STRUCTURAL STUDIES OF ALGAL PLASTOCYANINS

The primary structure of plastocyanins from

Chlorella pyrenoidosa and *Chlorella fusca*

by

JANICE MARGARET KELLY

A thesis presented for the degree of Master of Science of the University of Edinburgh.

Department of Molecular Biology,
University of Edinburgh. September, 1971.

PREFACE

This thesis consists of work carried out in the Department of Molecular Biology, University of Edinburgh, under the supervision of Dr. R.P. Ambler, between 1969 and 1971. I would like to thank Dr. Ambler for his help and constant encouragement throughout this project.

I would like to thank Professor M.R. Pollock for permission to present this thesis.

I am deeply grateful to Dr. B. Prokes and all the staff of the Department of Applied Algology, Trebon, Czechoslovakia, for supplying me with the dried algal cells, and for the continued interest they have shown in the development of my work. Without their assistance this work might never have materialised.

SUMMARY

Plastocyanin has been purified from three species of green algae: Chlorella pyrenoidosa, Chlorella fusca, and Scenedesmus obliquus.

Two molecular forms of C.pyrenoidosa plastocyanin (A & B) and S.obliquus plastocyanin (A & B), were separated by ion-exchange chromatography on DEAE-cellulose; the former proteins being found in approximately equal amounts, whilst the latter were found in approximately a 5:1 ratio. A single form of plastocyanin was purified from C.fusca. Each protein ran as a single band on starch gel electrophoresis.

The amino acid sequence of C.pyrenoidosa 'A' plastocyanin has been determined and proved to be a single chain of 101 residues. The partial amino acid sequences of C.pyrenoidosa 'B', and C.fusca plastocyanins have been determined, and aligned with the C.pyrenoidosa 'A' plastocyanin sequence.

With the sequences so aligned 64% of the residues are in an identical position. The sequences are discussed in relation to plastocyanin from french bean, and to azurins which are blue copper-containing proteins found in bacteria.

The C-terminal regions of the algal and french bean plastocyanins and the bacterial azurins appear to be at least analogous, if not homologous.The most conserved sequence is that

around the single cysteine residue, which is of the type Tyr-X-(Tyr, Phe)-Cys-. This is of interest since a sulphydryl group has been proposed as one of the ligands involved in the copper-binding site of both plastocyanin and azurin.

Algal plastocyanin sequences

```
    1                          10                            20                            30
1.  a. n. V. K. L. G. A. d. s. G. A. L. V. F. e. P. a. y. y. T. v. t. I. k. a. G. d. s. V. T. y. V. N. N. A. G. F. P. H.
2.  a. t.                b.(b.    )           z.(   g. -  -     n. s).   s. k.    e. s.       f.
3d. v. t.                d. s.                e.    s. -  -(    v. s).   k. a.    e. t.       f.

    40                      50                                  60                            70
1   t. w. N. I. V. F. n. a. D. E. D. a. P. v. v. G. A. n. A. L. S. h. d. D. Y. L. N. A. P. G. e. s. y. t. a. K. F. D. T.
2   -  -              n.(a.          s.    v. v.       q.    i.    r. e.                      z. z. f. t. v.    l.
3   t. w.             q. q.          v.    (a. s.      n.  ) l.    h. e.                      e. s. y. s. a.    f.

    80                   90                                     100
1   G. e. Y. G. Y. F. C. Q. P. H. Q. G. A. G. M. v. G. k. I. v. V. q.
2      e.                                        w.    r.    t.    n.
3      t.                                        k.    t.    t.    q.
```

1. *C.pyrenoidosa* 'A'
2. *C.pyrenoidosa* 'B'
3. *C.fusca*

CONTENTS

PAGE

CHAPTER 1 . INTRODUCTION

1. Examples of comparative amino acid sequence studies

A. General Introduction 1

B. Minimal mutational distance 2

C. Serine proteinases 4

D. Cytochromes C 6

E. Ferredoxins 7

2. A comparative study of azurin and plastocyanin

A. Classes of copper-containing proteins 9

B. Function and isolation of azurin and plastocyanin 10

C. Primary structure studies on azurin 12

D. Physical and chemical properties of azurin and plastocyanin 14

E. The origin of the chloroplast 20

F. Conclusions 21

CHAPTER 2. ABBREVIATIONS AND METHODS 23

1. Abbreviations

A. Chemicals 23

B. Amino acids 24

C. Peptide nomenclature 25

2. Methods

A. Purification of plastocyanin

(i) Detection of plastocyanin in solution 26

(ii) Preparation of acetone powder 26

(iii) Ion-exchange chromatography 26

(iv) Gel filtration 27

(v) Protein determination 28

PAGE

B. Criteria of purity of plastocyanins.
(i) Starch Gel Electrophoresis 29
(ii) Absorption Index 31

C. Experiments on plastocyanin.
(i) Carbohydrate determination 31
(ii) Tryptophan determination 32

D. Determination of amino acid sequence.
(i) Oxidation with performic acid 34
(ii) Preparation of proteolytic enzymes 34
(iii) Cyanogen bromide digestion 36
(iv) High Voltage Paper Electrophoresis 36
(v) Paper Chromatography 39
(vi) Detection and elution of peptides 40
(vii) Peptide mobilities 42
(viii) Acid hydrolysis 43
(ix) Qualitative amino acid analysis 44
(x) Quantitative amino acid analysis 44
(xi) The Dansyl-Chloride Method 46
(xii) Identification of DNS-amino acids 46
(xiii) Carboxypeptidase treatment of peptides 48
(xiv) Sequential degradation (DNS-PTC method) of peptides 49

CHAPTER 3. PURIFICATION OF PLASTOCYANIN 51

A. Strains of green algae 51
B. Growth of algae 51
C. Purification of plastocyanin 52
(i) Extraction of dry algal powders 53
(ii) Chromatography steps: (a) DEAE-11 cellulose columns 53
(b) Gel filtration (on Sephadex G-50) 55
(c) DEAE-52 cellulose columns 55

PAGE

CHAPTER 4. PROPERTIES OF PLASTOCYANIN 59

A. Criteria of purity of proteins

(i) Starch Gel Electrophoresis 59

(ii) Absorption Index 60

B. Absorption spectra 62

C. Carbohydrate content 63

D. Amino acid analysès 64

E. Tryptophan content 65

F. N- and C-terminal studies

(i) N-terminal studies 66

(ii) C-terminal studies 67

CHAPTER 5. DIGESTIONS PERFORMED ON C.PYRENOIDOSA 'A' PLASTOCYANIN 69

A. Tryptic digest

(i) Conditions for digestion 69

(ii) Fractionation of tryptic digest 69

(iii) Summary of tryptic peptides 70

(iv) Peptides formed from tryptic digestion of plastocyanin 72

Peptides TB1, TN1 and TA4 72

Peptides TN3, TB2, TA5, TA1 and TA2 73

Peptides TA3, TN2, TA6 75

B. Chymotryptic digest

(i) Fractionation of chymotryptic digest 76

(ii) Summary of chymotryptic peptides 77

(iii) Peptides formed from chymotrypsin digestion of plastocyanin 77

Peptides CB1 and CB2 77

Peptides CA4 abd CA2 78

PAGE

Peptides CA3 and CA5, 79

Peptides CN1, CA6 and CA8 80

Peptides CB3, CB5, CN2, CN3 and CB4 81

Peptides CA1 and CA7 83

C. Deduction of sequence 84

Discussion of sequence 88

CHAPTER 6. DIGESTIONS PERFORMED ON C.PYRENOIDOSA 'B' PLASTOCYANIN 91

A. Tryptic digest

(i) Fractionation of tryptic digest 91

(ii) Summary of tryptic peptides 92

(iii) Peptides formed from trypsin digestion 92

Peptide T1 92

Peptides T6, T2 and T3 93

Peptide T4 94

Peptide T5 95

B. Chymotryptic digest

(i) Fractionation of chymotryptic digest 96

(ii) Summary of chymotryptic peptides 96

(iii) Peptides formed by chymotrypsin digestion 97

Peptides CB2, CB3 and CB1 97

Peptides CA1 and CA2 98

Peptides CA3 and CA4 99

Peptides CA5, and CA6 100

Peptides CA7, CA8, CN1 and CN2 101

C. Deduction of sequence 103

Discussion of sequence 105

PAGE

CHAPTER 7. DIGESTIONS PERFORMED ON C.FUSCA PLASTOCYANIN 107

A. Chymotryptic digest

(i) Fractionation of chymotryptic digest 107

(ii) Summary of chymotryptic peptides 108

(iii) Peptides formed from chymotryptic digest 108

Peptides CB3, CB1 and CA6 109

Peptides CB2, CN2 and CB4 110

Peptides CA1, CA8 and CA9 111

Peptides CN1, CA2 and CA3 112

Peptides CA4 and CA7 113

Peptide CA5 114

B. Subtilisin digest

(i) Conditions of digestion 115

(ii) Fractionation of subtilisin 115

(iii) Discussion of peptide S5N 116

C. Deduction of sequence 117

Discussion of sequence

CHAPTER 8. PEPTIDES FORMED FROM CHYMOTRYPSIN DIGESTION OF S. OBLIQUUS 'A' PLASTOCYANIN 121

Amino acid sequences of chymotryptic peptides 122

Peptides CB2, CA1 and CN3 122

Peptides CB1, CA7, CA9 and CA2 123

Peptides CA8 and CA10 124

Peptides CA6, CB3, CN3, CA3 and CA4 125

Peptides CA5, CN1, CB4 and CB5 126

Discussion 127

PAGE

CHAPTER 9. DISCUSSION 129

A. Introduction 129

B. Elucidation of the plastocyanin sequences 133

C. Sequence homology between the algal plastocyanins 137

D. Sequence homology between the algal and *P.vulgaris* (french bean) plastocyanin 141

E. Sequence homology between algal plastocyanins and bacterial azurins 142

REFERENCES (i)

CHAPTER I

INTRODUCTION

1. Examples of Comparative amino acid sequence studies

A. General Introduction

The comparison of amino acid sequences as a method for determining the evolutionary relationships between organisms has been widely applied. In general, the proteins which have been used for such studies have been readily available (i.e. easily purified) and of low molecular weight. This enables quick determination of the primary structure of these proteins when one applies the modern techniques of protein chemistry. Two proteins which are functionally analagous and therefore, one hopes, are evolutionarily related, are the bacterial azurins and plastocyanins, which are associated with the chloroplast in plants. Comparison of the amino acid sequences of these two small molecular weight proteins, therefore, may elucidate the evolutionary connection between bacteria and chloroplasts, if any should exist.

Amino acid sequences of proteins are directly related to the mRNA and hence the DNA from which they are translated and transcribed, and as such are more sensitive records of the evolutionary development of an organism than are observations of gross morphological differences. Care must be exercised in using primary structure data for phylogenetic purposes, since Zuckerhandl and Pauling (1962) pointed out that in studying a protein, such as cytochrome c, one is only studying a part of the evolutionary product of the gene. The study does not include

the evolutionary precursors, protein(s) which eventually evolved into cytochrome c, nor the products of genes that may have evolved from cytochrome c.

B. Minimal Mutational Distance

From the determination of many amino acid sequences of structurally and/or functionally similar proteins, and the translation back to the nucleotide sequence of the DNA coding for the proteins, one is able to predict the possible alterations in the DNA sequences involved in the evolution of the proteins. Mutational events likely to occur in the DNA molecule during the evolution of a protein, may include duplication of the whole gene or part of the gene; deletion or insertion of one or more nucleotide pairs; or point mutations.

This approach to molecular evolution has been made more complete by the work of Nirenberg & Leder (1964) and Nishimura & Khorana (1965), in assigning triplet codons to each amino acid. Using the genetic code the minimal mutational distance between two proteins, being the minimum number of nucleotide base changes required to convert the gene for one sequence into that of the other, can be calculated. Prior to this, comparisons between sequences were incomplete, as there was no way of predicting how many mutational events were involved in the substitution of one amino acid for another.

The prediction of an evolutionary connection between proteins based largely on minimal mutational distances must be

treated in a sceptical manner. Cantor & Jukes (1966) have scanned peptide sequences in mammalian and bacterial cytochromes c and in globins, looking for cases where the number of base changes in the DNA required to convert one sequence into that of another is significantly less than that expected for random sequence comparisons. But, the genetic dode is degenerate and so in comparing two amino acid sequences the degeneracy is ignored to give the minimal mutational distance. The minimal mutational distance is therefore the lower limit as regards the actual number of mutations which have occurred in going from one sequence to another. Also, since the genetic code is of a conservative nature, such that the triplet codons are strongly correlated with the nature of the chemical side-chains of the amino acids, then chemical or functional similarity between regions of two otherwise unrelated proteins, will reveal a figure for the minimal mutational distance which would indicate homology between the proteins. However, determinations of minimal mutational distances is a useful method for substantiating homology evidence between proteins from primary or tertiary structural studies, but not as a method of detecting an otherwise invisible homology between proteins.

If one assumes that the primary structure of a protein is a major factor in determining its tertiary structure, it is pertinent to ask to what extent the primary structure needs to be altered in order to change the specificity of a protein; also, which is the

more meaningful guide to an evolutionary homology between proteins, amino acid sequence similarities or tertiary structure similarities, or do both carry equal weight in discerning an evolutionary homology?

C. Serine proteinases

Shotton & Hartley (1970) have studied the sequence homologies between the mammalian serine proteinases, trypsin, chymotrypsin A and B, and elastase. These serine proteinases have different specificities; trypsin is specific for basic residues, chymotrypsin for aromatic side-chains, and elastase for non-aromatic side chains. By aligning the sequences in a way that maximised the homologies between the proteinases, a 35-42% identity between the sequences was shown. Models of the elastase and trypsin molecules were built, based on the assumption that the three-dimensional folding of the two proteins was very similar to that of α-chymotrypsin (Blow, Birktoft & Hartley, 1969). These models suggested that the differences in specificities of the serine proteinases were due to relatively minor side-chain alterations on the enzyme surface. X-ray data on the elastase molecule (Shotton & Watson, 1970) confirms these findings.

Kendrew, Dickerson, Stranberg, Hart, Davies, Phillips and Shore (1960) showed, from the X-ray structure of sperm whale myoglobin, that nearly all the polar groups are found on the surface of the molecule, while nearly all the non-polar groups are found in the interior of the molecule. Shotton et al (1970)

found a similar distribution of amino acid side-chains in the serine proteinases, excepting that the hydrophobic amino acids are fairly evenly distributed between the inside and outside of the molecules. Of interest is that the hydrophobic residues on the surface of the serine proteinases are mostly in non-homologous positions, while, the ones in the interior are in similar, if not identical positions. Thus, it does seem that hydrophobic residues could play an important role in the maintenance of the very similar tertiary structures of elastase, trypsin, and chymotrypsin.

Probably the conclusion to draw from these studies is that a large number of changes in the primary structure of a protein can be accommodated without changing its function, as long as the overall tertiary structure is preserved.

Shotton et al (1970), also suggested that several families of serine proteinases may exist. The mammalian serine proteinases possess some homology, in terms of substrate kinetics and specificity, with serine proteinases from organisms including the spiny Pacific dogfish, the turtle, a sea anemone, a mould and a soil bacterium _Sorangium sp_. They have an aspartyl-seryl-glycyl sequence at the serine residue of the active site, suggesting a chymotrypsin-like specificity. Thus, an evolutionary connection between the two groups of serine proteinases is highly probable. In contrast, a group of bacterial extracellular proteinases, subtilisin BPN' from _Bacillus amyloliquofaciens_ and subtilisin Carlsberg from _Bacillus subtilis_, which have a "threonine-serine-methionine" sequence at

the active serine residue, while showing homology with one another, have no primary or tertiary structure homology with the mammalian serine proteinases. It appears that their analagous enzymic function must be due to convergent evolution.

D. Cytochromes c

Extensive comparative sequence studies have been performed on 'mammalian-type' cytochromes c, including sequences from such divergent eukaryotic organisms as the kangaroo and the yeast, Candida krusei. A striking feature of the amino acid sequences is the maintenance of the basic and acidic residues, clusters of hydrophobic residues, and the cysteine-X-Y-cysteine-histidine haem binding site with a methionine residue as the sixth haem ligand in all the cytochromes c. A sequence homology of over fifty per cent between all the cytochromes c, so far studied, suggested that the tertiary structures might also be similar (Margoliash & Nolan, 1968).

Dickerson (1971) has just completed the X-ray structure of horse and bonito oxidised cytochromes c and proposed that the folding for these two molecules is the same. Thus, it seems that the vertebrate-type cytochromes c constitute an homologous family of proteins which presumably evolved from a common precursoral gene.

Attempts have been made to show sequence homology between the bacterium Pseudomonas aeruginosa P6009 cytochrome c551, and the eukaryotic cytochromes c. They both have the typical haem-binding pentapeptide "cyteine-X-Y-cysteine-histidine". Dus,

Sletton & Kamen (1968) have proposed sequence homologies between the photosynthetic bacterium Rhodospirillum rubrum and eukaryotic cytochromes c. Dickerson (1971), has used this proposed homology to produce a model for the P.aerugionosa cytochrome c551 molecule based on the X-ray analyses of horse and fish cytochromes c. He proposes that most of the changes between the two classes of cytochromes c can be explained by the insertion of a sixteen-residue hair-pin loop into the horse cytochrome c. The proposed homology is a structural hypothesis, and there is still no real concrete evidence that P.aeruginosa P6009 cytochrome c551 and eukaryotic cytochrome c shared a common origin.

E. Ferredoxins

Matsubara, Sasaki & Chain (1967) have studied the sequence homology between bacterial and plant-type ferredoxins in an attempt to establish a common precursoral gene for ferredoxin, as has been proposed for algal, plant and animal cytochromes c (Margoliash, 1963). Ferredoxins are single-chain, non-haem iron-containing proteins present in some non-photosynthetic anaerobic bacteria and in all photosynthetic organisms. They are involved in electron-transfer processes.

From sequence data on five plant ferredoxins; spinach (Matsubara et al, 1967), alfafa (Keresztes-Nagy, Perini and Margoliash, 1969), L.glauca (Benson, Yasunobu, 1969), Scenedesmus obliquus (Sugeno, Matsubara, 1969), Colocasia esculenta (Krishna Rao,

Matsubara, 1970), there is 61% similarity in the primary structures. There is the conservation of residues 35 to 50 in all the five sequences, except for alanine-43 in Scenedesmus ferrodoxin. This segment includes three cysteines and suggests the possibility of the location of the functional groups in this segment. It is likely that the plant-type ferrodoxins evolved from a common ancestral gene.

The ferrodoxins from Chlostridium pasteurianum and butyricum, anaerobic bacteria that usually live in soil without any exposure to light, contain essentially a repeating sequence of a peptide twenty-nine residues long. The best sequence alignment between the bacterial and plant ferrodoxins is over a distance of nineteen consecutive amino acids, between residues 9 to 27 in Chlostridium and residues 42-60 in spinach, giving an average base difference per codon of 0.79 for this region. This might suggest evolution of the two types of ferrodoxins from a common ancestral gene by perhaps a recombinational event involving lengthening of the ancestral gene. But the bacterial and plant ferrodoxins have different absorption spectra, redox potentials, and iron labile sulphur contents, although in vitro experiments have shown that bacterial ferrodoxins can replace the native plant ferrodoxin in the photosynthetic reduction of NADP, by isolated spinach chloroplasts. However, work on the active centre of spinach ferrodoxin (Margoliash, Perini & Keresztes-Nagy, 1967), indicates very strongly

that the electron transfer mechanism and the molecular structure around the cysteine residues differs from those in the Chlostridial ferrodoxins.

Thus, it would seem that both sequence and tertiary structure homology is necessary before an evolutionary connection between two proteins can be proposed with conviction.

Also, the primary structure of a mammalian non-haem iron-containing protein, bovine adrenodoxin, shows no homology with the bacterial and plant ferrodoxins nor with mammalian haem-containing proteins. This may indicate the evolution of the mammalian non-haem iron-containing protein from a separate gene, which has no counterpart or has undergone so many mutations that the homology is no longer evident in the bacteria and plants. But, adrenodoxin contains a 'cysteine-X-Y-cysteine' segment which is also present in many of the non-haem iron-containing proteins sequenced so far. This might suggest that several families of non-haem iron-containing proteins exist which have different ancestral origins, and which may or may not show primary and tertiary structure homologies.

2. A comparative study of azurin and plastocyanin

A. Classes of copper-containing proteins

Another class of metallo-proteins that have been widely studied, is the copper-containing proteins. They fall into three main groups with respect to their biological function. They may participate in electron-transfer reactions, in the transport of oxygen, and in the transport or storage of the metal itself.

A list of the main classes of copper-containing proteins is shown in Table 1 (i). The storage and transport proteins play an important metabolic role in mammals, since their impairment results in a syndrome known as Wilson's disease, where excess copper ions (Cu^{2+}) are free to compete with metal-ions (e.g. Mg^{2+}) in other metallo-proteins, causing inhibition of these enzymes.

In all the above proteins, the copper is mainly present as the cupric (Cu^{2+}) ion. This is the most stable valency state for the copper in biological systems. Due to a gradual increase in ionization potential with ionic charge, copper, along with other transitional metals, exhibits a variable valency. Thus, Cu^{2+}-Cu^{1+} is a common oxidation-reduction system, an important feature in the biological function of copper.

Electron paramagnetic resonance (EPR) studies on copper complexes have classified three different forms of copper in proteins; an EPR-detectable cupric ion responsible for the intense blue colour of some copper-containing proteins, an EPR-detectable cupric ion that is termed "non-blue" in view of its low extinction coefficient relative to the "blue" cupric ion, and copper in an EPR non-detectable form.

B. Function and isolation of azurin and plastocyanin

Azurin and plastocyanin fall into the class of copper-containing proteins that function as electron-transfer proteins, and contain copper only in the "blue" cupric ion form.

TABLE 1(i): Main classes of copper-containing proteins.

PROTEIN	FUNCTION
Hemocyanin	Oxygen carriers.
Erythrocuprein Cerebrocuprein Hemocuprein Hepatocuprein	Storage or transport of copper.
Azurin Plastocyanin Stellacyanin	Electron carriers
Tyrosinase Dopamine-β-hydroxylase	Mixed function oxidases
Laccase Ascorbate oxidase [1]Ceruloplasmin	'Blue' oxidases'
Momoamine oxidase Diamine oxidase Galactose oxidase Uricase	Other oxidases
[2]Cytochrome c oxidase	Terminal oxidase.

[1]Protein with weak oxidase activity but unknown physiological function.

[2]Contains haem as well as copper.

From Malkin & Malstrom (1970).

The precise role of azurin has not been clearly defined, but it is widely accepted that it participates in a bacterial electron transport chain. Experiments with plastocyanin negative strains of Chlamydomonas rheinhardi (Gorman and Levine 1965) have placed plastocyanin as an essential component of the photosynthetic electron transport chain, acting in series with cytochrome f. Also several workers (including Arnon et al 1970), by isolating a number of chloroplast fragments, have shown that plastocyanin acts as an essential electron donor between the two photochemical reaction centres, PS1 and PS2, in plant and algal photosynthesis. The currently accepted electron pathway in photosynthesis is one in which the photosystems cooperate in a sequential manner. Reduction of NADP is catalysed by PS-1, while PS-2 is involved in the oxidation of water.

Azurin was first found in Pseudomonas (Verhoeven and Takeda, 1956; Horio, 1958), and has since been isolated in several other genera, including Bordetella, Alcaligenes and Agrobacterium (Sutherland and Wilkinson, 1963;, Sutherland, 1966). The different azurins are monomers with similar molecular weights (about 14,000) and all contain one copper atom per molecule.

Plastocyanin was first isolated from the green alga, Chlorella ellipsoida, by Katoh (1960), and has subsequently been found in many green plants and green algae. The plants from which it has been isolated

include, spinach and parsley (Katoh, 1961), the sap of Phaseolus vulgaris (french bean), (Wells, 1965), the leaves of Chenopodium album (Yakushiji, 1966). Lightbody and Krogman (1967) isolated plastocyanin from the blue-green alga, Anabaena variabilis, the first demonstration that plastocyanin is present in a procaryotic cell. It has not been found in any photosynthetic bacteria, nor in any non-photosynthetic organism.

The molecular weight and copper content for the different plastocyanins studied is variable. A figure of 21,000 is quoted for the molecular weight of native spinach plastocyanin and a copper content of 2 copper atoms/molecule (Katoh, 1961). However, the minimum chemical molecular weight determined from amino acid analyses and copper analyses gives a value around 11,000. This suggests that plastocyanin might exist as a dimer under certain conditions. The molecular weight of spinach is contrasted with a molecular weight of 11,500 and a copper content of 2 atoms/molecule for Chenopodium album, and a molecular weight of 10,690 and a copper content of 1 atom/molecule for french bean plastocyanin.

c. Primary structure studies on azurin

It is relevant here to consider the primary structure information on the bacterial azurins, and relate it to the physical studies that have been performed on the nature of the copper binding to azurin and plastocyanin, and on the elucidation of the nature of the environment around the copper binding site.

FIGURE 1(i)

AZURIN SEQUENCES (Part I)

```
                      1                   2                   3                   4                   5
    1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3

1   A.E.C.S.V.D.I.Q.G.N.D.Q.M.Q.F.N.T.N.A.I.T.V.D.K.S.C.K.Q.F.T.V.N.L.S.H.P.G.N.L.P.K.N.V.M.G.H.N.W.V.L.S.T.A.
2   A.E.  S.V.D.I.Q.G.N.  Q.  Q.F.S.T.N.A.I.T.V.D.  S.    T.    V.N.  S   P.  S.L.P.K.N.V.        W.  L.T.T.A.
3   A.E.  K.T.T.I.D.S.T.  Q.  S.F.N.T.K.A.I.E.I.D.  S.    T.    V.E.  S.  T.  S.L.P.K.N.V.        L.  I.S.K.O.
4   A.E.  K.V.T.[illegible].F.S.T.  Q.  S.F.D.T.K.A.I.E.I.D.  S.    T.    V.D.  K.  S.  N.L.P.K.N.V.        W.  L.T.T.Q.
5.  A.E.  K.V.D.V.D.S.T.  Q.  S.F.N.T.K.E.I.T.I.D.  S.    T.    V.N.  T.  S.  S.L.P.K.N.V.        W.  L.S.K.S.
6.  A.E.  S.V.D.I.A.G.T.  Q.  Q.F.D.K.K.A.I.E.V.S.  S.    Q.    V.N.  K.  T.  K.L.P.R.N.V.        W.  L.T.K.T.
7.  A.E.  E.A.T.I.E.S.N.  A.  Q.Y.D.L.K.E.M.V.V.D.  S.    Q.    V.H.  K.  V.  K.M.A.K.A.V.        W.  L.T.K.E.
8.  A.A.  D.V.S.I.E.G.N.  S.  Q.F.N.T.K.S.I.V.V.D.  T.    E.    I.N.  K.  T.  K.L.P.K.A.A.        V.  V.S.K.K.
9.  A.E.  S.V.D.I.A.G.N.  G.  Q.F.D.K.K.E.I.T.V.S.  S.    Q.    V.N.  K.  P.  K.L.A.K.N.V.        W.  L.T.K.Q.

                                                                                                    1
    5           6                   77                  8                   9                   0
    4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5

1   A.D.M.Q.G.V.V.T.D.G.M.A.S.G.L.D.K.D.Y.L.K.P.D.D.S.R.V.I.A.H.T.K.L.I.G.S.G.E.K.D.S.V.T.F.D.V.S.K.L.K.E.G.
2   A.  M.Q.G.V.V.T.    M.A.A.  L.D.K.N.  V.  D.G.  T.  V.I.      K.I.I.  S.    K.    V.  F.    S.    K.A.G.
3   A.  M.Q.P.I.A.T.    L.S.A.  I.D.K.N.  L.  E.G.  T.  V.I.      K.V.I.  A.    K.    L.  I.    S.    N.A.A.
4   A.  M.Q.P.V.A.T.    M.A.A.  I.D.K.N.  L.  E.G.  T.  I.I.      K.I.I.  A.    T.    V.  F.    S.    K.A.D.
5.  A.  M.A.G.I.A.T.    M.A.A.  I.D.K.D.  L.  P.G.  S.  V.I.      K.I.I.  S.    K.    V.  F.    S.    T.A.G.
6.  A.  M.Q.A.V.E.K.    I.A.A.  L.D.N.Q.  L.  A.G.  T.  V.L.      K.V.L.  G.    S.    V.  F.    A.    A.A.G.
7.  A.  K.E.G.V.A.T.    M.N.A.  L.A.Q.D.  V.  A.G.  T.  V.I.      K.V.I.  G.    S.    V.  F.    S.    T.P.G.
8.  S.  E.S.A.V.A.T.    M.K.A.  L.N.N.D.  V.  A.G.  E.  V.I.      S.V.I.  G.    T.    V.  F.    S.    K.E.G.
9.  A.  M.Q.G.A.V.N.    M.A.A.  L.D.N.N.  V.  K.D.  A.  V.I.      K.V.I.  G.    T.    V.  F.    S.    A.A.G.
```

Azurin sequences (Part 2)

	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129
1.	E.	Q.	Y.	M.	F.	F.	C.	T.	F.	P.	G.	H.	S.	A.	L.	M.	K.	G.	T.	L.	T.	L.	K.	
2	D.	A.		A.	F.			S.					S.	A.	M.		K.		T.	L.	T.		K.	
3	E.	K.		G.	F.			S.					I.	S.	M.		K.		T.	V.	T.		K.	
4.	G.	K.		M.	F.			S.					I.	A.	M.		K.		T.	V.	T.		K.	
5.	E.	S.		E.	F.			S.					N.	S.	M.		K.		A.	V.	V.		K.	
6.	D.	D.		T.	F.			S.					G.	A.	L.		K.		T.	L.	K.		V.	D.
7.	E.	A.		A.	Y.												K.		T.	L.	K.		S.	N.
8.	E.	D.		A.	F.			S.					W.	S.	I.		T.		E.	I.	K.		G.	S.
9.	E.	D.		A.	Y.			S.					F.	A.	L.		K.		V.	L.	K.		V.	D.

From Ambler, 1968; and personal communication (1971).

1. *Pseudomonas aeruginosa* P6009.
2. *Pseudomonas denitrificans* NCIB 9496
3. *Pseudomonas fluorescens* B.93 (ATCC 17467)
4. C-18 (ATCC 17400)
5. D-35 (ATCC 17414)
6. *Bordetella bronchiseptica* NCTC 8344.
7. *Alcaligines denitrificans* NCTC 8582
8. *Alacaligines faecalis* NCIB 8156
9. *Alcaligines spp*. (Iwasaki's *Pseudomonas denitrificans*).

The amino acid sequence of ninebacterial azurins, from the genera *Pseudomonas* (three different strains), *Bordetella*, and *Alcaligines* has been done (Ambler, 1968), in an attempt to elucidate the nature of the copper-binding site. Although the azurins share similar physical properties, it was apparent from peptide maps that there would be differences in the amino acid sequences. Figure 1(i) gives the amino acid sequences of the bacterial azurins. Approximately 70% similarities are found between azurins from different genera. The percentage similarities between the azurins is higher than that found between the penicillinases from *Bacillus licheniformis* and *Staphlococcus aureus*, which have about 40% homology and (Ambler & Meadway, 1969) is of the same order as that found between the subtilisins from two different strains of *Bacillus subtilis* (Smith, Markland, Kasper, Delange, Landon & Evans, 1966).

The most interesting features of the azurin sequences from a structural point of view are; the presence of a single cysteine residue at position 112 with a strongly hydrophobic environment around it, of the type - "tyrosine-X-Y-phenylalanine-cysteine" - where X is a variable neutral amino acid and Y is tyrosine or phenylalanine; the presence of two invariant tyrosine residues at positions 72 and 108; a single residue of tryptophan is found in all azurins, except *Pseudomonas fluorescens B-93*, but it is not in an invariant position.

It is found in position 48 in all the azurins that contain tryptophan, except in Alcaligines faecalis, where it is in position 118. The rest of the similarities and differences in the amino acid residues are distributed throughout the protein. A large number of the changes in the amino acid residues, in the highly variable positions are conservative changes, that is the changed amino acids have similar side-chains or net charge. Of great interest is that the amino acid compositions of the plastocyanin from both spinach and french bean indicate the presence of a single cysteine residue, the absence of tryptophan, and a low tyrosine content compared with that of phenylalanine.

D. Physical and chemical properties of azurin and plastocyanin

Azurins and plastocyanins exhibit very similar absorption spectra. They both have three absorption bands in the visible region; a band in the region of 450 nm., the major band responsible for the intense blue colour at approximately 600nm., and a band in the near infra-red region. The ultraviolet spectra of azurins and plastocyanins gives a major peak at 280 nm. and the occurrence of fine structure peaks reflects a low tryptophan and tyrosine content compared with that of phenylalanine. This is consistent with the amino acid sequence data for azurins, and the amino acid composition data for plastocyanins.

The reported oxidation-reduction potentials for azurins and plastocyanins vary within the range of 0.32 to 0.39 volts. Table 1(ii) gives the oxidation-reduction potentials for all the blue copper containing proteins.

TABLE 1 (ii) : Oxidation-reduction potentials of the "Blue" Cu^{2+} ion.

Protein	E'_o(pH) in volts	Reference
Laccase (Polyporus versicolor)	0.767 (6.25)	Fee and Malmström (1968)
Ceruloplasmin (Human plasma)	0.5-0.6 (6.25)	-do-
Laccase (Rhus vernicifera)	0.415 (6.8)	Nakamura (1958)
Plastocyanin (Spinach)	0.37-0.39 (7.0)	Katoh, Shiratori & Takamiya (1962).
Plastocyanin (Chlorella ellipsoida)	0.39 (7.0)	Katoh, (1961)
Azurin (Pseudomonas)	0.328 (6.4)	Yamanaka (1966)
Azurin (Bordetella)	0.395 (?)	Sutherland and Wilkinson (1963)
Stellacyanin (Rhus vernicifera)	0.3 (6.5)	Omura (1961)
From Malkin & Malstrom (1970)		

Of interest is that the values of all the oxidation-reduction potentials listed are higher than the potential of the Cu^{2+}/Cu^{+} couple (0.148 volts) in dilute acid (James and Williams, 1961). Several mechanisms have been proposed that lead to high oxidation-reduction potentials for Cu^{2+} complexes including: the involvement of sulphur or other reducing ligands, binding to unsaturated groups, and distortions in the Cu^{2+} site that do not allow a tetragonal symmetry about the metal. (Brill & Bryce, 1964). The involvement of sulphur as a ligand of copper in azurin is highly probable from the sequence studies (Ambler, 1968).

Optical Rotary Dispersion (ORD) and Circular Dichroism (CD) measurements have been used to some extent in describing the optical activity of the "blue" copper ion. The ORD curves of stellacyanin and plastocyanin (Chenopodium album) both consist of three Cotton effects in the visible region, and these agree well with the positions of the three absorption bands in the visible optical spectra. (Blumberg and Peisach, 1966). The ORD curves have been resolved into only three Gaussian components, and the curves calculated on the basis of these three components fit the experimentally obtained curves. The Cotton effects observed for the optical peaks indicate the asymmetry of the copper chromophoric site (Blumberg et al. 1966) Similar experiments with azurin have also indicated asymmetry of the copper binding site (Brill et al. 1967).

The EPR parameters of a copper complex (g and A values) are related to both the symmetry of the ligand atom arrangement around the metal ion and the covalency of the Cu^{2+}-ligand bonds. The EPR spectrum for Chenopodium album plastocyanin (Blumberg and Peisach, 1966) is similar to that reported for Pseudomonas blue protein (Mason, 1963) and for Bordetella azurin (Broman et al. 1963). Evidence from EPR studies confirms the ORD and CD work which indicated that both, azurin and plastocyanin, exhibit distortion from square planar symmetry at the cupric ion site. However, the actual geometry of the cupric ion site in these and other "blue" copper proteins remains unknown, so that the relationship of the optical and magnetic properties to the bonding of the "blue" cupric ion, will only be fully understood when the X-ray structure of a "blue" copper-protein is obtained. X-ray crystallography is proceeding slowly because of the difficulty in obtaining large enough crystals for analysis.

