Characterization of a cDNA clone for human NAD⁺-specific isocitrate dehydrogenase α -subunit and structural comparison with its isoenzymes from different species

Yong-Ou KIM,* Il-Ung OH,* Hee-Sung PARK,† Jiingjau JENG,‡ Byoung J. SONG‡ and Tae-Lin HUH§

*Department of Genetic Engineering, College of Natural Sciences, Kyungpook National University, Taegu 702-701, Korea,

†Department of Plant Breeding and Genetics, Hyosung Women's University, Kyungsan, Kyungpook 713-702, Korea,

and ‡Laboratory of Neurogenetics, NIAAA, 12501 Washington Avenue, Rockville, MD 20852, U.S.A.

A 0.6 kb cDNA fragment encoding the human NAD⁺-specific isocitrate dehydrogenase α -subunit (H-IDH α) was amplified by PCR using oligonucleotide primers synthesized on the basis of pig tryptic peptide sequences [Huang and Colman (1990) *Biochemistry* 29, 8266–8273]. With the amplified cDNA as a probe, cDNA clones for IDH α were isolated from a human heart λ gt11 cDNA library. The deduced protein sequence of the largest cDNA clone (2628 bp) rendered a precursor protein of 366 amino acids (39 591 Da) and a mature protein of 339 amino acids (36 640 Da). The deduced H-IDH α protein sequence is highly similar to the partial peptide sequences of the pig enzyme. It is 55, 43 and 44% identical with yeast NAD⁺-specific IDH2, yeast NAD⁺-specific IDH1 and monkey NAD⁺-specific IDH γ -subunit (IDH γ) respectively. However, it has less similarity (about 30%) to NADP⁺-specific IDH from *Escherichia coli* and bovine mitochondria. These results indicate that the structure of

IDH α closely resembles that of IDH2, the catalytic subunit of the yeast enzyme. Structural analysis of the deduced H-IDH α protein revealed that the amino acids responsible for the binding of isocitrate, Mg²⁺ and NAD⁺ are highly conserved. It also has two conserved motifs for the binding sites of ATP and ADP, but a canonical Ca²⁺-binding motif was not recognized. Unusual penta- (ATTTA) and tri- (TAA or ATT) nucleotides which are respectively believed to interact with RNA-binding proteins and be near the endonuclease cleavage sites were frequently recognized in its 3' untranslated region, indicating the possibility of an additional method of regulation of this enzyme. Northern-blot analysis suggests that one mRNA transcript (2.8 kb) exists in cultured HeLa cells. Genomic DNA Southern-blot analysis indicates that the IDH α gene is not closely related to that of the other IDH isoenzymes, and IDH α appears to be encoded by a single gene.

INTRODUCTION

Isocitrate dehydrogenases (IDHs; EC 1.1.1.41 and EC 1.1.1.42) catalyse the oxidative decarboxylation of isocitrate to 2-oxoglutarate and are considered to be key enzymes of the tricarboxylic acid cycle [1,2]. It is known that only one type of IDH (the NADP⁺-specific form) exists and is involved in isocitrate decarboxylation in *Escherichia coli* [2–5]. In yeast, the decarboxylation of isocitrate in the tricarboxylic acid cycle is primarily accomplished by NAD⁺-specific IDH [6] although both the NAD⁺- and NADP⁺-specific enzymes are present in the mitochondria. On the other hand, mammalian tissues contain three kinds of IDH isoenzymes: mitochondrial NAD⁺-specific IDH, mitochondrial NADP⁺-specific IDH and cytosolic NADP⁺-specific IDH [1,2]. Of these, NAD⁺-specific IDH is presumed to play a major role in the oxidative decarboxylation of isocitrate in the tricarboxylic acid cycle as its activity is modulated by numerous allosteric effectors [7–9]. For instance, it is positively regulated by AMP in yeast and ADP in mammals, but is inhibited by ATP, NADH or NADPH [10]. Bivalent cations [1] are also required for the formation of the substrate complex and subsequent catalytic activity (dehydrogenation and decarboxyl-

ation of isocitrate) [11]. The amino acid residues involved in the binding of substrate and cofactors in the *E. coli* enzyme have been determined by X-ray crystallography [12] and other biochemical studies [11–15].

In eukaryotic systems, IDHs have been isolated to homogeneity from yeast and other mammalian sources [16–18]. In contrast with the homodimeric structure of mitochondrial NADP⁺-specific IDH (IDPm), mitochondrial IDH consists of several polypeptides of non-identical subunits. The regulatory IDH1 subunit and catalytic IDH2 subunit [19] in yeast have an octameric structure, whereas the pig enzyme exists as a heterotetramer of 2α , 1β and 1γ subunits, all of which have similar molecular masses of about 39–41 kDa. The catalytically active form of IDH has been proposed to be a heterodimer consisting of either $\alpha\beta$ - or $\alpha\gamma$ -subunits [20,21]. Despite numerous biochemical and kinetic studies of this enzyme, its complete structure and subunit interactions have not been addressed, particularly in the mammalian system. In order to elucidate the biochemical interaction of the three subunits in the catalytic activity of IDH, we have isolated cDNA clones for the human IDH α -subunit (H-IDH α) using a DNA fragment amplified by PCR. Partial peptide sequences reported for the pig heart enzyme [21] were used to

