Mol Gen Genet (1990) 223:211–216

MGG
© Springer-Verlag 1990

Mitochondrial DNA of *Chlamydomonas reinhardtii*: The gene for apocytochrome *b* and the complete functional map of the 15.8 kb DNA

Georg Michaelis, Carola Vahrenholz, and Elke Pratje

Botanisches Institut der Universität Düsseldorf, Universitätsstrasse 1, D-4000 Düsseldorf, Federal Republic of Germany

Received March 1, 1990

Summary. We have sequenced the termini of the mitochondrial genome of *Chlamydomonas reinhardtii* and now present the DNA sequence of the gene for apocytochrome *b*. This gene is the thirteenth gene of the linear 15.8 kb DNA and appears to be the last one of the mt genome. The deduced protein sequence of 381 amino acid residues shows 56%, 48.6% and 48% identity with the apocytochrome *b* proteins of maize, *Drosophila yakuba* and mouse, respectively. RNA analysis reveals a transcript of about 1250 nucleotides. It is now possible to present the complete protein-coding capacity, the pattern of codon utilization for all eight protein genes, and the complete functional map of the mitochondrial 15.8 kb DNA of *C. reinhardtii*. One surprising feature is the absence of mitochondrial genes for ATPase and subunits II and III of cytochrome oxidase. No more than three tRNA genes appear to be present on the 15.8 kb mitochondrial DNA.

Key words: *Chlamydomonas reinhardtii* – Functional map of mtDNA – Apocytochrome *b* gene

Introduction

The unicellular green alga *Chlamydomonas reinhardtii* is a very suitable system for genetic and molecular studies (Harries 1989). In contrast to the very large circular mitochondrial genomes of higher plants (200–2400 kb), a linear DNA molecule of only 15.8 kb can be extracted from mitochondria of this green alga (Ryan et al. 1978). Extensive DNA sequence data for the mitochondrial genome of *C. reinhardtii* have been published (Pratje et al. 1984, 1989; Vahrenholz et al. 1985; Kück and Neuhaus 1986; Boer and Gray 1986, 1988a, b, c; Ma et al. 1988, 1989). Due to the small size of the mitochondrial genome, *C. reinhardtii* will provide the first complete mitochondrial DNA sequence from a photoautotrophic organism.

We have completed the DNA sequence of the two termini of the linear molecule and describe the sequence of the last protein gene here. This information allows us to summarize the protein coding capacity, the pattern of codon utilization in all eight protein genes, and the complete functional map of the 15.8 kb mitochondrial DNA of *C. reinhardtii*.

Materials and methods

Isolation of mitochondrial DNA. Mitochondria were isolated from the cell wall-less mutant CW15 (strain CCAP 11/32 CW15$^+$ from the Culture Centre of Algae and Protozoa, Cambridge, UK). The DNA was extracted as described by Vahrenholz et al. (1985).

*DNA cloning and sequencing. Sau*3AI and *Alu*I fragments cloned into pUC19 vector DNA were used for sequencing the apocytochrome *b* gene by the dideoxynucleotide chain termination procedure (Sanger et al. 1977). In addition *Hpa*I and *Dde*I subclones were constructed to sequence both DNA strands completely across all restriction sites.

Transcript analysis. Isolation of mitochondrial RNA and Northern hybridization have been described (Vahrenholz et al. 1985).

Results and discussion

Identification and characterization of the gene for apocytochrome b *on the 15.8 kb mitochondrial DNA of* C. reinhardtii

The DNA sequence determined is illustrated in Fig. 1. Translation of the DNA sequence according to the universal genetic code reveals a continuous open reading frame (ORF) of 1146 base pairs (bp) that starts at position 26 with an ATG initiation codon. The derived protein sequence of 381 amino acid residues has a molecular

