Volume 11 Number 13 1983 Nucleic Acids Research

Biogenesis of mitochondria: the mitochondrial gene (*aap1*) coding for mitochondrial ATPase subunit 8 in *Saccharomyces cerevisiae*

Ian G.Macreadie, Charles E.Novitski, Ronald J.Maxwell, Ulrik John, Beng-Guat Ooi, Gabrielle L.McMullen, H.B.Lukins, Anthony W.Linnane and Phillip Nagley*

Department of Biochemistry, Monash University, Clayton, Victoria 3168, Australia

Received 20 May 1983; Accepted 13 June 1983

ABSTRACT

A mitochondrial gene (denoted *aap1*) in *Saccharomyces cerevisiae* has been characterized by nucleotide sequence analysis of a region of mtDNA between the *oxi3* and *oli2* genes. The reading frame of the *aap1* gene specifies a hydrophobic polypeptide containing 48 amino acids. The functional nature of this reading frame was established by sequence analysis of a series of *mit*⁻ mutants and revertants. Evidence is presented that the *aap1* gene codes for a mitochondrially synthesized polypeptide associated with the mitochondrial ATPase complex. This polypeptide (denoted subunit 8) is a proteolipid whose size has been previously assumed to be 10 kilodaltons based on its mobility on SDS-polyacrylamide gels, but the sequence of the *aap1* gene predicts a molecular weight of 5,815 for this protein.

INTRODUCTION

The mitochondrial proton translocating, oligomycin sensitive ATPase complex (mtATPase) isolated from *Saccharomyces cerevisiae* contains between nine and fourteen polypeptides (1-3). The mtATPase complex contains a hydrophobic membrane sector denoted F_0 (4) and an F_1 sector consisting of five different relatively hydrophilic subunits. In addition to F_0 and F_1, components designated F_1 inhibitor protein and oligomycin sensitivity conferring protein (OSCP) have been identified (5,6). Most of the mtATPase subunits are encoded by nuclear genes and imported into mitochondria from the cytoplasm (4,7). Only the F_0 sector contains polypeptides coded by genes in mtDNA.

To date only two mitochondrial genes coding for mtATPase complex components have been well characterized; these specify subunits 6 and 9, which are F_0 components. The *oli1* gene codes for subunit 9, a 7.6 kilodalton (kd) proteolipid, as shown by a comparison of the amino acid sequence of the proteolipid (8) with that predicted by the nucleotide sequence of the *oli1* gene (9,10). The *oli2* gene specifies mtATPase subunit 6, a polypeptide

© IRL Press Limited, Oxford, England.

which has an apparent molecular weight of 20 kd based on its gel mobility and which is involved in the coupling of respiration and oxidative phosphorylation (11). This gene-protein relationship was deduced from analysis of a number of *oli2 mit*⁻ mutants (unable to grow on non-fermentable substrates) in which the 20 kd polypeptide is replaced by smaller polypeptides that can be immunoprecipitated with anti-holo-mtATPase antibody (12,13). An additional *oli2 mit*⁻ mutant contains a 20 kd ATPase subunit with altered pI (14). DNA sequencing studies of the *oli2* region in two wild-type strains of *S. cerevisiae* (15,16) have led to the assignment of the coding region to a 780 bp open reading frame with a predicted product of molecular weight 28,257.

The status of two other mitochondrially synthesized polypeptides, subunits 5 and 8, that were originally suggested to form part of the mtATPase complex in yeast (17), has been addressed recently. The protein band denoted subunit 5 (about 35 kd) has been shown to be a contaminant derived from other mitochondrial enzyme complexes (18,19). A second polypeptide denoted subunit 8 has an apparent size of about 10 kd based on its mobility in sodium dodecyl sulphate-polyacrylamide gels. It is seen in gel displays of total products of mitochondrial protein synthesis in wild-type cells. This protein is frequently found in immunoprecipitates made from Triton X-100 mitochondrial extracts using anti-holo-mtATPase antibody but in studies on mitochondrial proteins labelled in the presence of cycloheximide, the newly made subunit 8 is not seen in preparations of the mtATPase complex purified on glycerol gradients (18).

The present work is concerned with the definition of a third mitochondrial gene, the product of which is associated with the mtATPase complex. This gene is here denoted *aap1* (ATPase associated protein). The available evidence suggests strongly that it codes for subunit 8. We recently reported (20) the preliminary characterization of two novel mitochondrial *mit*⁻ mutants, M26-10 and M31, that are unable to grow on non-fermentable substrates. These mutants have been of interest because their *mit*⁻ mutations appeared to map outside known mitochondrial genes and because the mtATPase complex in such mutants is affected. In these mutants, subunit 8 was not observed either in sodium dodecyl sulphate lysates of mitochondria or in immunoprecipitates made from Triton X-100 mitochondrial extracts of these mutants treated with anti-holo-mtATPase antibody.

In this paper we report the identification of the *aap1* gene through analysis of mtDNA sequences of the wild-type, mutant, and revertant strains.

The predicted gene product of molecular weight 5815 is a hydrophobic protein, 48 amino acids in length, whose amino acid composition fits very well with that described (21) for a purified preparation of an ATPase associated proteolipid with an apparent size of 10 kd that corresponds to subunit 8 of the mtATPase complex.