An important finding which is almost certainly linked with their biological function is that both azurin and plastocyanin undergo auto-reduction. This was first studied by Katoh (1962). He showed that ascorbate readily reduced spinach plastocyanin. This same phenomenon has also been observed in all the plastocyanins so far isolated (including the plastocyanins studied in this thesis) and in all the azurins so far studied, and in Polyporus laccase. This is interesting

because laccase is a more complicated copper-protein than either azurin or plastocyanin, with respect to its copper content. It contains four copper atoms per molecule, one is a "blue" cupric ion, another is a "non-blue" cupric ion, and two are "EPR-non-detectable". EPR measurements have shown that the decolourization at high pH, in azurin, is due to a reduction of the "blue" cupric ion (Brill et al. 1967). With azurin, plastocyanin and laccase little denaturation occurs at high pH since the colour can be quantitatively recovered by the addition of a suitable oxidising agent, e.g. potassium ferricyanide.

The mechanism of the pH-dependent reaction is not clear. From experiments with Pseudomonas aeruginosa and fluorescens azurin (Brill et al, 1967), it was suggested that a protein bound sulphydryl group is responsible for the reduction of the cupric ion. The sulphydryl grouping from the cysteine residue in position 112, in the sequence of azurin, is the only reducing agent present on the protein, and the pH effect would reflect a change in the redox potential for sulphydryl oxidation upon ionization.

Thus, there is suggestive evidence that sulphydryl groups may be involved in the copper binding site of azurin and plastocyanin. Further evidence for a copper-sulphydryl linkage in azurin and plastocyanin has come from experiments with sulphydryl specific

reagents. The copper from azurin and plastocyanin can be removed by extensive dialysis (Yamanaka et al. 1963) against cyanide, and restored by titration with copper salts. Katoh (1964), showed that this reversible binding of copper with the apo-protein can be abolished by blocking the single sulphydryl group of the protein with a sulphydryl specific reagent such as mercuric ions or p-mercuribenzoate. Blumberg et al. (1966) showed, by a proton relaxation experiment, that the site of binding copper is not at the surface of the molecule. A similar conclusion was reached by Brill et al. (1967) for Bordetella azurin. Blumberg et al. (1966) argued that since sulphur has been indicated as a ligand of copper in plastocyanin and since sulphur has p and d electrons which can participate in the bonding with copper, the electron for reduction of plastocyanin could easily travel through these sulphur ligands to the copper.

Evidence from these studies suggests that the binding of copper is not as simple as a copper-sulphydryl linkage. Katoh (1964), studying the copper binding site of spinach plastocyanin, has proposed that two groups of amino acids are responsible for maintaining the protein in a spatial conformation which allows the copper to form a chelation complex.

Rotilio, Finazzi Agro, Avigliano, Lai, Conti, Franconi and Mondovi (1970) using evidence from fluorescence, absorption

and EPR spectra suggests, that in Pseudomonas fluorescens azurin, there is present at least one strongly hydrophobic site for copper binding, in addition to the very probable involvement of a sulphydryl group. They tentatively propose tryptophan as a ligand for the copper. However, this seems fairly unlikely as tryptophan is not found in all azurins sequenced so far, and is not in an invariant position in the azurins that do contain it. (Ambler, 1969). Also tryptophan appears to be absent from the plastocyanins so far looked at.

High resolution nuclear magnetic resonance studies are in progress as a further method for elucidating the nature of the ligands binding the copper in azurin and plastocyanin.

The accumulative evidence from physical and chemical studies is very indicative for azurin and plastocyanin belonging to a particular family of "blue" copper-containing proteins. The primary structure of azurin is highly characterised so it was decided to study the primary structure of algal plastocyanins both as a means of finding out more about the specific bonding of copper to copper-containing proteins, and to see if there was any sequence homology between plastocyanin and azurin and if so whether it was significant enough to predict an evolutionary connection between the two proteins. Algal plastocyanins were considered to be the simplest plastocyanins to study from an evolutionary point of view.

E. The origin of the chloroplast

It is of interest to consider the origin of the chloroplast within photosynthetic organisms as being pertinent to postulating an evolutionary connection between a bacterial and a chloroplast protein.

There are two main theories. The first, an episomal theory of origin, proposes that chloroplast DNA arose from a piece of nuclear DNA establishing and evolving in the cytoplasm of a cell. The second theory suggests that the DNA originated from a prokaryotic photosynthetic cell, e.g. a blue-green alga or photosynthetic bacterium, which established itself as an endosymbiont within a eukaryotic cell. The evidence for both theories is largely circumstantial. DNA-DNA and DNA-RNA hybridization experiments have shown that homologous base sequences do exist in nuclear and chloroplast DNA (Richards, 1967), but further experiments need to be done before a common origin of nuclear and chloroplast DNA can be suggested. The similarities between chloroplast DNA with DNA of bacteria (and blue-green alga) include; the similarity in appearance of DNA images seen in electron micrographs of chloroplasts, bacteria, and blue-green algae, the comparable amount of DNA in a chloroplast and bacterium, and the apparent absence of histones associated with DNA in chloroplasts and bacteria. Also it has been found that chloroplast ribosomes are similar to bacterial ribosomes. They both have a sedimentation

coefficient in the ultracentrifuge of 70S, share similar sensitivities to magnesium ion concentration and antibiotics, and the RNA components of the ribosomes are similar in size. However, reported figures for the sizes of the chloroplast ribosomal RNA components is not consistently the same from different laboratories and more detailed studies need to be done. Actually, mounting genetic evidence (Kirk, 1966) suggests that a large number of structural genes for individual chloroplast proteins do reside in the nucleus. This would make an endosymbiont theory of chloroplast origin more difficult to uphold. But either theories is non-contradictory to a possible evolutionary connection between bacteria and eukaryotes.

However, if the nuclear DNA is found to code for plastocyanin then one must consider the evolution of the whole eukaryotic cell from a bacterial origin, while if it is found that the chloroplast DNA codes for plastocyanin, then only the evolution of the chloroplast within the eukaryotic cell requires to be considered.

F. Conclusions

Although bacteria do contain some enzymes whose activities are analogous to those of enzymes in higher organisms, e.g. the *Pseudomonas* cytochrome c-551 (Ambler, 1963), *Rhodospirillum rubrum* cytochrome c (Dus et al. 1968) with animal cytochromes c, *Chlostridial*

ferredoxins with plant ferredoxins (Matsubara et al. 1967), bacterial subtilisins with the mammalian serine proteinases, only in the case of the soil bacterium, Sorangium sp., and the mammalian serine proteinases is it widely accepted that the bacterial protein and the animal protein are homologous (Smillie & Whitaker, 1967).

There remain many functionally analogous proteins common to both bacteria and higher organisms which have yet to be compared. Care must always be taken, however, when comparing proteins, to consider the possibility that analogous chemical structures may have arisen as a result of convergent evolution.

CHAPTER 2

ABBREVIATIONS & METHODS

1. Abbreviations

A. Chemicals

BAWP:	Butan-1-ol: acetic acid: water: pyridine (15:3:12:10 by volume)
CPA:	Carboxypeptidase A
C-t:	C-terminus
DEAE:	Diethylamino ethyl
DNP:	2,4 - dinitrophenyl
DNS:	1-dimethyl aminonaphthalene-5-sulphonyl
DNS-PTC:	Using PTC to degrade peptides and DNS-chloride to label the N-termini.
DPCC-:	Treated with diphenyl carbamyl chloride
HVPE:	High voltage paper electrophoresis
E-DNP Lys:	Epsilon 2,4-dinitrophenyl lysine
KCL:	Potassium chloride
N-t:	N-terminus
PTC:	Phenyl isothiocyanate
XCFF:	Xylene cyanol FF
m:	Electrophoretic mobility at pH 6.5
m':	Electrophoretic mobility at pH 3.5

B. Amino acids

Two types of abbreviations have been used for amino acids, the three letter and the single letter abbreviations Commission on Biochemical Nomenclature, 1967, and 1969, respectively).

Amino acid	Three letter	One letter
Lysine	Lys	k
Histidine	His	h
Arginine	Arg	r
Aspartic acid	Asp	d
Asparigine	Asn	n
Either Asp or Asn	Asx	b
Cysteic acid (Cysteine)	Cys	c
Methionine	Met	m
Methionine Sulphone	Mes	m
Glutamic acid	Glu	e
Glutamine	Gln	q
Either Glu or Gln	Glx	z
Proline	Pro	p
Glycine	Gly	g
Alanine	Ala	a
Valine	Val	v
Isoleucine	Ile	i
Leucine	Leu	l

Amino acid	Three letter	One letter
Threonine	Thr	t
Serine	Ser	s
Tyrosine	Tyr	y
Phenylalanine	Phe	f
Tryptophan	Trp	w

Residues joined by hyphen are in sequence, whereas those bracketed are in unknown order.

C. Peptide nomenclature

The first letter shows the primary digestion by which the peptide was formed: T, trypsin; C, chymotrypsin; S, subtilisin.

The second letter shows the mobility of the peptide at pH 6.5: B, basic; A, acidic; N, neutral.

The peptides in each series are then distinguished by numbers. Thus, CN1 is a neutral chymotryptic peptide.

Peptides produced by further digestion with another protease are distinguished by the same system. Thus, TA6CB1 is a basic chymotryptic peptide formed by chymotrypsin digestion of peptide TA6.

In the figures, peptides represented by a solid line (———) were sequenced by the Dansyl Edman procedure, those represented by a broken line (-------) were analysed for amino acid composition, and those represented by a solid line and partly by a broken line (—— --- ——) were partially sequenced. In the latter case the N-termini were usually determined by the DNS-chloride method.

2. Methods

A. Purification of plastocyanin

A(i) Detection of plastocyanin in solution

In crude extracts plastocyanin is present in the reduced form and is brownish in colour. It is non-auto oxidisable, but can be readily oxidised by the addition of 0.1 M potassium ferricyanide. On oxidation plastocyanin is converted to a pronounced blue colour.

A(ii) Preparation of acetone powder

The algal cells were homogenised in a Waring blendor with cold acetone, chilled for several hours in the deep-freeze. The cells were filtered free of excess acetone and dried in a dessicator using a water-pump.

A(iii) Ion-exchange Chromatography

DEAE-cellulose (Whatman DEAE-11 and DEAE-52 grades) has been used in the purification of plastocyanin. Both grades of DEAE-cellulose were equilibrated with 0.ol M K phosphate buffer pH 7.0 before use.

Elution of the proteins was achieved using both stepwise and gradient methods of elution. These are described in the text where appropriate (see Chapter 3).

Gradients of increasing salt concentrations were generated using a two bottle system, having a constant volume with respect to one another. The second resevoir contained the buffer of higher salt concentration, and this flowed into the first resevoir, containing the buffer of lower salt concentration, through a rubber tubing connecting the two resevoirs. The buffer in the first resevoir was stirred magnetically and pumped on to the top of the column.

The extinction of the effluent from columns at 254 nm was recorded automatically on an ultraviolet spectrophotometer.

A(iv) Gel filtration

Sephadex consists of dextran macromolecules which are cross-linked to give a three-dimensional array of polysaccharide chains. It has a high degree of hydroxyl ions and is therefore strongly hydrophilic.

The Sephadex types found in the G-series possess varying degrees of cross-linkage and swelling properties in water.

The Sephadex used in the purification of plastocyanin is Sephadex G-50, which has a fractionation range for peptides and globular proteins of between 1,500 and 30,000.

The Sephadex G-50 purification step described in the text (Chapter 3), was not used as a chromatographic method for the determination of the molecular weight of plastocyanin.

A single Sephadex G-50 column (of dimensions 83 x 5 cm.), equilibrated in 0.01 M K Phosphate buffer pH 7.0, was used for all the plastocyanins, and the extinction of the effluent at 254 nm. was recorded automatically on an Ultraviolet Spectrophotometer.

Sephadex G-25 (column of dimensions 1.5 x 135 cm. and pumped at 40 ml/hr), equilibrated in 5% formic acid, was used for the fractionation of protein digests. The digests (see later chapters) were dissolved in 5% formic acid (about 2.0 ml.) and 50 drop fractions (2.9 ml.) were collected. The extinction of the effluent at 254 nm. was recorded automatically on an Ultraviolet spectrophotometer.

In one instance (see 5A(ii)) Sephadex G-50 was used for the separation of a peptide mixture. Sephadex G-50 (column of dimensions 1.5 x 135 cm. and pumped at 40 ml./hr) equilibrated in 5% formic acid was used. The extinction of the effluent at 254 nm. was recorded as described above.

In all the gel filtration separations E-DNP Lysine (2.5 mg./ml.) was used as an external marker.

A(v) Protein determination.

Protein was determined by the Lowry (1951) method. The protein samples were made up to 0.5 ml. with water, and 3 ml. of a freshly made solution A was added, and the mixture shaken and left for 10 min.

Solution A

100 ml. of 2% (w/v) Na_2CO_3 in 0.1 M sodium hydroxide

1 ml. of 2% (w/v) sodium potassium tartrate

1 ml. of 1% (w/v) $CuSO_4\ 5H_2O$.

Then 0.3 ml. of Folin-Ciocalteau reagent (diluted with an equal volume of water) was added, the tubes shaken, and left to stand for 30 min. Standards were prepared containing up to 0.5 mg. of bovine serum albumin, and a standard curve drawn. The extinctions at 750 nm. were read after ½ to 2 hr., and the protein determined from the standard curve.

B. Criteria of purity of plastocyanins.

B(i) Starch Gel Electrophoresis.

Horizontal gels were prepared by the method of Smithies (1959). 19.25 g. of starch (Connaught) was added to 175 ml. 0.02 M. sodium cacodylate/HCL pH 6.5, made up to contain 1 mM. potassium ferricyanide, in a 1 litre round bottomed flask. It was mixed and swirled over a Bunsen flame until translucent, de-gassed, and poured into a perspex tray (14.8 x 9.0 x 1.5 cm. internally, with a plate 0.6 cm. thick fitted in the bottom). The gel was allowed to set, covered with polythene, and stored at 4°.

A slit was cut with a razor blade 7 cm. from one end and samples of protein, loaded on small pieces of 3 MM. Chromatography paper, were inserted. One piece of paper was loaded with XCCF (0.25 mg./ml) as an external marker. The protein was always run in the blue oxidised form, so that it was visible whilst running.

The electrode vessels were filled with 0.2 M sodium cacodylate/HCL pH 6.5, made up to 1mM potassium ferricyanide, and 30 m.amps. were allowed to flow through the gel (250 volts). The run was ended when the XCCF had moved about 3 cm. from the origin (approx. 4 hr.), and the gel left to cool. The distances moved by XCCF and the blue protein bands from the origin were recorded.

The gel was sliced, usually into three, by inserting increasing numbers of 0.1cm. thick spacer plates underneath the gel and false bottom of the perspex tray. The gel was then sliced with a piece of wire across the top of the tray.

The slices of gel were stained for protein by covering with 0.5% (w/v) Amido (Napthalene) Black in methanol: water:acetic acid (5:5:1 v/v) for ten minutes. The excess stain was removed by repeated washing with the same solvent. This provided a permanent record of the protein bands.

B(ii) Absorption Index

The spectral purity of plant and algal plastocyanins has been expressed as the $E_{278nm} : E^{ox}_{597nm}$ ratio (Katoh et al. (1962); Wells (1965) and Gorman & Levine (1966)).

The plastocyanin content of protein samples is obtained by measuring their absorbance at 597 nm., after ensuring complete oxidation with 0.1 M potassium ferricyanide. This value is corrected for absorption due to contaminating pigments, after dialysis of the protein to remove the ferri- and ferro-cyanides, followed by reduction with sodium ascorbate.

The criterion is that the absorbance at 278 nm. is assumed to be an approximate measure of the total protein and the absorbance at 597 nm. (oxidised minus reduced) a measure of the amount of plastocyanin. The absorption index is therefore inversely related to the purity of the plastocyanin.

C. Experiments on plastocyanin.

C(i) Carbohydrate determination

The method of Devor (1950) was used, which detects pentose and hexose, but not amino-sugars. These were looked for on the long column of the amino acid analyser.

Devor method

5 mg./ml. soltutions of the algal plastocyanins were made up in water and 0.2 ml. samples of each was hydrolysed with 6N HCL for 24 hrs. and analysed in order to quantitate the amount of protein present. 0.1 ml. samples of each solution (including duplicates) were taken in parallel with standards containing up to 0.1 mg. glucose. The samples were diluted to 2 ml. with water.

A solution of 0.4 g. 1-napthol/100 ml. conc.sulphuric acid was kept in the dark for several hours before use, and 5 ml. was added to each sample. After 10 mins. at 100°C the samples were cooled and their extinctions at 575 nm. were measured.

The protein samples gave faint red colours.

Galactosamine and glucosamine were looked for on the long column of the amino acid analyser (see section 2 D(x)). Using the method of Spackman, Stein & Moore (1958), the amino sugars are eluted approximately 2 hr. after the aromatic amino acids, tyrosine and phenylalanine.

C(ii) Tryptophan determination

The tryptophan content of the whole protein was determined by the Harrison & Hoffman (1961) modification of the method of Spies & Chambers (1958). The reason for using this method is that some proteins, instead of giving a blue colour (absorption maximum 590-600 nm.) like that given by free tryptophan, give a purple-pink colour (absorption maximum 545-560 nm.). Thus, no reliable estimate

of tryptophan content can be measured. However, stable blue chromophores, with essentially the same absorption as free tryptophan, can be obtained if the protein is first denatured and partially digested with pepsin.

5 mg./ml. solutions of C.pyrenoidosa 'A', C.fusca, and S.obliquus 'A', plastocyanins were made up in pH 2.0 buffer (formic acid:acetic acid:water 1:4:45 v/v). The protein solutions, plus two water blanks, were boiled for 15 mins. 0.015 ml. pepsin (10 mg./ml. solution) was added to each tube and incubated at 37°C for 5 hr. The digests were dried down and each one was taken up in 0.5 ml. water. 0.1 ml. samples (including duplicates) and 0.2 ml. samples (including duplicates) were taken from each digest and their volumes made up to 0.25 ml. with water. Standards containing up to 0.5 µmole L-tryptophan were taken and these were also made up to 0.25 ml. with water.

To each tube 2 ml. 65 % (v/v) H_2SO_4 plus 0.25 ml. 3% (w/v) p-dimethylamino-benzaldehyde in 2N HCL was added. The solutions were mixed thoroughly, left for 1 hr. in the dark, 0.025 ml. 0.04% (w/v) sodium nitrite added, mixed, and finally left for ½ hr. in the dark. The extinctions were read at 590 nm. and the tryptophan content determined from a standard tryptophan curve.

0.01 ml. samples (5 mg/ml.) of each plastocyanin solution were hydrolysed in 6N HCL, and analysed for protein content.

D. Determination of amino acid sequence.

D(i) Oxidation with performic acid.

The reagent was prepared, according to the method of Hirs (1956), by adding together 9.5 ml. concentrated formic acid and 0.5 ml. 30% (100 vol.) hydrogen peroxide. After 2 hr. at room temperature it was ready for use.

The protein was dissolved in concentrated formic acid (40 mg./ml.) and an equal volume of performic acid (equivalent to a ten fold excess over that of methionine) was added. The mixture was allowed to react for up to 2 hr. in an ice bucket, and then diluted with 25 volumes of cold water and freeze-dried.

D(ii) Preparation of proteolytic enzymes.

Solutions of trypsin and chymotrypsin were stored at -20°, whereas pepsin, subtilisin B, and carboxypeptidase A solutions were prepared immediately before use.

Trypsin was treated with diphenyl carbamychloride (DPCC), an inhibitor of chymotrypsin (Erlanger & Cohen, 1963). Trypsin was dissolved at 10 mg./ml. in 0.1 M tris-chloride, pH 8.0 containing 0.1 M $CaCl_2$. 0.005 ml. of DPCC (25 μmoles/ml. in acetone) ml. of trypsin solution was added and the mixture allowed to react for 30 min. at room temperature. If a precipitate formed it was centrifuged off.

Chymotrypsin (at 10 mg./ml.) was pre-incubated with soya-bean trypsin inhibitor (1:10 w/w) in 0.2 M ammonium acetate pH 8.5 at 37° for 1 hr.

An aqueous solution of pepsin (at 10 mg./ml.), and subtilisin B in 0.2 M ammonium acetate pH 8.5 were prepared just prior to using.

Carboxypeptidase A was prepared by the method of Fraenkel-Conrat et al. (1955). 1.25 mg. CPA was washed with water and centrifuged. The solution was decanted off and the enzyme suspended in 0.1 ml. 1% sodium bicarbonate, and cooled in ice. 0. 1 M NaOH was added drop by drop until the enzyme dissolved and then pH was brought down to approximately 8.5 with 0.1 M HCl. If the pH dropped lower than this the enzyme precipitated and had to be discarded. The CPA solution was diluted to 1.25 ml. with 0.2 M N-ethyl morpholine acetate pH 8.5, giving a 1 mg./ml. solution. The solution was kept on ice and used as soon as possible.

For tryptic, chymotryptic subtilisin B, and peptic digestions of plastocyanin, the enzyme to substrate ratio was 1:40 parts by weight. For digestion of peptides (0.02-0.05 µmole), 0.005-0.01 ml. of a freshly made 1 mg./ml. solution of the enzyme was added. For CPA digestion of peptides, 0.01-0.02 ml. from 1 mg./ml. solution of enzyme was used per 0.02-0.05 µmole peptide. The digestion procedure was based on the methods described by Ambler (1963b).

D(iii) Cyanogen Bromide Digestion

The method of Gross & Witkop (1961) was used. Cyanogen bromide specifically cleaves the protein chain at/the C-terminal side of methionine residues and the conditions used generally achieved over 90% cleavage. The methionine residues are converted to homoserine, which is in equilibrium with its lactone.

Unoxidised protein (20 mg.) was dissolved in 2 ml. 50% (v/v) formic acid and treated with an equal weight of cyanogen bromide for 20 hr. at room temperature. At the end of this time, the reaction mixture, which had remained in solution, was diluted with 10 volumes of water and freeze-dried.

D(iv) High-voltage paper electrophoresis.

A Michl (1951) Tank apparatus was used. The paper (57 cm. long and up to 46 cm. wide) is suspended in an organic solvent, cooled with water passing through a coil at the top of the tank. The top of the paper is held in a trough containing buffer and connected to one electrode, and the bottom dips into a layer of buffer at the bottom of the tank connected to the other electrode.

The buffer systems were pyridine-acetic acid-water (25:1:225 v/v), pH 6.5, pyridine-acetic acid-water (1:10:89 v/v), pH 3.5, formic acid-acetic acid-water (1:4:45 v/v), pH 2.0. Toluene (containing 8% pyridine v/v) was the coolant with the pH 6.5 buffer, and White Spirit with the pH 3.5 and pH 2.0 buffers.

Whatman 3MM paper was used except for amino acid separation and small quantities of peptides, when Whatman No.1 paper was used. At pH 6.5 the sample was loaded near the middle, and at pH 3.5 and pH 2.0 (preparative peptide separations) it was loaded 7 cm. from the bottom of the paper. For amino acid separations at pH 2.0, the sample was loaded 11 cm. from the bottom of the paper.

For preparative peptide separations and analytical peptide maps at pH 6.5, 3.5 and 2.0, the sample was dissolved in an appropriate volume of 0.1 N ammonia, and loaded on to the paper on between 0.1-0.5 µmole/cm. (for a digest of the whole protein). For separating small quantities of peptide, 0.02-0.1 µmole peptide/cm. was loaded on to the paper. For all peptide separations 0.7 cm. bands of peptide were placed 1 cm. from the main band on either side of it. Also, 1 cm. bands of a standard amino acid mixture (approx. 0.005 ml.) was placed 1 cm. from the 0.7 cm. bands of peptide. When it was necessary to load the peptide mixture on bands between 15 and 20 cm., in addition to the two 0.7 cm. peptide bands on either side of the main band, a 0.7 cm. band was marked in the middle of it. After the paper had been run the 0.7 cm. bands of peptide and 1 cm. band of standard mixture were cut off and stained to locate the peptides, and then the corresponding band was cut out from the main band and the peptides eluted.

Analytical 'peptide maps' of protein digests, fractionated on Sephadex G-25 were separated at pH 6.5 0.01 ml., samples from each fraction tube were successively applied 0.6cm. apart on the origin. The spots were dried with cold air from a hair drier, and the samples applied until 0.05 ml. of each had been spotted. 1 cm. bands of a standard amino acid mixture (about 0.01 ml.) were placed on either side of the peptide mixture.

The standard mixture ("Wondermix") contains 5 µmole/ml. of lysine, histidine, arginine, glycine, valine, glutamic acid, aspartic acid, alanyl-glycine, taurine, cysteic acid and the yellow marker E-DNP lysine, and the blue marker Xylene cyanol-FF. At pH 6.5 a mark of red pentol was placed on top of the standard mixture band and this was run to about 3 cm. from the bottom of the paper (about 1 hr. at 3 Kvolts). At pH 3.5 a red pentol mark was put on the edge of both sides of the paper about 7 cm. from the top, and the paper was run until the red pentol mark reached E-DNP lysine (about 1 hr. at 3 kvolts). At pH 2.0 E-DNP lysine was run to 12 cm. from the origin of the peptide sample (about 50 mins. at 3 Kvolts). The conditions used for separation of peptides are based on those described by Ambler (1963b).

For analytical amino acid separations at pH 2.0, the hydrolysed peptides were dissolved in 0.01 ml.0.1 M ammonia, and loaded on to 1 cm. bands (separated by 1 cm. from one another). Two standard amino acid mixtures (denoted 'R' and 'T') were loaded on 1 cm. bands, 'R' on one side and 'T' on the other side of the samples. Mixture 'T' contains E-DNP lysine, and this was run to 10.5 cm. from the origin (about 15 mins. at 7 Kvolts).

Amino acid markers for electrophoresis at pH 2.0

'R'	'T'
lysine	histidine
arginine	glycine
valine	alanine
leucine	serine
proline	isoleucine
phenylalanine	threonine
tyrosine	glutamic acid
methionine sulphone	aspartic acid
cysteic acid	E-Dinitrophenyl lysine

The order is the order of separation from cathode (-ve) to anode (+ve) on electrophoresis at pH 2.0.

D(v) Paper Chromatography (B.A.W.P.)(Waley & Watson, 1953).

Peptides were separated by descending chromatography, on Whatman 3MM paper with the solvent system butan-1-ol-acetic acid-water-pyridine (15:3:12:10 v/v) for 15 hr. at 20°. The sample was loaded 7 cm. from the top of the paper and the bottom edge of the paper was serrated. The sample load/cm. of paper was the same as that used for HVPE.

Peptides which had been run on paper chromatography were always purified by re-electrophoresis at pH 6.5 or pH 3.5, before their amino acid compositions were determined, since the bands obtained on paper chromatography were always very diffuse.

D(vi) Detection and elution of peptides

The paper was always examined for tryptophan-containing peptides by viewing under ultraviolet (365 nm.) light, before staining with ninhydrin.

Ninhydrin dip

Peptides were located by dipping the paper in 0.2% (w/v) ninhydrin in acetone, drying, and then heating at 80-100°C for a few minutes. The peptides mostly showed up as blue spots, but brown and yellow spots were also observed.

Histidine stain (Dent 1947)

This works well after ninhydrin. The paper is first sprayed on both sides with diazotised sulphanilic acid. This is prepared by mixing 1 % (w/v) sulphanilic acid in N HCl with an equal volume of ice-cold 5% (w/v) sodium nitrite, and leaving the mixture at 4° for 5 mins. to react. The paper is then sprayed, whilst still wet, with 10% (w/v) sodium carbonate. Histidine-containing peptides show up as cherry-red bands.

Tyrosine stain (Jepson & Smith, 1953).

The reagent works well after ninhydrin. The paper is dipped in 0.1% (w/v) α-nitroso β-napthol in acetone, allowed to dry, and then dipped in 10% (v/v) nitric acid in acetone. The paper is dried, and then heated. Tyrosine-containing peptides show up as red bands.

Tryptophan stain (Ehrlich reaction - Dagliesh, 1952).

The paper is dipped in 10% (v/v), of 1% (w/v) p-amino dimethyl benzaldehyde in acetone and conc. HCl. The reagent is prepared fresh. Tryptophan-containing peptides slowly show up as mauve spots. The reagent works well after ninhydrin.

Elution of peptides

The peptides were eluted with 0.1 M ammonia, by allowing the ammonia to run down the strips for about 3 hr. The peptides were collected in pyrex tubes (10 x 100 mm.) and dried under vacuum in a conc. H_2SO_4/NaOH dessicator. Caution was taken to ensure that bumping did not occur.

Occasionally the peptide strips, after elution, were dried and dipped in ninhydrin to check that all the peptide had been eluted. In every case no remaining peptide material was detected.

D(vii) Peptide mobilities.

The 'electrophoretic mobility' of peptides was measured at pH 6.5(m) and pH 3.5 (m'). The values given are not absolute figures, since different peptide mixtures were run on electrophoresis for varying lengths of time, and in the system used the endosmotic movement of uncharged substances was quite appreciable. The position of the 'neutral' (mono amino mono carboxylic) amino acids, separated in the standard amino acid mixture, run on every paper, was always taken as the true origin of the sample.

At pH 6.5 the mobility of basic peptides were measured relative to the distance moved by lysine from the origin, and acidic peptides relative to the distance moved by aspartic acid from the origin. Basic peptides were given a positive charge, and acidic peptides a negative charge. The mobility of the same peptide could alter between ± 0.02 to 0.05, as measured in different digests, but the mobility relative to adjacent peptides in the same digest remained unaltered.

At pH 3.5 the mobility of all peptides was measured relative to the distance moved by lysine from the origin. For some neutral and acidic peptides, electrophoresis at pH 3.5 was run for an extended length of time, when lysine was run off the top of the paper. In these cases the mobilities of the peptides were not recorded in any way.

D(viii) Acid hydrolysis.

Protein (usually 1 mg.) and peptide (0.02-0.05 μmoles) samples for quantitative amino acid analysis were hydrolysed in pyrex tubes (12 x 100 mm.) with 0.2 ml. Analar conc. HCl plus 0.2 ml. water, and the tubes were sealed under vacuum to prevent oxidation. The tube was drawn out, to give a constriction of about 1 mm. diameter, the contents frozen in a butanol/dry ice mixture, evacuated,usually until the contents thawed, and finally the tube was sealed at the constriction. The sample was hydrolysed at 105°C for 24 hr. Occasionally protein and/or peptides were hydrolysed for a longer time, e.g. 96 hr., to ensure complete hydrolysis of particularly resistant peptide bonds, e.g. Val-Val bonds.

Peptide (0.005-0.01 μmole) samples for qualitative amino acid analysis were dissolved in 0. 1 ml. 6 M HCl in a Durham tube (8 x 35 mm.). The tube was sealed and hydrolysed at 105°C for 12-24 hr.

For both quantitative and qualitative amino acid analysis of peptides, the acid was removed in a vacuum dessicator over sodium hydroxide pellets. Peptide samples for quantitative amino acid analysis were stored, stoppered, at 4°C.

D(ix) Qualitative amino acid analysis.

The acid hydrolysates were subjected to electrophoresis at pH 2.0 as described in Chapter 2 (section D(iv)). The paper was heat dried and examined under U.V. light (365 nm.) for fluorescence, characteristic of decomposition products of tryptophan.

The paper was dipped in 0.2 % (w/v) ninhydrin in acetone to which, just before use, about 1% (v/v) collidine is added (Levy and Chung, 1953). It was then dried and heated for about 1 min. at 105°C. Characteristic bright, transient colours are given by particular amino acids, e.g. proline is yellow, phenylalanine is greenish-blue, tyrosine is greyish-brown, histidine is brown, glycine is red, serine is grey, threonine is grey, and aspartic acid is turquoise. On standing all amino acids, except proline, give a uniform blue colour.

The amino acid compositions were recorded and the relative amount of each amino acid in a peptide was estimated visually by comparison of the colour intensities.

D(x) Quantitative amino acid analysis

The quantitative amino acid composition of protein and peptides was determined in acid hydrolysates by ion-exchange chromatography using an automatic amino acid analyser (Spackman et al. 1958). Two

amino acid analysers were used, an Evans Electroselenium Ltd. (EEL) model, where the columns were packed with pulverised and graded resin (Beckman 50B), and the Beckman 120C model, where the columns were packed with resin polymerised in beads (Beckman UR-30γ PA-35,γ Durrum Hi-Rez DC-1A) (Benson & Patterson, 1965). The two machines differed in their sensitivity, measured in μmole amino acid for 0.01 absorbance peak height. For the EEL, the sensitivity was 0.001 (Proline 0.003), and for the Beckman, 0.003 (Proline 0.012).

The samples were prepared for analysis as described in this chapter. Each sample was dissolved in 0.35 ml. citrate buffer pH 2.2, containing 1 μmole/ml. of two amino acid standards, DL-Norleucine and L-2-amino 3-guanido propionic acid (Walsh & Brown, 1962). Both the short and long columns were loaded with 0.15 ml. samples, leaving 0.05 ml. of sample in case one of the columns had to be repeated.

The amino acids were identified, and the amount of each was determined from the area of the peak and compared with a value for a known amount of that amino acid.

For whole protein samples the amount of each amino acid was expressed as μmole/mg. protein (by weight).

Contaminating amino acids are given in tables of the amino acid composition of peptides, when they are present in an amount of 0.2 mole/mole peptide or more.

D(xi) The Dansyl-Chloride Method. (DNS method)

(Gray, Hartley, 1963a,b; Gray, 1967).

The DNS method was used both for the determination of peptide N-termini and of the new N-terminus exposed in a peptide after the removal of an amino acid residue by the PTC degradation method.

The peptide (0.005-0.01 μmole) was dried down in a Durham tube (6 x 30 mm.). 0.01 ml. of 0.1 M sodium bicarbonate was added and the tube dried under vacuum. Then 0.01 ml. of water plus 0.01 ml. DNS chloride (2.5 mg./ml. in acetone) were added, the tube covered and incubated at 37°C for 2 hr., after which time the mixture was no longer yellow. The liquid was removed under vacuum, 0.05 ml. 6N HCl added and the tube sealed. The DNS-peptide was normally hydrolysed at 105°C for about 12 hr., but for longer (up to 96 hr.) if a stable N-terminal residue was suspected or for shorter (4 hr.) if N-terminal proline was suspected.

D(xii) Identification of DNS-amino acids.

DNS-amino acids were identified by Thin Layer Chromatography on polyamide plates (Woods & Wang, 1967), using three solvent systems (see Table 2(i).

TABLE 2(i) Solvent systems for separation of DNS-amino acids.

1st dimension: Conc.formic acid:water (3:200 v/v)

2nd dimension: Toluene: glacial acetic acid (9:1 v/v)

3rd dimension: Ethyl acetate:methanol:glacial acetic acid (20:1:1 v/v).

The DNS-amino acids were dissolved in 0.004 ml. 50% pyridine in water and applied to a spot on the plate marked 3.5cm. along the diagonal from the bottom left-hand corner of the plate. 0.002 ml. of a standard mixture of DNS-amino acids (1 μmole/ml.), was applied on the opposite side of the plate in a coincident position to the sample. The plates were placed in a support and run in the 1st solvent for 50 min., the 2nd solvent in a direction at right angles to the 1st dimension for about 1 hr., and in the 3rd solvent (in the same direction as the second)for about 1 hr. The plates were dried by gentle heating between runs.

