The EMBO Journal Vol.1 No.5 pp.565-571, 1982

The imported preprotein of the proteolipid subunit of the mitochondrial ATP synthase from *Neurospora crassa*. Molecular cloning and sequencing of the mRNA

Adelheid Viebrock, Angela Perz, and Walter Sebald*

Department of Stoffwechselregulation, GBF-Gesellschaft für Biotechnologische Forschung mbH., Mascheroder Weg 1, D-3300 Braunschweig-Stöckheim, FRG

Communicated by K. Beyreuther
Received on 6 May 1982

The proteolipid subunit of the mitochondrial ATP synthase from *Neurospora crassa* is an extremely hydrophobic protein of 81 amino acid residues, which is imported into mitochondria as a precursor of mol. wt. 15 000. The primary structure of the imported form has now been determined by isolating and analyzing cDNA clones of the preproteolipid mRNA. An initial cDNA clone was identified by hybridizing total polyadenylated RNA to pooled cDNA recombinant plasmids from an ordered clone bank and subsequent cell-free translation of hybridization-selected mRNA. Further preproteolipid clones were identified at a frequency of 0.2% by colony filter hybridization. One isolated cDNA represented the major part of the preproteolipid mRNA. The nucleotide sequence showed 243 bases corresponding to the mature proteolipid and, in addition, 178 bases coding for an amino-terminal presequence. Non-coding sequences of 48 bases at the 5' end and of 358 bases at the 3' end plus a poly(A) tail were determined. The long presequence of 66 amino acids is very polar, in contrast to the lipophilic mature proteolipid, and includes 12 basic and no acidic side chains. It is suggested that the presequence is specifically designed to solubilize the proteolipid for post-translational import into the mitochondria.

Key words: cloned mRNA/mitochondrial ATP synthase/nucleotide sequence/preproteolipid/protein import

Introduction

The biogenesis of mitochondria depends on the import of proteins, which are encoded in the nucleus and synthesized on cytosolic ribosomes (for review, see Schatz and Mason, 1974; Tzagoloff *et al.*, 1979). These imported proteins are initially synthesized as preproteins that are often much larger than the corresponding mature proteins of the organelle (for review, see Neupert and Schatz, 1981). Up to now, no information on the primary structure of such imported mitochondrial preproteins has existed. Cloning and sequencing of the corresponding mRNAs is one possible way to characterize imported mitochondrial proteins in more detail. Working with the eukaryotic microorganism *Neurospora crassa*, we have therefore constructed, from total polyadenylated RNA, an ordered cDNA bank cloned in *Escherichia coli*. These cDNA clones are presently screened for sequences encoding mitochondrial proteins by procedures based on the translation of hybridization-selected mRNA (Parnes *et al.*, 1981; Kvist *et al.*, 1981).

This paper describes the isolation of cDNA encoding the major part of the mRNA of the proteolipid subunit of the mitochondrial ATP synthase. The proteolipid subunit is an extremely hydrophobic polypeptide of 81 amino acid residues whose sequence is known (Sebald *et al.*, 1980). It occurs as an oligomer – most likely as a hexamer in the proton-conducting F_0 portion of the ATP synthase (Sebald *et al.*, 1979a). In *N. crassa* the proteolipid is synthesized as a larger preprotein with an apparent mol. wt. of 12 000–15 000 (Michel *et al.*, 1979). *In vitro* the preprotein could be transported into mitochondria post-translationally (Zimmermann *et al.*, 1981).

The nucleotide sequence of isolated proteolipid cDNA indicates the amino acid sequence of the whole preproteolipid, i.e., the 81 residues of the mature protein as well as a 66-residue extension at the amino-terminal end. This provides, for the first time, information on the structure of a preprotein that is imported into mitochondria. Furthermore, the highly polar and unusually long presequence gives a clue as to how the extremely hydrophobic proteolipid can be solubilized under cellular conditions and thus can be transported into the organelle after translation is completed. The purified cDNA will be a valuable tool for future studies of the import pathway as well as of the organization of the nuclear gene of the *N. crassa* proteolipid.

Results

Isolation of a cDNA clone encoding proteolipid mRNA

The proteolipid subunit comprises 10% of the total ATP synthase protein, and 0.2–0.4% of total cellular protein of *N. crassa* (Sebald *et al.*, 1979a). After cell-free translation of total mRNA, 0.4% of the total incorporated [^{3}H]leucine (Michel *et al.*, 1979) and 1% of incorporated [^{35}S]methionine (Figure 1) is recovered as preproteolipid. These data suggest that the preproteolipid mRNA is of middle abundance (Wong and Marzluf, 1980), and that a cDNA encoding proteolipid mRNA could be found when some hundreds of clones had been analyzed.

Two cDNA plasmids that hybridized specifically with proteolipid mRNA were identified among 288 screened cDNA clones that had been prepared from total polyadenylated RNA and had been ordered and conserved on microtitre plates. Each of the clones had been grown separately in a 50-ml culture. To shorten the screening, the cultures were combined together in groups of twelve for plasmid isolation and subsequent hybridization-selection of mRNA. Finally, single clones from a positive plasmid pool were analyzed. As demonstrated in Figure 1 (lanes B, C, D), total cell-free translation products synthesized in response to total mRNA, as well as to the mRNA selected by a positive plasmid pool (Pl/E) and by the single purified proteolipid cDNA plasmid (P1/E2), clearly indicate the progress of purification. The mRNA selected by the cDNA plasmid P1/E2 yields a translation product that can be immunoadsorbed to 60–80% with proteolipid antibodies.

The cDNA of clone P1/E2 was sequenced. The small insert in this plasmid corresponds to 160 bases of the 3' end plus ~20 bases of A of the proteolipid mRNA (Figure 2). Further proteolipid cDNA clones were obtained at a frequency of 0.2% by filter hybridization of ~20 000 colonies using an

*To whom reprint requests should be sent.

 0261-4189/82/0105-0565$2.00/0.

