The EMBO Journal vol.3 no.9 pp.2137-2143, 1984

Sequencing of the nuclear gene for the yeast cytochrome c_1 precursor reveals an unusually complex amino-terminal presequence

Ingrid Sadler, Kitaru Suda[1], Gottfried Schatz[1], Fritz Kaudewitz and Albert Haid

Institut für Genetik und Mikrobiologie, Maria-Ward-Str. 1a, D-8000 München 19, FRG, and [1]Biocenter, University of Basel, CH-4056 Basel, Switzerland

Communicated by G. Schatz

Cytochrome c_1 is a component of the mitochondrial respiratory chain in most eukaryotes. The protein is coded by nuclear DNA, synthesized as a larger precursor outside the mitochondria and then cleaved to the mature form in two successive steps during its import into the mitochondria. We have cloned the structural gene for yeast cytochrome c_1 by functional complementation of a cytochrome c_1-deficient yeast mutant with a yeast genomic library in the yeast-*Escherichia coli* 'shuttle' vector YEp 13. The complete nucleotide sequence of the gene and of its 5'- and 3'-flanking regions was determined. The deduced amino acid sequence of the yeast cytochrome c_1 precursor reveals an unusually long transient amino-terminal presequence of 61 amino acids. This presequence consists of a strongly basic amino-terminal region of 35 amino acids, a central region of 19 uncharged amino acids and an acidic carboxy-terminal region of seven amino acids. This tripartite structure of the presequence resembles that of the precursor of cytochrome c peroxidase and supports a previous suggestion on the import pathways of these two precursors.

Key words: cytochrome c_1/yeast/mitochondria/nucleotide sequence/protein import

Introduction

Studies of mitochondrial protein import in different organisms have revealed the existence of several different import pathways (Schatz and Butow, 1983; Hay *et al.*, 1984). Among these, the pathways for cytochrome b_2, cytochrome c peroxidase and cytochrome c_1 appear to be the most complex ones (Daum *et al.*, 1982; Ohashi *et al.*, 1982). The precursors to these three proteins have unusually long transient presequences of up to 7 kd which are removed in two successive steps. Studies with yeast suggest the following sequence of events: in the first step, the precursor is translocated across both mitochondrial membranes such that at least an amino-terminal domain of several kilodaltons is exposed to the matrix. An amino-terminal part of the presequence is then removed by a chelator-sensitive, matrix-localized protease, yielding a transmembrane intermediate form. A second proteolytic step (presumably at the outer face of the mitochondrial inner membrane) releases the mature protein. Maturation of the cytochrome c_1 precursor is even more complicated since the second proteolytic cleavage requires the presence of heme. Also, mature cytochrome c_1 remains firmly attached to the mitochondrial inner membrane presumably *via* its apolar carboxy terminus (Wakabayashi *et al.*, 1980).

To understand this import pathway at the molecular level, the amino acid sequences of the precursor proteins must be known. The availability of cytochrome c_1-deficient yeast mutants (Schwieters, 1981; Lang and Kaudewitz, 1982) has allowed us to clone the structural gene for the yeast cytochrome c_1 precursor by functional complementation and to determine the nucleotide sequence of this gene. The deduced amino acid sequence of the precursor protein, together with information on the amino-terminal sequence of the mature cytochrome, identifies an unusually long and complex presequence. Its primary structure supports the suggestion (Daum *et al.*, 1982) that import of the cytochrome c_1 precursor proceeds via a transmembrane intermediate whose bulk protrudes into the intermembrane space.

Results and Discussion

Isolation of the yeast gene for cytochrome c_1

The isolation of yeast mutants specifically lacking cytochrome c_1 (Schwieters, 1981; Lang and Kaudewitz, 1982) has opened the way for isolating the cytochrome c_1 gene by functional complementation (Hinnen *et al.*, 1978; Broach *et al.*, 1979). The mutants isolated so far represent two different complementation groups. Mutants of both groups lack respiration and the typical absorption band of cytochrome c_1; however, only some mutants of the second group completely lack any immunodetectable cytochrome c_1 polypeptide. This fact, together with additional observations (Haid *et al.*, in preparation), suggested that mutants of the second group carry a lesion in the structural gene for cytochrome c_1.

One mutant belonging to the second complementation group (mutant n a2006) was carried through appropriate crosses to render it easily transformable and to introduce a non-reversible leu-2 auxotrophic marker. This marker permits the selection of cells transformed by the yeast-*Escherichia coli* 'shuttle' vector YEp 13 since this vector carries the wild-type *LEU-2* gene (Broach *et al.*, 1979). The resulting yeast strain (n a2006/2) was then transformed with a clone bank consisting of *Bam*HI fragments of yeast genomic DNA inserted into the single *Bam*HI site of YEp 13. Transformants were selected based on their ability to grow in the absence of added leucine and screened further for respiratory competence (based on ability to grow on the non-fermentable carbon source glycerol). This two-step screening procedure has already been successfully used to isolate other genes for imported mitochondrial proteins (O'Malley *et al.*, 1982; Saltzgaber-Muller *et al.*, 1983). One of the leucine-independent, respiring transformants was randomly chosen for further characterization.

