The EMBO Journal Vol.17 No.5 pp.1423–1433, 1998

Prep1, a novel functional partner of Pbx proteins

Jens Berthelsen[1,2], Vincenzo Zappavigna[3], Fulvio Mavilio[3] and Francesco Blasi[1,2,4]

[1]Dipartimento di Genetica e Biologia dei Microrganismi dell'Università, [2]Molecular Genetics Unit and [3]TIGET, DIBIT, H.S. Raffaele, via Olgettina 58, 20132, Milan, Italy

[4]Corresponding author
e-mail: blasif@dibit.hsr.it

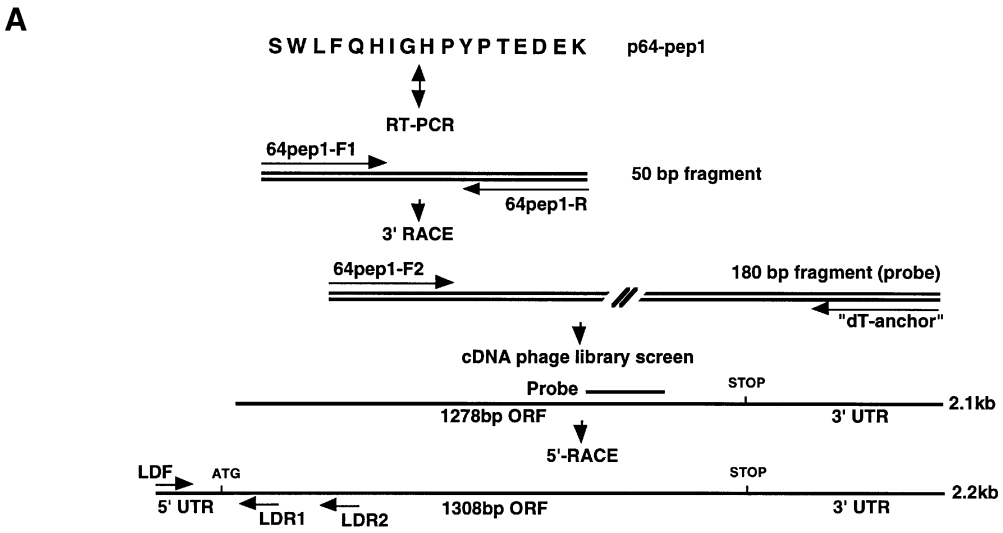
The human transcription factor, UEF3, is important in regulating the activity of the urokinase plasminogen activator (uPA) gene enhancer. The UEF3 DNA target site is a regulatory element in the promoters of several growth factor and protease genes. We reported previously that purified UEF3 is a complex of several subunits. In this paper we report the cloning of the cDNA of one of the subunits which encodes for a novel human homeodomain protein, which we have termed Prep1. The Prep1 homeodomain belongs to the TALE class of homeodomains, is most closely related to those of the TGIF and Meis1 proteins, and like these, recognizes a TGACAG motif. We further identify the other UEF3 subunit as a member of the Pbx protein family. Unlike other proteins known to interact with Pbx, Prep1 forms a stable complex with Pbx independent of DNA binding. Heterodimerization of Prep1 and Pbx results in a strong DNA binding affinity towards the TGACAG target site of the uPA promoter. Overall, these data indicate that Prep1 is a stable intracellular partner of Pbx *in vivo*.

Keywords: gene expression/homeobox/protein–protein interaction/transcription/urokinase

Introduction

Transcriptional regulation of extracellular proteases is in general a tightly controlled, and highly inducible, process. Urokinase plasminogen activator (uPA) is a serine protease and a key enzyme in the proteolytic cascade leading to fibrin degradation (Besser *et al.*, 1996). uPA is constitutively expressed in murine extra-embryonic tissue at the post-blastocyst stage, in embryonic and adult kidney and in the lung (Teesalu *et al.*, 1996), and urokinase synthesis is induced by a variety of factors in most cells both *in vivo* and *in vitro* (Besser *et al.*, 1996). Transcription of the *uPA* gene depends on the activity of an enhancer located ~2000 bp upstream of the transcriptional start site, requiring the co-operation of binding sites for three transactivator proteins: a combined PEA3–AP-1a and a downstream AP-1b site (Verde *et al.*, 1988; Rørth *et al.*, 1990; Nerlov *et al.*, 1991, 1992). The transactivator proteins ets-2 (Stacey *et al.*, 1995), Atf2–c-jun and c-jun–c-jun or c-jun–c-fos (De Cesare *et al.*, 1995) bind to these sites to activate transcription. Co-operation between these three transactivators is essential for enhancer function, and requires a 72 bp region, spanning from PEA3–AP-1a to the AP-1b site, termed COM for co-operation mediating (Nerlov *et al.*, 1992). Several nuclear proteins (urokinase enhancer factors: UEF-1 to 4) recognize partially overlapping binding sites in the COM region (Nerlov *et al.*, 1992; De Cesare *et al.*, 1996). The UEF3 binding site exists not only in the urokinase COM sequence, but also in the promoters of cytokine and chemokine genes, such as interleukin 3, *LD78/MIP-1a*, GM-CSF, of other protease genes, such as stromelysin (Berthelsen *et al.*, 1996; De Cesare *et al.*, 1997), and of the murine intracisternal A-particle (IAP) LTR (Basu *et al.*, 1997). All these genes undergo transcriptional activation by a variety of inducers, including the phorbol ester TPA, and contain AP-1 and PEA3 sites in their regulatory regions. We demonstrated previously that the purified UEF3 factor also binds to the homologous motives of the IL-3 and LD78/MIP-1α promoter (Berthelsen *et al.*, 1996). In the uPA enhancer, a mutation specifically destroying the UEF3 binding site, weakens the response of the enhancer to TPA (Nerlov *et al.*, 1992; De Cesare *et al.*, 1996). UEF3 specifically recognizes a TGACAG core sequence, and forms two retarded complexes in a electrophoretic mobility shift assay (EMSA). Purification of these complexes shows that each share a common 64 kDa protein. In addition, they contain a protein of either 50 or 40 kDa (Berthelsen *et al.*, 1996).

In this paper, we report the cloning of the cDNA encoding the 64 kDa subunit common to both complexes of UEF3, named Prep1, a novel homeodomain-containing, DNA-binding protein. The Prep1 homeodomain is structurally related to the PBC homeodomain class, contains an isoleucine residue in the third position of the conserved WF_N sequence in the homeodomain and displays a weak but specific affinity for the TGACAG sequence of the uPA promoter. The Prep1 homeodomain shows highest homology to the murine proto-oncogenic Meis protein family and to human TGIF. In addition, we also demonstrate that the 40 and 50 kDa subunits of HeLa-purified UEF3 complexes are products of the *Pbx* gene family, *Pbx1b* and *Pbx2*. These data suggest that intracellular stable Prep1–Pbx complexes exist *in vivo*. We also show the *in vitro* reconstitution of these complexes and the effect of Prep1 in modifying the affinity and DNA-binding specificity of Pbx: hence the name Prep1, for Pbx regulating protein 1. The Pbx–Prep1 complex displays strong affinity for the TGACAG motif by which it was purified, a target sequence also bound by the TGIF protein. In contrast to other homeodomain proteins known to interact with Pbx, Prep1 forms a strong and stable complex with Pbx independent of DNA binding, and does not contain the conserved YPWMX, or similar amino acid motif, required for the interaction of Hox proteins with



