Mol Gen Genet (1986) 203:117–128

MGG

Chloroplast ATP synthase of spinach contains nine nonidentical subunit species, six of which are encoded by plastid chromosomes in two operons in a phylogenetically conserved arrangement

Juliane Hennig and Reinhold G. Herrmann

Botanisches Institut der Universität Düsseldorf, D-4000 Düsseldorf, Federal Republic of Germany

Summary. We describe the nucleotide sequence of a 4.0 kilobase pair (kbp) region of the spinach plastid chromosome that encodes three subunits of the ATP synthase CF_0 sector proximal to the gene for CF_1 subunit alpha. The four genes are located on the same strand and form a transcriptional unit. Our study presents details about the genes and their products, resolves ambiguities in the constitution of the CF_0 sector, and establishes precise gene-polypeptide relationships as well as intersubunit homologies in ATP synthases from diverse organisms. The amino acid sequences deduced exhibit substantial homology (30%–40%) with published sequences for the F_0 subunits *a*, *c* (proteolipid) and *b* from *E. coli* and mitochondria, which therefore can be presumed to be structurally similar and functionally equivalent. Sequence comparison shows that the counterpart of CF_0–I is F_0–*b*. The homologue to the bacterial F_0–*a* is CF_0–IV, the product of a newly found gene, designated *atpI*. The data confirm our previous suggestion that the gene for subunit I, *atpF*, is split. The gene arrangement is *atpI* (247 codons) – 692 bp spacer – *atpH* (proteolipid, 81 codons) – 395 bp spacer – *atpF* (exon 1, 145 bp) – 764 bp intron class II – *atpI* (exon 2, 410 bp) – 65 bp spacer – *atpA* (alpha). The detection of *atpI* implies that the spinach CF_0 sector contains four, the entire ATP synthase (CF_0–CF_1) nine different subunit species. The extra subunit in the membrane sector is the nuclear-encoded component CF_0–II which has no known counterpart in *E. coli*. The six ATP synthase genes that are plastid-encoded are organized in two operons that map 40 kbp away from each other on the organelle chromosome. These transcriptional units, designated *atp* operons *A* and *B*, differ in their functional organisation. The plastid gene arrangement corresponds to the arrangement of genes in the *atp* (*unc*) operon of *E. coli* substantiating the theory that ATP synthase complexes have evolved from a common ancestor.

Key words: ATP synthase CF_0 genes – Plastid DNA – Class II intron – Evolution – Spinach

Introduction

The energy-transducing ATP synthase is one of the major multisubunit protein complexes of photosynthetic and respiratory membranes in chloroplasts, mitochondria and prokaryotic cells, coupling proton translocation across the membrane with the synthesis of ATP. The complex from all sources can be resolved into two functional parts, namely the membrane sector F_0 or CF_0 that forms the proton channel, and the F_1 or CF_1 assembly exposed to cytosol or organelle matrix that catalyses the terminal step in oxidative phosphorylation or photophosphorylation. ATP synthases have been extracted with mild detergents, purified by density gradient centrifugation and described as being composed of eight or more non-identical subunit species depending on the organism studied (reviewed in Nelson 1981; Senior and Wise 1983; Walker et al. 1985).

The structural genes for all eight subunits of the *Escherichia coli* complex have been sequenced and found to form a transcriptional unit, designated the *atp* or *unc* operon (Gay and Walker 1981a, b; Nielsen et al. 1981; Saraste et al. 1981). Eukaryotic ATP synthases are of dual genetic origin. The CF_1 of chloroplast ATP synthases, like bacterial and mitochondrial F_1-ATP synthases, comprises uniformly five subunits ($\alpha-\varepsilon$) with a likely subunit stoichiometry of 3:3:1:1:1 (Nelson 1981; Senior and Wise 1983; Walker et al. 1985). Three of these, namely α, β and ε are encoded in the organelle. Their genes have been mapped in two regions of the spinach chromosome, some 40 kbp away from each other (Westhoff et al. 1981) and sequenced (Zurawski et al. 1982; Rudolph, Herrmann and Bottomley, unpublished data). Subunits γ and δ derive from nuclear genes and are synthesized on cytosolic ribosomes as precursors, before import into the organelle (Westhoff et al. 1981, 1985; Herrmann et al. 1983). No species-specific differences in the distribution of the genes for CF_1 subunits have as yet been detected (cf. Huttly and Gray 1984) in contrast to those for mitochondrial ATP synthases (Leaver and Gray 1982).

The characterization of the CF_0 sector has encountered greater difficulties than that of CF_1, and consequently this protein complex has not yet been studied in comparable detail. The exact number of subunits, their stoichiometry, the functional equivalence to the F_0 subunits from other organisms and, to some extent, the intracellular location of the genes have remained largely obscure (Nelson 1981; Senior and Wise 1983; Walker et al. 1985). Four polypeptide species with approximate molecular masses of 19, 18 (CF_0–I), 16 (CF_0–II) and 8 kd (CF_0–III, proteolipid) have been reported (Pick and Racker 1979; Herrmann et al. 1983; Westhoff et al. 1985) but a structure composed of

Offprint requests to: R.G. Herrmann, Botanisches Institut der Ludwig-Maximilians-Universität, Menzinger Str. 67, D-8000 München 19, FRG

three types of subunit is generally preferred in analogy to the F_0 sector of the *E. coli* complex (Nelson 1981; Walker et al. 1985).

Studies involving compartment-specific inhibitors of protein synthesis have indicated that subunits I and III are synthesized on chloroplast ribosomes while CF_0–II originates in the cytosol (Nelson et al. 1980; Sebald and Hoppe 1981). We have been studying the organization of genes present on a 10.1 kbp *Xho*I fragment (*Xho*I-6) of spinach plastid DNA which bears the genes for the ATP synthase proteolipid (Alt et al. 1983), subunit CF_0–I (Herrmann et al. 1983; Westhoff et al. 1985) and subunit alpha (Westhoff et al. 1981) in the order given. We have also shown that subunit CF_0–II is made as a precursor from cytosolic (polyadenylated) RNA, 7 kd larger than the mature protein, and imported into the organelle (Herrmann et al. 1983; Westhoff et al. 1985). DNA-programmed transcription-translation indicated that the gene for CF_0–I in spinach is split (Herrmann et al. 1983), unlike most other plastid-encoded proteins in higher plants including ATP synthase subunits. Unfortunately, this finding could not be regarded as unequivocal since hybrid release translation experiments, like the DNA-programmed assays, exclusively yielded truncated CF_0–I polypeptide as well (Westhoff et al. 1985). Only when the cell-free assays were supplemented with charged bean chloroplast tRNAs was a virtually complete polypeptide obtained, indicating that proper decoding of this component might have been obscured by the lack of minor tRNA species in the heterologous systems.

To resolve this ambiguity we have investigated the fine structure of this DNA segment. In the course of this work a reading frame was detected, proximal to the proteolipid gene, which apparently encodes an additional ATP synthase subunit. We present here our analysis of a 4.0 kbp segment of that DNA region of the spinach plastid chromosome, adjacent to the 5′ end of the gene for subunit alpha, and give the complete nucleotide sequence and deduced amino acid sequence of the genes for three CF_0 subunits. Our findings help to clarify the organization of the gene for CF_0–I, the polypeptide composition of the CF_0 sector of spinach ATP synthase, as well as to establish precise relationships between subunits of the CF_0 complex and subunits of the corresponding bacterial and mitochondrial enzymes. Our analysis reinforces the idea that photosynthetic and respiratory electron transport chains in chloroplasts eubacteria and mitochondria possess common phylogenetic roots (cf. Heinemeyer et al. 1984; Cramer et al. 1985; Walker et al. 1985).

Materials and methods

All chemicals were of highest available purity. Restriction endonucleases were purchased from Boehringer (Mannheim) or New England Biolabs (Bad Schwalbach) and used according to the manufacturers specifications. Boehringer was the source for the DNA polymerase I large fragment (Maniatis et al. 1982) and Amersham (Braunschweig) for ^{32}P-αdNTP.

The recombinant plasmids pWHsp306 and pWHsp210, carrying the 10.1 kbp fragment *Xho*I–6 and the 2.4 kbp fragment *Sal*I–10 of spinach plastid DNA in pWHsp300 (Alt et al. 1983) and pBR322, respectively, were isolated from shot-gun libraries and have been described in detail (Alt et al. 1983; Westhoff et al. 1985). The *Eco*RI secondary fragments of *Xho*I–6 were subcloned in pBR325. Three of the resulting hybrid plasmids, designated pWHsp306/E1 (insert; the largest *Eco*RI subfragment of 5.2 kbp), pWHsp306/E2 (1.6 kbp), pWHsp306/E4 (580 bp) were used in this study. Plasmid DNA was prepared by the alkaline lysis procedure of Birnboim and Doly (1979) and purified by one or two cycles of isopycnic CsCl-ethidium bromide centrifugations.

