ABSTRACT OF THESIS

Name of Candidate ALASTAIR AITKEN

Address Department of Molecular Biology, University of Edinburgh.

Degree Ph.D. *Date* 1975

Title of Thesis PRIMARY STRUCTURES OF BLUE-GREEN ALGAL PROTEINS

Complete and partial primary structures of plastocyanin and cytochrome f from prokaryotic blue-green algae were determined. The amino acid sequences of these proteins were compared to the proteins from eukaryotic algae and higher plants in an attempt to clarify the mode of origin of photosynthesis in eukaryotes. In order to elucidate the relationship between the two main types of prokaryotic organism, the protein sequences were also compared to proteins from bacteria. The results were interpreted in terms of the possible common origin for photosynthetic and respiratory electron transfer chains and the limitations on the possibility of protein sequences as a means of determining the main events in Precambrian evolution.

Use other side if necessary.

PGS/ABST/74/5000

PRIMARY STRUCTURES OF BLUE-GREEN ALGAL PROTEINS

by

Alastair Aitken

A thesis presented for the degree of Doctor of Philosophy
of the University of Edinburgh

Department of Molecular
Biology
University of Edinburgh

1975

This thesis is the result of research into the primary structure of proteins from blue-green algae carried out by myself (except for the assistance acknowledged below) in the Department of Molecular Biology between 1972 and 1975. Part of this work has been published in the paper at the end of this thesis.

ACKNOWLEDGEMENTS

I thank my supervisor, Dr. R.P. Ambler for his advice, instruction and constructive criticism, and Professor M.R. Pollock for his interest in the work. This study was made possible by the award of a scholarship from the Medical Research Council. During my second year I benefited from the award of a short-term fellowship from the European Molecular Biology Organisation spent at the Pasteur Institute Paris, where I was the guest of Professor and Mrs. R.Y. Stanier. I also benefited from the advice on growing blue-green algae given to me during two days spent in Dr. N.G. Carr's laboratory at Liverpool University.

This thesis would not have been possible without generous gifts of material from the following people:-

From Dr. N.G. Carr I received an axenic strain of Anabaena variabilis; from Professor R.Y. Stanier I received cells of Synechococcus 6312 and from Dr. G. Cohen-Bazire some cytochrome f from this strain; from Professor D.W. Krogmann I received cytochrome f from A. variabilis.

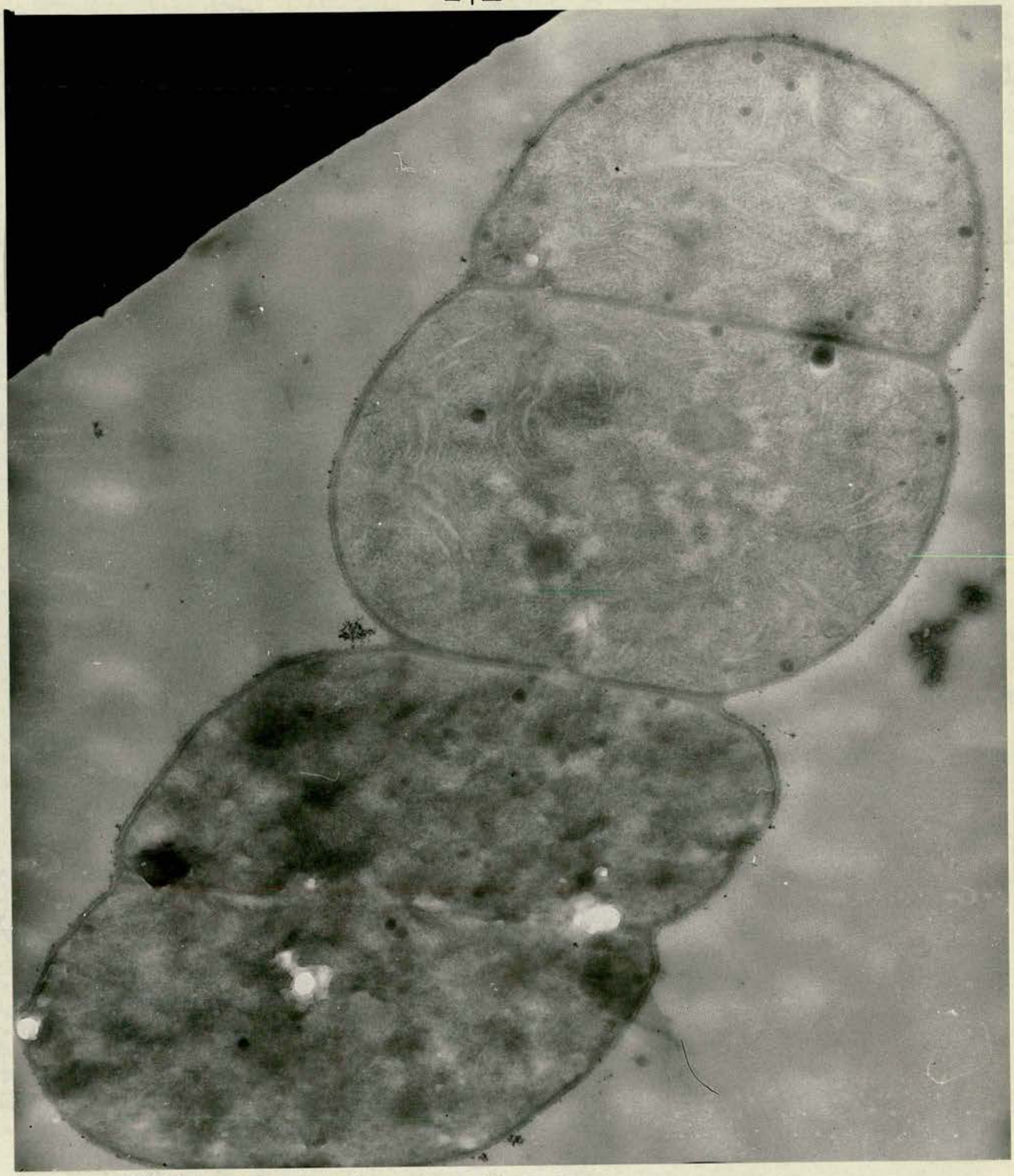
I would also like to thank Drs. D.R. Thatcher and G.W. Pettigrew for advice and instruction; Mrs. M. Daniel, Miss S. Murray and Dr. D. Pritchard for running sequenator

programmes; Mr. A. J. Savill for preparing enzymes; Mrs. E. Taylor, Mr. S. Hughes and Mr. T. Bruce for assistance in growing blue-green algae; Mr. R. Ferguson for taking electron micrographs and Mr. L. Morss for running S.D.S. gels.

SUMMARY

Complete and partial primary structures of plastocyanin and cytochrome f from prokaryotic blue-green algae were determined. The amino acid sequences of these proteins were compared to the proteins from eukaryotic algae and higher plants in an attempt to clarify the mode of origin of photosynthesis in eukaryotes. In order to elucidate the relationship between the two main types of prokaryotic organism, the protein sequences were also compared to proteins from bacteria. The results were interpreted in terms of the possible common origin for photosynthetic and respiratory electron transfer chains and the limitations on the possibility of protein sequences as a means of determining the main events in Precambrian evolution.

INDEX



Electron micrograph of a section of Anabaena variabilis.

(Taken by R.Ferguson)

CHAPTER 1

INTRODUCTION

(a) Plastocyanin

Plastocyanin is a copper containing protein of molecular weight around 11,000 daltons. The copper (1 atom per molecule protein) is in a type 1 ligand environment (Malkin & Malmström, 1970). The protein has an intense blue colour at about 600 nm when oxidised. The molar extinction coefficient of the copper chromophore is around 4,500. This is about 50 times as intense as a simple copper complex.

Other copper proteins of this type are azurins (bacterial respiratory proteins which do not occur in photosynthetic bacteria, Sutherland & Wilkinson, 1963) and stellacyanin (Malkin & Malmström, 1970). These are all electron transfer proteins containing one copper atom per molecule (which is detectable by electron paramagnetic resonance, EPR). There are other types of blue copper proteins which contain additional copper atoms in other ligand environments (Malkin & Malmström, 1970).

Plastocyanin occurs in higher plants (Katoh et al., 1961; Ramshaw et al., 1973) and in eukaryotic algae although it may be absent from the Euglenoids (Wood & Bendall, 1975; Nolan & Bishop, 1975). It has been isolated from filamentous blue-green algae (Lightbody & Krogmann, 1967; Biggins, 1967) and has been detected by EPR signals in the unicellular blue-green alga, Anacystis nidulans (Visser et al., 1974).

Attempts to isolate plastocyanin from Spirulina maxima, Anacystis nidulans and another unicellular blue-green

alga, Synechococcus 6312 have all been unsuccessful so far.

The intense blue colour is believed to be due to d-d transitions in a distorted square planar copper complex. The configuration may be halfway between the planar structure preferred by Cu(II) complexes and the tetrahedral configuration preferred by Cu(I). This would readily allow reduction and oxidation. The reduced symmetry leads to a much more intense chromophore (Malkin & Malmström, 1970). Spectroscopic studies on Cobalt (II) derivatives of blue copper proteins support the distorted tetrahedral structure but do not rule out the possibility of a five-coordinated complex (McMillin et al., 1974).

Evidence from laser Raman spectra of these proteins (Miskowsi et al., 1975) suggests that the coordination of the copper involves an approximately trigonal bipyramidal co-ordination with a sulphur and 2 nitrogen ligands in the equatorial plane and less strongly bound nitrogen or oxygen ligands at the axial positions (Fig. 1(a)).

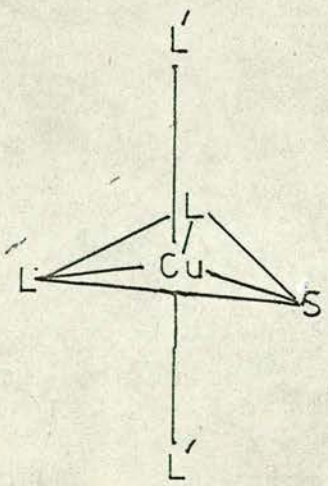


Fig 1(a)

Miskowsi et al. (1975) suggest that S → Cu(II) charge transfer is responsible for the chromophore at about 600 nm.

The single cysteine residue near the C-terminus in plastocyanins and azurins is in a highly conserved region (see results section) and Katoh et al. (1961) have suggested that the thiol group is one of the copper ligands.

(b) Cytochrome f

This is a c-type cytochrome containing a covalently bound haem prosthetic group, connected to the protein by thioether linkages. Cytochromes f are not oxidised by cytochrome aa_3. This type of cytochrome has been isolated from a wide range of eukaryotic algae (listed in Lemberg & Barrett, 1973, p.203) and from a number of blue-green algae (Ambler & Bartsch, 1975; Holton & Myers, 1967; Susor & Krogmann, 1966; Biggins, 1967). Molecular weights of the above cytochromes f are about 10,000 daltons.

Cytochrome f occurs in higher plants (Davenport & Hill, 1952) but here it is membrane bound and can be extracted with ethanol and detergent (Bendall et al., 1971). Studies on plant cytochrome f have suggested a much higher molecular weight (Bendall et al., 1971).

(c) Function of the proteins

Although plastocyanin and cytochrome f function in eukaryotic chloroplasts they may be both coded by nuclear genes and synthesised on cytoplasmic ribosomes (Haslett & Cammack, 1974; Boulter et al., 1972).

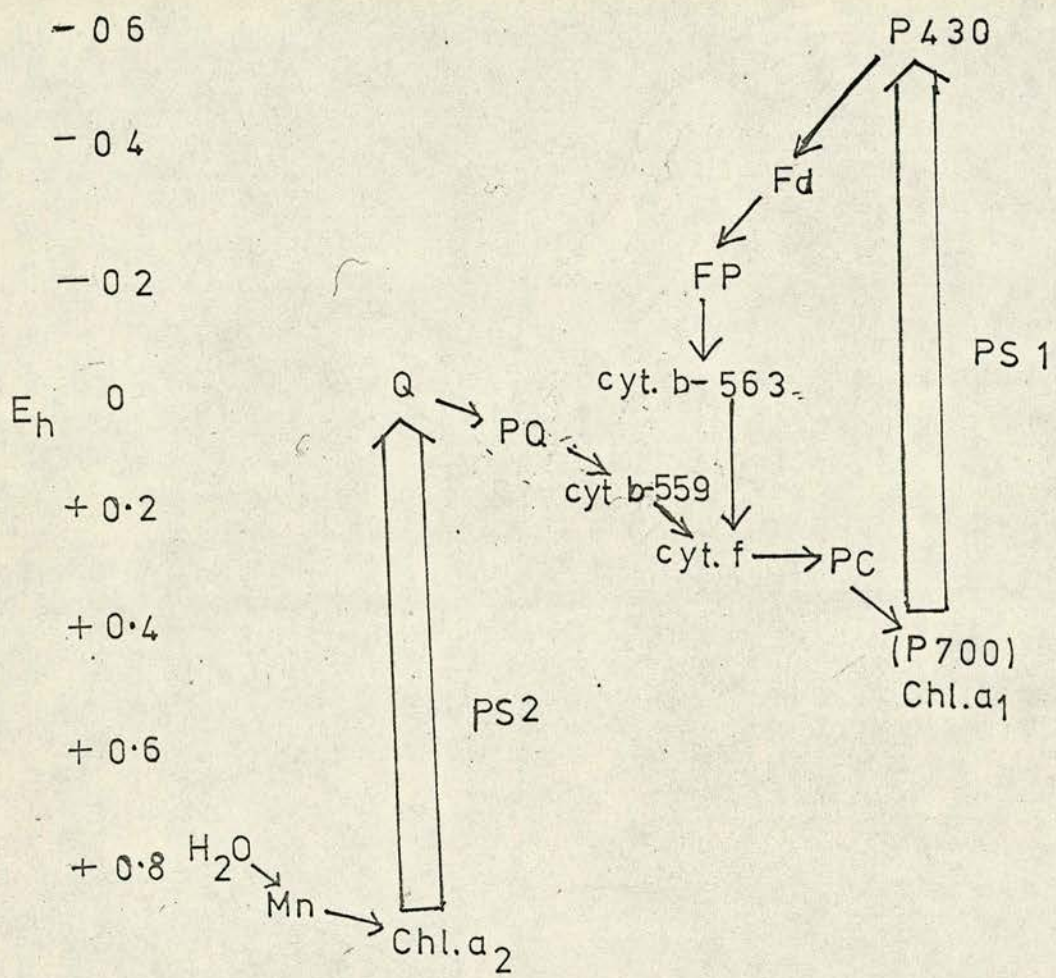


Fig. 1(b) Electron transfer scheme for oxygen-evolving photosynthesis.

Both proteins function in the photosynthetic electron transport chain (see Fig. 1(b)) but their precise order in electron transfer has been in dispute (Hind & Olson, 1968; Bendall & Hill, 1968; Kunert & Boger, 1975; Siedow *et al.*, 1973; Bearden & Malkin, 1974), but it appears that it is plastocyanin which donates electrons to P 700 (Plesnicar & Bendall, 1973; Gorman & Levine, 1965; Wood & Bendall, 1975).

The redox potentials of the 2 proteins are quite similar *in vitro* (about +.36V for cytochrome f and +.37V for plastocyanin) but they may differ *in situ* (Malkin *et al.*, 1973) and may be about +.34V for plastocyanin and +.385V for cytochrome f i.e. plastocyanin is on the reducing side of cytochrome f.

Knaff & Arnon (1969) have suggested a scheme involving 3 light reactions in photosynthesis where plastocyanin and cytochrome f are not in the same electron transfer chain.

There is also a suggestion (Bishop, 1971) that plastocyanin and cytochrome f have parallel and different donor sites to P 700.

(d) Use of amino-acid sequences

The amino acid sequence of a protein reflects the nucleotide base sequence of the structural gene coding for that particular protein. This is a direct reflection of the structure of the genome. Provided that the function of the protein has remained the same since divergence of the species under study and that no transfer of genetic material has occurred the topology of amino acid sequence comparisons

should represent the true ancestral relationships of organisms expressing that particular gene.

There are two main numerical methods of obtaining phylogenies - the ancestral sequence method of Dayhoff (1972) and the matrix method (Moore et al., 1973). The phylogenetic trees constructed by comparing the amino acid sequences of a number of proteins are more or less in agreement with the accepted phylogenies obtained by classical methods.

Some criticisms of molecular methods are that data based on a single structural gene cannot possibly reflect the evolution of the whole genome. There is no absolute way of quantifying the amount of back mutation or multiple mutation at the same site when comparing sequences. The assumption that rates of change of amino acid sequences are linear is only valid to a first approximation (Dickerson, 1971). Langley & Fitch (1974) have found that despite some variation in the rates of evolution of vertebrate haemoglobin, cytochrome c and fibrinopeptide, it is possible to obtain estimates of the relative times of divergence which correspond well with palaeontological evidence. An evaluation of these methods using amino acid sequence data from higher plant cytochrome c and plastocyanin (Peacock & Boulter, 1975) has revealed that careful choice of proteins leads to a calculated phylogeny which is a fairly accurate reflection of the true phylogeny of the protein.

Kimura (1968) and King & Jukes (1969) have suggested that many amino acid substitutions may be neutral in selective value. Others have disagreed with this (Clarke,

1970 and Richmond, 1970) and recent comparisons of amino acid sequences from bacterial cytochrome c-551 argue against neutral mutation playing a significant role in molecular evolution (Ambler, 1974).

The greatest advantage of molecular phylogenetic methods is their potential ability to be able to elucidate evolutionary relationships where classical methods do not give a convincing answer, i.e. if fossil evidence is controversial or if there is little morphological diversity in the organisms.

(e) Prokaryotic nature of blue-green algae

Blue-green algae (blue-green bacteria, cyanobacteria, Myxophycae, Cyanophycae) are prokaryotic organisms. Like other prokaryotes they lack the internal compartmentalisation of their cells including the membrane-bound nucleus characteristic of eukaryotic cells (Wolk, 1973; Hughes et al., 1970). Photosynthesis and respiration is carried out on membrane systems in prokaryotes (Echlin & Morris, 1965) and not in separate organelles (chloroplasts and mitochondria in eukaryotes). Blue-green algae lack Golgi bodies and endoplasmic reticulum (Allsopp, 1969). The gas vacuoles (Lang, 1968) which occur in cyanobacteria (and in a few widely separated genera of bacteria) are inclusions which are not bounded by unit membranes but by special single layers and are without counterpart in eukaryotes.

In eukaryotic cells, division of the nucleus occurs by mitosis (Allsopp, 1969). This may differ in some respects

in some simple eukaryotes such as euglenoids and dinoflagellates but the process is essentially the same. Nothing parallel occurs in prokaryotes. Histones are absent from blue-green bacteria (Makino & Tsuzaki, 1971) and although they may also be absent from dinoflagellates these latter organisms are eukaryotes by other criteria. The bacterial genophore is a much simpler structure than the eukaryotic chromosome and consists of a single linkage group (Allsopp, 1969). The DNA in the nuclear region of blue-green bacteria has been shown to be similar in structure (Leach & Herdman, 1973). In bacteria DNA is attached to mesosomes which have been shown to have a role in genophore segregation (Ryter, 1968). The mechanism may be similar in blue-green algae. Allen (1972) has shown the presence of mesosomes in the latter prokaryotes.

The ribosomes of blue-green algae are of the 70S type and have the same type of subunit size as those of bacteria (Taylor & Storck, 1964). Antibiotics such as chloramphenicol which specifically inhibit protein synthesis by 70S ribosomes inhibit the growth of cyanobacteria (Leach & Herdman, 1973). Cycloheximide which inhibits protein synthesis by 80S ribosomes, was ineffective.

A further similarity between the two types of prokaryote occurs in the viruses which infect blue-green algae. These cyanophage (Safferman, 1973; Padan & Shilo, 1973) contain DNA and are of the same size range and morphology as certain bacteriophages. Specific endonucleases have also been isolated from Anabaena variabilis (see experimental section) with similarities to the enzymes which restrict the

host range of infectivity of phages in bacteria.

The ability to utilise atmospheric nitrogen (Stewart, 1973); the use of poly-β-hydroxybutyrate as a storage substance (Carr, 1966) and the use of N-formylmethionine transfer RNA (also in eukaryotic organelles) as the initiator of protein synthesis (Wittman, 1970) are other features which distinguish prokaryotes from eukaryotes. Stanier (1970) considers that the capacity to harbour cellular endosymbionts, which has never been described in a symbiosis for which the host is a prokaryote may be a fundamental biological difference between contemporary prokaryotes and eukaryotes.

The cell wall structure in both types of prokaryote is basically the same and contains a peptidoglycan component (Drews, 1973). Cell walls are absent from many phytoflagellate and animal cells but the cell walls of those eukaryotes that have them are composed of polysaccharides [most frequently cellulose (Allsopp, 1969)]. Prokaryotes such as mycoplasmas and halobacteria do not have cell walls but the presence of peptidoglycan is considered to be an important bridging character between bacteria and blue-green algae (Murray, 1974).

(f) Are blue-green algae bacteria?

Echlin & Morris (1965) concluded their review of the relationship between bacteria and blue-green algae by stating that "it seems probable that bacteria and blue-green algae arose from a common ancestor, since it is unlikely that so many features common to both groups arose independently more than once". There is no dispute about the prokaryotic nature

of blue-green algae but considerable dissent remains as to whether or not blue-green algae should be called bacteria (Pringsheim, 1949; Soriano & Lewin, 1965; Stanier & Van Niel, 1962).

Blue-green algae are not flagellated but movement is by gliding (Castenholz, 1973). This is a characteristic which is shared with the Flexibacteria (Soriano & Lewin, 1965). There is evidence that the gliding and flagellar motilities in prokaryotes have a similar mechanism (Castenholz, 1973).

Blue-green algae such as Oscillatoria and Spirulina have close morphological similarities to gliding bacteria especially Beggiatoa, Thiothrix, Vitreoscilla and Leucothrix. The latter are classically considered "colourless blue-green algae" or "apochlorotic cyanophytes" (Pringsheim, 1949; Soriano & Lewin, 1965; Brock, 1973). Some gliding bacteria (e.g. Cytophaga and the fruiting bacteria) have less resemblance to blue-green algae, although it has been considered that the former gliding bacteria evolved from cyanobacteria by loss of photosynthetic pigments.

The cyanobacterial photosynthetic apparatus resembles that of the chloroplast much more than that of other photosynthetic bacteria. The blue-green algal and eukaryotic photosystems use water as the electron donor and have been shown to be very similar in structure and components (Krogmann, 1973; Bearden & Malkin, 1974). See Fig. 1(b). Proteins common to these prokaryotes' and eukaryotes' photosystems have been shown to have great similarity in amino acid sequence. These include cytochrome f (Ambler &

Bartsch, 1975), ferredoxin (Tanaka et al., 1975; Wada et al., 1975) and phycobiliproteins (in certain eukaryotic algae - Brown et al., 1975; Troxler et al., 1975; Williams et al., 1975). Considerable immunological crossreactivity also exists between the corresponding subunits of phycobiliproteins in prokaryotes and red algae (Berns, 1967; Glazer & Cohen-Bazire, 1971; Stanier, 1974).

The two important differences between photosynthesis in bacteria and cyanobacteria are that in the former anaerobic conditions are essential and the electron donor used is not water but a reduced sulphur compound, molecular hydrogen or an organic compound. Purple non-sulphur bacteria use organic compounds as their carbon source while purple sulphur and green sulphur bacteria use CO_2 as the sole carbon source (Pfennig, 1967).

It now appears that there are more similarities between the two types of photosynthesis than were previously supposed. There is much evidence that bacterial photosynthesis involves two photochemical reaction centres (Olson, 1970; Hind & Olson, 1968; Sybesma, 1969).

Green sulphur bacteria contain cytochromes c-555 and c-552 while purple sulphur and non-sulphur bacteria contain cytochromes c_2 and c' (Lemberg & Barrett, 1973). The high potential cytochromes c-555 and c_2 have been suggested as components of a cyclic electron transport system (Duysens, 1973) and the low potential cytochromes c' and c-552 may be components of substrate linked non-cyclic electron transport (Sybesma, 1969). Other cytochromes have been postulated as

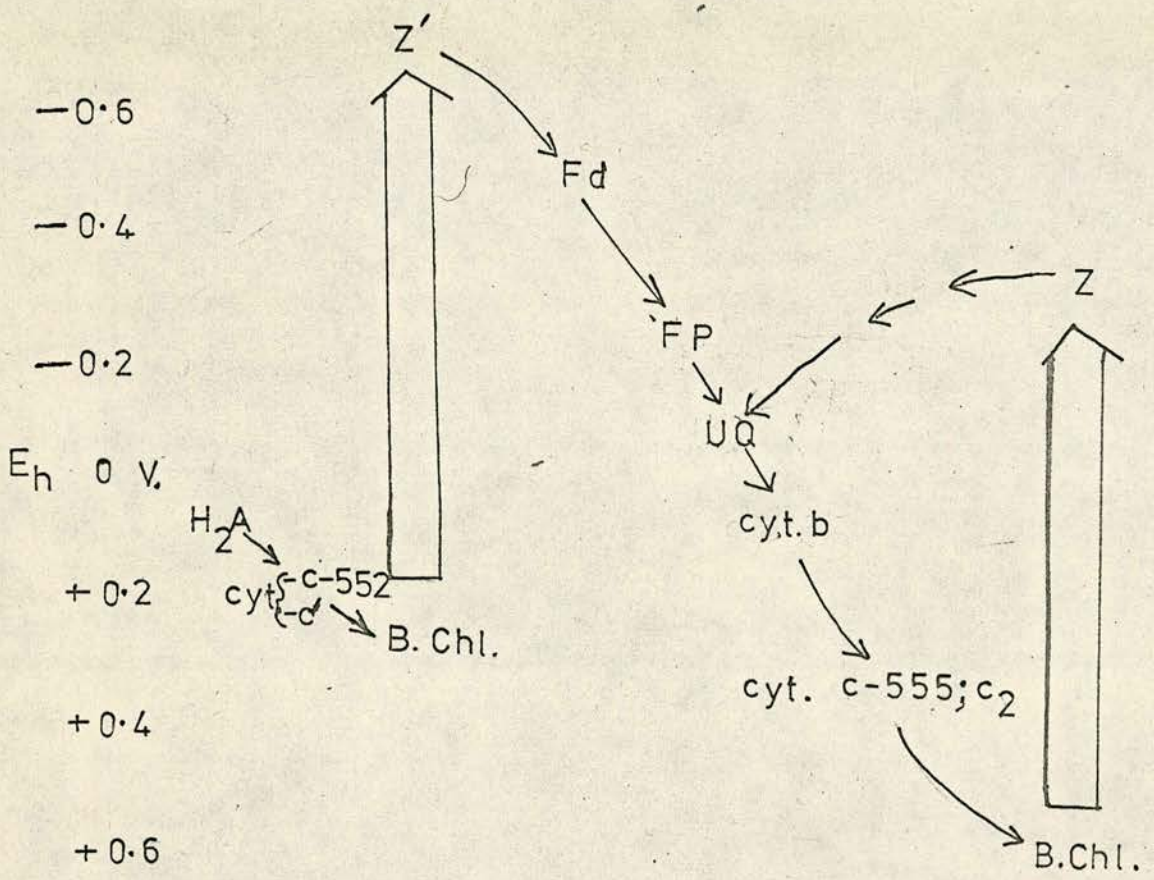


Fig. 1(c) Electron transfer scheme for bacterial photosynthesis.

occurring in the electron transfer schemes, e.g. cytochrome b (Prince & Dutton, 1975).

Parson (1974) dissents from the view of two photosystems. There is, for example, no apparent enhancement effect in bacterial photosynthesis.

If the scheme outlined in Fig. 1(c) is correct then the high potential cytochromes have a similar function to cytochrome f in the oxygenic type of photosynthesis.

In view of the probable common origin of both types of photosynthetic prokaryote from a heterotrophic ancestor (Olson, 1970) possible amino acid sequence homology between these cytochromes is discussed in the results section.

Other similarities between the two types of photosynthesis in the prokaryotes include resemblances between the photosynthetic lamellae of blue-green algae and the chromatophores of photosynthetic bacteria (particularly Athiorhodaceae, Echlin & Morris, 1965). In relation to respiration, increased uptake of oxygen by blue-green algae has been noticed in the dark (Carr, 1973). This has been observed in purple non-sulphur bacteria but not in eukaryotic algae. McFadden & Tabita (1974) have shown particular similarities in subunit size and composition of ribulose-1,5-diphosphate carboxylases from Thiorhodaceae, hydrogen bacteria and oxygenic photosynthesisers.

The recently discovered genus Chloroflexus is considered to be closer to blue-green algae than other

photosynthetic bacteria (Pierson & Castenholz, 1974). These are gliding, filamentous photosynthetic bacteria with base ratios similar to Oscillatoria and resemble blue-green algae in lipid and β-carotene content. They are not however, oxygenic photosynthesisers.

Oscillatoria limnetica, isolated from an H_2S-rich lake, has been shown to exhibit both "bacteria-like" and oxygenic photosynthesis (Cohen et al., 1975). O. limnetica can use Na_2S as electron donor for CO_2 photoassimilation in the presence of DCMU or 700 nm light. The energy for the photoassimilation is supplied by photosystem 1. The authors conclude that this blue-green alga could represent a secondary specialisation back to bacterial type photosynthesis or it could be a missing link between the two types of photosynthesis.

(g) The fossil record

Fossil remains from Swaziland rocks have shown that some form of life may have existed 3.2×10^9 years ago (Schopf, 1970; Swain, 1969). These fossils are of rod-shaped organisms in the Fig-tree Chert and spheroids and filaments in the Onverwacht rocks. These remains resemble extant bacteria but this is not surprising since the morphology of such simple cells as the first bacteria would have been limited. A variety of evidence (Engell et al., 1968) suggests that these organisms were physiologically rather advanced. This implies a substantial period of prior evolutionary development. It may not be possible therefore

to obtain direct evidence of the earliest evolution of life forms unless very ancient sedimentary rocks are discovered. The oldest known rocks are about 3.5×10^9 years old.

There is chemical evidence to suggest the existence of chlorophyll about 3×10^9 years ago (Calvin, 1969) and Brooks et al. (1973) have suggested the presence of photosynthetic organisms in Onverwacht rocks 3.4×10^9 years old. Ancestral forms of blue-green algae may have diverged from the bacterial photosynthesising line of descent more than 3×10^9 years ago (Olson, 1970). Recently, filamentous blue green algae have been identified from stromatolites in Transvaal Dolomite which is 2.3×10^9 years old (MacGregor et al., 1974) or 2.2×10^9 years old (Nagy, 1974). In the latter report heterocystous and coccoid forms were also recognised.

In the Gunflint Iron formations ($1.6 - 2 \times 10^9$ years old) many representatives of extant cyanophycean families have been recognised (Chroococcaceae, Oscillatoriaceae and Nostocaceae) as well as chemosynthetic bacteria but no nucleated organisms were identified.

The Bitter Springs formations (0.9×10^9 years old) contain many species of cyanophytes which have been identified at the specific or generic level with extant organisms, and Schopf (1970) states that "it seems evident therefore that blue-green algae were highly diversified in the Late Precambrian, and that certain cyanophytes have exhibited little or no evolution, at least in terms of organismal morphology, since Bitter Springs time". Fogg et al. (1973) have suggested that the unicellular blue-green algae evolved first,

and filamentous forms very shortly afterwards.

The oldest rocks containing fossils which are reasonably interpretable as eukaryotic microorganisms are from the Beck Springs Dolomite (about 1.3×10^9 years old) (Raff & Mahler, 1972). It is possible that nucleated organisms first appeared 1.7×10^9 years ago (Schopf, 1970), although the first eukaryotic type of organism would possibly be difficult to recognise as a fossil. However, it appears that eukaryotes evolved sometime between the Gunflint and Bitter Springs formations and Kimura & Ohta (1973) have suggested a time of 1.8×10^9 years ago for prokaryote-eukaryote divergence on the basis of 5S ribosomal RNA and cytochrome c sequences. This is compatible with the suggestion of Cloud (1968) that on the basis of the appearance of red beds of oxidised iron about 1.8×10^9 years ago these deposits marked the appearance of accumulations of oxygen in the atmosphere. Berkner & Marshall (1965), however, have predicted that the atmospheric oxygen concentration was too low to support aerobic respiration until about 600 million years ago.

It is probable that blue-green algae were the first oxygen evolving photosynthetic organisms and were initially responsible for producing atmospheric oxygen. Some blue-green algae can grow anaerobically (Stewart & Pearson, 1970) and atmospheric oxygen is a prerequisite for the evolution of eukaryotes, which normally depend on mitochondrial respiration for their energy requirements.

However well established the time of the origin of

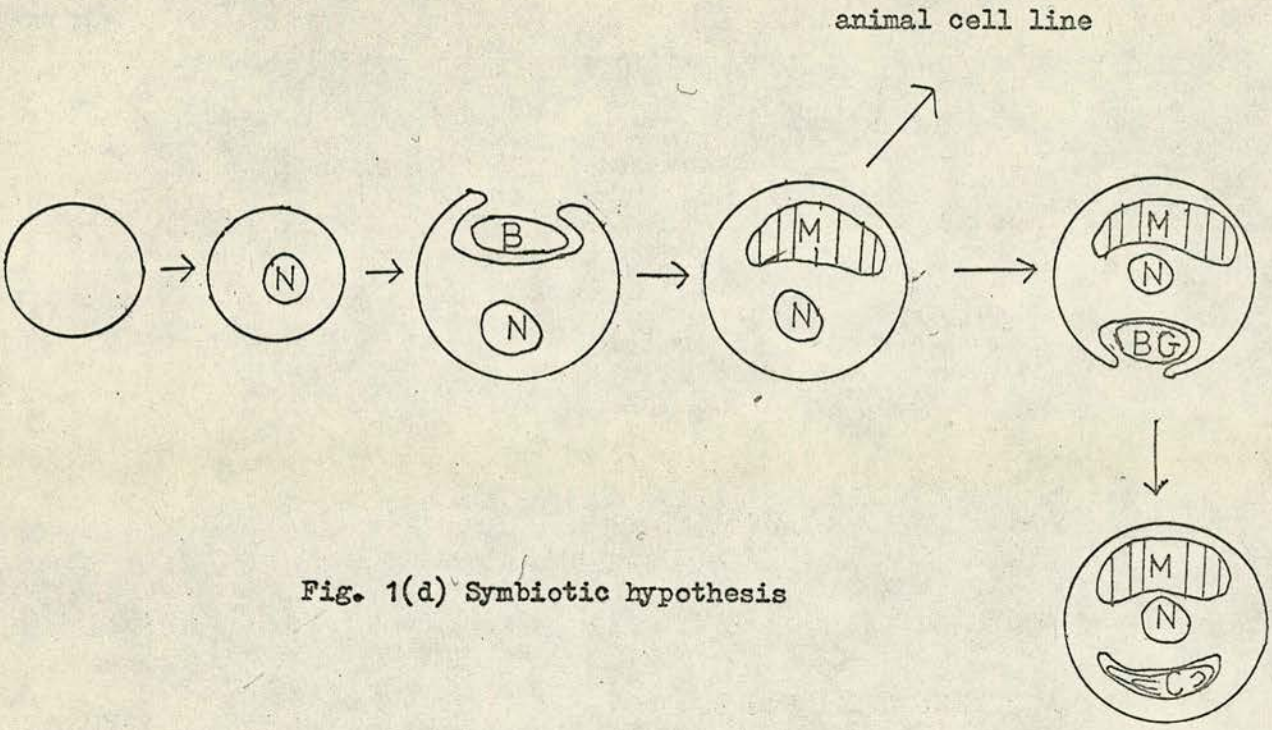


Fig. 1(d) Symbiotic hypothesis

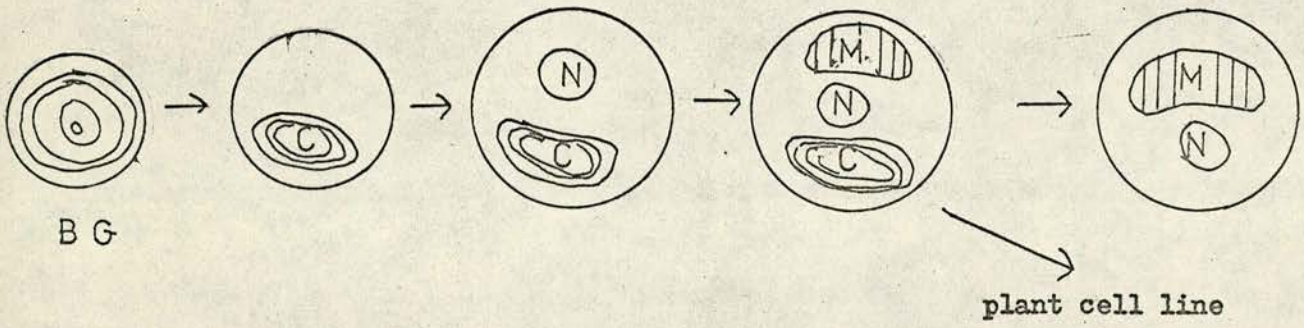


Fig. 1(e) Classical view

N nucleus

B aerobic bacterium

BG blue-green alga

M mitochondrion

C chloroplast

eukaryotic cells becomes, the actual mode of origin is a different matter and has been the subject of a great deal of speculation in recent years.

(h) Endosymbiotic origin of eukaryotic cells

The symbiotic hypothesis is summarised in Fig. 1(d) and compared to the classical view [Fig. 1(e)]. In brief, it is the hypothesis that the chloroplast and mitochondria of eukaryotic cells arose from an endosymbiotic association of a blue-green alga and an aerobic bacterium respectively (Marguilis, 1970; under the name of Sagan, 1967; Taylor, 1974), with an anaerobic amoeboid host cell.

