Aspartyl β-Hydroxylase (Asph) and an Evolutionarily Conserved Isoform of Asph Missing the Catalytic Domain Share Exons with Junctin*

The mouse aspartyl β-hydroxylase gene (Asph, BAH) has been cloned and characterized. The mouse BAH gene spans 200 kilobase pairs of genomic DNA and contains 24 exons. Of three major BAH-related transcripts, the two largest (6,629 and 4,419 base pairs) encode full-length protein and differ only in the use of alternative polyadenylation signals. The smallest BAH-related transcript (2,789 base pairs) uses an alternative 3′ terminal exon, resulting in a protein lacking a catalytic domain. Evolutionary conservation of this noncatalytic isoform of BAH (humbug) is demonstrated in mouse, man, and Drosophila. Monoclonal antibody reagents were generated, epitope-mapped, and used to definitively correlate RNA bands on Northern blots with protein species on Western blots. The gene for mouse junctin, a calnexin-binding protein, was cloned and characterized. The mouse junctin cDNA sequence and biochemical analyses, BAH can be divided into four distinct regions (12–14). BAH appears to be a type 2 integral membrane protein containing a short amino-terminal domain that projects into the cytoplasm. The amino-terminal domain is followed by a predicted transmembrane domain and a highly charged region that projects into the lumen of the endoplasmic reticulum. The COOH-terminal region of BAH, which corresponds to the 52- and 56-kDa polypeptides isolated from bovine liver, contains the aspartyl hydroxylase catalytic domain. This domain contains dibasic glycine and His2 motifs that have been shown to be critical for catalytic activity (13, 15).

Northern blotting of bovine liver mRNA with a 5′ bovine BAH cDNA probe revealed two mRNA species of ~2.6 and 6.6 kb (12). More recently, a 4.3-kb cDNA isolated by immunoscreening against highly expressed antigens in a human osteosarcoma cell line identified protein bands with masses ranging from 50 to 120 kDa (16), suggesting that smaller forms of human cancer cell lines identified protein bands with masses ranging from 50 to 120 kDa (16), suggesting that smaller forms

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF289199, AF289200, AF289205–AF289215, and AF289486–AF289494.

§ To whom correspondence should be addressed; Dept. of Applied Biotechnology, Experimental Station, E400/5446, Wilmington, DE 19880-0400. Tel.: 302-695-9180; Fax: 302-695-9401; E-mail: joseph.e.dinchuk@dupontpharma.com.

† The abbreviations used are: BAH, aspartyl β-hydroxylase; PCR, polymerase chain reaction; RT-PCR, reverse transcription-polymerase chain reaction; UTR, untranslated region; BAC, bacterial artificial chromosome; PBS, phosphate-buffered saline; RACE, rapid amplification of cDNA ends; kb, kilobase pair(s); bp, base pair(s); EST, expressed sequence tag; mAb, monoclonal antibody; oligo, oligonucleotide.
represent proteolytic degradation products of the full-length protein.

In order to better understand the organization and regulation of BAH, we have cloned and characterized the BAH genomic locus. Three major RNA forms of BAH-related sequences were cloned and analyzed. In addition to the full-length murine BAH cDNA, a novel, alternatively spliced, transcript was obtained. This transcript (humbug) encodes a protein identical to BAH through the NH₂-terminal half of the protein, but completely lacks the catalytic domain of BAH. The genomic organization of BAH is described, and the transcription start site is mapped. Monoclonal antibody reagents that allow specific recognition of the catalytically active version of human BAH were produced, characterized, and used to definitively assign bands observed on Northern blots to those observed on Western blots from human tumor cells.

A short region of homology between canine junctin, a 28-kDa calsequestrin-binding protein found in skeletal and cardiac muscle, and bovine BAH has been previously described (17). To understand the relationship between BAH, humbug, and junctin, a mouse junctin cDNA was cloned from a mouse heart cDNA library. The humbug Mouse BAH, humbug, and junctin were found to be identical in sequence over the 220 nucleotides spanning the exon 2 and 3. Further genomic organization revealed three additional junctin exons contained within the BAH genomic locus. Analysis of mouse heart mRNA definitively demonstrated that BAH/humbug is expressed as a chimera that results in fusion of the amino-terminal half of junctin to the lumenal domain of BAH/humbug. A possible function of this unusual chimera is discussed.

**Experimental Procedures**

cDNA Cloning and Characterization—A 580-bp fragment representing a partial murine BAH cDNA was obtained by RT-PCR using murine liver cDNA and primers designed from bovine BAH (12). The resulting clone was used to screen a murine liver cDNA library (18). The 5' end of mouse BAH was amplified from a liver cDNA library using PCR primers designed from murine BAH and the library vector. 5'- and 3'-RACE were performed using commercially available systems as described by the manufacturers (LifeTechnologies, Inc.; CLONTECH, Palo Alto, CA). Sequences were analyzed using Sequencher (Gene Codes Corp., Ann Arbor, MI) and the CGG sequence analysis package (Genetics Computer Group, Inc., Madison, WI).

**Cloning Drosophila Aspartyl (Asparaginyl) β-Hydroxylase—**Drosophila β-aspartyl hydroxylase sequences were obtained by EST data base searching and degenerate PCR using cDNA from embryo. Primers were designed from the human, bovine, and Caenorhabditis elegans sequences; optimum PCR amplification was seen with a degenerate sense primer (CARAGRTCNCTNTAYAA) and an antisense primer designed using Drosophila codon frequencies (CTCGTGCTCGAAGGAAT). Two of these antibodies (HBOH-1 and HBOH-2) were characterized and used for Western analysis of human cell lines.