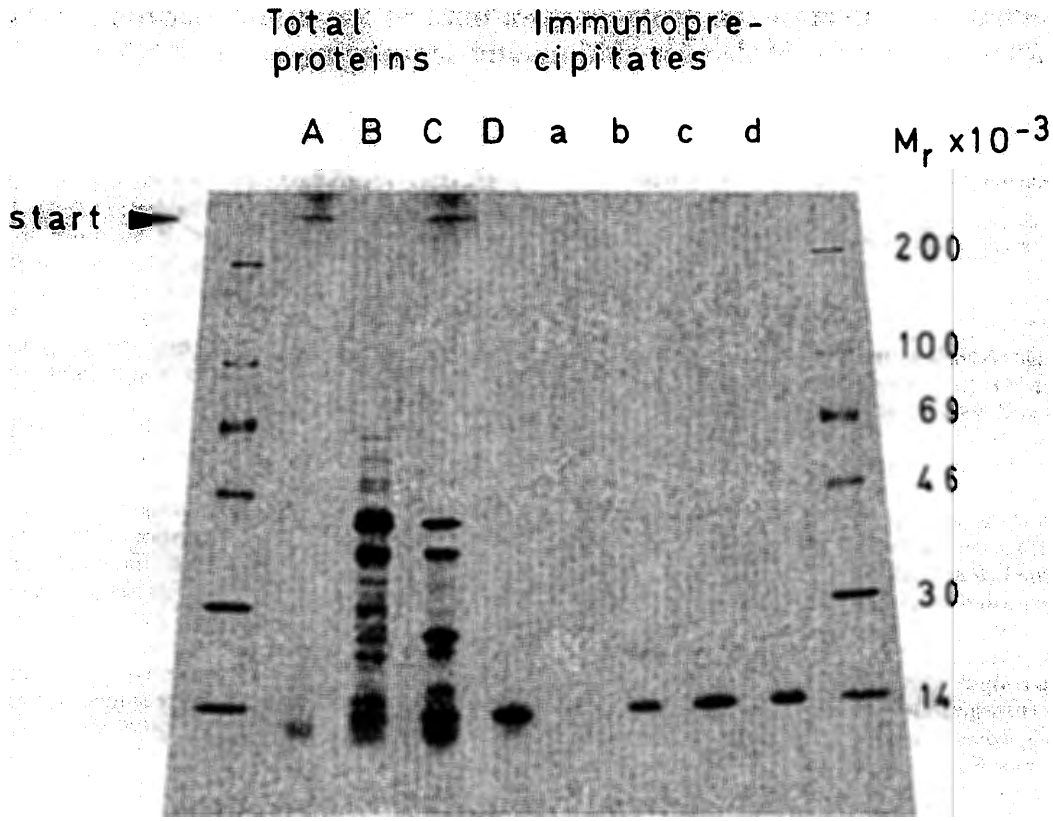

Fig. 1. Purification of a cDNA-pBR322 recombinant plasmid (P1/E2) hybridizing specifically to proteolipid mRNA. Recombinant plasmid DNA was covalently bound to paper and hybridized with total polyadenylated RNA as described in Materials and methods. The selected mRNAs were translated in parallel with total mRNA in 15 μl of a cell-free wheat germ system with 15 μCi [^{35}S]methionine for 60 min. An aliquot (80%) of each assay was incubated with antiproteolipid immunoglobulins. Total translation products **(lanes A, B, C, D)** and the immunoadsorbed proteins **(lanes a, b, c, d)** were separated by SDS-gel electrophoresis and visualized by fluorography. Radioactive marker proteins were lysozyme (mol. wt. 14 300), carbonic anhydrase (mol. wt. 30 000), ovalbumin (mol. wt. 46 000), bovine serum albumin (mol. wt. 69 000), phosphorylase b (mol. wt. 92 500), and myosin heavy chain (mol. wt. 200 000). The four translation assays contained no added RNA **(lanes A/a)**, total polyadenylated RNA **(lanes B, b)**, mRNA selected by 12 pooled plasmids P1/E1 to P1/E12 **(lanes C, c)**, and mRNA selected by purified plasmid P1/E2 **(lanes D, d)**.

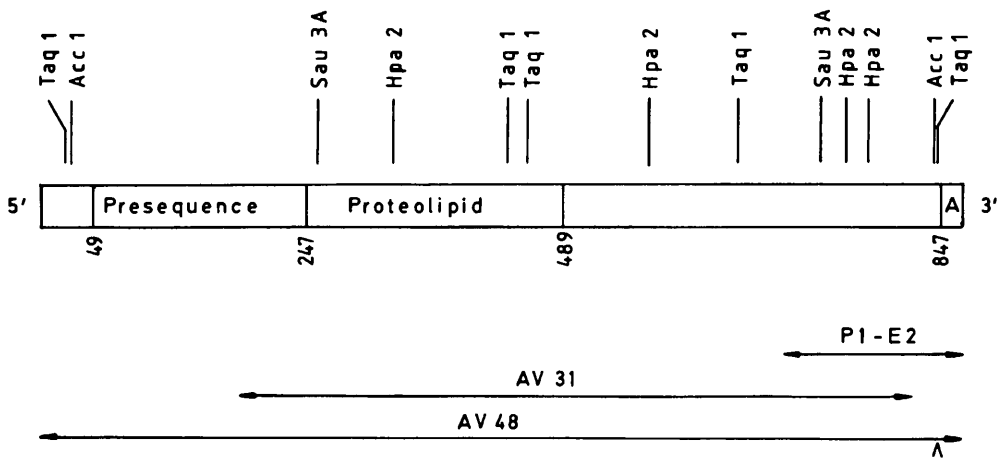

Fig. 2. Schematic representation of the proteolipid mRNA. The structure was deduced from the sequenced cDNA inserts of plasmids P1/E2, AV31, and AV48. Compared with P1/E2 the cDNA of plasmid AV48 has a deletion of at least two bases before the poly(A) tail (arrow). The restriction sites used for DNA sequence determination are also indicated.

end-labelled *Sau*3A-*Taq*I fragment as a probe. The cDNA inserts of two cloned plasmids consisting of 670 bp (AV31) and 865 bp (AV48) were then analyzed.

