The EMBO Journal vol.3 no.5 pp.1039–1043, 1984

The DNA sequence of the nuclear gene coding for the 17-kd subunit VI of the yeast ubiquinol-cytochrome *c* reductase: a protein with an extremely high content of acidic amino acids

Adolphus P.G.M.Van Loon[1], Raoul J.De Groot, Muus De Haan, Albert Dekker and Leslie A.Grivell*

Section for Molecular Biology, Laboratory of Biochemistry, University of Amsterdam, Kruislaan 318, 1098 SM Amsterdam, The Netherlands

[1]Present address: Department of Biochemistry, Biocenter, University of Basel, CH-4056 Basel, Switzerland
*To whom reprint requests should be sent
Communicated by L.A.Grivell

We have determined the DNA sequence of the nuclear gene coding for the 17-kd subunit VI of the ubiquinol-cytochrome c reductase. The reading frame found encodes a putative polypeptide of 17 394 daltons. This protein is highly unusual: 38% of its residues are acidic and 14% are basic amino acids. The most notable feature in the protein sequence is a stretch of 25 consecutive acidic amino acids. The polypeptide has homology with the 9-kd so-called 'hinge' protein of beef-heart complex III, which also has a cluster of acidic residues. Acidic amino acids are likely to be essential for the function of these proteins, since their degree of conservation is higher than that of other residues.
Key words: mitochondrial biosynthesis/genes for imported mitochondrial proteins/DNA sequence analysis

Introduction

The ubiquinol-cytochrome *c* reductase (complex III or cytochrome bc_1 complex) is part of the mitochondrial respiratory chain and is located in the mitochondrial inner membrane. It contains three polypeptides with known prosthetic groups: the cytochromes c_1 and *b* and an iron-sulphur protein. In addition, the complex contains a number of other subunits, whose exact number (between four and seven) is subject to debate, since their function is unknown. Of the eight putative subunits of the yeast complex, seven are usually found in fixed stoichiometric amounts. The eighth, a 17-kd protein, has been reported to be present in variable and substoichiometric amounts or even to be entirely lacking (Katan and Groot, 1976; Reed and Hess, 1977; Kreike, 1982), thus raising doubts as to whether it belongs to the complex. The 17-kd polypeptide is characterized by a high polarity (66.9%). It has a pI of 4.7 and is extremely rich in asparagine/aspartic acid (17.8%) and glutamine/glutamic acid (28.6%) (Katan and Groot, 1976). A 9-kd subunit of the beef complex – the so-called 'hinge' protein – has similar characteristics (Wakabayashi *et al.*, 1982), suggesting that it might be (partially) homologous to the 17-kd subunit. The 'hinge' protein is thought to mediate interaction between the cytochromes c_1 and *c* (King, 1983) and is found firmly associated in a 1:1 ratio with cytochrome c_1 in subcomplexes obtained from the purified beef enzyme (Kim and King, 1981; Schägger *et al.*, 1983).

To analyse the 17-kd subunit VI of the yeast complex in more detail, the DNA sequence of its gene has been determined. The encoded polypeptide indeed has homology with the beef-heart 9-kd protein and is, therefore, identified as a true subunit of the yeast complex. Its predicted properties are discussed with respect to function, complex assembly and topology.

Results

Precise localization of the gene coding for the 17-kd protein

We have demonstrated previously (Van Loon *et al.*, 1982) that plasmid 209 contains the cloned yeast nuclear gene for the 17-kd subunit of complex III. Figure 1 shows a restriction map of this plasmid. The position of the gene on the insert has been determined (Van Loon *et al.*, 1982). The region around the gene was characterized further by the construction of a detailed restriction map using the 660-bp *Rsa*I fragment as labelled probe on Southern blots. This fragment is capable of hybrid selection of the 17-kd mRNA (not shown); the resulting map is also shown in Figure 1.