Epitope Mapping of HBOH-1 and HBOH-2—Monoclonal antibodies HBOH-1 and HBOH-2 were biotin-labeled using NHS-LC-biotin (Pierce) as follows. Approximately 50 μg of antibody was incubated with 2 μl of 3.9 mM NHS-LC-biotin in PBS buffer. After 2 h, 1 ml of 1 M ethanolamine was added to block unreacted biotin reagent. To minimize losses during washing, 20 μl of 50 mg/ml bovine serum albumin were added and the labeled antibody washed three times with 2 ml of PBS using a Centricon concentrator (Millipore, Bedford, MA) with a 30-kDa cut-off. Phage displayed peptide libraries encoding seven or eight random amino acids (26, New England Biolabs, Beverly, MA) were screened to identify peptide-binding phage as follows. One microgram of biotinylated antibody was incubated with approximately 10⁷ plaque-forming units of the library at 4 °C overnight. The phage antibody mixture was incubated for 10 min on streptavidin-coated plates (0.5 mg/ml in 0.1 M NaHCO₃, pH 8.0) that had been blocked with bovine serum albumin (29 mg/ml in 0.1 M NaHCO₃). Plates were washed 10 times with 3 ml of Tris-buffered saline containing 0.5% Tween 20 over 30 min. Bound phage were eluted with glycine buffer, pH 2.2, neutralized with 2 μl of 1 M Tris, and titered for plaque-forming units on Xgal/LacZ likewise plates were amplified by infection of XL-1 Blue E. coli and used for additional rounds of selection. After three rounds of selection, individual plaques were picked and amplified to recover DNA for sequencing.

Epitope Mapping FB50 Monoclonal—FB50 was the kind gift of Dr. Jack R. Wands (16). A set of 90 peptides comprising the region between residues 82 and 354 of human BAH were synthesized by Chiron Mimotopes (Raleigh, NC). The fifteen residue peptides were synthe-
sized covalently attached to pins to facilitate antibody screening in a 96-well format. The pins were used according to the manufacturer’s instructions to identify peptide fragment(s) recognized by the antibodies in question.

**Antisense Studies and Western Analysis—** Oligonucleotides for antisense experiments were synthesized and purified using protocols previously described (24). A549 cells were plated at a density of $5.5 \times 10^5$ cells/6-cm plate and cultured in high glucose Dulbecco’s modified Eagle’s medium containing 10% fetal calf serum, 2 mM glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin, for 24 h. The following day, the oligonucleotides (0.4 μM final concentration) were mixed with lipofectin (20 μg/ml final concentration, Life Technologies, Inc.) and the cells were incubated with this complex for 4 h. The oligonucleotide-lipofectin complex was washed off, and fresh medium containing 10% fetal calf serum was added to the cells. A second addition of the complex was carried out 48 h later. Cells were harvested for Western analysis 24 h after this second addition and resuspended in SDS lysis/gel-loading buffer to a final concentration of $2 \times 10^6$ cells/ml buffer. Five μl of lysate solution was fractionated on a 7.5% SDS-polyacrylamide gel (Bio-Rad) for 1 h at 100 V in Tris-glycine buffer. Protein was transferred onto a nitrocellulose membrane by electrotransfer and probed with primary antibody (FB-50 or HBOH-1 at 1:5000 and 1:500 dilutions, respectively) at 4 °C overnight. Secondary goat anti-mouse antibodies conjugated to horseradish peroxidase (Vectastain ABC kit, Vector Laboratories, Burlingame, CA) were applied according to manufacturer’s instructions. Detection was accomplished using the Renaissance® Western blot chemiluminescence reagent (PerkinElmer Life Sciences).

**RESULTS**

**Murine BAH mRNA Is Present as Three Major Bands in Multiple Tissue Northern Blots—** Previous Northern analyses of human and bovine RNAs have demonstrated that multiple RNA species contain BAH-related sequences (12, 14, 16). To clarify the relationship of BAH to these transcripts, a detailed RNA, cDNA, and genomic analysis of the murine BAH gene was undertaken. Northern analysis of poly(A)$^+$ RNAs derived from multiple mouse tissues probed with a murine BAH cDNA, which includes the BAH intralumenal domain (probe A, Fig. 1D), demonstrated the presence of three different size transcripts (2.8, 4.5 and 6.6 kb). These transcripts are expressed in a wide variety of normal murine tissues (Fig. 1A).

**Characterization of the Multiple BAH Transcripts by cDNA Cloning and Northern Analysis—** Seven BAH cDNA clones were isolated from a murine liver cDNA library. The largest clone included a polyadenylation signal and extended 5′ to include the transmembrane region of BAH, but stopped short of the translation start site (FB-50 or HBOH-1 at 1:5000 and 1:500 dilutions, respectively) at 4 °C overnight. Secondary goat anti-mouse antibodies conjugated to horseradish peroxidase (Vectastain ABC kit, Vector Laboratories, Burlingame, CA) were applied according to manufacturer’s instructions. Detection was accomplished using the Renaissance® Western blot chemiluminescence reagent (PerkinElmer Life Sciences).

---

**FIG. 1. Mouse and human BAH cDNAs and relationship to Northern blots.** A, a normalized multiple tissue Northern blot from adult normal mouse using BAH probe A. B, a Northern blot of mouse liver hybridized with probe B (1), and then stripped and hybridized with probe C (2). C, Northern blot of mouse liver hybridized with probe B (1) and then stripped and hybridized with probe D (2). D, schematic diagrams of the four alternative spliced genes that are a part of the BAH genomic locus including the three isoforms of BAH seen in Northern blots (1–3) and mouse junctin (4). The four regions of BAH are indicated in color: green, NH$_2$-terminal cytoplasmic; yellow, transmembrane region; purple, highly charged luminal region; red, catalytic domain. Junctin-specific sequences are indicated in blue. nts, nucleotides.
middle transcript band detected in the Northern analysis. To
examine the 3’ end of the BAH transcript, a 3’-RACE experi-
ment was performed, priming from the 3’-UTR mouse BAH
close. Fragments isolated from this RACE experiment identi-
fied a second class of transcript which terminated after a sec-
ond polyadenylation signal 6,629 bp downstream from the
transcription start site (Fig. 1D, 2, accession no. AF289487). 
This transcript is consistent with the longest size transcript,
6.6 kb, identified by Northern analysis. Both the 4.5- and
6.6-kb clones are identical in sequence through the cytoplasmic
domain, transmembrane region, highly charged region and
catalytic domain. They share a common 3’-UTR up to the first
polyadenylation signal at 4.5 kb. The longer clone extends the
3’-UTR for another 2.1 kb until it terminates at a second
polyadenylation signal.

