A sixth subunit of ATP synthase, an F₀ component, is encoded in the pea chloroplast genome

Alison L.Cozens, John E.Walker, Andrew L.Phillips^{1,2}, Alison K.Huttly^{1,3} and John C.Gray¹

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, and ¹Botany School, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

²Present address: Rothamsted Experimental Station, Harpenden AL5 2JQ, UK

³Present address: Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, UK

Communicated by J.E.Walker

DNA from pea and tobacco chloroplasts hybridizes specifically with probes derived from the gene for a membrane component, the *a* subunit of ATP synthase of the cyanobacterium, *Synechococcus* 6301. DNA sequence of the hybridizing region of the pea plastid DNA has revealed that it encodes a protein of 247 amino acids related in sequence to the *a* subunits of ATP synthase of *Escherichia coli*, *Synechococcus* and mitochondria. This is the sixth component of chloroplast ATP synthase that is plastid coded. The gene is located upstream from the genes for three other ATP synthase subunits and a transcript of 6 kb contains coding sequences from each of these genes. Thus the subunit *a* gene is part of a co-transcribed cluster of four ATP synthase genes arranged in the order *a*:*c*(or *III*):*b*(or *I*):*α*. Two other ATP synthase genes, those for *β* and *ε* subunits, are known to form a separate cluster. These gene arrangements are most closely related to those found in the cyanobacterium, *Synechococcus* 6301. Hence, this finding provides strong evidence for a common origin for cyanobacteria and plant chloroplasts.

Key words: chloroplast/ATP synthase/F₀ subunit/plastid coded

Introduction

The ATP synthases of bacteria and chloroplasts (and also mitochondria) are very similar protein complexes. They are composed of an extrinsic domain, F₁-ATPase, and an intrinsic membrane domain, F₀, which is involved in proton-translocation across the membrane. The *Escherichia coli* F₁-ATPase contains five subunits named *α*, *β*, *γ*, *δ* and *ε*, and chloroplast enzymes are composed of homologues of these proteins (reviewed by Walker and Tybulewicz, 1986). The *E. coli* F₀ complex comprises three different subunits, *a*, *b* and *c*; all of them have been implicated in proton translocation (Klionsky *et al.*, 1983; Schneider and Alten-dorf, 1985; von Meyenburg *et al.*, 1985).

The composition of the chloroplast F₀ complex is less clearly defined. Pick and Racker (1979) detected four subunits in spinach chloroplast F₀, but subsequently it was suggested that chloroplast ATP synthase contains only three F₀ subunits (Nelson *et al.*, 1980; Pick, 1982); these were named subunits I, II and III. Subunit I (19 kd) has been shown to be the chloroplast equivalent of *E. coli* subunit *b* (Bird *et al.*, 1985) and subunit III (8 kd) is the homologue of subunit *c* (Sebald and Hoppe, 1981; Howe *et al.*, 1982a). Subunit II (15 kd) has not been characterized suf-

ficiently to allow comparison to be made with bacterial or mitochondrial subunits but the gene for this subunit is located in the nuclear genome (Westhoff *et al.*, 1985). The presence of a fourth subunit in the chloroplast F₀ complex has been demonstrated by electrophoresis in the presence of urea; it has a mol. wt of 19 kd (Pick and Racker, 1979; Süß, 1980; Westhoff *et al.*, 1985). The work described here shows that this subunit is encoded in chloroplast DNA and that it is equivalent to the *E. coli* subunit *a*. The experiments upon which these conclusions are based depend upon the availability of a gene for another *a* subunit that has been cloned