Abbreviations used: IDH, isocitrate dehydrogenase; IDH α , NAD⁺-specific IDH α -subunit; H-IDH α , human IDH α ; P-IDH α , pig IDH α ; IDH β , NAD⁺-specific IDH β -subunit; M-IDH γ , monkey NAD⁺-specific IDH γ -subunit; Y-IDH1, yeast IDH1; Y-IDH2, yeast IDH2; E-IDH, *Escherichia coli* NADP⁺-specific isocitrate dehydrogenase; AUBF, adenosine-uridine-binding factor; IDPm, mitochondrial NADP⁺-specific IDH; UTR, untranslated region.

§ To whom correspondence should be addressed.

The nucleotide sequence for H-IDH α will appear in EMBL/GenBank/DBJ Nucleotide Sequence Databases under the accession number U07681.

design the oligonucleotide primers used for PCR. We here report, for the first time, the complete nucleotide and deduced protein sequences for H-IDH α . A structural comparison of IDH isoenzymes from different species further reveals that the H-IDH α appears to represent the catalytic subunit of the mitochondrial NAD⁺-specific IDH.

MATERIALS AND METHODS

Materials

Restriction enzymes, T4 DNA ligase, nick translation kit and agarose were purchased from Bethesda Research Laboratories (Gaithersburg, MD, U.S.A.). GeneScreen, [α -³²P]dCTP (specific radioactivity 3000 Ci/mmol) and [α -³⁵S]thio]dATP (specific radioactivity 500 Ci/mmol) were from DuPont-New England Nuclear (Boston, MA, U.S.A.). Nitrocellulose filter was obtained from Schleicher & Schuell (Keene, NH, U.S.A.). Other chemicals were of the highest grade commercially available.

Amplification of cDNA

On the basis of two internal peptide sequences of pig heart IDH α [21], two oligonucleotide primers were synthesized using a DNA synthesizer (Applied Biosystems, Foster City, CA, U.S.A.): primer 1 (sense direction), 5'-GCCAACGTTCTCGTCCCTGTGTGTCTAT-3' corresponding to the peptide sequence ANVR-PCVSI, and primer 2 (anti sense direction), 5'-CTCTTCAGTGAAGTCTGAGCA-3' for the peptide sequence CSDFTTE. Amplification of DNA fragments was performed according to the manufacturer's protocol using the GeneAmp DNA-amplification kit (Perkin-Elmer-Cetus Instruments). PCR (50 μ l of reaction mixture containing 100 ng of total phage DNA prepared from human heart λ gt11 cDNA library) was conducted in a DNA thermal cycler (Perkin-Elmer-Cetus Instruments) as follows: 35 cycles of the amplification step (94 °C, 1 min; 55 °C, 1 min; 72 °C, 2 min) followed by the extension step for 10 min at 72 °C. Amplified PCR products were cloned into pCITE T-vector (Novagen, Madison, WI, U.S.A.), and their nucleotide sequences were determined by double-strand DNA sequencing [18]. Of three amplified cDNA fragments, one had a deduced amino acid sequence identical with the peptide sequences of the pig enzyme [21]. This was then used as a probe for further isolation of the full-length cDNA clones for H-IDH α .

Screening of the cDNA library and nucleotide sequence analysis

The amplified DNA was labelled with [α -³²P]dCTP by nick translation and used to screen the human heart λ gt11 cDNA library by plaque hybridization. Conditions of hybridization and washing were as previously described [18]. The DNA inserts of the positive cDNA clones were subcloned into plasmid pGEM7(+) and their nucleotide sequences were analysed from both ends by the dideoxynucleotide-chain-termination method with Sequenase V2.0 kit (United States Biochemicals, Cleveland, OH, U.S.A.). Subclones with gradually deleted lengths from the largest cDNA insert were constructed using a kit (Erase-a-Base from Promega, Madison, WI, U.S.A.); the protocol suggested by the manufacturer was followed. Nucleotide sequence data were assembled and analysed by using the computer software PC/GENE (IntelliGenetics, Mountain View, CA, U.S.A.).

Northern-blot analysis

Polyadenylated RNA (2 μ g) prepared from cultured HeLa cells was separated on 0.66 M formamide/1 % agarose gel, transferred to GeneScreen membrane and hybridized with ³²P-labelled

cDNA probe. Prehybridization, hybridization and washing conditions were as previously described [18].

Southern-blot analysis

Genomic DNA from human placenta was purified by the method of Davis et al. [22]. Human placental genomic DNA (10 μ g) was digested overnight with various restriction endonucleases. The DNA fragments were separated by 0.7 % agarose-gel electrophoresis, transferred to GeneScreen membrane, and hybridized with ³²P-labelled cDNA probe. The hybridization and washing conditions were the same as those described for Northern-blot analysis. Other methods were as previously described [18].