Restriction of DNA, polyacrylamide and agarose gel electrophoresis and recovery of DNA fragments from gels were performed as described previously (Herrmann et al. 1980; Alt et al. 1983; Westhoff et al. 1983). Appropriate fragments, terminally labelled by fill-in synthesis of 3′ protruding ends (Wu 1970), were sequenced using the procedure of Maxam and Gilbert (1980). The sequence data were stored and processed as described (Morris and Herrmann 1984; Sijben-Müller et al. 1985).

Results

The genes for the subunits alpha, proteolipid and CF_0–I of the thylakoid membrane ATP synthase have been mapped to a 3.6 kbp region within the 10.1 kbp fragment *Xho*I–6 of the plastid chromosome from spinach (Westhoff et al. 1981, 1985; Alt et al. 1983; Herrmann et al. 1983, 1985). They have been identified on the basis of specific immunoprecipitation of polypeptides synthesized in vitro in DNA-programmed transcription-translation or hybrid release translation experiments using appropriate recombinant DNAs. The relative locations and transcription polarities of these genes as well as relevant restriction sites are shown in Fig. 1.

The 4.0 kbp nucleotide sequence of the interval bordered by the restriction sites *Eco*RI and *Hph*I (Fig. 1) was established on both strands using the chemical modification and chain cleavage procedure of Maxam and Gilbert (1980). The sequencing flow diagram is shown in Fig. 1. The overall nucleotide sequence and the amino acid sequences deduced are depicted in Fig. 2. Computer-assisted translation of the sequences disclosed seven open reading frames, greater than 50 nucleotides, five on one and two on the other strand. In accord with the previous mapping data, the former could readily be identified by sequence homology with polypeptides of ATP synthases from *E. coli* and mitochondria.

Gene 1 begins with an ATG (methionine) codon at position 1 and terminates with a TGA-terminator at position 744 (Fig. 2). If translated, the product would be a 247 residue hydrophobic polypeptide of MW 27,060. The homology search and the fact that the methionine residue is preceded by an AG-rich sequence complementary to the 3′ end of the 16S rRNA of plastids (Schwarz and Kössel 1980; Tohdoh and Sugiura 1982) and *E. coli* (Shine and Dalgarno 1974) suggest that the tentatively assigned ATG codon is the most probable start of the structural gene. Comparison of the amino acid sequence deduced with published sequence data revealed significant homology with the F_0 subunit *a* from *E. coli* (Gay and Walker 1981 b) and the equivalent ATP synthase-6 gene product of mouse, yeast and *Aspergillus* mitochondria (Macino and Tzagaloff 1980; Bibb et al. 1981; Grisi et al. 1982). The alignment of these sequences with the spinach sequence is shown in Fig. 3. The homology persists throughout the protein chain. There are 28% identical residues, an additional 13% conservative re-

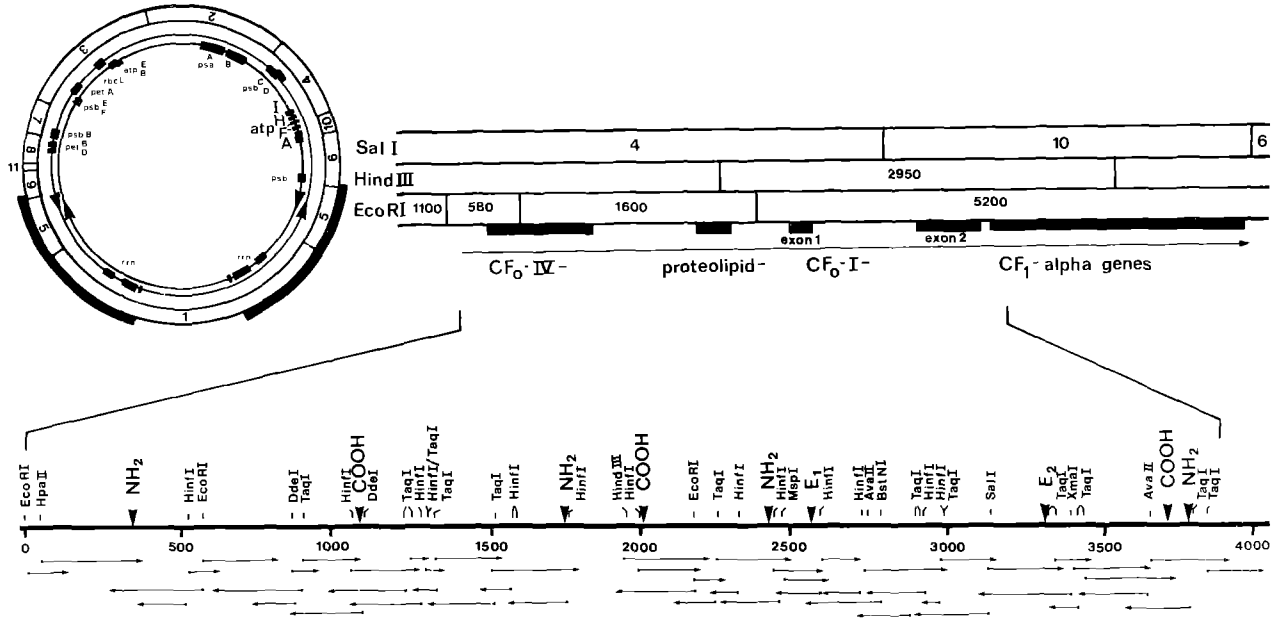

Fig. 1. Restriction map and sequence assay strategy of the region of the spinach plastid chromosome encoding ATP synthase subunits CF_0–I, –III, (proteolipid), –IV and CF_1-alpha. The positions of the genes and their direction of transcription (arrow) is indicated in the above map. *Arrows* commencing from *solid dots* in the *lower* map indicate direction and extent of the individual DNA sequence assays from the given restriction sites. Numbering is in base pairs. For sequencing of the proteolipid gene see Alt et al. (1983). The locations of the amino termini as well as of the carboxy termini are marked NH_2 and COOH, respectively, E_1 and E_2 refer to exon/intron boundaries of *atpF*. *Insert*: Simplified restriction cleavage site map of the spinach plastid chromosome showing the restriction sites of *Sal*I (Herrmann et al. 1980). The inverted repeat is marked by *bold lines*. The locations *atpA*, *F*, *H* and *I* as well as the positions of the genes for ribosomal RNAs (*rrn*; Whitfeld et al. 1978), ribulose bisphosphate carboxylase/oxygenase (*rbcL*, Herrmann and Possingham 1980; Zurawski et al. 1981) and various thylakoid membrane proteins including those for ATP synthase subunits beta and epsilon (*atpB* and *E*; Westhoff et al. 1981; Herrmann et al. 1985) are shown. The transcription polarities are indicated by arrow heads in the inner circles

placements when compared to the *E. coli* protein, and the predicted secondary structures (Chou and Fasman 1978; Garnier et al. 1978; Kyte and Doolittle 1982) are remarkably similar (Fig. 4). The hydropathy profile suggests at least four positionally conserved segments of high local hydrophobicity, each about 20 amino acid residues in length, which can traverse the membrane, and probably represent trans-membrane spans. Single or paired positively charged residues which flank these regions, and may interact with the surface charges of the lipid bilayer and stabilize spans, substantiate this view. We conclude that the product of this gene is the homologue of the *E. coli* subunit *a* of the spinach ATP synthase membrane sector. We designate this coding region *atpI* and its product CF_0 subunit IV (see Discussion).

Gene 2, separated by a 692 bp spacer, beginning at position 1437, and terminating at TAA position 1682 encodes a protein of molecular weight 8,000 (Fig. 2). The protein sequence deduced is identical to the amino acid sequence of the proteolipid (CF_0 subunit III or proton-translocating subunit, Sebald and Hoppe 1981). The data for this gene, *atpH*, have been published (Alt et al. 1983). This subunit plays a central role in proton translocation and is the counterpart to subunit *c* in the *E. coli* F_0 complex (Sebald and Hoppe 1981). More than 90% of its hydrophobic amino acids are clustered into two segments of 20–30 residues which possess the potential to form a trans-membrane span and include the glutamic acid residue at position 61 (nucleotides 1617–1689) involved in the catalytic process.