Closer association of these endosymbionts developed into the semi-autonomous organelles as they are known today. According to Marguilis (1970) the flagellar and mitotic apparatus including the centriole and chromosomal centromeres of eukaryotic cells originated by mutation of a symbiotic flagellated prokaryote. There is little firm evidence for this last proposal.

The evidence for the endosymbiotic origin of the chloroplast is extensive and will be dealt with briefly. Symbiotic associations involving cyanobacteria are very common (Fogg et al., 1973; Echlin, 1970). There are some which could represent stages in the origin of the chloroplast. In the fungus Geosiphon pyriforme the symbiotic blue-green alga is a Nostoc species and can be cultured separately from the host (Schnepf, 1964). Neither symbiont in Cyanophora paradoxa (Hall & Claus, 1963) nor that in Glaucocystis nostochinearum (Hall & Claus, 1967) can be cultured apart

from the host although the symbiont in C. paradoxa still has a thin cell wall (Hall & Claus, 1963).

The structure and components of the oxygenic prokaryotic and eukaryotic photosynthetic apparatus is basically the same (Krogmann, 1973). The amino acid sequences of proteins in blue-green algae and chloroplasts are very similar (see above). Kenyon & Stanier (1970) have shown that the lipids and fatty acids of blue-green algae are more similar than those in other photosynthetic prokaryotes to those in chloroplasts.

Chloroplasts contain their own protein synthesising machinery (Boulter, et al., 1972; Ellis & Hartley, 1974). This contains the 70S type of ribosome with the appropriate antibiotic specificities and N-formylated methionine transfer RNA as the initiator of protein synthesis (Lea & Norris, 1972). The DNA has been shown to be circular in many cases (Ellis & Hartley, 1974). This evidence suggests that the chloroplast may once have been a free-living prokaryote identifiable as a blue-green bacterium.

Evidence for the endosymbiotic origin of the mitochondrion follows similar lines. Mitochondria also have a degree of autonomy in protein synthesis although the size of the ribosomes is more varied (Raff & Mahler, 1972).

The mitochondrial electron transport chain is very uniform over the wide range of eukaryotes (Stanier, 1970) but respiratory electron transport in aerobic bacteria is more varied (Smith, 1961). This is not inconsistent with the symbiotic theory since bacterial electron transport may have diverged considerably since symbiosis occurred or the endosymbiotic bacterium contained only one of the types of

respiratory chain which were then extant. John & Whatley (1975) have suggested that a bacterium like Paracoccus denitrificans is a suitable candidate for the endosymbiont partly on the basis of the close similarity in components of its respiratory system with those in mitochondria.

Strong support for the endosymbiotic origin of mitochondria has been obtained when Steinman & Hill (1973) and Bridgen et al. (1975) showed a high degree of N-terminal amino acid sequence homology between superoxide dismutases from bacteria and mitochondria. In contrast these sequences showed no significant similarity to the N-terminal region of superoxide dismutase from bovine erythrocytes which is a smaller enzyme and may have had a different evolutionary origin (Steinmann & Hill, 1973).

Some authors have suggested that the original host was not an anaerobic amoeboid fermentative protoeukaryote as proposed by Marguilis (1970) but had some aerobic capabilities (Cohen, 1973; Raff & Mahler, 1972; de Duve, 1973). This is suggested on the grounds that among other things enzymes such as superoxide dismutases were present in the eukaryotic cytoplasm to protect cellular components from oxidation and the synthesis of components such as steroids and unsaturated fatty acids in the eukaryotic cytoplasm by aerobic pathways (Raff & Mahler, 1972). De Duve (1973) considers that the host cell into which the aerobic bacterium was incorporated had a peroxisomal type of respiration. The postulated aerobic nature of the host cell has been used as an argument against symbiosis on the grounds that no selective advantage could result from the acquisition of an aerobic

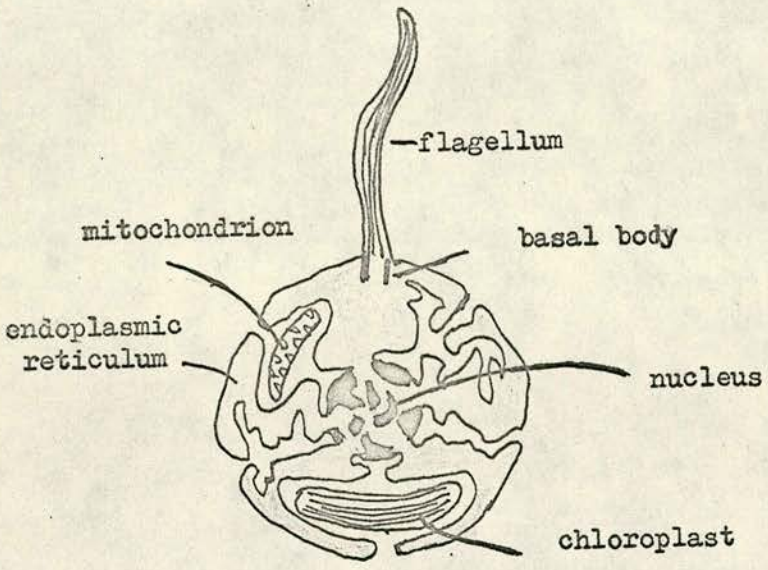


Fig. 1(f) Topological model of photosynthetic eukaryotic flagellate.

symbiont (Raff & Mahler, 1972). This is a fatuous argument because it would be necessary for the protoeukaryote to survive in a micro-aerobic environment in order to associate with an aerobic bacterium with which to form a symbiosis. The aerobic bacterium would also respire by the much more efficient phosphorylating electron transport chain and thus confer a selective advantage on the host cell.

Some mitochondrial and chloroplast proteins are synthesised on cytoplasmic ribosomes (Ellis & Hartley, 1974). This is explained on the basis of the endosymbiotic hypothesis by transfer of some genetic information to the nucleus (Boulter et al., 1972) which may be necessary for a regulatory purpose (perhaps in order that the host or symbiont does not outgrow the other partner). This is also used as another argument against the symbiotic hypothesis in that chloroplasts and mitochondria are not entirely autonomous and cannot be cultured separately from the rest of the eukaryotic cell. In fact many autonomous "organelles" are known today which perform photosynthesis and are also found free-living. In this case the association is called a symbiosis (see above). Bogorad (1975) has discussed the possibility that in different eukaryotes, different genes may have been transferred from mitochondria and chloroplasts to the nucleus. Taylor (1974) considers that mitochondria and chloroplasts are topologically outside the eukaryotic cell and the endosymbionts did not completely enter the cell but were "embraced" by the host cell. The centriole/basal body and other microtubular apparatus are however inside the eukaryotic cell (Fig. 1(f)).

Various authors differ on the order of the symbiotic events. Marguilis (1970) and Taylor (1974) are of the opinion that the protomitochondrion entered first. Mitochondria are very uniform and occur in all eukaryotes except in a few cases of secondary loss. Chloroplasts do not occur in all eukaryotes and are much more diverse. They also appear to be less completely integrated with the host cell. Stanier (1970) takes the view that symbiosis of the chloroplast occurred first because aerobic respiration appeared long after oxygenic photosynthesis was established. This was due to a long slow build-up of atmospheric oxygen before there was sufficient quantity to allow aerobic respiration.

A polyphyletic origin of chloroplasts has been proposed by Sagan (1967) and Raven (1970) on the basis of the similarity in type of chlorophyll and accessory pigments between blue-green algae, Rhodophyta, Cryptophyta and Cyanidium. The photosynthetic pigment content is quite different in other groups of eukaryotic algae. Raven (1970) suggests that many symbiotic events occurred and algae such as Euglenophyta and Chlorophyta (and through the latter eventually vascular plants) arose from endosymbiosis of a hypothetical "green prokaryote". The other groups of algae originated from a "yellow prokaryote". These hypothetical "green" and "yellow" prokaryotes are now extinct - or yet to be discovered.

The alternative hypothesis proposes that eukaryotic cells arose from a single prokaryote species by intracellular

differentiation (Allsopp, 1969; Raff & Mahler, 1972; Klein & Cronquist, 1967; Cavalier-Smith, 1975; Rejnders, 1975). Allsopp (1969) and Raff & Mahler (1972) propose that the mitochondrion evolved when a large increase in respiratory membrane surface became necessary when the size of the protoeukaryote increased. The membrane became invaginated and membrane bound vesicles formed which subsequently blebbed off from the cell membrane. This became a stable organelle which later acquired an outer membrane. According to Allsopp (1969) the organelle took with it some membrane bound genes coding for protein synthesising machinery and genes for some proteins. Alternatively (Raff & Mahler, 1972) the organelles later acquired a plasmid (of the type which occur widely in extant bacteria) which contained the appropriate genes. This degree of autonomy is necessary for synthesis of proteins which cannot pass the membranes of mitochondria (or chloroplasts - Boulter et al., 1972). Subsequent evolution of mitochondria and chloroplasts (which are supposed to have arisen similarly) has been much less than the rest of the cell thus the organelles have retained superficial resemblances to aerobic bacteria and blue-green algae respectively.

Rejnders (1975) has recently proposed a somewhat similar hypothesis for the origin of mitochondria where nuclear and mitochondrial DNA developed from compartmentalisation of duplicate DNA in a prokaryote. This followed the trapping of the DNAs in separate invaginations of the cell membrane. Redundant DNA was eliminated in one part (the protomitochondrion) which specialised its functions and

developed into the mitochondrion.

A hypothesis advanced for the origin of eukaryotic cells by Cavalier-Smith (1975) also includes an explanation for the origin of the nucleus. The eukaryotic cell is supposed to have evolved non-symbiotically when a blue-green alga without a cell wall developed phagocytosis. This claims to be far simpler than symbiotic hypotheses but does in fact require quite a large number of mostly hypothetical stages. The nucleus and organelles developed when it became necessary to compartmentalise the cell which had become very large and diverse in function. This partitioning of the cell into organelles which specialised in particular functions avoided dilution of components and inefficiency of reactions. Hall (1973) has proposed similar reasons for compartmentalisation of the protoeukaryote.

(i) Summary

From the evidence outlined above, the conclusion that the oxygen evolving photosynthesis had a single evolutionary origin seems inescapable. The general features of eukaryotic cell structure, biochemistry and organisation are too close to suggest that the eukaryotic cell had a polyphyletic origin. Therefore, if the symbiotic hypothesis is discarded, it is difficult to avoid the alternative conclusion that a blue-green alga (or ancestral form) gave rise to all eukaryotic cells - some of which subsequently lost photosynthetic capability.

Taylor (1974) is of the opinion that it is unreasonable to discuss the origin of mitochondria separately from that of

chloroplasts. It is less likely that only one of these organelles had a symbiotic origin (if the endosymbiotic hypothesis is accepted). Other proposals for the origin of eukaryotic cells have been put forward such as that of Goksoyr (1967) who suggests that coencytic fusion of similar prokaryotic cells occurred in order to obtain sufficient DNA for eukaryotic chromosomes and Nass (1969) has suggested fusion of a colonial prokaryote in which members had already differentiated.

The difference between prokaryotic and eukaryotic cells is undoubtedly the greatest evolutionary discontinuity, there are no known intermediate cellular forms (Stanier, 1970). This is itself an argument in favour of the endosymbiotic hypothesis. Suggestions such as that of Bisset (1973) that extant bacteria have evolved from nucleated eukaryotic ancestors do not seem reasonable. There is no evidence for this and this is at variance with the fossil record which suggests a prokaryotic type of cell in the oldest Precambrian fossils. It also seems illogical that the morphology of the first life forms could be much different from a simple prokaryote-like structure.

There is great difficulty in devising experiments which could prove which if any of the events outlined above actually took place. The origin of the eukaryotic cell is a historical series of events which happened about 1½ to 2 billion years ago and furnishing proof of such events is unlikely. It may be possible by means of RNA and amino acid sequences (such as Kimura & Ohta, 1973, placing the divergence

of prokaryotes and eukaryotes 1.8 billion years ago) to suggest which were the most likely occurrences. Phylogenetic trees constructed from sequences of gene products from different organelles may show different topologies if the endosymbiotic hypothesis is correct and may be able to suggest the sequence of events leading to the eukaryotic cell.

CHAPTER 2

Growth of Organisms and Protein Purifications

The reagents used were normally BDH or Fisons general purpose chemicals unless otherwise stated.

(a) Growth of Organisms

A. variabilis was generally grown in medium C of Kratz & Myers (1955) in 20 ℓ bottles bubbled with 5% CO_2 in air with the addition of sodium bicarbonate (0.8 g/ℓ) to maintain the pH above 7.

P. luridum was grown in the medium of Allen & Arnon (1955), medium C of Kratz & Myers and latterly in the medium of Cannon et al. (1971).

Cells of Synechococcus 6312, grown in medium BG-11, were obtained from R.Y. Stanier. One batch of Anacystis nidulans was grown in medium BG-11 (Stanier et al., 1971).

The culture medium was illuminated with an array of 6 x 20 W "grolux" lamps and 8 warm-white fluorescent tubes arranged evenly round the 20 ℓ bottles - generally 6 in number. The organisms were grown at a temperature of around 30°C except Anacystis nidulans which was grown at 37°C.

After growing for about 10 to 15 days, the cells were harvested in an Alfa-Laval centrifuge (model LAB102B) and stored frozen until required.

In an attempt to increase the yield of cytochrome f batches of A. variabilis were grown in the presence of 12 mg/ℓ diphenylamine (Ogawa & Vernon, 1971). This organic chemical proved difficult to dissolve in the media used - initially medium C of Kratz & Myers (1955) and later "modified Detmer's medium" (Watanabe, 1960) was used, which was the original medium for growth with diphenylamine.

A. variabilis did not grow well under these conditions and much lower yields of the organism were obtained, although the comparative yield of cytochrome f from unit weight of cells was increased.

Allen's medium (Allen, 1968) was used in surface culture for the maintenance of the strains on slopes. Double strength (1.5%) Difco Bacto agar was used which was autoclaved separately.

The trace element mix was added (1 ml/ℓ) to all the media although the constituents of the mixes listed in the original papers varied somewhat between different media.

(b) Strains of blue-green algae

Anabaena variabilis

Obtained from N.G. Carr, Department of Biochemistry, University of Liverpool. This is A. variabilis (Kützing) which is the strain used by Kratz & Myers (1955) in their study of nutrition and growth. The strain was isolated by R.C. Hecker as Cylindrospermum and the identification was revised by E.G. Pringsheim.

Phormidium luridum

Obtained from Culture collection of Algae and Protozoa, Cambridge (George, 1971) where it is listed as Plectonema boryanum 1462/2. This is the same strain used by Biggins (1967) under the name Phormidium luridum var. olivaceum. It is listed in the Indiana University culture collection (Starr, 1964) under this name. It is also called Schizothrix calcicola (George, 1971).

Synechococcus 6312

The cells were obtained from cultures grown by R.Y. Stanier from his isolate No. 6312 (Stanier et al., 1971).

Anacystis nidulans

Obtained from Culture collection of algae and protozoa (George, 1971) where it is listed as Synechococcus leopoliensis 1405/1 Kratz/Allen. It is most widely known as A. nidulans but has also been called many other names (e.g. Lauterbornia nidulans; Aphanothece nidulans; Phormidium mucicola).

(c) Blue-green algal media

Kratz & Myers medium C (Kratz & Myers, 1955) - Modified.

$Mg.SO_4.7H_2O$	0.25 g/l
K_2HPO_4.	1.00 g/l
$Ca(NO_3)_2.4H_2O$	0.025 g/l
KNO_3	1.00 g/l
sodium citrate.$2H_2O$	0.165 g/l
$Fe_2(SO_4)_3.6H_2O$	0.004 g/l
EDTA (disodium salt)	0.001 g/l

Allen & Arnon (1955)

$MgSO_4.7H_2O$	0.25 g/l
$CaCl_2$.	0.073 g/l
NaCl	0.002 g/l
K_2HPO_4	0.0035 g/l
KNO_3	2 g/l

Cannon et al. (1971)

$NaNO_3$	0.179 g/l
$Ca(NO_3)_2.4H_2O$	0.06 g/l
K_2HPO_4	.001 g/l
$MgSO_4.7H_2O$	0.025 g/l
Na_2CO_3	0.02 g/l
$Na_2SiO_3.5H_2O$	0.044 g/l
ferric citrate	0.35 g/l
citric acid	0.35 g/l

Modified Detmer's medium (Watanabe, 1960)

$Mg.SO_4.7H_2O$	0.25 g/l
K_2HPO_4	0.25 g/l
NaCl	0.1 g/l
KNO_3	1.0 g/l
$CaCl_2$	0.01 g/l
$FeSO_4.7H_2O$	0.02 g/l

Allen (1968)

$NaNO_3$	1.5 g/l
K_2HPO_4	0.039 g/l
$Mg.SO_4.7H_2O$	0.075 g/l
Na_2CO_3	0.02 g/l
$CaCl_2$	0.027 g/l
$Na_2SiO_3.9H_2O$	0.058 g/l
EDTA	0.001 g/l
citric acid	0.006 g/l
ferrous citrate	0.006 g/l

Trace element mix (Stanier et al., 1971)

H_3BO_3	2.86 g/l
$Mn.Cl_2.4H_2O$	1.81 g/l

$Zn.SO_4.7H_2O$	0.222 g/ℓ
$Na_2MoO_4.2H_2O$	0.39 g/ℓ
$CuSO_4.5H_2O$	0.079 g/ℓ
$Co(NO_3)_2.6H_2O$	0.0494 g/ℓ

(d) Purification of Proteins

The cells (generally about 200 g wet weight) were suspended in 1 ℓ potassium phosphate buffer (0.05 M., pH 7) and broken in a Manton-Gaulin homogenizer. Care was taken to keep the temperature below about 30°. The buffered suspension was cooled in an ice-bath before breakage and after each pass through the homogenizer. The beaker in which the suspension was stored was also kept ice-cold. After about 3 passes the solution became warm and was examined microscopically to determine the extent of cell breakage. The re-cooled suspension was broken again if judged necessary. The suspension was usually sonicated at this stage for 20-30 min. to ensure complete release of soluble proteins. Extensive sonication was not in itself sufficient to break the cell walls. Occasionally the whole breaking operation was carried out in the presence of 0.1% (v/v) triton X-100 (Borchert & Wessels, 1970) but this did not appear to make any difference in the yields of plastocyanin or cytochrome f from any of the strains investigated.

The cell debris was removed by centrifugation at 10,000 g for 10 minutes. The debris was not discarded but was stored and added to subsequent batches of cells. This was considered a better alternative to resuspending the debris in buffer and re-extracting for protein. Large

numbers of apparently intact cells were still seen under the microscope even in batches which had been passed through the homogenizer many times.

After removal of cell debris by centrifugation, the supernatant was subjected to ammonium sulphate fractionation. The 40-90% saturated ammonium sulphate fraction was dialysed or desalted by Sephadex G-25 (30 cm x 6 cm) gel filtration to 1 mM phosphate buffer, pH 7. All of the chlorophyll and most of the phycobiliproteins were removed in the 0-40% ammonium sulphate fraction. It occasionally proved difficult to reduce the ionic strength of the 40%-90% fraction to that of 1 mM potassium phosphate. In this case after gel filtration or dialysis, the volume of solution was reduced in a Millipore ultrafiltration cell (with a capacity of 1,600 ml) to about 30 ml. This small volume was diluted with distilled water to the correct conductivity after pH adjustment. If the conductivity was still very high the solution was diluted while remaining in the ultrafiltration apparatus and the volume was reduced again before final adjustment of pH and conductivity.

The solution was passed down DEAE-cellulose equilibrated with potassium phosphate, pH 7 where all the remaining phycobiliproteins as well as ferredoxin and other proteins were absorbed. Of the proteins in the present study, only Synechococcus 6312 cytochrome f was absorbed. The solutions containing plastocyanin and cytochrome f from the other strains were absorbed onto CM-cellulose under the same conditions of pH and conductivity after addition of sufficient potassium ferricyanide solution (0.01 M) to

oxidise the plastocyanin. The proteins were eluted with buffers of increased ionic strength. The individual proteins were further purified by the methods outlined in Table 2(a).

Chromatography on CM-23 at pH 3.9, ammonium acetate, 0.05 M (Ambler, 1963) was achieved by reducing the pH of the solution containing the cytochrome to 4.1, absorbing the protein onto the CM-cellulose and eluting with ammonium acetate buffers of increasing pH. The exact conditions of elution from DEAE- or CM-cellulose varied slightly between batches of cellulose or depending on the exact pH of the phosphate buffer. Approximate yields are given in Table 2(b). A gift of A. variabilis cytochrome f from Professor D.W. Krogmann was gratefully received.

After removal of the phycobiliproteins from the DEAE-cellulose column by washing with 0.05 M potassium phosphate, pH 7, containing 0.2 M NaCl, ferredoxin was eluted with 1 molar buffer (0.05 M phosphate and 0.95 M NaCl, pH 7). The ferredoxin fractions were sent to Professor D.O. Hall, Kings College, London, for purification and further study (Tel-or et al., 1975). The 0-40% ammonium sulphate fractions from Anabaena variabilis which contained the bulk of the phycobiliproteins were sent to Professor R.F. Troxler, Boston University. The phycobiliproteins from Synechococcus 6312 - which contain larger amounts of allophycocyanin than usual - have also been retained for future study.

Specific endonucleases (restriction enzymes) have been isolated from A. variabilis protein fractions and are being studied in this department.

A. variabilis plastocyanin was sent to Dr. R.J.P. Williams, Biochemistry Dept., Oxford University for nuclear magnetic resonance studies.

The purity of the plastocyanin and cytochrome f preparations was checked by polyacrylamide gel electrophoresis. Tris-glycine buffer, pH 8.7, was used with 10% gels using the system described by Ornstein (1964).

Several attempts were made to isolate plastocyanin from Spirulina maxima: Cytochrome f was prepared from this organism by methods similar to those used for the purification of A. variabilis cytochrome f but no trace of plastocyanin was detected. Small amounts of the cytochromes c-(549, 552 and 554) were isolated from Anacystis nidulans by the methods described by Holton & Myers (1967) but no plastocyanin was detected. Cells of Synechococcus 6312 became available from Professor R.Y. Stanier and this unicellular blue-green alga proved to be a better source of larger amounts of cytochrome f and the cytochrome from this strain was subsequently studied. Again no plastocyanin was detected in any of the preparations of cytochrome f from this last organism.

It is possible that these strains contain plastocyanin but it may be chromatographing with the phycocyanins during the initial stages where its blue colour would be completely masked by the much more intense colour of the phycocyanins. The phycobiliproteins are not usually completely removed even in a 60% ammonium sulphate precipitate. Experience has also shown that in the preparation of plastocyanin from Chlorella fusca it is difficult to prevent the

reduction of this plastocyanin in the initial stages of purification to its colourless form even in the presence of excess potassium ferricyanide. It is also the practise in some laboratories to carry out protein preparations with buffers containing mercaptoethanol (about 0.1 mM). Plastocyanin can not be oxidised in the presence of this reagent and may have remained undetected. Electron paramagnetic resonance studies have indicated the presence of plastocyanin in Anacystis nidulans (Visser et al., 1974).

The isoelectric point of some blue-green algal plastocyanins may be about 7 (cf. chromatography of A. variabilis plastocyanin at pH 7 on DEAE- and CM-cellulose). Some oxidised plant plastocyanins are unstable at their isoelectric point (Ramshaw et al., 1973)

Phormidium luridum cytochrome f was prepared in considerably higher yield from one batch of cells grown in the medium of Cannon et al. (1971). This medium contains 5.85 mg iron/ℓ compared to 0.1-0.6 mg/ℓ in most blue-green bacterial media. This success could not be repeated, however.

Table 2(a) Final protein purification conditions

A. variabilis plastocyanin	(a) Eluted from CM-52 at 5 mM phosphate buffer pH 7. (b) Gel filtration (occasionally) on Sephadex G75 (100 cm x 1 cm) in 0.05 M ammonium acetate pH 5.5.
A. variabilis cytochrome f	(a) Eluted from CM-52 at 25 mM phosphate buffer pH 7. (b) pH adjusted to 4.1 and absorbed onto CM-23, pH 3.9. Eluted from CM-23 with 50 mM ammonium acetate, pH 4.6. (c) Gel filtration on Sephadex G-50 (2.5 x 30 cm) in 5 mM phosphate buffer pH 7 (occasionally).
P. luridum plastocyanin	(a) Eluted from CM-52 at 25 mM phosphate buffer pH 7. (b) Small amounts of this protein were available and this plastocyanin was not further purified except for ammonium sulphate precipitation and desalting on Sephadex G-25 into volatile buffer for analyses or sequenator studies. (c) Occasionally step (a) was repeated on a smaller column to remove traces of cytochrome f.
P. luridum cytochrome f	(a) Eluted from CM-52 at 5 mM phosphate buffer pH 7 (oxidised form) (b) Eluted from CM-23 (ammonium acetate, pH 3.9) at pH 4.8.

Synechococcus 6312 cytochrome f	(a) Absorbed onto DEAE-cellulose (pH 7) and eluted with 10 mM phosphate buffer, pH 7. (b) Did not absorb onto CM-52 at pH 7 even at 1 mM phosphate buffer. (c) Eluted from CM-23 (ammonium acetate, pH 3.9) at pH 4.4.

Table 2(b) Yields of blue-green algae and protein from 120 ℓ culture

Strain	Wet weight of cells	Plastocyanin	Cytochrome f
A. variabilis	240 g	25 mg	1.5 mg
P. luridum	500-700 g	5 mg	2.5 (20 mg from one batch)
Synechococcus 6312	-	-	8 mg from 150 g wet packed cells

CHAPTER 3

Properties of the Proteins

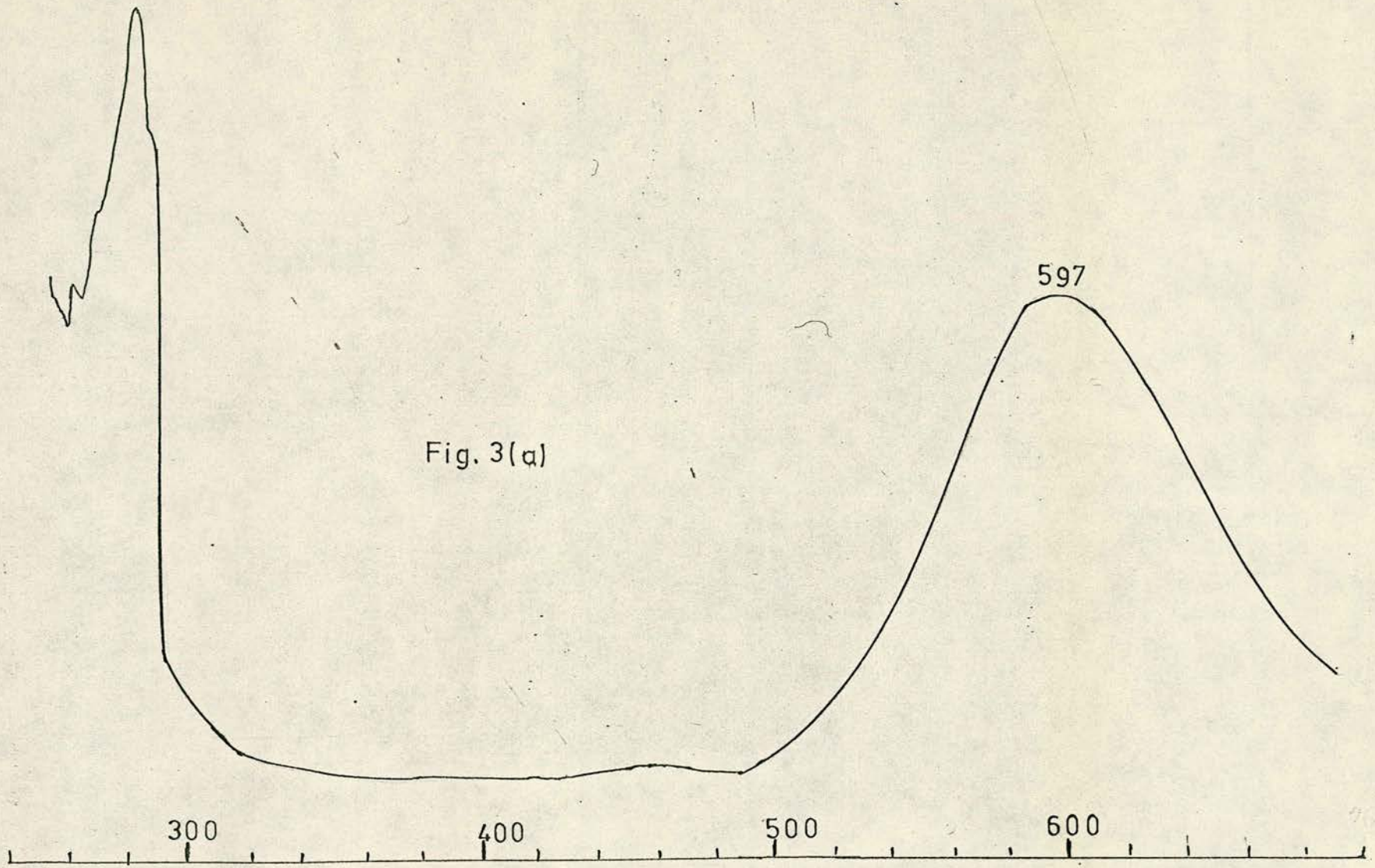


Fig. 3(a)

(a) Properties of blue-green algal plastocyanins

The spectral properties of A. variabilis plastocyanin and P. luridum plastocyanin were identical with those obtained by Lightbody and Krogmann (1967) and Biggins (1967) respectively.

The spectrum of A. variabilis plastocyanin is reproduced in Fig. 3(a). It shows a broad maximum at 597 mμ in the oxidised form and fine structure in the ultraviolet with a maximum at 278 nm. The best value of the absorption index (E_{278}/E_{597}; Katoh et al., 1962) obtained was 1.45. The absorption index of A. variabilis plastocyanin was usually around 1.5 - 1.6.

The absorption index of P. luridum plastocynin was 1.7 in one sample. A lower value might have been expected due to presence of 2 tyrosine residues instead of 3 in A. variabilis plastocyanin. Analysis and sequence results were satisfactory and loss of some copper would explain this discrepancy.

The extinction coefficient of the copper chromophore in plastocyanins is approximately 4.5×10^3 and the absorption spectrum in the ultraviolet region indicates lack of tryptophan in the blue-green algal plastocyanins (in common with higher plant plastocyanins (Ramshaw et al., 1973). The absorption maximum at 460 nm is pronounced in some higher plants and Chlamydomonas (Scawen et al., 1975) but is very weak in blue-green algae (and Chlorella).

The copper content of the plastocyanins, estimated with bathocuproine (Wharton & Rader, 1970), was an average

of 1.1 mole Cu/mole protein for 7 samples of Anabaena variabilis plastocyanin and 1.0 for a single sample of Phormidium luridum plastocyanin when the protein concentration was measured spectrally. Some plant plastocyanins are known to lose copper readily (Ramshaw et al., 1973) and the molar ratio of copper to protein for A. variabilis plastocyanin was 0.8 when the protein concentration was measured by hydrolysis followed by amino acid analysis. These results compare favourably with those of Ramshaw et al. (1973) for higher plant plastocyanin.

The molecular weight of A. variabilis plastocyanin was estimated from Sephadex G-75 gel filtration to be 12,000 daltons (Andrews, 1964). This compares with the molecular weight of 11,148 calculated from the amino acid sequence.

A. variabilis plastocyanin appeared to run as a dimer of 25,000 daltons in 8% SDS polyacrylamide gels (Weber & Osborn, 1969). This experiment was carried out by L. Morss of this department. Siegelman et al. (1975) have recently reported 40,000 dalton spinach and Scenedesmus plastocyanin species in SDS gels and have suggested that this is the active component. A molecular wt. of 21,000 has been reported for spinach plastocyanin (Katoh et al., 1962).

(b) Determination of Copper

The copper content of the plastocyanins was estimated with bathocuproine (Wharton & Rader, 1970). A volume of plastocyanin solution containing an estimated 20 to

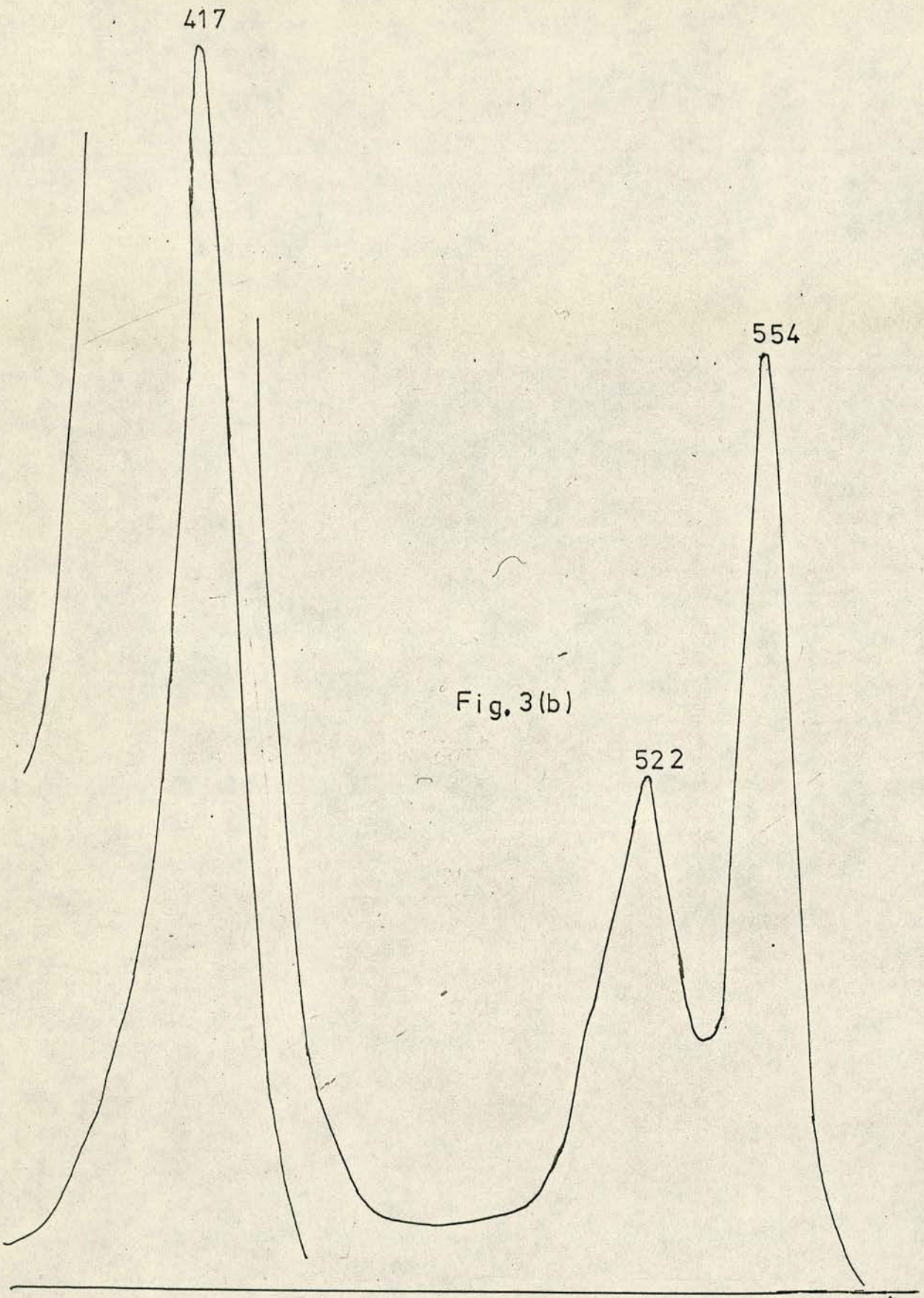


Fig. 3(b)

80 n mol copper was added to 12 ml glass centrifuge tubes and diluted to 0.4 ml with water. Mercaptoacetic acid (0.1 ml of 5% aqueous solution) was added to reduce the copper followed by 0.5 ml of 0.1% solution of bathocuproine (4,7-diphenyl-1,10-phenanthroline) in glacial acetic acid. Ethanol (0.5 ml, 95%) was added and the tube covered with parafilm. The cuprous-bathocuproine complex was extracted by mixing for 5 min. with a vortex mixer. The suspension was centrifuged for 10 minutes at top speed in a clinical centrifuge. The clear supernatant was transferred by Pasteur pipette to a spectrophotometer cuvette and the absorbance at 479 mμ was determined versus a blank containing identical sample minus bathocuproine. The amount of copper was estimated from a standard curve. Readings were taken on a Unicam SP500 Series 2 spectrophotometer.

(c) Cytochrome f Spectra

The absorption spectrum of Anabaena variabilis cytochrome f in the visible region is shown in Fig. 3(b). The absorption maxima of reduced blue-green algal cytochromes f are shown below. All the spectra are very similar in shape and absorption maxima although Synechococcus 6312 is a possible exception.

	Absorption peaks of reduced cytochromes f		
	α	β	Soret
A. variabilis (present study)	554	523	417
A. variabilis (Susor & Krogmann, 1966)	554	522	416
P. luridum (Biggins, 1967)	554	523	417
Synechococcus 6312	552	520	414
A. nidulans (Holton & Myers, 1967)	554	522.5	416.5
A. nidulans (Meyer, 1970)	553.3	522	416
Spirulina maxima	554	524	-

A. variabilis cytochrome f spectra were measured on a Unicam SP1800 spectrophotometer which was out of action when *P. luridum* cytochrome f spectra were run. The latter were measured on a Unicam SP800 instrument. *Synechococcus* 6312 was measured on a Cary spectrophotometer at the Pasteur Institute, Paris, and it is possible that the alignment of this instrument varied by about 2 nm from other instruments.

The spectral data given here are incidental to the main project but reduced cytochrome f spectra (and oxidised plastocyanin spectra) give a quick measure of quantity (and purity if other cytochromes are initially present). The extinction coefficient at 554 nm (ε_{554}) used for the cytochromes f was 29 (Lemberg & Barrett, 1973).