The DNS-amino acids were identified as yellow fluorescent spots under U.V. light, and the identification of the DNS-amino acid was determined by direct comparison with the running positions of the components of the standard mixture on the reverse side of the plate. The relative running position of each DNS-amino acid is shown in Figure 2(i).

This method enables one to identify unambiguously most DNS-amino acids commonly found in proteins, except DNS-arginine, DNS-histidine and -DNS-lysine which do not separate from one another very well.

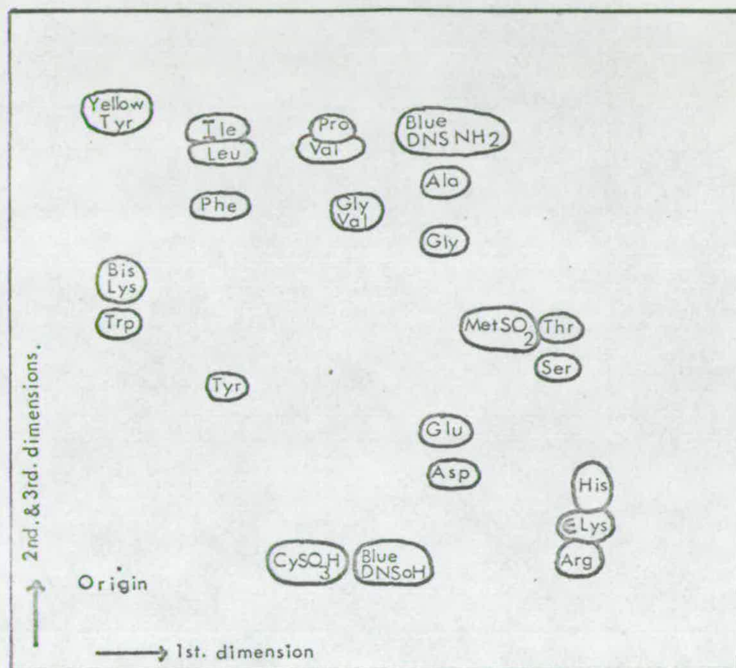


Figure 2(i). Separation of DNS-amino acids on polyamide thin layer plates.

D(xiii) Carboxypeptidase A treatment of peptides

The enzyme was prepared as described in this chapter.

The peptide (0.02-0,05 μmole) was dissolvediin about 0.1 ml. 0. 2 M N-ethylmorpholine adjusted to pH 8.5 with acetic acid, and 0.01-0.05 ml. of enzyme was added and the mixture incubated at 37°C for 4 hr. After incubation about 0.01 ml. 1 M acetic acid is added to stop the reaction and the solution dried under vacuum.

The products of digestion (free amino acids, unchanged peptide and residual peptides) were separated on paper electrophoresis either at pH 6.5, if the original peptide carried a net charge at this pH, or at pH 3.5 if the original peptide was electrophoretically neutral at pH 6.5.

After paper electrophoresis, at either pH, the neutral amino acid band was not stained, but was cut out, sewn on to another sheet of paper and separated on electrophoresis at pH 2.0, along with samples of 'R' and 'T' standard markers (see 2 D (iv)). The amino acids were identified by staining with ninhydrin solution.containing 1% collidine (v/v).

The remaining pieces of the original paper were stained with ninhydrin and the electrophoretic mobilities of the residual peptide (or peptides) was recordèd.

Water blanks were always carried through the procedure to check on any self-digestion by the enzyme. Also, samples of the undigested peptide were spotted on the same paper electrophoresis as the digested peptides to obtain an accurate electrophoretic mobility of the peptide.

D(xiv) Sequential degradation (DNS-PTC method) of peptides.

The method described by Edman (1950, 1956) was used. The peptide (0.01 - 0.02 μmole x number of cycles of degradation intended) was dried in a screw cap tube (12 x 65 mm.). 0.2 ml. 50% pyridine (v/v) and 0.1 ml. phenyl isothiocyanate solution (PTC) (5% v/v in pyridine) was added and the tube gassed with oxygen-free nitrogen, capped, and incubated at $37^{\circ}C$ for 1 hr. The solution was dried in a $60^{\circ}C$ vacuum dessicator over conc. H_2SO_4, the peptide cooled, and then dissolved in 0.2 ml. tri-fluoroacetic acid, gassed as before, capped, and incubated at $37^{\circ}C$ for ½ hr. The solution was dried under vacuum over sodium hydroxide pellets. 0.25 ml. of water was added and the PTC-amino acid(s) extracted three times with 2 ml. butyl acetate, the top phase being discarded after each extraction, and the solution was finally dried under vacuum.

The dried residue was dissolved in 0.2 ml. 50% (v/v) pyridine, and a sample (0.01-0.02 μmole peptide) removed for N-terminal group determination by the DNS-chloride method. The volume of the remainder was made up to 0.2 ml. and the next cycle of degradation performed.

In order to identify the position of asparigine and glutamine residues in some peptides, additional samples were taken out after the appropriate cycle of degradation, and the electrophoretic mobility of the residual peptide at pH 6.5 examined.

CHAPTER 3

PURIFICATION OF PLASTOCYANIN

A. Strains of green algae

The strain designations were as follows:

Chlorella pyrenoidosa:	CHICK strain SOROKIN-MYERS Tx71105
Chlorella fusca:	var.vacuolata (KRAUSS strain TOKUGAWA/TOKYO 27)
Scenedesmus obliquus:	(TURB) KUTZ strain LHOTSKY 1966/67

B. Growth of algae

The algae were grown in large outdoor culture units, which held a total volume of 40,000 litres. Such a unit required a massive innoculum (by laboratory standards) to initiate the algal growth, and to ensure a rapid and reliable population growth. The innoculum was provided by an indoor culture unit, capable of continuous operation, which working with an artificial radiation source, yields 0.5 kg. dry weight of algae per day. The daily yields were stored at approximately 0°C, in a viable state, without appreciable damage, until some 20 kg. of cells had been collected. The cells were then resuspended in media and used to innoculate the outdoor culture unit.

The suspension of algae flows on a slightly inclined plane glass surface, fitted with transverse baffles, which create a turbulent motion favourable for growth. The suspension is collected in a trough at the

EX BIBL UNIV. EDINBURGEN

bottom of the glass surface, and then pumped back up to the top. Thus, a continuous flow is maintained. The suspension of algae is kept circulating only during the hours of day when there is enough incident radiation to promote a reasonable growth. During the night hours the suspension is kept in a tank, where it is well aerated and stirred.

The algae is harvested by centrifugation and then subjected to spray-drying.

Some control on the algae population was made by removing samples from the outdoor culture unit, each day, for examination under the light microscope. Thus, if a massive contamination of an algal culture occurred it would be detected. But it is almost impossible to prevent the algal powders, obtained from such an outdoor culture unit, from being contaminated with a mixture of symbiotic bacteria and with material (e.g. pollen and dust), which can get into the cultures from the open air. This type of contamination did not prove to be a problem during the purification of plastocyanin.

C. Purification of plastocyanin

The method described for the purification of plastocyanin is a general one, which was used for all the plastocyanins. The only variable in the method of purification, that needed to be applied, was at the final DEAE-52 cellulose step where slightly different salt concentrations were required to achieve the best purification of the individual plastocyanins. All the procedures were carried out at room temperature, unless otherwise stated.

Several preparations from each algal species wereperformed and each time a batch size of 1 kg. dried cells was used.

(i) Extraction of dry algal powders

An acetone powder of the dried cells was prepared as described (2A(ii)), in order to ensure good breakdown of the cell walls.

The acetone powder was homogenised with 0.01 M potassium phosphate pH 7.0, and the suspension left overnight at 4°C, to ensure complete lysis of the chloroplasts.

The suspension was spun at 10,000 xg in an MSE HS18 centrifuge for 3 hr. This length of time of spinning was necessary to obtain a clear yellow supernatant free from any cell debris.

(ii) Chromatography steps: (a) DEAE-11 cellulose columns.

The crude extract was diluted to a conductivity of 1.0 mmhos. The volume of the extract at this stage was usually of the order of five litres, and was yellowish-brown in colour due to the presence of a high percentage of soluble pigments, e.g. carotenoids and xanthophyls, and also the plastocyanin is in the reduced form. Plastocyanin could not be detected at this stage, after addition of potassium ferricyanide, due to the reducing contaminants present in the crude extract.

The extract was absorbed on to a DEAE-11 cellulose column (4 x 32 cm.), equilibrated with 0.01 M potassium phosphate pH 7.0. The column was washed with 0.05 M potassium phosphate pH 7.0, and then 0.05 M potassium phosphate pH 7.0 + 0.15 M KCL was put through the column. This level of salt elutes a high percentage of the contaminating pigments. Finally, the plastocyanin was eluted with 0.05 M potassium phosphate pH 7.0 + 0.5 M KCl, and detected by adding a drop of potassium ferricyanide (see 2A(i)) to each tube. The solution in the tubes containing plastocyanin turned blue, and these were pooled, and the solution dialysed overnight against distilled water at 4°C.

The plastocyanin solution was then absorbed on to a smaller DEAE-11 cellulose column (2.5 x 10 cm.), equilibrated with 0.01 M potassium phosphate pH 7.0, both as a further purification step and as a method of concentrating the solution. The column was washed with 0.02 M potassium phosphate pH 7.0, and plastocyanin eluted with a salt gradient between 0.1 M KCL and 0.4 M KCL, in 0.02 M potassium phosphate pH 7.0. The plastocyanin was detected as before, the solution in the tubes containing it pooled, and the solution dialysed as above.

The solution was freeze-dried.

(ii)b Gel filtration (on Sephadex G-50)

The protein was dissolved in 0.01 M potassium phosphate pH 7.0, in as small a volume as possible. The final volume was usually about 25 ml. The plastocyanin was oxidised with potassium ferricyanide, and a small volume of 5 mg./ml. solution of E-DNP lysine was added, to act as an external marker on the column.

Plastocyanin was eluted almost free from any contaminating pigments and other proteins, and it usually remained in the oxidised form during and after gel filtration. The solutions in the tubes containing plastocyanin were pooled.

(ii)c DEAE-52 cellulose columns

Chlorella pyrenoidosa plastocyanin

The blue plastocyanin solution was absorbed on to a DEAE-52 cellulose column (2.5 x 2.5 cm.), equilibrated with 0.01 M potassium phosphate pH 7.0. It absorbed as a single broad blue zone at the top of the column. The column was washed with 0.02 M potassium phosphate pH 7.0. On elution with 0.02 M potassium phosphate pH 7.0 + 0.13 M KCL, the broad blue zone separated into two blue bands, one band (designated C.pyrenoidosa 'A' plastocyanin) which appeared to be the major band, moved down the column, whilst the other slightly narrower band (designated C.pyrenoidosa 'B'plastocyanin), remained at the top

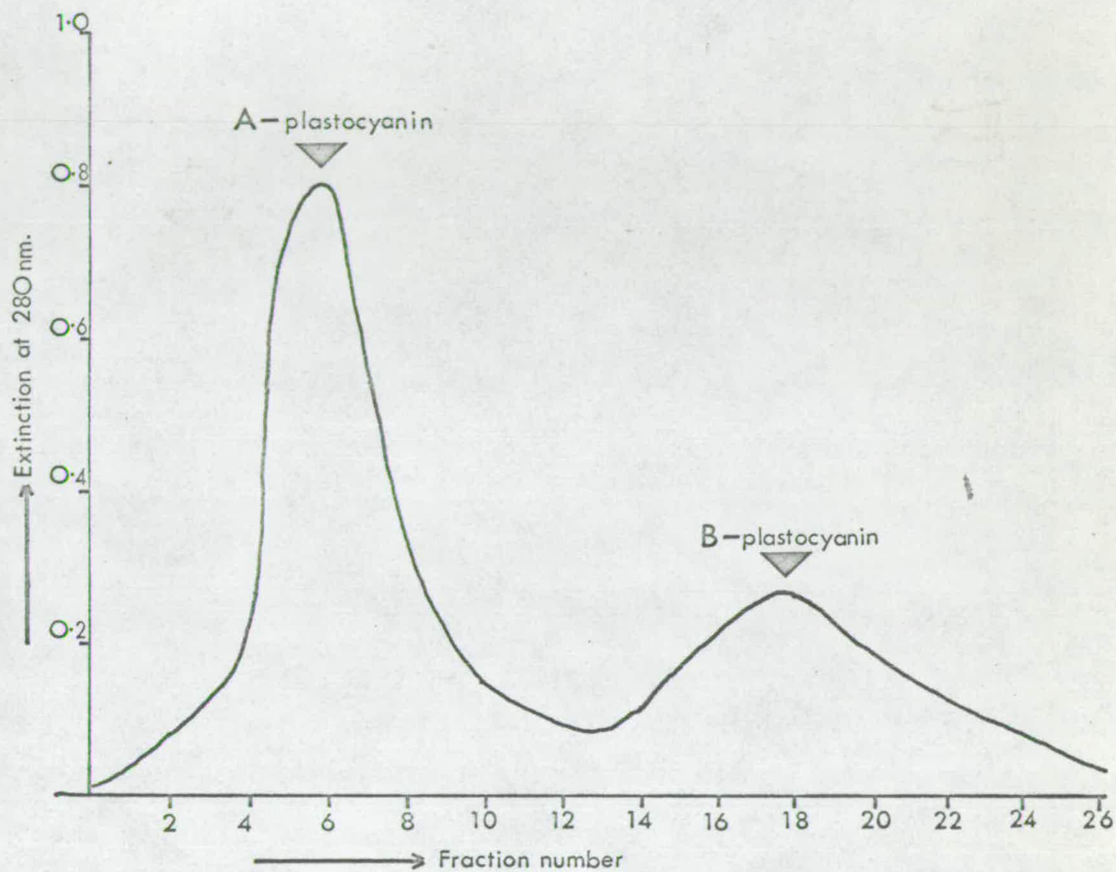


Figure 3(i) Separation of *Chlorella pyrenoidosa* A and B plastocyanins on DEAE-52 cellulose (2·5 x 2·5 cm. equilibrated in 0·01M potassium phosphate pH 7·0).

Stepwise elution: A plastocyanin with 0·13M KCL in 0·02M KPh. pH 7·0.

B plastocyanin with 0·18M KCL in 0·02M KPh. pH 7·0.

of the column. The second blue band was eluted with 0.02 M potassium phosphate pH 7.0 + 0.18 M KCL, and the separation of the two bands is shown in Figure 3(i). It was found that the eluates in the tubes between the two peaks turned blue on addition of potassium ferricyanide.

The eluates were pooled according to the 280 nm. absorption diagram, and the two solutions diluted to a conductivity of Immhos. They were then reabsorbed separately on to small DEAE-52 cellulose columns (2.5 x 2.5 cm.) and eluted with the appropriate concentration of KCl. In both cases, the two forms of plastocyanin eluted as a single major peak, separated from a very minor peak, corresponding to the other form. The two major peaks were pooled separately, dialysed and freeze-dried.

Scenedesmus obliquus plastocyanin

On absorption to a similar DEAE-52 cellulose column (2.5 x 2.5 cm., equilibrated in a 0.01 M potassium phosphate pH 7.0), the S.obliquus plastocyanin solution exhibited a similar phenomenon, to that observed with C.pyrenoidosa plastocyanin.

In this case, the best separation of the two blue bands was achieved using a salt gradient of 0.05 M KCl (100 ml.) to 0.2 M KCl (200 ml.) in 0.02 potassium phosphate pH 7.0. The two peaks were pooled separately, diluted, and rechromatographed on DEAE-52 cellulose columns (2.5 x 2.5 cm.). The eluates from the two major peaks were pooled separately, dialysed, and freeze-dried.

The blue band eluting first, from the original DEAE-52 cellulose column, was called S.obliquus 'A' plastocyanin, and that eluting second was called S.obliquus 'B' plastocyanin.

Chlorella fusca plastocyanin

On absorption of the blue plastocyanin solution to a DEAE-52 cellulose column (2.5 x 2.5 cm., equilibrated in 0.01 M potassium phosphate pH 7.0), a broad blue zone was obtained at the top of the column. Elution with a salt gradient of 0.02 M potassium phosphate pH 7.0 (100 ml.) to 0.02M potassium phosphaste pH 7.0 + 0.2 M KCl (200 ml.) separated the blue zone into an intense blue band that eluted between volume 85 to 165 ml. (fractions 17 to 33 on Figure 3(ii)), and a narrower brown band that eluted between 175 to 275 ml. (fractions 35 to 55 on Figure 3(ii)). The eluates in the latter group of fractions turned blue on addition of potassium ferricyanide.

It was shown later, both by running starch gels on the 'blue' and 'brown' bands and re-chromatography on further DEAE-52 cellulose columns that in fact the 'brown' band was largely the reduced form of the major blue plastocyanin. However, overloaded starch gels of the 'brown' band revealed a slight shadow of a band running faster than the main band. For this reason only plastocyanin purified in the 'blue' band was used for sequence studies, and lack of time made it impossible to further study the 'brown' band.

The plastocyanin was freeze-dried.

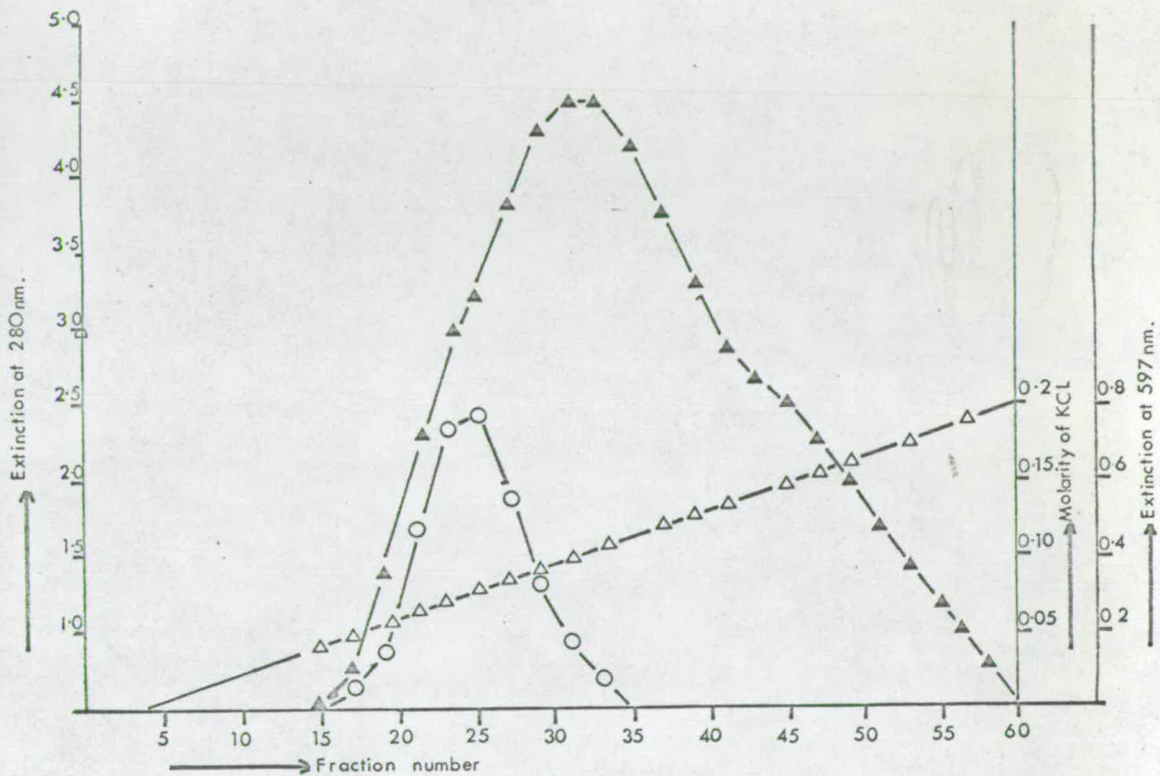


Figure 3(ii) Purification of Chlorella fusca plastocyanin on DEAE-52 cellulose (2·5 x 2·5cm.) equilibrated in 0·01M potassium phosphate pH 7·0.

Elution by a salt gradient (0·02M KPh. pH 7·0 to 0·02M KPh. pH 7·0 + 0·2 M KCL

Fraction volume = 5ml.

▲ Extinction $_{280\,nm.}$ ○ Extinction $_{597\,nm.}$ △ Molarity of KCL

TABLE 3(i) : Purification procedure for plastocyanin.

Chromatographic step	Volume (ml.)	Extinction at 280 nm.	Extinction at 597 nm.	Absorption index (E_{280nm}/E^{ox}_{597nm})	Yield of plastocyanin from 1 kg. dried cells (m.g.)
1st DEAE-11 cellulose column	1260	2.19	0.28	7.8	1209.5 (121 µmoles)
2nd DEAE-11 cellulose column	375	3.4	0.66	5.2	862.5 (86.3 µmoles)
Sephadex G-50 column	275	2.61	0.80	3.2	764.5 (76.5 µmoles)
DEAE-52 cellulose column	345	1.26	0.58	2.1	655.5 (65.6 µmoles)

The figures were not corrected for losses at each step.

D. Summary.

Table 3(i) summarises the purification procedure used for plastocyanin, with the yields of protein and the absorption index ($E_{280nm}:E^{ox}_{597nm}$) of the solution at each stage.

The figures given are representative of all the protein preparations, and were actually obtained from a C.fusca plastocyanin preparation using 1 kg. of dried cells. The final absorption indices were different for each plastocyanin (see Chapter 4).

CHAPTER 4

PROPERTIES OF PLASTOCYANIN

A. Criteria of purity of proteins

(i) Starch Gel Electrophoresis

The starch gels were prepared as described in Chapter 2 (section B(i)). The samples of protein run on the gels were all in the blue oxidised form. All the plastocyanins ran as single bands, except for S.obliquus "B" plastocyanin, which proved very difficult to purify completely free from contamination with S.obliquus "A" plastocyanin. The gels were always examined for non-plastocyanin protein, by staining with Amido Black, but in no case was any contaminating protein observed, even on heavily overloaded gels. Figure 4(i) shows a photograph of the separation of C.pyrenoidosa "A" and "B" plastocyanins on starch gel electrophoresis, and figure 4(ii) shows a diagram of the distances moved by all the plastocyanins relative to one another. Table 4(i) gives a list of mobilities of each plastocyanin, relative to the distance moved by a marker (XCFF) from the origin, of the gel. The mobilities were recorded from the same gel.

TABLE 4(i) Mobilities of the algal plastocyanins on Starch Gel Electrophoresis

Species of plastocyanin	Mobility (with respect to the distance moved by the blue dye Xylene cyanol-FF- from the origin).
C.pyrenoidosa "A"	1.29 ± 0.02
C.pyrenoidosa "B"	1.48
S.obliquus "A"	1.30
S.obliquus "B"	1.55
C.fusca	1.45

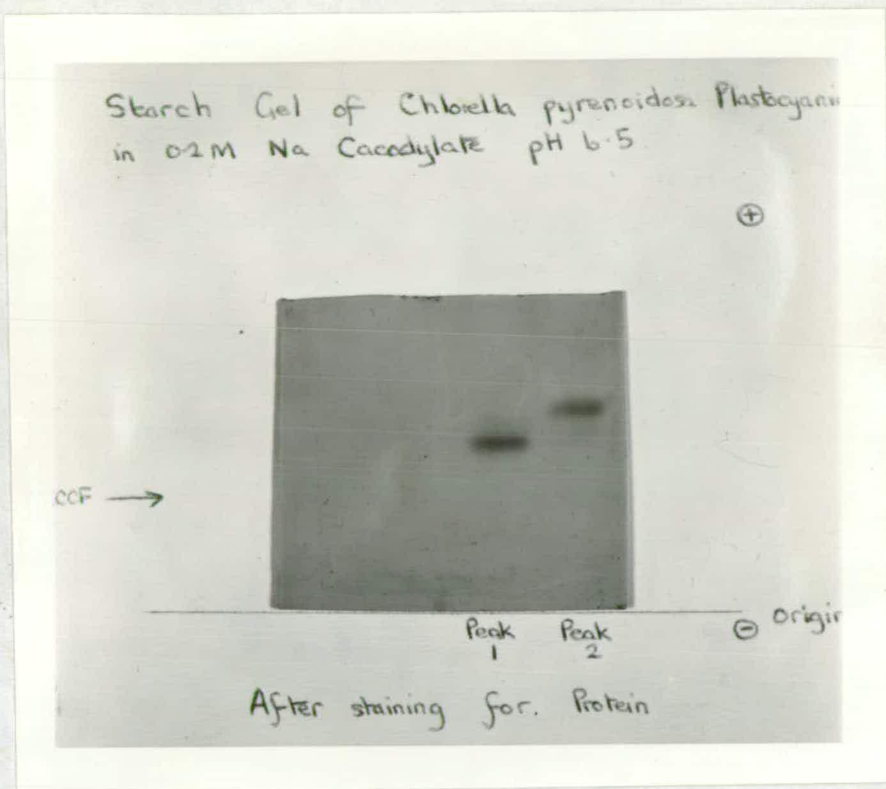


Figure 4(i) Starch gel pattern of C.pyrenoidosa A(peak I) and B(peak 2) plastocyanins.

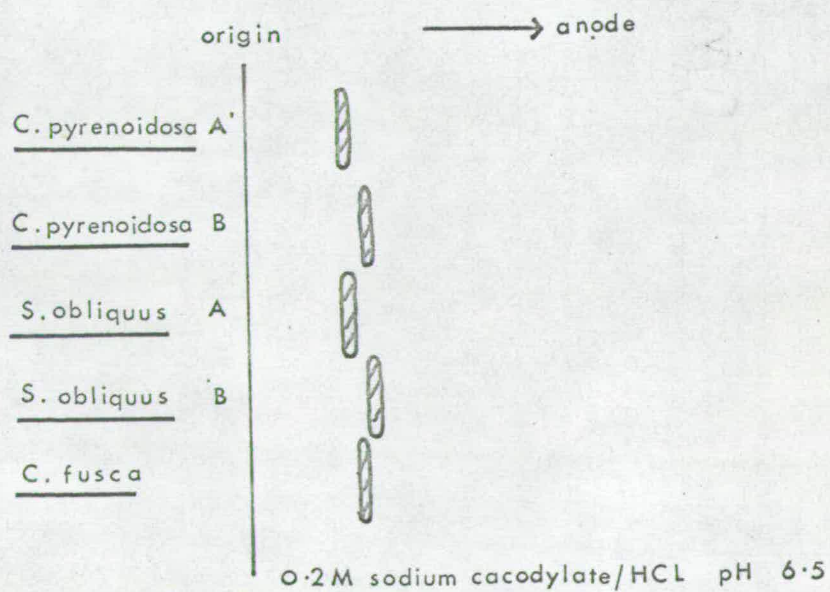


Figure 4(ii) Starch gel pattern of algal plastocyanins.

Discussion of starch gel pattern

The comparative mobilities for the algal plastocyanins show that C.pyrenoidosa 'A' and S.obliquus 'A' plastocyanins, have an almost identical net acidic charge at pH 6.5. C.fusca plastocyanin has a net acidic charge at pH 6.5, which is closer to that of the "B"-type plastocyanins than the "A"-type plastocyanins. The reproducibility of the mobilities for different preparations and for separate gels was consistently within $\pm$ 0.2.

(ii) Absorption index

The basis for this method as a criterion of purity is described in Chapter 2.

The highest purity obtained for spinach plastocyanin (Katoh et al. 1962), corresponded to an absorption index (E_{278nm}/E^{ox}_{597nm}) of 0.80, and for french bean plastocyanin (Wells, 1966) an absorption index of 1.1 is given. In contrast the absorption at 278 nm. for Chlamydomonas reinhardi plastocyanin (Gorman et al. 1966) is greater than the relative absorption at 597 nm., giving an absorption index for the 'purest' material equal to 2.0.

The absorption indices obtained for the algal plastocyanins studied in this thesis are listed in Table 4(ii).

TABLE 4(ii) Absorption indices of algal plastocyanins

Species of plastocyanin	Absorption index E_{278nm}/E^{ox}_{597nm}
C.pyrenoidosa A	2.4
C.pyrenoidosa B	1.9
S.obliquus A	1.1
S.obliquus B	1.3
C.fusca	2.9

Discussion

A disadvantage of using this spectral method as a criterion of purity is that one is ignoring the possibility of a percentage of the protein being present as the apo-protein. It is highly probable that the protein might lose some copper during the preparation.

Thus, the absorption at 597 nm., even making sure all the copper is in the oxidised form, may not be a true measure of all the plastocyanin present. Also, the absorption at 278 nm. is due to the three aromatic amino acids tyrosine, phenylalanine, and tryptophan, and so the extent of absorption in this region is directly related to the aromatic content of the protein being measured.

However, for chemical studies, such as those described in this thesis it is not important whether one starts with the apo- or native protein. The only requirement necessary for amino acid sequence studies is to ensure that there is not any non-plastocyanin

material present in the preparation, and this has been satisfied (Section 4A(i)).

B. Absorption spectra

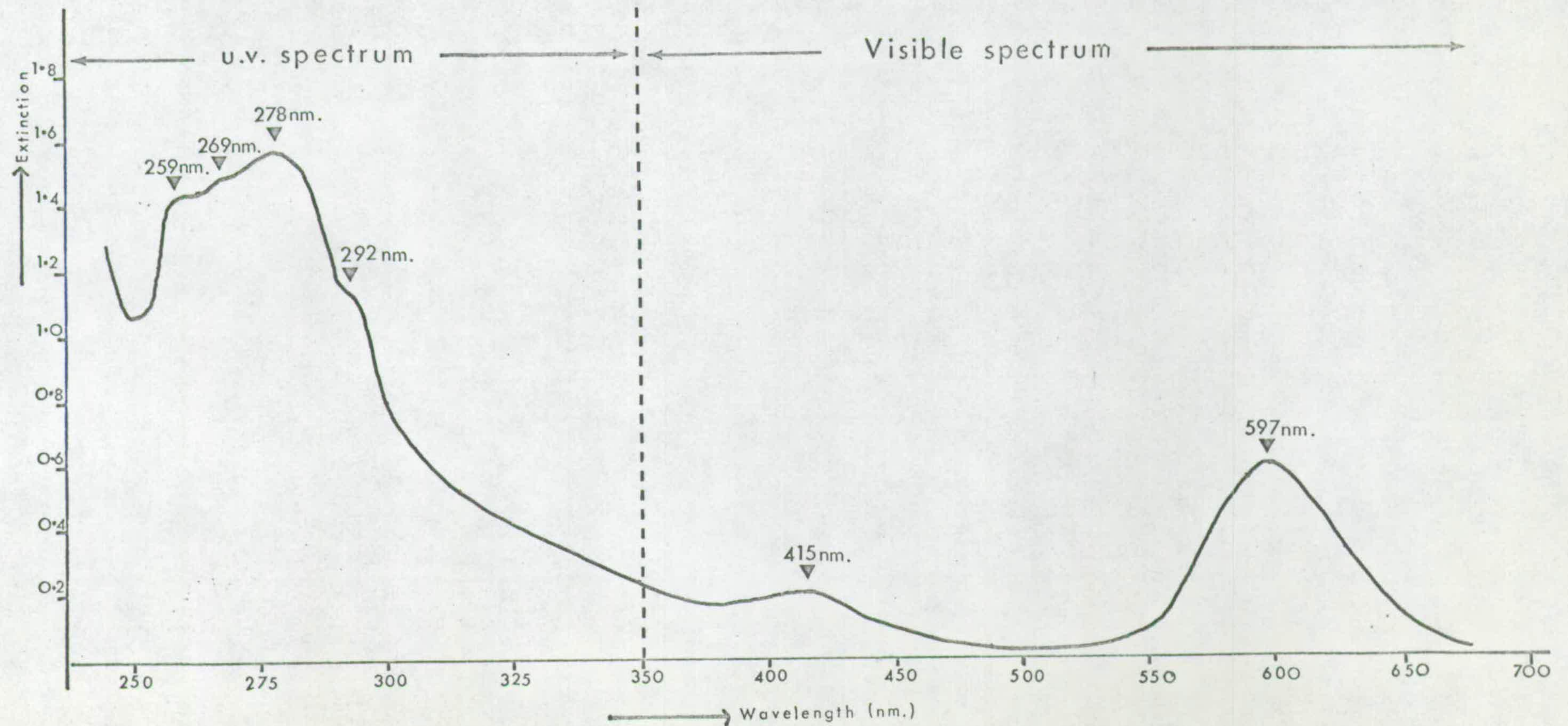
All the algal plastocyanins exhibited very similar absorption spectra both in the visible and ultraviolet regions.

Figure 4(iii) shows a typical spectrum obtained for all the algal plastocyanins. In the visible spectrum a single major peak was observed at around 597 nm., with fully oxidised protein, and which disappeared on reduction (section 2B(ii)). Most preparations showed a trace of a Soret band at around 415 nm. The amount of cytochrome present was calculated to be always less than 0.005 mole/mole, showing that all but a trace of the proteins very closely associated with plastocyanin during the preparation are removed.

As seen in figure 4(iii) the u.v. spectrum is fairly complex. The principal peak is at 278 nm., with two subsidiary peaks at round 259 nm. and 269 nm., corresponding to the fine structure bands of phenylalanine. Also, the distinct shoulder obtained at around 292 nm., corresponds to tryptophan, and was found in the spectra of all the plastocyanins excepting C.pyrenoidosa 'B' plastocyanin.

At alkaline pH (0.1 M NaOH) the u.v. spectrum changed. Two subsidiary peaks were observed in the region of 284 nm. and 288 nm., corresponding to tyrosine and tryptophan, respectively.

Figure 4(iii) Spectrum of plastocyanin.
1cm. path length.



Calculation of tyrosine and tryptophan content for C.fusca plastocyanin.

Using the known molar extinction coefficients of tryptophan (5,500) and tyrosine (1,200) at 280nm. (Beaven & Holiday, 1952), the values observed for E_{280nm} for 1 mg. (weighed protein)/ml. are consistent with the protein containing 1 residue of tryptophan and 4 residues of tyrosine per molecule.

Similar calculations were done for C.pyrenoidosa 'A' and S.obliquus A and in both cases the observed E_{280nm} was consistent with the proteins containing 1 residue of tryptophan per molecule.

C. Carbohydrate content

The method of Devor (1950) was used (section 2C(i)), for the detection of pentose and hexose. The reducing sugar content of the plastocyanin preparations is shown in Table 4(iii).

The amino-sugar (galactosamine and glucosamine) content was calculated from the amino acid analyser. Taking the colour yields to be about the same as for amino acids, the amount of amino-sugar present, was calculated very approximately, to be equivalent to 2 residues per molecule.

TABLE 4(iii) Reducing sugar content of plastocyanin preparations

Species of plastocyanin	mole sugar/mole protein
C.pyrenoidosa "A"	2.8
C.pyrenoidosa "B"	7.1
S.obliquus "A"	9.7
C.fusca	6.3

I have no evidence to indicate whether the sugar in the protein preparations is covalently-bound to the proteins or is an artefact of the preparation procedure.

D. Amino acid analyses.

10 mg./ml. solutions of each of the plastocyanins were prepared and 8 x 0.1 ml. samples of each were taken. These were hydrolysed in duplicate for 12, 24, 48 and 96 hr., as described in section 2 D(viii).

Internal standards (Walsh et al. 1962) were used to allow correction for pipetting errors in sample application. Each analysis was then normalised to the same amount of protein, by summing a few consistent amino acids and adjusting this figure to the same for each analysis. The best value for threonine and serine was found by extrapolation to zero time of hydrolysis, and the best value for valine and isoleucine by extrapolation to infinite time. The recoveries of the other amino acids did not vary systematically with time of hydrolysis, and the best values were obtained by averaging the values obtained for the different times of hydrolysis.

