# Bovine $NAD^+$-Dependent Isocitrate Dehydrogenase: Alternative Splicing and Tissue-Dependent Expression of Subunit 1†

Cipora Weiss, Yan Zeng,‡ Jianmin Huang,§ Malgorzata B. Sobocka, and Julie Ivory Rushbrook*

*Department of Biochemistry, State University of New York Downstate Medical Center, 450 Clarkson Avenue, Brooklyn, New York 11203*

*Received July 22, 1999; Revised Manuscript Received October 4, 1999*

ABSTRACT: $NAD^+$-dependent isocitrate dehydrogenase (IDH), a key regulatory enzyme in the Krebs cycle, is a multi-tetrameric enzyme. At least three of the subunits in the core tetramer of mammals are unique gene products. Subunits 1/$\beta$ and 2/$\gamma$ are considered to be regulatory, while subunits 3,4/$\alpha$, comprising half the tetramer, are catalytic. The full sequence was obtained for the major subunit 1 cDNA in bovine heart, IDH 1-A. A second cDNA, rare in heart, was also identified (IDH 1-B). Differences in the two were confined to the 3′-region, suggesting alternative splicing. Screening of brain, kidney, and liver RNA showed the presence of IDH 1-A and 1-B and a third major species, IDH 1-C. Amplification of bovine genomic DNA by PCR across the regions of difference produced a single product. Comparison of the genomic and mRNA sequences showed that IDH 1-A resulted from splicing of exon W to exon Y, eliminating intron w, exon X, and intron x. IDH 1-B was formed by splice junctions between exon W, exon X, and exon Y. IDH 1-C resulted from splicing of exon W to exon X and subsequent retention of intron x. The 2 proteins predicted from these 3 mRNAs are identical over their first 357 residues. Protein IDH 1-A, resulting from a termination codon within exon Y, contains an additional 26 residues. Proteins IDH 1-B and 1-C derive from a common termination codon within exon X and contain an additional 28 residues. The two C-terminal regions differ notably in the number and nature of charged residues, resulting in proteins with a charge difference of 3.2 at pH 7.0. Subunit 1 sequences previously reported from other species grouped with one or the other of the bovine proteins. No evidence was found for alternative splicing in subunit 3,4/$\alpha$. The results of the present study, together with recent work on the 2/$\gamma$ subunit [Brenner,V., Nyakatura, G., Rosenthal, A., and Platzer, M. (1998) *Genomics 44,* 8], indicate that the regulatory subunits of the enzyme, but not the catalytic, possess alternatively spliced forms varying in C-terminal properties with tissue-specific expression. The finding is suggestive of a mechanism for modulation of allosteric regulation tailored to the needs of different tissues.

$NAD^+$-dependent isocitrate dehydrogenase ($NAD^+$-IDH),[1] located in the mitochondrial matrix, catalyzes the conversion of isocitrate to $\alpha$-oxoglutarate and carbon dioxide, providing reducing equivalents for the electron transport chain. Additional enzymes for which biological roles are not well-defined catalyze the same reaction in both mitochondria and cytosol but utilize $NADP^+$. Prokaryotes such as *E. coli* have a single IDH with $NADP^+$ as a cofactor.

The $NAD^+$-dependent enzymes from bovine and pig hearts and from yeast have been studied extensively. Based on allosteric effects which modulate activity according to the energy requirements of the cell, the enzyme is believed to be a key regulatory point in the tricarboxylic acid cycle. The enzymes from bovine heart and yeast exhibit positive cooperativity for isocitrate (*1*−*3*). Negative cooperativity has been reported for $NAD^+$ for the enzyme from beef heart (*3*). In the mammalian enzymes, ADP acts as a positive heterotropic effector, increasing the affinity of the enzyme for isocitrate (*2*−*4*). NADH, a product of the reaction, and NADPH and ATP have negative heterotropic effects (*5*−*7*). Regulation of the enzyme from yeast is simpler, with AMP acting as a positive heterotropic effector for the binding of isocitrate (*1*). Both mammalian and yeast enzymes are activated by $Ca^{2+}$ in the presence of adenine nucleotides (*8*, *9*).

The $NAD^+$-dependent enzymes are multi-tetramers with subunits of about 40 kDa (*1*, *5*, *10*−*13*). At least three of the four subunits of the core mammalian tetramer are unique gene products. Beef heart subunits 1, 2, and 3, named in order of their increasing mobility on sodium dodecyl sulfate (SDS)−polyacrylamide gels, correlate by amino acid sequence homology with pig heart subunits $\beta$, $\gamma$, and $\alpha$, respectively (*14*). Bovine subunit 4, which migrates slightly faster than subunit 3 by SDS−gel electrophoresis, is identical

† This work was supported by a Grant-In-Aid from the American Heart Association, New York Affiliate, to J.I.R.

* Correspondence should be addressed to this author. Telephone: (718) 270-2647; fax: (718) 270-3316; e-mail: jrushbrook@netmail.hscbklyn.edu.

‡ Current address: Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO 63110.

§ Current address: Department of Microbiology, Weill Medical College of Cornell University, 1300 York Ave., New York, NY 10021.

[1] Abbreviations: IDH, isocitrate dehydrogenase; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; RT, reverse transcription; RACE, rapid amplification of cDNA ends; AP, adapter primer; UAP, universal amplification primer; AUAP, abridged universal amplification primer; UTR, untranslated region; bp, base pair(s).

10.1021/bi991691i CCC: $19.00 
Published on Web 01/25/2000

in sequence with subunit 3 through the 30 N-terminal residues characterized. In support of structural differences in subunits 3 and 4, the corresponding porcine subunit, α, also appears as two distinct bands on SDS−gel electrophoresis (*14*). The nature of the structural differences in subunits 3/4/α must yet be determined but appears to originate at the protein level (see Discussion). In the tetramer, the subunits are believed to be present in the ratio 1:1:1:1 (*β*:*γ*:2α) (*15, 16*). In the simpler enzyme from yeast, the core tetramer contains two unique gene products, IDH 1 and 2, present in equimolar amounts (*13*).

Three-dimensional studies of the distantly related $NADP^+$-dependent enzyme from *E. coli* have provided basic information relevant to the functions of the subunits of the complex eukaryotic enzymes (*17*−*21*). The *E. coli* enzyme is a homodimer, active only in the absence of phosphorylation (*22*−*24*). It lacks homology with eukaryotic $NADP^+$-dependent enzymes, but has low but significant identity (approximately 30%) with subunits of the $NAD^+$-dependent enzymes (*14*, *25*−*29*). The two *E. coli* subunits interact in such a way that each contributes to the two catalytic sites. Residues important for isocitrate, $Mg^{2+}$, and $NADP^+$ binding have been identified.

Sequence comparison of the *E. coli* enzyme with the subunits of the eukaryotic $NAD^+$-dependent enzymes indicates that yeast IDH-2 and mammalian subunits 3/4/α retain residues required for isocitrate, $Mg^{2+}$, and $NAD^+$ binding (*14*, *26*, *27*), suggesting that they are involved in catalysis. Yeast subunit IDH 1 and mammalian *β* and *γ* subunits lack residues for binding $NAD^+$ (*26*−*28*) and are considered to be involved in allosteric regulation rather than catalysis. These conclusions are supported by studies of interactions among the pig heart subunits together with their chemical modification (*30*, *31*), and the results of mutagenesis of the yeast subunits (*32*, *33*).

Whether unique subunit gene products exist which are differentially expressed in mammalian tissues has received little attention. While the major protein studies have been on the pig and beef heart enzymes, some work has been carried out on enzymes from beef brain and pig liver (*34*−*37*). The enzymes from the latter sources have many of the properties described above for the heart enzymes, although positive cooperativity for isocitrate was not demonstrated in the very early report on the liver enzyme. For a truly informative functional comparison, enzymes would have to be purified from tissues of the same species and compared directly.

Sequences from cDNAs of a particular subunit have been reported from different tissues, but these have been from different species, e.g., monkey testis and beef heart 3/4/α subunits (*14*, *27*), and rat epididymis and monkey testis 2/*γ* subunits (*28*). Supporting the idea that each subunit is encoded by a single gene, Southern blot analysis of human genomic DNA indicated that the 3/4/α subunit is encoded by a single gene (*29*).