MATERIALS AND METHODS

Yeast strains

Saccharomyces cerevisiae strain J69-1B α *ade1 his* [*rho*$^+$] (22) is the wild-type parent of *mit*$^-$ mutants M26-10, M31 and M68-5 which were derived by mutagenesis with manganese chloride (20,23). Strains M26-10-R1, M26-10-R7 and M26-10-R13 are spontaneous mitochondrial revertants of *mit*$^-$ strain M26-10. M26-10-R7 is temperature sensitive, with almost normal oxidative growth at 28°C but with no oxidative growth at 36°C and slow oxidative growth at 18°C. M26-10-R13 is also temperature sensitive, with slower oxidative growth than M26-10-R7 at 28°C and 18°C. Petite G5 (20) arose spontaneously from 70M *a ade1 lys2 trp1* [*rho*$^+$ *oli2-23r*] (previously denoted *oli23-r*) (22). Strain BT2-1 *a kar1-1 leu1* [*rho*^{o}] (20) was used in cytoduction experiments. Petite 620M was isolated from a cross between strains M26-10 and BT2-1 as described in the Results section.

Nucleotide sequencing

The source of DNA for sequencing mtDNA of strains J69-1B (wild-type), M31, M68-5, M26-10-R7 and M26-10-R13 was the EcoRI restriction fragment R7 (Fig. 1) of the particular strain cloned in pBR325 and propagated in E. *coli* ED8654 (24). The EcoRI R7 fragment of M31 was also cloned into vector M13mp9 (25). The mtDNA of petite G5, of petite 620M and of revertant M26-10-R1 was isolated from each yeast strain and purified on density gradients containing the Hoescht dye 33 258 (26). Specific restriction fragments (Fig. 1) were 3' end-labelled by a fill-in reaction using [α-^{32}P]-labelled deoxyribonucleoside triphosphates with Klenow fragment of DNA polymerase I (24). In some cases, the labelled restriction fragments were denatured in the presence of 0.05 M NaOH and strands were separated by electrophoresis through neutral 5% polyacrylamide gels. In other cases, the fragments were cleaved by a second restriction enzyme and the double stranded DNA fragments separated by electrophoresis through 1% or 2% agarose or 5% polyacrylamide gels. The DNA was electroeluted from the gels and in some instances purified by DEAE-cellulose (DE52) chromatography. The nucleotide sequence of these fragments was determined by the method of Maxam and

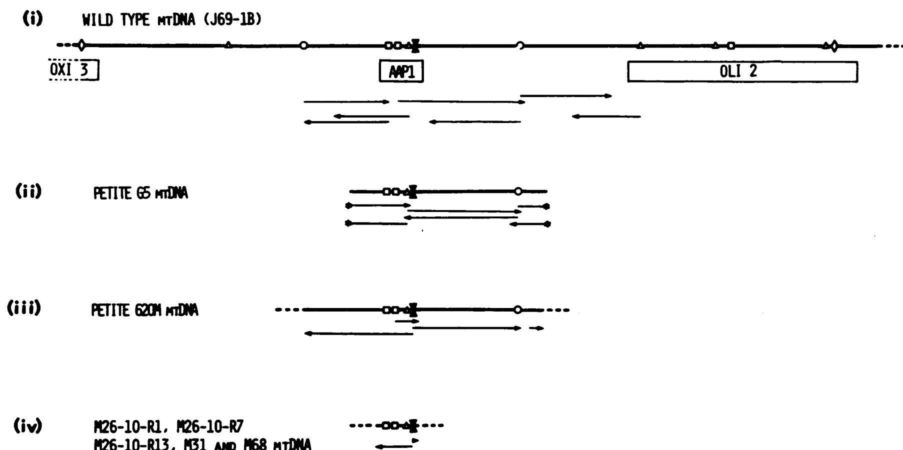

Figure 1. The nucleotide sequencing strategies for a region of the mitochondrial genome between the *oxi3* and *oli2* genes. The nucleotide sequence was determined from (i) wild-type (J69-1B) mtDNA EcoRI fragment R7 (2.57 kb) cloned into bacterial plasmid pBR325, (ii) petite G5 mtDNA, (iii) petite 620M mtDNA, (iv) mtDNA of *mit*⁻ strains M31 and M68-5 and of revertant strains M26-10-R1, M26-10-R7 and M26-10-R13. A restriction map is shown for a segment of each mtDNA. Restriction sites Sau3AI (Δ), HpaII (o), HinfI (□), XbaI (⧗) and EcoRI (◇) are indicated. Solid lines represent DNA segments for which nucleotide sequence was obtained and dashed lines indicate that there are unsequenced DNA regions flanking the sequenced regions. Below each restriction map the arrows indicate the restriction sites at which end labelling was carried out and the direction and extent of sequencing. The full sequence of the region between *oxi3* and *oli2* genes will be published separately (Novitski et al., in preparation). The restriction map of petite G5 is circular and therefore two of the fragments sequenced appear to contain discontinuities (stars) in this linear representation. The locations of the *aap1*, *oli2* and *oxi3* genes are indicated.

Gilbert (27) or, following cloning into M13mp9, by the Sanger dideoxy chain termination method (28) using as a primer the oligonucleotide (complementary to mtDNA sequence) 5'-CTAGATACATATAATCT-3' synthesized by the phosphotriester method (29) in this laboratory.

