Available online at www.sciencedirect.com
SCIENCE DIRECT®
Gene 323 (2003) 133–140
GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES
www.elsevier.com/locate/gene

# The mouse zinc-fingers and homeoboxes (ZHX) family; ZHX2 forms a heterodimer with ZHX3

Hiroko Kawata[a], Kazuya Yamada[a,b,*], Zhangfei Shou[b], Tetsuya Mizutani[a,b], Kaoru Miyamoto[a,b]

[a] *CREST, JST (Japan Science and Technology), Fukui 910-1193, Japan*
[b] *Department of Biochemistry, Fukui Medical University, Fukui 910-1193, Japan*

Received 19 May 2003; received in revised form 4 September 2003; accepted 10 September 2003

Received by T. Gojobori

**Abstract**

Human *z*inc-fingers and *h*omeobo*x*es (ZHX) 1, ZHX2 and ZHX3, members of the ZHX family, contain two $Cys_2$–$His_2$-type zinc-finger motifs and five homeodomains (HDs). These proteins not only form homodimers but heterodimers with ZHX1 as well and act as ubiquitous transcriptional repressors. The cloning of mouse ZHX2 and ZHX3 cDNAs and the corresponding genes from a 129 mouse genomic library are reported, along with an analysis of the heterodimerization of ZHX2 with ZHX3. The mouse ZHX2 and ZHX3 proteins consist of 836 and 951 amino acid residues, respectively. The similarity of amino acid sequences of each protein with those of human orthologue is 87.0% and 85.2%, respectively. An analysis of genomic clones revealed that an entire coding sequence and a portion of the 5′- and 3′-noncoding sequence of mouse ZHX2 cDNA are encoded by a single exon of the mouse *ZHX2* gene as well as the mouse *ZHX1* gene. In contrast, in the case of the mouse *ZHX3* gene, the coding sequences of ZHX3 cDNA are separated by an intron. A 4.5-kb ZHX2 transcript, and three ZHX3 transcripts, 9.5-, 6.5- and 4.4-kb, are ubiquitously expressed, although their levels vary. Lastly, in vitro and in vivo protein–protein interaction assays revealed that ZHX2 is able to form a heterodimer with ZHX3 via a region containing each HD1.


*Keywords:* cDNA cloning; Genome cloning; Transcriptional repressor; Dimerization; Zinc-finger motif; Homeodomain

## 1. Introduction

The *z*inc-fingers and *h*omeobo*x*es (ZHX) family consists of three members, ZHX1, ZHX2 and ZHX3 (Yamada et al., 1999a, 2003; Kawata et al., 2003). All these proteins contain two $Cys_2$–$His_2$-type zinc-finger (Znf) motifs and five homeodomains (HDs), and are ubiquitous transcriptional repressors that are localized in the nuclei of cells (Barthelemy et al., 1996; Yamada et al., 1999a,b, 2002, 2003; Hirano et al., 2002; Kawata et al., 2003). Mammalian ZHX1 consists of 873 amino acid residues and the similarity of the amino acid sequence of the human ZHX1 with that of the rat and mouse forms are 93% and 91%, respectively (Barthelemy et al., 1996; Yamada et al., 1999a; Hirano et al., 2002). Human ZHX2 and ZHX3 consist of 837 and 956 amino acid residues, respectively. The similarity of amino acid sequence of the human ZHX1 with that of ZHX2 and ZHX3 are 41.9% and 34.4%, respectively. The ZHX family proteins form a dimer with ZHX1 via a region containing the HD1.

Nuclear factor-Y (NF-Y), a ubiquitous transcriptional activator, consists of three subunits, NF-YA, NF-YB and NF-YC. The NF-YA subunit contains two activation domains (ADs), a glutamine-rich region and a serine/threonine-rich region. ZHX proteins interact with different ADs of NF-YA; ZHX1 interacts with the glutamine-rich region; and both ZHX2 and ZHX3 interact with the serine/threonine-rich region, respectively (Yamada et al., 1999b, 2003; Kawata et al., 2003). The interaction domain of ZHX1 with the glutamine-rich AD of the NF-YA is the amino acid

*Abbreviations:* ZHX, zinc-fingers and homeoboxes; Znf, zinc-finger; HD, homeodomain; NF-Y, nuclear factor-Y; AD, activation domain; NLS, nuclear localization signal; RT-PCR, reverse transcription–polymerase chain reaction; RACE, rapid amplification of the cDNA ends; GST, glutathione-*S*-transferase; DDBJ, The DNA Data Bank of Japan; DBD, DNA-binding domain.

* Corresponding author. Department of Biochemistry, Fukui Medical University, Fukui 910-1193, Japan. Tel.: +81-776-61-8316; fax: +81-776-61-8102.
*E-mail address:* kazuya@fmsrsa.fukui-med.ac.jp (K. Yamada).

0378-1119/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2003.09.013

sequence between residues 272 and 564, which contains the HD1 through HD2 region (Yamada et al., 1999b). The interaction domains of ZHX2 and ZHX3 with the serine/threonine-rich region of NF-YA are comprised of the amino acid sequences between residues 263 and 497, and between 242 and 488, respectively. These regions correspond to the HD1 through HD2 of ZHX2 and a region containing the HD1 of ZHX3, respectively (Kawata et al., 2003; Yamada et al., 2003).

ZHX1 contains a transcriptional repressor domain in the C-terminal acidic region, corresponding to the amino acid sequence between residues 831 and 873 (Yamada et al., 2002). The repressor domains of ZHX2 and ZHX3 are a region containing the HD1, which correspond to the amino acid sequence between residues 263 and 446, and 303 and 502, respectively (Kawata et al., 2003; Yamada et al., 2003). The dimerization with ZHX1, as well as each repressor domain of ZHX1 and ZHX3, are required for full repressor activities (Yamada et al., 2002, 2003). The nuclear localization signals (NLSs) of these proteins are different. In the case of ZHX1, it is located in an arginine-rich basic region, corresponding to the amino acid sequence between residues 734 and 768 (Yamada et al., 2002). The NLS of ZHX2 is located in the amino acid sequence between residues 317 and 446, which contains a proline-rich region (Kawata et al., 2003). In contrast, ZHX3 contains two NLSs that are located in the N-terminal Znf1 and HD2 region, and at least one nuclear export signal, located in the C-terminal region, in the molecule (Yamada et al., 2003).

