Am. J. Hum. Genet. 61:1318–1326, 1997

Mutations in *PDX1,* the Human Lipoyl-Containing Component X of the Pyruvate Dehydrogenase–Complex Gene on Chromosome 11p1, in Congenital Lactic Acidosis

Bernard Aral,[1] Chantal Benelli,[2] Ghania Ait-Ghezala,[1] Mohamed Amessou,[2] Françoise Fouque,[2] Catherine Maunoury,[1] Nicole Créau,[1] Pierre Kamoun,[1] and Cécile Marsac[1,3]

[1]Laboratoire de Biochimie Médicale B and Laboratoire d'Expression Génétique et Pathologie du Dévéloppement CNRS, URA1335, [2]Unité de Recherches sur le Mécanisme d'action cellulaire des hormones, INSERM, Unité 30, and [3]Laboratoire de Biochimie Pharmacologique et Métabolique, INSERM, Unité 75, IFREM, Institut Necker, Hôpital des Enfants Malades, Paris

Summary

We have identified and sequenced a cDNA that encodes an apparent human orthologue of a yeast protein-X component (*ScPDX1*) of pyruvate dehydrogenase multienzyme complexes. The new human cDNA that has been referred to as "*HsPDX1*" cDNA was cloned by use of the "database cloning" strategy and had a 1,506-bp open reading frame. The amino acid sequence of the protein encoded by the cDNA was 20% identical with that encoded by the yeast *PDX1* gene and 40% identical with that encoded by the lipoate acetyltransferase component of the pyruvate dehydrogenase and included a lipoyl-bearing domain that is conserved in some dehydrogenase enzyme complexes. Northern blot analysis demonstrated that the major *HsPDX1* mRNA was 2.5 kb in length and was expressed mainly in human skeletal and cardiac muscles but was also present, at low levels, in other tissues. FISH analysis performed with a P1-derived artifical chromosome (PAC)–containing *HsPDX1* gene sublocalized the gene to 11p1.3. Molecular investigation of PDX1 deficiency in four patients with neonatal lactic acidemias revealed mutations 78del85 and 965del59 in a homozygous state, and one other patient had no *PDX1* mRNA expression.

Introduction

The mammalian and yeast pyruvate dehydrogenase multienzyme complexes (PDHc) contain a tightly associated 50,000-kD polypeptide of unknown function—component X (PDX1)—in addition to its three constituent enzymes, pyruvate dehydrogenase (E1; E.C. 1.2.4.1), lipoate acetyltransferase (E2; E.C. 2.3.1.12), and lipoamide dehydrogenase (E3; E.C. 1.6.4.3), which are jointly responsible for production of acetyl-CoA and reduced nicotinamide adenine dinucleotide (Reed 1974; De Marcucci and Lindsay 1985; Jilka et al. 1986). It has been shown elsewhere that component X is an immunologically distinct entity (De Marcucci and Lindsay 1985; Jilka et al. 1986; Rahmatullah et al. 1989*b*) that is related to the E2 subunit, particularly in its N-terminal region; moreover, bovine component X resembles bovine E2 in that both proteins contain a lipoyl moiety that undergoes reductive acetylation and deacetylation (De Marcucci et al. 1986; Hodgson et al. 1986; Jilka et al. 1986; Rahmatullah and Roche 1987). There is also some sequence similarity in the amino-terminal segment of component X and E2 (Niu et al. 1988; Rahmatullah et al. 1989*a*). The role of component X has not yet been clearly demonstrated. Its primary role appears to be a structural one in mediating the binding and optimal orientation of E3 with regard to the E2 core assembly (Neagle and Lindsay 1991) interaction required for optimal catalytic efficiency of the complex (Powers-Greenwood et al. 1989). The importance of component X is highlighted by the severe consequences of its deficiency in humans. This autosomal recessive disorder manifests with a complex organic acidosis, a result of the accumulation of the α-keto acid substrates, and with varying degrees of neurological dysfunction that often result in death within the first few years of life (Robinson et al. 1990; Marsac et al. 1993; Geoffroy et al. 1996). The residual level of PDHc activity observed in deficient fibroblasts is <20% of normal. In accordance with previous in vitro studies of reconstituted mammalian PDHc when the component-X subunit has been selectively degraded by proteolysis (Sanderson et al. 1996), the immunoblotting studies, performed with antibodies against PDHc, clearly have demonstrated the absence of component X in some patients. The other components ap-

Received August 13, 1997; accepted for publication October 15, 1997; electronically published December 5, 1997.

Address for correspondence and reprints: Dr. Bernard Aral, Faculté de Médecine Necker, CNRS URA1335 156, rue de Vaugirard 75015, Paris. France. E-mail: aral@necker.fr

0002-9297/97/6106-0016$02.00

pear normal, compared with those control samples. In addition, SSCP analysis and sequencing of the whole coding region excludes the involvement of an X-linked PDH-E1α gene abnormality as an explanation of the low PDHc activity. Moreover, the fact that PDHc activities are normal in the mother and the father, with a normal component X on western immunoblot analysis, strongly suggests that these are primary defects in these families. We report here the identification of the human cDNA sequence for the X-lipoyl–containing component of the PDHc, its tissue-specific expression, and the chromosomal localization of its gene. Examination of four patients presenting with the neonatal form of lactic acidosis revealed three mutations and two frameshift deletions, and one patient had no detectable *HsPDX1* transcript.