Nucleotide sequence of the cloned proteolipid mRNA

Five regions are evident in the proteolipid mRNA sequence which should correspond to the nucleotide sequences of the

cloned cDNAs (Figure 2). (1) The presence of a poly(A) tail in AV48 and P1/E2 indicates that these plasmids included the whole 3′ end of the mRNA. The polyadenylation site differs in the two analysed inserts. At least the nucleotides dC 846 and dC 847 are absent in AV48 (see also Sasavage *et al.*, 1982). This leads to a loss of a *Sal*I restriction site. (2) A long non-translated segment of 355 bases occurs between the TAA stop codon and the polyadenylation site. The almost universal polyadenylation signal AAUAAA (Proudfoot and Brownlee, 1974; Fitzgerald and Shenk, 1981) is absent. The segment AAAAAGAAA (nucleotides 822–830) or a part of it could be an equivalent signal in *N. crassa*. (3) The mature proteolipid is coded for by nucleotides 247–489. This nucleotide sequence matches completely the known amino acid sequence of the wild-type proteolipid (Sebald *et al.*, 1980) with the exception of one nucleotide: the cloned cDNA has been derived from the mRNA of the oligomycin-resistant mutant AP-12 (Sebald and Hoppe, 1981). The mutant proteolipid has a phenylalanine (wild-type)-serine (mutant) substitution at amino acid position 61. Apparently, the mutational event was a T/C transition at nucleotide position 428 (Figure 3). (4) There exists a reading frame in phase with the proteolipid sequence which starts with an ATG at nucleotide 49. These 178 nucleotides encode the presequence of the preproteolipid, as confirmed by partial amino acid sequences of preprotein fragments (see below). (5) A non-translated segment of 48 nucleotides is found at the 5′ end. Analysis of the corresponding genomic DNA must be awaited to establish the true start(s) of the mRNA.

Amino acid sequence studies

Attempts to sequenate whole immunoadsorbed preproteolipid failed, since the amino-terminal end is blocked. When an acetyl-coenzyme A trap was added during cell-free translation (Palmiter, 1977), no methionine radioactivity was recovered during step 1 of the Edman degradation. Apparently, the preproteolipid is proteolytically processed under this condition. After cyanogen bromide cleavage of [^{3}H]leucine-labelled or [^{3}H]alanine-labelled preproteolipid, a fragment of ~15 amino acid residues could be isolated by gel chromatography. Sequence analysis of this fragment revealed (Figure 4) a partial sequence Ala xxxx Leu Ala xx Leu Ala xxx. Such a sequence of alanine and leucine residues is found at positions 2–15 of the presequence (Figure 3).

Properties of the presequence and the preproteolipid

The amino acid sequence deduced from the nucleotide sequence yields a preproteolipid with a mol. wt. of 15 348. During SDS-gel electrophoresis the preproteolipid migrates slightly slower than lysozyme (Figure 1) with an apparent mol. wt. of ~15 000. The amino acid composition of the presequence is remarkably polar (Table I). The polarity calculated from the content of hydrophilic side chains (Capaldi and Vanderkooi, 1972) is 53%. In contrast, the mature proteolipid, which is one of the most hydrophobic proteins known, has a polarity of only 25.9%. Because the presequence is unusually long, in the whole preproteolipid the contents of hydrophilic and lipophilic residues are approximately balanced. The polarity of 38.2% approaches that of a water-soluble protein. The presequence contains 12 basic residues, whereas no acidic residues are present. Because the mature proteolipid contains equal numbers of only four basic and four acidic residues, the preproteolipid exhibits an unusually high concentration of positive charges (see Figure 5). The few lipo-

