THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 14, Issue of April 6, pp. 11302–11309, 2001
Printed in U.S.A.

Subunit II of Cytochrome *c* Oxidase in Chlamydomonad Algae Is a Heterodimer Encoded by Two Independent Nuclear Genes*

Received for publication, November 9, 2000
Published, JBC Papers in Press, November 27, 2000, DOI 10.1074/jbc.M010244200

Xochitl Pérez-Martínez‡, Anaid Antaramian‡, Miriam Vázquez-Acevedo‡, Soledad Funes‡, Elena Tolkunova§, Jacques d'Alayer¶, Manuel G. Claros‖, Edgar Davidson§, Michael P. King§, and Diego González-Halphen‡

From the ‡Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Apartado Postal 70-243, México 04510, D.F. Mexico, §Department of Biochemistry and Molecular Pharmacology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, ¶Laboratoire de Microséquençage des Protéines, Département des Biotechnologies, Institut Pasteur, 75724 Paris Cedex 15, France, and ‖Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, Universidad de Málaga, E-29071 Málaga Spain

The mitochondrial genomes of Chlamydomonad algae lack the *cox2* gene that encodes the essential subunit COX II of cytochrome *c* oxidase. COX II is normally a single polypeptide encoded by a single mitochondrial gene. In this work we cloned two nuclear genes encoding COX II from both *Chlamydomonas reinhardtii* and *Polytomella* sp. The *cox2a* gene encodes a protein, COX IIA, corresponding to the N-terminal portion of subunit II of cytochrome *c* oxidase, and the *cox2b* gene encodes COX IIB, corresponding to the C-terminal region. The *cox2a* and *cox2b* genes are located in the nucleus and are independently transcribed into mRNAs that are translated into separate polypeptides. These two proteins assemble with other cytochrome *c* oxidase subunits in the inner mitochondrial membrane to form the mature multi-subunit complex. We propose that during the evolution of the Chlorophyte algae, the *cox2* gene was divided into two mitochondrial genes that were subsequently transferred to the nucleus. This event was evolutionarily distinct from the transfer of an intact *cox2* gene to the nucleus in some members the Leguminosae plant family.

Mitochondria are thought to descend from free-living α-proteobacteria (1), whose closest extants are bacteria of the genus *Rickettsia* (2). After the endosymbiotic event, there was a transfer of genes from the protomitochondrion to the nucleus such that few genes now remain in mitochondrial genomes (3). Those genes that remain encode only a sub-set of the mitochondrial proteins needed for oxidative phosphorylation and a portion of the factors necessary for their expression (4). The mtDNA-encoded respiratory chain subunits are highly hydrophobic polytopic proteins that contain two or more transmembrane stretches (5, 6). The genes for *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6* (encoding subunits 1, 2, 3, 4, 4L, 5, and 6 of NADH:ubiquinone oxidoreductase), *cob* (encoding cytochrome *b* of the bc_1 complex), *cox1*, *cox2*, and *cox3* (encoding subunits COX I, COX II, and COX III of cytochrome *c* oxidase), and *atp6* and *atp8* (encoding subunits *a* and A6L of the F_0 portion of ATP synthetase) are found in the mitochondrial genomes of most organisms.

The transfer of mitochondrial genes to the nucleus is an ongoing process, as shown by the presence of particular genes encoded in both the mitochondrial and the nuclear genomes in certain species. These are exemplified by COX II in some members of the family leguminosae (7–9) and by subunit 9 of ATP synthetase in *Neurospora crassa* (10).

The algae of the family Chlamydomonadaceae, including *Chlamydomonas reinhardtii* and *Polytomella* sp., lack some of the genes that are typically found in mitochondrial genomes, including *nad3*, *nad4L*, *cox2*, *cox3, atp6*, and *atp8* (11–13)[1]. We have shown that, in at least two members of this family, *C. reinhardtii* and *Polytomella* sp., the *cox3* gene was transferred to the nucleus, and the corresponding mitochondrial copy has been lost (14).

Several mitochondrial respiratory chain complexes have been isolated and characterized from the colorless alga *Polytomella* sp. (15–17), a close relative of *Chlamydomonas*, including an active, cyanide-sensitive cytochrome *c* oxidase preparation (14). In this work, we show that COX II is present as a heterodimer in this complex. All COX II sequences that have been described to date are single polypeptides encoded by one gene, normally in the mitochondrial genome. In both *Polytomella* sp. and *C. reinhardtii*, COX II is encoded by two nuclear genes that were named *cox2a* and *cox2b*. The *cox2a* gene encodes a protein, COX IIA, corresponding to the N-terminal half of a typical one-polypeptide COX II, and the *cox2b* gene encodes a protein, COX IIB, equivalent to the C-terminal half of the same subunit. We propose that these separate genes give rise to a heterodimeric COX II that results from the noncovalent assembly of the COX IIA and COX IIB polypeptides in cytochrome *c* oxidase of the inner mitochondrial membrane.

EXPERIMENTAL PROCEDURES

Strain and Culture Conditions—Polytomella sp. (198.80, E. G. Pringsheim), from the Sammlung von Algenkulturen (Gottingen, Germany), was grown as previously described (15).

* This work was supported by National Institutes of Health Grants TW01176 and HL59646, Consejo Nacional de Ciencia y Tecnologia, Mexico Grant 27754N, and Dirección General de Apoyo al Personal Académico, Universidad Nacional Autónoma de México Grant IN202598. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF305078 (Polytomella sp. cox2a cDNA), AF305079 (Polytomella sp. cox2b cDNA), AF305080 (C. reinhardtii cox2a cDNA), AF305540 (C. reinhardtii cox2b cDNA), AF305541 (Polytomella sp. genomic cox2a), AF305542 (Polytomella sp. genomic cox2b), and AF305543 (C. reinhardtii genomic cox2b).

** To whom correspondence should be addressed. Tel.: 011-525-622-5620; Fax: 011-525-622-5611 or 011-525-548-0387; E-mail: dhalphen@ifisiol.unam.mx.

[1] S. Funes, A. Antaramian, and D. González-Halphen, unpublished results.

Protein Analysis—Cytochrome *c* oxidase from *Polytomella* sp. was obtained as previously described (14). Polyacrylamide gel electrophoresis was performed as in Schägger *et al.* (18) using 16% acrylamide gels. For tryptic digestion analysis, gels were stained with Amido Black, and the polypeptide of interest was excised from the gel. Polypeptides were isolated as previously described (15) for N-terminal sequencing. Amino-terminal Edman degradation was carried out on an Applied Biosystems Sequencer at the Laboratoire de Microséquençage des Protéines, Institut Pasteur, Paris, France. An 18.6-kDa polypeptide was isolated from polyacrylamide gels and subjected to tryptic and endolysin-C digestion and separation on DEAE-C14 and DEAE-C18 HPLC[2] columns. Peaks eluted from the columns were subjected to N-terminal sequence analysis.

Nucleic Acids Preparation—Total DNA and total RNA from *Polytomella* sp. and *C. reinhardtii* were isolated as previously described (14). PCR fragments were cloned into pMOS blue-T (Amersham Pharmacia Biotech) or pGEM-T easy (Promega). cDNA was prepared from 1–2 μg of total RNA with Moloney murine leukemia virus reverse transcriptase (Promega) or Superscript II reverse transcriptase (Life Technologies). All standard molecular biology techniques were as described (19). Sequencing was carried out by the Kimmel Cancer Center DNA Sequencing Facility, Thomas Jefferson University, and at the Unidad de Biología Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México.

