The EMBO Journal vol.3 no.12 pp.2831 – 2837, 1984

Subunit IV of yeast cytochrome *c* oxidase: cloning and nucleotide sequencing of the gene and partial amino acid sequencing of the mature protein

Ammy C. Maarse[1], Adolphus P.G.M. Van Loon[1,3], Howard Riezman[4], Ivan Gregor[2], Gottfried Schatz[2] and Leslie A. Grivell[1]

[1]Section for Molecular Biology, Laboratory of Biochemistry, Kruislaan 318, University of Amsterdam, NL-1098 SM Amsterdam, The Netherlands and [2]Biocenter, University of Basel, CH-4056 Basel, Switzerland

Present addresses: [3]Biocenter, University of Basel, CH-4056 Basel, and [4]Swiss Institute for Experimental Cancer Research, CH-1066 Epalinges/ Lausanne, Switzerland
Communicated by L.A. Grivell

The six small subunits (IV – VII, VIIa, VIII) of yeast cytochrome *c* oxidase are encoded by nuclear genes and imported into the mitochondria. We have isolated the gene for subunit IV from a yeast genomic clone bank and determined its complete nucleotide sequence. We have also isolated subunit IV from purified yeast cytochrome *c* oxidase and determined most of its amino acid sequence which confirms the positioning of ~90% of the amino acid residues. The sequence comparison shows that the coding sequence of the gene lacks introns and that subunit IV is made as a precursor with an amino-terminal extension of 25 residues, five of which are basic and none of them acidic. Precursor processing involves cleavage of a Leu-Gln bond.
Key words: cytochrome *c* oxidase/nucleotide sequence/ amino acid sequence

Introduction

Cytochrome *c* oxidase of yeast consists of at least nine different subunits. The three largest of these (I – III) are mitochondrially encoded, while the remaining six (IV – VII, VIIa, VIII) are encoded in the nucleus (Schatz and Mason, 1974; Power *et al.*, 1984a). The dual origin of this enzyme complex poses important questions concerning the coordinate synthesis, import and assembly of the individual subunits. Attempts to answer some of these questions at the molecular level are likely to be facilitated by the isolation and characterization of the genes encoding the individual subunits. This has been achieved for subunits I – III, specified by mitochondrial DNA (Tzagoloff, 1982) and, partly, for the nuclearly-encoded subunits V and VI (Cumsky *et al.*, 1983; Wright *et al.*, 1984). Here we describe the isolation and the complete nucleotide sequence of the nuclear gene for subunit IV and an almost complete amino acid sequence of the mature protein. Since subunit IV is made as a larger precursor (Mihara and Blobel, 1980; Lewin *et al.*, 1980), a comparison between these two sequences provides information on the exact size and primary structure of the transient pre-sequence.

Results

Isolation of the gene encoding subunit IV of cytochrome c oxidase

The synthesis and steady-state levels of many yeast mitochondrial proteins are repressed by glucose. This property has already allowed us to clone the genes for several of these by competitive RNA-DNA hybridization (Van Loon *et al.*, 1982, 1983a; Riezman *et al.*, 1983a). The procedure described in these previous studies was used in an attempt to isolate the subunit IV gene from a clone bank consisting of *Escherichia coli* cells carrying random fragments of yeast genomic DNA in the yeast-*E. coli* 'shuttle vector' pFL-1. One recombinant plasmid (plasmid 19) hybrid-selected yeast mRNA directing the *in vitro* synthesis of a protein which is immunoprecipitated by an antibody against yeast cytochrome *c* ox-

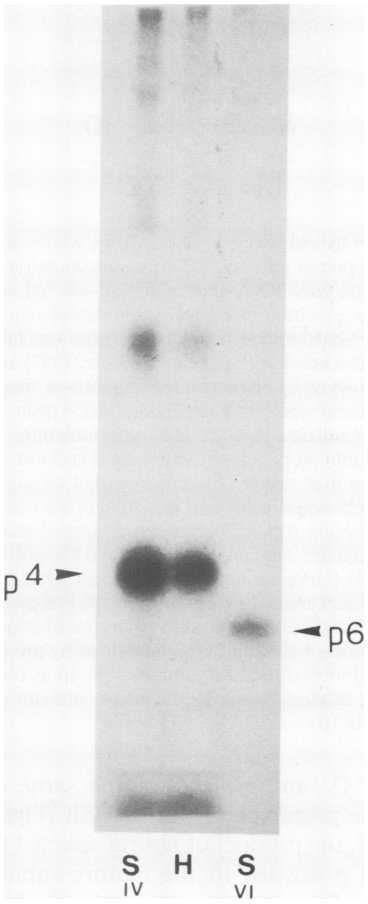

Fig. 1. The mRNA hybrid-selected by plasmid 19 codes for a polypeptide recognized by antiserum against subunit IV of yeast cytochrome *c* oxidase. Plasmid 19 was incubated with total cellular RNA from yeast; the hybrid-selected mRNA was eluted and used to program protein synthesis in a nuclease-pretreated reticulocyte lysate in the presence of [^{35}S]methionine. The labeled lysate was subjected to immunoprecipitation with a mixture of two rabbit antisera directed respectively against yeast cytochrome *c* oxidase subunits IV or VI. The labeled immunoprecipitate was analyzed by SDS-14% polyacrylamide gel electrophoresis and fluorography (**lane H**). Another aliquot of the nuclease-treated lysate was treated in a similar way except that it was programmed with total yeast RNA and then subjected to immunoprecipitation with either antiserum against cytochrome *c* oxidase subunit IV (S IV) or against cytochrome *c* oxidase subunit VI (S VI). p4 and p6 denote positions of the larger precursors to subunit IV and VI, respectively (Lewin *et al.*, 1980; Mihara and Blobel, 1980).