The following observations indicated that this transformed yeast clone carried the plasmid-borne gene for cytochrome c_1. (i) When the transformant was grown on rich medium (i.e., without selecting for the presence of the plasmid) respiration-competence and leucine-independence were lost in parallel. (ii) When the plasmid (termed YEp 13-41) was re-isolated from the yeast transformant, amplified in *E. coli*, and re-

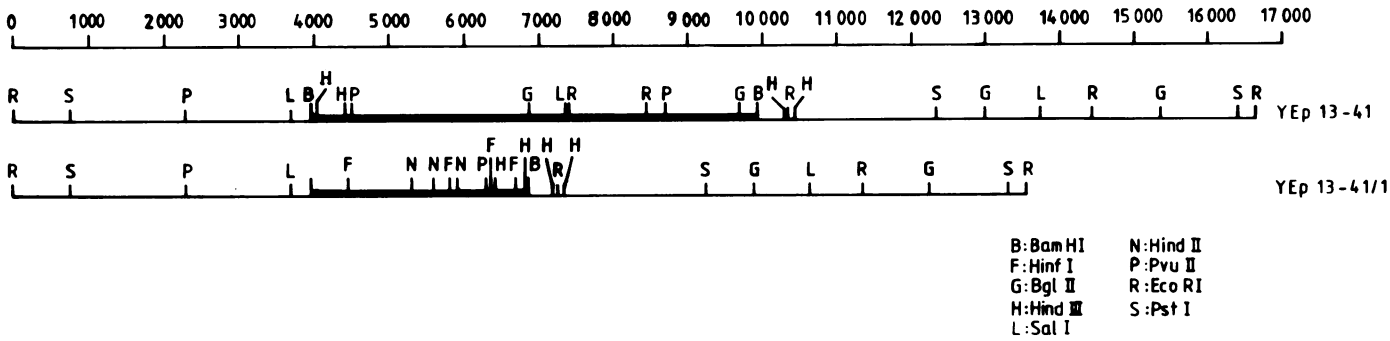

Fig. 1. Restriction maps of the two plasmids YEp 13-41 and YEp 13-41/1 carrying the gene for cytochrome c_1. The insert is indicated by a thick line, the vector YEp 13 by a thin line. The scale on top indicates the number of base pairs.

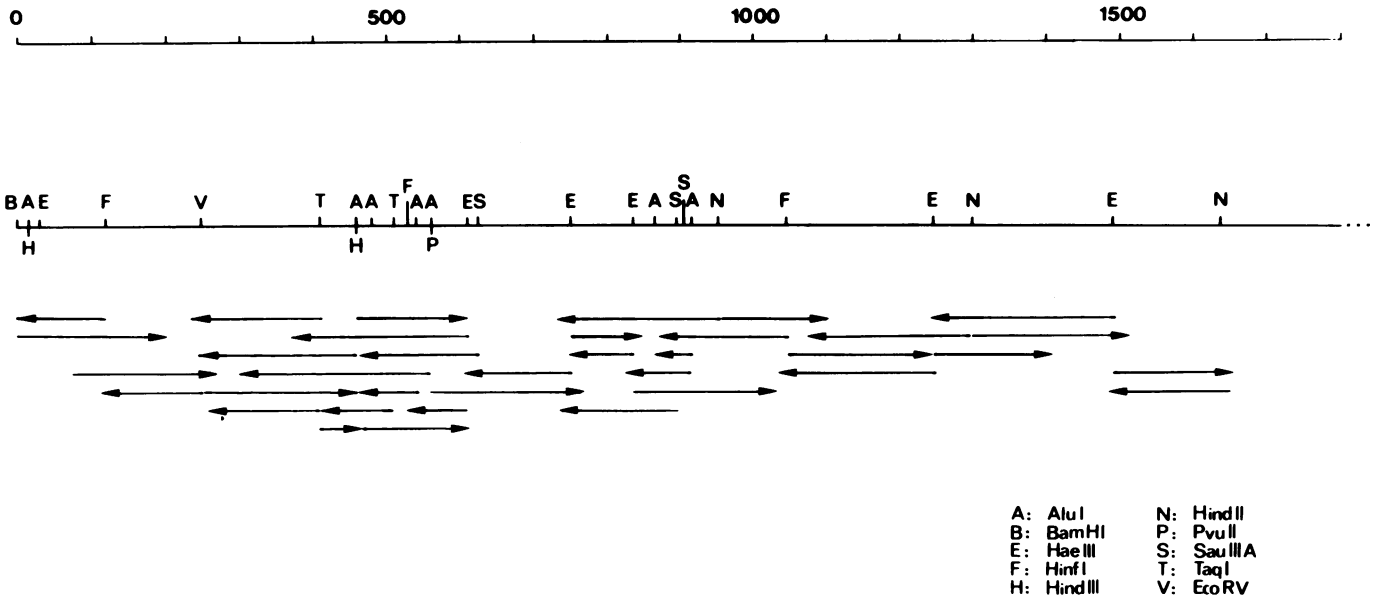

Fig. 2. Sequencing strategy. The DNA insert present in YEp 13-41/1 was prepared and cut into smaller pieces with the indicated restriction enzymes. The resulting fragments were cloned into the phage M13 derivatives mp8 and mp9, either by 'shotgun' cloning (*Sau*3A, *Alu*I, *Hae*III and *Taq*I) or by directed cloning. The top bar indicates the fragment sizes in base pairs, the middle bar the restriction sites used for subcloning, and the arrows at the bottom the direction and extent of sequencing.

introduced into the various cytochrome c_1-deficient yeast mutants, only mutants of the second complementation group always yielded respiring transformants. (iii) When plasmid YEp 13-41 was challenged with total yeast mRNA, it hybrid-selected an mRNA which directed the synthesis of yeast cytochrome c_1 precursor by a nuclease-treated reticulocyte lysate. The product was identified by immunoprecipitation with a monospecific rabbit antiserum and by its mobility upon SDS-polyacrylamide gel electrophoresis (cf. Suissa *et al.*, 1984). (iv) When plasmid YEp 13-41 was introduced into yeast cells carrying a wild-type cytochrome c_1 gene in the nucleus, the transformed cells accumulated increased levels of mature cytochrome c_1 as judged by immune blotting. Since YEp 13-derived plasmids exist in yeast in multiple copies (Botstein and Davis, 1982), the increased steady-state levels of cytochrome c_1 are most plausibly explained by an increased copy number of the corresponding structural gene.

The results strongly indicated that plasmid YEp 13-41 harbors the structural gene for the yeast cytochrome c_1 precursor; they are not documented in detail since the nucleotide sequence of the gene (cf. below) convincingly proves this conclusion.