Gene 3: The sequence of the segment spanning *atpH* and *atpA* (Fig. 1), which should contain the gene for CF_0–I, is expectedly complex (Westhoff et al. 1985). It includes four reading frames greater than 50 amino acid residues, two on each strand, and the potential to form extensive secondary structure between them (Fig. 2). Although hybrid select translation experiments using single-stranded probes have shown that proteolipid, CF_0–I and subunit alpha mRNA is transcribed from the same DNA strand (Westhoff et al. 1985), in the region corresponding to this fragment there is no open reading frame of length sufficient to encode the expected 18,000 dalton CF_0–I subunit of spinach chloroplast ATP synthase. Reading frame 3 on the same strand is 56 codons long, begins at position 2078 with an ATG codon, terminates with a TAG stop at 2248 and would code for a protein of 6.0 kd. The following (fourth) reading frame of 145 codons is separated by 710 bp of non-protein-coding DNA and would form a polypeptide of molecular weight 16,841. Comparison of the amino acid sequence derived with those of the *E. coli* ATP synthase proteins showed that *both* spinach sequences can be aligned with bacterial ATP synthase subunit *b* (Fig. 5). The *E. coli* protein sequence is shorter at the amino terminus by 21 residues than the spinach protein (cf. also Fig. 6), but the alignment is unequivocal. There are more than 25% identical amino acid residues. If conservative and compensating amino acid changes are allowed for, the homology is ca. 33%. We infer from this result that both reading frames encode the functional equivalent for the bacterial ATP synthase subunit *b* and, consequently that the gene is split (Hennig 1984). The facts (a) that the intragenic untranslated

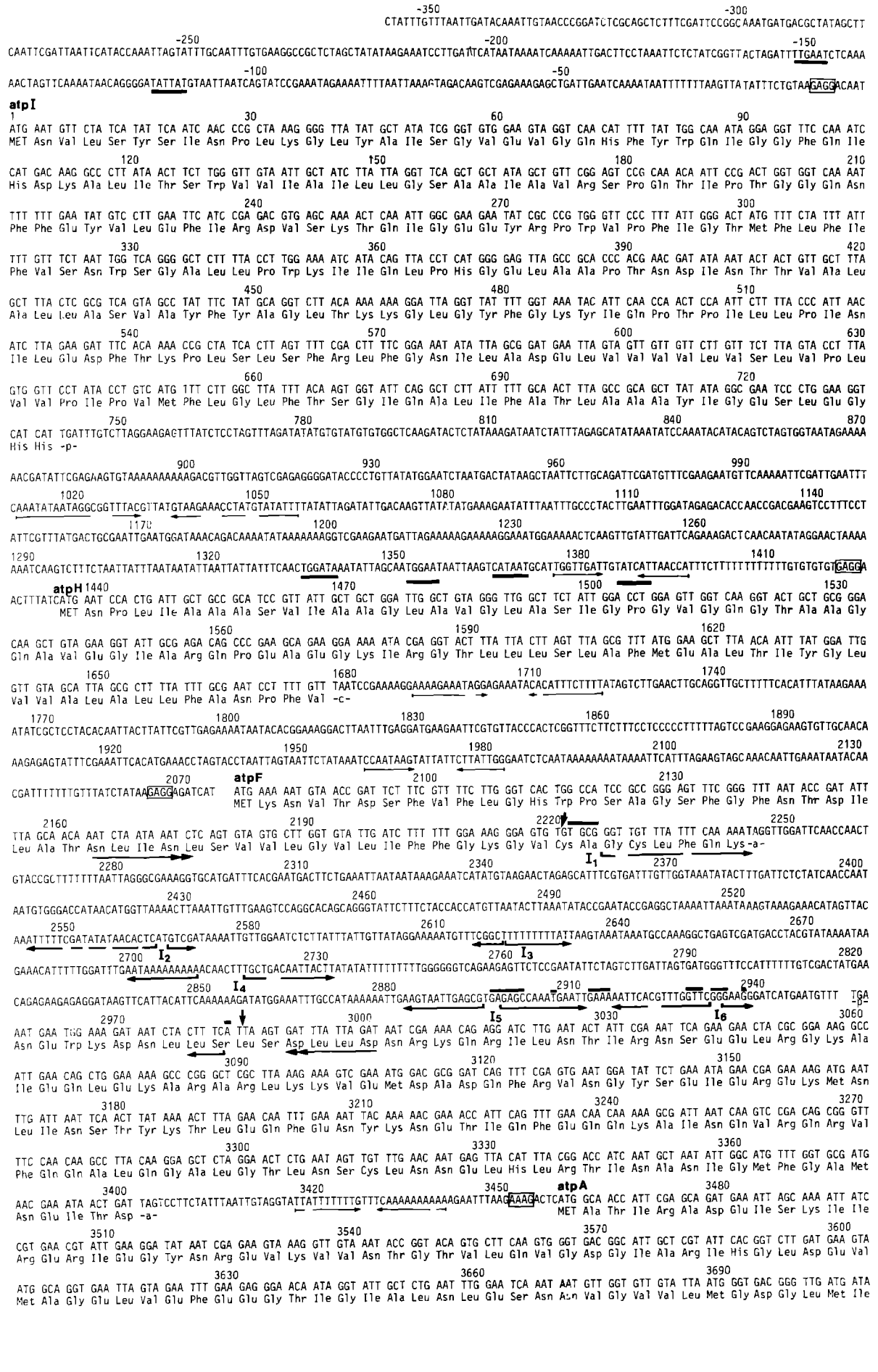
atpI
atpH
atpF
atpA
I1
I2
I3
I4
I5
I6

Fig. 3. Comparison of the amino acid sequence of the CF_0 subunit IV from spinach, F_0 subunit *a* from *E. coli* (Gay and Walker 1981a) and F_0 ATP synthase-6 subunit from *Aspergillus nidulans* (Grisi et al. 1982), *Saccharomyces cerevisiae* (Macino and Tzagaloff 1980) and mouse (Bibb et al. 1981) mitochondria. Sequences have been aligned to allow for maximal homology and are numbered according to the spinach protein. Homologous amino acid sequences are *boxed* and putative membrane spanning segments of the spinach protein are indicated by horizontal *brackets* (cf. Fig. 4). Conservatively substituted residues are not indicated

region displays all features typical of class II introns (Michel and Djon 1983, Figs. 2 and 7 and below), (b) that the recently determined N-terminal sequence of the spinach CF_0–I protein (Bird et al. 1985) matches the predicted amino acid sequence from exon 1 (starting at position 18), (c) that reading frame four (exon 2) does not begin with a methionine codon, (d) that a typical Shine-Dalgarno-like consensus sequence is not found in the region immediately upstream of or early within this frame (Fig. 2), and (e) the similarity of the primary and predicted secondary structures of the bacterial and chloroplast proteins in spite of substantial amino acid replacements (Fig. 6) are all consistent with this interpretation. Assuming translation initiates at the first available AUG codon of reading frame 3 (nucleotide 2078) and splicing occurs at the indicated positions (Figs. 2 and 7, see Discussion) the predicted protein has a molecular weight of 20.9 kd, similar to the estimated size of spinach CF_0–I (Westhoff et al. 1981, 1985). Hydropathy analysis suggests that the gene product, CF_0–I, would be a hydrophilic protein with a single N-terminal membrane span, as the *E. coli* subunit *b*. This coding region is designated *atpF* (cf. Hallick and Bottomley 1983).

The DNA sequence on the opposite strand displays two open reading frames (5 and 6) at 2510–2307 (68 codons) and 1845–1615 (77 codons overlapping *atpH* with 72 bp) which, if transcribed and translated, would code for proteins of molecular weights of 7.3 and 9.3 kd. Predictions of secondary structures suggest these proteins could not be intrinsic. We have not yet detected mRNA, the actual protein(s) or similarity to any other published component,

Fig. 2. DNA sequence of 4.0 kbp and derived amino acid sequences of the spinach plastid chromosome encoding three ATP synthase CF_0 subunits. The genes, designated *atpF*, *H*, *I* and *A* are believed to encode, respectively, CF_0-subunits I, III, IV (see text) and CF_1 subunit alpha. The sequence of the non-coding strand is presented. Protein coding regions are arranged in triplets. Numbering starts at the ATP triplet at which translation of subunit IV is probably initiated. Putative ribosome binding sites are boxed; possible "−10" and "−35" consensus sequences are underlined (*thick lines*); elements capable of forming stable hairpin structures analogous to those involved in transcription termination or processing in prokaryotic genes (Rosenberg and Court 1979) as well as short direct or inverted repeats are marked by *arrows*, respectively. Stem-loop structures I_1–I_6 (*thicker arrows*) designate secondary structures resembling those of class II introns (Michel and Dujon 1983). The *vertical arrows* at 2222 and 2986 mark the putative splicing site. Consensus sequences of class II introns are *overlined*; the 12 bp inverted repeat that may bring the splice site together is indicated by *double-headed arrows*

