SHORT COMMUNICATION

Molecular Cloning and Characterization of the Human Mitochondrial NADH:Oxidoreductase 10-kDa Gene (NDUFV3)

René F. M. de Coo,*,†,[1] Paul Buddiger,* Hubert J. M. Smeets,*,[2] and Bernard A. van Oost*,[3]

*Department of Human Genetics, University Hospital Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands; and †Department of Child Neurology, University Hospital Rotterdam, P.O. Box 2060, 3000CB Rotterdam, The Netherlands

Received February 13, 1997; accepted July 29, 1997

The human gene for the 10-kDa flavoprotein subunit of the mitochondrial NADH:ubiquinone oxidoreductase (Complex I) was completely cloned and sequenced. The so-called NDUFV3 gene contains three exons, spanning 20 kb. The open reading frame contains a 34-codon import sequence and a 74-codon mature protein sequence. A database search revealed close homology to bovine and rat protein sequence but not to any other known protein. Northern blot analysis showed that the NDUFV3 gene is ubiquitously expressed. The NDUFV3 gene was assigned by FISH to a single location on chromosome 21q22.3 and might contribute to the Down syndrome phenotype.

NADH:ubiquinone oxidoreductase (complex I; EC 1.6.5.3) is an inner mitochondrial membrane-bound multisubunit enzyme complex. Complex I consists of at least 41 subunits of which 7 are encoded by the mitochondrial genome. As one of the complexes of the mitochondrial respiratory chain, its function is the catalysis of the rotenone-sensitive oxidation of NADH and the reduction of ubiquinone. Using chaotropic agents complex I can be resolved into two hydrophilic fractions, the flavoprotein fraction and the iron–protein fraction, and a hydrophobic fraction. The flavoprotein fraction comprises the 51-, 24-, and 10-kDa subunits, which are all encoded by nuclear genes. This fraction plays a catalytic role in the oxidation of NADH as it is associated with flavoprotein and NAD binding. The tetranuclear Fe-S center in the 51-kDa subunit and the binuclear Fe-S center in the 24-kDa subunit are involved in electron transfer (12). The function of the 10-kDa protein is unknown as yet, despite the fact that the primary structure of the bovine 10-kDa subunit is known (14).

Specific or generalized deficiencies in subunits of complex I have been described in a subgroup of patients with a mitochondrial encephalomyopathy. In some cases mutations were detected in the mitochondrial DNA, but in most cases a defect in one of the nuclear coding genes was suspected (4), although as yet it has not been found. To elucidate the possible pathological role for nuclear-encoded complex I genes, the human genes must be characterized first. We have chosen to begin with the functionally important flavoproteins. We reported earlier on the 24-kDa gene (3), and we now report on the genomic structure, expression, and localization of the human 10-kDa gene.

To isolate the 10-kDa cDNA a human renal λgt11 cDNA library was screened with a 10-kDa bovine cDNA probe (positions 135 to 361; EMBL/GenBank Accession No. X59048) (14). After successive rounds of screening, we obtained two positive plaques, from which the same 1600-bp insert was isolated by PCR with T3 and T7 primers. The fragment was analyzed by cycle sequencing and contained a hybrid cDNA sequence of which 273 bp coded for the carboxy-terminal region of the 10-kDa protein (Fig. 1B) and 1300 bp coded for an unknown sequence. To obtain the 5′ end of the human 10-kDa cDNA, nested primers derived from the human cDNA sequence (see Fig. 2 for the positioning of the primers) were applied in a modified Rapid Amplification of cDNA Ends (RACE) procedure (1). Three fragments were obtained from a human brain cDNA library. The longest fragment of 308 bp was cloned in pCR-Script SK(+) and sequenced. From the overlapping renal cDNA and the brain cDNA fragments (273 and 308 bp, respectively), a full-length 484-bp human cDNA was inferred (Figs. 1B and 1C).