Subjects and Methods

Subjects

Patient S.D., diagnosed with a neonatal lactic acidemia, was a female newborn born from unrelated healthy nonconsanguineous parents after a normal pregnancy and delivery. A case report detailing the clinical course and biochemical investigation of this patient and her family has been described elsewhere (Geoffroy et al. 1996). The patient is under a physician-supervised regimen of a low-glucose, ketogenic diet, which has dramatically improved metabolic status, and she is presently doing well at 1 year of age. Patients R.D. and R.H. were brothers; the parents, of Portuguese origin, were first cousins. The brothers presented with encephalomyelopathy and chronic lactic acidemia and had a normal E1 subunit of PDHc but deficiency of PDX1 protein. Both patients went into a deep coma with irreversible severe lactic acidosis and died of cardiopulmonary arrest. These cases are described in the report by Marsac et al. (1993). The fourth case of neonatal lactic acidemia with a PDHc defect caused by the specific absence of component X is patient K.M., who was born after an uneventful pregnancy, from healthy nonconsanguineous parents. He was dysmorphic, with trigonocephaly, a frontal metopic ridge, and a supranasal lipoma. He was developmentally delayed from birth. He became severely quadriplegic and microcephalic. At the age of 10 mo high levels of serum lactate and pyruvate were found. A computed-tomography scan of the brain revealed a partial frontal corpus callosum agenesis with dilated lateral ventricles. The full case report has been published by De Meirleir et al. (in press).

Database Searches

Homology searches for various expressed sequence tags (ESTs) were performed on data from the DNA Database of Japan (http://www.nig.ac.jp/home.html), the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), and the European Molecular Biology Laboratories (EMBL) (http://www.ebi.ac.uk/); and comparative genomics and genome cross-referencing were performed on data from XREFdb (http://www.ncbi.nlm.nih.gov/XREFdb/) and on a database of precalculated comparison (http://protein.toulouse.inra.fr/.html).

cDNA Cloning and Sequence Analysis

We obtained I.M.A.G.E. Consortium cDNA clones, corresponding to human ESTs H58032 and AA159537, from Genome Systems. The composite human sequence derived from sequencing these EST clones extended from bp +1 (where +1 is the A of the initiation methionine) to the start of the poly(A) tail (bp +1744). The *HsPDX1* cDNA sequence (Y13145) that we deposited into EMBL is a full-length fetal liver/spleen 1NFLS clone (bp +1 to bp +1611) identical to one derived from an adult pancreatic-tissue clone. Verification of the 5′ end (bp +1 to bp +151) was obtained by 5′-RACE (rapid amplification of cDNA ends) analysis using the human Pancreas Marathon-Ready cDNA (Clontech-Ozyme, France) and according to the manufacturer's cDNA-amplification protocol. Sequencing was performed with the AutoRead™ DNA Sequencing Kit (Pharmacia Biotech), which uses T7 DNA polymerase and an A.L.F.™ automated DNA sequencer (Pharmacia Biotech). The cDNA clones were sequenced completely, in both directions. Probe labeling was performed by use of random hexamers or gene-specific primers as described elsewhere (Feinberg and Vogelstein 1983).

RNA Isolation, Northern Blot Analysis, and Chromosome Mapping

We isolated total RNA from cultured skin fibroblasts, either by guanidium thiocyanate extraction (Chirgwin et al. 1979) or with Trizol™ reagent (Life Technologies–BRL). Chromosomal assignment of the *HsPDX1* was performed by use of in situ hybridization as described elsewhere (Liu et al. 1986). Probe labeling was performed by use of random hexamers or gene-specific primers as described elsewhere (Feinberg and Vogelstein 1983). Standard procedures were used for P1-derived artificial chromosome (PAC) library screening (Sambrook et al. 1989). PAC DNAs were prepared according to the method of Wang et al. (1994). H58032 cDNA

(bp +1 to bp + 1744) was used to probe the northern blots (Clontech-Ozyme); poly(A) binding protein (PABP) and lipoate acetyltransferase component (i.e., E2) cDNAs were controls.

Western Blot Analysis

Fibroblast or β-lymphoblastoid–cell line mitochondria (5 μg) were subjected to 10% SDS-PAGE under reducing conditions. Proteins were electrotransferred onto nitrocellulose membranes at 100 V for 2 h in 25 mM Tris and 192 mM glycine buffer pH 8.3. Membranes were incubated for 1 h in PBS containing 2% nonfat dry milk, 1% ethanolamine, and 0.02% thimerosal; were washed in PBS and 0.1% Tween-20; and incubated overnight at 4°C with rabbit anti-PDHc antibody (1/2,000 dilution) diluted in PBS, 1% BSA, 0.05% Tween-20, and 0.02% thimerosal. After an extensive washing, membranes were incubated for 1 h at room temperature with horseradish peroxidase–conjugated goat anti-rabbit immunoglobulin G (1/3,000). Immunocomplexes were detected by the Amersham ECL detection reagents.

Reverse Transcriptase–PCR (RT-PCR) Analysis

RT-PCR was performed by use of the Superscript™ RT protocol (Life Technologies–BRL), 50 ng primer XR2 (located in the 3′ UTR), and 3 μg total fibroblast RNA. A 1–3-μl portion of the first-strand cDNA was amplified with primers F2/DW2 and UP2/XR (described below). The RNA isolated from each cell line was reverse transcribed, amplified, and sequenced a minimum of three times, to insure the authenticity of mutations found on the sequencing of amplified transcripts. The PCR products were gel purified and sequenced directly by use of a Sequitherm™ sequencing kit (Epicentre Technologies–TEBU). Unless otherwise specified, PCR reactions contained 75 ng template DNA, 10 pmol each primer, 200 mM each dNTP, 0.25 U *Taq* DNA polymerase, and 1 × PCR buffer (1.5 mM $MgCl_2$; Life Technologies) in 50 μl. Standard cycling conditions were used (Aral et al. 1996).

GenBank Accession Numbers

GenBank accession numbers are as follows: human ESTs—H58032, AA159537, and T16961; *HsPDX1* cDNA—Y13145; *ScPDX1*—M28222; and PDC-E2—Y00978 and X13969.