RESULTS

Amplification of cDNA for H-IDH α

Three different DNA fragments were amplified by PCR using primers 1 and 2. DNA sequencing of these amplified cDNAs and subsequent structural analysis (results not shown) revealed that only one DNA fragment (0.6 kb) encoded the pig peptide sequence (ANVRPCVSIIEGYK) [21]. Thus the amplified cDNA fragment represented a part of the cDNA clone for H-IDH α .

Isolation and characterization of cDNA clones for H-IDH α

In order to determine the primary structure of H-IDH α , we isolated cDNA clones using a 0.6 kb PCR-amplified cDNA as a probe. On screening more than 500 000 plaques from a human heart λ gt11 cDNA library, ten positive cDNA clones (inserts ranging from 0.9 to 2.7 kb) were identified and further plaques were purified until independent cDNA clones were obtained. The nucleotide structure of the largest cDNA clone (2.7 kb insert) was determined by double-strand sequencing. As shown in Figure 1, H-IDH α cDNA was 2628 bp long with an open reading frame (1106 bp) for the entire protein region of IDH α , starting with the initiation codon, ATG, and ending with the termination codon TAA. It also contained a long stretch (1522 bp) of 3' untranslated region (3'-UTR) with three potential sites of polyadenylation signal, AATAAA. In addition, two pentanucleotides ATTTA (AUUUA in the mRNA), and trinucleotides AAT or TAA, are frequently present in the 3'-UTR. These unusual polynucleotides are known to form a complex with AUBF [24] or generate secondary-loop structures, which can be cleaved by cellular endonucleases [25,26]. The presence of these unusual nucleotides in the 3'-UTR of H-IDH α mRNA suggests the possibility of another regulatory mechanism via its mRNA stability.

As shown in Figure 1, the N-terminus of the deduced H-IDH α protein sequence (Thr¹-Gln³⁵) is identical with the N-terminal 35-amino acid sequence of pig IDH α (P-IDH α) [21] except for the first and fifth amino acid residues. Therefore the precursor H-IDH α protein contains 366 amino acids (39 591 Da), 27 of these being a mitochondrial signal peptide, and the mature protein is made up of 339 amino acids (36 640 Da). These values agree well with the apparent size of the IDH α protein [20,21].

Structural analysis of H-IDH α protein

The deduced H-IDH α protein sequence was highly similar to the cysteinyl peptide sequences of P-IDH α [21]. However, on the basis of our cDNA sequence, the peptide sequence Va of the pig protein, CSDFTTEICR, should be reassigned to the C-terminal region of H-IDH α , not adjacent to the peptide sequence EVAENCKDIK as originally proposed [21]. The H-IDH α sequence showed 55, 43 and 44 % similarity to the IDH isoenzymes from