RESULTS

Localization of the *aap1-1* and *aap1-2* mutations from sequence analysis of G5 mtDNA

The mitochondrial *mit*⁻ mutants M26-10 and M31 (20) serve to define the *aap1* locus. The preliminary localization of the mutations in these strains was made possible by use of respiratory deficient petite G5 (Fig. 1), which when mated to respiratory deficient strain M26-10 or to M31 gave rise to

```
-150                                                        TATAATATTTTTTTTTATAAATATA TTATTTATATTAATAAATATATATA

-100 TTATTTTTATAATATATATATATTT TTATTAAATATTTATTAAATATTTA TTAAATTATTATAATGTTGTTATTA ATCTTATTAAAAAATATATATAAAA
     AAP1
     Met Pro Gln Leu Val Pro Phe Tyr Phe Met Asn Gln Leu Thr Tyr Gly Phe Leu Leu Met Ile Thr Leu Leu Ile Leu
   1 ATG CCA CAA TTA GTT CCA TTT TAT TTT ATG AAT CAA TTA ACA TAT GGT TTC TTA TTA ATG ATT CTA TTA TTA ATT TTA
                                            HinfI                                  HinfI

     Phe Ser Gln Phe Phe Leu Pro Met Ile Leu Arg Leu Tyr Val Ser Arg Leu Phe Ile Ser Lys Leu ***
  79 TTC TCA CAA TTC TTT TTA CCT ATG ATC TTA AGA TTA TAT GTA TCT AGA TTA TTT ATT TCT AAA TTA TAA TATATATTATT
                                  Sau3AI                    XbaI

 159 AATTTATTTATTCATATAAATATTA TTATTATATATAAATATTAATAATA TTTATACTTATTTAATAATAATAAT AAAATAAAAAAAATAATTATAATTT

 259 AATATATTTAATATATTTCCTTACA GACTATATATTTATATATATATATT AAATACAATTTAATTTAATTTAATT ATGTTATTTATTAAATAAAGTTATA

 359 TTATAATATAATAACAATATTATAT ATTATTATATAATTATAATATATTT TAATATAATTATCAAAAGAAATAAT AAAAAATATTAATAAGAATATAATT

 459 TAATAATTATTAAAAAAATTCTTAT TTATAGTCCGGTCCGCCCCCGCGGG GCGGACCCCAAAGGAGGAGTAATAA AAATTATTAAATACAAATATTATAT
                                     HpaII

 559 ATATATAATTCATTATATATATATA TATATAATAATTAATCTTATTTTTT
```

Figure 2. The nucleotide sequence of a portion of the wild-type (J69-1B) mtDNA between the *oxi3* and the *oli2* gene. The segment of mtDNA from which petite G5 mtDNA (681 bp) is derived consists of the sequence (-99 to +572) between the two boxed regions (5'-ATATATATAT-3') plus a single copy of 5'-ATATATATAT-3'. Petite G5 mtDNA sequence is identical to the corresponding J69-1B mtDNA sequence. The reading frame between nucleotides +1 and +147 which constitutes the *aap1* gene in wild-type (J69-1B) mtDNA is translated into amino acids using the yeast mitochondrial genetic code (see Table 3). Nucleotide +41 is C, correcting the previously reported T at that position (16). The A in the initiating ATG codon of the *oli2* gene is at nucleotide +854. Nucleotides involved in a base pairing within a G,C-rich stem and loop structure (→ ←) are discussed elsewhere (16). Numbers at the left refer to the leftmost nucleotide on each line.

respiratory competent diploid progeny. It was therefore concluded (20) that the sites of the mutations, *aap1-1* and *aap1-2*, which result in the *mit*⁻ phenotype of strain M26-10 and M31, respectively, lie within that segment of the wild-type mitochondrial genome which is also present in petite G5 mtDNA. The source of the mtDNA segment retained in petite G5 was initially determined to be located between the *oxi3* and the *oli2* genes, based on restriction mapping and hybridization of labelled G5 mtDNA to wild-type mtDNA restriction fragments (20).

The *aap1* gene was definitively located and the reading frame identified by analysis of mtDNA sequences of wild-type strain (J69-1B), *mit*⁻ strain M26-10, and strains M26-10-R1, M26-10-R7 and M26-10-R13(revertants of M26-10). Wild-type (J69-1B) nucleotide sequence (Fig. 2) of a segment of the mitochondrial genome between the *oxi3* and *oli2* genes was determined from J69-1B EcoRI fragment R7 (Fig. 1) cloned into the bacterial plasmid pBR325. The

complete 681 bp nucleotide sequence of petite G5 mtDNA (Fig. 2) was also determined and found to be identical to a continuous 681 bp segment of the J69-1B mitochondrial genome. The 681 bp G5 mtDNA sequence consists of 671 bp (-99 to +572, Fig. 2) plus the sequence 5'-ATATATATAT-3', which is identical to a directly repeated sequence flanking the unique 671 bp segment in J69-1B mtDNA (-109 to -100 and +573 to +582); these direct repeats are likely to represent the excision sequences (30) at which recombination in wild-type mtDNA occurred to give rise to petite G5 mitochondrial genome. The *mit*$^-$ mutation, *aap1-1*, in strain M26-10 is therefore localized to a mtDNA region between nucleotides -109 and +582. Consideration of the sequence (Fig. 2) included in the G5 mtDNA genome reveals two open reading frames covering the sequences +1 to +147, and -37 to +122. These open reading frames potentially code for polypeptides 48 and 53 amino acids in length, respectively. As we will show below, the sequence in mutant M26-10 and revertant M26-10-R1 localizes the *mit*$^-$ mutation in M26-10 to lie within these reading frames. Moreover, the sequence in revertant M26-10-R7 reveals that it is the 48 amino acid reading frame which constitutes the *aap1* gene and this conclusion is supported by amino acid composition of the gene product and considerations of codon utilization (see Discussion). The predicted polypeptide sequence of the *aap1* gene product is shown in Fig. 2.