While the human *ZHX1* and *ZHX2* genes are located on chromosome 8q, the *ZHX3* gene is located on chromosome 20q (Yamada et al., 1999a, 2003; Kawata et al., 2003). We recently isolated the mouse *ZHX1* gene from a 129 *svj* mouse library and determined the complete genomic structure (Shou et al., 2003). We also reported that transcription of the *ZHX1* gene is synergistically stimulated by PEA3 and Yin and Yang 1 (Shou et al., 2003). However, the mouse *ZHX2* and *ZHX3* genes remain to be analyzed.

In the present study, to obtain a better understanding of the evolution of the ZHX family and the mechanism of the transcriptional regulation, we cloned mouse ZHX2 and ZHX3 cDNAs and the corresponding genes from a 129 *svj* male mouse spleen genomic library. We also report on heterodimerization between ZHX2 and ZHX3 and the mapping of their minimal domains.

## 2. Materials and methods

### 2.1. Materials

F9 cells, a mouse embryonal carcinoma cell line, were obtained from the Japanese Collection of Research Bioresources. The TRIOZOL reagent and Superscript II were purchased from Invitrogen (Groningen, the Netherlands). The ExTaq DNA polymerase and BcaBest DNA labeling kit were obtained from Takara Biomedicals (Kyoto, Japan). The pGEM-T Easy vector and T7 TNT Quick-coupled transcription/translation system were purchased from Promega (Madison, WI). Mouse Liver Marathon-Ready cDNA, the Advantage 2 PCR kit, the mouse Multiple Tissue Northern blot and the ExpressHyb hybridization solution were purchased from CLONTECH (Palo Alto, CA). The Big Dye terminator FS cycle sequencing kit was purchased from Applied Biosystems Japan (Tokyo, Japan). $\alpha$-$^{32}$P dCTP (110 TBq/mmol), pGEX-5X-1, glutathione-sepharose 4B and $^{35}$S-methionine (37 TBq/mmol) were purchased from Amersham Biosciences (Cleveland, OH). The Qiagen lambda kit was purchased from QIAGEN (Hilden, Germany). The TOPP3 cells were obtained from Stratagene (La Jolla, CA).

### 2.2. Reverse transcription–polymerase chain reaction (RT-PCR) and rapid amplification of the cDNA ends (RACE)

F9 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum at 37 °C in a 5% $CO_2$ incubator. Total RNA was prepared from F9 cells using the TRIZOL reagent according to the manufacture's recommended protocol. RT-PCR was performed as described previously with minor modifications (Yamada et al., 1999b). Combinations of S-hrZHX2, 5′-GCCCG CCTGG TGACA GACAC-3′, and As-hrZHX2, 5′-GCGAC CCACC GTTTG GTCTA AG-3′, S-hrZHX3, 5′-CCAAT CATGA AGATA ATGAA AGGC-3′, and As-hrZHX3, 5′-GTGGG CTGAG GCACA GACTG-3′, S-mZHX2-Met, 5′-ATGGC AAGCA AACGG AAATC-3′, and As-mZHX2-665, 5′-GGCTG TCACC AGGCG GGC-3′, S-mZHX2-920, 5′-GAATC TGGTT TGCCA CCCAG CG-3′, and As-mZHX2-STOP, 5′-CTAGG CTTGG CCGGC CTCTG C-3′, S-mZHX3-Met, 5′-ATGGC CAGCA AAAGG AAGTC-3′, and As-mZHX3-541, 5′-CCTTT CATTA TCTTC ATGAT TGG-3′, and S-mZHX3-1148, 5′-CCAGT CTGTG CCTCA GCCCA C-3′, and As-mZHX3-STOP, 5′-TCAGT CTGCT TCGAG TTG-3′, were used as the primers. PCR products were subcloned into the pGEM-T Easy to give pGEM-T Easy hrZHX2, pGEM-T Easy hrZHX3, pGEM-T Easy mZHX2-Met-665, pGEM-T Easy mZHX2-920-STOP, pGEM-T Easy mZHX3-Met-541 and pGEM-T Easy mZHX3-1148-STOP, respectively.

To obtain 5′- and 3′-noncoding, and a part of the coding sequence of the mouse ZHX2 and ZHX3 cDNAs, we employed the 5′- and 3′-RACE method using Mouse Liver Marathon-Ready cDNA and the Advantage 2 PCR kit. The gene-specific primers of 5′- and 3′-RACE, mZHX2-5RACE-As1, 5′-GCCTC CTCCA GCATA TCTTG TTCC-3′, mZHX2-5RACE-As2, 5′-GTGAT GTCCG AACCA TGCAG GGAG-3′, mZHX3-5RACE-As1, 5′-CACAG GCTGA GGCTC TGTGC TGG-3′, mZHX3-5RACE-As2, 5′-GTCTT AACGG GGATC ATGCA TGGG-3′, mZHX2-3RACE-S1, 5′-CCAAG AGGAG AGTGC AAGTC AAACC-3′, mZHX2-3RACE-S2, 5′-GTAGC AGCTT

GGCTG CGGGA GAG-3′, mZHX3-3RACE-S1, 5′-CATGG GAGCC AAGTG CTCCT GAG-3′, and mZHX3-3RACE-S2, 5′-GAGCC TTTTG ACACT TCAAG TCCCC-3′, were used. The RACE procedure was carried out according to the manufacturer's recommended protocol. All amplified DNA fragments were also subcloned into the pGEM-T Easy vector.