	ATG GCT GGG CCC GCG TGG ATC TCC AAG GTC TCT CGG CTG CTG GGG GCA TTC CAC	AA GCG	-1
	M A G P A W I S K V S R L L G A F H N P	AAC CCA	60
-27	AAA CAG GTG ACC AGA GGT TTT ACT GGT GGT GTT CAG ACA GTA ACT TTA ATT CCA GGA GAT		120
	K Q V T R G F T G V Q T V T L I P G D		
-7	GGT ATT GGC CCA GAA ATT TCA GCT GCA GTT ATG AAG ATT TTT GAT GCT GCC AAA GCA CCT		180
	G I G P E I S A A V M K I F D A A K A P		
14	ATT CAG TGG GAG GAG CGG AAC GTC ACT GGC ATT CAA GGA CCT GGA GGA AAG TGG ATG ATC		240
	I Q W E E R N V T A I Q G P G G K W M I		
34	CCT TCA GAG GCT AAA GAG TCC ATG GAT AAG AAC AAG GGC TTG AAA GGC CCT TTG AAG		300
	P S E A K E S M D K N K M G L K G P L K		
54	ACC CCA ATA GCA GCC GGT CAC CCA TCT ATG AAT TTA CTG CTG CGC AAA ACA TTT GAC CTT		360
	T P I A A G H P S M N L L R K T F D L		
74	TAC GCG AAT GTC CGA CCA TGT GTC TCT ATC GAA GGC TAT AAA ACC CCT TAC ACC GAT GTA		420
	Y A N V R P C V S I		
94	AAI ATT GTG ACC ATT CGA GAG AAC ACA GAA GGA TAC GGA ATT GAG CAT GTG ATT		480
	N I V T I R E N T E G E Y S G I E H V I		
114	GTT GAT GGA GTC GTG CAG AGT ATC AAG CTC ATC ACC GAG GGG GCG AGC AAG CGC ATT GCT		540
	V D G V V Q S I K L I T E G A S K R I A		
134	GAG TTT GCC TTT GAG TAT GCC CGG AAC AAC CAC CGG AGC AAC GTC ACG GCG GTG CAC AAA		600
	E F A F E Y A R N N H R S N V T A V H K		
154	GCC AAC ATC ATG CGG ATG TCA GGT GGT TTT CTA CAA AAA TGC AGG GAA GTT GCA GAA		660
	A N I M R M S D G L F L Q K C R E V A E		
174	AGC TGT AAA GAT ATT AAA TTT AAT GAG ATG TAC CTT GAT ACA GTA TGT TTT AAT ATG GTA		720
	S C K D I K F N E M Y L D T V C L N M V		
194	CAA GAT CCT TCC CAA TTT GAT GTT CTT GTT ATG CCA AAT TTG TAT GGA GAC ATC CTT AGT		780
	Q D P S Q F D V L V M P N L Y G D I L S		
214	GAC TTG TGT GCA GGA TTG ATC GGA GGT CTC GGT GTG ACA CCA AGT GGC AAC ATT GGA GCC		840
	D L C A G L I G G L G V T P S G N I G A		
234	AAT GGG GTT GCA ATT TTT GAG TCG GTT CAT GGG ACG GCT CCA GAC ATT GCA GGC AAG GAC		900
	N G V A I F E S V H G T A P D I A G K D		
254	ATG GCG AAT CCC ACA GCC CTC CTG CTC AGT GCG GTG ATG ATG CTG CGC CAC ATG GGA CTT		960
	M A N P T A L L L S A V M M L R H M G L		
274	TTT GAC CAT GCT GCA AGA ATT GAG GCT GCG TGT TTT GCT ACA ATT AAG GAC GGA AAG AGC		1020
	F D H A A R I E A A C F A T I K D G K S		
294	TTG ACA AAA GAT TTG GGA GGC AAT GCA AAA TGC TCA GAC TTC ACA GAG GAA ATC TGT CGC		1080
	L T K D L G G N A K		
314	CGA GTA AAA GAT TTA GAT TAA CAC TTC TAC AAC TGG CAT TTA CAT CAG TCA CTC TAA ATG		1140
	R V K D L D		
334	GAC ACC ACA TGA ACC TCT GTT TAG AAT ACC TAC GTA TGT ATG CAT TGG TTT GCT TGT TTC		1200
	TTG ACA GTA CAT TTT TAG ATC TGG CCT TTT CTT AAC AAA ATC TGT GCA AAA GAT GCA GGT		1260
	GGA TGT CCC TAG GTC TGT TTT CAA AGA ACT TTT TCC AAG TGC TTG TTT TAT TTA TTA AGT		1320
	GTC TAC CTG GTA AAT GTT TTT TTT GTA AAC TCT GAG TGG ACT GTA TCA TTT GCT ATT CTA		1380
	AAC CAT TTT ACA CTT AAG TTA AAA TAG TTT CTC TTC AGC TGT AAA TAA CAG GAT ACA GAA		1440
	TTA ACA AGA GAA AAT GTC TAA CTT TTT AAG AAA AAC CTT ATT TTC TTC GGT TTT TGA AAA		1500
	ACA TAA TGG AAA TAA AAC AGG ATA TTG ACA TAA TAG CAC AAA ATG ACA CTC TTC TAA AAC		1560
	TAA ATG GGC ACA AGA GAA TTT TCC TGG GAA AGT TCA CAT CAA AAA GAG TGA ATT TGG TAT		1620
	ATT TCT AAA TGA TAT GGA AAA TAG AGA CAG ATT TGT CCT TTA CAG AAA TTA CTG AGT GTG		1680
	AAT AAA AAC TTC AGA TCC AAG AAA TAT ATA ATG AGA GAT ATA ATT TTT GTT AAT AAG ACA		1740
	AAG GTA ATA TAT TGG ATA CAA AGA CAC AAA TGT ATT GTG TGT TCA ATT ATT TTG TTG TCT		1800
	TGA GAT TTA ATA TTC TTT CCA AGA GCT TTT AAT GAA GCA GAG AGC TAG TAC TTC ATT TTC		1860
	ACT GGA TAC ATT TTC AGC ATC ATG AGT TGT CAC AGC CTC TGA GCC CCT GAT CTG AAG CCA		1920
	GAA GGG CTG AGT GTA TTG TAA ACT TAT TCT TGC ATG TTG CTG TCT GGG AAT GGA CCA CAC		1980
	TAC AGC AGG TAG TTC TGG GGG CGA TAC TGC CGA AAG GCC CGA ACA CAT GTA TTT TGG CTG		2040
	CAA TTG AGG AAC TTG GGA TGC TAT TAA TTT TGT ATT TCA GCA ACT GCC CCT TCT CCT ATC		2100
	CCA AAG CAC CAA TTA CTG CCC TCT GCC TCA GCA GTA CCA GTA TAA GAT GAC ATT CCA AAG		2160
	ACT GGA GGC AAC TCA GCC TGA GTT AAT TCA CAA AAT TAT GCC ATG CTG GGG CTT GAG CTT		2220
	GAG CTT GGG CTT AGG CTT GGG CTC AGC TTT TGA CCC TCA GGC ATC TCC TTT TCC TTC CTG		2280
	TCT TCC TCT CCC TTC TCC TCT GCT GCA TGA TTT TCT TAA TCT TCA GAC ACT CAC TAT		2340
	TTT CAT GAA CAG TTA CCC TCT GTC CCC ACA ACC AAA GAC AAC TCA TGG CCT CCT TTG GCC		2400
	CTT GTG TAA CAT TGC AAA CCT GTG GCT TTG CAA AAT GTA CCC AGG TCA CAA GGG GAT TTT		2460
	TTT TTT TTT AGC AAT GAT ATC CCT GTC TGG GTC ACT TTT TAA GCT TGT AAC CGC CCC CCC		2520
	AGA CTT ATA ATC TTA AAT GTA TTT TCC TTT GTT TAA GCT GCT GCT TCC TCT GTT TCA TTG		2580
	GAT TGT GCC AGT TAT CAG TGG CTC TTG GGT TCA AAG TAA A		2623

