A Homologue of a Nuclear-Coded Iron–Sulfur Protein Subunit of Bovine Mitochondrial Complex I Is Encoded in Chloroplast Genomes†

Alain Dupuis,‡,§ J. Mark Skehel, and John E. Walker*

The Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

Received August 23, 1990; Revised Manuscript Received October 29, 1990

ABSTRACT: The chloroplast genomes of *Marchantia polymorpha*, *Nicotiana tabacum*, and *Oryza sativa* contain open reading frames (ORFs or potential genes) encoding homologues of some of the subunits of mitochondrial NADH:ubiquinone oxidoreductase (complex I). Seven of these subunits (ND1–ND4, ND4L, ND5, and ND6) are products of the mitochondrial genome, and two others (the 49- and 30-kDa components of the iron–sulfur protein fraction) are nuclear gene products. These findings have been taken to indicate the presence in chloroplasts of an enzyme related to complex I, possibly an NAD(P)H:plastoquinone oxidoreductase, participating in chlororespiration. This view is reinforced by the present work in which we have shown that chloroplast genomes encode a homologue of the 23-kDa subunit, another nuclear-encoded component of bovine complex I. The 23-kDa subunit is in the hydrophobic protein fraction of the enzyme, the residuum after removal of the flavoprotein and iron–sulfur protein fractions. The sequence motif CysXXCysXXCysXXXCysPro, which provides ligands for tetranuclear iron–sulfur centers in ferredoxins, occurs twice in its polypeptide chain and is evidence of two associated 4Fe-4S clusters. This is the only iron–sulfur protein identified so far in the hydrophobic protein fraction of complex I, and so it is possible that one of these centers is that known as N-2, the donor of electrons to ubiquinone. The sequence of the 23-kDa subunit is closely related to potential proteins, which also contain the cysteine-rich sequence motifs, encoded in the *frx*B ORFs in chloroplast genomes. *Frx*B is found in a cluster of ORFs in tobacco, liverwort, and rice, with genes encoding other homologues of complex I, arranged in the order 49 kDa, ND1, 23 kDa (frxB), ND6, ND4L, and ND4. Two ORFs encoding homologues of 30-kDa and ND3 subunits cluster at a second separate locus. Thus, chloroplast genomes contain homologues of at least ten components of mitochondrial complex I, with eight of them clustered at two separate loci.

An unexpected feature of the chloroplast genomes of the liverwort, *Marchantia polymorpha*, and the tobacco plant, *Nicotiana tabacum*, evident from their DNA sequences, is that each contains seven potential genes, or open reading frames (ORFs), that code for homologues of seven subunits of mitochondrial NADH:ubiquinone oxidoreductase (complex I). In mammals [and also in other species; see Fearnley et al. (1989) for a summary] these proteins are encoded in mitochondrial DNA and are known as subunits ND1–ND6 and ND4L (Chomyn et al., 1985, 1986). In consequence, the chloroplast ORFs were named *ndh*1–*ndh*4, *ndh*4L, *ndh*5, and *ndh*6 and *ndh*A–*ndh*G in liverwort and tobacco, respectively. Homologues are also found in the rice chloroplast genome (Hiratsuka et al., 1989). These observations have led to the proposal that the putative chloroplast proteins encoded in the ORFs are subunits of a NADH or NADPH:plastoquinone oxidoreductase component of a chloroplast respiratory electron-transport activity and that this enzyme complex will be rather closely related to the mitochondrial NADH:ubiquinone oxidoreductase complex. Studies of the 25 or more nuclear-coded components of the mitochondrial enzyme, of which the present work is part, show that some of them also have homologues in the chloroplast genomes, and so lend support to this view. The first two examples to be described were the 49- and 30-kDa subunits of bovine complex I (Fearnley et al., 1989; Pilkington et al., 1991a), and the third one, the 23-kDa subunit, is described below. The 23-kDa subunit has been little studied and is a component of the "hydrophobic protein" (HP) fraction, a residue left behind from the intact complex after removal of the flavoprotein (FP) and iron–sulfur protein (IP) fractions [reviewed by Ragan (1987)]. Its sequence is related to those of ferredoxins and especially to that of frxB, a ferredoxin-like protein encoded in ORFs in liverwort, tobacco, rice, and wheat chloroplast DNAs. This family of proteins appears to contain two tetranuclear iron–sulfur centers, and it is possible that the 23-kDa subunit donates electrons to ubiquinone.

Materials and Methods

Isolation and Protein Sequencing of Subunits of Complex I. Complex I was purified from beef heart mitochondria (Hatefi, 1978; Ragan et al., 1987). Its activity was about 50 units/mg of protein (Fearnley et al., 1989), and its subunit composition was closely similar to a preparation of the enzyme from the laboratory of Dr. C. I. Ragan as judged by electrophoresis in a 10–20% polyacrylamide gradient gel. The subunits of the enzyme were separated by 2D isoelectric focusing and electrophoresis in polyacrylamide gels (O'Farrell, 1975) as follows. Complex I (100 μg in 50 μL of sample buffer) was layered on the top (acidic end) of an isoelectric focusing tube gel (170 mm × 0.75 mm) containing pH 3.5–10 ampholines. The gel was run under nonequilibrium conditions at 100 V for 10 min followed by 200 V for 15 min, 300 V for 30 min, and finally 400 V for 4 h and then placed on top of a 16% polyacrylamide slab gel (Laemmli, 1970) with a 3% polyacryl-

† The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05313.

* To whom correspondence should be addressed.

‡ A.D. was supported by INSERM and the Conseil Régional Rhône-Alpes, France.

§ Present address: Laboratoire de Biochimie, INSERM, BP 85X, Grenoble, France.