The nucleotide sequences of the inserts of these plasmids were determined by sequencing.

### 2.3. Genomic library screening

A 129 *svj* mouse male spleen genomic library was a generous gift from Dr. Makoto Satoh (Fukui Medical University, Japan). A 300-bp *Eco*RI fragment of the pGEM-T Easy hrZHX2 and a 600-bp *Eco*RI fragment of the pGEM-T Easy hrZHX3 were employed as probe DNAs for ZHX2 and ZHX3, respectively. These probes were labeled with α-$^{32}$P dCTP using a BcaBest DNA labeling kit. Five × $10^5$ independent clones were screened using the plaque hybridization method (Sambrook and Russell, 2001). Prehybridization, hybridization and washing conditions have been described previously (Shou et al., 2003). After drying, the filters were exposed to a Kodak BioMax film at − 80 °C with an intensifying screen. Lambda DNA was prepared from the positive clones using the Qiagen lambda kit. Their nucleotide sequences were determined by direct sequencing.

### 2.4. Poly(A)$^+$-RNA blot analysis

Mouse Multiple Tissue Northern blot was hybridized with $^{32}$P-labeled mZHX2 or mZHX3 probe. The ExpressHyb hybridization solution was used for prehybridization and hybridization. The prehybridization, hybridization and washing procedures were performed according to the manufacture's recommended protocol. The blot was exposed to a FUJIX imaging plate (Kanagawa, Japan). Hybridization signals were detected with the FUJIX BAS-2000 image analyzing system.

### 2.5. Glutathione-S-transferase (GST) pull-down assays

pGST-ZHX2 (1–837) and pGST-ZHX3 (1–956) have been described previously (Kawata et al., 2003; Yamada et al., 2003). TOPP3 cells were transformed with the pGEX-5X-1, pGST-ZHX2 (1–837) or pGST-ZHX3 (1–956) fusion protein expression plasmid. The preparation of the GST-fusion protein and in vitro-translated, $^{35}$S-labeled, full-length of ZHX2 and ZHX3 proteins have been described previously (Kawata et al., 2003; Yamada et al., 2003). Finally, the beads were resuspended in an equal volume of 2 × SDS sample buffer and each supernatant was subjected to a 10% SDS-PAGE, along with a prestained molecular weight marker. The gel was dried and exposed to a FUJIX imaging plate. Interaction signals were detected with the FUJIX BAS-2000 image analyzing system. The relative purity and amount of each fusion protein were determined by gel staining with Coomassie Brilliant Blue R-250.

### 2.6. Yeast two-hybrid system and liquid β-galactosidase assays

The pDBD, pDBD-ZHX2 (263–322), pDBD-ZHX2 (195–358) (previously referred to as the pDBD-L26), pDBD-ZHX2 (1–837), pACT2, pAD-ZHX3 (303–364), pAD-ZHX3 (242–488) and pAD-ZHX3 (242–615) (previously referred to as the pAD-G23) plasmids have been described previously (Kawata et al., 2003; Yamada et al., 2003). The yeast strain SFY526 that contains a quantifiable *lacZ* reporter gene, was transformed with pDBD, pDBD-ZHX2 (263–322), pDBD-ZHX2 (195–358) or pDBD-ZHX2 (1–837). These reporter yeasts were also transformed with pACT2, pAD-ZHX3 (303–364), pAD-ZHX3 (242–488) or pAD-ZHX3 (242–615), respectively. Quantitative β-galactosidase assays, using *o*-nitrophenyl-β-D-galactoside, were carried out on permeabilized cells, as described previously (Yamada et al., 1999a,b; Hirano et al., 2002).

## 3. Results

### 3.1. PCR cloning and determination of nucleotide sequence of mouse ZHX2 and ZHX3 cDNAs

To obtain mouse ZHX2 and ZHX3 cDNAs, RT-PCRs were performed using total RNA from mouse F9 cells. The RT-PCR primers for ZHX2 and ZHX3 were prepared from nucleotide sequences that are completely conserved between human and rat (Kawata et al., 2003; Yamada et al., 2003). A 300-bp product of mZHX2 and a 600-bp product of mZHX3 were initially obtained. After a determination of their nucleotide sequences, some combinations of RT-PCR primers were newly synthesized and RT-PCRs were then carried out. Furthermore, to clone the 5′- and 3′-noncoding sequences and a portion of the coding sequence of both mouse ZHX2 and ZHX3 cDNAs, 5′- and 3′-RACE methods were employed. Using combinations of gene-specific primers and adaptor primers, cDNA fragments were obtained by PCR using the Mouse Liver Marathon-Ready cDNA as a template. Finally, the sizes of the entire coding sequence of both mouse ZHX2 and ZHX3 cDNAs were determined to be 2508- and 2853-bp, respectively. The nucleotide sequences of their cDNAs have been submitted to the DNA Data Bank of Japan (DDBJ) Accession No. AB099526 and AB099527, respectively. The ZHX2 contains an open reading frame of 836 amino acid residues. The deduced amino acid sequence is compared with that of human ZHX2 (Fig. 1A). The mouse ZHX2 protein has a predicted molecular mass of 92.2 kDa and an isoelectric point of 7.12. In contrast,

the ZHX3 consists of 951 amino acid residues. The deduced amino acid sequence is compared with that of human ZHX3 (Fig. 1B). The mouse ZHX3 protein has a predicted molecular mass of 95.1 kDa and an isoelectric point of 6.12. The amino acid sequences of mouse ZHX1, ZHX2 and ZHX3 were examined using the Clustal W program to analyze their phylogenetic relationship. As shown in Fig. 1C, ZHX1 and ZHX2 are very similar but ZHX3 is different.