TABLE 4: (iv) Amino acid composition of plastocyanins (Experimental details are given in the text)

	C.pyrenoidosa 'A'			C.pyrenoidosa 'B'			S.Obliquus 'A'			C.fusca		
	*Best value (μmoles)	Residues/Mol. Best* value	Residues/Mol. From sequence	Best value (μmoles)	Residues/Mol. Best value		Best value (μmoles)	Residues/Mol Best value		Best value (μmoles)	Residues/Mol. Best value	Residues/Mol. From sequence
Gly	0.1088	10.9	11	0.1713	11.7	12	0.2618	10.9		0.2756	11.5	11
Ala	0.1371	13.7	14	0.1422	9.7	9	0.3349	14.0		0.2833	11.8	11
Val	0.0923	9.2	11	0.1457	9.9	10	0.2400	10.0		0.2220	9.2	9
Leu	0.0400	4.0	4	0.0602	4.1	4	0.1000	4.2		0.1000	4.1	4
Ile	0.0307	3.1	3	0.0752	5.2	4	0.0740	3.1		0.0760	3.1	3
Ser	0.0507	5.1	4	0.0771	5.3	5	0.1220	5.1		0.1820	7.6	7
Thr	0.0560	5.6	6	0.0790	5.4	6	0.1440	6.0		0.1840	7.7	9
Asp	0.1355	13.5	14	0.2192	15.0	14	0.3286	13.7		0.2695	11.2	11
Glu	0.0736	7.4	7	0.1415	9.7	10	0.1884	7.9		0.2565	10.7	10
Phe	0.0501	5.0	5	0.0882	6.0	6	0.1235	5.1		0.1212	5.1	6
Tyr	0.0468	4.7	7	0.0456	3.1	3	0.1222	5.1		0.0885	3.7	4
Trp	-	-		-	-	0	-	0.4		-	0.5	1
$CySO_3H$	0.0100	1.0	1	0.0176	1.0	1	0.0208	0.9		0.0218	0.9	1
$MetSO_2$	0.0100	1.0	1	0.0132	0.9	1	0.0196	0.8		0.0245	1.0	1
Pro	0.0500	5.0	5	0.0700	4.8	5	0.1185	5.0		0.1218	5.1	5
Lys	0.0396	3.9	4	0.0452	3.1	3	0.0923	3.9		0.0964	4.0	4
His	0.0295	2.9	3	0.0307	2.1	2	0.0675	2.8		0.0714	3.0	3
Arg	0.0000	0	0	0.0282	1.9	2	0.000	0		0.0000	0	0

TABLE 4(iv) Legend

*Best value: the amount of amino acid recovered (μmoles/mg. protein by weight).
These values are the average of 12 hr., 24 hr., 48 hr., and 96 hr., recoveries for amino acids other than serine and threonine, which are extrapolated to zero time of hydrolysis, and valine and isoleucine, which are extrapolated to infinite time of hydrolysis.

As cysteic acid and methionine sulphone in oxidised samples of protein.

Not estimated for C.pyrenoidosa 'A' and 'B' proteins, but estimated colorimetrically for S.obliquus 'A' and C.fusca proteins. See the text.

The amidated residues are included in these figures.

Determined in unoxidised protein samples.
The values are not corrected for the water content of the preparation.

From the sequence information the number of lysine and histidine residues per molecule for C.pyrenoidosa "A", S.obliquus "A", and C.fusca plastocyanins is 4 and 3, respectively, and for C.pyrenoidosa "B" 3 and 2, respectively. The quantities of these amino acids were summed, for each analysis, and divided by 7 in the case of the first group of plastocyanins and by 5 for C.fusca plastocyanin. Each best value was then divided by the resultant figure in both cases, giving the number of residues of each amino acid per molecule of protein. The amino acid compositions of the plastocyanin from each of the algal species is shown in Table 4(iv).

B. Tryptophan content

The u.v. spectra of all the plastocyanins excepting C.pyrenoidosa 'B' plastocyanin showed a distinct shoulder at 292 nm. (at pH 7.0) suggestive of tryptophan (section 4B).

Attempts to locate tryptophan-containing peptides by looking for fluorescent peptides, under u.v. light, after digestion of oxidised protein, did not prove fruitful. However, if the paper was stained with Ehrlich reagent (section 2D(iv)) mauve spots, characteristic of the presence of tryptophan, showed up. The presence of tryptophan in plastocyanin was studied further in the case of C.fusca plastocyanin and this will be described in Chapter 7.

A tryptophan determination on the whole protein (method of Harrison et al. (1961) was performed on *C.fusca* and *S.obliquus* 'A' plastocyanins. A figure of 0.40 residue of tryptophan/molecule was calculated for *S.obliquus* 'A' plastocyanin, and a figure of 0.52 residue tryptophan/molecule for *C.fusca* plastocyanin. Lack of protein made it impossible to do the determination on *C.pyrenoidosa* plastocyanins.

F. N- and C-terminal studies : (i) N-terminal studies.

The N-terminus of each plastocyanin was determined by the Dansyl Chloride method. 1 mg. of each protein was subjected to the Dansyl Edman procedure. In all cases two cycles of the method gave satisfactory results, but subsequent cycles yielded nothing. This was almost certainly due to the fact that the proteins became insoluble in the reaction mixtures, as the ε-amino groups of lysine residues reacted with PITC. Actually, that it worked at all, due to the low lysine content of the proteins was quite surprising.

Single residues were obtained for the N-terminus and 2nd and 3rd residues of each protein. The results are shown in Table 4(v).

TABLE 4(v) N-terminal amino acid residues of the algal plastocyanins

Species of plastocyanin	N-terminus	2nd residue	3rd residue
C.pyrenoidosa A	Ala	Asx	Val
C.pyrenoidosa B	Ala	Asx	Val
S.obliquus A	Ala	Asx	Val
C.fusca	Asx	Val	Thr

(ii) C-terminal studies

The C-terminal region was identified from a cyanogen bromide digest on 20 mg. of each protein, as it was known from the amino acid analyses that each protein contained a single methionine residue.

The digests were fractionated by gel filtration through Sephadex G-25 and a diagram of the extinction of the effluent at 254 nm. is shown in Figure 4(iv). The fractions containing the low molecular weight material was separated by High Voltage Paper Electrophoresis at pH 6.5 and further purified at pH 3.5. For each plastocyanin one small peptide was purified, in good yield, and on paper electrophoresis at pH 2.0 (section 2D(iv), it was shown to contain no homoserine/homoserine lactone. It should, therefore, be the COOH-terminal peptide, unless non-specific hydrolysis of peptide bonds has taken place in the acid reaction mixture. The amino acid composition of the C-terminal peptides (after 96 hr. hydrolysis) is given in Table 4(vi).

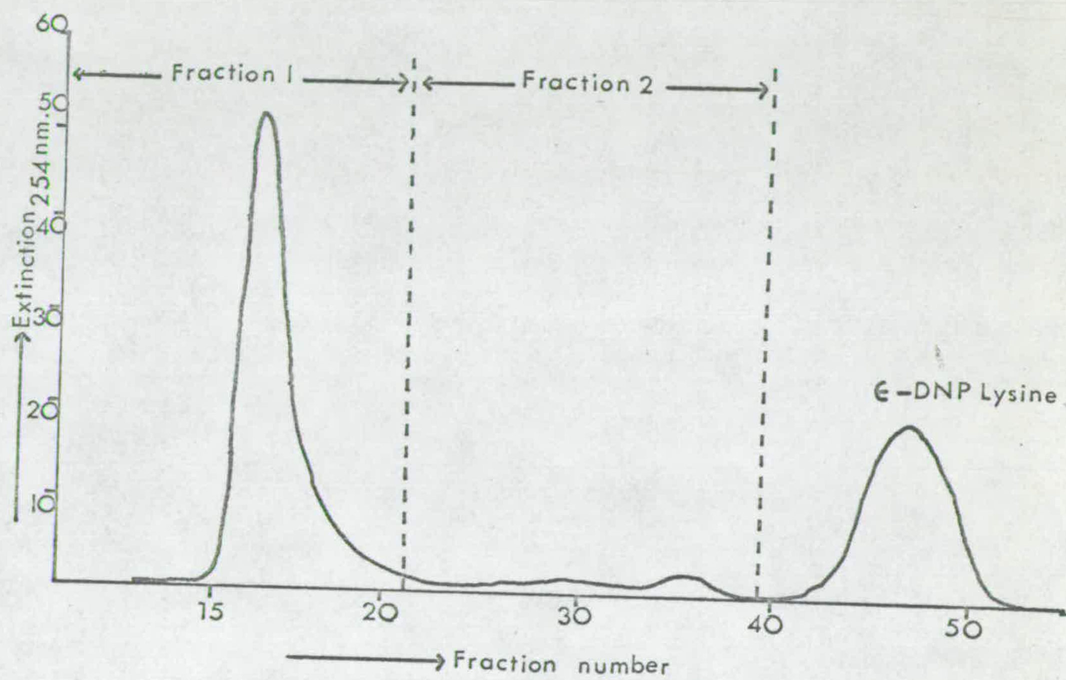


Figure 4(iv) Gel filtration (Sephadex G-25) of CNBR digest.

$E_{254nm.}$ is shown in arbitrary units.

Fraction size = 2·9ml.

TABLE 4(vi): Amino acid composition of C-terminal peptides.

Species of plastocyanin	Amino acid composition	Electrophoretic mobility		*Yield	Comments
		pH 6.5	pH 3.5		
C.pyrenoidosa 'A'	Lys(0.9),Glx(1.0),Gly(1.0),Val(2.5),Ile(0.7)	0.35	0.57	22	Gln not Glu
C.pyrenoidosa 'B'	Arg(1.2),Asx(1.0),Thr(0.9),Gly(1.1),Val(2.0),Ile(1.0)	0.40	0.59	21	Asn not Asp
S.obliquus 'A'	Lys(1.1),Glx(1.0),Gly(0.9),Val(3.2),Ile(0.9)	0.43	0.56	14	Gln not Glu
C.fusca	Lys(0.9),Thr(1.8),Glx(1.0),Gly(1.0), Val(1.0),Ile(1.0)	0.39	0.54	18	Gln not Glu

*Percentage μmole peptide/μmole protein digested.

The N-termini of the peptides were determined by the DNS-chloride method and the sequences by the Dansyl Edman procedure. The sequences are shown in Table 4(vii). From the presumed specificity of cyanogen bromide, a methionine residue must be expected on the N-terminal side of the C-terminal peptides.

TABLE 4 (viii) Sequence of C-terminal peptides

Species of plastocyanin	Sequence
C.pyrenoidosa 'A'	Val-Gly-Lys-Ile-Val-Val-Gln
C.pyrenoidosa 'B'	Val-Gly-Arg-Ile-Thr-Val-Asn
S.obliquus 'A'	Val-Gly-Lys-Ile-Val-Val-Gln
C.fusca	Lys-Gly-Thr-Ile-Thr-Val-Gln

CHAPTER 5

DIGESTIONS PERFORMED ON C.PYRENOIDOSA 'A' PLASTOCYANIN

Introduction

Two major digestions were performed on C.pyrenoidosa 'A' protein; a tryptic digest on 150 mg. (about 15 μmoles) and a chymotryptic digest on 70 mg. (about 7 μmoles) of protein. A further chymotryptic digest was done on 20 mg. of protein to isolate a single peptide (CB4), which gave a very weak ninhydrin reaction, and was overlooked in the major digest.

A. Tryptic digest

(i) Conditions for digestion

The conditions for digestion are described in Chapter 2. The protein was denatured by oxidation with performic acid, freeze-dried, and dissolved in 0.2 M ammonium acetate pH 8.5. Trypsin was added and the digestion mixture shaken at 37°C for 2½ hr., diluted, and freeze-dried.

(ii) Fractionation of tryptic digest

The digest was dissolved in 5% formic acid and a reddish-brown precipitate was spun off. The peptides in the soluble fraction were separated by gel filtration through Sephadex G-25, followed by paper electrophoresis at pH 6.5. Figure 5(i) shows the peptide 'map' obtained by electrophoresis of a portion of each fraction from the

gel filtration separation. The peptides were pooled into fractions T1 to T4, as shown, taking care as far as was possible to separate peptides of similar mobilities and to avoid splitting peptides in half. Each fraction was freeze-dried.

Fraction T1 was separated by gel filtration through Sephadex G-50. When samples of each fraction were separated on paper electrophoresis at pH 6.5, the only peptide material present, after staining the paper with ninhydrin, was stuck at the origin. This fraction was presumably a mixture of undigested protein, trypsin, and very large fragments of the protein, and no satisfactory work has been done on this fraction. However, a sample was quantitated on the amino acid analyser, and fraction 1 was found to comprise about 45% (by weight) of the total digest.

Fractions T2, T3 and T4 were separated by paper electrophoresis at pH 6.5 on 20 cm. strips (except T4 which was separated on a 15 cm. strip). The peptides were purified by paper electrophoresis at pH 3.5.

(iii) Summary of tryptic peptides.

Table 5(i) gives a list of the tryptic peptides purified, with their electrophoretic mobilities at pH 6.5 and pH 3.5. Most of the mobilities were obtained from the preparative separations, and will not be very accurate.

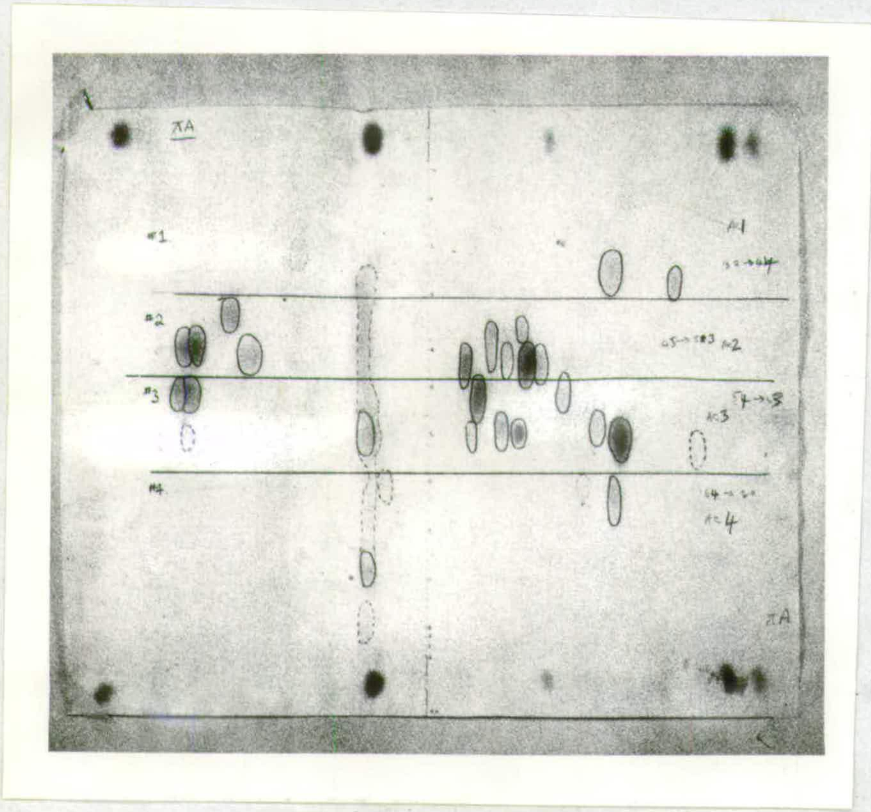


Figure 5(v) HVPE at pH 6·5 of chymotryptic peptides from *C.pyrenoidosa* A plastocyanin.

70
60
55
50
45
40
35
30
25
20
15
12
T4
T3
T2
T1
Lys
Arg
Markers
Neutrals
Glu
Asp
Fraction number

ε-DNP-LYS eluted after 204ml.(fraction 70)

Figure 5(i) Separation of tryptic peptides of *C.pyrenoidosa* A plastocyanin by Gel filtration followed by paper electrophoresis at pH 6·5.

TABLE 5(i): Purification procedure for tryptic peptides

Peptide	Sephadex fraction	*Purification procedure	Electrophoretic mobilities pH 6.5(m)	pH 3.5(m)
ATN1	3	63	0	+0.30
ATN2	2	63	0	+0.25
ATN3	2	63	0	+0.44
ATB1	2	63	+0.57	+0.78
ATB2	2	63	+0.48	+0.67
ATA1	2	63	-0.31	+0.11
ATA2	2	63	-0.22	+0.16
ATA3	2	63	-0.53	+0.06
ATA4	3	63	-0.18	+0.21
ATA5	2	63	-0.05	+0.29
ATA6	2	63	-0.05	+0.38

* Paper electrophoresis at pH 6.5 : 6

Paper electrophoresis at pH 2.5 : 3

TABLE 5(ii): Tryptic peptides formed from C.pyrenoidosa 'A' plastocyanin.

	TN1	TN2	TN3	TB1	TB2	TA1	TA2	TA3	TA4	TA5	TA6
His		1.0							0.5		1.0
Lys			1.0	1.1	0.9				0.6	0.8	1.0
Arg											
Asp		2.8	0.3	1.1		1.6	1.2	4.7	1.3	1.1	2.9
Thr		1.9	1.8		1.7				1.0	*1.6	1.2
Ser		1.1	0.3			1.0	0.8			1.2	2.0
Glu	1.3		1.1			1.1		1.0	3.0	0.9	1.3
Pro		0.9	1.1			1.0		0.9	1.0	0.9	0.5
Gly		2.2	0.3			1.9	1.7	1.0	4.7	1.8	1.5
Ala		2.0	1.2	1.0		3.0	1.8	2.8	2.1	2.8	2.8
Val	1.95	1.1	*1.5	0.9	1.0	1.1	0.3	*2.4	1.1	*1.6	
$MetSO_2$									0.7		
$CySO_3H$									0.9		
Ile	0.7	0.2	0.9		1.0			*0.7		0.9	
Leu		0.2				1.6	*1.3			1.7	1.9
Tyr		1.0	1.4			0.4			2.1	1.5	1.4
Phe		0.9	1.0			0.8		0.9	1.6	0.9	
N-terminus	Ile	Ala	Val	Ala	Thr	Leu	Leu	Asx	Phe	Leu	Ala
Yield	3.9	2.3	1.8	4.0	2.4	4.0	3.5	1.2	1.2	1.6	1.2

*Low yield due to partial hydrolysis of stable peptide bond

Peptide given 96 hr. hydrolysis.

Contaminating amino acids (amounts less than 0.2 mole/mole peptide not shown).

Table 5(ii) gives the amino acid compositions, N-termini, and yields of the peptides. Yields of peptides are expressed as percentages (μmole peptide/μmole of the protein digested). No attempts have been made to correct these yields for losses that inevitably occur during the purification, such as material lost on marker strips and peptide that is not eluted from the paper. Contaminating amino acids are given in Table 5(ii), when they are present in an amount of 0.2 mole/mole peptide or more. The quantitative determination of tyrosine in the large tryptic peptides was not very reliable, (see peptides TN3, TA1, TA5 and TA6). An explanation for this is that the protein was probably purified as the hydrochloride so that on oxidation (with performic acid) the tyrosines present were converted to a mixture of mono- and di-chloride tyrosines (Hirs, 1967). This was undesirable (and precautions should have been taken to remove the halide), since the column constant on the amino acid analyser for mono- and di-chloride tyrosine would not be the same as for tyrosine. The amount of tyrosine present was quantitated by adding in the amount of chloride-tyrosine (calculated from the column constant for tyrosine) to the amount of unmodified tyrosine present. Additional evidence from the peptides formed by chymotryptic digestion of the large tryptic peptides was used in the final deduction of tyrosine content.

(iv) Peptides formed from tryptic digestion of plastocyanin

The evidence for the sequence of the small peptides and the partial sequence of the large peptides is presented.

Peptide TB1

From the mobility of the peptide at pH 6.5, it is evident that the second residue is asparagine. The sequence, determined by the Dansyl Edman method is:

Ala-Asn-Val-Lys

Peptide TN1

Since the peptide runs in the neutral band at pH 6.5, the C-terminal amino acid must be glutamine. The sequence, determined by the Dansyl Edman method is:

Ile-Val-Val-Gln

Peptide TA4

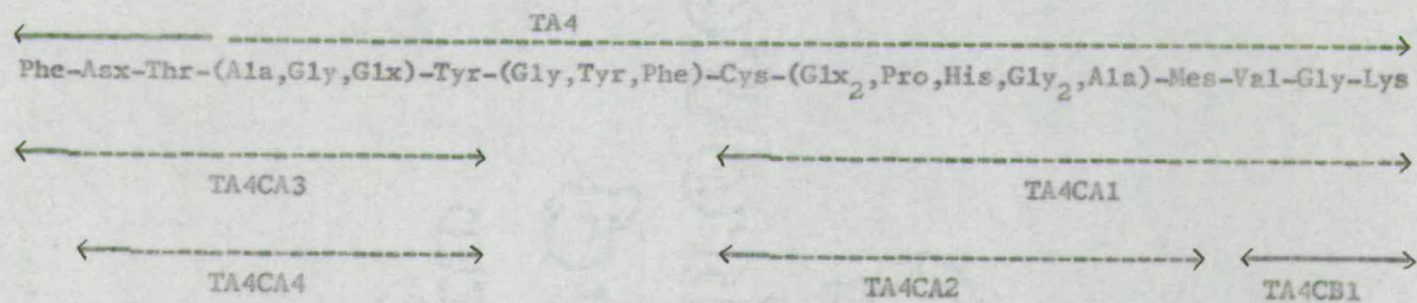
The N-terminal sequence of peptide TA4 was examined by the Dansyl Edman procedure.

When digested with chymotrypsin, peptide TA4 formed the peptides shown in Table 5(iii). Quantitative analyses and N-termini were obtained for all the chymotryptic peptides.

TABLE 5 (iii): Peptides formed from chymotrypsin digestion of peptide TA4.

	TA4CA1	TA4CA2	TA4CA3	TA4CA4	TA4CA5
His	0.9	1.0			
Lys	0.8				0.8
Arg					
$CySO_3H$	0.9	0.9			
$MetSO_2$	0.7	0.9			
Asp			1.1	1.0	
Thr			1.0	1.0	
Ser					
Glu	1.9	2.2	1.2	1.3	
Pro	1.1	1.0			
Gly	2.6	2.0	1.2	1.4	1.2
Ala	1.0	1.1	1.1	1.2	
Val	0.7				1.0
Ile					
Leu					
Tyr			1.0	0.9	
Phe			1.0		
N-terminus	$CySO_3H$	$CySO_3H$	Phe	Asx	Val
m	-0.01	-0.30	-0.50	-0.60	+0.66
m'	+0.36	+0.10	+0.11	+0.02	+0.80

Figure 5(ii) Partial sequence of peptide TA4



Definition of symbols is given in the text.

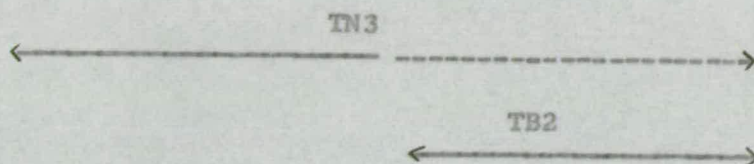
The evidence for the location of the chymotryptic peptides within the sequence of peptide TA4 was derived largely from amino acid compositions and from a knowledge of which bonds in the sequence are slightly susceptible to chymotrypsin. It was assumed that lysine is the C-terminus of peptide. TA4, from the presumed specificity of trypsin. Figure 5(ii) gives the partial sequence of peptide TA4. The peptide covering the region (Gly, Tyr, Phe) in the sequence was not found. It was most likely lost during the purification of the neutral band. Evidence for the insertion of the (Gly, Tyr, Phe) sequence was obtained from the chymotrypsin digestion of the whole protein (see chymotryptic peptide CA6).

Peptides TN3, TB2, TA5, TA1 and TA2.

Peptide TB2 was sequenced by the Dansyl Edman procedure, the sequence being:

Thr-Val-Thr-Ile-Lys

The N-terminal sequence of peptide TN3 was determined by the Dansyl Edman procedure, leaving a peptide that corresponded in amino acid composition to peptide TB2. Now the tyrosine content, from the amino acid analysis of peptide TN3, gives a figure of 1.4 residues. I propose that this figure suggests the presence of 2 residues of tyrosine (not 1), and since I have no evidence for the presence of tyrosine in peptide TB2, I postulate a Tyr-Tyr sequence as shown below:

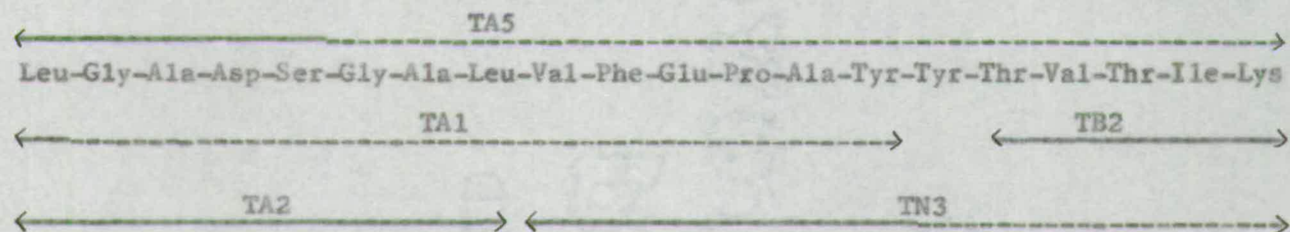


Peptide TB2 must have been formed by the trypsin possessing some chymotryptic-like activity, resulting in cleavage of the Tyr-Thr bond.

The sequence of peptide TA2 was determined by the Dansyl Edman method. From its mobility (m) at pH 6.5, the fourth residue must be aspartic acid.

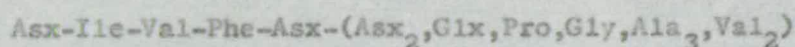
The N-terminal sequence of peptide TA5 was examined by the Dansyl Edman procedure, and it was evident that the sequence of peptide TA5 is comprised of peptides TA2 plus TN3. A tyrosine content of 1.5 residues (from the amino acid analysis) for peptide TA5 is in agreement with the presence of 2 residues of tyrosine in this region of the protein. Peptide TA1 has an amino acid composition that fits into the sequence of TA5. Thus, combining peptides TA1, TA2, TN3 and TB2 the postulated sequence of peptide TA5 is shown in Figure 5(iii).

Figure 5(iii): Postulated amino acid sequence of peptide TA5



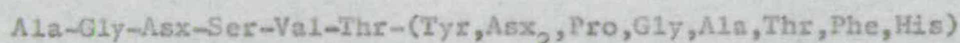
Peptide TA3

The N-terminal sequence was determined by the Dansyl Edman method, and the partial sequence of peptide TA3 is:


Asx-Ile-Val-Phe-Asx-(Asx_2,Glx,Pro,Gly,Ala_3,Val_2)

Peptide TN2

The N-terminal sequence of peptide TN2 was examined by the Dansyl Edman method and the partial sequence is shown below:


Ala-Gly-Asx-Ser-Val-Thr-(Tyr,Asx_2,Pro,Gly,Ala,Thr,Phe,His)

Peptide TA6

The N-terminal sequence of peptide TA6 was examined by the Dansyl Edman procedure.

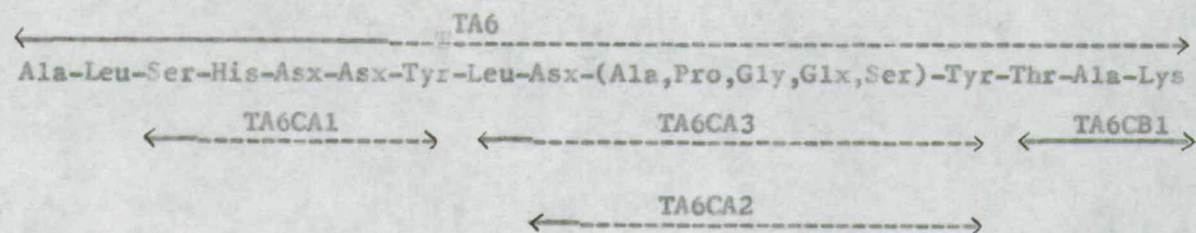
When digested with chymotrypsin, peptide TA6 formed the peptides shown in Table 5(iv). Qualitative analyses of peptides TA6, CB1 and TA6, CA3 were obtained, and quantitative analyses of peptides TA6 CA1 and TA6 CA2. Peptide TA6 CB1 was sequenced by the Dansyl Edman procedure, and it is assumed to constitute the C-terminal sequence of peptide TA6, since it has lysine at its C-terminus. The evidence for the location of the chymotryptic peptides in the sequence of TA6 was obtained from amino acid compositions and from a knowledge of the bonds in the sequence which are susceptible to chymotrypsin (see partial sequence of TA6). Figure 5(iv) gives the partial sequence of peptide TA6.

TABLE 5(iv): Peptides purified from chymotrypsin digestion of peptide TA6

Peptide	m	Amino acid composition	N-terminus
*TA6CB1	+0.63	Thr(+++),Ala(+++),Lys(+++)	Thr
*TA6CA1	-0.25	Ser,His,Asx_2,Tyr	Ser
TA6CA2	-0.18	Asx,Ser,Glx,Pro,Gly,Ala,Tyr	Asx
*TA6CA3	-0.13	Asx(++),Ser(+),Glx(++),Gly(++), Ala(++),Leu(++) Pro(++),Tyr(+)	Leu

*Qualitative analysis only. Definition of symbols is given in the text.

Figure 5(iv) Partial sequence of peptide TA6



The tyrosine content of peptide TA6 was calculated (from the amino acid analysis) as 1.4 residues, but the formation of peptides TA6 CA1 and TA6 CA3 from chymotrypsin digestion of TA6, shows that there are 2 residues of tyrosine. The glycine content is high, but I suspect that some 20% is contaminating amino acid, since glycine is a common amino acid present in chromatography paper.

B. Chymotryptic digest

The conditions used for the digestion were the same as those described for the trypsin digestion.

(i) Fractionation of chymotryptic digest

The digest was dissolved in 5% formic acid and the peptides separated by gel filtration through Sephadex G-25, followed by paper electrophoresis at pH 6.5. Figure 5(v) shows the peptide 'map' obtained by electrophoresis at pH 6.5, of a portion of each fraction from the gel filtration. The peptides were pooled into fractions C1 to C4, as shown, taking care to separate peptides of similar mobilities and to avoid splitting peptides in half. Each fraction was freeze dried.

Fractions C1 to C4 were separated by paper electrophoresis at pH 6.5 on 15 cm. strips. The peptides were further purified on (BAWP) paper chromatography and paper electrophoresis at pH 3.5.

(ii) Summary of chymotryptic peptides

Table 5(v) gives a list of the chymotryptic peptides, the purification procedure, with their electrophoretic mobilities at pH 6.5 and pH 3.5. Table 5(vi) gives the amino acid compositions, N-termini and yields (expressed as a percentage µmole/µmole protein digested) of the peptides. The yields were not corrected for any losses inevitably incurred during the purification of the peptides. Impurities are shown in Table 5(vi), when the amino acids are present in an amount 0.2 mole/mole peptide or more. Not all the expected chymotryptic peptides (from the presumed specificity of chymotrypsin) were recovered from the digest and these omissions will be pointed out in the discussion of the sequence.

(iii) Peptides formed from chymotrypsin digestion of plastocyanin

The evidence for the sequence of the chymotryptic peptides is presented.

Peptide CB1

The peptide was sequenced by the Dansyl Edman procedure and the sequence is:

Thr-Ala-Lys-Phe

Peptide CB2

From the electrophoretic mobility at pH 6.5 the second residue must be asparigine. The sequence, determined by the Dansyl-Edman method, is:

Ala-Asn-Val-Lys-Leu

TABLE 5(v): Purification procedure for chymotryptic peptides

Peptide	Sephadex fraction	*Purification procedure	Electrophoretic mobilities	
			pH 6.5	pH 3.5
CA1	1	63	-0.56	-
CA2	3	63	-0.57	-
CA3	3	63	-0.25	-
CA4	3	6B3	-0.28	+0.38
CA5	3	6B3	-0.16	+0.21
CA6	2	6B3	-0.17	+0.15
CA7	1	63	-0.74	-
CA8	2	6B3	-0.26	+0.15
CB1	3	63	+0.50	-
CB2	2	632	+0.48	+0.60
CB3	2	6B3	+0.47	+0.57
CB4	2	6B3	+0.35	+0.47
CB5	1	6B3	+0.25	+0.45
CN1	4	6B3	0	+0.31
CN2	2	6B3	0	+0.34

*Paper electrophoresis at pH 6.5 : 6

" " at pH 3.5 : 3

". " at pH 2.0 : 2

BAWP Chromatography : B

TABLE 5(vi): Chymotryptic peptides formed from *C.pyrenoidosa* 'A' plastocyanin.

	CA1	CA2	CA3	CA4	CA5	CA6	CA7	CA8	CB1	CB2	CB3	CB4	CB5	CN1	CN2	CN3
His				1.2		0.9		1.0				0.8	0.9			
Lys									++	++	0.9		1.0		0.9	0.5
Arg																
Asp	4.6	0.9	0.7	2.3	1.0		3.4			++		1.8	3.3		1.0	0.9
Thr		0.9							++			0.5	2.2		1.1	0.8
Ser			1.1	1.0	1.1							≠0.3	1.3		1.1	1.2
Glu	1.0	1.1	1.0		1.0	2.2	0.9	2.2			1.0	≠0.2				
Pro	0.8		1.0		1.0	1.0	1.1	1.0				0.9	1.0			
Gly	0.9	1.1	1.1		1.1	2.8	0.9	2.1			1.0	1.2	2.4	0.9	1.0	1.3
Ala	4.0	1.1	1.2		1.1	1.1	3.8	1.0	++	++		1.1	2.3		1.0	1.0
Val	*2.5						1.9			++	*2.4	*0.3	1.5		1.0	*0.7
$MetSO_2$						0.8		1.0								
$CySO_3H$						1.0		1.2								
Ile	*0.6										*0.7		0.9		0.8	
Leu	1.0				0.9		1.0			++						
Tyr		1.0	0.8	1.0	0.5	0.9							0.9	1.1	1.0	1.0
Phe	1.0					1.1			++			0.9	1.2			
N-terminus	Asx	Asx	Asx	Ser	Leu	Gly	Asx	$CySO_3H$	Thr	Ala	Val	Val	Ile	Gly	Ile	Lys
Yield	8.6	15.3	2.9	6.7	8.6	2.9	9.6	10.0	9.0	10.0	4.6	2.3	0.6	1.3	2.2	0.6

*Low yield due to partial hydrolysis of stable peptide bond.

≠Contaminating amino acids (amounts 0.2 mole/mole peptide not shown).

Peptide CA4

Peptide CA4 was sequenced by the Dansyl Edman procedure. Samples of peptide were removed after the 1st, 2nd and 3rd cycles of PTC degradation, and the electrophoretic mobilities at pH 6.5 were measured. Table 5(vii) gives the mobilities at pH 6.5 of the degradative products of peptide CA4, after the 1st, 2nd and 3rd cycles of PTC degradation.

TABLE 5 (vii) PTC degradation products of peptide CA4

PTC degradation cycle	Peptide Sequence	Electrophoretic Mobility at pH 6.5	Comments
0	Ser-His-Asx-Asx-Tyr	-0.44	
1	His-Asx-Asx-Tyr	-0.41	
2	Asx-Asx-Tyr	-0.66	
3	Asx-Tyr	-0.39	Both Asps (not Asns)

The sequence of peptide CA5 was shown to be:

Ser-His-Asp-Asp-Tyr

Peptide CA2

CPA released tyrosine (+++). The peptide was sequenced by the Dansyl Edman procedure. After the 1st cycle of PTC degradation, a fraction of the remaining part of the peptide (corresponding to the amino acid composition (Thr, Glx, Gly, Ala, Tyr)), was run out

on paper electrophoresis at pH 6.5. It retained a net acidic charge but it was greatly reduced when compared with that of the whole peptide. Thus the N-terminus of peptide CA2 is aspartic acid and the 5th residue is glutamic acid.

