Molecular cloning and nucleotide sequence of the nuclear *PET122* gene required for expression of the mitochondrial *COX3* gene in *S.cerevisiae*

Jeffrey D.Ohmen, Barbara Kloeckener-Gruissem¹⁺ and Joan E.McEwen*

Department of Microbiology and the Molecular Biology Institute, University of California at Los Angeles, Los Angeles, CA 90024 and ¹Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309, USA

Received April 29, 1988; Revised and Accepted October 24, 1988

ABSTRACT

The nuclear *PET122* gene from *S. cerevisiae* is necessary for translation of a single mitochondrial mRNA that encodes subunit III of cytochrome c oxidase. We report here the cloning and nucleotide sequence of *PET122*, and properties of the predicted protein product, which consists of 242 residues. Analysis of *PET122-lacZ* translational fusions confirms that the *PET122* coding region is translated *in vivo* and indicates that the *PET122* protein product is targeted to mitochondria. A 117 residue domain located in the carboxy-terminal half of the *PET122* protein, at least part of which is shown by characterization of mutants to be critical for *PET122* function, exhibits 24% identity and 59% similarity to a portion of the catalytic domain of *E. coli* alanyl-tRNA synthetase. However, *pet122* mutants are not defective in mitochondrial translation *per se*, as would be expected if *PET122* encoded a tRNA synthetase. Instead, the *PET122* protein may carry out one or more activities in common with tRNA synthetases, such as binding of ATP or RNA.

INTRODUCTION

Genetic analysis of mitochondrial biogenesis in *S. cerevisiae* has been a productive approach for identification of nuclear genes required for expression of mitochondrial genes (1, 2, 3, 4). Two classes of these nuclear genes have been identified. One class encodes products necessary for expression of all mitochondrial protein coding genes and the other class encodes products required for expression of a small subset of mitochondrial protein coding genes (2). So far, all of the genes in the latter class have been shown to act at the post-transcriptional level in expression of target mitochondrial genes (5). For example, a number of nuclear genes have been identified that are required for splicing of precursor transcripts for one or both of the split mitochondrial genes that encode apocytochrome b or

cytochrome c oxidase subunit I (2, 4, 6, 7). Alternatively, a number of nuclear genes that encode proteins necessary for translation of specific mitochondrial mRNAs have been characterized in detail. The CBS1, CBS2, and CBP6 genes are required for translation of the mRNA for apocytochrome b (8, 9, 10, 11); the PET111 gene is required for translation or stability of cytochrome c oxidase subunit II (12), and the PET494 and PET54 genes are required for translation of the mRNA for cytochrome c oxidase subunit III (13, 14, 15).

We recently identified a third nuclear protein coding gene, PET122, that is required for synthesis of cytochrome c oxidase subunit III (16). pet122 mutants produce normal amounts of correctly processed COX3 mRNA (16) and the activity of the PET122 gene product in expression of COX3 is mediated by the 5' untranslated leader of COX3 mRNA (17). Hence, as in the case of the PET494 and PET54 genes, the PET122 gene is required for translation of COX3 mRNA. We report here the cloning and nucleotide sequence of the nuclear PET122 gene. These results represent the first step towards a detailed understanding of the function of the PET122 gene product.

MATERIALS AND METHODS

Strains

Saccharomyces cerevisiae strains used in this study are D273-10B (α ; ATCC # 24657), AB320 (HQ, ade2-1, lys2-1, trp5-2, leu2-1, can1-100, ura3-1 and/or ural-17, met4-1; ref. 18), JM43 (α , leu2-3, leu2-112, ura3-52, trp1-289, his4-580; ref. 3), B122 (α , his4, opl, pet122-1; ref. 16), B120 (α , his4, opl, pet54-101, ref. 16), KG122-2A (α , ade2-1, ura3-52, his4-580, trp1-289, can^r, pet122-1; this study), KG122-12B (α , ade1, ade2-1, leu2-3, leu2-112, ura3-52, pet122-1; this study), GD122 (α , leu2-3, leu2-112, ura3-52, trp1-289, his4-580, pet122(Δ 961-1427)::LEU2; this study) and KBB122 (α , leu2-3, leu2-112, ura3-52, trp1-289, his4-580, pet122(Δ 447-730)::URA3; this study). Escherichia coli strains used in this study are DH5 α (endA1 hsdR17 (rk⁻, m_k⁺) supE44 thi-1 λ^- recA1 gyrA96 relA1 #80dlacZM15; obtained from Bethesda Research Laboratories), 71.18 (Δ lac-proAB thi supE [F' proAB⁺ lacI^q lacZM15], ref. 19), and DR458 (dam3 Δ lacX74 galOP308 rpsL Δ (tonB-trpA) trpR; obtained from Robert Simons, Dept. of Microbiology, UCLA).

Transformation of yeast and *E. coli*

Routine transformations of yeast were performed by the lithium acetate method (20) with 10 μ g supercoiled DNA. One-step gene disruption (21) was performed by the spheroplast transformation method (22) with 50 μ g of linear DNA. *E. coli* strains were made competent and transformed according to standard methods (23). Plasmid DNA was isolated by alkaline SDS lysis (24) and further purified by isopycnic centrifugation in CsCl gradients containing ethidium bromide (23).

Cloning and subcloning of the PET122 locus

Standard methods were used for restriction digestions, agarose gel electrophoresis, end-filling and ligations (23). All reagents were purchased from BRL (Bethesda, MD). The vectors used for construction of plasmids were YEpl3 (18), YCp50 (25), pUC18 (19), M13mp8 (26), and M13mp18 (19).

To clone the PET122 gene, strain KG122-12B (*pet122-1*) was transformed with DNA prepared from the yeast genomic DNA library of Nasmyth and Tatchell (18). This library consists of a partial Sau3A digest of yeast DNA cloned in the multicopy vector YEpl3. After screening 2200 LEU⁺ transformants, one respiration-proficient transformant was identified. The plasmid isolated from this strain was named p12B-4. To isolate the two additional plasmids p122.2a and p122.2b, which have inserts that overlap a portion of the insert in p12B-4, colony hybridization on the library (in *E. coli*) was carried out with a probe derived from the insert of p12B-4. Both p122.2a and p122.2b are able to complement the *pet122* mutation in strain KG122-12B.