C

Name		Identity	Similarity	Species
	Helix 1 *Helix 2* *Helix 3*			
Prep1	SKNKRGVLPKHATNVMRSWLFQHIGHPYPTEDEKKQIAAQTNLTLLQVNNWFINARRRILQPM	**63** (100%)	**63** (100%)	Human/murine
MRG2	R-K---IF--V---I--A-----LT----S-EQ---L-QD-G--I----------V---	**45** (71%)	**55** (87%)	Human
Meis1	RHK---IF--V---I--A-----LT----S-EQ---L-QD-G--I----------V---	**44** (70%)	**54** (86%)	Murine
TGIF	KRRR--N---ESVQIL-D--YE-RYNA--S-Q--ALLSQ--H-ST---C---------L-PD-	**31** (49%)	**45** (71%)	Human/murine
Kn-1	K-K-K-K---E-RQQLL--WD--YKW---S-TQ-VAL-ES-G-D-K-I------Q-K-HWK-S	**30** (48%)	**38** (60%)	Maize
Pbx1	ARR--RNFN-Q--EILNEYFYS-LSN---S-EA-EEL-KKCGI-VS--S---G-K-I-YKKN-	**22** (35%)	**41** (65%)	Human/murine
Exd	ARR--RNFS-Q-SEILNEYFYS-LSN---S-EA-EEL-RKCGI-VS--S---G-K-I-YKKNI	**20** (32%)	**40** (63%)	Drosophila
MATα2	KPYRGHRFT-ENVRILE--FAKN-EN--LDTKGLENLMKN-S-SRI-IK--VS-R--KEKTIT	**14** (22%)	**26** (41%)	Yeast
Hox Consensus	RRR--TAYTRYQLLELEKEFHFN R-L-RRRRIEL-HSL---ER--KI--Q-R-MKWKKEN	**14** (22%)	**21** (33%)	n/a

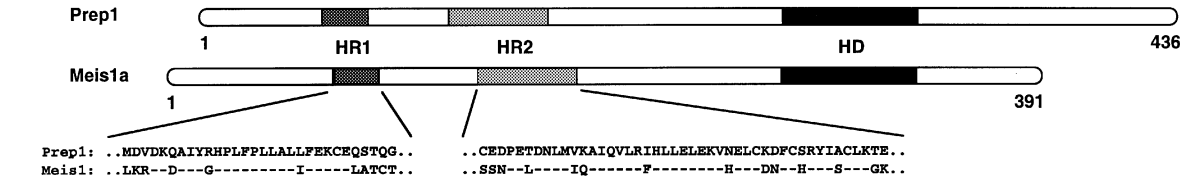


Fig. 1. Cloning, sequence and homologies of human Prep1. (**A**) Scheme of the cloning strategy of the Prep1 cDNA. From the originally obtained peptide sequence of Prep1 (p64-pep1), we designed two degenerate primers, 64pep1–F1 and 64pep1–R, corresponding to the flanking ends of p64-pep1, and used these to amplify the 50 bp cDNA fragment encoding for p64-pep1 by RT–PCR. A new primer, 64pep1–F2, identical to the middle part of the 50 bp fragment, was used together with an 'anchor' primer identical to an anchor sequence in the dT-primer used to prime the cDNA synthesis, in a 3′-RACE protocol. The resulting 180 bp fragment (primed from F2 and an internal poly(A) stretch) was used to screen a HeLa cDNA phage library. The longest clone from this screening was 2.1 kb long, but lacked the most 5′ of the ORF and 5′UTR sequences. An additional five screenings of five different cDNA phage libraries failed to reveal clones with intact 5′. We therefore used a modified 5′-RACE protocol, using primers LDR1 and -2, to clone the remaining 5′ 150 bp. The integrity of the whole cDNA was accessed by PCR on single-strand cDNA using primers LDF and a downstream primer complementary to the most 3′ sequence. (**B**) DNA sequence of Prep1 cDNA and extrapolated amino acid sequence of the 1308 bp ORF. The single letter code is employed. The underlined sequence shows the amino acid sequence of Pep1 originally obtained from the purified p64 protein band of UEF3 (Berthelsen *et al.*, 1996). (**C**) Amino acid homologies with the Prep1 homeodomain. Amino acid identity with respect to Prep1 is indicated by (–). Residue identities and similarities of the different homeodomain with respect to Prep1 are indicated as an absolute number (in bold) and in percentage (in brackets). Besides Meis1 and MRG2, the other proteins showed is no significant homology with Prep1 outside of the homeodomain. The following accession numbers or references were used: MRG2:U68285; Meis1:U33629; TGIF:X89750; KN-1:X61308; Exd:L19295; Pbx1:M86546; Matα2:P01367; Hox consensus sequence as compiled (Burglin, 1994). A sequence, termed Pknox1, encoding for a protein identical to Prep1, has recently been published as a gene mapping to chromosome 21 (U68727) (Chen *et al.*, 1997). (**D**) Alignment of full-length Prep1 and Meis1, showing the conserved positions of three homologous regions: (HR1 and HR2), homology region 1 and 2; and (HD), the homeodomain. Below: sequence alignments of HR1 and HR2. Residue identities between Prep1 and Meis1 are indicated by (–). All Meis family proteins contain these homology regions. Outside HR1, HR2 and HD there is no significant homology between the two proteins.