Cloning the cox2b Gene from Polytomella sp.—A genomic *Polytomella* sp. *cox2b* fragment was amplified by PCR using two degenerate primers: F1 (5′- CA(A/G) GA(C/T) AG(C/T) GC(C/T) AC(A/T) AG(C/T) CA(G/A) GC(C/T) CA(A/G) G-3′) based on internal sequence (QDSATSQAQA) of COX II, and B1 (5′-TG (G/A)TT (C/T)AA (A/G)CG (A/T)CC (A/T)GG (A/G)AT (A/G)GC (A/G)TC CAT-3′) based on the internal sequence MDAIPGRLNQ. The resulting 300-nt product was used to isolate genomic *cox2b* clones from a library of *Polytomella* DNA *Pst*I fragments of ~2 kilobases in length in pBluescript.

The *cox2b* cDNA sequence from *Polytomella* sp. was obtained by PCR and 5′-RACE (20) using primers based on the genomic sequence obtained above. A poly(dT) tail was added to the 5′ end of the cDNA with a terminal transferase (Roche Molecular Biochemicals). The forward primer was oligo(dT)/adapter primer: 5′-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTTT-3′, and the reverse primer (B2) was 5′-AGCTGTTTAAGACCATGACTTC-3′.

Cloning the cox2a Gene from Polytomella sp.—A *cox2a* cDNA fragment was amplified using the degenerate primers F2 (5′-GA(G/A) GC(T/C) CC(T/C) GT(T/C) GC(T/C) TGG CAG CT(T/G) GG-3′), based on the N-terminal protein sequence EAPVAWQLG, and B3 (5′-CCA (A/G)TA CCA CTG (A/G)TG (A/G/T)CC (A/G)AT (A/G)GC C-3′), based on the internal conserved sequence KAIGHQWYW. Nested PCR was done with degenerate primers F3 (5′-CAG GA(T/C) TC(T/C) GC(T/C) AC(T/C) TC(T/C) CAG GC(T/C) CAG G-3′), based on the N-terminal sequence of the protein QDSATSQAQA and B4 (5′-GA(A/G) TA(A/G) AG(A/G) AG(A/G) GC(A/G) AA(A/G)GA(A/G) GG-3′), based on the internal conserved sequence PSFALLYS. For 3′-end cDNA cloning, oligo(dT)/adapter primer and primer F4 (5′-TCCTCTACCACATCGCCACCC-3′) were used. For nested PCR, adapter (5′-GACTCGAGTCGACATCGA-3′) and primer F5 (5′-ACTACACTAAGCAAGCTCTCCCTG-3′) were used. For 5′-RACE, primers B5 (5′-TCAGGGAGAGCTTGCTTAGTGTAG-3′), with oligo(dT)/adapter primer, and B6 (5′-TTGGTGGCGATGTGGTAGAGG-3′), with adapter for nested PCR, were used. Primers F6 (5′-AATGCTCGCCCAGCGTATC-3′) and B7 (5′-AAACCTTCACACACCCATAGGC-3′), derived from the 5′- and 3′-RACE products, were used to amplify the full-length cDNA of the *cox2a* gene. The genomic sequence of the *Polytomella* sp. *cox2a* gene was obtained by PCR amplification of total *Polytomella* sp. DNA with the same primers.

Cloning the cox2a and cox2b Genes of C. reinhardtii—The cDNAs for *cox2a* and *cox2b* were obtained by screening a *C. reinhardtii* cDNA library in λgt10 (21) using the *Polytomella* sp. *cox2a* cDNA or *cox2b* genomic DNA as probes. A total of 17 positive clones were obtained from 5×10^4 plaque-forming units screened. PCR using two primers based on λgt10 sequences was used to identify the positive clones with the largest inserts. Phage DNA from these clones was isolated with the Qiagen Lambda mini kit. The 5′ end of the *cox2a* cDNA was completed by RACE using primers based on the cDNA sequence obtained.

A bacterial artificial chromosome clone containing *cox2b* was obtained from a BAC genomic library from *C. reinhardtii* (22) by Genome Systems using the *C. reinhardtii cox2b* cDNA obtained above. BAC DNA was sequenced directly using internal primers.

Protein Sequence Analysis—Mitochondrial targeting sequences were analyzed using MitoProt II (23, 24). The same program was used to calculate the segments with high local hydrophobicity (<H>) in a distance comprising 13 to 17 residues. The *meso*H was determined by scanning each sequence for a maximum average hydrophobicity measured in windows from 60 to 80 residues and averaging the values (24). More hydrophobicity scales were included to reduce the possibility of bias. Protein transmembrane stretches were predicted using the program TodPred II (25). Three-dimensional structure modeling was carried out using SWISS-MODEL (26).

Data Base Accession Numbers—The nucleotide sequences are in the DDBJ/EMBL/GenBank™ nucleotide sequence data base under the accession numbers AF305078 (*Polytomella* sp. *cox2a* cDNA), AF305079 (*Polytomella* sp. *cox2b* cDNA), AF305080 (*C. reinhardtii cox2a* cDNA), AF305540 (*C. reinhardtii cox2b* cDNA), AF305541 (*Polytomella* sp. genomic *cox2a*), AF305542 (*Polytomella* sp. genomic *cox2b*), and AF305543 (*C. reinhardtii* genomic *cox2b*).

RESULTS

Subunit II Is Present in the Cytochrome c Oxidase Complex from Polytomella sp.—An active cytochrome *c* oxidase from *Polytomella* sp. was isolated as previously described (14). The four largest polypeptides with apparent molecular masses of 54.6, 29.6, 18.6, and 14.5 kDa were subjected to Edman degradation. The 54.6-kDa polypeptide, not susceptible to Edman degradation, was identified as subunit I of cytochrome *c* oxidase (COX I) since its mass was similar to that predicted for the mitochondrial *cox1* gene sequence (54,781 Da (27)). The 29.6-kDa polypeptide (N-terminal sequence: SSDAGHHLSPRERYLV) was previously identified as subunit III (14). The N-terminal sequence of the 14.5-kDa polypeptide, DANSSELVLEPTRKYKAGLATRELW, did not show similarity with any sequences in GenBank™.

The N-terminal sequence of the 18.6-kDa polypeptide, EAPVAWQLGFQDSATSQAQA, was similar to COX II from other species, allowing identification of this polypeptide as COX II. A second sequence, MDAIPGR(R/L)NQIWLTINREG, was obtained during the Edman degradation of the 18.6-kDa polypeptide region at a yield of less than 50% that of the yield of the first sequence. This sequence also was similar to COX II from other species and was identified as an internal fragment of COX II that was obtained after partial cleavage of the protein during Edman degradation.

Cloning of the cox2b Gene from Polytomella sp. and C. reinhardtii—On the basis of the primary amino acid sequences obtained for COX II, two degenerate oligodeoxynucleotide primers were designed. Using these primers, a PCR amplification product of 300 nt was obtained using total DNA from *Polytomella* sp. as a template. This PCR product encoded the C-terminal portion of the *cox2* gene but lacked the region that encodes the N-terminal sequence of COX II. The absence of an N-terminal sequence was attributed to nonspecific annealing of the primer based on the N-terminal amino acid sequence.