Transcript mapping

Mapping of the mRNA for the 17-kd protein was carried out by S1 nuclease analysis, using the 720-bp *Sau*3A and 660-bp *Rsa*I fragments (see Figure 1). Since levels of translatable mRNAs coding for subunits of complex III have been estimated to be only 0.01–0.03% of total mRNA, even in derepressed conditions (Van Loon *et al.*, 1982), we have used uniformly-labelled restriction fragments for the analysis. As shown in Figure 2, the lengths of the fragments, protected by RNA from cells grown on lactate, were ~360-bp and 550-bp for the 720-bp *Sau*3A and 660-bp *Rsa*I fragments, respectively. Digestion of the labelled *Rsa*I fragment with *Bsp*RI prior to hybridization to the mRNA, resulted in a protected fragment of 300 nucleotides derived from the 410-bp *Bsp*RI-*Rsa*I fragment, while the 210-bp *Bsp*RI fragment was completely protected. A second 39-bp *Rsa*I-*Bsp*RI fragment is too short to be detected on the gel used. The positions of the *Bsp*RI sites were deduced from the DNA sequences (Figure 4). The results are summarized in Figure 2. They locate the

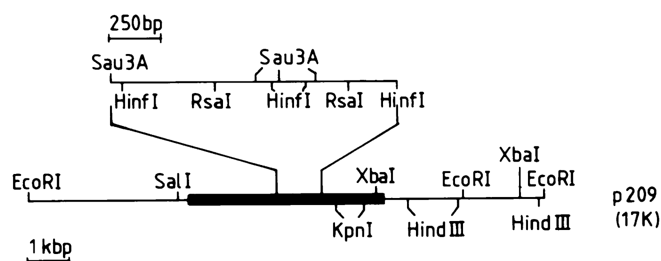

Fig. 1. Restriction map of plasmid 209 encoding the gene for the 17-kd protein. The vector (thin line) is pFL1, which consists of pBR322, a 1.1-kb *Hind*III fragment that carries the yeast URA3 gene and a 2.1-kb fragment of the yeast 2-μm circle (Chevallier *et al.*, 1980). The insert is indicated by a thick line. A detailed map of the region around the gene for the 17-kd protein was constructed by hybridization of Southern blots with the 660-bp *Rsa*I fragment. The position of the gene on the insert is deduced on basis of data presented in Van Loon *et al.* (1982).

© IRL Press Limited, Oxford, England.

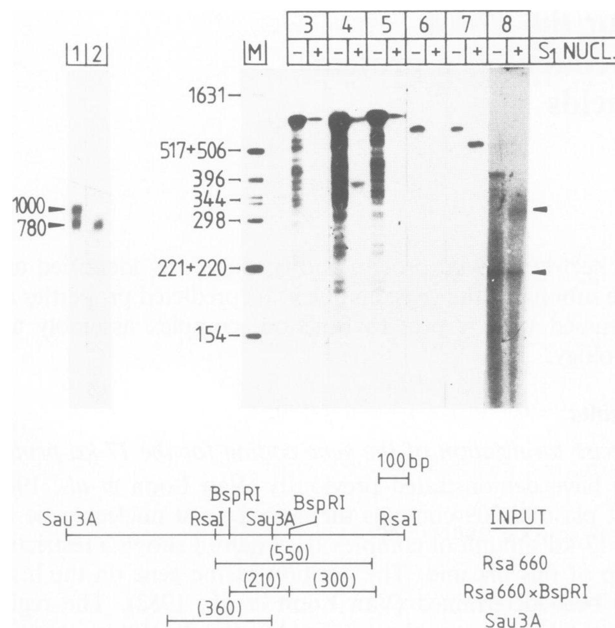

Fig. 2. Transcript mapping. The length of the mRNA that encodes the 17-kd protein was determined. Total glyoxylated yeast RNA was size-fractionated on an agarose gel, blotted to nitrocellulose filters and hybridized with the ^{32}P-labelled *SalI-HindIII* fragment that contains the entire cloned sequence (**lane 1**) or the 660-bp *RsaI* fragment in M13 mp7 (**lane 2**) (see Figure 1). S1 nuclease analysis was performed by use of uniformly-labelled DNA fragments. Total RNA from lactate-grown cells was used. The lengths of the fragments protected from S1 nuclease digestion were calculated by comparison with pBR322 digested with *HinfI* and 5′-labelled with [γ^{32}P]ATP and polynucleotide kinase. The fragments used are: **lanes 3–5**, 720-bp *Sau3A*; **lanes 6–8**, 660-bp *RsaI*. The *RsaI* fragment in **lane 8** was digested with *BspRI* prior to hybridization. The RNAs used are: **lanes 3** and **6**, *Escherichia coli* tRNA; **4**, **7** and **8**, total yeast RNA; **5**, total yeast RNA treated with ribonuclease prior to hybridization. The results of the S1 nuclease analysis are summarized in the map. The sizes of the protected fragments are indicated in brackets.