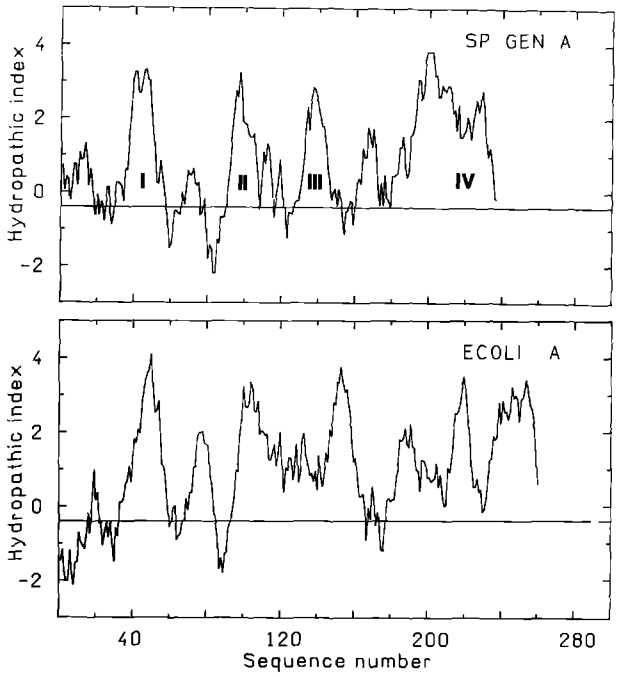

Fig. 4. Hydropathy blots of the spinach ATP synthase subunit CF_0–IV (*top*) and F_0 subunit *a* of *E. coli* (*bottom*) calculated according to Kyte and Doolittle (1982) with a 11-point moving window. Roman numbers indicate predicted membrane spans

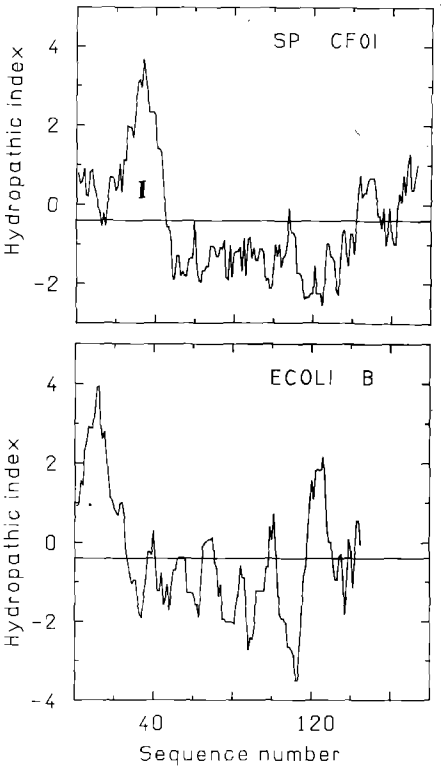

Fig. 6. Hydropathy blots of the spinach ATP synthase subunit CF_0–I (*top*) and the *E. coli* F_0 subunit *b* (*bottom*) calculated according to Kyte and Doolittle (1982) with a 11-point moving average. The single predicted membrane span is designated by the roman letter I. Note the *N*-terminal difference between the two proteins (see text)

and the reading frames are not preceded by typical ribosome binding sites. The role, if any, of this sequence is currently unknown.

Gene 4 (reading frame 7) begins at position 3462, 65 bp distal to the translational stop codon of *atpF*. This reading frame encodes the homologue of the *E. coli* α-subunit, confirming our previous mapping data (Westhoff et al. 1981; Alt et al. 1983). The sequence of this gene, *atpA*, will be detailed elsewhere (Rudolph, Herrmann and Bottomley, unpublished data).

Hybrid selection translation and Northern analysis, taken together with the DNA sequence data, suggest that *atpI*, *H*, *F* and *A* are expressed in vivo and lie in a single operon (cf. Alt et al. 1983; Westhoff et al. 1985). Blot analyses using size-fractionated spinach chloroplast RNA and appropriate strand-specific probes reveal multiple-banded patterns of overlapping RNA species ranging from more than 6 kb to less than 500 bp (cf. Fig. 7 in Westhoff et al. 1985 and unpublished data). The largest RNA species are

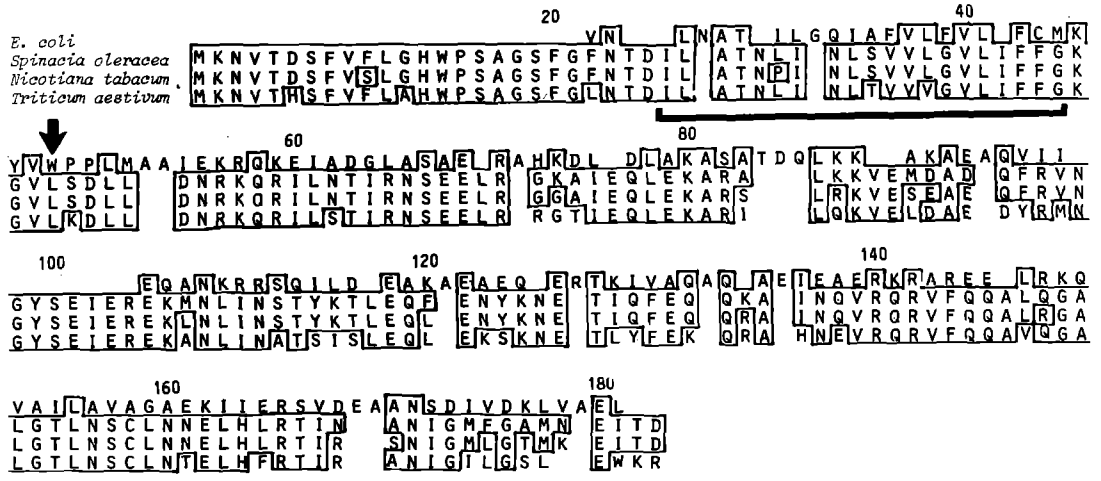

Fig. 5. Amino acid sequence alignment between CF_0–I deduced from the spinach wheat and tobacco plastid gene and F_0 subunit *b* of *E. coli*. Numbering is according to the spinach protein. The wheat sequence is from Bird et al. (1985), that from tobacco from Deno et al. (1984), and the *E. coli* sequence is from Gay and Walker (1981b). Boxes indicate perfect matches; conservatively replaced residues are not marked. The predicted single membrane span of the spinach protein is indicated by a horizontal bracket, the putative mRNA splice point by an arrow (cf. Fig. 6)