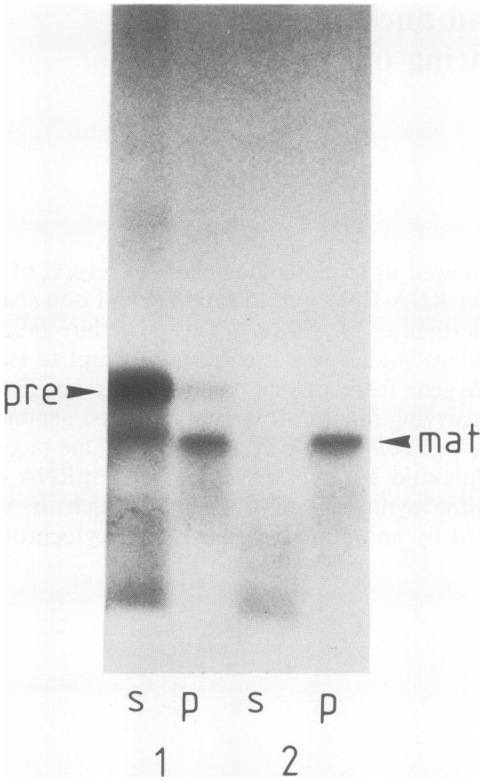

Fig. 2. Plasmid 19 hybrid-selects a yeast mRNA whose *in vitro* translation product can be imported into isolated yeast mitochondria. Plasmid 19 was incubated with total yeast RNA and the hybrid-selected mRNA was used to program protein synthesis in a nuclease-pretreated reticulocyte lysate in the presence of [^{35}S]methionine. The labeled lysate was incubated with energized yeast mitochondria (Gasser and Schatz, 1982) and the mixture was divided into two equal aliquots. One aliquot was treated with trypsin to digest accessible radiolabeled polypeptides and trypsin was then inhibited with phenylmethyl sulfonyl fluoride and trypsin inhibitor. The other aliquot received the inhibitors before the trypsin (Gasser *et al.*, 1982). Mitochondria were then pelleted by centrifugation through a sucrose cushion and mitochondrial pellet and supernatant were separately subjected to immunoprecipitation with rabbit antiserum against yeast subunit IV. The immunoprecipitates were analyzed by SDS-14% polyacrylamide gel electrophoresis and fluorography. (**1**) Supernatant (s) and mitochondrial pellet (p) from aliquot treated with trypsin in the presence of inhibitor; (**2**) supernatant (s) and pellet (p) from trypsin-treated aliquot. The precursor and mature form of subunit IV are identified by **pre** and **mat**. The slightly faster migration of mature subunit IV in 1p as compared with 1s is an electrophoretic artefact caused by the greater amount of unlabeled protein present in lane 1p.

idase subunit IV and which has the same electrophoretic mobility as the precursor of this subunit (Figure 1).

As expected, the product of hybrid-selected translation was imported and processed to the mature subunit by isolated yeast mitochondria: a nuclease-treated reticulocyte lysate was programmed with the hybrid-selected mRNA, incubated with yeast mitochondria in the presence of ATP and separated from mitochondria by centrifugation. Mitochondrial pellet and supernatant were separately analyzed for radiolabeled subunit IV by immunoprecipitation, SDS-polyacrylamide gel electrophoresis and fluorography.

As shown in Figure 2, lanes 1p and 2p, radiolabeled mature subunit IV was mainly found associated with the re-isolated mitochondria whereas most of the remaining radiolabeled precursor was found in the supernatant fraction (lane 1s). Only the radiolabeled mature subunit in the mitochondrial pellet was resistant to externally added protease (lane 2p), indicating that the radiolabeled polypeptide had been transported into the mitochondria. These results suggest that the subunit IV precursor can be imported into mitochondria in the absence of stoichiometric amounts of precursors to the other nuclear-coded cytochrome *c* oxidase subunits.

Table I. Yeast cells transformed with plasmid 19.4 contain increased steady-state levels of cytochrome *c* oxidase subunit IV

Cells transformed with	Intracellular concentration[a] of		Subunit IV
	Cytochrome *c* oxidase subunit IV	ATPase β-subunit	ATPase β-subunit
pFL-1 (vector only)	0.164	0.155	1.05
19.4	0.609	0.133	4.58

[a]Arbitrary units.
Yeast cells transformed with either the control vector pFL-1 or plasmid 19.4 were grown on minimal medium containing 3% ethanol, 3% glycerol, 20 μg/ml L-histidine and 20 μg/ml L-leucine to the late logarithmic phase. Total cell protein was extracted (Riezman *et al.*, 1983b) and equal amounts of extracted protein were analyzed for cytochrome *c* oxidase subunit IV and the F_1 ATPase β-subunit by SDS-polyacrylamide gel electrophoresis and quantitative immune blotting using rabbit antisera monospecific for these polypeptides. The autoradiograms were quantified by scanning and the resulting values are given in the Table in arbitrary units. Values obtained for subunit IV were normalized relative to the values obtained for the F_1-β subunit (cf. last vertical column).

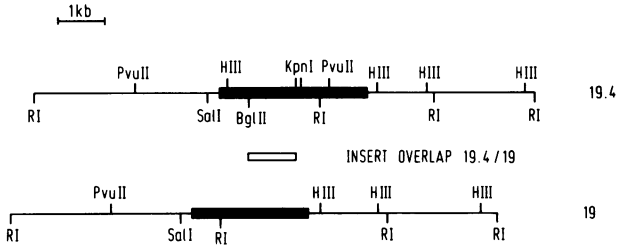

Fig. 3. Restriction maps of plasmid 19 and 19.4 and localization of the gene encoding subunit IV. The vector pFL-1 is indicated by a thin line and the cloned yeast inserts are indicated by filled bars. Only plasmid 19.4 contains an intact copy of the gene encoding subunit IV (see Table I and Results). The gene was localized on the insert as follows: plasmid 19 was linearized with *Pvu*II and plasmid 19.4 was linearized with *Sal*I. Heteroduplexes were allowed to form and analyzed in the electron microscope. Apart from vector sequences, only a 1-kb fragment internal to both cloned sequences (open bar) was found to form duplexes. This observation suggests that at least part of the gene for subunit IV is located on this fragment and that at least one of these recombinant plasmids carries an insert that had been rearranged during the cloning procedure.
RI = *Eco*RI; HIII = *Hind*III.