CHAPTER 4

Protein sequence methods

Materials were obtained from the sources listed by Ambler & Wynn (1973) unless otherwise stated.

(a) Denaturation of proteins

Before digestion, the proteins were first denatured to render them susceptible to proteolysis. A. variabilis plastocyanin was denatured by precipitation by addition of an equal volume of 10% (w/v) trichloracetic acid. This left the copper in solution. The precipitated protein was washed thoroughly with aliquots of cold acetone then dissolved in 0.1 M-NH_3 solution.

Cytochromes f were denatured by removal of the heme group with mercuric chloride. The cytochrome was passed into 5% (v/v) formic acid by gel filtration through Sephadex G-25 (20 cm x 1.5 cm) and freeze dried. The protein (about 1.0 μmol) was dissolved in 2.5 ml of a solution containing 0.1 M HCl and 8 M urea. 40 mg finely ground $HgCl_2$ was added with shaking on a vortex mixer. The solution was incubated overnight at 37°C on a shaker. The apoprotein was separated from haem and reagents by gel filtration through Sephadex G-25 (20 cm x 1.5 cm) equilibrated with 50% formic acid.

The apoprotein occasionally precipitated during removal of the haem. Alkaline pH was avoided to help prevent deamidation at -Asn-Gly- sequences and is the reason why the latter gel filtration step was not carried out with 0.1 M NH_3.

(b) Proteolytic digestion

Apoproteins were dissolved in 0.2 M ammonium acetate,

pH 8.5, and digested with $\frac{1}{40}$th by wt. trypsin, chymotrypsin or thermolysin at 37°C for about 4 h. A. variabilis plastocyanin was digested with Staphylococcus aureus proteinase under the same conditions, overnight (Houmard & Drapeau, 1972).

Papain subdigestion was carried out in 0.2 ml pyridine acetate, pH 5.5 containing 1% (w/v) dithiothreitol. An equal volume of a suspension of papain (1 mg/ml) in the same buffer was added and incubated at 37°C for 5 h.

Pseudomonas pyrollidonecarboxylyl peptidase (Doolittle, 1970) was used to remove the N-terminal pyrollidonecarboxylic acid residue from one tryptic peptide. Details are given in the results section.

Carboxypeptidase A digestion of proteins and peptides was carried out according to method (a) of Ambler (1972). The substrate was dissolved in 0.2 ml N-ethylmorpholine acetate buffer, pH 8.5. Carboxypeptidase A (10 µg, 0.1 Enzyme Commission unit) was added and the sample digested at 37°C for various times. 50 µg of carboxypeptidase A was sometimes used for whole proteins. At the end of the incubation period a few drops of glacial acetic acid was added to stop the reaction and the sample dried in a vacuum desiccator. The dried samples were then dissolved in pH 2.2 citrate buffer and analysed on an amino acid analyser. Enzyme blanks were run in parallel.

Enzymes were obtained from the following sources:-

Trypsin - Worthington Biochemical Corp. Freehold, N.J. U.S.A. and pretreated with diphenylcarbamyl chloride before use.

Chymotrypsin - Worthington Biochemical Corp. Freehold, N.J. U.S.A. and pretreated with soya bean trypsin inhibitor before use.

Thermolysin - Daiwa Kasei, K.K., Osaka, Japan.

Papain - from Sigma London Chemical Co., London, S.W.6, U.K.

Carboxypeptidase A - (DFP treated) from Sigma London Chemical Co., London, S.W.6, U.K.

Staphylococcus aureus proteinase and Pseudomonas pyrollidonecarboxylyl peptidase were prepared by A.J. Savill, this department.

(c) Cleavage by Hydroxylamine

The method of Bornstein (1970) was used to cleave P. luridum cytochrome f at -Asn-Gly- bonds. The apoprotein (12 mg) was dissolved in 2 ml freshly prepared 2 M hydroxylamine solution. Hydroxylamine hydrochloride (recrystallised from ethanol: H_2O, 3:1, v/v) was dissolved in the minimum amount of 5% KOH then the pH adjusted to 9.0 with K_2CO_3 solution. The volume was adjusted with distilled water to give a final molarity of 0.2 for hydroxylamine. The solution was incubated at 45° for 2 h. The products were passed down Sephadex G50 (150 cm x 1.5 cm) in 0.1 M ammonia.

Aliquots were taken from selected fractions and analysed for N-terminal amino acids by the dansyl method. The larger fragments (and probably some uncleaved protein) were pooled, freeze-dried and subdigested with thermolysin. Fractions 4 and 5 were freeze-dried and passed down

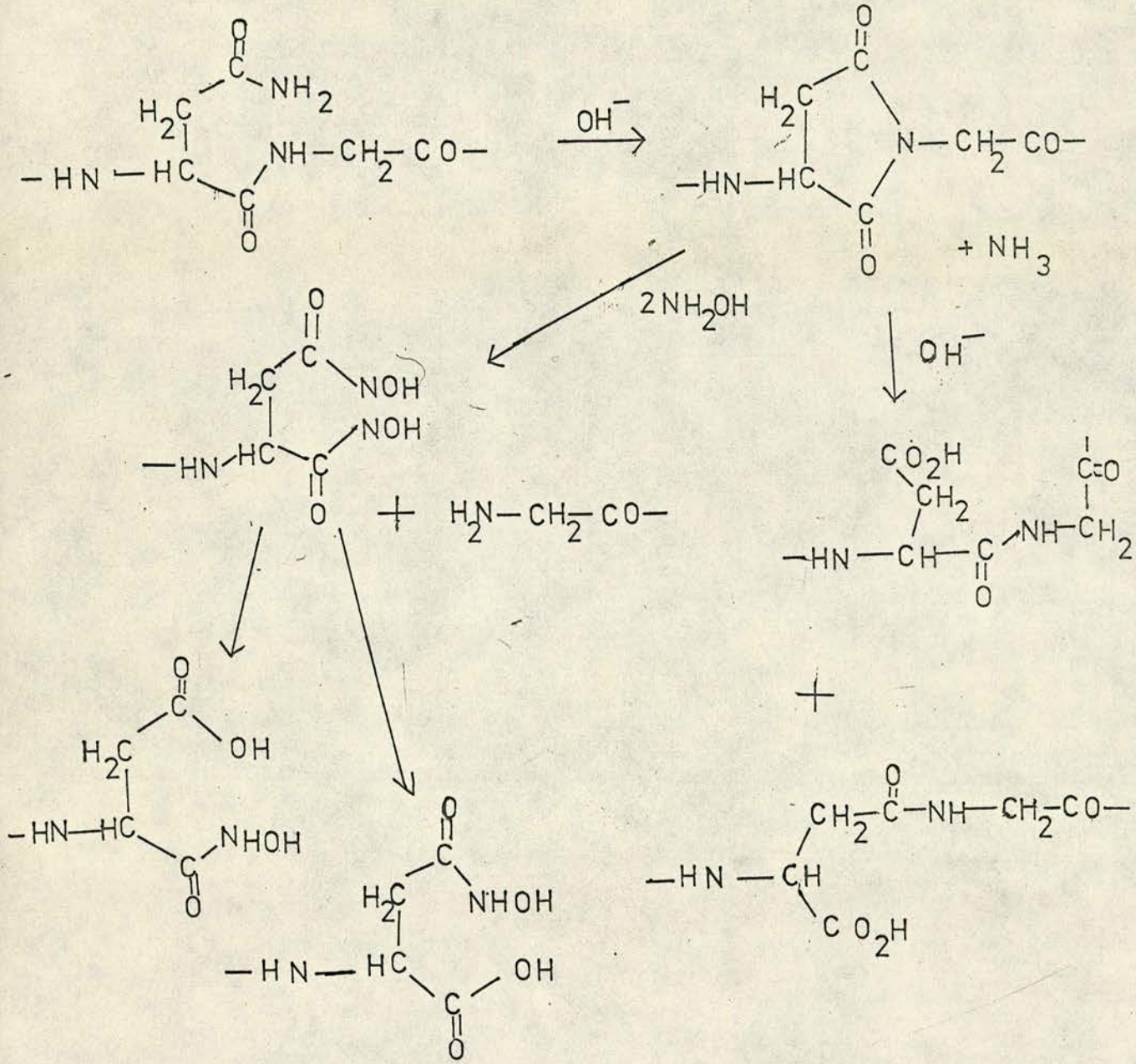


Fig.4 (a)

Sephadex G-25 superfine (85 cm x 1 cm) and the peptides purified by the usual methods.

The proposed mechanism for the cleavage is shown in Fig. 4(a) (Deselnicu et al., 1973).

(d) Separation of peptides by size

Peptides from digests of whole protein were initially fractionated on the basis of size by gel filtration on Sephadex G-25 superfine (80 to 90 cm x 1 cm) in 5% formic acid. The column effluent was monitored on an L.K.B. Uvicord at 254 nm and collected in 1.5 ml fractions. 50 - 100 µl aliquots of each fraction were spotted on Whatman 3 MM paper in 10 µl applications in successive 1 cm bands and peptides separated at right angles by electrophoresis at pH 6.5. After drying the paper was stained with ninhydrin (0.2% in acetone + 1% glacial acetic acid) and heated at 105°C for 5 min. The paper was subsequently stained for histidine (Dent, 1947) or alternatively stained for ninhydrin negative peptides with starch/KI (Rydon & Smith, 1952).

This peptide map was used as a basis for deciding which fractions to pool. Tryptophan containing peptides were often retarded to a greater extent on Sephadex G-25 and eluted after the small ion volume.

(e) Purification of peptides on paper

Whatman 3 MM paper was used for preparative work and Whatman No.1 paper for analytical work and occasionally for purification of peptides in low yield. Details of buffers

etc. are given in Table 4(a). Electrophoresis was carried out in Michl tank apparatus (Michl, 1951) with white spirit or toluene coolant. Maximum loading of peptide was 100 nmol/cm for 3 MM paper and 20 nmol/cm for No.1 paper.

Electrophoresis at pH 6.5 was the first stage in purification on paper. Occasionally extra peptide markers were applied to either side of the band of peptide. This enabled one, without intermediate elution, to sew on the peptide band with the extra markers to another piece of paper for paper chromatography or pH 3.5 electrophoresis. This was most frequently carried out with the neutral band of peptides. Peptides containing cysteine were identified at this stage by oxidising another peptide marker strip after pH 6.5 electrophoresis with performic acid vapour. The marker strip was placed in a desiccator containing performic acid (formic acid:hydrogen peroxide, 20:1) and the desiccator evacuated partly. After thorough airing the strip was sewn onto another piece of paper and electrophoresed at pH 6.5 at right angles to the first direction (Brown & Hartley, 1966). This produced a diagonal pattern of peptides with cysteic-acid-containing-peptides lying off the diagonal (to the acidic side). The band containing the cysteine peptide was then identified, oxidised as before and rerun in the electrophoresis tank at pH 6.5 after sewing onto another piece of paper.

Peptide bands were eluted with 0.1 M ammonia or in the case of cytochrome f peptides when rearrangement of Asn-Gly bonds was to be avoided, with 1 M acetic acid.

Aliquots of the peptides were spotted onto No.1 paper for pH 3.5 electrophoresis and paper chromatography (BAWP). On the basis of these analytical runs peptides were purified accordingly or if single spots only resulted on both papers the peptides were subjected to N-terminal analysis and qualitative amino acid analysis at pH 2.0. N-terminal analysis was by the dansyl method (see below) and qualitative amino acid analysis by hydrolysis of an estimated 10 nmol of peptide with 6 MHCl followed by electrophoresis at pH 2. The paper was stained with a solution of ninhydrin, 0.2%; acetic acid, 1%; collidine, 0.1%, in acetone. Characteristic colours were obtained from many amino acids.

If material was in short supply this qualitative analysis was missed out and 10 nmol (estimated from intensity of dansyl amino acid) was analysed on an LKB Biocal BC200 amino acid analyser using a single column programme. Qualitative amino acid analysis was occasionally carried out with hydrolysis followed by dansylation (Scawen et al., 1974).

Paper chromatography was generally avoided as the final purification step and was followed by an electrophoresis step. If additional purification was required after paper chromatography and pH 3.5 electrophoresis the peptide mixture was normally subjected to electrophoresis at pH 2. This proved particularly useful for small neutral peptides. Peptides are not clearly distinguished from free amino acids at this pH but extra markers containing mixtures of all the amino acids were spotted alongside the usual ones and the paper marker strips dipped in ninhydrin solution containing collidine as before. Also by this stage of purification

Table 4(a) Electrophoresis and chromatography buffers

Electrophoresis at pH 6.5	pyridine:acetic acid:water, 25:1:225 (v/v), toluene coolant, 3 kV.
Electrophoresis at pH 3.5	pyridine:acetic acid:water, 1:10:89 (v/v), white spirit coolant, paper wetted with half strength buffer, 3 kV.
Electrophoresis at pH 2	acetic acid:formic acid:water, 4:1:45 (v/v), white spirit coolant, 3 or 7 kV.
Paper chromatography (BAWP)	(Waley & Watson, 1953) n-butanol:acetic acid:water: pyridine, 15:3:12:10 (v/v), descending chromatography, 16 h or longer.

the N-terminal amino acids of the peptide mixture were probably known and free amino acids could be identified by viewing the characteristic colours which only the free amino acids form with collidine.

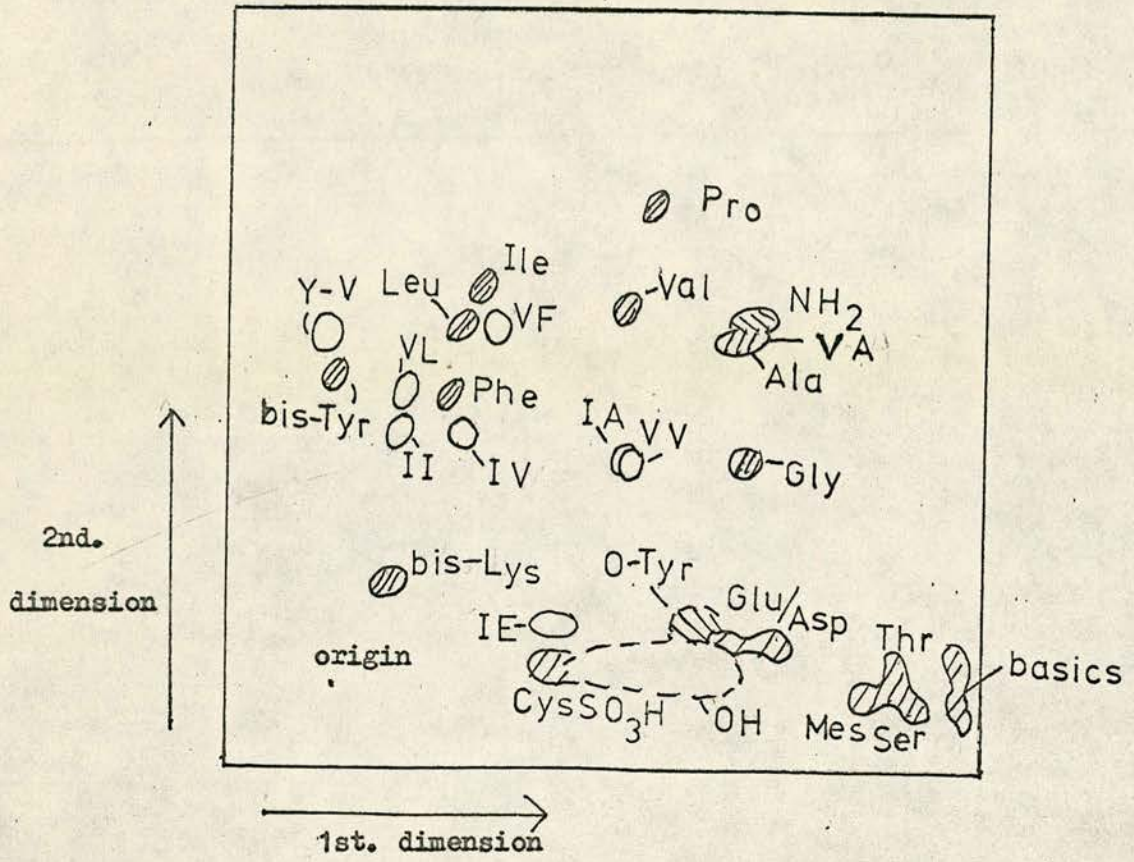
(f) Amino acid analysis

Routinely quantitative amino acid analysis was carried out using the Biocal analyser (see above). A Beckmann 120C amino acid analyser using a 2 column system was normally used for analysis of whole protein and occasionally for large peptides. The latter analyser was set on a lower sensitivity but gave a logarithmic print out which was useful when the sample contained widely different amounts of each constituent amino acid.

(g) Analysis of Tryptophan

Tryptophan was analysed separately after hydrolysis with 3 M mercaptoethanesulphonic acid (Penke et al., 1974), obtained from Pierce Chemicals, Rockford, Illinois, U.S.A. Proteins were hydrolysed for 96 h and peptides for 24 h. The solution after hydrolysis was neutralised with NaOH and placed directly onto the analyser resin bed after the addition of pH 2.2 citrate buffer containing internal standard. The analysis was carried out on the short column (6 cm x 1 cm) of the Beckmann 120C amino acid analyser at pH 5.28.

Tryptophan was also identified by analysis after carboxypeptidase A digestion.


Pro
Ile
Y-V
Leu
Val
NH2
VF
VA
VL
Ala
Phe
bis-Tyr
IA
VV
II
IV
Gly
2nd.
dimension
bis-Lys
O-Tyr
Glu/Asp
IE
Thr
basics
origin
CysSO3H
OH
Mes
Ser
1st. dimension

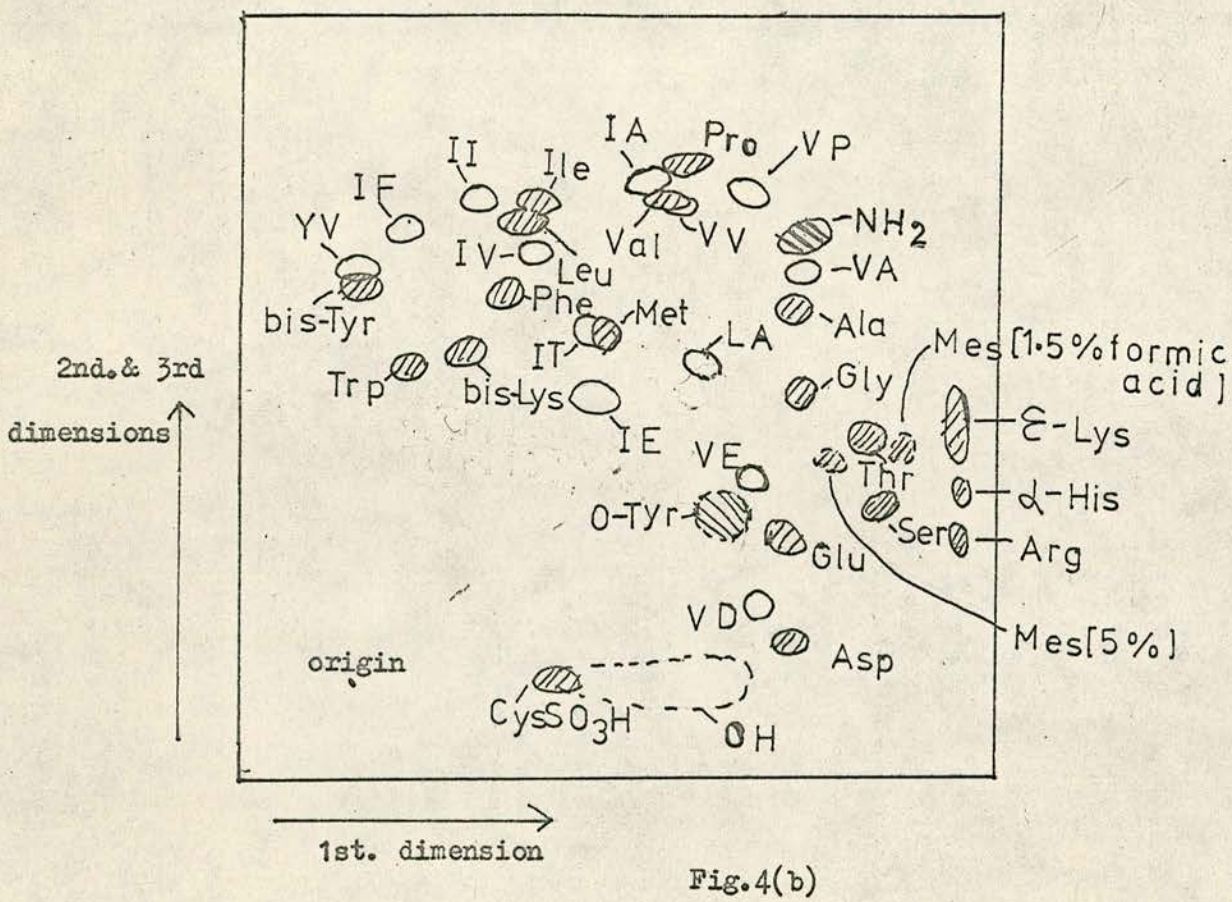

IA
Pro
VP
II
Ile
IF
YV
VV
NH2
Val
IV
Leu
VA
Phe
Met
bis-Tyr
Ala
LA
Mes[1.5% formic acid]
2nd.& 3rd
dimensions
IT
Trp
bis-Lys
Gly
ε-Lys
IE
VE
Thr
α-His
O-Tyr
Ser
Arg
Glu
VD
Mes[5%]
origin
Asp
CysSO3H
OH
1st. dimension

Fig.4(b)

(h) Identification of N-terminal groups

This was the dansyl method of Gray (1972a). The dansyl amino acids were identified by thin layer chromatography on polyamide layer sheets (Hartley, 1970) and viewed under ultra-violet light.

The dansyl derivatives were co-chromatographed with a standard mixture on the reverse side. After coupling with dansyl reagent the peptide was normally hydrolysed at 105°C overnight with 6 M HCl. Many acid-stable dipeptide derivatives can be identified [see Fig. 4(b)]. The presence of Dns-proline was usually confirmed by hydrolysis for 3-5 h.

Dns-Arg, -His, -α-Lys and -ε-Lys were often successfully identified on polyamide layers after chromatography in a 4th solvent, but if necessary they were identified by electrophoresis at pH 4.4 on a flat-plate. apparatus (Gray, 1972a).

(i) Subtractive dansyl - Edman degradation

This was essentially the method of Gray (1972b). The peptide (10 nmol x no. of cycles intended or as much material as was available) was dried in a 12 mm x 65 mm screw cap tube. Pyridine (50%, v/v; 0.2 ml) and phenylisothiocyanate (5%, v/v in pyridine; 0.1 ml) were added. The tube was gassed with oxygen-free nitrogen (OFN) and capped. The solution was incubated at 37°C for 1 h and dried at 60°C under vacuum for 1 h. Trifluoracetic acid (o.15 ml) was added, the tube gassed with OFN and incubated at 60°C for 5 min. The solution was dried under vacuum

over NaOH in a hot desiccator. Water (0.25 ml) was added and the solution extracted 3 times with 2 ml aliquots of butyl acetate. The top, organic phase after centrifugation was discarded. A sample was taken for N-terminal group determination and the next cycle of degradation begun.

Due to extraction losses, partial blockage of exposed N-terminal group, hydrolysis of the peptide by trifluoracetic acid etc., the yields of Dns-amino acids recovered begins to fall after a few cycles. If many cycles are required, more material can be used initially and small aliquots removed each time. Normally the new N-terminal amino acid was identified while the succeeding degradation was carried out and the amount of degraded peptide removed was adjusted to take sufficient for a good identification but allow one to conserve material (2-3 nmol is sufficient for identification on polyamide layers but more is required for electrophoresis). Aliquots of degraded peptide were also frequently taken and electrophoresed at pH 6.5 for location of amide residues.

(j) Automatic phenylisothiocyanate degradation

Native oxidised plastocyanin and native reduced cytochrome f (0.3 - 0.8 μmol) were degraded in a Beckmann, model 890A sequenator by using a 70 min Quadrol double cleavage programme (Edman & Begg, 1967).

Later, cytochromes f were degraded using a single cleavage peptide programme with dimethylbenzylamine replacing Quadrol (Hermodsen et al., 1972).

The phenylthiohydantoin derivatives produced were identified by thin layer chromatography on silica-gel layers (Eastman-Kodak, Rochester, N.Y., U.S.A.) containing fluorescent indicator by using the solvent system of Edman & Sjoquist (1956), by g.l.c. on SP4001 Chromosorb W and by hydrolysis followed by amino acid analysis. Hydrolysis was carried out with 6 M HCl for 24 h at 150°C (Van Orden & Carpenter, 1964) and 5.5 M HI for 24 h at 150°C (Inglis et al., 1971). The latter method was particularly useful for PTH-threonine and PTH-serine which are hydrolysed to α-amino-n-butyric acid and alanine respectively. HCl hydrolysis destroys both these derivatives completely. α-Amino-n-butyric acid elutes between alanine and valine during analysis.

(k) Assignment of amide or acidic side groups

These were assigned on the basis of electrophoretic mobilities at pH 6.5 of small peptides (Offord, 1966). In some cases the mobilities of residual peptides after a number of Edman degradation steps were used as an aid to distinguishing Asp from Asn and Glu from Gln.

Carboxypeptidase A digestions and results from automatic sequencing were also used. The latter was not used as the sole criterion for any acidic side-group owing to the possibility of deamidation of the phenylthiohydantoin derivatives of asparagine and glutamine before identification.

(1) Cyanogen bromide cleavage of P.luridum cytochrome f

This was undertaken on material from 2 sources.

(1) About 50 nmol native cytochrome f which was the pure "dregs" from previous experiments.

(2) The residual film recovered from the cup of the automatic sequencer after 26 cycles of the D MBA programme.

The combined material was freeze-dried and acetylated with acetic anhydride (Riordan & Vallee, 1967), in order to block amino-terminal groups.

The material was dissolved in 0.5 ml of 50% saturated sodium acetate solution in an ice-bath and 5 x 2 µl aliquots of acetic anhydride were added during 1 h. to the solution at 0°C. with stirring. The stirring was continued for an additional hour and the product was desalted on a column of Sephadex G-25 (medium, 15 cm x 1.5 cm) equilibrated with 5% (v/v) formic acid. The salt-free material was freeze-dried and dissolved in 1 ml of HCl (0.1 M). Cyanogen bromide (50 times excess; Gross, 1967) was added and the solution was incubated at 37°C for 24 h and freeze-dried. The products were dissolved in 0.1 M ammonia and applied directly to the spinning cup of the automatic sequencer.

RESULTS SECTION

CHAPTER 5

Anabaena variabilis plastocyanin

(a) Standard abbreviations used in tables and figures in results section

M — Electrophoretic mobility on paper at pH 6.5 relative to lysine (+ ve) or aspartic acid (- ve), corrected for endosmosis.

V/V_o — Relative elution volume on gel-filtration through Sephadex G-25 equilibrated with 5% (v/v) formic acid (except hydroxylamine cleavage peptides where Sephadex G-50 values are given).

Purn. — Purification methods, in order of use: G, Sephadex G-25 gel-filtration; paper electrophoresis, 6, pH 6.5; 3, pH 3.5; 2, pH 2; B, paper chromatography in butan-1-ol-acetic acid - water-pyridine (15:3:12:10, by vol.); ox, performic acid oxidation on paper before elution.

N-t — N-terminus identified by dansylation (observed Dns-dipeptides are occasionally included).

Yield — Absolute recovery of purified peptide (mol/100 mol protein) from primary digests. Sub-digestion and hydroxylamine cleavage peptide yields are given in nmol.

----/ — Indicates substandard Dns-phenyl isothiocyanate degradation results.

 Indicates Dns-dipeptide also observed.

( ) Bracketed sequence results are tentative.

X Indicates substandard values of amino acids in peptides.

tlc Thin layer chromatography.

glc Gas/liquid chromatography.

PTH- Phenylthiohydantoin derivative of an amino acid.

DMBA Dimethylbenzylamine automatic sequencer buffer.

VILP As a sequencer tlc result this indicates presence of Val, Ile, Leu, Pro or Phe which do not separate well in the solvent system used.

N.D. Not determined (usually due to lack of material).

Pauly stain - Sulphanilic acid/nitrous acid stain for His (Dent, 1947).

αABA α-Amino butyric acid, the product of HI hydrolysis of PTH-threonine.

(b) Amino acid compositions of proteins

Results are shown as residues per molecule. Analysis values were from 24 h hydrolyses, except for the following amino acids; valine and isoleucine were derived from 96 h hydrolyses, cysteine (as cysteic acid) and methionine (as

methionine sulphone) from performic acid - oxidised samples hydrolysed for 24 h. Values for serine and threonine were not corrected for destruction during hydrolysis. Tryptophan was determined by hydrolysis with mercaptoethane sulphonic acid [Chapter 4(g)].

(c) Criteria for satisfactory results (Kelly & Ambler, 1974)

Dansyl-phenylisothiocyanate degradation results were considered unsatisfactory when more than one dansyl-amino acid was identified as being present in a subjective visual estimate of more than 20% of the sequence residue at any step of the degradation. Accidents and shortage of material also contributed to unsatisfactory results.

Amino acid analyses were considered satisfactory when:- (a) no impurity was present in amounts larger than 0.2 mol/mol.

(b) the relative amounts of amino acids present, calculated on the basis that the average amount was integral, do not fall outside the limits 0.8-1.2, 1.8-2.2, 2.7-3.3 or 3.7-4.3. Values as low as 0.7(=1) for tyrosine were considered acceptable. If acid stable bonds containing valine and isoleucine are subsequently proved, low values for these amino acids are often acceptable. 96 h hydrolyses of such peptides were undertaken if amounts of material permitted. N-terminal amino acids of peptides are often recovered in reduced yield.

All values above 0.2 mol/mol were recorded on the data sheets.

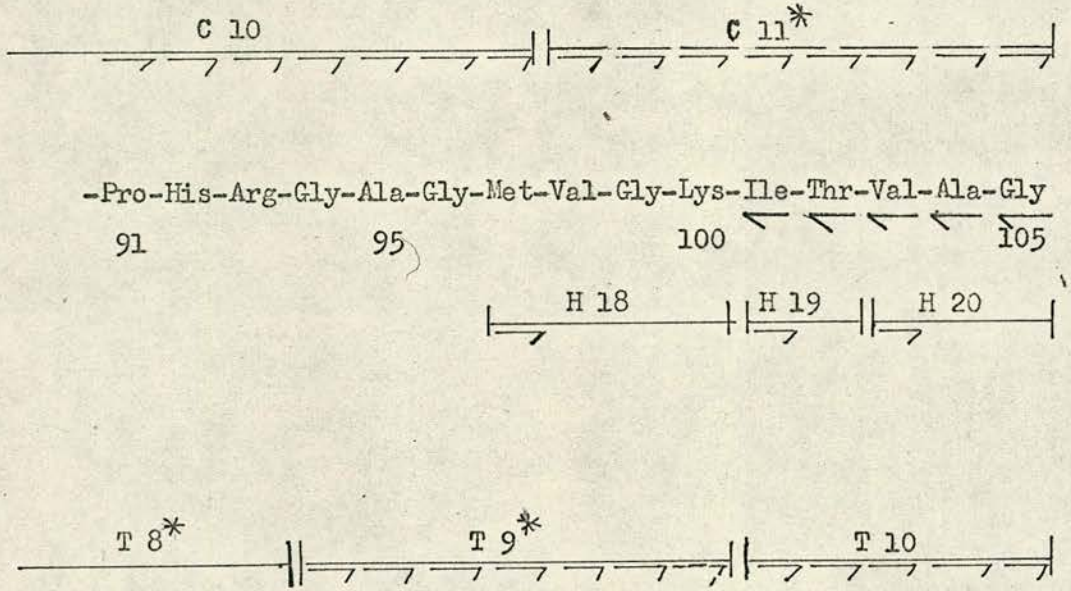

C 10
C 11*
-Pro-His-Arg-Gly-Ala-Gly-Met-Val-Gly-Lys-Ile-Thr-Val-Ala-Gly
91
95
100
105
H 18
H 19
H 20
T 8*
T 9*
T 10

Fig.5(a) contd.

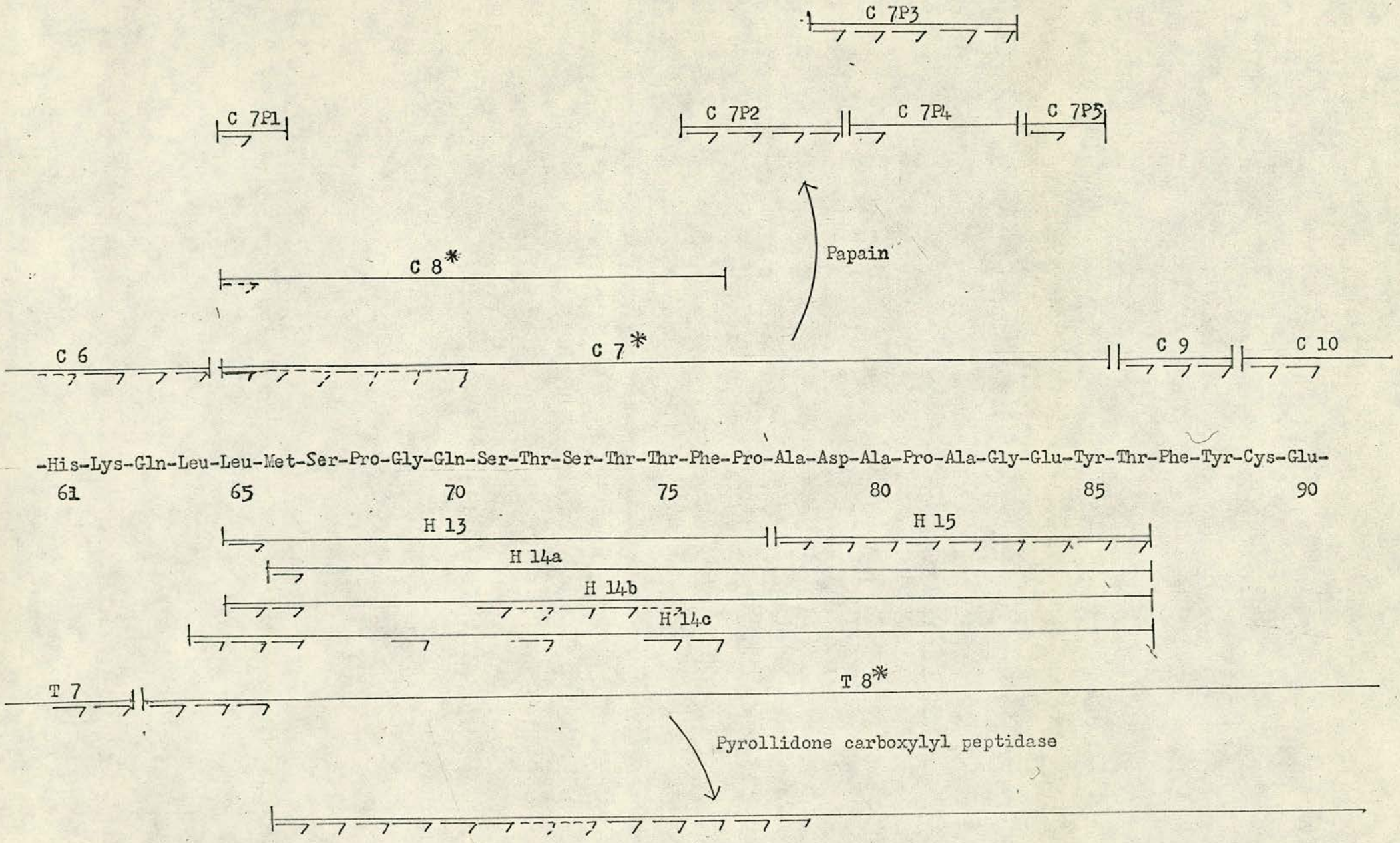

C 7P3
C 7P1
C 7P2
C 7P4
C 7P5
C 8*
Papain
C 6
C 7*
C 9
C 10
-His-Lys-Gln-Leu-Leu-Met-Ser-Pro-Gly-Gln-Ser-Thr-Ser-Thr-Thr-Phe-Pro-Ala-Asp-Ala-Pro-Ala-Gly-Glu-Tyr-Thr-Phe-Tyr-Cys-Glu-
61
65
70
75
80
85
90
H 13
H 15
H 14a
H 14b
H 14c
T 7
T 8*
Pyrollidone carboxylyl peptidase

Fig.5(a) contd.

