Characterization of PREP2, a paralog of PREP1, which defines a novel sub-family of the MEINOX TALE homeodomain transcription factors

C. Fognani, C. Kilstrup-Nielsen[1], J. Berthelsen, E. Ferretti, V. Zappavigna[1] and F. Blasi*

Unit of Molecular Genetics and [1]Unit of Transcriptional Regulation of Development, DIBIT, Department of Molecular Biology and Functional Genetics, Università Vita Salute San Raffaele, Via Olgettina 58, 20132 Milano, Italy

ABSTRACT

TALE (three amino acid loop extension) homeodomain proteins include the PBC and the MEINOX sub-families. MEINOX proteins form heterodimer complexes with PBC proteins. Heterodimerization is crucial to DNA binding and for nuclear localization. PBC–MEINOX heterodimers bind DNA also in combination with HOX proteins, thereby modulating their DNA-binding specificity. TALE proteins therefore play crucial roles in multiple developmental and differentiation pathways *in vivo*. We report the identification and characterization of a novel human gene homologous to *PREP1*, called *PREP2*. Sequence comparisons indicate that PREP1 and PREP2 define a novel sub-family of MEINOX proteins, distinct from the MEIS sub-family. *PREP2* is expressed in a variety of human adult tissues and displays a more restricted expression pattern than *PREP1*. PREP2 is capable of heterodimerizing with PBC proteins. Heterodimerization with PBX1 appears to be essential for nuclear localization of both PREP2 and PBX1. A comparison between the functional properties of PREP1 and PREP2 reveals that PREP2–PBX display a faster DNA-dissociation rate than PREP1–PBX heterodimers, suggesting different roles in controlling gene expression. Like PREP1, PREP2–PBX heterodimers are capable of forming ternary complexes with HOXB1. The analysis of some PREP2 *in vitro* properties suggests a functional diversification among PREP and between PREP and MEIS MEINOX proteins.

INTRODUCTION

The homeodomain is a three-helix DNA-binding domain formed by 60 amino acid residues and is present in a large number of transcription factors. Within this large group, the TALE (three amino acid loop extension) family of homeodomain proteins is characterized by the presence of three additional amino acids between the first and second helix of the homeodomain. The TALE family is divided into two sub-families, the PBC and the MEINOX. The PBC sub-family includes PBX1, 2, 3 and 4, the *Caenorhabditis elegans* CEH-20 and the *Drosophila* EXTRADENTICLE (EXD) gene products. The MEINOX sub-family includes the mammalian MEIS1, 2 and 3, the human and mouse PREP1, the *C.elegans* CEH-25, the *Drosophila* HOMOTHORAX (HTH) and other MEIS orthologous gene products from *Dario rerio*, *Xenopus laevis* and yeast (1). This sub-division has both a structural and functional basis, since members of one sub-family form stable DNA-binding heterodimers with members of the other. The interaction of MEINOX and PBC proteins is DNA independent, requires specific conserved regions in the respective N-termini, and is compatible with the interaction of the PBC homeodomain with anterior HOX proteins (2–6). The use of different interaction surfaces allows PBC proteins to act as a bridge between MEINOX and HOX proteins and to form transcriptionally active ternary complexes (4,7–9). In addition, some TALE proteins can also interact directly with posterior HOX (10,11).

Heterodimerization between MEINOX and PBC proteins is required to prevent the active export of the PBX1 and EXD proteins from the nucleus (12–14). However, while the MEIS gene product is able to directly reach the nucleus, the PREP1 protein is missing a nuclear localization signal and is only translocated to the nucleus by interaction with PBX (7).

The sequence similarities within the MEINOX sub-family are essentially confined to the homeodomain and to the two N-terminal regions, involved in protein–protein interactions with PBC proteins, termed MH or HR domains (3,5,7,15).

PREP1 and PBX form a stable complex in solution, interacting through sequences within the N-terminal part of both proteins; the interaction does not require the respective homeodomains. In addition, complex formation precedes DNA binding, as deletion of the N-terminal part of either protein disrupts co-operative DNA binding of the PREP1–PBX complex (5). Due to the different interaction surfaces used in the PBX–HOX and PREP1–PBX complexes, interaction of PREP1 with PBX does not exclude further interaction with HOX proteins and, indeed, a ternary PREP1–PBX–HOX

complex can be observed to bind a PBX–HOX DNA target. This ternary complex has a higher transcriptional activity compared to a PBX–HOX heterodimer (4,8). In contrast to PREP1, MEIS1 can also directly bind to some of the most posterior HOX proteins, like HOXA9 (10,11), through its C-terminal sequence. This apparently results in a functional difference between PREP1 and MEIS1, since only the latter is able to accelerate HOXA9-dependent acute myeloid leukemia in mice (16).

Prep1 is ubiquitously expressed in adult tissues, and is expressed in the embryos throughout development (6). Conversely, the MEIS-family proteins seem to have a specific pattern of expression during development. MEIS2 is expressed in the developing nervous system, limbs and face and in various viscera (17,18).

Here we describe the identification and cloning of a novel MEINOX family member displaying a high degree of similarity with *PREP1*, which we have termed *PREP2*. *PREP2* is located on chromosome 11 and is expressed in a variety of human adult tissues and displays a more restricted expression pattern than *PREP1*. The analysis of the protein sequence of PREP2 indicates that PREP1 and PREP2 form a distinct structural group within the MEINOX family defining a novel sub-family of proteins different from that including MEIS and HTH. PREP2 is capable of forming heterodimers with PBC proteins and heterodimerization with PBX1 appears to be essential for nuclear localization of both PREP2 and PBX1. PREP2, like PREP1, can form ternary complexes with PBX1 and HOXB1. A comparison between PREP1 and PREP2 functional properties reveals that PREP2–PBX complexes display a faster DNA-dissociation rate than PREP1–PBX heterodimers, suggesting different roles in controlling gene expression. Our results indicate the existence of two functional groups within the MEINOX family of TALE proteins, represented by the PREP and MEIS sub-families. Furthermore, our analysis of some of the PREP2 functional properties *in vitro* suggests a diversification of activities among PREP proteins.