Fig. 1. Physical and genetic maps of the pea chloroplast genome. (A) Map showing restriction sites and direction of transcription of genes (shown by arrows). Gene names: 16S and 23S, rRNAs (Bowman *et al.*, 1985; Palmer and Thompson, 1981); P700 is the P700/chlorophyll *a* apoprotein of photosystem I (Smith and Gray, 1984) 32 kd, 44 kd, D2, cyt b559 are components of photosystem II (see Dyer, 1985); cyt *f*, cyt b563 and 15 kd are subunits of the cytochrome *b-f* complex (Willey *et al.*, 1984; Phillips and Gray, 1984). LS is the large subunit of ribulose biphosphate carboxylase (Bowman *et al.*, 1981; Oishi and Tewari, 1983). *a*, I, III *α*, *β* and *ε* are components of ATP synthase (Huttly *et al.*, 1984). (B) Map showing *Hind*III and *Eco*RI sites in the vicinity of the ATP synthase genes. The gene for ATP synthase subunit I contains an intron (Bird *et al.*, 1985).

and sequenced from the cyanobacterium, *Synechococcus* 6301 (A.L.Cozens and J.E.Walker, unpublished work). Restriction fragments of chloroplast DNA from tobacco and pea hybridize specifically with probes derived from this gene. DNA sequence analysis shows that the hybridizing fragment in pea chloroplast DNA contains a sequence encoding a protein which is homologous to bacterial α subunits and to the equivalent subunit of the mitochondrial complex. The gene is located on the 5' side of the cluster of genes for ATP synthase subunits III, I and α , respectively, and is co-transcribed with these genes. The genes for the β and ϵ subunits cluster at another locus in the pea chloroplast genome (Huttly and Gray, 1984a, 1984b). Thus, six of the nine constituent proteins of the pea (and presumably at least tobacco) chloroplast ATP synthase are encoded in the plastid DNA.

Results and Discussion

DNA hybridizations

Digests of chloroplast DNA from tobacco and pea were hybridized with 'prime-cut' probes derived from the gene for the α subunit of ATP synthase of *Synechococcus* 6301. Specific hybridization was obtained with fragments of the tobacco and pea DNA, respectively. In pea, hybridization was obtained with the cluster of *Pst*I fragments of ~ 12 kb and with a 2.4-kb *Hind*III fragment.

From the restriction map of the pea chloroplast genome it seemed likely that the hybridization was to a region near the genes for subunits III, I and α (see Figure 1). Accordingly a 6.2-kb *Pst*I-*Sal*I fragment containing the 2.4-kb *Hind*III fragment was

Fig. 2. DNA sequence of a segment of the pea chloroplast genome. It encodes the α subunit and part of the c subunit of ATP synthase. The boxed sequences are potential ribosome binding sites. Asterisks denote termination codons. Also shown is an *Eco*RI site used to prepare probes for transcriptional mapping (see Figure 7).

isolated from a plasmid clone, pPscS6, containing the chloroplast fragment.

DNA sequence

The DNA sequence of the 6.2-kb fragment was investigated by the random strategy. It soon became apparent that the 3' end of this fragment contained part of the gene for ATP synthase subunit

Fig. 3. Comparison of the protein sequence of the *a* chloroplast subunit with those of equivalent subunits in bovine mitochondria, *E. coli* and *Synechococcus* 6301. The comparisons are made with the aid of the computer program DIAGON (Staden, 1982b) with a score of 290 and a window of 25 amino acids. One division on the axes represents 50 amino acids.

III, and also DNA sequences encoding a protein homologous to the *a* subunits from *E. coli* (Gay and Walker, 1981) and *Synechococcus* 6301 (A.L. Cozens and J.E. Walker, unpublished work). Hence, the sequencing experiments were focussed on this region of the fragment.