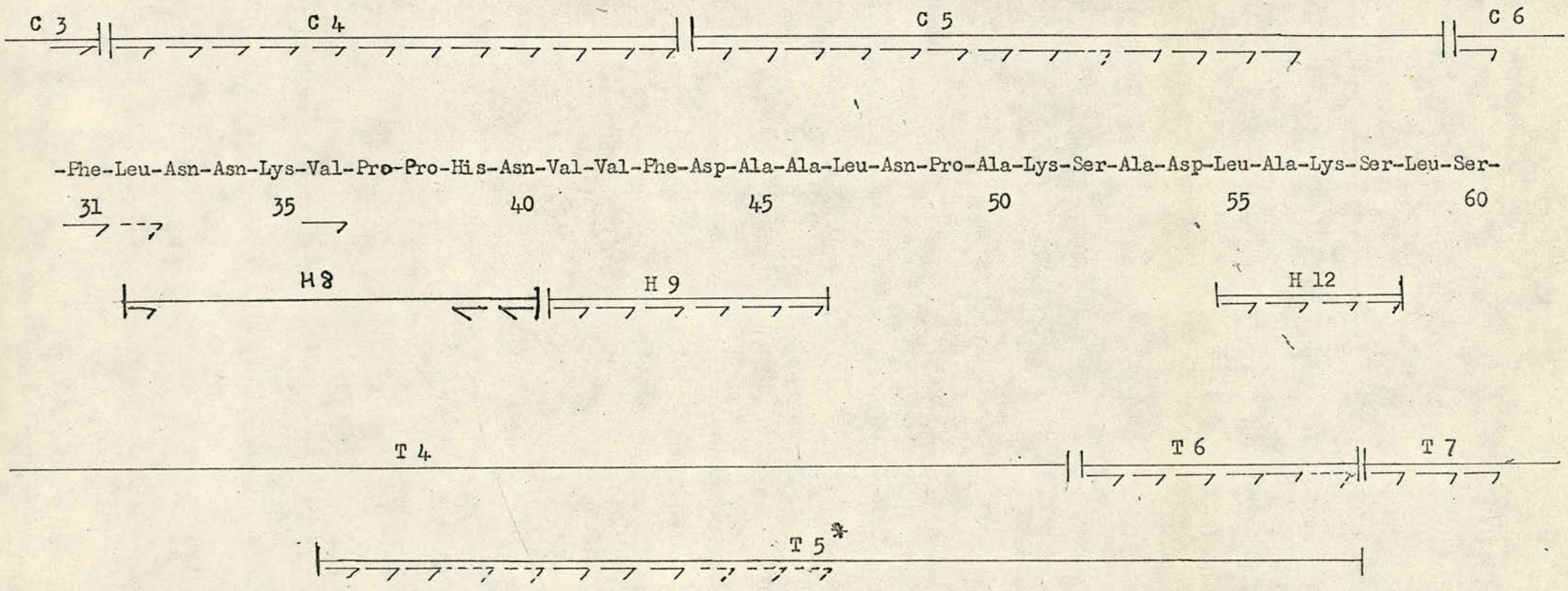

C 3
C 4
C 5
C 6
-Phe-Leu-Asn-Asn-Lys-Val-Pro-Pro-His-Asn-Val-Val-Phe-Asp-Ala-Ala-Leu-Asn-Pro-Ala-Lys-Ser-Ala-Asp-Leu-Ala-Lys-Ser-Leu-Ser-
31
35
40
45
50
55
60
H 8
H 9
H 12
T 4
T 6
T 7
T 5*

Fig.5(a) contd.

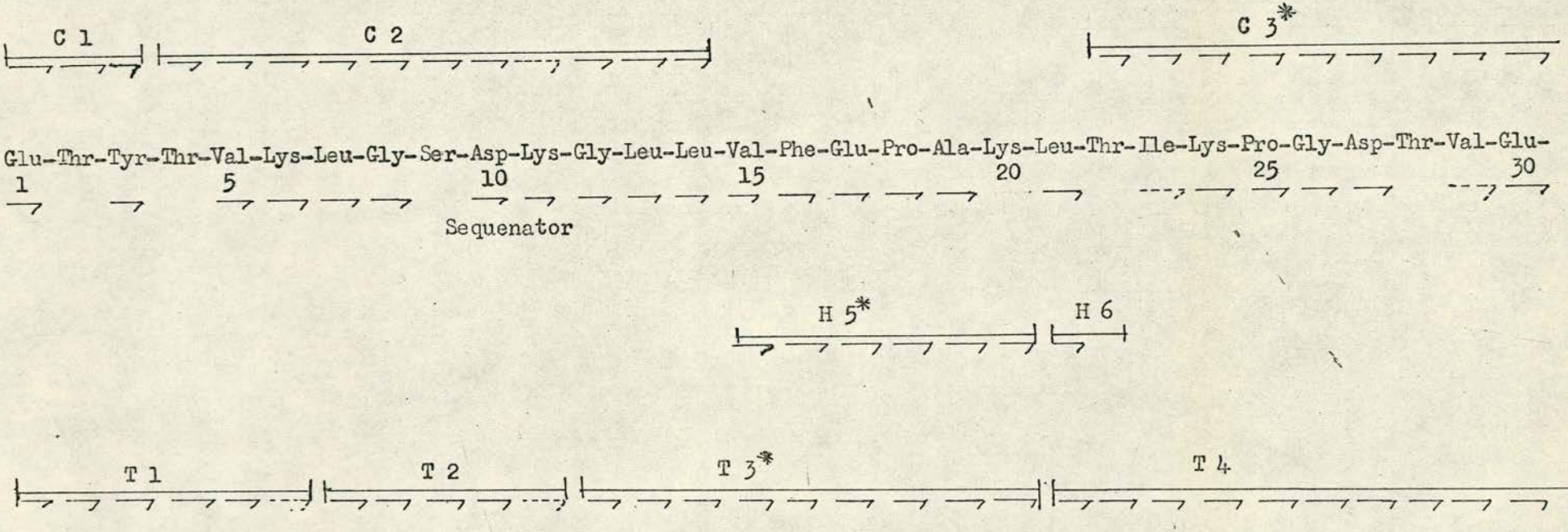

C 1
C 2
C 3*
Glu-Thr-Tyr-Thr-Val-Lys-Leu-Gly-Ser-Asp-Lys-Gly-Leu-Leu-Val-Phe-Glu-Pro-Ala-Lys-Leu-Thr-Ile-Lys-Pro-Gly-Asp-Thr-Val-Glu-
1
5
10
15
20
25
30
Sequenator
H 5*
H 6
T 1
T 2
T 3*
T 4

Fig.5(a)

Fig. 5(a)

Amino Acid Sequence of *Anabaena Variabilis* Plastocyanin

Peptides derived by chymotryptic digestion (C) are shown above the sequence and by tryptic (T) and thermolysin (H) digestion below the sequence. Vertical arrows show peptides formed by further digestion. Full lines indicate quantitative amino acid analyses (substandard if marked *). Under the sequence —↓ indicates residues identified after degradation of the protein in a sequenator. Under peptide lines, —↓ indicates end groups and subsequent residues revealed by phenylisothiocyanate degradation, and identified by the dansyl method (substandard if shown ---↓) and ↖— indicates residues released by carboxypeptidase A. For clarity, only those thermolysin peptides required to complete overlaps are shown on the diagram.

(d) Proposed sequence of A. variabilis plastocyanin

The sequence determination was carried out on material derived from different batches of cells and the products of different isolations of plastocyanin. Tryptic digestion was carried out on 1.8 μmol of denatured plastocyanin; chymotryptic digestion and thermolysin digestion on 2 μmol each of denatured apoprotein.

The amino acid compositions of A. variabilis plastocyanin is shown in Table 5 (a).

Automatic phenylisothiocyanate degradation on 0.8 μmol native oxidised plastocyanin unequivocally identified 24 of the first 36 residues [Fig. 5(a)]. Three of the remaining residues were tentatively identified (Ile - 23, Val - 29 and Leu - 32). Conventional sequence determination experiments showed these undetected residues to be threonine (4), serine (1), asparagine (2) and lysine (2). Where traces of subsidiary phenylthiohydantoins were detected they were consistent with ragged degradation and there was no suggestion of microheterogeneity at any position. The yields of phenylthiohydantoins were not quantitated.

The evidence for the amino acid sequence is summarised in Fig. 5(a). This shows all the peptides isolated from the tryptic and chymotryptic digests together with a sufficient number of thermolysin peptides to complete the overlaps. The peptides obtained by papain sub-digestion are also shown. Symbols indicate how much of the sequence of each peptide was determined by the

dansyl-phenylisothiocyanate method. Cases where dansyl-phenylisothiocyanate degradation results were considered not wholly satisfactory are indicated. The detailed properties of each peptide, dansyl-Edman sequence evidence from each peptide and carboxypeptidase A digestion results are given in tables at the end of this chapter.

The guidelines followed in considering dansyl-phenylisothiocyanate degradation results or amino acid analyses to be satisfactory are those stated in the previous part of this chapter.

The peptides which did not meet these analytical criteria are marked with an asterisk in Fig. 5(a) and failed for the following reasons:-

(1) Peptide T3 contained 0.4 serine and 0.2 threonine.

(2) Peptide T5 was an impure, minor tryptic peptide. Dansyl-phenylisothiocyanate degradations revealed no impurities in peptides T3 or T5.

(3) Peptide T8 contained 31 residues which is near the maximum size for satisfactory paper purification. The value for aspartic acid was high (1.35 = 1) and the value for proline was low (3.37 = 4). This peptide was oxidised on paper with performic acid and the value for tyrosine (1.2 = 2) is higher than might be expected. The values for glycine (2.55 = 2) and lysine (0.17 = 0) may be due to the presence of a small amount of a larger peptide which included Gln-63 to Lys-96.

(4) Peptide T9 appeared to contain tyrosine (0.38 = 0) but

no dansyl-tyrosine was detected on polyamide sheets or during dansyl-phenylisothiocyanate degradation.

(5) Peptide C3 had threonine (1.36 = 2) which may be due to the fact that threonine was N-terminal and that both Thr-Val and Thr-Ile bonds are present. After 96 h hydrolysis 1.46 threonine was recovered.

(6) Peptide C7 appeared to be pure except for the low value for leucine (0.55). The N-terminal amino acids were later proved to be leucine, methionine and serine in the subjective ratio of intensity 4:1:1. It would appear that chymotrypsin also cleaved after Leu-65 and Met-66. This peptide mixture was subdigested with papain. In a later digest 3 almost identical thermolysin peptides from the same region were successfully separated.

(7) Peptide C8 was a minor chymotryptic peptide with N-termini similar to C7. The values for threonine (2.27) and phenylalanine (0.54) indicate partial splitting between Thr - 74 and Thr - 75. This peptide is included as an aid to amide assignment.

(8) Peptide C11 appeared to contain serine (1.08) and proline (0.41) but no trace of serine or proline was detected after dansyl-phenylisothiocyanate degradation for 9 cycles. This was repeated with the same result. These amino acids are absent from the overlapping tryptic and thermolysin peptides. Degradation of the C-terminus of the protein with carboxypeptidase A revealed no serine or proline.

(9) Values for methionine are occasionally low - particularly for thermolysin peptides H14 where a large number of purification steps were required and partial oxidation has presumably occurred.

(10) Values for valine (in thermolysin peptides particularly) are often low where this amino acid is N-terminal - for example peptide H5 (valine = 0.64 and phenylalanine = 0.72) and peptide H9 with Val-Val N-terminal sequence. Sufficient material for a 96 h hydrolysis was not always available. In the cases of peptides H5 and H9 dansyl N-terminal study revealed the presence of considerable amounts of dansyl-Val-Phe and dansyl-Val-Val respectively, after 16 h hydrolysis.

The amino acid sequence of Anabaena variabilis plastocyanin was deduced from tryptic, chymotryptic and thermolysin digests. Homology with other plastocyanins was not used in the deduction of the sequence. Tryptic digestion of the protein was extremely useful in deducing the sequence due to the higher proportion of lysine and arginine (9 and 1 residues respectively) compared to other plastocyanins.

Evidence for the sequence is generally good but weaknesses in the proposed sequence are discussed below.

The amino acid composition of the protein determined by analysis of acid hydrolysates [Table 5(a)] is in good agreement with that of the deduced sequence. This is evidence that the protein is 105 residues long and that no regions have been deleted by non-discovery of peptides in any digest.

Overlap between residues 20 and 21 is not well established from purified peptides and automatic sequencing has been used to establish this overlap. A major chymotryptic peptide which probably mapped from residue 15 to residue 21 was lost during purification. Spots indicative of a peptide which reacted slowly with ninhydrin were seen on marker strips a considerable number of days after dipping in ninhydrin. The putative chymotryptic peptide would have had N-terminal valine. However, the results of chymotryptic digestion were consistent with the rest of the proposed sequence and no peptides, minor or major, were purified which did not fit into the sequence.

Some difficulty was experienced in dansyl-phenylisothiocyanate degradation between residues 71 to 75 inclusive. The proposed sequence is -Ser-Thr-Ser-Thr-Thr-.
71 75

The sequence is preceded by a Gly-Gln sequence and dansylphenylisothiocyanate degradation is known to become "ragged" on occasion after glycine residues (Ambler & Wynn, 1973). Degradations were carried out on two thermolysin peptides through this region and twice on tryptic peptide T8 after digestion with Pseudomonas pyrollidonecarboxylyl peptidase (Doolittle, 1970) which removed the N-terminal pyrollidone carboxylic acid residue which had formed during purification. In this instance the next two residues (both leucine) were also removed and a peptide with N-terminal methionine was produced. This peptide had a V/V_o ratio (Sephadex G25) and tyrosine content (bis-DNS tyrosine identified on polyamide sheets) which were consistent with

the parent peptide. The results of dansyl-phenylisothiocyanate degradation were in all instances consistent with the proposed sequence although in one or two cases the amount of contaminant dansyl-amino acid was present in greater amounts than desirable. Papain subdigestion on peptide C7 produced an apparently random assortment of neutral serine and threonine containing peptides in low yields - none of which could be satisfactorily purified. Pure subdigestion peptides were readily obtained from other parts of peptide C7. The N-terminal region of C7 was very "ragged" - discussed in (6) above.

The proposed sequence in this region is also consistent with the amino acid compositions of minor chymotryptic and thermolysin peptides.

Cyanogen bromide cleavage of the protein was not attempted as material was in short supply and the predicted result was another large fragment similar in size to the peptides C7, T8 and H14 already obtained, with similar problems of subdigestion. Digestion of plastocyanin was previously attempted with the extra-cellular proteinase from Staphylococcus aureus (Houmard & Drapeau, 1972) but the specificity of this batch of enzyme was not confined to cleavage on the C-terminal side of acidic residues. A considerable number of other bonds were cleaved (including proline linkages) and although a peptide was isolated whose sequence appeared to be Pro-Ala-Lys-Leu-Thr-Ile-Lys-Pro-Gly-Asp which would appear to provide overlap between residues 20 and 21, this digest is not given further consideration.

Assignment of amide or acidic groups in the sequence was on the basis of electrophoretic mobilities at pH 6.5 of small peptides (Offord, 1966) supplemented with information from carboxypeptidase A experiments. Acidic side groups were also assigned in the N-terminal region when phenylthiohydantoin derivatives of aspartic or glutamic acid were identified after degradation of the protein on the sequenator. See Table 5(b). This was not used as the sole criterion for any acidic residue owing to the possibility of deamidation of the PTH-derivatives of asparagine and glutamine during conversion from the thiazolinones (Niall, 1973).

Table 5(a)

Amino acid composition of A. variabilis plastocyanin

	Analysis	Sequence
Glycine	8.9	9
Alanine	11.0	11
Valine	7.2	8
Leucine	10.0	10
Isoleucine	1.8	2
Serine	7.1	7
Threonine	8.1	9
Aspartic acid	8.8	5
Asparagine		4
Glutamic acid	7.0	5
Glutamine		2
Phenylalanine	4.9	5
Tyrosine	2.7	3
Tryptophan	0	0
Cysteine	1.1	1
Methionine	1.9 (2.2 as sulphone)	2
Proline	8.5	9
Lysine	8.5	9
Histidine	3.0	3
Arginine	1.1	1

Table 5(b)

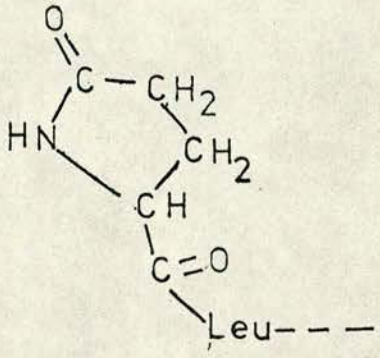
Evidence for Amide Assignments in Anabaena Variabilis Plastocyanin

Glu-1	Sequenator result. Acidic mobility of peptide C1 (-0.4, Glu-Thr-Tyr) neutral mobility of peptide T1 (Glu-Thr-Tyr-Thr-Val-Lys)
Asp-10	Sequenator result. Neutral mobility of peptide T2 (Leu-Gly-Ser-Asp-Lys)
Glu-17	Sequenator result. Neutral mobility of peptide T3 (Gly-Leu-Leu-Val-Phe-Glu-Pro-Ala-Lys)
Asp-27) Glu-30)	Sequenator results. Acidic mobility of peptide C3 (-0.24; Thr-Ile-Lys-Pro-Gly-Asp-Thr-Val-Glu-Phe)
Asn-33) Asn-34) Asn-40)	Basic mobilities of peptide C4 (+0.35; His_1, Lys) and peptide T4 (+0.11; His_1, Lys_3, Asp_2, Glu_1) Carboxypeptidase A digestion on peptide H8 indicated C-terminal Asn.
Asp-44	Acidic mobility of peptide N9 (-0.32; Val-Val-Phe-Asp-Ala-Ala)
Asn-48	Basic mobility of peptide H10 (+0.5; Leu-Asn-Pro-Ala-Lys-Ser)
Asp-54	Acidic mobility of peptide H11 (-0.59; Ala-Asp)

The neutral mobility of peptide C5 (Asp-Ala-Ala-Leu-Asn-Pro-Ala-Lys-Ser-Ala-Asp-Leu-Ala-Lys-Ser-Leu) confirms the presence of 2 Asp and 1 Asn.

Gln-63 Basic mobility of peptide C6 (+0.57; Ser-His-Lys-Gln-Leu)

Peptide T8 with N-terminal Gln-63 was ninhydrin negative as it readily (during purification) formed the pyrollidone carboxylyl residue:-



Gln-70 Neutral mobility of peptide C8 [Leu-Met-Ser-Pro-Gly-Gln-Ser .....]

Asp-79) Acidic mobility of peptide H15 (-0.64; Ala-Asp-
Glu-84) Ala-Gly-Glu-Tyr-Thr)

Acidic mobilities of papain subdigest peptides C7 P2 (-0.45), C7 P3 (-0.52) and C7 P5 (-0.60).

Glu-90 Basic mobility of peptide C10 (+0.08 - before oxidation; Cys-Glu-Pro-His-Arg-Gly-Ala-Gly-Met).

(e) Detailed evidence for the amino acid sequence of *A. variabilis* plastocyanin

(1) Tables showing the properties of all the peptides shown in Fig. 5(a) plus thermolysin peptides from the rest of the protein. The successive steps used in the purification of each peptide are given and where applicable the values of V/V_o [elution volume/void volume for gel filtration through Sephadex G25 in 5% (v/v) formic acid] and electrophoretic mobility or paper at pH 6.5. Absolute percentage yields are given for peptides from chymotryptic and tryptic digests. Amino acid analyses are given for all peptides.

(2) The individual sequence evidence for each peptide is on tables. This consists of the N-terminal residue analysis and the results of dansyl-phenylisothiocyanate degradation. Details are given of substandard results.

(3) Amino acid analysis obtained after degradation of the protein and peptide H8 by carboxypeptidase A.

Properties of tryptic peptides from Anabaena variabilis plastocyanin.

Peptide	m 6.5	V/Vo G-25	Purn	Yield	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
T 1	0	1·86	G6B32	7			1·01				2·03		1·09		0·76					1·11			6	Glu	1–6
T 2	0	2·1	G6B32	6	1·01			1·0		0·93		1·02								1·04			5	Leu	7–11
T 3	0	1·76	G6B32	5	1·15	1·11	0·98	2·02		0·43 ×	0·23 ×		1·16	0·86					0·97	1·23			9	Gly	12–20
T 4	+0·11	1·19	G63	14	1·1	2·98	3·31	2·76	1·01		1·73	6·07	1·14	2·02					4·09	2·85	0·95		31	Leu	21–51
T 5	0	1·48	G6B3	8·5	1·75 ×	2·64 ×	2·04	1·69	1·07 ×	1·26 ×	0·88 ×	3·62	2·00 ×	1·11					2·42 ×	1·57 ×	0·81		~22	Val	includes 36–57
T 6	0	1·9	G6B3	15		1·91		1·06		0·89		1·04								0·93			6	Ser	52–57
T 7	+0·78	1·67	G63	15				0·96		1·84										1·01	1·19		5	Ser	58–62
T 8	−0·15	1·14	G6ox6	20	2·55 ×	3·08		2·15		2·99	4·03	1·35 ×	4·08	1·85	1·23		0·76	0·91	3·37 ×	0·17 ×	0·98	0·96	31	Glx	63–93
T 9	+0·48	1·67	G63	14	2·90	1·19	0·88								0·38 ×			0·96		1·07			7	Gly	94–100
T 10	0	2·1	G6B3	33	1·02	1·06	1·06		0·88		1·00												5	Ile	101–105

Sequence experiments on tryptic peptides from A.variabilis plastocyanin.

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	Other expts.
T1	6	GLX	THR	THR	THR	VAL	---> LYS					
T2	5	LEU	GLY	SER	ASX	---> LYS						
T3	10	GLY	LEU	LEU	VAL	PHE	GLX	PRO	ALA	LYS		
T4	30	LEU	THR	ILE	LYS	PRO	GLY	ASX	THR	VAL	GLX	
T5	22	VAL	PRO	PRO	---> HIS	---> ASX	VAL	VAL	PHE	---> ASX	---> ALA	---> ALA
T6	6	SER	ALA	ASX	LEU	ALA	---> LYS					
T7	5	SER	LEU	SER	HIS	LYS						
T8	31	GLX	LEU	LEU								
T8 after digestion with pyrallidone carboxylyl peptidase:-		MET	SER	PRO	GLY	GLX	SER	---> THR	---> SER	THR	THR	-10 PHE -11 PRO -12 ALA
T9	7	GLY	ALA	GLY	MET	VAL	GLY	---> LYS				
T10	5	ILE	THR	VAL	ALA	GLY						

Properties of chymotryptic peptides from Anabaena variabilis plastocyanin.

Peptide	m 6.5	V/Vo G-25	Purn	Yield	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
C1	−0·4	2·65	G6	65							0·98		1·07		0·95								3	Glu	1-3
C2	+0·2	1·82	G63	30	2·06		1·05	2·94		1·00	0·95	1·02								1·94			11	Thr	4-14
C3	−0·24	1·64	G63	15	0·99		1·03		0·91		1·46	1·01	1·06	1·03					0·94	1·01			10	Thr	22-31
C4	+0·35	1·76	G63	20			2·17	0·81				2·94		1·02					1·97	1·09	1·01		12	Leu	32-43
C4(96h)							3·00	0·85				2·97		1·03					2·15						
C5	0	1·41	G63	35	× 0·22	4·90		3·18		1·89		2·94							1·11	1·98			16	Asp	44-59
C6	+0·57	2·43	G6	50				1·03		1·13			0·94							1·03	0·93		5	Ser	60-64
C7	−0·24	1·52	G63	30	2·02	3·00		× 0·55		2·90	2·94	1·03	2·08	1·09	0·92			0·92	3·08				21	Leu (Met,Ser)	65-85
C8	0	1·94	G632	10	1·06			× 0·41		2·92	× 2·27			× 0·54				0·91	1·09				10-12	Leu (Met)	65-74,76
C9	0	3·06	G63	30							0·94			1·10	0·96								3	Thr	86-88
C10	+0·08	1·41	G6ox6	20	2·04	0·96							1·00				0·83	0·82	0·99		0·96	1·04	9	Cys	89-97
C11	+0·35	1·76	G63	20	2·00	1·00	2·00		0·91	× 1·08	0·82								× 0·41	1·03			8	Val	98-105

Properties of C 7 papain subdigest peptides from A. variabilis plastocyanin.

Peptide	m 6.5	V/Vo G-25	Purn	Yield nmol	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N–t
C7P1	0			12	x 0.2			1.02		x 0.3								0.98					2	Leu
C7P2	-0.45			100		1.05						1.07		0.82					1.06				4	Phe
C7P3	-0.52			60	0.96	2.02						1.09							0.91				5	Asp
C7P4	0			25	1.22	1.90													0.88				4	Ala
C7P5	-0.60			240						x 0.24			1.14		0.86								2	glu(O-tyr)

Sequence experiments on chymotryptic (and papain subdigest) peptides from A.variabilis plastocyanin.

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	Other expts.
C1	3	GLU	THR	TYR										
C2	11	THR	VAL	LYS	LEU	GLY	SER	ASX	LYS	GLY	LEU	LEU		
C3	10	THR	ILE	LYS	PRO	GLY	ASX	THR	VAL	GLX	PHE			
C4	12	LEU	ASX	ASX	LYS	VAL	PRO	PRO	HIS	ASX	VAL	VAL	PHE	
C5	16	ASX	ALA	ALA	LEU	ASX	PRO	ALA	LYS	SER	ALA	ASX		
C6	5	SER	HIS	LYS	GLX	LEU								
C7	19-21	LEU	MET	SER	PRO	GLY	GLX							
C8	10-12	LEU	—	—										
C9	3	THR	PHE	TYR										
C10	9	CYS	GLX	PRO	HIS	ARG	GLY	ALA	GLY	MET				
C11	8	VAL	GLY	LYS	ILE	THR	VAL	ALA	GLY					
C7P2	4	PHE	PRO	ALA	ASX									
C7P3	5	ASX	ALA	PRO	ALA	GLY								

Properties of thermolysin peptides from Anabaena variabilis plastocyanin(a).

Peptide	m 6.5	V/Vo G-25	Putn	Yield	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
H1	-0.59		G 63								0.97		1.03										2	Glu	1-2
H2	-0.39		G 63								2.04		1.10		0.86								4	Glu	1-4
H3	+0.85		G 6				0.82													1.18			2	Val	5-6
H4	0		G 6 3 2		2.05			0.94		0.95		1.13								0.93			6	Leu	7-12
H5	0		G 6 3			1.21	× 0.64						1.02	× 0.72					0.78	0.98			6	Val+ V-F	15-20
H6	0		G 6 3 2					0.99			1.01												2	Leu	21-22
H7	-0.24		G 6 B 3 2		1.09		1.06		0.89		0.97	1.06	1.14	1.21					× 0.60	1.00			8	Ile	23-31
H8	+0.5		G 6 3 2				1.04	0.88				3.30							2.12	0.81	0.87		9	Leu	32-40
H9	-0.32		G 6			1.98	× 0.81					1.22		0.98									6	Val+ V-V	41-46
H9(96h)						1.97	× 1.28					1.15		0.89											
H10	+0.5		G 6 3			1.04		0.90		0.95		1.12							0.86	1.00			6	Leu	47-52
H11	-0.59		G 6 3			0.96						1.04											2	Ala	53-54
H12	+0.5		G 6 3 2			1.16		0.89		1.05										0.91			4	Leu	55-58
H13	0		G 6 3 2		1.13			1.18		2.62	2.79		1.06	0.78				0.68	1.97				13	Leu	65-77

Properties of thermolysin peptides from Anabaena variabilis plastocyanin(b).

Peptide	m 6.5	V/Vo G-25	Putn	Yield	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
H14a	-0·24		G6B32		2·22	3·23				2·86	3·83	1·12	2·10	1·04	0·93			0·87	2·83				21	Met	66-86
H14b	-0·24		G6B32		2·16	3·06		0·81		2·85	3·79	0·98	2·00	1·05	0·92			partly ox.	2·76				22	Leu	65-86
H14c	-0·24		G6B32		2·20	3·08		1·78		3·07	3·58	1·26	1·87	1·06	1·01			0·64	2·92				23	Leu	64-86
H15	-0·64		G63		1·03	2·83					1·09	1·22	1·08		0·99				1·13				9	Ala	78-86
H16	+0·33		G6B32		2·5	1·18				× 1·02	× 0·31	× 0·55	1·92	0·86	0·6		0·75	some	× 0·67		× 1·25	× 1·59	15	Gly	83-97
H17	0		G632		1·16	0·85																	2	Ala	95-96
H18	+0·5		G6B32		1·19		0·94											0·89		0·98			4	Met	97-100
H19	0		G63						0·95		1·05												2	Ile	101-102
H20	0		G632		1·15	1·02	0·83																3	Val	103-105

Sequence determination experiments on thermolysin peptides from *A.variabilis* plastocyanin.

(N.D.: No sample taken for dansyl determination due to shortage of material)

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	Other expts.
H14b	22	LEU	MET	N.D	N.D	N.D	N.D	SER	THR	SER	THR	THR			(At step 7 on H14b and step 8
H14c	23	LEU	LEU	MET	N.D	N.D	GLY	N.D	N.D	THR	N.D	N.D.	THR	PHE	on H14c some SER was
H15	9	ALA	ASX	ALA	PRO	ALA	GLY	GLX	TYR	THR					present - probably overlap
H9	6	VAL	VAL	PHE	ASX	ALA	ALA								from the adjacent residues)
H12	4	LEU	ALA	LYS	SER										
H5	6	VAL	PHE	GLX	PRO	ALA	LYS								

(3) Carboxypeptidase A Experiments on Anabaena Variabilis Plastocyanin and Peptide H8 (Ambler, 1972).

Figures after the amino acids indicate relative amounts of each amino acid recovered.

Denatured plastocyanin (---Lys-Ile-Thr-Val-Ala-Gly)
Gly 1.0, Ala 0.7, Val 0.7, Thr 0.5, Ile 0.4 (lysine was not determined; Analysis was done on long column of Beckmann 120C amino acid analyser.

Peptide H8 (---Pro-His-Asn) "Asn" 1.0, His 0.36.
("Asn" peak on analyser is presumed correct because the peptide contained neither Ser nor Gln).

CHAPTER 6

Amino-terminal sequence of plastocyanin from P. luridum

The amino acid composition of P. luridum plastocyanin (compared with that from A. variabilis) is shown in Table 6(a). The results obtained using a quadrol double cleavage programme are shown in Table 6(b). Identification of a residue was considered satisfactory when it was positively identified by at least 2 of the 3 methods. In some cases a good amino acid analysis result and tentative tlc and glc results were considered acceptable. Exceptions are residues 38 and 39 (steps 37 and 38). The amino acid analysis results were good, and PTH-Histidine is not extracted into ethyl acetate but remains in the aqueous layer (along with PTH-Cys and PTH-Arg) as the thiazolinone derivative. When the aqueous and remaining organic phases are combined and hydrolysed the thiazolinone derivatives are converted to free amino acids (Inglis et al., 1971). Tlc satisfactorily distinguishes Asp from Asn and Glu from Gln although Asn and Gln are deamidated to the extent of around 20-30% during conversion of the thiazolinones to the PTH derivatives and the deamidated PTH amino acids are detected on the plates (Niall, 1973).

The alignment of P. luridum plastocyanin with plastocyanins from A. variabilis, Chlorella fusca and higher plants is shown in Fig.10(c) The best fit is obtained if a deletion at position 13 is considered. 35 of the first 38 residues were unequivocally identified. Gly-34

(one of the tentative assignments), Val-33 and Met-7 are the only exceptions to what have previously been invariant residues in 41 partial plastocyanin sequences (D. Peacock, personal communication).

Table 6(a) Amino acid compositions of plastocyanins from *P. luridum* (a) and *A. variabilis* (b).

	(a)	(b)
Glycine	11.7	8.9
Alanine	12.7	11.0
Valine	11.8	7.2
Leucine	6.4	10.0
Isoleucine	3.7	1.8
Serine	3.4	7.1
Threonine	4.5	8.1
Aspartic acid and Asparagine	9.3	8.8
Glutamic acid and Glutamine	7.8	7.0
Phenylalanine	5.2	4.9
Tyrosine	2.1	2.7
Tryptophan	0	0
Cysteine	1.1	1.1
Methionine	1.7	2.2
Proline	6.5	8.5
Lysine	10.0	8.5
Histidine	3.2	3.0
Arginine	1.0	1.1
Total	102.1	

Table 6(b) P. luridum plastocyanin automatic sequencing (quadrol programme)

step	tlc	intensity	glc	nmol	hydrolysis/analysis	nmol	sequence residue
1	Asp	3+	Asp	124			Asp
2	Thr	+	Thr	-	α-ABA	70	Thr
3	VILP	2+	Val	128	Val	82	Val
4	Lys	4+	Lys	-			Lys
5	VILP	3+	Val	62	Val	82	Val
6	VILP	3+	Ile	94	Ile + alloIle	-	Ile
7	Met	3+	Met	167	(destroyed)		Met
8	Gly	3+	Gly	60	Gly	75	Gly
9	Gly	2+	Gly	143	Gly	(fault)	Gly
10	(Ser)	3+	?	-	? (20 nmol Lys)		(Ser)
11	Lys	3+	Lys	-			Lys
12	Gly	2+	Gly	120			Gly
13	VILP	2+	Leu	112	Leu	45	Leu
14	VILP	2+	Val	92	Val	(fault)	Val
15	VILP	2+	Phe	77	Phe	32	Phe
16	Glu	2+	Glu	26			Glu
17	VILP	½+	Pro	38	Pro	45	Pro
18	Ala	2+	Ala	120			Ala
19	VILP	2+	Val	108	Val	68	Val
20	VILP	+	Val	74	Val(tr. Gly & Ala)	77	Val
21	Asn	+	-		Asx	62	Asn
22	VILP	2+	Val	50	Val	57	Val
23	Lys	½+	-		Lys	32	Lys
24	Ala	+	Ala				Ala

(Contd.

Table 6(b) (contd.)

step	tlc	intensity	glc	nmol	hydrolysis /analysis	nmol	sequence residence
25	Gly	½+	-		Gly	64	Gly
26	Asp	2+	Asp	9			Asp
27	VILP	½+	Thr	7	α-ABA	18	Thr
28	VILP	2+	Ile	20	Ile & allo-Ile		Ile
29	Glu/ Gln	+ & +	some Glu & Gln		Glx	40	Gln
30	-		Phe	14	Phe	14	Phe
31	Glu	+	Glu	9			Glu
32	VILP	+	Val	20	Val	10	Val
33	-		-		Gly	22	(Gly)
34	Glu/ Gln	+ & +	-		Glx (and Gly)	32 (17)	Gln
35	VILP	+	-		Leu	5	(Leu)
36	VILP	½+	Pro	8	Pro	18	Pro
37	-	-	-		Pro	19	Pro
38	VILP	½+	-		His	10	His

CHAPTER 7

Proposed amino acid sequence of P. luridum cytochrome f

The amino acid composition of P. luridum cytochrome f is shown in Table 7(a). This is in reasonable agreement with the proposed sequence except for the value for Arg (1.7 = 1). Although tryptic digestion results were poor all peptide mixtures were subjected to qualitative analysis by electrophoresis at pH 2, some were analysed quantitatively and there was no sign of any Arg containing peptide other than PT8. Similarly, the only Arg containing peptides from hydroxylamine and thermolysin digests were those mapping through residue 68. Slight traces of proteinaceous material with high Arg content could account for this as there is only 1 residue of this amino acid in the protein. The value for Ser is low due to destruction to the extent of 6% per 24 h (Ambler & Brown, 1967). The corrected value for 24 h hydrolysis is 2.65 and for 96 h hydrolysis it is 2.73.

Analysis values for Glu are also low (5.3 and 5.2 = 6) but the proposed sequence contains Gln-Ile, Gln-Ile-Gln, Val-Glu and Gln-Val bonds. The values for Val and Ile show the expected increases after 96 h hydrolysis due to the acid stable Ile-Val bond.

Automatic phenylisothiocyanate degradation on 0.6 μmol protein using quadrol buffer [Table 7(b)] was not as successful as the programme on 0.25 μmol protein using DMBA buffer [Table 7(c)]. During the former programme the protein film

was gradually washed out of the spinning cup and there was practically none left after the heme was removed (at step 17). Criteria for satisfactory results from automatic sequencing were not as strict as for P. luridum plastocyanin since amino acid compositions and manual sequences were available from purified peptides. Elucidation of the sequence immediately after the haem group, which is always a difficult region, is discussed in more detail. Tlc was done on 20% of the PTH derivatives and glc on 10 or 20% of the sample. Glc was repeated if necessary on a sample silylated with bis(trimethylsilyl)acetamide which distinguished PTH-Ile and PTH-Leu (double peak with former). PTH derivatives of Asp, Glu and Gln were identified as silylated derivatives. PTH-Asn could not be identified on glc as silylation on or off the column was not successful. Yields of silylated derivatives of Asp, Glu and Gln are variable and silylation is not always successful (Niall, 1973). Conversion of the thiazolinone derivative of Gly to PTH-Gly is usually about 70% complete and derivatives of Gln and Asn are usually deamidated to the extent of about 20-30% under the usual conditions (0.1M HCl, 80°C, 10 min.; Niall, 1973). These values may be lower if the conversion time is reduced - see DMBA programme below. After tlc and glc the remaining sample (usually 40-70%) was hydrolysed with HCl or HI and yields of amino acids normalised to 100% sample and a correction made for partial destruction of some amino acids during hydrolysis (Inglis et al., 1971).

A dark colour was seen in the fraction collector tubes when the 2nd cysteine residue was removed at step 17. This

gives an indication that the haem group is being removed at this step.

After automatic sequencing using the DMBA programme only glc of a few PTH derivatives from residues 1-12 were run. This was in order to check the progress of the Edman degradation and the results were consistent with previous results obtained from degradation using the quadrol programme [Table 7(b)]. Results obtained from the remaining steps of the DMBA programme are shown in Table 7(c). Hydrolyses were carried out with 6M HCl at 150°C for 24 h (Van Orden & Carpenter, 1964). The aqueous phases from steps 14 and 17 (the putative cysteine residues) were oxidised with performic acid to form PTH-cysteic acid which is hydrolysed in good yield to cysteic acid. Cysteine is not recovered from hydrolysis of PTH-cysteine. The amino-acid analyser broke down during analysis on hydrolysed residues 14, 17, 18 and 26. During removal of the haem group at cycle No.17 the protein film detached itself from the spinning cup and the programme was stopped. The film was dissolved in heptafluorobutyric acid (HFBA) and reformed on the cup. The exposure of anhydrous HFBA to the atmosphere caused some subsequent hydrolysis of peptide bonds by the moist HFBA. The subsequent degradation products were contaminated with PTH derivatives of other amino acids notably Ala, Asx, Val and Gly. These background amounts of PTH derivatives were recovered in similar amounts at each cycle and it is apparent which is the sequence residue.