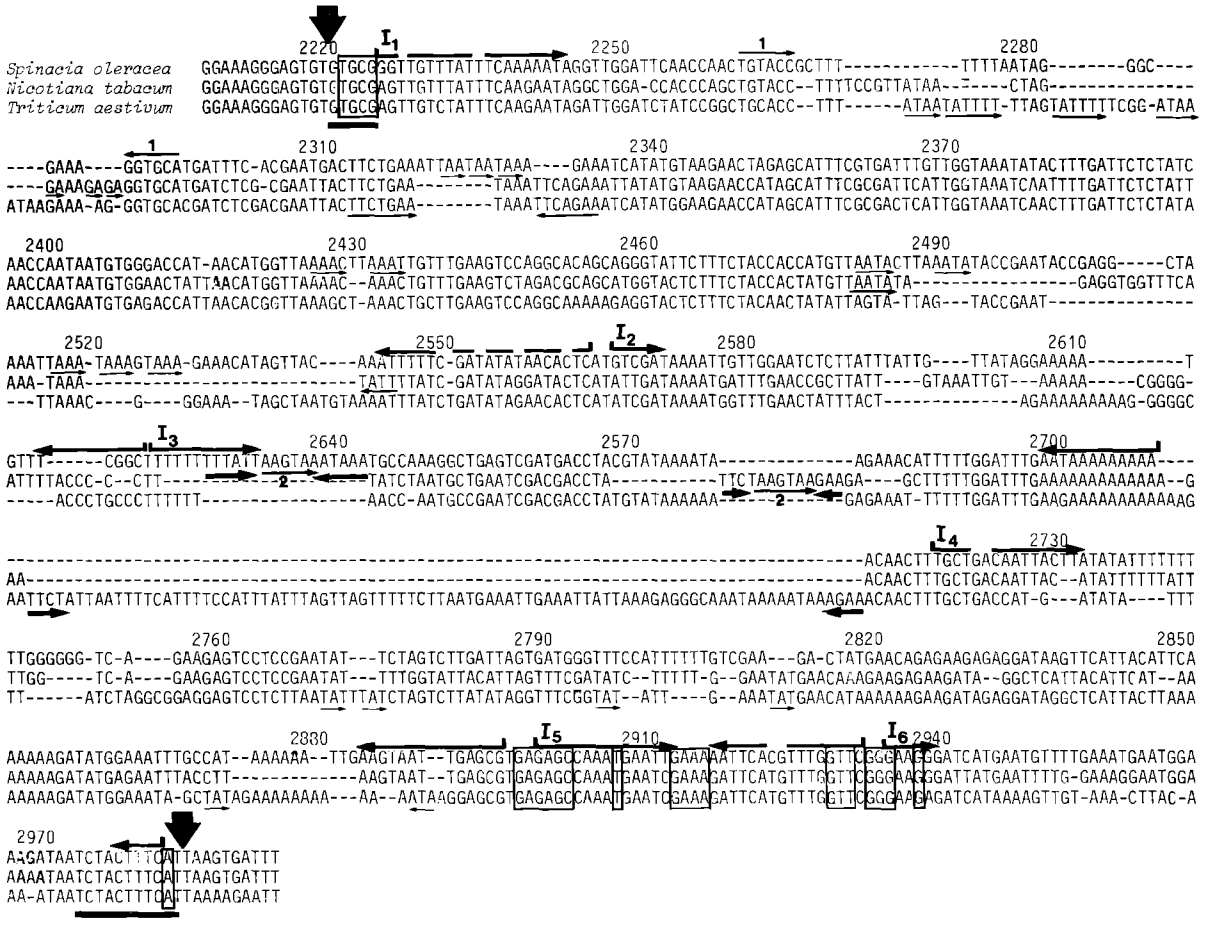

Fig. 7. Comparison of the intervening sequences in the *atpF* genes from spinach, tobacco (Deno et al. 1984) and wheat (Bird et al. 1985). The class II introns are aligned to maximize homology and are numbered according to the spinach sequence (cf. Fig. 2). *Thick vertical arrows* point at the putative exon/intron junctions. Intron consensus sequences common to all introns of the second family (Michel and Dujon 1983) are *boxed*; inverted repeats I_1–I_6 stand for the postulated stem-loop structures common to most introns of the same family. The 5′ and 3′ consensus sequences similar to those proposed for *Euglena* chloroplast genes (Karabin et al. 1984; Koller et al. 1984) are shown over a *solid line*. Some contiguous and non-contiguous repetitions that flank deletions/inversions (Efstradiatis et al. 1980) are indicated by *horizontal arrows*, distant repeats are numbered in addition

long enough to span all four genes. This and the finding that RNA selected with fragments containing sequences from any one of these genes also directs the synthesis in vitro of the other CF_0 and α subunit proteins, suggest that the synthesis of translationally active RNA for these ATP synthase proteins may involve polycistronic transcription and post-transcriptional modification (Herrmann et al. 1985).

The complexity of the Northern patterns finds its complement in the DNA sequence (Fig. 2). Regions resembling prokaryotic promoters ("−10" and "−35" motifs; Rosenberg and Court 1979) are recognizable in front of *atpI* at −111 to −116 and −149 to −154, respectively. The 3′ non-coding region of *atpA* is capable of forming a relatively stable stem-loop structure (not shown) analogous to transcription termination signals in prokaryotic genes (Rosenberg and Court 1979). It is worth noting that the inverted repeat region, 630 bp downstream of *atpI*, is followed by a T-rich stretch and thus exhibits features of transcription terminators. It coincides or is surrounded by potential "−10" (positions 1337–1342/1354–1361) and "−35" (1367–1372/1388–1393) elements preceding *atpH*. However, this region of dyad symmetry and various other intergenic sequences that can form secondary structures may function in RNA processing or attenuation rather than in transcription termination. The segment between the two *atpF* exons is capable of forming six hairpins (I_1–I_6 in Figs. 2 and 7) that share typical structural and nucleotide consensus motifs known from class II introns (Michel and Dujon 1983). The functional organization of this DNA segment deserves further study (see Discussion).

Discussion

The results outlined combined with the previous data (Westhoff et al. 1985) provide compelling evidence that the DNA sequence of the spinach plastid chromosome presented in Fig. 2 contains functional genes for three CF_0 subunits of the chloroplast ATP synthase. The comparison of the amino acid sequences deduced and of the predicted secondary structures with those of corresponding bacterial and mitochondrial components shows that all these sequences, as those for plastome-encoded CF_1 constituents, are structurally similar, functionally equivalent, and therefore probably homologous. Our analysis bears on four basic points:

(1) Composition of CF_0. The newly found gene, *atpI*, encoding the putative homologue of the *E. coli* F_0 subunit *a* resolves ambiguities about the actual constitution of CF_0. We infer that the membrane sector of spinach ATP synthase is composed of four subunit species, of which three, CF_0–I, –III, and –IV, are encoded in the chloroplast, while one, CF_0–II, originates in the nucleus. We suggest that this holds for chloroplast ATP synthases from many plant organisms. Several lines of evidence indicated that eukaryotic ATP synthase membrane sectors appear to be more complex than those from prokaryotes. More than three subunits have been detected from mitochondrial F_0 assemblies (reviewed in Senior and Wise 1983). Chloroplast ATP synthases with four CF_0 subunit species co-purifying with the ATP synthase holoenzyme in stoichiometric or near stoichiometric quantities have been described (Pick and Racker 1979; cf. Fig. 2 in Herrmann et al. 1983; Westhoff et al. 1985) although a fourth subunit has not been noted or consistently noted by other workers in the field. We have presented evidence recently that the number of detectable CF_0 subunit species could be dependent upon experimental conditions, particularly on the preparations of the complex and on the resolution of the chosen gel system (Westhoff et al. 1985).

AtpI is the 6th gene for constitutents of a chloroplast ATP synthase and the 17th for a thylakoid membrane protein to be located on the plastid chromosome from spinach (cf. Herrmann et al. 1985). Although we have not provided direct evidence that the atpI product is translated due to the lack of an appropriate antiserum, this seems likely. A polypeptide of 19 kd, immunologically not related to the products of atpE, F, G and H can be seen among the products made in vitro. This product co-migrates on SDS-polyacrylamide gels with the authentic protein and exhibits the characteristic changes in mobility in one- and two-dimensional gel systems (cf. Figs. 1 and 2 in Westhoff et al. 1985). It is therefore a potential candidate for CF_0 subunit IV. Subunit IV would appear from the DNA sequence to have a molecular weight of 27 kd, yet it runs on SDS gels with a considerably lower mobility. Apart from the fact that the N-terminal sequence of this protein is not known, so that the real molecular weight cannot be calculated, there are various indications that SDS gels greatly underestimate molecular weights, especially of membrane proteins (cf. references in Morris and Herrmann 1984). Prediction of its secondary structure suggests it could be an intrinsic membrane protein containing four trans-membrane helical segments (Figs. 3 and 4). A fifth hydrophobic domain (amino acid residues 199–218) is probably not a trans-membrane segment because of its proline content.

The detection of subunit IV creates nomenclature problems for CF_0 proteins and their genes (cf. Hallick and Bottomley 1983). Subunits and corresponding genes are frequently designated according to decreasing molecular weights. In addition, there is no generally accepted nomenclature to correlate corresponding CF_0 and F_0 components as yet, but equivalence of subunits with the same or comparable designation in chloroplasts, mitochondria and *E. coli* is often assumed (but see Walker et al. 1982; Ovchinnikov et al. 1984a, b). Apart from the fact that relative electrophoretic mobilities of individual CF_0 components may differ with the chosen gel system, our sequence analysis clearly shows that the nomenclatoric equivalents, subunits CF_0–I and F_0–*a*, are not related to each other, nor are CF_0–II and F_0–*b*. Instead, chloroplast CF_0–I is the homologue of bacterial *b*, and the subunit equivalent to bacterial *a* is CF_0–IV. We have not changed the original naming and hope that our designations *atpI* and CF_0–IV for gene and polypeptide, respectively, of the *largest* CF_0 component will not lead to confusion. Inconsistencies like these are an inherent drawback of the suggested nomenclature.

(2) The split CF_0–I gene. Our nucleotide sequence analysis is consistent with the previous identification and properties of the *atpF* gene and its product, CF_0–I (Herrmann et al. 1983; Westhoff et al. 1985). The gene is interrupted by a large, 764 bp, intervening sequence which resembles introns of the second family. It has the six predicted stem-loop structures in correct relative sizes including the conserved structure at the 3′ terminal region of class II introns (Michel and Dujon 1983), 5′ and 3′ boundaries GTGCG (positions 2223–2227) and TCTACTTTCAT (2975–2986) respectively, similar to those reported for the corresponding introns in protein-encoding regions in the *Euglena* and *Nicotiana debneyi* plastid chromosomes (Koller and Delius 1984; Karabin et al. 1984; Zurawski et al. 1984), and a 12 bp inverted repetition within the exons close to the splice sites (positions 2166–2177 and 2922–3003; Figs. 2 and 7). Essentially the same gene organization has been found independently in wheat (Bird et al. 1985) and the published sequence of the corresponding region from tobacco (Deno et al. 1984) can be interpreted in a similar way (Fig. 7).