The sequence of 1792 bp of DNA presented in Figure 2 ends at the *Pst*I site of pPscS6. This DNA has an A + T content of 67.45% in keeping with values of 61–67.3% for the pea chloroplast DNA overall (Herrmann and Possingham, 1980). The open reading frames (potential genes) could be readily detected in this sequence from the distribution of stop codons which occur frequently on both strands of the DNA in non-coding regions and in all three of the DNA phases. Analysis of this sequence using the positional base preference method (Staden, 1984) confirmed their locations. (This method depends upon uneven use of amino acids by known proteins and does not take codon preference into account. In this case it works particularly well as the genes detected, in common with other A:T-rich chloroplast genes, show a marked preference for A or T in the third position of codons.) The sequence contains an open reading frame encoding a hydrophobic protein of 247 amino acids with a mol. wt of 27.5 kd. It is followed at the 3' end by a partial gene (96 bp) also encoding a hydrophobic protein. This is subunit *c* (or III) of pea chloroplast ATP synthase (Huttlly and Gray, 1984b). Passavant and Hallick (1985) have reported 38 codons of an open reading frame upstream from the chloroplast gene for subunit III in *Euglena gracilis*. Part of the *Euglena* open reading frame corresponds to nucleotides 904–932 of the pea open reading frame, suggesting that a similar gene is present in *Euglena* chloroplast DNA.

Identity of the encoded protein

Comparison of the sequence of the protein of 247 amino acids

SUBUNIT <i>a</i>		1	50
Pea chloroplast		...MNVLLCY	INTLNRFYDI SAVEVGQHFY WQIGDFQVHA
<i>Synechococcus</i> 6301		MGSAT LPSDLMSNPT	LLELSSVLPL AELEVGQHFY WQIGNYRLHG
<i>E. coli</i>		MASENMTPQD	YIGHHLNNLQ LDLRTFSLVD PQNPATFWT INIDSMFFSV
Beef mitochondrion			..MNEMLTFSF ITPVILGLP.
Pea chloroplast	51	QVLITSWVVI	AILLISTILV VRNPQTIPS GQNFPEYVLE FIRDVSKTQI
<i>Synechococcus</i> 6301		QVFLTSWFVI	AALVVLSELLA NENLQRIPSG LQNFMEYVLD FIRNLARTQI
<i>E. coli</i>		VLGLLFLVLF	RSVAKKATSG VPGKFQTAIE LVIGFVNGSV KDMYHGKSKL
Beef mitochondrion		..LVTILIVLF	P...SLLFPTS NRLVSNRFVT LQQWMLQLVS KQWMSIHNSK
Pea chloroplast	101	GE.EYGPWVP	FIGTLFLFIF VSNWSGALLP WKIIKLPHGE LAAPTNDINT
<i>Synechococcus</i> 6301		GEKEYRPWVP	FIGTLFLFIF LSNWSGALIP WKLIKLPSEGE LAAPTSNDINT
<i>E. coli</i>		IA.PLALTIF	VWVFLNNLMD LLPIDLLPYI AEHVLGLPAL RVVPSADVNV
Beef mitochondrion		GQ.TWTL...	MLNSLILFIG STNLLGLLPH SFTPTTQLSM NLGMAIPLWA
Pea chloroplast	151	TVALALLTSV	AYFYAGISKE GLAYFGKYI. .QPT..... PILLPINILE
<i>Synechococcus</i> 6301		TVALALLTSL	AYFYAGFSEK GLGYFGNYV. .HPT..... PVMLPFKILE
<i>E. coli</i>		TLSMALGVFI	LILFYSIKMK GIGGFTKELT LQPFNHWAFI PVNLILEGVS
Beef mitochondrion		GA.....	. VITGFRRK TKASLANFLP QGTPTPLIPM LVIIETISLF
Pea chloroplast	201	DFTKPLSLSF	RLFGNILLDE LVVVVLVSL. VPLVV
<i>Synechococcus</i> 6301		DFTKPLSLSF	RLFGNILLDE LVVAVLVLL. VPLFV
<i>E. coli</i>		LLSKPVSGL	RLFGNMYAGE LIFILIAGL. LPWVS
Beef mitochondrion		IQPM...AV	RLTANITAGE LILHIGGAT LALWSISTTT ALITFTILIL
Pea chloroplast	251	PIPVNMLGL.	FTS...GIRA LIFATLAAAT IGESMEGHH
<i>Synechococcus</i> 6301		PLPANILGL.	FTS...AIRA LIFATLAAAT IGEAVEEHGE EHAE
<i>E. coli</i>		QWILNVPWAI	PHILIIITLQA FIFMVLITIVY LSMASEEH
Beef mitochondrion		LT...ILE..	FAY...AMIRA YVFLLVSLY LHDNT
			294