Characterization of additional murine BAH related cDNA
clones identified a class whose homology with the full-length
BAH clone terminated 1,094 bp downstream of the transcrip-
tional start site (Fig. 1D, 3, accession no. AF289488). This
cDNA is consistent in size with the smallest (2.8 kb) transcript
identified on the mouse tissue Northern blots (Fig. 1, A and D
(3)). Comparison of the sequence of this truncated clone to BAH
revealed that they are identical in the cytoplasmic and trans-
membrane domains, and diverge near the beginning of the
BAH catalytic domain. The coding sequence of the truncated
close terminates at a stop codon two amino acids downstream
of the start codon (Fig. 2). This truncated form of the BAH
transcript (humbug) does not include the catalytic domain and
therefore is predicted to be catalytically inactive as an aspartyl
hydroxylase.

The three BAH-related transcripts observed in Northern
blots were directly related to the three classes of mouse liver
cDNAs by further Northern analysis. Northern blots of adult
mouse liver RNA probed with a 5’ fragment (probe B) that
was contained within all three BAH-related transcripts (Fig. 1, B(lane 1) and C (lane 1) identified all three bands (2.8, 4.5, and
6.6 kb) previously seen on Northern analysis (Fig. 1A). Re-
hybridization of this Northern with a BAH catalytic domain
probe (probe C) recognized both of the larger transcripts (4.5
and 6.6 kb), but not the 2.8-kb band (Fig. 1B, lane 2). In
contrast, when the Northern blot depicted in Fig. 1C (lane 1)
was re-probed with a fragment from the 3’-UTR of the trun-
cated clone (probe D), only the 2.8-kbp transcript was identified
(Fig. 1C, lane 2). These results are consistent with the 2.8-kb
transcript representing humbug, the truncated catalytically
inactive form of BAH, and the two larger transcripts represent-
ing the full-length catalytically active version of BAH, differing
only in their use of alternative polyadenylation signals.

The humbug Transcript Exists in Drosophila, Mouse, and
Man—To determine if the catalytically inactive form of BAH
(humbug) exists in humans, 3’-RACE was performed on a
human liver cDNA library using a primer chosen from the
highly charged region of human BAH. A 2.8-kb human ortholog
of mouse humbug was obtained (accession no. AP289489). The
encoded amino acid sequences of the mouse and human hum-
bug transcripts indicate that protein truncation occurs at the
identical position, 2 amino acids downstream of the break in
homology with the full-length BAH sequence (Fig. 2).

With the identification of a major alternatively spliced hum-
bug transcript in the mouse and human, and the knowledge
that BAH activity has previously been demonstrated in insect
cells (25), the existence of humbug in Drosophila was explored.
Drosophila BAH sequences were obtained by degenerate PCR,
EST data base searching, and 5’- and 3’-RACE. Two different
cDNAs were identified. The first contained a contiguous open
reading frame of 2,358 bp (786 amino acids) that included a
putative cytoplasmic domain, transmembrane domain, a
highly charged region, and a catalytic domain. The second
cDNA, however, encoded a truncated open reading frame of
1149 bp (383 amino acids) whose structure and sequence are
consistent with it being the Drosophila ortholog of humbug
(Figs. 4C and 2, accession nos. AF289493 and AF289494).

Mouse junctin Shares 100% Sequence Identity over a 220-bp
Region with Mouse BAH and humbug—Previously, a cross-
species comparison of bovine BAH and dog junctin has shown
that these genes share 95% amino acid identity over a 73-
residue region, which includes the putative transmembrane
domain of BAH (17). To determine if this high degree of homol-
gy is derived from two closely related genes or whether junctin
shares exons with BAH and humbug, a mouse junctin cDNA
was cloned (accession no. AP289490) from a mouse heart cDNA
library. When the sequences of mouse BAH/humbug derived
from mouse liver and mouse cardiac junctin are compared, a
region of 100% identity spanning 220 nucleotides is observed
(Fig. 3). The region of identity includes the BAH/humbug trans-
membrane region and the first 42 amino acids of the protein on
the endoplasmic reticulum lumenal side of the membrane. The
sequence of the cytoplasmic domain of junctin and the down-
stream lumenal sequence diverge from BAH/humbug (Figs. 1D
and 3).

Mouse BAH Genomic DNA Spans over 200 kb and Encodes
Three Distinct Proteins: BAH, humbug, and junctin—To char-
acterize the genomic organization of BAH, humbug, and junc-
tin, a 129/SvJ mouse ES cell BAC genomic library was
screened. Probes were used from both the 5’ and 3’ ends of
mouse BAH cDNA to maximize the chances of obtaining the
genomic regions spanning the entire open reading frame. Three
separate BAC genomic clones were isolated, shotgun-cloned,
and sequenced (Fig. 4). This analysis indicated that BAH is
encoded by 24 exons and extends over 200 kb of genomic DNA
(accession nos. AF289205–AF289215). Interestingly, exons 1
through 14 splice in a codon position that would permit alter-
native splicing to occur without disrupting the reading frame
(Table I).

The sequence homology between full-length BAH and hum-
bug ends precisely at the end of BAH exon 13 (Table I, Fig. 2).
Sequencing of the region between BAH exons 13 and 14 re-
vealed the presence of the humbug alternatively spliced exon
(exon 14a) downstream of BAH exon 13 (Fig. 4B, 7, accession
no. AF289207). The distance between exon 14a and exon 14
was determined to be ~23 kb. Translation of an exon 13 to exon
14a alternative splice indicates that exon 14a encodes for only
two additional amino acids followed by a stop codon (Fig. 2,
Table I), in agreement with the sequence of the previously
isolated humbug cDNA. This demonstrates that the humbug
transcript is produced by alternative splicing of BAH exon 13 to
humbug exon 14a.