Increased steady-state levels of subunit IV are found in yeast cells carrying multiple copies of the complete gene

pFL-1 and plasmids derived therefrom exist in yeast in multiple copies (Van Loon *et al.*, 1983b); if they carry a complete yeast gene, they often cause increased steady-state levels of the corresponding gene product. However, introduction of plasmid 19 into wild-type yeast cells did not result in increased steady-state levels of subunit IV, nor were shorter forms of the subunit seen (data not shown). This finding suggested that the plasmid does not carry a complete copy of the gene. Additional recombinant plasmids carrying sequences coding for subunit IV were therefore isolated from the pFL-1 clone bank, using the insert of plasmid 19 as a labeled hybridization probe. One of these additional plasmids (19.4) is described in more detail. As shown in Table I, introduction of plasmid 19.4 into yeast is followed by a 4-fold increase in the steady-

```
   1  GAATTCGATGCGAAATATTCCCTGTCACTTTCCTTTAGTCAACGTATTCTTCTCCCTGAAGAAACAGTATACTAACAATACTCACCCATTTCGATTTTGATGTTGCCATACAAATAGATA

                M  L  S  L  R  Q  S  I  R  F  F  K  P  A  T  R  T  L  C  S  S  R  Y  L  L  Q  Q  K  P  V  V  K  T  A  Q  N  L
 121  ACAAGCACAATGCTTTCACTACGTCAATCTATAAGATTTTTCAAGCCAGCCACAAGAACTTTGTGTAGCTCTAGATATCTGCTTCAGCAAAAACCCGTGGTGAAAACTGCCCAAAACTTA

       A  E  V  N  G  P  E  T  L  I  G  P  G  A  K  E  G  T  V  P  T  D  L  D  Q  E  T  G  L  A  R  L  E  L  L  G  K  L  E  G
 241  GCAGAAGTTAATGGTCCAGAAACTTTGATTGGTCCTGGTGCTAAAGAGGGTACCGTTCCAACAGACCTAGATCAAGAAACTGGTTTAGCTAGGTTAGAATTATTGGGTAAATTAGAGGGT

       I  D  V  F  D  T  K  P  L  D  S  S  R  K  G  T  M  K  D  P  I  I  I  E  S  Y  D  D  Y  R  Y  V  G  C  T  G  S  P  A  G
 361  ATCGATGTTTTCGACACCAAACCATTAGATTCGTCCAGGAAGGGTACCATGAAAGATCCGATCATCATTGAATCTTATGATGATTATCGTTATGTCGGTTGTACGGGTTCTCCCGCTGGT

       S  H  T  I  M  W  L  K  P  T  V  N  E  V  A  R  C  W  E  C  G  S  V  Y  K  L  N  P  V  G  V  P  N  D  D  H  H  H  *
 481  TCACATACTATTATGTGGTTAAAACCAACTGTTAACGAAGTCGCCAGATGCTGGGAATGTGGTTCTGTTTACAAACTAAACCCTGTTGGTGTTCCAAATGATGACCACCATCACTAATCT

 601  TATCATTCAAGTTGCCCTTCTGTTTCTCTTTTACTTTTTACCTTCTTCTTTATTTATATTTTTTTTTTTGGAAACCTAGCGAATATTCTTACCAGGTCTGAGCTTTTACACACGTCATGT

                                                                     M  K  S  Q  K  I  H  N  Q  K  D  R  E  K  V  E  I  Y  F
 721  TAACAAAGTCCACAAGAAAGTTCTTTGCAGGAAGGTTAAGGCGACAGCTTTGCTCTATCTATAATGAAGAGCCAAAAAATACATAACCAAAAAGATCGAGAAAAGGTCGAGATATATTTT

       Y  L  K  I  L  I  N  I  S  T  I  H  V  I  T  D  Y  Y  F  Y  L  C  Q  R  W  L  A  R  G  C  C  E  V  T  S  K  R  F  *
 841  TATTTAAAAATTCTTATTAATATTAGTACTATTCATGTCATAACTGATTACTATTTCTATCTCTGTCAGAGATGGCTAGCTAGAGGTTGTTGCGAAGTTACTTCCAAAAGATTCTAGACT

 961  GT88TTACAGCATCCATACACCCACCCATACATACTGATTTCTTTTTTTAAAGACGTCGATTTTTCGAAAAAAGTAAATTCTCGGCACAGGGAGTTGAATTGAACTTCCCTGCCCCGACG

1081  GTAAGCAGCTTACCGGTATTGCTTCGTTCTCCTTGGGAGATGTTCTCGGCTCTGGAAGGAAAAAACCTTCGTGGGGGGAGGGCTCATATCCAGTAACATAGGCGGAACTCGAAGTGTCAG

1201  CTTACACCGCTTCGTTCTCATTGAGTGTTGAGGGATTACTTGGTATTTGAAATACCTACTAGATTTAATGTTCGTTATAGTAAATGATTTAATTTGTTCGCTATTACAGATAAAAGAACC

1321  ATAGTCTTAAGTAGTATGTTAACGACACAAAAGTGTGAAAGTAGAGAAGGGAAGAACGATGAGATCTTTGAATTGGGAGAAAGTAATAGTGACAAGATATTATTAAAACATHGCAAATGT

1441  AACTTGTTTTCTGAAAGGAAACCAGTTAATCACTAGAAGTGTGCCAATCATGAGTTCACACTAACTAAAGAAAAAGAAGGAAGTCGCCCATCGTATGCAATGTACTTATTATTGAAGAAT

1561  ATTGTACCGAGAAAATTTCAAATGTTACGTA
```

Fig. 4. Nucleotide sequence of the gene encoding subunit IV of yeast cytochrome *c* oxidase. The sequence shown starts from the *Eco*RI site of the 1900-bp *Eco*RI-*Hind*III fragment of 19.4 (Figure 3). The gene encoding subunit IV starts at position 130 and ends at position 597. Another open reading frame starts at position 784 and ends at position 957. Any evidence for the expression of this reading frame is lacking however. The uncertainty code (8) at positions 963 and 964 indicates that the sequence is either CG or GC and the H at position 1431 indicates the sequence G or GG. The deduced amino acid sequence is indicated in the one-letter code: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

state level of subunit IV as detected by immune blotting. Thus, plasmid 19.4 appears to carry a complete copy of the subunit IV gene. The accumulated subunit IV was present as the mature form of the subunit (data not shown), suggesting that the subunit had been imported into mitochondria. Restriction maps of plasmids 19 and 19.4 are presented in Figure 3. The size of the insert in plasmid 19.4 is ~3000 bp.