0006-2960/91/0430-2954$02.50/0

amide stacking gel. After electrophoresis, the separated proteins were transferred onto a poly(vinylidene difluoride) membrane as described before (Pilkington et al., 1991b), except that the transfer time was 1 h. The transferred proteins, including the 23-kDa subunit, were detected with Coomassie Blue dye, excised, and subjected to protein sequence analysis (Fearnley et al., 1989).

Oligonucleotide Synthesis. The synthesis and purification of synthetic oligonucleotides have been described before (Pilkington et al., 1991b). Oligonucleotide primers for use in polymerase chain reactions were synthesized with a linker sequence containing various restriction enzyme sites (*Eco*RI, *Hin*dIII, or *Bam*HI) on their 5′ ends, to facilitate cloning of the products into M13 vectors.

DNA and RNA Preparations. The preparation of bovine genomic DNA and poly(A+) RNA from bovine heart and the syntheses of single- and double-stranded cDNA have been described earlier (Walker et al., 1987b; Pilkington et al., 1991b). Single-stranded cDNA for use in polymerase chain reactions to produce cDNAs extending to the 5′ extremity of the corresponding mRNA was made from poly(A+) RNA (1.5 mg) with random primers (cDNA synthesis kit; Amersham International PLC, Amersham, U.K.). The product was freed from salt with a Centricon 30 (Amicon, Danvers, MA). A run of A residues was added to its 5′ end by incubation at 37 °C for 30 min with terminal transferase (36 units; Pharmacia, Milton Keynes, U.K.) in a buffer (final volume 15 μL) containing 100 mM potassium cacodylate, pH 6.9, 0.1 mM dithiothreitol, 1 mM $CoCl_2$, 1 mM dATP, and bovine serum albumin (0.5 mg).

Polymerase Chain Reactions. Conditions for polymerase chain reactions, for the fractionation of products by gel electrophoresis, for the detection of products by hybridization with synthetic oligonucleotides, and for the recovery and cloning of cDNAs in M13 vectors have been described previously (Pilkington et al., 1991b; Runswick et al., 1990). When degenerate primers or poly(T) primers were employed (see PCR 1, 2, and 4 in Figure 2), the specificity of the polymerase chain reaction was increased by carrying out the first 30 cycles at a low primer concentration (10–25 nM), followed by an additional 30 cycles at a higher primer concentration (1 mM).

The sequences of cloned products of polymerase chain reactions were determined according to the modified dideoxy method (Sanger et al., 1977; Biggin et al., 1983) using the universal LMB2 primer (Duckworth et al., 1981) and four other unique synthetic oligonucleotide primers. All sequences were determined completely in both senses of the DNA, and compressions were resolved by the use of deazaguanidine or dITP in sequencing reactions. DNA sequences were compiled and analyzed with the computer programs DBUTIL and ANALYSEQ (Staden, 1982a, 1985).

Data Analysis. The protein sequence of the 23-kDa subunit of bovine mitochondrial complex I was compared with those in the PIR and SWISSBANK databases by using the program FASTA (Lipman & Pearson, 1985). Further analysis of the homologies and sequence alignments were performed with DIAGON (Staden, 1982b).

Hybridization with Genomic DNA. Samples of bovine genomic DNA (20 μg) were digested separately for 2 h at 37 °C with each of the restriction endonucleases *Bam*HI, *Sac*I, *Eco*RI, and *Pst*I (20 units in each case, in a final volume of 400 μL). After addition of a second and equal amount of restriction enzyme, the digestion was continued for 2 h more. The digestion products were precipitated with ethanol, redissolved in electrophoresis buffer (30 μL), and fractionated on a 0.6% high melting point agarose gel. After hybridization, the filter was washed at 65 °C with a solution containing 0.2 × SSC (1 × SSC is 150 mM sodium chloride and 15 mM sodium citrate) and 0.1% sodium dodecyl sulfate. Conditions for transfer of fragments to membranes, for the preparation of radioactive "prime-cut" probes and for hybridization and autoradiography have been given elsewhere (Walker et al., 1989).

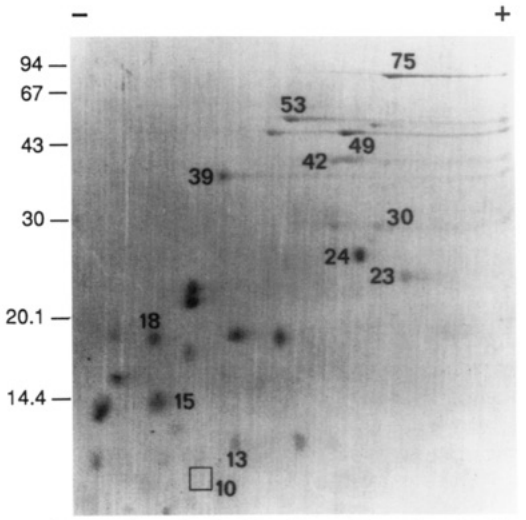

FIGURE 1: Separation of subunits of complex I from bovine heart mitochondria by two-dimensional isoelectric focusing and polyacrylamide gel electrophoresis. The plus and minus signs indicate the acidic and basic ends, respectively, of the isoelectric focusing gel. The positions of subunits of the FP and IP fractions of complex I are indicated, together with the 23-kDa component of the HP fraction.