Results

"Database Cloning" of HsPDX1 *cDNA*

The yeast *Saccharomyces cerevisiae* component X, ScPDX1, has been cloned and sequenced elsewhere (Behal et al. 1989). Furthermore, study of the *ScPDX1* gene has demonstrated that its disruption results in the dissociation of the E3 component from the mutant complex, concomitant with a loss of pyruvate-oxidation capacity (Lawson et al. 1991). Deletion analysis also has shown that the putative subunit-binding domain (residues 144–180) plays an important role in the binding of E3 (Lawson et al. 1991). To identify the molecular basis of PDHc deficiency with a missing component X, we used XREFdb (Bassett et al. 1995, 1997) to probe the database of ESTs by use of the *S. cerevisiae* PDX1 protein sequence. Initially, we identified three overlapping human partial-length candidate cDNAs (GenBank accession numbers H58032, AA159537, and T16961, from fetal liver/spleen, pancreas, and infant-brain libraries, respectively), with *P* values of 1.2^{-8}, 4.2^{-8}, and 1.8^{-10}, respectively. The composite sequence had a *P* value of 1.0^{-24} and contained an open reading frame (ORF) that we estimated to be ~100% of the full length, on the basis of comparison to both *ScPDX1* and the human lipoate acetyltransferase (PDC-E2) cDNA. From the 2.5-kb *HsPDX1* cDNA (GenBank accession number Y13145), we sequenced 1,744 bp and found no 5′ UTR sequence, a 238-bp 3′ UTR, and a 1,506-bp ORF encoding a 501-amino-acid protein (fig. 1). The initiation methionine cannot be predicted Kozak (1992), because of quasi-absent 5′ UTR sequence in all sequenced cDNA clones (clones 205084 and 591560). Therefore, using the EST sequence information (H58032 and AA159537, respectively) from these clones, we set out to clone the 5′ end of the full-length *HsPDX1* cDNA, by use of the 5′ RACE (Frohman et al. 1988; Frohman 1993) strategy, from a pancreas Marathon-Ready™ (Clontech-Ozyme) cDNA library (Chenchik et al. 1995). We were able to amplify a unique cDNA fragment by use of the two gene-specific primers (1R and 2R), on the one hand, and 5′ adaptor-specific primers, provided with the library, on the other hand (data not shown). PCR-amplified 5′-end cDNA fragments were subjected to direct cycle sequencing by the primer-walking strategy, from the known sequences at the 3′ end of the cDNA toward the unknown sequences at the 5′ end of the cDNA. The sequence data from the 5′-RACE fragment gave no additional sequence information for the 5′ UTR. Since PDX1 is a mitochondrial protein, it must undergo a proteolytic removal of its mitochondrial leader peptide. According to Hendrick et al. (1989), a highly conserved 3-amino-acid motif is common to this class of leader peptides. This motif includes an arginine at position −10, a hydrophobic res-