transcript roughly between positions −130 to −120 and 570 to 580 of the DNA sequence (see Figure 4). The size of the protected region (700 bp) is consistent with a total length of 780 nucleotides as determined for the transcript on RNA blots (Figure 2, lane 2), if one assumes that this mRNA – like other mRNAs in yeast – has a poly(A) tail of 50–100 nucleotides. The major body of the mRNA coding sequence is thus continuous, although splice points close to an end cannot be excluded.

DNA sequence analysis and predicted primary structure of the 17-kd protein

The DNA sequence of fragments subcloned in the bacteriophage M13 mp7 was determined by use of the dideoxy method (Sanger *et al.*, 1980). The sequencing strategy is summarized in Figure 3 and the DNA sequence is presented in Figure 4. The longest reading frame found is 441 nucleotides long, it starts at the methionine at sequence position 1–3 and is terminated by a single TAG codon at position 442–444. The reading frame encodes a putative protein of 147 amino acids with a calculated mol. wt. of 17 394 daltons. This is much shorter than expected on the basis of the migration behaviour of the precursor protein (26 kd) on SDS-polyacrylamide gels (see Discussion). The encoded protein is clearly unusual. It is extremely rich in charged residues (52%), most of which are acidic (38%). The most striking feature, however, is a continuous stretch of 25 acidic amino

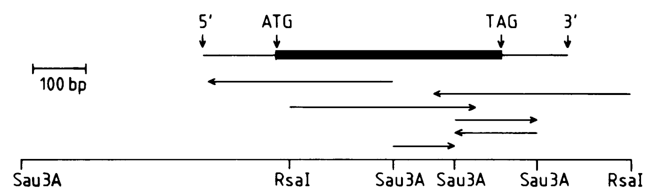

Fig. 3. Sequencing strategy. The DNA sequence of the region encoding the 17-kd protein was determined by use of the dideoxy chain termination method on fragments sub-cloned in the phage M13 mp7. The arrows indicate the parts of the sequence determined by use of various *Sau3A* and *RsaI* fragments. The length of the mRNA coding region, as determined by S1 nuclease analysis (see Figure 2) is indicated by a thin line, while the position of the gene, as deduced from the DNA sequence (see Figure 4), is indicated by a thick line.

acids (sequence positions 145–219). The amino acid composition of the purified, mature 17-kd subunit (Katan and Groot, 1976) and that predicted for the precursor protein from the DNA sequence are listed in Table I. Precise comparison cannot be made, since the exact terminal residues of the mature protein are not known. Both amino acid compositions are highly similar and reveal the same characteristics (especially the very high content of Asx and Glx, see Table I). We conclude, therefore, that the sequence presented is that of the gene for the 17-kd protein. Other open reading frames are absent from this area, apart from a frame of 162 nucleotides that is entirely internal to the gene for the 17-kd protein (sequence positions 79–244). This reading frame does not correspond to an individual transcript and there is no evidence for expression of the highly positively-charged protein which it encodes.

Discussion

Homology between the 17-kd protein and the 9-kd 'hinge' protein of beef complex III

The amino acid sequence deduced from the DNA sequence of the gene for the 17-kd subunit of the ubiquinol-cytochrome *c* reductase is unusual. Acidic residues make up 38% of the sequence and the uninterrupted stretch of 25 acidic amino acids is, to our knowledge, the longest found so far in a naturally-occurring protein. The acidic nature of some of these residues might be diminished by post-translational modification of the 17-kd subunit. However, direct amino acid sequence analysis of the analogue of this subunit in the beef complex gives no evidence for this (see below).

Figure 5 shows that homology is found between the predicted sequence of the 17-kd protein and that of the 9-kd, so-called 'hinge' protein of beef complex III (Wakabayashi *et al.*, 1982). A continuous block of acidic residues is found in both proteins, although its length differs in the two organisms. The beef protein has a block of eight glutamates, while the yeast 17-kd subunit has a much longer stretch (25 residues) consisting of both glutamate and aspartate.