### 3.2. Cloning of the mouse ZHX2 and ZHX3 genes

We then screened $5 \times 10^5$ independent clones from the 129 *svj* mouse spleen genomic library to obtain clones encoding the mouse *ZHX2* and *ZHX3* genes, respectively. We obtained 10 and 2 positive clones for each gene and then determined nucleotide sequence of these clones. The nucleotide sequences of the mouse *ZHX2* and *ZHX3* genes have been submitted to the DDBJ Accession No. AB099704, AB099702, and AB099703, respectively. As shown in Fig. 2, the entire coding sequence and a portion of the 5′- and 3′-noncoding sequence of mouse ZHX2 cDNA were encoded by a single exon as well as the mouse *ZHX1* gene (Shou et al., 2003). In contrast, in the case of the mouse *ZHX3* gene, a portion of the 5′-noncoding sequence and nearly the full-length coding sequence were encoded by an exon, but the resultant coding sequence (9-bp) and 3′-noncoding sequence were encoded by another exon. Thus, the coding sequence of ZHX3 was disrupted by an intron. GT and AG residues are present at the 5′- and 3′-boundaries of the intron of the

**A)**

```
Mouse    1:MASKRKSTTP CMVRTSQVLE QDMLEEADRA KDKGAGMPQS DVTKDSWAAE PEHSSKETEV 60
Human    1:---------- --------V- --VP-V---- -E--I-T--P -A-------- L-N----N-- 60

                           +++ ++++++++++ ++++++++++ +         + ++++++++++
Mouse   61:VEVKSMGENP SKKLQGGYEC KYCPYSTQNL NEFTEHVDMQ HPNVILNPLY VCAECNFTTK 120
Human   61:I-------SQ ---------- ---------- ---------- ---------- ---------- 120

          ++++++++++ ++++++++
Mouse  121:KYDSLSDHNS KFHPGETNFK LKLIKRNNQT VLEQSIEATN HVVPITASGP GSSDNDPGVS 180
Human  121:---------- ---------- ---------- -------T-- ---S--T--- -TG-S-S-I- 180

Mouse  181:VGKTPMTKTG KLKADAKKVP KKPDEAAPEN HMEGTARLVT DTAEILARLG SVELLQDSLG 240
Human  181:-S---IM-P- P--------- ---E-IT--- -V-------- ------S--- G------T-- 240

Mouse  241:HVMPSVQLPP NINLVPKVPV PLNTTKYNSA LDTNATMINS FNKFPYPTQA ELSWLTAASK 300
Human  241:---------- ---------- ---------- ---------- ---------- ---------- 300

Mouse  301:HPEEHIRIWF ATQRLKHGIS WSPEEVEEAR KKMFNGTIQS VPPTITVLPA QLTPTKVSQP 360
Human  301:---------- ---------- ---------- ---------- ---------- --A----T-- 360

Mouse  361:ILQTALPCQI LGQPSLVLTQ VTSGSTTVSC SPITLAVAGV TNHGQKRPLV TPQAAPEPKR 420
Human  361:---------- ---T------ ---------- ---------- ---------- ---------- 420

Mouse  421:PHIAQVPEPP PKVANTPLTP ASDRKKTKLQ IAHLKASFLQ SQFPDDAEVY RLIEVTGLAR 480
Human  421:---------- -----P---- --------E- ---------- ---------- ---------- 480

Mouse  481:SEIKKWFSDH RYRCQRGIVH ITSESLAKDQ MAITGTRHGR TYHVYPDFAP QKFKEKSQGQ 540
Human  481:---------- ---------- ---------- L-AAS----- ---A------ ------T--- 540

Mouse  541:LKTLEDSFLK SSFPTQAEVE RLRVETKLSR REIDSWFSER RKLRDSMEQA VLDSMGSGKK 600
Human  541:V-I------- --------LD ---------- ---------- ---------- ---------- 600

Mouse  601:GSDAVAPNGA LSRLDQLSGA QLAGSLPSPS SAIVQNQEQV HLLRSTFART QWPTPQEYDQ 660
Human  601:-Q-VG----- ---------- --TS------ P-AKS----- ---------- ---------- 660

Mouse  661:LAAKTGLVRT EIVRWFKENR CLLKTGTLSW LEQYQRHHLS DDRGRDAVSR KVAKQVAESP 720
Human  661:---------- ---------- -------VK- M----HQPMA --H-Y---A- -AT-PM---- 720

Mouse  721:KNGSEAAHQY AKDPKALSEE DSEKLVPRMK VGGDPTKDCL AGKPSEATSD RSEG*SRDGQ 779
Human  721:----DVVP-- -----K-C-- -L----T-V- --SE-A---- PA-------- ----S----- 780

Mouse  780:GSEENEESGI VDFVEVTVGE EDAISEKWGS WSRRVAEGTV ERADSDSDST PAEAGQA    836
Human  781:--D-----SV -Y-------- -----DRSD- --QAA---VS -L-E----CV -------    837
```

Fig. 1. Deduced amino acid sequences of mouse ZHX2 and ZHX3 and comparisons with each human orthologue. The amino acid sequences of mouse ZHX2 (A) and ZHX3 (B) are shown and compared with those of human ZHX2 and ZHX3, respectively. The dashes indicate amino acid identity. The two Znf motifs and five HDs are indicated by plus signs and underlines, respectively. The asterisks indicate a gap when no corresponding amino acid sequence between each protein was found. (C) Phylogenetic tree of members of the ZHX family. Scores were obtained from an analysis using the Clustal W program at the DDBJ.