B

```
AGGAATTCTAGGGCACTGCGGGTCCCGATTGCTGCAGCCGCTTGTCAGTGTGATGAAGAT   60
TGGCACCCAGACACCATTCGCTTTTCACCCAAGATGATTTGATGTCTTATAAAACTCTGA  120
TGAACCATGATGGCTACACAGACATTAAGTATAGACAGCTATCAAGATGGGCAACAGATG  180
    1       M  M  A  T  Q  T  L  S  I  D  S  Y  Q  D  G  Q  Q  M
CAAGTAGTAACAGAGTTAAAGACAGAACAAGATCCAAACTGCTCTGAACCCGATGCAGAA  240
   18   Q  V  V  T  E  L  K  T  E  Q  D  P  N  C  S  E  P  D  A  E
GGAGTGAGCCCTCCCCCTGTGGAGTCTCAGACCCCGATGGATGTGGACAAGCAGGCCATT  300
   38   G  V  S  P  P  P  V  E  S  Q  T  P  M  D  V  D  K  Q  A  I
TATAGGCATCCACTATTTCCATTATTAGCTTTGTTGTTTGAAAAATGTGAACAATCTACA  360
   58   Y  R  H  P  L  F  P  L  L  A  L  L  F  E  K  C  E  Q  S  T
CAGGGCTCTGAAGGCACAACTTCTGCCAGTTTTGATGTAGACATCGAAAATTTTGTAAGA  420
   78   Q  G  S  E  G  T  T  S  A  S  F  D  V  D  I  E  N  F  V  R
AAGCAAGAGAAGGAAGGGAAACCTTTCTTTTGTGAAGATCCAGAAACCGATAATTTAATG  480
   98   K  Q  E  K  E  G  K  P  F  F  C  E  D  P  E  T  D  N  L  M
GTAAAAGCAATCCAGGTTTTGCGCATTCATCTTCTTGAGCTGGAAAAGGTTAACGAACTC  540
  118   V  K  A  I  Q  V  L  R  I  H  L  L  E  L  E  K  V  N  E  L
TGCAAAGATTTCTGCAGTCGATACATTGCTTGTCTGAAAACAGAAATGAACAGTGAAACT  600
  138   C  K  D  F  C  S  R  Y  I  A  C  L  K  T  E  M  N  S  E  T
CTGTTGAGTGGAGAGCCTGGAAGCCCGTACTCACCAGTGCAGTCCCAGCAGATTCAAAGT  660
  158   L  L  S  G  E  P  G  S  P  Y  S  P  V  Q  S  Q  Q  I  Q  S
GCCATCACAGGCACCATCAGCCCTCAGGGAATTGTGGTGCCGGCGTCCGCGCTGCAGCAG  720
  178   A  I  T  G  T  I  S  P  Q  G  I  V  V  P  A  S  A  L  Q  Q
GGAAACGTAGCCATGGCGACGGTGGCAGGTGGCACAGTGTATCAGCCTGTCACGGTCGTC  780
  198   G  N  V  A  M  A  T  V  A  G  G  T  V  Y  Q  P  V  T  V  V
ACTCCCCAAGGCCAAGTGGTCACACAGACATTGTCGCCTGGGACAATTAGGATCCAGAAC  840
  218   T  P  Q  G  Q  V  V  T  Q  T  L  S  P  G  T  I  R  I  Q  N
TCCCAGCTTCAGTTACAGTTAAACCAAGATCTCAGCATCTTGCATCAAGATGATGGTTCA  900
  238   S  Q  L  Q  L  Q  L  N  Q  D  L  S  I  L  H  Q  D  D  G  S
TCTAAGAACAAGAGGGGCGTCCTGCCAAAGCATGCCACGAACGTGATGCGGTCCTGGCTC  960
  258   S  K  N  K  R  G  V  L  P  K  H  A  T  N  V  M  R  S  W  L
TTCCAGCACATCGGGCATCCCTACCCAACAGAGGATGAGAAAAAACAGATTGCTGCTCAG 1020
  278   F  Q  H  I  G  H  P  Y  P  T  E  D  E  K  K  Q  I  A  A  Q
ACAAATTTGACACTACTCCAAGTCAACAACTGGTTCATCAATGCCAGAAGACGAATTCTT 1080
  298   T  N  L  T  L  L  Q  V  N  N  W  F  I  N  A  R  R  R  I  L
CAGCCAATGTTGGATTCAAGTTGTTCAGAGACCCCCAAAACAAAGAAAAAAACTGCTCAG 1140
  318   Q  P  M  L  D  S  S  C  S  E  T  P  K  T  K  K  K  T  A  Q
AACCGGCCAGTTCAGAGGTTTTGGCCTGATTCTATTGCATCAGGAGTCGCACAGCCACCG 1200
  338   N  R  P  V  Q  R  F  W  P  D  S  I  A  S  G  V  A  Q  P  P
CCGAGCGAGCTCACCATGTCGGAAGGAGCTGTTGTCACCATCACCACGCCCGTGAACATG 1260
  358   P  S  E  L  T  M  S  E  G  A  V  V  T  I  T  T  P  V  N  M
AACGTGGACAGCCTTCAGTCTCTGTCCTCGGACGGGGCCACCCTGGCGGTGCAGCAGGTC 1320
  378   N  V  D  S  L  Q  S  L  S  S  D  G  A  T  L  A  V  Q  Q  V
ATGATGGCAGGGCAGAGCGAGGACGAGTCTGTGGACAGCACAGAGGAGGATGCGGGTGCC 1380
  398   M  M  A  G  Q  S  E  D  E  S  V  D  S  T  E  E  D  A  G  A
CTGGCCCCTGCCCACATCAGCGGGCTGGTCTTGGCGAACAGTGACTCCCTGCAGTAGGGG 1440
  418   L  A  P  A  H  I  S  G  L  V  L  A  N  S  D  S  L  Q
CAGGAGCAGACGCACCTGACTTTTTGGAGTTTGCACAGCAAACATTTTACACAGTTTTAT 1500
TTCTAATATGTTTTTATATGTAGATATAGAAGAGTGCACTTTTGTATTTCATAGTAAGCTT 1560
AAAGCGCGTCTTTGCCGGTGCAGCGACTTCTTTCAAGTGTGTGTGTGTGTGCGTGTGT 1620
GCGTCTGTTGTGGATTTTTAAAGAAATTCTTTAAAGGTTTAACGCTAGATTGTGAGGAATG 1680
ACACACCACTCCCTCCCCACCTTGAATCCCTAATTAGATTAAGGAATAGCGCTGCCATTT 1740
TCTAAACCGTGATGCGGTTGTCACTTAGTCCTGNGGTCCCAGAAGATCTCAGTGGGCTGG 1800
TTGATTTGTTTTGGCCCATGGANNTNGAAAGAACTNCTTCAACCCGAAACCTGCCAGTGT 1860
GCGGTTGCCCNCGGCACACCGCCTAGACTGAGTGTGGTTTCCGTCGTGAGTGGATGGNCG 1920
GCAAGCTTAGCAAGCCTAAGTCCCCTCATGTTCAGTGAGCCTGTTTCATTTGCTATATAG 1980
AAAAAGAAACTCCTATTTTTACCTTGCTGGAATTATTGGATAAAAAAGCTATTTTTATAA 2040
ATTCGTTATGAATTGGATGATGACTATATTGAGGAAAAAATTTCTAGAGAAGAAACAATA 2100
CATGCTTGCTATTAATATTTCAATTTGGAATGTTCTGAATTGACCAAATTTAACGAACCT 2160
GCCCAAAGTTAGCTACCGTTCCATGGTTCTTTGCTCTCCCCCGAATTC
```

Pbx. Based on these data, we therefore propose that Prep1 is an intrinsic partner for Pbx protein function.

Results

Prep1 belongs to the PBC-like class of homeodomain proteins with extended homology to the murine proto-oncogene Meis1

UEF3 consists of three subunits, p64, p50 and p40. In a previous study, upon purification of UEF3 from HeLa cells, we obtained a peptide sequence from the p64 subunit (p64pep1, Berthelsen *et al.*, 1996). Using this 17 amino acid sequence, we generated a DNA probe of 180 bp by RT–PCR (Figure 1A; see Materials and methods) using as a template poly(A)$^+$ RNA from HeLa cells. This probe was used to screen a HeLa cell cDNA library, which, together with a final 5′-RACE screening, led to the isolation of a 2.2 kb cDNA clone for Prep1. As shown in Figure 1B, the cDNA clone contains 126 nucleotides of a 5′-untranslated sequence, after which an ATG codon is followed by an uninterrupted ORF of 1308 nucleotides, coding for a predicted 436 amino acid residue polypeptide, including the 17 residues sequenced originally (Berthelsen *et al.*, 1996). After the stop codon, the clone continues with >1000 bp of a 3′-untranslated sequence. An homology search in the DDBJ/EMBL/GenBank and NCBI data banks revealed Prep1 as a novel protein with the 60 amino acids from residues 257–317 being a typical homeodomain. The Prep1 homeodomain resembles those of the divergent TALE class of homeodomain proteins which includes the PBC class, KN-1 and MATα2 proteins (Figure 1C), and is characterized by a conserved insertion of three amino acids between helix 1 and 2 (Burglin, 1994). The Prep1 sequence has 14/60 identical residues with the homeodomain consensus (Burglin, 1994), 14/60 with the yeast Matα2 protein, 21/60 with Pbx–exd and 30/60 with plant KN-1. The highest homologies are found with the homeodomain of Meis1 (44/60), its related human and murine homologues MRG2 (45/60), and the human TGIF protein (31/60). Meis1 is a murine oncogenic protein involved in murine lymphoma formation (Moskow *et al.*, 1995), while TGIF is a repressor of RXR-dependent activation in several promoters (Bertolino *et al.*, 1995). The homology of Prep1 with Meis1 is spread all over the homeobox, but is mainly concentrated (16/18 residues) in helix 3, i.e. the DNA-recognition helix. Finally, in Prep1 the third position of the conserved WF_N sequence contains an isoleucine (as in KN-1, Meis1 and TGIF), while PBC proteins contain a glycine and the Hox consensus a glutamine. The amino acid present in this position has been shown to be important in determining the DNA-binding sequence specificity (Treismann *et al.*, 1989); thus Prep1 and PCB proteins may have different DNA specificity. Outside the homeodomain two Prep1 regions of 22 and 40 amino acids, respectively HR1 and HR2, display strong homology to similar regions in Meis-related proteins (Figure 1D). In addition, the position of HR1 and HR2 relative to the homeodomain is conserved between Prep1 and Meis proteins. These regions are not found in any of the other proteins listed in Figure 1C, while the dbEST database at NCBI contains several EST sequences containing HR1 and/or HR2 homologies. Prep1, however, is not the human homologue of Meis proteins, as we have also cloned a murine homologue of Prep1 which shows 100% homology in the homeodomain (Ferretti, Schulz, Blasi and Berthelsen, unpublished).