Subclones of p12B-4 (see Fig. 1) were generated as follows. Subclone A resulted from digestion of p12B-4 with HindIII and self-ligation of the 13.7 kb fragment, which contains most of YEpl3 and the right hand part of the insert DNA as diagramed in Fig. 1. Subclone B was constructed by first digesting the original isolate p12B-4 with PvuII and self-ligation of the 13.5 kb fragment. Subclone C contains a 2.4 kb ClaI - SalI fragment inserted at the BamHI site of YEpl3. This cloning was accomplished after blunting the ends by end-filling. Subclone D contains the 2.0 kb HindIII - PvuII fragment inserted into the BamHI site of YEpl3. As described for subclone C, ends were blunted by end-filling. Subclone E was constructed by digestion of subclone A with XbaI. The resulting 3.65 kb fragment containing 0.7 kb of YEpl3 DNA and 2.95 kb of insert DNA was ligated into

XbaI digested YEpl3. pJD08 was constructed by cloning a 2.25 kb HindIII - XhoI fragment derived from pl22.2a (see Fig. 1) into the HindIII and SalI sites of YCp50.

Disruption of the PET122 gene

To generate the disrupted strain GD122, plasmid DNA containing the PET122 gene was digested with PvuII, resulting in deletion of PET122 DNA between nucleotides 961 and 1427 (see Fig. 5). The LEU2 gene, on a 2.9 kb BglII fragment, was inserted at the position of the PvuII deletion and the disrupted PET122 locus (pet122(Δ961-1427)::LEU2) was placed in the yeast genome by the one-step gene replacement method (21). The disrupted strain KBB122 was constructed in a similar manner, except an NruI fragment (nucleotides 447-730) was deleted and the URA3 gene, on a 1.1 kb HindIII fragment, was inserted at the position of the deletion. To attain complete digestion of the PET122 gene with NruI, it was necessary to prevent methylation of the site centered at nucleotide 730. This was achieved by propagating plasmid DNA in a Dam⁻ *E. coli* host.

Construction of PET122-lacZ translational fusions

Two PET122-lacZ translational fusions were constructed in the vector pSEY101 (27). pJD014 contains a 1.5 EcoRI-NruI fragment from the PET122 locus cloned into the EcoRI and SmaI sites of pSEY101, while pJD018 contains a 1.9 kb EcoRI-PvuII fragment from the PET122 locus cloned into the EcoRI and SmaI sites of pSEY101. These constructions contain about 1.45 kb of DNA upstream of PET122 and should produce fusion proteins with 15 (pJD014) or 187 (pJD018) amino-terminal residues of the PET122 protein, respectively, fused in frame to β-galactosidase.

Sequencing of PET122

DNA sequencing was performed by the dideoxynucleotide chain termination method (28) with α-³⁵S-thio-dATP (500 Ci/mmol; New England Nuclear) as the label (29). Two regions of sequence were initially unreadable due to the presence of GC compressions. In one instance, the compression was relieved by sequencing in the presence of 7-deazaguanidine (Boehringer-Mannheim; used according to manufacturer's instructions) instead of dGTP, while in the other instance, the compression was relieved by including 25% (v/v) formamide in the sequencing gel.

The DNA templates for sequencing were derived from M13 clones containing part or all of the 1.9 kb ClaI - XbaI PET122 fragment. These clones were generated as follows. The HindIII site of pBKG18 (pUC18 containing the ClaI - XbaI fragment of PET122) was converted to an EcoRI site by

insertion of an EcoRI linker. This resulted in a plasmid containing the 1.9 kb PET122 fragment flanked by EcoRI sites and allowed insertion of the fragment, in either orientation, into the EcoRI site of the sequencing vector M13mp8. Clone mp481 has the XbaI site of the insert adjacent to the sequencing primer site, whereas clone mp483 has the ClaI site at this position. To facilitate sequencing of other internal regions of the 1.9 kb insert, systematic deletions were generated by the method of Lin *et al* (30). Representative deletions were sequenced as indicated in Fig. 6. For some regions of DNA, deletion endpoints were too far apart to obtain continuous sequence. In these cases, custom oligonucleotides were used to prime inside of the 1.9 kb ClaI - XbaI insert of either M13 mp481 or mp483 or specific subclones were constructed in M13 mp18 and sequenced with the universal primer (see Fig. 5).

Northern and Southern analysis

Mitochondria, prepared as described previously (4), were lysed in 4M guanidium thiocyanate, 25 mM sodium acetate, and 140 mM 2-mercaptoethanol and the RNA was purified by ultracentrifugation through a 0.5 ml cushion consisting of 5.7 M CsCl, 100 mM EDTA. Ultracentrifugation was performed at 17° in a Beckman tabletop ultracentrifuge in the TLS55 rotor at 44,000 rpm for 13 hr. Mitochondrial RNA was analyzed by Northern blotting as described previously (4).

Total yeast DNA was isolated from 75 ml cultures (31). Restriction digestion, gel electrophoresis of DNA fragments, Southern transfer to Nylon membrane (Hybond M, Amersham) and hybridization were performed by standard methods (23).

Radioactive probes for hybridization to Northern or Southern blots were prepared either from purified restriction fragments, by the random primer method (32), or from single stranded DNA from the M13mp481 clone (33).

