The sequence of a small subunit of cytochrome *c* oxidase from *Crithidia fasciculata* which is homologous to mammalian subunit IV**

Dave Speijer[a], Cornelis K.D. Breek[a], Anton O. Muijsers[a], Paulette X. Groenevelt[a], Henk Dekker[b], Annett de Haan[a], Rob Benne[a,*]

[a]*E.C. Slater Institute, University of Amsterdam, Academic Medical Centre, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands*
[b]*E.C. Slater Institute, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands*

Received 8 January 1996

Abstract The sequence of subunit 8 of cytochrome *c* oxidase from *Crithidia fasciculata* was determined by sequencing cDNA and N-terminus of the mature protein (M_r = 15.7 kDa). The (inferred) protein is homologous to mammalian cox IV and the corresponding cox subunits from yeast, *Neurospora crassa* and *Dictyostelium discoideum*, which is reflected in a very similar hydropathy profile. Elements that are conserved in the *C. fasciculata* sequence include (i) an N-terminal (D/E)-(K/R)-X-K-(X_2)-W-(X_2)-(I/L) motif, (ii) a putative membrane-spanning region in the middle portion of the protein, and (iii) a C-terminal W-(X_{13})-(N/D)-P motif. The *C. fasciculata* protein is synthesized with a cleavable presequence.

Key words: RNA editing; Trypanosome; Mitochondrion; Cytochrome *c* oxidase; Kinetoplast; Cox IV homolog

1. Introduction

Cytochrome *c* oxidase (cox) is the terminal enzyme of the respiratory chain located in the mitochondrial inner membrane [1–3]. It catalyzes the transfer of electrons from reduced cytochrome *c* to molecular oxygen, while at the same time translocating protons across the inner membrane. Eukaryotic cox is a complex composed of three large mitochondrially encoded subunits, which form the catalytic core, and up to 10 smaller subunits encoded by the nuclear genome. The function of the small subunits is still unknown, although it has been speculated that they are involved in complex assembly and (in higher eukaryotes) tissue-specific regulation [1].

Trypanosomes are unicellular parasites belonging to the order of the kinetoplastida which diverged from the main eukaryotic lineage more than 600 million years ago [4]. Due to their parasitic lifestyle, these organisms have been studied extensively, leading to the discovery of several biochemical peculiarities. One of these is the remarkable process of RNA editing in which mitochondrial mRNAs (including the mRNAs for cox subunits II and III in *Crithidia fasciculata*) are altered posttranscriptionally by Uridylate insertion and deletion (for reviews, see [5,6]). In order to study the effects of this process at the protein level, we have purified cox from *C. fasciculata* (Speijer et al., unpubl. results) and have initiated sequence analysis of its subunits. In this report, we describe the identification of one of the small subunits of cox from *C. fasciculata* as the homologue of the mammalian subunit IV (see e.g. [7]), which corresponds to *Neurospora crassa* cox V [8], yeast cox Va [9], and *Dictyostelium discoideum* cox VI [10].

2. Materials and methods

Cox from *C. fasciculata* was purified from mitochondrial vesicles obtained, as described in [11], from aerated 10-l cultures (containing 1×10^{12} cells) of *C. fasciculata*. Lysis with 0.1% Triton X-100 was performed to separate the matrix proteins from the membrane-bound proteins. For further purification, the membrane-bound complexes were solubilized with 3% lauryl maltoside and cox was purified with the aid of a methyl hydrophobic interaction chromatography (HIC) column (BioRad) (Speijer et al., unpubl. results). Upon blotting of the complex onto PVDF membranes (BioRad), N-terminal sequences were obtained from the small subunits of the complex. With the aid of two oligonucleotides, a specific 255-nt cox subunit 8 fragment was generated from total RNA from *C. fasciculata* by RT-PCR (30 cycles at a hybridization temperature of 60°C). The first oligonucleotide, 5′-GGAATTCTGCATGTGSAGSGGRATCTCGTCCCA-3′, was derived from the sequence WDEIPLHM (see Fig. 2) (S = C,G and R = G,A); the second oligonucleotide, 5′-GGAATTCCGCTATATA-AGTATCAGTTTCTGTAC-3′, was based on the mini-exon sequence present at the 5′ end of all nuclear mRNAs of *C. fasciculata* [12]. For cloning purposes, both oligonucleotides were equipped with an *Eco*RI site at the 5′ end (underlined). The N-terminal sequence GGDMHSSDRFKAAWDEIPLHM of subunit 8 shows a low but distinct homology with subunit V from *N. crassa* (see [8], conserved residues are in bold; see Fig. 2). The PCR fragment was digested with *Eco*RI, cloned in PUC 19 (digested with *Eco*RI and treated with CIP), and the insert was sequenced according to [13]. The insert was used as a probe to screen a *C. fasciculata* cDNA library, as described in [14]. One positive clone was sequenced with the aid of the vector-derived sequence primers and two internal primers (5′-GGGGTCGTCTCGCAAGCAAA-3′, nt 200–219 and 5′-TTTGA-GAAGGAGGCATGTTG-3′, nt 666–647). The complete cDNA sequence and the inferred amino acid sequence have been deposited in GenBank under accession number U44442. Hydropathy plots were generated with the MacVector program using the programmed Hopp–Woods algorithm with a hydrophilicity window of seven residues.