Figure 1 Nucleotide and deduced amino acid sequences of H-IDH α cDNA insert

The entire nucleotide and deduced amino acid sequences of the H-IDH- α cDNA insert are shown. Nucleotide numbers are indicated on the right and amino acid numbers on the left. The amino acid residues corresponding to the two PCR primers are denoted by the horizontal arrows. The vertical arrow indicates the cleavage site of the mitochondrial signal peptide. The conserved motif for the binding site of [23] is indicated by bold type. Three polyadenylation signals, AATAAA, are underlined and adenosine-uridine-binding factor (AUBF) motifs [24] are indicated by italicized bold type.

yeast {mitochondrial IDH2 (Y-IDH2) [27] and IDH1 (Y-IDH1) [28]} and monkey (M-IDH γ) [29] respectively. However, it was less similar to the *E. coli* NADP⁺-specific IDH (30.6 %) (E-IDH) [30] and bovine IDPm (24.8 %) [18].

A comparison of the structures of the IDH isoenzymes from different species (Figure 2) revealed that five of eight amino acids (Ile³⁷, Gly³²¹, His³³⁹, Ala³⁴² and Asn³⁵²) responsible for the binding of the adenine moiety of NAD⁺ in E-IDH [11] are conserved in H-IDH α (Ile¹⁵, Gly²⁴⁴, His²⁶³, Ala²⁶⁶ and Asn²⁷⁶) and the yeast

catalytic subunit, Y-IDH2. In contrast, only three and two amino acids were conserved in M-IDH γ and the yeast regulatory subunit, Y-IDH1, respectively, indicating that the latter two enzymes may not be involved in the interaction with NAD⁺. All ten amino acids (Ser¹¹³, Asn¹¹⁵, Arg¹¹⁹, Arg¹²⁹, Arg¹⁵³, Tyr¹⁶⁰, Lys²³⁰, Asp²⁸³, Asp³⁰⁷ and Asp³¹¹) responsible for binding to either isocitrate or Mg²⁺-isocitrate in E-IDH [11–15] were also conserved in the H-IDH α protein. These results strongly suggest that H-IDH α possesses two separate binding sites for NAD⁺

Figure 2 Structural comparison of IDH isoenzymes from different species

The deduced amino acid sequence of IDH isoenzymes from H-IDH α , Y-IDH2 [27], Y-IDH1 [28], M-IDH γ [29] and E-IDH [30] and partial tryptic peptide sequences from P-IDH α [21] were aligned using PC/GENE software. Spaces inserted for alignment purposes are indicated by hyphens. Invariant amino acids in all six IDH isoenzymes are marked by darker shading. The amino acids conserved (> 80%) in IDH isoenzymes are represented in stippled boxes. The asterisks and closed circles indicate the amino acids involved in the binding of isocitrate and Mg^{2+} in E-IDH [11–15] respectively. The conserved amino acids responsible for binding of NAD^+ in E-IDH [11] are indicated by open circles.

and Mg^{2+} –isocitrate, and that it is assembled in a similar conformation to that of Y-IDH2, the catalytic subunit of the yeast enzyme [19].

A consensus sequence, Ala³⁰⁶–Xaa–Xaa–Xaa–Xaa–Gly³¹¹–Lys³¹²–Ser³¹³, found in a typical ATP/GTP-binding motif (Ala/Gly–Xaa–Xaa–Xaa–Xaa–Gly–Lys–Ser/Thr) [23], was observed near the C-terminus of the IDH α protein (Figure 2). In addition, the presence of an ATP-binding motif in M-IDH γ [29] suggests that the negative regulator ATP can bind to both subunits. Furthermore, a putative ADP-binding site (DGLFLQKCR) was recognized in the H-IDH α protein. Aspartate-181 of the above

sequence is highly conserved in other eukaryotic enzymes including P-IDH β [21] and P-IDH γ [21,29,31]. Our structural analysis of H-IDH α provides evidence that it may contain the binding site for the allosteric inactivator (ATP) as well as the activator (ADP).