Miscellaneous methods

Mitochondrial translation products were radiolabeled *in vivo* in the presence of cycloheximide as described previously (4). Proteins were separated by electrophoresis in SDS polyacrylamide gels containing 12.5 % acrylamide, 3.5 M urea and 12.5 % glycerol (34) and analyzed by autoradiography. Preparation of whole cell extracts of yeast and assay of β -galactosidase activity was performed as described in ref. 35. Fractionation of yeast cells for localization of the PET122-lacZ fusion protein was carried out as described by Daum *et al.* (36). To assay the protease resistance of mitochondrial-associated β -galactosidase activity, mitochond-

ria (32 μ g protein) were incubated with or without 70 μ g trypsin and with or without 2 % (w/v) Triton X-100 in a final volume of 50 μ l of 0.6 M Mannitol, 2 mM EDTA (pH 8.0) for 30 min. at 22°. Then β -galactosidase activity was measured in the presence of 2 mM PMSF, to inactivate trypsin and inhibit any further digestion of protein. Cytochrome c oxidase activities were determined by the spectrophotometric assay (3). The FASTP program (37) was used to search the NBRF database (Release 14.0, Sept. 30, 1987) for proteins homologous to the PET122 and ORF2 reading frames. The RELATE program (38) was used to assess the statistical significance of the pairing between PET122 and alanyl tRNA synthetase from *E. coli*. RELATE was run with the Mutation Data Matrix, a segment length of 25, and with 100 randomizations of the sequences. To calculate the hydrophobic moment of segments of protein sequence, the MOMENT program from the University of Wisconsin Computer Group was used.

RESULTS

Cloning of DNA that complements a pet122 mutation

We isolated plasmids p12B-4, YEpl3-122.2a and YEpl3-122.2b (Fig. 1), which are able to complement the pet122 mutant KG122-12B (pet122-1). These plasmids are derivatives of the multicopy vector YEpl3 and carry inserts of 7.5, 6.6 and 7.8 kb of yeast genomic DNA, respectively. The restriction maps of these inserts (Fig. 1) indicate that a 3.4 kb region of DNA, between the left junction of YEpl3-122.2a (-1.2 kb on the scale) and the right junction of p12B-4 (+2.2 kb on the scale), is held in common by these three plasmids. This region must therefore contain the gene that complements the pet122-1 mutation.

To further define the region of DNA that complements the pet122-1 mutation, we constructed and tested the complementation activity of several subclones. As indicated in Fig. 1, subclones A, C, E and pJD08 complement the mutation. The area of overlap between the inserts of these subclones is a 1.3 kb region between the ClaI and XhoI sites at 0 and 1.3 kb on the scale. The insert of subclone D, which does not complement the pet122-1 mutation, contains the HindIII - PvuII portion of the insert of subclone E. Together, these data suggest that the gene that complements the pet122-1 mutation is encoded within the 1.3 kb ClaI - XhoI fragment of DNA and that information between the PvuII endpoint

Fig. 1. Cloning and subcloning of DNA that complements *pet122* mutations. The restriction maps of plasmids p12B-4, YEp13-122b, and YEp13-122a, which were derived from the genomic library, and additional plasmids that resulted from subcloning of plasmid p12B-4, are shown. The scale (kb) is indicated on the bottom of the figure; 0 on the scale was assigned to the *Cla*I site of the insert of p12B-4, in order to correspond to the numbering of the DNA sequence (Fig. 6). The restriction sites are abbreviated as R (*Eco*RI), C (*Cla*I), H (*Hind*III), P (*Pvu*II), *X*_b (*Xba*I), and S (*Sal*I). The insert of plasmid p12B-4 carries two pieces of DNA that are not contiguous in the genome of yeast (see text). These pieces are indicated as open or cross-hatched areas, and the cross-hatched areas of the inserts indicate DNA that is linked to the *PET122* locus.

of subclone D and the *Xho*I endpoint of subclone E is essential for complementation.

To assess the colinearity of plasmid and chromosomal restriction maps, particularly in the region that is sufficient for complementation of *pet122*, we carried out Southern hybridization analysis on yeast genomic DNA. *Hind*III, *Eco*RI, *Xba*I, *Cla*I, or *Cla*I plus *Xba*I digests of DNA from strain D273-10B yielded the expected number and sizes of bands when the 1.9 kb *Cla*I-*Xba*I fragment (between 0 and 1.9 kb on the scale, Fig. 1) was used as the probe (data not shown). Hence, these data confirm the colinearity of plasmid and chromosomal restriction maps within the region bounded by the *Eco*RI sites at -1.1 kb and +6.2 kb on the scale in Fig. 1. In contrast, the results of Southern analysis (data not shown) of *Pvu*II-digested genomic DNA suggest that plasmid p12B-4 arose by the insertion of two non-contiguous

pieces of genomic DNA, as indicated in Fig. 1 by the differential shading of the two parts of the insert. It is important to note that the apparent scrambling of DNA in plasmid p12B-4 occurs outside of the region that is sufficient for complementation of pet122 and therefore does not affect our characterization of this DNA.

Evidence that the cloned DNA carries the PET122 gene

Two lines of evidence indicate that we have cloned the PET122 gene. First, we demonstrated complementation of the pet122-1 mutation by plasmid pJD08, which carries a 2.25 kb HindIII - XhoI fragment inserted in the single copy (CEN) vector YCp50 (see Fig. 1). Complementation by a single copy plasmid suggests that the cloned DNA contains the bona fide PET122 gene rather than a "suppressor" capable of bypassing PET122 function when overexpressed from a high copy plasmid. Second, we disrupted the chromosomal copy of the cloned gene in order to demonstrate by genetic complementation tests that the resulting mutant strains possess pet122 mutations. The gene disrupted strain GD122 was constructed so that the 0.5 kb PvuII fragment is deleted and replaced by the LEU2 gene (see Fig. 2). Similarly, strain KBB122 was constructed so that the 0.3 kb NruI fragment was deleted and replaced by the URA3 gene (Fig. 2). In both cases, Southern analysis of genomic DNA (not shown) confirmed that the chromosomal PET122 locus was replaced by the appropriate gene disruption.

We performed complementation tests to determine whether the disrupted genes and the original pet122-1 mutation are in the same complementation group, as would be expected if the cloned gene is PET122. When strains GD122 or KBB122 were crossed to a pet122-1 strain, we found that the diploids could not grow on media containing the non-fermentable carbon sources ethanol and glycerol. Hence, the gene disruptions are not complemented by the pet122-1 mutant, as expected. As a control, we demonstrated that the pet122 gene disruptions are complemented in crosses to a strain with a mutation in the PET54 gene required for expression of COX3 (15, 16).