There is an apparent discrepancy between the observed molecular mass of Prep1 purified from HeLa cells (64 kDa) and the cDNA ORF sequence which suggests a protein of 436 residues, corresponding to 49 kDa. To investigate this point, the cDNA of Prep1 was cloned downstream of bacterial T7 promoter (pBS–Prep1), transcribed *in vitro*, and the RNA translated *in vitro* using rabbit reticulocyte lysate. As shown in Figure 2A, the protein band obtained migrates with an apparent molecular mass of 64 kDa. The additional 40 kDa band (*) appears to be due to the utilization of an internal ATG codon at position 731 in the cDNA, as shown by site-directed mutagenesis (data not shown). The apparent molecular mass of 64 kDa for Prep1 was confirmed by transient transfection of NIH 3T3 cells with a construct containing the Prep1 cDNA driven by a eukaryotic promoter (pBOSPrep1). Immunoblotting experiments with anti-Prep1 antibodies (Figure 2B) showed a weak band of ~64 kDa in the nuclear extract of untransfected NIH 3T3 cells, co-migrating with that of HeLa cells. NIH 3T3 cells transfected with pBOSPrep1 plasmid showed an increase in the appearance of the

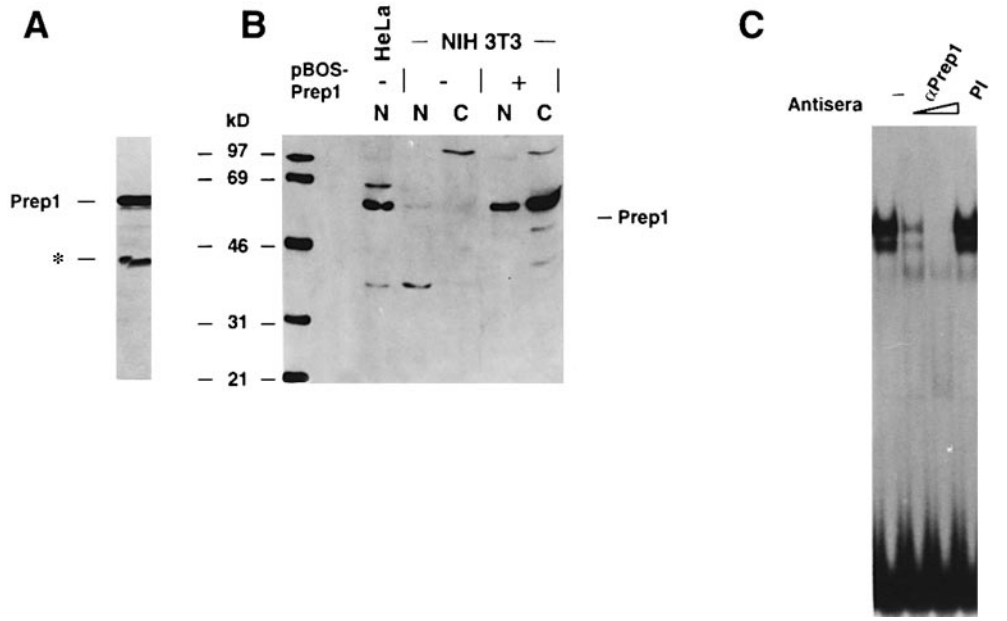


Fig. 2. Analysis of Prep1 cDNA expression. (**A**) SDS–PAGE analysis of ^{35}S-labelled *in vitro* transcribed and translated Prep1 products. The major product run as a 64 kDa polypeptide. A minor product (marked with *) is observed as a 40 kDa polypeptide. (**B**) Immunoblotting analysis with anti-Prep1 antibodies on transfected NIH 3T3 cells lysates separated by SDS–PAGE. N, nuclear extracts; C, cytosolic extracts. The lane labelled HeLa shows the results obtained with crude nuclear extract from HeLa cells. Cells were either mock transfected, or transfected with 10 μg of pBOSPrep1 plasmid. A major product appears in transfected NIH 3T3, both in the nuclear and cytosolic fraction, with the same mobility as endogenous HeLa cell Prep1 (64 kDa). (**C**) Effect of addition of anti-Prep1 serum (αPrep1) or pre-immune serum (PI) on the DNA binding of endogenous UEF3 complexes. EMSA binding reactions with HeLa nuclear extract and o-1 oligonucleotide were assembled with the addition of 0.05 μl or 0.5 μl αPrep1 or 0.5 μl PI, as indicated.

same band in the nucleus but, interestingly, most of the exogenous protein remained in the cytoplasm.

As a final test of the presence of the Prep1 product in UEF3, we used an anti-Prep1 serum raised against a GST–Prep1 fusion protein, and tested its effect on the DNA-binding activity of UEF3 in crude nuclear extract. As shown in Figure 2C, the addition of Prep1 antiserum strongly and specifically inhibited the formation of both bands of UEF3 while pre-immune serum had no effect. From all these data, we conclude that Prep1 cDNA does indeed code for the 64 kDa subunit of both complexes of UEF3.

Prep1 is a DNA-binding protein with specificity for the TGACAG sequence

To investigate the DNA-binding capacity of Prep1, we tested the DNA-binding activity of a recombinant, bacterially expressed GST–Prep1 fusion protein with a labelled DNA sequence from the human urokinase gene containing a TGACAG motif that specifically binds the UEF3 factor (Berthelsen *et al.*, 1996), (o-1 oligonucleotide). Figure 3 shows the comparison between the binding of the HeLa cells nuclear extract (NE) and the purified recombinant protein (GST–Prep1). By EMSA, the nuclear extract gave two bands which were specifically competed by unlabelled o-1, but not by the mutated o-1m oligonucleotide (see Materials and methods). The GST–Prep1 protein preparation gave rise to two slower migrating bands (possibly due to the presence of the GST tail) which were specifically competed by unlabelled o-1, not by o-1m, oligonucleotide. Binding of the recombinant protein to o-1 was specific, as no activity was found with the GST control. These results indicated that the recombinant

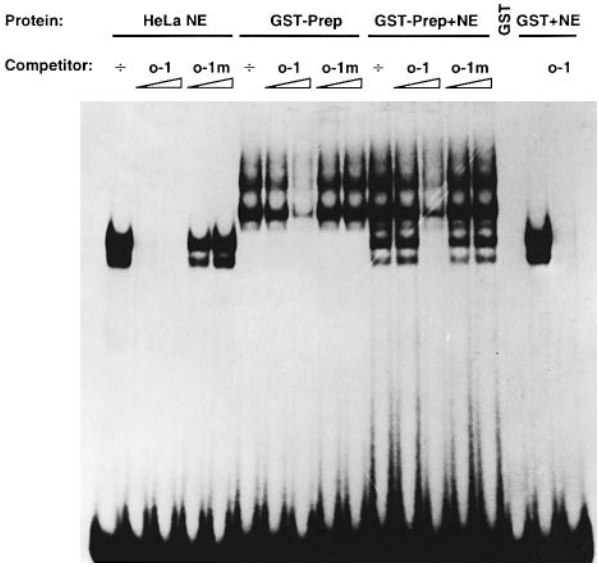


Fig. 3. EMSA analysis of recombinant Prep1 protein. Labelled o-1 oligonucleotide was incubated with either 10 μg of HeLa crude nuclear extract (NE), or 100 ng of purified recombinant GST–Prep1 fusion protein, or a combination of both (NE+GST–Prep1). As a control we included purified recombinant GST protein (GST) and the combination of nuclear extract and recombinant GST (NE+GST). DNA-binding specificity was tested by competing with unlabelled o-1 or mutated o-1 (o-1m) oligonucleotides, at 50- and 500-fold molar excess over labelled oligonucleotide.