Identification of the subunit IV gene carried by plasmid 19.4

To locate the subunit IV gene on plasmid 19.4 we searched for sequence homologies between plasmids 19 and 19.4 by heteroduplex analysis using the electron microscope. The result showed that the gene is located within a 1900-bp *Hind*III-*Eco*RI fragment of plasmid 19.4 (see Figure 3). The DNA sequence of part of this fragment revealed two open reading frames exceeding 50 codons (Figure 4). The first open reading frame starts at position 130 and could encode a polypeptide of 155 amino acids. The second reading frame starts at position 784, ends at position 957 and could encode a polypeptide of 57 amino acids.

To identify one of these open reading frames as the structural gene for cytochrome *c* oxidase subunit IV, the predicted protein sequence was compared with that determined for the mature subunit IV isolated from yeast cytochrome *c* oxidase. Sequence analysis of the purified protein has not been completely finished, some ambiguities in the amino acid sequencing experiments were left unresolved. Nevertheless, the position of most of the amino acids, predicted from the DNA sequence could be unambiguously confirmed (Figure 5).

The directly determined amino acid sequence agrees well with the nucleotide sequence of the 465-bp open reading

```
                                          (e)q k p v
M L S L R Q S I R F F K P A T R T L C S S R Y L L Q Q K P V
                 10                  20                 30

                (q)
v k t a q n l a e v n g p e t l i g p g a k e g t v p t d l
V K T A Q N L A E V N G P E T L I G P G A K E G T V P T D L
                 40                  50                 60

d q e t g l a r l e l l g k l e g i d v f d t k p l d s s r
D Q E T G L A R L E L L G K L E G I D V F D T K P L D S S R
                 70                  80                 90

                                                   x
k g t m k d p i x i e s x d x y r y v g x t g s p a g s (s)t
K G T M K D P I I I E S Y D D Y R Y V G C T G S P A G S H T
                 100                 110                120

i m w l k p t v n e v a r c w e ( )g v s y k l n p v g v p n
I M W L K P T V N E V A R C W E C G S V Y K L N P V G V P N
                 130                 140                150

d d h h ( )

D D H H H
```

Fig. 5. Partial amino acid sequence of mature subunit IV isolated from yeast cytochrome *c* oxidase. Residues determined by amino acid sequencing are given in lower case letters, residues deduced from the nucleotide sequence in upper case letters. Amino acid residues which were not reliably identified are indicated by 'x'; those that are in disagreement with the nucleotide sequence of the gene are bracketed (see Results for further explanation). The underlined residues 139 and 140 occur in the protein sequence predicted from the DNA sequence in the opposite order. The Gln placed at position 45–46 and the unidentified residue placed at position 116–117 lack codons in the corresponding nucleotide sequence and are therefore placed above the amino acid sequence. The differences between both protein sequences are most likely due to inaccuracies in the protein sequencing, since this sequencing has not been fully completed. The sequence presented here agrees with the partial amino acid sequence of the yeast subunit IV presented by Power *et al.* (1984c), with the exception of the amino acids 26 (Q *versus* E) and 73 (G *versus* C) in the precursor protein (positions 1 and 48 of the mature protein sequence).

