*Nucleic Acids Research, 1994, Vol. 22, No. 5* ***749–755***

# The expression pattern of *Id4*, a novel dominant negative helix-loop-helix protein, is distinct from *Id1*, *Id2* and *Id3*

Veit Riechmann, Ingeborg van Crüchten and Fred Sablitzky*
Max-Delbrück-Laboratorium in der Max-Planck-Gesellschaft Carl-von-Linné-Weg 10, 50829 Köln, Germany

Received December 20, 1993; Revised and Accepted February 3, 1994 EMBL accession no. X75018

## ABSTRACT

**Molecular interaction between transcription factors containing an basic-helix-loop-helix (bHLH) domain is known to regulate differentiation in several cellular systems including myogenesis, neurogenesis and haematopoiesis. DNA-binding activity of the bHLH proteins is mediated via the basic region and is dependent upon formation of homo- and/or heterodimers of these transcription factors. Dominant negative (dn) HLH proteins (*Id1, Id2, Id3* and *emc)* also contain the HLH-dimerization domain but lack the DNA-binding basic region. Formation of heterodimers between dnHLH and bHLH proteins abolishes the DNA-binding activity of the latter. Concordantly, it was shown that the dnHLH protein *Id1* inhibits differentiation of muscle and myeloid cells *in vitro*. Therefore, it was postulated that dnHLH proteins serve as general antagonists of cell differentiation. We have isolated and characterized a novel mouse dnHLH gene, designated *Id4*. The *Id4* protein contains a HLH domain highly conserved among the dnHLH proteins from mouse and drosophila. Outside of the HLH domain, three additional short regions of the dnHLH proteins show some degree of homology. DNA-binding of *E47* homo- as well as *E47/MyoD* heterodimers is inhibited by *Id4*. Transcription of the *Id4* gene results in three RNA molecules of 3.7, 2.0 and 1.7 kb which are presumably a result of differential splicing and/or alternatively used polyadenylation sites within the 3′ untranslated region. During embryogenesis, *Id4* expression is up-regulated between day 9.5 and 13.5 of gestation. The highest expression in adult tissues was detected in testis, brain and kidney. Comparison of the expression patterns of the four mouse dnHLH genes revealed that *Id4* expression differs from the more restricted expression of *Id2* as well as from the widespread expression of *Id1* and *Id3*.**

## INTRODUCTION

The helix-loop-helix (HLH) domain, was initially identified as a structural motif by comparing the amino acid sequences of the drosophila neurogenic proteins encoded by the *achaete–scute* (*ac–sc*) complex (1), *daughterless* (2) and the mammalian *c-myc* (3) and *MyoD* (4, 5) proteins. Murre *et al.* (6) proposed that amino acids conserved within the region of homology between the various proteins adopted a common secondary structure: two amphipathic $\alpha$-helices separated by an intervening loop. They and others demonstrated that the HLH domain mediates protein dimerization (6) and that dimeric proteins bind DNA via a group of positively charged amino acids [the basic (b) region] directly adjacent to the HLH domain (7, 8). The DNA sequences that these bHLH proteins bind are the so called E-boxes which are represented by the core sequence *CANNTG* (9, 10). bHLH proteins are transcription factors which have been implicated in oncogenesis and/or cell determination and differentiation in several cellular systems including myogenesis [*MyoD* (3), *myf-5* (11), *MRF4* (13–14) and *myogenin* (15–17)], neurogenesis [*ac–sc* (1), *MASH1* (18)] and haematopoiesis [*SCL/TCL-5/tal-1* (19–23) and *lyl-1* (24–26)]. It was shown (4, 25) that these cell-type-restricted bHLH proteins form heterodimers with the ubiquitously expressed bHLH proteins which are encoded by the *E2A* gene [*E12*, *E47* (6) and *E2-5/ITF1* (26)] and the *E2-2* gene [*ITF2* (26)] in mouse and by the *daughterless* gene (2) in drosophila. Furthermore, formation of heterodimers is essential for DNA-binding *in vivo* (27). Transcriptional activity of E-box containing promoters and/or enhancers might therefore exclusively be dependent upon the concentration of different bHLH proteins within the cells.

However, subgroups of HLH proteins were described which antagonize the function of the above bHLH proteins. The drosophila *hairy* (28) and *Enhancer of split* (29) gene products and the rat protein, *HES-1* (30), are distinct from the bHLH proteins in that their basic regions contain a proline residue altering their DNA-binding capacity. Proteins of this subgroup do not bind the E-box efficiently but they do bind the so called N-box motif [*CACNAG* (29–31)]. Another subgroup of HLH proteins was described [*emc* (32, 33), *Id1* (34), *HLH 462* (35), and *Id2* (36)] which lack the basic domain. These proteins still dimerize but they cannot bind DNA. Moreover, dimerization with bHLH proteins block DNA-binding of the resulting heterodimers (32–34). Therefore, these proteins are referred to as dominant negative (dn) HLH proteins (37). It was shown in drosophila for example, that *emc* represses proneural gene expression of the *ac–sc* complex by sequestering the latter proteins in complexes

*To whom correspondence should be addressed

unable to interact with DNA (38, 39). In controlling the spatial pattern of *ac−sc* transcription, *emc* is involved in the regulation of pattern formation of functional proneural clusters and the subsequent bristle formation in the wing imaginal disc (38). For the mammalian dnHLH gene products, *Id1* and/or *Id2*, it was demonstrated that their expression is inversely correlated with cell differentiation in myogenesis (42) and neurogenesis (58) as well as myelopoiesis (43) and lymphopoiesis (36, 40, 44). Concordantly, overexpression of the *Id1* protein inhibits differentiation of muscle cells (42) as well as myeloid cells (43). In addition, ectopic expression of the *Id1* protein in B cells represses the activity of the immunoglobulin intron enhancers of heavy and κ-light chains (40) as well as the 3′ κ enhancer (44). Hence, a model was proposed in which the *Id1* protein serves as a general antagonist of cell differentiation by inhibiting bHLH proteins specifically required for developmental programs (34, 37, 45). Although not yet shown directly, it is likely that in analogy to *Id1*, mouse (36) and human (46) *Id2* as well as mouse [HLH462 (35)] and human [HLH1R21/Heir-1 (47, 48)] *Id3* are important regulators of differentiation in various cellular systems.