The 300-nt *cox2* gene fragment hybridized to a 2-kilobase *Pst*I fragment of *Polytomella* sp. total DNA in Southern analyses. To obtain a full-length gene, a mini-library was constructed from *Pst*I fragments of ~2 kilobases, and a positive clone was isolated and sequenced. This genomic sequence contained a 462-nt open reading frame that encoded a 153-amino acid protein homologous to the C terminus of COX II. The gene was named *cox2b*. There was no open reading frame corresponding to the N-terminal portion of COX II in the 960 nt preceding the 462-nt open reading frame.

Using primers based on the genomic sequence, a portion of a cDNA corresponding to the *cox2b* open reading frame was

[2] The abbreviations used are: HPLC, high precision liquid chromatography; *meso*H mesohydrophobicity; mtDNA mitochondrial DNA; nt, nucleotide(s); PCR, polymerase chain reaction; <H>, maximum local hydrophobicity of a sequence segment; μH, hydrophobic moment; RACE, rapid amplification of cDNA ends; BAC, bacterial artificial chromosome.

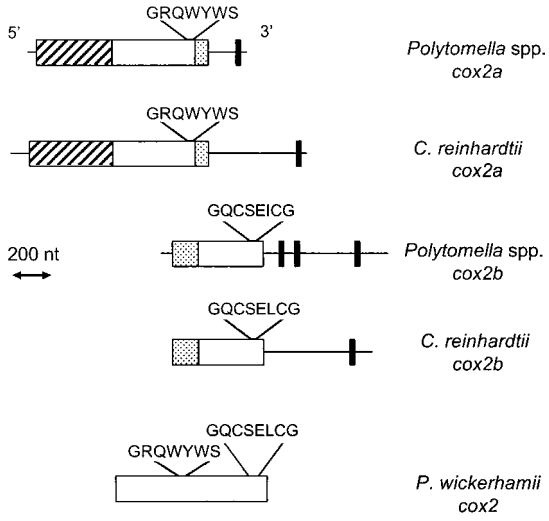

FIG. 1. **Organization of the cox2a and cox2b genes of Chlamydomonad *algae.*** The coding regions of *cox2a* and *cox2b* cDNAs from *Polytomella* sp. and *C. reinhardtii* are indicated by a *box*. The regions that are conserved with other COX II proteins are *white*, the putative mitochondrial-targeting sequences are *striped*, and the highly charged sequences unique to the Chlamydomonad *cox2* genes are *stippled*. Nontranslated regions are shown as a *black line*. Putative polyadenylation signals are indicated as *vertical black bars*. The locations of the highly conserved COX II sequences GRQWYWSY and GQCSE(I/L)CG are indicated above the boxes. The single *cox2* mitochondrial gene of *P. wickerhamii,* representing a canonical *cox2* gene, is included for comparison.

amplified from *Polytomella* sp. total RNA using reverse transcription-PCR. The remainder of the 5′-cDNA sequence was obtained by 5′-RACE. The full-length cDNA obtained contained a 65-nt 5′-untranslated region and a 462-nt open reading frame identical to that of the genomic clone. This showed that this gene encoded only the C-terminal portion of COX II and did not contain introns. The overall organization of the *cox2b* gene is shown in Fig. 1. The predicted protein contained the MDAIPGRLNQIWLTINREG internal sequence that was obtained from direct protein sequencing as well as the sequence GQCSEICG known to be the COX II binding site for binuclear copper (28). The major N-terminal sequence of COX II, determined by Edman degradation, was not present in the deduced protein sequence. The N-terminal 43 amino acids lacked homology to any COX II proteins. The remaining sequence was homologous with the C-terminal half of many COX II proteins. The highest similarity was with COX II from the alga *Prototheca wickerhamii* (29). Altogether, these data suggested that the *cox2* gene had been split into two genes in *Polytomella* sp.

The PCR amplification product of the *cox2b* gene of *Polytomella* sp. was also used to isolate a *cox2b* cDNA from a λgt10 cDNA library of *C. reinhardtii*. A comparison of this *cox2b* cDNA with the genomic sequence (see below) showed that this *cox2b* cDNA sequence lacked the first four codons of the *cox2b* gene and a 5′-untranslated region. It exhibited 85% identity with the *Polytomella* sp. *cox2b* (Fig. 2). The *cox2b* cDNA from *C. reinhardtii* was used to isolate *cox2b* from a *C. reinhardtii* BAC genomic library. The genomic *cox2b* sequence contained the complete coding region and was identical to the cDNA except for the presence of one intron of 187 nt, located 21 nucleotides upstream of the stop codon. There was no open reading frame encoding a known protein in the 1.7 kilobases preceding the start codon for the *C. reinhardtii cox2b* gene. The predicted COX IIB protein was 153 amino acids long and was 85% identical and 92% similar to COX IIB of *Polytomella* sp. Both *C. reinhardtii* and *Polytomella* sp. are therefore likely to use two genes to code for COX II.

Cloning the cox2a Gene from Polytomella sp. and C. reinhardtii—To clone the gene that encoded the N-terminal region of COX II, nested PCR was performed with primers derived from the major N-terminal sequence obtained from the protein, and internal sequences (KAIGHQWYW and PSFALLYS) were conserved among COX II proteins. Using *Polytomella* sp. cDNA as a template, a 250-nt PCR product was obtained that exhibited similarity with other *cox2* genes. The full-length cDNA, obtained using 5′- and 3′-RACE (Fig. 1), contained an open reading frame of 816 nt predicted to encode a protein of 271 amino acids including the sequence EAPVAWQLGF determined for the N terminus of the 18.6-kDa polypeptide. This gene was named *cox2a*, since it encoded the N-terminal portion of the COX II protein (COX IIA).

The N terminus of the mature *Polytomella* sp. COX IIA protein, derived from direct sequencing, corresponded to Glu-131 of the predicted sequence, indicating that this protein contains a mitochondrial targeting sequence of 130 amino acids. The mature protein is predicted to be 141 amino acids long and to contain two putative transmembrane stretches, from Ile-28 to Thr-48 and from Val-69 to Leu-89, and the highly conserved sequence GRQWYWSY present in all sequences of COX II subunits known to date (Fig. 2). The sequence from Glu-131 to Glu-246 was homologous with the N-terminal portion of many COX II proteins. This sequence was most similar to COX II from the alga *P. wickerhamii*. The COX IIA protein contained a C-terminal 20-amino acid region, lacking similarity to conventional COX II proteins, that had a high density of charged amino acids.

Primers corresponding to the 5′ and 3′ ends of the *cox2a* cDNA sequence of *Polytomella* sp. were used to amplify a portion of the *cox2a* nuclear gene from total DNA. This 1773-nt PCR product was cloned and sequenced. The genomic *cox2a* gene contained 6 introns, ranging in size from 84 to 136 nt.

The *Polytomella* sp. *cox2a* cDNA was used to isolate a partial cDNA clone of *cox2a* from a *C. reinhardtii* cDNA library. The complete sequence of *C. reinhardtii cox2a* was obtained by 5′-RACE (Fig. 1). The *C. reinhardtii cox2a* cDNA contained a 5′- untranslated region of 30 nt, an open reading frame of 855 nt, encoding a protein of 284 amino acids, and a 3′-untranslated region of 201 nt. The predicted *C. reinhardtii* COX IIA mature polypeptide exhibited 72% identity and 81% similarity with the sequence predicted for the COX IIA protein from *Polytomella* sp. (Fig. 2). The gene sequence predicts an extension of 21 residues at the C-terminal end that lacked homology to COX II proteins but that was highly similar to the extension predicted by the *cox2a* gene from *Polytomella* sp.