---

Offprint requests to: G. Michaelis

```
   1   TAAACTTTGTTATTTATTATTTACAATGCGTATGCATAACAAAATTCAATTGTTGAGTGTACTAAACACTCATTTGGTAGCCTAC
   1                            M  R  M  H  N  K  I  Q  L  L  S  V  L  N  T  H  L  V  A  Y

  86 CCAACTCCAATGAACCTAAACTATTCTTGGAACGGTGGTTCTCTAGCTGGTATGATGCTAGCTAGTCAAATGCTTACTGGTATTCTA
  21  P  T  P  M  N  L  N  Y  S  W  N  G  G  S  L  A  G  M  M  L  A  S  Q  M  L  T  G  I  L

 173 CTAGCCATGCACTATGTAGGTCACGTAGACTACGCTTTTGCTAGCGTACAACACCTAATGACTGATGTACCTTCTGGTATGATCTTG
  50  L  A  M  H  Y  V  G  H  V  D  Y  A  F  A  S  V  Q  H  L  M  T  D  V  P  S  G  M  I  L

 260 CGTTACGCTCACGCTAACGGCGCCAGCTTGTTCTTTATTGTAGTCTATTTGCACGTATTGCGTGGTATGTACTACGGTAGCGGCGCT
  79  R  Y  A  H  A  N  G  A  S  L  F  F  I  V  V  Y  L  H  V  L  R  G  M  Y  Y  G  S  G  A

 347 CAGCCACGTGAGATCGTCTGGATCAGTGGTGTCGTTATCTTGTTGGTAATGATTATCACCGCCTTCATTGGTTATGTACTACCATGG
 108  Q  P  R  E  I  V  W  I  S  G  V  V  I  L  L  V  M  I  I  T  A  F  I  G  Y  V  L  P  W

 434 GGCCAAATGTCTTTCTGGGGTGCTACCGTAATTACTAGTTTGGCTACTGCCATTCCAGTAGTAGGTAAACACATCATGTACTGGTTG
 137  G  Q  M  S  F  W  G  A  T  V  I  T  S  L  A  T  A  I  P  V  V  G  K  H  I  M  Y  W  L

 521 TGGGGTGGTTTCAGTGTTGATAACCCAACCTTGAACCGCTTCTACAGCTTCCACTACACTCTACCATTCATCTTGGCTGGTTTGAGC
 166  W  G  G  F  S  V  D  N  P  T  L  N  R  F  Y  S  F  H  Y  T  L  P  F  I  L  A  G  L  S

 608 GTATTCCACATTGCCGCCTTGCACCAATACGGTAGTACTAACCCACTAGGTGTTAACAGCCAAAGCAGCCTAATTTCTTTCGGTTCT
 195  V  F  H  I  A  A  L  H  Q  Y  G  S  T  N  P  L  G  V  N  S  Q  S  S  L  I  S  F  G  S

 695 TACTTTGGTGCTAAAGACTTGGTCGGTGCTTTGTTCTTGGCTCTTGTGTTCAGCATTCTAGTCTTCTTCTACCCAGACTTGTTGGGT
 224  Y  F  G  A  K  D  L  V  G  A  L  F  L  A  L  V  F  S  I  L  V  F  F  Y  P  D  L  L  G

 782 CACCCAGACAACCTAATCCCAGCTAACCCATATAGCACCCCACAACACATTGTACCAGAGTGGTACTTCTTGTGGGTATACGCTATT
 253  H  P  D  N  L  I  P  A  N  P  Y  S  T  P  Q  H  I  V  P  E  W  Y  F  L  W  V  Y  A  I

 869 CTACGTTCCATTCCAAACAAAGCTATGGGCGTATTGGCTATTGGTCTAGTCTTCGCTAGTTTGTTTGCTATGCCATTCATCGGTTTG
 282  L  R  S  I  P  N  K  A  M  G  V  L  A  I  G  L  V  F  A  S  L  F  A  M  P  F  I  G  L

 956 GGCGGTGGTAAATTCCGCATCATCACTGAGTGGCTATACTGGACTTTCCTTGCTGATGTATTGCTATTGACCTGGTTGGGTGGTAAC
 311  G  G  G  K  F  R  I  I  T  E  W  L  Y  W  T  F  L  A  D  V  L  L  L  T  W  L  G  G  N

1043 GAGATTACTCCAATTACCTCTTTCGTCGGACAGTGCTGCACTGCTTACCTATTCTTCTACCTACTTGTTTGTCAACCACTTGTAGGT
 340  E  I  T  P  I  T  S  F  V  G  Q  C  C  T  A  Y  L  F  F  Y  L  L  V  C  Q  P  L  V  G

1130 TACTTGGAGACTCAGTTTGCCCACGGTACTCAAACCAACTAATATTTTGCATTG
 369  Y  L  E  T  Q  F  A  H  G  T  Q  T  N
```

Fig. 1. Nucleotide sequence of the apocytochrome *b* gene and its amino acid sequence translated according to the universal genetic code. The numbering of the DNA starts at the TAA termination codon of the preceding *nad4* gene. Initiation and termination codons of the *cytb* gene are *underlined*