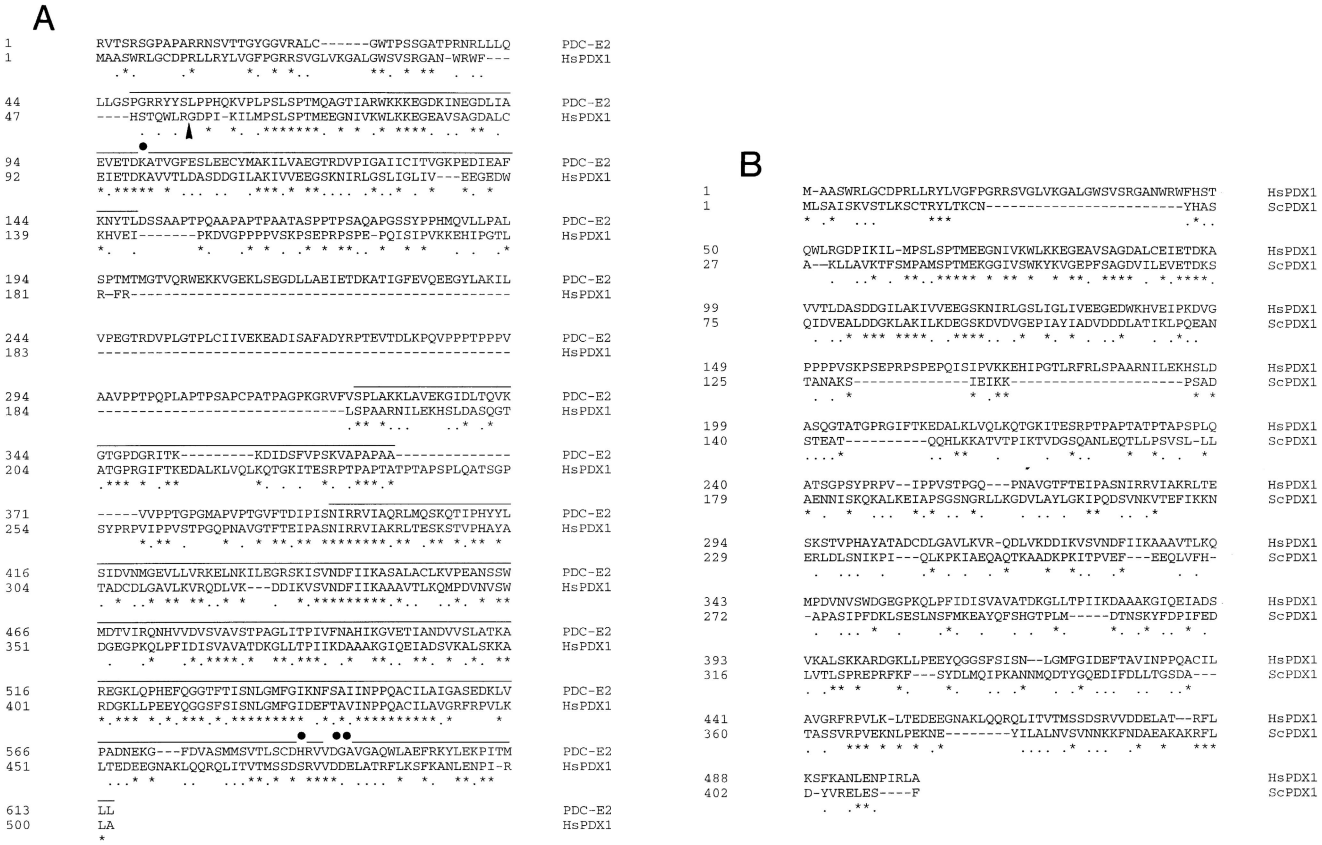

Figure 1 Comparison of amino acid sequence (single-letter code) of HsPDX1 with that of PDC-E2 (*A*) and with that of ScPDX1 (*B*). The deduced sequences of HsPDX1, PDC-E2, and ScPDX1 are aligned for maximum similarities, by the sequence-alignment tool Clustal V. Identical amino acids are indicated by an asterisk (*); and similar amino acids are indicated by periods (.). The arrowhead indicates the putative cleavage site of the HsPDX1 leader peptide. Both the lipoyllysine residue in the lipoyl domain and the residues in the putative active site of PDC-E2 are denoted by black dots. The lipoyl domains (amino-terminal), E_3-binding domains, and catalytic inner-core domains (carboxy-terminal) are denoted by unbroken lines.

idue at position −8, and serine, threonine, or glycine at position −5, relative to the mature amino terminus. Using this criteria, we hypothesized that the putative cleavage site of the PDX1's leader peptide, by matrix-processing protease, must occur between residues at 53 and 54, such that arginine at position −10, threonine at position −5, and phenylalanine, a hydrophobic residue, at position −8 may account for the conserved 3-amino-acid motif. The presequence is rich in basic and hydroxylated amino acid residues and is devoid of acidic amino acid residues. This feature, similar to the putative presequence of PDC-E2 (Thekkumkara et al. 1988), is characteristic of presequences of mitochondrial-matrix/inner-membrane proteins (Von Heijne 1986). Comparison of the deduced amino acid sequences of human PDX1 (HsPDX1) and experimentally (partially sequenced) determined N-terminal amino acid sequences of bovine heart PDX1 (Rice et al. 1992), reveals 100% sequence identity. Comparison of the deduced amino acid sequences of HsPDX1 and PDC-E2 reveals that both the amino-terminal part (residues 55–131) and the carboxyl-terminal part, including catalytic inner-core domains (residues 258–501) of HsPDX1 resemble E2 (fig. 1*A*). The two proteins exhibit 50% sequence identity in the amino-terminal segment that corresponds to the putative lipoyl-bearing domain of E2 (residues 77–111). This extensive homology indicates that HsPDX1 and PDC-E2 evolved from a common ancestor. In contrast to PDC-E2, ScPDX1 and HsPDX1 share sequence identity only in the lipoyl-bearing domain (fig. 1*B*). The overall amino acid identity between HsPDX1 and PDC-E2 is 40%, and that between HsPDX1 and ScPDX1 is 20%.

Secondary-Structure Predictions of HsPDX1 Protein

In addition to "live" on-line homology searching, we probed precalculated databases with the primary amino acid sequence of the HsPDX1. We obtained multiple

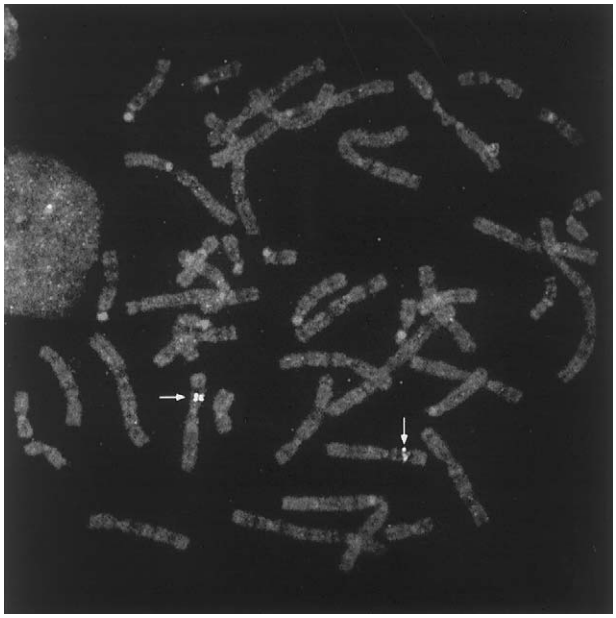

Figure 2 FISH to human metaphase chromosomes with the PAC clone F1172, containing the human gene for PDX1.The human gene encoding PDX1 is assigned to human chromosome 11p1.3.