The following procedures aided in clarifying the genomic structure and chromosomal assignment of the 10-kDa subunit gene. Southern blot analysis of an *Eco*RI digest of human genomic DNA with a nearly full-length cDNA probe corresponding to nucleotides 1 to 399 revealed two bands of 20 and 1.8 kb (Fig. 3). The human cDNA harbors an *Eco*RI restriction site at position 183 (Fig. 1). The 1.8-kb genomic fragment is

The human 10-kDa subunit gene has been assigned the symbol NDUFV3 by the Human Gene Mapping Workshop Nomenclature Committee.

[1] To whom correspondence should be addressed. Telephone: 31-104636342. Fax: 31-104636345. E-mail: decoo@neur.azr.nl.

[2] Present address: Division of Genetics, University of Limburg, P.O. Box 1475, 6201 BL Maastricht, The Netherlands.

[3] Present address: Department of Clinical Sciences of Companion Animals, University of Utrecht, P.O. Box 80154, 3508 TD Utrecht, The Netherlands.

GENOMICS **45,** 434–437 (1997)
ARTICLE NO. GE974930
0888-7543/97 $25.00

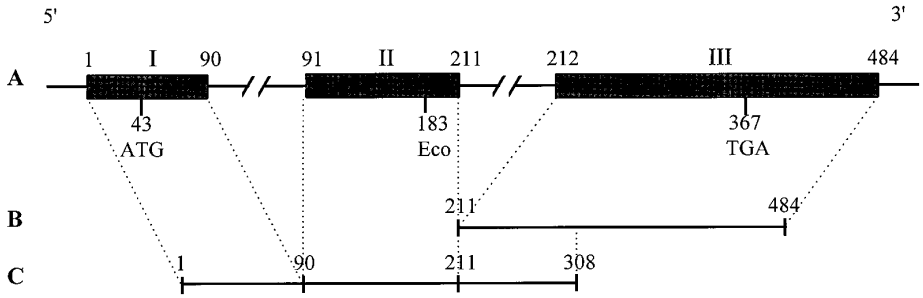

FIG. 1. Structure of the human 10-kDa subunit gene (NDUFV3). (**A**) Genomic structure. Nucleotide 1 is the first nucleotide according to the cDNA (See Fig. 2). Exons I–III, the positions of the intron–exon boundaries are given. ATG, initiation codon; TGA, termination codon; Eco, *Eco*RI restriction site cuts at position 183. (**B**) Partial 10-kDa cDNA sequence derived from the kidney cDNA library. (**C**) Partial 10-kDa cDNA sequence derived from brain cDNA by a RACE procedure.

situated 5′ of this *Eco*RI site of the gene and the 20-kb genomic fragment downstream, as a 3′ cDNA probe (nucleotides 212–484) showed hybridization to the 20-kb *Eco*RI fragment only (data not shown). Based on the Southern blot analysis the length of the gene is between 8 and 20 kb (Fig. 3). Attempts to perform a long PCR spanning intron 2 to refine the length estimation failed. The length of Intron 1 is less than 1.5 kb.