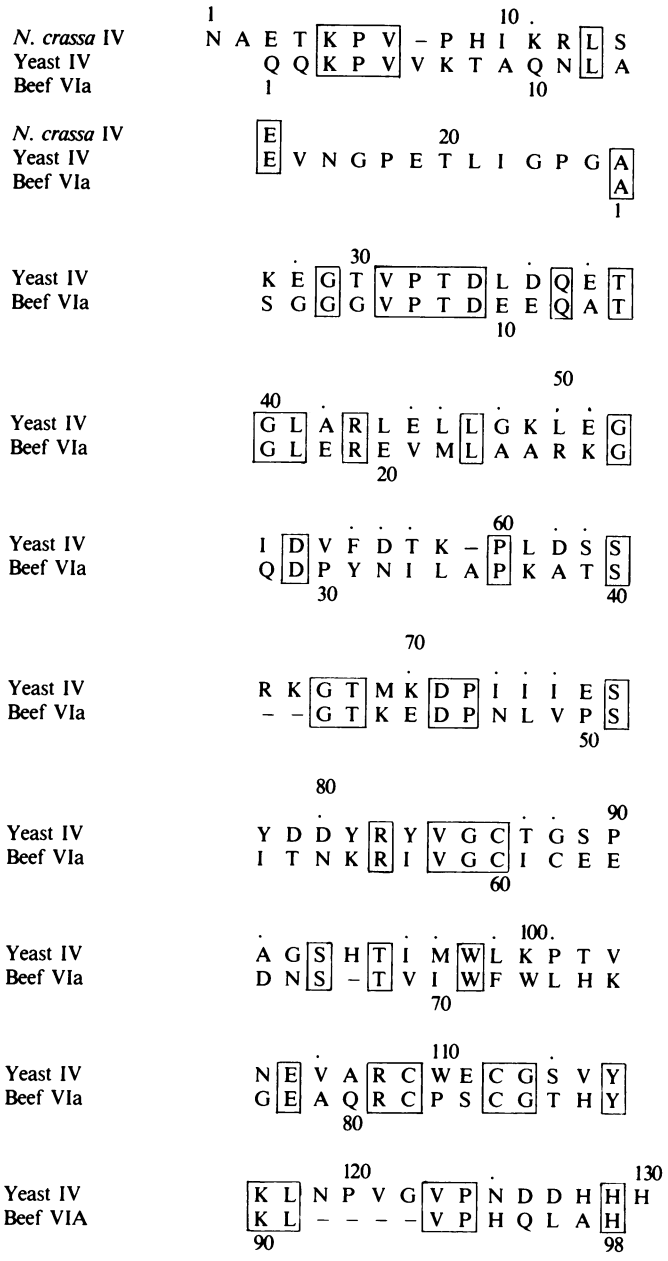

Fig. 6. Cytochrome *c* oxidase subunits IV from yeast and *Neurospora crassa* and VIa from beef are homologous. The sequence of the bovine subunit VIa was determined by Biewald and Buse (1982), the amino-terminal sequence of subunit IV of the *N. crassa* complex by Werner *et al.* (1980). The dots above the protein sequence indicate changes in amino acid sequence possibly due to a single base pair mutation. The amino acids of the yeast subunit IV have been renumbered compared with Figure 5. 1 = first residue of each mature protein.

frame. Six out of the 130 amino acids (indicated by brackets or underlining) do not agree. Comparison of the nucleotide with the amino acid sequence also shows that subunit IV is synthesized as a precursor protein with a transient amino-terminal extension of 25 amino acids. The mol. wts. of the precursor and the mature subunit are predicted to be 17 563 and 14 570, respectively. These values are in good agreement with those determined by SDS-polyacrylamide gel electrophoresis of the mature subunit IV (Poyton and Schatz, 1975) and its precursor (Lewin *et al.*, 1980).

Discussion

Cytochrome c oxidase subunits IV of yeast and Neurospora crassa and subunit VIa of beef are homologous

Cytochrome *c* oxidase from mammalian sources appears to contain significantly more subunits than the yeast enzyme (Kadenbach *et al.*, 1983). So far, amino acid and DNA sequencing has revealed extensive homology between the mitochondrially made subunits I–III from mammals and yeast (Tzagoloff, 1982), and between several of the nuclear-encoded subunits in yeast and subunits in the beef complex (Gregor and Tsugita, 1982; Power *et al.*, 1984b, 1984c). The amino acid sequence of yeast subunit IV reported here shows a weak, but significant homology to that of the bovine subunit VIa (Biewald and Buse, 1982) (see Figure 6). Power *et al.* (1984c) also show homology between both proteins on the basis of a partial amino acid sequence [the same bovine subunit VIa is called Vb in their publication, according to Kadenbach's nomenclature (Kadenbach *et al.*, 1983)]. The bovine subunit VIa has a length of only 98 amino acids, while the mature subunit IV of yeast is 130 amino acids long (its precursor has a length of 155 residues). An extra, N-terminally located stretch of 25 amino acids in the yeast subunit IV accounts for most of the difference in length between both mature proteins. Subunit IV of the *N. crassa* complex also shows sequence homology to the yeast subunit IV (Figure 6). The *Neurospora* subunit, however, is two amino acids longer at the amino terminus than the yeast subunit IV.

No homology was detected between the amino acid sequence of the yeast subunit IV and those of the other nuclear-encoded bovine subunits sequenced so far (IV: Sacher *et al.*, 1979; V: Tanaka *et al.*, 1979; VII: Steffens *et al.*, 1979; VIIIa: Buse and Steffens, 1978; VIIIb: Meinecke *et al.*, 1984).

The mature yeast subunit IV contains 19 acidic and 13 basic amino acid residues uniformly distributed throughout the sequence. The overall polarity (Capaldi and Vanderkooi, 1972) is 47% and thus in the range of water-soluble proteins. These facts agree with previous observations that the mature subunit has an isoelectric point of 5.5 (Cabral and Schatz, 1978) and that it is one of the most hydrophilic subunits of yeast cytochrome *c* oxidase (Poyton and Schatz, 1975).

The transient pre-sequence

Cytochrome *c* oxidase subunit IV from yeast is one of the few imported mitochondrial proteins for which both the nucleotide sequence of the gene and the amino acid sequence of the mature protein are known (see Figure 5). It is thus possible to identify the exact length of the transient pre-sequence of subunit IV and the bond which is cleaved during the import process. The transient pre-sequence consists of 25 amino acids, five of them basic and none of them acidic. Most amino-terminal regions of imported mitochondrial proteins share this property (Reid, 1984).

During import into the mitochondria, the yeast subunit IV precursor must be cleaved at the Leu_{25}-Gln_{26} bond, presumably by the chelator-sensitive matrix-localized protease (Böhni *et al.*, 1980, 1983; McAda and Douglas, 1982). Comparison of this cleavage site with those of three other mitochondrial precursors indicates that the protease does not recognize a specific peptide bond, but rather cleaves after an uncharged amino acid (Figure 7).

A similar cleavage pattern has been found for the membrane-bound 'signal peptidases' of the bacterial plasma

Mitochondrial precursor polypeptide	Sequence surrounding cleavage site	Reference(s)
Yeast cytochrome *c* oxidase subunit IV	...Arg-Tyr-Leu-Leu↓-Gln-Gln-Lys...	This paper
Yeast cytochrome *c* oxidase subunit VI	...Ser-Arg-Lys-Tyr↓-Ser-Asp-Ala...	Wright *et al.*, 1984 Gregor and Tsugita, 1982
Yeast superoxide dismutase	...Ala-Arg-Arg-Thr↓-Lys-Val-Thr...	Marres *et al.*, in preparation Harris and Steinman, 1977
Neurospora ATPase subunit IX	...Gln-Lys-Arg-Ala↓-Tyr-Ser-Ser...	Viebrock *et al.*, 1982 Sebald *et al.*, 1980

Fig. 7. Peptide bonds cleaved by the mitochondrial processing protease. The cleavage site is indicated by the arrow.