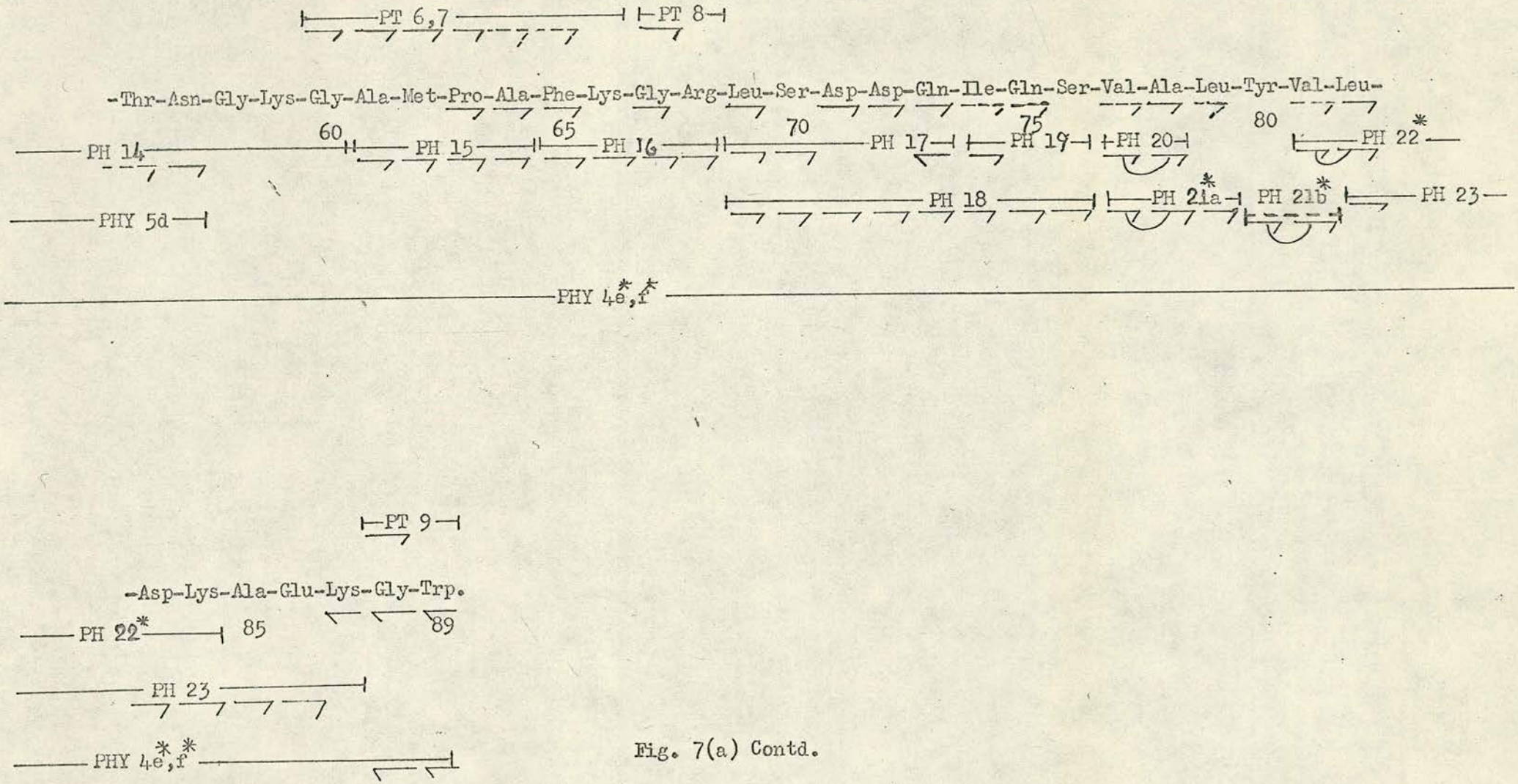

PT 6,7
PT 8
-Thr-Asn-Gly-Lys-Gly-Ala-Met-Pro-Ala-Phe-Lys-Gly-Arg-Leu-Ser-Asp-Asp-Gln-Ile-Gln-Ser-Val-Ala-Leu-Tyr-Val-Leu-
60
65
70
75
80
PH 14
PH 15
PH 16
PH 17
PH 19
PH 20
PH 22*
PH 18
PH 21a*
PH 21b*
PH 23
PHY 5d
PHY 4e*,f*
PT 9
-Asp-Lys-Ala-Glu-Lys-Gly-Trp.
PH 22*
85
89
PH 23
PHY 4e*,f*

Fig. 7(a) Contd.

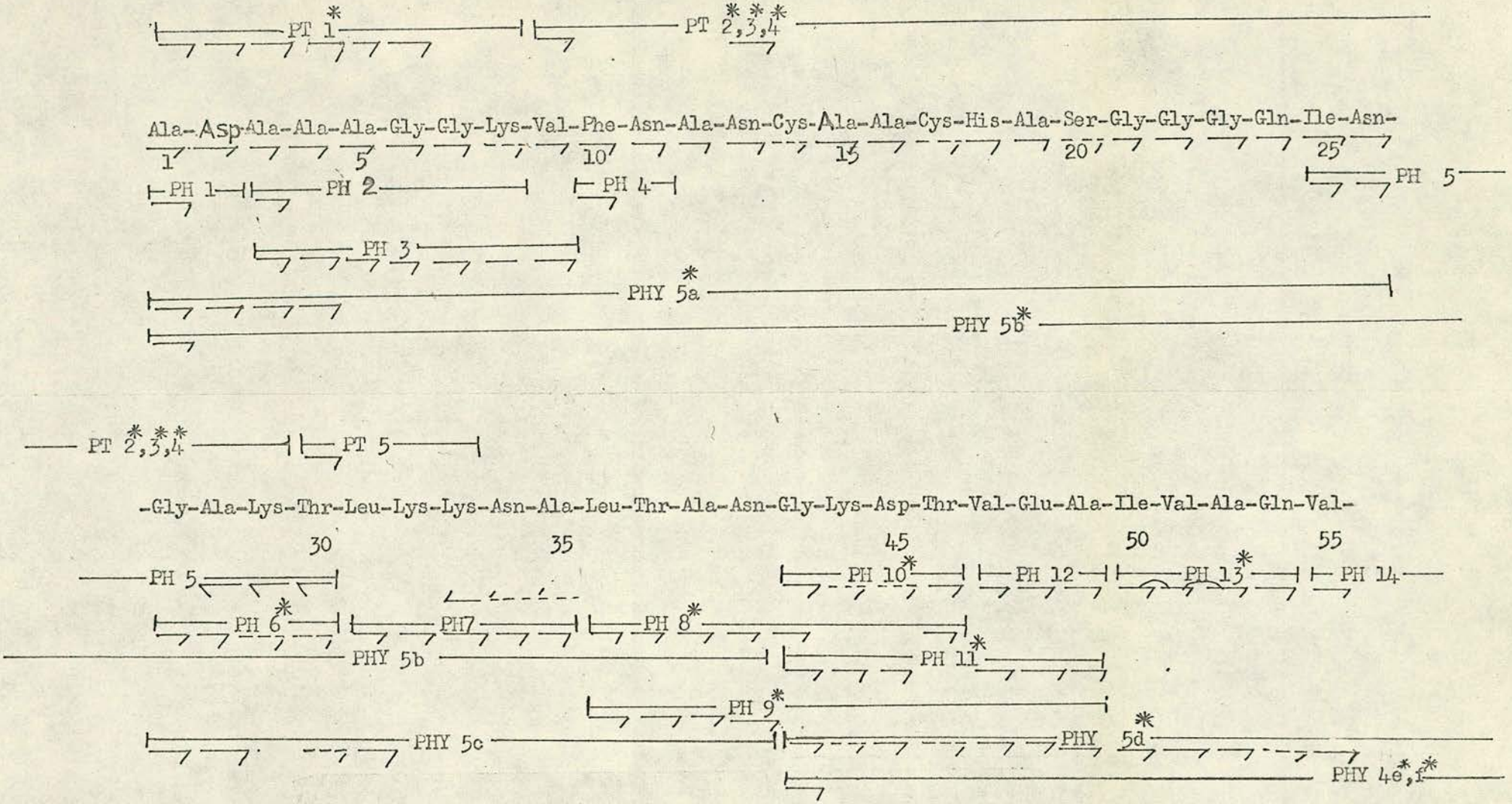


Fig.7(a) Amino acid sequence of P.luridum cytochrome f

[symbols as for fig.5(a), HY=hydroxylamine peptides]

Fractions collected at cycles 19-24 inclusive were converted for 2-3 min instead of the usual 10 min. This was done in an attempt to recover identifiable amounts of PTH-serine which is very unstable. The presence of serine somewhere between residues 19-24 was known from peptide compositions and some PTH-Ser was subsequently identified on tlc at position 20. This short conversion time was responsible for low recoveries of PTH-Gly in cycles 21, 22 and 23. These amounts are much higher than in steps 19, 20 and 24. When the aqueous and remaining ethyl acetate solubles were hydrolysed unconverted thiazolinone derivatives were also converted to free amino acids (Inglis et al., 1971), which explains the subsequent high recoveries of Gly in amino acid analyses.

The DMBA programme was stopped after step 26. Partial blockage was expected anyway. The residual material was subjected to acetylation followed by CNBr cleavage [Chapter 4(1)] and the results of automatic sequencing of the CNBr fragment are shown in Table 7(d). The CNBr cleavage reaction did not go anywhere near to completion and due to the small amount of cleaved material there was only sufficient PTH derivative at each step for one identification method. The PTH derivatives were identified by putting all the material on tlc or glc, whichever was thought to provide a more positive identification for the expected residue. The results provide tentative overlaps in the region where few were previously established, see Fig. 7(a). Although the yields of PTH derivatives are very low (glc sensitivity

was turned up high) the sequence residues at any step are usually more than 50% above any other derivative recorded on that glc trace. It was only the very low level of "background" PTH amino acids throughout this sequencer run which enabled identification of the later residues. This was probably due to virtually complete acetylation of free amino groups before the CNBr cleavage.

The cytochrome (1.0 μmol) was digested with trypsin. Peptides could not be purified from large regions of the sequence [see Fig. 7(a)] presumably due to rearrangement at the Asn-Gly bonds. Some peptides missing from the C-terminal region may have undergone deamidation at other susceptible residues or peptides containing deamidated Asn-Gly bonds may have co-electrophoresed with some of these C-terminal peptides. Both P. luridum and Synechococcus cytochromes f precipitated during removal of the heme group and the apoprotein was not subsequently desalted cleanly to remove urea. Some salt was observed on subsequent peptide maps and some dehaemed protein was lost in the later Sephadex G-25 fractions which contained the bulk of the urea. A large number of acidic tryptic peptides with N-terminal amino acids which included combinations of varying amounts of Asx and Leu were identified but could not be separated. Some of these mixtures were analysed quantitatively after hydrolysis of almost the total sample (a few nmol). Detailed properties and sequence determination experiments on the purified tryptic peptides are given in tables at the end of this chapter.

The protein (1.2 μmol) was also cleaved with hydroxylamine [see Chapter 4(c)] and the larger fragments and uncleaved protein subdigested with thermolysin. Two smaller hydroxylamine fragments from the Sephadex G-50 column were separately gel filtered on Sephadex G-25 (90 cm x 1 cm). After pooling appropriate fractions these peptide mixtures (fractions 4 and 5 from Sephadex G-50) were subjected to electrophoresis at pH 6.5. Many more peptides than expected were prepared. The compositions of those which were considered reasonably pure (tables at end of chapter) and those which were not very pure suggest that some deamidation of Asn-Gly bonds had occurred under the alkaline conditions as well as cleavage. See mechanisms in Fig. 4(a).

Thermolysin digestion on the larger fractions proved very useful in determining the sequence. Some thermolysin peptides overlapped rearranged Asn-Gly bonds. Manual Edman degradations were almost always completely blocked when this bond was reached (except peptide PH8). These thermolysin peptides (and some others) were digested with carboxypeptidase A [Table 7(e)].

Assignments of amide or acidic side groups are explained in Table 7(f).

Peptides which failed to meet the analytical criteria in Chapter 5(c) were the following:-

(1) Most hydroxylamine cleavage peptides. Yields of each peptide were so low that it was not practical to purify them further.

(2) Values for Gly and Ala in peptides PH 6, 8 and 9 are high, but in view of the high levels of these amino acids in

the protein it is not surprising that this is a common occurrence due to low levels of contamination with other peptides.

(3) Peptide PH10 contained Leu (0.26 = 0). N-terminal analysis revealed the presence of some Dns-Leu. The occurrence of free leucine is expected (released from Leu-79).

(4) Peptides PHY5d, PHY4e, PHY4f and PH13 contain an acid-stable Ile-Val bond and the values for these amino acids are low in these peptides, Ile (0.33 = 1) in PH13 is also N-terminal which frequently reduces the recovery of an amino acid in this position. This is also the probable reason for the value for Ala (3.5 = 4) in peptide PT1.

(5) Peptide PH22 was a minor thermolysin cleavage peptide, and although it was impure it gave a good dansyl-Edman degradation result (only sufficient material for 1 cycle). Dns-Val-Leu was also identified on polyamide plates. This provides overlap between residues 81 and 82.

(6) The tryptic peptides PT2, 3 and 4 all mapped from residue 9-29 and their amino acid compositions are consistent with the proposed sequence, although they were not surprisingly recovered in low yield. The recovery of cysteine containing peptides is usually low due to low solubility and partial oxidation before identification and stabilisation by performic acid oxidation. Added to this was the presence of an Asn-Gly bond which caused further mobility changes due to deamidation.

(7) Peptide mixture PH21a and b was not completely separated after 5 purification steps. The mixture is proposed to contain peptides with the sequences Val-Ala-Leu

and Tyr-Val for the following reasons:-

(a) After Sephadex G-25 gel filtration the composition of peptide PH21a corresponded to Val-Ala-Leu except that the peptide had become contaminated with 0.7 residues of glycine which was not there previously (Dns-Gly was identified as present in N-terminal analysis only after paper chromatography and gel filtration).

(b) Peptide mixture PH21a and b contained 20% of peptide PH15 (Ala-Met-Pro-Ala) - the only Met and Pro residues in the protein are in the latter peptide. PH21 and PH15 peptide bands overlapped and were not completely separated by electrophoresis.

(c) The mobility and colour of PH21a and b after electrophoresis at pH 2 were very different from free tyrosine. The electrophoretic mobility at pH 2 was also completely different from peptide PH20 (Val-Ala).

(d) The N-terminal dansyl derivatives were identified as Dns-Val, Dns-Val-Ala, bis-Dns-Tyr and an orange fluorescent spot corresponding to bis-Dns-Tyr-Val (there was no trace of the very acid stable derivative Dns-Val-Val after overnight hydrolysis).

(e) By a subjective visual estimate there were about equimolar amounts of the dansyl derivatives produced after each step of Edman degradation on the peptide mixture before paper chromatography and gel filtration. This fits the analysis data which indicates the composition:- Val(2), Tyr(1), Ala(1), Leu(1).

(f) Some bis-Dns-Tyr was identified on N-terminal analysis after gel filtration of the mixture but amino acid analysis

revealed no trace of tyrosine in PH21a.

There was no suggestion of the existence of any missing peptides or amino acids which belong in the 2 gaps inserted in this sequence compared to the sequence of Spirulina maxima cytochrome f (see discussion section). No peptides were isolated which did not fit the proposed sequence. Overlaps were obtained for all parts of the sequence from peptide sequences and compositions and from the results of automatic sequencing.

Peptide PH18 was subjected to automatic PITC-degradation in an attempt to obtain additional information for amide assignments. This small peptide readily washed out of the spinning cup and the results obtained were poor. However at step 4 of this degradation (Asp-72) the PTH derivative ran much faster on tlc (higher Rf value) than PTH-Asp standards although it had the characteristic pink colour and inverted "V" shape of PTH-Asp. This phenomenon has been noticed previously but not to such a pronounced extent. Exactly the same happened with the PTH derivative from step 10 on automatic sequencing of the CNBr fragment. In both cases standard PTH-Asp and PTH-Asp derived from residue 71 ran with normal Rf values. PTH-Asn runs more slowly on the plates than PTH-Asp and is yellow in colour. The cause of this is not known, but it is unlikely that the condition of the tlc plate and age of the solvent would be responsible for such an extreme effect. It is possible that a labile group may be attached at residue 72 to the carboxyl group. This could account for the fact that a peptide (AT8) was recovered in high yield from

A. variabilis cytochrome f but the expected tryptic peptide from the same region of the P. luridum cytochrome f sequence could not be purified, although a peptide of apparently similar size, electrophoretic mobility at pH 6.5, with N-terminal Leu and containing tyrosine was detected in mixtures. It was initially suspected that there was an Asn-Gly bond in this region because peptides from this region could not be purified. Thermolysin peptides were readily purified which had properties consistent with the proposed sequence in this region [Fig. 7(a)].

The weakest part of the proposed sequence is the evidence for Tyr-80. There is however only 1 tyrosine in the protein and it is almost certainly not at position-53 which is assigned as a deletion in P. luridum cytochrome f. Automatic sequencing on the CNBr fragment was done on such a small amount of material that PTH-Tyr did not show on t l c (O-acetyl-tyrosine is labile). There was no sign of any other residue at this position. Evidence from peptides PH21a and b (although they could not be satisfactorily separated) and the compositions of peptides PHY4e and 4f suggest that the assignment of Tyr to position 80 is correct. Tyrosine is also a conserved residue at position 80 in most of the other cytochromes f.

The material used in the sequence determination represents the results of many months growths of P. luridum, and although there are weaknesses in the proposed sequence, one must consider the additional time, effort, and expense involved in growing many more batches of cells, preparing

more protein and carrying out additional experiments to obtain confirmation of the less well established parts of the sequence. If additional work is considered necessary, treatment with hydroxylamine to stabilise (and partly cleave) Asn-Gly bonds followed by chymotryptic digestion on the whole of the products would yield peptides which would confirm the proposed sequence.

Table 7(a) Amino acid composition of P. luridum cytochrome f

Analysis	24 h	96 h	Sequence
Glycine	11.0	11.0	11
Alanine	17.8	15.9	17
Valine	4.8	5.5	6
Leucine	4.8	4.7	5
Isoleucine	2.5	2.8	3
Serine	2.5	2.2	3
Threonine	3.8	3.4	4
Aspartic acid	12.0	11.0	5
Asparagine	-	-	6
Glutamic acid	5.3	5.2	2
Glutamine	-	-	4
Phenylalanine	2.2	1.9	2
Tyrosine	0.9	-	1
Tryptophan	-	0.9	1
Cysteine	1.6	-	2
Methionine	1.0 (as sulphone)	-	1
Proline	1.0	1.25	1
Lysine	8.6	8.4	9
Histidine	1.0	1.0	1
Arginine	1.7	1.7	1
		Total =	85

Table 7(b) P. luridum cytochrome f automatic sequencing (quadrol programme)

step	tlc	intensity	glc	nmol	hydrolysis/ dansylation	sequence residue
1	Ala	3+	Ala	86		Ala
2	Asp	2+	-			Asp
3	Ala	2+	Ala	58		Ala
4	Ala	+	-			Ala
5	Ala	+			?	Ala
6	Gly	+	Gly	76		Gly
7	Gly	+	Gly	28		Gly
8	Lys	+				(Lys)
9	VILP	+	Val	29		Val
10	VILP	½+	Phe	17		Phe
11	Asn	+			?	Asn
12			Ala			Ala
13	Asn	+			?	Asn
14						(Cys)
15			Ala	17		Ala
16			Ala	14		Ala
17					(haem removed)	(Cys)
18					(Pauly test positive)	His
19					?	
20					Gly	(Gly)

Table 7(c) P. luridum cytochrome f automatic sequencing (DMBA programme)

Step	tlc	intensity	glc results (followed by yield in nmol)	hydrolysis/analysis (followed by yield in nmol)	sequence residue
13	Asn	2+			Asn
14				analyser breakdown	
15					
16					
17				analyser breakdown	
18				analyser breakdown	
19	Ala	2+	Ala,21 (Val,5 and Gly,5)	Ala,16 (Val,7; Gly,30; Asx,16;Glx,13)	Ala
20	Ser (trace Ala and possibly Gly)		(Ala,16; Gly,7.5; Val,4.3)	(Gly,36; Ala,18; Asx,14; Glx,24)	(Ser)
21	Gly (trace Asn or Gln)	2+	Gly, 10 (Ala,10; Gln, 12; Val, 7)	Gly,62 (Ala,15; Val,5; Asx,15; Glx,19)	Gly
22	Gly	2+	Gly, 13 (Gln, 0; Ala, 10; Val, 7)	Gly, 37 (Ala, 9; Val, 5; Asx, 14; Glx, 12)	Gly
23	Gly	2+	Gly,13 (Gln, 17; Glu, 7; Ala, 13; Val, 11)	Gly, 46 (Ala, 21; Val, 6; Asx, 16; Glx, 26)	Gly
24	Gln	+	Gln, 10; Glu, 16 (Ala, 8; Val, 6; Gly, 3)	tube broke during hydrolysis Glx and Gly mostly	Gln
25			Ile, 32 (Gln, 19; Glu, 8; Ala,25; Val,12; Gly,9; Asp,10)	analyser fault - some Ile and allo-Ile (Ala,19; Glx,14; Val,11; Gly,36; Asx,17)	Ile
26	Asn (Gly +)	+		analyser breakdown	Asn

Table 7(d) Automatic sequencing on CNBr fragment from P. luridum cytochrome f

Step	tlc	intensity	glc	nmol	sequence residue
1			Pro	4.2	Pro
2			Ala	10.2	Ala
3			Phe	3.4	Phe
4	trace Gly and Ala	(about ½+)			(PTH-ε-acetyl-Lys will be present)
5	Gly (60% sample)	⅓-½+	traces Gly, Val, Ala, Pro, Leu (0.5-0.8 nmol)		Gly
6	trace Gly	(about ½+)			(Thiazolinone of Arg will remain in aqueous phase)
7			Leu, 1.8 (Ala, 1.3; Val, 0.4)		Leu
8	possible trace Asp				-
9	Asp	½+			Asp
10	Asp (Rf high)	+			Asp
11	Glu (trace Gln)	+			Gln
12			Ile/Leu, 0.9 (Ala, 0.5)		Ile/Leu
13	trace at Glu (too little to reveal Gln)				Glx
14	-				-
15			Val, 0.6 (Ala, 0.4; Leu, 0.5; Pro, 0.4)		(Val)
16			Ala, 0.75 (Val, 0.5; Leu, 0.4; Pro, 0.3)		Ala
17			Leu, 0.3 (Val, 0.2; Pro, 0.1)		(Leu)
18	(nothing identifiable)				-
19			Val, 0.34 (Pro, 0.3; Leu, 0.27)		(Val)
20			Leu, 0.5 (Pro, 0.2; Val, 0.2)		Leu

Table 7(e) Carboxypeptidase A experiments on P. luridum cytochrome f

Results are given in nmol after the amino acid.

(1) Native cytochrome f (50 nmol protein, 50 μg enzyme, 4h)

Trp(11); Gly(20); Lys(6); traces Ser,Thr,Val,Ile ("blank" levels)

Interpretation:-

--- Lys-(Gly,Trp)
(89)

Residue 86 is Glu which is normally removed very slowly by carboxypeptidase A and glycine is released very slowly and lysine slowly (Ambler, 1972) at pH 8.5.

(2) Peptide PHY4f (25 nmol peptide, 10 μg enzyme, 4h)

Trp(11); Gly(11); Ser,Ala,Ile, Leu,Val (3-5 nmol, about twice "blank" level); Lys (1.4, "blank" level).

Interpretation:-

--- Gly-Trp
(89)

The peptide was not completely pure and the much lower concentration of carboxypeptidase A failed to remove lysine.

(3) Peptide PH5 (30 nmol peptide each time, 10 μg enzyme)

time, 2h Thr(11); Lys(6); Ser,Ala,Gly (less than 1)

time, 4h Thr(18); Lys(8); Ala(4); Ser,Gly (about 2)

Interpretation:-

--- Ala-Lys-Thr
(30)

(4) Peptide PH14 (50 nmol each time, 10 μg enzyme)

time, 4h Gly(9); Lys(6); Thr,Ser (3, slightly above "blank" value)

overnight Gly(8); Lys(6); Thr(3); Ser,Ala (1).

Interpretation:

-- (α- and β-Asp-Gly)-Lys-Gly
(57) (60)

Lysine is released slowly and glycine is released very slowly with carboxypeptidase A and the peptide was probably blocked at position 57.

(5) Peptide PH7 (30 nmol peptide each time, 10 µg enzyme)

time, 1h Lys(5); possible dipeptide (broad peak eluting in 3rd buffer at about His position, about 13 nmol); Asp,Gly,Ala (2-3 nmol).

time, 4h Lys(12); possible dipeptide (21 nmol); Asp,Ser,Gly,Ala (2-3 nmol).

Asn-Ala bonds can be affected by hydroxylamine similarly to Asn-Gly bonds (but to a much lesser extent, Bornstein, 1970).

Possibility:- --(Lys)-Lys-(Asn-Ala) ?
(32) (35)

(6) Peptide PH17 (40 nmol peptide, 10 µg enzyme, 2h and 4h)

After incubation the mixtures were electrophoresed at pH 6.5 with appropriate amino acid marker mixtures alongside and the neutral band cut out. It was sewn onto another piece of paper and electrophoresed at pH 2. The paper was stained with ninhydrin/collidine solution.

After 2 h incubation about 8 nmol glutamine released (subjective estimate).

After 4h, about 15 nmol glutamine released (subjective estimate).

No trace of aspartic or glutamic acid was seen after staining the acidic region of the pH 6.5 electrophoresis paper. The rate of release of glutamine was probably much slower than expected due to the presence of aspartic acid as the penultimate residue (Ambler, 1972).

Table 7(f) Assignment of amide side-groups in *P. luridum* cytochrome f

Asp-2 Acidic mobility of peptide PH1 (-0.75); Ala-Asp.
Neutral mobility of peptide PT1; Ala-Asp----Lys.
Automatic sequencing result.

Asn-11, Asn-13, Gln-24, Asn-26: Automatic sequencing results.
Basic mobility of peptides PT2; PT3; PT4 (+0.18) which contain 1 Lys (charge = +1) and 1 His (partial positive charge). The mobility suggests about 1.5 positive charges present.

Asn-11 Neutral mobility of peptide PH4; Phe-Asn.

Asn-26 The Asn-Gly bond is cleaved by hydroxylamine
(26) (27)
Neutral mobility of peptide PH5;
Ile-(α- and β-Asp)-Gly-Ala-Lys-Thr

Asn-34 Basic mobilities of peptides PH7 (+0.48);
Leu-Lys-Lys-Asn-Ala
and PHY5c (+0.49);
Gly-Ala-Lys-Thr-Leu-Lys-Lys-Asn-Ala-Leu-Thr-Ala-
(Asp-CONHOH)

Asn-42 Asn-Gly is cleaved by hydroxylamine
(42) (43)
Mobilities of peptides PHY5c (above), PH8 (-0.31);
Leu-Thr-Ala-(α- and β-Asp)-Gly-Lys-Asp-Thr
(42) (45)
and peptide PH9 (-0.44);
Leu-Thr-Ala-(α- and β-Asp)-Gly-Lys-Asp-Thr-Val-Glu-Ala
(42) (45) (48)

Asp-45 Neutral mobility of peptide PH10; Gly-Lys-Asp-Thr
Acidic mobility of peptide PH11 (-0.31);
Gly-Lys-Asp-Thr-Val-Glu-Ala

Glu-48 Acidic mobility of peptide PH12 (-0.59); Val-Glu-Ala

Gln-54 Neutral mobility of peptide PH13; Ile-Val-Ala-Gln

Asn-57 Hydroxylamine cleaves Asn-Gly
(57) (58)
Neutral mobility of peptide PH14;
Val-Thr-(α- and β-)-Gly-Lys-Gly

Asp-71, Asp-72, Gln-73: Automatic sequencing results (on CNBr fragment) Acidic mobilities of peptides PH17 (-0.75) and PH18 (-0.59) indicate 2 negative charges in this region and carboxypeptidase A digestion removed C-terminal Gln from PH17.
Mobility at pH 6.5 of residual peptide after 3 Edman degradation cycles on PH17:- Asx-Glx = -0.7
and after 4 Edman cycles:- Glx = 0

Gln-75: Neutral mobility of peptide PH19.

The sequence in this part of the protein is therefore:-

Leu-Ser-Asp-Asp-Gln-Ile-Gln-Ser
(71) (72)(73) (75)

Asp-83, Glu-86: Neutral mobilities of peptides PH22 and PH23; Val-Leu-Asp-Lys and Leu-Asp-Lys-Ala-Glu-Lys

Properties of tryptic peptides from P.luridum cytochrome f.

Peptide	m 6.5	V/Vo G-25	Purn	Yield %	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
PT1	0	1·7	G62	6·0	2·19	3·5						1·2								1·11			8	Ala	1 → 8
PT2	+0·18	1·6	G6ox62	1·1	4·35	4·82	1·0		0·75	0·92		3·47	0·62	0·55			1·52			fault 1	0·97		21	Val	9 → 29
PT3	+0·18	1·6	G6ox6	1·8	4·28	4·82	0·72		0·88	fault 1		3·13	0·7	0·72			fault 1·3			0·94	1·04		21	Val	9 → 29
PT4	+0·18	1·6	G6ox62	1·4	4·04	5·52	0·88		0·70	0·98		3·89	0·89	0·63			1·78			0·92	0·86		21	Val	9 → 29
PT5	+0·83	1·8	G63	4·0				0·84			0·91									2·2			4	Thr	30 → 33
PT6	+0·3	1·9	G63	3·0	1·03	2·16								0·89				1·13	0·85	0·95			7	Gly	60 → 66
PT7	+0·4	1·9	G63	2·4	0·81	2·17								0·92				0·7	1·06	1·1			7	Gly	60 → 66
PT8	+0·8	2·5	G632	8·0	0·93																	1·07	2	Gly	67 → 68
PT9	0	4·7	G62	4·4	1·1											0·9							2	Gly	88 → 89

Properties of hydroxylamine cleavage peptides from P.luridum cytochrome f.

Peptide	m 6.5	V/Vo G50	Purn	Yield n mol	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
PHY5a	+0.21	2.5	G6ox6	51	5.61	9.15	1.03	0.53	0.94	0.42	0.6	4.91	0.42	0.97			1.52			1.11	0.82		26	Ala	1→26
PHY5b	-0.02	2.5	G6	36	6.68	11.78	1.6	0.88	1.03	0.81	0.9	6.55	1.08	1.3			2nd ox.			2.48	0.62		39	Ala	1→39
PHY5c	+0.49	2.5	G6	25	1.08	3.84		1.87			2.25	2.06								3.08			13	Gly	27→42
PHY5d	-0.17	2.5	G6	69	2.43	3.91	2.38	0.41	0.65	0.66	1.74	3.19	2.15	0.41	0.21			0	0	0.91			14	Gly	43→57
PHY4e	-0.04	2.3	G6	42	5.92	8.47	3.35	3.34	1.35	2.14	2.66	6.29	3.76	1.13	0.78	-		1.25	0.52	4.27		0.86	46	Gly	43→89
PHY4f	-0.07	2.3	G6	34	5.74	7.94	2.78	2.88	1.28	1.51	1.38	6.36	2.63	1.11	0.58	-		0.93	0.81	3.95	0.36	0.79	46	Gly (trace Ala)	43→89

Sequence determination experiments on tryptic and hydroxylamine peptides from P.luridum cytochrome f

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	Other expts.
PHY 5a	(26)	Ala	Asx	Ala	Ala							
PHY 5c	13	Gly	Ala	lost	Thr	Leu						-10 -11
PHY 5d	14	Gly	Lys	Asx	Thr	Val	Glx	Ala	Ile I-V	Val	Ala	Glx Val
PT1	8	Ala	Asx	Ala	Ala	Ala	Gly					
PT3+PT4	21	Val	N.D.	N.D.	N.D.	Asx						
PT6	7	Gly	Ala	Met	Pro	Ala	Phe					

Properties of thermolysin peptides from P. luridum cytochrome f. Sheet I.

Peptide	m 6.5	V/Vo G-25	Purn	Yield n mol	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
PH1	-0·75		63	189		0·95						1·05											2	Ala	1→2
PH2	+0·43		63	22	2·19	2·8														1·01			6	Ala	3→8
PH3	+0·43		63	61	2·22	2·93	0·98													0·86			7	Ala	3→9
PH4	0		63B2	37								1·1		0·9									2	Phe	10→11
PH5	0		632	123	1·11	1·06			0·84		1·12	1·02								0·84			6	Ile I-Glx	25→30
PH6	+0·6		6	69	×1·23	1·03					×0·7									1·03			4	Gly	27→30
PH7	+0·48		63	212		0·97		0·96				1·17								1·91			5	Leu	31→35
PH8	-0·31		63	39	×1·28	×1·31		×0·72			1·71	2·15								0·82			8	Leu	36→46
PH9	-0·44		63	56	×1·54	2·24	×0·63	0·86			1·78	1·92	1·18							0·91			11	Leu	36→49
PH10	0		632	78	0·88			×0·26			0·88	1·19								1·06			4	Gly (trace Leu)	43→46
PH11	-0·31		63	31	0·97	1·01	×0·68				0·97	1·1	1·04							0·91			7	Gly	43→49
PH12	-0·59		63	124		1·11	0·87						1·02										3	Val	47→49
PH13	0		632B	53		1·02	×0·48		×0·33				0·98										4	Ile I-V	50→5
PH14	0		632	208	1·98		0·9				1·11	1·16								0·88			6	Val	55→6
PH15	0		632	101		1·99												0·97	1·05				4	Ala	61→6

Properties of thermolysin peptides from P.luridum cytochrome f. Sheet 2.

Peptide	m 6.5	V/Vo G-25	Purn	Yield n mol	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
PH16	+0·86		6	410	1·14									1·01						0·85		0·99	4	Phe	65 → 68
PH17	-0·75		63	181				0·85		0·89		2·11	1·15										5	Leu	69 → 73
PH18	-0·59		63	137				0·91	0·99	1·77		2·19	2·15										8	Leu	69 → 76
PH19	0		632B	60					0·94	0·91			1·14										3	Ile	74 → 76
PH20	0		632	181		1·07	0·94																2	Val V-A	77 → 78
PH21a and b	0		632	70	0·2 ×	1·33 ×	1·72	0·85							0·63 ×			0·24 ×	0·19 ×				3 and 2	Val V-A Tyr	77 → 79 80 → 81
PH21a	0		632BG	18	(0·7)	1·11	1·15	0·73 ×							0								3	Val V-A	77 → 79
PH22	0		632B	20		0·3 ×	0·98	0·77		0·5 ×		1·18	0·4 ×							1·09			4	Val V-L	81 → 84
PH23	0		632	48		0·89		1·14				0·97	1·13							1·89			6	Leu	82 → 87

Sequence determination experiments on thermolysin peptides from P.luridum cytochrome f. Sheet I

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	Other expts.
PH3	7	Ala	Ala	Ala	Gly	Gly	?	Val				
PH5	6	Ile	Asx	(blocked)								
PH6	4	Gly	Ala	Lys	Thr							
PH7	5	Leu	Lys	Lys	Asx	Ala						
PH8	8	Leu	Thr	Ala	Asx	Gly	?	Lost	Thr			
PH9	11	Leu	Thr	Ala	Asx	(blocked)						
PH10	4	Gly	Lys	Asp	Thr							
PH11	7	Gly	Lys	Asp	?	Val	Glx	Ala				
PH12	3	Val	Glx	Ala								
PH13	4	Ile I-V	Val V-A	Ala	Glx							
PH14	6	Val	Thr	Asx	(blocked)							
PH15	4	Ala	Met	Pro	Ala							
PH16	4	Phe	Lys	Gly	Arg							
PH17	5	Leu	Ser									
PH18	8	Leu	Ser	Asx	Asx	Glx	Ile	Glx	Ser			

Sequence determination experiments on thermolysin peptides from P. luridum cytochrome f. Sheet 2.

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	Other expts.
PH21a and b	3	Val +V-A	Ala	Leu	blank							
PH21a and b	2	Tyr +Y V	Val	—	blank							
PH21a (after Sephadex G-25)	3	Val (Gly+Tyr)	Ala (a. Val)	Leu								
PH22	4	Val +V-L	Leu									
PH23	6	Leu	Asx	Lys	Ala	Glx						

CHAPTER 8

Partial amino acid sequence of A. variabilis cytochrome f

The partial sequence is shown in Fig. 8(a), with tryptic peptides aligned by homology with sequences of other cytochromes f (see discussion section). Tryptic digestion (on 0.8 μmol dehaemed protein) was much more successful with this cytochrome than with either P. luridum or Synechococcus cytochromes f. Peptides were isolated and sequenced (see Tables at end of chapter) from all parts of the proposed sequence including the N-terminus (peptide AT1 contains an Asn-Gly bond) except where Asn-Gly bonds are known to occur in other parts of cytochromes f. Automatic sequencing on A. variabilis cytochrome f was however plagued with problems. The results (on 0.5 μmol native protein) using a quadrol programme are shown in Table 8(a). During automatic PITC degradation using DMBA buffer the eluant tubing to the fraction collector and waste bottle became blocked after a few cycles and the experiment had to be abandoned. Later, automatic sequencing on 0.2 μmol dehaemed protein was attempted but was not successful due to a large number of spurious amino acid derivatives appearing in the first few cycles.

The amino acid composition is given in Table 8(b) on the basis of about 85 residues. An additional hydrolysis for 24 h was done on a sample of dehaemed protein.