Construction of petite strain 620M carrying the *aap1-1* mutation

The mtDNA sequence change at the site of the *aap1-1* mutation was deduced from analysis of the mtDNA of petite 620M mtDNA. This petite mutant arose during a cross between *mit*$^-$ strain M26-10 and *rho*^{o} strain BT2-1 (*a leu1 kar1-1*). Petite 620M is a haploid derived by cytoduction (31,32) and which carries the nuclear genetic markers of BT2-1 and a 5 kb segment of the mitochondrial genome from M26-10. The following experiments indicated that the *aap1-1* mutation is present in 620M mtDNA. First, when 620M was mated to either *oxi3* or *oli2 mit*$^-$ strains, respiratory competent progeny were obtained, indicating that the *oxi3* and *oli2* regions (and probably the mtDNA segment of interest between them) are present in the mtDNA of this petite. This interpretation was supported by the presence in EcoRI digests of 620M of a fragment corresponding in size to the 2.57 kb wild-type EcoRI fragment R7. Finally, when 620M was mated to the *rho*$^+$ strain J69-1B, a particular class of respiratory deficient haploid cells, carrying the J69-1B nuclear genotype (α *ade1 his*) and showing characteristic *mit*$^-$ colony morphology, was produced at a moderate frequency (about 1.5% of the cells in the mixture after this type of mating). Thirteen of these *mit*$^-$ cytoductants were picked,

Table 1. Nucleotide sequence differences between wild-type (J69-1B) mtDNA and specific mutant and revertant strains derived from J69-1B.

Strain		Insertion	Deletion	Substition
M26-10	Mutant		T (+50/+53)	C → T (+41)
M26-10-R1	Revertant	T (+47/+51)	T (+50/+53)	C → T (+41)
M26-10-R7	Revertant	T (+23/+28)	T (+50/+53)	C → T (+41)
M26-10-R13	Revertant	TT (+23/+28)	T (+50/+53) T (+18/+23)	C → T (+41)
M68-5	Mutant		T (+50/+53)	
M31	Mutant	T (+9/+12) T (-6/-8)		C → T (+ 7) C → T (+51)

For each strain, the nature of the change and the nucleotide location of the change, with reference to the nucleotide numbering in Fig. 2 is given. T (X/Y) means that a single base T is inserted or deleted between nucleotides numbered X and Y.

mated to petite G5, and in all cases respiratory competent progeny were obtained. This indicated that the cytoductants carried the *aap1-1* *mit*⁻ allele in otherwise wild-type mtDNA. Such mtDNA molecules arose through recombination between the mitochondrial genomes of strains J69-1B and 620M followed by segregation of the recombinant *mit*⁻ mtDNA genome into haploid cells carrying the J69-1B nucleus. It was therefore concluded that petite strain 620M mtDNA carries the *aap1-1* mutation.

The nucleotide sequence changes in the *mit*⁻ mutants and mitochondrial revertants

In order to determine the nucleotide sequence change at the site of the *aap1-1* mutation, 620M mtDNA nucleotide sequence through more than 95% of the region from nucleotides -109 to +582 was obtained (Fig. 1). The only nucleotide differences noted between 620M and wild-type sequence resulted from the deletion of a T between nucleotides +50 and +53 and from the base change of C to T at nucleotide +41 (Table 1). These data correct a previous report (16) that the two sequences were identical at +41. The effect of the frameshift mutation on the 48 amino acid reading frame is to change the 18th codon to a missense codon (TTA[leu]→TAT[tyr]) and the 19th codon to a termination codon (TAA); the predicted *aap1* gene product in M26-10 is therefore a polypeptide only 18 amino acids long (Fig. 3).

Strain	Status	Predicted Amino Acid Sequence				
J69-1B	Wild-type	MPQLVPFYFM	NQLTYGFLLM	ITLLILFSQF	FLPMILRLYV	SRLFISKL
M26-10	Mutant	MPQLVPFYFM	NQLMYGFY			
M26-10-R1	Revertant	MPQLVPFYFM	NQLMYGFTLM	ITLLILFSQF	FLPMILRLYV	SRLFISKL
M26-10-R7	Revertant	MPQLVPFYFY	ESINMWFTLM	ITLLILFSQF	FLPMILRLYV	SRLFISKL
M26-10-R13	Revertant	MPOLVPFIFY	ESINMWFTLM	ITLLILFSQF	FLPMILRLYV	SRLFISKL
M68-5	Mutant	MPQLVPFYFM	NQLTYGFY			
M31	Mutant	MP				

Figure 3. Predicted *aap1* gene product from wild-type, mutant and revertant strains. The predicted sequence in each strain is given in one letter code and differences from J69-1B sequence are underlined.

Analysis of the mtDNA sequence changes in revertants M26-10-R1, M26-10-R7 and M26-10-R13 confirms that it is the deletion at +52 that leads to respiratory deficiency. The only difference observed between M26-10-R1 sequence in the vicinity of the *aap1* mutation (Fig. 1) and M26-10 mtDNA sequence was the insertion of a T between nucleotides +47 and +51 (Table 1). The prematurely terminated *aap1* reading frame of M26-10 is reopened by the insertion of the T residue, restoring the predicted wild-type 48 amino acid product except for the replacement of leucine by threonine at residue 18 and the replacement of threonine by methionine at residue 14 due to the base substitution at nucleotide +41 (Fig. 3). In the case of revertant strain M26-10-R7, the insertion of T between nucleotides +23 and +28 restores the reading frame of the *aap1* gene and changes amino acid residues 10 to 16 and residue 18 in the gene product (Fig. 3). In the case of revertant strain M26-10-R13, the deletion of T between nucleotide +18 and +23 and the insertion of TT between nucleotides +23 and +28 restores the reading frame of the *aap1* gene and changes amino acid residues 8, 10-16 and residue 18 (Fig. 3). These extensive substitutions correlate with the temperature sensitivity of revertant strains M26-10-R7 and M26-10-R13 (see Discussion).