B)

```
Mouse     1:MASKRKSTTP CMIPVKTVVL PGASTEPQPV ESLPEGPQQD LPSEAPDASS EAAPNPSSTD 60
Human     1:---------- ---------- QD----A--A -T-------- --P-SA---- ---Q------ 60

                       ++++ ++++++++++ ++++++++++          ++ ++++++++++
Mouse    61:GSALANGHRS TLDGYVYCCK ECEFRSQDVT HFIGHMNSEH TDFNKDPTFV CTGCSFLAKN 120
Human    61:--T------- -----L-S-Y -D----H-M- Q-V------- ---------- -S-------T 120

          ++++++++++ ++++++++
Mouse   121:PEGLSLHNAK CHSGEASFLW NVTKPDNHVV VEQSVPDSAS SSVLAGE*ST **TEG*TEII 176
Human   121:---------T --------V- --A------- ----I-E-T- TPD----PSA EGAD-QA--- 180

Mouse   177:ITKTPIMKIM KGKAEAKKIH MLKENAPNQP GSEALPKPLA GEREVKEGDH TFINGAAPGS 236
Human   181:---------- ---------- T----V-S4V G------LST -M---R---- S-----V-V- 240

Mouse   237:QASAKSTKPP PAANGPLIGT VPVLPAGIAQ FLSLQQQPPV HAQHHTHQPL PTSKTLPKVM 296
Human   241:----S-A-N- H--------- ---------- ---------- -----V---- --A-A----- 300

Mouse   297:IPLSSIPTYN AAMDSNSFLK NSFHKFPYPT KAELCYLTVV TKYPEEQLKI WFTAQRLKQG 356
Human   301:---------- ---------- ---------- ---------- ---------- ---------- 360

Mouse   357:ISWSPEEIED ARKKMFNTVI QSVPQPTITV LNTPLVASAG NVQHLIQATL PGHAVGQPEG 416
Human   361:---------- ---------- ---------- ---------- --------A- ---V------ 420

Mouse   417:TAGGLLVTQP LMANGLQASS SSLPLTTASV PK*PTVAPIN TVCSNSASAV KVVNAAQSLL 475
Human   421:-G-------- --------T- P-----VT-- --Q-G----- -----TT--- ---------- 480

Mouse   476:TACPSITSQA FLDANIYKNK KSHEQLSALK GSFCRNQFPG QSEVEHLTKV TGLSTREVRK 535
Human   481:---------- ----S----- ---------- ---------- ---------- ---------- 540

Mouse   536:WFSDRRYHCR NLKGSRAMMP GEHGSVLIDS VPEVPFPLAS KVPEVTCIPT ATSLVSHPAT 595
Human   541:---------- --------I- -D-S-II--- ----S-SPS- ---------- TAT-AT--SA 600

Mouse   596:KRQSWHQTPD FTPTKYKERA PEQLRVLENS FAQNPLPPEE ELDRLRSETK MTRREIDGWF 655
Human   601:---------- ---------- -----A-S-- -------LD- ---------- -------S-- 660

Mouse   656:SERRKKVNTE ETKKADGHMP KEEEEGAEQE GRDEELANEL RVPGENGSPE MFLSHALAER 715
Human   661:--------A- -----EENAS Q----A--D- -GE-D--S-- --S-----L- -PS--I---- 720

Mouse   716:KVSPIKINLK NLRVTEASGK SEFPGMGVCE PEEDGLNKLV EQPPSKVSYK KTAQQRHLLR 775
Human   721:---------- -------N-R N-I--L-A-D --D-ES---A ----G---C- ---------- 780

Mouse   776:QLFVQTQWPS NQDYDSIMAQ TGLPRPEVVR WFGDSRYALK NGQLKWYEDY KRGNFPPGLL 835
Human   781:---------- ---------- ---------- ---------- ---------- ---------- 840

Mouse   836:VIAPGNRELL QDYYMTHKML CEEDLQTLCD KTQMSAQQVK QWFAEKMGEE TRAVADISSE 895
Human   841:---------- ---------- Y-----N--- -----S---- ---------- -------G-- 900

Mouse   896:DQGPRNGEPV AVHKVLGDAY SELSENSESW EPSAPEASSE PFDTSSPQSG RQLEAD     951
Human   901:----GT--LT ----GM--T- --V------- --RV------ --------A- ----T-     956
```

C)

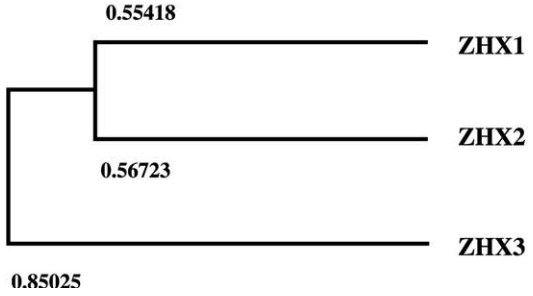


Fig. 1 (*continued*).

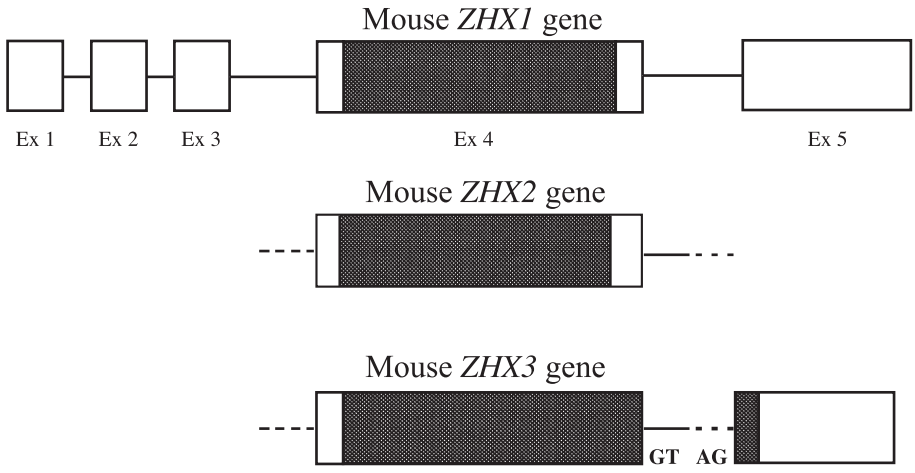


Fig. 2. Schematic diagram of the mouse *ZHX* genes. The exons and introns are indicated by squares and lines, respectively. The open and solid squares indicate noncoding and coding sequences of ZHX cDNAs, respectively. The dotted line indicates undetermined nucleotide sequences. Ex indicates the exon.