Recently the gene for the human 2/*γ* subunit was identified and characterized (*38*). A search of Expressed Sequence Tag databases using 2/*γ* intronic sequences revealed several human sequences involving alternative splice sites at the 3′-end. Reverse transcription polymerase chain reaction (RT-PCR) examination of mRNAs from human tissues indicated differential expression of alternatively spliced species.

In the present work, alternative splicing is reported for subunit 1/*β* from bovine tissues. Three mRNA species resulting in two predicted protein products are expressed variously in heart, brain, kidney, and liver, with one mRNA being characteristic of heart muscle.

## MATERIALS AND METHODS

*Materials.* Beef heart ventricular tissue, brain (segments from the cortical lobes of both hemispheres), kidney (segments including cortical and pyrimidal layers), and liver were frozen and stored as previously described (*14*). Reagents, enzymes, and cells were as previously described (*14*).

*3′-Rapid Amplification of cDNA Ends (RACE) for Full-Length Sequencing of Heart Subunit 1.* mRNA was obtained from adult beef heart muscle by LiCl extraction followed by oligo(dT) chromatography as previously described (*14*). All PCRs utilized the GeneAmp PCR System 9600 (Perkin-Elmer, Norwalk, CT). First-strand cDNA was synthesized from poly(A)$^+$ RNA using the adapter primer AP (see *Primers*) and the 3'-RACE System from GIBCO BRL (Grand Island, NY) as previously described (*14*). A small amount of the cDNA was amplified by PCR using the 3′-universal amplification primer, UAP, and the 5′-gene-specific primer 605S (see *Primers*) as previously described (*14*) with the following modifications: after the addition of Taq polymerase and cDNA, the tubes were heated at 95 °C for 1.5 min. Five cycles of 95 °C for 15 s, 37 °C for 30 s, and 72 °C for 3.5 min were followed by 30 cycles of 95 °C for 15 s, 55 °C for 30 s, and 72 °C for 3.5 min and a final extension of 14 min at 72 °C. Products were evaluated by agarose gel electrophoresis. The PCR products was extracted using QIAEX gel (Qiagen, Inc., Chatsworth, CA), precipitated, and cut with *Mlu*I (UAP has an *Mlu*I site) and *Eco*RI (605S has an *Eco*RI site) in *Mlu*I buffer overnight at 37 °C. DNA between 1100 and 1350 base pairs (bp) in length was isolated after agarose gel electrophoresis using QIAEX, ligated to pGEM7Zf(+), and introduced into JM83 *E. coli* cells. Colonies were tested for the presence of inserts by miniprep plasmid DNA preparations and acrylamide gel analysis of restriction digests. Two full-length clones, IDH 1-A and IDH 1-B, were sequenced using the strategies outlined in Figure 1.

*5′-RACE for Full-Length Sequencing of Heart Subunit 1.* mRNA was obtained using the PolyATract System 1000 (Promega, Madison, WI) as previously described (*14*). cDNA for PCR was prepared using the 5′-AmpliFINDER RACE kit (Clontech, Palo Alto, CA) as previously described (*14*). Ligation to the AmpliFINDER Anchor employed T4 RNA ligase in a 20 h incubation at room temperature. A portion of the cDNA was amplified by the hot-start PCR procedure (Perkin Elmer) using anchor primer and SUB1P2, an IDH subunit 1 specific primer (see *Primers*). The conditions were 95 °C for 20 s, 35 cycles of 94 °C for 15 s, 60 °C for 15 s, and 72 °C for 75 s, and finally 7 min at 72 °C. Agarose gel electrophoresis using ethidium bromide staining revealed a major band of about 650 bp in the PCR mixture. Following QIAEX gel extraction, the product was introduced into the pDIRECT vector and cloned using the nonligation PCR-DIRECT cloning kit (Clontech). Three clones were analyzed: P1-7, P1-9, and P1-24. Each was sequenced with M13 forward and M13 reverse primers. P1-7 and P1-9 were

additionally sequenced in the reverse direction with primers C4 and C5 (Figure 1) (see *Primers*).

Sequencing of the 3′- and 5′-RACE products described above was carried out as previously reported using the silver sequencing system from Promega (*14*). Sequencing of internal regions of both 3′- and 5′-RACE inserts utilized synthetic primers based on the newly obtained sequence.

*3′-RACE To Investigate Differences in the 3′-Ends of IDH 1 in Various Tissues.* Reverse transcription was carried out with total RNA using AP. PCR utilized the Advantage-HF PCR kit (Clontech), the sequence-specific primer C2 located at bases 824−840 (Figure 1C), and the abridged universal amplification primer AUAP. Following incubation at 94 °C for 15 s, PCR was carried out as follows: 35 cycles comprising 94 °C for 5 s, 58 °C for 15 s, and 68 °C for 3 min. Extension took place at 68 °C for 3 min. PCR products were vizualized by 1.5% agarose gel electrophoresis and immediately inserted into pT-Adv vector (Clontech) and cloned in TOP10F′ *E. coli* competent cells using the AdvanTAge PCR Cloning Kit (Clontech).

Plasmid DNA was prepared from positive clones using either Wizard Minipreps (Promega) or QUIAprep Miniprep (Qiagen) kits. Inserts released by *Eco*RI digestion were analyzed by agarose gel electrophoresis. Selected clones were sequenced either by the silver sequencing method or by MacConnell Research Corp., San Diego, CA.

*PCR To Investigate Potential Differences in the 5′-Ends of IDH 1 in Various Tissues.* Reverse transcription was carried out with total RNA from heart, brain, kidney, and liver using primer AP. PCR utilized the Advantage-HF PCR kit (Clontech) and sequence-specific primers C11 and C14, located at bases 58−74 and 856−835, respectively (Figure 1C). PCR was carried out as follows: 5 cycles comprising 94 °C for 5 s, 68 °C for 2 min, followed by 30 cycles comprising 94 °C for 5 s, 55 °C for 15 s, and 68 °C for 2 min. Extension took place at 68 °C for 4 min. PCR products were vizualized by agarose gel electrophoresis and immediately inserted into pT-Adv vector (Clontech) and cloned in TOP10F′ *E. coli* competent cells using the AdvanTAge PCR Cloning Kit (Clontech).

Plasmid DNA was prepared from positive clones using either Wizard Minipreps (Promega) or QUIAprep Miniprep (Qiagen) kits. Inserts released by *Eco*RI digestion were analyzed by agarose gel electrophoresis. Sequencing was carried out by MacConnell Research Corp.

*Purification of Genomic DNA and Preparation of Sequence Containing the Requisite 3′-End of Subunit 1.* Genomic DNA was prepared from beef liver using the Promega DNA Purification Kit (Promega). PCR was carried out using the sequence-specific primers K1 and K2 (see *Primers*) and the Advantage-HF PCR Kit (Clonetech) as follows: 94 °C for 15 s, 3 cycles comprising 97 °C for 10 s, 68 °C for 6 min, followed by 27 cycles comprising 94 °C for 10 s, 68 °C for 6 min. Extension took place at 68 °C for 7 min. Analysis of the products by 1% agarose gel electrophoresis revealed a single band of 1.1 kb. Following insertion into pT-Adv vector and cloning, the insert was sequenced in both directions by MacConnell Research Corp.

*Sequence Analysis*. Sequences were analyzed using programs in PC/Gene (Oxford Molecular Group, Inc., Campbell, CA).

*Primers.* 605S (5′-TCGAATTC ACC CAA/G GGN GAA/G GAT/C GT-3′) codes for residues 5−10 in the first 22 residues of the N-terminus of subunit 1, TQGEDV, determined by protein sequencing (*14*), with the most commonly used codon for T (*39*), and all possible codons for Q, G, E, and D. Underlined is an added *Eco*RI site. 605S, which was used as the sequence-specific 5′-primer in 3′-RACE for full-length sequencing, was synthesized by the DNA Synthesis and Purification Laboratory of the Department of Microbiology and Immunology of SUNY Downstate Medical Center.