MATERIALS AND METHODS

Cloning of human PREP2

PREP2 cDNA was cloned by PCR using oligonucleotide primers generated as described in the Results section. Oligo 2 and oligo 3 were used to amplify (94°C 30 s, 60°C 1 min, 72°C 3 min for 35 cycles) a fragment of 740 bp from a double-stranded cDNA reverse-transcribed from a human fetal mRNA (Clontech 7438-1) in a 50 µl reaction containing 0.5 ng cDNA, 10 pmol each primer, 0.2 mM dNTPs and 1.5 U high fidelity *Taq* polymerase (Boehringer-Mannheim). The 740 bp fragment was cloned directly in a T-vector (Invitrogen K4500-01) and sequenced. Based on this sequence we designed two specific primers (oligos 4 and 5) to clone the full length *PREP2* cDNA by 5′- and 3′-rapid amplification of cDNA ends (5′- and 3′-RACE) according to the manufacturer's protocol (Clontech 7438-1). The 3025 bp long *PREP2* cDNA was cloned into pCR2.1 expression vector for transcription–translation studies and into vectors containing a C-terminal FLAG for transfection in mammalian (pSG-FLAG-Prep2) and *Drosophila* cells (pPAC-FLAG-Prep2).
Primers list:
oligo 2: 5′-TGT GAA CAG GCC ACC AGG GGC-3′

oligo 3: 5′-ACT AAG CAT GGG CTG CAG GAT-3′
oligo 4: 5′-TGG CGG AGG TGA TGC ACT CAG-3′
oligo 5: 5′-CAA CTG GTT CAT CAA TGC CCG-3′

In vitro transcription–translation

All pCR2.1 expression vectors were translated *in vitro* using the coupled TNT transcription/translation system (Promega) as described previously (5). PREP2 and PBX or PREP1 and PBX were co-translated (plasmids in equimolar amounts).

Electrophoretic mobility shift assay (EMSA)

EMSA was performed as described previously (5). One micro-liter of nuclear extract (5 µg of proteins) or 2 µl of reticulocyte lysate, containing a combination of *in vitro* translated PREP1, PREP2, PBX1 and HOXB1, were mixed with 18 µl of PPH binding buffer (10 mM Tris–HCl pH 7.5, 75 mM NaCl, 1 mM EDTA, 6% glycerol, 3 mM spermidine, 1 mM DTT and 0.5 mM PMSF) in the presence of 0.5 µg poly(dI·dC), 30 000 c.p.m. ^{32}P-labeled oligonucleotides and, when necessary, antibody or unlabeled competitors. After 30 min on ice, the reactions were separated by a 5% polyacrylamide gel in 0.5× TBE. For the analysis of the dissociation rate, labeled probe was incubated with *in vitro* translated proteins for 10 min on ice, then a 500-fold molar excess of unlabeled competitor oligonucleotide was added to the reactions. Samples from the same binding reactions were taken at different times and immediately loaded on native gels. The oligonucleotide probes used in EMSA were: O1, 5′-CACCTGAGAGTGACAGAAGGAAGGCAGG-GAG-3′; b2PP2, 5′-GGAGCTGTCAGGGGGCTAAGATTG-ATCGCCTCA-3′. O1 contains one binding site for PREP1–PBX complex from the urokinase enhancer (19). b2PP2 contains both a PREP–MEIS and a PBX–HOX site from the *Hoxb2* r4 enhancer (8).

Northern blot analysis

Two commercially available human northern blots (Clontech 7759-1 and 7760-1) were hybridized with a ^{32}P-labeled probe generated by PCR, from the expression vector pCR2.1 containing the *PREP2* full length cDNA. The probe included only the 3′-untranslated region from nucleotide 1380 to 3025. Hybridization was carried out according to the manufacturer's protocol.

Transfections

NIH 3T3 fibroblasts were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% newborn calf serum and antibiotics. Cells were transfected with 5 or 10 µg of pSG-Pbx1a and pSG-FLAG-Prep2 by CaPO$_4$ precipitation in 10 cm dishes. Cells were reseeded after 24 h on Chambers Slides (Nunc) and fixed after 16 h for immuno-cytochemistry. *Drosophila* SL-2 Schneider cells were grown in *Drosophila* medium (Gibco) at 25°C and transfected with calcium phosphate, using 5 µg each of pPAC-FLAG-Prep2 and pPAC-Pbx1a. The *PBX1a* vectors have been described before (13). COS-7 cells were maintained in DMEM, supplemented with 10% fetal calf serum (Gibco), 100 IU/ml of penicillin and 100 µg/ml streptomycin. In a typical transfection experiment, 8 µg of reporter plasmid, 4–8 µg of expression construct and 0.2 µg of pCMV-β-gal (Clontech) as an internal control, were used per 10 cm dish. Forty-eight hours after transfection, cells were washed and lysed directly on the plate with

a solution containing 1% Triton X-100, 25 mM Glycil-Glycine pH 7.8, 15 mM $MgSO_4$, 4 mM EGTA and 1 mM DTT. Extracts were collected, centrifuged to clear the supernatant and assayed for luciferase and β-galactosidase expression as described by Di Rocco *et al.* (20).

Antibodies and immunocytochemistry

For immunocytochemistry, cells were fixed in 100% methanol, rehydrated in PBS, blocked in PBS and 1% BSA, incubated for 1 h with the primary antibody, washed in PBS and 1% BSA, and mounted for examination with an Olympus Provis fluorescence microscope. Rabbit anti-PBX antibodies C20 were obtained from Santa Cruz Biochemicals (Santa Cruz, CA), FLAG epitope-tagged PREP2 was visualized with mouse anti-Flag monoclonal M2 (Sigma Chemical Co., St Louis, MO). Secondary antibodies were fluorescein isothiocyanate antimouse and TRITC-anti rabbit IgG from Jackson Immuno Research (West Grove, PA).

RESULTS

Identification and structural characterization of a novel *PREP1* paralogous gene

We have searched for *PREP1* homologous genes in the human genome databank. To avoid finding a member of the *MEIS* sub-family, we looked for homologies not confined to specific domains, but extending throughout the entire length of *PREP1*. The *PREP1* gene maps on the sequenced chromosome 21, and its exon–intron structure is known (21) and conserved in the mouse gene (our unpublished data). This information allowed us to circumvent the problem represented by the possible presence of introns in the genomic sequence of putative *PREP1* homologs. We searched the data bank for genomic sequences having a high level of identity to individual exons of *PREP1*. The similarities found covered the entire *PREP1* cDNA, with the exception of the C-terminus, and were all present on a sequenced, but not totally mapped, 200 000 bp fragment of human chromosome 11. This putative gene was called *PREP2*. Interestingly, all the homologies were compatible with the exon–intron structure of the *PREP1* gene. We reconstructed a putative *PREP2* cDNA sequence that we used to synthesize suitable oligonucleotide primers (see Materials and Methods) to amplify a double stranded cDNA reverse transcribed from human fetal mRNA. We next isolated the full-length *PREP2* cDNA by 5′- and 3′-RACE. The *PREP2* cDNA was completely sequenced and found to be 3025 bp long (deposited in the EMBL data bank, accession no. HX1003667103). In parallel, we also found that a 3175 bp cDNA sequence, AK023136, had been deposited in the EMBL data bank. This sequence is almost identical to that of our *PREP2* cDNA, except for a significant sequence divergence in the coding region, since it lacks a CT dinucleotide (position 895–896) causing a premature termination of the open reading frame (data not shown).