Characterization of the plasmid carrying the cytochrome c_1 gene

Plasmid YEp 13-41 carried a single *Bam*HI insert of ~6 kb whose restriction map is given in Figure 1. When each of the two fragments obtained upon cleaving the insert with *Bgl*II was recloned into the *Bam*HI site of YEp 13, only one of the two fragments (that shown on the left in Figure 1) restored respiration to the yeast mutants of the second complementation group. The plasmid carrying this 2.9-kb *Bam*HI/*Bgl*II fragment was denoted YEp 13-41/1 (Figure 1). The orientation of the fragment in YEp 13-41 proved to be opposite to that in YEp 13-41/1 (not shown). Since the orientation of the DNA fragment does not affect its ability to complement the cytochrome c_1-deficient yeast mutants, the fragments probably carry all the necessary *cis*-acting information for controlling expression of the cytochrome c_1 gene.

Nucleotide sequence of the cytochrome c_1 gene

The 2.9-kb *Bam*HI/*Bgl*II fragment carried by plasmid YEp 13-41/1 was cut into smaller pieces with different restriction enzymes; individual pieces were cloned into the phage M13 derivatives mp8 and mp9 and sequenced by the dideoxy

```
ATCCATAGACTATCTAAGCTTTTAGTAGAGGCCATTCGTTTTCTTATTAATCATCGGTTTCACAATCTTTCTTTGTTAAG
-366

CAACATTTTATACATTTTTGTACTTTATCTTCAAATATTAAGGACAAGAAAGAAAAAGAAAAAGAATCCATATTCTATAT

TAGTGTAACACTATTTTTCTTAAATATTTTTTCCAGTACTCCCCCAGTTTGATAAAAAGCATTCATTGAATTTCCCCACA

TTTGTGATATCTTCAATTGATTAGTTTGAACTAGTTCTGAAAAATAATATTTTACAATTTGCATTTTCATTACACTATAT

                                                    1
CATCTACTATTTTTTCTCAGAAGCGGAAGTTATAACTAATTTGACA MET PHE SER ASN LEU SER LYS ARG
                                               -1 ATG TTT TCA AAT CTA TCT AAA CGT
                                                  +1
     10
 TRP ALA GLN ARG THR LEU SER LYS SER PHE TYR SER THR ALA THR GLY ALA ALA SER LYS
 TGG GCT CAA AGG ACC CTC TCG AAA AGT TTC TAC TCT ACC GCA ACA GGT GCT GCT AGT AAA

     30
 SER GLY LYS LEU THR GLN LYS LEU VAL THR ALA GLY VAL ALA ALA ALA GLY ILE THR ALA
 TCT GGC AAG CTT ACT CAA AAG CTA GTT ACA GCG GGT GTT GCT GCC GCC GGT ATC ACC GCA

     50
 SER THR LEU LEU TYR ALA ASP SER LEU THR ALA GLU ALA MET THR ALA ALA GLU HIS GLY
 TCG ACT TTA CTC TAT GCA GAC TCA TTA ACT GCC GAA GCT ATG ACC GCA GCT GAA CAC GGA

     70
 LEU HIS ALA PRO ALA TYR ALA TRP SER HIS ASN GLY PRO PHE GLU THR PHE ASP HIS ALA
 TTG CAC GCC CCA GCA TAT GCT TGG TCC CAC AAT GGG CCT TTT GAA ACA TTT GAT CAT GCA

     90
 SER ILE ARG ARG GLY TYR GLN VAL TYR ARG GLU VAL CYS ALA ALA CYS HIS SER LEU ASP
 TCC ATT AGA AGA GGT TAC CAG GTT TAC CGT GAA GTT TGT GCC GCC TGC CAT TCT CTT GAC

     110
 ARG VAL ALA TRP ARG THR LEU VAL GLY VAL SER HIS THR ASN GLU GLU VAL ARG ASN MET
 AGA GTT GCT TGG AGA ACT TTG GTT GGT GTT TCT CAT ACC AAC GAA GAG GTT CGT AAT ATG

     130
 ALA GLU GLU PHE GLU TYR ASP ASP GLU PRO ASP GLU GLN GLY ASN PRO LYS LYS ARG PRO
 GCC GAA GAA TTT GAA TAC GAT GAC GAA CCT GAT GAA CAA GGT AAC CCT AAA AAG AGA CCA

     150
 GLY LYS LEU SER ASP TYR ILE PRO GLY PRO TYR PRO ASN GLU GLN ALA ALA ARG ALA ALA
 GGT AAG TTG TCC GAT TAC ATC CCT GGC CCA TAC CCA AAC GAA CAG GCT GCA AGA GCT GCT

     170
 ASN GLN GLY ALA LEU PRO PRO ASP LEU SER LEU ILE VAL LYS ALA ARG HIS GLY GLY CYS
 AAT CAA GGT GCC TTG CCA CCT GAT CTA TCT TTG ATC GTG AAA GCT AGA CAC GGT GGT TGT

     190
 ASP TYR ILE PHE SER LEU LEU THR GLY TYR PRO ASP GLU PRO PRO ALA GLY VAL ALA LEU
 GAC TAC ATT TTC TCT TTG TTG ACC GGT TAT CCT GAT GAA CCT CCT GCT GGT GTG GCT TTA

     210
 PRO PRO GLY SER ASN TYR ASN PRO TYR PHE PRO GLY GLY SER ILE ALA MET ALA ARG VAL
 CCA CCA GGT TCT AAT TAT AAC CCT TAC TTC CCA GGT GGT TCC ATT GCA ATG GCA AGA GTC

     230
 LEU PHE ASP ASP MET VAL GLU TYR GLU ASP GLY THR PRO ALA THR THR SER GLN MET ALA
 TTG TTT GAT GAC ATG GTT GAG TAC GAA GAT GGT ACC CCC GCA ACG ACA TCT CAA ATG GCA

     250
 LYS ASP VAL THR THR PHE LEU ASN TRP CYS ALA GLU PRO GLU HIS ASP GLU ARG LYS ARG
 AAG GAC GTT ACC ACC TTT TTA AAC TGG TGT GCC GAA CCT GAA CAT GAC GAA AGA AAG AGA

     270
 LEU GLY LEU LYS THR VAL ILE ILE LEU SER SER LEU TYR LEU LEU SER ILE TRP VAL LYS
 TTG GGT TTG AAA ACG GTG ATA ATC TTA TCA TCT TTG TAT TTG CTA TCT ATC TGG GTG AAG

     290
 LYS PHE LYS TRP ALA GLY ILE LYS THR ARG LYS PHE VAL PHE ASN PRO PRO LYS PRO ARG
 AAG TTC AAA TGG GCC GGT ATC AAA ACC AGA AAA TTC GTT TTC AAT CCA CCA AAA CCA AGA

 309
 LYS *** TTGACAAAAAAAATGAAAACCACTATTATTTAAATTACGGAAAAAAAAAAAACCAATAAACAAAAAACATAAT
 AAG TAG
      +930

AAAAGAAATGAAAAAAAAGAATATTAGTAGAGCATACTGAATTGTTTTCAGAAAGAGAGTGAAAGGTAACAAAAATTTAC

CTACTTTTCCTTTTTATTTCTCTGGTTTCGTTTCTTTTTTTTCGGCCACCCTATTTATTTTTTCACTGCTTGTAAACAAC

GTCTCACTTCAATAGTGTCTCCTTCATGTACTGAGATACTGCTGC
                                              +1572
```