Adapter primer [AP, 5′-GGCCACGCGTCGACTAGTAC-$(T)_{17}$-3′] and universal amplification primer (UAP, 5′-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3′) are included in the GIBCO 3′-RACE kit. AP was used as the reverse transcription primer for full-length 3′-RACE and UAP as the 3′-primer in the PCR step of 3′-RACE for directional cloning. Z10 (5′-AGTGGCCATCATTGGAA-3′, sense, bases 378−394; see Figure 1C), Z11 (5′-TCG AGT CCG TAC CTT GC-3′, antisense, bases 1121−1137; Figure 1C, IDH 1-B), C1 (5′-TCTTAGTGGCATAGTCA-3′, antisense, bases 660−676; Figure 1C), C2 (5′-TGGTGCAGAATCCTTAC-3′, sense, bases 824−840; Figure 1C), and C3 (5′-AGCTTCCAACATGCTGC-3′, sense, bases 1038−1054; Figure 1C) were based on newly obtained sequence. These oligonucleotides were trityl-specific from Ransom Hill Bioscience, Ramona, CA.

For 5′-RACE, Clontech provided the AmpliFINDER anchor 5′-CACGAATTCACTATCGATTCTGGAACCTTCAGAGG-$NH_3$ (the sequence complementary to the anchor primer is underlined) and the anchor primer 5′-**CTGGTTCGGCCCA**CCTCTGAAGGTTCCAGAATCGATAG-3′ (the sequence complementary to the anchor is underlined; the unique sticky end for pDIRECT cloning is in boldface type).

SUB1P2 (5′-**CTCGCTCGCCCA**CTCAATCACACCCCTCGCACT-3′, antisense, bases 586−606, Figure 1C; the unique sticky end added for pDIRECT cloning is in boldface type) was used as the sequence-specific primer in 5′-RACE.

C4 (5′-TGACTTCACATGGACCA-3′, antisense, bases 479−495; Figure 1C) and C5 (5′-GCCTTGAACACCTCCTT-3′, antisense, bases 256−272; Figure 1C) were used in sequencing of 5′-RACE inserts. SUB1P2, C4, and C5 were trityl-specific sequences from Ransom Hill Bioscience.

For analysis of sequence differences at the 3′-end of IDH 1 messages by 3′-RACE, AP was used as the reverse transcription primer. PCR was carried out with the abridged universal amplification primer (AUAP, 5′-GGCCACGCGTCGACTAGTAC-3′, antisense, for nondirectional cloning) and C2 (see above).

For analysis of possible sequence differences at the 5′-end of IDH 1, PCR was carried out using C11 (5′-GCTGCCCTCAGCCGTGT-3′, sense, bases 58−74; Figure 1C) and C14 (5′-CCAGCACGTCAAACTGGT-3′, antisense, bases 856−839; Figure 1C).

Genomic DNA was amplified by PCR using primers K1 (5′-GCGGCATCTCAATCTCGAGCATCACT-3′, sense, bases 1053−1078) and K2 [5′-ACCTATGGGTTGGACAGGGCACCAAT-3′, antisense, bases 1412−1387; numbered according to IDH 1-B (Figure 1C)].

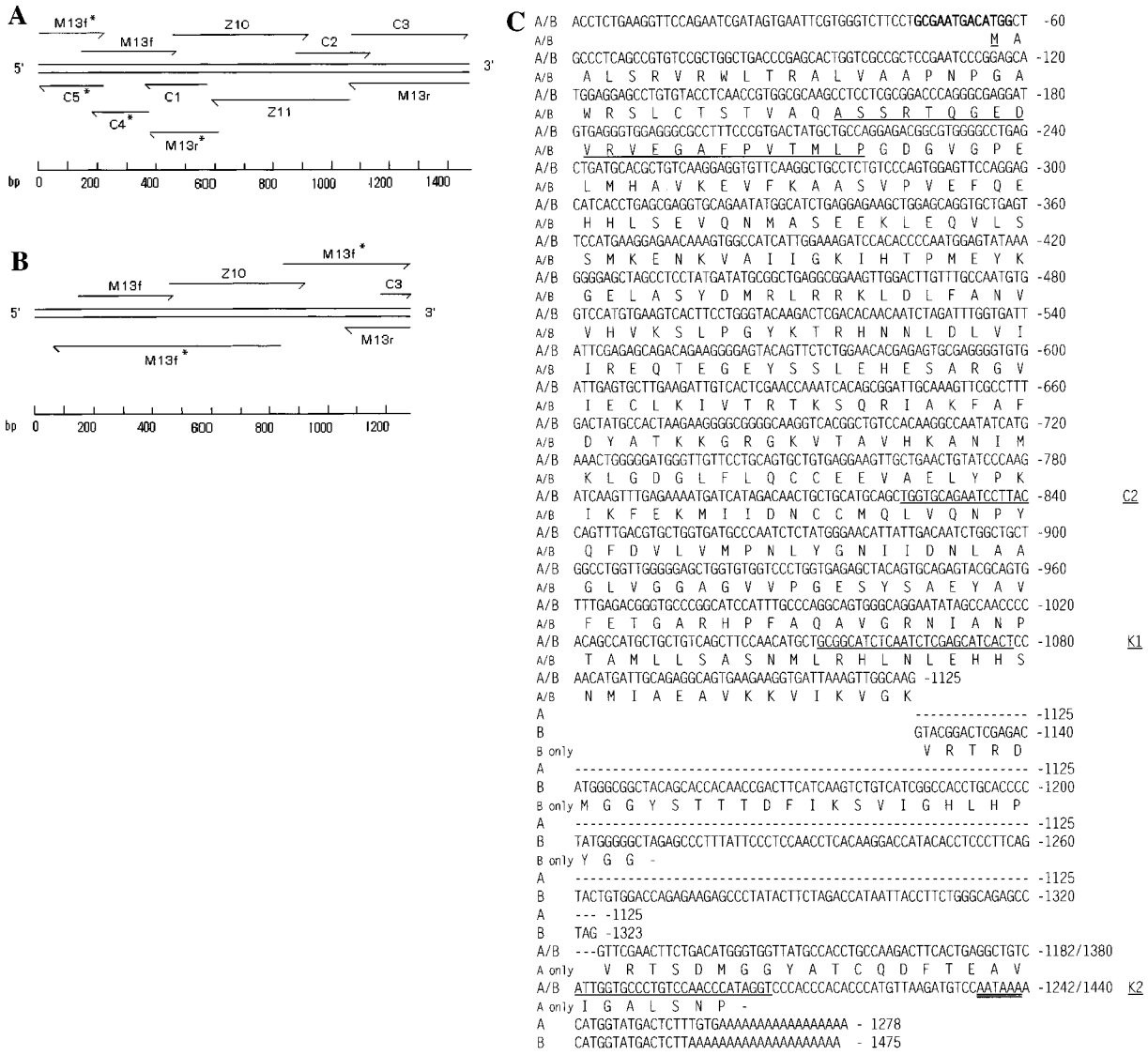


FIGURE 1: Sequences of bovine subunit 1 species corresponding to clones 1-A and 1-B. (A) Strategy for sequencing of clone 1-B and the common 5′-RACE product. Sequences marked with an asterisk were obtained from 5′-RACE products of heart mRNA after initial analysis of clone 1-B. (B) Sequence strategy for clone 1-A. Asterisked sequences were derived from heart mRNA after initial analysis of clone 1-A. The asterisked 5′-sequence was derived from cloned RT-PCR products obtained using primers C11 and C14 (encompassing bases 75−838). The asterisked 3′-sequence was derived from 3′-RACE products obtained using primer C2. (C) cDNA sequences for subunit 1 corresponding to 1-A and 1-B clones (5′-untranslated region, regions coding for mitochondrial presequence and mature protein, and 3′-untranslated region) and deduced protein sequences. A/B in large and small type denote cDNA and protein sequences, respectively. The N-terminal amino acid sequence of the mature protein, determined independently, is underlined, as are the locations of primers C2, K1, and K2. The polyadenylation signal "AATAAA" is doubly underlined. The translation initiation site consensus sequence (*40*) is in boldface type. The DNA sequences have been deposited in the GenBank Database under accession numbers AF090321 (1-A) and AF090322 (1-B).