A comparison of junctin sequence to the 24 BAH exons re-
veals that the 220 nucleotides of identity exactly match exons 2
and 3 of BAH. This demonstrates that the common region of
junctin and BAH/humbug arise by exon sharing and suggests
that the upstream and downstream sequences of junctin are
encoded in other exons that are distinct from the 24 BAH
exons. Sequence analysis of the genomic region between BAH
exons 3 and 4 revealed two additional exons, 4a and 5a (42 and
1311 bp, respectively) that encode the remaining part of the
lumenal domain of junctin (accession no. AP289200). Sequence
analysis of the BAH genomic region between exons 1 and 2
identified one small exon, 1a, which encoded the cytoplasmic
domain of junctin and includes the initiator methionine (acces-
sion no. AP289199). Exon 1a is located approximately 8 kb
downstream of exon 1 and 4 kb upstream of exon 2. Thus,
mouse junctin is encoded in 5 exons spread across ~25 kb,
where exons 2 and 3 are shared with BAH/humbug (Fig. 4, A and B (10)).

Alternative Splicing and Conservation of BAH Structure from Insect to Man—Northern analysis of BAH and humbug expression in mouse cardiac RNA suggested that the transcripts of these genes in cardiac tissues were slightly smaller than in other mouse tissue RNAs (Fig. 1A). To investigate this observation, 5′-RACE experiments utilizing BAH/humbug
exon 5 and 6 primers (chosen to be outside the exons shared with junctin) were performed on mouse heart tissue RNA. Twelve BAH/humbug 5'-RACE cDNAs were analyzed, and all were found to use the first exon of junctin (Fig. 4B, 3–6 and 8–9). This suggests that in heart tissue, where junctin is highly expressed (see below), alternative splicing can occur such that use of the promoter associated with exon 1a leads to fusion of this exon into transcripts including BAH/humbug downstream exons. In fact, detailed analysis of the 12 isolated cDNAs revealed multiple alternative splicing patterns that included

Fig. 3. A partial nucleotide comparison of mouse BAH (top), mouse humbug (middle), and mouse junctin (bottom). Start codons are indicated for the respective sequences by outlined boxes. Both BAH and humbug share the same start codon, while the junctin open reading frame starts just upstream of the area of 100% identity (220 nucleotides) common to all three genes. The stop codon of junctin is indicated.
transcripts in which junctin exon 4a was fused to BAH/humbug exon 4. These changes do not alter the reading frame of downstream BAH sequences (Table I). In summary, the genomic loci of junctin (<25 kb, 5 exon gene), humbug (>60 kb, 14 exon gene), and BAH (200 kb, 24 exon gene) overlap and share exons. Further, in heart tissue BAH/humbug utilizes junctin exon 1 (and sometimes exon 4) without altering its reading frame (accession nos. AF289491 and AF289492), while junctin shares exons 2 and 3 with BAH/humbug in all transcripts studied.

A comparison of the deduced protein coding sequences of mouse, bovine, human and Drosophila BAH is presented in Fig. 2. The putative transmembrane domain shows complete identity between mouse and human and is limited to a portion of exon 2 in the mouse. One of the major differences noted between the bovine and human sequence in the 5’ end was the absence of a string of 15 amino acids in the human sequence (14). This additional sequence (Fig. 2), derived from a bovine brain cDNA, is inserted at the junction of exons 2 and 3, suggesting that an additional exon and splicing event accounts for this sequence. A search of the current EST data bases for human BAH sequences revealed one human EST clone (accession no. F07451) isolated from infant brain that contains this apparently missing exon and thus increases the overall homology between the orthologues. Because both clones that contain this additional sequence were derived from brain RNA, it is possible that this additional sequence represents a brain-specific alternatively spliced form. Other regions where evidence of alternatively spliced forms exists occurs at the ends of exons 10 and 13 where differences exist among several species (Fig. 2).

Overall, the highest degree of homology shared between the various forms of BAH are sequences that encompass exons 2–3 and 14–24 of mouse BAH, areas that encode the transmembrane and catalytic regions, respectively. Specifically, the catalytic domain of mouse BAH shares 44% identity with Drosophila BAH and greater than 95% identity with human or bovine BAH. The critical dibasic Gly and His2 motifs of BAH are encoded on exons 22 and 23, respectively.

Drosophila BAH Genomic Structure—Comparison of the Drosophila BAH cDNA (accession no. AF289493) to sequences from genomic P1 clone DS03910 (accession no. AC004248) indicates that the gene contains 11 exons spanning 6.5 kb (Fig. 4C). The Drosophila humbug cDNA (accession no. AF289494) results from alternative splicing into exon 8, generating a protein that terminates 4 amino acids after the alternative splice (Fig. 2).

Expression Profiling of BAH, humbug, and junctin by Real Time PCR—To determine the tissue distribution of BAH/humbug/junctin transcription, quantitative real time RT-PCR profiling was performed across a panel of 14 murine tissues. Probe sets were generated to permit the specific detection of each of the three transcripts. These probe sets were derived from 1) the

---

**Fig. 4. Genomic organization of mouse BAH, mouse junctin and Drosophila BAH.** A, mouse BAH, humbug, and junctin cDNA structure in the context of the BAH genomic locus. B, splicing patterns of BAH, humbug, and junctin. Exons are shown as boxes and introns as lines. Colored boxes denote protein coding regions while untranslated regions are shown as gray boxes. The splicing pattern that gives rise to the full-length and truncated forms are indicated. cDNAs 1, 2, and 7 represent the three major versions of BAH/humbug cloned from liver tissue while 3–8 and 8–9 represent cDNAs present in heart tissue. cDNA 10 depicts mouse junctin. The schematics are color-coded as in Fig. 1. C, Drosophila BAH/humbug genomic structure. Exons are shown as boxes and introns as lines. White boxes denote protein coding regions, and untranslated regions are shown as gray boxes. The splicing pattern that gives rise to the full-length and truncated forms are indicated.