Fig. 4. Alignment of the protein sequences of ATP synthase *a* subunits from various sources. The bovine sequence is from Anderson *et al.* (1982), the *E. coli* sequence from Gay and Walker (1981) and the *Synechococcus* sequence from Cozens and Walker (unpublished work).

Fig. 5. Comparison of the hydrophobicity profiles of pea chloroplast ATP synthase *a* subunit with equivalent subunits from other sources. The calculation was made with the computer program HYDROPLLOT (Kyte and Doolittle, 1982) with a window of 11 amino acids. The horizontal bar indicates an average hydrophobicity.

with sequences of subunits *a* of ATP synthase from mitochondria (Anderson *et al.*, 1982) and *E. coli* (Gay and Walker, 1981) shows that the pea protein is significantly related to these sequences. The regions of greatest homology are confined to the C-terminal half of the protein (see Figure 3) as noticed previously with other *a* subunits (Walker *et al.*, 1984). The protein is much more closely related to the *Synechococcus* 6301 *a* subunit of ATP synthase, 70% of residues being identical (A.L.Cozens and J.E. Walker, unpublished work, see Figure 4). A more extensive relationship between these proteins is shown by comparison of their hydrophobic profiles (Figure 5). These are very similar and fit the suggested secondary structure of six or seven transmembrane helices (Walker *et al.*, 1984). These sequences and secondary structure relationships strongly suggest that the chloroplast protein plays a similar role to its homologues. The relative molecular masses of the *a* subunits from mitochondria and *E. coli* are ~30 kD, based on the gene sequences, but the bacterial protein migrates in SDS-polyacrylamide gel electrophoresis with mobilities corresponding to a polypeptide of ~25 kD. The significant homologies between the chloroplast, mitochondrial and bacterial *a* subunits (Figures 3–5) suggest that the chloroplast *a* subunit will also show abnormal electrophoretic mobility. It seems probable that the 20-kD polypeptide synthesized in the coupled transcription-translation system (Figure 6) is the chloroplast *a* subunit, and that this corresponds to the fourth F_0 subunit (~19 kD) separated by electrophoresis in the presence of urea (Pick and Racker, 1979; Süß, 1980; Westhoff *et al.*, 1985).

Transcription

Complex patterns of hybridization were obtained when gene-specific 'prime-cut' probes for the pea genes for subunits *a*, II and α were hybridized to pea chloroplast RNA that had been subjected to agarose gel electrophoresis and transferred to nylon membranes (Figure 7A). The probes for all three genes hybridized to a 6-kb RNA species indicating that the genes were co-transcribed. In addition, transcripts of 3.4, 2.9 and 2.3 kb hybridized only to the probe for the *a* subunit gene. Major

Fig. 6. Products of coupled transcription-translation of the 2.4-kb *Hind*III fragment in a cell-free system from *E. coli*. (lane a) no added DNA; (lane b) total products of pUC8; (lane c) total products of 2.4-kb *Hind*III fragment in pUC8; (lane d) immunoprecipitation of (c) with non-immune serum; (lane e) immunoprecipitation of (c) with antibodies to a 17.5-kD polypeptide of wheat chloroplast ATP synthase. The sizes (in kD) of marker proteins are shown on the left.