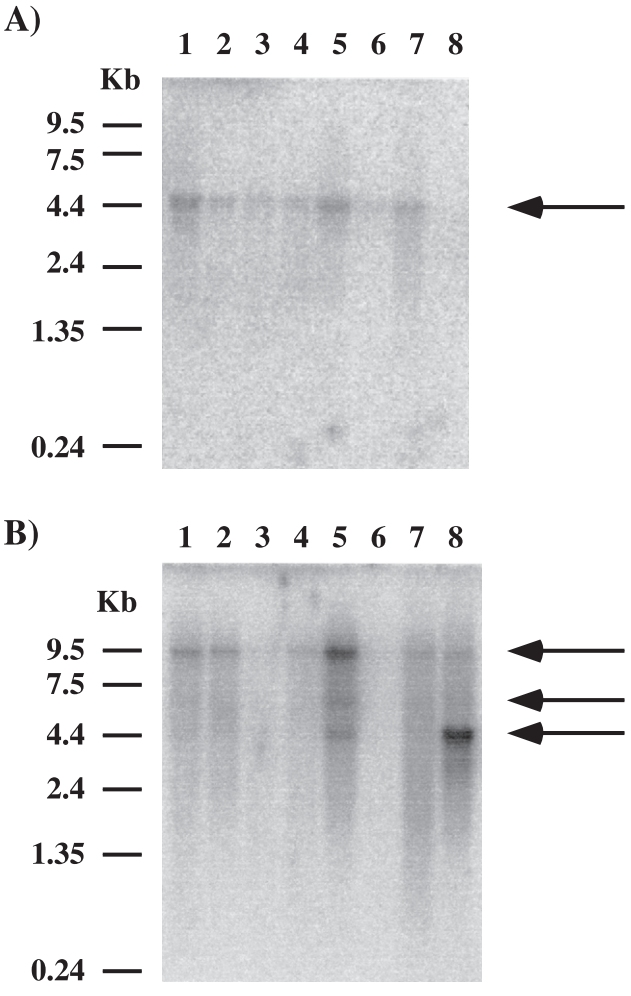


Fig. 3. Tissue distribution of mouse ZHX2 and ZHX3 mRNAs. A mouse Multiple Tissue Northern blot was hybridized with $^{32}$P-labeled mouse ZHX2 (A) or ZHX3 cDNA (B). Each lane contains 2 μg of poly A$^+$-RNA isolated from the indicated tissues. Size markers are shown on the *left* in kb. Lane 1, heart; lane 2, brain; lane 3, spleen; lane 4, lung; lane 5, liver; lane 6, skeletal muscle; lane 7, kidney; lane 8, testis.

mouse *ZHX3* gene, consistent with the consensus sequence for the splicing of eukaryotic mRNA (Mount, 1982).

### 3.3. Tissue distribution of mouse ZHX2 and ZHX3 mRNAs

The tissue distribution of mouse ZHX2 and ZHX3 mRNAs were determined by northern blot analysis (Fig. 3). A 4.5-kb ZHX2 transcript and three ZHX3 transcripts, approximately 9.5-, 6.5- and 4.4-kb, were observed in all mouse tissues examined, although the intensity of these transcripts varied among tissues. These results indicate that mouse ZHX2 and ZHX3 mRNAs are expressed ubiquitously.

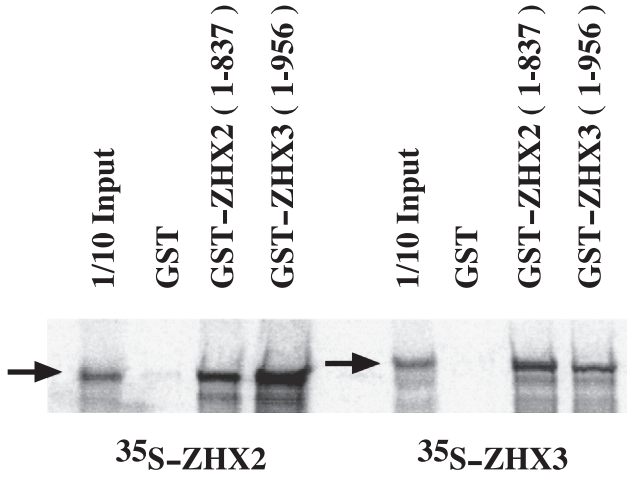


Fig. 4. Heterodimerization between ZHX2 and ZHX3 using GST pull-down assays. In vitro-translated, $^{35}$S-labeled, ZHX2 or ZHX3 was incubated with glutathione-sepharose beads containing bound GST, GST-ZHX2 or GST-ZHX3 fusion protein, respectively. The beads were washed thoroughly and the bound protein was eluted and analyzed by 10% SDS-PAGE. The interaction signal was detected by the FUJIX BAS-2000 image analyzing system. The signal in the lane marked "1/10 Input" denotes 10% of the protein added to the reactions shown in the other lanes.