Results and Discussion

Protein Sequence Analysis of the 23-kDa Subunit. The subunits of complex I from bovine heart mitochondria were resolved under denaturing conditions by two-dimensional isoelectric focusing and electrophoresis in a polyacrylamide gel. Then they were transferred to a poly(vinylidene difluoride) membrane and stained with dye (see Figure 1), and the spots were excised. Several known subunits including the 75-, 53-, 49-, 42-, 39-, 30-, 24-, 18-, 15-, and 13-kDa subunits were recognized from their N-terminal sequences (Runswick et al., 1989; Pilkington et al., 1991a,b; Fearnley et al., 1989; Pilkington & Walker, 1989; J. M. Skehel and J. E. Walker, unpublished work), and the sequence of amino acids 1–23 of the 23-kDa subunit was determined. Subsequently, the same sequence was detected in complex I fractionated on one-dimensional gels, and the protein appears to be an authentic component of the complex.

cDNA Encoding the 23-kDa Subunit of Complex I. The isolation of cDNA clones for the 23-kDa subunit was carried out by a strategy based on the polymerase chain reaction that has been developed in the course of sequence studies on complex I and on other mitochondrial proteins (Runswick et al., 1990; Pilkington et al., 1991b). The first step in this procedure is to determine the sequence of a stretch of at least 17 consecutive amino acids, either at the N-terminus or at an internal site in the protein of interest. The hexapeptide sequences at the N- and C-terminal ends of the known sequence are used to design degenerate oligonucleotides 17 bases in length to be used as forward and reverse primers, respectively, in a polymerase chain reaction with total bovine cDNA as template. The protein sequence between the extremities is used to design a third mixed oligonucleotide, which is then employed as a hybridization probe in the identification of the short cDNA

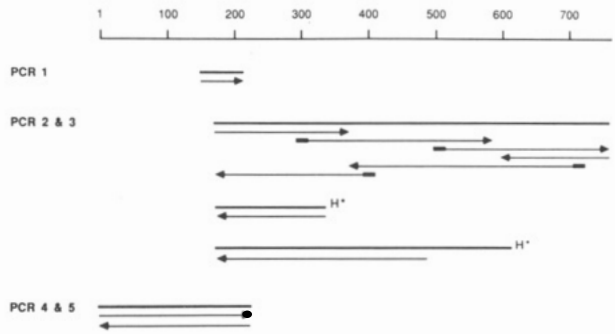

FIGURE 2: Generation of clones for the 23-kDa subunit of bovine mitochondrial complex I and sequence analysis strategy. The heavy horizontal lines represent the fragments of DNA produced by polymerase chain reactions, and the arrows show the direction and extent of determined sequences. H* indicates clones that arose in consequence of the so-called "star" activity of *Hin*dIII by cleavage in the sequence AAGCTG or AAGCT. The scale is in bases.

encoding the segment of protein sequence from the products of the polymerase chain reaction. The sequence of this short cDNA is determined. Only the segment of sequence between the priming sites is accurate, and it is used in further polymerase chain reactions to produce cDNAs that extend to the 3′ and 5′ extremities of the mRNA. In both cases, oligo(dT) is used as the second primer, and, in the latter, cDNA tailed with A residues at its 5′ end is used as template.

In the case of the 23-kDa subunit of complex I, a cDNA representing amino acids 1–23 was synthesized by using mixed primers 17 bases long based on residues 1–6 (forward primer) and 18–23 (reverse primer); both primers contained 128 different sequences (see PCR 1 in Figure 2). The cDNA was identified by hybridization with a mixed oligonucleotide probe containing 1024 sequences based on amino acids 10–15. The sequence of the cloned fragment encompassed 68 bases of the cDNA sequence (Figures 2 and 3), and because of the degeneracy of the primers, only bases 18–51 from the central part of this sequence were unambiguous. In a second experiment, the polymerase chain reaction was used to synthesize a cDNA extending from the known sequence to the 3′ poly(A) with primers 2F and 2R (see Figure 3 and PCR 2 in Figure 2). The product of this reaction was about 600 bases long and hybridized strongly with a unique probe containing bases 188–204. A cDNA extending to the 5′ extremity was synthesized with the use of primers 4R and oligo(dT), the latter hybridizing with the poly(A) tail added to the 5′ end of the cDNA population with terminal transferase (PCR 4 in Figure 2). This product was recognized by hybridization with the complement of the forward primer used in experiment PCR 2. Independent verification of sequences extending in both 3′ and 5′ directions was obtained by synthesizing cDNAs covering these regions with primers at the extremities of the known sequence (experiments PCR 3 and PCR 5).

The complete cDNA presented in Figure 3 is 766 nucleotides in length. It is terminated by a poly(A) tail 17 nucleotides to the 3′ side of the sequence AATAAA, a typical polyadenylation signal (Proudfoot & Brownlee, 1976).

Sequence of the 23-kDa Subunit of Bovine Complex I. The mature 23-kDa subunit encoded in the cDNA sequence is 176 amino acids long and has a calculated molecular weight of 20 196, in reasonable agreement with the value of 23 kDa estimated by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. Three in-phase ATG codons are found in the DNA sequence to the 5′ of the sequence coding for the mature subunit, and in the absence of any other information it is assumed that the one nearest to the 5′ extremity is the translational initiation codon. If this is so, a 36 amino acid presequence is attached to the N-terminus of the mature protein and presumably would serve as a signal for mitochondrial import (Nicholson & Neupert, 1988). The absence of acidic amino acids and the presence of four arginine residues in this region support this interpretation. The presence of the sequence RXLXXS at amino acids −10 to −5 indicates that