The amino acid sequence of PREP2 deduced from the cDNA is reported in Figure 1A in comparison with PREP1. PREP2 is longer than PREP1 (461 versus 436 residues), mainly because of small insertions at different positions. Overall, the amino acid residues in the two proteins are 52% identical with some conservative substitutions raising the overall similarity to 60%

(Table 1). The homeodomains and the two regions of homology with MEIS (HR1 and HR2) share the highest level of similarity. The N-terminal 50 residues share 36% identity, whereas the extreme C-termini are largely divergent. Interestingly, there is a 10 amino acid insertion in PREP2 between residues 172 and 176 of PREP1. Since this general region is involved in the formation of heterodimers with PBX proteins, this difference might lead to qualitative and/or quantitative differences in heterodimer formation.

Alignment of the *PREP2* cDNA sequence with the previously identified 200 kb chromosome 11 region allowed us to define the intron–exon organization and to determine the size of most human *PREP2* introns (Fig. 1B). As expected, *PREP1* and *PREP2* share the same exon–intron organization. In Figure 1B, in some cases the size of the introns in the *PREP2* gene is not indicated because the sequence of some subfragments of the 200 kb region lacked overlapping verification.

The two PREP proteins share amino acid sequence similarity with MEIS1, MEIS2, MEIS3 and HTH. However, PREP1 and PREP2 lack similarity to any MEIS family members in the 55 most N-terminal residues, in the region between HR2 and the homeodomain (residues 244–264 and 328–407 of PREP1), and in the C-terminus from amino acid 339 to the end. Conversely, in these same three regions, PREP1 and PREP2 display a high degree of similarity. Generation of a phylogenetic tree using the CLUSTALX and NJPLOT software (22) confirms that PREP1 and PREP2 define a novel different sub-family distinct from that including MEIS and HTH (Fig. 1C).

The *PREP2* gene is expressed in a variety of human adult tissues

We next tested whether *PREP2* mRNA was expressed in human tissues by performing a northern blot hybridization on a commercially available (Clontech) array of RNAs extracted from several different human tissues as indicated in Figure 2. A single band of ~4 kb was observed in some, but not all tissues (Fig. 2). In particular, the expression was strong in mRNA extracted from heart, brain, skeletal muscle and ovary. A lower expression level was observed in placenta, lung, pancreas, prostate, testis and small intestine. No band was visible in mRNA extracted from liver, kidney, spleen, thymus, colon and peripheral blood leukocytes. The amount of $poly(A)^+$ RNA on this blot was adjusted by the manufacturer (Clontech) so that the β-actin hybridization signal was of comparable intensity in every lane.

PREP2 displays similarities and differences in DNA-binding and transcriptional regulating activity compared to PREP1

In vitro synthesized PREP2 has a migration rate corresponding to an apparent molecular weight of 70 kDa, thus somewhat slower than PREP1, in agreement with its higher molecular weight (463 versus 436 amino acids of PREP1), but also slower than expected from its calculated molecular weight (data not shown). PREP2 behaves in this respect as PREP1, displaying an apparent molecular weight higher than expected from its amino acid composition (5).

To assess the DNA-binding activity of PREP2 and to compare it with PREP1, we used the O1 oligonucleotide. We previously showed that O1 contains a single binding site for PREP1–PBX complexes (5). We performed binding studies in the presence and absence of PBX1a. Neither PREP2 nor
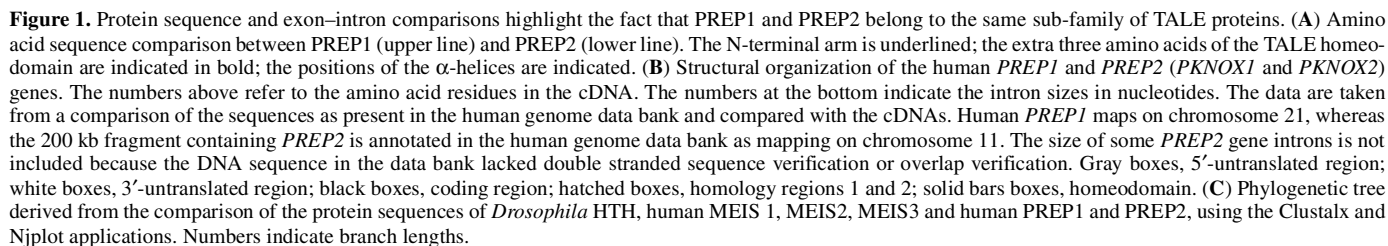
A

```
  1 MMATQTLSID SYQDGQQM_Q VVTELKTEQD PNCSEPDAEG VSP_P__PVE SQTPMDVDKQ  56
    ^|||     |||  ||         |   | |    | |  |   |      |
  1 MMATQNVPPP PYQDSPQMTA TAQPPSKAQA VHISAPSAAA STPVPSAPID PQAQLEADKR  60


 57 AIYRHPLFPL LALLFEKCEQ STQGSEGTTS ASFDVDIENF VRKQEKEGKP FFCEDPETDN 116
    | |||||||| | |||||||| |||||  ||  |||||||||| | || ||  ||  ||| ||
 61 AVYRHPLFPL LTLLFEKCEQ ATQGSECITS ASFDVDIENF VHQQEQEHKP FFSDDPELDN 120


117 LMVKAIQVLR IHLLELEKVN ELCKDFCSRY IACLKTEMNS ETLLSGEPGS PYSPVQSQQI 176
    |||||||||| |||||||||| |||||||  | | |||| | |   ||    |  ||||| |
121 LMVKAIQVLR IHLLELEKVN ELCKDFCNRY ITCLKTKMHS DNLLRNDLGG PYSPNQPSIN 180


177 QSAITGTISP QGIVVPASAL QQGNVAMATV AGGTVYQPVT VVTPQGQVV_ ____TQ__T_ 228
              |         ||| |    |        |   |    |  |||          |
181 LHSQDLLQNS PNSMSGVSNN PQGIVVPA_S ALQQGNIAMT __TVNSQVVS GGALYQPVTM 237

                                                            HELIX 1
229 _LSPG___TI RI_Q_NSQLQ _LQLNQDL_S ILHQDDGSSK NKRGVLPKHA TNVMRSWLFQ 280
     | |     |   | |   |  | |  | ||  |     ||  |||||||||| || |||||||
238 VTSQGQVVTQ AIPQGAIQIQ NTQVNLDLTS LLDNEDKKSK NKRGVLPKHA TNIMRSWLFQ 297

            HELIX 2                    HELIX 3
281 HIGHPYPTED EKKQIAAQTN LTLLQVNNWF INARRRILQP MLD_SSCSET PKTKKKTAQN 339
    |  |||||||| || |||||||  |||||||||| |||||||||| |||  |  |   || |
298 HLMHPYPTED EKRQIAAQTN LTLLQVNNWF INARRRILQP MLDASNPDPA PKAKKIKSQH 357


340 RPVQRFWPDS IASGVAQPPP SELTMSEGAV VTITTPVNMN VDSLQSLSSD GATLAVQQVM 399
    || ||||| | || ||      |     ||  |        |  | |||||||| || | ||
358 RPTQRFWPNS IAAGV_____ _LQQQGGAP GT_NPDGSIN LDNLQSLSSD SATMAMQQAM 409


400 MAGQSEDESV DSTEE___D_ _____ __AGALAPAH ISGLVLANSD SLQ        436
    ||     | |  | |||   |              |        | | |  || ||
410 MA__AHDDSL DGTEEEDEDE MEEEEEEELE EEVDELQTTN VSDLGLEHSD SLEZ       461
```