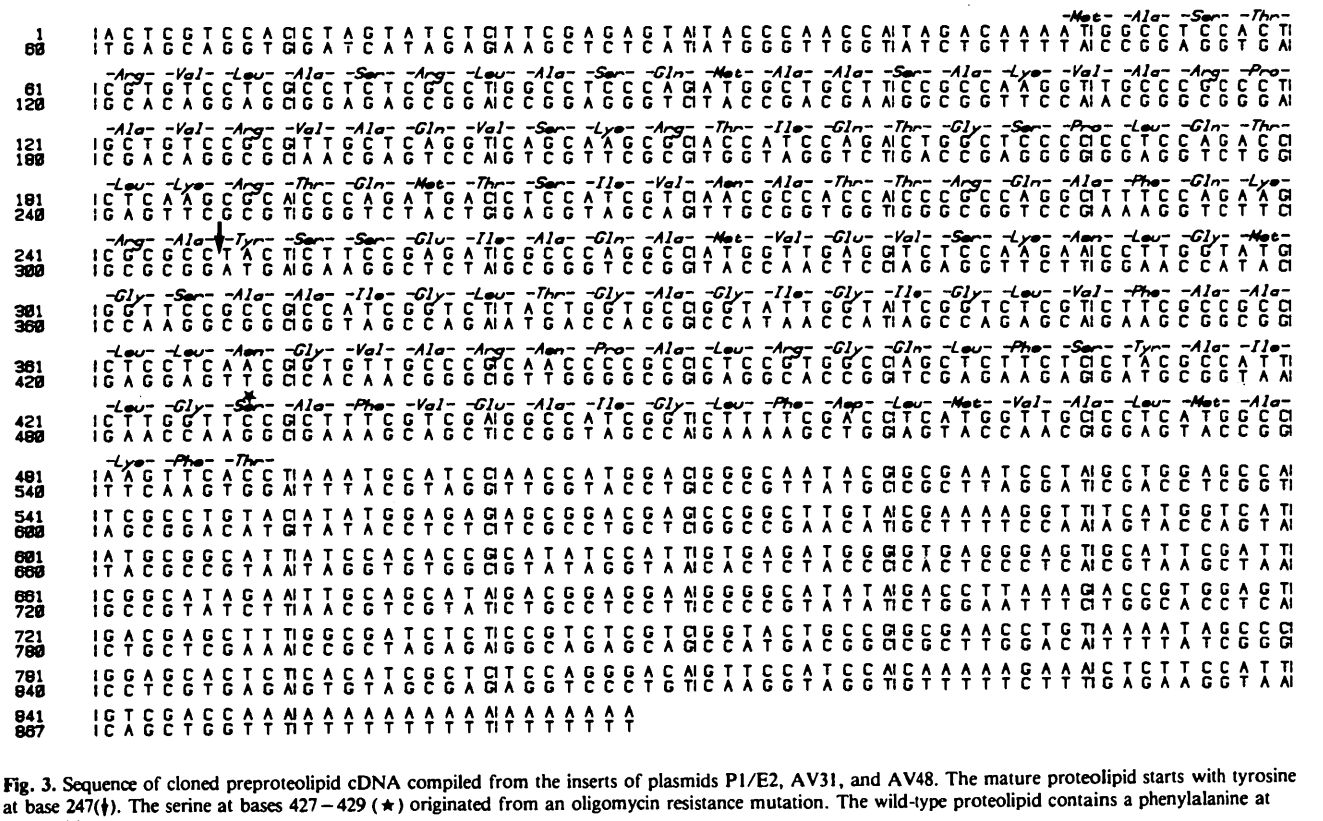

Fig. 3. Sequence of cloned preproteolipid cDNA compiled from the inserts of plasmids P1/E2, AV31, and AV48. The mature proteolipid starts with tyrosine at base 247(ϟ). The serine at bases 427–429 (★) originated from an oligomycin resistance mutation. The wild-type proteolipid contains a phenylalanine at this position.

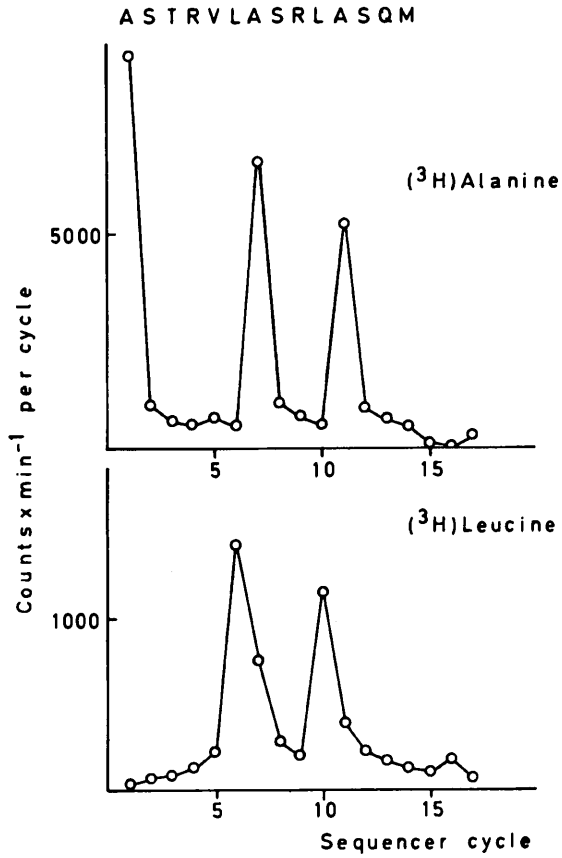

Fig. 4. Partial amino acid sequences of a cyanogen bromide fragment of the preproteolipid. The preproteolipid was synthesized in the presence of either [^{3}H]alanine or [^{3}H]leucine. The cyanogen bromide fragment was purified by gel chromatography (see Materials and methods) and submitted to 17 cycles of automated Edman degradation. The sequence of the released alanines and leucines is found only in a 14-residue fragment representing amino acids 2–15 of the presequence.

philic side chains of the presequence are randomly distributed. In contrast, the mature proteolipid clearly exhibits two segments each of ~25 residues where the lipophilic residues are clustered (Sebald and Hoppe, 1981). Consequently, parameters indicating the free energy gain for a side chain during transition from a random coil in water to a membrane-embedded α-helix (von Heijne, 1981) predict membrane-solubility for two segments of the mature proteolipid only, but not for the presequence (Figure 5).

Various prediction methods (Chou and Fasman, 1978; Nagano, 1977; Robson and Suzuki, 1976; Maxfield and Scheraga, 1976) were applied to calculate secondary structure elements of the preproteolipid from the amino acid sequence. α-Helical regions and β-turns predicted by all four programs are included in the upper part of Figure 5. Turns are predicted for residues 37–40 and 67–70. The latter turn would immediately follow the proteolytically cleaved peptide bond of the preproteolipid. Consistently, α-helical conformation is assigned to residues 1–33 and 56–66 of the presequence. The confirmation of the rest of the presequence is predicted ambiguously.

Discussion

Increasing evidence has been accumulating that the transport of proteins into organelles of possibly endosymbiotic origin (Margulis, 1981), such as chloroplasts and mitochondria, differs in important aspects from the transmembrane transport of secretory proteins in eukaryotic and prokaryotic cells (Chua and Schmidt, 1978; Blobel, 1980; Neupert and Schatz, 1981). The most notable difference is that the mitochondrial and chloroplast preproteins are released from the ribosome into the cytosol before import into the organelle takes place. Moreover, most mitochondrial and chloroplast preproteins possess much longer presequences than the secretory proteins. The present work permits a more precise definition of the differences between the presequence of a protein that is transported into mitochondria and the presequence of secretory proteins. It also allows us to consider aspects of a post-translational import process in relation to the primary structure of the preproteolipid.

Formerly, it has been difficult to understand how an intrinsic membrane protein, and especially one as hydrophobic as the proteolipid, can exist in solubilized form in the cytosol (Neupert and Schatz, 1981). The finding of a long polar presequence that balances the hydrophobic sequences of the mature protein and could serve as a 'swimming belt' gives a satisfactory answer to this question, so that it is no longer necessary to postulate the existence of a cytosolic carrier system.