Fig. 3. Nucleotide sequence of the cytochrome c_1 gene and deduced amino acid sequence of the yeast cytochrome c_1 precursor protein. Numbers underneath the lines refer to the nucleotide sequence and numbers on top of the lines to the amino acid sequence. The arrow indicates the border between the transient presequence and the mature sequence.

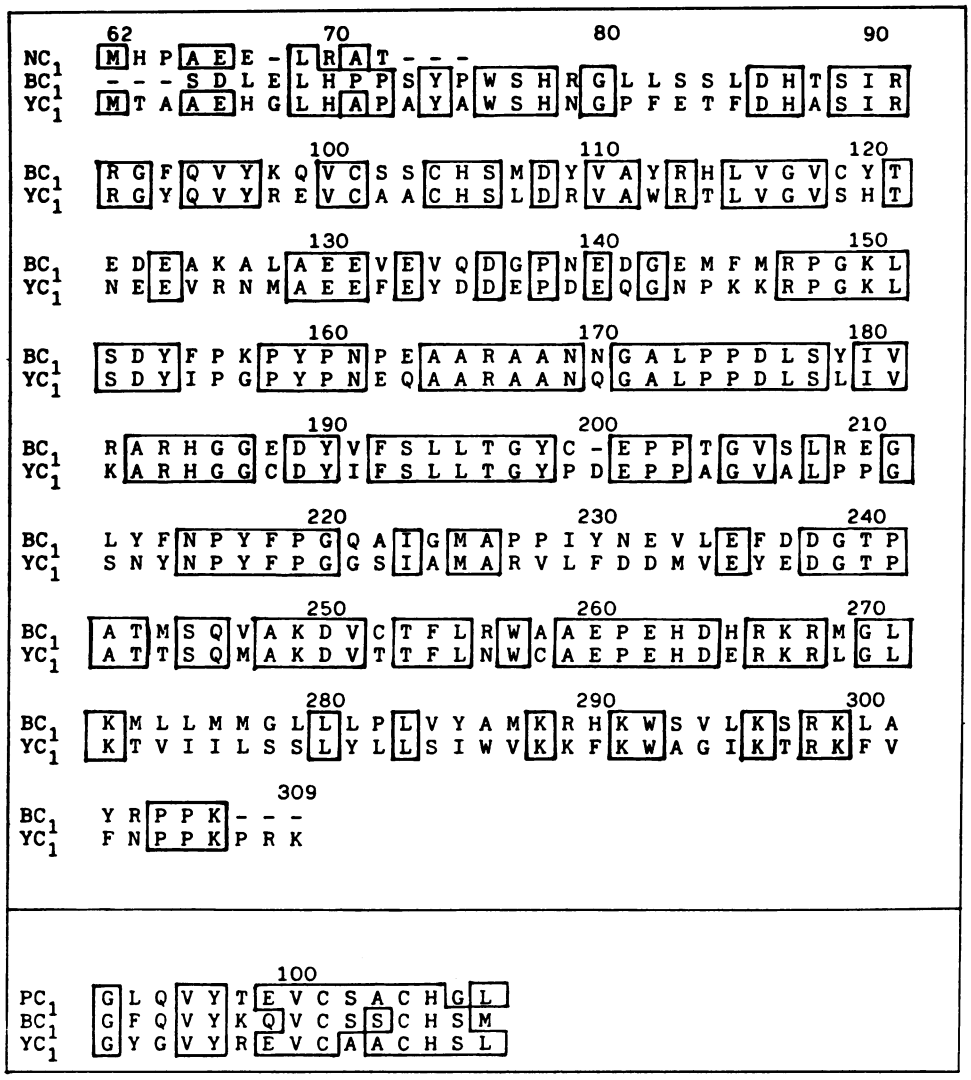

Fig. 4. Comparison between the amino acid sequences of mature yeast cytochrome c_1 (YC_1) (this paper), of mature bovine cytochrome c_1 (BC_1) (Wakabayashi *et al.*, 1980) and the amino terminus of *N. crassa* mature cytochrome c_1 (NC_1) (Tsugita *et al.*, 1979). The lower panel of the Figure compares the heme-binding domains of the cytochrome c_1 species from yeast and beef-heart (cf. above) with that of *Paracoccus denitrificans* cytochrome c_1 (PC_1) (Ludwig *et al.*, 1983). Regions of homology are boxed. Positions are numbered starting from the amino terminus of the yeast cytochrome c_1 precursor.