Fig. 7. Transcription of pea gene for subunit *a*. **(A)** Northern blot analysis of transcripts of genes for subunits *a*, and α III, in pea chloroplast RNA. **(a)** Hybridization with 248-bp fragment (nucleotides 443–690) internal to gene for subunit *a*; **(b)** hybridization with 136-bp *Hae*III-*Hind*III fragment internal to gene for subunit III; **(c)** hybridization with 327-bp *Hind*III fragment internal to gene for subunit α . Sizes of bands are in kb. **(B)** S1 nuclease protection of a 580-bp *Eco*RI fragment by pea chloroplast RNA. A 'prime-cut' probe was produced by *Hind*III digestion of double-stranded DNA synthesized from the 580-bp *Eco*RI fragment inserted in M13 tg 131. This contains an additional 33 bp of primer and vector sequence at the 5' end and an extra 44 bp of vector sequence at the 3' end. **(a)** Probe without 5' nuclease treatment; **(b)** S1 nuclease treatment of probe without hybridization to RNA; **(c)** S1 nuclease treatment of probe hybridized to pea chloroplast RNA.

transcripts of 5.2, 4.2 and 3.5 kb hybridized to probes for the genes for subunits III and α , similar to the transcripts for the corresponding wheat genes (Bird *et al.*, 1985).

To investigate if the 5' end of any of these RNA species was located in the region of the gene for subunit *a*, S1 nuclease protection experiments were carried out. A 'prime-cut' probe containing the 580-bp *Eco*RI fragment was hybridized to pea chloroplast RNA and DNA protected from S1 nuclease digestion was separated by electrophoresis (Figure 7B). The presence of a band of ~430 bp indicated a 5' end at position 26–30 of Figure 2. However, the absence of identifiable promoter sequences in this region suggests that this is an RNA processing site. Full-length protection of the probe by chloroplast RNA also suggests that the major transcript of the subunit *a* gene initiates upstream of the sequence shown in Figure 2.

Gene order

The *E. coli unc* operon encodes all eight subunits of its ATP synthase (Walker *et al.*, 1984). Within the operon are two sub-clusters of genes corresponding to the F_1 and F_0 components of the enzyme (see Figure 8). Clusters of genes for F_1 subunits have been found in purple non-sulphur photosynthetic bacteria and have the same order as in the *E. coli* F_1 subcluster (Tybulewicz *et al.*, 1984; Falk *et al.*, 1985). By analogy with genes in the morphogenetic loci of coliphages it was proposed that this gene order might influence the assembly of the enzyme complex (Walker *et al.*, 1984). It now seems more likely that these clusters are related to the evolution of the enzyme complex, and that gene orders in these loci can be used to trace evolutionary relationships between micro-organisms (Walker and Cozens, 1986). It is thus of particular interest in relation to the endosymbiotic origin of chloroplasts that the order of the six ATP synthase genes in the pea chloroplast is clearly related to these gene orders. It is particularly significant that the chloroplast gene orders are most closely related to the gene order of ATP synthase genes in *Synechococcus* 6301, a finding that strongly supports a common origin for plant chloroplasts and cyanobacteria.

Materials and methods

Isolation of chloroplast DNA fragment

Plasmid pPscS6 containing the 10.6-kb *Sal*I fragment of pea chloroplast DNA, inserted in pBR322 (Hutty and Gray, 1984a) was digested with *Sal*I and *Pst*I. A 6.2-kb fragment derived from the inserted DNA was purified by electrophoresis in LGT agarose (Wieslander, 1979).