The cox2a and cox2b Genes of Polytomella sp. and C. reinhardtii Are Nuclear-localized, Single-copy Genes—The nuclear location of both *cox2a* and *cox2b* was confirmed by Southern analysis. After electrophoresis of *Polytomella* sp. total DNA through agarose gels, the mtDNA was detected as a discrete band below the bulk of the nuclear DNA (Fig. 3*A*, *left panel*). This was confirmed with a *cox1* gene probe (27) that hybridized with the mtDNA and a β-tubulin gene probe (30) that hybridized with the nuclear DNA (Fig. 3*A*, *left panel*). The *cox2a* and *cox2b* genes of *Polytomella* sp. hybridized with the major DNA fraction and not with the mtDNA band, confirming their nuclear localization. A similar analysis was carried out with total DNA from *C. reinhardtii* using a *cox1* gene probe that hybridized with the mtDNA (31) and a cytochrome *c* gene probe (32) that hybridized with the nuclear DNA (Fig. 3*A*, *left panel*). Those results confirmed that the *cox2a* and *cox2b* genes were also nuclear-encoded in *C. reinhardtii*.

To determine whether *cox2a* and *cox2b* were present as single-copy genes in the genomes of *Polytomella* sp. and *C. rein-*

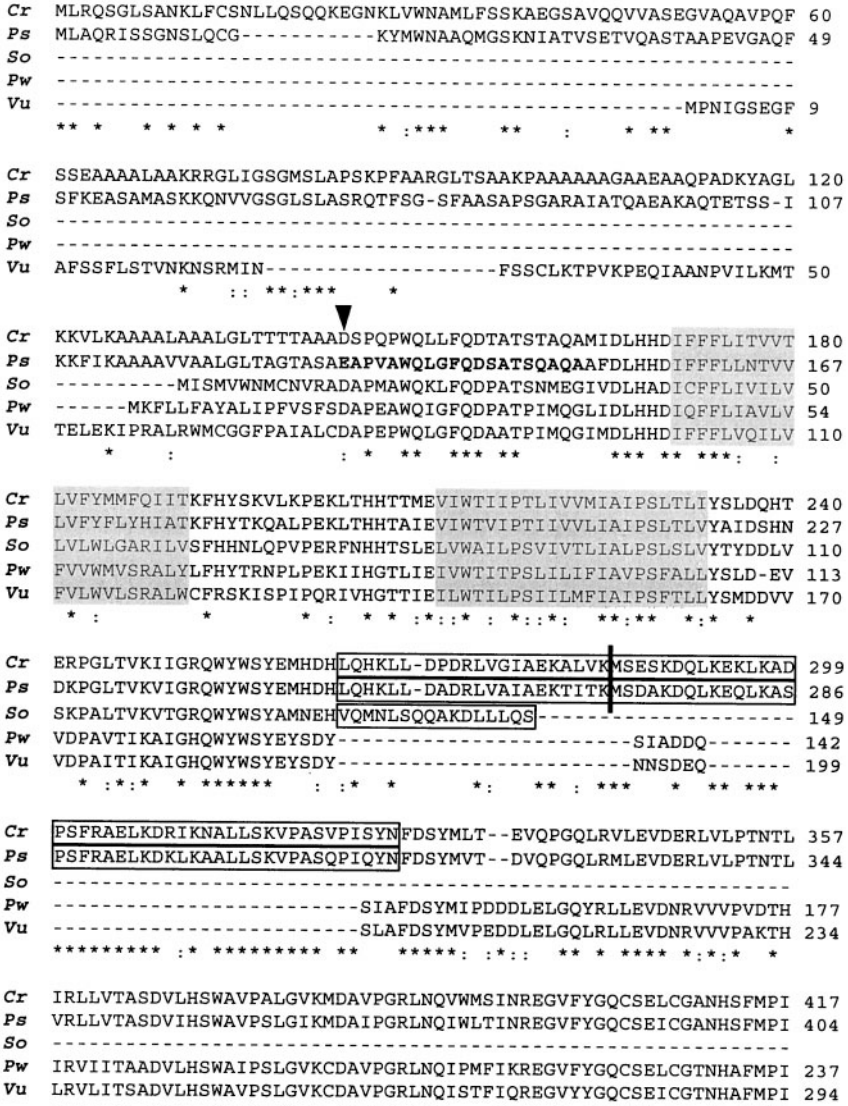

FIG. 2. **Sequence comparison of COX II, COX IIA, and COX IIB polypeptides.** Shown are the sequences of COX II of *P. wickerhamii* (*Pw*), *Vigna unguiculata* (*Vu*), and *S. obliquus* (*So*). Also shown are COX IIA and COX IIB of *C. reinhardtii* (*Cr*), and *Polytomella* sp. (*Ps*). A *black bar* indicates the end of the COX IIA protein of *C. reinhardtii* and *Polytomella* sp. and the N terminus of the COX IIB protein. An *asterisk* denotes an identical amino acid, and a *colon* indicates a similar amino acid. The 20-amino acid extension at their C terminus of COX IIA and the 42-amino acid extension at the N terminus of COX IIB are *boxed*. These unique extensions of COX IIA and COX IIB are hypothesized to interact for the assembly of the two COX II subunits in the cytochrome *c* oxidase complex. The putative transmembrane stretches are indicated within *gray boxes*. The *black triangle* indicates the N terminus of the mature COX IIA that was determined for *Polytomella* sp. The highly conserved COX II regions G(R/H)QWYWSY present within COX IIA and GQCSE(I/L)CG present within COX IIB are *underlined*. *Numbers* indicate the last amino acid for each protein.

hardtii, additional Southern blot analyses were performed. Single hybridization bands were obtained for the *cox2a* and *cox2b* genes of *Polytomella* sp. and *C. reinhardtii* with several restriction enzymes (Fig. 3, *panel B* and data not shown). This suggested that both genes were present in only one copy in their respective genomes.

The expression of *cox2a* and *cox2b* of *Polytomella* sp. and *C. reinhardtii* was examined by Northern analyses. Probes derived from the *cox2a* and *cox2b* genes hybridized to independent transcripts of sizes consistent with the corresponding cDNA sequences for both algae (Fig. 3, *panel C*). The presence of a larger, mature transcript that could suggest a transpliced product was not observed in any case. The *cox2b* gene of *Polytomella* sp. exhibited a double band. Since the genomic sequence of this *cox2b* gene contained three putative polyadenylation sites in the 3′-noncoding region, it is possible that these bands correspond to mRNAs that have different sites of polyadenylation.