Prep1 bound specifically to the same sequence as the HeLa cell UEF3 factor, but with an apparently lower affinity, as high concentrations of GST–Prep1 are necessary for binding, and higher concentrations of unlabelled

o-1 oligonucleotide are required to compete for binding. When we added the recombinant protein to the nuclear extract (NE+GST–Prep1 in Figure 3), the binding pattern was the sum of that obtained with the nuclear extract and the recombinant protein, separately. The presence of the recombinant protein reduced the binding by the nuclear extract, while the control GST had no effect.

We have shown previously, using mutation analysis and methylation interference analysis, that UEF3 preferentially binds the TGACAG core motif (Berthelsen *et al.*, 1996). The data in Figure 3 thus indicate that the Prep1 component of UEF3 is sufficient for specific recognition of the TGACAG motif, and suggest that the p50 and p40 subunits may be important in raising the binding affinity. In several cases, heterodimerization of homeodomain proteins increases DNA-binding affinity (Vershon, 1996). Interestingly, the TGIF protein, which contains a homeodomain highly homologous to Prep1 and almost identical in the third helix (Figure 1C), has been found by PCR-mediated binding site selection to bind to a TGACA$^{A}/_{G}$ motif, identical to the UEF3 binding site (Bertolino *et al.*, 1995). This type of recognition sequence is seen rarely for homeodomains, thus Prep1, TGIF, Meis1 (Chang *et al.*, 1997) and similar homeodomains may constitute a novel class of homeodomains recognizing this type of binding motif.

The 40 and 50 kDa subunits of UEF3 complexes are members of the Pbx protein family

Active transcription factor UEF3 has been purified from HeLa nuclear extracts by employing a variety of column chromatography procedures, including DNA-affinity chromatography. The two complexes of UEF3, UC and LC, co-purify throughout the purification procedure, but could be partially separated by a final Mono-S chromatography. Such purified preparations analysed by SDS–PAGE, displayed two bands: a common 64 kDa band (Prep1) and either a p40 or a p50 polypeptide, respectively (Berthelsen *et al.*, 1996). We set out to determine the nature of the p40 and p50 bands. As discussed above, since Prep1 is a homeodomain protein, its partners in the UEF3 complexes might be other similar homeodomain proteins. We noticed that the molecular weights of p40 and p50 were compatible with those of members of the Pbx family. To investigate this possibility, we performed immunoblotting experiments with anti-Prep1 and different anti-Pbx antibodies using both a crude nuclear fraction of HeLa cells and purified UEF3 containing the p64, p50 and p40 polypeptides. Overall, anti-Prep1 antibodies detected the 64 kDa band from both fractions, while anti-Pbx antibodies recognized the 50 and 40 kDa bands (Figure 4A). In particular, antibodies specific for Pbx1 identified the p40 band in the purified UEF3 preparation, corresponding in size with the Pbx1b splice variant. Additional bands were detected in the crude extract. Antibodies specifically reactive with the 50 kDa splice-variant product of all Pbx family genes (αPbx1/2/3) or with αPbx2 only, recognized the p50, and not the p40, band in both purified and crude preparations. Antibodies to Pbx3 did not react with either fraction. Overall these results indicate that the two complexes of HeLa UEF3 are heterodimers of Prep1 with either Pbx2 or Pbx1b.

We tested the same antibodies in DNA-binding assays using crude UEF3 (Figure 4B). In EMSA, formation of both UC and LC complexes was abolished by anti-Prep1 antibodies. Incubation with anti-Pbx1 supershifted all LC but not UC, while αPbx2 or αPbx1/2/3 inhibited DNA binding of UC only. Both αPbx3 and pre-immune sera had no effect. Overall, these data confirm that crude UEF3 from HeLa cells contains Prep1 complexed with either Pbx1b or Pbx2, and that the migration of the two complexes seen by EMSA is in agreement with the size of the Prep1-associated Pbx subunits.

Association with Pbx increases the affinity of Prep1 for DNA

In view of the difference in DNA-binding affinity of recombinant Prep1 versus UEF3, heterodimerization with Pbx might lead to increased DNA-binding affinity. To test for this possibility, we carried out DNA-binding studies with *in vitro* translated products of the Prep1, Pbx1a and Pbx2 cDNAs, using oligonucleotide o-1 as the binding site. *In vitro* translated Prep1, Pbx1a, Pbx1b or Pbx2 did not give rise to any retarded complex in EMSA when translated individually (Figure 5A and B). However, retarded bands were observed when Prep1–Pbx1b, Prep1–Pbx2 and also Prep1–Pbx1a were co-translated. In contrast, mixing singly translated Prep1 and Pbx proteins did not lead to efficient DNA binding (not shown). Thus, co-translation of Prep1 with one of the Pbx proteins led to the reconstitution of the high-affinity DNA-binding UEF3 complex. The migration of the retarded complexes again agreed with the size of the Pbx subunits. The *in vitro* data show that co-translation of Prep1 with Pbx1a or Pbx2 (both 50 kDa polypeptides) gave raise to the UC slow-migrating complex of UEF3 (UC), while Prep1 and Pbx1b (40 kDa) gave raise to the fast migrating complex, LC. The complexes exhibited the same DNA-binding specificity, as shown by the specific competition with excess unlabelled UEF3 binding site (Figure 5B, lanes 1–3 versus lanes 4–6). Thus, co-translated Prep1–Pbx displays co-operative DNA binding to the TGACAG motif and DNA-binding specificity indistinguishable from crude and purified UEF3. Importantly, co-translation of Prep1 and one of the Pbx family members was sufficient to reconstitute UEF3 binding activity *in vitro*.