weight of 42303 daltons. Its identity with the apocytochrome *b* gene is evident by its strong similarity with the apocytochrome *b* genes from animals, fungi, trypanosomes, and higher plants. The alignment of the amino acid sequences (Fig. 2) indicates a sequence similarity of 56.2% with maize, 48.6% with *Drosophila yakuba* and 48% with mouse apocytochrome *b* gene products. The homology between the various apocytochrome *b* polypeptides extends over the complete sequence indicating that translation starts at the first ATG codon of the *C. reinhardtii* ORF. The gene is not interrupted by intervening sequences in contrast to the homologous genes from fungi, such as *Saccharomyces cerevisiae* (Nobrega and Tzagoloff 1980), *Aspergillus nidulans* (Waring et al. 1981), and *Neurospora crassa* (Helmer-Citterich et al. 1983). The apocytochrome *b* gene of the closely related green alga *Chlamydomonas smithii* is almost identical to that of *C. reinhardtii* except for a 1 kb insertion (see accompanying paper Colleaux et al. 1990). Sequence comparison of the two apocytochrome *b* genes reveals 7 nucleotide differences within the ORF, one of which alters the derived amino acid sequence. The apocytochrome *b* protein of *C. reinhardtii* contains the four invariant histidine residues (His-83, 97, 183, 197) that form the putative heme attachment sites. The protein shows a hydropathy profile similar to those of apocytochrome *b* proteins from other organisms (data not shown) and fits the eight transmembrane helices model (Rao and Argos 1986; di Rago et al. 1989).

The apocytochrome *b* gene of *C. reinhardtii* is transcribed from the same DNA strand as the *nad4* and *nad5* genes. An internal probe for the *cytb* gene hybridizes to a single major transcript (Fig. 3) of about 1250 nucleotides, as estimated with reference to the two mitochondrial ribosomal RNA species of *Saccharomyces cerevisiae* (3.27 and 1.66 kb). The calculated size is probably overestimated due to the high AT content of the yeast rRNA. The size of the apocytochrome *b* transcript corresponds well to the length of the gene and the mRNA seems to have few 5′ and 3′non-coding sequences.

Stable mutants with a deleted apocytochrome *b* gene have been described recently (Matagne et al. 1989). These mutant cells cannot grow on acetate medium in the dark and need light for survival, which demonstrates that cytochrome *b* is dispensable in the light.

Complete protein coding capacity of the 15.8 kb mitochondrial DNA of C. reinhardtii