## RESULTS

*Sequence of Subunit 1 in Bovine Heart Tissue: Indication of Alternative Splicing.* Beef heart poly(A)$^+$ RNA was reverse-transcribed using the oligo(dT) region of the adapter primer (AP) and the resultant cDNA amplified by 3′-RACE using 605S, a sequence-specific primer coding for amino acids 5−10 (TQGEDV) of the N-terminus of the protein (*14*), as sense primer. PCR products in the size range 1100−1350 bp were isolated by agarose gel electrophoresis and cloned into *E. coli*. Three clones, with two unique sequences at the 3′-end, were obtained, represented by clones 1-A and 1-B. mRNA coding for the mitochondrial presequence and the 5′-untranslated region (UTR) for subunit 1 in heart was obtained following cloning of 5′-RACE products as described under Materials and Methods. No heterogeneity was apparent in this part of the sequence.

The cDNA sequences corresponding to the full-length mRNAs derived from clones 1-A and 1-B, their common 5′-UTRs, and their deduced amino acid sequences are shown in Figure 1. For the remainder of the report, the corresponding full-length mRNA will be referred to by the name of the 3′-clone. The reading frame contains the 22 residues previously shown to be present at the N-terminus of the mature beef heart subunit 1 (*14*). A typical polyadenylation signal, "AATAAA", is located near the 3′-end. The sequence GCG AAT GAG ATG G, enclosing the initiation codon ATG at bases 56−58, is identical at 7/13 positions with the translation initiation consensus sequence GCC GCC G/ACC ATG G (*40*).

The 2 mRNAs are identical over their first 1125 nucleotides. At this point, IDH 1-B contains an insert of 199 nucleotides, with an in-frame stop codon at bases 1210−

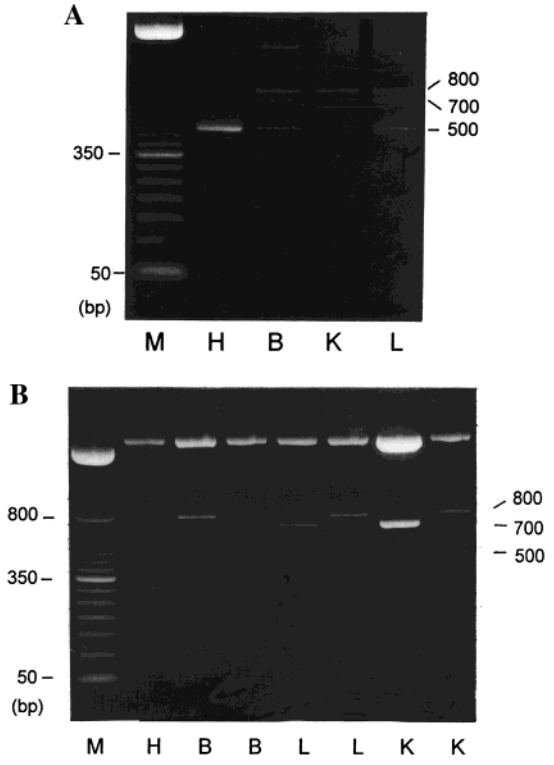


FIGURE 2: Subunit 1 species present in various tissues. (A) Agarose gel electrophoresis of PCR products from heart (H), brain (B), kidney (K), and liver (L) following RT-3′-RACE using primer C2. M, 50 bp DNA ladder. (B) Inserts from selected cloned products of (A) showing the three major PCR products. Visualization is with ethidium bromide.

1212. IDH 1-A, lacking this insert, is identical from base 1126 until its last 4 nucleotides with the remainder of IDH 1-B, starting from base 1324. IDH 1-A utilizes a stop codon at bases 1204−1206 of its sequence. The differences in the C-terminal amino acids encoded by the two mRNAs will be discussed below.

Further experiments (see below) were carried out on RNA from heart muscle to determine the relative abundance of the 1-A and 1-B species. In these experiments, there was evidence only for IDH 1-A. Thus, IDH 1-B mRNA occurs rarely in heart muscle.

*Subunit 1 mRNA Species Expressed in Heart, Brain, Kidney, and Liver*. RT-3′-RACE using 5′-sense primer C2 positioned prior to the differences in IDH 1-A and 1-B (Figure 1C) was carried out on RNA from heart, brain, kidney, and liver. With this primer, IDH 1-A and 1-B would be expected to yield PCR products of 478 and 672 bp, respectively [including estimated 20-base poly(A) stretches and the AUAP sequence].

Species of approximately 500, 700, and 800 bp were present in the initial PCR products of brain, kidney, and liver, with the 500 bp product very prominent in heart (Figure 2A). Representative inserts of each size following cloning are shown in Figure 2B. The 500 bp species was the only form detected in 30 such clones from heart in the 3′-RACE analysis. All three forms were found in clones from kidney, liver, and brain.

Sequence analysis confirmed that the 500 bp product corresponded to IDH 1-A and the 700 bp product to IDH 1-B. The 800 bp product had the sequence of 1-B with an additional 96 nucleotides inserted immediately after the 199

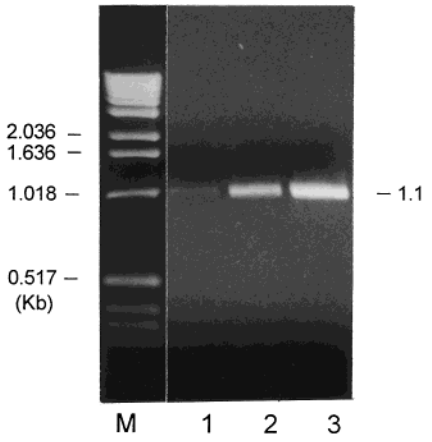


FIGURE 3: Agarose gel electrophoresis of the product obtained following PCR of genomic DNA using primers K1 and K2. Visualization is with ethidium bromide. Lanes: M, 1 kb DNA extension ladder; lanes 1, 2, and 3, PCR mixtures for these lanes contained 1, 10, and 100 ng of DNA template, respectively.

nucleotides, which differentiate 1-B from 1-A (Figure 1C). The mRNA corresponding to the 800 bp product is termed IDH 1-C. The sequence for IDH 1-C, starting at base 824 (numbered as in Figure 1C) through the polyA sequence, has been deposited in the GenBank Database under accession number AF090323. These sequences will be presented and discussed further below.

Three sequences, -ACT C$(A)_n$, -ACT CTT $(A)_n$, and -ACT CTT TGT G $(A)_n$, were found at the 3′-end of the cDNAs. -ACT C$(A)_n$ was found on only one 600 bp species (see below). -ACT CTT $(A)_n$ was present on 500, 700, and 800 bp species, -ACT CTT TGT G $(A)_n$ on 500 and 800 bp species. In sixteen 500 bp species from heart, -ACT CTT $(A)_n$ accounted for seven, while -ACT CTT TGT G $(A)_n$ was present on the remaining nine.

In addition to the three major mRNAs represented by the 500, 700, and 800 bp species, cloning of the RT-3′-RACE products yielded several minor species, six larger than 800 bp, three between 500 and 700 bp. Three large and one small insert (600 bp) were sequenced. The large inserts were unrelated to subunit 1. The small insert (not shown) was similar to the 800 bp species in sequence (see below), but lacked bases 1071−1242 of the IDH 1-B sequence shown in Figure 1C. Its significance was not investigated further.

*PCR To Investigate Potential Differences in the 5′-Ends of IDH 1 in Various Tissues.* RT-PCR was carried out on total RNA from heart, brain, kidney, and liver using primers C11 and C14, which encompass bases 58−856 (Figure 1C), to determine whether there were any differences at the 5′-end of IDH 1 (C14 overlaps with primer C2 used in investigating the 3′-end of the mRNA). A major PCR product of approximately 800 bp was obtained from each tissue, which, on sequence analysis, was identical to the corresponding segment of IDH 1-A/B. Variability was thus confined to the 3′-region of the mRNA.