Alignment of the amino acid sequences reveals a homology of 35–40% at the carboxy-terminal side of the block of acidic amino acids. This alignment requires assumption of six insertions in the 17-kd polypeptide and four in the 9-kd protein. Of these, six consist of one or two amino acids.

The beef 9-kd protein is predicted to be entirely water-soluble by calculation of the gain in free energy for a water to membrane phase transition, according to Von Heijne (1981) (data not shown). A similar analysis for the precursor of the 17-kd subunit shows that only a short amino-terminal seg-

```
-109  CATATTGGTTGGCGCTTATTTGCACATATTGTATACACGCACTCACATTAACAGGAGCAC

                                                                M   D   M   L
 -49  ACATATACACTTACACCTACACACACGGATAAAGAAAAAGAAATAGAAAATGGACATGTT

        E   L   V   G   E   Y   W   E   Q   L   K   I   T   V   V   P   V   V   A   A
  12  GGAACTAGTTGGTGAGTACTGGGAACAACTAAAGATAACCGTTGTGCCTGTTGTGGCCGC
                          ^
        A   E   D   D   D   N   E   Q   H   E   E   K   A   A   E   G   E   E   K   E
  72  GGCCGAAGATGACGATAACGAGCAGCATGAAGAAAAGGCAGCAGAAGGAGAAGAAAAAGA

        E   E   N   G   D   E   D   E   D   E   D   E   D   E   D   D   D   D   D   D
 132  AGAAGAAAATGGGGATGAAGATGAGGATGAAGACGAAGACGAAGATGATGATGATGATGA

        D   E   D   E   E   E   E   E   E   V   T   D   Q   L   E   D   L   R   E   H
 192  CGACGAAGATGACGAAGAAGAGGAAGAAGTCACTGATCAGTTGGAAGATTTGAGAGAACA

        F   K   N   T   E   E   G   K   A   L   V   H   H   Y   E   E   C   A   E   R
 252  TTTCAAGAACACGGAGGAGGGTAAGGCCCTTGTGCACCACTACGAGGAGTGTGCTGAGAG

        V   K   I   Q   Q   Q   Q   P   G   Y   A   D   L   E   H   K   E   D   C   V
 312  AGTCAAGATACAGCAACAACAACCCGGCTACGCGGATCTTGAACACAAGGAGGACTGTGT

        E   E   F   F   H   L   Q   H   Y   L   D   T   A   T   A   P   R   L   F   D
 372  GGAGGAGTTTTTCCATCTACAGCACTATTTGGACACTGCCACGGCACCTAGATTATTTGA

        K   L   K   *
 432  CAAATTAAAGTAGTCTGTTTATCCGTCATATTCCCTATATTCATATTTATTCTGAGTCTT

 492  GTTACAATTTCTTCAAGAAACCGATCCTTATTTAACCTACTGTATTTAGAAAGTAAATAA

 552  ATAAACTAAAATCAAAGCAAAAATAATTTTATACACTATTCTTATTTTTATTTTTTAATT

 612  TTTAATTTTTTTTATAACGGAGCACAAGTATACAGACGGAAAAGGGATTAGGCTTTTTTT

 672  ATGCGCTCTGTCTAGTT
```

Fig. 4. DNA sequence of the region encoding the 17-kd protein. The transcript maps roughly between sequence positions −130 and −120 and 570–580 of the DNA sequence (see Figure 2). Putative consensus sequences (Dobson *et al.*, 1982) are underlined: the CACACA sequence which is usually found close to the initiation codon in yeast genes occurs twice. The A at −3 and purine at +4, which have been proposed to function in binding of the 40S ribosomal unit (Kozak, 1981) are found, as is also the pyrimidine at +6; the hexanucleotide AATAAA which might function in transcript termination or poly(A) addition in higher eukaryotes (Fitzgerald and Shenk, 1981) is present. Transcript termination or polyadenylation sequences, found by Zaret and Sherman (1982) and Henikoff *et al.* (1983), are absent. Amino acids are indicated in the one-character amino acid code: A, alanine; 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.