In the case of secretory proteins it has been hypothesized (von Heijne, 1981) that the presequence, containing a sequence of ~20 uncharged and predominantly lipophilic residues, partitions directly into the membrane interior, thereby making further translocation of the growing nascent chains possible. It is evident that the presequence of the proteolipid, which contains only a few separated lipophilic side chains, is unable to fulfil this role. It might be postulated however, that membrane contacts and trans-membrane

Table I.

Amino acid residues	Amino acid frequencies in		
	Presequence	Proteolipid	Preproteolipid
	mol/mol		
Alanine	12	14	26
Arginine	8	2	10
Asparagine	1	3	4
Aspartic acid	–	1	1
Cysteine	–	–	–
Glutamine	7	2	9
Glutamic acid	–	3	3
Glycine	1	11	12
Histidine	–	–	–
Isoleucine	2	6	8
Leucine	4	11	15
Lysine	4	2	6
Methionine	3	4	7
Phenylalanine	1	5	6
Proline	2	1	3
Serine	7	6	13
Threonine	8	2	10
Tryptophan	–	–	–
Tyrosine	–	2	2
Valine	6	6	12
Total residues	66	81	147
Polarity	53%	25.9%	38.2%

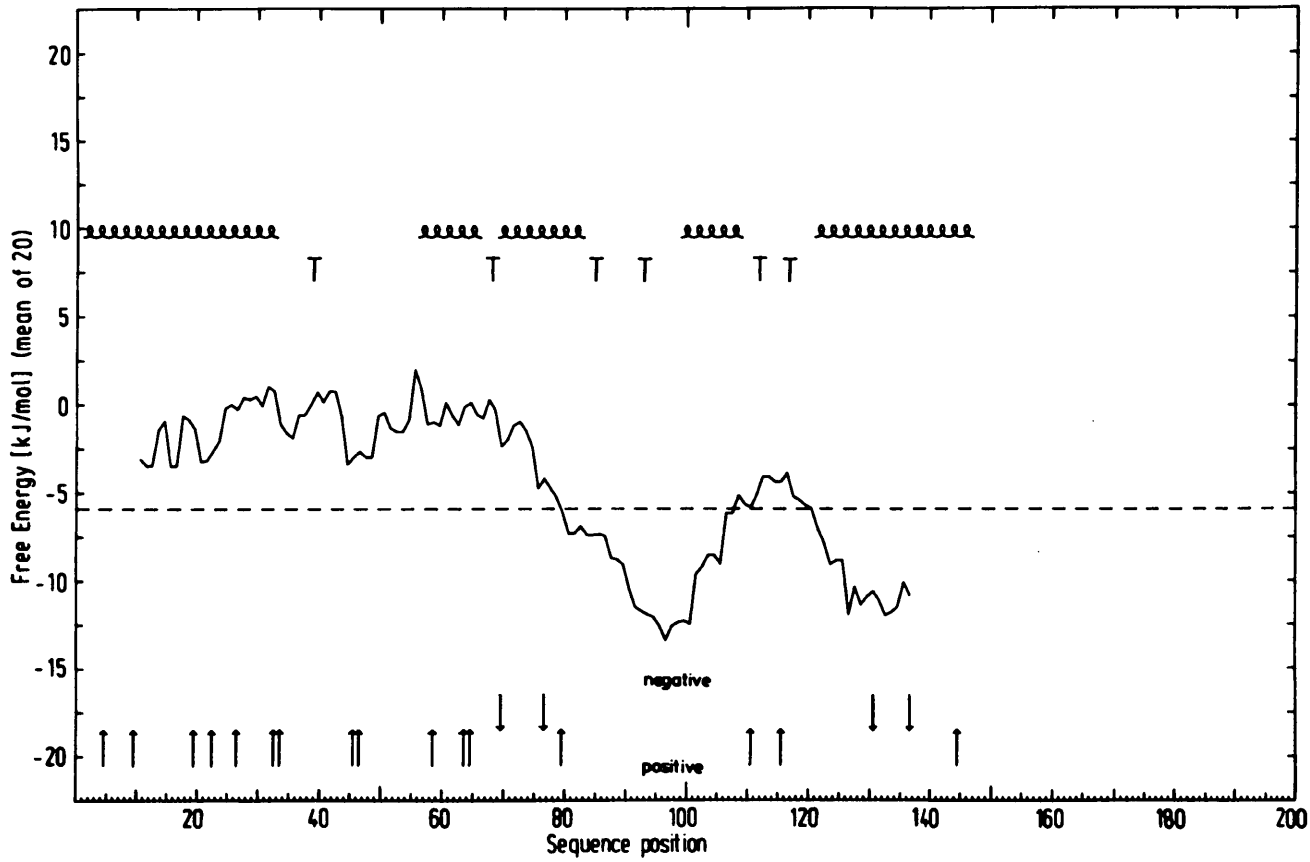

Fig. 5. Predicted distribution of lipophilic segments, secondary structure elements, and charges along the amino acid sequences of the preproteolipid. **Top:** The indicated α-helical segments ( ) and β-turns (T) are consistently predicted by four different methods (see Results). **Middle:** The gain of free energy during transition of a 20-residue segment from a random coil in water to an α-helix in the membrane is calculated for all sequence positions according to von Heijne (1981). **Bottom:** The position of basic residues (positive charge) and acidic residues (negative charge) is indicated by arrows.

movement of the preproteolipid are facilitated and perhaps guided by the two lipophilic 25-residue segments occurring in that part of the preprotein which ultimately forms the mature proteolipid. One reasonable hypothesis is that during post-translational import the location of hydrophobic segments in the polypeptide chain is of secondary importance.