3. Results and discussion

We have purified cox from *C. fasciculata* with a methyl–HIC column in a rapid procedure (not shown). The resulting cox preparation was identified as such on the basis of (i) the presence of Cu_A in electron paramagnetic resonance analysis, (ii) a characteristic 605-nm peak in reduced-minus-oxidized optical spectroscopy, and (iii) the capacity to efficiently oxidize homologous, but not heterologous, cytochrome *c*.

Fig. 1 shows the result of a typical 2D PAGE experiment

*Corresponding author. Fax: (31) (20) 6915519.

**The nucleotide sequence reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accesion number U44442.

Abbreviations: cox, cytochrome *c* oxidase; nt, nucleotide; HIC, hydrophobic interaction chromatography; PAGE, polyacrylamide gel electrophoresis.

SSDI S0014-5793(96)00095-6

[15] with purified cox. From close inspection of this and numerous other gels, we conclude that cox from *C. fasciculata* consists of approximately 10 subunits, the same complex being present in a crude mitochondrial extract. We obtained N-terminal sequences of the six smallest subunits. The N-terminus of subunit 8 (arrowhead in Fig. 1) showed a low but significant similarity to that of *N. crassa* subunit V [8]. With the aid of oligonucleotides derived from the amino acid sequence and from the mini-exon sequence present in all *C. fasciculata* mRNAs [12], we isolated the corresponding cDNA (see section 2).

The 874-nt cDNA contained an ORF of 157 amino acids, the N-terminal amino acid of the mature protein being G 28 (indicated by a dot in Fig. 2). This results in a predicted MW of 15.7 kDa and a pI of 7.4 for the mature protein. Fig. 2 shows alignment of the *C. fasciculata* cox 8 sequence to that of *N. crassa* cox V [8] and to that of (putative) *N. crassa* cox V homologues from other organisms such as mouse subunit IV [7], yeast subunit Va [9] and *D. discoideum* subunit VI [10].