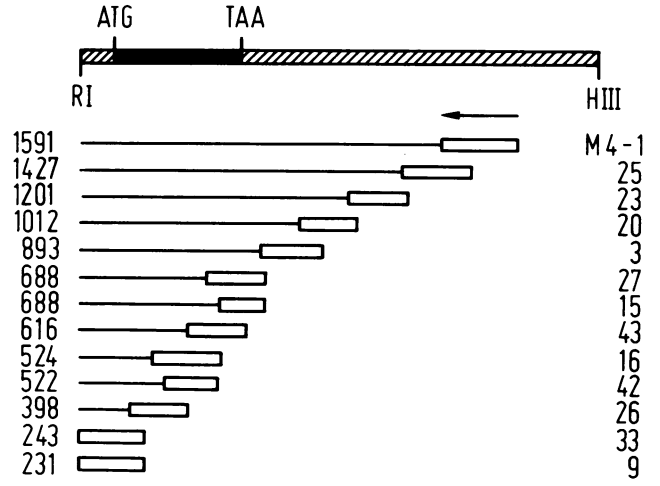

Fig. 8. DNA sequence analysis of the gene encoding subunit IV. Sequencing strategy. The 1900-bp *Eco*RI-*Hind*III (RI-HIII) fragment present in the cloned insert of plasmid 19.4 (upper bar) was cloned into M13-mp8. A collection of M13 phages containing inserts of decreasing lengths was constructed using the exonuclease *Bal*31 as described in Materials and methods. The numbers of the recombinant clones used for DNA sequence analysis are indicated (e.g. M4-20). The arrow shows the direction of DNA sequencing. The nucleotide at which DNA sequence obtained from each clone starts is indicated on the left side (cf. Figure 4 for numbering). The total area sequenced with each subclone is indicated by the open bars. The position of the subunit IV gene on the 1900-bp fragment is deduced on the basis of the DNA and protein sequence and is indicated by the filled bar on top.

membrane and the endoplasmic reticulum (Kreil, 1981; Von Heijne, 1983). However, the sequences upstream from the cleavage sites suggest some differences between the mitochondrial protease and its counterparts in bacteria and endoplasmic reticulum. On analyzing 78 different signal sequences of eukaryotic secretory proteins, Von Heijne (1983) noted that none of them contained a charged or bulky amino acid at position −3 (i.e., the third position upstream from the cleavage site) and that only a few of them contained a small, uncharged amino acid at position −2. A similar pattern was found for bacterial secretory proteins. In contrast, three of the four mitochondrial pre-sequences listed in Figure 7 feature a basic amino acid at position −3 and one of them contains a small, uncharged amino acid at position −2. These differences are not unexpected since the signal peptidases associated with the bacterial plasma membrane or the endoplasmic reticulum are integral membrane proteins (Kreil, 1981) whereas the mitochondrial processing protease is a protein releasable with the matrix contents (Böhni *et al.*, 1980). Since the mitochondrial processsing protease does not cleave denatured precursors (Böhni *et al.*, 1983; Ohta and Schatz, 1984), it presumably recognizes a domain structure which is shared by most (and perhaps all) larger precursor polypeptides which enter the matrix space during their import into the mitochondria.

Possible function of subunit IV

Is the polypeptide studied here a *bona fide* subunit of cytochrome *c* oxidase or merely a tightly and reproducibly bound contaminant? This question is relevant since no function has yet been assigned to any of the imported cytochrome *c* oxidase subunits. A yeast mutant containing an insertion in the nuclear gene coding for subunit IV has been recently constructed. This mutant lacks cytochrome *c* oxidase activity as well as the absorption band at 605 nm typical of cytochrome aa_3 (Dowhan *et al.*, in preparation). Therefore, the subunit IV gene product is required for oxidase activity.

Materials and methods

Recombinant DNA clone bank

The yeast clone bank contained partially *Sau*3A-digested genomic DNA from the *Saccharomyces cerevisiae* strain FL 100 inserted in the single *Bam*HI site of the *E. coli*-yeast 'shuttle vector' pFL-1. The vector contains plasmid pBR322, a 1.1-kb *Hind*III fragment carrying the yeast *URA 3* gene and the 2.1-kb *Eco*RI fragment of the yeast 2 μm plasmid (Chevallier *et al.*, 1980). This bank was a generous gift from Professor F. Lacroute, Strasbourg.

Cloning in bacteriophage M13 mp8 and preparation of a Bal31-clone bank

Plasmid 19.4 (see Figure 3) was digested with *Hind*III and *Eco*RI. A 1900-bp fragment (which electron microscopic heteroduplex analysis had suggested to contain the gene for subunit IV) was purified by electrophoresis through a 1% (w/v) low-melting agarose gel and ligated into the replicative form of phage M13 mp8 cut with *Hind*III and *Eco*RI. About 20 μg DNA from the resulting recombinant phage was digested with *Hind*III, diluted 5-fold and digested with 5 units of nuclease *Bal*31 in the presence of 12 mM $CaCl_2$, 12 mM $MgCl_2$, 200 mM NaCl, 20 mM Tris-Cl (pH 8.0) and 1 mM EDTA at 30°C in a final volume of 250 μl. Samples of 10, 20, 40, 80 and 100 μl, were withdrawn after 2, 4, 6, 10 and 15 min, respectively. Reactions were stopped by addition of TNS [10 mM Tris-Cl, pH 7.5, 500 mM NaCl, 0.1% (w/v) SDS] to a final volume of 200 μl. An equal volume of phenol, equilibrated with TNS, was then added and the reactions were stored on ice.

After phenol extraction and ethanol precipitation, DNA samples were digested with *Eco*RI and half of each was analyzed by electrophoresis through a 1.7% (w/v) agarose gel. The remainder was re-extracted with phenol and

ethanol precipitated. About 10% of this DNA was ligated into M13 mp8, restricted with *Eco*RI and *Hind*II. Plaque-lift and hybridization with DNA from clone 19.4, labelled with ^{32}P by nick-translation (Jeffreys and Flavell, 1977) was used to select transformants containing part of the 19.4 insert. Insert sizes were determined by restriction of double-stranded DNA followed by gel electrophoresis.