The mtDNA sequence of *mit*$^-$ strain M68-5 indicates only the deletion of a T between nucleotides +50 and +53. This is the same as *mit*$^-$ strain M26-10, but in this case the base change at nucleotide +41 has not occurred. The mtDNA sequence of *mit*$^-$ strain M31 reveals four nucleotide sequence changes with respect to J69-1B (Table 1). The base change of C to T at nucleotide position +7 creates a TAA as the third codon and thus causes premature termination. The insertion of a T between nucleotides +9 and +10 is a frameshift mutation in such a position that if T reverted to C at position +7, the *aap1* gene product sequence would still be incorrect after

the third N-terminal residue. The predicted requirement of at least two mutational events to lead to reversion is consistent with our failure to detect any revertants in cultures of strain M31.

As mentioned above there are in fact two overlapping reading frames in this region of mtDNA, potentially coding for polypeptides 48 amino acids long and 53 amino acids long, respectively. It can be formally considered that the deletion of T from between nucleotides +50 and +53 in *mit*$^-$ mutant M26-10 could result in a frameshift mutation in either of these two reading frames. Analysis of the sequence changes in revertants allows us to exclude the longer of these overlapping reading frames as being required for oxidative growth. The *mit*$^-$ mutation in M26-10 would lead to an extended polypeptide 60 amino acids long, including 30 of the original hypothetical N-terminal residues. In revertant M26-10-R7, the 53 amino acid reading frame is not restored, but rather truncated to 25 amino acids, including only 19 of the original hypothetical N-terminal residues. The reversion in strain M26-10-R13 has a similar effect. The failure to restore an intact open reading frame in the revertants thus eliminates the possibility that this reading frame codes for a protein indispensible for respiratory function.

DISCUSSION

The *aap1* gene

The data presented here demonstrate that the *mit*$^-$ mutations in *S. cerevisiae* strains M26-10, M31 and M68-5 are located in a previously unidentified mitochondrial gene, designated *aap1*. The *aap1* gene product is a polypeptide 48 amino acids long with a predicted molecular weight of 5815. Two-thirds of the residues are non-polar. Based on one index of hydrophilicity (33) with a scale of -3.4 (hydrophobic) to +3.0 (hydrophilic), the predicted 5.8 kd *aap1* gene product is more hydrophobic (index = -1.1) than either mtATPase subunit 6 (index = -0.9) or subunit 9 (index = -0.8).

Evidence that the *aap1* gene product is mtATPase subunit 8

We have previously reported that the *mit*$^-$ mutations in strains M26-10 and M31 appear to have an effect on the assembly of the mtATPase (20). In immunoprecipitates of Triton X-100 mitochondrial extracts from wild-type (J69-1B) made with anti-holo-mtATPase antibody, subunit 6, subunit 9 and subunit 8 (apparent molecular weight 10,000) are present. In mutants M26-10 and M31, the subunit 8 is consistently missing both from sodium dodecyl

Table 2. Comparison of the amino acid composition of the predicted *aap1* gene product with that of the mtATPase proteolipid of apparent size 10kd[a].

Amino acid	*aap1* gene product	mtATPase proteolipid[b] (apparent 10kd)
*Ala	0	0.6
Arg	2	2.1
*Asn	1	1.9
*Asp	0	
Cys	0	0.3
*Gln	3	3.9
*Glu	0	
Gly	1	1.4
His	0	0
Ile	4	4.1
Leu	12	11.7
Lys	1	1.3
Met	4	3.7
Phe	7	7.1
Pro	3	nd
Ser	3	3.0
Thr	2	1.7
Trp	0	nd
Tyr	3	2.6
Val	2	2.5

[a] Data shown are number of residues per 48 amino acids. Asterisks indicate unmatching residues.

[b] These numbers are derived from the data of Velours et al. (21). Data for gln and glu are combined (glx), as are data for asn and asp (asx); nd signifies no data provided.

sulphate lysates of mitochondria and from immunoprecipitates of Triton X-100 extracts (20). Subunit 6, although not found in M26-10 and M31 immunoprecipitates, is present among the total mitochondrial translation products in these strains (L. Watkins, W. Choo, S. Marzuki and A.W. Linnane, unpublished data). These results suggest that the *mit*⁻ mutations occur in the gene coding for subunit 8. The absence of subunit 8 itself may result in the absence of subunit 6 in the mtATPase complex. Amino acid composition data support the conclusion that subunit 8 is the *aap1* gene product. The amino acid composition of the predicted *aap1* gene product fits very well with the amino acid composition described (21) for a purified preparation of an ATPase associated proteolipid with an apparent size of 10 kd that has similar properties to those of subunit 8 of the mtATPase complex. The observed number of each residue per 48 residues is in striking

agreement with the predicted *aap1* gene product except for a difference of 0.6 ala residues, 0.9 asx residues and 0.9 glx residues (Table 2). This strong similarity in amino acid composition between these two unrelated *S. cerevisiae* strains is notable, particularly in view of the observation that the *aap1* gene product can tolerate amino acid substitutions at several positions in its sequence, as is discussed below. It is not clear why this protein (whose predicted size is 5.8 kd) has the mobility on SDS-polyacrylamide gels expected for a 10 kd protein.