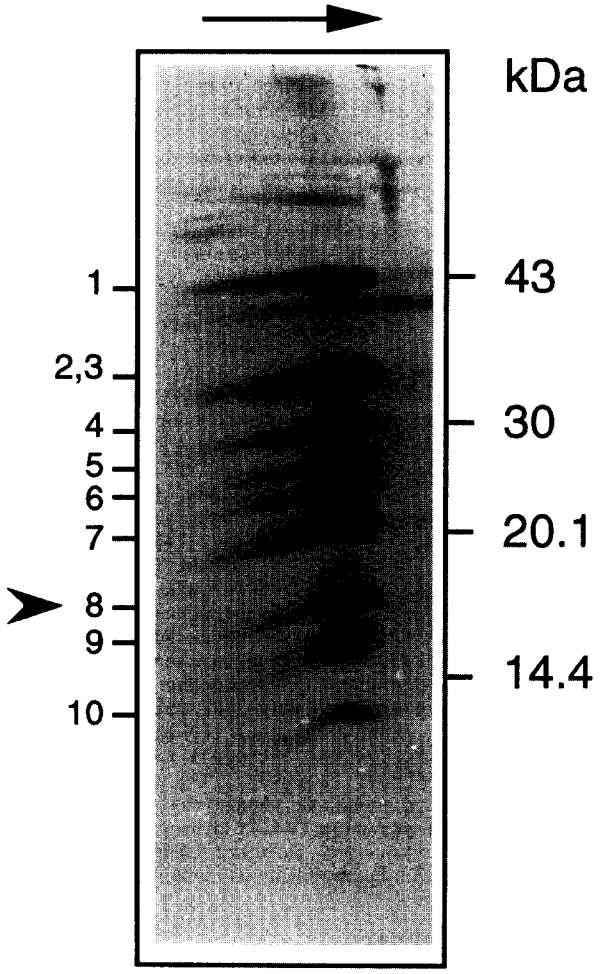

Fig. 1. 2D PAGE analysis of purified *C. fasciculata* cox. Purified *C. fasciculata* cox (7.5 μg) was run in a first dimension 5% 'blue native' polyacrylamide gel and a second dimension 16% Tris/tricine polyacrylamide gel (see [15]). Molecular weights are indicated in kDa. Subunit numbers of *C. fasciculata* cox are given on the left of the panel; subunit 8 is indicated by an arrowhead. The arrow above the Fig. indicates the direction of electrophoresis in the first dimension. Staining was done with Coomassie brilliant blue. The figure only shows the protein-containing part of the gel.

The overall degree of identity is low: e.g. 16% of the amino acids in the *C. fasciculata* and the yeast sequence are identical if conservative substitutions are included (see Fig. 2). In spite of that, a number of conserved elements that are present in (most of) the other sequences are also found in *C. fasciculata*: (i) a (D/E)-(K/R)-X-K-(X_2)-W-(X_2)-(I/L) motif (amino acids 35–44; see Fig. 2) in the N-terminal part, (ii) a putative membrane-spanning region (amino acids 84–104; see Fig. 2), and (iii) a W-(X_{13})-(N/D)-P motif (amino acids 132–147; see Fig. 2), previously identified in [10], in the C-terminal part. We conclude, therefore, that *C. fasciculata* cox 8 is the homologue of mammalian cox IV, the yeast cox Va, *N. crassa* cox V and *D. discoideum* cox VI. This conclusion is fully supported by a comparison of the hydropathy plots of the *C. fasciculata* subunit 8 and yeast subunit Va, as shown in Fig. 3. The profiles are very similar, and only in the C-terminal section some minor differences can be observed. The alignment further shows that the homologues differ in size and that in addition to length differences at the N- and C-termini, *C. fasciculata* cox 8 contains two extra stretches of about 10 amino acids each, flanking the putative membrane-spanning region, which are absent from the other sequences (see Figs. 2 and 3).

It is unclear which of the three in-frame AUG's upstream of the (codon encoding the) N-terminal amino acid of the mature *C. fasciculata* protein is used as a translation start site (Fig. 2). All of the possible presequences have several basic residues, but only the two that would result from initiation at the second or third methionine can be folded into the amphipathic helix found in the import signal of mitochondrial proteins of other organisms. These presequences would be very short (eight or nine amino acids, respectively), but it has been proposed that these exceptionally short signal peptides are functional in kinetoplastids (see [16]). Our results show that some of the conserved amino acid motifs of a small cox subunit are also present in the trypanosomatid sequence. Given the large evolutionary distance between trypanosomatids and other eukaryotes and the overall low degree of conservation of the small cox subunits [7–10], these motifs must be functionally important.

Acknowledgements: We thank J. van den Burg for skilled technical assistance, Dr. P. Sloof and D. Blom for stimulating discussions and W. van Noppen for expert help in the preparation of the manuscript. This research is supported by The Netherlands Foundation for Chemical Research (SON) and the Foundation for Medical and Health Research (MW), which are subsidized by The Netherlands Foundation for Scientific Research (NWO).

References

[1] Capaldi, R.A. (1990) Arch. Biochem. Biophys. 280, 252–262.
[2] Calhoun, M.W., Thomas, J.W. and Gennis, R.B. (1994) TIBS 19, 325–330.
[3] Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T. et al. (1995) Science 269, 1069–1074.
[4] Fernandes, A.P., Nelson, K. and Beverley, S.M. (1993) Proc. Natl. Acad. Sci. USA 90, 11608–11612.
[5] Hajduk, S.L., Harris, M.E. and Pollard, V.W. (1993) FASEB J. 7, 54–63.
[6] Benne, R. (1994) Eur. J. Biochem. 221, 9–23.
[7] Carter, R.S. and Avadhani, N.G. (1991) Arch. Biochem. Biophys. 288, 97–106.
[8] Sachs, M.S., David, M., Werner, S. and RajBhandary, U.L. (1986) J. Biol. Chem. 261, 869–873.
[9] Cumsky, M.G., Trueblood, C.E., Ko, C. and Poyton, R.O. (1987) Mol. Cell. Biol. 7, 3511–3518.