The availability of the DNA sequence of the apocytochrome *b* gene allows us to summarize the complete protein coding information of the 15.8 kb DNA of *C. reinhardtii* (Table 1). The DNA sequence of the remaining 532 bp region present at both termini in inverted orientation does not contain protein coding information (Vahrenholz et al. in preparation). Eight protein encoding genes have been identified: the genes for subunit I of cytochrome oxidase, apocytochrome *b*, 5 subunits of NADH dehydrogenase (*nad1, nad2, nad4, nad5, nad6*), and a reverse transcriptase-like protein (for references see Table 1).

```
Zm   MTIRNQRFSLLKQPIYSTLNQHLIDYPTPSNLSYWWGFGCLAGICLVIQIVTGVFLAMHYTPHVD
                :    : :: ::  :::: :: : : :  :::  :  :  ::  :::::  :::
Cr         MRMHNKIQLLSVLNTHLVAYPTPMNLNYSWNGGSLAGMMLASQMLTGILLAMHYVGHVD   59
               :   :    :  ::  : : :    :: ::: :  :  : :::  :::::   :
Dy       MHKPLRNSHPLFKIANNALVDLPAPINISSWWNFGSLLGLCLIIQILTGLFLAMHYTADVN
              :  :::::: :    ::::: ::::::::::::: ::  :: :::::::::: :
Mm        MTNMRKTHPLFKIINHSFIDLPAPSNISSWWNFGSLLGVCLMVQIITGLFLAMHYTSDTM

                           ↓             ↓
Zm   LAFNSVEHIMRDVEGGWLLRYMHANGASMFLIVVHLHIFRGLYHASYSSPREFVWCLGVVIFLLM
      :: :: : : ::  :  ::: :::::: : ::: ::  :: :  :   ::: ::  :::: : :
Cr   YAFASVQHLMTDVPSGMILRYAHANGASLFFIVVYLHVLRGMYYGSGAQPREIVWISGVVILLVM   124
      :: :: :   ::  :  ::  :::::: :::  :::  :: ::::        :  :: ::
Dy   LAFYSVNHICRDVNYGWLLRTLHANGASFFFICIYLHIGRGIYYGSYLFTP--TWLVGVIILFLV
      :: :: ::::::::::: :  :::::: ::::  :: ::: ::::: :    ::  ::  :: :
Mm   TAFSSVTHICRDVNYGWLIRYMHANGASMFFICLFLHVGRGLYYGSYTFME--TWNIGVLLLFAV

                                                               ↓
Zm   IVTAFIGYVPPWGQMSFWGATVITSLASAIPVVGDTIVTWLWGGFSVDNATLNRFFSLHHLLPLI
     : ::::::: ::::::::::::::::: ::::::  :  :::::::::: ::::: : :  :: :
Cr   IITAFIGYVLPWGQMSFWGATVITSLATAIPVVGKHIMYWLWGGFSVDNPTLNRFYSFHYTLPFI   189
       ::: :::::::::::::::::: :  :::  :     :::::: ::: :: ::  ::  ::::
Dy   MGTAFMGYVLPWGQMSFWGATVITNLLSAIPYLGMDLVQWLWGGFAVDNATLTRFFTFHFILPFI
     : :::::::::::::::::::::::::::::: : :: : :::: :: ::::::: ::::::::
Mm   MATAFMGYVLPWGQMSFWGATVITNLLSAIPYIGTTLVEWIWGGFSVDKATLTRFFAFHFILPFI

            ↓
Zm   LAGASLLHLAALHQYGSNNPLGVHSEMDKIASYPYFYVKDLVGRVASAIFFSIWIFFAPNVLGHP
     ::: :  : :::::::: ::::: :    :    ::  :::::    :  :::  :: :  ::::
Cr   LAGLSVFHIAALHQYGSTNPLGVNSQSSLISFGSYFGAKDLVGALFLALVFSILVFFYPDLLGHP   254
            :   ::: :: :: : ::    : :  ::  :: ::          ::   : ::: :
Dy   VLAMTMIHLLFLHQTGSNNPIGLNSNIDKIPFHPYFTFKDIVGFIVMIFILISLVLISPNLLGDP
       :    :::::: :::::: ::::  :::::::: : ::: :   :  ::  :::  :  ::::
Mm   IAALAIVHLLFLHETGSNNPTGLNSDADKIPFHPYYTIKDILGILIMFLILMTLVLFFPDMLGDP

Zm   DNYIPANPMPTPPHIVPEWYFLPIHAILRSIPDKAGGVAAIAPVFISLLALPFFKEMYVRSSSFR
     :: :::::  :: :::::::::   ::::::: :: :: ::  :: :: : ::          ::
Cr   DNLIPANPYSTPQHIVPEWYFLWVYAILRSIPNKAMGVLAIGLVFASLFAMPFIGL---GGGKFR   316
     :: :::::  :: :: ::::::  ::::::::::  :: :  :  : :   ::  :    :  :
Dy   DNFIPANPLVTPAHIQPEWYFLFAYAILRSIPNKLGGVIALVLSIAILMILPFYNLSKFRGIQFY
     ::  ::::: :: :: :::::::::::::::::::::: :: ::: ::   ::   :: :   :
Mm   DNYMPANPLNTPPHIKPEWYFLFAYAILRSIPNKLGGVLALILSILILALMPFLHTSKQRSLMFR

Zm   PIHQGIFWLLLADCLLLGWIGCQPVEAPFVTIGQISSFFFFLFFAITPIPGRVGRGIPKYYTE
      :     :  ::: ::: : :           ::      :                    :
Cr   IITEWLYWTFLADVLLLTWLGGNEITPITSFVGQCCTAYLFFYLLVCQPLVGYLETQFAHGTQTN   381
      :   : :  :  : :::: :           ::  :   : : :   :::
Dy   PINQILFWSMLVTVILLTWIGARPVEEPYVLIGQILTIIYFLYYLI-NPLVTKWWDNLLN
     :: ::: :        :::::  ::: :   :::   : ::   ::  :      :  :
Mm   PITQILYWILVANLLILTWIGGQPVEHPFIIIGQLASISYFSIILILMPISGIIEDKMLKLYP
```

Fig. 2. Alignment of the predicted apocytochrome *b* protein from *Chlamydomonas reinhardtii* (Cr) with that of *Zea mays* (Zm; Dawson et al. 1984), *Drosophila yakuba* (Dy; Clary et al. 1984), and mouse (Mm; Bibb et al. 1981). Numbering refers to the *Chlamydomonas* sequence. Amino acid identities are indicated by *colons*(:). The four invariant histidine residues of the presumed heme attachment sites are indicated by *arrows*

The mitochondrial genomes of all organisms examined so far encode the genes for apocytochrome *b* and cytochrome oxidase subunits I, II, and III. On the 15.8 kb mtDNA of *C. reinhardtii*, only the genes for apocytochrome *b* and subunit I of cytochrome oxidase are present, whilst the genes for cytochrome oxidase subunits II and III are lacking. Another unique feature of this DNA is that mitochondrial genes for ATPase are also missing. The subunits of this enzyme complex thus have to be imported into *Chlamydomonas* mitochondria. The question whether the NAD3 protein is encoded in the nucleus or is absent from *C. reinhardtii* remains to be investigated. The function of the reverse transcriptase-like gene (*rtl*; Boer and Gray 1988a) in *Chlamydomonas* is unknown. A homologous gene does not exist in mammalian mitochondrial DNA.