---

Asph Isoforms Share Exons with Junctin
The exon/intron boundary structure of mouse BAH, humbug, and junctin

| Exon no. | Intron Exon  | cDNA Exon | bp | bp | kb |
|----------|--------------|-----------|----|----|----|
| 1        | gcagcctaagga | GTCGCAGGCTGGGC. . .GCT CGG AGA Ggtcagagttcccg | 1–304 | 304 | >2.3 |
| 2        | cttccacacag  | AG GCA AAG CAG. . .GAA GTT CTA G gtaagaactctggg | (1–140) | 140 | 4.4 |
| 3        | ctctgtgtgta | GA AAA CTA GGA. . .GTT TTA TTA G gtaaagtactctttc | 305–454 | 455–523 | >2.8 |
| 4        | ataaatagcag | GC CTT AAA GAA. . .GAG GGG GCA gtaaggtctactg | 524–613 | 524–613 | >2.2 |
| 5        | aataacacacag | AA GGA CCT GGT. . .AAG GCT AAA G gtatattttctg | (360–401) | 42 | 2.9 |
| 6        | taattatatttgaag | AC ATC CAG AAT. . .ACA GAC CAT Ggtatgaattcagt | 614–688 | 614–688 | 2.5 |
| 7        | tctctgttaag | AC CTT GAA GCA. . .GTT CAT GAA G gtagaatacctt | 689–802 | 689–802 | 2.2 |
| 8        | tttttctctag | AA ATC GAG CAT. . .GAA GAC ACA G tgtagatctttat | 803–835 | 803–835 | 2.0 |
| 9        | cttctcatttag | CA TCG CAG AAC. . .GAA AAT TCA G tatttttataaat | 836–892 | 836–892 | 1.1 |
| 10       | tttttctcttcag | AG GAA GTA AGA. . .GAT GAA CCA G gtttgaatatatt | 941–973 | 941–973 | 1.5 |
| 11       | tttctctctatag | TA TAT GAA CCT. . .GAG ATT TCA G gtaggtttttctctt | 974–1012 | 974–1012 | 0.68 |
| 12       | ttctttctctag | AT AAC ACC ATA. . .ATC TCA GAA G gtaagttttttttt | 1013–1048 | 1013–1048 | >3.5 |
| 13       | gttaatttttag | AT ATA AAT GTC. . .ACA CCA CCA G gtatgcattttc | 1049–1093 | 1049–1093 | 4.5 |
| 14A      | ttatttttttag | AT ACT TAA. . .TTAAGCGAATAA | 1094A–2789A | 1094A–2789A | 1696 |
| 15       | cttctctctcag | TT AAG AAG AAG. . .CGG AAA AGG gtaggtgtgtatt | 1094–1179 | 1094–1179 | 4.5 |
| 16       | ggcacctctcag | TTG GAA GAT GAC. . .CAG TTT CTA G gtagagatcagtgc | 1180–1266 | 1180–1266 | 5.8 |
| 17       | ttccacacctag | GTT CCA ACT GGT. . .GCA TAC GAA G gtaagcttcctgc | 1267–1417 | 1267–1417 | 1.8 |
| 18       | ctttgatcctag | GTG CTA ATC GTG. . .TAC TTA AAG gtagctcagcttc | 1418–1554 | 1418–1554 | 2.0 |
| 19       | gtcttttctacag | GAA GGA ATC GAA. . .AAC AAA GAG gtagaatattaact | 1555–1653 | 1555–1653 | 3.0 |
| 20       | cttctctctcag | GGA GAT GAC. . .CGG AAG CAG gtagaagctttg | 1654–1743 | 1654–1743 | 6.0 |
| 21       | ttctttttcag | GA TAT AAG TGG. . .CTA GTG AAG gtagtggtacttc | 1744–1881 | 1744–1881 | >3.9 |
| 22       | ctttctctctcag | GG AAG AAG GAC. . .CCA CAA GAG gtaaagctgtctgc | 1882–2017 | 1882–2017 | >4.0 |
| 23       | ctttctctctcag | ATC AAA TAC TCC. . .ATG CAG ACC AG gtaggttttttc | 2018–2109 | 2018–2109 | >3.5 |
| 24       | ctttctctctcag | G ACG TGG GAA. . .AGCAATAAA . . .CCAAATAA | 2110–2243 | 2110–2243 | 4386 |

* Introns are given in lowercase letters; exons are listed in uppercase.
Asph Isoforms Share Exons with Junctin

DISCUSSION

Three major BAH-related transcripts have been identified by Northern analysis on RNA from multiple mammalian species (2.8, 4.5, and 6.6 kb in the mouse). cDNA analysis of these transcripts from mouse liver was used to define the structure of these transcripts. The two largest transcripts, 4,419 and 6,629 nucleotides, encode full-length BAH and have identical coding regions, which include cytoplasmic, transmembrane, highly charged luminal region, and COOH-terminal catalytic domain. These transcripts differ only in the 3′-UTR, where the longer terminal to the start of the proteolytically processed 56- and 52-kDa forms of BAH isolated from bovine liver. This suggests that FB-50 should not bind to these proteolytically generated truncated catalytic forms.

Epitope Mapping mAbs HBOH-1 and HBOH-2—Phage displayed peptide libraries were screened to identify the BAH epitopes recognized by mAbs HBOH-1 and HBOH-2. During three rounds of selection for each antibody, the percentage of bound phage increased more than 3 orders of magnitude indicating enrichment of specific peptide binding phage. Sequencing of the random peptide inserts of selected phage indicated consensus sequences that correlate well with residues 573–579 (QPWWTPK) for HBOH-1 and a peptide comprising residues 613–620 (LPEDENLR) of human BAH for HBOH-2 (Figs. 2 and 6A). These antibodies are specific for the full-length, catalytically active form of BAH and would not be expected to recognize humbug or junctin.