DNA sequence analysis

DNA sequence analysis with chain elongation inhibitors was performed essentially as described by Sanger *et al.* (1977, 1980). The sequencing strategy is summarized in Figure 8.

Amino acid sequencing

The haploid *S. cerevisiae* strain D 273-10B (ATCC 25657) was grown in 100 litre batches on 1% yeast extract – 2% peptone-2% glucose (Gregor and Tsugita, 1982). Mitochondria were isolated from the cells (yield: 6–11 g protein/kg cells wet weight; Deters *et al.*, 1978), the mitochondria were converted to submitochondrial particles (Deters *et al.*, 1978) and cytochrome *c* oxidase was prepared essentially as described by Birchmeier *et al.* (1976). The yield was usually 4–5 mg cytochrome *c* oxidase (~8–11 nmol heme *a*/mg protein) per g submitochondrial particle protein.

Subunit IV was isolated from cytochrome *c* oxidase as described by Poyton and Schatz (1975). In the final step of this procedure, subunit IV was eluted at a NaCl concentration of between 55 and 75 mM. The yield was ~40–80 μg subunit IV/mg cytochrome *c* oxidase. An aliquot of the purified subunit was always checked for homogeneity by SDS-polyacrylamide gel electrophoresis. When contaminants were detected, the subunit was repurified by repeating the DEAE-cellulose chromatography step. Amino acid sequencing of proteolytic fragments by manual Edman degradation and identification of phenylthiohydantoin-derivatized amino acids by h.p.l.c. were done exactly as described earlier (Gregor and Tsugita, 1982). Carboxy-terminal analysis was performed by digestion with the carboxypeptidases A and B coupled with identification of the released amino acids in a Durrum amino acid analyzer.

Immunochemical methods

Rabbit antisera against isolated yeast cytochrome *c* oxidase subunits IV and VI (Poyton and Schatz, 1975; Gregor and Tsugita, 1982) were raised as described (Daum *et al.*, 1982) and tested by immune blotting (Towbin *et al.*, 1979) with SDS-dissociated yeast mitochondria as test antigen. Immunoprecipitation of ^{35}S-labeled polypeptides was performed according to Glasser (1984).

Electron microscopic heteroduplex analysis

Plasmids 19 and 19.4 (see Figure 3) were cut with *Pvu*II and *Sal*I, respectively. Heteroduplexes were formed by hybridization at 35°C in 70% formamide, 10 mM Tris-HCl (pH 8.5), 0.3 M NaCl, 1 mM EDTA. Nucleic acids were analyzed by electron microscopy according to Davis *et al.* (1968, 1971).

Miscellaneous methods

Published methods were employed for isolation of recombinant clones (Van Loon *et al.*, 1983a) hybrid-selected translation (De Haan *et al.*, 1984) isolating total yeast mRNA (Maccechini *et al.*, 1979), preparing a nuclease-pretreated reticulocyte lysate (Pelham and Jackson, 1976), SDS-polyacrylamide gel electrophoresis (Laemmli, 1970), fluorography (Chamberlain, 1979), determining protein (Lowry *et al.*, 1951) measuring heme *a* (Rieske, 1967) isolating yeast mitochondria (Daum *et al.*, 1982), assaying import of *in vitro* synthesized precursor polypeptides into isolated yeast mitochondria (Gasser, 1984), transformation and growth of yeast strain SF 747-19D (Riezman *et al.*, 1983a and isolation of M13 double-stranded DNA and recombinant phages (Sanger *et al.*, 1980).

Acknowledgements

We wish to thank W. Oppliger,G. Gasser, S. Straumann, K. Suda and W. Jager for excellent technical assistance and J. Horlacher for typing the manuscript, Dr. A.C. Arnberg and J.H. Vermaat (University of Groningen, The Netherlands) for performing the electron microscopic heteroduplex analysis, Professor F. Lacroute (University of Strasbourg, France) for the gift of the yeast clone bank and Drs. G. Buse and G. Steffens (Technische Hochschule, Aachen, FRG) for their help in comparing amino acid sequences of different cytochrome *c* oxidase subunits, Dr. R.O. Poyton, for communication of sequence data before publication and Drs. R. Hay, T. Hase and S. Ohta for critical reading of the manuscript. This study was supported in parts by grants (to L.A.G.) of the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), by grants 3.394–0.83 and 3.037–1.81 (to G.S.) from the Swiss National Science Foundation, by a fellowship (to H.R.) from the Jane Coffin Childs Memorial Fund for Medical Research and by a short-term EMBO fellowship to H.R.