There is a significant bias in codon usage in these genera of algae (30), and this bias differs between mitochondrial and nuclear genes (14). Analysis of the codon usage for the *cox2a* and *cox2b* genes of *Polytomella* sp. and *C. reinhardtii* indicated that the codon usage was consistent with their nuclear localization (data not shown). In addition, the conserved poly-

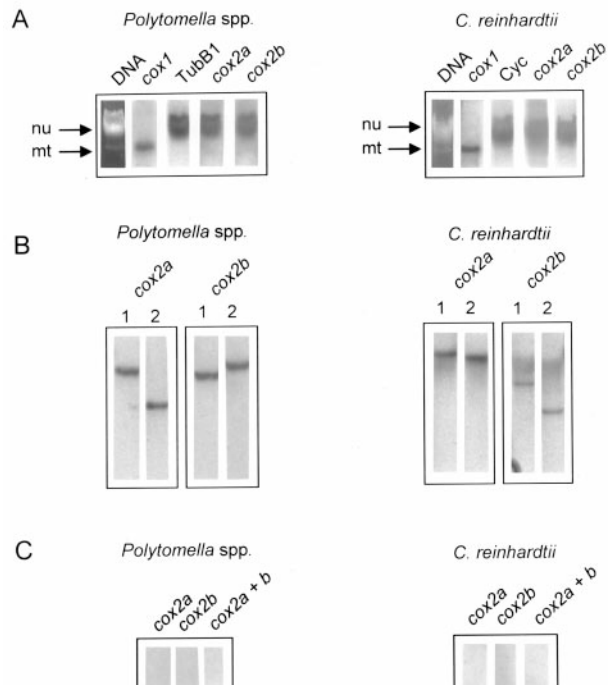

FIG. 3. **The *cox2a* and *cox2b* genes are single-copy, nuclear-localized, and functionally expressed in Chlamydomonad algae.** *A*, nuclear localization of the *cox2a* and *cox2b* genes. Thirty μg of total DNA from *Polytomella* sp. (*left panel*) or *C. reinhardtii* (*right panel*) was electrophoresed through a 0.7% agarose gel. The first *lanes* (DNA) of both *panels* are photographs of the ethidium bromide-stained sample after electrophoresis. The positions of the nuclear DNA (*nu*) and mtDNA (*mt*) are shown by *arrows*. The remaining *lanes* show autoradiograms after Southern analysis. The *left panel* shows *Polytomella* sp. DNA hybridized with probes for mtDNA-encoded *Polytomella* sp. *cox1* (second *lane*), nuclear DNA-encoded *Polytomella agilis* β-tubulin (*TubB1*, third *lane*), and *Polytomella* sp. *cox2a* (fourth *lane*) and *cox2b* (*cox2b*, fifth *lane*) genes. The *right panel* shows Southern analyses of the *C. reinhardtii* DNA hybridized with mtDNA-encoded *C. reinhardtii cox1* (second *lane*), nuclear DNA-encoded cytochrome *c* of *C. reinhardtii* (*Cyc*, third *lane*), and *C. reinhardtii cox2a* (fourth *lane*) and *cox2b* (*cox2b*, fifth *lane*) genes. *B*, *cox2a* and *cox 2b* are single-copy genes. Total DNA from *Polytomella* sp. (*left*) was digested with *Eco*RV (*lane1*) or *Pvu*II (*lane 2*) for *cox2a* and *Eco*RV (*lane1*) or *Eco*RI (*lane 2*) for *cox2b*. Total DNA from *C. reinhardtii* (*right*) was digested with *Kpn*I (*lane1*) or *Xba*I (*lane2*) for *cox2a* and *Kpn*I (*lane 1*) or *Hin*dIII (*lane2*) for *cox2b*. Shown are autoradiograms of the Southern blot analyses of the restricted DNAs hybridized with probes for *cox2a* or *cox2b*. *C*, Northern blot analysis. Shown are autoradiograms of Northern blots of total RNA isolated from *Polytomella* sp. (*left*) or *C. reinhardtii* (*right*) hybridized with probes specific for *cox2a* and *cox2b* genes from *Polytomella* sp. and *C. reinhardtii*. The sizes calculated for the hybridizing RNAs are indicated to the left of the autoradiograms. *bp*, base pairs.

adenylation signals TGTAA (33), present in the vast majority of nuclear genes in the Chlamydomonad family, were present at the 3′ ends of the cDNA sequences of *cox2a* and *cox2b* for both algae.

Two COX II Polypeptides Are Present in Polytomella sp. Cytochrome c Oxidase—Since the primary protein sequences derived from the 18.6-kDa region of a polyacrylamide gel corresponded to the predicted amino acid sequences of both COX IIA and COX IIB from *Polytomella* sp., it is likely that this region of the gel contained both subunits. To confirm that *Polytomella* sp. contained two independent COX II polypeptides, the purified cytochrome *c* oxidase complex of this alga was subjected to matrix-assisted laser desorption-time of flight mass spectrometry analysis. The complex contained two major

TABLE I
Mass spectrometry analysis of the tryptic fragments obtained from the COX IIB subunit of Polytomella sp

The experimental values were obtained by matrix-assisted laser desorption-time of flight mass spectrometry analysis of tryptic digestion fragments of COX IIB purified by reverse transcription-HPLC from isolated cytochrome *c* oxidase. These masses are compared with the predicted molecular masses for the tryptic fragments from the *cox2b* gene product.

Sequence of the tryptic peptide	Theoretical molecular mass	Experimental molecular mass
	Da	*Da*
AFLTEYVK	970.12	970.448
LVLPTNTLVR	1125.37	1125.649
LNQIWLTINR	1270.48	1270.702
LLVTASDV. . .AVPSLGIK	2106.47	2123.172
VPASQPIQ. . .DVQPGQLR	2854.18	2854.383
EGVFYGQC. . .VVEAISPR	3141.56	3140.432

polypeptides in the expected mass range for COX IIA and COX IIB, one of 15,984 Da (theoretical 16,222 Da for mature COX IIA) and one of 17,169 Da (theoretical 17,219 Da for full-length COX IIB). The differences between the predicted and observed masses are likely due to post-translational modifications of the proteins. The 17,169-Da polypeptide was isolated by reverse phase-HPLC, trypsinized, and analyzed by matrix-assisted laser desorption-time of flight mass spectrometry analysis. The trypsin digestion products exhibited molecular masses that were consistent with the theoretical molecular masses expected for COX IIB peptides (Table I).

We wished to determine whether COX IIB contained a cleavable mitochondrial-targeting sequence or if the novel N-terminal charged domain was present in the mature protein. The observed band of 18.6 kDa (containing COX IIA and COX IIB) was purified from polyacrylamide gels and subjected to digestion by trypsin or endolysin followed by HPLC separation and Edman degradation (Fig. 4). Amino-terminal sequences were obtained for all the endolysin C-derived fragments, except for the 16.0-kDa polypeptide, which exhibited a blocked N terminus. The sequence was obtained from COX IIB starting at Asp-5 of the full-length sequence (tryptic fragment DQLK). That the mass spectrometry analysis of the mature protein gave a mass consistent with an intact N terminus suggests mature COX IIB is the full-length protein. Several overlapping sequences were obtained for the tryptic products and for the endolysin C fragments of the protein. A total of 154 amino acids out of the 294 residues predicted by the *cox2a* and *cox2b* genes were identified in the 18.6-kDa region of the gel by direct amino acid sequence analysis. This analysis, using a different cytochrome *c* oxidase preparation from that first used for Edman degradation, confirmed that Glu-131 was the N terminus of the mature COX IIA. No difference was found between the predicted sequences and the sequences obtained by Edman degradation.

DISCUSSION

COX II Is Encoded by Two Distinct Nuclear Genes in the Algae of the Family Chlamydomonadaceae—COX II is conserved in cytochrome *c* oxidases throughout all phyla. It constitutes the main core of the enzyme along with COX I and COX III (34). Although the *cox2* gene is absent from the mtDNA of *Polytomella* sp.,[1] the presence of COX II in cytochrome *c* oxidase from *Polytomella* sp. was confirmed by N-terminal and internal amino acid sequence analysis of polypeptides with an apparent molecular mass of 18.6 kDa from the purified complex. This size is consistent with a 14.5-kDa polypeptide that was identified immunochemically as COX II in cytochrome *c* oxidase isolated from *C. reinhardtii* (35).

