and PECK, H.D. (1971). Biochem. Biophys. Res. Commun. 45, 452.

VAUGHN, J.G. (1968a). Sci. Prog. Oxford. 56, 205.

VAUGHN, J.G. (1968b). In: Chemotaxonomy and Serotaxonomy (Hawkes, J.G. Ed.) P. 93. Systematics Association Special Vol. No. 2.

WADA, K., HASE, T., TOKUNAGA, H. and MATSUBARA, H. (1975). FEBS Letts. 55, 102-104.

WAKABAYASHI, S., HASE, T., WADA, K., MATSUBARA, H., SUZUKI, K. and TAKAICHI, S. (1978). J. Biochem. (Tokyo). 83, 1305-1319.

WILSON, A.C., CARLSON, S.S., WHITE, T.J. (1977). Ann. Rev. Biochem. 46, 573-639.

WOODS, K.R. and WANG, K.T. (1967). Biochim. Biophys. Acta. 133, 369-370.

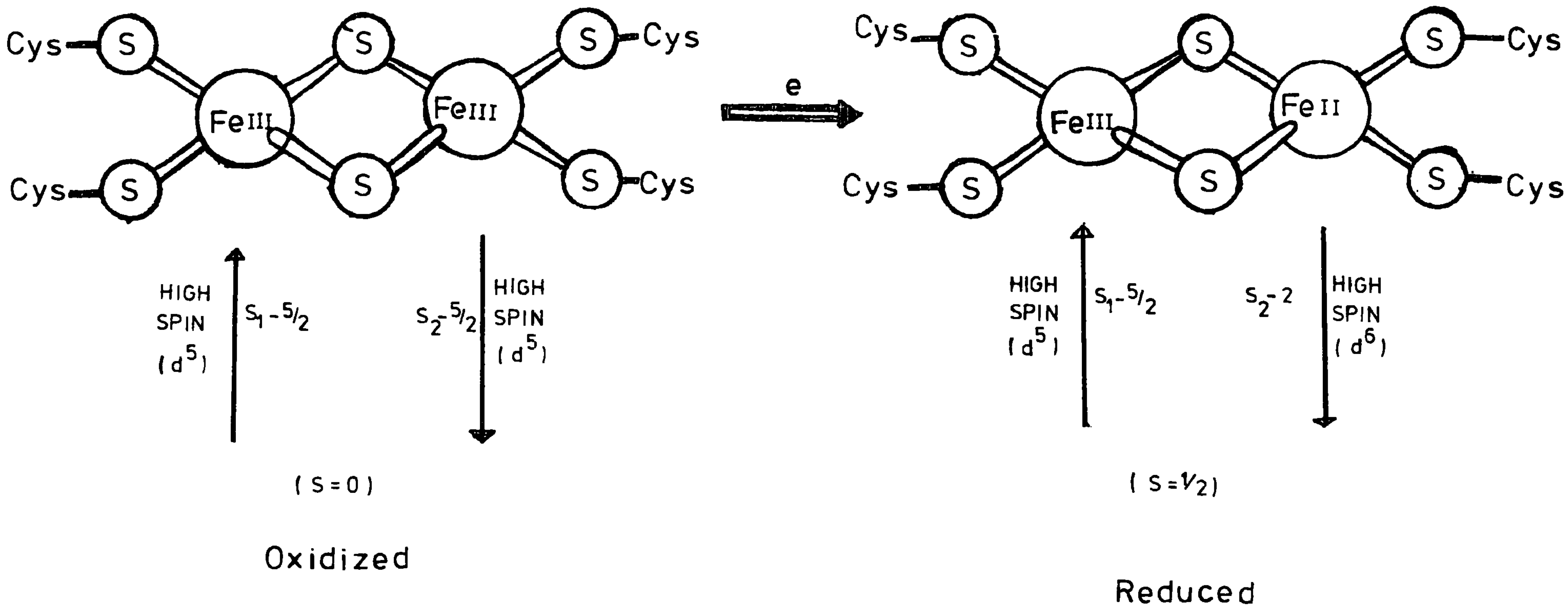
WRIGLEY, C. (1968). Sci. Tools 15, 17-23.