Here we report the isolation, functional characterization and expression pattern of a novel mouse dnHLH protein, designated *Id4*. The *Id4* protein contains a HLH motif which is highly conserved among the dnHLH proteins. Three additional regions of the *Id4* protein are also similar to the other members of the dnHLH protein family, but conservation is less pronounced. *Id4*, which lacks a basic region, forms heterodimers with *E47* and inhibits DNA-binding of *E47* homodimers as well as *E47/MyoD* heterodimers *in vitro*. Expression of *Id4* is up-regulated during embryonic development and is highest in testis, brain and kidney of adult mice. This expression pattern of *Id4* differs from the more restricted expression of *Id2* as well as from the widespread expression of *Id1* and *Id3*.

## MATERIALS AND METHODS

### Isolation and sequence analysis of *Id4* cDNA clones

A degenerate oligonucleotide (5′CTGMAGGTCSMIG-ATGTARTCKATKACRYGCTGSAKGATYTCC3′) corresponding to the reverse complementary sequence of the second helix of *Id1* (34), human and mouse *Id2* (36, 46), HLH462 (35) and *emc* (32, 33) was used to screen a λgt11 cDNA library made from mouse bone marrow cells (Clontech). The filters were hybridized for 43 hr at 44°C with the 43-mer, which had been radiolabeled by T4 polynucleotide kinase (Boehringer) in the presence of $\gamma$-$^{32}$P-ATP (Amersham) in 1% bovine serum albumin/1 mM EDTA/0.5 M $NaHPO_4$ (pH 7.2)/7 % SDS. Washing was performed two times with 1 mM EDTA/40 mM $NaHPO_4$ (pH 7.2)/5 % SDS and eight times with 1 mM EDTA/40 mM $NaHPO_4$ (pH 7.2)/1 % SDS at RT for 5 min each (49). Two cDNA clones (*VR4, VR18*) were subcloned in pGEM 7 (Promega) and sequenced (sequenase kit; United States Biochemical). In order to obtain full length cDNA clones, *VR4* was used to screen a λgt10 cDNA library from RNA of 12.5 days old mouse embryos (generously supplied by Drs M.Hanks and A.Joyner, Toronto). Three overlapping clones (*B*, *F* and *I*; see Fig. 1) were isolated, subcloned in pGEM 7 and sequenced. Sequence analysis of GC rich regions of the *Id4* cDNA were performed in parallel reactions using dITP instead of dGTP.

### *In vitro* transcription and translation

*In vitro* transcription of *Id4* (clone *I* linearized with XbaI which cuts in the 3′ untranslated region), *MyoD* [plasmid pEMCIIs (4) generously provided by Dr H.Weintraub, Seattle] and *E47* [plasmid pBS-ATG-E47S (6) generously provided by Dr C.Murre, La Jolla] was performed under standard conditions using 100 u of SP6 or T3 polymerases (Boehringer) in a total volume of 100 μl. RNA was purified, ethanol precipitated and stored at −70 °C in diethylpyrocarbonate treated destilled water. *In vitro* translation of the RNAs was performed in 50 μl rabbit reticulocyte lysate (Promega) using 5−45 % of the synthesized RNA as a substrate. To generate radioactively labeled proteins, parallel reactions were performed, using [$^{35}$S] methionine and the resulting proteins were analyzed on a 12 % discontinuous SDS-PAGE gel (data not shown).

### Electrophoretic mobility shift assay

A 25 bp double stranded oligonucleotide (34) containing the *E47* and/or *E47/MyoD* binding site from the *MCK* enhancer (5) was labeled as described (34). 20 μl of *in vitro* translation reaction was heated to 37 °C for 20 min. DNA-binding reactions with 20 μl DNA-binding cocktail including 0.1 ng double stranded oligonucleotide probe and analysis on a 5 % PAGE were performed as described (34).

### Expression of *Id4*

polyA$^+$ RNA was isolated from various mouse tissues and cell lines using the Fast Track kit (Invitrogen). Northern blot analysis were done under standard conditions (50). Probe *a* corresponds to *VR4* and probe *b* corresponds to *F2* (Fig. 1).

First strand cDNA was synthesized using oligo(dT)$_{18}$ and M-MLV Reverse Transcriptase for 60 min at 37 °C in a total volume of 50 μl (according to the protocol of GIBCO BRL).