Codon utilization in all eight genes that encode proteins

The pattern of codon utilization for all protein coding genes of the 15.8 kb DNA of *C. reinhardtii* is summarized in Table 2. The seven genes for respiratory chain components are very similar in their codon usage, but the *rtl* gene differs. Overall, a significant bias in the codon frequencies is found for each codon family. Nine codons are not used at all and four codons occur exclusively in the *rtl* gene. As discussed earlier (Pratje et al. 1984; Vahrenholz et al. 1985) a deviation from the universal genetic code seems most unlikely. TGA cannot specify tryptophan in *Chlamydomonas* mitochondria because the codon is entirely absent. The CGG codon, which is changed to a UGG codon by RNA editing in higher plants (Covello and Gray 1989; Gualberto et al. 1989), is also absent from *Chlamydomonas* mitochondria.

Complete genetic map of the 15.8 kb mitochondrial DNA from C. reinhardtii

The functional map of the 15.8 kb mitochondrial genome of *C. reinhardtii* is presented in Fig. 4. The genes are located on both strands. *Cytb, nad4* and *nad5* are clustered in the left part of the map and are transcribed from one strand. All the other genes for cytochrome

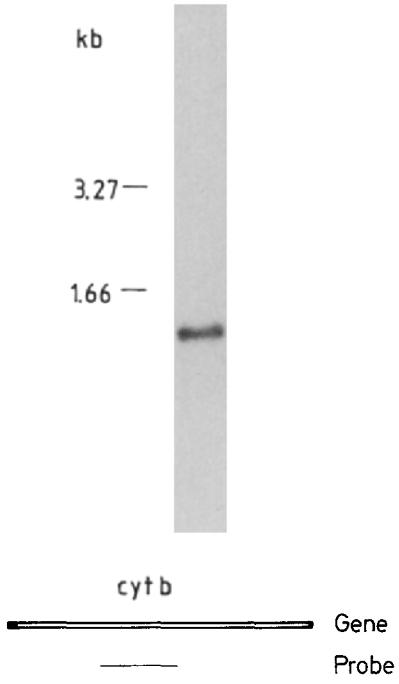

Fig. 3. Transcript of the apocytochrome *b* gene. Approximately 5 µg mitochondrial RNA were fractionated on a 1.4% agarose gel in the presence of formaldehyde, transferred to nitrocellulose and hybridized to an [α-^{32}P]-labelled *Sau*3AI restriction fragment containing internal sequences of the apocytochrome *b* gene. The map position of the hybridization probe is shown. The mitochondrial 21S (3.27 kb) and 15S (1.66 kb) ribosomal RNAs of *Saccharomyces cerevisiae* were used as size markers

oxidase subunit I, three subunits of the NADH dehydrogenase (*nad1, nad2* and *nad6*), the reverse transcriptase-like protein, 3 tRNAs and the two scrambled ribosomal RNAs are expressed from the opposite strand.

Table 1. Protein genes encoded by mitochondrial DNA from different organisms

Genes and functions	Human	Fungi	Higher plants	*Chlamydomonas reinhardtii*
Cytochrome bc_1 complex				
cytb	+	+	+	+
Cytochrome oxidase complex				
coxI	+	+	+	+
coxII	+	+	+	−
coxIII	+	+	+	−
ATPase complex				
atpA	−	−	+	−
atp6	+	+	+	−
atp8	+	+	+	−
atp9	−	(+)	+	−
NADH-Q1 complex				
nad1	+	(+)	+	+
nad2	+	(+)	nd	+
nad3	+	(+)	+	−
nad4	+	(+)	+	+
nad5	+	(+)	+	+
nad6	+	(+)	nd	+
Ribosome-associated protein(s)				
rps	−	+	+	−
Reverse transcriptase-like protein				
rtl	−	(+)[1]	+[1]	+[1]

Data were taken from Chomyn et al. (1985) for human, from Grivell (1983) for fungi, from Lonsdale (1988) for higher plants, and for *C. reinhardtii* from this study, Boer and Gray (1988a, b, c), Pratje et al. (1984, 1989), and Vahrenholz et al. (1985). +, the protein is encoded by mtDNA; −, the respective gene is absent from mtDNA; (+), gene present in mtDNA of some fungi, but absent from others; nd, not yet identified;

[1] The *rtl* gene of *C. reinhardtii* does not show homology to any mitochondrial genes of fungi or higher plants

A minimum of 32 tRNAs is needed to recognize all codons of the universal genetic code according to the wobble hypothesis of base pairing. In animal mitochondria a relaxed "2 out of 3" codon-anticodon pairing is assumed because only 22 tRNAs are encoded by mammalian mitochondrial DNA. In the 15.8 kb mtDNA of *C. reinhardtii* only 3 tRNA genes have been detected. These 3 tRNAs (Ma et al. 1988, 1989; Boer and Gray 1988c) for tryptophan, glutamine, and methionine do not represent the most frequently used codons. The TGG (Trp), CAA/CAG (Gln) and ATG (Met) codons occur with a frequency of 2.4%, 2.9% and 4.1% (without counting the eight initiation codons) respectively, whereas the TTG (Leu, 8.6%), GCT (Ala, 6%) and GGT (Gly, 6%) codons are used most frequently (Ta-