Antisense Oligonucleotide Studies on A549 Cells Correlate Northern and Western Data for BAH and humbug—In the mouse, three transcripts corresponding to two deduced protein isoforms (BAH and humbug) have been identified. To determine whether humbug is expressed as a protein and to correlate the transcripts to protein species, Northern and Western analyses were performed on RNA and protein prepared from cultured human A549 cells. In contrast to the mRNA sizes observed in mouse Northern (2.8, 4.5, and 6.6 kb, Fig. 1), the three BAH/humbug RNA transcripts in human A549 cells migrate at approximately 2.8, 4.5, and 5.2 kb (Fig. 6E). When FB-50 antibody (should recognize both BAH and humbug) is used to probe protein extracts on a Western gel from A549 cells, one band at 60 kDa and a doublet at 120–130 kDa are identified (Fig. 6B). In contrast, monoclonal antibodies HBOH-1 and HBOH-2, which were epitope-mapped to the COOH-terminal region of human BAH, recognize only the upper (120–130 kDa) bands observed on Western blots of A549 cells. These results demonstrate that the alternative splice into exon 14a identified in mouse and human humbug results in production of a humbug protein that migrates at 60 kDa in size while the full-length version of BAH is represented by a doublet in the 120–130-kDa range. A549 cells were treated with antisense oligos complementary to either the NH₂-terminal (oligo 583) or COOH-terminal regions (oligo 1302), and cell extracts were subjected to Western analysis. Oligo 583 treatment, which is complementary to BAH and humbug, reduced all three transcripts observed in Northern blots of A549 cells (Fig. 6E) and reduces both the 60- and 120–130-kDa bands by Western analysis (Fig. 6C). In contrast, oligo 1302, which is complementary only to the BAH transcript, resulted in a reduction of only the upper two transcripts on Northern blots of A549 cells (Fig. 6E) and the 120–130-kDa protein doublet on Western analysis in A549 lysates (Fig. 6D). These results are consistent with the conclusion that the 2.8-kb transcript encodes the humbug protein which migrates at 60 kDa in Western blots of A549 cells, whereas the larger (4.5 and 5.2 kb) human BAH transcripts result in a protein with a relative molecular mass of approximately 120–130 kDa.

Epitope Mapping Monoclonal Antibody FB-50—A set of noncleavable 15 residue peptides attached to pins were screened to identify the region of human BAH recognized by mAb FB-50 (16). Strongly bound antibody was indicated for four of the immobilized peptides with signal ranging from 1.8 to 3.7 optical density at 405 nm. The peptides that bound FB-50, peptides 69–72, comprise residues 277–300 of BAH. The common core peptide found in each of the 15-mers and, therefore, the epitope for FB-50 includes residues 286–291 (NPVEDS) (Figs. 2 and 6A). Based on the position of this epitope, FB-50 mAb would be expected to bind to both full-length BAH and humbug, but not junctin. Further, the epitope recognized by FB-50 is amino-
transcript uses a downstream poly(A) addition signal. cDNA clones (2,789 nucleotides) representing the shortest transcript revealed the existence of a truncated form that is identical to the BAH transcript from the NH2 terminus through the highly charged lumenal region. This “humbug” transcript lacks the defining 56-kDa catalytic region of BAH. humbug is conserved in other species. For example, humans express humbug RNA in a wide range of tissues. Further, in studies presented above, a humbug ortholog exists in Drosophila. Conservation of this truncated transcript suggests a biologically significant function.

Using a combination of antisense and Western analyses with epitope mapped monoclonal antibodies, we have shown that the full-length human BAH protein runs as a doublet with an apparent mass of 120–130 kDa. In addition, this approach demonstrated that human humbug, the truncated transcript lacking the catalytic domain, is translated into protein with an apparent mass of 60 kDa. It is interesting to note that the masses of BAH and humbug protein, predicted from translation of the cDNA sequences, are 85 and 35 kDa, respectively. This differs significantly from the sizes seen in Western blots. The difference between the predicted mass and that seen on gels may be due to the highly charged nature of these two proteins or glycosylation.

Previous Northern analysis (16) has been carried out to examine the expression of BAH in human tumor cell lines. It is now clear that the probe utilized in those experiments will hybridize with both BAH and humbug. Those studies revealed the presence of a major 2.8 kb transcript and a minor transcript of ~5.0 kb in a wide range of human tumor cell lines, but not normal human liver. Our results imply that the major transcript observed by Lavaissier et al. (16) corresponds to humbug and the minor transcript to BAH. Because the two genes share significant homology, use of selective and specific reagents will be required to identify the effects of BAH and humbug and their relationship to malignancy. For example, a recent report suggests that overexpression of BAH may be associated with malignant transformation (26).
that study, immunoreactivity with the FB-50 antibody in human cholangiocarcinoma was convincingly demonstrated. Based on the results presented above, which demonstrate that FB-50 recognizes both BAH and humbug protein, this reactivity could be caused by binding to BAH, humbug or both.

Jones et al. (17) have previously shown that canine junctin shares significant homology with bovine BAH over a small region. To determine if the homology of junctin with BAH/humbug is derived from two different genes or whether they share common exons, a mouse junctin cDNA was isolated and sequenced. A comparison of the mouse junctin sequence to mouse BAH/humbug demonstrated that the two sequences were identical for 220 nucleotides. This sequence was internal to the overall coding sequence of junctin and included the putative transmembrane domain. These results strongly suggest that junctin and BAH/humbug are encoded using common exons.

Clarification of the relationship of BAH to humbug and junctin was sought by the cloning and isolation of the BAH genomic locus from mouse. These studies demonstrated that BAH is encoded by 24 exons spanning over 200 kb of DNA. The humbug sequence was shown to be derived from the first 13 exons of the BAH locus with the break in homology arising precisely at the end of exon 13. Searching the DNA residing between BAH exon 13 and 14 revealed another exon, 14a, which completed the coding region of humbug and provided a stop codon and poly(A) addition signal. A blast search of GenBank with the 3'-UTR of human humbug revealed one sequence-tagged site (STS WT-11767) that confirms that humbug maps to the same chromosomal position as BAH (8q12) (16).