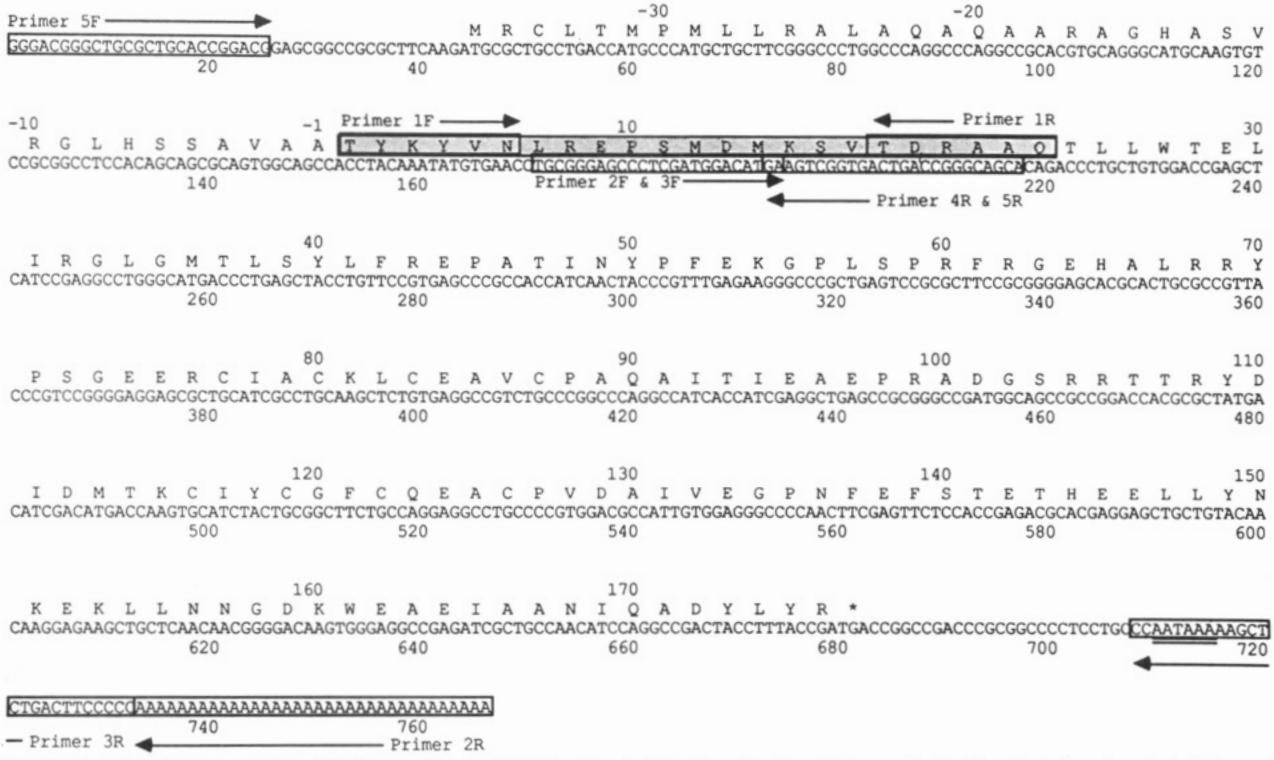

FIGURE 3: Compiled sequence of the cDNA encoding the 23-kDa subunit of bovine mitochondrial complex I. The shaded region of the N-terminus of the protein sequence was determined by sequence analysis of the protein. The boxed protein sequences were used to design the oligonucleotide primers used in experiment PCR 1 (see Figure 2). Boxed nucleotide sequences were used as unique primers in experiments PCR 2, 3, 4, and 5 and are numbered accordingly; F and R indicate forward or reverse primers, respectively. A potential polyadenylation signal is underlined.

```
                              1          11         21         31         41         51
bovine CI-23 kd subunit       TYKYVNLREP SMDMKSVTDR AAQTLLWTEL IRGLGMTLSY LFREPATINY PFEKGPLSPR
tobacco chloroplast frxB         MLPMITE FINYGQQTIR AARYI----- GQGFMITLSH ANRLPVTIQY PYEKLITSER
liverwort chloroplast frxB       MFSIING LKNYNQQAIQ AARYI----- GQGFLVTLDH MNRLPTTIQY PYEKLIPSER
rice chloroplast frxB            MFPMVTG FM--GQQTIR AARYI----- GQSFIITLSH TNRLPITIHY PYEKSITSER
wheat chloroplast frxB        MNMFPMVTG FMSYGQQTIR ATRYI----- GQSFITTLSH TNRLPITIHY PYEKSITPER

                             61         71         81         91        101        110
bovine CI-23 kd subunit       FRGEHALRRY PSGEERCIAC KLCEAVCPAQ AITIEAEPRA DGSR-RTTRY DIDMTKCIYC
tobacco chloroplast frxB      FRG-----RI HFEFDKCIAC EVCVRVCPID LPVVDWKLET DIRKKRLLNY SIDFGICIFC
liverwort chloroplast frxB    FRG-----RI HFEFDKCIAC EVCVRVCPIN LPVVDWELKK TIKKKQLKNY SIDFGVCIFC
rice chloroplast frxB         FRG-----RI HFEFDKCIAC EVCVRVCPID LPLVDWRFEK DIKRKQLLNY SIDFGVCIFC
wheat chloroplast frxB        FRG-----RI HFEFDKCIAC EVCVGVCPID LPVVDWRFEK DIKKKQLLNY SIDFGVCIFC

                            120        130        140        150        160        170
bovine CI-23 kd subunit       GFCQEACPVD AIVEGPNFEF STETHEELLY NKEKLLNNGD KWEAEIAANI QADYLYR
tobacco chloroplast frxB      GNCVEYCPTN CLSMTEEYEL STYDRHELNY NQIALGRLPM SVIDDYTIRT ISNLPQIKNE
liverwort chloroplast frxB    GNCVEYCPTN CLSMTEEYEL STYNRHELNY DQIALGRLPI SIIEDSTIEN IFNLTSLPKG
rice chloroplast frxB         GNCVEYCPTN CLSMTEEYEL STYDRHELNY NQIALSRLPI SIMGDYTIQT IRNSTQSKID
wheat chloroplast frxB        GNCVEYCPTS CLSMTEEYEL STYDRHELNY NQIALSRLPI SIMGDYTIQT IRNSSESKIN

                            180
liverwort chloroplast frxB    K
rice chloroplast frxB         EEKSWNSRTIT
wheat chloroplast frxB        KEKSSNS
```