Table I. Comparison of the amino acid composition of the 17-kd protein (A) with that predicted from the DNA sequence (B)

Amino acid	A (mol %)*	B (mol %)
Asx	17.8	18.3
Thr	4.0	3.5
Ser	3.4	0.0
Glx	28.60	30.9
Pro	2.9	2.1
Gly	6.0	3.5
Ala	6.0	7.0
Cys	n.d.	–
Val	4.4	6.3
Met	n.d.	–
Ile	1.9	1.4
Leu	6.2	7.8
Tyr	2.6	2.8
Phe	3.3	2.8
Trp	n.d.	–
Lys	6.6	6.3
His	3.8	4.9
Arg	2.9	2.1

Cys, Met and Trp have not been determined (n.d.) and were, therefore, neglected in the calculation of the mol % of amino acids on the basis of the DNA sequence.
*Determined after 24 h hydrolysis of the purified protein (Katan and Groot, 1976).

```
                    1        10        20        30        40
17 kDa    protein   MDMLELVGEYWEQLKITVVPVVAAAEDDDNEQHEEKAAEGEEK
"hinge" protein

                          50        60        70
17 kDa    protein   EEENGDEDEDEDEDEDDDDDDEDEEEEEE-VTD---------
"hinge" protein         GDPK--------------EEEEEEEELV-DPLTTVREQC
                        1                       10         20

                        80         90        100       110
17 kDa    protein   -QLEDL---REHFKNTEEGKALVHHYEECAERVKIQQQQPGYA
"hinge" protein     EQLEKCVKARERL--------------ELCDERVSSRSQT----
                         30                        40        50

                        120        130       140
17 kDa    protein   DLEHKEDCVEEF--FHLQHYLDTATAPRLFDKLK
"hinge" protein     --E--EDCTEELLDF-L-HARDHCVAHKLFNSLK
                                  60         70
```

Fig. 5. Comparison between the predicted sequence of the 17-kd protein and that of the 'hinge' protein of beef complex III. The sequence of the 'hinge' protein has been determined by Wakabayashi *et al.* (1982). Sequences are aligned on the basis of the assumption that insertions and/or deletions (−) have occurred, but that the order of residues has not changed during evolution. Identical amino acids are underlined.

ment of this protein may be membrane-soluble (Figure 6). This part may well be absent from the mature protein, however (see below).

We therefore suggest that the 17-kd and 9-kd polypeptides have similar functions in the yeast and beef complexes, re-

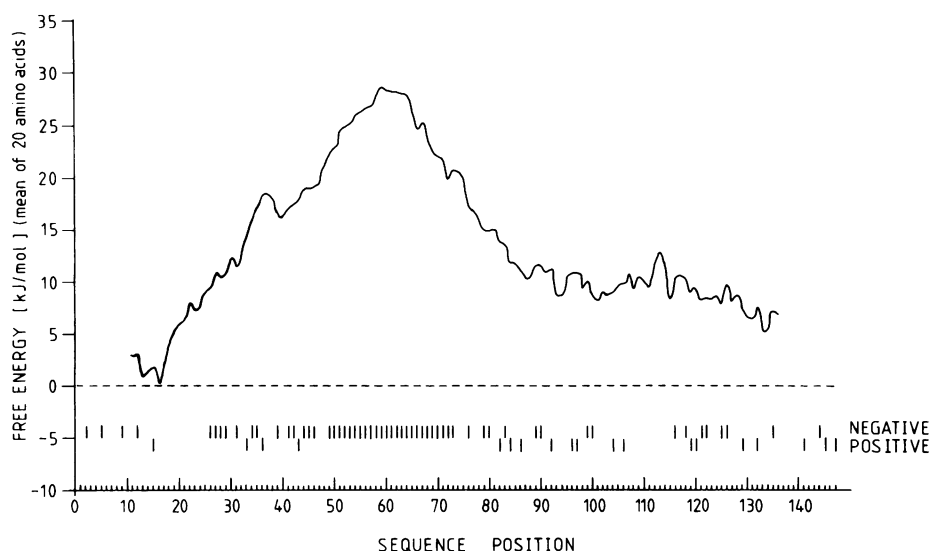

Fig. 6. Predicted distribution of lipophilic segments and charges in the amino acid sequence of the 17-kd protein. The gain in free energy during transition of 20-residue segments from an α-helical structure in water to an α-helix in the membrane is calculated for all sequence positions according to Von Heijne (1981). The choice of segments of 20 amino acids is based on the number of residues required to span the bilayer. The positions of the basic residues Arg, Lys and His (positive charge) and of the acidic residues Glu and Asp (negative charge) are indicated by vertical lines.

spectively. The 9-kd protein is found in purified 'two-band' cytochrome c_1 (Kim and King, 1981) and is thought to mediate the interaction between the cytochromes c_1 and c (King, 1983). A protein with a calculated mol. wt. of 18.5-kd, which might be the 17-kd protein described here, has been found associated with isolated yeast cytochrome c_1 (Ross and Schatz, 1976).