A comparison of mouse junctin sequence to BAH exon sequences revealed that the 220 nucleotides of identity between junctin and BAH were defined precisely by the coding boundaries of BAH exons 2 and 3. Characterization of the DNA between BAH exon 1 and 2 led to the identification of junctin exon 1a, which encoded the region of junctin upstream of the BAH exon 2 homology. Examination of the DNA downstream of BAH exon 3 permitted the localization of two additional exons, 4a and 5a, which completed the exon structure of junctin. Based on these studies, it is clear that the genomic coding DNA for junctin is contained in the BAH locus and, when expressed, junctin shares exons 2 and 3 with BAH and humbug.

The close coding relationship between BAH, humbug, and junctin led us to study the expression of these three transcripts in RNA derived from multiple tissues. These studies showed that both BAH and humbug were widely expressed with the highest levels occurring in heart, ovary, and adrenal gland. The distribution of junctin expression was more limited, with the highest levels of expression occurring in heart and skeletal muscle and lower levels in ovary and adrenal. The overlap of high relative levels of expression of junctin, BAH, and humbug in heart led to the examination of the structure of BAH/humbug transcripts in RNA derived from this tissue. Of 12 cDNA clones examined, all of the BAH/humbug transcripts in the heart began with exon 1a of junctin, not exon 1 of BAH/humbug. This suggests that the promoter associated with exon 1a provides strong tissue-specific transcription, which produces expression of BAH/humbug in this tissue as well as junctin. Because the coding region of exon 1a of junctin is smaller than exon 1 of BAH/humbug, the cardiac forms of BAH/humbug should be at least 3.2 kDa smaller and perhaps smaller still since exon 1 of BAH/humbug contains a putative N-linked glycosylation site (Fig. 2). Analysis of the downstream sequences for these cDNAs revealed that, although most of the BAH/humbug clones examined spliced from exon 3 to exon 4, a significant number of clones spliced exon 4a of junctin to exon 4 of BAH/humbug. These data demonstrate that in heart, BAH/humbug shares more similarity to junctin, which not only includes exon 2 and 3, but also frequently includes exon 1a and 4a of junctin. Thus, in heart, BAH/humbug is expressed as a transcript that includes 3 or 4 of the 5 exons of junctin. This leads to the hypothesis that transcription of BAH, humbug, and junctin may initiate with exon 1a in heart, and perhaps skeletal muscle, and exon 1 in most other tissues. According to this proposal, the level of expression of BAH, humbug, and junctin in a particular tissue would be controlled by both tissue-specific transcription initiation and tissue-specific RNA splicing. This idea is supported by the observation that, although junctin, BAH, and humbug are highly expressed in heart using exon 1a, only junctin appears to be highly expressed in skeletal muscle. This raises the possibility that, in skeletal muscle, both transcription initiation control and tissue-specific splicing of junctin, BAH, and humbug may be differentially regulated. Because junctin RNA is easily detected in ovary and adrenal gland, tissues where BAH and humbug are relatively highly expressed, junctin RNA can be studied in these tissues to determine if junctin downstream sequences are fused to BAH/humbug exon 1. Detailed analysis of junctin, BAH, and humbug transcription initiation and splicing in other tissues is being investigated.

The fusion of junctin sequences to BAH/humbug in murine heart leads to a discussion of their functional role and relationship in this tissue. Junctin is known to be an important part of a complex of proteins in heart junctional sarcoplasmic reticulum that includes the ryanodine receptor, triadin, and calsequestrin. These proteins act coordinately to release Ca$^{2+}$ from intraluminal stores. Interestingly, examination of triadin, calsequestrin, and junctin reveals the presence of intraluminal domains that contain a large number of charged amino acids (rat triadin: 47% charged, 27% positive, 20% negative; mouse junctin: 49% charged, 29% positive, 20% negative). It has been suggested that the intraluminal charged regions of triadin and junctin are essential for forming this protein complex (27). The repeating KEKE motifs of the charged regions of calsequestrin appear critical for protein-protein interactions and may also play a role in binding of Ca$^{2+}$ (mouse calsequestrin: 38% charged, 11% positive, 27% negative). Interestingly, BAH/humbug also contain a region of highly charged amino acids on the putative luminal side of the membrane, but this region is not derived from the junctin coding exons. Instead it is encoded in exons 4–13 of BAH/humbug and is significantly more negatively charged (40% charged, 8% positive, 32% negative) than junctin. Exon sharing with junctin and the presence of a BAH/humbug encoded highly charged domain proximal to the membrane on the luminal side of the junctional sarcoplasmic reticulum leads to the proposal that, in heart, BAH/humbug may be associated with the ryanodine receptor complex in heart. Junctin overexpression experiments done on canine cardiac junctional sarcoplasmic reticulum by Zhang et al. (27) revealed the presence of calsequestrin and triadin. In addition to these proteins there are bands in this overlay with sizes similar to humbug and BAH. Zhang et al. (27) also used calsequestrin affinity chromatography to examine this protein complex, and junctin and triadin were co-purified by this method. On this stained gel, proteins consistent with the sizes of humbug and BAH were identified. These results suggest that it is worth directly testing whether BAH and humbug are associated with this complex.

Alternative splicing produces humbug, which lacks the 56-kDa catalytic domain of BAH. There have been a number of published accounts of protein isoforms that result from alternative splicing interfering with the catalytic activity of their...
catapultly active partners either by competing for substrate molecules or by direct interaction/inhibition of the full-length isoforms (28–31). It is possible that humbug acts as a modifider of BAH function by altering substrate-enzyme interactions, directly inhibiting enzyme activity or competing with BAH for assembly into larger protein complexes such as those seen associated with the ryanodine receptor.

The widespread expression of BAH and humbug in normal mouse tissue and human tumors and cancer cell lines, coupled with conservation of these genes back to Drosophila, suggests that BAH and humbug are likely to have an important functional role in vivo. To approach the possible role of BAH and humbug in normal and disease processes in the mouse, the mouse BAH gene has been carefully characterized at the cDNA and genomic levels. This information is being used to develop transgenic and knockout models of these genes with the aim of carefully dissecting what is likely to be a highly complex process of protein modification, assembly, and signaling in vivo.