*Origin of the Different Subunit 1 mRNAs.* To investigate the origin of the three major mRNAs, bovine genomic DNA was subjected to PCR utilizing primers K1 and K2 (Figure 1C), which are common to all forms and enclose the region of difference. A PCR product of 1.1 kb was obtained (Figure 3). The sequence obtained following cloning of this species is compared to the relevant regions of the 500, 700, and 800 bp 3′-RACE products in Figure 4.

```
500(K1/K2)    CCAACATGATTGCAGAGGCAGTGAAGAAGGTGATTAAAGTTGGCAAG---   47
700(K1/K2)    CCAACATGATTGCAGAGGCAGTGAAGAAGGTGATTAAAGTTGGCAAG---   47
800(K1/K2)    CCAACATGATTGCAGAGGCAGTGAAGAAGGTGATTAAAGTTGGCAAG---   47
GENE(K1/K2)   CCAACATGATTGCAGAGGCAGTGAAGAAGGTGATTAAAGTTGGCAAGgta   50

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   agtttagaaaggctgtggggtttagggtcttgggcactaatttctgggac    100
                                                                 a
500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   actgagttggggcctcctttccctccaggtccccagagcttctgctcttc    150
              c
500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   ccccaatcctgtcagggctgtctcattctcttgatcatgtttgcctctta    200

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   gtttcccatctgctatttcctctttcatccactggcagaacttgtagctt    250

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   cctatctagcgtgggggagatggagcagcagagggagatagctgacagg     300

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   atgaggcaccactttctccacacctggaagcctatccttatctgccaccc    350

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   tgcagttaggtatcttccttctgtctgacttgcatttcccatctcactca    400

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   tgctgagcttaacctgctgaggctgcagcagagcatgaatcgtggctctg    450

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   gccctcagaagacaggagcgcacaggtttgtgcttttccaaacctggctg    500

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   cctccttgcatctctgctgcatcgcatcttgttttcctggccttgtcacc    550

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   47
800(K1/K2)    ----------------------------------------------------   47
GENE(K1/K2)   ctctttggattccatttccatcacattgccctgcccctcaccttcatggg    600

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------GTACG    52
800(K1/K2)    ----------------------------------------------GTACG    52
GENE(K1/K2)   ctctcttctcttcttcaccacggttattgccggtggtggctgtagGTACG    650

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    GACTCGAGACATGGGCGGCTACAGCACCACAACCGACTTCATCAAGTCTG    102
800(K1/K2)    GACTCGAGACATGGGCGGCTACAGCACCACAACCGACTTCATCAGGTCTG    102
GENE(K1/K2)   GACTCGAGACATGGGCGGCTACAGCACCACAACCGACTTCATCAAGTCTG    700

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    TCATCGGCCACCTGCACCCCTATGGGGGCTAGAGCCCTTTATTCCCTCCA    152
800(K1/K2)    TCATCGGCCACCTGCACCCCTATGGGGGCTAGAGCCCTTTATTCCCTCCA    152
GENE(K1/K2)   TCATCGGCCACCTGCACCCCTATGGGGGCTAGAGCCCTTTATTCCCTCCA    750

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ACCTCACAAGGACCATACACCTCCCTTCAGTACTGTGGACCAGAGAAGAG    202
800(K1/K2)    ACCTCACAAGGACCATACACCTCCCTTCAGTACTGTGGACCAGAGAAGAG    202
GENE(K1/K2)   ACCTCACAAGGACCATACACCTCCCTTCAGTACTGTGGACCAGAGAAGAG    800

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    CCCTATACTTCTAGACCATAATTACCTTCTGGGCAGAGCCTAG-------    245
800(K1/K2)    CCCTATACTTCTAGACCATAATTACCTTCTGGGCAGAGCCTAGGTTGTTG    252
GENE(K1/K2)   CCCTATACTTCTAGACCATAATTACCTTCTGGGCAGAGCCTAGgttgttg    850

500(K1/K2)    ----------------------------------------------------   47
700(K1/K2)    ----------------------------------------------------   245
                                                          G
800(K1/K2)    GGAGCTGGCTTCCTTAGGGAACAGTTAGGTGGTGGGGGGTTAGGGATGGG    302
GENE(K1/K2)   ggagctggcttccttagggaacagttaggtggtgggggttagggatggg    900

500(K1/K2)    ----------------------------------GTTCGAACTTCT    59
700(K1/K2)    ----------------------------------GTTCGAACTTCT    257
800(K1/K2)    GCCCAGGCCACAGAGATGATAACAATTCTCCCCTACAGGTTCGAACTTCT    352
GENE(K1/K2)   gcccaggccacagagatgataacaattctcccctacagGTTCGAACTTCT    950

500(K1/K2)    GACATGGGTGGTTATGCCACCTGCCAAGACTTCACTGAGGCTGTC   104
700(K1/K2)    GACATGGGTGGTTATGCCACCTGCCAAGACTTCACTGAGGCTGTC   302
800(K1/K2)    GACATGGGTGGTTATGCCACCTGCCAAGACTTCACTGAGGCTGTC   397
GENE(K1/K2)   GACATGGGTGGTTATGCCACCTGCCAAGACTTCACTGAGGCTGTC   995
```

FIGURE 4: Comparison of the 500, 700, and 800 bp cDNA sequences, in the region between primers K1 and K2 (see Figure 1C), with the corresponding genomic DNA sequence. The genomic sequence has been deposited in the GenBank Database under accession number AF090324.

The exon which ends at gene-47 (numbered as in Figure 4) may be spliced to gene-939, yielding the 500 bp species corresponding to IDH 1-A. The same exon may instead be spliced to gene-646 and the subsequent intron retained, yielding the 800 bp species corresponding to IDH 1-C. In some instances, splicing of gene-47 to gene-646 is accompanied by the splicing of gene-843 to gene-939, resulting in the 700 bp species reflected in IDH 1-B. The various mRNAs thus originate from a single gene.

The splicing events are summarized in Figure 5. The sequence ending at gene-47 is designated exon W; that from gene-48 through gene-645, intron w. Exon X comprises gene-646 through gene-843; and the adjacent intron x, gene-844 to gene-938. Exon Y begins at gene-939.

*Sequence Comparison of the Proteins Predicted from Bovine mRNAs with Those from Monkey and Human Heart*. Two proteins differing at their C-terminal ends are predicted from the bovine mRNA species. The first is derived from IDH 1-A, while IDH 1-B and IDH 1-C, which differ in mRNA sequence following a common translation stop codon (bases 1210−1212, Figure 1C), will produce the second. These proteins are compared in Figure 6 with the published sequence based on the mRNA from the monkey testis $\beta$ subunit (*27*), and three human $\beta$ sequences predicted from mRNAs deposited in GenBank. The human mRNAs are described simply as isoform A, isoform B, and heart (*41*−*43*).

Except for the mitochondrial presequence, the adjacent 15 or so amino acids, and the final approximately 15 amino acids, the sequences are remarkably identical. The differences from the beginning until the last 15 amino acids are all conservative. With just 3 exceptions they fall along species lines with the bovine sequences differing from human/monkey. The exceptions occur at monkey amino acid 209, which differs from the other forms, and at human heart at residues 267 and 268, which differ from the other forms.