Acidic amino acids are conserved in the 17-kd and 9-kd proteins

Further comparison of the primary sequences of the 17-kd and 9-kd proteins (Figure 5), shows that, apart from a non-interrupted stretch of acidic amino acids, acidic residues are also preferentially conserved at the expense of neutral amino acids in the remainder of these proteins. Acidic residues form 36–37% of the conserved amino acids in this area, while only 22–23% of the non-conserved residues are glutamic or aspartic acid. Basic amino acids occur with similar frequencies in both the conserved and non-conserved parts. Several functions may be proposed for these acidic residues.

(i) They may be responsible for the proposed interaction between cytochromes c and c_1. In cytochrome c_1 most basic residues are clustered within the amino- and carboxy-terminal parts of the mature protein (Wakabayashi *et al.*, 1980). Analysis of the primary (Smith *et al.*, 1979) and of the three-dimensional structure (Dickerson, 1980) of cytochrome c also showed clustering of basic amino acids in this protein. In addition, 'three band' cytochrome c_1 – a subcomplex obtained from purified complex III of beef – contains an additional basic protein and interaction of this with acidic residues in the 9-kd protein has also been proposed (Schägger *et al.*, 1983).

(ii) Clustering of acidic amino acids is likely to change the p*K*s of the individual residues, resulting in a gradient within the acidic area. This part of the protein might, therefore, function as a 'proton sink'.

(iii) A pH-dependent change in secondary structure of the cluster of acidic amino acids might occur. This area is predicted to have an α-helical structure according to the rules of Chou and Fasman (1978). This seems reasonable if the residues are in non-dissociated form. However, after dissociation of the carboxyl groups, a random structure is more likely, due to the repulsion of the charged amino acids. The activity of the polypeptide could, therefore, be regulated by varying the pH in the environment of the protein.

Discrepancy between the lengths determined for the precursor of the 17-kd protein and that predicted from the DNA sequence

In yeast, the 17-kd protein is initially synthesized as a longer precursor which has an apparent mol. wt. of 26-kd, as determined from its migration on SDS-polyacrylamide gels (Van Loon *et al.*, 1983). However, the length of the precursor predicted from the DNA sequence is only 17 394 daltons. The discrepancy is not due to the presence of intervening sequences, since the S1 nuclease mapping of the mRNA excludes existence of extra exon sequences. We suggest that the lengths of both the precursor and of the mature protein have been substantially over-estimated on SDS-polyacrylamide gels. Similar anomalous behaviour is seen for the 9-kd protein. Based on its electrophoretic mobility, a mol. wt. of 15-kd has been calculated, while only 9175 daltons is predicted from the amino acid sequence (Wakabayashi *et al.*, 1982). The discrepancies in both cases are probably attributable to reduced binding of SDS due to the high content of acidic residues.

The amino- and carboxy-terminal sequences of the mature 17-kd yeast protein are not known. Attempts to determine the amino-terminal sequence, using the 17-kd subunit eluted from SDS-polyacrylamide gels have failed, possibly because the amino-terminal residue is blocked (G.J.Steffens, personal communication). It is, therefore, not possible to determine which part of the predicted protein sequence corresponds to the precursor extension. Compared with the beef protein, the yeast protein extends 47 amino acids further at the amino terminus, while the carboxy-termini are very similar (four out of six residues identical). One possible interpretation of this difference is that the additional 47 residues in the yeast protein represent an amino-terminal precursor extension and that the mature protein is only 100 amino acids long. One prediction of this interpretation is that the mature protein should not be

radio-labelled by [^{35}S]methionine, since the only methionine residues occur at positions 1 and 3 of the pre-sequence. Labelling of both precursor and mature protein by [^{35}S]methionine is in fact observed with low efficiency. This might be the result of leakage of ^{35}S into cysteine residues, as observed in the case of subunit VI of yeast cytochrome *c* oxidase. This latter protein also lacks methionine (Gregor and Tsugita, 1982), but nevertheless is still labelled by [^{35}S]methionine (G.Schatz, personal communication).