Codon usage

Analysis of revertant M26-10-R7 mtDNA sequence excludes the possibility that the reading frame potentially capable of coding for a 53 amino acid product (-37 to +122, Fig. 2) is required for respiratory competency. Furthermore, the following analysis of codon usage suggests that the 53 amino acid reading frame has no function. Among the seven known *S. cerevisiae* mitochondrial protein coding genes expressed at high levels (*oli1, oli2, oxi1, oxi2, oxi3, cyb, var1*), there is a marked preference for the utilization of a limited number of codons (Table 3). As a result of the 2119 usages listed for these genes in Table 3, 33 codons occur at a frequency of less than 0.8%. The frequency of utilization of these rarely used codons in yeast mitochondrial genes and published unassigned reading frames (URF) has been compared with the expected codon frequency determined from the base composition of the coding regions. The reading frame of the *aap1* gene identified here, capable of coding for a polypeptide 48 amino acids long, uses the less often used codons at a frequency significantly lower than expected by random chance (χ^2 = 8.9, p<0.005 (Table 3)), as is the case for each of the highly expressed genes and eight of the published yeast mitochondrial URF's which we have analyzed. In contrast, the overlapping reading frame, with the potential to code for a 53 residue product, has a usage of less commonly used codons not statistically different from what would be expected by random chance (χ^2 = 0.4,p>0.5).

Presence of an *aap1*-like gene in *Aspergillus*

The *aap1* gene in yeast strain J69-1B is located upstream of the *oli2* gene; the *oli2* AUG initiation codon starts at nucleotide +854 using the nucleotide numbering systems of Fig. 2. In the mtDNA of *Aspergillus nidulans* there is an unassigned reading frame, URFx, terminating about 100 nucleotides upstream of the gene for mtATPase subunit 6 (41). URFx could code for a hydrophobic protein 48 amino acids long; as yet there are no reported mutants nor data on possible function. The predicted poly-

Table 3. Codon utilization in the *aap1* gene and other yeast mitochondrial genes.

Amino acid	Codon	*aap1* gene	Other genes[a]	Amino Acid	Codon	*aap1* gene	Other genes[a]
Ala	GCA	0	53	Met	AUA	0	38
	U	0	66		AUG	4	67
	C	0	6 *				
	G	0	4 *	Phe	UUU	4	76
					C	3	60
Arg	AGA	2	40				
	G	0	0 *	Pro	CCA	2	35
	CGA	0	0 *		U	1	43
	U	0	1 *		C	0	4 *
	C	0	0 *		G	0	1 *
	G	0	1 *				
				Ser	UCA	1	74
Asn	AAU	1	186		U	2	38
	C	0	11 *		C	0	1 *
					G	0	0 *
Asp	GAU	0	51		AGU	0	24
	C	0	2 *		C	0	0 *
Cys	UGU	0	13 *	Thr	ACA	1	52
	C	0	1 *		U	0	41
					C	0	1 *
Gln	CAA	3	33		G	0	0 *
	G	0	4 *		CUA	1	16 *
					U	0	5 *
Glu	GAA	0	39		C	0	0 *
	G	0	3 *		G	0	2 *
Gly	GGA	0	28	Trp	UGA	0	39
	U	1	96		G	0	0 *
	C	0	1 *				
	G	0	7 *	Tyr	UAU	3	101
					C	0	13 *
His	CAU	0	51				
	C	0	4 *	Val	GUA	1	77
					U	1	47
Ile	AUU	3	189		C	0	4 *
	C	1	28		G	0	6 *
Leu	UUA	12	265	Stop	UAA	1	7
	G	0	2 *		G	0	0
Lys	AAA	1	66				
	G	0	3 *				

[a] The total number of times each codon is used in the following yeast mitochondrial genes: *oxi1* (34,35), *oxi2* (36), *oxi3* (37), *oli1* (9,10), *oli2* (15), *cyb* (38) and *var1* (39). The yeast mitochondrial genetic code differs from the *E. coli* code: UGA codes for tryptophan (41), CUN codes for threonine (40) and AUA codes for methionine (39). Asterisks indicate those codons referred to in the Discussion as less commonly used.

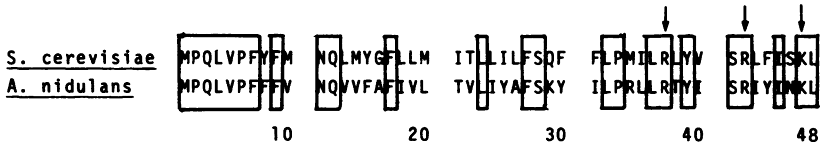

Figure 4. Comparison of the predicted amino sequence of the polypeptide specified by the yeast mitochondrial *aap1* gene in strain J69-1B with that of the URFx sequence in *Aspergillus nidulans* mtDNA (41). Arginine and lysine residues in the *aap1* gene product are indicated by vertical arrows (see text).

peptide product of the URFx reading frame is 50% homologous to the predicted yeast *aap1* gene product (Fig. 4), with the most notably conserved regions the N-terminal seven residues and the three positively charged residues found near the C-terminal end of the yeast product. Based on our genetic and biochemical data on the *aap1* gene, it may be expected that this unassigned reading frame in *Aspergillus* will turn out to be functional, probably coding for a proteolipid associated with the mtATPase complex.

The effect of amino acid substitution on the functional *aap1* gene product

A consideration of the predicted gene products of the various forms of the *aap1* gene (including the URFx of *Aspergillus*) suggests which amino acid residues are responsible for the temperature sensitivity of revertant M26-10-R7. This is achieved by noting differences in the chemical characteristic of the residue side chains. The yeast *aap1* gene product appears to be able to tolerate a number of changes in amino acid sequence. In addition to the wild-type form of the *aap1* gene product, there appear to be two other functional forms of the gene product in the yeast strains we have sequenced. One form in revertant M26-10-R1 has two replacements, met for thr at residue 14 and thr for leu at residue 18 (Fig. 5). The other form, in temperature sensitive revertant M26-10-R7, has four residues different in chemical characteristics from wild-type, at residues 10, 11, 15 and 18. Threonine at residue 18 is carried as well in M26-10-R1, which is not temperature sensitive. The *Aspergillus* residues 9 to 20 have similar chemical characteristics to those in yeast wild-type, except for val and phe as residues 14 and 15, respectively (Fig. 5); the 15th residue in *Aspergillus* (phe) is non-polar as is the 15th residue in M26-10-R7 (met). Therefore the prime candidates for the residues causing temperature sensitivity in M26-10-R7 are glu at position 11 and tyr at position 10.