First strand cDNA was amplified by PCR (51) using the primers depicted below:

*Id1*: sense: 5′-GGTGGATCCACCATGAAGGTCGCCAGTG-3′
antisense: 5′-GGTGGATCCGTCCATCTGGTCCCTCAGTGC-3′
nucleotide position 84−560 (34) resulting in the amplification of a 476 bp cDNA fragment

*Id2*: sense: 5′-GGTGGATCCACCATGGCAATTCAGGGATGC-3′
antisense: 5′-GGCGGATCCTTATTTAGCCACAGAGTAC-3′
nucleotide position 39−603 (36) resulting in the amplification of a 564 bp cDNA fragment

*Id3*: sense: 5′-AAGGCGCTGAGCCCGGTGC-3′
antisense: 5′-TCGGGAGGTGCCAGGACG-3′
nucleotide position 61−441 (35) resulting in the amplification of a 380 bp cDNA fragment

*Id4*: sense: 5′-GCGATATGAACGACTGCTAC-3′
antisense: 5′-TCACCCTGCTTGTTCACGGC-3′
nucleotide position 268−538 of *Id4* sequence resulting in the amplification of a 270 bp cDNA fragment

β-actin: sense: 5′-TGGAATCCTGTGGCATCCATGAAAC-3′
antisense: 5′-TAAAACGCAGCTCAGTAACAGTCCG-3′
nucleotide position 728−1076 (52) resulting in the amplification of a 348 bp cDNA fragment

Amplification of first strand cDNAs was performed as described (53) using 1.5 mM $MgCl_2$ and 3 u of Taq polymerase (GIBCO BRL). In the cases of *Id1*, *Id3* and *Id4* 9 % DMSO was added. The amount of input cDNA was adjusted using the β-actin amplification for standardization. To assure linear

amplification the optimal number of cycles was determined for each prime pair (33 cycles *Id1*, 36 cycles *Id2*, 23 cycles *Id3*, 29 cycles *Id4* and 19 cycles β-actin). All PCR amplifications were carried out using the following conditions for denaturation (92°C for 85 s), annealing (59°C for 85 s) and polymerization (72°C for 150 s). Whenever possible (*Id1*, *Id2* and β-actin) primers were designed to amplify sequences including introns to be able to discriminate between genomic and cDNA.

Southern blots of the amplified PCR products were hybridized (50) with probes specific for *Id1*, *Id2* [subclones of amplified genomic DNA (I.van Crüchten and F.Sablitzky, unpublished)], *Id3* [a 0.8 kb EcoRI−cDNA fragment (V.Riechmann and F.Sablitzky, unpublished)] and *Id4* (probe *a* in Fig. 1).

### Nucleotide sequence accession number

X75018

## RESULTS

### Isolation and sequence analysis of Id4 cDNA

Sequence comparison of the dnHLH genes *Id1* (34), human (46) and mouse *Id2* (36), *HLH 462* [(35); hereafter called mouse *Id3*] and drosophila *extramacrochaetae* [*emc* (32, 33)] revealed a strong homology within the HLH region of these genes (see Fig. 2). Helix 2 is especially strongly conserved among dnHLH genes but distinct from all members of the DNA-binding bHLH gene family (6). In order to clone novel members of the dnHLH gene family, a degenerate oligonucleotide complementary to the helix 2 region of the dnHLH genes was designed (see Materials and Methods) and used to screen a cDNA library made from mouse bone marrow cells. Twenty out of 22 clones initially obtained were identified by hybridization as either *Id1* or *Id3*, respectively. The two remaining clones (*VR4* and *VR18)* which did not hybridize to any of the dnHLH genes turned out to be identical. Sequence analysis revealed that these two cDNA clones encode for a novel member of the dnHLH gene family (Fig. 1). To obtain full length cDNA, *VR4* was used to screen a cDNA library prepared from 12.5 days old mouse embryos. The sequence analysis of three overlapping clones (*B*, *F* and *I*) is summarized in Fig. 1. Clone *F* contained two EcoRI fragments (1.5 and 0.7 kb) the larger one hybridizing to the helix 2 specific oligonucleotide and *VR4* (data not shown). However, the 0.7 kb fragment hybridizes with two of three *Id4* mRNAs ( see below and Fig. 3) indicating that it is most likely part of the 3′ untranslated regions of two *Id4* mRNA species. The 5′ end of clone *B* started with 51 A residues. Most likely, a short, unrelated cDNA containing a polyA tail was ligated to the 5′ end of the *Id4* cDNA reflecting a cloning artifact (data not shown). Therefore the composite nucleotide sequence of *VR4*, *B*, *F1* and *I* starts at the 5′ end of clone *I*. The sequenced 1.659 kb *Id4* cDNA fragment contains an open reading frame (nucleotide position 6 to 554) with a putative AUG start codon at nucleotide position 72, predicting an *Id4* protein of 161 amino acids (data not shown).

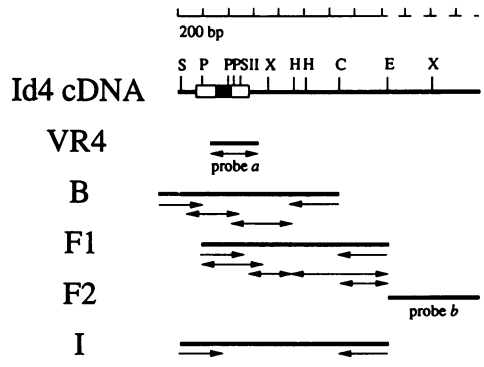


**Figure 1.** Structure of the partial *Id4* cDNA. A restriction map of the *Id4* cDNA is shown at the top (S: SacI; P: PstI; SII: SacII; X: XbaI; H: HindIII; C: ClaI; E: EcoRI). The open box represents the putative *Id4* coding region. The dnHLH region is shown as a black box. Solid lines indicate the four cDNA clones isolated. Arrows depict the sequence strategy used.