B

C

Figure 1. Protein sequence and exon–intron comparisons highlight the fact that PREP1 and PREP2 belong to the same sub-family of TALE proteins. (**A**) Amino acid sequence comparison between PREP1 (upper line) and PREP2 (lower line). The N-terminal arm is underlined; the extra three amino acids of the TALE homeodomain are indicated in bold; the positions of the α-helices are indicated. (**B**) Structural organization of the human *PREP1* and *PREP2* (*PKNOX1* and *PKNOX2*) genes. The numbers above refer to the amino acid residues in the cDNA. The numbers at the bottom indicate the intron sizes in nucleotides. The data are taken from a comparison of the sequences as present in the human genome data bank and compared with the cDNAs. Human *PREP1* maps on chromosome 21, whereas the 200 kb fragment containing *PREP2* is annotated in the human genome data bank as mapping on chromosome 11. The size of some *PREP2* gene introns is not included because the DNA sequence in the data bank lacked double stranded sequence verification or overlap verification. Gray boxes, 5′-untranslated region; white boxes, 3′-untranslated region; black boxes, coding region; hatched boxes, homology regions 1 and 2; solid bars boxes, homeodomain. (**C**) Phylogenetic tree derived from the comparison of the protein sequences of *Drosophila* HTH, human MEIS 1, MEIS2, MEIS3 and human PREP1 and PREP2, using the Clustalx and Njplot applications. Numbers indicate branch lengths.

Table 1. Sequence comparison of PREP1 and PREP2 proteins

Region of PREP1 considered	% identity[a]	% similarity[a]
Total protein	225/436 (52)	262/436 (60)
HR1 (50–83)[b]	24/34 (70.5)	29/34 (85)
HR2 (108–153)[b]	39/45 (87)	41/45 (91)
HR1+HR2 (50–153)[b]	82/103 (80)	91/103 (89)
Homeodomain (222–285)[b]	54/63 (86)	56/63 (89)
N-terminus (1–50)[b]	18/50 (36)	34/50 (62)
C-terminus (285–415)[b]	75/130 (58)	104/130 (80)
Extreme C-terminus (416–436)[b]	4/21 (19)	5/21 (23)

[a]Numbers refer to the number of amino acid residues of PREP2 identical or similar to those of PREP1. Numbers in parentheses express percentage identity or similarity.
[b]Numbers in parentheses indicate the amino acid residues in PREP1 and the regions considered.

Figure 2. Northern blot analysis of *PREP2* expression in human tissues. A commercial poly(A)+ mRNA filter (Clontech 7438-1) was hybridized with a specific *PREP2* probe (untranslated *PREP2* region from nucleotide 1380 to 3025). Tissues are indicated at the top.

Figure 3. PREP2 forms heterodimers with PBX proteins. (**A**) EMSA analysis with O1 oligonucleotide and *in vitro* co-translated PREP2 and PBX1a or PREP1 and PBX1a, as indicated. O1 probe contains a single binding site for PREP1–PBX. The binding was competed by a 500-fold molar excess of cold oligonucleotide and was challenged with antibodies against PREP1, PBX1a or preimmune serum (PI) as control. The reculocyte lysate contains a non-specific endogenous activity, marked by Lys. The arrows indicate the DNA-binding complexes. (**B**) Dissociation rate analysis of the dimeric PREP1−PBX1a or PREP2–PBX1a complexes. Labeled O1 oligonucletide was incubated for 10 min on ice with *in vitro* translated proteins, as indicated at the top. Then a 500-fold molar excess of unlabeled O1 was added and aliquots analyzed by EMSA after different times of incubation, as indicated. Endogenous activity is indicated (Lys). The arrows show the DNA binding complexes.

PBX1a bound DNA alone. However, PREP2 formed retarded bands with the O1 DNA when co-translated with PBX1a (Fig. 3A). The PBX1–PREP2 complex ran slightly slower than the PBX1–PREP1 included as a control, in accordance with the higher molecular weight of PREP2. The binding was competed by a 500-fold molar excess of cold oligonucleotide and was inhibited by anti-PBX1a and partially by anti-PREP1 antisera. PREP1 antibodies totally inhibited the binding of the PREP1–PBX1a heterodimer as described previously (5). A pre-immune serum had no effect on either PREP1–PBX or PREP2–PBX binding. The high amino acid similarity between PREP1 and PREP2 and the fact that the antiserum was raised against the full-length PREP1 protein justify the cross-inhibition of the anti-PREP1 antiserum on PREP2. The anti-PREP1 antiserum recognized PREP2 also in immunoprecipitation of the *in vitro* translated protein (data not shown).

Since the sequences of the PREP1 and PREP2 homeo-domains somewhat diverge in the first helix and in the first loop, but the two proteins seem to bind equally well to the O1 site, we tested whether we could point out differences in the binding dissociation rate of PREP1–PBX1a versus PREP2–PBX1a complexes. The labeled O1 oligonucleotide
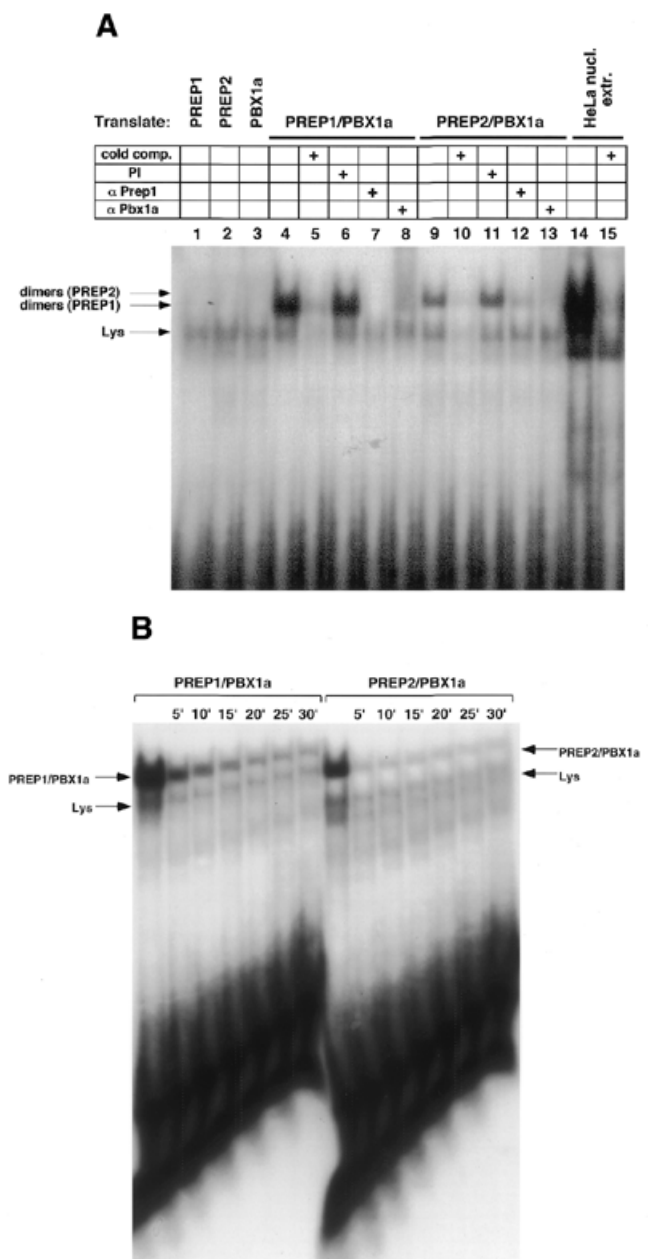
was incubated with the *in vitro* co-translated PBX1a–PREP1 or PBX1a–PREP2 protein complexes for 10 min at 4°C. Subsequently, an excess of unlabeled O1 oligonucleotide was added and the incubation was continued for various times as indicated in Figure 3B. Aliquots of the incubation mixture