Phenotype of the pet122 gene disruptions

We reported previously that the pet122-1 mutant B122 possesses apparently normal levels of properly processed COX3

Fig. 2. **Gene Disruptions.** The structures of the gene disruptions in strains GD122 and KBB122 are schematically illustrated. The LEU2 and URA3 inserts are not drawn to scale. The arrows within the inserts indicate the direction of transcription of the inserted gene. The restriction sites are abbreviated as C (ClaI), N (NruI), P (PvuII), X (XhoI), and Xb (XbaI).

mRNA (16). This strain bears a COX3 allele that produces a 3.6 kb COX3 mRNA. Another wild type COX3 allele that produces a 1.6 kb COX3 mRNA has been identified in some strains of yeast (39, 40). It is of interest to find whether a pet122 mutation has the same effect on expression of the "1.6 kb" allele as we observed previously for the "3.6 kb" allele. For that purpose, we constructed the strains GD122 and KBB122 so that each carries the gene disruption as well as the 1.6 kb COX3 allele. As shown in Fig. 3 (mitochondrial translation products), strains GD122 (lane 3) and KBB122 (lane 4) are unable to accumulate cytochrome c oxidase subunit III, as we observed previously for a pet122-1 strain (ref. 16, and Fig. 3, lane 2). To determine if the 1.6 kb COX3 mRNA is produced and stable in a pet122 mutant, we compared the steady state levels of COX3 mRNA in the wild type (JM43) and one of the mutant (GD122) strains. As shown in Fig. 4 (Northern analysis of COX3 transcripts), strain GD122 (lane 2), in comparison to the parental strain JM43 (lane 1), exhibits an apparently normal level of the 1.6 kb COX3 mRNA. As an internal control for equal loading of RNA in lanes 1 and 2 of the gel in Fig. 4, we also examined the levels of COX2 mRNA in a duplicate blot of these RNA samples, and found them to be equivalent (data not shown). Hence, the PET122 gene product is required for translation of both the 3.6 kb and 1.6 kb forms of COX3 mRNA.

Nucleotide sequence of PET122

The 1.9 kb ClaI - XbaI region (between 0 and 1.9 kb on the scale in Fig. 1) was sequenced according to the strategy shown in

Fig. 3. Mitochondrial translation products. Strains JM43 (lane 1), KGI22-2A (lane 2), GDI22 (lane 3) and KBB122 (lane 4) were labeled with ^{35}S -Methionine in the presence of cycloheximide (4). The specific activities of labeling of mitochondrial proteins are similar in all four strains. Mitochondrial proteins were extracted and analyzed by SDS gel electrophoresis and autoradiography. The mitochondrial-encoded polypeptides are labeled as VarI, COI (cytochrome c oxidase subunit I), Cytb (apocytochrome b), COII (cytochrome c oxidase subunit II), COIII (cytochrome c oxidase subunit III), ATPase6 (subunit 6 of ATPase) and ATPase9 (subunit 9 of ATPase).

Fig. 5, and the nucleotide sequence is presented in Fig. 6. We then searched the nucleotide sequence for potential protein coding regions. As shown in Fig. 5 and Fig. 6, the two longest potential protein coding open reading frames (ORF1 and ORF2) are encoded on the same strand of DNA. ORF1 is predicted to encode a protein of 242 amino acids. ORF2 begins 90 nucleotides downstream of ORF1 and is truncated by the XbaI site at the end of the fragment. Hence, ORF2 is predicted to encode a protein of at least 231 amino acids. Northern blot analyses with probes specific for ORF1 or ORF2 (unpublished data) indicate that both ORFs are transcribed, which suggests that both are translated into proteins. These ORFs do not appear to be exons of a single gene because the consensus splicing signals GTATGT and TACTAAC (41) are not present in the nucleotide sequence. Two experiments described above demonstrate that ORF1 is PET122. First, plasmid pJD08, which contains all of ORF1 and only 14 amino acids of ORF2, is able to complement pet122-1 in strain KGI22-12B.

Fig. 4. Northern blot of COX3 transcripts. Mitochondrial RNAs from strains JM43 (lane 1), and GD122 (lane 2), with VC37 (lane 3; a derivative of strain JM22) included as an example of a strain that produces a 3.6 kb COX3 mRNA, were hybridized to a radiolabeled COX3 probe. The sizes (kilobases) of the radioactive bands were deduced from a standard curve drawn from the migrations of RNA size markers (RNA ladder from BRL).

Second, the gene disruption in strain KBB122 (Fig. 2) disrupts only ORF1 and is pet122⁻ by both genetic and phenotypic criteria.

Expression and localization of the PET122 gene product. Two considerations of the nucleotide sequence suggest that the PET122 gene is not highly expressed. First, the coding region of PET122 does not exhibit a strong codon bias, as has been observed in highly expressed yeast genes (42, 43). Second, the sequence context of the initiator AUG codon in PET122 (5'-CGUCAAUGUUG-3') does not closely resemble the yeast consensus context 5'-(A/Y)A(A/U)AAUGUCU-3' (44), particularly in the -3 position, for which an A has been shown by mutational analysis to contribute to efficiency of translation of other yeast mRNAs (45, 46). Expression of PET122 *in vivo* does indeed appear to be low in view of the relatively low β -galactosidase activities expressed *in vivo* from the PET122-lacZ translational fusion plasmids pJD014 (4.0 ± 0.4 units of β -galactosidase, average of 3 experiments) and pJD018 (11.9 ± 1.7 units, average of 3 experiments).