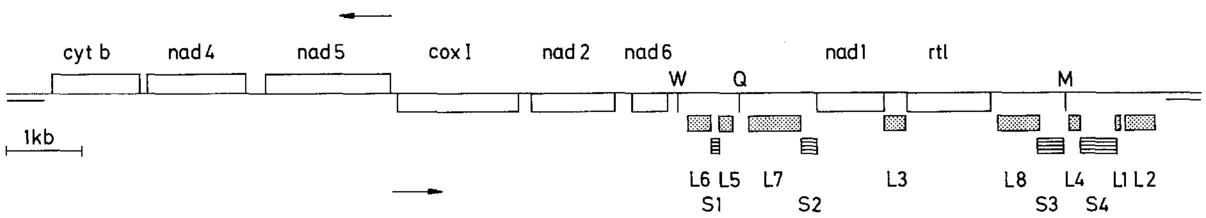

Fig. 4. Functional map of the 15.8 kb mitochondrial genome of *C. reinhardtii.* Genes for proteins: *cytb* encodes apocytochrome *b* (this study), *nad1* (Boer and Gray 1988a), *nad2, nad4* (Pratje et al. 1984, 1989), *nad5* (Vahrenholz et al. 1985) and *nad6* (Ma et al. 1988) encode subunits of NADH dehydrogenase, *coxI* (Vahrenholz et al. 1985) encodes subunit 1 of cytochrome oxidase, and rtl (Boer and Gray 1988a) encodes a reverse transcriptase-like protein. The letters W (tryptophan), Q (glutamine), and M (methionine) denote tRNA genes (Boer and Gray 1988c; Ma et al. 1988, 1989). Regions of the large (L, ▨) and small (S, ▤) scrambled rRNA genes are indicated (Boer and Gray 1988b). The inverted repeats at both linear termini are characterized by a *double line.* Directions of transcription are shown by *arrows*

Table 2. Codon usage in the mitochondrial genes of *C. reinhardtii*

Codon	Amino acid	*coxI*	*cytb*	*nad1*	*nad2*	*nad4*	*nad5*	*nad6*	*rtl*	Total	%
TTT	Phe	12	5	7	15	9	19	7	8	82	2.7
TTC	Phe	28	21	16	18	26	16	8	10	143	4.6
TTA	Leu	0	0	0	0	0	0	0	0	0	0
TTG	Leu	49	28	24	49	44	41	16	13	264	8.6
CTT	Leu	1	5	2	3	10	9	0	8	38	1.2
CTC	Leu	0	0	0	0	0	0	0	2	2	0.06
CTA	Leu	18	19	12	21	23	25	11	18	147	4.8
CTG	Leu	4	0	3	1	3	2	2	10	25	0.8
ATT	Ile	24	16	12	15	19	19	11	10	126	4.1
ATC	Ile	8	11	6	5	11	12	3	6	62	2
ATA	Ile	0	0	0	0	0	0	0	1	1	0.03
ATG	Met	23	15	14	14	20	28	8	12	134	4.3
GTT	Val	9	4	8	15	13	23	4	6	82	2.7
GTC	Val	13	7	4	2	5	8	0	3	42	1.4
GTA	Val	25	19	11	10	25	18	8	8	124	4
GTG	Val	0	1	0	1	1	2	1	6	12	0.4
TCT	Ser	8	7	8	10	9	9	8	4	63	2
TCC	Ser	0	1	3	2	0	8	0	5	19	0.6
TCA	Ser	0	0	0	0	0	0	0	0	0	0
TCG	Ser	0	0	0	0	0	0	0	0	0	0
CCT	Pro	3	1	4	0	9	1	1	7	26	0.8
CCC	Pro	1	0	0	0	0	2	0	2	5	0.2
CCA	Pro	19	18	6	10	11	10	5	2	81	2.6
CCG	Pro	0	0	3	1	0	0	0	1	5	0.2
ACT	Thr	23	14	6	12	17	19	6	13	110	3.6
ACC	Thr	10	7	5	8	7	12	3	4	56	1.8
ACA	Thr	0	0	0	0	0	0	0	2	2	0.06
ACG	Thr	0	0	0	0	0	0	0	0	0	0
GCT	Ala	31	21	28	23	29	31	10	12	185	6
GCC	Ala	15	8	9	13	10	20	3	9	87	2.8
GCA	Ala	0	0	0	2	0	0	1	4	7	0.2
GCG	Ala	0	0	0	1	0	1	0	4	6	0.2
TAT	Tyr	5	5	5	6	5	7	1	5	39	1.3
TAC	Tyr	10	17	5	6	9	11	5	7	70	2.3
TAA	*	1	1	1	0	1	1	1	0	6	0.2
TAG	*	0	0	0	1	0	0	0	1	2	0.06
CAT	His	2	2	1	4	5	6	1	6	27	0.9
CAC	His	16	12	0	7	8	9	0	12	64	2.1
CAA	Gln	8	9	5	7	5	5	1	24	64	2.1
CAG	Gln	1	3	2	5	0	3	0	12	26	0.8
AAT	Asn	0	0	0	3	1	2	3	8	17	0.6
AAC	Asn	14	15	4	2	5	12	3	8	63	2
AAA	Lys	4	5	7	6	6	8	0	16	52	1.7
AAG	Lys	0	0	0	1	1	1	1	6	10	0.3
GAT	Asp	6	3	4	3	4	11	0	9	40	1.3
GAC	Asp	6	4	3	3	2	6	1	4	29	0.9
GAA	Glu	0	0	0	0	0	0	0	0	0	0
GAG	Glu	7	5	9	8	11	10	5	8	63	2
TGT	Cys	1	1	3	5	3	3	0	2	18	0.6
TGC	Cys	4	2	2	6	7	6	1	3	31	1
TGA	*	0	0	0	0	0	0	0	0	0	0
TGG	Trp	14	11	5	10	9	13	1	11	74	2.4
CGT	Arg	10	5	4	4	7	10	6	6	52	1.7
CGC	Arg	3	2	5	2	3	5	0	5	25	0.8
CGA	Arg	0	0	0	2	0	1	0	2	5	0.2
CGG	Arg	0	0	0	0	0	0	0	0	0	0
AGT	Ser	9	7	10	12	9	19	2	3	71	2.3
AGC	Ser	12	10	6	7	12	8	2	4	61	2
AGA	Arg	0	0	0	0	0	0	0	0	0	0
AGG	Arg	0	0	0	0	0	0	0	0	0	0
GGT	Gly	45	29	14	14	22	42	10	8	184	6
GGC	Gly	4	5	6	6	4	13	1	3	42	1.4
GGA	Gly	0	1	1	2	4	0	2	3	13	0.4
GGG	Gly	0	0	0	0	0	0	0	3	3	0.1
		506	382	293	383	444	547	163	369	3087	100