DNA hybridization

Pea or tobacco chloroplast DNA (1 μ g) was digested for 1 h with *Pst*I or *Hind*III. The digest was loaded onto an agarose gel (200 ml) and was subjected to electro-

Fig. 8. Arrangements of genes encoding ATP synthase subunits in bacteria and chloroplasts. The scale is in kb. The letters *a*, *b*, *c*, α , β , δ , and ϵ indicate the ATP synthase subunit encoded in the gene. *I* in the *E. coli unc* operon and *X* in the *R. blastica* operon encode proteins of unknown function (Gay and Walker, 1981; Tybulewicz *et al.*, 1984). Chloroplast *c* subunit is subunit III and chloroplast *b* is subunit I; *b* contains an intron (Bird *et al.*, 1985). *b'* in *Synechococcus* 6301 is a duplicated and diverged form of *b* (A.L.Cozens and J.E.Walker, unpublished work). The dashed lines in *Synechococcus* and chloroplasts signify that the two gene clusters are far apart and separately transcribed.

phoresis (30 mA, 14 h). DNA was transferred from the gel to nitrocellulose according to Southern (1975). Filters were pre-hybridized at 65°C for 30 min in a solution (pH 6.2) containing 0.7 M sodium chloride, 0.2 M disodium hydrogen phosphate, 7 mM EDTA, Ficoll (1 g/l), bovine serum albumin (fraction V, 1 g/l), polyvinylpyrrolidone (1 g/l), 0.005% sodium azide, 0.5% Sarkosyl and yeast RNA (2 mg/l). Hybridizations were carried out overnight at 65°C in more of the same solution, to which had been added 10% dextran sulphate and single-stranded 'prime-cut' probes (Farrell *et al.*, 1983).

Hybridization probes were prepared from three M13mp8 phages as described by Biggin *et al.* (1983). These contained fragments of the *Synechococcus* 6301 *a* subunit gene (A.L.Cozens and J.E.Walker, unpublished). They cover amino acid residues corresponding to the pea *a* protein as follows: (i) 13–95, (ii) 101–105, (iii) 162–244. Filters were washed three times, for 20 min each, at 65°C in a solution containing 0.3 M sodium chloride and 0.03 M sodium citrate and then autoradiographed for 48 h.

DNA sequencing

The nucleotide sequence of the 6.2-kb *Pst*I-*Sal*I fragment was determined using the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) as modified by Biggin *et al.* (1983). Random fragments of 300–800 bp long were generated by sonication and cloned into M13mp8 that had been previously cut with *Sma*I (Bankier and Barrell, 1983). To complete the sequence, a 583-bp *Pvu*II fragment (nucleotides 1205–1786) was also cloned into the *Hind*III site of M13mp18.

Data analysis

DNA sequences were entered into a database with DBAUTO (Staden, 1982a). Each nucleotide position in the sequence was determined six times on average and at least once on each strand of the DNA. The DNA sequence was analysed with ANALYSEQ (Staden, 1984); a variety of methods were used to predict protein-coding regions. The predicted protein sequence was analysed with ANALYSEP (Staden, unpublished work), and compared with known ATP synthase subunit sequences with the aid of DIAGON (Staden, 1982b).

Transcriptional analysis

Chloroplasts were prepared from 8-day-old pea plants by sucrose density gradient centrifugation of crude homogenates as described by Phillips (1985), except that ethidium bromide was replaced in all solutions by 10 mM vanadyl ribonucleoside complexes. RNA was extracted from the chloroplasts (Koller *et al.*, 1982) and was further purified by precipitation with 2 M lithium chloride. The purified chloroplast RNA was redissolved in 0.15 M sodium acetate, pH 6.0, containing 0.1% SDS. It was precipitated with 2.5 vol of ethanol and stored at –20°C.

RNA samples for gel electrophoresis were prepared as follows. A sample of an ethanol suspension of RNA was centrifuged. The pellet was washed with 70% ethanol, dried *in vacuo* and redissolved at a concentration of 2 µg/µl in 100 mM sodium phosphate, pH 7.0, containing 100 µM aurin tricarboxylic acid. Then the RNA was reacted with glyoxal by the addition of 4 vol of glyoxal/formamide (30% deionized glyoxal: deionized formamide, 1:3.5, by vol) followed by incubation at 50°C for 15 min. Samples were applied to a 1% agarose gel in 10 mM sodium phosphate, pH 7.0, containing 1.5 mM iodoacetic acid. The gel was run submerged in 10 mM sodium phosphate, pH 7.0, for 3–4 h at 250 V. RNA was transferred from the gel to Biotodyne A nylon membranes (Thomas, 1980).