Analysis of the C-termini reveals two groups of proteins, A and B, the members of which align across species lines. The C-terminal sequences, beginning at the amino acid after the first splice junction in the bovine proteins, are grouped

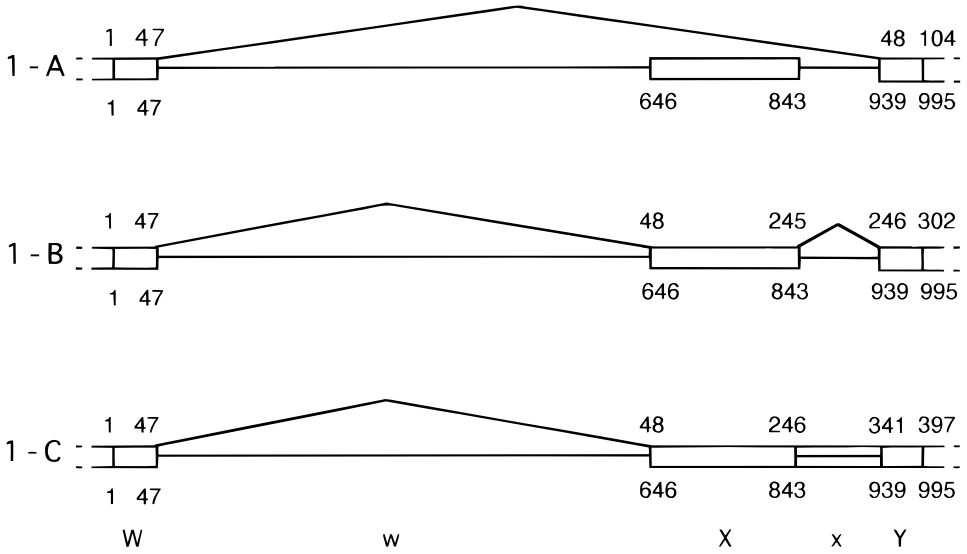


FIGURE 5: Summary of the splicing events of Figure 4. Exons/introns are named as described in the text. The numbers below each mRNA locate the beginning and end of exons in the corresponding gene sequence of Figure 4, while the numbers above identify base positions in that mRNA as numbered in Figure 4.

```
Bovine 1-A      MAALSRVRWLTRALVAAPNPGAWRSLCTSTVAQASSRTQGEDVRVEGAFP  100
Bovine 1-B/C    MAALSRVRWLTRALVAAPNPGAWRSLCTSTVAQASSRTQGEDVRVEGAFP  100
Monkey testis   --ALSGVRWLTRALVSAGNPGAWRGLSTSAAAHAASRSQAEDVRVEGSFP   98
Human A         MAVLSGVRWLTRALVSAGNPGAWRGLSTSAAAHAASRSQAEDVRVEGSFP  100
Human B         MAVLSGVRWLTRALVSAGNPGAWRGLSTSAAAHAASRSQAEDVRVEGSFP  100
Human heart     MAVLSGVRWLTRALVSAGNPGAWRGLSTSAAAHAASRSQAEDVRVEGSFP  100
                  .** *********.* ******.*.**..*.*.**.*.*******.**

Bovine 1-A      VTMLPGDGVGPELMHAVKEVFKAASVPVEFQEHHLSEVQNMASEEKLEQV  150
Bovine 1-B/C    VTMLPGDGVGPELMHAVKEVFKAASVPVEFQEHHLSEVQNMASEEKLEQV  150
Monkey testis   VTMLPGDGVGPELMHAVKEVFKAAAVPVEFQEHHLSEVQNMASEEKLEQV  148
Human A         VTMLPGDGVGPELMHAVKEVFKAAAVPVEFQEHHLSEVQNMASEEKLEQV  150
Human B         VTMLPGDGVGPELMHAVKEVFKAAAVPVEFQEHHLSEVQNMASEEKLEQV  150
Human heart     VTMLPGDGVGPELMHAVKEVFKAAAVPVEFQEHHLSEVQNMASEEKLEQV  150
                ************************.************************

Bovine 1-A      LSSMKENKVAIIGKIHTPMEYKGELASYDMRLRRKLDLFANVVHVKSLPG  200
Bovine 1-B/C    LSSMKENKVAIIGKIHTPMEYKGELASYDMRLRRKLDLFANVVHVKSLPG  200
Monkey testis   LSSMKENKVAIIGKIHTPMEYKGELASYDMRLRRKLDLFANVVHVKSLPG  198
Human A         LSSMKENKVAIIGKIHTPMEYKGELASYDMRLRRKLDLFANVVHVKSLPG  200
Human B         LSSMKENKVAIIGKIHTPMEYKGELASYDMRLRRKLDLFANVVHVKSLPG  200
Human heart     LSSMKENKVAIIGKIHTPMEYKGELASYDMRLRRKLDLFANVVHVKSLPG  200
                **************************************************

Bovine 1-A      YKTRHNNLDLVIIREQTEGEYSSLEHESARGVIECLKIVTRTKSQRIAKF   50
Bovine 1-B/C    YKTRHNNLDLVIIREQTEGEYSSLEHESARGVIECLKIVTRTKSQRIAKF   50
Monkey testis   YMTRHNNLDLVIIREQTEGEYSSLEHESARGVIECLKIVTRAKSQRIAKF   48
Human A         YMTRHNNLDLVIIREQTEGEYSSLEHESARGVIECLKIVTRAKSQRIAKF   50
Human B         YMTRHNNLDLVIIREQTEGEYSSLEHESARGVIECLKIVTRAKSQRIAKF   50
Human heart     YMTRHNNLDLVIIREQTEGEYSSLEHESARGVIECLKIVTRAKSQRIAKF   50
                * ************************************** ********

Bovine 1-A      AFDYATKKGRGKVTAVHKANIMKLGDGLFLQCCEEVAELYPKIKFEKMII  250
Bovine 1-B/C    AFDYATKKGRGKVTAVHKANIMKLGDGLFLQCCEEVAELYPKIKFEKMII  250
Monkey testis   AFDYATKKGRSKVTAVHKANIMKLGDGLFLQCCEEVAELYPKIKFETMII  248
Human A         AFDYATKKGRGKVTAVHKANIMKLGDGLFLQCCEEVAELYPKIKFETMII  250
Human B         AFDYATKKGRGKVTAVHKANIMKLGDGLFLQCCEEVAELYPKIKFETMII  250
Human heart     AFDYATKKGRGKVTAVHKANIMKLGDGLFLQCCEEVAELYPKIKFETMII  250
                ********** ********************************** ***

Bovine 1-A      DNCCMQLVQNPYQFDVLVMPNLYGNIIDNLAAGLVGGAGVVPGESYSAEY  300
Bovine 1-B/C    DNCCMQLVQNPYQFDVLVMPNLYGNIIDNLAAGLVGGAGVVPGESYSAEY  300
Monkey testis   DNCCMQLVQNPYQFDVLVMPNLYGNIIDNLAAGLVGGAGVVPGESYSAEY  298
Human A         DNCCMQLVQNPYQFDVLVMPNLYGNIIDNLAAGLVGGAGVVPGESYSAEY  300
Human B         DNCCMQLVQNPYQFDVLVMPNLYGNIIDNLAAGLVGGAGVVPGESYSAEY  300
Human heart     DNCCMQLVQNPYQFDVVNMPNLYGNIIDNLAAGLVGGAGVVPGESYSAEY  300
                **********.  ********************************

Bovine 1-A      AVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLNLEHHSNMIAEAVK  350
Bovine 1-B/C    AVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLNLEHHSNMIAEAVK  350
Monkey testis   AVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLNLEYHSNMIADAVK  348
Human A         AVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLNLEYHSSMIADAVK  350
Human B         AVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLNLEYHSSMIADAVK  350
Human heart     AVFETGARHPFAQAVGRNIANPTAMLLSASNMLRHLNLEYHSSMIADAVK  350
                ************************************ ** *** ***

Bovine 1-A      KVIKVGKVRTSDMGGYATCQDFTEAVIGALS--NP--  383
Bovine 1-B/C    KVIKVGKVRTRDMGGYSTTTDFIKSVIGHLHPYGG--  385
Monkey testis   KVIKVGKVRTSDMGGYATCHDFTEAVIAAL----PHP  381
Human A         KVIKVGKVRTSDMGGYATCHDFTEAVIAAL----PHP  383
Human B         KVIKVGKVRTRDMGGYSTTTDFIKSVIGHLQTKGSNL  387
Human heart     KVIKVGKVRTRDMGGYSTTTDFIKSVIGHLQTKGS--  385
                ********** ***** *  **   **  *
```

FIGURE 6: Comparison of the protein sequences predicted from the bovine subunit 1 mRNA species with those of the corresponding $\beta$ subunits from monkey (*27*) and human origins (*41*−*43*). Residues that are perfectly conserved are denoted by an asterisk and those well conserved by a boldface point. Mitochondrial presequences, determined (bovine) or predicted (monkey, human), are in boldface type. Carboxy-terminal sequences beyond the splice point difference of bovine sequences 1-A and 1-B/C are underlined.

in Figure 7. Group A contains bovine 1-A, monkey testis, human $\beta$ form A, and the pig heart $\beta$ C-terminal peptide [obtained at the protein level when Cys-containing peptides were identified (*44*)]. Group B consists of bovine 1-B, human $\beta$ form B, and human heart $\beta$.