### *Id4* is a member of the dnHLH protein family

Comparison of the amino acid sequences of *Id4* with *Id1* (34), human (46) and mouse (36) *Id2*, human [*Heir-1/HLH 1R21* (47, 48)] and mouse (35) *Id3* and drosophila *emc* (32, 33) indicated that *Id4* is a new member of the dnHLH protein family (Fig. 2). The members share a highly conserved HLH motif (box 3 in Fig. 2) with the highest degree of identity in helix 2 [11/16 amino acids are identical including a valine which substitutes for an alanine residue conserved among bHLH proteins (6)]. The length of the loop (10 amino acids) and proline residues at two positions within the loop are conserved. *Id4*, like the other

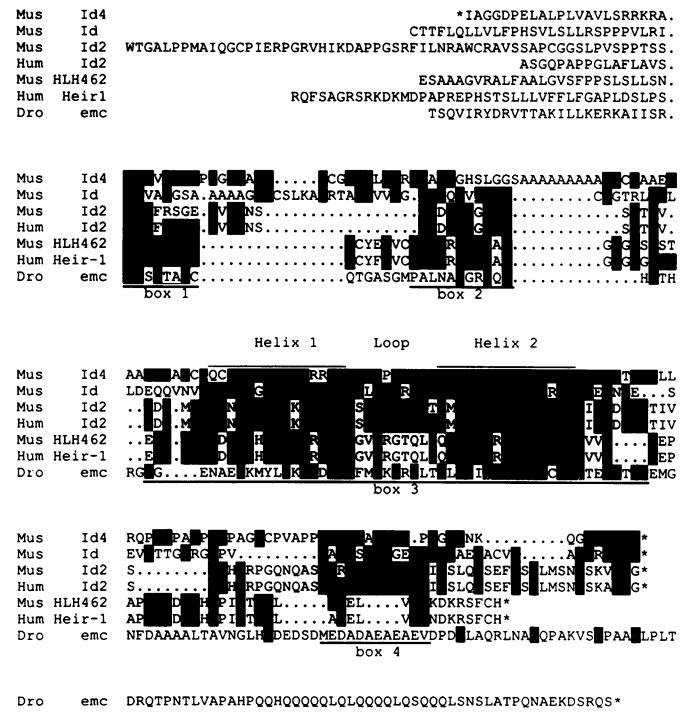


**Figure 2.** Comparison of the deduced amino acid sequences of mouse (*Id4*, *Id1*, *Id2*, *HLH462*), human (*Id2*, *Heir-1/HLH1R21*) and drosophila (*emc*) dnHLH encoding genes, respectively. Sequences were aligned to maximize homology. Identical amino acids found in at least two different *Id* genes are shown as white letters on black background. Four regions of homology are underlined and marked as box 1 to 4. Deduced amino acid sequences 5′ of the putative N-terminal ends of the *Id* proteins are shown at the top. * indicates stop codons. References are as indicated in the text.

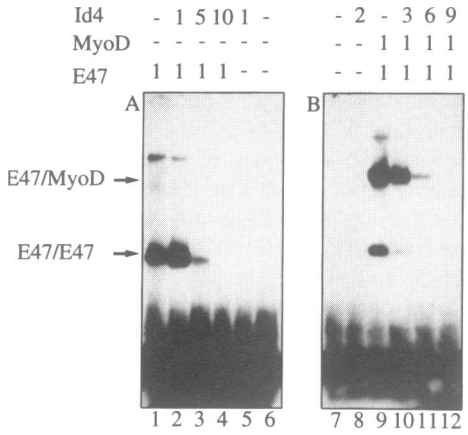


**Figure 3.** Inhibition of binding of *E47* homodimers **(A)** and *E47/MyoD* heterodimers **(B)** to the E-box of the *MCK* enhancer by *Id4*. Electrophoretic mobility shift assay were performed using unlabeled *in vitro* translated proteins mixed with a double stranded, $^{32}P$ labeled *MCK* enhancer oligonucleotide. The numbers at the top indicate the relative amounts of RNAs added to translation reactions. Reticulocyte lysate alone did not give rise to a band shift (lanes 6 and 7).

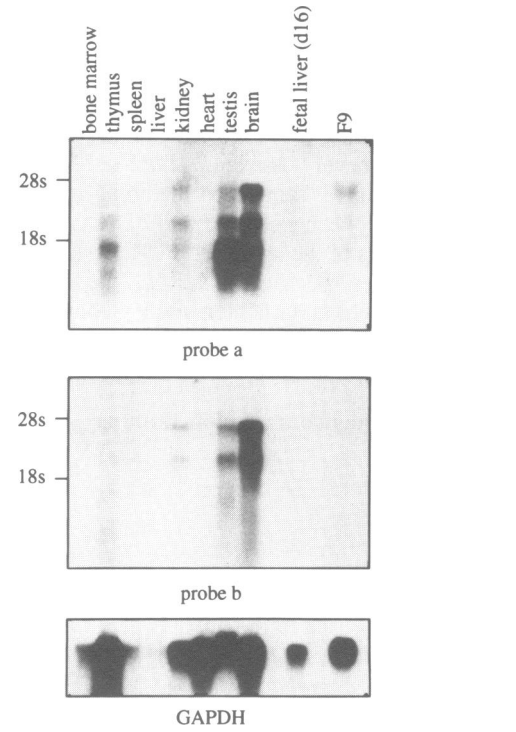


**Figure 4.** Tissue distribution of Id4 mRNA. Northern blot analysis of polyA$^+$ RNA isolated from adult and fetal organs as indicated. **(A)** Probe *a* (corresponding to *VR4*; see Figure 1) hybridizes with three mRNAs of 3.7, 2.0 and 1.7 kb, respectively. **(B)** Probe *b* (corresponding to *F2*; see Figure 1) hybridizes only with the two large mRNAs. **(C)** To control the amount of polyA$^+$ RNA in each lane, the same filter was hybridized with a *GAPDH* probe (55).