FIGURE 4: Alignment of the protein sequences of the mature 23-kDa subunit of bovine mitochondrial complex I with chloroplast frxB proteins from tobacco, liverwort, rice, and wheat. Identical amino acids in the aligned sequences are shaded.

the precursor may be processed in two steps (Hendrick et al., 1989), first, by a matrix specific protease after position −9 and, second, by an "intermediate-specific" protease to generate the mature protein.

Homology of the 23-kDa Subunit with Chloroplast Gene Products. The most obvious feature of the sequence of the 23-kDa subunit of complex I is the sequence motif CysXX-CysXXCysXXXCysPro, which occurs twice. In ferredoxins the cysteine residues in such sequences provide the ligands for 4Fe-4S clusters, and they are likely to have a similar role in the 23-kDa subunit of complex I. By analogy with a clostridial ferredoxin, which contains two 4Fe-4S centers (Tanaka et al., 1974), it would be expected that the ligands for one cluster would be provided by cysteines 77, 80, 83, and 126 and for the second cluster by cysteines 116, 119, 122, and 87. The sequence relationship between the 23-kDa subunit and the clostridial ferredoxin is confined to the immediate regions of the cysteine clusters, but a more extensive relationship is found with potential proteins encoded in unidentified open reading frames in the chloroplast genomes of higher plants (Shinozaki et al., 1986; Ohyama et al., 1986, 1988; Hiratsuka et al., 1989; Dunn & Gray, 1988). Because of the presence of the cysteine motifs in these proteins they were called the frxB, or ferredoxin-like proteins, but no function was attributed to them. The proteins from tobacco, liverwort, rice, and wheat are 167, 183, 178, and 176 amino acids long, respectively, and are highly conserved. They are related to the 23-kDa subunit of complex I through much of their sequences (see Figure 4).

In liverwort, tobacco, and rice chloroplast DNAs, *frx*B is part of a cluster of genes that encode five other homologues of subunits of complex I (Figure 5). Four of these are homologues of ND1, ND4, ND4L, and ND6, all mitochondrial gene products in a wide range of species [see Fearnley et al. (1989) for a summary]. The fifth protein, NDH392 in liverwort, NDH393 in tobacco and rice, is homologous to the 49-kDa subunit of the mitochondrial complex I, which is encoded in the nucleus (Fearnley et al., 1989). A homologue of another mitochondrial gene product, ND3, is found at a second locus in the chloroplast genomes of liverwort and tobacco, in association with a gene encoding a protein that is homologous to part of the 30-kDa subunit of complex I, which like the 49- and 23-kDa subunits is coded in the nucleus (Pilkington et al., 1991a). In view of their relationships with the 49-, 23-, and 30-kDa subunits of bovine mitochondrial complex I, it is proposed that ORF393 (or *ndh*393), ORF169, and *frx*B in tobacco should be renamed *ndh*H, *ndh*I, and *ndh*J, respectively, and that the same names should be applied to the equivalent genes in chloroplast genomes in other species.

The clustering of chloroplast genes with related functions, belonging to the same assembly, has been noted previously, for example, in relation to chloroplast-coded components of the ATP synthase complex. In the latter case, attention has been drawn to the similarity of the arrangements of genes at these loci to the arrangements of equivalent genes in bacterial operons, and this has been interpreted as evidence of an endosymbiotic origin of chloroplasts and also of evidence of separate modular evolution of the F_1 and F_o structural domains of the ATP synthase complex (Walker et al., 1987a; Cozens & Walker, 1987; Falk & Walker, 1988). In mitochondria from *Neurospora crassa*, evidence has been presented for the evolution of complex I from two ancestral complexes; one of them appears to contain the equivalents of the soluble IP and FP fractions of the bovine complex and to have contributed the first part of the electron pathway from NADH to intermediate acceptors. Independently, evidence that the 75-, 53-, and 24-kDa subunits of bovine complex I constitute a subcomplex involved in the first part of the electron pathway has been provided by the finding that sequences closely related to all three are present in the α and γ subunits of a soluble bacterial NAD^+-reducing hydrogenase (Pilkington et al., 1991b). The $\alpha\gamma$ dimer is an NADH oxidoreductase and contains one molecule of FMN (Tran-Betcke et al., 1990). The other *Neurospora* complex appears to contain all of the mitochondrially coded subunits and the N-2 iron–sulfur center and is suggested to have given rise to the later steps of electron transfer, ultimately to ubiquinone (Tuschen et al., 1990). There appears to be considerable overlap between the constituents of this latter complex and the complement of complex I homologues detected in the chloroplast genomes, and so a reasonable interpretation of the existing information is that