alignments with consensus sequences derived for two protein domains (data not shown). The first domain of homology on the HsPDX1 lies between residues 77 and 111 in the N-terminal moiety, and it is shared among three proteins—namely, the dihydrolipoamide acetyltransferase component (i.e., E2) of the pyruvate dehydrogenase complex, the dihydrolipoamide succinyltransferase component (i.e., E2) of the 2-oxoglutarate dehydrogenase complex precursor (E2K; E.C. 2.3.1.61), and the biotin carboxyl–carrier protein of acetyl-CoA carboxylase (BCCP; E.C. 6.4.1.2). This domain bears two highly conserved amino acid residues, K_{97} and G_{108}, and corresponds perfectly to the putative lipoyl-bearing domain. The second domain of homology lies between the amino acid residues 258 and 501 in the HsPDX1 C-terminal moiety that contains the highly conserved sequence D-H-R-X-X-D-G, which is thought to be part of the putative catalytic site of all dihydrolipoamide acetyltransferases and succinyltransferases, such as the lipoamide acetyltransferase component (i.e., E2) precursor of the branched-chain α-keto acid dehydrogenase complex (BCKAD E2 subunit; E.C. 2.3.1.–), the dihydrolipoamide acetyltransferase component (i.e., E2) of the pyruvate dehydrogenase complex (E.C. 2.3.1.12), and the dihydrolipoamide succinyltransferase component (i.e., E2) of the 2-oxoglutarate dehydrogenase complex precursor (E2K; E.C. 2.3.1.61). The HsPDX1 has, instead, the D_{474}-S-R-X-X-D-D_{480} motif, which suggests that HsPDX1, as its yeast orthologue, in contrast to PDC-E2, is not able to catalyze an acetyl transfer between the protein-bound S-acetyldihydrolipoyl moiety and coenzyme A. But this assertion has yet to be determined experimentally. It should also be noted that, like its yeast orthologue, HsPDX1 protein does contain an E3-binding domain (fig. 1*A*, residues ≈326–361 of PDC-E2), as in any other α-keto acid dehydrogenase complexes, in which E3 binds directly to the E2 core.

Tissue-Specific Expression Pattern of HsPDX1 *mRNA*

Because mitochondria are present in virtually all mammalian cells, we utilized a multitissue northern blot as an initial test of the tissue-expression pattern of a presumed *PDX1* gene. As expected for a gene involved in the intermediary metabolism of a ubiquitous organelle, an *HsPDX1* cDNA probe detects a transcript (≈2.5 kb) in all tissues examined, in a very low abundance (data not shown). Tissues with the highest expression are skeletal muscle and heart.

Chromosomal Localization of the HsPDX1 *Gene*

We next attempted to determine the chromosomal localization of the gene that encodes the HsPDX1. As a first step, we used a 535-bp PCR probe from the *HsPDX1* cDNA in order to screen and isolate human PAC clones containing the human gene encoding HsPDX1, from a whole-genome PAC library spotted on high-density filters (Lehrach et al. 1990). We isolated five positive clones (LLNLP704C05279Q13, LLNLP704F1172Q13, LLNLP704L10149Q13, LLNLP704L10167Q13, and LLNLP704G24136Q13). Two of these five clones, F1172 and G24136, were used to prepare PAC DNA probes for FISH to metaphase chro-

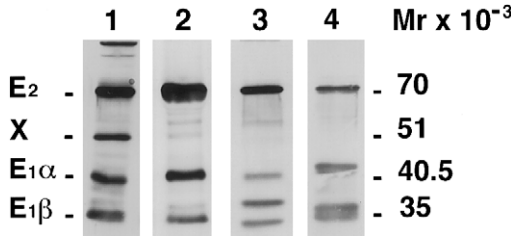

Figure 3 Western blotting of cultured skin fibroblast mitochondria with anti–PDH complex antibodies. Mitochondria (≈5 μg protein) from each cell line were solubilized in sample buffer and were resolved on a 10% SDS–polyacrylamide gel, under reducing conditions. The proteins were then electroblotted onto a nitrocellulose support matrix and were probed with antibodies that reacted with all subunits of the purified PDHc, except for the E3 subunit, which is poorly immunogenic. Immunoreactive proteins were visualized by electrogenerated chemiluminescence-detection reagents. Lane 1, Control cell lines. Lane 2, Patient S.D. Lane 3, Patient K.M. Lane 4, Patient R.D.