The sequence is:

Asp-Thr-Ala-Gly-Glu-Tyr

Peptides CA3 and CA5

CPA released tyrosine (++) from both peptides. The Dansyl Edman procedure gave the same sequence for peptides CA3 and CA5, except that peptide CA5 is one residue longer, having leucine at its N-terminus. The electrophoretic mobilities (at pH 6.5) of the corresponding peptides obtained after the 1st and 2nd cycles of PTC degradation of peptide CA5 were determined and are shown in Table 5(viii).

TABLE 5(viii) PTC degradation products of peptide CA5

PTC degradation cycle	Peptide sequence	Electrophoretic Mobility at pH 6.5	Comments
0	Leu-Asx-Ala-Pro-Gly-Glx-Ser-Tyr	-0.16	
1	Asx-Ala-Pro-Gly-Glx-Ser-Tyr	-0.22	
2	Ala-Pro-Gly-Glx-Ser-Tyr	-0.21	Asn and Glu

Thus, the sequence of peptide CA5 is:

Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr

CA5 (←———→ spanning Leu to Tyr)

CA3 (←———→ spanning Asn to Tyr)

Peptides CN1, CA6 and CA8

Peptide CN1 is a dipeptide with the sequence:

Gly-Tyr

CPA released methionine (+++), glycine (++), alanine (++) and glutamine (++) from peptide CA8. The peptide was sequenced by the Dansyl Edman procedure. After the 1st cycle of PTC degradation (i.e. after removal of the cysteic acid residue) the remaining peptide (with amino acid composition, (Glx_2, Pro, His, Gly_2, Ala, Mes)) ran in the neutral band on paper electrophoresis at pH 6.5. Thus, in the sequence of peptide CA8, shown below, the second and fifth residues are both glutamine residues.

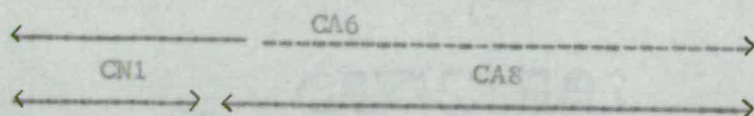
Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes

Peptide CA6 contains a single residue of cysteic acid. Since the amino acid analyses show that there is only one cysteine residue in the entire molecule, peptides CA6 and CA8 must come from the same region of the protein. The N-terminal sequence of peptide CA6 was determined by the Dansyl Edman method and the partial sequence is shown below:

Gly-Tyr-Phe-Cys-(Glx_2,Pro,His,Gly_2,Ala,Met)

Peptide CN1 is assumed to have come from the N-terminal region of peptide CA6, since no other Gly-Tyr sequences were found in the protein. Thus, the sequence around the cysteic acid residue is:

Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes



Peptide CB3

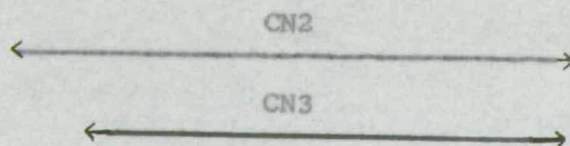
CPA released glutamine (+++), valine (+++), and isoleucine (++) from peptide CB3. The sequence was determined by the Dansyl Edman procedure and was shown to be:

Val-Gly-Lys-Ile-Val-Val-Gln

Peptides CB5, CN2, CN3 and CB4

CPA released tyrosine (+++) from peptides CN2 and CN3. The Dansyl Edman procedure gave the same sequence for peptides CN2 and CN3,except that CN2 is one residue longer possessing an isoleucine as its N-terminal residue. Since both peptides run in the neutral band on paper electrophoresis at pH 6.5, the 5th residue (for peptide CN2) and the 4th residue (for peptide CN3) is aspartic acid. The sequence of peptide CN2 is:

Ile-Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr

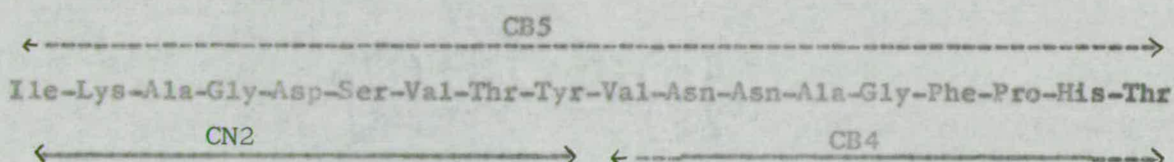


The cleavage of the Ile-Lys bond, giving rise to the formation of peptide CN3, is inconsistent with the presumed specificity of chymotrypsin. However, peptide CN3 was recovered in very low yield from the digest, so that cleavage of the Ile-Lys bond must have occurred very slowly.

Since peptide CB4 is slightly basic on paper electrophoresis at pH 6.5, it is assumed that the basicity is due to the histidine residue, and that the two residues of aspartic acid are amidated in the peptide. Now as shown in Table 5(vi) peptide CB4 was not purified completely free of contaminating amino acids, namely, serine, glutamic acid, and possibly valine. The determination of the N-terminus (DNS-chloride method) gave two spots on the TLC plate, corresponding to Val (+++) and Asx (++). The level at which Val was recovered (0.3 residue), after hydrolysis of peptide CB4, could suggest that it was a contaminating amino acid. However, after examination of peptide CB4 by the Dansyl Edman method, I propose the following sequence for CB4 (shown below), but it must be pointed out that this is weak sequence data based entirely on the reliability of the Dansyl Edman method.

The amino acid composition of peptide CB5 is equal to the sum of peptides CN2 plus CB4. Again the Valine content is fairly low (1.5 residues) for there to be 2 residues of valine, but it is difficult to estimate the level of impurity of each amino acid in the peptide. The N-terminus of peptide CB5 is isoleucine, so

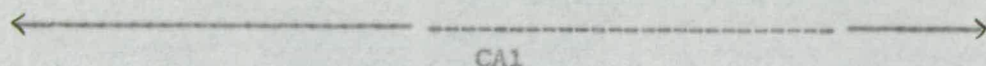
that the order of peptides in this region of the protein is that peptide CN2 is N-terminal to CB4, as shown below:



The formation of peptides CB5 and CB4 represents non-specific cleavage of the protein (after threonine), by chymotrypsin. Low recoveries of these peptides made further characterisation of them impossible. CB4 is homologous with CB1 from C.pyrenoidosa 'B' and CB2 from C.fusca plastocyanins, which both have N-terminal valine.

Peptides CA1 and CA7

CPA released leucine (+++) and alanine (+) from both peptides. The N-terminal sequence of peptide CA1 was determined by the Dansyl Edman procedure, and the partial sequence of CA1 is:



It was evident from a consideration of the amino acid compositions of peptides CA1 and CA7, that they must be derived from the same region in the sequence of plastocyanin. Peptide CA7 has been formed by chymotryptic cleavage at the C-terminal side of Phe (residue 4) in CA1 sequence. However, a peptide with the sequence Asx-Ile-Val-Phe was not isolated.

TABLE 5(ix): PTC degradation products of peptide CA7

PTC degradation cycle	Electrophoretic mobility (m) Predicted 2 Amides	1 Amide	Observed	Proposed Sequence	N-terminus (DNS)
0	-0.79	-0.94	-0.74	Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu	Asx
2	-0.46	-0.69	-0.66	Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu	Asx
3	-0.28	-0.48	-0.49	Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu	Glx
4	0	-0.25	-0.23	Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu	Asx
5		0	0	Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu	Ala

Experimental details are given in the text. The predicted mobilities were derived from graphs (electrophoretic mobility (m) against molecular weight of peptide) published by Offord (1966).

The sequence of peptide CA7 was determined by the Dansyl Edman procedure. The electrophoretic mobilities (at pH 6.5) of the PTC degradative products of peptide CA7 are given in Table (ix), and they are compared with the predicted mobilities (from Offord (1966)). The electrophoretic mobility of peptide CA1 is -0.56, which is less acidic than peptide CA7 (m = -0.74). Thus, the N-terminus of peptide CA1 is Asn and not ASp. This is in agreement with the predicted mobility (m = -0.60) for a peptide of the size of CA1 and containing 3 amide groups. The sequence data on peptide CA7 may be considered weak, since it is based on the reliability of the Dansyl Edman procedure.

The sequence covered by peptides CA1 and CA7 is:

Asn-Ile-Val-Phe-Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Gly-Ala-Asn-Ala-Leu

←―――――――――――― CA1 ――――――――――――→

←― ― ―――― CA7 ―――― ― ―→

C. Deduction of sequence

The overall sequence was deduced from the tryptic peptides, a list of which is given in Table 5(x). Some of the sequences were established from the peptides themselves, and this evidence has been noted, whereas others required a consideration of the chymotryptic peptides (see Table 5(xi)). These tryptic peptides are discussed.

TABLE 5(x) : List of tryptic peptides

TN1: Ile-Val-Val-Gln

TN2: Ala-Gly-Asp-Ser-Val-Thr-Tyr-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr

...... CN2 → ←-- CB4 →

TN3: Val-Phe-Glu-Pro-Ala-Tyr-Tyr-Thr-Val-Thr-Ile-Lys

TB1: Ala-Asn-Val-Lys

TB2: Thr-Val-Thr-Ile-Lys

TA1: Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu-Val-Phe-Glu-Pro-Ala-Tyr

← TA2 → ← TN3 ....

TA2: Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu

TA3: Asn-Ile-Val-Phe-Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn

← CA1 ....

← --- CA7 ....

TABLE 5(x): - cont'd

TA4: Phe-Asp-Thr-Ala-Gly-Glu-Tyr-Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Lys

TA2: Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu-Val-Phe-Glu-Pro-Ala-Tyr--Tyr-Thr-Val-Thr-Ile-Lys

TA6: Ala-Leu-Ser-His-Asp-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Thr-Ala-Lys

TABLE 5(xi): List of chymotryptic peptides.

CA1: Asn-Ile-Val-Phe-Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu

CA2: Asp-Thr-Ala-Gly-Glu-Tyr

CA3: Asn-Ala-Pro-Gly-Glu-Ser-Tyr

CA4: Ser-His-Asp-Asp-Tyr

CA5: Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr

CA6: Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes

CA7: Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu

CA8: Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes

CB2: Thr-Ala-Lys-Phe

CB2: Ala-Asn-Val-Lys-Leu

CB3: Val-Gly-Lys-Ile-Val-Val-Gln

CB4: Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr

CB5: Ile-Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr

CN1: Gly-Tyr

CN2: Ile-Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr

CN3: Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr

It is evident from the partial sequence of TN2, that the C-terminal region corresponds in amino acid composition to CB4, except that the recovery of valine, as in the case of CB4, is extremely low. However, although the evidence for the presence of 2 valine residues in TN2 and 1 in CB4 is inconclusive, it is thought that by homology of CB4 with peptides CB1 from C.pyrenoidosa 'B' and CB2 from C.fusca plastocyanins, that this is the case. The N-terminal sequence of TN2 corresponds to residues 3-9 of CN2.

The sequence of TA1 can be deduced from TA2 and the N-terminal region (residues 1-6) of TN3. Peptide TA2 begins at the N-terminus of TA1.

The sequence of TA3 can be deduced from aconsideration of CA1 and CA7. TA3 is two residues shorter (at the C-terminus) of CA7.

TA4 was partially sequenced (Figure 5(ii)), fromthe peptides formed by chymotrypsin digestion of it. The total sequence can be deduced from a consideration of peptides CA2, CA6 and CA8.

The sequence of TA5 was established from the trypsin digest. Chymotryptic peptides were expected from this region of the protein, but they were not recovered. They would be acidic peptides, and must have been lost in the complex mixture of acidic peptides formed in the digestion.

The partial sequence of TA6 is given in Figure 5(iv). The entire sequence can be deduced from a consideration of CA4 and CA5.

I shall begin the deduction of sequence of C.pyrenoidosa 'A' plastocyanin from the N-terminus.

The N-terminus was determined by the DNS-chloride method and the 2nd and 3rd residues by the Dansyl Edman method. TB1 and CB2 have their first three residues in common with the first three residues of the whole protein. They must both be at the N-terminus (see Figure 5C(i)).

TA5 has leucine as its N-terminus and is the only candidate for residue 5 in the protein, since TA1 and TA2, which also have N-terminal leucine are derived from TA5. The sequence between residues 4 to 24 is shown in Figure 5C(ii).

CN2 has the N-terminal sequence Ile-Lys-Ala..., which fits into the sequence at residue 23 (Figure 5C(ii)). TN2 has alanine as its N-terminal residue, and it is evident from the N-terminal sequence of TN2, that it fits into the sequence of the protein at the C-terminus of TA5. The sequence between residues 23 to 40 is shown in Figure 5C(iii).

No overlapping sequence has been found for residues 40 to 41, since both trypsin and chymotrypsin cleave the protein chain at the C-terminal side of threonine (residue 40). This is

inconsistent with the presumed specificity of both enzymes. However, TA3 must extend from residue 41 to 56, and this will become more obvious later on in the discussion. CA1 has the same amino acid sequence as TA3, except that CA1 has an extra Ala-Leu sequence at the C-terminus. TA6 has an Ala-Leu as its N-terminal sequence. This evidence orders TA6 at the C-terminus of TA3. The sequence between residues 41 to 58 is shown in Figure 5C(iv).

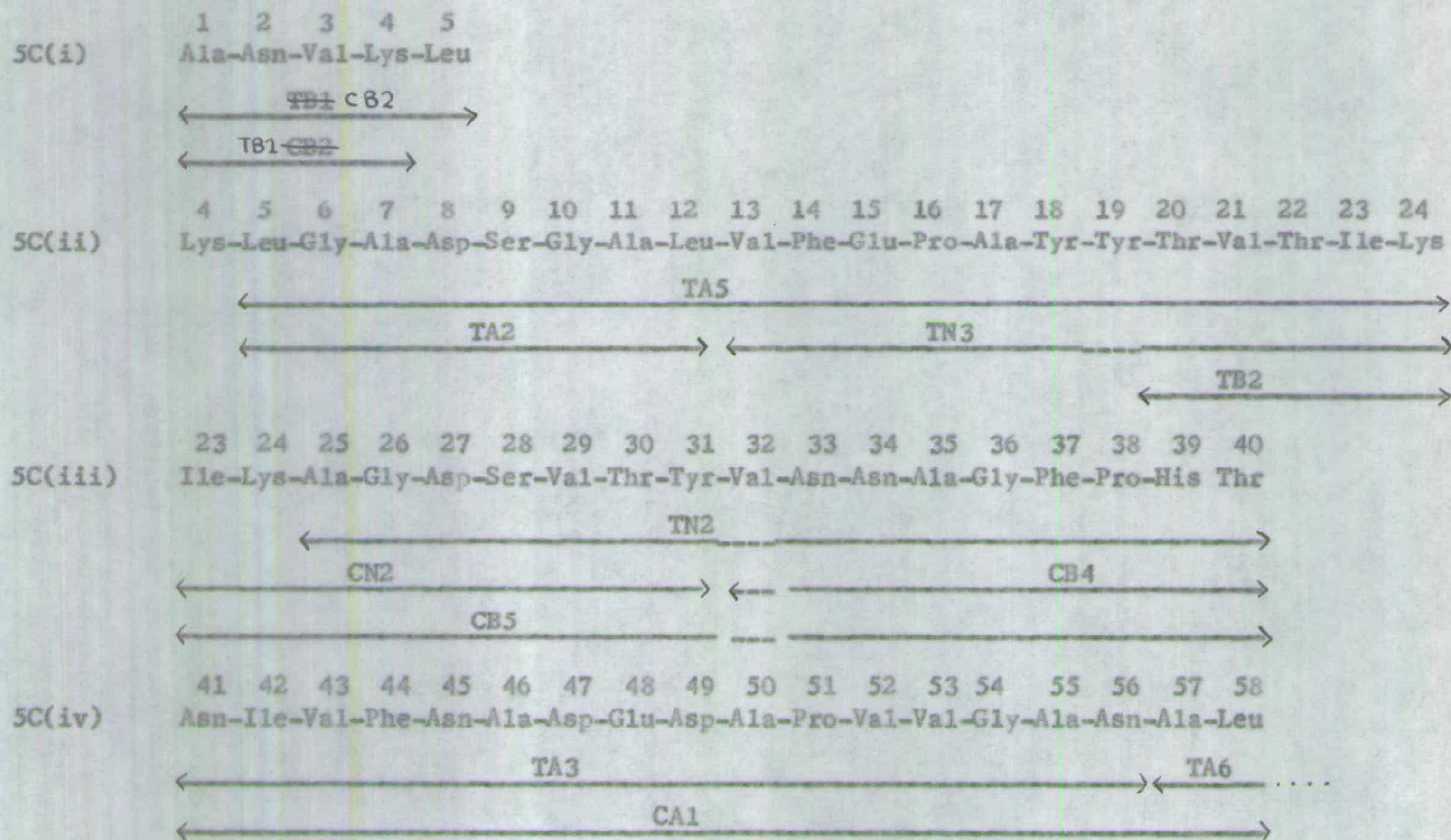
The C-terminal sequence of TA6 is -Thr-Ala-Lys. CB1 has the amino acid sequence Thr-Ala-Lys-Phe, and provides the overlapping sequence for TA6 and TA4, since TA4 has phenylalanine as its N-terminal residue, and it is the only candidate for this position in the protein. Figure 5C(v) gives the sequence between residues 57 to 75.

The C-terminal sequence of TA4 is -Val-Gly-Lys. The C-terminal peptide, of the whole protein, was isolated from a cyanogen bromide digest. It was identified as the C-terminal peptide, because it contained no trace of homoserine/homoserine lactone, and it has the sequence:

Val-Gly-Lys-Ile-Val-Val-Gln

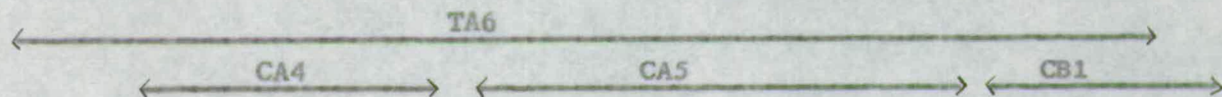
CB3 has the same sequence. From this evidence TN1 must be the C-terminal tryptic peptide. Figure 5C(vi) shows the sequence of the protein between residues 75 to 100.

Figures 5C(i) to (vi) : Overlapping sequences for tryptic peptides.



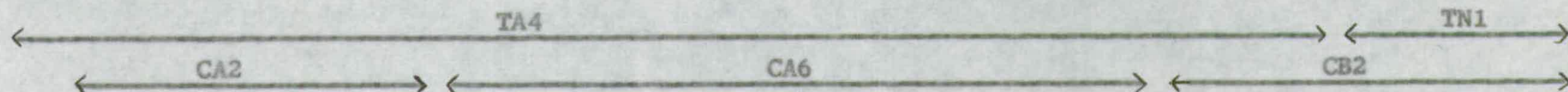
57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75

5C(v) Ala-Leu-Ser-His-Asp-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Thr-Ala-Lys-Phe



75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

5C(vi) Phe-Asp-Thr-Ala-Gly-Glu-Tyr-Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Lys-Ile-Val-Val-Gln



Discussion of sequence

The amino acid sequence of C.pyrenoidosa 'A' plastocyanin is given in Figures 5C (i) to (vi), where only sufficient peptides to establish the overall sequence are shown. All the peptides purified have been accounted for in the sequence. Overlapping sequences for the tryptic peptides have been found, excepting between TN2 and TA3 (residues 40 to 41 in the sequence). That trypsin cleaved the protein at all between TN2 and TA3 (a Thr-Asn bond) is rather surprising, and this will be referred to again in the final discussion. Non-specific cleavage by trypsin also occurred between TA3 and TA6; this time an Asn-Ala bond (residues 56-57) was cleaved. This type of cleavage is not chymotrypsin-like, and it can only be suggested that it represents a secondary activity of trypsin. Non-specific cleavage of the protein chain by chymotrypsin occurred in the formation of CN2, which meant the cleavage of a Thr-Ile bond (residues 22-23).

No chymotryptic peptides were isolated between residues 6 to 22 in the sequence, but fortunately tryptic peptides, formed by chymotryptic-like cleavages, were isolated in this region of the protein.

The most difficult region of the protein to sequence was between residues 32 to 56, where the sequence data is based entirely on the reliability of the Dansyl Edman method. The method worked well on peptides CA1 and CA7 (residues 40 to 58 in sequence), although

this is a long region to sequence. Special care was taken at each cycle to remove as completely as possible the phenylthionydanto in (PTH)-amino acid(s) from the reaction mixture. However, there was a fairly high level of impurity in CB4 (residues 32-40 in sequence) which made the identification of the newly exposed N-terminus, after each cycle, more difficult than in the case of a pure peptide. Attempts to split TA3 into smaller peptides proved fruitless, since the peptides formed were in too low a yield to purify satisfactorily.

As mentioned earlier the yields of tyrosine, in the large tryptic peptides, were low, probably due to halogenation of tyrosine during oxidation. But only in one case (residue 19) is there any doubt as to the placement of the tyrosine residues in the sequence. A Tyr-Tyr sequence (residues 18 and 19) in peptide TN3 can be readily deduced from the sequence data on this peptide.

The postulated sequence is in satisfactory agreement with the total amino acid composition of the protein (see Chapter 4), with three discrepancies. The first is in the serine content, but this is a corrected figure in the amino acid analyses, for decomposition of serine with increased time of hydrolysis, so to be out by one residue is not unprecedented. Two residues of tyrosine are unaccounted for in the total amino acid composition compared with the postulated sequence. This is rather surprising, especially since the tyrosine content was determined from unoxidised protein

samples, and no appreciable amount of mono- or di-chloro-tyrosine was observed on the amino acid analyses. The figure for valine (from total composition) is 2 residues below that determined in the sequence. This may be explained by incomplete hydrolysis of valine bonds, during the time allowed for hydrolysis.

CHAPTER 6

DIGESTIONS PERFORMED ON C.PYRENOIDOSA 'B' PLASTOCYANIN

Introduction

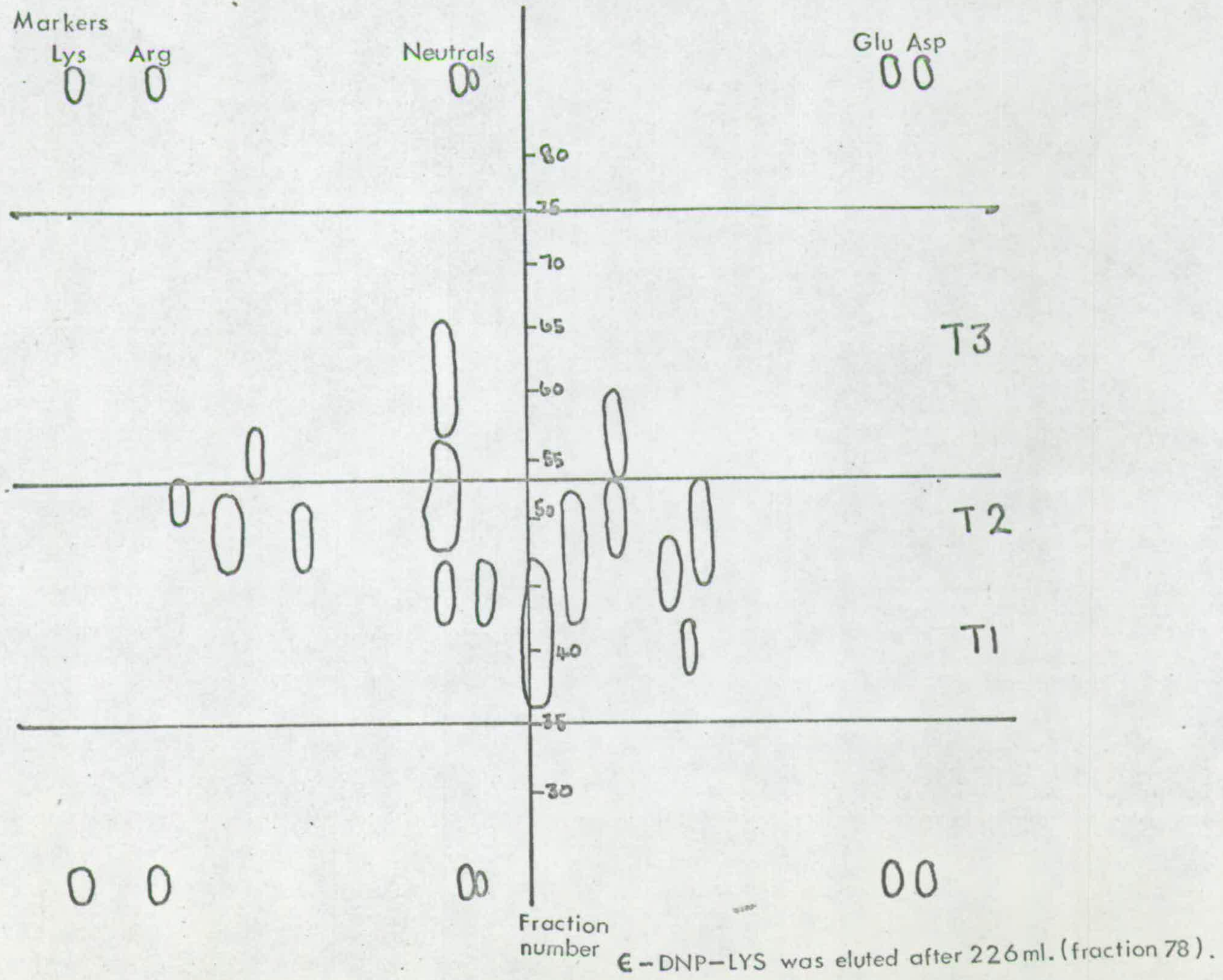
Two major digestions were performed on C.pyrenoidosa 'B' plastocyanin; a tryptic digest on 100 mg. (about 10 μmoles) and a chymotryptic digest on 70 mg. (about 7 μmoles) of protein. The conditions for oxidation and digestion of the protein were the same as those described for C.pyrenoidosa 'A' plastocyanin (Chapter 5).

A. Tryptic digest

(i) Fractionation of tryptic digest

The digest was dissolved in 5% formic acid and separated by gel filtration through Sephadex G-25, followed by paper electrophoresis at pH 6.5. Figure 6(i) shows the peptide 'map' obtained by paper electrophoresis of a portion of each fraction from the gel filtration separation. The peptides were pooled into fractions T1 to T3, as shown, taking care to separate peptides of similar mobilities and to avoid splitting peptides in half. Each fraction was freeze dried.

Fractions T1 to T3 were separated on paper electrophoresis at pH 6.5, on 20 cm. strips. The peptides were further purified on (BAWP) paper chromatography and paper electrophoresis at pH 3.5.



ε-DNP-LYS was eluted after 226 ml. (fraction 78).

Figure 6(i) Separation of tryptic peptides of C.pyrenoidosa B plastocyanin by Gel filtration followed by paper electrophoresis at pH 6·5.

(ii) Summary of tryptic peptides

This was not a very satisfactory digest, since several peptides were obtained as a result of chymotryptic-like cleavage of the protein. This made purification of the acidic peptides, which comprised the majority of the peptides formed, very difficult. Only those peptides that were pure and obtained in sufficient yield for sequence work are discussed.

Table 6(i) gives a list of the tryptic peptides purified, with their electrophoretic mobilities at pH 6.5 and pH 3.5. All the mobilities were obtained from the preparative separations, and will not be very accurate.

The amino acid compositions, N-termini and yields of the peptides are presented in Table 6(ii). The yields (expressed as percentages: µmole/µmole protein digested) are not corrected for any losses that must have occurred during the purification of the peptides.

(iii) Peptides formed from trypsin digestion.

The evidence for the sequence of the small tryptic peptides and the partial sequence of the large peptides is discussed.

Peptide T1

The sequence was determined by the Dansyl Edman method, the sequence being:

Ile-Thr-Val-Asn

TABLE 6(i): Purification procedure for tryptic peptides

Peptide	Sephadex fraction	*Purification procedure	Electorphoretic mobilities	
			pH 6.5	pH 3.5
T1	3	6B3	0	+0.32
T2	2	63	-0.34	+0.10
T3	2	63	stuck	stuck
T4	3	63	-0.36	+0.22
T5	2	63	-0.13	+0.10
T6	3	63	+0.60	+0.74

*Paper electrophoresis at pH 6.5: 6

Paper electrophoresis at pH 3.5: 3

BAWP paper chromatography: B

TABLE 6(ii): Tryptic peptides formed from C.pyrenoidosa 'B' plastocyanin

	T1	T2	T3	T4	T5	T6
His					0.7	
Lys			0.7	0.9	≠0.2	1.0
Arg		0.8			0.8	
Asp	1.1	4.8	3.0	2.0	1.0	
Thr	1.1		0.9	0.9	1.4	0.9
Ser	≠0.3	1.4	1.8			
Glu		2.1	0.7	3.0	2.9	
Pro		1.2	1.1	1.0	1.0	
Gly		1.4	2.4	1.1	4.9	
Ala		2.3	2.2	1.0	2.6	1.0
Val	1.2	*2.6	*1.5	1.1	1.2	1.1
$MetSO_2$					*0.6	
$CySO_3H$					0.7	
Ile	0.8	*1.5	1.7			
Leu			1.5	1.0	1.1	
Tyr				0.9	2.0	
Phe		1.2	1.3	0.9	1.2	
N-terminus	Ile	Asx	Leu	Glx	Leu	Ala
Yield	8.0	9.0	8.0	35	9.3	18.0

*Low yield due to partial hydrolysis of stable peptide bond.

≠ Contaminating amino acids (amounts 0.2 mole/mole peptide not shown).

Peptide T1 runs in the neutral band at pH 6.5 (paper electrophoresis), which means that the C-terminal residue is asparigine.

Peptide T6.

The sequence, determined by the Dansyl Edman method, is:

Ala-Thr-Val-Lys

Peptide T2

It is assumed that arginine is the C-terminal residue of T2, from the presumed specificity of trypsin. The N-terminal sequence was examined by the Dansyl-Edman method, the partial sequence being:

Asx-Ile-Val-Phe-Asx-(Asx_3,Ser_{1-2},Glx_2,Pro,Gly_{1-2},Ala_2,Val_2,Ile)-Arg

The observed mobility at pH 6.5 (m = -0.34) is in accord with the predicted mobility (from Offord, 1966) for a peptide of this size with a net electrostatic charge = 2, and therefore containing four amide groups. In Table 6(xi) the amide groups have been put in the same positions as they are found in the homologous region in C.pyrenoidosa 'A' plastocyanin sequence.

Peptide T3

The N-terminal sequence of peptide T3 was determined by the Dansyl Edman procedure.

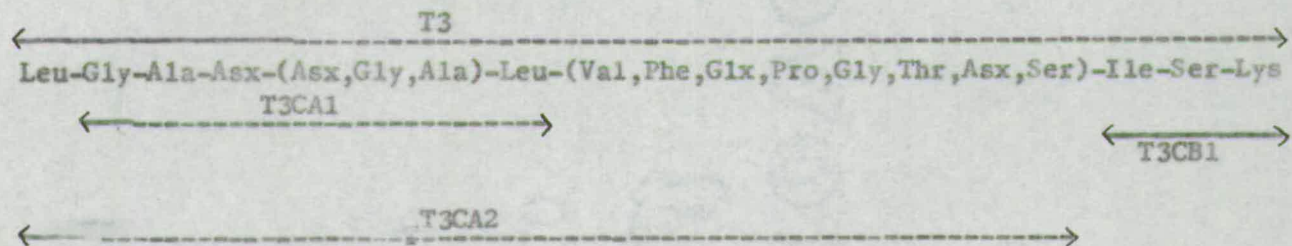
TABLE 6(iii): Peptides formed from chymotrypsin digestion of peptide T3.

Peptide	Amino acid composition	m	N-terminus
T3CB1	Ile(0.9), Ser(1.1), Lys(0.9)	+0.61	Ile
T3CA1	Asx(2.0), Gly(2.0), Ala(2.1),Leu(0.9)	-0.28	Gly
T3CA2	Asx(2.9),Thr(0.5),Ser(0.9),Glx(1.0), Pro(0.5),Gly(2.7),Ala(2.3),Val(1.2), *Ile(0.3),Leu(1.8),Phe 1.0	-0.46	Leu

*Contaminating amino acid.

Definition of symbols given in the text.

Figure 6(ii): Partial sequence of peptide T3



When digested with chymotrypsin, T3 formed the peptides shown in Table 6(iii). Quantitative analyses and N-termini were obtained for the chymotryptic peptides. Peptide T3 CB1 was sequenced by the Dansyl Edman method, and it is assumed to constitute the C-terminal sequence of T3, since it has lysine as its C-terminal residue. The evidence for the location of the chymotryptic peptides in the sequence of T3 was obtained from amino acid compositions and from a knowledge of the bonds in the sequence slightly susceptible to chymotrypsin. The partial sequence of T3 is shown in Figure 6 (ii).

Peptide T4

The N-terminal sequence of T4 was determined by the Dansyl Edman procedure.

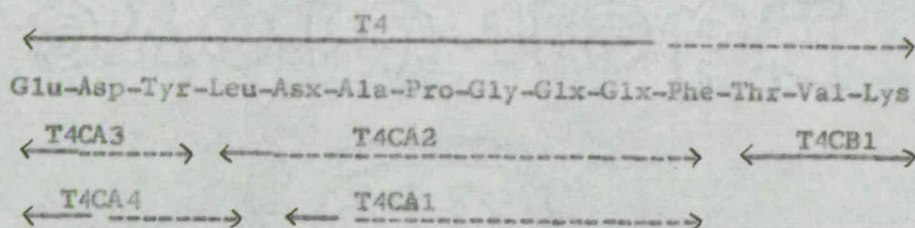
When digested with chymotrypsin, peptide T4 formed the peptides shown in Table 6(iv). Quantitative analyses and N-termini were obtained for the chymotryptic peptides. Peptide T4 CB1 was sequenced by the Dansyl Edman procedure and it is assumed to be the C-terminal peptide in the sequence of T4, from the presumed specificity of trypsin. Peptides T4 CA1 and T4 CA2 have the same amino acid compositions, except that peptide T4 CA2 is one residue longer having leucine at the N-terminus. The evidence for the location of the chymotryptic peptides was obtained from amino acid compositions, a knowledge of the bonds in the sequence which are susceptible to chymotrypsin, and from the sequence data on peptide T4. Figure 6(iii) gives the sequence of peptide T4.

TABLE 6(iv): Peptides formed from chymotrypsin digestion of peptide T4.

Peptide	Amino acid composition	m	N-terminus
T4CB1	Lys(0.9),Val(1.0),Thr(0.7).	+0.58	Thr
T4CA1	Asx(1.0),Glx(2.2),Pro(0.9),Gly(1.2),Ala(1.0),Phe(0.8)	-0.46	Asx
T4CA2	Asx(1.0),Glx(1.8),Pro(1.0),Gly(1.1),Ala(1.0),Leu(0.8) Phe(0.9)	-0.37	Leu
T4CA3	Asx(1.1),Glx(1.0),Tyr(1.0)	-0.82	Glx
T4CA4	Asx(1.1),Glx(1.1),Leu(1.0),Tyr(0.9)	-0.70	Glx

Definition of symbols and experimental details are given in the text.

Figure 6(iii) Amino acid sequence of peptide T4.



The observed mobility at pH 6.5 (m = -0.82) of T4 CA3 is in agreement with the predicted mobility (m = -0.90) for a peptide of this size with a net electrostatic charge = 2. Thus, T4 CA3 contains Glu and Asp.

Peptide T5

The N-terminal sequence of peptide T5 was determined by the Dansyl Edman procedure.