A Southern blot panel of human/hamster somatic cell hybrids (5) was analyzed using the 10-kDa subunit cDNA as a probe. There was a 100% concordance for a chromosome 21 location (data not shown). Subsequently, filters containing a chromosome 21-specific cosmid library (11) were screened with the human cDNA probe (nt 1–399). Two clones (ICRFc103C0564(10kdL) and ICRFc103G087 (8)) were found to be positive. These two cosmid clones were virtually identical by restriction mapping. Fluorescence *in situ* hybridization was performed to map the C0564 cosmid more specifically. The hybridization signal appeared consistently on chromosome 21q22.3 (Fig. 4). No significant signals were seen on other chromosomes or on other parts of chromosome 21. The C0564 cosmid was used as a template in a cycle-sequencing protocol with 10 different primers derived from the cDNA sequence. Analysis of the sequence obtained revealed the presence of three exons contiguous with the cDNA, varying in length between 90 and 273 bp and divided by two

```
  1 GAAGCTGCTG TGGCCCTGCT TGGTGCGCCC GCTGTCACCG CCATGGCTGC CCCGTGTTTG
                                                M  A   A  P  C  L
                                                                S  L

 61 CTGAGGAAAG GACGAGCCGG GGCGCTGAAG ACTATGCTCC AGGAAGCCCA GGTGTTTCGA
      L  R  K   G10 R  A   G  A  L  K   T  M  L   Q20 E  A   Q  V  F  R
            Q                                  V      L          G

121 GGACTTGCTT CTACGGTTTC TTTGTCTGCG GAATCAGGGA AGAGTGAAAA GGGTCAGCCA
      G  L  A   S30 T  V   S  L  S  A   E  S  G   K40 S  E   K  G  Q  P
         V      P  A                →                N              L

181 CAGAATTCCA AGAAGCAAAG TCCACCAAAA AAGCCAGCCC CA---GTGCC TGCTGAGCCG
      Q  N  S   K50 K  Q   S  P  P  K   K  P  A   P60    V   P  A  E  P
      P     P                                V   S  A  A       T

238 TTTGACAACA CTACCTACAA GAACCTGCAG CATCATGACT ACAGCACGTA CACCTTCTTA
      F  D  N   T  T70 Y   K  N  L  Q   H  H  D   Y  S80 T   Y  T  F  L

298 GACCTCAACC TCGAACTCTC AAAATTCAGG ATGCCTCAGC CCTCCTCAGG CCGGGAGTCA
      D  L  N   L  E90 L   S  K  F  R   M  P  Q   P  S100S   G  R  E  S
                  D

358 CCTCGACACT GAGGGCCCTC GGTGTGAAGA TGAACCTTCC ACCGTCTTCA CTGCATCCTG
      P  R  H108

418 GAGTGCAAAA ATAAAATCCA CTCAAGAGTC ACAAGGCCCG CTGTGCATAA TCGGTTTCAC

478 TTTTACC
```

FIG. 2. Human 10-kDa cDNA sequence and comparison of the deduced protein sequences of the precursors of the human and bovine 10-kDa proteins. The human cDNA sequence is shown in the upper row and the deduced human protein sequence in the middle; where the bovine protein sequence differs is shown below the human protein sequence in the bottom row. The human cDNA sequence and protein sequence are numbered. (→) The N terminus of the mature protein. The ATG iniator codon, the TGA termination codon, and the polyadenylation signal are shown in boldface. The primers used for the nested PCR in the RACE procedure are shaded. The human sequence shows a one-codon deletion compared to the bovine sequence at position 60.

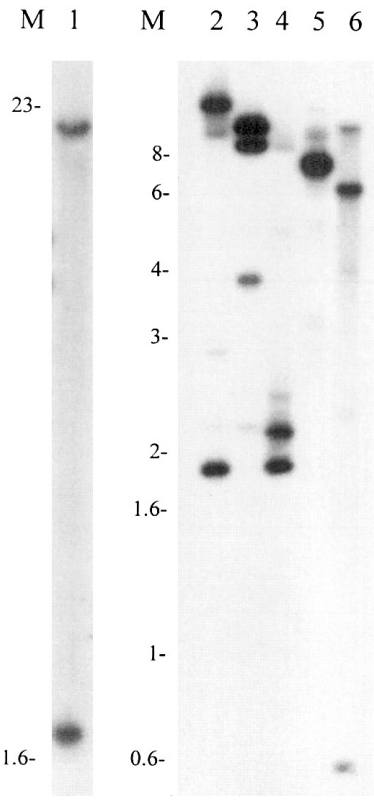

FIG. 3. Hybridization of *Eco*RI-restricted genomic DNA with the human 10-kDa subunit cDNA as a probe. This probe contains the entire coding region. Lane **1,** human genomic DNA; lanes **2–6,** cosmid clone C0564 restricted with, respectively, *Eco*RI (lane **2**), *Hin*dIII (lane **3**), *Pst*I (lane **4**), *Sst*I (lane **5**), and *Pvu*II (lane **6**). The size marker used is the 1-kb ladder from Gibco BRL.