Polytomella sp. and *C. reinhardtii* each contain two distinct

N-terminal region of the Ps-COX II subunit (encoded by the *cox2a* gene):

EAPVAWQLGFQDSATSQAQA AFDLHHDIFFFLLNTVVL

VFYFLYHIATK↓***FHYTK***↓***QALPE***K↓LTHHTAIEVIWTVIPTII

VVLIAIPSLTLVYAIDSHNDK↓PGLTVK↓VIGR↓***QWYWSYE***

MHDHLQHK↓LLDADR↓LVAIAEK↓TITK

C-terminal region of the Ps-COX II subunit (encoded by the *cox2b* gene):

MSADK↓***DQLK***↓***EQLK***↓***ASPSFR***↓AELK↓DK↓LK↓***AALLSK***↓***VPA***

SQPIQYNFDSYMVTDVQPGQLR↓***MLEVDER***↓***LVLPTNTL***

VR↓***LLVTASDVIHSW***AVPSLGIK↓***MDAIPGR***↓LNQIWLTIN

R↓***EG******VFYGQ***CSEICGANHSFMPIVVEAISPR↓***AFLTEYV***

K↓***K***↓***WIQ***

FIG. 4. **Protein sequence analysis of the COX IIA and COX IIB polypeptides of *Polytomella* sp.** Shown are the amino acid sequences of COX IIA and COX IIB. The polypeptides present in the region of the polyacrylamide gel (containing COX IIA and COX IIB) were isolated and subjected to digestion by trypsin or endolysin, HPLC separation, and Edman degradation. The N-terminal sequences were obtained for the endolysin C-derived fragments and tryptic digestion products of the proteins. The amino acids directly identified by these amino acid sequence analyses are in *bold italic* characters. The regions identified by N-terminal Edman degradation analysis of the 18.6-kDa region of the polyacrylamide gel are *boxed*. A total of 154 amino acids out of the 294 amino acids were identified in this analysis. No difference was found between the predicted sequences and the sequences obtained by Edman degradation.

nuclear genes that encode two COX II homologs. *cox2a* encodes a protein, COX IIA, corresponding to the N-terminal half of a canonical COX II. *cox2b* encodes a protein, COX IIB, corresponding to the C-terminal half of a canonical COX II. The *cox2a* and *cox2b* genes from *Polytomella* sp. and *C. reinhardtii* were shown to be nuclear-localized by Southern hybridization analyses. These genes also displayed other features typical of nuclear genes in Chlamydomonad algae: a biased codon usage, introns in the genomic sequences, and a TGTAA polyadenylation signal.

COX IIA and IIB were unresolved by SDS-polyacrylamide gel electrophoresis and migrated with an apparent molecular mass of 18.6 kDa. Mass spectrometry and reverse phase HPLC analyses confirmed that distinct COX IIA and IIB proteins existed in the purified cytochrome *c* oxidase of *Polytomella* sp. The COX IIB polypeptide may have a blocked N terminus (MSADKDQL), since only the major N-terminal sequence EAPVAWQLGF of COX IIA was detected by direct N-terminal sequence analysis of the 18.6-kDa polypeptide.

COX IIA and IIB Have Been Adapted to Facilitate Import into Mitochondria—Most nuclear-encoded mitochondrial proteins contain a mitochondrial targeting sequence that directs the import of the protein into mitochondria. Such sequences have been identified in the nuclear-encoded *cox2* genes from cowpea (7) and soybean (8). The predicted full-length COX IIA proteins each has a mitochondrial targeting sequence of 130 amino acids, since the amino acid sequence obtained for the mature *Polytomella* sp. COX IIA commences at Glu-131 of the predicted polypeptide. These sequences are 37% identical (Fig. 2). The COX III proteins from these organisms also contain unusually long (>100 amino acids) mitochondrial-targeting sequences (14). It is possible that during import into mitochondria the targeting sequences may act as chaperones for the import of the COX subunits and the assembly of the enzyme complex. Alternatively, after cleavage they may be maintained as components of the cytochrome *c* oxidase complex, as was observed with the targeting sequence of the Rieske subunit of beef heart mitochondrial cytochrome bc_1 complex (36).

In contrast to COX IIA, the COX IIB proteins lack a canonical, cleavable, mitochondrial-targeting sequence, since the COX IIB sequence was obtained starting at Asp-5 of the full-length sequence (Fig. 4). Additionally, mass spectrometry analysis of the mature protein gave a mass consistent with an intact N terminus. However, the N-terminal 43 amino acids in both predicted COX IIB proteins show no similarity with other known COX II sequences and contains a high density of charged residues. The numerous negatively charged amino acids are not consistent with a mitochondrial-targeting sequence, normally characterized by a paucity of acidic residues and abundance of basic residues (37). It is possible that the 43-amino acid stretch may function as a noncleavable mitochondrial-targeting sequence. The N-terminal region of *Polytomella* sp. COX IIB possesses a positively charged amphiphilic α-helix from position 19 to 44 with a high hydrophobic moment (μH ranging from 5.98 to 7.42) and a very hydrophobic face (maximum hydrophobicity ranging from 3.55 to 6.35). Similarly, the N-terminal region of *C. reinhardtii* COX IIB possesses a putative amphiphilic helix from residues 14 to 32, with μH ranging from 7.93 to 8.92 and a hydrophobic face (maximum hydrophobicity ranging from 4.70 to 6.44). The amphiphilicity is essential for the function of a mitochondrial targeting sequence (38). Alternatively, there may be no conventional N-terminal, cleavable mitochondrial targeting sequence, as in eukaryotic cytochrome *c*. Like COX IIB, cytochrome *c* is imported into the inner membrane space, but a minimum length of N-terminal region rather than a specific sequence is necessary for efficient mitochondrial import (39).

The highest average hydrophobicity over 60–80 amino acids of a polypeptide chain (termed mesohydrophobicity, or *meso*H) along with the maximum local hydrophobicity (<H>) of likely transmembrane segments are useful indicators of the likelihood that a protein could be imported into the mitochondrion (23). Both *Polytomella* sp. and *C. reinhardtii* mature COX IIA proteins, despite their two transmembrane domains, exhibited <H> and *meso*H patterns significantly lower than mtDNA-encoded COX II polypeptides from other organisms but still higher than most proteins imported into mitochondria (data not shown). The nuclear-encoded COX II proteins from legumes were also less hydrophobic than their mtDNA-encoded counterparts and more similar to the COX IIA proteins of the Chlamydomonad algae.

The low hydrophobicity parameters exhibited by the COX IIB proteins may also facilitate their import into mitochondria. These proteins contain no transmembrane stretches, exhibit very low <H> and *meso*H values, and are predicted to be imported readily into mitochondria (data not shown).

Assembly of COX IIA and COX IIB in Cytochrome c Oxidase—The normal COX II subunit exhibits two transmembrane stretches and a hydrophilic region containing the consensus copper binding site. We predict that the two mature Chlamydomonad COX II polypeptides would assemble to give a heterodimeric COX II with an overall structure similar to that determined by x-ray crystallography of the bacterial and mammalian monomeric COX II subunits (28, 40).