chain-termination method (Figure 2). Once the first partial sequences were available, we used the published amino acid sequence of bovine cytochrome c_1 (Wakabayashi *et al.*, 1980) to locate the cytochrome c_1 gene within the left region of the 2.9-kb insert. From then on, we concentrated our sequencing efforts on that region. Figure 3 lists the nucleotide sequence of the gene, the nucleotide sequence of the 5'- and 3'-flanking regions and the deduced amino acid sequence of the yeast cytochrome c_1 precursor protein. The coding region consists of 927 nucleotides, corresponding to a precursor polypeptide of 309 amino acids (mol. wt. 34 016). Codon usage (54 out of 61 possible codons) is that of a weakly expressed yeast gene (Bennetzen and Hall, 1982); this is in line with the low abundance of cytochrome c_1 mRNA in most yeast strains (A. Ohashi, unpublished). Starting at amino acid 62, the deduced amino acid sequence of the mature yeast cytochrome c_1 is 56.5% homologous to that which had been determined by amino acid sequence analysis of mature bovine cytochrome c_1 (Figure 4; Wakabayashi *et al.*, 1980). This argues against the presence of a long intron in the DNA sequence coding for the 'mature' region of yeast cytochrome c_1. The amino terminus of mature yeast cytochrome c_1 has recently been determined by amino acid analysis of purified yeast cytochrome c_1 (van Loon, personal communication); as shown in Figure 4, this amino terminus is closely homologous to that of mature cytochrome c_1 from *Neurospora crassa* (Tsugita *et al.*, 1979). Thus, the presequence of the yeast cytochrome c_1 precursor contains 61 amino acids. The mol. wt. of the presequence (6597) agrees well with the earlier estimate of 6000 which had been based on SDS-polyacrylamide gel electrophoretic analysis of immunoprecipitated precursor and mature form of yeast cytochrome c_1 (Ohashi *et al.*, 1982). Again this concordance argues against the presence of long introns in the corresponding DNA sequence; the absence of typical yeast splicing signals (Mount, 1982; Langford *et al.*, 1984) in the entire cytochrome c_1 coding sequence further supports this view. Compared with bovine cytochrome c_1, yeast cytochrome c_1 contains three additional amino acids at the amino terminus and three additional amino acids at the carboxy terminus. It cannot be excluded, however, that the three 'extra' carboxy-terminal amino acids are lost during maturation of the precursor (e.g., Suchanek and Kreil, 1977). Yeast cytochrome c_1 also contains a characteristic sequence for the covalent attachment of heme (Cys_{101}-