DNA sequence analysis of the genes encoding subunit 9 of the *Neurospora crassa* ATPase complex (Viebrock *et al.*, 1982) and the yeast cytochrome *c* peroxidase (Kaput *et al.*, 1982) has shown that the precursor extensions of both proteins contain basic but no acidic residues. It has been proposed that these basic amino acids function in transport of the protein across the mitochondrial inner membrane. In contrast, the amino-terminal part of the precursor of the yeast 17-kd protein contains a high number of acidic residues. Complete absence of acidic residues from the amino-terminal part of mitochondrial precursor proteins is thus not a prerequisite for import into the organelle.

Materials and methods

Strain and growth conditions

The *Saccharomyces cerevisiae* strain FL100 was used for DNA sequence analysis and for isolation of total yeast RNA. Cells were grown on lactate medium (Ohnishi *et al.*, 1966).

DNA sequence analysis

Plasmid 209 (see Figure 1) was digested with *Sal*I and *Hind*III and the fragment that contains the entire cloned sequence was isolated from a low-melting agarose gel. After digestion with *Sau*3A, the fragments were ligated into the *Bam*HI site of the bacteriophage M13 mp7 (Messing *et al.*, 1981). The 660-bp *Rsa*I fragment, which contains part of the mRNA-coding sequence (see Results), was purified from agarose gels and ligated into the *Eco*RI site of the bacteriophage by use of synthetic *Eco*RI linkers. Single-stranded DNA was prepared from insert-containing phages (Sanger *et al.*, 1980) and spotted onto nitrocellulose filters. These were incubated for 5 min in 0.5 M NaOH, 1.5 M NaCl, followed by two washes for 3 min each in 0.3 M Tris-HCl (pH 7.7), 3 M NaCl. The filters were air-dried and heated for 2 h at 80°C. They were then hybridized to a ^{32}P-labelled mRNA fraction enriched for the mRNAs coding for the 17-, 14- and 11-kd subunits ('fraction 12') as described (Van Loon *et al.*, 1982). The DNA sequence of fragments that hybridize to the labelled mRNA was determined by use of the dideoxy method (Sanger *et al.*, 1980) as described (De Haan *et al.*, 1984).

Transcript mapping

Fragments ligated into the bacteriophage M13 mp7 were labelled by synthesis of the complementary strand in the presence of α-^{32}P-labelled dNTPs (De Haan *et al.*, 1984). After incubation for 30 min at 65°C the DNA was digested with *Eco*RI, size-fractionated on a 5% polyacrylamide gel and the double-stranded insert fragment eluted into 500 mM ammonium acetate, 1 mM EDTA, 0.1% SDS (pH 8) during 18 h at 37°C. The DNA was alcohol-precipitated, washed once with 70% ethanol and dissolved in 200 μl 10 mM Tris-HCl (pH 7.5), 0.5 mM EDTA. S1 nuclease digestion was carried out according to Berk and Sharp (1977) as described in De Haan *et al.* (1984). $1-2 \times 10^4$ c.p.m. of the isolated fragment were used.

Miscellaneous

Published procedures were used for: (i) isolation of total yeast RNA (Van Loon *et al.*, 1982); (ii) electrophoresis of glyoxal-denatured RNA (McMaster and Carmichael, 1977); (iii) blotting to nitrocellulose filters of RNA (Thomas, 1980) or DNA (Southern, 1975); and (iv) purification of *Sau*3A (Sussenbach *et al.*, 1976).

Other restriction enzymes were purchased from Biolabs, the large fragment of DNA polymerase (Klenow fragment) from Boehringer Mannheim; S1 nuclease from Sigma and T4 polynucleotide kinase from New England Nuclear. Enzymes were used as recommended by the manufacturers. Radioactive materials were obtained from New England Nuclear.

Acknowledgements

We thank Dr G.J.Steffens and Professor G.Buse (Technische Hochschule, Abteilung Physiologische Chemie, Neues Klinikum, Forckenbeckstrasse, D-5100 Aachen, Germany) for their attempts to determine the amino-terminal protein sequence, Dr C.A.M.Marres for help and stimulating discussions, Professor P.Borst and Drs H.F.Tabak, R.Benne and P.Sloof for critical reading of the manuscript and Mrs M.A.ten Bosch-Tilanus for typing the manuscript. This work was supported in part by a grant to P.Borst-L.A.Grivell from The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for the Advancement of Pure Research (ZWO).