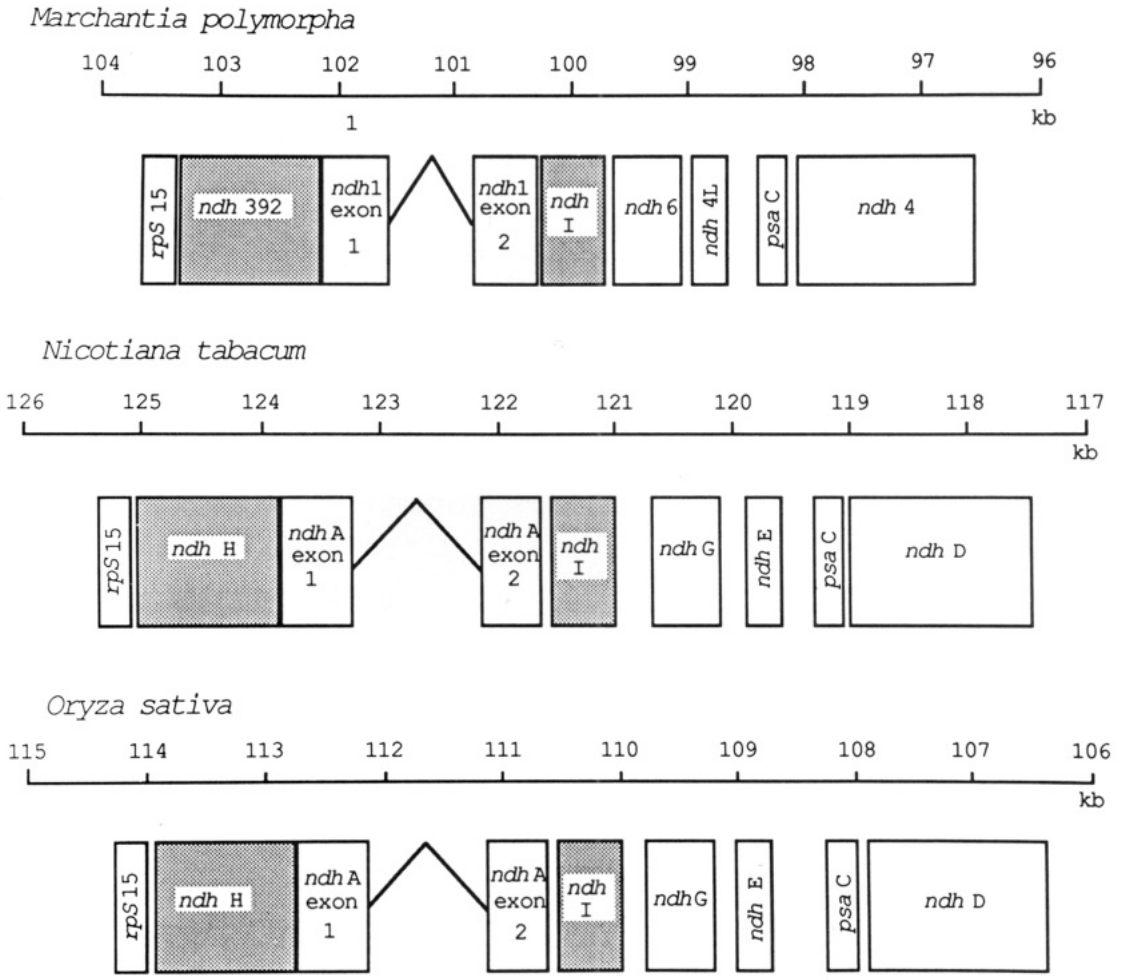

FIGURE 5: Arrangement of chloroplast genes and open reading frames in the vicinity of *frx*B. The diagram is based upon data published by Shinozaki et al. (1986), Kohchi et al. (1988), Wolfe and Sharp (1988), and Hiratsuka et al. (1989). The shaded regions designated *ndh*H (formerly known as ORF392 or *ndh*392 and ORF393 or *ndh*393) are homologous to the nuclear encoded 49-kDa subunit of bovine mitochondrial complex I (Fearnley et al., 1989), and *ndh*A (*ndh*1), *ndh*6, and *ndh*E (*ndh*4L) encode homologues of subunits of NADH dehydrogenase that are encoded in mitochondrial DNA. As described in the text, *ndh*I genes (also shaded; formerly *frx*B) are now shown to be the chloroplast equivalents of the 23-kDa subunit of mitochondrial complex I. The *rps*15 genes are homologues of the *E. coli* gene for ribosomal protein S15, and *psa*C (formerly *frx*A) is part of photosystem I (Oh-oka et al., 1987; Høj et al., 1987).

this group of chloroplast genes represent some or all of the components of a structural unit which forms part of mitochondrial complex I. It has been suggested that they represent part of an enzyme thought to be an NADH:plastoquinone oxidoreductase involved in a chlororespiratory pathway, and such an enzyme has been identified in *Chlamydomonas reinhardtii* chloroplasts (Bennoun, 1982; Maione & Gibbs, 1986). However, it should not be forgotten that homologues of some of the mitochondrially encoded subunits of complex I have also been shown to be part of formate hydrogenlyase in *Escherichia coli* (Böhm et al., 1990), and so it is possible that the chloroplast subunits are part of some other as yet unknown activity.

Location and Function of the 23-kDa Subunit in Bovine Mitochondrial Complex I. The 23-kDa subunit is part of the HP, or hydrophobic, subfraction of bovine mitochondrial complex I, the residuum after the removal of IP and FP fractions with chaotropes. This fraction is rather heterogeneous and poorly characterized. It probably contains the seven subunits of complex I that are encoded in mitochondrial DNA (ND1–6 and ND 4L) and at least twelve other proteins. It has been estimated that 6–7 of the 22–24 iron atoms present in complex I are in this cluster, possibly in a 4Fe-4S cluster, designated N-2, and in one or two other uncharacterized 2Fe-2S clusters (Ragan, 1987). The hydrophobic fraction is believed to contain the terminal part of the electron-transfer pathway in complex I. Because of its high redox potential, cluster N-2 is likely to be the donor to ubiquinone, and in keeping with operation in a hydrophobic environment, its midpoint potential depends on the phospholipid content of complex I. None of the seven subunits of complex I encoded in mitochondrial DNA contain cysteine-rich sequence motifs that are associated with iron–sulfur centers, nor are they evident in seven other components of the HP fraction that we have characterized (unpublished work). Therefore, the 23-kDa subunit is a strong candidate for the protein that bears the N-2 cluster, although from its sequence, it appears to contain two tetranuclear centers. This indicates that at least one other previously unidentified 4Fe-4S cluster is present in the hydrophobic fraction. The question of whether the 23-kDa subunit is an intrinsic membrane protein remains to be investigated. Its hydrophobic profile (not shown), together with the requirements for liganding of the proposed iron–sulfur centers, suggests that it probably is not.