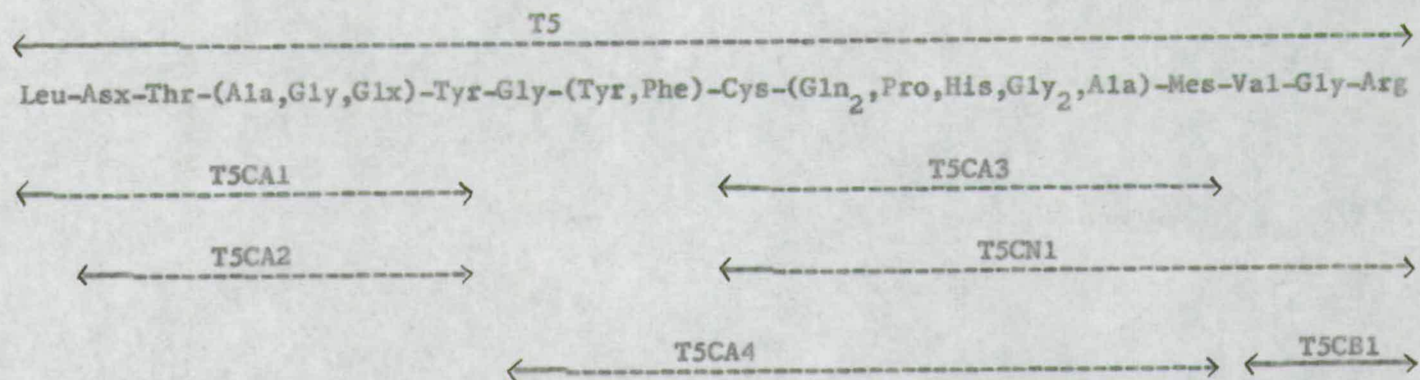
When digested with chymotrypsin, peptide T5 formed the peptides shown in Table 6(v). Peptide T5 CB1 was sequenced by the Dansyl Edman method and it was assumed to be the C-terminal peptide in the sequence of peptide T5, from the presumed specificity of trypsin. Peptides T5 CA1 and T5 CA2mshare the same amino acid compositions, except that peptide T5 CA1 is one residue longer and has leucine as its N-terminal residue. The evidence for the location of the chymotryptic peptides within the sequence of peptide T5 comes from quantitative amino acid compositions and from a knowledge of the bonds in the sequence slightly susceptible to chymotrypsin. The partial

TABLE 6(v): Peptides formed from chymotrypsin digestion of peptide T5.

	T5CA1	T5CA2	T5CA3	T5CA5	T5CB1	TBCN1
His			1.0	0.9		0.9
Lys						
Arg					0.5	0.8
Asp	1.1	1.0				
Thr	1.0	1.0				
Ser						
Glu	1.2	1.3	2.2	2.2		1.9
Pro			1.0	1.0		1.1
Gly	1.1	1.3	2.0	2.9	1.0	2.6
Ala	1.1	1.2	1.1	1.1		1.0
Val					0.9	*0.7
$MetSO_2$			0.9	0.9		*0.7
$CySO_3H$			0.9	0.9		0.9
Ile						
Leu	0.9					
Tyr	1.0	0.9		0.9		
Phe				1.1		
m	-0.50	-0.60	-0.30	-0.18	+0.66	0
m'	+0.11	+0.10	+0.10	+0.15	+0.80	+0.36
N-terminus	Leu	Asx	$CySO_3H$	Gly	Val	$CySO_3H$

* Low yield due to partial hydrolysis of stable peptide bond.

Figure 6(iv): Partial sequence of peptide T5



sequence of peptide T5 is shown in Figure 6(iv). Since peptide T5 CN1 runs in the neutral band at pH 6.5 (paper electrophoresis) it must have two residues of glutamine (not glutamic acid) in its sequence.

B. Chymotryptic digest

(i) Fractionation of Chymotryptic digest

The digest was dissolved in 5% formic acid and fractionated by Gel filtration through Sephadex G-25. Figure 6(v) shows the peptide 'map' obtained by paper electrophoresis (at pH 6.5) of a portion of each fraction from the gel filtration separation. The peptides were pooled into fractions C1 to C3, as shown, and each fraction was freeze dried.

Fractions C1 to C3 were separated on paper electrophoresis at pH 6.5 on 15 cm. strips. The peptides were further purified on (BAWP) paper chromatography and paper electrophoresis at pH 3.5 and 2.0.

(ii) Summary of chymotryptic peptides

Table 6(vi) gives a list of chymotryptic peptides purified, showing their electrophoretic mobilities at pH 6.5 and pH 3.5.

The amino acid compositions, N-termini, and yields (percentage µmole/µmole protein digested) are presented in Table 6(vii). The yields are not corrected in any way for losses incurred during the purification procedure.

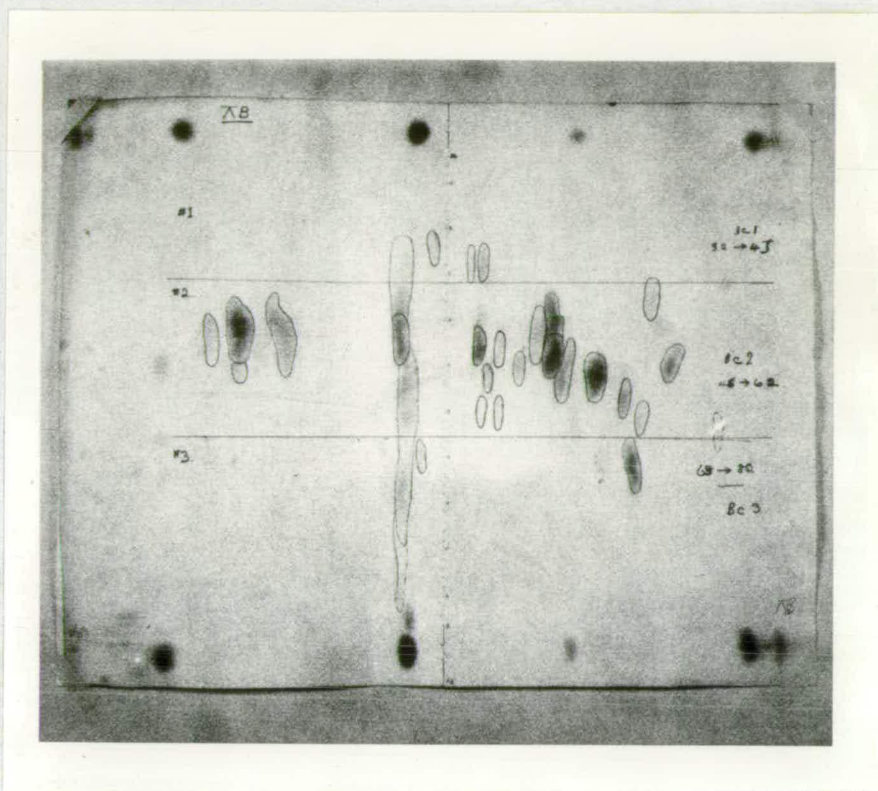


Figure 6(v) HVPE at pH 6·5 of chymotryptic peptides from *C.pyrenoidosa* B plastocyanin.

TABLE 6(vi): Purification procedure for chymotryptic peptides.

Peptide	Sephadex fraction	*Purification procedure	Electrophoretic mobilities pH 6.5	pH 3.5
CA1	2	63	-0.50	+0.19
CA2	2	63	-0.40	+0.21
CA3	1	63	-0.16	+0.27
CA4	2	63	-0.50	+0.07
CA5	2	63	-0.12	≠ -
CA6	2	6B3	-0.27	≠ -
CA7	2	6B3	-0.27	≠ -
CA8	2	6B3	-0.27	+0.12
CB1	2	632	+0.40	+0.47
CB2	2	632	+0.40	+0.68
CB3	2	632	+0.40	≠ -
CN1	2	6B3	0	≠ -
CN2	2	6B3	0	≠ -

*Paper electrophoresis at pH 6.5:6
" at pH 3.5:3
" at pH 2.0:2
BAWP paper chromatography: B

≠ Paper electrophoresis at pH 3.5 run for extended length of time (i.e. Lysine run off the paper) (see text).

TABLE 6(vii): Chymotryptic peptides formed from C.pyrenoidosa 'B' plastocyanin.

	CA1	CA2	CA3	CA4	CA5	CA6	CA7	CA8	CB1	CB2	CB3	CN1	CN2
His								0.9	0.9				
Lys					0.9					0.8		0.9	0.8
Arg						1.0					1.0		
Asp	1.0	1.0	1.2	4.6	1.0	1.0	1.7		2.0		1.0		
Thr			1.1		1.7					0.9	0.9	1.1	0.9
Ser			1.6	0.9		0.9						1.8	1.0
Glu	2.1	2.2	0.7	1.0	1.1	2.2	≠0.3	2.1				1.1	
Pro	0.9	1.0	0.5	1.0				1.1	0.9				
Gly	1.1	1.0	0.7	1.0	1.0		2.0	2.0	1.0	≠0.3	1.1	1.0	1.0
Ala	1.1	1.1	≠0.2	≠0.2	1.0	1.9	1.8	1.1	1.0	0.6			
Val			1.1	*2.4	0.9				1.1	0.9	2.0	1.0	1.1
$MetSO_2$								0.9					
$CySO_3H$								0.9					
Ile			1.1	*0.6		1.1					1.0	0.9	
Leu		1.0			1.0		1.3			1.2			
Tyr					1.0	1.0							
Phe	0.9	1.0	1.0	1.0					1.0			0.9	0.9
N-terminus	Asx	Leu	Val	Asx	Thr	Ala	Gly	$CySO_3H$	Val	Ala	Val	Ile	Lys
Yield	1.4	6.3	2.0	2.8	3.0	1.5	4.1	1.2	4.8	1.2	3.0	1.4	1.0

*Low yield due to partial hydrolysis of stable peptide bond.
≠Contaminating amino acids (amounts 0.2 mole/mole peptide not shown).

(iii) Peptides formed by chymotrypsin digestion

The evidence for the sequence of the chymotryptic peptides is presented.

Peptide CB2.

CPA released leucine (+++). The peptide was sequenced by the Dansyl Edman procedure, the sequence being:

Ala-Thr-Val-Lys-Leu

Peptide CB3.

CPA released asparigine (+++), threonine (++), valine (++), and isoleucine (++). The sequence was determined by the Dansyl Edman procedure, the sequence being:

Val-Gly-Arg-Ile-Thr-Val-Asn

Peptide CB1.

CPA released traces of histidine (+). The sequence of CB1 was determined by the Dansyl Edman procedure.CB1 is identical in sequence to the proposed sequences for CB4 from *C.pyrenoidosa* 'A' plastocyanin and FCB2 from *C.fusca* plastocyanin, except that it lacks a threonine residue at the C-terminus. The quantitative analysis of CB1 indicated that it was a pure peptide, and the results obtained from the Dansyl Edman are reliable. This sequence data helps to confirm the proposed sequences of CB4 and FCB2 (from *C.pyrenoidosa* A and *C.fusca* plastocyanins respectively).

Since CB1 is basic on paper electrophoresis at pH 6.5, it is assumed that the basicity is due to the histidine residue and that the 2nd and 3rd residues are both asparigines. The sequence of CB1 is:

Val-Asn-Asn-Ala-Gly-Phe-Pro-His

The formation of CB1 represents non-specific cleavage of the protein (after histidine), by chymotrypsin.

Peptides CA1 and CA2

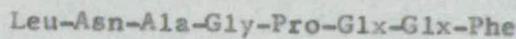
CPA released Phe (++) from both peptides. CA1 and CA2 were sequenced by the Dansyl Edman method, and were shown to have the same sequence, except that CA2 is one residue longer possessing leucine as its N-terminal residue.

The observed electrophoretic mobility of CA2 at pH 6.5 (m = - 0.40) is in agreement with the predicted mobility for a peptide of this size and with a net electrostatic charge = 1. A sample of peptide was removed after the 2nd cycle of the PTC degradative method, and the electrophoretic mobility (at pH 6.5) of the peptide measured. There was insufficient material to ascertain which of the two glutamic acid residues was amidated. Table 6 (viii) gives the electrophoretic mobility of the degradative product of CA2.

Table 6(viii) PTC degradative product of peptide CA2

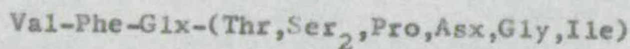
PTC degradation cycle	Peptide sequence	Electrophoretic mobility at pH 6.5	Comments
0	Leu-Asx-Ala-Gly-Pro-Glx-Glx-Phe	-0.40	
2	Ala-Gly-Pro-Glx-Glx-Phe	-0.43	Asn (not Asp)

The sequence of peptide CA2 was shown to be:


Leu-Asn-Ala-Gly-Pro-Glx-Glx-Phe

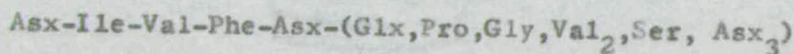
Peptide CA3

The N-terminal sequence was examined by the Dansyl Edman method, the partial sequence being:


Val-Phe-Glx-(Thr,Ser_2,Pro,Asx,Gly,Ile)

Peptide CA4

The N-terminal sequence of CA4 was examined by the Dansyl Edman procedure, the partial sequence being:


Asx-Ile-Val-Phe-Asx-(Glx,Pro,Gly,Val_2,Ser, Asx_3)

The observed mobility at pH 6.5 (m = -0.50) is in satisfactory agreement with the predicted mobility (m = -0.59) for a peptide of this size with a net electrostatic charge = 3, and thus containing 3 amide groups. CA4 is homologous with TA3 from C.pyrenoidosa 'A' plastocyanin, except that it is two residues shorter at the C-terminus. In Table 6(x) the amide groups

in CA4 have been placed in the same positions as they are found in TA3 (C-pyrenoidosa 'A' protein).

Peptide CA5

CPA released tyrosine (+++). The sequence was determined by the Dansyl Edman method. Since, CA5 has a net acidic charge (on paper electrophoresis) at pH 6.5, the 5th and 9th residues must be aspartic acid and glutamic acid, respectively. The sequence is:

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

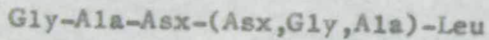
Peptide CA6

CPA released tyrosine (+++). The sequence was determined by the Dansyl Edman procedure. The observed mobility (m = -0.27) at pH 6.5 is in accord with the predicted mobility for a peptide of this size with a net electrostatic charge = 1. Now the last three residues in CA6 are identical in sequence to the N-terminal sequence (Glu-Asp-Tyr) in the tryptic peptide T4 (this Chapter). Therefore, the second residue in CA6 must be glutamine, if the peptide is to have a net electrostatic charge of 1. Thus, the proposed sequence for CA6 is:

Ala-Gln-Ala-Ile-Ser-Arg-Glu-Asp-Tyr

Peptide CA7

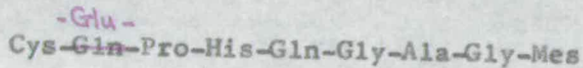
CPA released leucine (++). The N-terminal sequence was determined by the Dansyl Edman method. The partial sequence is:



Gly-Ala-Asx-(Asx,Gly,Ala)-Leu

Peptide CA8

CPA released methionine sulphone (+++), glycine (++), alanine (++), glutamine (++), and a trace of histidine (+). The sequence was determined by the Dansyl Edman method. After the 1st cycle of PTC degradation the resulting peptide (amino acid composition (Glx_2, Pro, His, Gly_2, Ala, Mes)) ran in the neutral band on paper electrophoresis at pH 6.5, showing that the 2nd residue in the sequence is glutamic acid. The sequence of CA8 is:

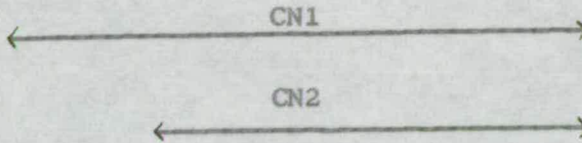


Cys-Glu-Pro-His-Gln-Gly-Ala-Gly-Mes

Peptides CN1 and CN2

CPA released phenylalanine (++) from both peptides. Peptide CN2 was sequenced by the Dansyl Edman method. Peptide CN1 has the same amino acid composition as peptide CN2, except that it has an extra Ile-Ser sequence at the N-terminal end. Since both peptides run in the neutral band (on paper electrophoresis) at pH 6.5, then the 5th residue in peptide CN1 and the 3rd residue in peptide CN2 is glutamic acid. The sequence covered by peptides CN1 and CN2 is:

Ile-Ser-Lys-Gly-Glu-Ser-Val-Thr-Phe



The formation of CN2 is an example of non-specific cleavage by chymotrypsin of a Ser-Lys band. However, CN2 was recovered in very low yield, so that the hydrolysis of the bond must have taken place slowly.

TABLE 6(ix): LIST OF TRYPTIC PEPTIDES

T1: Ile-Thr-Val-Asn

T2: Asn-Ile-Val-Phe-Asn-(Asn,Asp,Glu,Asp,Ser,Pro,Val,Val,Gly)-Ala-Gln-Ala-Ile-Ser-Arg

T2

CA4 CA6

T3: Leu-Gly-Ala-Asx-(Asx,Gly,Ala)-Leu-Val-Phe-Glx-(Pro,Gly,Thr,Asx,Ser)-Ile-Ser-Lys

CA7 CA3

T3

T4: Glu-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glx-Glx-Phe-Thr-Val-Lys

CA2

T4

T5: Leu-Asp-Thr-Ala-Gly-Glu-Tyr-Gly-(Tyr,Phe)-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Arg

CA5 CA8

T5

T6: Ala-Thr-Val-Lys

TABLE 6(x) List of chymotryptic peptides

CA1:	Asn-Ala-Gly-Pro-Glx-Glx-Phe
CA2:	Leu-Asn-Ala-Gly-Pro-Glx-Glx-Phe
CA3:	Val-Phe-Glx-(Pro,Gly,Thr,Asx,Ser,Ile,Ser)
CA4:	Asn-Ile-Val-Phe-Asn-(Asn,Asp,Glu,Asp,Ser,Pro,Val,Val,Gly)
CA5:	Thr-Val-Lys-Leu-Asp-Thr-Ala-Gly-Glu-Tyr
CA6:	Ala-Gln-Ala-Ile-Ser-Arg-Glu-Asp-Tyr
CA7:	Gly-Ala-Asx-(Asx,Gly,Ala)-Leu
CA8:	Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes
CB1:	Val-Asn-Asn-Ala-Gly-Phe-Pro-His
CB2:	Ala-Thr-Val-Lys-Leu
CB3:	Val-Gly-Arg-Ile-Thr-Val-Asn
CN1:	Ile-Ser-Lys-Gly-Glu-Ser-Val-Thr-Phe
CN2:	Lys-Gly-Glu-Ser-Val-Thr-Phe

Figure 6(v) to 6(x): Alignment of *C.pyrenoidosa* 'B' plastocyanin against the sequence of *C.pyrenoidosa* 'A' plastocyanin.

6(v)

1 2 3 4 5

A Ala-Asn-Val-Lys-Leu

B Ala-Thr-Val-Lys-Leu

CB2 (residues 1–5)

T6 (residues 1–4)

6(vi)

4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

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

B Lys-Leu-Gly-Ala-Asx-(Asx,Gly,Ala)-Leu-Val-Phe-Glx-(Pro,Gly ___ ___,Thr,Asx,Ser)-Ile-Ser-Lys

T3

CA7 CA3 CN2 .....

6(vii)

23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40

A Ile-Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr

B Ile-Ser-Lys-Gly-Glu-Ser-Val-Thr-Phe-Val-Asn-Asn-Ala-Gly-Phe-Pro-His- —

CN1 CB1

CN2

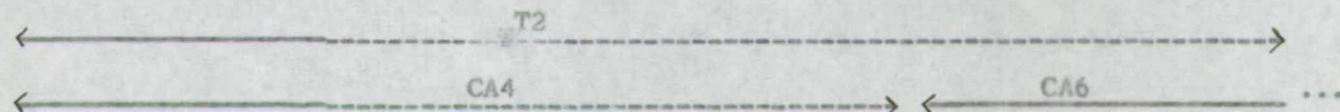
Figure

6(viii)

41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

A Asn-Ile-Val-Phe-Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu-Ser-His

B Asn-Ile-Val-Phe-Asn-(Asn,Asp,Glu,Asp,Ser,Pro,Val,Val,Gly)-Ala-Gln-Ala-Ile-Ser-Arg



6(ix)

55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74

A Ala-Asn-Ala-Leu-Ser-His-Asp-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Thr-Ala-Lys

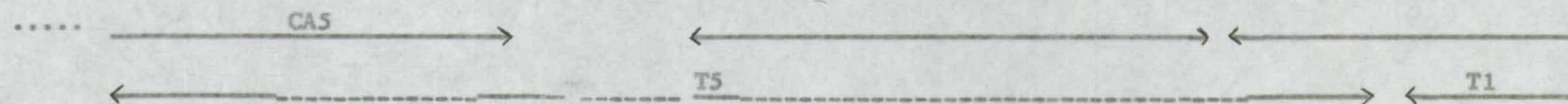
B Ala-Gln-Ala-Ile-Ser-Arg-Glu-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glx-Glx-Phe-Thr-Val-Lys

6(x)

75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

A Phe-Asp-Thr-Ala-Gly-Glu-Tyr-Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Lys-Ile-Val-Val-

B Leu-Asp-Thr-Ala-Gly-Glu-Tyr-Gly-(Tyr,Phe)-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Arg-Ile-Thr-Val-



C. Deduction of sequence

The overall sequence was deduced from the tryptic peptides, a list of which is given in Table 6(ix). The sequences of the small peptides were established from the peptides themselves, and this evidence has been noted, while others required a consideration of the chymotryptic peptides (see Table 6(x)). These tryptic peptides are discussed.

The sequence data derived from the tryptic and chymotryptic digests on C.pyrenoidosa 'B' plastocyanin (B), is insufficient, in itself, to deduce a sequence. A sequence has been proposed by homology with the sequence of C.pyrenoidosa 'A' plastocyanin (A). The amino acid residues in the bracketed regions have been ordered to maximise the alignment of the two sequences.

The partial sequence of T2 can be deduced from a consideration of CA4 and the N-terminal region (residues 1-6) of CA6. T2 is homologous with CA1 from A, except that it is longer at the C-terminus.

The partial sequence of T3 can be deduced from a consideration of CA7 and CA3.

The partial sequence of T5 can be derived from a consideration of CA8 and the C-terminal sequence (residues 4-10) of CA5.

I shall begin the deduction of sequence at the N-terminus. The N-terminus was determined by the Dansyl Chloride method and the 2nd and 3rd residues by the Dansyl Edman method. CB2 and T6 have their first three residues in common with the whole protein. They must both be at the N-terminus (see Figure 6(v)). Two tryptic peptides B T3 and

B T5, have leucine as their N-termini, but it is evident, from a knowledge of the C-terminus of the protein, that B T5 is positioned near the C-terminus. B T3 is, therefore, the only candidate for residue 5 in the sequence. It is homologous with A TA4 (residues 5 to 24 in A sequence), except that Tyr-18 and Tyr-19 appear to be deleted in the B sequence. Figure 6(vi) shows the sequence homology in this region of the protein. B CN1 has the N-terminal sequence Ile-Ser-Lys-Gly, which fits into the sequence at residue 22 (Figure 6(vi)). No tryptic peptide was isolated from B with N-terminal glycine to fit into residue 25 in the sequence. However, B CN1 is homologous with A CN2 (residues 23-31 in sequence of A) and B CB1 is homologous with A CB4 (residues 32 to 40 in A sequence). I have tentatively placed B CB1 in the sequence at residue 32 (Figure 6(vii)) and proposed a deletion at residue 40 to obtain the best alignment of the rest of the sequence with the A sequence. No overlapping sequence was found between residues 39 and 41. However, B T2 is homologous with A CA1 (residues 41 to 58 in A sequence, except that B T2 is two residues longer at the C-terminus. BCA6 provides the overlapping sequence for the C-terminus of B T2 and the N-terminus of B T4 (Figure 6(ix)). B T4 has the C-terminal sequence -Thr-Val-Lys, which corresponds to the N-terminal sequence of B CA5. B CA5 provides the overlapping sequence for B T4 and B T5 (Figure 6(x)). The C-terminal sequence of B T5

is -Val-Gly-Arg. The C-terminal peptide, of the whole protein, was isolated from a cyanogen bromide digest, and was identified as such because it contained no trace of homoserine/homoserine lactone. It has the sequence:

Val-Gly-Arg-Ile-Thr-Val-Asn

B CB3 has the same sequence, and it is homologous with the C-terminal peptide of A. B T1 must be the C-terminal tryptic peptide. Figure 6(x) gives the sequence between residues 75 and the C-terminus.

Discussion of sequence

The alignment of the tryptic and chymotryptic peptides formed from B against the postulated sequence of A is shown in Figures 6(v) to 6(x). All the peptides purified have been accounted for in the sequence. A tryptic peptide covering the region of the protein corresponding to residues 26-29, expected from the presumed specificity of trypsin, was not isolated. No overlapping sequence was found between residues 39-41, and this was also true for the A sequence. Non-specific cleavage by chymotrypsin occurred at Serine-24, and must also have occurred in the formation of B CN1 and B CA6, but the preceding residues were not sequenced so that I cannot say which bonds were hydrolysed by chymotrypsin.

The proposed sequence is in satisfactory agreement with the total amino acid composition of the protein (Chapter 4). The amino acid changes between the sequences of A and B will be discussed in the final chapter.

CHAPTER 7

DIGESTIONS PERFORMED ON C.FUSCA PLASTOCYANIN

Introduction

A chymotryptic digest was performed on 100 mg. (about 10 μmoles) of protein. A 15 mg. (about 1.5 μmoles) subtilisin digest was carried out on unoxidised protein in an attempt to search for a tryptophan-containing peptide, if tryptophan is present in the molecule. Subtilisin was chosen for the digestion because it results in relatively non-specific cleavage of peptide bonds, and hence many small peptides should be formed during the digestion. Also, tryptophan-containing peptides are known to be anomalously retarded by gel filtration through Sephadex, and so such a peptide should separate quite well from other peptides (Determann & Walter, 1968).

A. Chymotryptic digest

The digestion was performed on oxidised protein and the conditions for digestion were the same as those described for the tryptic digestion of C.pyrenoidosa 'A' plastocyanin.

(i) Fractionation of chymotryptic digest

The digest was dissolved in 5% formic acid and separated by gel filtration through Sephadex G-25, followed by paper electrophoresis at pH 6.5. Figure 7(i) shows a peptide 'map' obtained by

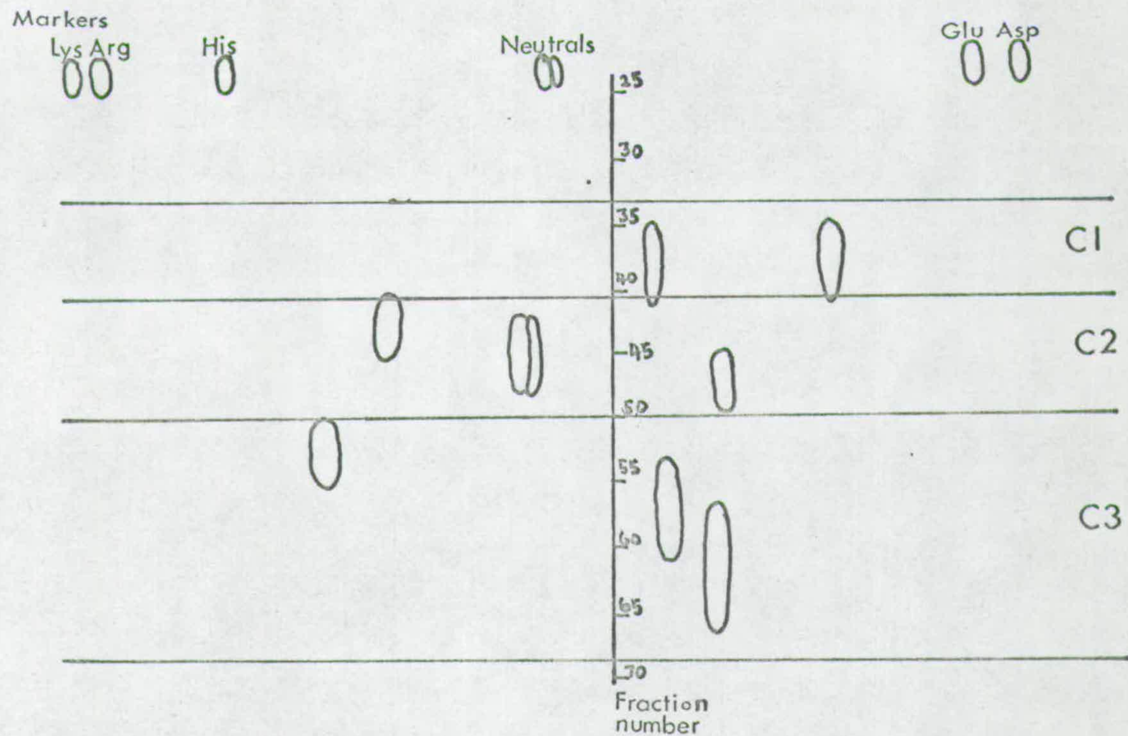


Figure 7(i) Separation of chymotryptic peptides of *C. fusca* plastocyanin by Gel filtration followed by paper electrophoresis at pH 6·5.

paper electrophoresis of a portion of each fraction from the gel filtration separation. The peptides were pooled into three fractions (C1 to C3) as shown, and each fraction was freeze dried.

Fractions C1 to C3 were separated on 16 cm. strips, on paper electrophoresis at pH 6.5. The peptides were further purified on (BAWP) paper chromatography and paper electrophoresis at pH 3.5.

(ii) Summary of chymotryptic peptides

Table 7(i) gives a list of the chymotryptic peptides and their electrophoretic mobilities at pH 6.5 and pH 3.5. Most of the mobilities were recorded from the preparative separation and will not be very accurate.

Table 7(ii) gives the amino acid compositions, N-termini, and yields of the peptides. The yields (expressed as percentage µmole/ µmole protein digested) are approximate and are not corrected for losses incurred during the purification procedure.

(iii) Peptides formed from chymotryptic digest

The evidence for the sequence of the small peptides and the partial sequence of the large peptides is presented. Where the peptides are identical in sequence (or obviously homologous) with the corresponding chymotryptic peptide formed from *C.pyrenoidosa* 'A' plastocyanin, it is stated.

TABLE 7(i): Purification procedure for chymotryptic peptides

Peptide	Sephadex fraction	*Purification procedure	Electrophoretic mobilities	
			pH 6.5	pH 3.5
CA1	2	63	-0.22	+0.11
CA2	1	63	-0.11	+0.21
CA3	3	63	-0.26	0
CA4	3	63	-0.16	≠ -
CA5	1	63	-0.60	+0.24
CA6	3	63	-0.51	+0.36
CA7	2	63	-0.14	≠ -
CA8	2	63	-0.14	≠ -
CA9	2	63	-0.29	+0.13
CB1	1	63	+0.42	+0.60
CB2	2	63	+0.39	+0.46
CB3	2	63	+0.52	+0.70
CB4	1	63	+0.29	+0.47
CN1	2	6B3	0	+0.39
CN2	2	6B3	0	0.39

*Paper electrophoresis at pH 6.5 : 6

≠ " " at pH 3.5 : 3

BAWP paper chromatography : B

≠ Paper electrophoresis at pH 3.5 run for extended length of time (i.e. lysine run off the paper) (see text)

TABLE 7(ii): Chymotryptic peptides formed from C.fusca plastocyanin.

	CA1	CA2	CA3	CA4	CA5	CA6	CA7	CA8	CA9	CB1	CB2	CB3	CB4	CN1	CN2
His	0.9					1.0		0.8	1.0		1.0	1.0	1.0		
Lys		0.9								0.9			0.9	0.9	1.0
Arg															
Asp	/0.2	1.9	1.0	1.0	3.9	1.0					2.0		1.9	0.9	
Thr	/0.2	1.0	1.8				1.0			1.8	0.5		1.9	1.0	1.8
Ser		1.0		1.0	0.9	0.7	1.9	/0.3				0.8			
Glu	2.1			1.1	3.2	1.0	1.0	2.2	2.2	1.0	/0.2		1.0		1.0
Pro	1.0			0.9	1.0		1.0	1.0	1.0		1.0		1.0		
Gly	2.1	2.1	1.0	1.1	1.0			2.7	2.3	1.0	1.2		2.0		1.0
Ala	1.2	2.0	1.1	1.0	3.1			1.1	1.3		1.1	1.2	2.1		1.0
Val		1.9			*1.5		1.9	/0.2		1.0	1.1		1.9	2.1	1.0
$MetSO_2$	0.8							0.9	0.9						
$CySO_3H$	1.0							1.0	1.1						
Ile					*0.5		*0.5			1.0	/0.2		*0.7	1.1	
Leu		1.8		0.9	1.0		/0.2								
Tyr			1.0			1.0		1.0							
Phe	1.0				1.0		0.9	1.0			1.0	0.9	1.6		0.9
N-terminus	Phe	Asx	Asx	Leu	Asx	Ser	Val	Gly	$CySO_3H$	Lys	Val	Ser	Ile	Asx	Lys
Yield	4.5	4.0	32.0	14.2	19.8	8.4	8.0	2.1	1.6	14.4	5.4	10.7	1.2	5.7	2.0

*Low yield due to partial hydrolysis of stable peptide bond

/Contaminating amino acids (amounts 0.2 mole/mole peptide not shown)

Peptide CB3

Peptide CB3 was sequenced by the Dansyl Edman method, the sequence being:

Ser-Ala-Lys-Phe

Peptide CB1

Peptide CB1 was sequenced by the Dansyl Edman method. Since the peptide has a net basic charge (on paper electrophoresis) at pH 6.5, the C-terminal residue must be glutamine. The sequence of peptide CB1 is:

Lys-Gly-Thr-Ile-Thr-Val-Gln

Peptide CA6

Peptide CA6 was sequenced by the Dansyl Edman procedure. It is clearly homologous with peptide CA4 derived from C.pyrenoidosa 'A' plastocyanin (Chapter 5). Both peptides have a very similar electrophoretic mobility at pH 6.5; peptide CA6 (m = -0.51) and peptide CA4 (m = -0.44). It is, therefore, assumed that the 2nd and 3rd residues in the sequence of peptide CA6 are glutamic acid and aspartic acid, respectively. The sequence of peptide CA6 is:

Ser-His-Glu-Asp-Tyr

Peptides CB2, CN2 and CB4

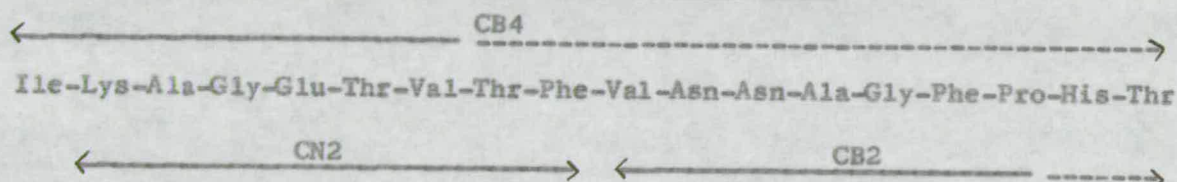
Peptide CN2 was sequenced by the Dansyl Edman method, and since it runs in the neutral band at pH 6.5, the 4th residue must be glutamic acid (see the sequence below). The N-terminal sequence of CB4 was examined by the Dansyl Edman method, and it was evident that CB4 corresponded in amino acid composition to the sum of CN2 and CB2, and in addition has an isoleucine residue at the N-terminus.

Peptide CB2 is homologous with CB4 from C.pyrenoidosa 'A' plastocyanin (denoted A), except that the amino acid composition clearly indicates the presence of a valine residue, the presence of which there was some doubt in A CB4. The sequence of CB2 was examined by the Dansyl Edman procedure, and the method gave fairly conclusive results except for the last two residues (His, and Thr). The yield of threonine from the quantitative analysis was low (0.5 residue), and it could be argued that it is a contaminating amino acid. The total amino acid composition of CB4 is consistent with there being 2 residues of threonine in this region of the protein, and since CN2 contains 2 residues of threonine in its sequence, the presence of threonine in CB2 is inconclusive. Since CB2 is slightly basic (on paper electrophoresis) at pH 6.5,

it is assumed that the basicity is due to histidine and that the 2nd and 3rd residues are asparigines.