Chloroplast DNA fragments (internal to ATPase genes), inserted in M13mp8, mp18 or mp19, were labelled with [³²P]dATP by primer extension from the hybridization primer (Huttlly and Gray, 1984a). Hybridization to Northern blots was carried out according to Huttlly and Gray (1984a) except that probes were not heat-denatured before use. Excess labelled DNA was removed from blots by washing three times, each for 5 min at 15°C with 2 × SSC containing 0.1% SDS, and twice, each for 15 min at 50°C in 0.1 × SSC containing 0.1% SDS (SCC, 150 mM NaCl/15 mM sodium citrate buffer, pH 8.5).

S1 nuclease analysis was carried out as described by Bird *et al.* (1985), using pea chloroplast RNA and a 'prime-cut' probe produced by *Hind*III digestion of double-stranded DNA synthesized from the 580-bp *Eco*RI fragment inserted in M13 tg 131.

Coupled transcription-translation

Coupled transcription-translation of the 2.4-kb *Hind*III fragment of pea chloroplast DNA, inserted in pUC8, in a cell-free system from *E. coli* strain PR7 was carried out as described by Howe *et al.* (1982b). Immunoprecipitations with antibodies to [³⁵S]methionine-labelled ATP synthase polypeptides and protein A-Sepharose and electrophoresis on 15% polyacrylamide gels in the presence of SDS were carried out as described by Howe *et al.* (1982b). Radioactive bands were detected by fluorography. Antibodies to polypeptides of the wheat and spinach ATP synthases isolated as described by Bird *et al.* (1985) were a gift from C.R. Bird.

Acknowledgements

We thank D.G.Zimmerman and S.Purton for gifts of chloroplast DNA. A.L.C. is supported by an MRC Research Studentship, A.K.H. was supported by an SERC

research grant and A.L.P. was supported by a Fellowship from the Broodbank Fund.