Translation of the PET122 protein is predicted to initiate at the first ATG of the open reading frame (Fig. 6), and the expression of β -galactosidase from pJD014 (see above), which bears the translational fusion with only the first 15 residues of PET122, suggests that this is the case. However, depending on

Fig. 5. Sequencing strategy. Both strands of the 1909 base pairs ClaI - XbaI fragment were sequenced. The arrows that begin with solid squares indicate sequence obtained from priming at the universal primer site of either the full length mp481 and mp483 clones or deletion derivatives of these clones. The arrows that begin with open circles indicate sequence obtained from priming with synthetic oligonucleotides specific for interior segments of the inserts of mp481 and mp483. The filled bars at the bottom of the figure indicate open reading frames.

the locations of the 5' ends of PET122 mRNA (not determined), it is possible that an ATG codon located downstream, for example at codon number 26, is utilized in addition to the first one. To demonstrate that expression of a functional PET122 gene product is dependent upon translation from the beginning of the open reading frame, we constructed a mutation in PET122 that should lead to termination of translation after codon number 10, which is located upstream of the next in frame ATG, at codon number 26. Briefly, this mutation was generated in plasmid pJD08 by filling in and reclosing the BstEII restriction site at nucleotide 426 (Fig. 6). This mutation is predicted to change the nucleotide sequence from G GTG ACC to G GTG ACG TGA CC, resulting in generation of a TGA translation termination codon within the PET122 reading frame. Upon transformation of a pet122 mutant with the mutated plasmid, we observed no complementation of the respiration-deficient phenotype, in contrast to the result obtained with the wild-type pJD08 plasmid (see Fig. 1). Hence, the function of PET122 in mitochondrial respiration depends upon translation of the protein from the beginning of the reading frame.

To examine the intracellular location of the PET122 protein, we examined the distribution of β -galactosidase within subcel-

```

1  CGATGGAGGCACTGGGTTGGATTCCGAGACATCGTTGGCATTGGGCCGTCGAATTAAATCTTTGGCCTGAAAGAGATCCATGAOSGATGGG
101  CCGGGGCAATACTATGTTCCGAGCGGTGCCAGTCTGGAAGAGCGAGCAAACTTGACGTGACGAGTGGAGAGTGGAATTTGAACATGTCGGGGAGGTT
201  ATTCTGTGGCTCCGCTTGTACGTGAACAGATACGTATAGAGGGGAGGCCACTGGTTAAATTTTCATGGCTCGGATTACTTCGTACTGCTGGCTAAAT
301  CGAAATCTGGGCTGCTGAGAGTGTTTGGAGCAATCAGGGAAACATCTGAACGTGGAAGAGCAGAGGCGCATAGCTCGAACATAAAGACGGACAGTC
402  ATG TTG ACT ATC ACG AAA AGA CTG GTG ACC ACC GAT GTG CGG TCG CGA ATA CTG TTA AGC AGT TTA AAC GGG AAA
    met leu thr ile thr lys arg leu val thr thr asp val arg ser arg ile leu leu ser ser leu asn gly lys
477  ATG TCC GAT GAA CTG CGG CTG CTG CGT CAG CAG CAG CAG ACC AGC GTG GAT GTG GAG CTG CTG CAC ACG ATG CTA
    met ser asp ala leu ala leu leu arg gln gln gln gln thr ser val asp val glu leu leu his thr met leu
552  GCG CGA GGC GCT GCG CTT GGC CAT GGC GAC ACT ATA GCA TAC ATG TGG TAT CAG CAT GTG ATG CCA CGC CGG TTG
    ala arg ala ala ala leu ala his ala asp thr ile ala tyr met trp tyr gln his val met pro arg arg leu
627  CCA GTA GAG GGC GCG CTG CTA TGT GAA ATG GCT GGC GTA GCA TTG TAC CAG GAC AGG CTC TTC TTA CCC GCG CAG
    pro val glu gly arg leu leu cys glu met ala gly val ala leu tyr gln asp arg leu phe leu pro ala gln
702  TTC CTC CAG CAC TAC CAG GCG ATG AAT CGC GAT CGT CGC ACC AGC CCA GAA GAT GAA CTG ATT GAG TAT GAG CTT
    phe leu gln his tyr gln ala met asn arg asp arg thr ser pro glu asp glu leu ile glu tyr glu leu
777  AGA CGG ATT AAA GCG GCG TTT GCG CGT GGC ACA ATG CAC TCC ACG GCG CTC AGG GAA AGG TGG AGG GTA TTC
    arg arg ile lys val glu ala phe ala arg gly thr met his ser thr ala leu arg glu lys trp lys val phe
852  TTG CAG GAG ATG GAT ACG CTA CCA GCG CAG CGG CCA TTA AGG CTG CGC GAC TTC CGG CAA ATG ACC AAG GCT ATG
    leu gln glu met asp thr leu pro gly gln pro pro leu arg leu arg asp phe pro gln met thr lys ala met
927  GGC ATA GCA TTG ATG CAG CAA GAT GAG CAA GCA GCT GGC CTG GCG TTG TTT GGA CGA CAG CAC CCA CTG AAG AAA AGA
    gly ile ala leu met gln gln asp glu gln ala ala leu ala leu phe gly arg gln pro leu val lys arg
1002  ACG AAT GGT CAC TAC CGC TAC TAC TGG CTG GTG TCC TTT GGC ATG TTC CGG GCG CAG CGC AGG CGC GAC GTG TGC
    thr asn gly his tyr arg tyr tyr trp leu val ser phe gly met phe pro ala arg thr arg arg asp val cys
1077  TGG CGG AGT TCC GTC AAA GTT ATC GCG GGC TGC CGC TGC TGG ATG CGC AAC TAG TGA TAA
    trp arg ser ser val lys val ile ala gly cys arg cys trp met pro asn * * *
1137  AGAGAAAGAGGATTGAAATCAACACATAAATCTGGGTGGAGCATCGCTGTAAACAGGAACAACGGGTCTAGCAAGCGGTA
1218  ATG AAA TAC AAG GAA ATC AAT TTC TTC AAG GGC CAT CGG AGC TCG AGG TTG CTG CCT CGA GAA GCA GTA ATT CAA
    met lys tyr lys gln leu ile asn phe phe lys gly his pro ser ser arg leu pro arg glu ala val ile gln
1293  GCG ACT GCG GCT ATA TTG GGG CCC GAG ACC AGG GAG TAC GAT AAC GAC CCC TAT AAC AGG CAT CGG CTG ACG TAC
    ala thr ala ala ala leu gly pro glu thr arg glu tyr asp asn asp pro tyr asn arg his pro leu thr tyr
1368  GGT TCG GAC GAA GGT GGC CTG TGG GTG CGA GAG CAG ATT TGT ACG TTT CTG AAT GAT CAG CTG TTT AAG TTC GAA
    gly ser asp glu gly ala leu trp val arg glu gln ile cys thr phe leu asn asp gln leu phe lys phe glu
1443  AAT GGG GCT CGG ACG AGG ACA CGG GCA GAC TAT TTG AAT CTG AAT AGC GGC GCT TCG TAT GGC ATG CTG AAC ATC
    asn gly ala arg ser arg thr arg ala asp tyr leu asn leu asn ser gly ala ser tyr gly met leu asn ile
1518  CTT CTG CAA ACA ACC TTG CCA CAT AAC GGG TAT ACC AGG CAG GCG TTC ATC ATC ACG CCA ACA TAT TTC TTG ATC
    leu leu gln thr thr leu pro his asn gly tyr thr arg gln ala phe ile thr pro thr tyr phe leu ile
1593  AAC AAT TGC TTC ACA GAT GCG GGA TTC AAG GGG AAA ATG ACC GGC ATC AAC GAG CAG GGC CAC GAC TCG ATT GAT
    asn asn cys phe thr asp ala gly phe lys gly lys met thr ala ile asn glu gln gly his asp ser ile asp
1668  TTC GAG TCG TTG ATT TCT GGC CTT GAG CAG CAC GAG GCG GAG CGG CAG CCC CAT AGT ACC ACA GAG ATG ATT CAG
    phe glu ser leu ile ser ala leu glu gln his glu ala glu pro gln pro his ser thr thr glu met ile gln
1743  GGG CCA AAG TTG ACC AAG AAG GTC TAC AGG TAC GTT ATG TAC TGC ATC CGG ACG TTT GCA ACA CCA TCG GGA AAC
    gly pro lys leu thr lys lys val tyr arg tyr val met tyr cys ile pro thr phe ala asn pro ser gly asn
1818  ACA TAC TCG CTT GAG ACC AGA CGC AGA CTT ATC GAC ATC GCT CGG AAG TAC GAC ATG CTG ATA ATC ACT GAT GAC
    thr tyr ser leu glu thr arg arg leu ile asp ile ala arg lys tyr asp met leu ile ile thr asp asp
1893  GTG TAC GAT ATT CTA GA
    val tyr asp ile leu