Acknowledgment—We thank Dr. Nicki Neff for thoughtful discussions and careful review of the manuscript.

REFERENCES

1. Rees, D. J. G., Jones, I. M., Handford, P. A., Walter, S. J., Esmouf, M. P., Smith, K. J., and Brownlee, G. G. (1988) EMBO J. 7, 2053–2061
2. Sunnerhagen, M. S., Person, E., Dahlyvist, J., Drakenberg, T., Stenflo, J., Mayhew, M., Robin, M. Handford, P., Tilley, J. W., Campbell, I. D., and Brownlee, G. G. (1993) J. Biol. Chem. 268, 23339–23344
3. Campbell, I. D., and Bork, P. (1993) Curr. Opin. Struct. Biol. 3, 385–392
4. Davis, C. G. (1990) New Biol. 2, 410–419
5. Stenflo, J. (1991) Blood 78, 1637–1651
6. Becay, P. F., Fleming, R. J., Fehon, R. G., Cherbas, L., Cherbas, P., and Artavanis-Tsakonas, S. (1991) Cell 67, 687–699
7. Sun, B. S., Zu, X. Clayton, M. M., Pan, J., and Fettelson, M. A. (1998) Hepatology 27, 228–239
8. Nakamura, T., Ruiz-Lezano, P., Lindner, V., Yabe, D., Taniwaki, M., Furukawa, Y., Kobuke, K., Tashiro, K., Lu, Z., Andon, N., Schaub, R., Matsumura, A., Sasayama, S., Chien, K. R., and Honjo, T. (1999) J. Biol. Chem. 274, 22443–22450
9. Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D., and Handford, P. A. (1996) Cell 85, 597–605
10. Goruppi, S., Yamane, H., Marcaudali, F., Garcia, A., Clogston, C., Gorissatza, M., Varnum, B., and Schneider, C. (1997) FEBS Lett. 415, 59–63
11. Wang, Q., VanDusen, W. J., Petroksi, C. J., Garsky, V. M., Stern, A. M., and Friedman, P. A. (1991) J. Biol. Chem. 266, 14004–14010
12. Jia, S., VanDusen, W. J., Diehl, R. E., Kohl, N. E., Dixon, R. A. F., Elliston, K. O., Stern, A. M., and Friedman, P. A. (1992) J. Biol. Chem. 267, 14322–14327
13. Jia, S., McGinnis, K., VanDusen, W. J., Burke, C. J., Kuo, A., Griffin, P. R., Sardana, M. K., Elliston, K. O., Stern, A. M., and Friedman, P. A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 7227–7231
14. Koriith, F., Gieffers, C., and Frey, J. (1994) Gene (Amst.) 150, 385–399
15. McGinnis, K., Ku, G. M., Van Dusen, W. J., Fu, J., Garsky, V. M., Stern, A. M., and Friedman, P. A. (1996) Biochemistry 35, 3957–3962
16. Lavaissier, L., Jia, S., Nishiyama, M., de la Monte, S., Stern, A. M., Wands, J. R., and Friedman, P. A. (1996) J. Clin. Invest. 98, 1313–1323
17. Jones, L. R., Zhang, L., Sanborn, K., Jorgensen, A. O., and Kelley, J. (1995) J. Biol. Chem. 270, 30787–30796
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
19. Povinelli, C. M., and Gibbs, R. A. (1993) Anal. Biochem. 210, 16–26
20. Gibbon, U. E., Heid, C. A., and Williams, P. M. (1996) Genome Res. 6, 995–1001
21. Abbassade, I., Liu, R.-Q., Yang, F., Rosenfeld, S. A., Ross, O. H., Link, J. R., Ellis, D. M., Tortorella, M. D., Pratta, M. A., Hollis, J. M., Wynn, R., Duke, J. L., George, H. J., Hillman, M. C., Murphy, K., Wiswall, B. H., Copeland, R. A., DeCicho, C. P., Bruckner, R., Nagase, H., Itoh, Y., Newton, R. C., Magolda, R. L., Trzaskos, J. M., Hollis, G. F., Arner, E. C., and Burn, T. C. (1999) J. Biol. Chem. 274, 23443–23450
22. Harlow, E., and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
23. O’Neill, K. T., Hoess, R. H., Jackson, S. A., Ramachandran, N. S., Mouse, S. A., and DeGrado, W. F. (1992) Proteins Struct. Funct. Genet. 14, 509–515
24. Kiss, J. P., Britton, D. H., Stone, B. A., Behren, D. A., Leffet, L. M., Hobbs, F. W., Miller, J. A., and Trainer, G. L. (1996) Nucleic Acids Res. 24, 1901–1907
25. Monkovic, D. D., VanDusen, W. J., Petroksi, C. J., Garsky, V. M., Sardana, M. K., Zavadovsky, P., Stern, A. M., and Friedman, P. A. (1992) Biochem. Biophys. Res. Commun. 189, 233–241
26. Ince, N., de la Monte, S. M., and Wands, J. R. (2000) Cancer Res. 60, 1261–1266
27. Zhang, L., Kelley, J., Schmeisser, G., Kobayashi, Y. M., and Jones, L. R. (1997) J. Biol. Chem. 272, 23389–23397
28. Scott, G. K., Robles, R., Park, J. W., Montgomery, P. A., Daniel, J., Holmes, W. E., Lee, J., Keller, G. A., Li, W. L., Feeny, B. M., Wood, W. I., Shepard, H. M., and Ben, C. C (1995) Mol. Cell. Biol. 15, 2247–2257
29. Nagano, K., Fukami, K., Minagawa, T., Watanabe, Y., Ozaki, C., and Takenawa, T. (1999) J. Biol. Chem. 274, 2872–2879
30. Ross, R. J. M., Kostoplos, M., Shen, X. Y., Van Laue, S., Chew, S. L., Dobson, P. R. M., Pastel-Vinay, M.-C., and Finidori, J. (1997) Mol. Endocrinol. 11, 265–273
31. Doherty, J. K., Bond, C., Jardim, A., Adelman J. P., and Clinton, G. M. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 10869–10874