YCAS, M. (1976). J. Mol. Evol. 7, 215-244.

YOCH, D.C., CARITHERS, R.P. and ARNON, D.I. (1977). J. Biol. Chem. 252, 7453-7460.

YASUNOBU, K.T. and TANAKA, M. (1973). The Types, Distribution in Nature, Structure-function and Evolutionary Data of the Iron-Sulfur Proteins. In: Iron-Sulfur Proteins (Lovenberg, W. Ed) p. 27 Academic Press, New York and London.

ZUCKERKANDL , E. and PAULING, L. (1965). J. Theoret. Biol. 8, 357.



Model of active centre of 2 Fe Ferredoxins (Rao, *et al.*, 1971).

Appendix 1

APPENDIX 2

Table of Abbreviations and Symbols for Amino acids

A single letter code for the nomenclature of amino acids as given in the tentative rules of the IUPAC-IUB Commission on Biochemical Nomenclature.

Amino acid	Abbreviation	Code letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

The following abbreviations have been used in Figure 18.

PE I & II	Pea I & II	Pisum sativum I & II
TA	Taro	Colocasia esculenta
AL	Alfalfa	Medicago sativa
EL	Elder	Sambucus nigra
KO	Koa	Leucaena glauca
SP	Spinach	Spinacia oleracea
HO	Hogweed	Heracleum mantegazzianum
WH	Wheat	Triticum aestivum
RA	Rape	Brassica napus
P.am I & II	Phytolacca americana I & II	
E.te I & II	Equisetum telmateia I & II	
E.ar I & II	Equisetum arvense I & II	
SC	Scenedesmus	
PO	Porphyra umbilicalus	
POR	Porphyridium	
NO I & II	Nostoc strain MAC I & II	
A.sa I & II	Aphanothece sacrum I & II	
MA	Mastigocladus laminosus	
S.ma	Spirulina maxima	
S.pl	Spirulina platensis	
CH	Chlorogloeopsis	
N.mu	Nostoc muscorum	

Appendix 3

Matrix of amino acid sequence difference of 20 ferredoxins

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
1.	Taro	0																			
2.	Koa	20	0																		
3.	Spinach	18	20	0																	
4.	Alfalfa	16	23	19	0																
5.	Wheat	20	20	21	19	0															
6.	Elder	17	16	21	13	19	0														
7.	Rape	22	24	24	24	20	21	0													
8.	Phytolacca americana (I)	28	35	30	27	32	29	24	0												
9.	Phytolacca americana (II)	29	37	31	27	34	31	31	23	0											
10.	Equisetum telmateia (I)	36	36	39	39	32	38	34	35	39	0										
11.	Equisetum telmateia (II)	43	39	45	45	42	40	44	43	47	29	0									
12.	Equisetum arvense (I)	37	37	40	40	34	37	35	36	40	1	30	0								
13.	Equisetum arvense (II)	44	40	46	46	44	46	44	44	48	30	1	31	0							
14.	Scendesmus	26	33	30	29	30	28	25	19	25	30	42	31	43	0						
15.	Porphyra	35	42	37	37	37	39	37	38	41	39	43	40	44	31	0					
16.	A. sacrum (I)	38	36	34	37	30	34	34	33	35	42	41	42	42	30	41	0				
17.	A. sacrum (II)	46	47	40	45	36	46	38	42	48	50	56	51	55	40	29	37	0			
18.	S. platensis	37	40	36	38	33	36	29	28	36	42	44	43	45	25	28	32	31	0		
19.	S. maxima	36	39	34	37	33	34	26	28	35	40	44	41	45	25	28	31	31	4	0	
20	N. muscorum (I)	36	32	34	34	27	31	31	32	38	38	40	39	41	28	37	27	34	21	21	0

APPENDIX 4.