Previous data [17,21,31] suggested that two ADP molecules bind to the tetrameric IDH enzyme *in vivo* and that both IDH β and IDH γ contain ADP-binding sites. Our data (Figure 2) suggest that IDH α also has an ADP-binding motif, indicating that the tetrameric IDH enzyme composed of $2\alpha 1\beta 1\gamma$ subunits contains at least four ADP-binding sites. It is unclear which IDH

Figure 3 Northern-blot analysis of human IDH α mRNA

Polyadenylated RNA samples (2 μ g/well) from cultured HeLa cells were subjected to denaturing agarose-gel electrophoresis, transferred to GeneScreen membranes and then hybridized with the ³²P-labelled H-IDH α cDNA probe as described in the Materials and methods section.

Figure 4 Southern-blot analysis of H-IDH α gene

Human genomic DNA (10 μ g/well) isolated from placenta was digested with six restriction endonucleases as indicated. The DNA fragments were separated on a 0.8% agarose gel, transferred to GeneScreen membrane and then hybridized with ³²P-labelled H-IDH α cDNA probe. Size markers are indicated on the left.

subunit is responsible for the two ADP-binding sites *in vivo* [17,32]. Further experimental evidence is needed to clarify this point.

Expression of IDH α mRNA

One species of H-IDH α mRNA (2.8 kb) was detected in HeLa cells (Figure 3), but no bovine IDPm [18] mRNA was observed (results not shown). One IDH α mRNA transcript (2.8 kb) was also observed in rat heart and brain, but it was expressed 10-fold less abundantly than the IDPm mRNA (results not shown). These results further suggest that the expression of the IDH α and IDPm genes is differentially regulated in a cell- or tissue-specific manner similar to those of IDH isoenzymes [1,33].

Genomic Southern-blot analysis of H-IDH α gene

Southern-blot analysis was performed to determine whether the IDH α gene has a closely related gene family member (Figure 4). The simple patterns and the relatively small size of the DNA fragments that hybridized with the H-IDH α cDNA suggest that the H-IDH α gene is derived from a single gene family.

DISCUSSION

Three distinct isoenzymes of IDH exist in the mammalian system. Their subcellular localization, cofactor specificity and tissue-specific expression are well characterized [1,2,10,33]. Despite the extensive biochemical studies that have been carried out on these isoenzymes, the current information on their structure is not complete. The entire nucleotide and deduced amino acid sequences of the cDNA clones have been reported for mitochondrial Y-IDH [27,28], M-IDH γ [29] and NADP⁺-specific IDHs of mammalian origin [18,34] (Y.-O. Kim, K.-L. Oh, I.-L. Huh and B. J. Song, unpublished work). In the present paper, we report the isolation of a full-length cDNA clone for heart H-IDH α . This was achieved by PCR amplification of DNA using two oligonucleotide primers resulting in a specific 0.6 kb fragment. Several attempts using different sets of primers corresponding to other peptide sequences of P-IDH α [21] were unsuccessful. Using the sequence-verified PCR fragment, we were able to isolate cDNA clones for heart H-IDH α .

The mature H-IDH α protein sequence was highly similar to those of tryptic peptides of the pig enzyme [21]. Although we do not know the actual N-terminus of the H-IDH α protein, it is reasonable to assume that threonine-1 (Figures 1 and 2) is the first amino acid of the mature protein for the following reasons. First, alignment of the deduced amino acid sequence of H-IDH α with the N-terminal peptide sequence of the pig enzyme [21] indicates that this residue corresponds to the first amino acid, alanine-1 of P-IDH α , as shown in Figure 2. Secondly, extremely high structural similarity (about 98%) was observed between the IDPms of cow, pig and human origin [18,34] (Y.-O. Kim, I.-L. Oh, T.-L. Huh and B. J. Song, unpublished work). Furthermore, the sizes of their mitochondrial signal peptides and mature proteins were identical. Thirdly, threonine-1 is the third amino acid residue from the basic amino acid (in this case arginine) which could represent the cleavable presequence, as proposed by Hartl and Neupert [35].

AUBF binding to the AUUUA motif in the 3'UTR of the labile mRNA transcript [36] leads to an increased half-life of certain mRNAs, thus increasing the level of translated protein. Such a mechanism has also been reported for interleukin 3, c-Fos, v-Myc, interferon γ and granulocyte-macrophage colony-stimulating factor [37,38]. In addition, the numerous trinucleotides in the 3'-UTR may form secondary-loop structures which may be cleavage sites for endonucleases; this may result in changes in the stability of the mRNA as demonstrated for cholesterol 7 α -hydroxylase [26]. The presence of these AUBF motifs and trinucleotides in IDH α mRNA but not in that of IDH γ [29] or IDPm [18,34] may indicate regulation of IDH α mRNA stability. This method of regulation would be in addition to the already known allosteric control of IDH activity [1,2,10]. Further experiments need to be performed to test this hypothesis.

