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 structure and absence of a nuclear membrane, blue-green algae are prokaryotic organisms (Stanier *et al.*, 1971). Their photosynthetic apparatus is, however, oxygenic and resembles that of the chloroplast much more closely than that of photosynthetic bacteria. There is much fossil evidence to indicate that blue-green algae are an ancient group of organisms (Schopf, 1970) which probably evolved from a photo-heterotrophic ancestor common to all types of photosynthetic bacteria (Olson, 1970). It has been suggested that blue-green algae were the first oxygen-evolving photosynthetic organisms and were initially responsible for producing atmospheric O₂ (during the Precambrian period). Some blue-green algae can grow anaerobically (Stewart & Pearson, 1970) and atmospheric O₂ is a prerequisite for the evolution of eukaryotes, which normally depend on mitochondrial respiration for their energy requirements.

The present study is an attempt to obtain some insight into the evolutionary connexion between prokaryotes and eukaryotes by comparing the amino acid sequence of plastocyanin from the blue-green alga *Anabaena variabilis* with plastocyanin sequences from eukaryotes.

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

Plastocyanin is widely distributed in the chloroplasts of plants (Katoh *et al.*, 1961). The protein has been isolated from green algae (Katoh, 1960) and two 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* (Kützinger) was a gift from N. G. Carr, Department of Biochemistry, University of Liverpool, Liverpool, U.K.

Pyroglutaminyl carboxylase 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% CO₂ in air. NaHCO₃ (0.8 g/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 200 g wet wt. of packed cells.

The cells (200 g wet wt.) were suspended in 1 litre of potassium phosphate buffer (0.05 M, pH 7) and broken in a Manton-Gaulin homogenizer. The suspension was made 40% (w/v) in (NH₄)₂SO₄ and the precipitate was centrifuged. This step removed much of the phycobiliproteins. The supernatant was made 90% (w/v) in (NH₄)₂SO₄ and left at 4°C for 2 days to ensure complete precipitation. The precipitate was centrifuged at 10000g for 30 min, redissolved in the minimum volume of distilled water, and then passed down a column (30 cm × 6 cm) of Sephadex G-25 in 1 M-potassium phosphate (pH 7). The plastocyanin was further purified by the methods of Lightbody & Krogmann (1967) which entails DEAE- and CM-cellulose chromatography. The yield of plastocyanin from 200 g 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 0.1 M-NH₃.

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 deriva-

tives identified on polyamide layer sheets (Hartley, 1970). Identification of histidine and arginine dansylated derivatives was by high-voltage paper electrophoresis at pH 4.4. 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 1.8 μmol of denatured plastocyanin; chymotryptic digestion and thermolysin digestion on 2 μmol each of denatured apoprotein.

Automatic phenylisothiocyanate degradation. Native oxidized plastocyanin (0.8 μmol) was degraded in a Beckmann, model 890A, sequenator by using a 70 min 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 5.5 M-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		5
Glutamine	7.0	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 *O*- 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 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 given in the form of Tables. This consists of the *N*-terminal residue analysis and the results of dansyl-phenylisothiocyanate 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

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* pyroglutaminidase (Doolittle, 1970) which removed the *N*-terminal pyroglutaminidase-resistant 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-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-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-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-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-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).

	<i>Anabaena</i>	<i>Chlorella</i>	Higher plants
<i>Anabaena</i>	—	52 or 55	41–45
<i>Chlorella</i>	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 (Margulis, 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
- Margulis, 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., Ames, J. & Van Gelder, B. F. (1974) *Biochim. Biophys. Acta* **333**, 279–287