References

Berk,A.J. and Sharp,P.A. (1977) *Cell*, **12**, 721-732.

Chevallier,M.-R., Bloch,J.C. and Lacroute,F. (1980) *Gene*, **11**, 11-19.

Chou,P.Y. and Fasman,G.D. (1978) *Adv. Enzymol.*, **47**, 45-148.

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.

Dickerson,R.E. (1980) *Sci. Am.*, **242**, 98-110.

Dobson,M.J., Tuite,M.F., Roberts,N.A., Kingsman,A.J., Kingsman,S.M., Perkins,R.E., Conroy,S.C., Dunbar,B. and Fothergill,L.A. (1982) *Nucleic Acids Res.*, **10**, 2625-2637.

Fitzgerald,M. and Shenk,T. (1981) *Cell*, **24**, 251-260.

Gregor,I. and Tsugita,A. (1982) *J. Biol. Chem.*, **257**, 13081-13087.

Henikoff,S., Kelly,J.D. and Cohen,E.H. (1983) *Cell*, **33**, 607-614.

Kaput,J., Goltz,S. and Blobel,G. (1982) *J. Biol. Chem.*, **257**, 15054-15058.

Katan,M.B. and Groot,G.S.P. (1976) in Bücher,Th., Neupert,W., Sebald,W. and Werner,S. (eds.), *Genetics and Biogenesis of Chloroplasts and Mitochondria*, North-Holland, Amsterdam, pp. 273-280.

Kim,C.H. and King,T.E. (1981) *Biochem. Biophys. Res. Commun.*, **101**, 607-614.

King,T.E. (1983) *Adv. Enzymol.*, **54**, 267-366.

Kozak,M. (1981) *Nucleic Acids Res.*, **9**, 5233-5252.

Kreike,J. (1982) *The Biosynthesis of Complex III in Yeast*, PhD Thesis, Rodopi, Amsterdam.

McMaster,G.K. and Carmichael,G.G. (1977) *Proc. Natl. Acad. Sci USA*, **74**, 4835-4838.

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309-321.

Ohnishi,T., Kawaguchi,K. and Hagihara,B. (1966) *J. Biol. Chem.*, **241**, 1797-1806.

Reed,J. and Hess,B. (1977) *Hoppe-Seyler's Z. Physiol. Chem.*, **358**, 1119-1124.

Ross,E. and Schatz,G. (1976) *J. Biol. Chem.*, **251**, 1991-1996.

Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.

Schägger,H., Von Jagow,G., Borchart,U. and Machleidt,W. (1983) *Hoppe-Seyler's Z. Physiol. Chem.*, **364**, 307-311.

Smith,M., Leung,D.W., Gillam,S., Astell,C.R., Montgomery,D.L. and Hall,B.D. (1979) *Cell*, **16**, 753-761.

Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.

Sussenbach,J.S., Montfoort,C.H., Schiphof,R. and Stobberingh,E.E. (1976) *Nucleic Acids Res.*, **3**, 3193-3202.

Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.

Van Loon,A.P.G.M., De Groot,R.J., Van Eijk,E., Van der Horst,G.T.J. and Grivell,L.A. (1982) *Gene*, **20**, 323-337.

Van Loon,A.P.G.M., Kreike,J., De Ronde,A., Van der Horst,G.T.J., Gasser,S.M. and Grivell,L.A. (1983) *Eur. J. Biochem.*, **135**, 457-463.

Viebrock,A., Perz,A. and Sebald,W. (1982) *EMBO J.*, **1**, 565-571.

Von Heijne,G. (1981) *Eur. J. Biochem.*, **120**, 275-278.

Wakabayashi,S., Matsubara,H., Kim,C.H., Kawai,K. and King,T.E. (1980) *Biochem. Biophys. Res. Commun.*, **97**, 1548-1554.

Wakabayashi,S., Takeda,H., Matsubara,H., Kim,C.H. and King,T.E. (1982) *J. Biochem. (Tokyo)*, **91**, 2077-2085.

Zaret,K.S. and Sherman,F. (1982) *Cell*, **28**, 563-573.

Received on 21 December 1983; revised on 5 March 1984