### 3.4. Heterodimerization between ZHX2 and ZHX3

Both ZHX2 and ZHX3 form heterodimers with ZHX1 as well as homodimers (Kawata et al., 2003; Yamada et al., 2003). We then investigated the issue of whether ZHX2 and ZHX3 form a heterodimer with each other. We used in vitro GST pull-down assays to test for the heterodimerization of ZHX2 with ZHX3. We employed three plasmids, the pGEX-5X-1, which expresses GST alone, pGST-ZHX2 (1–837) and pGST-ZHX3 (1–956), which expresses the entire coding region of human ZHX2 and ZHX3 fused to GST, respectively. In vitro-translated, $^{35}$S-labeled full-length ZHX2 or ZHX3 was incubated with purified GST, GST-ZHX2 (1–837) or GST-ZHX3 (1–956) protein bound to glutathione-sepharose. Both ZHX2 and ZHX3 proteins strongly bound to the GST-ZHX2 (1–837) and GST-ZHX3 (1–956) fusion proteins, but not to GST alone (Fig. 4). In contrast, an unprogrammed reticulocyte lysate failed to bind to either protein (data not shown). These results indicate that both ZHX2 and ZHX3 are able to form a heterodimer in vitro.

We then determined each minimal heterodimerization domain of ZHX2 and ZHX3 using an yeast two-hybrid system. Four SFY526 yeast strains harboring the pDBD, pDBD-ZHX2 (263–322), pDBD-ZHX2 (195–358) or pDBD-ZHX2 (1–837), were used as reporter yeasts. We prepared four prey plasmids, pACT2, pAD-ZHX3 (303–364), pAD-ZHX3 (242–488) and pAD-ZHX3 (242–615). These plasmids were transformed into the reporter yeasts and β-galactosidase activities were then determined (Table 1). When the yeast SFY526 strain harboring pDBD or pDBD-ZHX2 (263–322) was transformed with the four prey plasmids, the β-galactosidase activities were very low. Yeasts harboring combinations of the pDBD-ZHX2 (195–358) or pDBD-ZHX2 (1–837), and pACT2 or pAD-ZHX3 (303–364) also showed low β-galactosidase activities. In contrast, yeasts harboring combinations of pDBD-ZHX2 (195–358) or pDBD-ZHX2 (1–837), and pAD-ZHX3 (242–488) or pAD-ZHX3 (242–615), expressed the highest β-galactosidase activity of all the yeasts tested. While pDBD-ZHX2 (195–358) expresses the amino acid sequence between residues 195 and 358 of human ZHX2 fused to the C-terminal of the GAL4 DNA-binding domain (DBD), pAD-ZHX3 (242–488) expresses the amino acid sequence between residues 242 and 488 of human ZHX3 fused to the C-terminal of the GAL4 AD. Each region encodes a region containing the HD1. These results indicate that both ZHX2 and ZHX3 form a heterodimer via a region containing each HD1 in vivo.

Table 1
Mapping of the heterodimerization domains between ZHX2 and ZHX3 using a yeast two-hybrid system

| | pDBD | pDBD-ZHX2 (263–322) | pDBD-ZHX2 (195–358) | pDBD-ZHX2 (1–837) |
|---|---|---|---|---|
| pACT2 | 0.1 ± 0.1 | 0.1 ± 0.1 | 0.2 ± 0.1 | 0.2 ± 0.1 |
| pAD-ZHX3 (303–364) | 0.2 ± 0.1 | 0.3 ± 0.1 | 0.1 ± 0.1 | 0.1 ± 0.1 |
| pAD-ZHX3 (242–488) | 0.1 ± 0.1 | 0.1 ± 0.1 | 18.0 ± 7.6 | 2.0 ± 1.0 |
| pAD-ZHX3 (242–615) | 0.3 ± 0.1 | 0.2 ± 0.1 | 5.6 ± 3.3 | 1.8 ± 0.3 |

Each value represents the mean and standard error of β-galactosidase activities (U) in at least three experiments.

## 4. Discussion

Mouse ZHX2 and ZHX3 cDNAs and the corresponding genes were cloned. We also demonstrated the heterodimerization of ZHX2 with ZHX3 proteins. The number of amino acid residues of mouse ZHX1, ZHX2 and ZHX3 are 873, 836 and 951, respectively. The similarities in the amino acid sequence between mouse and human ZHX2, and between mouse and human ZHX3 were found to be 87.0% and 85.2%, respectively (Fig. 1A and B). The number of amino acid residues of mouse ZHX2 is one amino acid residue shorter than that of human ZHX2 (Kawata et al., 2003). In contrast, mouse ZHX3 is 5 amino acid residues shorter than that of human ZHX3 (Yamada et al., 2003). The two Znf-and five HD-motifs of ZHX2 and ZHX3 are conserved among human, rat and mouse (Kawata et al., 2003; Yamada et al., 2003). The amino acid sequence of the Znf-motifs, HD1, the proline-rich region and HD2 of ZHX2 were identical between these species, but that of HD3 and HD4 were slightly different. The similarity of HD5 of mouse ZHX2 with human was 76.7%. In the case of ZHX3, the HD1, HD2 and HD4 were identical for two species. Similarities in the Znf-motifs, HD3 and HD5 were slightly different from each other. The dimerization domains and the interaction domains with the ADs of NF-YA of the ZHX family proteins are conserved between these two species (Yamada et al., 1999b, 2003; Kawata et al., 2003). The similarity in the amino acid sequence of ZHX2 and ZHX1 is higher than that of ZHX3 and ZHX1. In addition, an analysis of their phylogenetic relationship also shows that ZHX1 and ZHX2 are very similar but that ZHX3 is different (Fig. 1C). We recently reported that the mouse *ZHX1* gene consists of five exons, and that the fourth exon encodes a portion of the 5′-noncoding sequence, an entire coding sequence, and a portion of the 3′-noncoding sequence (Fig. 2 and (Shou et al., 2003)). In the case of the mouse *ZHX2* gene, the entire coding sequence is also encoded by a single exon, as is the case for the mouse *ZHX1* gene. In contrast, the coding sequences of ZHX3 were intervened by an intron. In addition, the human *ZHX1* and *ZHX2* genes are located on chromosome 8q region and the two genes are separated by only 396-kb, but the human *ZHX3* gene is located in the chromosome 20q region (Kawata et

al., 2003; Yamada et al., 2003). Therefore, a common ancestral gene was duplicated to two genes, for example, the *ZHX1* and *ZHX3* genes. The *ZHX1* gene was then duplicated again to the *ZHX1* and *ZHX2* genes. Since only the *ZHX3* gene of the ZHX family is intervened by an intron within the coding sequence, the *ZHX3* gene may be evolutionally older than other two genes.