The mobility of peptide AT8 (-0.35) at pH 6.5 and the mobility after 6 cycles of Edman degradation (-0.38)

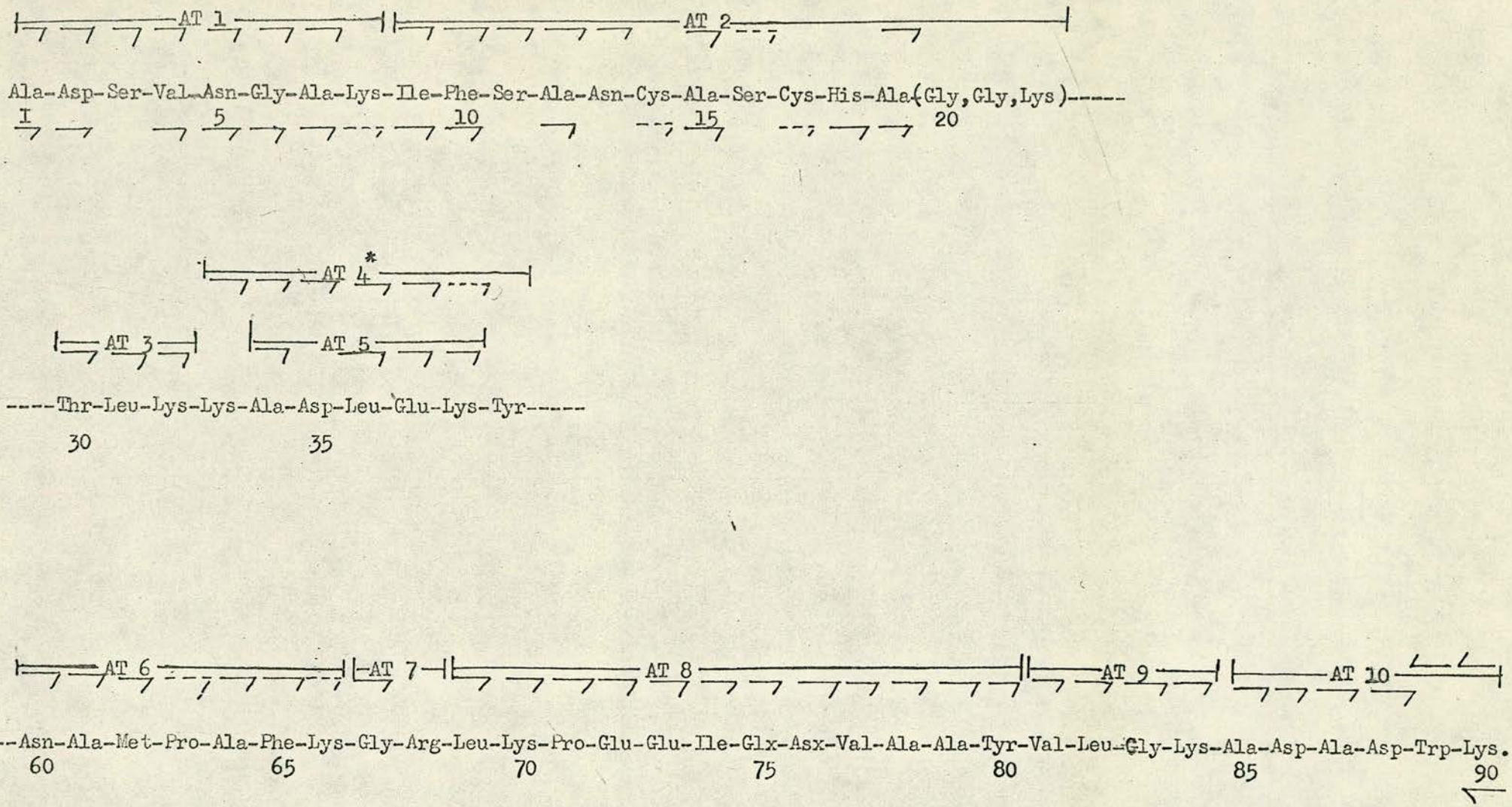


Fig.8(a) Partial amino acid sequence of *A. variabilis* cytochrome f
(symbols as in fig.5(a))

indicates the presence of an amide side group at position 75 or 76.

Carboxypeptidase A digestion on 50 nmol native cytochrome f released only a trace of lysine, and similar digestion of peptide AT10 (10 μg enzyme, 100 nmol peptide, 4h) released 42 nmol lysine and 35 nmol tryptophan. The total digest was analysed on the short column of the Beckmann analyser. Carboxypeptidase A releases lysine slowly and tryptophan quickly (Ambler, 1972) thus the slow release of lysine was followed by rapid release of tryptophan. The amount of aspartic acid released (not measured) would be very little. Taking into account the fact that AT10 was generated by tryptic digestion the sequence of this peptide and the proposed C-terminal sequence of the protein is:-

----Ala-Asp-Ala-Asp-Trp-Lys
(85) (90)

Table 8(a) A. variabilis cytochrome f automatic sequencing (quadrol programme)

Step	tlc	nmol	glc	nmol	hydrolysis/ dansylation	sequence residue
1	Ala	2+	Ala	15		Ala
2	Asp	2+	(Asp, possibly)		?	Asp
3	?	+	-		?	?
4	VILP	2+	Val	34	(Val)	Val
5	Asn	+				Asn
6	Gly	+	Gly	41		Gly
7	Ala	+	Ala	40		Ala
8	Lys	trace	-			(Lys)
9	VILP	+	Ile	30		Ile
10	VILP	½+	Phe	23		Phe
11	-		-			?
12	Ala	½+	Ala	25		Ala
13					?	?
14						(Cys)
15			Ala	16		Ala
16					?	?
17					(haem removed)	(Cys)
18					(Pauly stain positive)	His
19			Ala	7	?	(Ala)

Table 8(b) Amino acid composition of A. variabilis cytochrome f

	24 h (dehaemed)	24 h (native)	96 h
Glycine	7.0	6.9	6.7
Alanine	15.4	15.1	15.4
Valine	4.5	4.6	4.6
Leucine	4.8	5.1	4.6
Isoleucine	2.5	2.5	3.1
Serine	3.4	3.6	2.9
Threonine	2.1	2.1	2.1
Aspartic acid	9.5	9.3	-
Glutamic acid	8.2	7.5	8.6
Phenylalanine	1.8	1.7	2.0
Tyrosine	2.7	2.7	2.8
Tryptophan	-	-	1.0
Cysteine	-	2.0	-
Methionine	1.9(2.4 as sulphone)	1.7	-
Proline	1.7	2.0	1.7
Lysine	10.6	-	-
Histidine	1.0	-	0.9
Arginine	1.1	-	1.0

Properties of tryptic peptides from A. variabilis cytochrome f.

Peptide	m 6.5	V/Vo G-25	Purn	Yield %	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
AT1	0	1·64	G632	20	1·02	2·15	1·08			0·97		1·87								0·92			8	Ala	1→8
AT2		1·69	G60x	6	2·27	2·94			0·77	2·04		1·08		0·77			1·99			0·97	1·19		14	Ile	9→22
AT3	+0·59	1·81	G63	15				1·13			0·91									0·96			3	Thr	30→32
AT4	0	1·64	G632	6·4		0·87		1·11				0·95	1·07		0·57×					1·51×			7	Lys	33→39
AT5	-0·35	1·64	G6	6		0·95		1·14				1·02	1·09							0·8			5	Ala	34→38
AT6	+0·36	1·7	G60x3	12		2·06						1·05		0·81				0·88	0·99	1·16			7	Asx	60→66
AT7	+0·78	2·0	G6	40	0·83																	1·16	2	Gly	67→68
AT8	-0·35	1·64	G63	16·6		1·95	0·99	0·95	1·0			1·12	3·01		0·92				1·07	1·04			12	Leu	69→80
AT9	+0·59	1·81		11·8	1·13		0·9	1·04												0·93			4	Val	81→84
AT10	-0·3	2·05	G63	32		1·97						2·03				0·84				0·98			6	Ala	85→90

Sequence determination experiments on tryptic peptides from A.variabilis cytochrome f.

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	Other expts.
AT 1	8	Ala	Asx	Ser	Val	Asx	Gly	Ala	Lys			-10
AT 2	14	Ile	Phe	Ser	Ala	Asx	?	Ala	Ser	N.D.	N.D	Ala
AT 3	3	Thr	Leu	Lys								
AT 4	8	Lys	Ala	Asx	Leu	Glx	Lys/Tyr					
AT 5	5	Ala	lost	Leu	Glx	Lys						
AT 6	7	Asx	Ala	Met	Pro	Ala	Phe	Lys				-10 -11
AT 8	12	Leu	Lys	Pro	Glx	Glx	Ile	Glx	Asx	Val	Ala	Ala Tyr
AT 9	4	Val	Leu	Gly	Lys							
A10	6	Ala	Asx	Ala	Asx							

CHAPTER 9

Partial amino acid sequence of Synechococcus 6312 cytochrome f

The partial sequence is shown in Fig. 10(g). Problems similar to those encountered with tryptic digestion of P. luridum cytochrome f were also encountered with this protein. The properties of the tryptic peptides from a digest of 0.8 μmol protein which were purified are shown in the tables at the end of this chapter, along with results of sequence determination experiments. The results of automatic PITC degradation on the native protein (0.6 μmol) using quadrol buffer are given in Table 9(a), and results of automatic sequencing using 0.6 μmol native protein with DMBA buffer are given in Table 9(b). The amino acid composition of the protein (on the basis of about 90 residues) is given in Table 9(c).

Table 9(a) Synechococcus 6312 cytochrome f automatic sequencing (quadrol programme)

Step	tlc	intensity	glc	n mol	hydrolysis/ dansylation	sequence residue
1	Ala	2+				Ala
2	Asp	2+				Asp
3	VILP	2+	Ile	24		Ile
4	Ala	2+				Ala
5	Asp	2+				Asp
6	Gly	+	Gly	36		Gly
7	Ala	+				Ala
8	2 spots at about Lys position ½+					(Lys)
9	VILP	½+	Val	17		Val
10	VILP	½+	Phe	14		Phe
11	?	¼+				-
12	Ala	¼+	Ala	5		Ala
13	-		-			-
14						
15			(Ala)	3 (well above background)		(Ala)
16			(Ala)	2.5 (well above background)		(Ala)
17					(haem collected)	(Cys)
18					(Pauly test positive)	His
19						-
20					Gly	(Gly)

Table 9(b) Synechococcus 6312 cytochrome f automatic sequencing (DMBA programme)

Step	tlc	intensity	glc	nmol	hydrolysis (with HCL- except No.30)	nmol	sequence residue
1	Ala	4+					Ala
2	Asp	4+					Asp
3	VILP	4+	Ile	240			Ile
4	Ala	3+					Ala
5	Asp	3+					Asp
6	Gly	2+	Gly	118			Gly
7	Ala	2+					Ala
8	Lys	3+					Lys
no.11	Ser		(traces Ala,Val,Lys)			(No.11)	Ser
12	Ala	+					Ala
13	Asn	+	(some Ala)				Asn
14							(Cys)
15	Ala	2+	Ala	75			Ala
16	Ala	2+	Ala	91			Ala
17					(haem removed)		(Cys)
18					(Pauly stain positive)		His
19			Met	50	(destroyed)		Met
20	Gly	+	Gly	50 (20 Met)			Gly
21	Gly	2+	Gly	55			Gly
22	Gly	2+	Gly	59			Gly

(Contd.

Table 9(b) (contd.)

Step	tlc	intensity	glc	nmol	hydrolysis (with HCL- except No.30)	nmol	sequence residue
23	Asn,2+;Gly+		(blank)		fault on analyser		(Asn)
24	VILP	2+	Val(no Gly)	24	Gly,14 (background amount of Val)		(Val)
25	VILP	2+	(Val, 19)		(Val and Gly; background amounts)		?
26	Met	2+	Met	9 (17 Val)	(destroyed)		Met
27	Ala	+	Ala	13 (8 Val)	fault - Gly or Ala		Ala
28	Asn	+			Asx	10	Asn
29	Lys	+					(Lys)
30	?		traces Pro or Thr and Ala, Val, Leu, Gly		αABA	4 (HI hydrolysis)	Thr
31	?		Leu	15	?		(Leu)

Table 9(c) Amino acid composition of Synechococcus 6312 cytochrome f

	24 h	96 h
Glycine	9.2	9.9
Alanine	15.9	16.0
Valine	4.7	5.8
Leucine	4.2	4.1
Isoleucine	2.9	3.0
Serine	5.0	4.3
Threonine	1.7	1.4
Aspartic acid	12.3	12.9
Glutamic acid	9.5	10.2
Phenylalanine	3.2	3.1
Tyrosine	1.5	1.5
Tryptophan	-	0.5
Cysteine	1.4	-
Methionine	4.2 (as sulphone)	-
Proline	1.2	1.3
Lysine	7.0	6.3
Histidine	1.2	1.0
Arginine	1.4	0.9

Properties of tryptic peptides from Synechococcus 6312 cytochrome f.

Peptide	m 6.5	V/Vo G-25	Purn	Yield %	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	Cys	Met	Pro	Lys	His	Arg	TOT	N-t	Position
6T1	-0.33	1.8	G63	6.5	0.96	2.85			0.99			2.19											8	Ala	1→8
6T2	+0.16	1.5	G60x2	2	3.50	3.20	1.18		1.96	1.20		2.91	0.67	0.53			1.29	1.61		fault	1.05		21	Val	9→29
6T3	+0.16	1.5	G60x2	5.4	2.16	2.03						0.94		0.90				0.95	1.10			0.95	9	Asx	61→69
6T4	-	1.8	G63	2.4	1.77	2.13			0.81	0.5		1.94	1.08							2.26			12	Ser	?

Sequence determination experiments on tryptic peptides from Synechococcus 6312 cytochrome f.

Peptide	TOT	N-t	-1	-2	-3	-4	-5	-6	-7	-8	-9	Other expts.
6T1	8	Ala	Asx	Ile I-A	Ala	Asx	Gly	Ala				
6T2	21	Val	N.D.	Ser								
6T3	9	Asx	Ala	Met	?	Gly						

CHAPTER 10

DISCUSSION

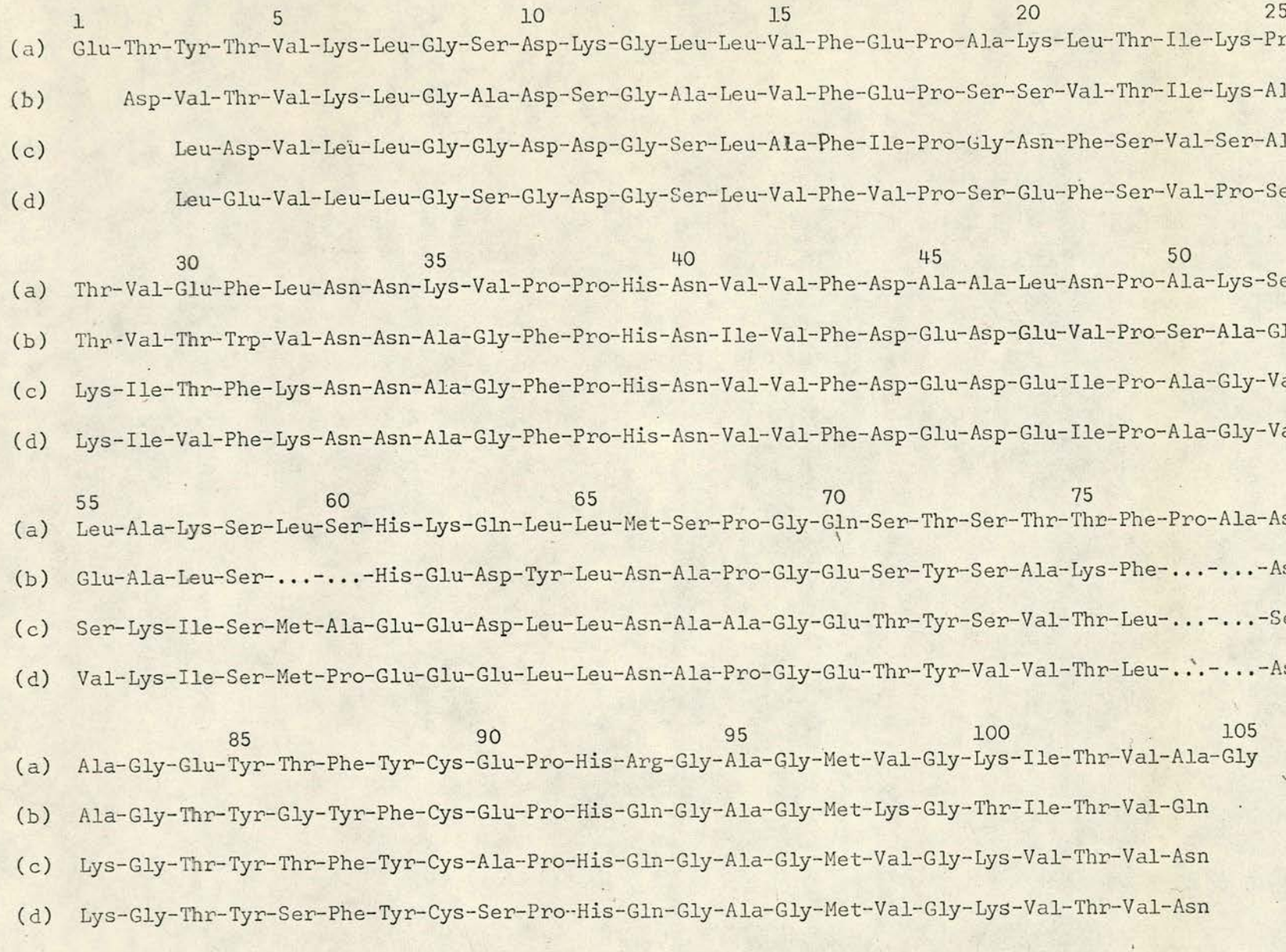


Fig.10(d) The alignment of (a) Anabaena variabilis plastocyanin with (b) Chlorella fusca plastocyanin (Kelly & Ambler, 1974), (c) potato plastocyanin (Ramshaw et al, 1974) & (d) French bean plastocyanin (Milne et al, 1974).

(a) Sequence similarity between prokaryotic and eukaryotic plastocyanins

The alignment of Anabaena variabilis plastocyanin with eukaryotic plastocyanins is shown in Fig. 10(a). The best alignment of the prokaryotic plastocyanin is obtained by placing the protein sequence two residues longer at the N-terminus than plant plastocyanins (and one longer than Chlorella plastocyanin). The Anabaena plastocyanin is also one residue longer at the C-terminus and has insertions at positions 77, 78 and 81.

An alternative alignment with a gap in the Chlorella sequence at residues 50 and 51 instead of 59 and 60 would generate 3 more identities between Chlorella and Anabaena plastocyanins. This would result in 2 [or 3 in the case of elder Sambucus nigra L plastocyanin (Scawen et al., 1974)] fewer identities between Chlorella plastocyanin and the plant plastocyanins.

The highly acidic region from residues 44 to 47 in the eukaryotic plastocyanins is not present (except ASP-44) in Anabaena variabilis plastocyanin. A similarity matrix of identities between Anabaena plastocyanin, Chlorella plastocyanin and the published plant plastocyanins is given, Fig. 10(b).

Kelly and Ambler(1974) found some evidence for dimorphism at residue 84 in Chlorella fusca plastocyanin. The sequence amino acid is given as threonine (the incumbent residue in most plant sequences published) but a small proportion of glutamic acid may have been present. Residue 84 is glutamate in Anabaena plastocyanin.

Fig 10(c) AMINO ACID SEQUENCES OF PLASTOCYANINS

(A) *Phormidium luridum* 1462/2 (= *Schizothrix calcicola*, = *Plectonema boryanum*)
(B) *Anabaena variabilis* (Kützing)
(C) *Chlorella fusca* (vacuolata) (Kelly & Ambler, 1974)
(D) Potato (*Solanum tuberosum*) (Ramshaw *et al.*, 1974)
(E) French bean (*Phaseolus vulgaris*) (Milne *et al.*, 1974)

```
      1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
(A)  Asp-Thr-Val-Lys-Val-Ile-Met-Gly-Gly-Ser-Lys-Gly-----Leu-Val-Phe-Glu-Pro-Ala-Val-Val-Asn-Val-Lys-Ala-Gly-Asp-Thr-Ile-Gln-
(B)  Glu-Thr-Tyr-Thr-Val-Lys-Leu-Gly-Ser-Asp-Lys-Gly-Leu-Leu-Val-Phe-Glu-Pro-Ala-Lys-Leu-Thr-Ile-Lys-Pro-Gly-Asp-Thr-Val-Glu-
(C)      Asp-Val-Thr-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu-Val-Phe-Glu-Pro-Ser-Ser-Val-Thr-Ile-Lys-Ala-Gly-Glu-Thr-Val-Thr-
(D)          Leu-Asp-Val-Leu-Leu-Gly-Gly-Asp-Asp-Gly-Ser-Leu-Ala-Phe-Ile-Pro-Gly-Asn-Phe-Ser-Val-Ser-Ala-Gly-Glu-Lys-Ile-Thr-
(E)          Leu-Glu-Val-Leu-Leu-Gly-Ser-Gly-Asp-Gly-Ser-Leu-Val-Phe-Val-Pro-Ser-Glu-Phe-Ser-Val-Pro-Ser-Gly-Glu-Lys-Ile-Val-

      31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
(A)  -Phe-Glu-Val-Gly-Gln-Leu-Pro-Pro-His-
(B)  -Phe-Leu-Asn-Asn-Lys-Val-Pro-Pro-His-Asn-Val-Val-Phe-Asp-Ala-Ala-Leu-Asn-Pro-Ala-Lys-Ser-Ala-Asp-Leu-Ala-Lys-Ser-Leu-Ser-
(C)  -Trp-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Ile-Val-Phe-Asp-Glu-Asp-Glu-Val-Pro-Ser-Ala-Gly-Asn-Ala-Glu-Ala-Leu-Ser---------
(D)  -Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Val-Val-Phe-Asp-Glu-Asp-Glu-Ile-Pro-Ala-Gly-Val-Asp-Ala-Ser-Lys-Ile-Ser-Met-Ala-
(E)  -Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Val-Val-Phe-Asp-Glu-Asp-Glu-Ile-Pro-Ala-Gly-Val-Asp-Ala-Val-Lys-Ile-Ser-Met-Pro-

      61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
(A)
(B)  -His-Lys-Gln-Leu-Leu-Met-Ser-Pro-Gly-Gln-Ser-Thr-Ser-Thr-Thr-Phe-Pro-Ala-Asp-Ala-Pro-Ala-Gly-Glu-Tyr-Thr-Phe-Tyr-Cys-Glu-
(C)  -His-Glu-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Ser-Ala-Lys-Phe---------Asp-Thr-----Ala-Gly-Thr-Tyr-Gly-Tyr-Phe-Cys-Glu-
(D)  -Glu-Glu-Asp-Leu-Leu-Asn-Ala-Ala-Gly-Glu-Thr-Tyr-Ser-Val-Thr-Leu---------Ser-Glu-----Lys-Gly-Thr-Tyr-Thr-Phe-Tyr-Cys-Ala-
(E)  -Glu-Glu-Glu-Leu-Leu-Asn-Ala-Pro-Gly-Glu-Thr-Tyr-Val-Val-Thr-Leu---------Asp-Thr-----Lys-Gly-Thr-Tyr-Ser-Phe-Tyr-Cys-Ser-

      91  92  93  94  95  96  97  98  99 100 101 102 103 104 105
(A)
(B)  -Pro-His-Arg-Gly-Ala-Gly-Met-Val-Gly-Lys-Ile-Thr-Val-Ala-Gly
(C)  -Pro-His-Gln-Gly-Ala-Gly-Met-Lys-Gly-Thr-Ile-Thr-Val-Gln
(D)  -Pro-His-Gln-Gly-Ala-Gly-Met-Val-Gly-Lys-Val-Thr-Val-Asn
(E)  -Pro-His-Gln-Gly-Ala-Gly-Met-Val-Gly-Lys-Val-Thr-Val-Asn
```

Residues shown in *italics* are tentative

	Anabaena	Chlorella	Higher plants
Anabaena	-	52 or 55	41-45
Chlorella	52 or 55	-	52-58
Higher plants	41-45	52-58	-

Fig. 10(b) Matrix of amino acid identities between plastocyanins

The figures for identities with plant plastocyanins are averages for all the published plant plastocyanin sequences. The number of identities between Anabaena and Chlorella plastocyanin sequences depends on the alignment - see text.

(b) Sequence variation between blue-green algal plastocyanins

The amino-terminal sequence of (A) Phormidium luridum plastocyanin is shown in Fig. 10(c) aligned with the complete amino acid sequences of plastocyanin from (B) A. variabilis, (C) Chlorella fusca and two examples of higher plant plastocyanins (D) Potato and (E) French bean.

The partial sequence of plastocyanin from P. luridum is seen to have considerable similarity to the same regions in the other plastocyanins, if a deletion at position 13 is considered. Gly-34 (one of the tentative assignments), Val-33 and Met-7 are the only exceptions to what have previously been invariant residues in 41 partial plastocyanin sequences (D. Peacock, personal communication). It is seen from the similarity matrix, Fig. 10(d), that C. fusca plastocyanin is slightly

(A) Tyr-Thr-Phe-Tyr-Cys-Glu-----Pro-----His-Arg-Gly-Ala-Gly-Met-Val-Gly-Lys-Ile-Thr-Val-Ala-Gly

(B) Tyr-Met-Phe-Phe-Cys-Thr-Phe-Pro-Gly-His-----Ser-Ala-Leu-Met-Lys-Gly-Thr-Leu-Thr-Leu-Lys

Fig. 10(e) (A) A. variabilis plastocyanin

(B) P. aeruginosa azurin

closer than the blue-green algal plastocyanins to the higher plant plastocyanins and there are large differences in N-terminal sequence between the two prokaryotic plastocyanins. A. variabilis plastocyanin is closer to C. fusca plastocyanin than to plastocyanin from Phormidium luridum. These findings are valid even if the changes in the previously invariant residues are treated with caution.

The implications of these findings are discussed in Section (g) below.

Fig. 10(d) Similarity Matrix for Plastocyanins

	A	B	C	D	E
A	-	49	41	36	33
B	49	-	56	36	38
C	41	56	-	49	46
D	36	36	49	-	69
E	33	38	46	69	-

A, P. luridum; B, A. variabilis; C, Chlorella fusca; D, Potato; E, French bean.

The values are expressed as a percentage identity for residues 1-39. The gap in the Phormidium sequence has been assigned as a 21st amino acid.

(c) Sequence similarity to bacterial azurins

The sequence similarity round the single cysteine residue in eukaryotic plastocyanin to the sequence round the free cysteine in bacterial azurins (Kelly & Ambler, 1974) is also "conserved" in A. variabilis plastocyanin

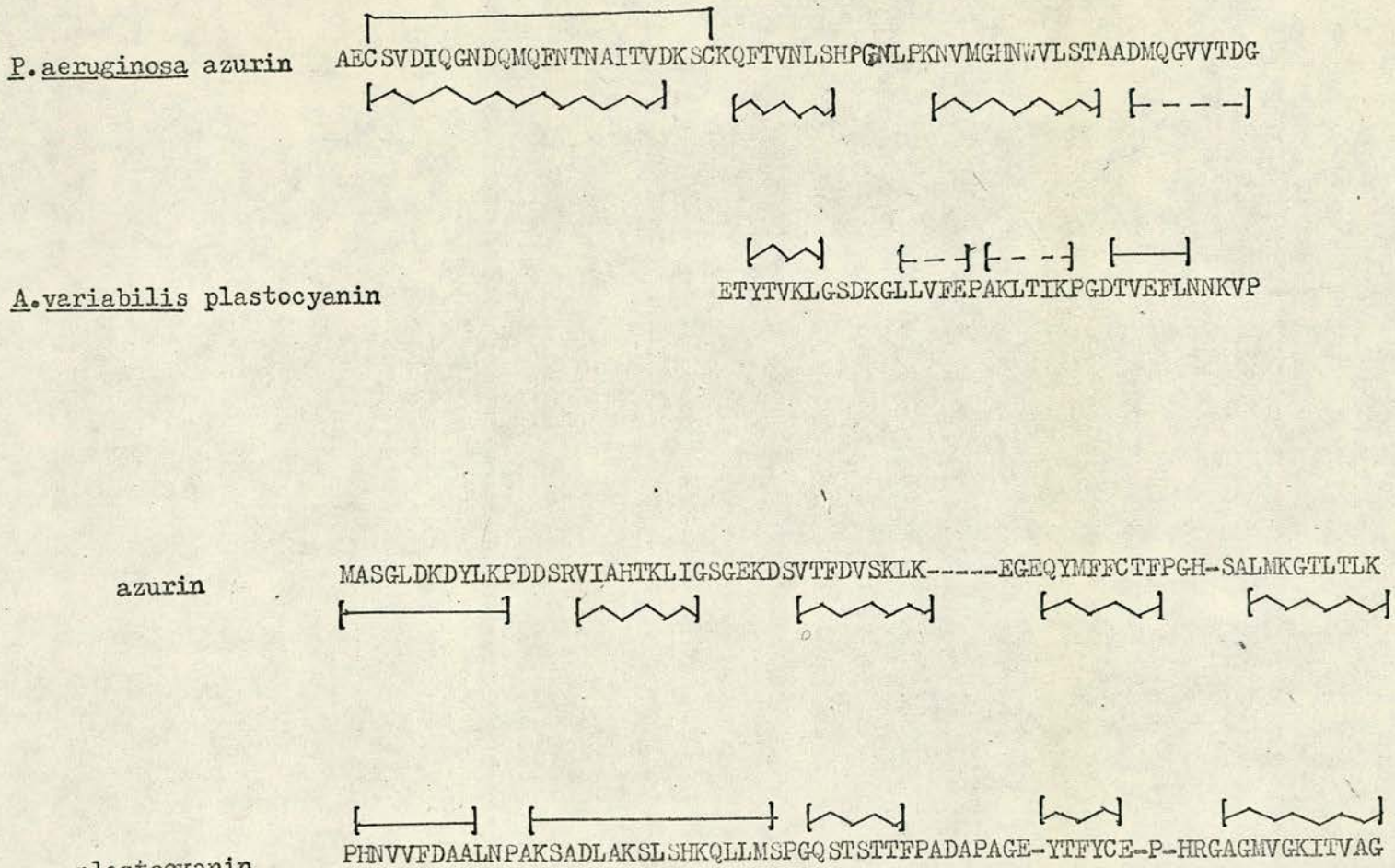

P. aeruginosa azurin
AECSVDIQGNDQMQFNTNAITVDKSCKQFTVNLSHPGNLPKNVMGHNWVLSTAADMQGVVTDG
A. variabilis plastocyanin
ETYTVKLGSDKGLLVFEPAKLTIKPGDTVEFLNNKVP
azurin
MASGLDKDYLKPDDSRVIAHTKLIGSGEKDSVTFDVSKLK-----EGEQYMFFCTFPGH-SALMKGTLTLK
plastocyanin
PHNVVFDAALNPAKSADLAKSLSHKQLLMSPGQSTSTTFPADAPAGE-YTFYCE-P-HRGAGMVGKITVAG
α —
β —
uncertain secondary
structure —

Fig. 10(f)

[see Fig. 10(e)]. The degree of similarity is not however any greater than with the eukaryotic plastocyanins [see also Ramshaw et al. (1974) for alignment of plant plastocyanins with azurin]. The weight of evidence supports the view that these sequence similarities are due to similarity in function. Bacterial azurins also contain copper in a type I ligand environment and Katoh et al. (1961) have suggested that the thiol group is one of the copper ligands.

The primary structures of the two classes of protein give little indication of divergence from a common ancestral sequence, although there is another part of the alignment [Fig. 10(f)] which shows close similarity. Residues 1-7 in the plastocyanin sequence match closely with residues 27-33 in the azurin sequence. In addition, residue 28 is threonine in four of nine azurin sequences (Ambler, 1971). Predictions of secondary structure suggest that there may be an evolutionary relationship.

(d) Comparison of secondary structures of plastocyanins and azurins

Fig. 10(f) shows the predicted regions of α-helix and β-pleated sheet in A. variabilis plastocyanin and Pseudomonas aeruginosa P6009 azurin (Ambler & Brown, 1967) calculated by the method of Chou & Fasman (1974).

The method does not always clearly distinguish which type of secondary structure is to be expected and sometimes

overpredicts the existence of β-sheet. If the choice between α-helical and β-sheet secondary structure was close in the above examples of plastocyanin and azurin then the sequences of other proteins of these types were examined in order to make a decision. Ambler (1971) gives the alignment of the variations in amino acid sequence in nine azurins. The single letter notation for amino acids is used in Fig. 10(f) [Biochem. J. (1969) 113, 1-4] and the gaps in the alignment allow the matching round the cysteine residues shown in Fig. 10(e).

Although azurins are respiratory electron transfer proteins occurring in non-photosynthetic bacteria (Sutherland & Wilkinson, 1963) and plastocyanins function in the photosynthetic electron transfer chain, it is possible that the two types of protein evolved from a common ancestor. Timkovich & Dickerson (1973) have suggested that the photosynthetic and respiratory electron transfer chains had a common origin [see Chapter 10(f)].

Azurins have been proposed to function in a respiratory chain in the position shown below (Dickerson & Timkovich, 1975).

e^- → Cytochome b → C-type cytochromes → azurin → (Cu protein)

→ "oxidase" (nitrite reductase)

There are strong similarities between this respiratory chain and the photosynthetic electron transfer chain of oxygenic photosynthesisers [Fig. 1(b)] in which the following electron transfer process occurs:-

e^- → plastoquinone → cytochrome b → cytochrome f → (C-type cytochrome)

→ plastocyanin → P700. (Cu protein)

The functions and the primary structures of these copper proteins may have altered radically since divergence but secondary and tertiary structures of proteins are highly conserved.

If we consider the close similarity in secondary structure over the whole alignment of the proteins and the areas of matching primary structure at the terminal regions it is quite reasonable to suggest that plastocyanin and azurin shared a common ancestor. The large time span since divergence and the change in function preclude any use of comparisons between the sequences of these copper proteins for phylogenetic purposes.

(e) Comparison of prokaryotic and eukaryotic algal cytochromes f

The amino acid sequences of cytochromes f are aligned in Fig. 10(g). The alignment is adapted from Ambler & Bartsch (1975). The sequences show great similarity at the terminal regions particularly round the haem attachment site (residues 14-17) and around Met-62 [the sixth haem ligand (Dickerson et al., 1975]. There are regions in the middle of the sequences (positions 34-48) which are very varied and are of different lengths (including the two complete prokaryote sequences).