Possible *aap1*-like gene in mammalian species

There is suggestive evidence for a gene analogous to *aap1* in organisms

Strain	Amino Acid Sequence													
J69-1B	F	Y	F	M	N	Q	L	T	Y	G	F	L	L	M
M26-1-R1	F	Y	F	M	N	Q	L	M	Y	G	F	T	L	M
M26-10-R7	F	Y	F	Y	E	S	I	N	M	W	F	T	L	M
M26-10-R13	F	I	F	Y	E	S	I	N	M	W	F	T	L	M
A. nidulans	F	F	F	V	N	Q	V	V	F	A	F	I	V	L
Residue No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20

Figure 5. Amino acid sequences of residues 7-20 in the *aap1* gene product of various respiratory competent *S. cerevisiae* strains (see Fig. 3) and of the predicted product of the URFx sequence of *Aspergillus nidulans* mtDNA (41). Residues which are chemically dissimilar to the corresponding residue in J69-1B are underlined.

other than yeast and *Aspergillus*. In the frugally organized mitochondrial genome of mammalian species (42-44) there is a reading frame (URF A6L) upstream of, and overlapping with the coding region for subunit 6. In addition to an extensive hydrophobicity in each of the predicted products, the most notable features among all of these eukaryotes is the common sequence MPQL at the N-terminus of each polypeptide and the presence of positively charged residues in the C-terminal portion of the polypeptide. The mammalian A6L URFs exhibit the strongest divergency of any mammalian mitochondrial gene or URF, there being only 51.5% homology between human and bovine polypeptide sequences, including homology at three of the positively charged residues in the C-terminal portion of the predicted product.

Further analogies among these mitochondrial systems occur in the transcriptional organization of the ATPase subunit 6 genes. In yeast mitochondria the most abundant *oli2* transcript (24), presumably the *oli2* mRNA, includes the sequence of the *aap1* gene. In HeLa cells, the nucleotide sequence of the mRNA (RNA 14) transcribed from the gene for mtATPase subunit 6 has been determined at the 5'-end (45) and at the 3'-end (46). These sequence data indicate that the mRNA contains the upstream overlapping URF A6L sequence. The location of a functional reading frame on the mRNA upstream of the ATPase subunit 6 coding region in these organisms could reflect the need for a hydrophobic polypeptide whose synthesis at a particular time and place is specifically required for the assembly of subunit 6 into the ATPase complex.

The role of the *aap1* gene product

We interpret the absence of mtATPase subunit 8 from the *aap1* mit^- mut-

ants (20) and the striking similarity in observed and predicted amino acid composition to indicate that the *aap1* gene product is mtATPase subunit 8. The observation that anti-holo-mtATPase precipitates subunit 8 suggests that subunit 8 is closely associated with the mtATPase complex. The function of subunit 8 within the complex is not clear at present. The hydrophobicity of the protein suggests that it is a membrane protein, and may be part of the mtATPase F_0 sector. This polypeptide, as analyzed by Guerin and Napias (47), has a phosphate binding property which could reflect a role in binding one of the substrates for the ATP synthetase reaction, or a role in transport of phosphate across the inner mitochondrial membrane, or both. The position of the *aap1* gene upstream of the *oli2* gene coding for mtATPase subunit 6 may be due to a temporal relationship in the assembly of subunits 6 and 8. The possibility that the *aap1* coding region is an exon of the *oli2* gene is inconsistent with preliminary data showing that the frame shift mutation in M26-10 does not prevent the appearance of subunit 6 among the mitochondrial translation products (data not shown).

Acknowledgements

This is paper number 59 in the series "Biogenesis of Mitochondria". Paper 58 is Cobon et al. (24). This work was supported by the Australian Research Grants Scheme. We thank Mr. Con Neri for his competent assistance in computer programming. A preliminary report of this work was presented to the 11th International Conference on Yeast Genetics in Montpellier, France, in September 1982.

*To whom correspondence should be sent

REFERENCES

1. Tzagoloff, A. and Meagher, P. (1971) J. Biol. Chem. 246, 7328-7336.
2. Ryrie, I.J. and Gallagher, A. (1979) Biochim. Biophys. Acta 545, 1-14.
3. Todd, R.D., Griesenbeck, T.A. and Douglas, M.G. (1980) J. Biol. Chem. 255, 5461-5467.
4. Tzagoloff, A., Macino, G. and Sebald, W. (1979) Ann. Rev. Biochem. 48, 419-441.
5. Ebner, E. and Maier, K.L. (1977) J. Biol. Chem., 252, 671-676.
6. Tzagoloff, A. (1970) J. Biol. Chem. 245, 1545-1551.
7. Schatz, G. and Mason, T.L. (1974) Ann. Rev. Biochem. 38, 51-87.
8. Sebald, W. and Wachter, E. (1978) in Energy Conservation in Biological Membranes, Schafer, G. and Klingenberg, M. Eds., pp.228-236, Springer-Verlag, Berlin.
9. Macino, G. and Tzagoloff, A. (1979) J. Biol. Chem. 254, 4617-4623.