Relevant information about the nature and function of this group of proteins has emerged from studies of a protein isolated from thylakoid membranes of *C. reinhardtii*, which N-terminal sequence analysis shows to be an frxB homologue. It can be released from the thylakoids by salt washes and is soluble in aqueous buffers. Its electron paramagnetic resonance spectra are similar to those of the clostridial 2[4Fe-4S] ferredoxin, and so this provides direct evidence for the presence of iron–

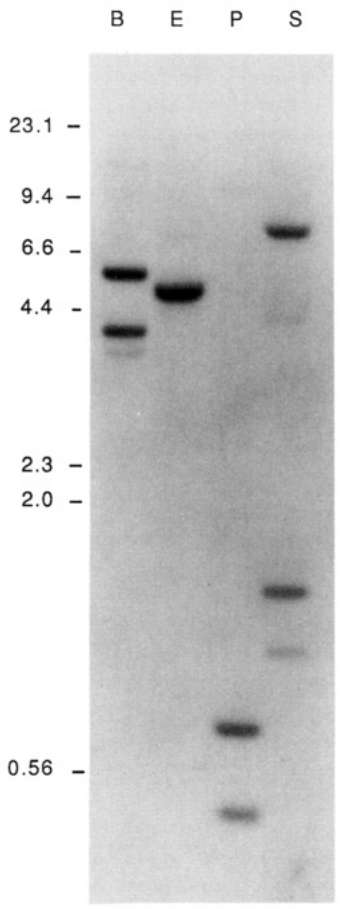

FIGURE 6: Hybridization of restriction digests of bovine genomic DNA with the cDNA for the 20-kDa protein. The letters B, E, P, and S denote the restriction enzymes *Bam*HI, *Eco*RI, *Pst*I, and *Sac*I used in the various digests. For experimental details see Materials and Methods. On the left-hand side are shown the sizes in kilobases of DNA markers produced by digestion of DNA from bacteriophage λ with *Hin*dIII. The probe was nucleotides 171–732 of the cDNA sequence.

sulfur centers in this family. Apart from being in the thylakoid membrane where it may be part of a NADH:plastoquinone oxidoreductase, the *C. reinhardtii* protein also binds to the chloroplast DNA replication origin and could act, for example, in reduction of ribonucleotide primers at the moment of initiation of DNA replication or in coordination of initiation of DNA replication with available reducing power (Wu et al., 1989). It remains to be shown that these hypotheses are correct and whether the 23-kDa subunit of complex I fulfills a similar role in mitochondria.

Number of Bovine Genes for the 23-kDa Subunit. Nuclear-encoded mitochondrial proteins such as the c subunit (Gay & Walker, 1985; Dyer et al., 1989; Dyer & Walker, 1990) and the α subunit (Walker et al., 1989) of ATP synthase as well as the ADP/ATP translocator (Battini et al., 1987; Walker et al., 1987a; Neckelman et al., 1987; Houldsworth & Attardi, 1988; Powell et al., 1989) have been shown to be encoded by more than one gene. Often, the different copies of the genes are expressed in a tissue-specific manner (Gay & Walker, 1985). The biological significance of such a duplication is not always clear, but it has been proposed in some instances to allow the molding of the mitochondrial enzyme to the metabolic specificity of a given tissue (Lomax & Grossman, 1989).

In order to address the question of the number of genes coding for the bovine 23-kDa subunit of complex I, a fragment of its cDNA was hybridized with digests of bovine genomic DNA. Only one or two strongly hybridizing bands were found in each digest (Figure 6). In the case of the *Sac*I digest, a minimum of two hybridizing bands was expected since there is a *Sac*I restriction site at position 236 in the cDNA and two strongly hybridizing bands were observed. A single strongly hybridizing band was found in the *Eco*RI digest, and so these results suggest that the 23-kDa subunit of bovine complex I is probably encoded by a unique gene and that there is not a large gene family as in other cases.

Acknowledgments

We are grateful to Mr. T. V. Smith and Mrs. J. Fogg for the preparation of synthetic oligonucleotides and to Dr. G. Gibb for the gift of a sample of complex I.