The conservation of the *atpI* introns in the three plant species (Fig. 7) is remarkable in two respects. Apart from the fact that intervening sequences are rare in higher plant plastid chromosomes (Koller and Delius 1984), contrary to those of lower organisms (Stiegler et al. 1982; Montadon and Stutz 1983; Erickson et al. 1984; Karabin et al. 1984), the only other intron reported to date from a plastid gene in higher plants coding for a protein, the 666 bp intron in the gene for the ribosomal protein L2 of *N. debneyi*, is not present in the corresponding spinach gene (Zurawski et al. 1984). The *atpI* intervening sequence of wheat is 823 bp (Bird et al. 1985) and that of tobacco 694 bp (Deno et al. 1984). We have compared the three introns, and analysed the nucleotide sequences surrounding the apparent 8–10 gaps, using the sequence evolution models of Efstradiatis et al. (1980) and Shen et al. (1981; Figs. 2 and 7). The comparison indicates that these regions are more than 80% conserved, diverge by deletion and addition, and by base substitution, which leave most stem structures unaltered. The gaps are associated with contiguous and non-contiguous repeats, or AT-rich sequences. The significance of these deletions and insertions, which are remarkably similar to those observed within the genes for the two P_{700} chlorophyll *a* apoproteins, (Kirsch et al., submitted) is unclear.

On the basis of S1 nuclease analysis and comparison with the amino acid sequence of the *E. coli* polypeptide, Bird et al. (1985) have determined the exon/intron boundaries in the wheat gene corresponding to nucleotide positions 2222 and 2986 in spinach (Fig. 2). We have preliminary data that this assignment holds for spinach (data not shown). The organization outlined indicates that the previously observed immunoprecipitable, truncated spinach products obtained with both DNA- and RNA programmed cell-free assays (see Introduction, Westhoff et al. 1985) are initiated at one of the later in-frame AUG codons. A poten-

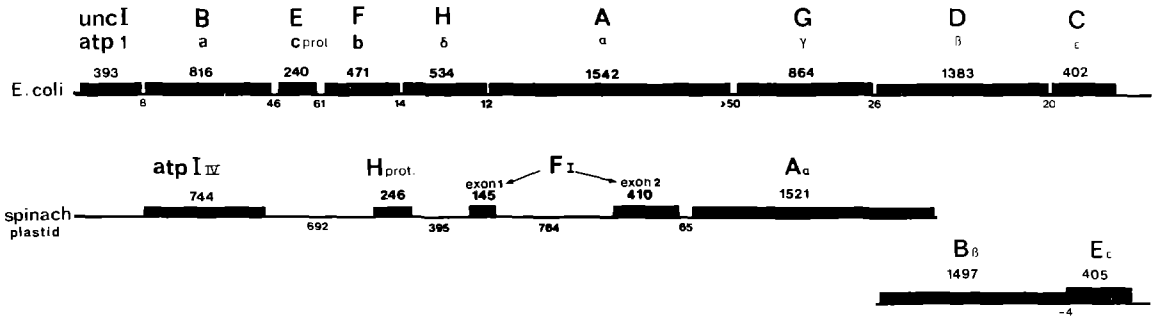

Fig. 8. Gene order in the *E. coli atp* (*unc*) operon (modified from Saraste et al. 1981; Senior and Wise 1983) and in the *atp* operons *A* and *B* of the spinach plastid chromosome. $\alpha-\varepsilon$ designate F_1 and CF_1 subunits, a, b and c F_0 subunits. Genes *atpI* (CF_0–IV), *H* (proteolipid) and *F* (CF_0–I) encode polypeptides for the CF_0 sector of ATP synthase. The structural genes were aligned based on the sequence data; homology between the intergenic region is not so great and in general it is not possible to align these sequences

tial ribosome binding site, AAAG, preceds Met 51 in reading frame 4 (exon 2), and the preceding secondary structure of the 12 bp inverted repeat sequence (see above) of the processed mRNA may prevent correct decoding in vitro. We have repeatedly realized that heterologous cell-free systems give rise to various translation artifacts (see discussion in Westhoff et al. 1985; Sears and Herrmann 1985).

The molecular weight of the CF_0–I protein calculated from the deduced amino acid sequences of the combined open reading frames (20.9 kd) agrees reasonably with the values determined by gel electrophoresis (18 kd). Based on the experimentally determined N-terminal amino acid sequence of the spinach protein, Bird et al. (1985) have raised the possibility that the wheat protein is 17 residues (2 kd) shorter than that predicted from the DNA sequence, suggesting that some processing might occur after translation. (The undetermined residue at amino acid position 23 is Asn; Fig. 2). Most of the corresponding codons are conserved in spinach (Fig. 5). As in *E. coli*, the sequence of this protein is quite unique. Residues 26–45 are hydrophobic and possess the potential of forming a single transmembrane span (Fig. 6, Kyte and Doolittle 1982) suggesting that this amino terminal region of the protein is embedded in the lipid bilayer while the rest would protrude from the membrane, probably into the stroma. (Note that both domains are encoded by different exons). Significant roles in proton translocation from the proton channel to the catalytic sites of ATP synthesis and in topogenesis, that is to form a membrane anchor around which F_1 and CF_1 polypeptides could assemble, have been envisaged for this protein (Nelson 1981; Senior and Wise 1983; Hoppe and Sebald 1984; Walker et al. 1985).

It is interesting to note in this context that the predicted secondary structure for subunit I (Fig. 6) is reminiscent of that of cytochrome *f*, the only known plastome-encoded protein synthesized in chloroplasts with an N-terminal extension (Alt and Herrmann 1984; Willey et al. 1984). Both proteins share a single terminal trans-membrane span, a putative transitory hydrophobic segment and a relatively large hydrophilic segment. Yet, they differ in the position of the trans-membrane span within the polypeptide chain, which is C-terminal in cytochrome *f*, but N-terminal in CF_0–I. While it is conceivable that the cytochrome *f* signal sequence is required for directing (co-translational?) transfer of the hydrophilic part across the thylakoid membrane, the role, if any, of the proposed leader sequence for CF_0–I remains to be established. Unlike that of cytochrome *f*, this sequence does not conform the characteristics of signal sequences (von Heijne 1983).

(3) Phylogeny. Probably the most striking finding of this study besides the detection of *atpI* and the interrupted gene for *atpF* is the positional conservation of the ATP synthase structural genes in the spinach plastid chromosome when compared with the arrangement of genes in the *atp* operon of *E. coli*. This is obvious from Fig. 8, which schematically presents the organisation of the ATP synthase genes from both organisms. Using the plastid gene nomenclature for comparison, the order of genes (products) in the *E. coli atp* operon is 5′ 1-*atpI*(a)-*atpH*(c)-*atpF*(b)-*atpD*(δ)-*atpA*(α)-*atpC*(γ)-*atpB*(β)-*atpE* (ε; Gay and Walker 1981a, b; Nielsen et al. 1981; Saraste et al. 1981). Gene 1 constitutes an unidentified reading frame. The occurrence of the closely spaced genes *atpI-atpH-atpF-atpA*, and separated by 40 kbp, of *atpB-atpE* in spinach (Fig. 1) is reminiscent of a partial bacterial *atp* operon. In fact, each of these clusters forms a transcriptional unit. This arrangement in two operons holds for plastid chromosomes of a variety of higher plants even if intrachromosomal rearrangement of transcriptional units has occurred (cf. discussion in Huttly and Gray 1984; Herrmann et al. 1985). We therefore designate these transcription units *atp* operons *A* and *B*.

ATP synthase CF_1 subunits gamma and delta derive from nuclear genes. The absence of these genes in operon A (Fig. 8) and the positional conservation of loci in the operons A and B support the notion of gene transfer during the evolution of the intergenomic integration of eukaryotic genomes and argue against convergent evolution of ATP synthase genes in eukaryotes and eubacteria. If correct, this transfer must have been attended or followed by structural and functional rearrangements within plastid chromosomes, since the A and B operons are 40 kbp apart. Both operons are transcribed from the same strand (Fig. 1), which implies that at least two such events should have occurred. There are other examples of plastid genes being encoded in operons that are organized as in *E. coli*, but with some genes being deleted or rearranged. Recently found examples are the alpha, strep and S10 operons (for references see Sijben-Müller et al. 1985) which even may maintain their relative position in the *E. coli* and plastid chromosomes.