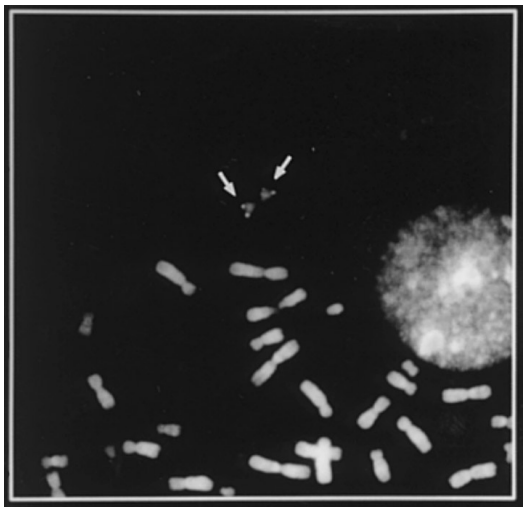

FIG. 4. Chromosome localization of NDUFV3 by fluorescence *in situ* hybridization on chromosome 21. Arrows, paired signals present at 21q22.3.

introns that follow the conventional intron splice donor and acceptor sequence rules. The in-frame ATG codon at position 43 is most likely the start codon, because the sequence upstream contains the consensus sequence for a eukaryotic translation initiation site, CCGCCATG (7). Exon I codes for the first 16 amino acids, exon II for amino acids 17 to 56, and exon III for amino acids 57 to 108. The 5′ untranslated region (UTR) is 42 nucleotides long and the 3′ UTR 115 nucleotides. The poly(A) tail is located 52 nucleotides downstream from the polyadenylation signal (Fig. 2). Immediately following the 3′ UTR an *Alu*-repeat sequence was detected. The exact size of the introns was not determined. In intron 2 we found four copies of a 75-bp repeat (data not shown). The sequence is deposited with the EMBL Nucleotide Sequence Database under Accession Nos. X99726–X99728.

The 10-kDa subunit mRNA is expressed ubiquitously and shows alternative splicing. Northern blot analysis of human poly(A)$^+$ RNA from 16 different human tissues (Fig. 5) showed a single fragment of about 500 nucleotides in each lane. A relatively strong signal was seen in skeletal muscle and heart. RNA from testis showed a single slightly larger fragment. A reverse transcriptase (RT)-PCR from human fibroblast RNA and from human placenta RNA with primers 904f and 821r showed two cDNA fragments, one fragment of full length and one fragment lacking exon II, as was found by direct sequencing of the RT-PCR product. These fragments were not visible on the Northern blot (Fig. 1).

In our ongoing research to clone nuclear genes that encode for mitochondrial complex I proteins (3), we have now characterized the human 10-kDa subunit gene. We describe here that there is only one 10-kDa gene in the human genome, which is ubiquitously expressed.

The human 10-kDa cDNA appeared to be quite similar to the bovine cDNA in a number of aspects. Both the human and bovine cDNA have an AATAAA polyadenylation signal, which is located, respectively, at 52 and 18 bases upstream of the poly(A) tail. The 5′ regions of both cDNAs contain a preponderance of C and G residues, indicative of CpG islands associated with the 5′ ends of the genes. The bovine coding region is 3 bases longer. Over the region of alignment, the nucleotide sequence differed in 55 of the 324 bases. Of these, 37 were within the mature protein coding region (aa 35–108) and 18 in the region encoding the import protein sequence (aa 1–34). This 83% identity with the bovine nucleotide sequence for the coding region is less than was found in other mitochondrial genes like the 24-kDa gene with 93% homology (3). The 3′ and the 5′ UTR are completely divergent, including a 307-bp insertion in the bovine 3′UTR compared to the human 3′UTR.

The cDNA divergence corresponds with the protein divergence, as the human and bovine protein sequences are 82% homologous. The mature bovine and human 10-kDa proteins both have the same N-terminal sequence (S A E S) (14). The 74-amino-acid mature protein sequence differs in 10 positions from the bovine protein with, as mentioned above, the bovine protein sequence being 1 amino acid longer. A further 9, mostly conservative, amino acid substitutions are present within the 34-amino-acid N-terminal presequence (Fig. 2).

As expected, the NDUFV3 gene is ubiquitously expressed. The different RT-PCR products from the human fibroblast and human placental RNA suggest alternative splicing of the NDUFV3 pre-mRNA, although this is not a major quantitative effect since only the full-length mRNA fragment is observed on the Northern blot. However, in testis tissue the major fragment appeared to be slightly longer than the 10-kDa mRNA in the other tissues. With a RT-PCR for the coding region and 32 nucleotides of the 3′UTR of the 10-kDa mRNA, no evidence for a length divergence was obtained. Further analysis of the 5′ and 3′ ends of the testis cDNA are warranted as a length difference may also be caused by a second polyadenylation signal situated more downstream.