10. Hensgens, L.A.M., Grivell, L.A., Borst, P. and Bos, J.L. (1979) Proc. Natl. Acad. Sci. USA 76, 1663-1667.
11. Murphy, M., Roberts, H., Choo, W.M., Macreadie, I., Marzuki, S., Lukins, H.B. and Linnane, A.W. (1980) Biochim. Biophys. Acta 592, 431-444.
12. Roberts, H., Choo, W.M., Murphy, M., Marzuki, S., Lukins, H.B. and Linnane, A.W. (1979) FEBS Lett. 108, 501-504.
13. Linnane, A.W., Astin, A.M., Beilharz, M.W., Bingham, C.G., Choo, W.M., Cobon, G.S., Marzuki, S., Nagley, P. and Roberts, H. (1980) in The Organization and Expression of the Mitochondrial Genome, Kroon, A.M. and Saccone, C. Eds., pp.253-263, Elsevier/North-Holland Biomedical Press, Amsterdam.
14. Stephenson, G., Marzuki, S. and Linnane, A.W. (1981) Biochim. Biophys. Acta 636, 104-112.
15. Macino, G. and Tzagoloff, A. (1980) Cell 20, 507-517.
16. Novitski, C.E., Macreadie, I.G., Maxwell, R.J., Lukins, H.B., Linnane, A.W. and Nagley, P. (1983) in Manipulation and Expression of Genes in Eukaryotes, Nagley, P., Linnane, A.W., Peacock, W.J. and Pateman, J.A. Eds., pp.257-268, Academic Press, Sydney.
17. Tzagoloff, A. and Meagher, P. (1972) J. Biol. Chem. 246, 594-603.
18. Orian, J.M., Murphy, M. and Marzuki, S. (1981) Biochim. Biophys. Acta 652, 234-239.
19. Orian, J.M. and Marzuki, S. (1981) J. Bact. 146, 813-815.
20. Macreadie, I.G., Choo, W.M., Novitski, C.E., Marzuki, S., Nagley, P., Linnane, A.W. and Lukins, H.B. (1982) Biochem. Int. 5, 129-136.
21. Velours, J., Esparza, M. and Guerin, B. (1982) Biochem. Biophys. Res. Commun. 109, 1192-1199.
22. Murphy, M., Choo, K-B., Macreadie, I., Marzuki, S., Lukins, H.B., Nagley, P. and Linnane, A.W. (1980) Arch. Biochem. Biophys. 203, 260-270.
23. Putrament, A., Baranowska, H. and Prazmo, W. (1973) Molec. Gen. Genet. 126, 357-366.
24. Cobon, G.S., Beilharz, M.W., Linnane, A.W. and Nagley, P. (1982) Curr. Genet. 5, 97-107.
25. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
26. Bingham, C.G. and Nagley, P. (1983) Biochim. Biophys. Acta. In press.
27. Maxam, A.M. and Gilbert, W. (1980) Meth. Enzymol. 65, 499-580.
28. Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
29. Duckworth, M.L., Gait, M.J., Goelet, P., Hong, G.F., Singh, M. and Titmas, R.C. (1981) Nucleic Acids Res. 9, 1691-1706.
30. Bernardi, G., Baldacci, G., Bernardi, G., Faugeron-Fonty, G., Gaillard, C., Goursot, R., Huyard, A., Mangin, M., Marotta, R. and de Zamaroczy, M. (1980) in The Organization and Expression of the Mitochondrial Genome, Kroon, A.M. and Saccone, C., Eds., pp.21-31, Elsevier/North Holland Biomedical Press, Amsterdam.
31. Conde, J. and Fink, G.R. (1976) Proc. Natl. Acad. Sci. USA 73, 3651-3655.
32. Nagley, P. and Linnane, A.W. (1978) Biochem. Biophys. Res. Commun. 85, 585-592.
33. Hopp, T.P. and Woods, K.R. (1981) Proc. Natl. Acad. Sci. USA 78, 3824-3828.
34. Coruzzi, G. and Tzagoloff, A. (1979) J. Biol. Chem. 254, 9324-9330.
35. Fox, T.D. (1979) Proc. Natl. Acad. Sci. USA 76, 6534-6538.
36. Thalenfeld, B.E. and Tzagoloff, A. (1980) J. Biol. Chem. 255, 6173-6180.
37. Bonitz, S.G., Coruzzi, G., Thalenfeld, B.E., Tzagoloff, A. and Macino, G. (1980) J. Biol. Chem. 255, 11927-11941.

38. Nobrega, F.G. and Tzagoloff, A. (1980) J. Biol. Chem. 255, 9828-9837.
39. Hudspeth, M.E.S., Ainley, W.M., Shumard, D.S., Butow, R.A. and Grossman, L.I. (1982) Cell 30, 617-626.
40. Bonitz, S.G., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F.G., Nobrega, M.P., Thalenfeld, B.E. and Tzagoloff, A. (1980) Proc. Natl. Acad. Sci. USA 77, 3167-3170.
41. Grisi, E., Brown, T.A., Waring, R.B., Scazzocchio, C. and Davies, R.W. (1982) Nucleic Acids Res. 10, 3531-3539.
42. Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457-465.
43. Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. (1982) J. Mol. Biol., 156, 683-717.
44. Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167-180.
45. Montoya, J., Ojala, D. and Attardi, G. (1981) Nature 290, 465-470.
46. Ojala, D., Montoya, J. and Attardi, G. (1981) Nature 290, 470-474.
47. Guerin, M. and Napias, C. (1978) Biochem. 17, 2510-2516.