A remarkable result emerges from the percentage similarity matrix [Fig. 10(h)] which compares the completed sequences of cytochromes f. Although there is quite a spread in the degree of similarity of the prokaryotic cytochromes f to the eukaryotic algal cytochromes f, the difference in amount of similarity between either P. luridum or S. maxima cytochrome f

Fig. 10(g). Alignment of cytochromes f

A *Synechococcus* 6312

B *Anabaena variabilis*

C *Phormidium luridum*

D *Spirulina maxima* (Ambler & Bartsch, 1975)

E *Monochrysis lutheri* (chrysophytan alga; Laycock, 1972)

F *Porphyra tenera* (red alga; unpublished, aligned in Ambler & Bartsch, 1975)

G *Euglena gracilis* (Pettigrew, 1974)

H *Alaria esculenta* (brown alga; Laycock, 1975)

1 10 20 30

(A) Ala-Asp-Ile-Ala-Asp-Gly-Ala-Lys-Val-Phe-Ser-Ala-Asn-Cys-Ala-Ala-Cys-His-Met-Gly-Gly-Gly-Asn(Val)(Ile)Met-Ala-Asn-Lys-Thr-Leu-

(B) Ala-Asp-Ser-Val-Asn-Gly-Ala-Lys-Ile-Phe-Ser-Ala-Asn-Cys-Ala-Ser-Cys-His-Ala-Gly-Gly-Lys Thr-Leu-

(C) Ala-Asp-Ala-Ala-Ala-Gly-Gly-Lys-Val-Phe-Asn-Ala-Asn-Cys-Ala-Ala-Cys-His-Ala-Ser-Gly-Gly-Gly-Gln-Ile-Asn-Gly-Ala-Lys-Thr-Leu-

(D) Gly-Asp-Val-Ala-Ala-Gly-Ala-Ser-Val-Phe-Ser-Ala-Asn-Cys-Ala-Ala-Cys-His-Met-Gly-Gly-Arg-Asn-Val-Ile-Val-Ala-Asn-Lys-Thr-Leu-

(E) Gly-Asp-Ile-Ala-Asn-Gly-Glu-Gln-Val-Phe-Thr-Gly-Asn-Cys-Ala-Ala-Cys-His-Ser-Val-Glx-Glx-Glx-Mml-Thr-Leu-Glu-Leu-Ser-Ser-Leu-

(F) Ala-Asp-Leu-Asp-Asn-Gly-Glu-Lys-Val-Phe-Ser-Ala-Asn-Cys-Ala-Ala-Cys-His-Ala-Gly-Gly-Asn-Asn-Ala-Ile-Met-Pro-Asp-Lys-Thr-Leu-

(G) Gly-Gly-Ala-Asp-Val-Phe-Ala-Asp-Asn-Cys-Ser-Thr-Cys-His-Val-Asn-Gly-Gly-Asn-Val-Ile-Ser-Ala-Gly-Lys-Val-Leu-

(H) Ile-Asp-Ile-Asp-Asn-Gly-Glu-Asp-Ile-Phe-Thr-Ala-Asn-Cys-Ser-Ala-Cys-His-Ala-Gly-Gly-Asn-Asn-Val-Ile-Met-Pro-Glu-Lys-Thr-Leu-

40 50 60

(A) -Asn-Ala-Met-

(B)-Lys-Lys-Ala-Asp-Leu-Glu-Lys-Tyr -Asn-Ala-Met-

(C)-Lys-Lys-Asn-Ala-Leu-Thr-Ala----------Asn-Gly-Lys-Asp-Thr-Val-Glu-Ala-Ile-Val-Ala-----Gln-Val-Thr-Asn-Gly-Lys-Gly-Ala-Met-

(D)-Ser-Lys-Ser-Asp-Leu-Ala-Lys-Tyr-Leu-Lys-Gly-Phe-Asp-Asp-Asp-Ala-Val-Ala-Ala-Val-Ala-Tyr-Gln-Val-Thr-Asn-Gly-Lys-Asn-Ala-Met-

(E)-Trp-Lys----------Ala-Lys-Ser-Tyr-Leu-Ala-Asn-Phe-Asn-Gly-Asp-Glu-Ser-Ala-Ile-Val------Tyr-Gln-Val-Thr-Asn-Gly-Lys-Asn-Ala-Met-

(F)-Lys-Lys-----Asp-Val----------Leu-Glu-Ala-Asn-Ser-Met-Asn-Thr-Ile-Asp-Ala-Ile-Thr-----Tyr-Gln-Val-Gln-Asn-Gly-Lys-Asn-Ala-Met-

(G)-Ser-Lys-Thr-Ala-Ile-Glu-Glu-Tyr-Leu-Asp-Gly-Gly-Tyr------Thr-Lys-Glu-Ala-Ile-Glu------Tyr-Gln-Val-Arg-Asn-Gly-Lys-Gly-Pro-Met-

(H)-Lys-Lys-----Asp-Ala----------Leu-Ala-Asp-Asn-Lys-Met-Val-Ser-Val-Asn-Ala-Ile-Thr-----Tyr-Gln-Val-Thr-Asn-Gly-Lys-Asn-Ala-Met-

70 80 90

(A)-Pro-Gly-Phe-Ala-Gly-Arg-

(B)-Pro-Ala-Phe-Lys-Gly-Arg-Leu-Lys-Pro-Glu-Glu-Ile-Glx-Asx-Val-Ala-Ala-Tyr-Val-Leu-Gly-Lys-Ala-Asp-Ala-Asp-Trp-Lys

(C)-Pro-Ala-Phe-Lys-Gly-Arg-Leu-Ser-Asp-Asp-Gln-Ile-Gln-Ser-Val-Ala-Leu-Tyr-Val-Leu-Asp-Lys-Ala-Glu-Lys-Gly-Trp

(D)-Pro-Gly-Phe-Asn-Gly-Arg-Leu-Ser-Pro-Lys-Gln-Ile-Glu-Asp-Val-Ala-Ala-Tyr-Val-Val-Asp-Gln-Ala-Glu-Lys-Gly-Trp

(E)-Pro-Ala-Phe-Gly-Gly-Arg-Leu-Glu-Asp-Asp-Glu-Ile-Ala-Asx-Val-Ala-Ser-Tyr-Val-Leu-Ser-Lys-Ala-Gly

(F)-Pro-Ala-Phe-Gly-Gly-Arg-Leu-Val-Asp-Glu-Asp-Ile-Glu-Asp-Ala-Ala-Asn-Tyr-Val-Leu-Ser-Gln-Ser-Glu-Lys-Gly-Trp

(G)-Pro-Ala-Trp-Glu-Gly-Val-Leu-Ser-Glu-Asp-Glu-Ile-Val-Ala-Val-Thr-Asp-Tyr-Val-Tyr-Thr-Gln-Ala-Gly-Gly-Ala-Trp-Ala-Asn-Val

(H)-Pro-Ala-Phe-Gly-Ser-Arg-Leu-Ala-Glu-Thr-Asp-Ile-Glu-Asp-Val-Ala-Asn-Phe-Val-Leu-Thr-Gln-Ser-Asp-Lys-Gly-Trp-Asp

Fig. 10(g)

	C	D	E	F	G	H
C	-	56	46	54	43	47
D	56	-	47	53	42	49
E	46	47	-	48	35	45
F	54	53	48	-	35	67
G	43	42	35	35	-	39
H	47	49	45	67	39	-

Fig. 10(h) Percentage similarity matrix for cytochromes f

C *P. luridum*

D *Spirulina maxima*

E *Monochrysis lutheri*

F *Porphyra tenera*

G *Euglena gracilis*

H *Alaria esculenta*

For the comparison Asx is taken as identical to both Asp and Asn, and Glx to both Glu and Gln.

and any eukaryotic algal sequence is very small, e.g., the percentage similarities of P. luridum and S. maxima sequences to Alaria esculenta cytochrome f are 47% and 49% respectively (a difference of two residues).

The numbers of identities in prokaryote cytochrome f sequences to each of the other eukaryotic algal sequences differ by only one residue. The prokaryotic cytochromes f have themselves only 56% identity in sequence (and two deletions in P. luridum cytochrome f).

The blue-green algal cytochromes f are only about 10% closer to each other in sequence than to the eukaryotic sequences (except for the red alga, Porphyra tenera). The partial sequences of the other prokaryotic cytochromes f, particularly that of A. variabilis, suggests that the relationships discussed above may be fairly general.

In view of the particularly close similarity of the red algal cytochrome f to the prokaryotic sequences this may lend support to the hypothesis of polyphyletic origin of chloroplasts [Chapter 1(h)] or a modification of this proposed by Taylor (1974) that the existence of photosynthetic bodies in eukaryotes which seem intermediate between endosymbiotic blue-green algae and chloroplasts suggests that some eukaryotic algae may have had a more recent origin than the Precambrian era. This implies that organisms such as red algae diverted more recently from blue-green algae than other eukaryotic algal orders. Rhodophyta have been described by Allsopp (1969) and Klein & Cronquist (1967) as "primitive eukaryotes" but these authors accept the

```
 1  GDVEKGKKIF-VQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFTYTDANKN-K
 2 EGDAAAGEKVSK-K-CLACHTFDQGGANKVGPNLFGVFENTAAHKDNYAYSESYTEMK
 3     ATPAELATKAGCAVCHQPTA--K-GLGPSYqeIAKK----------------
 4  ADAAAGGKVF-NANCAACHAS-GGG----QINGAKLTKK----------------
 5  YDAAAGKATYDAS-CAMCHKT-GMG----APK-VG--DKA---------------
6 AASAGGGARSADDII-AKHCNACHGAGVLG----APK-------I-----------
 7     EDPEVLFKNKGCVACHAIDTKM---VGPAYKDVAAK----------------
 8     ADGAALY--KSCIGCHSADGGK---AMMTNA-VKGKYSD-------------
```

```
1 ---GITWKEETLMEYLENPKKYI------PG--TK- M-IF--AGIKKKTEREDLIAYLKKATNE
2 AKGLTWTEANLAAYVKDPKAFVLEKSGDPKAKSK- M-TF---KLTKDDEIENVIAYLKTLK
3    YKGQAGapamLAERVRKGSVGI------FGKLPMTPTppa-RISDADIKLVIDWILKTP
4 NALTANGK---DTVEAIV-A-QVTN-------GKGAMPAFKG--RLSDDQIQSVALYVLDKAEKGW
5 ----AWAP-HIAKGMNVM----VANSIKGYKGTKGMMPAKGGNPKLTDAQVGNAVAYMVGQSK
6 DGTAAWKERADHQGGLDG-ILAKAIS------GINAMPPKGT--CADCSDDELREAIQKMSGL
7    FAGQAGAEAELAQRIKNGSQGV------WGPIPMPPNAV----SDDEAQTLAKWVLSQK
8 EELKALADYMKAAMGSAK-PVKGQGAE-----ELYKMKGYAD--GSYGGERKAMSKL
```

1 Equus caballus cytochrome c

2 Rhodospirillum rubrum cyt. c_2

3 Rhodopseudomonas gelatinosa cyt. c_2 (R.P.Ambler, personal communication, tentative residues in lower case)

4 P.luridum cytochrome f

5 Chlorobium thiosulphatophilum c-555

6 Pseudomonas mendocina cyt. c_5

7 Pseudomonas aeruginosa cyt. c-551

8 Desulphovibrio vulgaris cyt. c-553

Fig. 10(i)

existence of a large gap between the cytology of prokaryotic blue-green algae and the simplest member of the red algae. The greater sequence similarity between blue-green and red algal cytochromes f may be due to a more recent time of divergence or due to closer functional similarities between the proteins in blue-green algae and red algal chromatophores than between blue-green algae and other chloroplasts. Although the cytochromes f from the red and brown algae are very close, as might be expected, the blue-green algal cytochromes f are not particularly close to the brown algal sequence.

The greater distinctiveness of the Euglena gracilis cytochrome f has been noted by Ambler & Bartsch (1975), in keeping with many other differences between euglenoids and other algae.

(f) Comparison of cytochromes f with bacterial cytochromes

Dickerson et al. (1975) have suggested that various c-type cytochromes, aligned in Fig. 10(i), may have had a common evolutionary origin. This is based on their conclusions reached from X-ray structural analysis of respiratory cytochrome c from eukaryotes, photosynthetic cytochrome c_2 from purple non-sulphur bacteria and cytochrome c-550 from nitrate respiring Paracoccus denitrificans. Low resolution X-ray analysis on Pseudomonas cytochrome c-551 suggests that it has the same overall folding pattern (the cytochrome fold). Tertiary structural differences between the cytochromes occur in the form of additional loops of chain on the molecular surface. The large numbers of gaps

seen in the alignment in Fig. 10(i) are based on the suggested tertiary structures in the other cytochromes.

The amino acid sequence of P. luridum cytochrome f is aligned with cytochromes from bacteria and horse cytochrome c. The cytochromes shown are examples of the types of cytochrome which have the standard haem attachment site near the N-terminus and a methionine residue (the sixth iron ligand) further along the sequence towards the C-terminus. In some cytochromes the methionine is followed by a chain-bending proline residue. The methionine in cytochrome c-551 is bracketed by 3 proline residues:-

——— C x x C H ——————————(P) M (P)(P)———

Multihaem cytochromes and cytochrome c' which is high-spin and has the haem attachment site near the C-terminus are excluded from this comparison.

The cytochrome f from P. luridum appears to be particularly close to cytochrome c-555 from Chlorobium thiosulphatophilum. The blue-green algal cytochromes f have 23-25 residues identical to this c-555 and eukaryotic cytochromes f have about 20 residues identical in the alignment shown [Fig. 10(i)].

There are about 19 identities between the prokaryotic cytochromes f and the cytochromes c_2. The eukaryotic cytochromes f have about 17 residues in common with the published cytochromes c_2 in the extensive comparison by Dickerson et al. (1975).

The similarity between cytochrome c-555 and the prokaryotic and eukaryotic algal cytochromes f is in agreement with the observations of Yamanaka & Okunuki (1968) who showed

similarities between cytochrome c-555 and eukaryotic algal cytochromes f in their Soret/α-peak ratios and in reactions with oxidase systems. The cytochromes f are less similar in primary structure to cytochromes c_2 although the redox potentials of these two types of cytochrome are much more similar (c_2 = +0.31 V and f = +0.36 V). The redox potential of cytochrome c-555 is +0.14 V (Lemberg & Barrett, 1973).

The structural similarities between the various c-type cytochromes discussed here support the proposal that the electron transfer chains of photosynthesis and respiration shared a common ancestry (Timkovich & Dickerson, 1973). It is probable that respiration evolved from bacterial photosynthesis by the development of new reducing donors and oxidising acceptors at the ends of the common ancestral chain. Dickerson et al. (1975) have suggested that the divergence occurred in the ancestors of the purple non-sulphur bacteria as these are the only photosynthetic bacteria capable of both respiration and photosynthesis. In addition evidence suggests that the same electron transfer chain is used in both photosynthesis and respiration in these photosynthetic bacteria.

The blue-green algal cytochromes f may not be significantly closer to the bacterial cytochromes compared above, than the cytochromes f from eukaryotic algae. Without the aid of a computer rigid numerical comparisons cannot be made between the large numbers of sequences available. It is possible that the group of photosynthetic electron transfer cytochromes (f, c-555 and c_2) have retained greater similarity to each other because of greater functional similarity.

Photosynthetic cytochrome c_2 has been included in the phylogenetic tree relating respiratory cytochromes c because of the comparatively close sequence similarities between the two types of cytochrome (Dayhoff, 1972). When cytochrome c_2 (from Rhodospirillum rubrum) is included, the cytochrome c sequence data support one of the alternatives in the endosymbiotic hypothesis that the initial eukaryotic cell line was non-photosynthetic (i.e. endosymbiosis of protomitochondrion occurred first) because the amino acid sequence of cytochrome c from the protozoan Crithidia oncopelti (Pettigrew, 1972) is closer to the prokaryotic sequence than to those from the photosynthetic eukaryotes (McLaughlin & Dayhoff, 1973).

(g) Origin of chloroplast and eukaryotic cell

The G + C contents of the DNA from the blue-green algae in this study are:- Anabaena variabilis, 44%; Phormidium luridum, 47%; Synechococcus 6312, 50.2%; Spirulina maxima, 36.6%. The first two values are from Edelman et al. (1967), Synechococcus from Stanier et al. (1971) and the last is a personal communication from M.Herdman measured on dried cells from the source in Ambler & Bartsch (1975). Chloroplast DNA from a wide range of organisms has a narrow range of base composition between 37% and 39% G + C although Euglena gracilis (25% G + C) and Chlorella ellipsoidea (33% G + C) chloroplast DNAs are markedly different (Ellis & Hartley, 1974).

The blue-green algae in the present study are fairly diverse with respect to DNA base composition. The range of composition for blue-green algae is 34% to 71% (Leach &

Herdman, 1973) but most fall within much narrower limits. Fossil evidence indicates that filamentous and unicellular forms diverged in the early Precambrian era and many blue-green algae in the Bitter Springs microflora (0.9×10^9 years ago) were comparable at specific or generic level with extant Cyanophytes e.g. Spirulina and the other members of Oscillatoriaceae (Schopf, 1970). It seems likely therefore that most of the sequence comparisons made here will prove generally valid for blue-green algae as a whole.

The amino acid sequences of ferredoxin from the blue-green algae Spirulina maxima (Tanaka et al., 1975), Spirulina platensis (Wada et al., 1975) and from Aphanothece sacrum (Wada et al., 1974) have been compared to the sequences of ferredoxins from eukaryotic algae and plants. Similar results to those in the present study were obtained. The blue-green algal protein sequences are either not closer or only slightly closer to each other than to the corresponding eukaryotic algal sequences. The amino acid sequences of plastocyanin and ferredoxin from blue-green algae do appear to be closer to each other than to the corresponding higher plant sequences. Ferredoxin sequence comparisons must be treated with caution because multiple ferredoxins with quite distinct amino acid compositions have been shown to be present in some plants (Tanaka et al., 1975) and in a blue-green alga, Noctoc strain MAC (Hutson & Rogers, 1975).

These sequence studies have added considerable confirmation to the hypothesis that blue-green algae are evolutionarily related to photosynthetic eukaryotes. The amino acid sequence similarities are too great to consider the possibility

of convergent evolution. The sequence comparisons do not clearly distinguish whether eukaryotes descended as a whole from a blue-green algal ancestor or if the chloroplast alone evolved from such a prokaryote as outlined in Chapter 1(h).

The rate of evolution of cytochrome f is probably comparable to that of mitochondrial cytochrome c (three "accepted point mutations" per 100 million years, Dayhoff, 1972) and the rate of evolution of higher plant plastocyanin is about twice that of cytochrome c (Peacock & Boulter, 1975). It seems probable that the rates of evolution of these proteins in blue-green algae is much less than in eukaryotes. It is fairly certain that blue-green algae have evolved very little in morphological terms since the Precambrian era (see above) and probably have evolved little in physiology since the oxygenic mode of photosynthesis was established. The concept of linear rates of evolution for a particular protein (Dickerson, 1971) may be acceptable for proteins which function in a similar organelle in a range of eukaryotic organisms which are evolving. The rate of evolution of proteins such as mitochondrial cytochrome c appears to be approximately constant where the various factors involved such as generation time of the particular organism, constraints on mutation in the protein due to surface interactions with membranes or other proteins etc. balance out over the time scale and range of organisms. In blue-green algae there is probably very little selective value for changes in primary structure of photosynthetic proteins but chloroplasts have evolved and diverged in a wide variety of eukaryotes during the time that blue-green algae have continued

to populate similar ecological habitats. This would account for the finding that the blue-green algal proteins have about the same diversity amongst themselves as they have to eukaryotic algal proteins although blue-green algae diverged much longer ago than eukaryotic algal orders. By this interpretation the findings are very much against the theory of neutral mutation in proteins [Chapter 1(d)].

Another possible explanation of the results is that the pattern of divergence in the amino acid sequences has been obscured by transfer of genetic material coding for the photosynthetic proteins. This does not satisfactorily explain the approximately equal sequence variations between prokaryotic and eukaryotic algae (the latter belonging to different orders). It is difficult to envisage a mode of gene transfer between prokaryotes and eukaryotic algae of different orders although this could have occurred some time in the past evolutionary history of these photosynthetic organisms when the structural organisation of the cells and cell walls was not as different.

There is only a small probability that in this study, by chance, only blue-green algae which comparatively recently diverged were chosen. More amino acid sequences of proteins from a diverse range of eukaryotic algae and filamentous and unicellular blue-green algae would establish whether or not the conclusions are widely justified. The amino acid sequence of plastocyanin from the lower red alga, Porphyridium aerugineum (Visser et al., 1974) could prove particularly interesting. The amino acid sequences of bacterial cytochromes suggest that substantial gene transfer has occurred

between distantly related bacteria (Ambler, 1973). If this proves true for blue-green algae it would be less likely that molecular methods could distinguish which, if any, of the hypotheses discussed in this thesis were correct.

The comparison of amino acid sequences of proteins involved in eukaryotic cytoplasmic and mitochondrial functions to the corresponding blue-green algal proteins may prove useful in eliminating some of the possibilities. That is, if the topologies of phylogenetic trees for proteins such as superoxide dismutase (Lumsden & Hall, 1975), mitochondrial cytochrome c and other proteins which function in different organelles are seen to be substantially different.

REFERENCES

Allen, M.B. & Arnon, D.I. (1955) Plant Physiol. 30, 366-372.

Allen, M.M. (1968) J. Phycol. 4, 1-4.

Allen, M.M. (1972) Arch. Mikrobiol. 84, 199-206.

Allsopp, A. (1969) New Phytol. 68, 591-612.

Ambler, R.P. (1963) Biochem. J. 89, 341-349.

Ambler, R.P. (1971) in "Recent Developments in the Chemical Study of Protein Structures" (Previero, A., Pechere, J.-F. & Coletti-Previero, M.-A., eds.) pp. 289-305, Inserm, Paris.

Ambler, R.P. (1972) Methods Enzymol. 25, 262-272.

Ambler, R.P. (1973) Syst. Zool. 22, 554-565.

Ambler, R.P. (1974) Biochem. J. 137, 3-14.

Ambler, R.P. & Bartsch, R.G. (1975) Nature (London) 253, 285-288.

Ambler, R.P. & Brown, L.H. (1967) Biochem. J. 104, 784-825.

Ambler, R.P. & Wynn, M. (1973) Biochem. J. 131, 485-498.

Amesz. J. & Duysens, L.N.M. (1962) Biochim. Biophys. Acta 64, 261-278.

Andrews, P. (1964) Biochem. J. 91, 222-233.

Barghoorn, E. (1971) Sci. Am. 224, 30-42.

Bearden, A.J. & Malkin, R. (1974) Quart. Rev. Biophys. 7, 131-178.

Bendall, D.S., Davenport. H.E. & Hill, R. (1971) Methods Enzymol. 23, 327-344.

Bendall, D.S. & Hill, R. (1968) Annu. Rev. Plant Physiol. 19, 167-186.

Berkner, L.V. & Marshall, L.C. (1965) Proc. Nat. Acad. Sci. U.S. 53, 1215-1226.

Berns, D.S. (1967) Plant Physiol. 42, 1569-1586.

Biggins, J. (1967) Plant Physiol. 42, 1447-1456.

Bishop, N.I. (1971) Annu. Rev. Biochem. 40, 197-226.

Bisset, K. (1973) New Scientist 57, 296-298.

Bogorad, L. (1975) Science 188, 891-898.

Borchert, M.T. & Wessels, J.S.C. (1970) Biochim. Biophys. Acta 197, 78-83.

Bornstein, P. (1970) Biochemistry 9, 2408-2421.

Boulter, D., Ellis, R.J. & Yarwood, A. (1972) Biol. Rev. 47, 113-175.

Bridgen, J., Harris, J.I. & Northrop, F. (1975) FEBS Letters 49, 392-395.

Brock, T.D. (1973) in "The Biology of the Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) pp. 487-500, Blackwell, Oxford.

Brooks, J., Muir, M.D. & Shaw, G. (1973) Nature (London) 244, 215-217.

Brown, A.S., Foster, J.A., Voynow, P.V., Franzblau, C. & Troxler, R.F. (1975) Biochemistry, in the press.

Brown, J.R. & Hartley, B.S. (1966) Biochem. J. 101, 214-228.

Calvin, M. (1969) Chemical Evolution, Oxford Univ. Press, Oxford.

Cannon, R.E., Shane, M.S. & Bush, V.N. (1971) Virology 45, 149-153.

Carr, N.G. (1966) Biochim. Biophys. Acta 120, 308-310.

Carr, N.G. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) pp. 39-65 (Blackwell, Oxford).

Castenholz, R.W. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) pp. 320-339 (Blackwell, Oxford).

Cavalier-Smith, T. (1975) Nature (London) 256, 463-468.

Chou, P.Y. & Fasman, G.D. (1974) Biochemistry 13, 222-245.

Clarke, B. (1970) Nature (London) 228, 159-160.

Cloud, P.E. (1968) Science 160, 729-736.

Cohen, S.S. (1973) Amer. Sci. 61, 437-445.

Cohen, Y., Padan, E. & Shilo, M. (1975) J. Bacteriol. - in the press.

Davenport, H.E. & Hill, R. (1952) Proc. R. Soc., Lond., B139, 327-345.

Dayhoff, M.O. (1972) Atlas of Protein Sequence & Structure (pub. National Biomedical Research Foundation).

Dent, C.E. (1947) Biochem. J. 41, 240-253.

Deselnicu, M., Lange, P.M. & Heidemann, E. (1973) Z. Physiol. Chem. 354, 105-116.

Dickerson, R.E. (1971) J. Molec. Evolution 1, 26-45.

Dickerson, R.E. & Timkovich, R. (1975) "Cytochromes c" in The Enzymes, ed. Boyer, P. (Academic Press, New York), 3rd Edition, Vol. 11, Chap. 7.

Dickerson, R.E., Timkovich, R. & Almassy, R.J. (1975) Proc. Nat. Acad. Sci. U.S. - in the press.

Doolittle, R.F. (1970) Methods Enzymol, 19, 555-569.

Drews, G. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) pp. 99-116 (Blackwell, Oxford).

de Duve, C. (1973) Science 182, 85p.

Echlin, P. (1970) Symp. Soc. Genl. Microbiol. 20, 221-248.

Echlin, P. & Morris, I. (1965) Biol. Rev. 40, 143-187.

Edman, P. & Begg, G. (1967) Eur. J. Biochem. 1, 80-91.

Edman, P. & Sjoquist, J. (1956) Acta Chem. Scand. 10, 1507-1509.

Edelmann, M., Swinton, D., Schiff, J.A., Epstein, H.T. & Zeldin, B. (1967) Bacteriol. Rev. 31, 315-331.

Ellis, R.J. & Hartley, M.R. (1974) Nucleic Acids of Chloroplasts in MTP International Review of Science, Biochemistry Series One, Vol. 6 "Biochemistry of Nucleic Acids (Burton, K. ed.), Butterworths & University Park Press.

Engel, A.E.U., Nagy, B., Nagy, L.A., Engel, C.G., Kremp, C.W.W. & Drew, C.M. (1968) Science 161, 1005-1008.

Fogg, G.E., Stewart, W.D.P., Fay, P. & Walsby, A.E. (1973) in "The Blue-Green Algae". Academic Press (London).

George, E. (1971) "Culture Collection of Algae & Protozoa. List of Strains". Culture Centre of Algae & Protozoa, Cambridge, U.K.

Glazer, A.N. & Cohen-Hazire, G. (1971) Proc. Nat. Acad. Sci. U.S. 68, 1398-1401.

Goksoyr, J. (1967) Nature (London) 214, 1161p.

Gorman, D.S. & Levine, R.P. (1965) Proc. Nat. Acad. Sci. U.S. 54, 1665-1669.

Gray, W.R. (1972a) Methods Enzymol. 25, 121-143.

Gray, W.R. (1972b) Methods Enzymol. 25, 333-344.

Gross, E. (1967) Methods Enzymol. 11, 238-255.

Hall, J.B. (1973) J. Theor. Biol. 38, 413-418.

Hall, W.T. & Claus, G. (1963) J. Cell. Biol. 19, 551-564.

Hall, W.T. & Claus, G. (1967) J. Phycol, 3, 37-51.

Hartley, B.S. (1970) Biochem. J. 119, 805-822.

Haslett, B.G. & Cammack, R. (1974) Biochem. J. 144, 567-572.

Hermodsen, M.A., Ericsson, L.H., Titani, K, Neurath, H. & Walsh, K.A. (1972) Biochemistry 11, 4493-4502.

Hind, G. & Olson, J.M. (1968) Annu. Rev. Plant Physiol. 19, 249-282.

Holton, R.W. & Myers, J. (1967) Biochim. Biophys. Acta 131, 362-374.

Houmard, J. & Drapeau, G.R. (1972) Proc. Nat. Acad. Sci. U.S. 69, 3506-3509.

Hughes, D.E., Llyod, D. & Brightwell, R. (1970) Symp. Soc. Genl. Microbiol. 20, 295-322.

Hutson, K.G. & Rogers, L.J. (1975) Biochem. Soc. Trans. 3, 377-379.

Inglis, A.S., Nicholls, P.W. & Roxburgh, C.M. (1971) Aust. J. Biol. Sci. 24, 1247-1250.

John, P. & Whatley, F.R. (1975) Nature (London) 254, 495-498.

Katoh, S., Suga, I., Shiratori, I. & Takamiya, A. (1961) Arch. Biochem. Biophys. 94, 136-141.

Katoh, S., Shiratori, I. & Takamiya, A. (1962) J. Biochem. (Tokyo) 51, 32-40.

Kelly, J. & Ambler, R.P. (1974) Biochem. J. 143, 681-690.

Kanyon, C.N. & Stanier, R.Y. (1970) Nature (London) 227, 1164-1166.

Kimura, M. (1968) Nature (London) 217, 624-626.

Kimura, M. & Ohta, T. (1973) Nature New Biol. 243, 199-200.

King, J.L. & Jukes, T.H. (1969) Science 164, 788-798.

Klein, R. & Cronquist, A. (1967) Quart. Rev. Biol. 42, 105-296.

Knaff, D.B. & Arnon, D.I. (1969) Proc. Nat. Acad. Sci. U.S. 64, 715-722.

Kratz, W.A. & Myers, J. (1955) Am. J. Bot. 42, 282-287.

Krogmann, D.W. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) 80-98 (Blackwell, Oxford).

Kunert, K.J. & Böger, P. (1975) Z. Naturforsch 30c, 190-200.

Lang, N.J. (1968) Annu. Rev. Microbiol. 22, 15-56.

Langley, C.H. & Fitch, W.M. (1974) J. Molec. Biol. 3, 161-177.

Laycock, M.V. (1972) Can. J. Biochem. 50, 1311-1325.

Laycock, M.V. (1975) Biochem. J. 149, 271-279.

Lea, P.J. & Norris, R.D. (1972) Phytochem. 11, 2897-2920.

Leach, C.K. & Herdman, M. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) pp. 186-200, (Blackwell, Oxford).

Lemberg, R. & Barrett, J. (1973) "The Cytochromes" Academic Press, London.

Lightbody, J.J. & Krogmann, D.W. (1967) Biochim. Biophys. Acta 131, 508-515.

Lumsden, J. & Hall, D.O. (1975) Nature (London) 257, 670-671.

MacGregor, I.M., Truswell, J.F. & Eriksson, K.A. (1974) Nature (London) 247, 538-540.

Makino, F. & Tsuzaki, J. (1971) Nature (London) 231, 446-447.

Malkin, R., Knaff, D.B. & Bearden, A.J. (1973) Biochim. Biophys. Acta 305, 675-678.

Malkin, R. & Malmström, B.G. (1970) Adv. Enzymol. Relat. Areas Mol. Biol. 33, 177-244.

Marguilis, L. (1970) Origin of Eukaryotic Cells, Yale University Press, New Haven.

McFadden, B.A. & Tabita, F.R. (1974) Biosystems 6, 93-112.

McLaughlin, P.J. & Dayhoff, M.O. (1973) J. Molec. Evolution 2, 99-116.

McMillin, D.R., Rosenberg, R.C. & Gray, H.B. (1974) Proc. Nat. Acad. Sci. U.S. 71, 4760-4762.

Michl, H. (1951) Monatsh. Chem. 82, 489-493.

Milne, P.R., Wells, J.R.E. & Ambler, R.P. (1974) Biochem. J. 143, 691-701.

Meyer, T.E. (1970) Ph.D. thesis, Univ. California, San Diego.

Miskowski, V., Tang, S.-P.W., Spiro, T.G., Shapiro, E. & Moss, T.H. (1975) Biochemistry 14, 1244-1250.

Moore, W.G., Goodman, M. & Barnabas, J. (1973) J. Theor. Biol. 38, 423-457.

Murray, R.G.E. (1974) in "Bergey's Manual of Determinative Bacteriology" (Buchanan, R.E. & Gibbons, N.E. eds.) p. 7, 8th Edn. (Williams & Wilkins, Baltimore).

Nagy, L.A. (1974) Science 183, 514-516.

Nass, S. (1969) Int. Rev. Cytol. 25, 55-129.

Nolan, W.G. & Bishop, D.G. (1975) Arch. Biochem. Biophys. 166, 323-329.

Niall, H.D. (1973) Methods Enzymol. 27, 942-1010.

Offord, R.E. (1966) Nature (London) 211, 591-593.

Ogawa, T. & Vernon, L.P. (1971) Biochim. Biophys. Acta 226, 88-97.

Olson, J.M. (1970) Science 168, 438-445.

Ornstein, L. (1964) Ann. N.Y. Acad. Sci. 121, 321-349.

Padan, E. & Shilo, M. (1973) Bacteriol. Rev. 13, 47-98.

Parson, W.W. (1974) Annu. Rev. Microbiol. 41, 41-59.

Peacock, D. & Boulter, D. (1975) J. Molec. Biol. 95, 513-527.

Penke, B., Frenczi, R. & Kovacs, K. (1974) Anal. Biochem. 60, 45-50.

Pettigrew, G.W. (1972) FEBS Letters 22, 64-66.

Pettigrew, G.W. (1974) Biochem. J. 139, 449-459.

Pfennig, N. (1967) Annu. Rev. Microbiol. 21, 285-324.

Pierson, B.K. & Castenholz, R.W. (1974) Arch. Mikrobiol. 100, 5-24.

Plesnicar, M. & Bendall, D.S. (1973) Eur. J. Biochem. 34, 483-488.

Prince, R.C. & Dutton, P.L. (1975) Biochim. Biophys. Acta 387, 609-613.

Pringsheim, E.G. (1949) Bacteriol. Rev. 13, 47-98.

Raff, R.A. & Mahler, H.R. (1972) Science 177, 575-582.

Ramshaw, J.A.M., Brown, R.H., Scawen, M.D. & Boulter, D. (1973) Biochim. Biophys. Acta 303, 269-273.

Ramshaw, J.A.M., Scawen, M.D., Bailey, C.J. & Boulter, D. (1974) Biochem. J. 139, 583-592.

Raven, P.H. (1970) Science 169, 641-646.

Reijnders, L. (1975) J. Molec. Evolution 5, 167-176.

Richmond, R.C. (1970) Nature (London) 225, 1025-1028.

Riordan, J.F. & Vallee, B.L. (1967) Methods Enzymol. 11, 570-576.

Rydon, H.N. & Smith, P.W.G. (1952) Nature (London) 169, 922-923.

Ryter, A. (1968) Bacteriol. Rev. 32, 39-54.

Safferman, R.S. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) 214-237, (Blackwell, Oxford).

Sagan, L. (1967) J. Theoret. Biol. 14, 225-274.

Scawen, M.D. & Boulter, D. (1974) Biochem. J. 143, 257-264.

Scawen, M.D., Ramshaw, J.A.M., Brown, R.H. & Boulter, D. (1974) Eur. J. Biochem. 44, 299-303.

Scawen, M.D., Hewitt, E.J. & James, D.M. (1975) Phytochem. 14, 1225-1233.

Schnepf, E. (1964) Arch. Mikrobiol. 49, 112-131.

Schopf, J.W. (1970) Biol. Rev. 45, 319-352.

Siedow, J.N., Curtis, V.A. & San Pietro, A. (1973) Arch. Biophys. Biochem. 158, 889-897.

Siegelman, M.H., Rasched, I. & Boger, P. (1975) Biochim. Biophys. Res. Commun. 65, 1456-1463.

Smith, L. (1961) in "The Bacteria" Vol. 2, 365-396 (edit. by Gunsalus, I.C. & Stanier, R.Y., Academic Press, London).

Soriano, S. & Lewin, R.A. (1965) Antonie von Leeuwenhoek 31, 66-80.

Stanier, R.Y. & van Niel, C.B. (1962) Arch. Mikrobiol. 42, 17-25.

Stanier, R.Y. (1970) Symp. Soc. Gen. Microbiol. 20, 1-38.

Stanier, R.Y., Kunisawa, R., Mandel, M. & Cohen-Bazire, G. (1971) Bacteriol. Rev. 35, 171-205.

Stanier, R.Y. (1974) Symp. Soc. Gen. Microbiol. 24, 219-240.

Starr, R.C. (1964) Am. J. Bot. 51, 1013-1044.

Steinman, H.M. & Hill, R.L. (1973) Proc. Nat. Acad. Sci. U.S. 70, 3725-3729.

Stewart, W.D.P. (1973) in "The Biology of Blue-Green Algae" (edit. by Carr, N.G. & Whitton, B.A.) 260-278 (Blackwell, Oxford).

Stewart, W.D.P. & Pearson, H.W. (1970) Proc. R. Soc. B175 293-311.

Susor, W.A. & Krogmann, D.W. (1966) Biochim. Biophys. Acta 120, 65-72.

Sutherland, I.W. & Wilkinson, J.K. (1963) J. Gen. Microbiol. 30, 105-112.

Swain, F.M. (1969) Annu. Rev. Microbiol. 23, 455-472.

Sybesma, C. (1969) Biochim. Biophys. Acta 172, 177-179.

Tanaka, M., Haniu, M., Zeitlin, S., Yasunobu, K.T., Evans, M.C.W., Rao, K.K. & Hall, D.O. (1975) Biochim. Biophys. Res. Comm. 64, 399-407.

Taylor, F.J.R. (1974) Taxon 23, 229-258.

Taylor, M.M. & Storck, R. (1964) Proc. Nat. Acad. Sci. U.S. 52, 958-965.

Tel-or, E., Cammack, R. & Hall, D.O. (1975) FEBS Letters 53, 135-138.

Timkovich, R. & Dickerson, R.E. (1973) J. Molec. Biol. 79, 39-56.

Troxler, R.F., Foster, J.A., Brown, A.S. & Franzblau, C. (1975) Biochemistry 14, 268-274.

Van Orden, H.O. & Carpenter, F.M. (1964) Biochim. Biophys. Res. Comm. 14, 399-403.

Visser, J.W.M., Amesz, J. & Van Gelder, B.F. (1974) Biochim. Biophys. Acta 333, 279-287.

Wada, K., Kagamiyama, H., Shin, M. & Matsubara, H. (1974) J. Biochem. (Tokyo) 76, 1217-1225.

Wada, K., Hase, T., Tokunaga, H. & Matsubara, H. (1975) FEBS Letters 55, 102-104.

Waley, S.G. & Watson, J. (1953) Biochem. J. 57, 529-538.

Watanabe, A. (1960) J. Gen. Appl. Microbiol. 6, 283-292.

Weber, K. & Osborn, M. (1969) J. Biol. Chem. 244, 4406-4412.

Wharton, D.C. & Rader, M. (1970) Anal. Biochem. 33, 226-229.

Williams, V.P., Freidenreich, P. & Glazer, A.N. (1974) Biochem. Biophys. Res. Commun. 59, 462-466.

Wittman, H.G. (1970) Symp. Soc. Gen. Microbiol. 20, 55-76.

Wolk, C.P. (1973) Bacteriol. Rev. 37, 32-101.

Wood, P.M. & Bendall, D.S. (1975) Biochim. Biophys. Acta 387, 115-128.

Yamanaka, K. & Okunuki, K. (1968) J. Biochem. (Tokyo) 63, 341-346.

chem. J. (1975) 149, 675-683
ted in Great Britain

Prokaryote–Eukaryote Relationships and the Amino Acid Sequence of Plastocyanin from *Anabaena variabilis*

By ALASTAIR AITKEN
Department of Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, U.K.

(*Received* 26 *February* 1975)

The amino acid sequence of plastocyanin from the prokaryotic blue-green alga *Anabaena variabilis* was determined. The protein consists of a single polypeptide chain of 105 residues. The amino acid sequence of the plastocyanin was compared with that of the eukaryotic green alga *Chlorella fusca* and with those of higher-plant plastocyanins. The considerable similarity between the prokaryotic and eukaryotic plastocyanins is discussed. Detailed evidence for the sequence of the protein has been deposited as Supplementary Publication SUP 50051 (13 pages) at the British Library (Lending Division), Boston Spa, Wetherby, W. Yorkshire LS23 7BQ, U.K., from whom copies may be obtained on the terms given in *Biochem J.* (1975) **145**, 5.