members of the dnHLH family, does not contain a basic amino acid region upstream of the HLH motif which is conserved in DNA-binding bHLH proteins (6). Outside of the HLH motif the dnHLH proteins are distinct except for three regions which show some degree of homology (boxes 1, 2 and 4 in Fig. 2). Box 1 comprises the first 8 amino acids of the dnHLH proteins. Upstream of the putative AUG start codons the amino acid sequences are diverged. This is also true for the otherwise highly conserved homologs of human and mouse *Id2* and *Id3*, respectively. We therefore predict that *Id4*, which has a stop codon 22 residues upstream of the putative AUG codon, shares a conserved N-terminus with the other dnHLH proteins. It has been outlined before (36) that the stretches of homology in the N-terminal (box 2) and C-terminal region (box 4) are rich in serine and threonine which could be targets for phosphorylation. In addition, the N-terminus of *Id4* is rich in alanine and the C-terminus is rich in proline residues. The latter feature is shared by the human and mouse *Id3* proteins.

### *Id4* inhibits DNA-binding of *E47* homodimers and *E47/MyoD* heterodimers

To test whether *Id4* is also functionally related to the dnHLH proteins we asked whether *Id4* can inhibit binding of *E47* homodimers or *E47/MyoD* heterodimers to the E-box in the *MCK* (muscle creatine kinase) enhancer oligonucleotide (34). *In vitro* transcripts of *Id4*, *E47* (6) and *MyoD* (4) expression vectors were cotranslated as described (34, 35) and DNA-binding of the resulting protein complexes were determined in an electrophoretic mobility shift assay (Fig. 3). DNA-binding of *E47* homodimers (Fig. 3A) as well as *E47/MyoD* heterodimers (Fig. 3B) was inhibited by the *Id4* protein and this inhibition was dependent upon the amount of *Id4* RNA added to the *in vitro* translation reactions (Fig. 3 lanes 2−4 and 9−12). As expected, *Id4* itself did not generate a band shift indicating that it does not bind the *MCK* enhancer oligonucleotide (Fig. 3 lanes 5 and 8). In parallel reactions, appropriate *in vitro* translation of radiolabeled HLH proteins was monitored by SDS-PAGE analysis (data not shown).

### *Id4* is differentially expressed in adult organs and is up-regulated during embryogenesis

Northern blot analysis, shown in Fig. 4, demonstrates that expression of *Id4* in adult organs of the mouse is highest in brain, testis and kidney and moderate in thymus. Using probe *a* which contains most of the *Id4* coding region (see Fig. 1), three transcripts (3.7, 2.0 and 1.7 kb, respectively) are present (Fig. 4A). Probe *b*, however, hybridizes only with the two larger transcripts (Fig. 4B). We therefore believe that probe *b* corresponds to the 3′ untranslated region of the two larger *Id4* mRNAs and that the three *Id4* transcripts differ in their 3′ ends due to alternatively used polyadenylation sites and/or alternatively spliced untranslated exons. In brain and kidney the three transcripts seemed to be equally abundant contrasting the situation in testis and thymus where much more of the 1.7 kb transcript is present. These differences in the abundance of the three transcripts among the organs suggest a tissue specific regulation of *Id4* expression.

Although not obvious from the Northern blot analysis, amplification of cDNA by RT-PCR revealed, that *Id4* is expressed in bone marrow and spleen, albeit much lower than in testis, brain, and kidney (Fig. 5; lanes 8 to 13). *Id4* RNA is hardly detectable in liver and absent in an endothelial cell line [sEnd.1 (54)].

During embryogenesis *Id4* expression is rather low at day 9.5 of gestation. A dramatic increase of *Id4* expression occurs within the following 4 days of development to reach a plateau at day

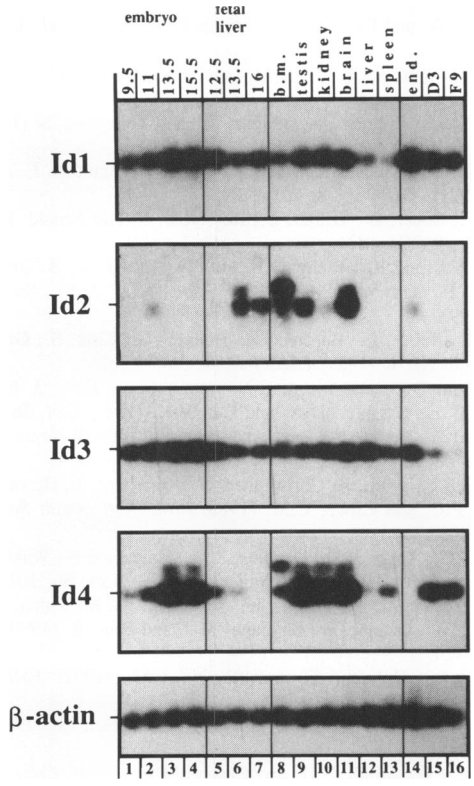


**Figure 5.** Expression patterns of the mouse dnHLH genes *Id1*, *Id2*, *Id3* and *Id4*. cDNA from various stages of embryogenesis (whole embryos from day 9.5, 11, 13.5 and 15.5 of gestation), fetal liver (day 12.5 13.5 and 16.5 of gestation), adult tissues as indicated [bone marrow (b. m.)] and cell lines [endothelial (end.) cell line *sEnd.1* (54), embryonic stem cell line *D3* (56) and carcinoma cell line *F9* (57)] were amplified using primers specific for the dnHLH genes *Id1*, *Id2*, *Id3* and *Id4*. The PCR products were blotted and hybridized with appropriate probes. Amplification of β-actin was performed to control the amount of input cDNA (for details see Material and Methods).