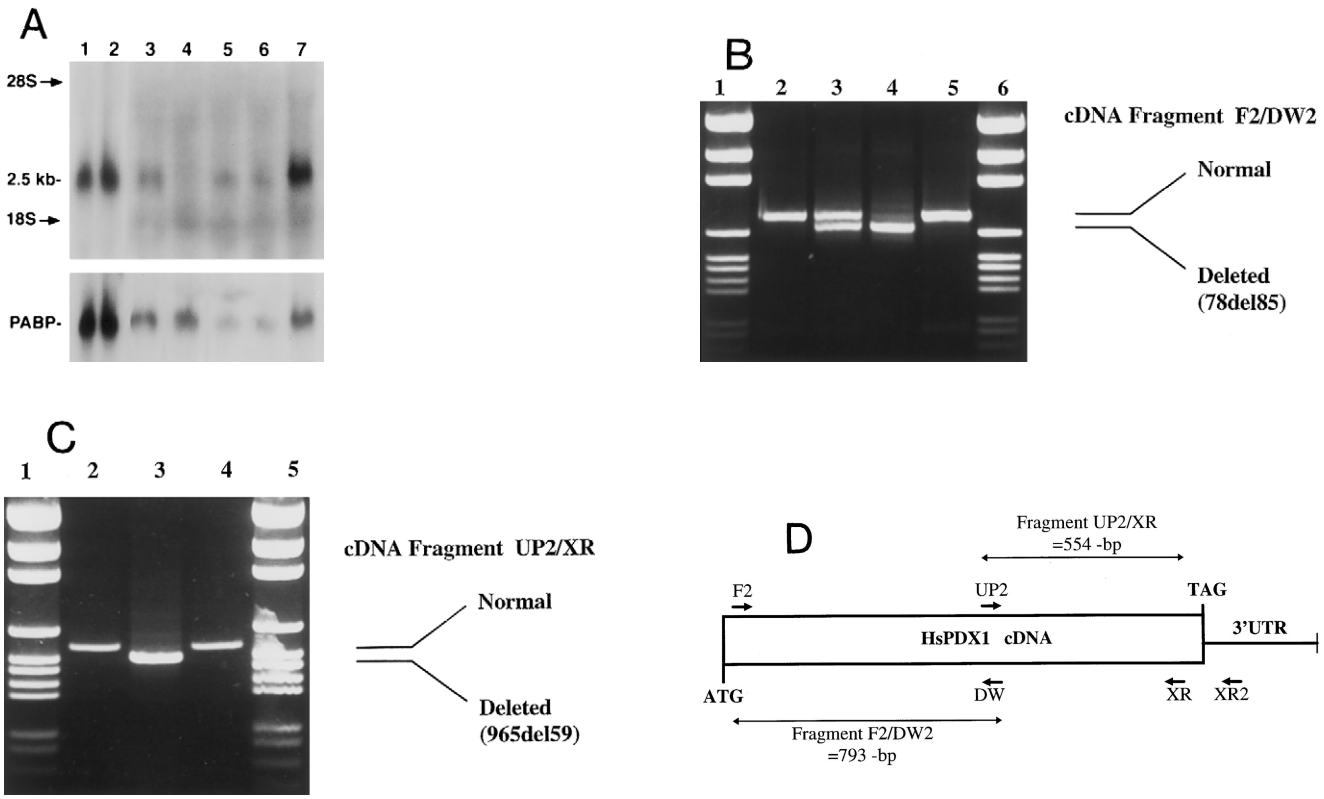

Figure 4 Identification of mutations in the HsPDX1 mRNA transcripts in PDH-deficient patients. *A, top,* Northern blot of total RNA (20 mg total RNA/lane) from controls' and patients' fibroblasts probed with [^{32}P]-*HsPDX1* cDNA. A 2.5-kb PDX1 transcript is present in six of the seven lanes, which show results for control β-lymphoblastoid cell lines (lanes 1 and 2) and for fibroblast cell lines (lane 3, patient K.M.'s mother; lane 5, patient S.D.; and lanes 6 and 7, controls); the exception was patient K.M. (lane 4), whose PDX1 mRNA was undetectable. *A, bottom,* Same blot, probed with a PABP cDNA probe to control for the quality and quantity of the RNA. *B,* Detection of deleted RT-PCR–amplified cDNA fragments F2/DW2 by agarose gel electrophoresis. RT-PCR (using primers F2 and DW2; for the primer sequences, see the text) was used to amplify a 793-nt-long cDNA fragment (between bp +37 and bp +828) from two unaffected individuals (lanes 2 and 5), a 708-nt-long cDNA fragment from patient S.D. (lane 4), and, in a heterozygous state, both a 793-nt-long fragment and a 708-nt-long cDNA fragment from patient S.D.'s mother (lane 3). Standards of known size also were run (lanes 1 and 6). *C,* Detection of abnormal RT-PCR–amplified cDNA fragments UP2/XR by agarose gel electrophoresis. RT-PCR (using primers UP2 and XR; for the primer sequences, see the text) was used to amplify a 554-nt-long cDNA fragment (between bp +808 and bp +1362) from two unaffected individuals (lanes 2 and 4) and a 495-nt-long cDNA fragment from patient R.D. (lane 3). Standards of known size also were run (lanes 1 and 5). *D,* Schematic representation of *HsPDX1* cDNA, along with the primers' positions for RT-PCR amplifications.

mosomes from a healthy male. Under these experimental conditions, a specific signal appearing as symmetrical spots on both chromatids of chromosome 11p1.3 was obtained in >90% of the 50 metaphases examined (fig. 2), with both F1172 and G24136.

HsPDX1-*Gene Mutation Analysis*

We confirmed the functional impairment of the HsPDX1 in two unrelated patients and two male siblings, by immunoblotting studies (fig. 3) performed with antibodies against PDHc components, which clearly demonstrated the absence of component X. This result also implicated *HsPDX1* as the gene responsible for the clinical and biochemical signs of lactic acidemia in a preliminary molecular investigation of these deficiencies. For the mutation analysis of *HsPDX1* in a series of two unrelated patients and in two male siblings with neonatal lactic acidemias, we first determined, by northern blot analysis, the expression levels of *HsPDX1* mRNA in patients' fibroblasts probed with a complete *HsPDX1* cDNA probe. The size and amount of *HsPDX1* mRNA in patient cells was similar to that of the controls when sample loading differences were accounted for, except for patient K.M., who has little if any detectable transcript, suggesting a transcriptional regulation abnormality (fig. 4*A*). Also, this reduction supports the prediction of involvement of PDX1 in neonatal lactic acidemia and predicts heterogeneity in the responsible mutations. We then utilized RT-PCR to amplify *HsPDX1* cDNA from fibroblast RNA isolated from patients' fibroblasts, by using, for the first half of the

cDNA (cDNA fragment F2/DW2), primers F2 (sense, 5′-CTGCTGCGTTATCTTGTGGGC T-3′) and DW2 (antisense, 5′-TCAGTGAATGTGCCCACTGCATTG-3′), which amplify a 793-bp fragment (bp +37 to bp +828) and, for the second half of the cDNA (cDNA fragment UP2/XR), UP2 (sense, 5′-CAATGCAGTGGGCACATTCACTGA-3′) and XR (antisense, 5′-CCTCTTCATCCTCAGTGAGCTTCA-3′), which amplify a 554-bp fragment (bp +808 to bp +1362); the total length of the amplified cDNA includes ~90% of the ORF (fig. 4*D*). First, we found that the yield of RT-PCR product amplified from the fibroblasts belonging to patient K.M. was reduced compared with that in control fibroblasts, although the starting amounts of total RNA material were equalized in all individuals (data not shown). Moreover, the *HsPDX1* cDNA fragments amplified from the cells of patient S.D. and of the two male siblings R.D. and R.H were decreased in size, compared with those in controls. We sequenced the products directly and found that patient K.M. had a normal cDNA sequence, whereas patient S.D. was homozygous for an 85-bp deletion between bp +77 and bp +162, producing the 78del85 mutation (data not shown); patient S.D.'s mother was heterozygous for this mutation (fig. 4*B*). In the two male siblings, R.D. and R.H., the size of the amplified cDNA fragment was also decreased compared with that in controls (fig. 4*C*). Direct sequencing of this cDNA revealed a 59-nt deletion between bp +964 and bp +1024, producing a 965del59 mutation (data not shown). Neither the 85-bp-deletion allele nor the 965del59 allele was detected in the 30 controls.