The primary structure gives no information on the 'address label' which determines the recognition of the preproteolipid by the mitochondria. The highly positive charge of the presequence is reminiscent of cytochrome c and may contribute to the recognition process. Considering the only superficial similarity of the presequences of secretory proteins (von Heijne and Blomberg, 1979; von Heijne, 1981), as well as the specificity of the import into chloroplasts of the small subunit of the ribulose-1,5-bisphosphate carboxylase in pea and *Chlamydomonas* (Schmidt *et al.*, 1979; Bedbrook *et al.*, 1980; Broglie *et al.*, 1981), it is evident that information on the conformation of preproteins is required for identification of the address label.

An electrochemical potential across the inner mitochondrial membrane is necessary for the import of the preproteolipid (Zimmermann *et al.*, 1981) and of other mitochondrial proteins (Gasser *et al.*, 1982). It is tempting to speculate that the potential (negative inside, positive outside) drives the highly charged preproteolipid into and across the membrane. It will be interesting to see whether or not a large positive charge is a property of other imported mitochondrial proteins.

By virtue of the isolation of proteolipid cDNA, the corresponding mRNA can be purified, and thus a radiochemically pure preproteolipid can be synthesized in a cell-free system. In the future this will allow a closer analysis of the proteolytic processing during import, since the fate of the presequence can be observed. In principle, the presequence might be removed in the mitochondrial inter-membrane space, and then the proteolipid could be inserted from the outside into the inner membrane. However, this would not be expected on the basis of the process of assembly of the ATP synthase proteolipid in other systems. In bacteria (Kanazawa *et al.*, 1981), chloroplasts (Nelson *et al.*, 1980; Sebald and Wachter, 1980), and remarkably also in yeast mitochondria (Macino and Tzagoloff, 1979; Hensgen *et al.*, 1979), the proteolipid is inserted into the membrane from the F_1 side, which corresponds in mitochondria to the matrix space. No proteolytic processing of the proteolipid is necessary in these systems, because only the mature form is genetically determined and synthesized. Thus, it is a reasonable hypothesis that in *Neurospora* mitochondria the preproteolipid is processed in the matrix space to the functional size, and that it then integrates spontaneously into the inner membrane. Such a processing pathway would be in accordance with the observation that in yeast mitochondria a processing protease is located in

the matrix space. Cytochrome c_1, which is located at the outer surface of the inner membrane, and even cytochrome b_2, which is located in the inter-membrane space, are both processed by this matrix protease (Gasser *et al.*, 1982).

Materials and methods

Strains and materials

N. crassa wild-type SL74A (FGSC stock number 987) and an oligomycin-resistant mutant AP-12 (Sebald and Hoppe, 1981) were used for RNA and protein isolations. cDNA recombinant plasmids were cloned in *E. coli* 5 K (Hubacek and Glover, 1970). Avian myeloblastosis virus reverse transcriptase was a gift from W. Beard. Calf intestinal alkaline phosphatase and *E. coli* DNA polymerase I were from Boehringer (Mannheim, FRG). Terminal deoxynucleotidyl transferase, T4 polynucleotide kinase, and restriction endonucleases were from BRL (Rockville, MD) or a gift from H. Mayer (Stöckheim). [γ-^{32}P]ATP (>5000 Ci/mmol), [α-^{32}P]dGTP (400 Ci/mmol), [^{3}H]dCTP (19 Ci/mmol), [^{3}H]leucine (40–60 Ci/mmol), [^{3}H]alanine (53 Ci/mmol), and [^{35}S]methionine (1000 Ci/mmol) were from Amersham Buchler (Braunschweig, FRG). Nitrocellulose filters of 8.5 cm diameter were from Schleicher and Schüll (Dassel, FRG). They were washed with 0.1 M EDTA and H_2O at 100°C and sterilized between filter paper before use. Formamide, reagent grade (Merck, Darmstadt, FRG), was used without further purification.

Preparation of total polyadenylated RNA

Hyphae of *N. crassa* were grown under aeration in Vogel's minimal medium plus 2% sucrose at 30°C to a wet weight of 10–15 g/l (Sebald *et al.*, 1979b). They were harvested by filtration, washed with water, immediately frozen in liquid nitrogen, and pulverized under liquid nitrogen in a Waring blender. From 100 g wet weight of cells, 750 mg total RNA was extracted (Michel *et al.*, 1979). After chromatography on poly(U)-cellulose (Sheldon *et al.*, 1972) ~ 3 mg of enriched polyadenylated RNA was obtained and stored in 70% ethanol at −20°C.

Cell-free protein synthesis

A cell-free protein-synthesizing system was prepared from wheat embryos (Roberts and Paterson, 1973) using an acetate medium (Davies *et al.*, 1977). Assay conditions were as described (Michel *et al.*, 1979) employing 1 mCi [^{35}S]methionine/ml. Using 15-μl assays the system corresponded linearly to added polyadenylated RNA up to 100 μg/ml. Incorporation into hot trichloroacetic acid-insoluble material was $2-3 \times 10^6$ c.p.m./μg added RNA corresponding to a 20- to 30-fold stimulation over endogenous incorporation.

Immunoadsorption

Antibodies against the purified *N. crassa* ATP synthase proteolipid were raised in rabbits (Sebald *et al.*, 1979a). Proteolipid-specific immunoglobulins were purified by affinity chromatography (Werner and Sebald, 1981). The preproteolipid was immunoadsorbed from 15 μl of solubilized wheat germ assays (Goldman and Blobel, 1978) using 20 μg immunoglobulin and 2 mg protein A-Sepharose CL-4B (Pharmacia, Uppsala, Sweden).