In mammals, it is not known which IDH subunit contains the catalytic activity, but it is thought to be the α -subunit for the following reasons. First, Ramachandran and Colman and Huang and Colman [20,21] originally proposed that catalytically active tetrameric enzyme can be divided into two catalytically active heterodimers of $\alpha\beta$ - and $\alpha\gamma$ -subunits, suggesting that the catalytic activity resides in the common subunit, IDH α in this case. Secondly, the structural comparison (Figure 2) demonstrated that H-IDH α is more similar to Y-IDH2, the catalytic subunit [19,27], than Y-IDH1, the regulatory subunit [28]. Thirdly, catalytic activity requires binding of NAD⁺ and Mg²⁺-isocitrate [19–21], both of which are found in H-IDH α but not in M-IDH γ [29].

To try to demonstrate the catalytic role of the α -subunit, we

overproduced the H-IDH α protein in *E. coli* and purified it by affinity column chromatography (results not shown). However, the purified H-IDH α protein was catalytically inactive in the presence or absence of the allosteric activator ADP (results not shown). This negative result could be due to unnatural protein folding of the overproduced H-IDH α protein. Alternatively, participation of another subunit, either β or γ , may be needed for IDH catalytic activity, as originally proposed [21]. Similar results have recently been reported for the yeast IDH system [19], other mitochondrial enzymes such as branched-chain α -keto acid dehydrogenase [41] and pyruvate dehydrogenase [40]. In the latter two cases, the respective regulatory E1 β subunits are needed for the catalytic E1 α subunits to function properly. Co-expression of the α - and γ -subunits of IDH in *E. coli* is currently being performed and the resulting proteins will be tested for IDH catalytic activity.

Denton and his associates [39] reported that three mitochondrial enzymes (pyruvate dehydrogenase, 2-oxoglutarate dehydrogenase and isocitrate dehydrogenase) are activated by physiological concentrations of Ca²⁺. The stoichiometry of Ca²⁺ binding has been determined to be 0.9 mol of Ca²⁺/mol per IDH tetramer. Binding of the Mg²⁺-isocitrate complex to the enzyme is a prerequisite for Ca²⁺ binding [9]. The consensus Ca²⁺-binding motif, D-X-D(or N)-X-X-G-X-I-D(or E)-X-X-E, proposed by Marsden et al. [42] is not found in either H-IDH α or M-IDH γ protein. Therefore IDH β may contain the putative Ca²⁺-binding site. Alternatively, IDH may have another type of Ca²⁺-binding site created by the presence of an additional IDH subunit. This type of Ca²⁺ binding has been proposed for 2-oxoglutarate dehydrogenase [9]. This hypothesis needs to be confirmed by isolation and sequence determination of the IDH β cDNA clone.

However, a consensus ATP/GTP-binding motif [23], Ala³⁰⁶-Xaa-Xaa-Xaa-Xaa-Gly³¹¹-Lys³¹²-Ser³¹³, was identified in the H-IDH α protein by the PROSITE search. The predicted secondary structure of this motif in H-IDH α is different from the typical ATP/GTP-binding motif referred to as the P-loop [43] which forms a flexible loop between a β -strand and an α -helix. The secondary structure for H-IDH α protein predicted by the Gascuel and Golmard basic statistical method [44] revealed that alanine-306 was located near the end of a long α -helix and the following Gly³¹¹-Lys³¹²-Ser³¹³ sequence was located on a long coil structure before another α -helix. Thus the absence of a β -strand domain before the putative ATP/GTP-binding sequence and the location of alanine-306 on an α -helix domain are different from the typical P-loop type [43]. On the other hand, the predicted secondary structure for the postulated ATP-binding site in M-IDH γ protein [29] was also similar to that of H-IDH α protein (results not shown). Therefore the actual ATP-binding site in H-IDH α still remains to be elucidated.

We are grateful to Mr Seok-Hyung Kim and Ms. Mi-Soon Park for help with DNA sequence analysis. We thank Dr. James Lee for critical reading of and helpful comments on the manuscript. Thanks also go to David Landsman (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, U.S.A.) for the protein databank search and sequence analysis. This work was partially supported by a Genetic Engineering Program Grant (1993) from the Ministry of Education, Republic of Korea.