13.5 (Fig. 5; lanes 1 to 4). This up-regulation of *Id4* is not a consequence of the rapidly growing fetal liver during this time of development since expression of *Id4* in the fetal liver is moderate and rather decreases between day 12.5 and 16.5 of gestation (Fig. 5; lanes 5 to 6).

### The expression pattern of *Id4* differs from the restricted expression of *Id2* as well as from the widespread expression of *Id1* and *Id3*

The RT-PCR analysis shown in Fig. 5 revealed that the four dnHLH genes are differentially expressed during mouse development. The widespread expression pattern of *Id3* and *Id1* is in contrast to the restricted expression pattern of *Id4* and *Id2*. In addition, the relative abundance of the transcripts for each dnHLH gene varies in the analyzed tissues and cell lines.

During embryogenesis (day 9.5 to 15.5 of gestation) the level of *Id3* expression is high and stays constant. *Id1* expression seems to increase slightly whereas the expression of *Id4* is dramatically up-regulated. In contrast, *Id2* transcripts could only be detected at day 11 of gestation (Fig. 5; lanes 1 to 4). In embryonic stem cells (*D3*) or carcinoma cells (*F9*) expression of *Id1* and *Id4* is abundant, of *Id3* is moderate and of *Id2* is not detectable (Fig. 5; lanes 15 and 16). In the developing fetal liver (day 12.5 to 16 of gestation) *Id1* and *Id3* are constantly expressed. In contrast, *Id2* is up-regulated and *Id4* is down-regulated suggesting a specific regulatory role of these dnHLH proteins in the development of the fetal liver (Fig. 5; lanes 5–6).

We also determined the expression pattern of the dnHLH genes in adult organs (Fig. 5; lanes 8–13). In brain, testis and bone marrow all four dnHLH genes are expressed. In kidney expression of *Id2* is weak compared to the one of *Id1*, *Id3* or *Id4*. In liver and spleen, however, only *Id3* expression is high. *Id1* expression is moderate in liver and barely detectable in spleen whereas *Id4* shows the opposite expression pattern in these two organs. *Id2* transcripts are neither detectable in liver nor in spleen. In the endothelial cell line *sEnd.1* (Fig. 5; lane 14) expression of *Id1* and *Id3* is high, of *Id2* low and of *Id4* undetectable.

## DISCUSSION

We have isolated and functionally characterized a fourth member (*Id4*) of the mammalian dnHLH gene family and compared its expression pattern with the ones of the other mouse dnHLH genes. Structurally, the dnHLH proteins are highly conserved within the HLH domain resulting in their ability to form heterodimers *in vitro* with ubiquitously expressed bHLH proteins (like *E47*) sequestering the cell-type restricted bHLH proteins (like *MyoD*) and, subsequently, to block DNA-binding. Outside of the HLH domain the dnHLH proteins are less conserved. Three short regions (boxes 1, 2 and 4 in Fig. 2) however, show some degree of homology. Box 2 and 4 are rich in serines and threonines suggesting a common post-translational regulation via phosphorylation (36).

Expression of the *Id4* gene results in three transcripts (3.7, 2.0 and 1.7kb) which are presumably a result of differential splicing or alternatively used polyadenylation sites within the 3′ untranslated region. This transcriptional modification seemed to be regulated in a tissue specific manner since the abundance of each transcript varies between the organs. Such complex regulation of expression has not been described so far for the other members of the dnHLH gene family. *Id1*, *Id2* and *Id3* seemed to be expressed as single transcripts of 1.3 kb (58), 1.6 kb (36) and about 1.0 kb (35), respectively. During mouse development the *Id4* gene expression is also differentially regulated. In undifferentiated embryonic stem cells (D3) and carcinoma cells (F9) *Id4* is highly expressed. In the course of embryogenesis, however, *Id4* expression is low at day 9.5 and increases dramatically within the next 4 days of development. During the same time of embryonic development, expression of *Id1* is slightly increasing, *Id3* is constantly expressed whereas *Id2* transcripts were only detectable by PCR at day 11 of gestation.

The latter result is somewhat surprising since *Id2* transcripts are detectable during embryonic development by Northern Blot analysis (I.v.Crüchten and F.Sablitzky; unpublished). One possible explanation for this discrepancy could be that the *Id2* gene is differentially spliced in the 5′ region with the consequence that the sequence complementary to the *Id2*-sense-primer is missing in those transcripts in the embryos detected by Northern Blot analysis. Alternatively spliced mRNA could also explain the presence of additional PCR products for *Id2* seen in fetal liver and bone marrow (Fig. 5; lanes 6 and 8) and for *Id4* seen in some tissues including embryo, bone marrow, testis, kidney and brain (Fig. 5; lanes 3,4 and 8–11).

*In situ* hybridization indicated that *Id1* is highly expressed in almost all regions of the mouse embryo upon gastrulation (59).