References

- Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. (1982) *J. Mol. Biol.*, **156**, 638–717.
- Bankier, A.T. and Barrell, B.G. (1983) In Flavell, R.A. (ed.), *Techniques in Nucleic Acid Biochemistry*. Elsevier Scientific Publishers, Ireland, pp.B508/1–B508/31.
- Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.
- Bird, C.R., Koller, B., Auffret, A.D., Huttlly, A.K., Howe, C.J., Dyer, T.A. and Gray, J.C. (1985) *EMBO J.*, **4**, 1381–1388.
- Bowman, C.M., Koller, B., Delius, H. and Dyer, T.A. (1981) *Mol. Gen. Genet.*, **183**, 93–101.
- Falk, G., Hampe, A. and Walker, J.E. (1985) *Biochem. J.*, **228**, 391–407.
- Farrell, P.J., Deininger, P.L., Bankier, A. and Barrell, B.G. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1656–1659.
- Gay, N.J. and Walker, J.E. (1981) *Nucleic Acids Res.*, **9**, 3919–3926.
- Herrmann, R.G. and Possingham, I.V. (1980) In Reinert, J. (ed.), *Chloroplasts*. Springer, Berlin, pp. 45–96.
- Howe, C.J., Auffret, A.D., Doherty, A., Bowman, C.M., Dyer, T.A. and Gray, J.C. (1982a) *Proc. Natl. Acad. Sci. USA*, **79**, 6903–6907.
- Howe, C.J., Bowman, C.M., Dyer, T.A. and Gray, J.C. (1982b) *Mol. Gen. Genet.*, **186**, 525–530.
- Huttlly, A.K. and Gray, J.C. (1984a) *Mol. Gen. Genet.*, **194**, 402–409.
- Huttlly, A.K. and Gray, J.C. (1984b) In Sybesma, C. (ed.), *Advances in Photosynthesis Research*, Martinus Nijhoff/Dr W.Junk, The Hague, Vol.IV, pp. 563–566.
- Huttlly, A.K., Howe, C.J., Bowman, C.M., Dyer, T.A. and Gray, J.C. (1984) *Biochem. Soc. Trans.*, **12**, 270–271.
- Klionsky, D.J., Brusilow, W.S.A. and Simoni, R.D. (1983) *J. Biol. Chem.*, **258**, 10136–10143.
- Koller, B., Delius, H. and Dyer, T.A. (1982) *Eur. J. Biochem.*, **122**, 17–23.
- Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.
- Nelson, N., Nelson, H. and Schatz, G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1361–1364.
- Oishi, K.K. and Tewari, K.K. (1983) *Mol. Cell Biol.*, **3**, 587–595.
- Palmer, J.D. and Thompson, W.F. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 5533–5537.
- Passavant, C.W. and Hallick, R.B. (1985) *Plant Mol. Biol.*, **4**, 347–354.
- Phillips, A.L. (1985) *Curr. Genet.*, **10**, 147–152.
- Phillips, A.L. and Gray, J.C. (1984) *Mol. Gen. Genet.*, **194**, 477–484.
- Pick, U. (1982) In Edelman, M., Hallick, R.B. and Chua, N.-H. (eds), *Methods in Chloroplast Molecular Biology*. Elsevier, Amsterdam, pp. 873–880.
- Pick, U. and Racker, E. (1979) *J. Biol. Chem.*, **254**, 2793–2799.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Schneider, E. and Altendorf, K. (1985) *EMBO J.*, **4**, 515–518.
- Seibald, W. and Hoppe, J. (1981) *Curr. Top. Bioenerget.*, **12**, 1–64.
- Southern, E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
- Smith, A.G. and Gray, J.C. (1984) *Biochem. Soc. Trans.*, **12**, 272–273.
- Staden, R. (1982a) *Nucleic Acids Res.*, **10**, 4731–4751.
- Staden, R. (1982b) *Nucleic Acids Res.*, **10**, 2951–2961.
- Staden, R. (1984) *Nucleic Acids Res.*, **12**, 521–538.
- Süss, K.-H. (1980) *FEBS Lett.*, **112**, 255–259.
- Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5204.
- Tybuliewicz, V.L.J., Falk, G. and Walker, J.E. (1984) *J. Mol. Biol.*, **179**, 185–214.
- von Meyenburg, K., Jorgensen, B.B., Michelson, O., Sorensen, L. and McCarthy, J.E.G. (1985) *EMBO J.*, **4**, 2357–2363.
- Walker, J.E. and Cozens, A.L. (1986) *Chem. Scrip.*, in press.
- Walker, J.E., Saraste, M. and Gay, N.J. (1984) *Biochim. Biophys. Acta*, **768**, 164–200.
- Walker, J.E. and Tybuliewicz, V.L.J. (1986) In Arntzen, C., Bogorad, L., Bonitz, S. and Steinback, K. (eds), *The Molecular Biology of the Photosynthetic Apparatus*. Cold Spring Harbor Laboratory Press, New York, in press.
- Westhoff, P., Alt, J., Nelson, N. and Herrmann, R.G. (1985) *Mol. Gen. Genet.*, **199**, 290–299.
- Wieslander, L. (1979) *Anal. Biochem.*, **98**, 305–309.
- Wiley, D.L., Phillips, A.L. and Gray, J.C. (1984) *Biochem. Soc. Trans.*, **12**, 271–272.

Received on 18 November 1985; revised on 13 December 1985

Note added in proof

The rest of the 6.2-kb *Sal*I-*Pst*I fragment encodes homologues of *E. coli* ribosomal subunit S2 and part of the β'-subunit of RNA polymerase [Cozens, A.L. and Walker, J.E. (1986) *Biochem. J.*, in press.]