The postulated sequence of CB4 is shown in Figure 7(ii).

Figure 7(ii) Postulated amino acid sequence of CB4

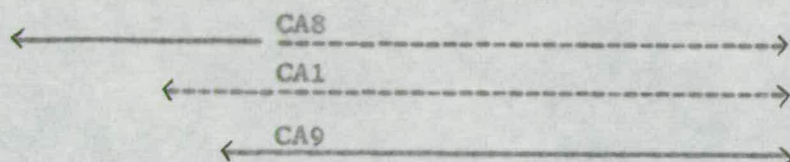


Peptides CA1, CA8 and CA9

All three peptides contain 1 residue of cysteic acid, and must come from the same region in the protein, since it is known from the amino acid analyses that the whole protein contains a single residue of cysteine.

Peptide CA9 was sequenced by the Dansyl Edman method, and it has the same sequence and very similar electrophoretic mobility at pH 6.5 to peptide CA8 formed from C.pyrenoidosa 'A' plastocyanin, and it is, therefore, assumed that the 2nd and 5th residues in peptide CA9 (from C.fusca plastocyanin) are both glutamines. Peptide CA1 has the same amino acid composition as peptide CA9, except that it is one residue longer and has phenylalanine as its N-terminal residue. The N-terminal sequence of peptide CA8 was examined by the Dansyl Edman procedure, and the sequence around the cysteic acid residue is shown below:

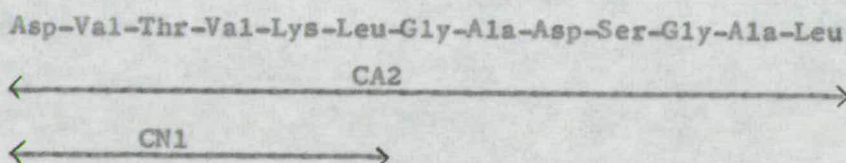
Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes



Peptides CN1 and CA2

The amino acid sequences of peptides CN1 and CA2 were determined by the Dansyl Edman method. It was evident from the sequences that peptide CA2 overlapps peptide CN1 at its C-terminal residue. Since peptide CA2 has a net acidic charge at pH 6.5 (paper electrophoresis), the N-terminal residue and the 9th residue must both be aspartic acid residues. The sequence of peptide CA2 is:

Asp-Val-Thr-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu



Peptide CA3

Peptide CA3 was sequenced by the Dansyl Edman method. The peptide possesses a net acidic charge (on paper electrophoresis) at pH 6.5, and so the N-terminal residue is aspartic acid. The sequence is:

Asp-Thr-Ala-Gly-Thr-Tyr

Peptide CA4

Peptide CA4 was sequenced by the Dansyl Edman procedure. It has the same sequence and similar electrophoretic mobility as peptide CA5 formed from C.pyrenoidosa 'A' plastocyanin. It is, therefore, assumed that the 2nd residue in the sequence is asparigine. The postulated sequence is:

Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr

Peptide CA7

The N-terminal sequence was determined by the Dansyl Edman method. Since CA7 has a net acidic charge (on paper electrophoresis) at pH 6.5, the 3rd residue in the sequence must be glutamic acid. CA7 is homologous with TN3 formed from C.pyrenoidosa 'A' plastocyanin, except that I have no evidence for the presence of any tyrosine residues in CA7. If the tyrosines were present as mono- or di-chloro tyrosine, I would not have observed them on the amino acid analyser, since they are eluted from the short column (used to determine basic amino acids, on Beckman 120C model) of the analyser, and I only ran a long column, since no basic amino acids or tyrosine had been suggested from the qualitàtive analysis of CA7.

The postulated sequence is:

Val-Phe-Glu-Pro-Ser-(Thr,Val,Ser,Ile)

From the amino acid composition of CA7, it has been formed as a result of non-specific cleavage by chymotrypsin.

Peptide CA5

The N-terminal sequence was examined by the Dansyl Edman method, the partial sequence being:

Asx-Ile-Val-Phe-Glx-Glx-Asx-Glx-Asx-Val-Pro-(Ala_3, Gly, Asx, Ser)-Leu

Since Phe was placed in the sequence as the 4th residue, it was assumed that leucine was the C-terminal residue. The observed mobility at pH 6.5 (m = -0.60) is in accord with the predicted mobility (from Offord, 1966) for a peptide of this size with a net electrostatic charge = 3, and therefore containing four amide groups.

Peptide CA5 is homologous with CA1 from C.pyrenoidosa 'A' plastocyanin (A), both peptides being 18 residues long. If the acid residues and amide groups were in the same positions in the sequence as they are in   A CA1 then the postulated partial sequence of CA5 is:

Asn-Ile-Val-Phe-Gln-Gln-Asp-Glu-Asp-Val-Pro- (Ala_3, Gly, Asn, Ser)-Leu

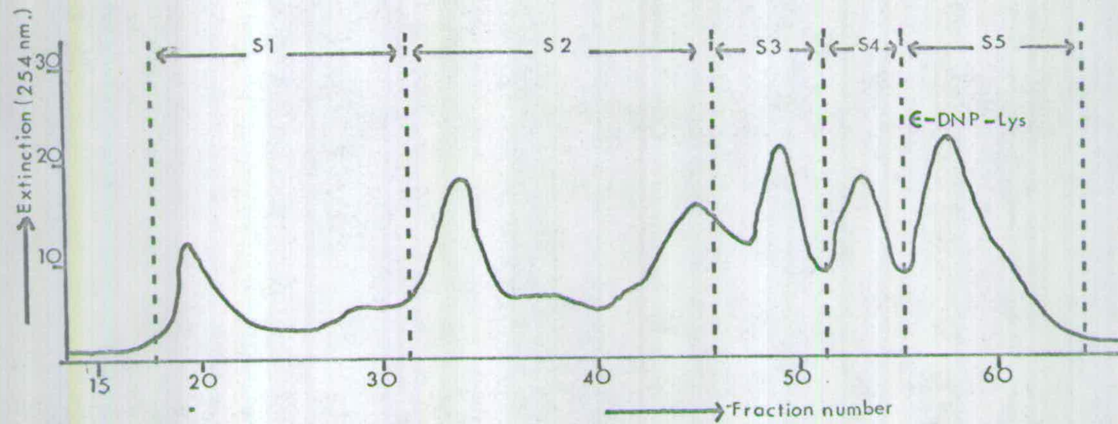


Figure 7(iii) Gel filtration of subtilisin digest from C.fusca plastocyanin.

E_{254nm} is shown in arbitrary units.

Fraction size = 2·9 ml.

B. Subtilisin digest

(i) Conditions of digestion

The protein was precipitated with TCA (5% v/v) and the precipitate washed 3 times with absolute alcohol and dissolved in 0.2M ammonium acetate pH 8.5. Subtilisin was added and the digestion mixture shaken at 37°C for 3 hr., diluted, and freeze dried.

(ii) Fractionation of subtilisin digest

The digest was dissolved in 5% formic acid and separated by gel filtration through Sephadex G-25. An extinction (at 254 nm.) profile of the effluent is shown in Figure 7(iii). The peptides were pooled into 5 fractions (S1 to S5), as shown, and each fraction was freeze dried.

Fractions S1 to S5 were run on 8 cm. strips on paper (Whatman No.1) electrophoresis at pH 6.5. One marker strip (from each fraction) was stained with ninhydrin, and after the peptides had been located, the same marker strips were stained for tryptophan (Ehrlich reagent). A mauve spot (i.e. tryptophan positive) peptide S5N, showed up close to the neutral band of fraction S5. The peptides on the second marker strip for fraction S5 were oxidised with performic acid in a vacuum dessicator for about 4 hr. The strip was allowed to dry and then examined under u.v. light (at 365 nm.). A strongly fluorescent spot was

observed corresponding to the position of peptide S5N. This is suggestive of the presence of tryptophan since oxidised tryptophan is more fluorescent under u.v. light than is unoxidised tryptophan.

Peptide S5N was cut out and sewn on to another sheet of paper and run out on electrophoresis at pH 3.5.

(iii) Discussion of peptide S5N

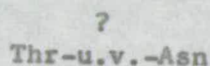
The qualitative amino acid composition was determined on paper electrophoresis at pH 2.0. Before staining for amino acids the paper was examined under u.v. light for fluorescent degradative products of tryptophan. The N-terminus was determined by the Dansyl Chloride method, and the 2nd and 3rd residues by the Dansyl Edman procedure. The results are shown in Table 7(iii).

TABLE 7(iii) Characterisation of peptide S5N

Amino acid composition	N-terminus	N-terminus (after 1st cycle PTC degradation)	N-terminus (after 2nd cycle PTC degradation)
Thr(+++),Asx(+++), u.v.(+++),*Phe(+), *Val(+)	Thr(+++), *Val(+)	Asx(+)	Asx(+++), *Phe (+)

* Contaminating amino acids (?)

If one assumes, from the qualitative amino acid analysis, that the amino acids Thr and Asn (peptide S5N is neutral at pH 6.5) are present as single residues, then the sequence of peptide S5N is:



I propose that there is a residue of tryptophan in position 2 of peptide S5N.

The evidence is suggestive (rather than confirmatory) of the presence of tryptophan in peptide S5N. Lack of material prevented me from determining the quantitative amino acid composition of the peptide, and doing a quantitative estimation of tryptophan content.

C. Deduction of sequence

The sequence of C.fusca plastocyanin was deduced by homology with C.pyrenoidosa 'A' plastocyanin. Some of the sequences of the chymotryptic peptides were established from the peptides themselves, by the Dansyl Edman procedure, while others required a consideration of the peptides formed from C.pyrenoidosa 'A' plastocyanin. A list of the chymotryptic peptides is given in Table 7(iv).

TABLE 7(iv): List of chymotryptic peptides formed from *C.fusca* plastocyanin.

CA1: Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes

CA2: Asp-Val-Thr-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu

CA3: Asp-Thr-Ala-Gly-Thr-Tyr

CA4: Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr

CA5: Asn-Ile-Val-Phe-Gln-Gln-Asp-Glu-Asp-Val-Pro-(Ala_3,Gly,Asn,Ser)-Leu

CA6: Ser-His-Glu-Asp-Tyr

CA7: Val-Phe-Glu-Pro-Ser-(Thr,Val,Ser,Ile)

CA8: Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes

CA9: Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes

CB1: Lys-Gly-Thr-Ile-Thr-Val-Gln

CB2: Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr

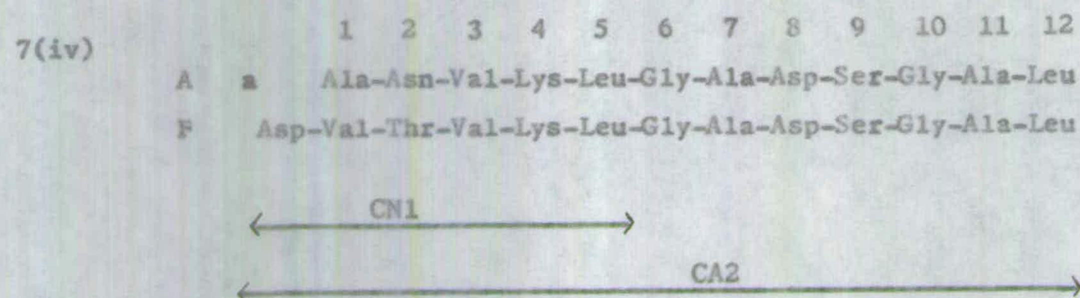
CB3: Ser-Ala-Lys-Phe

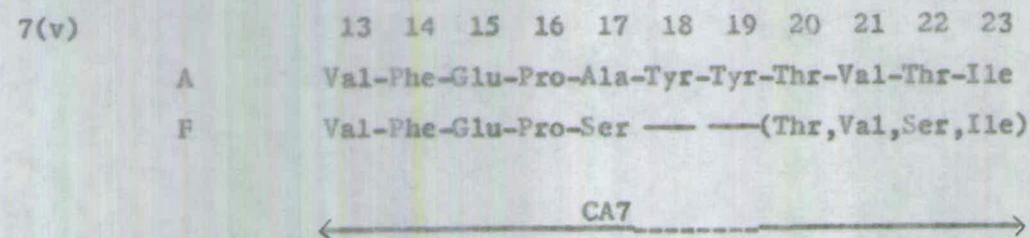
CB4: Ile-Lys-Ala-Gly-Glu-Thr-Val-Thr-Phe-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr

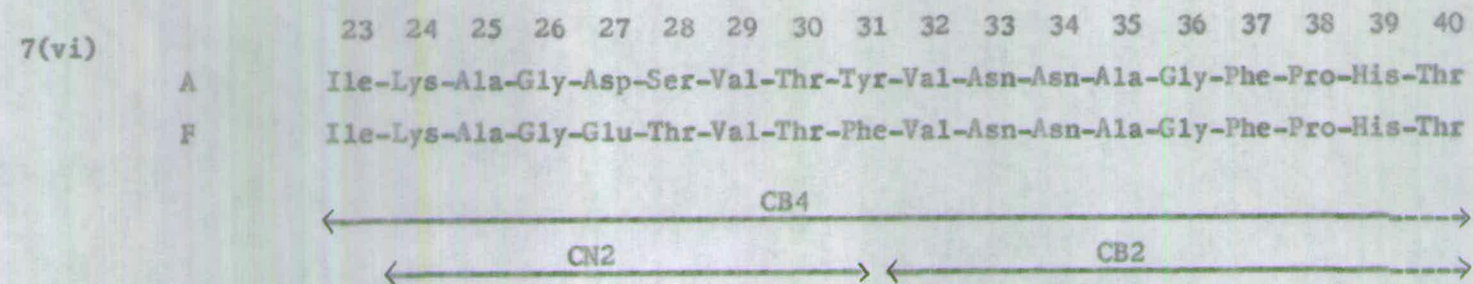
CN1: Asp-Val-Thr-Val-Lys-Leu

CN2: Lys-Ala-Gly-Glu-Thr-Val-Thr-Phe

Figures 7(iv) to 7(ix). Alignment of *C.fusca* plastocyanin chymotryptic peptides against the sequence of *C.pyrenoidosa* 'A' plastocyanin.







Figure

7(vii)

41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58

A Asn-Ile-Val-Phe-Asn-Ala-Asp-Glu-Asp-Ala-Pro-Val-Val-Gly-Ala-Asn-Ala-Leu

F Asn-Ile-Val-Phe-Gln-Gln-Asp-Glu-Asp-Val-Pro-(Ala,Ser,Gly,Ala,Asn,Ala)-Leu

⟵ CA5 ⟶

7(vii)

59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75

A Ser-His-Asp-Asp-Tyr-Leu- Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Thr-Ala-Lys-Phe

F Ser-His-Glu-Asp-Tyr-Leu- Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Ser-Ala-Lys-Phe

⟵ CA6 ⟶ ⟵ CA4 ⟶ ⟵ CB3 ⟶

7(ix)

76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

A Asp-Thr-Ala-Gly-Glu-Tyr-Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Lys-Ile-Val-Val-Gln

F Asp-Thr-Ala-Gly-Thr-Tyr-Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Lys-Gly-Thr-Ile-Thr-Val-Gln

⟵ CA3 ⟶ ⟵ CA8 ⟶ ⟵ CB1 ⟶

⟵ CA1 ⟶

⟵ CA9 ⟶

The amino acid residues of the bracketed regions of CA5 and CA7 are placed in the best possible alignment with the homologous regions in C.pyrenoidosa 'A' plastocyanin. From now on in the discussion let the symbol A denote C.pyrenoidosa 'A' plastocyanin and the symbol F be C.fusca. I shall begin the deduction of sequence at the N-terminus.

The N-terminus was determined by the DNS-chloride method and the 2nd and 3rd residues by the Dansyl Edman method. F CN1 and F CA2 have their first three residues in common with the first three residues of the whole protein. They must both be at the N-terminus (see Figure 7(iv)). To obtain the best alignment of A and F sequences it is necessary to postulate that F has an extra residue stuck on at the N-terminus of the protein. F CA2 overlaps F CN1 at its C-terminal residue. F CA7 has its N-terminal sequence in common with A TN3 (residues 13-24 in A sequence), and is clearly homologous with this region in the A sequence, except that there appears to be a deletion of two tyrosine residues in F CA7 (see Figure 7(v)). F CB4 corresponds in amino acid composition to residues 23 to 40 in the A sequence. F CB4 is equivalent to the sum of peptides F CB2 plus F CN2, where F CB2 is thought to be identical in sequence to A CB4 and F CN2 is clearly homologous with A CN2 (see Figure 7(vi)). F CA5 is clearly homologous with A CA1 (residues 41 to 58 in sequence). The amide groups in F CA5 have been

aligned in the same positions as they are found in A CA1, but I have no evidence that this is so. Figure 7(vii) shows the sequence homology between F and A in this region of the proteins.

F CA6 is homologous with A CA4 (residues 59 to 63 in sequence), F CA4 is identical in sequence to A CA5 (residues 64 to 71 in A sequence), and F CB3 is homologous with A CB1 (residues 72 to 75 in A sequence). The sequence homology between these peptides is shown in Figure 7(viii). F CA3 is homologous with A CA2 (residues 76-81 in A sequence) and F CA8 is identical in sequence to A CA6 (residues 82-93 in A sequence), The C-terminal peptide of F was isolated from a cyanogen bromide digest. It was identified as such, because it contained no trace of homoserine/homoserine lactone. Also, peptide CB1 has the same sequence. F CB1 is homologous with the C-terminal peptide of A (residues 94 to 100 in A sequence). Figure 7(ix) shows the C-terminal sequence.

Discussion of sequence

The alignment of the chymotryptic peptides formed from F against the postulated sequence of A is shown in Figures 7(iv) to 7(xi). All the peptides purified have been accounted for in the sequence.

From a consideration of the proposed sequence for F, the only place that the peptide S5N (formed from subtilisin digestion of F), with the postulated sequence Thr-Trp-Asn, could fit into the sequence is at Threonine-40. This would suggest that the true C-terminus of F CB2 is tryptophan, which is in agreement with the presumed specificity of chymotrypsin. This argument could also hold for the placement of a tryptophan residue in the same position in A, but this is purely speculative. The question of the plausibility and possible significance of the presence of a tryptophan residue in the algal plastocyanins will be discussed in the final chapter.

Non-specific cleavage of the protein chain by chymotrypsin occurred at Ile-23, in the hydrolysis of an Ile-Lys bond. This bond was also cleaved in A(by chymotrypsin), but very slowly.

The proposed sequence is in satisfactory agreement with the total amino acid composition of the protein (Chapter 4). The amino acid changes between the sequences of A and F will be discussed in the final chapter.

CHAPTER 8

PEPTIDES FORMED FROM CHYMOTRYPSIN DIGESTION OF S.OBLIQUUS 'A' PLASTOCYANIN

Introduction

A chymotryptic digest was performed on 100 mg. (about 10 μmoles) of protein. The conditions used for oxidation and digestion of the protein were the same as those described for chymotrypsin digestion of C.pyrenoidosa 'A' plastocyanin. The digest was fractionated, in the usual manner, by gel filtration through Sephadex G-25 and the peptides purified by paper electrophoresis at pH 6.5 and pH 3.5. Some peptides were also purified on paper chromatography (BAWP), and paper electrophoresis at pH 2.0.

Summary of chymotryptic peptides

Table 8(i) gives a list of the chymotryptic peptides with their amino acid compositions, N-termini, and electrophoretic mobidities at pH 6.5 (m) and pH 3.5 (m'). The yields, expressed as percentages (μmole peptide/μmole protein digested) are not corrected for any losses that most certainly occurred during the purification procedure. Contaminating amino acids are given when they are present in an amount of 0.2 mole/mole peptide or more.

TABLE 8(i): Chymotryptic peptides formed from S.obliquus 'A' plastocyanin.

	CA1	CA2	CA3	CA4	CA5	CA6	CA7	CA8	CA9	CA10	CB1	CB2	CB3	CB4	CB5	CN1	CN2	CN3
His			0.9	0.7		1.1								1.1	0.9		1.0	
Lys											1.0	1.0	1.0	1.0		0.8	1.1	0.9
Arg																		
Asp	1.0	1.0			3.7	2.3	0.9		0.8			1.0		3.0	2.0	1.0		1.9
Thr	*0.2						0.9≠	1.4		≠1.4	0.8			2.3	0.7	1.1		
Ser	0.9	1.0				1.0			0.9				1.1	1.3		0.9		1.0
Glu	*0.2	1.1	2.2	2.1	1.0		1.0	1.1	1.1	1.0							2.9	
Pro		1.0	0.9	1.0	1.1			1.0	0.9	1.0				1.2	1.0		1.0	
Gly	2.0	1.0	3.3	2.1	1.0		1.1		1.0				1.1	2.2	1.2	1.0	4.0	2.0
Ala	2.1	1.0	1.2	1.2	3.8		1.1	1.1	1.1	1.1	1.1	1.0		2.3	1.1	0.9	1.3	2.9
Val					1.9			1.8		0.9		1.0	2.7	≠1.4	0.4	0.9	≠2.5	0.9
$MetSO_2$			0.8	0.6													0.8	
$CySO_3H$			1.0	0.8													0.9	
Ile													0.8	1.0		0.8	≠0.7	
Leu	1.0	0.8			1.0							1.0						1.9
Tyr		0.6	0.7			0.8	0.5	1.2	0.8	+++				1.1		0.4	0.7	
Phe			1.0	1.0	0.8			0.9			1.1			1.3	0.9		0.9	

TABLE 8(i) - cont'd

N-terminus	Gly	Leu	Gly	Phe	Asx	Ser	Asx	Val	Asx	Glx	Thr	Ala	Val	Ile	Val	Ile	Gly	Ala
m	-0.40	-0.25	-0.16	-0.33	-0.72	-0.44	-0.34	-0.16	-0.34	-0.46	+0.39	+0.55	+0.33	+0.20	+0.41	0	0	0
m'	+0.20	+0.18	+0.12	+0.12	0	+0.33	0	+0.26	-	+0.22	+0.71	+0.66	+0.56	+0.40	-	+0.34	+0.29	+0.39
Yield	14.7	11.6	4.2	1.6	1.6	8.1	3.9	7.0	4.0	1.8	1.4	10.0	0.8	1.2	6.0	1.8	2.0	3.6

*Contaminating amino acids (amounts 0.2 mole/mole peptide are not shown)

≠Low yield due to partial hydrolysis of stable peptide bond.

Amino acid sequences of the chymotryptic peptides.

Most of the peptides were sequenced by the Dansyl Edman procedure. No peptide was formed that either did not have the same sequence as the corresponding peptide formed from C.pyrenoidosa 'A' plastocyanin (A) or would not fit into the sequence of the plastocyanin. The latter statement refers to peptides which were formed from S.obliquus 'A' plastocyanin (ϕ), and which had also been expected from the chymotrypsin digestion of C.pyrenoidosa 'A', but were not recovered. Each peptide is discussed.

Peptides CB2, CA1 and CN3.

From the mobility of CB2 on paper electrophoresis at pH 6.5, it is evident that the 2nd residue is asparigine. The sequence, determined by the Dansyl Edman procedure, is:

Ala-Asn-Val-Lys-Leu

Peptides CA1 and CN3 were sequenced by the Dansyl Edman procedure, and it was shown that CN3 is equivalent to the sum of CB2 and CA1, CB2 being N-terminal to CA1. Since CN3 runs in the neutral band at pH 6.5, the 8th residue in the sequence must be aspartic acid. The sequence of CN3 is:

Ala-Asn-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu

←——————— CN3 ———————→

←—— CB2 ——→ ←—— CA1 ——→

Peptide CB1

CPA released phenylalanine (+++). The sequence, determined by the Dansyl Edman method, is:

Thr-Ala-Lys-Phe

Peptide CA7

CA7 was sequenced by the Dansyl Edman method. After the 1st cycle of PTC degradation the remaining peptide with the amino acid composition (Thr, Ala,Gly, Glx,Tyr) retained a net acidic charge (m= 0.20) on paper electrophoresis at pH 6.5, but it was less than that of the CA7 (m = -0.70). Thus, in the sequence the N-terminus is aspartic acid and the 5th residue is glutamic acid. The sequence is:

Asp-Thr-Ala-Gly-Glu-Tyr

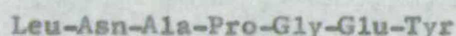
Peptides CA9 and CA2

CA9 and CA2 have the same amino acid compositions except that CA2 is one residue longer with leucine at the N-terminus. The observed electrophoretic mobility at pH 6.5 (m = -0.16) of CA2 is consistent with the predicted mobility (m = -0.18) for a peptide of this size with an electrostatic charge of 1 (Offord, 1966). CA2 was sequenced by the Dansyl Edman method, and Table 8 (ii) gives the electrophoretic mobilities at pH 6.5 of the corresponding peptides obtained after the 1st and 2nd cycles of PTC degradation.

TABLE 8(ii): Electrophoretic mobilities (m) of PTC degradation products of peptide CA2.

PTC degradation cycle	Peptide Sequence	m	Comments
0	Leu-Asx-Ala-Pro-Gly-Glx-Tyr	-0.16	
1	Asx-Ala-Pro-Gly-Glx-Tyr	-0.23	
2	Ala-Pro-Gly-Glx-Tyr	-0.21	Asn and Glu

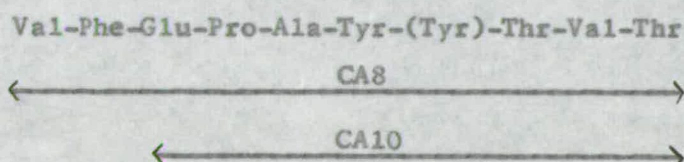
Thus, the sequence of CA2 is:


Leu-Asn-Ala-Pro-Gly-Glu-Tyr

Peptides CA8 and CA10

Peptides CA8 and CA10 were sequenced by the Dansyl Edman procedure. It was found that they have the same sequence except that CA8 is two residues longer having an extra Val-Phe sequence at the N-terminus. From the mobility on paper electrophoresis at pH 6.5 the 3rd residue in the sequence of CA8 (shown below) is glutamic acid.

Val-Phe-Glu-Pro-Ala-Tyr-(Tyr)-Thr-Val-Thr



The tyrosihe content of CA8 is an estimated value from the amount of chloro-tyrosine recorded on the amino acid analyser. It is rather surprising that chymotrypsin did not hydrolyse the tyrosine bonds, but it may be that it does not cleave the protein chain at modified

tyrosine residues. The yield of CA10 was low, suggesting that the phenylalanine bond was hydrolysed slowly, probably due to the close proximity of a proline residue.

Peptide CA6

CA6 was sequenced by the Dansyl Edman procedure. After two cycles of PTC degradation the resulting peptide with the amino acid composition (Asx,Asx,Tyr) had an observed electrophoretic mobility at pH 6.5 (m = -0.68) consistent with the predicted mobility (m = 0.63) for a peptide of this size with an electrostatic charge of 2. Thus, in the sequence (shown below) the third and fourth residues are both aspartic acid. The sequence of CA6 is:

Ser-His-Asp-Asp-Tyr

Peptides CB3,CN2,CA3, and CA4

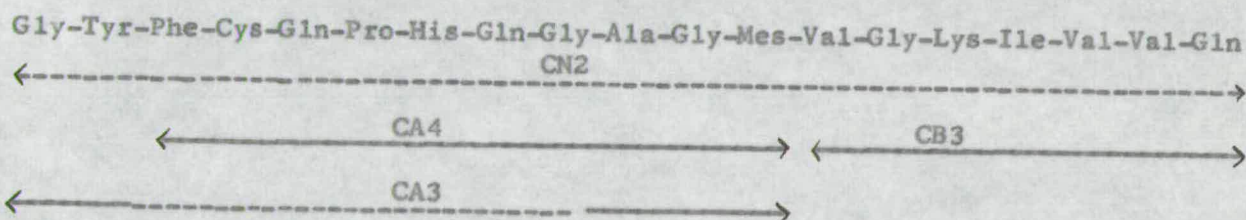
Peptide CB3 was sequenced by the Dansyl Edman method. From the mobility on paper electrophoresis at pH 6.5 the C-terminus must be glutamine. The sequence is:

Val-Gly-Lys-Ile-Val-Val-Gln

Peptides CN2, CA3 and CA4 all contain a single residue of cysteic acid, and since the protein contains only one cysteine residue, all three peptides must come from the same region of the protein. CA4 was sequenced by the Dansyl Edman method, showing the C-terminus to be methionine sulphone. CPA released methionine

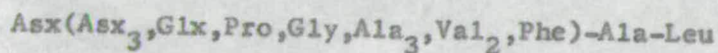
sulphone (++), glycine (+) and alanine from CA3. Since, CA4 and CA3 both have methionine sulphone at the C-terminus, the extra tyrosine residue that CA3 contains compared with CA4, must be placed in the 2nd position in the peptide, glycine being the N-terminal residue.

Peptide CN2 comprises the sum of the amino acid compositions of CA3 plus CB3, thus providing the overlapping sequence for these two peptides. Since CN2 runs in the neutral band on paper electrophoresis at pH 6.5 there must be three glutamine residues in the peptide. The sequence of CN2, as deduced from CA3,CA4 and CB3 is:



Peptide CA5

CPA released leucine (++) and alanine (+). CA5 was recovered in insufficient yield for sequence determination. The partial sequence is:

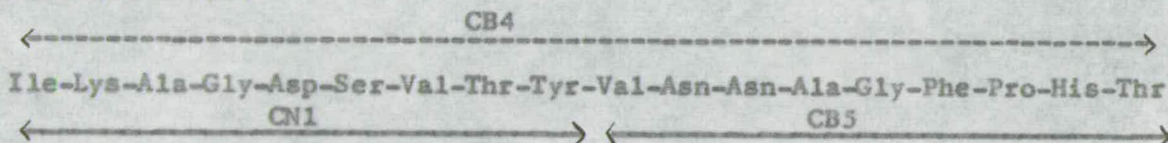


Peptides CN1, CB4 and CB5

Peptides CN1 and CB5 were sequenced by the Dansyl Edman method. Since CN1 runs in the neutral band on paper electrophoresis at pH 6.5,

the peptide contains aspartic acid (not asparigine). The net basic charge at pH 6.5 of CB5 is thought to be due to the histidine residue and that the 2nd and 3rd residues are both asparigines CB5 gave a positive test for tryptophan (using the Ehrlich reagent) on paper, but the tryptophan could not be located in the peptide since the protein had been oxidised.

CB4 is equivalent to the sum of the amino acid compositions of CN1 plus CB5, showing that CN1 is N-terminal to CB5 in the sequence shown below:



Since CB5 was sequenced through to the threonine residue, with a DNS-amino acid being clearly identified after each cycle of PTC degradation, if tryptophan is present in the peptide it must be at the C-terminus. This would be in agreement with the presumed specificity of chymotrypsin.

Discussion

A list of chymotryptic peptides formed from S.obliquus 'A' plastocyanin with the equivalent chymotryptic peptides formed from C.pyrenoidosa 'A' plastocyanin is given in Table 8(iii). Where no equivalent peptide was formed from chymotrypsin digestion of

C.pyrenoidosa 'A' plastocyanin, the residue numbers in the sequence (of C.pyrenoidosa 'A' plastocyanin), where the peptides would fit, are given. The S.obliquus 'A' plastocyanin peptides can account for the sequence of C.pyrenoidosa 'A' plastocyanin, except that the equivalent peptide to A CA1 (residues 41 to 58) was not found, although ØCA5 equivalent to ACA7 (residues 45-58) was found. Thus, the only region of the sequence of C.pyrenoidosa 'A' plastocyanin not accounted for by the S.obliquus 'A' plastocyanin peptides is between residues 41 to 44. However, the presence of residues 41 to 44 can be accommodated in the total amino acid composition of S.obliquus 'A' plastocyanin. I, tentatively, conclude that the sequences of the two plastocyanins are identical. The only discrepancy lies in the possible presence of a residue of tryptophan in S.obliquus 'A' plastocyanin (see peptide CB5), and this will be discussed in the final chapter.

TABLE 8(iii) Chymotryptic peptides formed from S.obliquus 'A' plastocyanin with the equivalent chymotryptic peptides from C.pyrenoidosa 'A' plastocyanin.

	S.obliquus 'A' plastocyanin peptides	C.pyrenoidosa 'A' plastocyanin peptides	Residue numbers in sequence
CA1	Gly-Ala-Asp-Ser-Gly-Ala-Leu		6-12
CA2	Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr	CA5	64-71
CA3	Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes	CA6	82-93
CA4	Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes		84-93
CA5	Asx(Asx$_3$,Gly,Glx,Pro,Ala$_3$,Val$_2$)-Ala-Leu	CA7	45-58
CA6	Ser-His-Asp-Asp-Tyr	CA4	59-63
CA7	Asp-Thr-Ala-Gly-Glu-Tyr	CA2	78-81
CA8	Val-Phe-Glu-Pro-Ala-Tyr-(Tyr)-Thr-Val-Thr	-	13-22
CA9	Asn-Ala-Pro-Gly-Glu-Ser-Tyr	CA3	65-71
CA10	Glu-Pro-Ala-Tyr-(Tyr)-Thr-Val-Thr	-	15-22
CB1	Thr-Ala-Lys-Phe	CB1	72-75
CB2	Ala-Asn-Val-Lys-Leu	CB3	1-5
CB3	Val-Gly-Lys-Ile-Val-Val-Gln	CB3	94-100
CB4	Ile-Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr-Val-Asn-Ala-Gly-Phe-Pro-His-Thr-(Trp)	CB5	23-40
CB5	Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Thr-(Trp)	CB4	32-40
CN1	Ile-Lys-Ala-Gly-Asp-Ser-Val-Thr-Tyr	CN2	23-31
CN2	Gly-Tyr-Phe-Cys-Gln-Pro-His-Gln-Gly-Ala-Gly-Mes-Val-Gly-Lys-Val-Val-Gln	-	82-100
CN3	Ala-Asn-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu	-	1-12

CHAPTER 9

DISCUSSION

A. Introduction

The sequence homology between the plastocyanins from the two species of Chlorella, and the proposed sequence homology between algal plastocyanins and bacterial azurins, will be discussed, with the possible consequences in terms of both molecular evolution and the evolutionary development of algae (or chloroplasts) from a bacterial source. The sequence homology between the algal plastocyanins and the partial sequence of Phaseolus vulgaris plastocyanin (Milne, 1967) will also be discussed.

Several points arose out of the work described in this thesis, notably the purification of two forms of plastocyanin, in differing ratios, in both the Chlorella species and in Scenedesmus. The relative mobilities, on starch gel electrophoresis, of each of the plastocyanins has been shown (Figure 4(ii)). The two distinct forms of C.pyrenoidosa plastocyanin, termed 'A' and 'B' plastocyanin, were well separated by DEAE-cellulose chromatography, and were isolated in approximately equal amounts. On purification of plastocyanin from both C.fusca and S.obliquus, a second minor band blue band was observed on DEAE-cellulose. This proved exceedingly difficult to

purify from the major blue component. In the case of S.obliquus plastocyanin, the minor form was separated fairly well from the major form, but the starch gel pattern showed that it was still contaminated (5-10%) with the major form. It was estimated that S.obliquus 'B' plastocyanin represented approximately 16% of the total plastocyanin purified from Scenedesmus. It was not possible to quantitate the minor form of C.fusca plastocyanin.