were loaded on a non-denaturing polyacrylamide gel at different times. As shown in Figure 3B, the PREP2–PBX1a complex dissociated faster than the PREP1–PBX1a complex, being completely abolished in <5 min of incubation with the cold competitor. These results suggest that the sequence divergence between PREP2 and PREP1 within the homeodomain determines different functional properties of the two proteins at the DNA-binding level.

A characteristic of the PREP or MEIS proteins is to form specific ternary complexes with PBX and HOX family members on given DNA binding sites which are functionally relevant (7–9). We therefore analyzed whether PREP2 was capable of forming a ternary complex with PBX1 and HOXB1. For this, we used the b2PP2 oligonucleotide from the *Hoxb2* gene enhancer, which contains both a PBX–HOX and a PREP1–MEIS binding site and was previously shown to specifically bind a ternary PREP1–PBX–HOXB1 complex (8). As shown in Figure 4A, incubation of PREP2, HOXB1 and PBX1a with the b2PP2 oligonucleotide formed a retarded band, corresponding to a PREP2–PBX–HOXB1 ternary complex. Binding was totally competed by the unlabeled oligonucleotide. In addition, formation of the ternary complex was inhibited by antisera against PREP1, PBX1a or HOXB1. As shown in Figure 4B, PREP2 was also able to form a ternary complex with HOXB1 and PBX1b, the shorter alternatively spliced form of PBX1. Ternary complexes with PREP2 (lane 9) co-migrated with the endogenous ternary complexes found in retinoic acid treated P19 cells (lane 11) (8).

To verify possible differences between PREP1 and PREP2 in modulating the transcriptional activity of a HOX–PBX heterodimer, we chose the transactivation of the HOXB1–PBX1 complex on the *Hoxb1* autoregulatory element (*Hoxb1*-ARE) as a model system (20). We have previously shown that co-expression of PREP1 was able to further enhance the transcriptional activation of the HOXB1–PBX1 heterodimer on the *Hoxb1*-ARE (5). We therefore analyzed the effect of co-expression of PREP2 on the HOXB1–PBX1-dependent transactivation through the *Hoxb1*-ARE enhancer in transient transfections of COS cells. HOXB1 and PBX1 co-expression induced a 5–6-fold activation of transcription (20). Co-expression of PREP1 with PBX1 or HOXB1, as previously reported (5), significantly further stimulated the *Hoxb1*-ARE reporter activity (Fig. 4C). However, co-expression of PREP2 did not enhance the activation by HOXB1–PBX1, but instead led to a slight, although consistent, reduction of the HOXB1–PBX1 activity (Fig. 4C). These results further suggest the existence of differences in the functional properties of these two closely related MEINOX proteins.

PREP2 is localized to the cytoplasm and requires PBX1 for its nuclear translocation

We have previously reported that PREP1, if exogenously expressed in cultured cells, is found mainly in the cytoplasm. We found that PREP1 lacks a nuclear localization signal, and is carried into the nuclei of expressing cells by heterodimerization with PBX1 (13). We have tested for the subcellular localization of PREP2 by expressing it exogenously in NIH 3T3 mouse fibroblast and in SL2 *Drosophila* cells. As previously reported, PBX1 is found in the nuclei of exogenously expressing NIH 3T3 cells, whereas it is cytoplasmic in SL2 cells and is transported to the nucleus only upon co-expression with PREP1 (13). We

expressed a FLAG epitope-tagged PREP2 and used the anti-FLAG™ antibody to follow the distribution of the protein by immunofluorescence upon cell transfection. As shown in Figure 5, in cells transfected with a FLAG-Prep2 expression vector, PREP2 was mostly found in the cytoplasm of both NIH 3T3 and SL2 transfected cells. However, when cells were co-transfected with an expression vector for the PBX1a protein, both PBX1a and PREP2 were found exclusively in the nucleus in both cell lines. These data indicate that as in the case of PREP1, PREP2 is capable of forming heterodimer complexes with PBX1a, to be transported to the cell nucleus, and to be able to induce nuclear localization of PBX1a in SL2 cells.

DISCUSSION

In the presence of DNA, PBC and MEINOX TALE sub-class of homeoproteins interact with HOX proteins to modulate their functions (20,23–28). HOX–PBC interaction occurs only in the presence of the target DNA and requires the homeodomains and the YPWM peptide stretch N-terminal to the homeodomain (29). On the other hand, the interaction between MEIS and HOX requires the C-terminus of the MEIS proteins (11). In addition to these DNA-dependent interactions, MEINOX proteins can form heterodimers with PBC proteins (4,30,31) that are stable enough to be purified from cell extracts (19). MEINOX–PBC interactions have functional consequences since dimerization prevents the nuclear export of PBX and EXD PBC proteins (12–15). Moreover, PBX can bind at the same time both HOX proteins and TALE members, forming complexes like the heterotrimeric HOXB1–PBX1–PREP1/MEIS1 complex which appears to be essential for HOXB1-dependent HOXB2 expression in the mouse hindbrain (rhombomere 4) (7–9).

Here we report the identification and cloning of a novel MEINOX homeodomain family member which we have named PREP2. A CLUSTALX (22) comparison of the amino acid sequence of PREP2, PREP1, MEIS1, MEIS2, MEIS3 and *Drosophila* HTH reveals that PREP1 and PREP2 belong to a sub-family of MEINOX proteins different from MEIS (Fig. 1C). From now on we therefore name these sub-familes PREP and MEIS, respectively. PREP1 and PREP2, in addition to the similarities in HR1, HR2 and the homeodomain, shared with MEIS, also share three more regions of homology: one at the N-terminus, one between HR2 and the homeodomain and a third downstream of the homeodomain. PREP1 and PREP2 sequences mainly differ in the extreme C-terminus.