The two groups differ notably in the number of positively charged residues in this short C-terminal region. Including the C-terminus, group A proteins have four negatively charged groups, and one (bovine 1-A), two (pig heart), or three (monkey and human A) positively charged side chains. Group B proteins have three negatively charged groups, and each has five positively charged side chains. Comparison of the bovine 1-A and 1-B/C proteins (without the mitochondrial presequences) indicates that they differ markedly in pI due to the charge differences concentrated in this one region. The pI of 1-A is 6.95, that of 1-B/C is 8.13, and the predicted charges at pH 7.0 are −0.13 and +3.09, respectively.

The above classifications of the subunit 1 proteins raise the question of whether the mRNA processing at the 3′-end is the same within each group. Analysis of the corresponding mRNAs (Figure 8) shows that the proteins within group A are derived from mRNAs which are processed identically. Within Group B, human heart mRNA is processed in the same way as bovine 1-C mRNA. Splicing of human form B mRNA in the region of interest is initiated from an internal exon "W" 5′-splice site, an AG corresponding to a position five bases before the TAG stop codon utilized in the other Group B mRNAs. The "A-1228" of the splice site is present in both human B and human heart mRNAs, but a "T" is present at this position in the bovine sequences. The 3′-splice

```
A.

Bovine 1-A      VRTSDMGGYATCQDFTEAVIGALSNP      383
Monkey testis   VRTSDMGGYATCHDFTEAVIAALPHP      381
Human A         VRTSDMGGYATCHDFTEAVIAALPHP      383
Pig heart         TSDMGGYATCQDFTEAVIGALSHP
                  +  -       +-  -       + -

B.

                  + +-          -  +    + + +  -
Bovine 1-B/C    VRTRDMGGYSTTTDFIKSVIGHLHPYGG    385
Human B         VRTRDMGGYSTTTDFIKSVIGHLQTKGSNL  387
Human heart     VRTRDMGGYSTTTDFIKSVIGHLQTKGS    385
```

FIGURE 7: Comparison of carboxy-terminal sequences beyond the bovine splice point difference from bovine, monkey (*27*), human (*41*−*43*), and pig (*44*) sources. The sequences are grouped according to the two sequence types, focusing on charge differences. Underlined charges are unique to all members of one group. Charges not underlined are present in some but not all members of one group.

site for this junction in human isoform B is that used for all Group A proteins and for bovine 1-B. However, splicing from the internal exon W 5′-site in human form B changes the reading frame for translation at this point with respect to other predicted proteins and results in a new termination codon shortly after the splice site. The last 4 amino acids of the predicted human isoform B protein are thus coded by 1 base prior to the splice junction and 11 bases after. Nonetheless, the resulting C-terminal protein sequence for the human B form is the same as that predicted for the human heart mRNA, with the addition of two final amino acids.

The identity in mRNA sequence among the different species is remarkable at the coding exon level, consistent with the homology at the protein level. However, even introns "x" in cow and human, which can be compared in bovine 1-C and human heart mRNAs, have approximately 70% identity.

## DISCUSSION

Three versions of IDH subunit 1/$\beta$ mRNA, varying in splicing in the 3′-region from exon W to exon Y, have been identified in bovine tissues. The form with the most straightforward splicing process, retaining exon X and eliminating introns w and x, results in mRNA 1-B and is found in brain, kidney, and liver. Form 1-C, in which intron w is eliminated but both exon X and intron x are retained, is also present in these tissues. In form 1-A, exon W is spliced directly to exon Y, eliminating intron w, exon X, and intron x. While present along with the other forms in brain, kidney, and liver, this is the mRNA characteristic of heart muscle.

While the factors which regulate production of the various mRNAs remain to be determined, certain features are notable. The 5′-splice site of exon W which is used in all mRNAs conforms exactly to the consensus sequence AG|GURAGU (*45*). The X exon appears in two of the three mRNAs, 1-B and 1-C. While the 3′-splice site of exon X conforms to the consensus sequence NYAG|, the polypyrimidine tract associated with it is not adjacent as is found usually (*46*), but 24 bases upstream. The 5′-splice site of the X exon differs at three positions (AG|GUUGUU) from the accepted 5′ consensus sequence. The unusual features involving the X exon may contribute to the different splicing patterns. The 3′-splice site for exon Y and its preceding polypyrimidine tract have no unusual features.

The finding of multiple polyA addition sites using the same AATAAA polyadenylation signal is unusual (-ACTCA$_n$, -ACTCTTA$_n$, -ACTCTTTGTGA$_n$). The three sites are located 14−20 bases downstream of the cleavage signal, within the usual range of 10−30 bases (*47*). The preference for the base immediately prior to cleavage is A > U > C ≫ G (*48*). Although it is possible that the longest product actually is cleaved after A, the first two clearly are not. A GU-rich motif is usually found 20−40 bases downstream of the cleavage site (*47*). Providing that this is present in the IDH subunit 1 gene, it is possible that there is simply not one strong site for cleavage/polyadenylation and that the several sites found result. One example has been located of several cleavage products resulting from a single AATAAA signal during an *in vitro* study of a wild-type pre-mRNA (*49*). It is also possible that multiple cleavage sites are often found in transcripts isolated directly from tissues but are not reported.

The three mRNAs result in two protein forms, 1-A and 1-B/C, differing notably in charge in their C-terminal peptides. The extensive mRNA analysis of bovine heart in this study indicates that the 1-A form of the subunit is the preferred form for this tissue. In porcine heart also, protein analysis clearly shows that the 1-A form is the expressed protein (*44*). In bovine brain, kidney, and liver, the mRNA results from the present study indicate that both proteins are to be expected. In the monkey testis study, there is insufficient documentation to indicate whether the mRNA reported is the major or only form present (*27*), and the mRNA

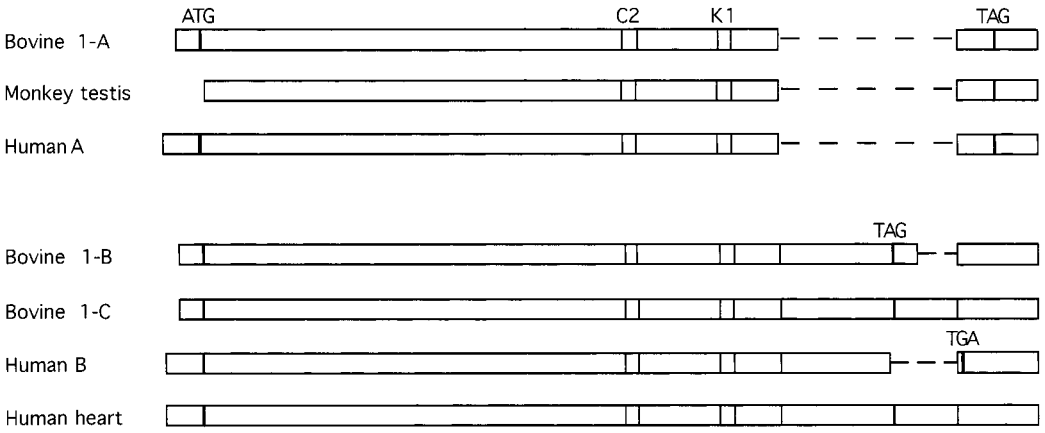


FIGURE 8: Schematic of the sequence comparison of the mRNAs giving rise to the proteins compared in Figure 6.

sequences reported from the human have little or no documentation (*41*−*43*). However, the fact that the predicted proteins from these mRNAs group across species lines with the two bovine proteins gives further credence to their physiological significance.