Preparation of cDNA recombinant plasmids

Single-stranded cDNA was synthesized from total polyadenylated RNA according to Friedman and Rosbash (1977) omitting actinomycin D and using 1 U reverse transcriptase/μg RNA. Yields of cDNA were between 3 and 5%. Double-stranded cDNA was synthesized following two different protocols: (1) After second strand synthesis for 4 h at 25°C in the presence of 1 mM of each of the four dNTPs (Woods *et al.*, 1980) the DNA was treated with S1 nuclease. (2) Single-stranded cDNA was tailed with dC (Land *et al.*, 1981), and the synthesis of the second strand was primed with annealed oligo(dG). The double-stranded cDNA obtained by both procedures was tailed with dC (Land *et al.*, 1981). Plasmid pBR322 was cleaved with *Pst*I, and the linearized form was purified by CsCl gradient centrifugation. After tailing with 15–20 bases of dG (Hoeijmakers *et al.*, 1980) the plasmid DNA was annealed with dC-tailed double-stranded cDNA at a weight ratio of 5 to 1. Transformation of *E. coli* 5K was performed according to Dagert and Ehrlich (1979). Employing 12 ng of the annealed DNAs/0.1 ml *E. coli* cells, 800–1600 tetracycline-resistant cells were obtained after plating directly on nitrocellulose filters (Hanahan and Meselson, 1980). This corresponds to 400–800 colonies/ng of input cDNA. The tailed plasmid annealed in parallel in the absence of cDNA yielded 10–20 tetracycline-resistant cells, i.e., <2%. Cells of single colonies were inoculated in 50 μl LB medium plus tetracycline in the 96 wells of a microtitre plate. After overnight growth, 25 μl glycerol was added, and the sealed plates were stored at −20°C.

Isolation of plasmids

Plasmids were isolated from 2–1000 ml cultures grown overnight at 37°C in LB medium plus tetracycline (5 μg/ml) following a modified protocol (Grosveld *et al.*, 1981) of Birnboim and Doly (1979). For large-scale preparations, supercoiled plasmid DNA corresponding to a 1-l culture was purified by banding once or twice in 11-ml CsCl gradients.

Binding of DNA to paper

Plasmid DNA (50–100 μg) submitted once to CsCl-centrifugation was partially depurinated by an incubation for 10 min at 25°C in 200 μl 50 mM HCl. The DNA was broken and dissociated by adding 0.2 ml 0.5 N NaOH for 30 min at 37°C. The DNA was precipitated with 1 ml ethanol plus 0.08 ml 3 M potassium acetate, pH 4.6. After a 70% ethanol wash, the pellet was desiccated and dissolved in 10 μl water. The solution was heated for 2 min at 100°C. Then 40 μl redistilled dimethyl sulfoxide plus 2.5 μl 3 M potassium acetate, pH 4.6, was added. The mixture was incubated with two 1-cm^2 circles of diazobenzyloxymethyl paper (Alwine *et al.*, 1979) overnight at room temperature. The circles were washed four times with 0.5 M NaOH, rinsed with water until the pH was neutral, and stored at 4°C in hybridization buffer (50% formamide, 5 x SSC = 750 mM NaCl plus 75 mM sodium citrate, 0.1% SDS, 100 μg poly(A) and 100 μg yeast tRNA/ml).

Hybridization selection of mRNA

Two circles of DNA paper were soaked with 80 μl of 1.25-fold concentrated hybridization buffer, and 50 μg polyadenylated RNA dissolved in 20 μl H_2O was added. Hybridization proceeded for 15 h at 37°C. The circles were washed twice with hybridization buffer (5 ml), twice with 50% formamide, 0.2 x SSC, 30 mM sodium phosphate, pH 7.3, 0.1% SDS at 37°C. The bound RNA was eluted at 65°C with four times 0.1 ml 98% formamide containing 5 μg tRNA/ml, 1 mM EDTA, 0.1% SDS, and 10 mM Tris-HCl, pH 7.5. Particles – mainly disintegrated paper – were removed by centrifugation for 5 min at 18 000 *g*. RNA was precipitated with 1 ml ethanol after addition of 40 μl 2 M sodium acetate, pH 6. RNA was reprecipitated from 100 μl water and dissolved in 10 μl water. 5-μl aliquots were analysed in a 15-μl wheat germ assay.

Filter colony hybridization

Transformed cells, at a density of 400–800/85-mm diameter nitrocellulose filter, were grown to a colony size of ~0.5 mm. DNA was amplified by a chloramphenicol treatment and fixed to the nitrocellulose filter as described (Grunstein and Hogness, 1975). The filters were hybridized (Jeffreys and Flavell, 1977) with end-labelled cDNA fragments prepared during DNA sequence analysis (see Figure 2). Single cells were isolated from positive colonies and submitted to a second round of colony hybridization.

DNA sequence analysis

Suitable DNA fragments (see Figure 2) were sequenced according to Maxam and Gilbert (1980), performing chemical modification at G, C plus T, C (Maxam and Gilbert, 1980), and at A + G (Burton, 1967).

Protein sequence analysis

Immunoadsorbed preproteolipid labelled in the wheat germ system with either [^{3}H]leucine or [^{3}H]alanine was cleaved with cyanogen bromide, and the resulting fragments were separated by Bio-Gel P-30 chromatography in the presence of 80% formic acid (Sebald *et al.*, 1980). Automated Edman degradation was performed after binding a fragment to aminated glass by the homoserine lactone coupling procedure (Laursen, 1977).

Analytical procedures

Proteins were separated electrophoretically on 12–20% concave gradients of SDS-polyacrylamide gels (Laemmli, 1970), and radioactivity was visualized by fluorography (Bonner and Laskey, 1974). ^{32}P, ^{35}S, and ^{3}H radioactivities were measured in a Prias liquid scintillation spectrometer.

Acknowledgements

We thank J. Hoppe and D. Lincoln for adapting the computer programs and H.U. Schairer, H. Mayer, H. Hauser, G. Gross, J.E.G. McCarthy, and H. Feldmann for helpful discussions.

References

Alwine,J.C., Kemp,D.J., Parker,B.A., Reiser,J., Renart,J., Stark,G.R., and Wahl,G.M. (1979) *Methods Enzymol.*, **68**, 220-242.
Bedbrook,J.R., Smith,S.M., and Ellis,R.J. (1980) *Nature*, **287**, 692-697.
Birnboim,A.C., and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.
Blobel,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1496-1500.
Bonner,W.M., and Laskey,R.A. (1974) *Eur. J. Biochem.*, **46**, 83-88.
Broglie,R., Bellemare,G., Bartlett,S.G., Chua,N.-H., and Cashmore,A.R. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7304-7308.
Burton,K. (1967) *Methods Enzymol.*, **12A**, 222-224.