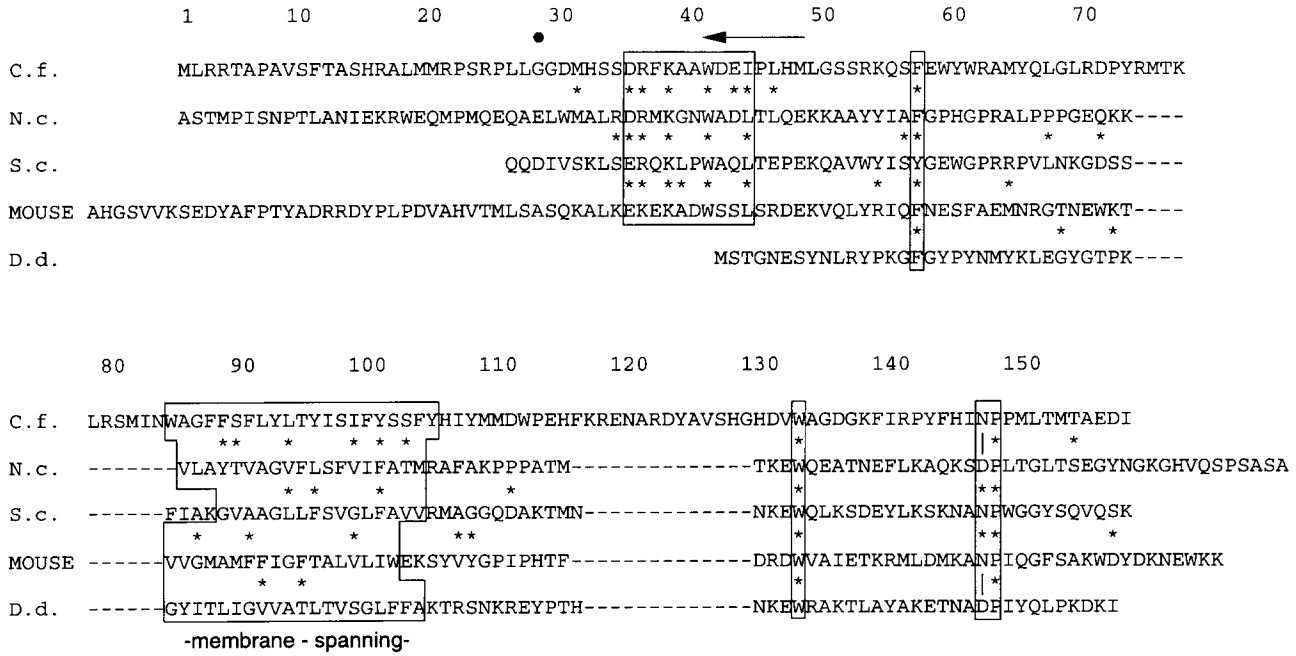

Fig. 2. Sequence alignment of cox subunit 8 from *C. fasciculata* with different homologues. Alignment of the sequence of *C. fasciculata* (C.f.) cox subunit 8 with subunit V from *N. crassa* (N.c.) [8], subunit Va from *Saccharomyces cerevisiae* (S.c.) [9], subunit IV from mouse [7] and subunit VI from *D. discoideum* (D.d.) [10]. An asterisk indicates conservation with respect to the *C. fasciculata* sequence, including conservative substitutions: D,E; K,R; F,Y; S,T; I,L,V. A conserved alignment of D/N in the C-terminal part is indicated by I. Highly conserved regions or residues are boxed. The *C. fasciculata* sequence starts with the methionine encoded by the first AUG, the sequences from the other organisms start with the N-terminal amino acid of the mature protein. The black dot indicates the N-terminal residue of the mature *C. fasciculata* protein, the arrow indicates the part used for the design of the PCR oligo. Insertions on both sides of the putative membrane-spanning regions to maximize homology are indicated by dashed lines.