Hydropathy profile analysis predicted two transmembrane stretches for the COX IIA polypeptides from both algae (data not shown), similar to conventional COX II proteins. In addition, both contain a C-terminal 20-amino acid region that has a high density of charged amino acids and is not homologous to known COX II proteins. The two COX IIB polypeptides contain a 42-amino acid extension at their N terminus that also exhib-

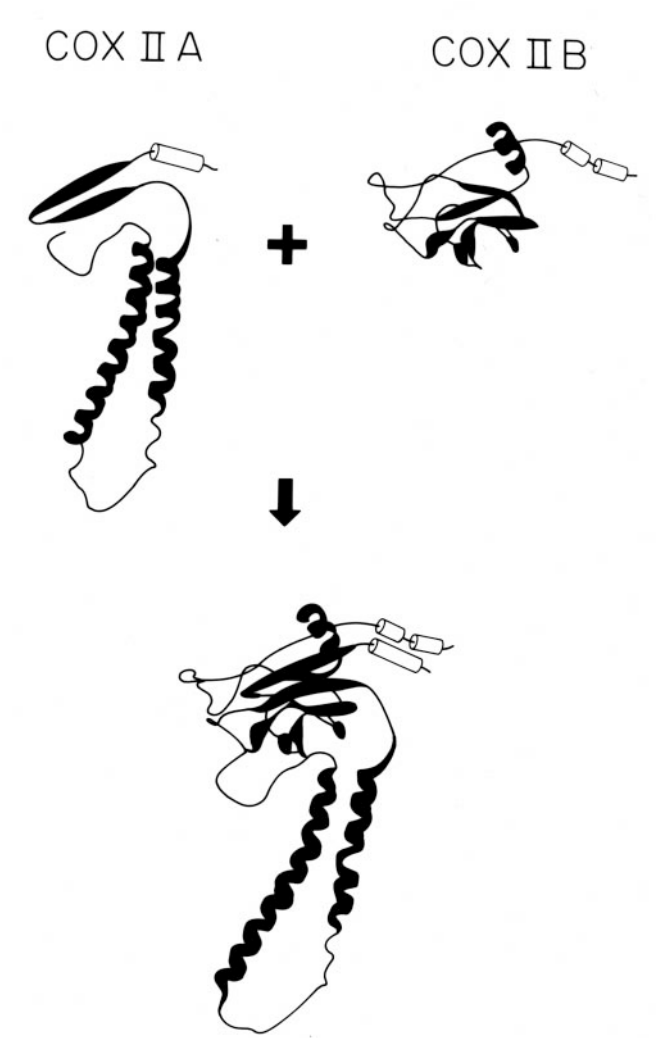

FIG. 5. **Model of the interaction of COX IIA with COX IIB to form the heterodimeric COX II of Chlamydomonad algae.** A three-dimensional structure of *Polytomella* sp. COX IIA and COX IIB was modeled with SWISS-MODEL based upon the three-dimensional structure of COX II from the bovine enzyme (28). This model proposes an interaction between the unique COX IIA C-terminal domain and the highly charged COX IIB N-terminal domain. We hypothesize that the loop formed by the interaction of the N-terminal and C-terminal extensions is important for the proper assembly of the two COX II subunits in cytochrome *c* oxidase in the mitochondrial inner membrane.

its a high density of charged amino acids and that is also not homologous to known COX II proteins. We propose that the C-terminal extension of COX IIA interacts with the N-terminal extension of the COX IIB protein. This interaction may stabilize the two COX II subunits in the cytochrome *c* oxidase complex (Fig. 5). According to this model, the extra loop formed by the interaction of the N-terminal and C-terminal extensions is topologically distant from the site of interaction of soluble cytochrome *c* with the COX IIB subunit.

Evolutionary Considerations—The transfer of *cox2* genes from the mitochondria to the nucleus in the Chlamydomonad algae satisfies many of the criteria proposed by Brennicke *et al*. (41) and Claros *et al*. (42) for this event: acquisition of a region encoding a mitochondrial-targeting sequence (at least for *cox2a*), altered codon usage, acquisition of a polyadenylation signal, and diminished <H> and *meso*H of the protein products. In addition, any mitochondrial *cox2* genes that presumably existed have been eliminated, suggesting that this transfer occurred early in evolution.

There are several precedents for mitochondrial genes being split into two. The mitochondrial *nad1* genes of *Tetrahymena pyriformis* and *Paramecium aurelia* (43) and rapeseed mitochondrial gene encoding a homologue of the bacterial protein Ccl1 (44) have also been divided into two independently transcribed reading frames, although retained on the mtDNA.

It is possible that in some organisms the transfer of *cox2a* and *cox2b* to the nucleus is ongoing, as in some legumes, or has been arrested at an intermediate stage. *cox1*, *cox2*, and *cox3* were present in the mtDNA of the Chlorophycean alga *Scenedesmus obliquus* (45, 46). However, the *cox2* gene was truncated and was predicted to be a pseudogene. The putative *Scenedesmus* COX II protein exhibits high similarity with COX IIA from *Polytomella* sp. and *C. reinhardtii* (Fig. 2), although it lacks a mitochondrial-targeting sequence, as would be expected for a mtDNA-encoded protein. The C-terminal region of the *Scenedesmus* COX II is also similar to the unique C-terminal regions of COX IIA of *Polytomella* sp. and *C. reinhardtii*. We suggest that the mitochondrial *cox2* gene from *S. obliquus* is likely to be an active *cox2a* gene and that the missing fragment of the gene (*i.e. cox2b*) has been transferred to the nucleus.

The gene content and fragmentation pattern of the ribosomal RNA genes on the mtDNA of *S. obliquus* suggest that this mitochondrial genome represents an intermediate stage between the *Prototheca*-like green algae and the *Chlamydomonas*-like green algae (45). *S. obliquus* may represent a stage of green algal evolution in which the *cox2b* gene (encoding for a highly hydrophilic polypeptide) has been transferred to the nucleus, whereas the *cox2a* gene (encoding a more hydrophobic polypeptide) is retained in the mitochondrial genome. Since the mitochondrial genetic code of *S. obliquus* has diverged from the standard genetic code (45, 46), it is likely that the *cox2a* gene has been retained on the mtDNA since it could not be functionally transferred to the nucleus in its current form. We hypothesize that two independent *cox2a* and *cox2b* genes existed in the mitochondrial genome of the common ancestor of Chlamydomonad algae and that these genes were transferred from the mitochondria to the nucleus before the separation of the colorless genera of *Polytomella* from the main photosynthetic lineage of *Chlamydomonas*.

The functional transfer of intact *cox2* genes to the nucleus has been reported in detail for members of the legume subfamily Papilionoideae (7, 8, 9). Our analysis of the Chlamydomonad *cox2a* and *cox2b* genes suggests that the division of the mitochondrial *cox2* gene into two nuclear genes that encode two independent proteins is an evolutionary event unrelated to the functional gene transfer of intact mitochondrial *cox2* genes to the nucleus in some legumes.

A number of organisms, particularly protists, have mitochondrial genomes that lack the *cox2* gene, including the apicomplexan protozoan *Plasmodium falciparum* (47) and the Chlorophyte alga *Pedinomonas minor* (1), whose mtDNA has been sequenced. It is likely that these organisms have transferred their *cox2* genes to the nucleus and that, in the green alga *P. minor*, this has also involved the transfer of *cox2a* and *cox2b* to the nucleus. Further analyses of such organisms will determine whether the process outlined here has an evolutionary lineage outside of the Chlorophyta.