While no heterogeneity has been found at the mRNA level in the bovine 3/4 α subunit (Rushbrook and Weiss, unpublished observations), alternatively spliced mRNAs resulting in three protein forms with differential tissue expression have been reported for the 2/γ subunit in human tissues (*38*, *50*). One protein contains a 12 residue insert located 33 residues from the C-terminus of the second. The third form does not contain the insert, and the final 33 residues of the first 2 forms are replaced by a different 20 residue sequence. Forms 1 and 2 have considerably more charged residues than version 3.

Thus, both the 1/β and 2/γ subunits, considered to be involved in allosteric regulation, and not the 3/4 α subunits involved in direct catalysis, have alternatively spliced forms varying in their C-terminal properties with tissue-specific expression. This is suggestive of a mechanism for modulation of allosteric regulation according to the needs of different tissues. The conservation, in the eukaryotic subunits, of key residues involved in substrate binding by the *E. coli* subunits indicates some important degree of three-dimensional structure retention. Nonetheless, the 1-A and 1-B/C protein sequences did not readily model on the *E. coli* three-dimensional template using the program ProModII through the Automated Protein Modelling Server SWISS-MODEL (Sobocka and Rushbrook, unpublished observation). At the primary structure level, a striking difference between the *E. coli* and all the eukaryotic subunit sequences is the lack, in the eukaryotic subunits, of most of the residues involved in *E. coli* dimer formation. Thus, the way in which the eukaryotic subunits interact to form the core tetramer is unknown, as are the structural features and interactions important for allosteric regulation. Three-dimensional structural studies of the eukaryotic subunits are clearly indicated.

## NOTE ADDED IN PROOF

While this paper was in press, a cDNA from human uterus encoding an NAD+-IDH subunit 1/β with the C-terminus of Group B proteins in Figure 7 was submitted to GenBank (accession no. AL050094).

## REFERENCES


1. Barnes, L. D., Kuehn, G. D., and Atkinson, D. E. (1971) *Biochemistry 10*, 3939.
2. Plaut, G. W. E., Schramm, V. L., and Aogaichi, T. (1974) *J. Biol. Chem. 249*, 1848.
3. Rushbrook, J. I. (1978) Ph.D. Dissertation, Rutgers, The State University of New Jersey.
4. Cohen, P. F., and Colman, R. F. (1972) *Biochemistry 11*, 1501.
5. Chen, R. F., and Plaut, G. W. E. (1963) *Biochemistry 2*, 1023.
6. Ehrlich, R. S., and Colman, R. F. (1982) *J. Biol. Chem. 257*, 4769.
7. Harvey, R. A., Heron (Rushbrook), J. I., and Plaut, G. W. (1972) *J. Biol. Chem. 247*, 1801.
8. Denton, R. M., Richards, D. A., and Chin, J. G. (1978) *Biochem. J. 176*, 899.
9. Nichols, B. J., Rigoulet, M., and Denton, R. M. (1994) *Biochem. J. 303*, 461.
10. Giorgio, N. A. J., Yip, A. T., Fleming, J., and Plaut, G. W. E. (1970) *J. Biol. Chem. 245*, 5469.
11. Shen, W. C., Mauck, L., and Colman, R. F. (1974) *J. Biol. Chem. 249*, 7942.
12. Fan, C. C., Lin, J. P. F., and Plaut, G. W. E. (1975) *J. Biol. Chem. 250*, 2022.
13. Illingworth, J. A. (1972) *Biochem. J. 129*, 1119.
14. Zeng, Y., Weiss, C., Yao, T.-T., Huang, J., Siconolfi-Baez, L., Hsu, P., and Rushbrook, J. I. (1995) *Biochem. J. 310*, 507.
15. Rushbrook, J. I., and Harvey, R. A. (1978) *Biochemistry 17*, 5339.
16. Ramachandran, N., and Colman, R. (1980) *J. Biol. Chem. 255*, 8859.
17. 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.
18. Hurley, J. H., Dean, A. M., Koshland, D. E., Jr., and Stroud, R. M. (1991) *Biochemistry 30*, 8671.
19. Hurley, J. H., Dean, A. M., Thorsness, P. E., Koshland, D. E., Jr., and Stroud, R. M. (1990) *J. Biol. Chem. 265*, 3599.
20. Stoddard, B. L., and Koshland, D. E., Jr. (1993) *Biochemistry 32*, 9317.
21. Stoddard, B. L., Dean, A., and Koshland, D. E., Jr. (1993) *Biochemistry 32*, 9310.
22. Garnak, M., and Reeves, H. C. (1979) *J. Biol. Chem. 254*, 7915.
23. Garnak, M., and Reeves, H. C. (1979) *Science 203*, 1111.
24. Borthwick, A. C., Holms, W. H., and Nimmo, H. G. (1984) *Eur. J. Biochem. 141*, 393.
25. Cupp, J. R., and McAlister-Henn, L. (1991) *J. Biol. Chem. 266*, 22199.
26. Cupp, J. R., and McAlister-Henn, L. (1992) *J. Biol. Chem. 267*, 16417.
27. Nichols, B. J., Perry, A. C., Hall, L., and Denton, R. M. (1995) *Biochem. J. 310*, 917.
28. Nichols, B. J., Hall, L., Perry, A. C. F., and Denton, R. M. (1993) *Biochem. J. 295*, 347.
29. Kim, Y. O., Oh, I. U., Park, H. S., Jeng, J., Song, B. J., and Huh, T. L. (1995) *Biochem. J. 308*, 63.
30. Ehrlich, R. S., and Colman, R. (1983) *J. Biol. Chem. 258*, 7079.
31. Ehrlich, R. S., and Colman, R. (1984) *J. Biol. Chem. 259*, 11936.
32. Cupp, J. R., and McAlister-Henn, L. (1993) *Biochemistry 32*, 9323.
33. Zhao, W.-N., and McAlister-Henn, L. (1997) *J. Biol. Chem. 272*, 21811.
34. Plaut, G. W. E., and Aogaichi, T. (1963) *J. Biol. Chem. 243*, 5572.
35. Willson, V. J. C., and Tipton, K. F. (1979) *J. Neurochem. 33*, 1239.
36. Willson, V. J. C., and Tipton, K. F. (1980) *Eur. J. Biochem. 109*, 411.
37. Willson, V. J. C., and Tipton, K. F. (1981) *Eur. J. Biochem. 117*, 65.
38. Brenner, V., Nyakatura, G., Rosenthal, A., and Platzer, M. (1998) *Genomics 44*, 8.
39. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds. (1998) *Current Protocols in Molecular Biology*, pp A.1.6−A.1.9, John Wiley and Sons, Inc., New York.
40. Kozak, M. (1987) *Nucleic Acids Res. 15*, 8125.
41. Ko, H., Kim, Y., Park, H., Oh, I., Yeo, S., Hong, S., Chae, B., Kim, S., Son, B., Lee, Y., Yeo, H., and Huh, T.-L. (1996) GenBank accession no. U49283, submitted 2/96.
42. Ko, H. J., Kim, Y. O., and Huh, T. L. (1999) GenBank accession no. AF023265, submitted 1/99.
43. Ko, H. J., Kim, Y. O., and Huh, T. L. (1999) GenBank accession no. AF023266, submitted 1/99.
44. Huang, Y.-C., and Colman, R. F. (1990) *Biochemistry 29*, 8266.
45. Kramer, A. (1996) *Annu. Rev. Biochem. 65*, 367.
46. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986) *Annu. Rev. Biochem. 55*, 1119.
47. Colgan, D. F., and Manley, J. L. (1997) *Genes Dev. 11*, 2755.

48. Chen, F., MacDonald, C. C., and Wilusz, J. (1995) *Nucleic Acids Res. 23*, 2614.
49. Sheets, M. D., Ogg, S. C., and Wickens, M. P. (1990) *Nucleic Acids Res. 18*, 5799.
50. Huh, T.-L., Kim, Y.-O., Kim, S.-H., Jo, S.-H., and Lim, K.-E. (1995) GenBank accession no. 4758581, submitted 10/95.

BI991691I