The 10-kDa protein sequence is only moderately conserved between human and cow, compared to the highly homologous 24-kDa protein sequence (4) and the 51-kDa subunit (de Coo *et al.*, manuscript in prepara-

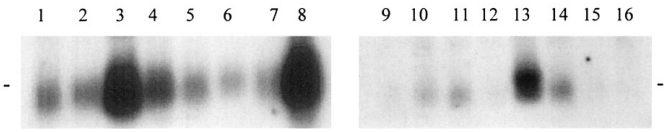

FIG. 5. Northern blot analysis. Hybridizations of the cDNA probe (nucleotides 1–399) from the NDUFV3 gene to 2 μg poly(A)$^+$ RNA from human tissues. Lanes **1,** pancreas; **2,** kidney; **3,** skeletal muscle; **4,** liver; **5,** lung; **6,** placenta; **7,** brain; **8,** heart; **9,** peripheral blood leukocyte; **10,** colon; **11,** small intestine; **12,** ovary; **13,** testis; **14,** prostate; **15,** thymus; **16,** spleen. RNA size marker band 0.5 kb are indicated at both sides.

tion). A search in the current databases did not reveal any proteins homologous to the 10-kDa subunit. The import presequence (Fig. 2) for the human protein is not extensively related in primary structure to other mitochondrial import sequences (10). The mitochondrial presequences are in general quite diverse, but have a net positive charge and rarely contain acidic amino acids. These two latter features are present in the putative import sequence of the 10-kDa protein.

The 10-kDa protein is a protein with unknown function and structure and is present in the mammalian complex I only (9). Within complex I it is situated in close proximity to the 24- and 51-kDa subunits (15), suggesting a function in the electron transport. However, the mature protein sequence does not contain any cysteine residues so there is no evidence that it participates in the liganding of a 2Fe-2S center necessary for the electron transfer, like the 24- and 51-kDa subunits. Probably it has a modifying or regulatory role in the electron transport. A defect in this subunit could thus be involved in mitochondrial encephalopathies, in particular in patients with a partial complex I deficiency. Isolation and characterization of the NDUFV3 gene allow us now to investigate these patients in detail.

The localization of the NDUFV3 gene to a single locus at chromosome 21q22.3 may provide another clue for its function since Down syndrome has been postulated to be a contiguous gene syndrome (6). Prince *et al.* (13) reported a mitochondrial enzyme defect in Down syndrome. This defect is present at an early stage in Down syndrome and might be related to the development of dementia. The 10-kDa subunit might contribute in a dosage-dependent manner to the phenotype of Down syndrome, similar as was suggested recently (2) to the mitochondrial ATP50 protein, which was located near the AML1 gene at the border between 21q22.1 and 21q22.2.

ACKNOWLEDGMENTS

The authors thank John E. Walker, MRC, Cambridge, UK for providing the bovine 10-kDa probe. We gratefully acknowledge The Reference Library, ICRF for providing the ICRFc103C0564 and ICRFc103G087 cosmid clones (8) and Ad Geurts van Kessel and Gerard Merx, Department of Human Genetics, University Hospital Nijmegen for performing the FISH analysis. This research was supported by Prinses Beatrix Fonds Grant 93-1105 to B.v.O. and R. Sengers, Department of Pediatrics, University Hospital Nijmegen and was also supported by the Ter Meulen Fonds.