Codon usages were determined from the DNA sequences of *coxI* (Vahrenholz et al. 1985), *cytb* (this study), *nad1, rtl* (Boer and Gray 1988a), *nad2* (Pratje et al. 1984), *nad4* (Pratje et al. 1989), *nad6* (taken from Boer and Gray 1988c), and *nad5* (Vahrenholz et al. 1985; two corrections are included: a T is inserted at position 602 and at position 572 the T is replaced by a C in the non-coding strand in Fig. 4). Results of transcript mapping (Boer and Gray 1986) are taken into account to define translation initiation codons. Triplets that are indicated by *asterisks* code for the termination of polypeptide chain synthesis ble 2). Because the AGA/AGG codon family is not used at all (Table 2), a minimum number of 22 or 23 tRNAs should be sufficient for translation, depending on the number of $tRNA^{Met}$. If protein synthesis takes place within the mitochondria of *C. reinhardtii* most of the tRNAs have to be imported. Evidence for the import of tRNAs into mitochondria of *Tetrahymena* has been reported by Chiu et al. (1975) and recently the presence of nuclear-encoded tRNAs was demonstrated in mitochondria of higher plants (Marechal-Drouard et al. 1988). Due to the low number of mitochondrial tRNA genes, *Chlamydomonas* could represent a good model system for the study of tRNA import into mitochondria.

The 15.8 kb DNA appears to represent the entire mitochondrial genome of *C. reinhardtii*, because no other DNA species can be extracted from isolated mitochondria using various procedures. This linear genome is small, displays a compact organization of thirteen genes and is transmitted predominantly by the mt^- mating type. Several unexpected features of this mitochondrial genome are unique and will form the basis for future investigations of mitochondrial import, replication, and interactions with nuclear, chloroplast, and basal body DNA.

Acknowledgements. We thank Dr. B. Dujon for exchange of unpublished data, Dr. W. Ecke for help with the computer analysis, and Miss S. Bühler for technical assistance.