Northern blot analysis of *PREP2* mRNA expression in human tissues reveals that *PREP2* is expressed at higher levels in heart, brain, skeletal muscle and ovary, at lower levels in placenta, lung, pancreas, prostate, testis and small intestine, and is absent in liver, kidney, spleen, thymus, colon and peripheral blood leukocytes (Fig. 2). Comparison with published data shows that the expression domains of *PREP1* and *PREP2* are similar, but not superimposable, *PREP2* displaying a more restricted tissue and organ expression (2). Moreover, in testis an additional band of 7–8 kb is observed, which could represent a tissue-specific transcript arising from alternative splicing and/or from a different tissue-specific transcription start site.

The molecular analysis of PREP2 has shown properties that are similar, but not identical to those of PREP1. PREP2 shares
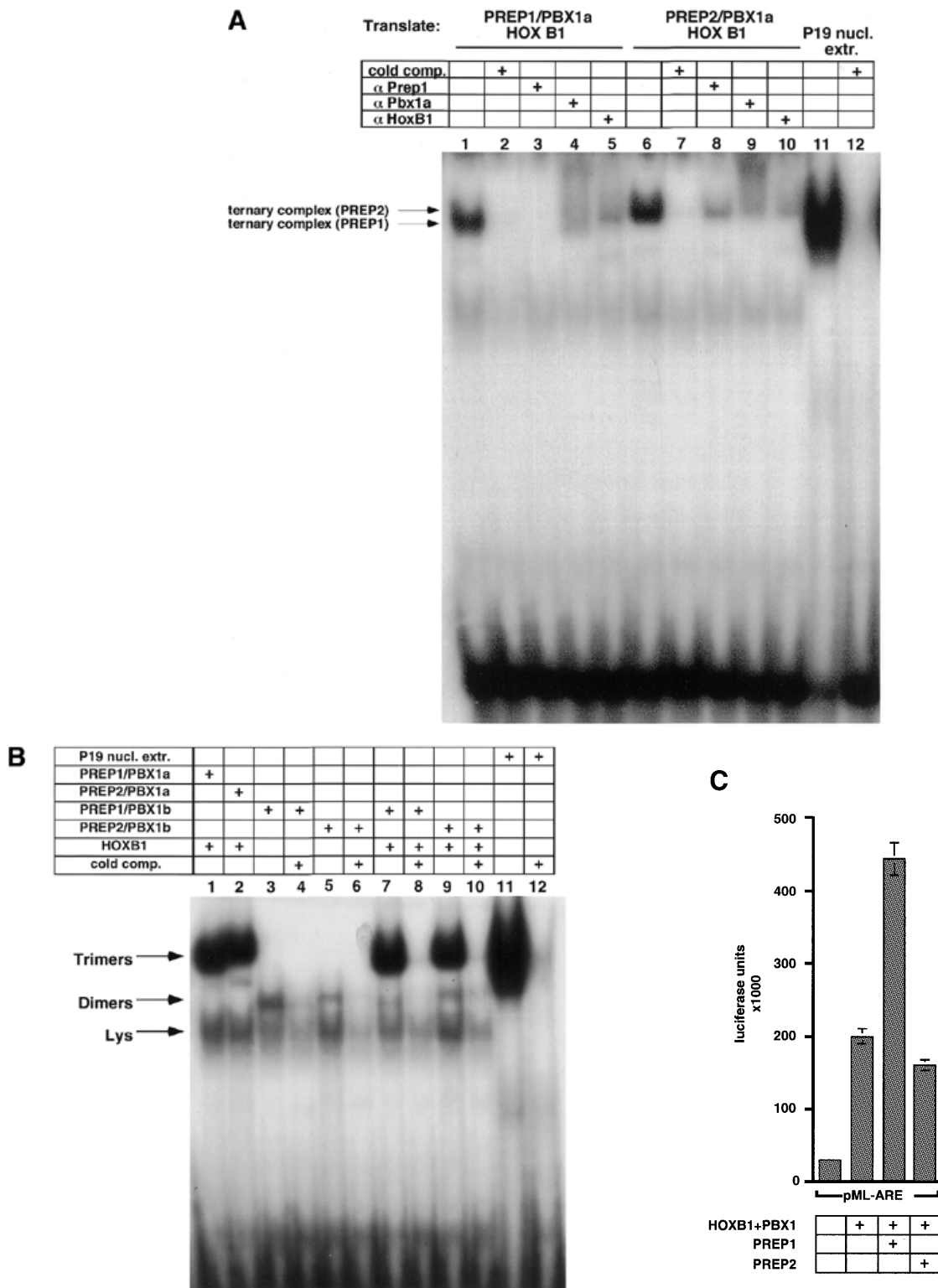
Figure 4. PREP2 can form a heterotrimeric complex with HOXB1 and PBX1. (**A**) EMSA analysis of combinations of the PREP1 or PREP2, PBX1a and HOXB1 *in vitro* translated proteins, as indicated, using b2PP2 oligonucletide as a probe. The probe contains both a PREP1–MEIS and a PBX–HOX binding site from the *Hoxb2* r4 enhancer. The binding was competed by a 500-fold molar excess of cold oligonucleotide and inhibited by anti-PREP1, anti-PBX or anti-HOXB1 antisera. Endogenous activity is indicated (Lys) and the arrows show the ternary complexes. (**B**) EMSA analysis of combinations of the PREP1 or PREP2, PBX1b and HOXB1 *in vitro* translated proteins, as indicated. b2PP2 oligonucleotide was used as probe. The EMSA was performed under similar conditions as in (A). The dimers and the ternary complexes are indicated by arrows. (**C**) Luciferase activity assayed from extracts of transiently transfected COS cells. Cells were transfected with 8 μg of the pAdMLARE reporter construct, with 4 μg of the pSGHOXB1 and 8 μg of the pSGPbx1a expression constructs as indicated. Cells were also co-transfected with 8 μg of pSGPrep1 or pSGPrep2 as indicated. An aliquot of 0.2 μg of the pCMVb-gal plasmid was co-transfected as an internal standard. Bars represent the mean luciferase activity ± SEM of at least four independent experiments.
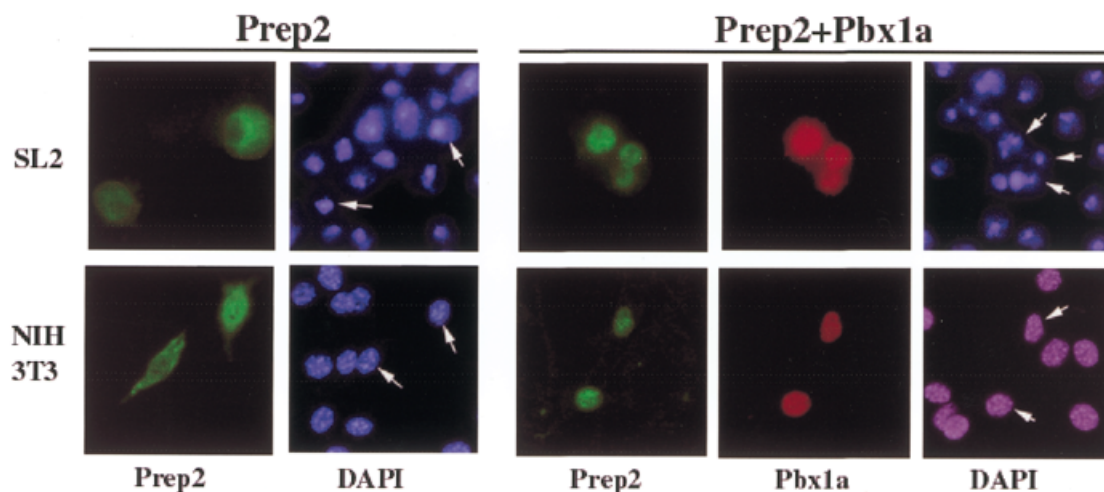
Figure 5. PREP2 requires PBX1a for nuclear localization in transfected *Drosophila* Schneider (SL2) and mouse NIH 3T3 cells. *Drosophila* Schneider (SL2) or NIH 3T3 cells were transiently transfected with a FLAG epitope-tagged PREP2 expression construct and with a PBX1a expression vector, and processed for indirect immunofluorescence with anti-FLAG and anti-PBX1a antisera. PREP2 (green) is found in the cytoplasm of expressing Schneider or NIH 3T3 cells in the absence of PBX1a. Co-expression of PREP2 and PBX1a causes nuclear localization of PREP1 (green) and PBX1a (red).