Sequence alignment of ferredoxins. The sequences are aligned relative to the invariant cysteine residues. Sequence data as cited in Dayhoff (1972) with the addition of:

Phytolacca americana I and II (Wakabayashi, et al., 1978)

Equisetum telmateia I and II (Hase, et al., 1977a)

Equisetum arvense I and II (Hase, et al., 1977b)

Porphyra (Takruri, et al., 1978)

Spirulina platensis (Wada, et al., 1975)

Spirulina maxima (Tanaka, et al., 1975)

Nostoc muscorum I (Hase, et al., 1976a)

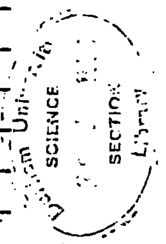
Aphanothece sacrum I (Hase, et al., 1976b)

Aphanothece sacrum II (Hase, et al., 1978a)

Mastigocladus laminosus (Hase, et al., 1978b)

Appendix 4   Ferredoxin Sequences

Species	1 · 10 · 20 · 30 · 40 · 50 · 60 · 70 · 80 · 90 · 100
Rape	AT-YKVKFIT-PEG-EQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSCAGKVVSGSV-DQSDQSFLDDDQIAEGFVLTCAAYPTSDVTIETHKEEELV-
Elder	AS-YKVKLIT-PDG-PQEFECPDDVYILEHAEELGIDIPYSCRAGSCSSCAGKLVAGSV-DQSDQSFLDDEQIEEGWVLTCVAYPKSDVTIETHKEEELTA
Wheat	AT-YKVKLIT-PEG-EVELEVPDDVYILDQAEEEGIDLPYSCRAGSCSSCAGKLVSGEI-DQSDQSFLDDDQMEAGWVLTCHAYPKSDIVIETHKEEELTA
Koa	-A-FKVKLLT-PDG-PKEFEVPDDVYILDQAEELGIDLPYSCRAGSCSSCAGKLVEGDL-DQSDQSFLDDEQIEEGWVLTCAAYPRSDVVIETHKEEELTG
Spinach	AA-YKVTLVT-PTG-NVEFQCPDDVYILDAAEEEGIDLPYSCRAGSCSSCAGKLKTGSL-NQDDQSFLDDDQIDEGWVITCAAYPVSDVTIETHKEEELTA
Alfalfa	AT-YKVKLVT-PEG-TQEFECPDDVYILDHAEEEGIDLPYSCRAGSCSSCAGKVAAGEV-NQSDGSFLDDDQIEEGWVLTCVAYAKSDVTIETHKEEELTA
Taro	AT-YKVKLVT-PSG-QQEFQCPDDVYILDQAEEVGIDLPYSCRAGSCSSCAGKVKVGDV-DQSDGSFLDDEQIGEGWVLTCVAYPVSDGTIETHKEEELTA
Phytolacca americana (II)	AS-YKVTFVT-PSG-TNTITCPADTYVLDAAEESGLDLPYSCRAGACSSCAGKVTAGAV-NQEDGSFLEEEQMEAGWVLTCVAYPTSDVTIETHKEEDLTA
Equisetum telmateia (I)	-A-YKTVLKT-PSG-EFTLDVPEGTTILDAAEEAGYDLPFSCRAGACSSCLGKVVSGSV-DQSEGSFLDDGQMEEGFVLTCIAIPESDLVIETHKEEELF-
Equisetum telmateia (II)	-A-YKVTLKT-PDG-DITFDVEPGERLIDIASEKA-DLPLSCQAGACSTCLGKIVSGTV-DQSEGSFLDDEQIEQGYVLTCIAIPESDVVIETHKEDEL--