By the criteria of cell-wall composition, ribosomal ucture and absence of a nuclear membrane, blue- een algae are prokaryotic organisms (Stanier *et al.*, 71). Their photosynthetic apparatus is, however, ygenic and resembles that of the chloroplast much ore closely than that of photosynthetic bacteria. ere is much fossil evidence to indicate that blue- een algae are an ancient group of organisms chopf, 1970) which probably evolved from a photo- terotrophic ancestor common to all types of photo- nthetic bacteria (C'son, 1970). It has been suggested at blue-green algae were the first oxygen-evolving otosynthetic organisms and were initially respons- le for producing atmospheric O_2 (during the Pre- mbrian period). Some blue-green algae can grow aerobically (Stewart & Pearson, 1970) and mospheric O_2 is a prerequisite for the evolution of karyotes, which normally depend on mitochondrial spiration for their energy requirements.

The present study is an attempt to obtain some sight into the evolutionary connexion between okaryotes and eukaryotes by comparing the amino id sequence of plastocyanin from the blue-green ga *Anabaena variabilis* with plastocyanin sequences om eukaryotes.

Plastocyanin is a low-molecular-weight protein ntaining copper in a type 1 ligand environment Malkin & Malmström, 1970). The protein partici- ates in photosynthetic electron transport, linking otosystems 1 and 2 at a site near cytochrome *f* Bishop, 1971). Evidence also indicates that plasto- anin may be involved in a cyclic electron pathway round photosystem 1 (Visser *et al.*, 1974).

Plastocyanin is widely distributed in the chloro- lasts of plants (Katoh *et al.*, 1961). The protein has een isolated from green algae (Katoh, 1960) and vo members of the prokaryotic, filamentous blue-green algae *A. variabilis* (Lightbody & Krogmann, 1967) and *Phormidium luridum* (Biggins, 1967). Plastocyanin has also been detected by electron-paramagnetic-resonance signals in the unicellular blue-green alga *Anacystis nidulans* and in the red alga *Porphyridium aerugineum* (Visser *et al.*, 1974).

Attempts to isolate plastocyanin from *Anacystis nidulans*, *Spirulina maxima* (filamentous blue-green alga) and another unicellular blue-green alga, *Synechococcus* 6312 (Stanier *et al.*, 1971) have all been unsuccessful so far.

The amino acid sequence of plastocyanin from the green alga *Chlorella fusca* (Kelly & Ambler, 1974) and those from a number of plant plastocyanins (e.g. Milne *et al.*, 1974; Scawen & Boulter, 1974) have been determined. In the present paper evidence is presented for the amino acid sequence of plastocyanin from the prokaryote *A. variabilis*. The sequence is compared with eukaryotic plastocyanin sequences.

Experimental

Materials

Organism. The culture of *A. variabilis* (Kützing) was a gift from N. G. Carr, Department of Biochemistry, University of Liverpool, Liverpool, U.K.

Pyrollidonecarboxylyl peptidase and *Staphylococcus aureus* proteinase were prepared by A. J. Savill, Department of Molecular Biology, University of Edinburgh, Edinburgh, U.K.

Enzymes and materials. These were obtained from the sources listed by Ambler & Wynn (1973). Polyamide layer sheets were purchased from BDH Chemicals Ltd., Poole, Dorset, U.K.

Methods

Growth of organism. A. variabilis was grown on modified medium C of Kratz & Myers (1955) in 20-litre bottles illuminated with 20W Grolux lamps and bubbled with 5% CO_2 in air. $NaHCO_3$ (0.8g/litre) was added to the medium to keep the pH above 7. The cells were harvested in an Alfa-laval centrifuge (model LAB102B) and stored frozen until required. The yield of *A. variabilis* from 100 litres of culture was about 200g wet wt. of packed cells.

The cells (200g wet wt.) were suspended in 1 litre of potassium phosphate buffer (0.05M, pH7) and broken in a Manton–Gaulin homogenizer. The suspension was made 40% (w/v) in $(NH_4)_2SO_4$ and the precipitate was centrifuged. This step removed much of the phycobiliproteins. The supernatant was made 90% (w/v) in $(NH_4)_2SO_4$ and left at 4°C for 2 days to ensure complete precipitation. The precipitate was centrifuged at 10000*g* for 30min, redissolved in the minimum volume of distilled water, and then passed down a column (30cm×6cm) of Sephadex G-25 in 1mM-potassium phosphate (pH7). The plastocyanin was further purified by the methods of Lightbody & Krogmann (1967) which entails DEAE- and CM-cellulose chromatography. The yield of plastocyanin from 200g wet wt. of packed cells was about 20 mg.

Properties of A. variabilis plastocyanin. The spectral properties of the purified protein were identical with those obtained by Lightbody & Krogmann (1967). The best value of the absorption index (E_{278}/E_{597}; Katoh *et al.*, 1962) obtained was 1.56. The molar extinction coefficient of the copper chromophore in plastocyanins is approx. 4.5×10^3 litre·mol^{-1}·cm^{-1} and E_{278} for *A. variabilis* plastocyanin indicates lack of tryptophan in this protein (in common with higher-plant plastocyanins).

Amino acid-sequence determination. Before digestion the copper was removed by precipitation of the protein by the addition of an equal volume of 10% (w/v) trichloroacetic acid at 0°C which left the copper in solution. The white precipitate that formed was washed thoroughly with cold acetone and then dissolved in 0.1M-NH_3.

The amino acid sequence was determined by a study of the peptides produced by enzymic hydrolysis of the plastocyanin with trypsin, chymotrypsin and thermolysin. The *N*-terminal sequence was also investigated by automatic phenyl isothiocyanate degradation and the *C*-terminal sequence by carboxypeptidase A digestion.

The peptides were fractionated by gel filtration followed by high-voltage paper electrophoresis and paper chromatography. In some cases peptides were further degraded by other proteolytic enzymes. The peptides were examined for sequence by the dansyl–phenyl isothiocyanate method and the dansyl derivatives identified on polyamide layer sheets (Hartley, 1970). Identification of histidine and arginine dansylated derivatives was by high-voltage paper electrophoresis at pH4.4. The methods used for sequence determination were otherwise those described previously (Ambler & Brown, 1967; Ambler & Wynn, 1973).

Peptides were analysed quantitatively on an LKB Biocal BC200 amino acid analyser or a Beckman 120C amino acid analyser. The latter was used for analysis of whole protein or large peptides.

Quantities of material. The sequence determination was carried out on material derived from different batches of cells and the products of different isolations of plastocyanin. Tryptic digestion was carried out on 1.8μmol of denatured plastocyanin; chymotryptic digestion and thermolysin digestion on 2μmol each of denatured apoprotein.

Automatic phenylisothiocyanate degradation. Native oxidized plastocyanin (0.8μmol) was degraded in a Beckmann, model 890A, sequenator by using a 70min Quadrol double-cleavage programme (Edman & Begg, 1967). The phenylthiohydantoin derivatives produced were identified by t.l.c. on silica-gel layers (Eastman-Kodak, Rochester, N.Y., U.S.A.) containing fluorescent indicator (no. 13181) by using the solvent system of Edman & Sjoquist (1956), by g.l.c. on SP.4001 Chromosorb W and by hydrolysis with 5.5M-HI (Inglis *et al.*, 1971) followed by amino acid analysis. A sequenator result was considered reliable when the derivative was positively identified by two of the three above methods.

Results

The amino acid composition of *A. variabilis* plastocyanin is shown in Table 1.

Automatic phenyl isothiocyanate degradation unequivocally identified 24 of the first 36 residues (Fig. 1). Three of the remaining residues were tentatively identified (isoleucine-23, valine-29 and leucine-32). Conventional sequence determination experiments showed the undetected residues to be threonine$_{(4)}$, serine$_{(1)}$, asparagine$_{(2)}$ and lysine$_{(2)}$. The yields of phenylthiohydantoin derivatives were not quantified.

The evidence for the amino acid sequence is summarized in Fig. 1. This shows all the peptides isolated from the tryptic and chymotryptic digests together with a sufficient number of thermolysin peptides to complete the overlaps. The peptides obtained by papain subdigestion are also shown. Symbols indicate how much of the sequence of each peptide was determined by the dansyl–phenyl isothiocyanate method. Cases where dansyl–phenyl isothiocyanate degradation results were considered not wholly satisfactory are indicated.

The guidelines followed in considering dansyl–phenyl isothiocyanate degradation results or amino

Table 1. *Amino acid composition of A. variabilis plastocyanin*

Results are shown as residues per molecule. Analysis values were from 24h hydrolysis, except for the following amino acids: valine and isoleucine were derived from a 96h hydrolysis, and cysteine (as cysteic acid) and methionine (as methionine sulphone) from a performic acid-oxidized sample hydrolysed for 24h. Values for serine and threonine were not corrected for destruction during hydrolysis.

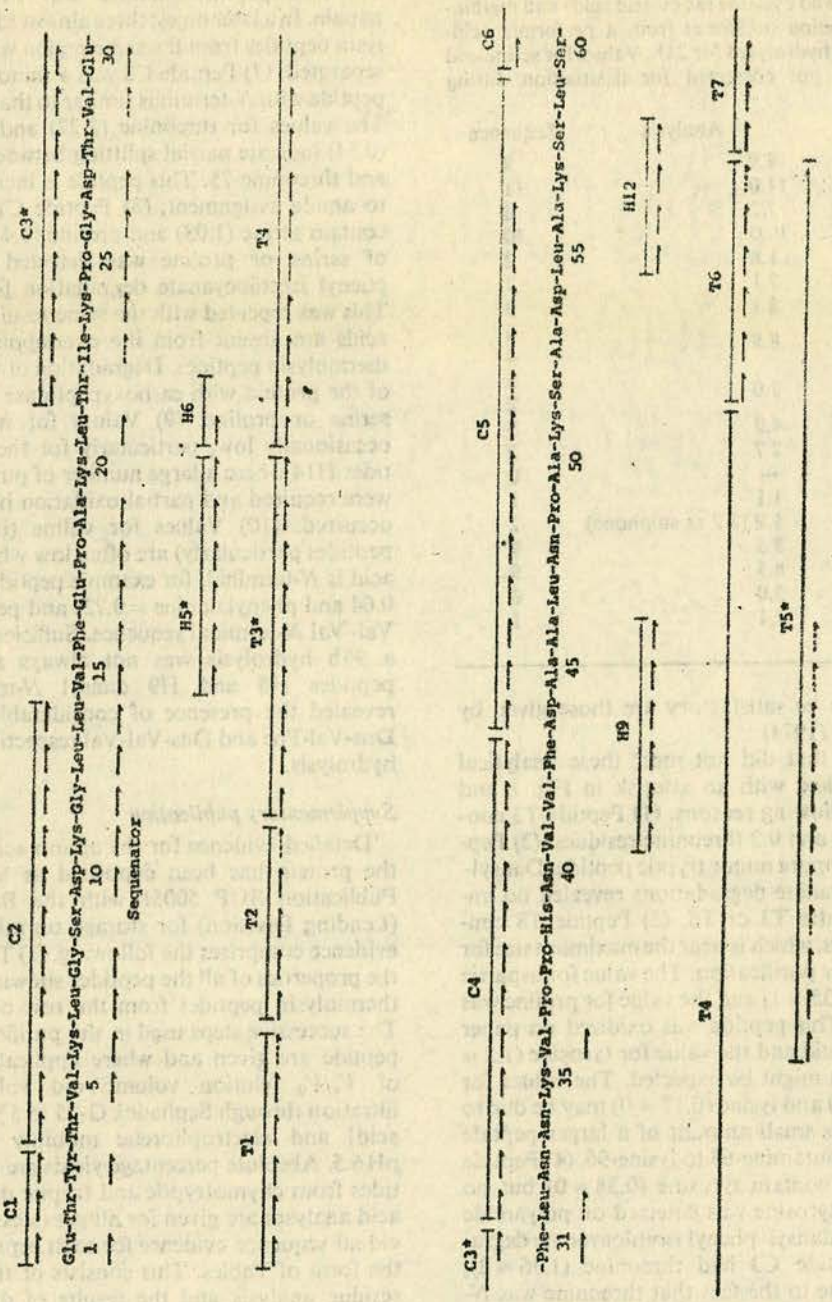
	Analysis	Sequence
Glycine	8.9	9
Alanine	11.0	11
Valine	7.2	8
Leucine	10.0	10
Isoleucine	1.8	2
Serine	7.1	7
Threonine	8.1	9
Aspartic acid	8.8	5
Asparagine		4
Glutamic acid	7.0	5
Glutamine		2
Phenylalanine	4.9	5
Tyrosine	2.7	3
Tryptophan	—	0
Cysteine	1.1	1
Methionine	1.9 (2.2 as sulphone)	2
Proline	8.5	9
Lysine	8.5	9
Histidine	3.0	3
Arginine	1.1	1

acid analyses to be satisfactory are those given by Kelly & Ambler (1974).

The peptides that did not meet these analytical criteria are marked with an asterisk in Fig. 1 and failed for the following reasons. (1) Peptide T3 contained 0.4 serine and 0.2 threonine residues. (2) Peptide T5 was an impure minor tryptic peptide. Dansyl–phenyl isothiocyanate degradations revealed no impurities in peptides T3 or T5. (3) Peptide T8 contained 31 residues, which is near the maximum size for satisfactory paper purification. The value for aspartic acid was high (1.35 = 1) and the value for proline was low (3.37 = 4). This peptide was oxidized on paper with performic acid and the value for tyrosine (1.2 = 2) is higher than might be expected. The values for glycine (2.55 = 2) and lysine (0.17 = 0) may be due to the presence of a small amount of a larger peptide which included glutamine-63 to lysine-96. (4) Peptide T9 appeared to contain tyrosine (0.38 = 0) but no 3- or bis-dansyltyrosine was detected on polyamide sheets or during dansyl–phenyl isothiocyanate degradation. (5) Peptide C3 had threonine (1.36 = 2), which may be due to the fact that threonine was *N*-terminal and that both Thr-Val and Thr-Ile bonds are present. After 96h hydrolysis 1.46 threonine residues were recovered. (6) Peptide C7 appeared to be pure except for the low value for leucine (0.55). The *N*-terminal amino acids were later proved to be leucine, methionine and serine in the subjective proportion of intensity 4:1:1. It appears that chymotrypsin also cleaved after leucine-65 and methionine-66. This peptide mixture was subdigested with papain. In a later digest three almost identical thermolysin peptides from the same region were successfully separated. (7) Peptide C8 was a minor chymotryptic peptide with *N*-terminus similar to that of peptide C7. The values for threonine (2.27) and phenylalanine (0.54) indicate partial splitting between threonine-74 and threonine-75. This peptide is included as an aid to amide assignment. (8) Peptide C11 appeared to contain serine (1.08) and proline (0.41) but no trace of serine or proline was detected after dansyl–phenyl isothiocyanate degradation for nine cycles. This was repeated with the same result. These amino acids are absent from the overlapping tryptic and thermolysin peptides. Degradation of the *C*-terminus of the protein with carboxypeptidase A revealed no serine or proline. (9) Values for methionine are occasionally low, particularly for thermolysin peptides H14 where a large number of purification steps were required and partial oxidation has presumably occurred. (10) Values for valine (in thermolysin peptides particularly) are often low where this amino acid is *N*-terminal, for example peptide H5 (valine = 0.64 and phenylalanine = 0.72) and peptide H9 with Val–Val *N*-terminal sequence. Sufficient material for a 96h hydrolysis was not always available. For peptides H5 and H9 dansyl *N*-terminal study revealed the presence of considerable amounts of Dns-Val-Phe and Dns-Val-Val respectively, after 16h hydrolysis.

Supplementary publication

Detailed evidence for the amino acid sequence of the protein has been deposited as Supplementary Publication SUP 50051 with the British Library (Lending Division) for storage on microfiche. This evidence comprises the following. (1) Tables showing the properties of all the peptides shown in Fig. 1 plus thermolysin peptides from the rest of the protein. The successive steps used in the purification of each peptide are given and where applicable the values of V_e/V_0 [elution volume/void volume for gel filtration through Sephadex G-25 in 5% (v/v) formic acid] and electrophoretic mobility on paper at pH6.5. Absolute percentage yields are given for peptides from chymotryptic and tryptic digests. Amino acid analyses are given for all peptides. (2) The individual sequence evidence for each peptide is given in the form of Tables. This consists of the *N*-terminal residue analysis and the results of dansyl–phenyl-isothiocyanate degradation. Details are given of substandard results. (3) Evidence for the presence or absence of amide groups on each aspartic acid and glutamic acid residue is given. Electrophoretic mobility


Glu-Thr-Tyr-Thr-Val-Lys-Leu-Gly-Ser-Asp-Lys-Gly-Leu-Leu-Val-Phe-Glu-Pro-Ala-Lys-Leu-Thr-Ile-Lys-Pro-Gly-Asp-Thr-Val-Glu-
1 5 10 15 20 25 30
Sequenator
-Phe-Leu-Asn-Asn-Lys-Val-Pro-Pro-His-Asn-Val-Val-Phe-Asp-Ala-Ala-Leu-Asn-Pro-Ala-Lys-Ser-Ala-Asp-Leu-Ala-Lys-Ser-Leu-Ser-
31 35 40 45 50 55 60
C1
C2
C3*
C4
C5
C6
T1
T2
T3*
T4
T5*
T6
T7
H5*
H6
H9
H12

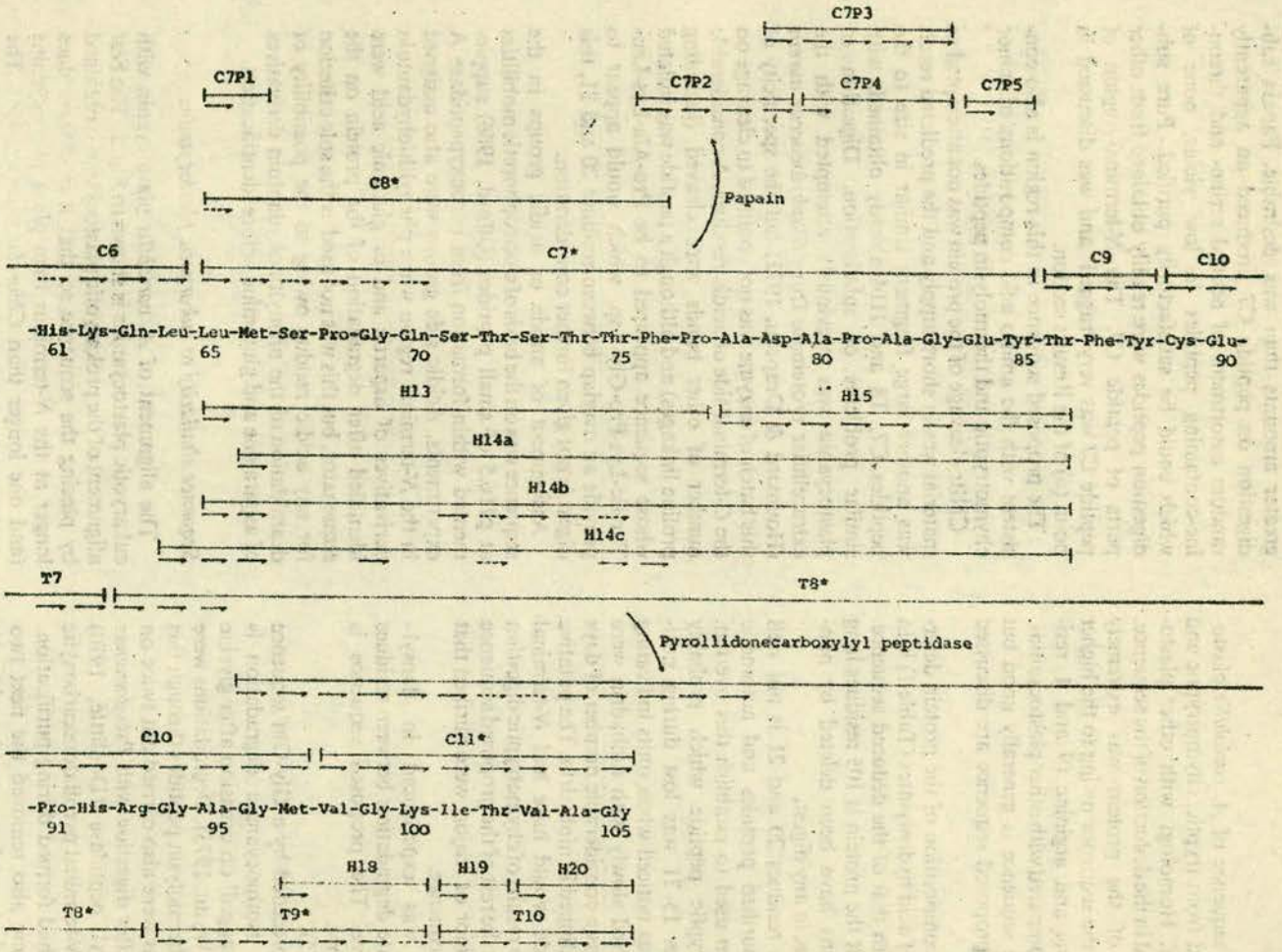


Fig. 1. ***Amino acid sequence of A. variabilis plastocyanin***

Peptides derived by chymotryptic digestion (C) are shown above the sequence and by tryptic (T) and thermolysin (H) digestion below the sequence. Vertical arrows show peptides formed by further digestion. Full lines indicate quantitative amino acid analyses (substandard if marked *). Under the sequence —› indicates residues identified after degradation of the protein in a sequenator. Under peptide lines —› indicates end groups and subsequent residues revealed by phenyl isothiocyanate degradation, and identified by the dansyl method (substandard if shown ···›). For clarity, only those thermolysin peptides required to complete overlaps are shown on the diagram.

evidence is based on Offord (1966). (4) Amino acid analysis obtained after degradation of the protein by carboxypeptidase A (Ambler, 1972) is given.

Discussion

The amino acid sequence of *A. variabilis* plastocyanin was deduced from tryptic, chymotryptic and thermolysin digests. Homology with other plastocyanins was not used in the deduction of the sequence. Tryptic digestion of the protein was extremely useful in deducing the sequence owing to the higher proportion of lysine and arginine (9 and 1 residues respectively) compared with other plastocyanins.

Evidence for the sequence is generally good but weaknesses in the proposed sequence are discussed below.

The amino acid composition of the protein determined by analysis of acid hydrolysates (Table 1) is in good agreement with that of the deduced sequence. This is evidence that the protein is 105 residues long and that no regions have been deleted by non-discovery of peptides in any digest.

Overlap between residues 20 and 21 is not well established from purified peptides and automatic sequencing has been used to establish this overlap. A major chymotryptic peptide which probably represented residues 15–21 was lost during purification. The loss was noticed when spots indicating a peptide that reacted slowly with ninhydrin were seen on marker strips a considerable number of days after they had been dipped in ninhydrin. The putative chymotryptic peptide would have had *N*-terminal valine. However, the results of chymotryptic digestion were consistent with the rest of the proposed sequence and no peptides, minor or major, were purified that did not fit into the sequence.

Some difficulty was experienced in dansyl–phenyl isothiocyanate degradation between residues 71 and 75 inclusive. The proposed sequence is

-Ser-Thr-Ser-Thr-Thr-.
(71) (75)

The sequence is preceded by a Gly-Gln sequence and dansyl–phenyl isothiocyanate degradation is known to become 'ragged' on occasion after glycine residues (Ambler & Wynn, 1973). Degradations were carried out on two thermolysin peptides through this region. Degradations were also carried out twice on tryptic peptide T8 after digestion with *Pseudomonas* pyrollidonecarboxylyl peptidase (Doolittle, 1970) which removed the *N*-terminal pyrollidonecarboxylic acid residue which had formed during purification. This batch of enzyme also removed the next two residues (both leucine), and a peptide with *N*-terminal methionine was produced. This peptide had a V_e/V_0 ratio (Sephadex G-25) and tyrosine content (bis-Dns-tyrosine identified on polyamide sheets) which were consistent with the parent peptide. The results of dansyl–phenyl isothiocyanate degradation were in all instances consistent with the proposed sequence although in one or two cases the amount of contaminant dansyl-amino acid was present in greater amounts than was desirable. Papain subdigestion on peptide C7 produced an apparently random assortment of neutral serine- and threonine-containing peptides in low yields, none of which could be satisfactorily purified. Pure subdigestion peptides were readily obtained from other parts of peptide C7. The *N*-terminal region of peptide C7 was very 'ragged' and was discussed in point (6) of the Results section.

The proposed sequence in this region is also consistent with the amino acid compositions of minor chymotryptic and thermolysin peptides.

CNBr cleavage of the protein was not attempted as material was in short supply and the predicted result was another large fragment similar in size to the peptides C7, T8 and H14 already obtained, with similar problems of subdigestion. Digestion of plastocyanin was previously attempted with the extracellular proteinase from *Staphylococcus aureus* (Houmard & Drapeau, 1972) but the specificity of this batch of enzyme was not confined to cleavage on the *C*-terminal side of acidic residues. A considerable number of other bonds were cleaved (including proline linkages) and although a peptide was isolated whose sequence appeared to be Pro-Ala-Lys-Leu-Thr-Ile-Lys-Pro-Gly-Asp, which would appear to provide an overlap between residues 20 and 21, this digest is not given further consideration.

Assignment of amide or acidic groups in the sequence was on the basis of electrophoretic mobilities at pH 6.5 of small peptides (Offord, 1966) supplemented within formation from carboxypeptidase A experiments. Acidic side groups were also assigned in the *N*-terminal region when phenylthiohydantoin derivatives of aspartic and or glutamic acid were identified after degradation of the protein on the sequenator, but this was not used as the sole criterion for any acidic residue owing to the possibility of deamidation of the phenylthiohydantoin derivatives of asparagine and glutamine before identification.

Sequence similarity to eukaryotic plastocyanins

The alignment of *A. variabilis* plastocyanin with eukaryotic plastocyanins is shown in Fig. 2. The best alignment of the prokaryotic plastocyanin is obtained by placing the sequence so that it is two residues longer at the *N*-terminus than plant plastocyanins (and one longer than *Chlorella* plastocyanin). The *Anabaena* plastocyanin is also one residue longer at the *C*-terminus and has insertions at positions 77, 78 and 81.

An alternative alignment with a gap in the *Chlorella* sequence at residues 50 and 51 instead of 59 and 60

```
                1                   5                  10                  15                  20                  25
(a) Glu-Thr-Tyr-Thr-Val-Lys-Leu-Gly-Ser-Asp-Lys-Gly-Leu-Leu-Val-Phe-Glu-Pro-Ala-Lys-Leu-Thr-Ile-Lys-Pro-Gly-Asp-
(b)     Asp-Val-Thr-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu-Val-Phe-Glu-Pro-Ser-Ser-Val-Thr-Ile-Lys-Ala-Gly-Glu-
(c)         Leu-Asp-Val-Leu-Leu-Gly-Gly-Asp-Asp-Gly-Ser-Leu-Ala-Phe-Ile-Pro-Gly-Asn-Phe-Ser-Val-Ser-Ala-Gly-Glu
(d)         Leu-Glu-Val-Leu-Leu-Gly-Ser-Gly-Asp-Gly-Ser-Leu-Val-Phe-Val-Pro-Ser-Glu-Phe-Ser-Val-Pro-Ser-Gly-Glu

                30                  35                  40                  45                  50
(a) Thr-Val-Glu-Phe-Leu-Asn-Asn-Lys-Val-Pro-Pro-His-Asn-Val-Val-Phe-Asp-Ala-Ala-Leu-Asn-Pro-Ala-Lys-Ser-Ala-Asp
(b) Thr-Val-Thr-Trp-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Ile-Val-Phe-Asp-Glu-Asp-Glu-Val-Pro-Ser-Ala-Gly-Asn-Ala
(c) Lys-Ile-Thr-Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Val-Val-Phe-Asp-Glu-Asp-Glu-Ile-Pro-Ala-Gly-Val-Asp-Ala
(d) Lys-Ile-Val-Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Val-Val-Phe-Asp-Glu-Asp-Glu-Ile-Pro-Ala-Gly-Val-Asp-Ala

                55                  60                  65                  70                  75                  80
(a) Leu-Ala-Lys-Ser-Leu-Ser-His-Lys-Gln-Leu-Leu-Met-Ser-Pro-Gly-Gln-Ser-Thr-Ser-Thr-Thr-Phe-Pro-Ala-Asp-Ala-Pro
(b) Glu-Ala-Leu-Ser-...-...-His-Glu-Asp-Tyr-Leu-Asn-Ala-Pro-Gly-Glu-Ser-Tyr-Ser-Ala-Lys-Phe-...-...-Asp-Thr-...
(c) Ser-Lys-Ile-Ser-Met-Ala-Glu-Glu-Asp-Leu-Leu-Asn-Ala-Ala-Gly-Glu-Thr-Tyr-Ser-Val-Thr-Leu-...-...-Ser-Glu-...
(d) Val-Lys-Ile-Ser-Met-Pro-Glu-Glu-Glu-Leu-Leu-Asn-Ala-Pro-Gly-Glu-Thr-Tyr-Val-Val-Thr-Leu-...-...-Asp-Thr-...

                85                  90                  95                 100                 105
(a) Ala-Gly-Glu-Tyr-Thr-Phe-Tyr-Cys-Glu-Pro-His-Arg-Gly-Ala-Gly-Met-Val-Gly-Lys-Ile-Thr-Val-Ala-Gly
(b) Ala-Gly-Thr-Tyr-Gly-Tyr-Phe-Cys-Glu-Pro-His-Gln-Gly-Ala-Gly-Met-Lys-Gly-Thr-Ile-Thr-Val-Gln
(c) Lys-Gly-Thr-Tyr-Thr-Phe-Tyr-Cys-Ala-Pro-His-Gln-Gly-Ala-Gly-Met-Val-Gly-Lys-Val-Thr-Val-Asn
(d) Lys-Gly-Thr-Tyr-Ser-Phe-Tyr-Cys-Ser-Pro-His-Gln-Gly-Ala-Gly-Met-Val-Gly-Lys-Val-Thr-Val-Asn
```

Fig. 2. *Alignment of* (*a*) *A. variabilis plastocyanin with* (*b*) *Chlorella fusca plastocyanin* (*Kelly & Ambler*, 1974), (*c*) *potato plastocyanin* (*Ramshaw et al.*, 1974) *and* (*d*) *French-bean plastocyanin* (*Milne et al.*, 1974)

Numbering of residues is not that used in the other publications. The best alignment of *A. variabilis* plastocyanin requires extra residues at the *N*- and *C*-termini, deletions in the eukaryotic sequences at positions 77, 78 and 81 and deletions in *Chlorella fusca* plastocyanin at positions 59 and 60.

Table 2. *Matrix of amino acid identities between plastocyanins*

The values for identities with plant plastocyanins are the range for all the published plant plastocyanin sequences. The number of identities between *Anabaena* and *Chlorella* plastocyanin sequences depends on the alignment (see the text).

	Anabaena	*Chlorella*	Higher plants
Anabaena	—	52 or 55	41–45
Chlorella	52 or 55	—	52–58
Higher plants	41–45	52–58	—

would generate three more identities between *Chlorella* and *Anabaena* plastocyanins. This would result in two [or three for elder (*Sambucus nigra*. L) plastocyanin (Scawen *et al.*, 1974)] fewer identities between *Chlorella* plastocyanin and the plant plastocyanins,

The highly acidic region from residue 44 to 47 in the eukaryotic plastocyanins is not present (except aspartate-44) in *A. variabilis* plastocyanin. A similarity matrix of identities between *Anabaena* plastocyanin, *Chlorella* plastocyanin and plant plastocyanins is given (Table 2).

Kelly & Ambler (1974) found some evidence for dimorphism at residue 84 in *Chlorella fusca* plastocyanin. The sequence amino acid is given as threonine (the incumbent residue in most plant sequences published) but a small proportion of glutamic acid may have been present. Residue 84 is glutamate in *Anabaena* plastocyanin.

Sequence similarity to bacterial azurins

The sequence similarity round the single cysteine residue in eukaryotic plastocyanin to the sequence round the free cysteine in bacterial azurins (Kelly & Ambler, 1974) is also 'conserved' in this prokaryotic plastocyanin. The degree of similarity is not, however, any greater than with the eukaryotic plastocyanins [see also Ramshaw *et al.* (1974) for alignment of plant plastocyanins with azurin]. Bacterial azurins also contain copper in a type I ligand environment and Katoh *et al.* (1961) have suggested that the thiol group is one of the copper ligands. The sequences of the two classes of protein do not show any real sign of divergence from a common ancestral sequence.

It has not been possible to detect phylogenetic relationships between blue-green algae and other bacteria in this study although the view is held (Olson, 1970) that blue-green algae and other photosynthetic bacteria shared a photoheterotrophic ancestor. The components necessary for oxygenic photosynthesis appear to have evolved only along the blue-green-algal line of descent. Blue-green algae and the gliding filamentous non-photosynthetic flexibacteria have many features in common (Soriano & Lewin, 1965) and the recently discovered 'photosynthetic flexibacteria' *Chloroflexus* are considered to be closer to blue-green algae than other groups of photosynthetic bacteria (Pierson & Castenholz, 1974).

The similarity in amino acid sequence between the prokaryotic plastocyanin and the eukaryotic plastocyanins is too great to reasonably suggest convergent evolution. The plastocyanin and cytochrome *f* (Ambler & Bartsch, 1975) sequence comparisons both strongly support an evolutionary relationship between prokaryotic blue-green algae and eukaryotic algae. The plastocyanin sequence comparisons presented here enable one to consider the position of higher plants in prokaryote–eukaryote relationships, although it is not possible to say whether or not the implied evolutionary relationships relate solely to the chloroplasts of the eukaryotes. Table 2 shows that the prokaryote plastocyanin sequence is closer to the eukaryotic algal sequence than to the plant plastocyanin sequences. Before too much is postulated from these data, it will be necessary to evaluate the amount of sequence variation between different extant blue-green algal groups. Preliminary sequencing studies on plastocyanin from *Phormidium luridum* (A. Aitken, unpublished work) suggest that the amino acid sequences of both prokaryotic plastocyanins are highly conserved.

Sequence studies on cytochromes *f* from a wide range of blue-green algae (Ambler & Bartsch, 1975; A. Aitken, unpublished work) suggest that these photosynthetic proteins have, *inter alia*, about 70–80% of their residues identical.

It remains to be seen if sequence comparisons can distinguish whether these photosynthetic eukaryotes have descended as a whole from a blue-green-algal ancestor or if the chloroplast alone has evolved from such a prokaryote as proposed by the endosymbiotic theory (Marguilis, 1970). Sequence comparisons of proteins involved in eukaryotic cytoplasmic and mitochondrial functions with the corresponding blue-green-algal proteins would be useful in eliminating some of the possibilities. Superoxide dismutase may prove a useful candidate in this respect

I am indebted to Dr. R. P. Ambler for invaluable advice and assistance and thank Miss S. Murray and Dr. D. Pritchard for performing sequenator runs, Dr. N. Carr for supplying the blue-green algal strain, Mr. T. Savill for preparing enzymes and The Medical Research Council for the award of a Scholarship.

References

Ambler, R. P. (1972) *Methods Enzymol.* **25**, 262–272
Ambler, R. P. & Bartsch, R. G. (1975) *Nature (London)* **253**, 285–288
Ambler, R. P. & Brown, L. H. (1967) *Biochem. J.* **104**, 784–825

Ambler, R. P. & Wynn, M. (1973) *Biochem. J.* **131**, 485–498
Biggins, J. (1967) *Plant Physiol.* **42**, 1447–1456
Bishop, N. I. (1971) *Annu. Rev. Biochem.* **40**, 197–226
Doolittle, R. F. (1970) *Methods Enzymol.* **19**, 555–569
Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
Edman, P. & Sjoquist, J. (1956) *Acta Chem. Scand.* **10**, 1507–1509
Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822
Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3506–3509
Inglis, A. S., Nicholls, P. W. & Roxburgh, C. M. (1971) *Aust. J. Biol. Sci.* **24**, 1247–1250
Katoh, S. (1960) *Nature (London)* **186**, 533–534
Katoh, S., Suga, I., Shiratori, I. & Takamiya, A. (1961) *Arch. Biochem. Biophys.* **94**, 136–141
Katoh, S., Shiratori, I. & Takamiya, A. (1962) *J. Biochem. (Tokyo)* **51**, 32–40
Kelly, J. & Ambler, R. P. (1974) *Biochem. J.* **143**, 681–690
Kratz, W. A. & Myers, J. (1955) *Am. J. Bot.* **42**, 282–287
Lightbody, J. J. & Krogmann, D. W. (1967) *Biochim. Biophys. Acta* **131**, 508–515
Malkin, R. & Malmström, B. G. (1970) *Adv. Enzymol. Relat. Areas Mol. Biol.* **33**, 177–244
Marguilis, L. (1970) *Origin of Eukaryotic Cells*, Yale University Press, New Haven
Milne, P. R., Wells, J. R. E. & Ambler, R. P. (1974) *Biochem. J.* **143**, 691–701
Offord, R. E. (1966) *Nature (London)* **211**, 591–593
Olson, J. M. (1970) *Science* **168**, 438–445
Pierson, B. K. & Castenholz, R. W. (1974) *Arch. Mikrobiol.* **100**, 5–24
Ramshaw, J. A. M., Scawen, M. D., Bailey, C. J. & Boulter, D. (1974) *Biochem. J.* **139**, 583–592
Scawen, M. D. & Boulter, D. (1974) *Biochem. J.* **143**, 257–264
Scawen, M. D., Ramshaw, J. A. M., Brown, R. H. & Boulter, D. (1974) *Eur. J. Biochem.* **44**, 299–303
Schopf, J. W. (1970) *Biol. Rev.* **45**, 319–352
Soriano, S. & Lewin, R. A. (1965) *Antonie van Leeuwenhoek* **31**, 66–80
Stanier, R. Y., Kunisawa, R., Mandel, M. & Cohen-Bazire, G. (1971) *Bacteriol. Rev.* **35**, 171–205
Stewart, W. D. P. & Pearson, H. W. (1970) *Proc. R. Soc. Ser. B* **175**, 293–311
Visser, J. W. M., Amesz, J. & Van Gelder, B. F. (1974) *Biochim. Biophys. Acta* **333**, 279–287