Capaldi,R.A., and Vanderkooi,G. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 930-932.
Chou,P.Y., and Fasman,G.D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol.*, **47**, 45-148.
Chua,N.-H., and Schmidt,G.W. (1979) *J. Cell Biol.*, **81**, 461-483.
Dagert,M., and Ehrlich,S.D. (1979) *Gene*, **6**, 23-28.
Davies,J.W., Aalbers,A.M.J., Stuik,E.J., and Van Kammen,A. (1977) *FEBS Lett.*, **77**, 265-269.
Fitzgerald,M., and Shenk,Th. (1981) *Cell*, **24**, 251-260.
Friedman,E.G., and Rosbash,M. (1977) *Nucleic Acids Res.*, **4**, 3455-3471.
Gasser,S.M., Ohashi,A., Daum,G., Böhni,P.C., Gibson,J., Reid,G.A., Yonetani,T., and Schatz,G. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 267-271.
Goldman,B.M., and Blobel,G. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 5066-5070.
Grosveld,F.G., Dahl,H.M., de Boer,E., and Flavell,R.A. (1981) *Gene*, **13**, 227-237.
Grunstein,M., and Hogness,D.S. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961-3965.
Hanahan,D., and Meselson,M. (1980) *Gene*, **10**, 63-67.
Hensgen,L.A.M., Grivell,L.A., Borst,P., and Bos,J.L. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1663-1667.
Hoeijmakers,J.H.H., Borst,P., Van den Burg,J., Weissmann,C., and Cross, G.A.M. (1980) *Gene*, **8**, 391-417.
Hubacek,J., and Glover,S.W. (1970) *J. Mol. Biol.*, **50**, 111-127.
Jeffreys,A.J., and Flavell,R.A. (1977) *Cell*, **12**, 429-439.
Kanazawa,H., Mabuchi,K., Kayano,T., Tamura,F., and Futai,M. (1981) *Biochem. Biophys. Res. Commun.*, **100**, 219-225.
Kvist,S., Bregegere,F., Rask,L., Carni,B., Garoff,H., Daniel,F., Wiman,K., Larhammar,D., Abastado,J.P., Gachelin,G., Peterson,P.A., Dobberstein,B., and Kourilsky,P. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2772-2776.
Laemmli,U.K. (1970) *Nature*, **227**, 680-685.
Land,H., Grez,M., Hauser,H., Lindenmeier,W., and Schütz,G. (1981) *Nucleic Acids Res.*, **9**, 2251-2266.
Laursen,R.A. (1977) *Methods Enzymol.*, **47**, 277-288.
Macino,G., and Tzagoloff,A. (1979) *J. Biol. Chem.*, **254**, 4617-4623.
Margulis,L. (1981) *Symbiosis in Cell Evolution*, published by Freeman and Company, San Francisco, CA, pp. 1-419.
Maxam,A.M., and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
Maxfield,F.R., and Scheraga,H.A. (1976) *Biochemistry (Wash.)*, **15**, 5138-5153.
Michel,R., Wachter,E., and Sebald,W. (1979) *FEBS Lett.*, **101**, 373-376.
Nagano,K. (1977) *J. Mol. Biol.*, **109**, 251-274.
Nelson,N., Nelson,H., and Schatz,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1361-1364.
Neupert,W., and Schatz,G. (1981) *Trends Biochem. Sci.*, **6**, 1-4.
Palmiter,R.D. (1977) *J. Biol. Chem.*, **252**, 8781-8783.
Parnes,J.R., Velan,B., Felsenfeld,A., Ramanathan,L., Ferrini,U., Appela,E., and Seidman,J.G. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2253-2257.
Proudfoot,N.J., and Brownlee,G.G. (1974) *Nature*, **252**, 359-362.
Roberts,B.E., and Paterson,B.M. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 2330-2334.
Robson,B., and Suzuki,E. (1976) *J. Mol. Biol.*, **107**, 327-356.
Sasavage,N.L., Smith,M., Gittman,S., Woychik,R.P., and Rottman,F.M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 223-227.
Schatz,G., and Mason,T.L. (1974) *Annu. Rev. Biochem.*, **43**, 51-87.
Schmidt,G.W., Devillers-Thiery,A., Desmisseaux,H., Blobel,G., and Chua, N.H. (1979) *J. Cell Biol.*, **83**, 615-622.
Sebald,W., Graf,Th., and Lukins,H.B. (1979a) *Eur. J. Biochem.*, **93**, 587-599.
Sebald,W., Neupert,W., and Weiss,H. (1979b) *Methods Enzymol.*, **55**, 144-148.
Sebald,W., and Wachter,E. (1980) *FEBS Lett.*, **122**, 307-311.
Sebald,W., Machleidt,W., and Wachter,E. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 785-789.
Sebald,W., and Hoppe,J. (1981) *Curr. Top. Bioenerg.*, **12**, 1-64.
Sheldon,R., Jurale,C., and Kates,J. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 417-421.
Tzagoloff,A., Macino,G., and Sebald,W. (1979) *Annu. Rev. Biochem.*, **48**, 419-441.
von Heijne,G., and Blomberg,C. (1979) *Eur. J. Biochem.*, **97**, 175-181.
von Heijne,G. (1981) *Eur. J. Biochem.*, **120**, 275-278.
Werner,S., and Sebald,W. (1981) *Methods Biochem. Anal.*, **27**, 109-170.
Wong,L.-J., and Marzluf,G.A. (1980) *Biochim. Biophys. Acta*, **607**, 122-135.
Woods,D., Crampton,J., Clarke,B., and Williamson,R. (1980) *Nucleic Acids Res.*, **8**, 5157-5168.
Zimmermann,R., Henning,B., and Neupert,W. (1981) *Eur. J. Biochem.*, **116**, 455-460.