REFERENCES

1. Apte, A., and Siebert, P. D. (1993). Anchor ligated cDNA libraries: A technique for generating a cDNA library for the immediate cloning of the 5′ ends of mRNAs. *Bio Techniques* **15:** 890–893.
2. Chen, H. M., Morris, M. A., Rossier, C., Blouin, J. L., and Antonarakis, S. E. (1995). Cloning of the cDNA for the human ATP synthase OSCP subunit (ATP50) by exon trapping and mapping to chromosome 21q22.1–q22.2. *Genomics* **28:** 470–476.
3. De Coo, I. F. M., Buddiger, P., Smeets, H., Geurts van Kessel, A., Morgan Hughes, J., Weghuis, D. O., Overhauser, J., and Van Oost, B. (1995). Molecular cloning and characterization of the active human mitochondrial NADH:ubiquinone oxidoreductase 24-kDa gene (NDUFV2) and its pseudogene. *Genomics* **26:** 461–466.
4. De Vries, D. D., De Coo, I. F. M., Buddiger, P., Ruitenbeek, W., Albrecht, R., and Van Oost, B. (1994). Isolated respiratory chain enzyme deficiency in patients with a mitochondrial (encephalo)-myopathy: Sequence analysis of the mitochondrial complex I and IV genes. *Am. J. Hum. Genet.* A358.
5. Geurts van Kessel, A., Tetteroo, P., Borne von dem, A., Hagemeijer, A., and Bootsma, D. (1983). Expression of human myeloid-associated surface antigens in human–mouse myeloid-associated surface antigens in human-mouse myeloid cell hybrids. *Proc. Natl. Acad. Sci. USA* **80:** 3748–3751.
6. Korenberg, J. R., Chen, X.-N., Schipper, R., Sun, Z., Gonsky, R., Gerwehr, S., Carpenter, N., Daumer, C., Dignan, P., Disteche, C., Graham, J. M., Jr., Hugdins, L., McGillivray, B., Miyazaki, K., Ogasawara, N., Park, J. P., Pagon, R., Pueschel, S., Sack, G., Say, B., Schuffenhauer, S., Soukup, S., and Yamanaka, T. (1994). Down syndrome phenotypes: The consequences of chromosomal imbalance. *Proc. Natl. Acad. Sci. USA* **91:** 4997–5001.
7. Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNA's. *Nucleic Acids Res.* **12:** 857–872.
8. Lehrach, H. (1990). Genetic and physical mapping. *In* "Genome Analysis" (K. E. Davies and S. M. Tilghman, Eds.), pp. 39–81, Cold Spring Harbor Laboratory Press, Cold Spring Habor, NY.
9. Masui, R., Wakabayashi, S., Matsubara, H., and Hatefi, Y. (1991). The amino acid sequence of the 9 kDa polypeptide and partial amino acid sequence of the 20 kDa polypeptide of mitochondrial NADH:ubiquinone oxidoreductase. *J. Biochem. Tokyo* **110:** 575–582.
10. Nicholson, D. W., and Neupert, W. (1988). "Protein Transfer and Organelle Biogenesis" (R. C. Das and P. W. Robbins, Eds.), pp. 677–746, Academic Press, New York.
11. Nizetic, D., Zehetner, G., Monaco, A. P., Gellen, L., Young, B. D., and Lehrach, H. (1991). Construction, arraying, and high-density screening of large insert libraries of human chromosomes X and 21; their potential use as reference libraries. *Proc. Natl. Acad. Sci. USA* **88:** 3233–3237.
12. Ohnishi, T., Ragan, C. I., and Hatefi, Y. (1985). EPR studies of iron-sulfur clusters in isolated subunits and subfractions of NADH-ubiquinone oxidoreductase. *J. Biol. Chem.* **260:** 2782–2788.
13. Prince, J., Jia, S., Båve, U., Annerén, G., and Oreland, L. (1994). Mitochondrial enzyme deficiencies in Down's syndrome. *J. Neural Transmission* **8:** 171–181.
14. Skehel, J. M., Pilkington, S. J., Runswick, M. J., Fearnley, I. M., and Walker, J. E. (1991). NADH:ubiquinone oxidoreductase from bovine heart mitochondria. Complementary DNA sequence of the import precursor of the 10 kDa subunit of the flavoprotein fragment. *FEBS Lett.* **282:** 135–138.
15. Yamaguchi, M., and Hatefi, Y. (1993). Mitochondrial NADH:ubiquinone oxidoreductase (complex I): Proximity of the subunits of the flavoprotein and the iron-sulfur protein subcomplexes. *Biochemistry* **32:** 1935–1939.