Between day 9.5 and 16.5 of gestation *Id1* was specifically expressed by undifferentiated neuronal precursors of the ventricular zone, but *Id1* transcripts could not be detected in their differentiated derivatives (58). We are currently analyzing the expression pattern of *Id4* during embryogenesis by *in situ* hybridization to determine whether *Id4* expression is similarly restricted to undifferentiated cells.

In adult organs *Id4* is highly expressed in testis, brain and kidney and moderately in thymus, bone marrow and spleen. This expression pattern of *Id4* is again different from the other dnHLH genes. All organs tested contain similar amounts of *Id3* transcripts. This is in agreement with previous data (35) except that we detected *Id3* transcripts also in testis. As was shown before (58), *Id1* is also expressed in all tissues, but *Id1* transcripts are less abundant in liver and in spleen. *Id2* transcripts on the other hand, could be amplified in bone marrow, testis and brain and to a much lesser extent in kidney.

As mentioned before, all four mammalian dnHLH proteins form heterodimers with bHLH transcription factors *in vitro* and block their DNA-binding capacity. Our comparative expression data indicate that in many tissues two or more dnHLH genes are expressed. So far, it is unknown whether such overlap of expression is true at the single cell level. In addition, we note that the presence of RNA is not necessarily indicative of the presence of functional protein. However, assuming that some differentiated cells express functional dnHLH proteins the model outlined above that dnHLH proteins serve as general antagonists of cell differentiation should be extended. It could be that in particular differentiation pathways bHLH transcription factors are required only transiently. It could also be that mature cells which maintain a particular differentiated state independent of the ubiquitous bHLH transcription factors express dnHLH proteins to inhibit their function. Finally, dnHLH and bHLH proteins could be co-expressed resulting in a balanced network of positive and negative regulators. Hence, dnHLH proteins might also be involved in the manifestation of a differentiated cell.

## ACKNOWLEDGEMENTS

We thank G.Heimeroth for expert technical help, E.Rüpping, M.Hotfilder and Dr R.Dildrop for reagents and/or valuable advice and U.Ringeisen for excellent photographic work. We are grateful to Drs H.Weintraub and C.Murre for generously providing expression vectors of *E47* and *MyoD* and Drs M.Hanks and A.Joyner and for generously providing a cDNA library of mouse embryo. We also thank Dr J.Dangl for critically reading this manuscript. This work was supported through the Bundesministerium für Forschung und Technologie.