ZHX2 and ZHX3 form a heterodimer both in vivo and in vitro via the HD1-containing region not HD1 alone (Fig. 4 and Table 1). Thus, we conclude that the ZHX family proteins, ZHX1, ZHX2 and ZHX3 form not only a homodimer but also a heterodimer (Hirano et al., 2002; Kawata et al., 2003; Yamada et al., 2003). The dimerization of ZHX1 with ZHX1 or ZHX3 is important for the full activity of the transcriptional repression of ZHX1 or ZHX3 (Yamada et al., 2002, 2003; Kawata et al., 2003). The heterodimerization of ZHX2 with ZHX3 may affect the repressor activity of both ZHX2 and ZHX3. The ZHX family proteins also interact with the ADs of NF-YA. We recently reported that ZHX2 intrinsically interacts with NF-YA in HEK293 cells and represses the promoter activity of the *cdc25C* gene, which is stimulated by NF-Y in *Drosophila* SL2 cells (Kawata et al., 2003). It has been reported that the cdc25C protein regulates cell-cycle progression into the M phase and the *cdc25C* gene transcription occurs late in the S/G2 phase (Gautier et al., 1991). ZHX proteins are expressed in various tissues of human, rat and mouse (Fig. 3 and (Barthelemy et al., 1996; Yamada et al., 1999a, 2003; Hirano et al., 2002; Kawata et al., 2003)). Therefore, ZHX proteins may participate in the expression of a number of NF-Y-regulated genes and be involved in the cell growth and differentiation of a variety of tissues. In addition, some co-factors and transcription factors were cloned as ZHX1-interacting proteins from rat liver and granulosa cell cDNA libraries (Yamada et al., 2003). These data suggest that ZHX proteins also interact with other transcription factors to play important roles in the cell.

In summary, we report on the cloning of mouse ZHX2 and ZHX3 cDNAs and the corresponding genes and the heterodimerization of ZHX2 with ZHX3 proteins by in vivo and in vitro assays. Which genes are regulated by ZHX proteins, which proteins interact with the monomer or homo- or heterodimer of ZHX proteins, and which mechanism regulates the formation of these complexes? These interesting issues will be clarified in future studies.

## Acknowledgements

We are grateful to Kaoru Matsuura and members of the Miyamoto Laboratory for helpful discussions and technical assistance. This investigation was supported by grants from the Ichiro Kanehara Foundation, Ministry of Education, Science, Sports and Culture of Japan, and CREST of JST (Japan Science and Technology).

## References

Barthelemy, I., Carramolino, L., Gutierrez, J., Barbero, J.L., Marquez, G., Zaballos, A., 1996. zhx-1: a novel mouse homeodomain protein containing two zinc-fingers and five homeodomains. Biochem. Biophys. Res. Commun. 224, 870–876.

Gautier, J., Solomon, M.J., Booher, R.N., Bazan, J.F., Kirschner, M.W., 1991. cdc25 is a specific tyrosine phosphatase that directly activates p34cdc2. Cell 67, 197–211.

Hirano, S., Yamada, K., Kawata, H., Shou, Z., Mizutani, T., Yazawa, T., Kajitani, T., Sekiguchi, T., Yoshino, M., Shigematsu, Y., Mayumi, M., Miyamoto, K., 2002. Rat zinc-fingers and homeoboxes 1 (ZHX1), a nuclear factor-YA-interacting nuclear protein, forms a homodimer. Gene 290, 107–114.

Kawata, H., Yamada, K., Shou, Z., Mizutani, T., Yazawa, T., Yoshino, M., Sekiguchi, T., Kajitani, T., Miyamoto, K., 2003. Zinc-fingers and homeoboxes (ZHX) 2, a novel member of the ZHX family, functions as a transcriptional repressor. Biochem. J. 373, 747–757.

Mount, S.M., 1982. A catalogue of splice junction sequences. Nucleic Acids Res. 10, 459–472.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shou, Z., Yamada, K., Inazu, T., Kawata, H., Hirano, S., Mizutani, T., Yazawa, T., Sekiguchi, T., Yoshino, M., Kajitani, T., Miyamoto, K., 2003. Genomic structure and promoter analysis of the mouse zinc-fingers and homeoboxes 1 (*ZHX1*) gene. Gene 302, 83–94.

Yamada, K., Printz, R.L., Osawa, H., Granner, D.K., 1999a. Human ZHX1: cloning, chromosomal location, and interaction with transcription factor NF-Y. Biochem. Biophys. Res. Commun. 261, 614–621.

Yamada, K., Osawa, H., Granner, D.K., 1999b. Identification of proteins that interact with NF-YA. FEBS Lett. 460, 41–45.

Yamada, K., Kawata, H., Matsuura, K., Shou, Z., Hirano, S., Mizutani, T., Yazawa, T., Sekiguchi, T., Yoshino, M., Kajitani, T., Miyamoto, K., 2002. Functional analysis and the molecular dissection of zinc-fingers and homeoboxes 1 (ZHX1). Biochem. Biophys. Res. Commun. 297, 368–374.

Yamada, K., Kawata, H., Shou, Z., Hirano, S., Mizutani, T., Yazawa, T., Sekiguchi, T., Yoshino, M., Kajitani, T., Miyamoto, K., 2003. Analysis of zinc-fingers and homeoboxes (ZHX) 1-interacting proteins: molecular cloning and characterization of member of the ZHX family, ZHX3. Biochem. J. 373, 167–178.