Discussion

Evolutionary conservation of the enzymes belonging to the various metabolic pathways, together with extensive molecular studies of the mutant variants in yeast and the rapidly expanding database of expressed mammalian genes, provides the resources necessary to identify the human orthologues of these enzymes' genes (Bassett et al. 1995, 1997). Strategies based on these resources have already proved successful (Bronner et al. 1994; Fishel et al. 1994; Aral et al. 1996; Dodt et al. 1996). In the present study, we utilized a "database cloning" strategy with *ScPDX1* to identify orthologous human cDNAs. The *HsPDX1* gene has been localized to human chromosome 11p1.3 by in situ hybridization; the *PDC-E2* gene has also been localized to human chromosome 11 (Genome Database accession number 750238). So, from a molecular-evolutionary point of view, it may be of interest that two components of the same enzyme complex, having strong similarities in their amino acid sequences, are on the same chromosome. Secondary-structure predictions for HsPDX1 protein clearly demonstrated its role in maintaining the three-dimensional structure of the PDHc, by predicting the correct association of E3 to E2 as a core assembly; these interactions between E2 and PDX1 and E3 are required for an optimal catalytic efficiency of the complex. We identified two major deleterious *pdx1* mutant alleles—namely, 78del85 and 965del59—and one other patient had a severely reduced level of *HsPDX1* mRNA. The 78del85 causes an 85-nt deletion starting with nucleotide 78 in cDNA, leading to a frameshift on translation and predicting production of a truncated PDHX1 of 25 amino acids, lacking the remaining 476 amino acids. The 965del59 causes a 59-nt deletion starting with nucleotide 965 in cDNA, leading to a frameshift on translation and predicting production of a truncated PDHX1 of 321 amino acids, lacking the remaining 180 amino acids. Further in vitro expression studies will be necessary to confirm the impact of 78del85 and 965del59 on the function of HsPDX1 inside the PDHc. Futhermore, we have not completed a comprehensive survey of the *HsPDX1*-gene mutations and expect to find more mutations in our cohort of 30 patients who have documented neonatal lactic acidemia and who, on the basis of immunoblot analysis, have PDHc deficiency with a normal E1α along with normal size and amount of HsPDX1. For example, the mutation(s) responsible for the severe reduction of *HsPDX1* mRNA in patient K.M. is unknown. The identification of *HsPDX1* as the gene deficient in neonatal lactic acidemia in one group of PDHc-deficient patients allows prenatal diagnosis and carrier detection in at-risk families. Furthermore, this approach shows, once more, the power, in the search for human disease genes, of the combined efforts resulting from the various investigations that are as diverse as the clinical biochemistry, yeast genetics, and homology screening of the continuously expanding databases of ESTs.

Acknowledgments

We thank the patients, families, and clinicians—in particular Drs. J. Cousin, J. M. Saudubray, and D. Lindhent—who have contributed to this project and without whom this study would not have been possible. We thank The Resource Centre/Primary Database of the German Human Genome Project (Berlin), for providing human PAC clones; J. G. Lindsay (University of Glasgow, Glasgow), for providing antibodies against pyruvate dehydrogenase complex purified from pig heart; and W. Lissens and L. De Meirleir (Vrije Universiteit Brussel, Brussels), for performing SSCP screening in the E1α gene of patients presented in this study. We thank F. Degoul, M. Diry, and C. Cepanec for their technical assistance. This work was supported by the Association Française contre les Myopathies and

by Programme Hospitalier de Recherche Clinique contract CRC 94055.

References

Aral B, Schlenzig JS, Liu G, Kamoun P (1996) 'Database cloning' human Δ^1-pyrroline 5-carboxylate synthetase (P5CS) cDNA: a bifunctional enzyme catalyzing the first two steps in proline biosynthesis. C R Acad Sci III 319:171–178

Bassett DE Jr, Boguski MS, Spencer F, Reeves R, Goebl M, Hieter P (1995) Comparative genomics, genome cross-referencing and XREFdb. Trends Genet 11:372–373

Bassett DE Jr, Boguski MS, Spencer F, Reeves R, Kim SH, Weaver T, Hieter P (1997) Genome cross-referencing and XREFdb: implications for the identification and analysis of genes mutated in human disease. Nat Genet 15:339–344

Behal RH, Browning KS, Hall TB, Reed LJ (1989) Cloning and nucleotide sequence of the gene for protein X from *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 86: 8732–8736

Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, Kane M, et al (1994) Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 368:258–261

Chenchik A, Moqadam F, Siebert P (1995) Marathon cDNA amplification: a new method for cloning full-length cDNAs. CLONTECHniques 1:5–8

Chirgwin JM, Przbyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochem J 18:5294–5299

De Marcucci OG, Hodgson JA, Lindsay JG (1986) The M_r-50 000 polypeptide of mammalian pyruvate dehydrogenase complex participates in the acetylation reactions. Eur J Biochem 158:587–594

De Marcucci O, Lindsay JG (1985) Component X: an immunologically distinct polypeptide associated with mammalian pyruvate dehydrogenase multi-enzyme complex. Eur J Biochem 149:641–648