Prep1 and Pbx form a strong DNA-independent complex

The co-purification of Prep1 and Pbx subunits of UEF3 throughout several chromatographic steps (Berthelsen *et al.*, 1996), the composition of the individual EMSA complexes, the need for co-translation in order to reconstitute DNA binding *in vitro* and the effect of specific antibodies on DNA binding of crude UEF3, indicate that a Prep1–Pbx complex assembles intracellularly independent of DNA binding. We confirmed this by immunoprecipitation of *in vitro* translated Prep1 and Pbx1a. As shown in Figure 6, co-translated Prep1 and Pbx1a were readily co-precipitated, using either anti-Prep1 or anti-Pbx1 antibodies (lanes 7 and 8); translation of single proteins showed the absence of cross-reactivity of the antibodies (lanes 1–4). In addition, precipitation of singly translated Prep1 and Pbx1a which were mixed after translation did not lead to co-precipitation (lanes 5 and 6), showing that complex formation only occurs during

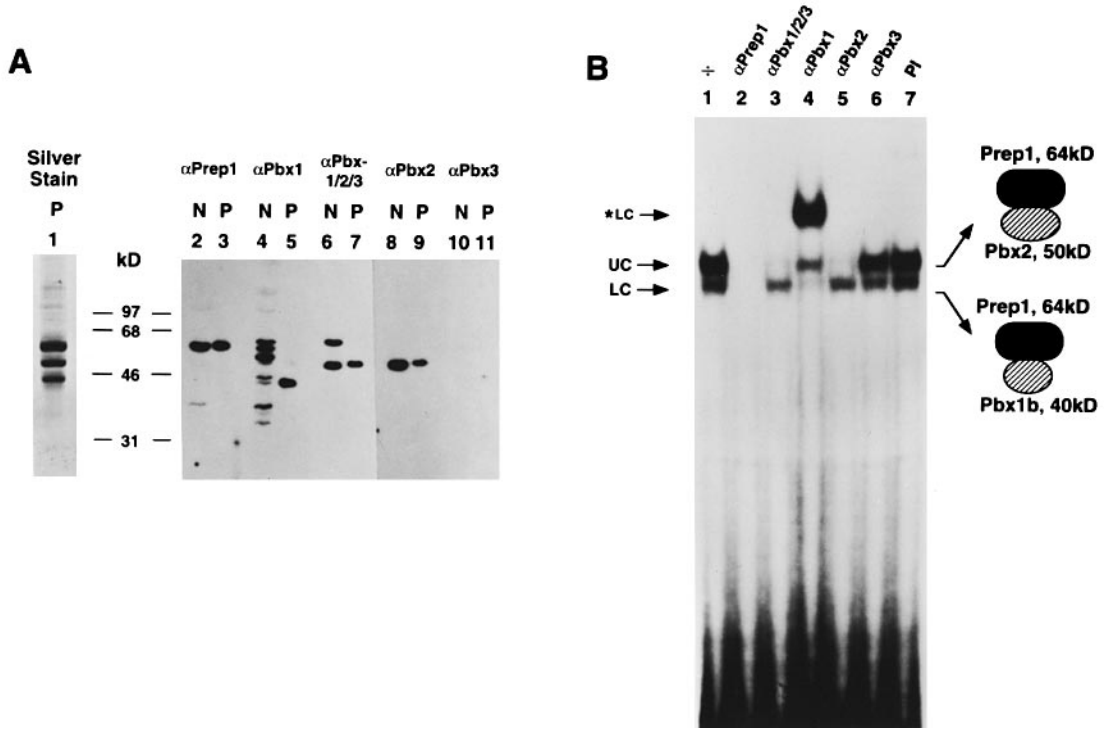


Fig. 4. Subunit composition of purified and crude UEF3. (**A**) Analysis after SDS–PAGE separation of purified UEF3 (P) and crude nuclear extract (N) from HeLa cells. Lane 1: Silver-staining of purified UEF3, showing the presence of the three co-purifying polypeptides, p64, p50 and p40. Molecular weight markers are shown. Lanes 2–11: Immunoblotting analysis on purified UEF3 and nuclear extracts with antibodies against Prep1 or against the various Pbx family members (see Materials and methods). The antibodies employed are indicated on top of each lane. (**B**) Effect of the same antibodies on the DNA binding of endogenous UEF3 upper complex (UC) or lower complex (LC). EMSA binding reactions with HeLa nuclear extract and o-1 oligonucleotide were assembled with the addition of 0.5 μl of αPrep1 or 3 μl of the various anti-Pbx antibodies, as indicated. As a control, we used 0.5 μl of PI. A supershifted LC is indicated by *LC. To the right: A model of our conclusion on the subunit composition of HeLa cell UEF3 UC and LC.

co-translation, in agreement with DNA-binding experiments (see above). Prep1 and Pbx1a contain comparable amounts of labelled methionine (13 versus 14), and thus the relative intensity of the bands reflects their molar ratio. This experiment therefore shows that all available Pbx1a was complexed with Prep1. Thus, during co-translation, the Prep1–Pbx complex formed highly efficiently in solution, and was resistant to stringent washes (600 mM NaCl) during precipitation. *In vitro* translated Prep1 preparations always contained a 40 kDa truncated product (marked by an *) due to translation from an internal ATG codon (see Figure 2 and comments above); this band, however, did not associate with Pbx1a. We conclude that the UEF3 factor must be a Prep1–Pbx complex; our data also indicate that this complex has the properties necessary to exist intracellularly as a stable entity.

Anti-Prep1 antibodies do not cross-react with Meis1

A recent paper shows that Meis1 interacts with Pbx proteins, and that Meis1–Pbx complexes can be detected in the nuclear extracts of several cell lines (Chang *et al.*, 1997). Since Prep1 and Meis1 share three short homologous domains (Figure 1D), we tested whether the anti-Prep1 polyclonal antiserum, raised against the entire Prep1 protein, was able to cross-react with the Meis proteins. Prep1 and Meis1a were produced by *in vitro* translation (Figure 7, I.T., lanes 1,2). The proteins were then subjected to either immunoblot analysis (I.B., lanes 3–5) or to immunoprecipitation (I.P., lanes 6–8) using anti-Prep1 antibodies. In immunoblot, faint cross-reacting background bands can be observed in lysates containing Prep1, Meis1 or non-programmed lysate (lanes 3–5). A strong band co-migrating with Prep1 can be observed in Prep1-containing lysate only (lane 3), while no band co-migrating with Meis1a can be observed in Meis1a-containing lysate (lane 4). In immunoprecipitation, Prep1 is precipitated by anti-Prep1 (lane 6) while Meis1a is not (lane 7). These data show that the anti-Prep1 antibodies are specific for Prep1 and do not cross-react with Meis1 proteins. This is also supported by the observation that anti-Prep1 antibodies do not recognize bands with migration-rates corresponding to Meis proteins, in immunoblot analysis on extracts from HeLa or NIH 3T3 cells (see Figures 2B and 4A).

Discussion

Prep1 is a novel homeodomain protein

We have identified a novel homeodomain protein component of the UEF3 DNA-binding activity. The protein has been named Prep1 for Pbx regulating protein-1. UEF3 is a transcription factor binding to the TGACAG core motif of the human and murine urokinase enhancer, located within a 17 bp sequence (uCOM) which is essential for enhancer function (Nerlov *et al.*, 1992; Berthelsen *et al.*, 1996). Yet, the uCOM sequence cannot substitute for the enhancer and does not mediate transactivation, but appears to allow the co-operation of two flanking transactivators (De Cesare *et al.*, 1996). In this paper, we have reported the cloning of the novel homeodomain protein, Prep1,