with PREP1 the ability to form complexes in solution with PBX1a and 1b, to bind DNA in association with PBX proteins and to form ternary complexes with PBX1 and HOXB1 on a *Hoxb2* promoter regulatory element. Apart from these similarities, PREP1 and PREP2 display some functional differences at the level of DNA-binding and transcriptional regulation. PREP2–PBX1 complexes show a significantly faster dissociation rate than PREP1–PBX complexes on a classical MEINOX–PBC binding consensus. This difference in DNA binding might be related to the sequence differences between PREP1 and PREP2 within the homeodomain. The third helix of the PREP1 and PREP2 homeodomains are essentially identical and in fact the two proteins recognize the same TGACAG sequence. The main differences are found in the N-terminal arm, in the first helix and in the first loop including the three extra amino acids characteristic of TALE proteins. Residues within the N-terminal arm of the homeodomain make contacts with DNA in the minor groove, whereas the three extra amino acids in the loop between the first and second helix in TALE proteins are usually required for protein–protein interactions with other homeodomain proteins (e.g. the HOX) (32,33). These sequence differences between PREP1 and PREP2 may therefore in different ways affect the stability of the DNA-bound protein complex, thus explaining the difference in dissociation rates between PREP2– and PREP1–PBX complexes. Moreover, we observed that while PREP1 increased the transcriptional activation of a HOXB1–PBX1 complex, PREP2 did not stimulate, but instead led to a slight, reproducible, reduction of the HOXB1–PBX1 hetorodimer activity. Since we could not detect differences between PREP2 and PREP1 in the efficiency of PREP–PBX or HOX–PREP–PBX complex formation *in vitro*, we believe it is unlikely that the differential capabilities of PREP1 versus PREP2 to enhance HOXB1–PBX1 activity reside on the lack of heterotrimer formation by PREP2. The observed different behaviour may rather depend on a different capability to stabilize HOX–PBX complexes on DNA, or on the lack of a transactivation domain in PREP2 which could be conversely present in PREP1. These results, together with the observation that *PREP2* expression

displays a narrower tissue distribution than the expression of *PREP1*, which was previously reported to be nearly ubiquitous (2,6), suggest that *PREP2* and *PREP1* may perform different functions. Further experiments, including the targeted inactivation of the two genes, are needed to clarify this issue.

We previously reported that in cultured cells exogenously expressed PREP1 is found mainly in the cytoplasm. Formation of a heterodimeric complex with PBX is required to induce the nuclear localization of PREP1, which lacks a nuclear localization signal (13). We have observed that upon transfection of NIH 3T3 mouse fibroblast and in SL2 *Drosophila* cells PREP2 also remains largely cytoplasmic. Only upon co-expression with PBX1, PREP2 is relocalized to the nucleus, as observed for PREP1 (13). As previously reported, PBX1 is found in the nuclei of exogenously expressing NIH 3T3 cells, whereas it is cytoplasmic in SL2 cells and is transported to the nuclei of these only upon co-expression with PREP1 (13). Similarly PREP2 expression in SL2 cells induces nuclear localization of PBX1, indicating that as in the case of PREP1, PREP2 is capable of forming heterodimer complexes with PBX1, which are subsequently transported to the cell nucleus.

Since PREP and MEIS belong to separate sub-families of the MEINOX proteins, they may play different functional roles. The molecular analysis revealed that both MEIS and PREP proteins form heterodimers with PBX, form ternary complexes with HOXB1 and are required to prevent PBX export from the nucleus. In *Drosophila*, mutations that alter the homeodomain of the *Drosophila* MEIS paralog, HTH, cause only partial homeotic transformation, whereas mutations that alter the HR1/HR2 site modify the nuclear localization of EXD, prevent the formation of ternary complexes and have more profound effects on fly development (15,34). In *Drosophila*, the function of HTH can be complemented by mammalian MEIS1 (15,31). The same experiment has not been carried out with PREP; however, overexpression of PREP1 in *Drosophila* induces an homeotic phenotype, similar but not superimposable to HTH overexpression. The latter is characterized by the absence of the eye and by leg and antenna alterations, whereas in the over-expression of *PREP1* the effect on the eye is weaker (14).

Furthermore, a major functional difference between MEIS and PREP proteins observed so far is in the acceleration of HOX-dependent leukemogenesis, as *PREP1* was found unable to substitute for *Meis1* to reinforce *HoxA9* transforming activity in bringing about leukemic transformation (16). This result was interpreted to be due to the inability of *PREP1* to bind posterior HOX proteins, even though this has not yet been directly tested. In fact, PREP proteins have a totally divergent sequence at the C-terminus, which in MEIS1 is involved in the binding of the HOXA9 protein (11). Further experiments *in vitro* and *in vivo* are required to fully understand the different functional roles played by the two MEINOX sub-families.