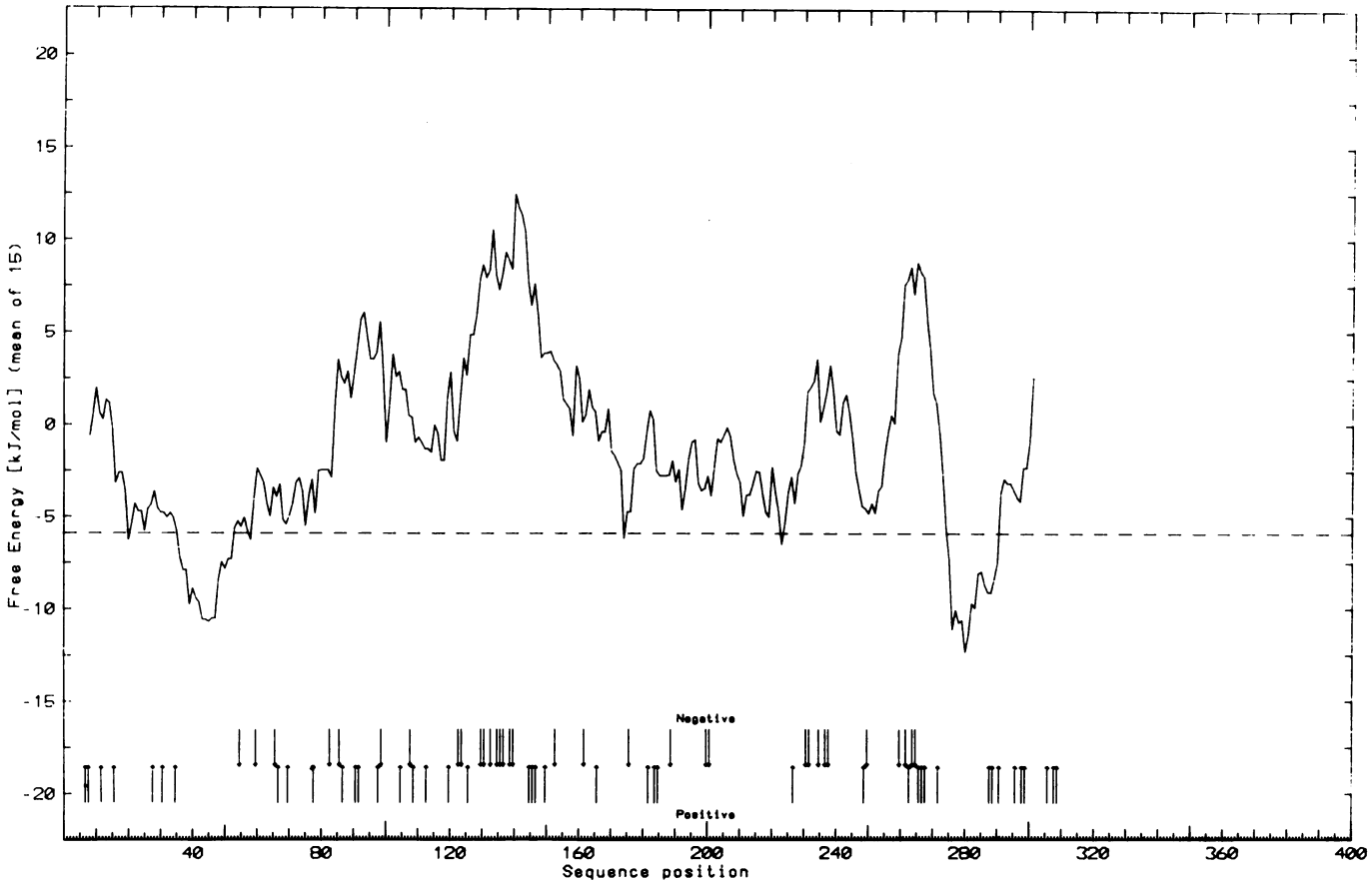

Fig. 5. Distribution of charged and lipophilic amino acid residues within the yeast cytochrome c_1 precursor. The gain of free energy during transition of a 15-residue segment from a random coil in water to a helix in the membrane is calculated for all sequence positions according to von Heijne (1981). The position of basic residues (positive charge) and acidic residues (negative charge) is indicated by arrows, the mean hydrophobicity per residue (−5.9 kJ/mol) by the dotted line.

Ala-Ala-Cys-His) near the amino terminus of the mature cytochrome; this sequence is similar to that in bovine and *Paracoccus denitrificans* cytochrome c_1 (Figure 4; Ludwig *et al.*, 1983). In addition, yeast cytochrome c_1 contains an uninterrupted stretch of 15 uncharged amino acids near the carboxy terminus. This uncharged region is flanked by positively charged amino acids and thus resembles a membrane-spanning polypeptide segment. It has already been suggested that this segment might anchor the mature cytochrome to the mitochondrial inner membrane (Wakabayashi *et al.*, 1980). The distribution of charged and lipophilic amino acids within the yeast cytochrome c_1 precursor is depicted in Figure 5. The most interesting feature of the precursor sequence is clearly the amino-terminal transient presequence. It will be considered in the following section.

The 366 bp of non-coding sequence upstream from the initiating ATG codon contains several eukaryotic consensus sequences which may function in the initiation of transcription and translation and perhaps other processes as well. However, the upstream region lacks any sequences resembling those which have been implicated in the heme and catabolite regulation of the yeast iso-1-cytochrome c gene (Guarente and Mason, 1983; Guarente *et al.*, 1984).

The presequence

The transient presequences of mitochondrial precursor polypeptides are essential for binding these polypeptides to the mitochondrial surface (Riezman *et al.*, 1983a) and for translocating them into the mitochondria (Gasser *et al.*, 1982). However, it is not known whether these presequences carry information on the intramitochondrial destination of the corresponding mature polypeptide.

This question is particularly intriguing for those proteins which, like pre-cytochrome c_1, reach their final location only after a 'detour' to the matrix and after two successive proteolytic cleavages (Ohashi *et al.*, 1982). This question could perhaps be answered by fusing the presequence of cytochrome c_1 to the mature sequence of a protein that is normally transported to another intramitochondrial location. Such studies are underway.

Like most mitochondrial presequences (Reid, 1984), that of pre-cytochrome c_1 is strongly basic (seven basic and no acidic amino acids within the first 54 residues) and rich in threonine and serine. It contains two uninterrupted stretches of uncharged amino acids. The first, an 11-residue stretch starting at Ser_{17} is probably too short to span a membrane. The second, a 19-residue stretch starting at Leu_{36}, could be a transmembrane segment. These features suggest that the precursor can assume a transmembrane orientation which is schematically illustrated in Figure 6. Such an orientation is in agreement with the two-step import model that had been suggested earlier, based on *in vitro* experiments as well as pulse-chase studies with intact yeast cells (Ohashi *et al.*, 1982; Gasser *et al.*, 1982). Similar observations have been made in *Neurospora crassa* (Teintze *et al.*, 1982). The structural information now available adds further details to this model

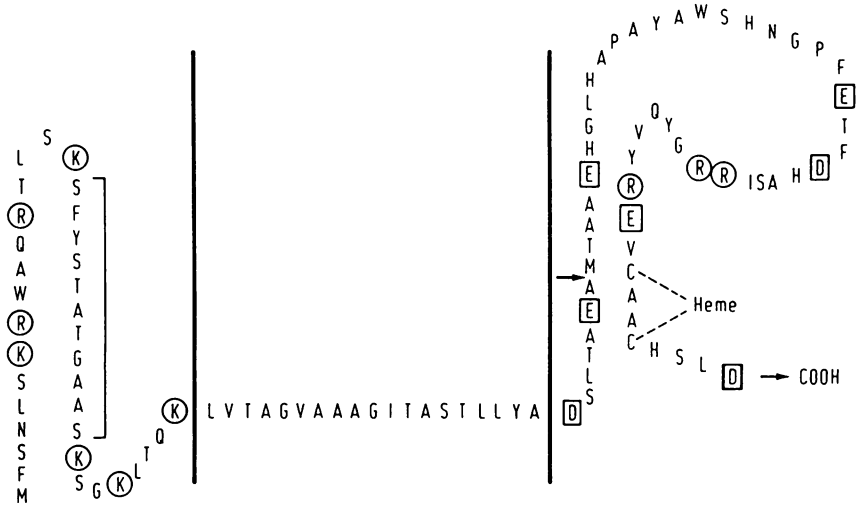

Fig. 6. Suggested orientation of the yeast cytochrome c_1 precursor across the mitochondrial inner membrane. Basic and acidic amino acids are indicated by circles and squares, respectively. The two vertical lines outline the boundaries of the mitochondrial inner membrane, the arrow the second cleavage site, and the vertical bracket the probable region of the first cleavage by the chelator-sensitive matrix protease. The initial part of the mature sequence is arbitrarily arranged such that the heme-binding sites are in close vicinity to the second proteolytic cleavage site.