References

Biewald,R. and Buse,G. (1982) *Z. Physiol. Chem.*, **363**, 1141-1153.
Birchmeier,W., Kohler,C.E. and Schatz,G. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 4334-4338.
Böhni,P., Gasser,S., Leaver,C. and Schatz,G. (1980) in Kroon,A.M. and Saccone,C. (eds.), *The Organization and Expression of the Mitochondrial Genome*, Elsevier/North Holland, Amsterdam, pp. 423-433.
Böhni,P.C., Daum,G. and Schatz,G. (1983) *J. Biol. Chem.*, **258**, 4937-4943.
Buse,G. and Steffens,G.J. (1978) *Z. Physiol. Chem.*, **359**, 1005-1009.
Cabral,F. and Schatz,G. (1978) *J. Biol. Chem.*, **253**, 4396-4401.
Capaldi,R.A. and Vanderkooi,G. (1972) *Proc. Natl. Acad, Sci. USA*, **69**, 930-932.
Chamberlain,J.P. (1979) *Anal. Biochem.*, **98**, 132-135.
Chevallier,M.-R., Bloch,J.C. and Lacroute,F. (1980) *Gene*, **11**, 11-19.
Cumsky,M.G., McEwen,J.E., Ko,C. and Poyton,R.O. (1983) *J. Biol. Chem.*, **258**, 13418-13421.
Daum,G., Gasser,S.M. and Schatz,G. (1982) *J. Biol. Chem.*, **257**, 13075-13080.
Davis,R.W. and Davidson,N. (1968) *Proc. Natl. Acad. Sci. USA*, **60**, 243-250.
Davis,R.W., Simon,M. and Davidson,N. (1971) *Methods Enzymol.*, **21**, 413-428.
De Haan,M., Van Loon,A.P.G.M., Kreike,J., Vaessen,R.T.M.J. and Grivell,L.A. (1984) *Eur. J. Biochem.*, **138**, 169-177.
Deters,D., Müller,U., Homberger,H. (1978) in Prescott,D.M. (ed.), *Methods Cell Biol.*, Vol. **20**, Academic Press, NY, pp. 107-112.
Gasser,S.M. and Schatz,G. (1982) *J. Biol. Chem.*, **258**, 3427-3430.
Gasser,S.M., Daum,G. and Schatz,G. (1982) *J. Biol. Chem.*, **257**, 13034-13041.
Gasser,S.M. (1984) *Methods Enzymol.*, **97**, 329-336.
Gregor,I. and Tsugita,A. (1982) *J. Biol. Chem.*, **257**, 13081-13087.
Harris,J.I. and Steinman,H.M. (1977) in Michelson,A.M., McCord,J.M. and Fridovich,I. (eds.), *Superoxide and Superoxide Dismutases*, Academic Press, NY, pp. 225-230.
Jeffreys,A.J. and Flavell,R.A. (1977) *Cell*, **12**, 429-439.
Kadenbach,B., Jarausch,J., Hartmann,R. and Merle,P. (1983) *Anal. Biochem.*, **129**, 517-521.
Kreil,G. (1981) *Annu. Rev. Biochem.*, **50**, 317-348.
Laemmli,U.K. (1970) *Nature*, **227**, 680-685.
Lewin,A., Gregor,I., Mason,T.L., Nelson,N. and Schatz,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3998-4002.
Lowry,O.H., Rosebrough,N.J., Farr,A.L. and Randall,R.J. (1951) *J. Biol. Chem.*, **193**, 265-275.
Maccechini,M.-L., Rudin,Y., Blobel,G. and Schatz,G. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 343-347.
McAda,P.C. and Douglas,M.G. (1982) *J. Biol. Chem.*, **257**, 3177-3182.
Meinecke,L., Steffens,G.J. and Buse,G. (1984) *Z. Physiol. Chem.*, **365**, 313-320.
Mihara,K. and Blobel,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4160-4164. 4164.
Ohta,S. and Schatz,G. (1984) *EMBO J.*, **3**, 651-657.
Pelham,H.R.B. and Jackson,R.J. (1976) *Eur. J. Biochem.*, **67**, 247-256.
Power,S.D., Lochrie,M.A., Sevarino,K.A., Patterson,T.E. and Poyton,R.O. (1984a) *J. Biol. Chem.*, **259**, 6564-6570.
Power,S.D., Lochrie,M.A., Patterson,T.E. and Poyton,R.O. (1984b) *J. Biol. Chem.*, **259**, 6571-6574.
Power,S.D., Lochrie,M.A. and Poyton,R.O. (1984c) *J. Biol. Chem.*, **259**, 6575-6578.
Poyton,R.O. and Schatz,G. (1975) *J. Biol. Chem.*, **250**, 752-761.
Reid,G.A. (1984) in Martonosi,A. (ed.), *Current Topics in Membrane Transport*, Academic Press, in press.
Rieske,J.S. (1967) *Methods Enzymol.*, **10**, 488-493.
Riezman,H., Hay,R., Gasser,S., Daum,G., Schneider,G., Witte,C. and Schatz,G. (1983a) *EMBO J.*, **2**, 1105-1111.
Riezman,H., Hase,T., Van Loon,A.P.G.M., Grivell,L.A., Suda,K. and Schatz,G. (1983b) *EMBO J.*, **2**, 2161-2168.
Sacher,R., Steffens,G.J. and Buse,G. (1979) *Z. Physiol. Chem.*, **360**, 1385-1392.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
Schatz,G. and Mason,T.L. (1974) *Annu. Rev. Biochem.*, **43**, 51-87.
Sebald,W., Machleidt,W. and Wachter,E. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 785-789.

Steffens,G.C.M., Steffens,G.J. and Buse,G. (1979) *Z. Physiol. Chem.*, **360**, 1641-1650.
Tanaka,M., Haniu,M., Yasunobu,K.T., Yu,C.-A., Yu,L., Wei,Y.-H. and King,T.E. (1979) *J. Biol. Chem.*, **254**, 3879-3885.
Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354.
Tzagoloff,A. (1982) *Mitochondria*, published by Plenum Press, New York.
Van Loon,A.P.G.M., de Groot,R.J., Van Eyk,E., Van der Horst,G.T.J. and Grivell,L.A. (1982) *Gene*, **20**, 323-337.
Van Loon,A.P.G.M., Maarse,A.C. Riezman,H. and Grivell,L.A. (1983a) *Gene*, **26**, 261-272.
Van Loon,A.P.G.M., Van Eyk,E. and Grivell,L.A. (1983b) *EMBO J.*, **2**, 1765-1776.
Von Heijne,G. (1983) *Eur. J. Biochem.*, **133**, 17-21.
Viebrock,A., Perz,A. and Sebald,W. (1982) *EMBO J.*, **1**, 565-571.
Werner,S., Machleidt,W., Bertrand,H. and Wild,G. (1980) in Kroon,A.M. and Saccone,C. (eds.), *The Organization and Expression of the Mitochondrial Genome*, Elsevier/North Holland, Amsterdam, pp. 399-411.
Wright,R.M., Ko,C., Cumsky,M.G. and Poyton,R.O. (1984) *J. Biol. Chem.*, in press.

Received on 12 July 1984