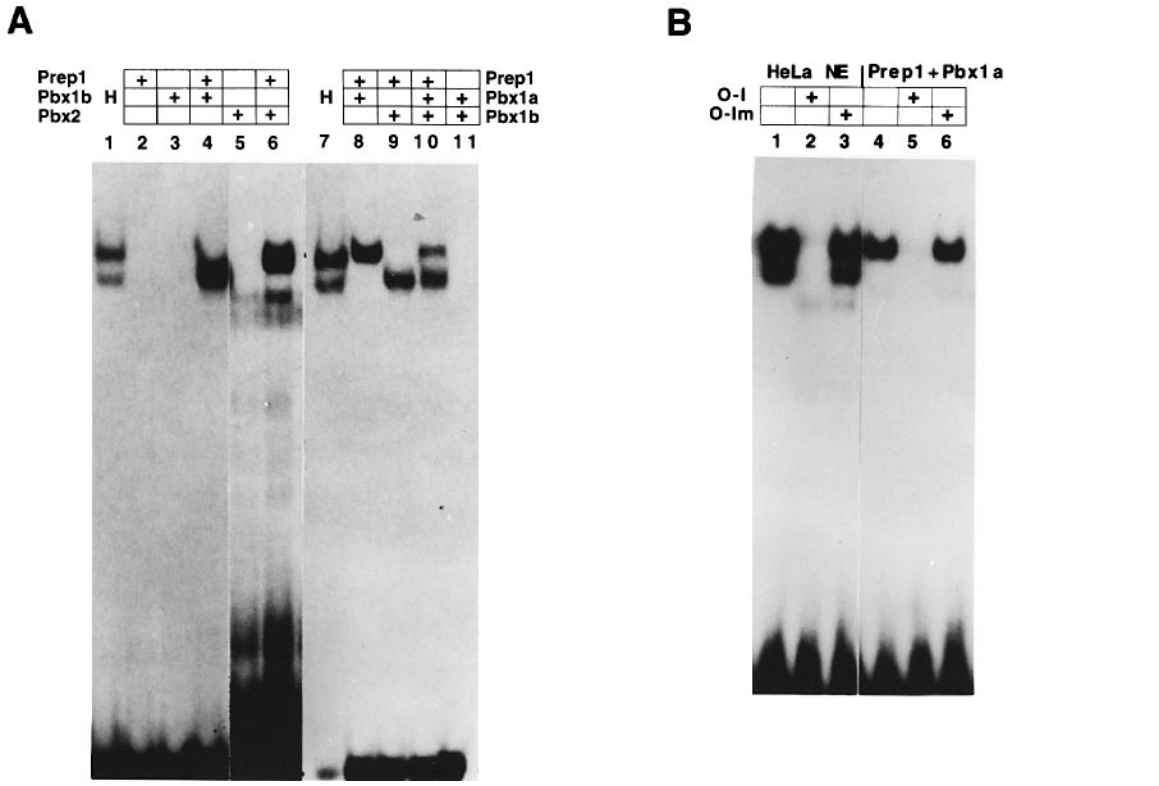


Fig. 5. Reconstitution of UEF3 *in vitro*. (**A**) DNA-binding ability of *in vitro* translated or co-translated Prep1 and Pbx products. EMSA with o-1 oligonucleotide and HeLa nuclear extracts (H, lanes 1 and 7), *in vitro* translated Prep1, Pbx1b or Pbx2 (lanes 2, 3 and 5) or co-translated Prep1–Pbx1b (lane 4) or Prep1–Pbx2 (lane 6). Lanes 1–6 are all from the same gel, but lanes 5–6 have been exposed for three times as long due to the weaker binding of the Prep1–Pbx2 complex. Right-hand side: comparison of DNA-binding complexes formed by co-translating Prep1 with either of the two splice variants of Pbx1, either Pbx1a (lane 8), Pbx1b (lane 9) or both together with Prep1 (lane 10) or without Prep1 (lane 11). (**B**) DNA-binding specificity of endogenous and *in vitro* reconstituted UEF3 complexes. EMSA with o-1 oligonucleotide and either HeLa nuclear extract (NE) or *in vitro* co-translated Prep1–Pbx1a complex. Binding was competed by addition of 500-fold molar excess of unlabelled o-1 or o-1m oligonucleotides.

which is part of the transcription factor UEF3. In nuclear extracts UEF3 forms two retarded bands in EMSA, named upper (UC) and lower complex (LC), when assayed with the oligonucleotide (o-1) containing the TGACAG motif of the human urokinase enhancer. Mutation of this sequence together with methylation interference data have shown that the upper and lower complexes have the same DNA specificity, recognizing the sequence TGACAG present in and important for the activity of the urokinase enhancer (Nerlov *et al.*, 1992; Berthelsen *et al.*, 1996; De Cesare *et al.*, 1996). During purification, the two complex-forming activities remain together through five subsequent chromatographic procedures involving ion exchange, heparin-affinity and DNA-affinity chromatography, but can be separated somewhat in the last stage by Mono-S ion exchange chromatography. The two complexes share a 64 kDa band and, in addition, UC contains a 50 kDa polypeptide, while LC contains a 40 kDa polypeptide (Berthelsen *et al.*, 1996). In this paper we report the cloning of the cDNA for the common 64 kDa protein (Prep1), the identification of the 40 and 50 kDa subunits and also demonstrate their functional interaction.

Prep1 cDNA encodes a novel protein of 436 amino acid residues containing a homeodomain resembling structurally those of the PBC class of divergent homeodomains, a group comprising human and murine Pbx family, the *Drosophila* Exd and the *Caenorhabditis elegans* Ceh-20

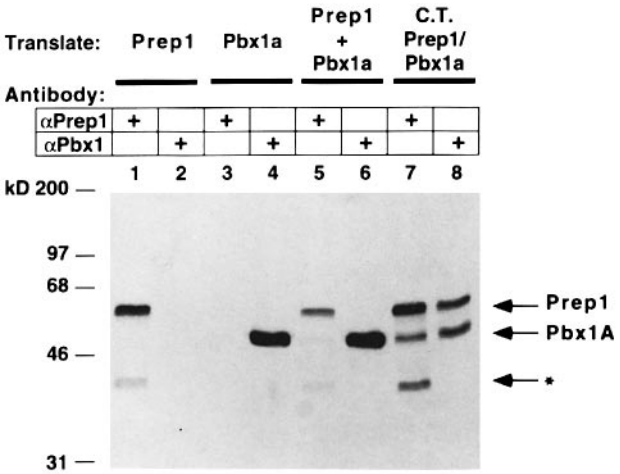


Fig. 6. Prep1–Pbx1a form a DNA-independent complex. Immunoprecipitation of *in vitro* translated Prep1 (lanes 1–2), Pbx1a (lanes 3–4), Prep1 and Pbx1a mixed after translation (lanes 5–6) or co-translated Prep1–Pbx1a (lanes 7–8). Translation products were precipitated using either anti-Prep1 antiserum (αPrep1, lanes 1, 3, 5 and 7) or anti-Pbx1 antibodies (αPbx1, lanes 2, 4, 6 and 8) and protein A–Sepharose. The precipitates were resolved by SDS–PAGE. A minor product (*) is observed when Prep1 is translated *in vitro*.

homeodomain proteins, that in turn also share homeodomain homology with the maize KN-1, yeast Matα2 and others (Burglin and Ruvkun, 1992). The highest homeodomain homology was found with members of human, murine and *Xenopus* Meis proto-oncogene family, and with human TGIF. In addition, Prep1 contains two short regions, HR1 and HR2, which display strong homology with all members of the Meis protein family.

Prep1 displays a specific, albeit low, affinity for the TGACAG motif which is identical to that bound by human TGIF (Bertolino *et al.*, 1995). Prep1, TGIF and Meis1 are almost identical in the third helix of their homeodomains, that is, DNA recognition, and in particular they have a isoleucine at the third position of the highly conserved homeodomain motif, WF_N_N, of the third helix. Isoleucine is rarely found at this position in homeodomains (Burglin, 1994), although searching the EST database at NCBI revealed numerous different EST sequences from yeast to human which contain homeodomains with a third helix that is highly similar to that of Prep1 and with isoleucine in this third position. Thus, Prep1, TGIF and Meis1 represent a novel emerging subclass of homeodomains that shows binding preference for a TGACAG-like motif.