(Figure 6). We suggest that the cytochrome c_1 precursor polypeptide initially translocates across both mitochondrial membranes such that its 35 amino-terminal residues protrude into the matrix, residues 36–54 are within the inner membrane, and residues 55–309 protrude into the intermembrane space. The matrix-localized, chelator-sensitive protease (Böhni *et al.*, 1980; McAda and Douglas, 1982) then removes part of the amino terminus; the cleavage site is unknown, but is probably within the region indicated by brackets since cleavage reduces the mol. wt. by ~2000 (Ohashi *et al.*, 1982). The resulting transmembrane intermediate form then binds one molecule of protoheme IX via thioether bridges to cysteine residues 101 and 104. Attachment of heme induces a conformational change in the polypeptide which renders the Ala_{61}-Met_{62} bond susceptible to cleavage by a second protease.

The transmembrane orientation of pre-cytochrome c_1 suggested by the amino acid sequence is strikingly similar to that which had been suggested by Kaput *et al.* (1982) for the precursor of cytochrome c peroxidase. The two precursors have several characteristics in common: (i) the corresponding mature proteins are hemoproteins functioning at the outer face of the mitochondrial inner membrane, (ii) both precursors are cleaved in two successive steps, the first of them catalyzed by the matrix-localized protease, (iii) their presequences are unusually long (61 and 66 residues for pre-cytochrome c_1 and pre-cytochrome c peroxidase, respectively) and have at least one stretch of uncharged amino acids sufficiently long to span a membrane. In contrast, the equally long presequence of ATPase subunit IX of *N. crassa* (a protein imported into the inner membrane) lacks any recognizable transmembrane sequence (Viebrock *et al.*, 1982).

The import mechanisms and the transmembrane orientations of pre-cytochrome c_1 and pre-cytochrome c peroxidase discussed above are still largely speculative. However, it is remarkable that the presequences of the two precursors have closely similar domain structures even though their amino acid sequences are not significantly homologous. This fact adds further weight to our earlier proposal on how these polypeptides reach their correct intramitochondrial location.

Materials and methods

Strains and growth media

The cytochrome c_1-deficient strain n a2006 (*a, pet⁻, ura 1⁻*) as well as the other strains of the two complementation groups (Schwieters, 1981; Lang and Kaudewitz, 1982) had been derived by mutagenesis from the respiratory-competent strain SM 11-6c (*a, ura 1⁻, met⁻*). Crossing n a2006 to AH 33 (*α, leu 2-3, leu 2-112, his 4⁻, ade 2⁻*) yielded the strain n a2006/2 (*a, pet⁻, leu 2-3, leu 2-112, his 4⁻, ura 1⁻, ade 2⁻*).

For constructing the yeast genomic library, strain D 273-10 B (α; ATCC 25657) was used. The strains were grown at 30°C in rich medium (1% yeast extract, 2% peptone) containing either 2% glucose (YPD) or 3% glycerol and 1% ethanol (YPEG), or in minimal medium (SD; 0.67% Difco yeast nitrogen base, 2% glucose) with or without 20 μg L-leucine/ml. The *E. coli* strain SF 8 (*lop11, lig, supE, recB⁻C⁻, rk⁻, mk⁻*) was used as a host for plasmid transformation and the *E. coli* strain JM 101 (Kedes *et al.*, 1975) for growth of M13-derived phages. Both *E. coli* strains were grown in LB-medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl); for growing cells transformed with YEp 13-derived plasmids, the medium was supplemented with 100 μg ampicillin/ml. Solid media contained 1.5% agar (*E. coli*) or 2% agar (yeast).

Preparation of the yeast genomic clone bank

DNA was isolated from the *Saccharomyces cerevisiae* strain D 273-10B (Cryer *et al.*, 1975) and partially digested with *Bam*HI. Size-selected fragments (10–25 kb) were ligated into the single *Bam*HI site of the yeast-*E. coli* 'shuttle' vector YEp 13 (Broach *et al.*, 1979) and the ligation mixture was used to transform *E. coli* strain SF 8. The resulting ampicillin-resistant transformants were pooled and plasmid DNA was isolated from them by the method of Birnboim and Doly (1979) modified for large-scale preparation.

Miscellaneous methods

Published methods were used for the small-scale isolation of plasmid DNA from transformed yeast cells (Nasmyth and Reed, 1980) and *E. coli* (Birnboim and Doly, 1979) as well as for the transformation of yeast (Hinnen *et al.*, 1978) and *E. coli* (Mandel and Higa, 1970). For selecting plasmids carrying the yeast cytochrome c_1 gene, the yeast mutant n a2006/2 was transformed as mentioned above with the following modifications: (i) yeast cells from a 500 ml culture and ~40 μg of pooled plasmid DNA were used; (ii) the regeneration agar containing the transformed yeast cells was first poured into empty Petri dishes and incubated at 30°C; after colonies had started to grow (usually after 3 days), the agar was dispersed into 100 ml of SD-medium by forcing it through a hypodermic syringe. This suspension was shaken for 1 h at 30°C and aliquots were then plated onto SD-plates.

The following methods were performed according to published reports: isolation of replicative forms of phage M13 derivatives mp8 and mp9 and cloning of DNA fragments to be sequenced into mp8 and mp9 (Messing *et al.*, 1981), DNA sequencing by the dideoxy chain termination method (Sanger *et al.*, 1977, 1980), isolation of yeast mRNA (Maccechini *et al.*, 1979), hybrid-selected translation (Suissa *et al.*, 1984), immunoprecipitation and SDS-polyacrylamide gel electrophoresis (Gasser, 1984), fluorograpy (Chamberlain,

1979), determination of protein (Lowry *et al.*, 1951), protein extraction from whole yeast cells (Riezman *et al.*, 1983b), immune blotting (Towbin *et al.*, 1979; Haid and Suissa, 1983) and preparation of rabbit antisera against yeast cytochrome c_1 (Ohashi *et al.*, 1982).