Of interest at this stage in the discussion is the finding of Katoh et al (1962). During the purification of spinach plastocyanin they observed that, on absorption of the plastocyanin to DEAE-cellulose the blue band separated into a major band, which moved faster down the column, from a minute but distinct band that stuck to the top of the column, and could only be eluted with a higher salt concentration. The nature of the minor blue band has not yet been studied. This observation parallels the conditions in which the two forms of plastocyanin from both C.pyrenoidosa and Scenedesmus were separated. The minor forms were only eluted with a higher salt concentration than that necessary for the elution of the major forms.

One would like to be able to offer some explanation as to why the ratio of the two forms of plastocyanin is species variable. A situation does exist in wild type Baker's yeast where two distinct forms of cytochrome c are present, a major component iso-1-cytochrome c and a minor component iso-2-cytochrome c

(approx. 5% of total). Sherman, Taber & Campbell (1965) showed that the concentration of cytochrome c in yeast is under complex genetic control. The examination of fifteen mutants (which had varying degrees of cytochrome c deficiencies) revealed six unlinked genes that decrease the absolute concentration of cytochrome c and its concentration relative to other cytochromes. A similar situation could exist in green algae, such that the relative amounts of A-type and B-type plastocyanin, produced in the cell, was under complex genetic control.

A point of interest that poses a real problem is the finding that the sequences of C.pyrenoidosa 'A' plastocyanin and S.obliquus plastocyanin appear to be identical. Their mobilities on starch gel electrophoresis were the same (within the limits of experimental error) and from chymotrypsin digestion of both proteins I was unable to detect any amino acid differences. It is very surprising to find two plastocyanins, purified from different species of alga, to have apparently identical amino acid sequences, when several differences have been located between C.pyrenoidosa 'A' and 'B' plastocyanins and between these two and C.fusca plastocyanin. Also, this is inconsistent with sequence studies on other proteins, for example the percentages differences in sequence between subtilisins from related strains of B.subtilis is approximately 30% (Smith et al. (1966)). However, I have no idea how close Scenedesmus and Chlorella are in evolutionary terms.

One could suggest from these results that in fact the original cultures of C.pyrenoidosa and S.obliquus were contaminated and that C.pyrenoidosa 'A' plastocyanin was produced from contaminating Scenedesmus cells. It would mean that the level of contamination with Scenedesmus cells would have had to be very high. Electron microscope studies of rehydrated cells of each of the cultures obtained from Czechoslovakia, showed that a single cell type was present in each of the cultures (Hobbs, D. personal communication).

An attempt was made to eliminate the possibility that the algal cultures were contaminated with a second species of alga, by growing a laboratory culture of C.pyrenoidosa (211/8h Emerson, No.3), and to see if this strain of Chlorella produced two kinds of plastocyanin molecule. Unfortunately this proved abortive, since it was impossible to grow the algal cultures up to a sufficient density for harvesting, partly because the cultures were prone to bacterial contamination (even under stringent sterile conditions), the bacteria eventually outgrowing the algae. Thus contamination being the reason for the presence of two forms of plastocyanin from each of the species of alga, has not been eliminated. It would offer the simplest explanation for the finding, but considering the proposed biological function of plastocyanin, as an electron carrier in the photosynthetic electron transport chain, it may be possible for such a system to accomodate two plastocyanins, acting either in series or in parallel. If they do not act in series it might be expected that they would differ in their redox potentials, which could be readily tested.

B. Elucidation of the plastocyanin sequences

The general experimental approach to the elucidation of each plastocyanin sequence was the same; the proteins being examined by trypsin and chymotrypsin digestion. The sequences of the large tryptic peptides were established by degrading the peptides with chymotrypsin, positioning the chymotryptic peptides in the sequence from a knowledge of the N-terminal sequence of the parent peptide, and of the bonds in the sequence slightly susceptible to chymotrypsin. The Dansyl Edman procedure was used to elucidate sequences whenever possible.

Cyanogen bromide digestion (Gross et al, 1961) was used as a convenient method of determining the C-terminal sequence of each plastocyanin, making use of the knowledge that each protein contained only a single residue of methionine per molecule.

The yields of tryptic and chymotryptic peptides were generally fairly low and variable, and no attempts were made to correct the figures for losses that certainly occur during the purification procedure, such as material lost on marker strips and material not eluted from paper. Because of these losses the yields calculated are of limited value. Low and variable yields of peptides seem to be characteristic of paper purification techniques.

Amide residues were recognised, where possible, by the electrophoretic mobilities of peptides at pH 6.5. With the eighteen residue peptide CA1 from C.pyrenoidosa A plastocyanin (A), which yielded five residues of aspartic acid and one residue

of glutamic acid on hydrolysis, the number and position of the amide groups were determined by following the change in electrophoretic mobility (at pH 6.5) of the peptide, as residues were successively removed by PTC degradation. For other peptides (e.g. A CA5 and A CA6) the region containing the residue was degraded to as small a peptide as possible (containing only one acidic or amide residue), and the assignment made on the basis of whether the electrophoretic mobilities at pH 6.5 were positive, negative, or neutral (i.e. very near zero).

The position of the amide residues in C.pyrenoidosa A plastocyanin have been located by the above mentioned methods. Only those amide residues in C.fusca plastocyanin (F), which are evident from electrophoretic mobilities of peptides, have been included in the sequence, except for the amide groups in peptides FCA4 and FCA9, which have been made on the basis that they have identical sequences and similar electrophoretic mobilities as peptides A CA6 and A CA9, respectively. Only those amide residues evident from electrophoretic mobilities of peptides have been located in C.pyrenoidosa B plastocyanin (B).

No evidence of any labile amide groups was found in any of the plastocyanins, a property which appears to be common to Asn-Gly sequences (Ambler, 1963b; Haley & Corcoran, 1967; Meadway, 1969).

In the tables of amino acid compositions of peptides, contaminating amino acids are shown when they are present in the ratio of 0.2 mole/mole of peptide or more. The majority of peptides contained a level of impurity of less than 0.08 mole/mole of peptides.

The main difficulty in the sequence determinations was the satisfactory purification of the complex mixture of closely related acidic peptides which were formed, especially in the chymotryptic digests. They were generally recovered in fairly low yields, the major loss probably occurring in trying to separate peptides of very similar electrophoretic mobilities from one another, ready for elution.

The most difficult region of the proteins to elucidate was between residues 40 to 58 in C.pyrenoidosa 'A' and C.fusca plastocyanins, and residues 39 to 59 in C.pyrenoidosa 'B' plastocyanin. These regions extend over homologous sequences. The reason for the difficulty is that these regions were not satisfactorily split up into smaller peptide sequences. There is, therefore, no confirmation of the sequence data obtained from the Dansyl Edman procedure. My experience is that this method can be trusted on pure peptides, and I feel reasonably confident that the peptides were pure.

Specificity of proteases used

The trypsin used for the digestions, produced peptides from C.pyrenoidosa 'A' and 'B' proteins as a result of chymotryptic-like cleavage. The yields of these peptides were not appreciably lower than the peptides formed from the splitting of lysyl and arginyl bonds. In addition, trypsin cleaved the C.pyrenoidosa 'A' chain at Asn-57 (an Asn-Ala bond) (see Figure 9(i)). This bond was not cleaved by chymotrypsin, so it would seem that the splitting of the aspariginyl bond was not caused by contaminating chymotrypsin.

Chymotrypsin hydrolysed the expected aromatic and leucyl bonds. In addition, the Thr-Ile bond residues (22 to 23) in the C.pyrenoidosa 'A' plastocyanin sequence was hydrolysed fairly slowly, and the Ile-Lys bond (position 23 to 24), extremely slowly. The same bonds were cleaved in C.fusca plastocyanin. The Ser-Lys bond (position 24 to 25) in C.pyrenoidosa 'B' plastocyanin was hydrolysed slowly, and the methionine sulphone-valine bond (residues 94 to 95 in Figure 9(i() was hydrolysed in all three proteins. The splitting of these bonds, although surprising, is not unprecedented since many cases have been described of chymotryptic cleavage of several bonds originally thought to be resistant to chymotrypsin (Hill, 1965). The expected chymotryptic peptides, were isolated from the proteins, excepting for C.pyrenoidosa 'A' plastocyanin where the expected chymotryptic peptides between position 6 to 22 were not recovered. Splitting at leucine-12, tyrosine-18 and -19 and phenylalanine-14 should have occurred, from the presumed specificity of chymotrypsin.

Figure 9(i) Algal plastocyanin sequences

```
                              1                             2                             3
   1                          0                             0                             0
1  a. n. V. K. L. G. A. d. s. G. A. L. V. F. e. P. a. y. y. T. v. t. I. k. a. G. d. s. V. T. y. V. N. N. A. G. F. P. H.

2  a. t.                b.(b.    ).          z.(   g. -  -     n. s).   s. k.    e. s.       f.

3d.v. t.                d. s.                e.    s. -  -(    v. s).   k. a.    e. t.       f.

   4                      5                                   6                             7
   0                      0                                   0                             0
1  t. w. N. I. V. F. n. a. D. E. D. a. P. v. v. G. A. n. A. L. S. h. d. D. Y. L. N. A. P. G. e. s. y. t. a. K. F. D. T.

2  -  -              n.(a.          s.    v. v.       q.    i.    r. e.                      z. z. f. t. v.    1.

3  t. w              q. q.          v.   (a. s.       n.  ) 1.    h. e.                      e. s. y. s. a.    f.
                                                               1
   8                   9                                      0
   0                   0                                      0
1  G. e. Y. G. Y. F. C. Q. P. H. Q. G. A. G. M. v. G. k. I. v. V. q.

2     e.                                        v.    r.    t.    n.

3     t.                                        k.    t.    t.    q.
```

1. C. pyrenoidosa 'A'
2. C. pyrenoidosa 'B'
3. C. fusca

C. Sequence homology between the algal plastocyanins

The sequences of C.pyrenoidosa 'A' and 'B', and C.fusca plastocyanins have been deduced (Figure 9(i)), and the sequences discussed in the chapters dealing with each protein. With the sequences aligned as shown in Figure 9(i), 64% of the residues are in identical positions, and this percentage might be higher when all the amide groups in the C.pyrenoidosa 'B' and C.fusca sequences have been positioned. Most amino acid substitutions can be explained on the basis of single base changes in the codon, a few require double base changes, but none require three base replacements. In the main the amino acid substitutions are conservative changes, except that in the alignment it was found that the positions of lysine residues were variable within a confined region of the sequence. This was also observed by Harris & Perham (1968) in their study of the sequence homology between mammalian and lobster glyceraldehyde 3-phosphate dehydrogenase.

One region of the sequence (residues (17-41) may have undergone more complex changes since deletions of tyrosine-18 and -19, threonine-40, and tryptophan-41 from the C.pyrenoidosa 'A' sequence have been postulated for C.pyrenoidosa 'B' protein, and deletions of tyrosine-18 and -19 for C.fusca plastocyanin sequence.

The absence of tyrosine from peptides B CA3 (residues 13 to 24) and FCA7 (residues 13-23), on the quantitative amino acid analysis, was not altogether convincing, since determination of the tyrosine content of peptides had proved difficult, due to the

modification of some of the tyrosines, by halogenation. However, there is no evidence for the presence of tyrosine in these two peptides. They did not stain positive for tyrosine on paper (Jepson et al. 1953), while other peptides on the same marker strip did, which rules out the possibility that the reagent was ineffective. If tyrosine were present in positions 18 and 19, then one would have expected chymotrypsin to have hydrolysed the tyrosine bonds, as was the case in the trypsin digestion of C.pyrenoidosa 'A' plastocyanin (see peptides A TN3, TB2, and TA1, Chapter 5); or it could be that chymotrypsin does not split very readily after modified tyrosine residues. With no evidence for the presence of tyrosine it is proposed that tyrosine-18 and -19 (in the A sequence) have been deleted in the B and F sequences. One explanation for the deletion could be that amino acid residues substituted for tyrosine-18 and -19 changed the conformation of the protein such that it was inactive, and that by deletion of the two residues the original or closely related conformation could be resumed.

Tryptophan content

Tryptophan has been reported absent from spinach (Katoh et al, 1962) and Phaseolus vulgaris (Milne & Wells, 1970) plastocyanins, the evidence being amino acid compositions, tryptophan determinations, and u.v. spectra.

In the early stages of the investigation, on the algal plastocyanins, the presence of tryptophan was suggested from the subsidiary maximum at 292 nm., observed in the u.v. spectra of all the native proteins, excepting C.pyrenoidosa 'B' plastocyanin (Figure 4(iii)). For C.fusca protein the absorption at 280 nm. was consistent with the protein containing 4 residues of tyrosine and one of tryptophan. Determination of tryptophan content of S.obliquus 'A' and C.fusca plastocyanins, by the Spies & Chambers (1948) method, detected 0.4 residue/molecule and 0.52 residue/molecule, respectively, when the proteins were first digested with pepsin. A peptide (FS5N, see Chapter 7) was purified from subtilisin digestion of C.fusca plastocyanin that stained positive for tryptophan and was strongly fluorescent (on oxidation) under ultraviolet light. Although not conclusively proven it was proposed that the sequence of the peptide is Thr-Trp-Asn, which can only fit into the sequence at threonine-40.

Indirect evidence for the presence of tryptophan in position 41 in C.pyrenoidosa 'A' sequence has been obtained. It has been suggested (Chapter 8) that the true C-terminus of peptide CB5 from S.obliquus 'A' plastocyanin (Ø), is tryptophan and not threonine. ØCB5 is identical in sequence to peptides A CB4 (residues 32 to 40 in sequence) and FC B2, so if tryptophan is also the C-terminus of these peptides (and not threonine), then this would position tryptophan as residue 41 in the A sequence. Unfortunately peptides

ACB4 and FCB2 were not stained for tryptophan, and neither were their C-termini determined, using CPA, due tollack of material. The significance of the presence of tryptophan in the algal plastocyanins will be discussed in the section dealing with bacterial azurins.

It might beeof significance that C.pyrenoidosa 'B' plastocyanin not only has threonine-40 deleted, but also tryptophan-41. C.pyrenoidosa 'B' plastocyanin also differs from the other two algal plastocyanins, spinach, and french bean plastocyanins, in containing two residues of arginine.

There is no overlap of the peptide bond joining residues 41 to 42 (Trp-Asn bond), in the sequences of C.pyrenoidosa 'A' and C.fusca proteins, or residues 39 to 40 (His-Asn- bond) in C.pyrenoidosa 'B' protein.

D. Sequence homology between the algal and P.vulgaris (french bean) plastocyanin.

The partial sequence of P.vulgaris plastocyanin has been determined (Milne, 1971). The sequences of the plastocyanins have been aligned (Figure 9(ii)), the amino acids in the bracketed regions of the french bean sequence bemg re-ordered to maximise the alignment.

Figure 9(ii) Sequence homology between algal plastocyanin and Phaseolus vulgaris plastocyanin

```
   1                          10                            20                            30
1. a. n. V. k. L. G. a. d. s. G. a. L. V. F. e. P. a. y. y. t. v. t. i. k. a. g. d. s. v. t. y. v. n. n. A. G. F. P. H.

2. a. t.    k.       a. d.(n.   a)          e.    g. -  -  t. n. s. i. s. k. g. e. s. v. t. f. v. n. n.

3. v. t.    k.       a. d. s.    a.          e.    s. -  - (t. v. s).i. k. a. g. e. t. v. t. f. v. n. n.

4. l. e.    l.       s. g. d.    s.          v.    s. e. f. s. v. p. s. g. e. k. i. -  -  v. f. k. b. b.

   40                            50                               60                                  70
1. t. w. n. I. V. F. n. a. d. e. d. a. P. v. v. G. A. n. A. -. l. S. -. -. h. d. d. y. L. n. A. P. g. e. s. y. t. a. k.

2. -. -. n.          n.(a. d. e. d. s.    v. v. )     q.    -. i.    -. -. r. e. d. y.    n.       g. z. z. f. t. v. k.

3. t. w. n.          q. q. d. e. d. v.    (a. s.       n.    ) -. l.    -. -. h. e. d. y.    n.       g. e. s. y. s. a. k.

4.(-. -. z.          b. b. b. z. b. v.    v. v.       z.    ) k. i.    m. p. z. z. z. l.    b.    .  z. z. t. y. v. v. t.

               80                            90                            100
1. F. D. T.a. G. e. Y. g. y. F. C. q. P. H. q. G. A. G. M. v. G. k. i. v. V. q.

2. l.       a.    e.    g.(y. F.    q.       q.             v.    r. i. t.    n.

3. F.       a.    t.    g. y. F.    q.       q.             k.    t. i. t.    q.

4. l.       k.    t.    s. F. y.    s.       z.             v.    k. v. t.    n.
```

1. C. pyrenoidosa 'A'
2. C. pyrenoidosa 'B'
3. C. fusca
4. Phaseolus vulgaris (from Milne, 1971).

With the inclusion of a tryptophan residue at position 41 in the C.pyrenoidosa 'A' sequence, the algal plastocyanin chain is 101 residues, compared with 100 residues for french bean plastocyanin.To obtain the maximum homology it was necessary to postulate four deletions from the C.pyrenoidosa 'A' sequence at positions 28, 29, 40 and 41, and three insertions at positions 59, 62 and 63. It should be pointed out that the deletion of residues 40 and 41 from the C.pyrenoidosa 'A' sequence occurs in a region of the sequence in which no overlap has been found. This match means that 37% of the residues are in identical positions; this being a minimum figure since very few amide groups have been located in the french bean sequence. It would be unwise to draw too many conclusions from the proposed sequence homology, except to draw attention to the observation that, in the sequence comparison, there are large areas which have obviously been open to considerable change, while other smaller regions of the proteins (e.g. near the C-terminus) have been strongly conserved. The most interesting conserved region is that around the single cysteine residue, of the type Tyr-X-(Tyr,Phe)-Cys-, which is also close to a histidine residue. It might be that this region of the protein provides one of the copper-binding ligands, since the involvement of a sulphydryl group in the copper-binding site of plastocyanin has been proposed (Katoh et al, 1964).

The percentage similarity between the algal and french bean plastocyanin sequences is lower than that found between the sequences of the algal plastocyanins. It would be interesting to study

the structure of plastocyanin from blue-green algae (Lightbody et al., 1967), to see if the plastocyanin sequences from green algae, provide a link between plastocyanin from blue-green algae and that from higher plants (e.g. french bean).

E. Sequence homology between algal plastocyanins and bacterial azurins.

It has been established in the Introduction (Chapter 1) that plastocyanin and azurin are functionally analagous proteins. The question to be asked now is are there sufficient similarities in the amino acid sequences of the two proteins to suggest that they may have been derived from a common ancestor, or is any similarity found merely a result of convergence, rather than divergence.

Because of the difference in lengths of the two peptide chains (128 to 129 residues for azurin and 101 to 99 for plastocyanin) it is difficult to align the sequences without the help of a computer program to search for matchingpositions, and to assess the significance of the matches chosen. Computer techniques for establishing sequence homology between related proteins has been used extensively (Fitch, 1966; Dus, Sletton & Kamen, 1968; Needleman & Blair, 1969; Ambler & Meadway, 1969). By inspection, the sequences of plastocyanin and azurin (except for the C-terminal region), are not obviously similar, and it is felt that further plastocyanin sequences need to be determined, from more diverse

Figure 9(iii) Proposed homology between C-terminal regions of bacterial azurins and algal plastocyanins

		107			110										120									
	1	q.	Y.	m.	f.	F.	C.	t.	F.	P.	g.	h.	s.	a.	l.	M.	k.	G.	t.	l.	t.	l.	k.	
	2	a.	.	a.	f.	.	.	s.	f.	.	g.	h.	s.	a.	m.	.	k.	.	t.	l.	t.	l.	k.	
	3	k.	.	g.	f.	.	.	s.	f.	.	g.	h.	i.	s.	m.	.	k.	.	t.	v.	t.	l.	k.	
	4	k.	.	m.	f.	.	.	s.	f.	.	g.	h.	i.	a.	m.	.	k.	.	t.	v.	t.	l.	k.	
Azurins	5	s.	.	e.	f.	.	.	s.	f.	.	g.	h.	n.	s.	m.	.	k.	.	a.	v.	v.	l.	k.	
	6	d.	.	t.	f.	.	.	s.	f.	.	g.	h.	g.	a.	l.	.	k.	.	t.	l.	k.	l.	v.	d.
	7	a.	.	a.	y.	.	.	.	f.	.	g.	h.					k.	.	t.	l.	k.	l.	S.	n.
	8	d.	.	a.	f.	.	.	s.	f.	.	g.	h.	w.	s.	i.	.	t.	.	e.	i.	k.	l.	q.	s.
	9	d.	.	a.	y.	.	.	s.	f.	.	g.	h.	f.	a.	l.	.	k.	.	v.	l.	k.	l.	v.	d.
	10	e.	.	g.	y.	.	.	q.	-.	.	h.	q.	g.	a.	g.	.	v.	.	k.	i.	v.	v.	q.	
Plastocyanins	11	e.	.	g.	y.	.	.	q.	-.	.	h.	q.	g.	a.	g.	.	v.	.	r.	i.	t.	v.	n.	
	12	t.	.	g.	y.	.	.	q.	-.	.	h.	q.	g.	a.	g.	.	k.	.	t.	i.	t.	v.	q.	
		81										90										100		

1. Pseudomonas aeruginosa P6009)
2. Pseudomonas denitrificans NCIB 9496)
3. Pseudomonas fluorescens B-93 (ATCC 17467)) (from Ambler, 1968 and personal communication
4. C-18 (ATCC 17400))
5. D-35 (ATCC 17414))
6. Bordetella bronchiseptica NCTC 8344)
7. Alcaligines denitrificans NCTC 8582)
8. Alcaligines faecalis NCIB 8156)
9. Alcaligines spp. (Iwasaki's Pseudomonas denitrificans))
10. C.pyrenoidosa 'A'
11. C.pyrenoidosa 'B'
12. C.fusca

photosynthetic organisms, before it would be worthwhile using a computer program to attempt to match them.

The C-terminal sequences (see Figure 9(iii)) of the two proteins do have certain characteristics in common, particularly the homologous sequences around the single cysteine residue present in both azurin (position 112) and plastocyanin (position 86). The cysteineresidue is present in a strongly hydrophobic environment, which is of the type Tyr-X-Tyr-Phe-Cys-, where X is a 'neutral' amino acid. Also, a histidine residue lies close to the cysteine residue.

The involvement of a sulphydryl group in the copper-binding site of both plastocyanin (Katoh & Takimiya, 1964; Blumberg et al. 1966) and azurin (Brill et al. 1968; Finnazzi Agro et al, 1970) has been repeatedly proposed. If this were so then the cysteine residue in the two proteins would be the only candidate.

The tetrapeptide sequence Thr-Ala-Gly-Glu found in one of the azurins, (Pseuodomonas fluorescens D-35 (ATCC17414) residues 103 to 106 in Figure 1(i)) (Ambler, personal communication), so far studied, is also present in C.pyrenoidosa 'A' and 'B' plastocyanins (residues 78 to 81 in Figure 9(i)).

Finnazzi Agro et al (1970) and Rotilio et al (1970), have tentatively proposed that in addition to the cysteinyl residue, a tryptophanyl residue is also involved in the copper-binding site of azurin. Tryptophan is present in all azurins, so far studied, except one (Pseuodomonas fluorescens B-93 (ATCC 17467))(Ambler, 1968 and personalcommunication), but it is not in an invariant position. However, in seven out of nine bacterial azurins studied, the tryptophan is in position 48 (Figure 1(i)). It may be significant

that the proposed sequence (His-Thr-Trp-Asn), around the tryptophan residue in C.pyrenoidosa 'A' and C.fusca plastocyanin sequences is similar to the sequence (His-Asn-Trp-Val) around tryptophan-48 in azurin. It is unfortunate that the presence of tryptophan in some plastocyanins has not been conclusively proven, but indirect evidence for the presence of tryptophan has been obtained from the U.V. spectra, since the subsidiary maximum at 292 nm., attributed to the fine structure of tryptophan, is present. This is also a common feature of the U.V. spectrum of azurin (Ambler, 1967).

If tryptophan and cysteine are involved in the binding of copper to azurin and plastocyanin, then it might be suggestive that the two proteins may have a similar tertiary structure, which would support a structure-function relationship. However, a number of X-ray crystallographers, working on metallo-proteins, have recently reported at the 1971 Cold Spring Harbor Symposia on Quantitative Biology (in press) the difficulty in relating structure and mechanism, since it has been found for zinc-containing proteins (e.g. carbonic anhydrase C, carboxypeptidase A, and insulin) that the geometry around the active sites is not the same.

Thus, although similarities in specific regions of the two proteins can be found, it would be unwise to suggest that azurin and plastocyanin are derived from a common ancestor; the similarities

might merely be a feature of their function, and may have arisen by convergent evolution. Again, the determination of the sequence of a blue-green algal plastocyanin could be very interesting, as an approach to finding a link between bacterial azurin and green algal (or higher plant) plastocyanin. This type of experiment might help to test the endosymbiotic theory of chloroplast development, which proposes that the chloroplast originated from inclusion of a prokaryotic (e.g. bacterium) organism into a eukaryotic cell (Smillie & Steele Scott, 1969).

REFERENCES

AMBLER, R.P. (1963b). Biochem.J., 89, 349.

AMBLER, R.P. & BROWN, L.H. (1967). Biochem.J. 104, 784.

AMBLER, R.P. (1968). In "Chemotaxonomy and Serotaxonomy", vol.2, pp.57-64. Ed. Hawkes, J.G. London: Academic Press Inc.Ltd.

AMBLER, R.P. & MEADWAY, R.J. (1969). Nature, Lond. 222, 24.

ARNON, D.I., CHAIN, R.K., McSWAIN, B.D., TSUJIMOTO, H.Y., KNAFF, D.B. (1970), Proc.Nat.Acad.Sci., U.S.A., 67, 1404.

BEAVEN, G.H. & HOLIDAY, E.R. (1952). Advanc.Protein Chem. 7, 320.

BENSON, A.M. & YASUNOBU, K.T. (1969). J.biol.Chem., 244, 955.

BENSON, J.V. & PATTERSON, J.A. (1965). Analyt.Chem., 37, 1108.

BLOW, D.M., BIRKTOFT, J.J. & HARTLEY, B.S. (1969). Nature, Lond. 221,337.

BLUMBERG, W.E. & PEISACH, J. (1966). Biochim.biophys.Acta. 126, 269.

BRILL, A.S., BRYCE, G.F. & MARIA, H.J. (1968). Biochim.biophys.Acta., 154, 342.

BROMAN, L., MALMSTRÖM, B.G., AASA, R. & VÄNNGARD, T., (1963). Biochim. biophys.Acta., 75, 365.

CANTOR, C.R. & JUKES, T.H. (1966). Proc.natn.Acad.Sci. U.S.A., 56, 177.

COMMISSION ON BIOCHEMICAL NOMENCLATURE (1967). Biochem.J.102, 23.

COMMISSION ON BIOCHEMICAL NOMENCLATURE (1969). Biochem.J.113, 1.

DALGLIESH, C.E. (1952). Biochem.J. 52, 3.

DENT, C.E. (1947). Biochem.J., 41, 240.

DETERMANN, H. & WALTER, I. (1968). Nature, Lond. 219, 604.

DEVOR, A.W. (1950). J.Am.chem.Soc., 72, 2008.

DICKERSON, R.E. (1971). J.molec.Biol., 57, 1.

DUS, K., SLETTON, K. & KAMEN, M.D. (1968). J.biol.Chem., 243, 5507.

EDMAN, P. (1950). Acta.chem.Scand., 4, 283.

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

ERLANGER, B.F. & COHEN, W. (1963). J. Am.chem.Soc., 85, 348.

FEE, J.A. & MALMSTRÖM, B.G. (1968). Biochim.biophys.Acta, 153, 299.

FINAZZI AGRO, A., ROTILIO, G., AVIGLIANO, L., GUERRIERI, P.,
BOFFI, V., & MONDOVI, B. (1970). Biochemistry, 9, 2009.

FITCH, W.M. (1966). J.molec.Biol., 16, 9.

FRAENKEL-CONRAT, H., HARRIS, J.I. & LEVY, A.L. (1955). Meth.biochem.
Anal. 2, 339.

GORMAN, D.S. & LEVINE, R.P. (1965). Proc.natn.Acad.Sci., U.S.A., 54, 1665.

GORMAN, D.S. & LEVINE, R.P. (1966). Plant Physiol. 41, 1637.

GRAY, W.R. & HARTLEY, B.S. (1963a). Biochem.J., 89, 379.

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

GRAY, W.R. (1967). In Methods in Enzymology, Vol.11, p.139. Ed. Hirs,C.H.W.
New York: Academic Press Inc.

GROSS, E. & WITKOP, B. (1961). J.Amer.chem.Soc., 83, 1510.

HALEY, E.E., CORCORAN, B.J., DORER, F.E. & BUCHANAN, D.L. (1966). Biochemistry, 5, 3229.

HARRIS, J.I. & HINDLEY, J. (1965). J.molec.Biol., 13, 894.

HARRIS, J.I. & PERHAM, R.N. (1968). Nature, Lond. 219, 1025.

HARRISON, P.M. & HOFFMAN, T. (1961). Biochem.J., 80, 38P.

HILL, R.E. (1965). Advanc.Protein Chem., 20, 37.

HIRS, C.H.W. (1956). J.biol.Chem., 219, 611.

HIRS, C.H.W. (1967). In Methods in Enzymology, vol. 11, p.197. Ed. Hirs, C.H.W. New York: Academic Press Inc.

HORIO, T. (1958). J.Biochem., Tokyo, 45, 195.

JAMES, B.R. & WILLIAMS, R.J.P. (1961). J.Chem.Coc., 2007.

JEPSON, J.B. & SMITH, I. (1953). Nature, Lond. 172, 1100.

KATOH, S. (1960). Nature, Lond. 186, 533.

KATOH, S., SUGA, I., SHIRATORI, I. & TAKAMIYA, A. (1961). Arch. Biochem.Biophys., 94, 136.

KATOH, S., SHIRATORI, I. & TAKAMIYA, A. (1962). J.Biochem., Tokyo, 51, 32.

KATOH, S. & TAKAMIYA, A. (1964). J.Biochem., Tokyo, 55, 378.

KENDREW, J.C, DICKERSON, R.E., STRANDBERG, B.E., HART, R.G., DAVIES, D.R., PHILLIPS, D.C., & SHORE, V.C. (1960). Nature, Lond. 185, 422.

KERESZTES-NAGY, S., PERINI, F., & MARGOLIASH, E. (1969). J.biol.Chem., 244, 981.

KIRK, J.T.O. (1966). In Biochemistry of Chloroplasts, vol.1, p.319. Ed. Goodwin, T.W. London & New York: Academic Press.

KRISHNA RAO, K. & MATSUBARA, H. (1970). Biochem.biophys.Res.Commun., 38, 500.

LEVY, A.L. & CHUNG, D. (1953). Analyt.Chem., 25, 396.

LIGHTBODY, J.J. & KROGMANN, D.W. (1967). Biochim.biophys.Acta., 131, 508.

MALKIN, R. & MALMSTROM, B.G. (1970). In Advances in Enzymology, Vol.33, p.177, Interscience Publishers.

MALMSTROM, B.G., REINHAMMER, B. & VANNGARD, T. (1970). Biochim.biophys. Acta, 205, 48.

MARGOLIASH, E. (1963). Proc.natn.Acad.Sci., U.S.A., 50, 672.

MARGOLIASH, E. & NOLAN, C. (1968). Ann.Rev.Biochem., 37, 727.

MARGOLIASH, E., PERINI, F. & KERESZTES-NAGY, S. (1967). Proc. 7th Int.Congr.Biochem., Tokyo, General Session A.68, p.595.

MATSUBARA, H., SASAKI, R.M. & CHAIN, R.K. (1967). Proc.natn.Acad.Sci., U.S.A., 57, 439.

MASON, H.S. (1963). Biochem.biophys.Res.Commun., 10, 11.

MEADWAY, R.J. (1969). Ph.D. Thesis: University of Edinburgh.

MICHL, H. (1951). Mh.Chem., 82, 489.

MILNE, P.R. (1971). Ph.D.Thesis: University of Adelaide.

MILNE, P.R. & WELLS, J.R.E. (1970). J.biol.Chem., 245, 1566.

NAKAMURA, T. (1958). Biochim.biophys.Acta., 30, 44.

NEEDLEMAN, S.B. & BLAIR, T.T. (1969). Proc.natn.Acad.Sci., U.S.A., 63, 1227.

NIRENBERG, M. & LEDER, P. (1964). Science, N.Y., 145, 1399.

NISHIMURA, S., JONES, D.S. & KHORANA, H.G. (1965). J.molec.Biol., 13, 302.

OMURA, T. (1961). J.Biochem., Tokyo, 50, 394.

RICHARDS, O.C. (1967). Proc.natn.Acad.Sci., U.S.A. 57, 156.

ROTILIO, G., FINAZZI AGRO, A., AVIGLIANO, L., LAI, A., CONTI, F., FRANCONI, C. & MONDOVI, B. (1970). FEBS Letters, 12, 114.

SHERMAN, F., TABER, H. & CAMPBELL, W. (1965). J.molec.Biol., 13, 21.

SHOTTON, D.M. & HARTLEY, B.S. (1970). Nature, Lond., 225, 802.

SHOTTON, D.M. & WATSON, H.C. (1970). Nature, Lond. 225, 802.

SHOTTON, D.M. & WATSON, H.C. (1970). Nature, Lond. 225, 811.

SMILLIE, R.M. & STEELE SCOTT, N. (1969). In Progress in Molecular and Subcellular Biology vol.1, p.136. Ed.Hahn, F.E., Puck, T.T., Springer, G.F. & Wallenfels, K. Springer-Verlag, Berlin: Heidelberg: New York.

SMILLIE, L.B. & WHITAKER, D.R. (1967). J.Am.Chem.Soc., 89, 3350.

SMITH, E.L., MARKLAND, F.S., KASPER, C.B., DELANGE, R.J., LANDON, M., & EVANS, W.H. (1966). J.biol.Chem., 241, 5974.

SMITHIES, O. (1959). Advanc.Protein Chem., 14, 65.

SPACKMAN, D.H., STEIN, W.H. & MOORE, S. (1958). Analyt.Chem. 30, 1190.

SPIES, J.R. & CHAMBERS, D.C. (1948). Analyt.Chem., 20, 30.

SUGENO, K. & MATSUBARA, H. (1969). J.biol.Chem. 244, 2979.

SUTHERLAND, I.W. (1966). Ark.Mikrobiol., 54, 350.

SUTHERLAND, I.W. & WILKINSON, J.F. (1963). J.Gen.Microbiol., 30, 105.

TANAKA, M., HANIU, M., & YASUNOBU, K.T. (1970). Biochem.biophys. Res.Commun., 39, No.6, 1182.

VERHOEVEN, W. & TAKEDA, Y. (1956). In "Inorganic Nitrogen Metabolism", p.159. Ed. McElroy, W.D. & Glass, W.B., John Hopkins Press, Baltimore.

WALEY, S.G. & WATSON, J. (1953). Biochem.J., 55, 328.

WALSH, K.A. & BROWN, J.R. (1962). Biochim.biophys.Acta, 58, 596.

WELLS, J.R.E. (1965). Biochem.J., 97, 228.

WOODS, K.R. & WANG, K.T. (1967). Biochim.biophys.Acta, 133, 369.

YAKUSHIJI, F. (1966). In Blumberg, W.E. & PEISACH, J., (1966). Biochim.biophys.Acta. 126, 269.

YAMANAKA, T. (1966). In The Biochemistry of Copper, p.275. Ed. Peisach, J., Aisen, P. & Blumberg, W.E. Academic Pres: New York.

YAMANAKA, T., KIJIMOTO, S. & OKUNUKI, K. (1963). J.Biochem., Tokyo, 53, 256.

ZUCKERHANDL, E. & PAULING, L. (1962). In Horizons in Biochemistry, p.189. Ed. Kashna, M. & Pullman, B. Academic Press: New York.