The present work highlights a novel strategy utilized by some Chlorophyte algae that allowed functional transfer of a mitochondrial gene to the nucleus. This strategy has allowed the functional transfer to the nucleus of a mitochondrial gene encoding an essential, hydrophobic subunit of a crucial respiratory complex. It is remarkable that a protein with such specific functional requirements (*e.g.* a cytochrome *c* docking site, binding site for metals involved in catalysis) should have

evolved into two components containing added domains that maintain the ability of the protein to perform its roles.

Acknowledgments—We thank Yun-Lu and T. A. Neubert (New York University) for mass spectroscopy analysis, M. W. Gray (Dalhousie University) and J. D. Palmer (Indiana University) for encouragement at the starting point of this project, L.-G. Franzén (Goteborg University) for the kind gift of the cDNA library of *C. reinhardtii*, L. Ongay, G. Codiz, and M. Sosa (Unidad de Biología Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México) for the synthesis of various oligonucleotides, and several colleagues and researchers who kindly provided DNA probes: T. D. Fox (Cornell University), M. Goldschmidt-Clermont (University of Geneva), E. H. Harris (Duke University), J. D. Palmer (Indiana University), and C. D. Silflow (University of Minnesota). We are also indebted to A. Atteia and A. Gómez-Puyou (Instituto de Fisiología Celular, Universidad Nacional Autónoma de México) and D. W. Krogmann (Purdue University) for critical review of the manuscript.

REFERENCES

1. Gray, M. W., Lang, B. F., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T. G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, Y., and Burger, G. (1998) *Nucleic Acids Res.* **26,** 865–878
2. Andersson, S. G. E., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C. M., Podowski, R. M., Nslund, A. K., Eriksson, A.-S., Winkler, H. H., and Kurland, C. G. (1998) *Nature* **396,** 133–140
3. Gray, M. W. (1992) *Int. Rev. Cytol.* **14,** 233–357
4. Attardi, G., and Schatz, G. (1988) *Annu. Rev. Cell Biol.* **4,** 289–333
5. von Heijne, G. (1986) *FEBS Lett.* **198,** 1–4
6. Popot, J. L., and de Vitry, C. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 369–403
7. Nugent, J. M., and Palmer, J. D. (1991) *Cell* **66,** 473–481
8. Covello, P. S., and Gray, M. W. (1992) *EMBO J.* **11,** 3815–3820
9. Adams, K. L., Song, K., Roessler, P. G., Nugent, J. M., Doyle, J. L., and Palmer, J. D. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13863–13868
10. van den Boogaart, P., Samallo, J., and Agsteribbe, E. (1982) *Nature* **298,** 187–189
11. Michaelis, G., Vahrenholz, C., and Pratje, E. (1990) *Mol. Gen. Genet.* **223,** 211–216
12. Denovan-Wright, E. M., Nedelcu, A. M., and Lee, R. W. (1998) *Plant Mol. Biol.* **36,** 285–295
13. Kroymann, J., and Zetsche, K. (1998) *J. Mol. Biol.* **47,** 431–440
14. Pérez-Martínez, X., Vázquez-Acevedo, M., Tolkunova, E., Funes, S., Claros, M. G., Davidson, E., King, M. P., and González-Halphen, D. (2000) *J. Biol. Chem.* **275,** 30144–30152
15. Gutiérrez-Cirlos, E. B., Antaramian, A., Vázquez-Acevedo, M., Coria, R., and González-Halphen, D. (1994) *J. Biol. Chem.* **269,** 9147–9154
16. Atteia, A., Dreyfus, G., and González-Halphen, D. (1997) *Biochim. Biophys. Acta* **1320,** 275–284
17. Brumme, S., Kruft, V., Schmitz, U.-K., and Braun, H.-P. (1998) *J. Biol. Chem.* **273,** 13143–13149
18. Schägger, H., Link, T. A., Engel, W. D., and von Jagow, G. (1986) *Methods Enzymol.* **126,** 224–237
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
20. Frohman, M. A. (1993) *Methods Enzymol.* **218,** 340–356
21. Franzén, L.-G., and Falk, G. (1992) *Plant Mol. Biol.* **19,** 771–780
22. Lefebvre, P. A., and Silflow, C. D. (1999) *Genetics* **151,** 9–14
23. Claros, M. G. (1995) *Comput. Appl. Biosci.* **11,** 441–447
24. Claros, M. G., and Vincens, P. (1996) *Eur. J. Biochem.* **241,** 779–786
25. Claros, M. G., and von Heijne, G. (1994) *Comput. Appl. Biosci.* **10,** 685–686
26. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723
27. Antaramian, A., Coria, R., Ramírez, J., and González-Halphen, D. (1996) *Biochim. Biophys. Acta* **1273,** 198–202
28. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., and Yoshikawa, S. (1996) *Science* **272,** 1136–1144
29. Wolff, G., Plante, I., Lang, B. F., Kück, U., and Burger, G. (1994) *J. Mol. Biol.* **237,** 75–86
30. Conner, T. W., Thompson, M. D., and Silflow, C. (1989) *Gene* **84,** 345–358
31. Gray, M. W., and Boer, P. H. (1988) *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **319,** 135–147
32. Amati, B. B., Goldschmidt-Clermont, M., Wallace, C. J., and Rochaix, J. D. (1988) *J. Mol. Evol.* **28,** 151–160
33. Silflow, C. (1998) *The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas* (Rochaix, J.-D., Goldschmidt-Clermont, M., and Merchant, S., eds) pp. 25–40, Kluwer Academic Publishers Group, Dordrecht, Netherlands
34. Saraste, M. (1990) *Q. Rev. Biophys.* **23,** 331–366
35. Bennoun, P., Atteia, A., Pierre, Y., and Delosme, M. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 10202–10206
36. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S., and Jap, B. K. (1998) *Science* **281,** 64–71
37. Claros, M. G., Brunak, S., and von Heijne, G. (1997) *Curr. Opin. Struct. Biol.* **7,** 394–398
38. Roise, D., Theiler, F., Horvath, S. J., Tomich, J. M., Richards, J. H., Allison, D. S., and Schatz, G. (1988) *EMBO J.* **7,** 649–653
39. Wang, X., Dumont, M. E., and Sherman, F. (1996) *J. Biol. Chem.* **271,** 6594–6604
40. Iwata, S., Ostermeier, C., Ludwig, B., and Michel, H. (1995) *Nature* **376,** 660–669
41. Brennicke, A., Grohmann, L., Hiesel, R., Knoop, V., and Schuster, W. (1993) *FEBS Lett.* **325,** 140–145
42. Claros, M. G., Perea, J., Shu, Y., Samatey, F. A., Popot, J. L., and Jacq, C. (1995) *Eur. J. Biochem.* **228,** 762–771
43. Edqvist, J., Burger, G., and Gray, M. W. (2000) *J. Mol. Biol.* **297,** 381–393
44. Handa, H., Bonnard, G., and Grienenberger, J.-M. (1996) *Mol. Gen. Genet.* **252,** 292–302
45. Nedelcu, A. M., Lee, R. W., Lemieux, C., Gray, M. W., and Burger, G. (2000) *Genome Res.* **10,** 819–831
46. Kück, U., Jekosch, K., and Holzamer, P. (2000) *Gene* **253,** 13–18
47. Feagin, J. E. (1992) *Mol Biochem. Parasitol.* **52,** 145–148