Acknowledgements

We wish to thank Marlene Redl, Wolfgang Oppliger and Stefanie Smit for excellent technical assistance, Drs. Graeme Reid, Toshiharu Hase and Mordechai Suissa for valuable advice, Dr. Albert Hinnen for samples of the yeast and *E. coli* strains, Dr. Mordechai Suissa for his help in preparing the yeast genomic DNA bank and Dr. Walter Sebald for supplying the lipophilicity plot. This study was supported by grants from the Deutsche Forschungsgemeinschaft (HA 1297/1-1, to A. Haid) and from the Swiss National Science Foundation (3.394-0.83 and 3.037-1.81, to G. Schatz), by an EMBO long-term fellowship (to A. Haid) and by short-term fellowships from EMBO and the Deutsche Akademische Austauschdienst (to I. Sadler).

References

Bennetzen,J.L. and Hall,B.D. (1982) *J. Biol. Chem.*, **257**, 3026-3031.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.
Böhni,P.C., Gasser,S.M., Leaver,C.J. and Schatz,G. (1980) in Kroon,A.M. and Saccone,C. (eds.), *Organization and Expression of the Mitochondrial Genome*, North Holland/Elsevier, Amersterdam, pp. 423-433.
Botstein,D. and Davis,R.W. (1982) in Strathern,J.N., Jones,E.W. and Broach,J.R. (eds.), *The Molecular Biology of the Yeast Saccharomyces cerevisiae*, Cold Spring Harbor Laboratory Press, NY, pp. 607-636.
Broach,J.R., Strathern,J.N. and Hicks,J.B. (1979) *Gene*, **8**, 121-133.
Chamberlain,J.P. (1979) *Anal. Biochem.*, **98**, 132-135.
Cryer,D.R., Eccleshall,R. and Marmur,J. (1975) *Methods Cell Biol.*, **12**, 39-44.
Daum,G., Gasser,S.M. and Schatz,G. (1982) *J. Biol. Chem.*, **257**, 13075-13080.
Gasser,S.M. (1983) *Methods Enzymol.*, **97**, 329-336.
Gasser,S.M., Daum,G. and Schatz,G. (1982) *J. Biol. Chem.*, **257**, 13034-13041.
Guarente,L. and Mason,T.L. (1983) *Cell*, **32**, 1279-1286.
Guarente,L., Lalonde,B., Gifford,P. and Alani,E. (1984) *Cell*, **36**, 503-511.
Haid,A. and Suissa,M. (1983) *Methods Enzymol.*, **96**, 192-205.
Hay,R., Böhni,P.C. and Gasser,S.M. (1984) *Biochim. Biophys. Acta.*, **779**, 65-87.
Hinnen,A., Hicks,J.B. and Fink,G.R. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 1929-1933.
Kaput,J., Goltz,S. and Blobel,G. (1982) *J. Biol. Chem.*, **257**, 15054-15058.
Kedes,L.H., Chang,A.C.Y., Houseman,D. and Cohen,S.N. (1975) *Nature*, **255**, 533-538.
Lang,B.F. and Kaudewitz,F. (1982) *Curr. Genet.*, **6**, 229-235.
Langford,C.L., Klinz,F.-J., Donath,C. and Gallwitz,D. (1984) *Cell*, **36**, 645-653.
Lowry,O.H., Rosebrough,N.J., Farr,A.L. and Randall,R.J. (1951) *J. Biol. Chem.*, **193**, 265-275.
Ludwig,B., Suda,K. and Cerletti,N. (1983) *Eur. J. Biochem.*, **137**, 597-602.
Maccechini,M.-L., Rudin,Y., Blobel,G. and Schatz,G. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 343-347.
Mandel,M. and Higa,A. (1970) *J. Mol. Biol.*, **53**, 159-162.
McAda,P. and Douglas,M.G. (1982) *J. Biol. Chem.*, **257**, 3177-3182.
Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309-321.
Mount,S. (1982) *Nucleic Acids Res.*, **10**, 459-472.
Nasmyth,K. and Reed,S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2119-2123.
Ohashi,A., Gibson,J., Gregor,I. and Schatz,G. (1982) *J. Biol. Chem.*, **257**, 13042-13047.
O'Malley,K., Pratt,P., Robertson,J., Lilly,M. and Douglas,M.G. (1982) *J. Biol. Chem.*, **257**, 2097-2103.
Reid,G.A. (1984) in Martonosi,A. (ed.), *Current Topics in Membranes and Transport*, Academic Press, NY, in press.
Riezman,H., Hay,R., Witte,C., Nelson,N. and Schatz,G. (1983a) *EMBO J.*, **2**, 1113-1118.
Riezman,H., Hase,T., van Loon,A.P.G.M., Grivell,L.A., Suda,K. and Schatz,G. (1983b) *EMBO J.*, **2**, 2161-2168.
Saltzgaber-Muller,J., Kunapuli,S.P. and Douglas,M.G. (1983) *J. Biol. Chem.*, **258**, 11465-11470.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Sanger,F., Coulson,A.R., Barrell,B.F., Smith,A.Z.H. and Roe,B. (1980) *J. Mol. Biol.*, **143**, 161-178.
Schatz,G. and Butow,R.A. (1983) *Cell*, **32**, 316-318.
Schwieters,M. (1981) Ph.D. Thesis, University of Munich, FRG.
Suchanek,G. and Kreil,G. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 975-978.
Suissa,M., Suda,K. and Schatz,G. (1984) *EMBO J.*, **3**, in press.
Teintze,M., Slaughter,M., Weiss,H. and Neupert,W. (1982) *J. Biol. Chem.*, **257**, 10364-10371.
Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354.
Tsugita,A., Gregor,I., Kubota,I. and v.d. Broek,R. (1979) in King,T.E., Orii,Y., Chance,B. and Okunuki,K. (eds.), *Cytochrome Oxidase*, Elsevier/North Holland, NY, pp. 67-77.
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.

Received on 30 May 1984; revised on 14 June 1984