```

Fig. 6. Nucleotide sequence of the 1.9 kb *Cla*I - *Xba*I fragment. Only the sequence of the coding strand is shown. The predicted amino acid sequences of **PET122** and ORF2 are shown below the nucleotide sequence.

lular fractions of cells that are expressing the **PET122-lacZ** translational fusion containing the 187 amino-terminal residues of **PET122** (pJD018). As shown in Table I, part A, a substantial fraction of β -galactosidase activity was associated with

Table I. The amino-terminal 187 amino acids of PET122 direct β -galactosidase into mitochondria**A. Distribution of β -galactosidase activity in cell fractions**

Cell Fraction	cytochrome oxidase activity (%)	β -galactosidase activity (%)
Mitochondria	96	53
Post-mitochondrial supernatant	4	47

B. Resistance of mitochondrial-associated β -galactosidase activity to protease digestion.

Pretreatment conditions	β -Galactosidase Activity (%)
none	100
Triton X-100	82.8
Trypsin	92.2
Triton X-100 and Trypsin	10.7

mitochondria. The post-mitochondrial supernatant fraction from this experiment was only minimally contaminated by mitochondria as indicated by the distribution of the mitochondrial marker, cytochrome *c* oxidase. We conducted two additional experiments to show that the mitochondrial-associated β -galactosidase activity in these cells is genuine rather than adventitious. First, we found that the β -galactosidase activity comigrates with cytochrome *c* oxidase activity during sucrose gradient centrifugation (data not shown). Second, we showed that the mitochondrial β -galactosidase activity is protected against protease digestion (Table I, Part B). The fact that β -galactosidase levels are resistant to protease digestion unless the detergent Triton-X-100 is present to solubilize mitochondrial membranes indicates that the fusion protein is sequestered within mitochondria, rather than bound to the mitochondrial outer surface. Since β -galactosidase itself exhibits no inherent ability to associate with mitochondria (ref. 27; and our unpublished results with the

shorter PET122-lacZ fusion in pJD014), these results suggest that the amino-terminal 187 residues of PET122 in the fusion protein possess mitochondrial targeting activity, and that the PET122 protein itself is localized within mitochondria.

DISCUSSION

We cloned the nuclear PET122 gene from S. cerevisiae by selecting for recombinant plasmids, derived from a yeast genomic library, that complement a chromosomal pet122 mutation. We then proved, by showing complementation in single copy and by gene disruption, that the cloned gene is PET122 rather than a suppressor of pet122 mutations. The nucleotide sequence of a 1.9 kb region of DNA revealed the presence of two long open reading frames (ORFs), which are encoded by the same strand. We conclude from subcloning experiments and characterization of the gene disruption strain KBB122 that ORF1, which begins at nucleotide 402 and ends at nucleotide 1127, represents the PET122 coding region. A nonsense mutation generated at a BstEII site near the beginning of the PET122 reading frame blocks expression of a functional PET122 protein, which demonstrates that translation initiates at the first ATG of the reading frame, as predicted.

In the course of generating the pet122 gene disruptions in strains GD122 and KBB122, we also demonstrated that the PET122 gene product is necessary for translation of both the 3.6 kb and the 1.6 kb forms of COX3 mRNA. The 3.6 kb and the 1.6 kb mRNAs differ at both 5' and 3' ends (40), but they evidently retain a common structural element that mediates the effect of the PET122 gene product. This element is presumably located in the 5' untranslated leader of the 1.6 kb mRNA, since it is this region of the 3.6 kb mRNA that has been shown to be important for translational activation by the PET494, PET54, and PET122 gene products (13, 14, 17). Comparison of the nucleotide sequence of the 5' upstream DNA from the "3.6 kb" COX3 allele (47) with that from the "1.6 kb" allele (not yet determined) should provide information about the element that mediates the effect of the PET122 gene product.