References

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

Boer PH, Gray MW (1986) The *URF5* gene of *Chlamydomonas reinhardtii* mitochondria: DNA sequence and mode of transcription. EMBO J 5:21–28

Boer PH, Gray MW (1988a) Genes encoding a subunit of respiratory NADH dehydrogenase (ND1) and a reverse transcriptase-like protein (RTL) are linked to ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. EMBO J 7:3501–3508

Boer PH, Gray MW (1988b) Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell 55:399–411

Boer PH, Gray MW (1988c) Transfer RNA genes and the genetic code in *Chlamydomonas reinhardtii* mitochondria. Curr Genet 14:583–590

Chiu N, Chiu A, Suyama Y (1975) Native and imported transfer RNA in mitochondria. J Mol Biol 99:37–50

Chomyn A, Mariottini P, Cleeter MWJ, Ragan CI, Doolittle RF, Yagi AM, Hatefi Y, Attardi G (1985) Functional assignment of the products of the unidentified reading frames of human mitochondrial DNA. In: Quagliariello et al. (eds) Achievements and perspectives of mitochondrial research, vol 2. Elsevier, Amsterdam, pp 259–275

Clary DO, Wahleithner JA, Wolstenholme DR (1984) Sequence and arrangement of the gene for cytochrome b, *URF1, URF4L, URF4, URF5, URF6* and five tRNAs in *Drosophila* mitochondrial DNA. Nucleic Acids Res 12:3747–3762

Colleaux L, Michel-Wolwertz M-R, Matagne RF, Dujon B (1990) The apocytochrome *b* gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Mol Gen Genet (in press)

Covello PS, Gray MW (1989) RNA editing in plant mitochondria. Nature 341:662–666

Dawson AJ, Jones VP, Leaver CJ (1984) The apocytochrome *b* gene in maize mitochondria does not contain introns and is preceded by a potential ribosome binding site. EMBO J 3:2107–2113

Grivell LA (1983) Mitochondrial gene expression 1983. In: Schweyen RW et al. (eds), Mitochondria 1983. de Gruyter, Berlin New York, pp 25–45

Gualberto JM, Lamattina L, Bonnard G, Weil J-H, Grienenberger J-M (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature 341:660–662

Harries EH (1989) The *Chlamydomonas* sourcebook. Academic Press, San Diego

Helmer-Citterich MH, Morelli G, Macino G (1983) Nucleotide sequence and intron structure of the apocytochrome b gene of *Neurospora crassa* mitochondria. EMBO J 2:1235–1242

Kück U, Neuhaus H (1986) Universal genetic code evidenced in mitochondria of *Chlamydomonas reinhardtii*. Appl Microbiol Biotechnol 23:462–469

Lonsdale DM (1988) Plant mitochondrial genes and sequences. Plant Mol Biol Reporter 6:266–273

Ma DP, Yang YW, Hasnain SE (1988) Nucleotide sequence of *Chlamydomonas reinhardtii* mitochondrial genes coding for subunit 6 of NADH dehydrogenase and $tRNA^{Trp}$. Nucleic Acids Res 16:11373

Ma DP, Yang YW, Hasnain SE (1989) Nucleotide sequence of *Chlamydomonas reinhardtii* mitochondrial genes coding for $tRNA^{Gln}$ (UUG) and $tRNA^{Met}$ (CAU). Nucleic Acids Res 17:1256

Marechal-Drouard L, Weil J-H, Guillemaut P (1988) Import of several tRNAs from cytoplasm into the mitochondria in bean *Phaseolus vulgaris*. Nucleic Acids Res 16:4777–4788

Matagne RF, Michel-Wolwertz MR, Munaut C, Duyckaerts C, Sluse F (1989) Induction and characterization of mitochondrial DNA mutants in *Chlamydomonas reinhardtii*. J Cell Biol 108:1221–1226

Nobrega FG, Tzagoloff A (1980) Assembly of the mitochondrial membrane system. DNA sequence and organization of the cytochrome b gene in *S. cerevisiae* D273-10B. J Biol Chem 255:9828–9837

Pratje E, Schnierer S, Dujon B (1984) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the DNA sequence of a region showing homology with mammalian *URF2*. Curr Genet 9:75–82

Pratje E, Vahrenholz C, Bühler S, Michaelis G (1989) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the ND4 gene encoding a subunit of NADH dehydrogenase. Curr Genet 16:61–64

di Rago J-P, Coppee J-Y, Colson A-M (1989) Molecular basis for resistance to myxothiazol, mucidin (strobilurin A), and stigmatellin. J Biol Chem 264:14543–14548

Rao JKM, Argos P (1986) A conformational preference parameter to predict helices in integral membrane proteins. Biochim Biophys Acta 869:197–214

Ryan R, Grant D, Chiang K-S, Swift H (1978) Isolation and characterization of mitochondrial DNA from *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 75:3268–3272

Sanger F, Nicklen S, Coulsen AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Vahrenholz C, Pratje E, Michaelis G, Dujon B (1985) Mitochondrial DNA of *Chlamydomonas reinhardtii*: sequence and arrangement of *URF5* and the gene for cytochrome oxidase subunit I. Mol Gen Genet 201:213–224

Waring RB, Davies RW, Lee S, Grisi E, Berks MM, Scazzocchio C (1981) The mosaic organization of the apocytochrome b gene of *Aspergillus nidulans* revealed by DNA sequencing. Cell 27:4–11

Communicated by R.G. Herrmann