REFERENCES

- 1 Plaut, G. W. E. (1970) *Curr. Top. Cell. Regul.* **2**, 1–27
- 2 Koshland, D. E., Jr., Walsh, K. and LaPorte, D. C. (1985) *Curr. Top. Cell. Regul.* **27**, 13–22
- 3 Garnak, M. and Reeves, H. C. (1979) *Science* **203**, 1111–1112
- 4 Nimmo, G. A., Borthwick, A. C., Holms, W. H. and Nimmo, H. G. (1984) *Eur. J. Biochem.* **141**, 401–408
- 5 LaPorte, D. C., Thorsness, P. E. and Koshland, D. E., Jr. (1985) *J. Biol. Chem.* **260**, 10563–10568
- 6 Keys, D. A. and McAlister-Henn, L. (1990) *J. Bacteriol.* **172**, 4280–4287
- 7 Hathaway, J. A. and Atkinson, D. E. (1963) *J. Biol. Chem.* **238**, 2875–2881
- 8 Cohen, P. F. and Colman, R. F. (1972) *Biochemistry* **11**, 1501–1508
- 9 Rutter, G. A. and Denton, R. M. (1989) *Biochem. J.* **263**, 445–452
- 10 Plaut, G. W. E. and Gabriel, J. L. (1983) in *Biochemistry of Metabolic Process* (Lennon, D. L. F., Stratman, F. W. and Zahiten, R. N., eds.), pp. 285–301, Elsevier Science Publishing Co., New York
- 11 Hurley, J. H., Dean, A. M., Koshland, D. E., Jr. and Stroud, R. M. (1991) *Biochemistry* **30**, 8671–8678
- 12 Hurley, J. H., Thorsness, P. E., Ramalingam, V., Helmers, N. H., Koshland, D. E., Jr. and Stroud, R. M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8635–8639
- 13 Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E., Jr. and Stroud, R. M. (1990) *Science* **249**, 1012–1016
- 14 Stoddard, B. L. and Koshland, D. E., Jr. (1993) *Biochemistry* **32**, 9317–9322
- 15 Stoddard, B. L. and Koshland, D. E., Jr. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1146–1153
- 16 Barnes, L. D., Kuehn, G. D. and Atkinson, D. E. (1971) *Biochemistry* **10**, 3939–3944
- 17 Ehrlich, R. S. and Colman, R. F. (1981) *J. Biol. Chem.* **256**, 1276–1282
- 18 Huh, T.-L., Ryu, J. H., Huh, J.-W., Sung, H.-C., Oh, I.-L., Song, B. J. and Veech, R. L. (1993) *Biochem. J.* **292**, 705–710
- 19 Cupp, J. R. and McAlister-Henn, L. (1993) *Biochemistry* **32**, 9323–9328
- 20 Ramachandran, N. and Colman, R. F. (1980) *J. Biol. Chem.* **255**, 8859–8864
- 21 Huang, Y.-C. and Colman, R. F. (1990) *Biochemistry* **29**, 8266–8273
- 22 Davis, L. G., Dibner, M. D. and Battey, J. F. (1986) *Basic Methods in Molecular Biology*, Elsevier Science Publishing Co., New York
- 23 Walker, J. E., Saraste, M., Runswick, M. J. and Gay, M. J. (1982) *EMBO J.* **1**, 945–951
- 24 Gillis, P. and Malter, J. S. (1991) *J. Biol. Chem.* **266**, 3172–3177
- 25 Binder, R., Hwang, S.-P. L., Ratnasabapathy, R. and Williams, D. L. (1989) *J. Biol. Chem.* **264**, 16910–16918
- 26 Noshiro, M., Nishimoto, M. and Okuda, K. (1990) *J. Biol. Chem.* **265**, 10036–10041
- 27 Cupp, J. R. and McAlister-Henn, L. (1991) *J. Biol. Chem.* **266**, 22199–22205
- 28 Cupp, J. R. and McAlister-Henn, L. (1992) *J. Biol. Chem.* **267**, 16417–16423
- 29 Nichols, B. J., Hall, L., Perry, A. C. F. and Denton, R. M. (1993) *Biochem. J.* **295**, 347–350
- 30 Thorsness, P. E. and Koshland, D. E., Jr. (1987) *J. Biol. Chem.* **262**, 10422–10425
- 31 Huang, Y.-C. and Colman, R. F. (1989) *J. Biol. Chem.* **264**, 12208–12214
- 32 Ehrlich, R. S. and Colman, R. F. (1982) *J. Biol. Chem.* **257**, 4769–4774
- 33 Stein, A. M., Stein, J. H. and Kirkman, S. K. (1967) *Biochemistry* **6**, 1370–1379
- 34 Haselbeck, R. J., Colman, R. F. and McAlister-Henn, L. (1992) *Biochemistry* **31**, 6219–6223
- 35 Hartl, F. U. and Neupert, W. (1990) *Science* **247**, 930–938
- 36 Malter, J. S. (1989) *Science* **246**, 664–666
- 37 Wodnar-Filipowicz, A. and Moroni, C. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 777–781
- 38 Shaw, G. and Kamen, R. (1986) *Cell* **46**, 659–667
- 39 Davie, J. R., Wynn, R. M., Cox, R. P. and Chuang, D. T. (1992) *J. Biol. Chem.* **267**, 16601–16606
- 40 Jeng, J., Huh, T.-L. and Song, B. J. (1994) *Biochem. Biophys. Res. Commun.* **203**, 225–230
- 41 Denton, R. M., McCormack, J. G. and Edgell, N. E. (1980) *Biochem. J.* **190**, 107–117
- 42 Marsden, B. J., Shaw, G. S. and Sykes, B. D. (1990) *Biochem. Cell Biol.* **68**, 587–601
- 43 Saraste, M., Sibbald, P. R. and Wittinghofer, A. (1990) *Trends Biochem. Sci.* **15**, 430–434
- 44 Gascuel, O. and Golmard, J. L. (1988) *Comput. Appl. Biosci.* **4**, 357–365