Since COX3 mRNA translation, which occurs within mitochondria, is dependent upon the function of the PET122 protein, it is

pertinent to ask where the PET122 protein is localized within the cell. A cytoplasmic location would suggest an indirect function for PET122 in COX3 mRNA translation, while a mitochondrial location would be consistent with a direct interaction between the PET122 protein and either COX3 mRNA or the mitochondrial translation apparatus. The results of analysing the intracellular location of a PET122-lacZ fusion protein (Table I) support the contention that the PET122 protein itself is transported into mitochondria and performs a direct role in promoting translation of COX3 mRNA.

The PET122 protein, as predicted from the nucleotide sequence (Fig. 6), is probably soluble rather than membrane bound. It contains 22 acidic residues and 34 basic (arginine and lysine) residues, and the longest stretch of uncharged amino acids is 13, occurring near the carboxyl terminus of the protein. Furthermore, no highly hydrophobic domains (value greater than 2) can be identified by hydropathy analysis (48). Since most mitochondrial proteins are targeted to mitochondria via amino terminal leader peptides that exhibit a characteristic amino acid composition (rich in ser, thr, lys and arg and usually devoid of acidic residues) (49) and are capable of forming an amphiphilic structure (50, 51), we inspected the amino terminal residues of the predicted PET122 protein for these features. Within the first 25 residues of the protein, there are 7 thr or ser residues and 5 lys or arg residues. Within this region, there is also one acidic residue (asp), which is not typical for a targeting signal but is also not incompatible with targeting activity (49). A portion of this amino terminal sequence is also predicted to be amphiphilic when in a β -sheet conformation (residues 7-16 have an average per residue hydrophobic moment of 0.55). Hence, these characteristics support the hypothesis that some portion of the first 25 residues of the PET122 protein is critical for targeting the protein to mitochondria.

The predicted amino acid sequences of both PET122 and ORF2 were compared with 4721 protein sequences in the NBRF data bank. No homologous matches over the entire lengths of the proteins were identified in either case. However, the C-terminal half of the PET122 protein can be aligned with a region of the E. coli

Fig. 7. Alignment of similar domains of yeast **PET122** and *E. coli* alanyl tRNA synthetase. The amino acid sequences were aligned by the FASTP program. Identities are indicated by | and conservative replacements by *. Conservative replacements are defined as any pair with a score of 8 or better, according to the Mutation Data Matrix of Dayhoff, as modified by Doolittle (ref. 52).

alanyl tRNA synthetase. After alignment and introduction of two gaps, the appropriate regions of the proteins exhibit 24% identity, or 59% similarity when both identical residues and conservative substitutions are included (Fig. 7). (A conservative replacement in this case is defined as any pair with a score of 8 or greater according to the mutation data matrix of Dayhoff, as modified by Doolittle (52)). The Z value (segment comparison score in standard deviation units) for the similarity of **PET122** and *E. coli* alanyl tRNA synthetase was computed with the RELATE program to be 3.228. This value indicates that the probability of obtaining this degree of similarity between two unrelated proteins is about 0.9×10^{-3} (38).

The similarity between the carboxyl-terminal half of the **PET122** protein and a region of *E. coli* alanyl tRNA synthetase suggests that the respective domains perform similar functions. Through analysis of mutants with carboxyl-terminal deletions of *E. coli* alanyl tRNA synthetase, Schimmel and coworkers have delineated the portions of the protein essential for each partial activity of the synthetase. They showed that a region of the

protein between amino acid residues 368 and 385 is essential for binding of tRNA^{ala}, while a region of the protein between amino acid residues 257 and 368 is essential for alanyl-adenylate formation (53, 54). Hence, the carboxyl-terminal half of the PET122 protein (residues 116 to 232, see Fig. 7) is similar to a region of the E. coli alanyl tRNA synthetase (residues 300 to 419, see Fig. 7) that is directly implicated in all of the catalytic activities necessary for aminoacylation of tRNA. Moreover, we know from the gene disruption that we generated in strain GD122 that a portion of the carboxyl terminal half of the PET122 protein between residues 185 and 242 (the end) is essential for PET122 function. However, since pet122 mutants do not exhibit a generalized defect in either cytoplasmic or mitochondrial protein synthesis, we do not believe that the PET122 protein is a tRNA synthetase. Instead, we speculate that the PET122 protein carries out a function similar to one of the partial activities of alanyl tRNA synthetase, such as binding of ATP or RNA. It is interesting to note that the PET122 protein, which resembles a tRNA synthetase, is necessary for translation of a specific mitochondrial mRNA, while at least three bona-fide mitochondrial tRNA synthetases (for threonine, tyrosine and leucine) have recently been implicated in splicing of specific mitochondrial mRNAs (55, 56, 57). Hence, recruitment of activities associated with tRNA synthetases for other types of RNA transactions appears to be prevalent in the expression of mitochondrial genes in yeast.

ACKNOWLEDGEMENTS

We are especially grateful to Tom Fox (Cornell Univ.) for helpful comments on our data and to Donald Nierlich (UCLA) and Robert Simons (UCLA) for helpful comments on the manuscript. We also acknowledge the expert technical assistance of Karen Beer. This work was supported by the Public Health Service (grant GM36192 from the National Institutes of Health) and the American Cancer Society (grant MV-258). JDO is a predoctoral trainee of the Public Health Service, grant GM07185.

*To whom correspondence should be addressed

†Present address: Department of Genetics, University of California, Berkeley, CA 94720, USA

REFERENCES

1. Michaelis, G., Mannhaupt, G., Pratje, E., Fisher, E., Naggert, J., Schweizer, E. (1983) In Slonimski, P., Borst, P., and Attardi, G.