The origin and role of the third nuclear encoded ATP synthase subunit, CF_0–II (Nelson et al. 1980; Herrmann et al. 1983; Westhoff et al. 1985), is obscure at present, since

it appears to lack a counterpart in *E. coli*. This component may be an adjunct of eukaryotic ATP synthases and therefore belong to a *phylogenetically different category* of nuclear-coded plastid proteins. It may well be the analogue of the bovine mitochondrial ε subunit, which has no known functional equivalent in *E. coli* (Walker et al. 1982). The relationships with the respective bacterial and mitochondrial proteins will probably be resolved when the sequences for the nuclear-encoded subunits become available.

The individual ATP synthase subunits must have evolved at different rates. The amino acid sequence of spinach CF_0 subunits I, III (Alt et al. 1983) and IV, and subunit ε (Zurawski et al. 1982) show lower homology (25%–38%) with the corresponding subunits of *E. coli* and mitochondrial complexes than the subunit β or α (50%–70%, Zurawski et al. 1982; Deno et al. 1983). This difference in primary structure conservation may be caused by the different functional roles they perform. The β subunit is known to be involved in the catalytic reaction (ATP synthesis), the other subunits may exert primarily structural or regulatory functions. Constraints on their sequence, therefore, may be less demanding. Predictions of secondary structure, which indicate structural conservation of corresponding homologues, support this suggestion. It will be interesting to see if structures of CF_0 subunits have drifted in parallel with those of CF_1 domains with which they associate.

(4) Expression. The expression of the ATP synthase genes to achieve final subunit stoichiometries constitutes a central, unsolved question. As far as organelle-located genes are concerned there are only single loci per subunit in each plastid chromosome, and yet some subunits are present in multiple copies, others appear to be present in single copies only. Purified CF_1 assembly has a proposed stoichiometry of 3:3:1:1:1 in analogy to the F_1 sector of the *E. coli* complex. Stoichiometries of I:II:III (a:b:c) of 1:2:6–15 that have been proposed for CF_0 in analogy to the F_0 section of *E. coli* are not definitive and, in addition, need to be re-evaluated in view of the fourth CF_0 subunit described in this paper.

We have noted recently that expression of the *atp* operons *A* and *B* in spinach is remarkably different (Alt et al. 1983). The bicistronic transcript of operon *B* in which the beta and epsilon reading frames overlap with four bases (Zurawski et al. 1982) is processed at the 5′ terminus but probably not dissected into two subspecies prior to translation (Krebbers et al. 1982; Zurawski et al. 1982; Alt et al. 1983 and unpublished results, but see Shinozaki et al. 1983). An essential control in decoding subunit β and ε in a ratio of 3:1 must, therefore, be at the translational level. The fused β/ε product found in the *Euoenothera* ATP synthase plastome mutant iota in vivo (but not after heterologous translation in vitro) indicates an impaired signal in this decoding process (Sears and Herrmann 1985).

Northern blot analysis and hybrid select translation suggest that the *atp* operon *A* has also a single major promotor, which initiates upstream from *atpI*. However, synthesis of mature, translationally active RNA involves extensive post-transcriptional modification. It is possible that the largest, polycistronic RNA species (ca. 6 kbp) represents an unprocessed component which is rapidly dissected, resulting in a complex RNA pattern with major RNA species of 4.5, 2.5, 1.6, 0.9 and 0.5 kb (cf. Fig. 7 in Westhoff et al. 1985). The latter component hybridises specifically with subunit III probes and about 200 bp preceding sequences. Since the stationary concentration of the giant primary transcript is quite low, its overall half-life is probably short. This RNA maturation includes nucleolytic and splicing activity, and there is evidence for *inter*-cistronic splicing as well (cf. Westhoff et al. 1985).

The significance of the complex RNA patterns derived from operon *A* and the different functional organization of the two operons is, as yet, not clear. In view of the outlined β/ε decoding mechanism it is difficult to envisage how gene arrangements such as those could function to ensure the subunit requirements of ATP synthase assembly. In *E. coli* the major factor underlying the synthesis of appropriate amounts of F_0 subunits appears to be the efficiency of translational initiation rather than local mRNA secondary structures or codon usage (McCarthy et al. 1985). The codons of the ATP synthase genes conform to the types commonly found in the tylakoid membrane genes (see Herrmann et al. 1985), in that they have acquired A or T in the third position (*atpI*: 74%, *atpH*: 80%, *atpF*: 66.3%). The relatively long transcribed intergenic regions in spinach that are found in processed RNA species suggest that leader sequences could be involved in decoding efficiency. On the other hand, the remarkable difference in the stationary concentrations for CF_0–IV and –III mRNAs (cf. Westhoff et al. 1985) and the potential promotor motifs prior to *atpH* could indicate a separate transcriptional control for both loci and/or that their mRNAs differ in stability.

The regulation of the expression of the ATP synthase genes is an interesting puzzle at the moment and its study can yield basic information on gene expression in eukaryotes. Our data suggest that post-transcriptional processes, translational controls and/or various transcription initiation and termination points constitute essential elements in the biogenesis of ATP synthase and other protein complexes of the thylakoid membrane (cf. Herrmann et al. 1985). The functional organization of this DNA segment is currently under study.

Acknowledgements: The skilfull technical assistence of Ms. Barbara Caspers is greatfully acknowledged. This work was supported by grants of the Stiftung Volkswagenwerk and Forschungsmittel des Landes Nordrhein-Westfalen.

References

Alt J, Winter P, Sebald W, Moser JG, Schedel R, Westhoff P, Herrmann RG (1983) Localization and nucleotide sequence of the gene for the ATP synthase proteolipid subunit on the spinach plastid chromosome. Current Genet 7:129–138

Alt J, Herrmann RG (1984) Nucleotide sequence of the gene for pre-apocytochrome *f* in the spinach plastid chromosome. Current Genet 8:551–557

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. Cell 26:167–180

Bird CR, Koller B, Auffret AD, Huttly AK, Howe CJ, Dyer TA, Gray JC (1985) The wheat chloroplast gene for CF_0 subunit I of ATP synthase contains a large intron. EMBO J 4:1381–1388

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl Acids Res 7:1513–1523

Chou PY, Fasman GD (1978) Prediction of the secondary structure

of proteins from their amino acid sequence. Advances Enzymol 47:45–148

Cramer WA, Widger WR, Herrmann RG, Trebst A (1985) Topography and function of thylakoid membrane proteins. Trends Biochem Sci 10:125–129

Deno H, Shinozaki K, Sugiura M (1983) Nucleotide sequence of tobacco chloroplast gene for the α subunit of proton-translocating ATPase. Nucl Acids Res 11:2185–2191

Deno H, Shinozaki K, Sugiura M (1984) Structure and transcription pattern of a tobacco chloroplast gene coding for subunit III of proton-translocating ATPase. Gene 32:195–201

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Sprik RA, DeRiel JK, Forget BG, Weissman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of the human β-globin gene family. Cell 21:653–668

Erickson JM, Rahire M, Rochaix JD (1984) *Chlamydomonas reinhardii* gene for the 32000 mol. wt. protein of photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. EMBO J 3:2753–2762

Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120:97–120

Gay NJ, Walker JE (1981a) The atp operon: nucleotide sequence of the region encoding the α-subunit of *Escherichia coli* ATP synthase. Nucl Acids Res 9:2187–2194

Gay NJ, Walker JE (1981b) The atp operon: nucleotide sequence of the promotor and the genes for the membrane proteins, and the δ subunit of *Escherichia coli* ATP synthase. Nucl Acids Res 9:3919–3926

Grisi E, Brown TA, Waring RB, Scazzacchion C, Davies RW (1982) Nucleotide sequence of a region of the mitochondrial genome of *Aspergillus nidulans* including the gene for ATPase subunit 6. Nucl Acids Res 10:3531–3539

Hallick RB, Bottomley W (1983) Proposals for the naming of chloroplast genes. Plant Mol Biol Reporter 1:38–43

Heinemeyer W, Alt J, Herrmann RG (1984) Nucleotide sequence of the clustered genes for apocytochrome *b*6 and subunit 4 of the cytochrome *b*/*f* complex in the spinach plastid chromosome. Current Genet 8:543–549

Hennig J (1984) Feinstrukturanalyse am ATP Synthase-Operon I von Spinat (*Spinacia oleracea*), Diplomarbeit Universität Düsseldorf, p 108

Herrmann RG, Possingham JV (1980) Plastid DNA – the plastome. In: Reinert J (ed) Results and problems in cell differentiation, vol 10. Springer, Berlin Heidelberg New York, pp 45–96