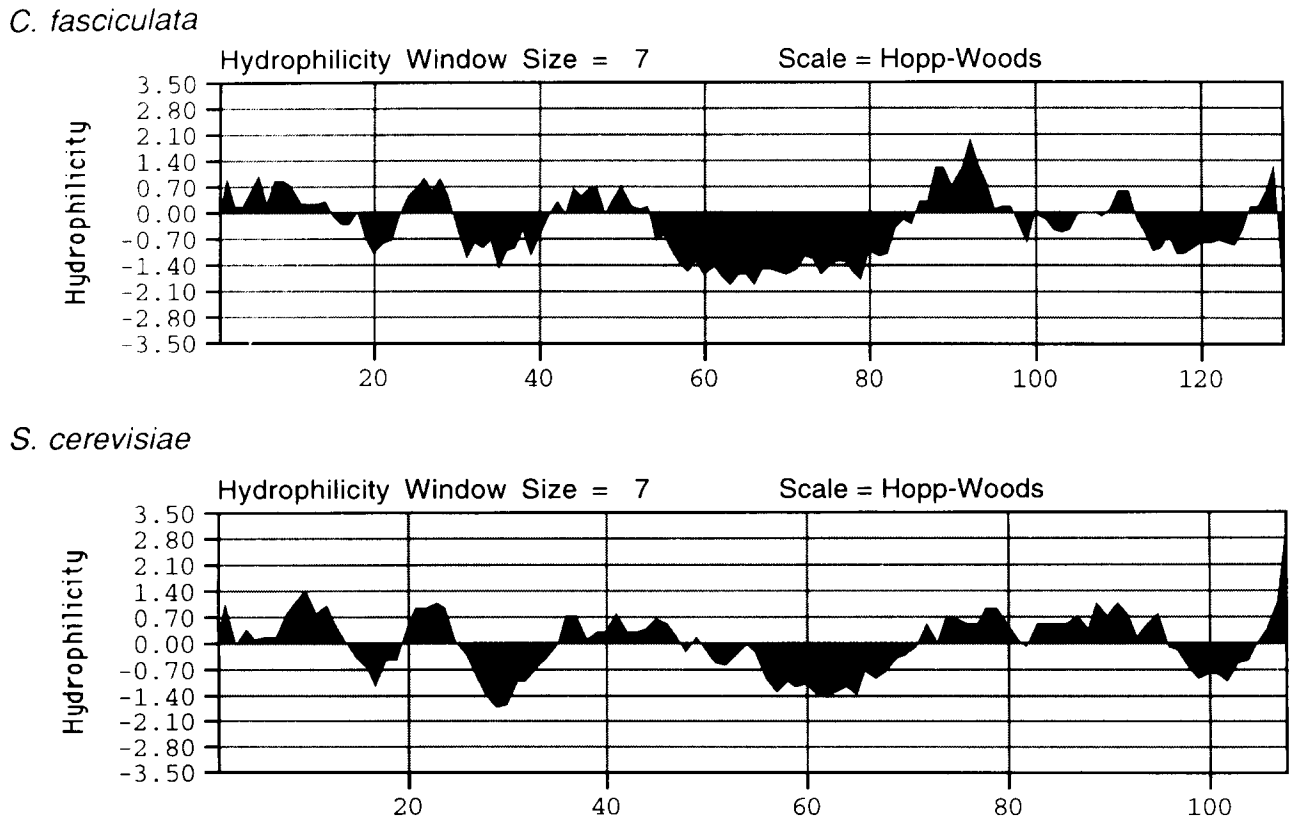

Fig. 3. Hydropathy plots of cox subunit 8 from *C. fasciculata* (top) and cox subunit Va from *Saccharomyces cerevisiae* (bottom) [9]. The plots were obtained by the method of Hopp–Woods as programmed in the MacVector software. The line at 0 value divides hydrophobic regions (below) from hydrophilic regions (above).

[10] Rizutto, R., Sandona, D., Capaldi, R.A. and Bisson, R. (1991) Biochim. Biophys. Acta 1089, 386–388.
[11] Birkenmeyer, L. and Ray, D.S. (1986) J. Biol. Chem. 261, 2362–2368.
[12] Muhich, M.L., Hughes, D.E., Simpson, A.M. and Simpson, L. (1987) Nucleic Acid Res. 15, 3141–3153.
[13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[14] van der Spek, H., van den Burg, J., Croiset, A., van den Broek, M., Sloof, P. and Benne, R. (1988) EMBO J. 7, 2509–2514.
[15] Schägger, H. and von Jagow, G. (1991) Anal. Biochem. 217, 220–230.
[16] Clayton, C., Häusler, T. and Blattner, J. (1995) Microbiol. Rev. 59, 325–344.