References

Battini, R., Ferrari, S., Kaczmarek, L., Calabretta, B., Chen, S. T., & Baserga, R. (1987) *J. Biol. Chem. 262*, 4355–4359.
Bennoun, P. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4352–4356.
Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.
Böhm, R., Sauter, M., & Böck, A. (1990) *Mol. Microbiol. 4*, 231–244.
Chomyn, A., Mariottini, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R. F., & Attardi, G. (1985) *Nature (London) 314*, 592–597.
Chomyn, A., Cleeter, M. W. J., Ragan, C. I., Riley, M., Doolittle, R. F., & Attardi, G. (1986) *Science 234*, 614–618.
Cozens, A. L., & Walker, J. E. (1987) *J. Mol. Biol. 194*, 359–383.
Duckworth, M. L., Gait, M. J., Goelet, P., Hong, G. F., Singh, M., & Titmas, R. (1981) *Nucleic Acids Res. 9*, 1691–1706.
Dunn, P. P. J., & Gray, J. C. (1988) *Nucleic Acids Res. 16*, 348.
Dyer, M. R., & Walker, J. E. (1990) *J. Mol. Biol.* (in press).
Dyer, M. R., Gay, N. J., & Walker, J. E. (1989) *Biochem. J. 260*, 249–258.
Falk, G., & Walker, J. E. (1988) *Biochem. J. 254*, 109–122.
Fearnley, I. M., Runswick, M. J., & Walker, J. E. (1989) *EMBO J. 8*, 665–672.
Gay, N. J., & Walker, J. E. (1985) *EMBO J. 4*, 3519–3524.
Hatefi, Y. (1978) *Methods Enzymol. 53*, 11–14.
Hendrick, J. P., Hodges, P. E., & Rosenberg, L. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4056–4060.
Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C. R., Meng, B. Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., & Sugiura, M. (1989) *Mol. Gen. Genet. 217*, 185–194.
Høj, P. B., Svendsen, I., Scheller, H. V., & Møler, B. L. (1987) *J. Biol. Chem. 262*, 12676–12684.
Houldsworth, J., & Attardi, G. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 377–381.
Kohchi, T., Shirai, H., Fukuzawa, H., Sano, T., Komano, T., Umesono, K., Inokuchi, H., Ozeki, H., & Ohyama, K. (1988) *J. Mol. Biol. 203*, 353–372.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Lipman, D. J., & Pearson, W. R. (1985) *Science 227*, 1435–1441.
Lomax, M. I., & Grossman, L. I. (1989) *Trends Biochem. Sci. 14*, 501–503.
Maione, T. E., & Gibbs, M. (1986) *Plant Physiol. 80*, 364–368.
Neckelmann, N., Li, K., Wade, R. P., Shuster, R., & Wallace, D. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7580–7584.

Nicholson, D. W., & Neupert, W. (1988) in *Protein Transfer and Organelle Biogenesis*, pp 677–746, Academic Press, New York.
O'Farrell, P. H. (1975) *J. Biol. Chem. 250*, 4007–4021.
Oh-oka, H., Takahashi, Y., Wada, K., Matsubara, H., Ohyama, K., & Ozeki, H. (1987) *FEBS Lett. 218*, 52–54.
Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., & Ozeki, H. (1986) *Nature (London) 322*, 571–574.
Ohyama, K., Fukuzawa, H., Kohchi, T., Sano, T., Sano, S., Shirai, H., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., & Ozeki, H. (1988) *J. Mol. Biol. 203*, 281–298.
Pilkington, S. J., & Walker, J. E. (1989) *Biochemistry 28*, 3257–3264.
Pilkington, S. J., Skehel, J. M., & Walker, J. E. (1991a) *Biochemistry 30*, 1901–1908.
Pilkington, S. J., Skehel, J. M., Gennis, R. B., & Walker, J. E. (1991b) *Biochemistry 30*, 2166–2175.
Powell, S. J., Medd, S. M., Runswick, M. J., & Walker, J. E. (1989) *Biochemistry 28*, 866–873.
Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature (London) 263*, 211–214.
Ragan, C. I. (1987) *Curr. Top. Bioenerg. 15*, 1–36.
Ragan, C. I., Wilson, M. T., Darley-Usmar, V. M., & Lowe, P. N. (1987) in *Mitochondria: A Practical Approach* (Darley-Usmar, V. M., Rickwood, D., & Wilson, M. T., Eds.) pp 79–112, IRL Press, Oxford, Washington.
Runswick, M. J., Gennis, R. B., Fearnley, I. M., & Walker, J. E. (1989) *Biochemistry 28*, 9452–9459.
Runswick, M. J., Walker, J. E., Bisaccia, F., Iacobazzi, V., & Palmieri, F. (1990) *Biochemistry 29*, 11033–11040.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., & Sugiura, M. (1986) *EMBO J. 5*, 2043–2049.
Staden, R. (1982a) *Nucleic Acids Res. 10*, 4731–4751.
Staden, R. (1982b) *Nucleic Acids Res. 10*, 2951–2961.
Staden, R. (1985) in *Genetic Engineering: Principles and Methods* (Setlow, J. K., Ed.) pp 67–114, Plenum Publishing Corporation, New York, London.
Stout, C. D. (1982) in *Iron Sulfur Proteins* (Spiro, T. G., Ed.) pp 97–146, J. Wiley & Sons, New York.
Tanaka, R., Haniu, M., Yasunobu, K. T., Jones, J. B., & Stadtman, T. C. (1974) *Biochemistry 13*, 5284–5289.
Tran-Betcke, A., Warnecke, U., Böcker, C., Zabarosch, C., & Friedrich, B. (1990) *J. Bacteriol. 172*, 2920–2929.
Tuschen, G., Sackmann, U., Nehls, U., Haiker, H., Buse, G., & Weiss, H. (1990) *J. Mol. Biol. 213*, 845–857.
Walker, J. E., Cozens, A. L., Dyer, M. R., Fearnley, I. M., Powell, S. J., & Runswick, M. J. (1987a) *Chem. Scr. 27B*, 97–105.
Walker, J. E., Gay, N. J., Powell, S. J., Kostina, M., & Dyer, M. R. (1987b) *Biochemistry 26*, 8613–8619.
Walker, J. E., Powell, S. J., Viñas, O., & Runswick, M. J. (1989) *Biochemistry 28*, 4702–4708.
Wolfe, K. H., & Sharp, P. M. (1988) *Gene 66*, 215–222.
Wu, M., Nie, Z. Q., & Yang, J. (1989) *Plant Cell 1*, 551–557.