ACKNOWLEDGEMENTS

REFERENCES

1. Burglin,T. (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res.*, **25**, 4173–4180.
2. Chen,H.M., Rossier,C., Nakamura,Y., Lynn,A., Chakravarti,A. and Antonarakis,S.E. (1997) Cloning of a novel homeobox-containing gene, PKNOX1, and mapping to human chromosome 21q22.3. *Genomics*, **41**, 193–200.
3. Knoepfler,P.S., Calvo,K.R., Chen,H., Antonarakis,S.E. and Kamps,M.P. (1997) Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. *Proc. Natl Acad. Sci. USA*, **94**, 14553–14558.
4. Berthelsen,J., Zappavigna,V., Mavilio,F. and Blasi,F. (1998) Prep1, a novel functional partner of Pbx proteins. *EMBO J.*, **17**, 1423–1433.
5. Berthelsen,J., Zappavigna,V., Ferretti,E., Mavilio,F. and Blasi,F. (1998) Prep1, a novel partner of Pbx proteins, modifies Pbx-Hox protein cooperativity. *EMBO J.*, **17**, 1434–1445.
6. Ferretti,E., Schulz,H., Talarico,D., Blasi,F. and Berthelsen,J. (1999) The PBX-regulating protein PREP1 is present in different PBX-complexed forms in mouse. *Mech. Dev.*, **83**, 53–64.
7. Jacobs,Y., Schnabel,C.A. and Cleary,M.L. (1999) Trimeric association of Hox and TALE homeodomain proteins mediates *Hoxb2* hindbrain enhancer activity. *Mol. Cell. Biol.*, **19**, 5134–5142.
8. Ferretti,E., Marshall,H., Pöpperl,H., Maconochie,M., Krumlauf,R. and Blasi,F. (2000) A complex site including both Pbx-Hox and Prep-Meis-responsive elements and binding a retinoic acid-inducible ternary Hoxb1-Pbx-Prep1 complex is required for HOXB2 rhombomere 4 expression. *Development*, **127**, 155–166.
9. Ryoo,H.D., Marty,T., Casares,F., Affolter,M. and Mann,R.S. (1999) Regulation of Hox target genes by a DNA bound homothorax/Hox/extradenticle complex. *Development*, **126**, 5137–5148.
10. Shanmugam,K., Green,N.C., Rambaldi,I., Saragovi,H.U. and Featherstone,M.S. (1999) PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. *Mol. Cell. Biol.*, **19**, 7577–7588.
11. Shen,W.F., Rozenfeld,S., Kwong,A., Köm ves,L.G., Lawrence,H.J. and Largman,C. (1999) HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. *Mol. Cell. Biol.*, **19**, 3051–3061.
12. Abu-Shaar,M., Ryoo,H.D. and Mann,R.S. (1999) Control of the nuclear localization of extradenticle by competing nuclear import and export signals. *Genes Dev.*, **13**, 935–945.
13. Berthelsen,J., Kilstrup-Nielsen,C., Blasi,F., Mavilio,F. and Zappavigna,V. (1999) The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. *Genes Dev.*, **13**, 946–953.
14. Jaw,T.J., You,L.R., Knoepfler,P.S., Yao,L.-C., Pai,C.Y., Tang,C.-Y., Chang,L.-P., Berthelsen,J., Blasi,F., Kamps,M.P. and Sun,Y.H. (2000)
Direct interaction of two homeoproteins, HTH and EXD, via the conserved MH domain in HTH, is essential for EXD nuclear localization and function. *Mech. Dev.*, **91**, 279–291.
15. Pai,C.-Y., Kuo,T.-S., Jaw,T.J., Kurant,E., Chen,C.-T., Bessarab,D.A., Salzberg,A. and Sun,Y.H. (1998) The homothorax protein activates the nuclear localization of another homeoprotein, extradenticle, and suppresses eye development in *Drosophila*. *Genes Dev.*, **12**, 435–446.
16. Thorsteinsdottir,U., Kroon,E., Jerome,L., Blasi,F. and Sauvageau,G. (2001) Defining role for Hox genes and their common collaboration with MEIS1 in the induction of AML. *Mol. Cell. Biol.*, **21**, 224–234.
17. Cecconi,F., Proetzel,G., Alvarez-Bolado,G., Jay,D. and Gruss,P. (1997) Expression of Meis2, a Knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. *Dev. Dyn.*, **210**, 184–190.
18. Oulad-Abdelghani,M., Chazaud,C., Bouillet,P., Sapin,V., Chambon,P. and Dolle,P. (1997) Meis2, a novel mouse Pbx-related homeobox gene induced by retinoic acid during differentiation of P19 embryonal carcinoma cells. *Dev. Dyn.*, **210**, 173–183.
19. Berthelsen,J., Vandekerkhove,J. and Blasi,F. (1996) Purification and characterization of UEF3, a novel factor involved in the regulation of the urokinase and other AP-1 controlled promoters. *J. Biol. Chem.*, **271**, 3822–3830.
20. Di Rocco,G., Mavilio,F. and Zappavigna,V. (1997) Functional dissection of a transcriptionally active, target-specific HOX/PBX complex. *EMBO J.*, **16**, 3644–3654.
21. Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K. *et al.* (2000) The DNA sequence of human chromosome 21. *Nature*, **405**, 311–319.
22. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876–4882.
23. Phelan,M.L., Rambaldi,I. and Featherstone,M.S. (1995) Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol. Cell. Biol.*, **15**, 3989–3997.
24. Shen,W.-F., Chang,C.-P., Rozenfeld,S., Sauvageau,G., Humphries,R.K., Lu,M., Lawrence,H.J., Cleary,M.L. and Largman,C. (1996) HOX homeodomain proteins exhibit selective complex stabilities with Pbx and DNA. *Nucleic Acids Res.*, **24**, 898–906.
25. Shen,W.-F., Montgomery,J.-C., Rozenfeld,S., Lawrence,H.J., Buchberg,A. and Largman,C. (1997) The Abd-B-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol. Cell. Biol.*, **17**, 6448–6558.
26. Chan,S.K., Jaffe,L., Capovilla,M., Botas,J. and Mann,R.S. (1994) The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell*, **78**, 603–615.
27. Pöpperl,H., Bienz,M., Studer,M., Chan,S.K., Aparicio,S., Brenner,S., Mann,R.S. and Krumlauf,R. (1995) Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon EXD/Pbx. *Cell*, **81**, 1031–1042.
28. Azpiazu,N. and Morata,G. (1998) Functional and regulatory interactions between Hox and extradenticle genes. *Genes Dev.*, **12**, 261–273.
29. Knoepfler,P.S. and Kamps,M.P. (1995) The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. *Mol. Cell. Biol.*, **15**, 5811–5819.
30. Chang,C.P., Jacobs,Y., Nakamura,T., Jenkins,N., Copeland,N.G. and Cleary,M.L. (1997) Meis proteins are major *in vivo* DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol. Cell. Biol.*, **17**, 5679–5687.
31. Rieckhof,G.E., Casares,F., Ryoo,H.D., Abu-Shaar,M. and Mann,R.S. (1997) Nuclear translocation of extradenticle requires homothorax, which encodes an extra-denticle related homeodomain protein. *Cell*, **91**, 171–183.
32. Passner,J.M., Ryoo,H.D., Shen,L., Mann,R.S. and Aggarwal,A.K. (1999) Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature*, **397**, 714–719.
33. Piper,D.E., Batchelor,A.H., Chang,C.P., Cleary,M.L. and Wolberger,C. (1999) Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell*, **96**, 587–597.
34. Kurant,E., Eytan,D. and Salzberg,A. (2001) Mutational analysis of the *Drosophila* homothorax gene. *Genetics*, **157**, 689–698.