Equisetum arvense (I)	-A-YKTVLKT-PSG-EFTLDVPEGTTILDAAEEAGYDLPFSCRAGACSSCLGKVVSGSV-DESEGSFLDDGQMEEGFVLTCIAIPESDLVIETHKEEELF-
Equisetum arvense (II)	-A-YKVTLKT-PDG-DITFDVEPGERLIDIGSEKA-DLPLSCQAGACSTCLGKIVSGTV-DQSEGSFLDDEQIEQGYVLTCIAIPESDVVIETHKEDEL--
Scenedesmus	AT-YKVTLKT-PSG-DQTIECPDDTYILDAAEEAGLDLPYSCRAGACSSCAGKVEAGTV-DQSDQSFLDDSQMDGGFVLTCVAYPTSDCTIATHKEEDLF-
Porphyra	AD-YKIHLVSKEEGIDVTFDCSEDTYILDAAEEEGIELPYSCRAGACSTCAGKVTEGTV-DQSDQSFLDDEQMLKGYVLTCIAYPESDCTILTHVEQELY-
Spirulina platensis	AT-YKVTLINEAEGINETIDCDDDTYILDAAEEAGLDLPYSCRAGACSTCAGTITSGTI-DQSDQSFLDDDQIEAGYVLTCVAYPTSDCTIKTHQEEGLY-
Spirulina maxima	AT-YKVTLISEAEGINETIDCDDDTYILDAAEEAGLDLPYSCRAGACSTCAGKITSGSI-DQSDQSFLDDDQIEAGYVLTCVAYPTSDCTIQTHQEEGLY-
Nostoc muscorum (I)	AT-FKVTLINEAEGTKHEIEVPDDEYILDAAEEEGYDLPFSCRAGACSTCAGKLVSGTV-DQSDQSFLDDDQIEAGYVLTCVAYPTSDVVIQTHKEEDLY-
Aphanothece sacrum (I)	AS-YKVTLKT-PDG-DNVITVPDDEYILDVAEEEGLDLPYSCRAGACSTCAGKLVSGPAPD-EDQSFLDDDQIQAGYILTCVAYPTGDCVIETHKEEALY-
Aphanothece sacrum (II)	AT-YKVTLINEEEGINAILEVADDQTILDAGEEAGLDLPSSCRAGSCSTCAGKLVSGAAPNQDDQAFLDDDQLAAGWVMTCVAYPTGDCTIMTHQESEVL-
Mastigocladus laminosus	AT-YKVTLINEAEGLNKTIEVPDDQYILDAAEEAGIDLPYSCRAGACSTCAGKLISGTV-NQSDQSFLDDDQIEAGYVLICVAYPTSDCVIETHKEEELY-
Phytolacca americana (I)	AT-YKVTLVT-PSG-TQTIDCPDDTYVLDAAEEAGLDLPYSCRAGSCSSCTGKVTAGTV-DQEDQSFLDDDQIEAGFVLTCVAFPKGDVTIETHKEEDIV-
Pea I	AS-YKVKIVT-PDG-TQEFECPSDVYILDHAEEVGLDLPYSCR..
Pea II	AT-YNIKLIT-PEG-TKEITCSDSEYILDAAEEKGLDLPYS..
Hogweed	AT-YKVKLIT-PDG-EVEFDCDDDVYVLDQAEEEGLDLPYS..
Rhodymenia	AVKYTVTLST-PGG-VEEIEGDETSYVLDSAEDQGIDLPYSC...
Porphyridium	AKKYKVRLLSEAEGIDVTIDCADDVYILDAAEEQGIDLPY...
Chlorogleopsis	AT-YKVTLINDAEGLNQTIEVDDDTYILDAAEEAGLDLPYS..
Nostoc MAC I	AT-YKVDLFNAAEGLDETIEVPDDEYILDAAEE..
Nostoc MAC II	ATVYKVTLVD-QEGTETTIDVPDDEYILDIAEDQGLDL...


Durham University
SCIENCE
SECTION
Library