## REFERENCES

1. Villares, R. and Cabrera, C.V. (1987) *Cell* , **50**, 415–424.
2. Caudy, M., Vässin, H., Brand, M., Tuma, R., Jan, L.Y. and Jan, Y.N. (1988) *Cell* , **55**, 1061–1067.
3. Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G. and Leder, P. (1983) *Cell*, **34**, 779–787.
4. Davis, R.L., Weintraub, H. and Lassar, A.B. (1987) *Cell* , **51**, 987–1000.
5. Lasser, A.B., Buskin, J.N., Lockshon, D., Davis, R.L., Apone, S., Hauschka, S.D. and Weintraub, H. (1989) *Cell* , **58**, 823–831.
6. Murre, C., Schonleber McCaw, P. and Baltimore, D. (1989) *Cell* , **56**, 777–783.
7. Davis, L., Cheng, P.-F., Lassar, A.B. and Weintraub, H. (1990) *Cell*, **66**, 733–746.
8. Voronova, A. and Baltimore, D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4722–4726.
9. Ephrussi, A, Church, G.M., Tonegawa, S. and Gilbert, W. (1985) *Science*, **227**, 134–140.
10. Church, G.M., Ephrussi, A., Gilbert, W. and Tonegawa, S. (1985) *Nature (London)*, **313**, 798–801.
11. Braun, T., Buschhausen-Denker, G., Bober, E., Tennich, E. and Arnold, H.H. (1991) *EMBO J.*, **8**, 701–709.
12. Braun, T., Bober, E., Winter, B.; Rosenthal, N. and Arnold, H.H. (1990) *EMBO J.*, **9**, 821–831.
13. Rhodes, S.J. and Konieczny, S.F. (1989) *Genes Dev.*, **3**, 2050–2061.
14. Miner, J.H. and Wold, B. (1991) *Proc. Natl. Acad. Sci. USA*, **87**, 1089–1093.
15. Braun, T., Bober, E., Buschhausen-Denker, G., Kotz, S., Grzeschik, K. and Arnold, H.H. (1989) *EMBO J.*, **8**, 3617–3625.
16. Edmondson, D.G. and Olson, E.N. (1989) *Genes Dev.*, **3**, 628–640.
17. Wright, W.E., Sassoon, D.A. and Lin, V.K. (1989) *Cell*, **56**, 607–617.
18. Johnson, J.E., Birren, S.J. and Anderson, D.J. (1990) *Nature (London)* , **346**, 858–861.
19. Finger, L.R., Kagan,, J., Christopher, G., Kurtzberg, J., Hershfield, M.S., Nowell, P.C. and Croce, C.M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5039–5043.
20. Begley, C.G., Aplan, P.D., Denning, S.M., Haynes, B.F., Waldmann, T.A. and Kirsch, I.R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 10128–10132.
21. Chen, Q., Cheng, J.-T., Tsai, L.-H., Schneider, N., Buchanan, G., Carroll, A.; Crist, W., Ozanne, B., Siciliano, M.J. and Baer, R. (1990) *EMBO J.*, **9**, 415–424.
22. Green, A.R., DeLuca, E. and Begley, C.G. (1991) *EMBO J.*, **10**, 4153–4158.
23. Aplan, P.D., Nakahara, K., Orkin, S.H. and Kirsch I.R. *EMBO J.*, **11**, 4073–4081.
24. Mellentin, J.D., Smith, S.D. and Cleary, M.L. (1989), *Cell*, **58**, 77–83.
25. Visvader, J., Begley, C.G. and Adams, J.M. (1991) *Oncogene*, **6**, 187–194
26. Kuo, S.S., Mellentin, J.M., Copeland, N.G., Gilbert, D.J., Jenkins, N.A. and Cleary, M.L. (1991) *Oncogene*, **6**, 961–968.
25. Murre, C., Schonleber McCaw, P., Vässin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., Weintraub, H. and Baltimore, D. (1989) *Cell*, **58**, 537–544.
26. Henthorn, P., Kiledjian, M. and Kadesch, T. (1990) *Science*, **247**, 467–470.
27. Lasser, A.B., Davis, R.L., Wright, W.E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D. and Weintraub, H. (1991) *Cell*, **66**, 305–315.
28. Ingham, P.W., Howard, K.R. and Ish-Horowicz, D. (1985) *Nature*, **318**, 439–445.
29. Klämbt, C., Knust, E., Tietze, K and Campos-Ortega, J.A. (1989) *EMBO J.*, **8**, 203–210.
30. Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R. and Nakanishi, S. (1992) Genes Dev., **6**, 2620–2634.
31. Tietze, K., Oellers, N. and Knust, E. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6152–6156.
32. Ellis, H.M., Spann, D.R. and Posakony, J.W. (1990) *Cell*, **61**, 27–38.
33. Garrell, J. and Modolell, J. (1990) *Cell*, **61**, 39–48.
34. Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L. and Weintraub, H. (1990) *Cell*, **61**, 49–59.
35. Christy, B.A., Sanders, L.K., Lau, L.F., Copeland, N.G., Jenkins, N.A. and Nathans, D. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1815–1819.
36. Sun, X.-H., Copeland, N.G., Jenkins, N.A. and Baltimore, D. (1991) *Mol. Cell. Biol.*, **11**, 5603–5611.
37. Kadesch, T. (1992) *Immunol. Today*, **13**, 31–36.
38. Van Doren, M., Powell, P.A., Pasternak, D., Singson, A. and Posakony, J.W. (1992) *Genes Dev.*, **6**, 2592–2605.
39. Martinez, C., Modolell, J., and Garrell, J. (1993) *Mol. Cell. Biol.*, **13**, 3514–3521.
40. Wilson, R.B., Kiledjian, M., Shen, C.-P., Benezra, R., Zwollo, P., Dymecki, S.M., Desiderio, S.V. and Kadesch, T. (1991) *Moll. Cell. Biol.*, **11**, 6185–6191.
41. Kawaguchi, N., DeLuca, H.F.and Noda, M. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 4569–4572.
42. Jen, Y. Weintraub, H. and Benezra, R. (1992) *Genes Dev.*, **6**, 1466–1479.
43. Kreider, B.L., Benezra, R., Rovera, G. and Kadesch, T. (1992) *Science*, **255**,1700–1702.
44. Pongubala, J.M.R. and Atchison, M.L. (1991) *Moll. Cell. Biol.*, **11**, 1040–1047.
45. Kadesch, T. (1993) *Cell Growth Differ.*, **4**, 49–55.
46. Biggs, J., Murphy, E.V.and Israel, M.A. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1512–1517.

47. Ellmeier, W., Aguzzi, A., Kleiner, E., Kurzbauer, R. and Weith, A. (1992) *EMBO J.*, **11**, 2563–2571.
48. Deed, R.W., Bianchi, S.M., Atherton, G.T., Johnston, D., Santibanez-Korf, M., Murphy, J.J. and Norton, J.D. (1993) *Oncogene*, **8**, 599–607.
49. Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
50. Sambrook, J., Fritsch, E.F and Maniatis, T. (1989) Molecular Cloning, Cold Spring Harbor Laboratory Press
51. Saiki, R.K., Faloona, F.A., Mullis, K.B., Horn, C.T., Erlich, H.A. and Arnheim, N. (1985) *Science*, **230**, 1350–1354.
52. Nudel, U., Zakut, R., Shani, M., Neuman, S. Levy, Z., Yaffe, D. (1983) *Nucl. Acids Res.*, **3**, 1759–1771.
53. Brady, G., Barbara, M. and Iscove, N.N. (1990) *Methods Mol. Cel. Biol.*, **2**, 17–25.
54. Williams, R.L., Courtneidge, S.A. and Wagner, E.F. (1988) *Cell*, **52**, 121–131.
55. Piechaczyk, M., Blanchard, J.M., Marty, L., Dani, C., Panabieres, F., Sabouty, S.E., Fort, P. and Jeanteur, P. (1984) *Nucl. Acids Res.*, **12**, 6951–6963.
56. Doetschman, T.C., Eistetter, M., Katz, M., Schmidt, W. and Kemler, R. (1985) *J. Embryol. Exp. Morph.*, **87**, 27–45.
57. Bernstine, E.G., Hooper, M.L., Grandchamp, S. and Ephrussi, B. (1973) *Proc. Nat. Acad. Sci., USA* **70**, 3899–3903.
58. Duncan, M., DiCicco-Bloom,E.M., Xiang, X., Benezra, R. and Chada, K. (1992) *Dev. Biol.*, **154**, 1–10.
59. Wang, Y, Benezera, R. and Sassoon, D.A. (1992) *Dev. Dynamics*, **194**, 222–230.