- (eds), Mitochondrial Genes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 311-321.
2. Tzagoloff, A., Myers, A.M. (1986) *Ann. Rev. Biochem.* 55, 249-285.
  3. McEwen, J.E., Ko, C., Kloeckener-Gruissem, B., Poyton, R.O. (1986) *J. Biol. Chem.* 261, 11872-11879.
  4. Kloeckener-Gruissem, B., McEwen, J.E., Poyton, R.O. (1987) *Curr. Genet.* 12, 311-322.
  5. Fox, T.D. (1986) *TIG* 2, 97-100.
  6. Pillar, T., Lang, B.F., Steinberger, I., Vogt, B., Kaudewitz, F. (1983) *J. Biol. Chem.* 258, 7954-7959.
  7. Simon, M., Faye, G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 8-12.
  8. Rödel, G., Kaudewitz, F. (1985) *Curr. Genet.* 9, 641-648.
  9. Rödel, G. (1986) *Curr. Genet.* 11, 41-45.
  10. Rödel, G., Michaelis, U., Forsbach, V., Kreike, J., Kaudewitz, F. (1986) *Curr. Genet.* 11, 47-53.
  11. Dieckmann, C.L., Tzagoloff, A. (1985) *J. Biol. Chem.* 260, 1513-1520.
  12. Poutre, C.G., Fox, T.D. (1987) *Genetics* 115, 637-647.
  13. Müller, P.P., Reif, M.K., Zonghou, S., Sengstag, C., Mason, T., Fox, T.D. (1984) *J. Mol. Biol.* 175, 431-452.
  14. Costanzo, M.C., Fox, T.D. (1986) *Mol. Cell. Biol.* 6, 3694-3703.
  15. Costanzo, M.C., Seaver, E.C., Fox, T.D. (1986) *EMBO J.* 5, 3637-3641.
  16. Kloeckener-Gruissem, B., McEwen, J.E., Poyton, R.O. (1988) *J. Bacteriol.* 170, 1399-1402.
  17. Costanzo, M.C., Fox, T.D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 2677-2681.
  18. Nasmyth, K.A., Tatchell, K. (1980) *Cell* 19, 753-764.
  19. Yanisch-Perron, C., Vieira, J., Messing, J. (1985) *Gene* 33, 103-119.
  20. Ito, H., Fukuda, Y., Murata, K., Kimura, A. (1983) *J. Bacteriol.* 153, 163-168.
  21. Rothstein, R.J. (1983) *Methods in Enzymology* 101, 202-209.
  22. Beggs, J.D. (1978) *Nature* 275, 104-109.
  23. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
  24. Birnboim, H.C., Doly, J. (1979) *Nucleic Acids Res.* 7, 1513-1523.
  25. Kuo, C., Campbell, J.L. (1983) *Mol. Cell. Biol.* 3, 1730-1737.
  26. Messing, J., Vieira, J. (1982) *Gene* 19, 269-276.
  27. Douglas, M.G., Geller, B.L., Emr, S.D. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 3983-3987.
  28. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., Roe, B.A. (1980) *J. Mol. Biol.* 143, 161-178.
  29. Biggin, M.D., Gibson, T.J., Hong, G.F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 3963-3965.
  30. Lin, H., Lei, S., Wilcox, G. (1985) *Anal. Biochem.* 147, 114-119.
  31. Nasmyth, K.A., Reed, S.I. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77, 2119-2123.
  32. Feinberg, A.P., Vogelstein, B. (1983) *Anal. Biochem.* 132, 6-13.
  33. Hu, N., Messing, J. (1982) *Gene* 17, 271-277.
  34. Merle, P., Kadenbach, B. (1980) *Eur. J. Bioch.* 105, 499-507.
  35. Rose, M., Botstein, D. (1983) *Methods in Enzymology* 101, 167-180.
  36. Daum, G., Böhm, P.C., Schatz, G. (1982) *J. Biol. Chem.* 257, 13028-13033.
  37. Lipman, D.J., Pearson, W.R. (1985) *Science* 227, 1435-1441.
  38. Dayhoff, M.O., . (1978) *Atlas of Protein Sequence and Structure* vol. 5 (supp. 3), 1-8.
  39. Thalenfeld, B.E., Hill, J., Tzagoloff, A. (1983) *J. Biol. Chem.* 258, 610-615.
  40. Fox, T.D., Reif, M.K. (1983) In Schweyen, R.J., Wolf, K. and Kaudewitz,

- F. (eds), *Mitochondria 1983 - Nuclear-Cytoplasmic Interactions*, Walter de Gruyter & Co., Berlin, pp. 441-448.
41. Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M.C., Mager, W.H., Planta, R.J., Schulz, L. (1984) *Nucl. Acids Res.* 12, 8295-8312.
  42. Bennetzen, J.L., Hall, B.D. (1982) *J. Biol. Chem.* 257, 3026-3031.
  43. Sharp, P.M., Tuohy, T.M.F., Mosurski, K.R. (1986) *Nucl. Acids Res.* 14, 5125-5143.
  44. Cigan, A.M., Donahue, T.F. (1987) *Gene* 59, 1-18.
  45. Werner, M., Feller, A., Messenguy, F., Pierard, A. (1987) *Cell* 49, 805-813.
  46. Zitomer, R.S., Wathall, D.A., Rymond, B.C., Hollenberg, C.P. (1984) *Mol. Cell. Biol.* 4, 1191-1197.
  47. Thalenfeld, B.E., Tzagoloff, A. (1980) *J. Biol. Chem.* 255, 6173-6180.
  48. Kyte, J., Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105-132.
  49. Baker, A., Schatz, G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 3117-3121.
  50. Roise, D., Theiler, F., Horvath, S.J., Tomich, J.M., Richards, J.H., Allison, D.S., Schatz, G. (1988) *EMBO J.* 7, 649-653.
  51. von Heijne, G. (1986) *EMBO J.* 5, 1335-1342.
  52. Doolittle, R.F. (1986), *Of URFS and ORFS. A Primer on how to analyze derived amino acid sequences*, University Science Books, Mill Valley, CA.
  53. Regan, L., Bowie, J., Schimmel, P. (1987) *Science* 235, 1651-1653.
  54. Schimmel, P. (1987) *Ann. Rev. Biochem.* 56, 125-158.
  55. Pape, L.K., Koerner, T.J., Tzagoloff, A. (1985) *J. Biol. Chem.* 260, 15362-15370.
  56. Akins, R.A., Lambowitz, A.M. (1987) *Cell* 50, 331-345.
  57. Herbert, C.J., Labouesse, M., Dujardin, G., Slonimski, P.P. (1988) *EMBO J.* 7, 473-483.