Herrmann RG, Seyer P, Schedel R, Gordon K, Bisanz C, Winter P, Hildebrandt JW, Wlaschek M, Alt J, Driesel AJ, Sears BB (1980) The plastid chromosomes of several dicotyledons. In: Bücher Th, Sebald W, Weiss H (eds) Biological chemistry of organelle formation. Springer, Berlin Heidelberg New York, pp 97–112

Herrmann RG, Westhoff P, Alt J, Winter P, Tittgen J, Bisanz C, Sears BB, Nelson N, Hurt E, Hauska G, Viebrock A, Sebald W (1983) Identification and characterization of genes for polypeptides of the thylakoid membrane. In: Cifferi O, Dure L III (eds) Structure and function of plant genomes. Plenum Publ Corp, New York, pp 143–153

Herrmann RG, Westhoff P, Alt J, Tittgen J, Nelson N (1985) Thylakoid membrane proteins and their genes. In: van Vloten-Doting L, Groot GSP, Hall TC (eds) Molecular form and function of the plant genome, Plenum Publ Corp, New York, pp 233–256

Hoppe J, Sebald W (1984) The proton conducting F_0-part of bacterial ATP synthases. Biochim Biophys Acta 768:1–27

Huttly AK, Gray JC (1984) Localisation of genes for four ATP synthase subunits in pea chloroplast DNA. Mol Gen Genet 194:402–409

Karabin GD, Farley M, Hallick RB (1984) Chloroplast gene for Mr 32000 polypeptide of photosystem II in *Euglena gracilis* is interrupted by four introns with conserved boundary sequences. Nucl Acids Res 12:5801–5812

Kirsch W, Seyer P, Herrmann RG (1986) Nucleotide sequence of the clustered genes for *two* P_{700} chlorophyll *a* apoproteins of the photosystem I reaction center and the ribosomal protein S14 in the spinach plastid chromosome. Current Genet, submitted

Koller B, Delius H (1984) Intervening sequences in chloroplast genomes. Cell 36:613–622

Krebbers ET, Larrinua IM, McIntosh L, Bogorad L (1982) The maize chloroplast genes for the β and ε subunits of the photosynthetic coupling factor CF_1 are fused. Nucl Acids Res 10:4985–5002

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Leaver CJ, Gray MW (1982) Mitochondrial genome organization and expression in higher plants. Annu Rev Plant Physiol 33:373–402

Macino G, Tzagoloff A (1980) Assembly of the mitochondrial membrane system: sequence analysis of a yeast mitochondrial ATPase gene containing the oli-2 and oli-4 loci. Cell 20:507–517

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor Laboratory, NY, pp 545

Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65:499–560

McCarthy JEG, Schairer HU, Sebald W (1985) Translational initiation frequency of atp genes from *Escherichia coli*: identification of an intercistronic sequence that enhances translation. EMBO J 4:519–526

Michel F, Dujon B (1983) Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. EMBO J 2:33–38

Montandon PE, Stutz E (1983) Nucleotide sequence of a *Euglena gracilis* chloroplast genome region coding for the elongation factor Tu; evidence for a spliced mRNA. Nucl Acids Res 11:5877–5892

Morris J, Herrmann RG (1984) Nucleotide sequence of the gene for the P_{680} chlorophyll *a* apoprotein of the photosystem II reaction center from spinach. Nucl Acids Res 12:2837–2850

Nelson N (1981) Proton-ATPase of chloroplasts. Current Topics Bioenerget 11:1–33

Nelson N, Nelson H, Schatz G (1980) Biosynthesis and assembly of the proton-translocating adenosine triphosphatase complex from chloroplasts. Proc Natl Acad Sci USA 77:1361–1364

Nielsen J, Hansen FG, Hoppe J, Friedl P, Meyenburg K (1981) The nucleotide sequence of the *atp* genes coding for the F_0 subunits a, b, c and the F_1 subunit δ of the membrane bound ATP synthase of *Escherichia coli*. Mol Gen Genet 184:33–39

Ovchinnikov YA, Modyanov NN, Grinkevich VA, Aldanova NA, Trubetskaya OE, Nazimov JV, Hundal T, Ernster L (1984a) Amino acid sequence of the oligomycin sensitivity-conferring protein (OSCP) of beef-heart mitochondria and its homology with the δ-subunit of the F_1-ATPase of *Escherichia coli*. FEBS Lett 166:19–22

Ovchinnikov YA, Modyanov NN, Grinkevich VA, Aldanova NA, Koestetsky PV, Trubetskaya OE, Hundal T, Ernster L (1984b) Oligomycin sensitivity-conferring protein (OSCP) of beef-heart mitochondria. FEBS Lett 175:109–112

Pick U, Racker E (1979) Purification and reconstitution of the N, N′-dicyclohexycarbodiimide-sensitive ATPase complex from spinach chloroplasts. J Biol Chem 254:2793–2799

Rosenberg M, Court D (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. Annu Rev Genet 13:319–353

Saraste M, Gay NJ, Eberle A, Runswick MJ, Walker JE (1981) The atp operon: nucleotide sequence of the genes for the γ, β, and ε subunits of *Escherichia coli* ATP synthase. Nucl Acids Res 9:5287–5296

Schwarz Z, Kössel H (1980) The primary structure of 16S rDNA from *Zea mays* chloroplast is homologous to *E. coli* 16S rRNA. Nature 283:739–742

Sears BB, Herrmann RG (1985) Plastom mutation affecting the chloroplast ATP synthase involves a post-transcriptional defect. Current Genet 9:521–528

Sebald W, Hoppe J (1981) On the structure and genetics of the proteolipid subunit of the ATP synthase complex. Curr Top Bioenerg 12:1–64

Senior AE, Wise JG (1983) The proton-ATPase of bacteria and mitochondria. J Membr Biol 73:105–124

Shen Sh, Slightom JL, Smithies O (1981) A history of the human fetal globin gene duplication. Cell 26:191–203

Shine J, Dalgarno L (1974) The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA 71:1342–1346

Shinozaki K, Deno H, Kato A, Sugiura M (1983) Overlap and cotranscription of the genes for the beta and epsilon subunits of tobacco chloroplast ATPase. Gene 24:147–155

Sijben-Müller G, Hallick RB, Alt J, Westhoff P, Herrmann RG (1986) Spinach plastid genes coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase α-subunit. Nucl Acid Res 14:1029–1044

Stiegler GL, Matthews HM, Bingham SE, Hallick RB (1982) The gene for the large subunit of ribulose-1,5-bisphosphate carboxylase in *Euglena gracilis* chloroplast DNA: location, polarity, cloning and evidence for an intervening sequence. Nucl Acids Res 10:3427–3444

Tohdoh N, Sugiura M (1982) The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. Gene 17:213–218

von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133:17–21

Walker JE, Runswick MJ, Saraste M (1982) Subunit equivalence in *Escherichia coli* and bovine heart mitochondrial F_1F_0 ATPases. FEBS Lett 146:393–396

Walker JE, Fearnley JM, Gay NJ, Gibson BW, Northrop ED, Powell SJ, Runswick MJ, Saraste M, Tybulewicz VLJ (1985) Primary structure and subunit stoichiometry of F_1-ATPase from bovine mitochondria. J Mol Biol 184:677–701

Westhoff P, Nelson N, Bünemann H, Herrmann RG (1981) Localization of genes for coupling factor subunits on the spinach plastid chromosome. Current Genet 4:109–120

Westhoff P, Alt J, Nelson N, Herrmann RG (1985) Genes and transcripts for the ATP synthase CF_0 subunits I and II from spinach thylakoid membranes. Mol Gen Genet 199:290–299

Whitfeld PR, Herrmann RG, Bottomley W (1978) Mapping of the ribosomal RNA genes on spinach chloroplast RNA. Nucl Acids Res 5:1741–1751

Willey DL, Howe CJ, Auffret AD, Bowman CM, Dyer TA, Gray JC (1984) Location and nucleotide sequence of the gene for cytochrome *f* in wheat chloroplast DNA. Mol Gen Genet 194:416–422

Wu R (1970) Nucleotide sequence analysis of DNA. J Mol Biol 51:501–521

Zurawski G, Perrot B, Bottomley W, Whitfeld P (1981) The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucl Acids Res 9:3251–3270

Zurawski G, Bottomley W, Whitfeld PR (1982) Structures of the genes for the β and ε subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. Proc Natl Acad Sci USA 79:6260–6264

Zurawski G, Bottomley W, Whitfeld PR (1984) Junctions of the large single-copy region and the inverted repeats in *Spinacia oleracea* and *Nicotiana debneyi* chloroplast DNA: sequence of the genes for $tRNA^{His}$ and the ribosomal proteins S19 and L2. Nucl Acids Res 12:6547–6558

Communicated by C.P. Hollenberg

Received December 23, 1985