De Meirleir L, Lissens W, Benelli C, Marsac C, De Klerck H, Scholte J, van Diggelen O, et al. Pyruvate dehydrogenase complex deficiency and absence of subunit X. J Inherit Metab Dis (in press)

Dodt G, Braverman N, Valle D, Gould SJ (1996) From expressed sequence tags to peroxisome biogenesis disorder genes. Ann N Y Acad Sci 804:516–523

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M, et al (1994) The human mutator gene homolog MSH2 and its association with hereditary non-polyposis colon cancer. Cell 75:1027–1038 (erratum, Cell 77:167)

Frohman MA (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary cDNAs: thermal RACE. Methods Enzymol 218:340–358

Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNA from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85:8998–9002

Geoffroy V, Fouque F, Benelli C, Poggi F, Saudubray JM, Lissens W, Meirleir LD, et al (1996) Defect in the X-lipoyl-containing component of the pyruvate dehydrogenase complex in a patient with neonatal lactic acidemia. Pediatrics 97:267–272

Hendrick JP, Hodges PE, Rosenberg LE (1989) Survey of amino-terminal proteolytic cleavage sites in mitochondrial precursor proteins: Leader peptides cleaved by two matrix proteases share a three-amino acid motif. Proc Natl Acad Sci USA 86:4056–4060

Hodgson JA, De Marcucci OG, Lindsay JG (1986) Lipoic acid is the site of substrate-dependent acetylation of component X in ox heart pyruvate dehydrogenase multienzyme complex. Eur J Biochem 158:595–600

Jilka JM, Rahmatullah M, Kazemi M, Roche TE (1986) Properties of a newly characterized protein of the bovine kidney pyruvate dehydrogenase complex. J Biol Chem 261: 1858–1867

Kozak M (1992) Regulation of translation in eukaryotic systems. Annu Rev Cell Biol 8:197–225

Lawson JE, Behal RH, Reed LJ (1991) Disruption and mutagenesis of the *Saccharomyces cerevisiae* PDX1 gene encoding the protein X component of the pyruvate dehydrogenase complex. Biochemistry 30:2834–2839

Lehrach H, Drmanac R, Hoheisel J, Larin Z, Lennon G, Monaco AP, Nizetic D, et al (1990) Hybridization fingerprinting in genome mapping and sequencing. In: Davies KE, Tilghman SM (eds) Genome analysis. Vol 1: Genetic and physical mapping. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 39–81

Liu G, Maunoury C, Kamoun P, Aral B (1996) Assignment of the human gene encoding the Δ^1-pyrroline-5-carboxylate synthetase (P5CS) to 10q24.3 by in situ hybridization. Genomics 37:145–146

Marsac C, Stansbie D, Bonne G, Cousin J, Jehenson P, Benelli C, Leroux JP, et al (1993) Defect in the lipoyl-bearing protein X subunit of the pyruvate dehydrogenase complex in two patients with encephalomyelopathy. J Pediatr 123:915–920

Neagle JC, Lindsay JG (1991) Selective proteolysis of the protein X subunit of the bovine heart pyruvate dehydrogenase complex: effects on dihydrolipoamide dehydrogenase (E3) affinity and enzymic properties of the complex. Biochem J 278:423–427

Niu XD, Browning KS, Behal RH, Reed LJ (1988) Cloning and nucleotide sequence of the gene for dihydrolipoamide acetyltransferase from *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 85:7546–7550

Powers-Greenwood SL, Rahmatullah M, Radke GA, Roche TE (1989) Separation of protein X from the dihydrolipoyl transacetylase component of the mammalian pyruvate dehydrogenase complex and function of protein X. J Biol Chem 264:3655–3657

Rahmatullah M, Gopalakrishnan S, Andrews P, Chang CL, Radke GA, Roche TE (1989*a*) Subunit associations in the mammalian pyruvate dehydrogenase complex: structure and

role of protein X and the pyruvate dehydrogenase component binding domain of the dihydrolipoyl transacetylase component. J Biol Chem 264:2221–2227

Rahmatullah M, Gopalakrishnan S, Radke GA, Roche TE (1989*b*) Domain structures of the dihydrolipoyl transacetylase and the protein X components of mammalian pyruvate dehydrogenase complex: selective cleavage by protease Arg C. J Biol Chem 264:1245–1251

Rahmatullah M, Roche TE (1987) The catalytic requirements for reduction and acetylation of protein X and the related regulation of various forms of resolved pyruvate dehydrogenase kinase. J Biol Chem 262:10265–10271

Reed LJ (1974) Multienzyme complexes. Acta Chem Res 7: 40–46

Rice JE, Dunbar B, Lindsay JG (1992) Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. EMBO J 11:3229–3235

Robinson BH, MacKay N, Petrova-Benedict R, Ozalp I, Coskun T, Stacpoole W (1990) Defects in the E2 lipoyl transacetylase and the X-lipoyl containing component of the pyruvate dehydrogenase complex in patients with lactic acidemia. J Clin Invest 85:1821–1824

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanderson SJ, Khan SS, McCartney RG, Miller C, Lindsay JG (1996) Reconstitution of mammalian pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase complexes: analysis of protein X involvement and interaction of homologous and heterologous dihydrolipoamide dehydrogenases. Biochem J 319:109–116

Thekkumkara TJ, Ho L, Wexler ID, Pons G, Liu TC, Patel MS (1988) Nucleotide sequence of a cDNA for the human dihydrolipoamide acetyltransferase component of human pyruvate dehydrogenase complex. FEBS Lett 240:45–48

Von Heijne G (1986) Mitochondrial targeting sequences may form amphiphilic helices. EMBO J 5:1335–1342

Wang LF, Voysey R, Yu M (1994) Simplified large-scale alkaline lysis preparation of plasmid DNA with minimal use of phenol. Biotechniques 17:26–28