Prep1 cDNA encodes a protein with a calculated molecular mass of 49 kDa, that migrates as a 64 kDa polypeptide in SDS–PAGE. The slow migration of Prep1 possibly results from an intrinsic, SDS-resistant folding of the protein, since the protein migrates much faster in SDS–PAGE in the presence of 8 M urea (data not shown). An antibody raised against recombinant Prep1 is able to specifically inhibit binding of both UEF3 complexes to the o-1 oligonucleotide, proving that Prep1 is part of both complexes, as originally shown by purification (Berthelsen *et al.*, 1996).

UEF3 complexes are heterodimers of Prep1 and Pbx proteins

Immunoblotting, *in vitro* co-translation and functional inhibition by specific antibodies have shown that the p40 and p50 subunits of UEF3 purified from HeLa nuclear extract belong to the Pbx family: Pbx2 (50 kDa) and Pbx1b (40 kDa). Because of the composition of the respective polypeptides (Berthelsen *et al.*, 1996), the UC of UEF3 is composed of Prep1–Pbx2, and the LC is composed of Prep1–Pbx1b. The fact that the proteins co-purify through a series of five different chromatographic columns suggests that Prep1 and Pbx are associated in an intracellular complex. In fact, antibodies specific for Prep1 or specific for individual members of Pbx inhibit or supershift binding of UEF3 complexes in EMSA even in crude nuclear extracts.

We are able to reconstitute the DNA-binding activity of UEF3 *in vitro*, by co-translation of the Prep1 and the Pbx components. The association between the two proteins increases the DNA-binding affinity towards the TGACAG motif strongly, and the reconstituted activity exhibits DNA binding specificity indistinguishable from UEF3. Furthermore, by using the different splice variant products of Pbx1 in association with Prep1 we are able to reproduce the migration of the two UEF3 complexes in EMSA. Overall, Prep1 and Pbx are necessary and sufficient to form the UEF3 activity.

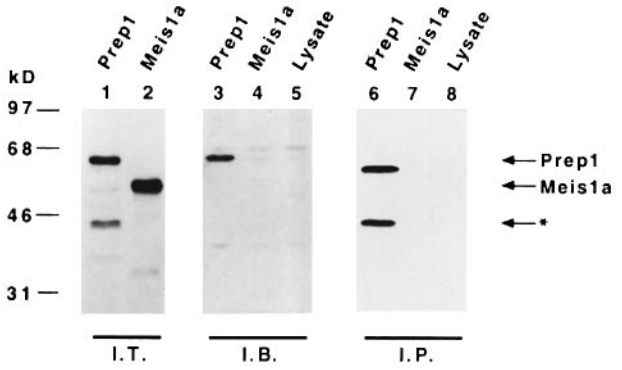


Fig. 7. Anti-Prep1 antibodies do not recognize Meis1. Lanes 1 and 2: Prep1 and Meis1 were produced by *in vitro* translation (I.T.). Lanes 3–5: Immunoblot (I.B.) analysis of *in vitro* produced proteins or of non-programmed reticulocyte lysate (lysate) with anti-Prep1 antibodies. Lanes 6–8: Immunoprecipitation (I.P.) of the different lysate with anti-Prep1 antibodies. Molecular weight markers are indicated. The migration of Prep1, Meis1 or of minor product of Prep1 (*) are shown.

So far, no differential function for the various members of the Pbx family has been described. The HeLa cell nuclear extracts from which UEF3 was purified are negative for Pbx3 protein. However, the high similarity between Pbx1, 2 and 3 would also argue that Pbx3 can interact with Prep1, but this is yet to be demonstrated. In this view, the subunit composition of UEF3 in a given cell would depend on, and reflect, the expression of the various genes and splice variants of the Pbx family. Preliminary data from our laboratory seem to support this view (not shown).

Prep1 and Pbx interaction is DNA-binding independent

Pbx and their *Drosophila* homologue Exd interact with Hox and HOM-C proteins (Mann and Chan, 1996). Although Pbx can interact with various Hox proteins in a yeast two-hybrid analysis, *in vitro* the interaction seems to occur only in the presence of DNA target (Chang *et al.*, 1995). Thus, it is not clear whether *in vivo* Pbx–Exd interacts with the Hox–HOM-C proteins in solution, or whether the interaction only occurs on specific DNA sequences. *In vitro* data, on the one hand, suggest that DNA binding promotes complex formation. Prep1, on the other hand, forms a strong complex with Pbx in solution independent of DNA binding. All homeodomain proteins known to interact with Pbx contain a conserved pentapeptide sequence YPWMX, or a similar ANW amino acid motif, N-terminal to the homeodomain (Chang *et al.*, 1995, 1996; Johnson *et al.*, 1995). Together with the homeodomain, this motif is essential for the heterodimerization of HOX–HOM-C proteins with Pbx–Exd (Chang *et al.*, 1995, 1996; Johnson *et al.*, 1995; Knoepfler and Kamps, 1995; Lu and Kamps, 1996; Lu *et al.*, 1995; Neuteboom *et al.*, 1995; Phelan *et al.*, 1995; van Dijk *et al.*, 1995; Chan and Mann, 1996; Chan *et al.*, 1996; Peltenburg and Murre, 1996). Interestingly, Prep1 does not contain any such motif. In fact, Prep1 does not contain any tryptophan residue N-terminal to the homeodomain. This, together with the strength of the Prep1–Pbx complex, classifies the Prep1–Pbx interaction as different to that of

Pbx–HOX, and probably involves structurally different dimerization motives.

It has been reported recently that Meis1–Pbx complexes can be immunoprecipitated from NIH 3T3 extracts, that Meis1–Pbx DNA-binding activities are present in extracts from several cell lines, and can be reconstituted with *in vitro* translated Meis1 and Pbx1 (Chang *et al.*, 1997). Moreover, Homothorax has been shown to be the *Drosophila* homologue of Meis and is able to interact with extradenticle, the *Drosophila* homologue of Pbx (Rieckhof *et al.*, 1997). The amino acid sequence of Prep1 and the specificity of anti-Prep1 antibodies show that Prep1–Pbx is a novel intracellular complex, as the DNA-binding activities described in this paper are completely dependent on the presence of Prep1.

Prep1–Pbx complexes are involved in regulating protease and growth factor gene expression

UEF3 (i.e. the Prep1–Pbx heterodimers) is known to be involved in regulating the activity of the enhancer of the human urokinase gene. This enhancer requires the co-operative function of three protein-binding sites, a PEA3 (binding ets-2) an octameric CRE–AP-1 site (binding the cJun–ATF2) and a classical heptameric AP-1 site binding cFos–cJun dimers (Nerlov *et al.*, 1992; De Cesare *et al.*, 1995). Transcription factor UEF3 is one of the factors which allows this co-operation (Nerlov *et al.*, 1992; De Cesare *et al.*, 1996), although the mechanism remains unknown. UEF3 also binds specifically a negative regulatory element (NIP) of the phorbol ester-responsive stromelysin, interleukin-3, LD78 cytokine and GM-CSF gene promoters and of the murine intracisternal A-particle (IAP) LTR (Berthelsen *et al.*, 1996; De Cesare *et al.*, 1996, 1997). In these genes, the UEF3-binding sequence (TGACAG consensus) has a negative regulatory function and, like in the urokinase gene, resides in the immediate neighbourhood of AP-1 and PEA-3 sites (Nomiyama *et al.*, 1993). Thus, growth factors or proteases are putative target genes for the Prep1–Pbx complex.

In conclusion, we have uncovered the molecular nature of a transcription factor which co-operates in activating the urokinase enhancer, as well as in the transcriptional regulation of other genes. The demonstration of an interaction between the Hox co-factor Pbx and the novel homeodomain protein Prep1 is of the utmost interest, as the properties of this heterodimer may interfere with the function of Pbx in regulating target selectivity, binding affinity and transcriptional activation of Hox proteins.

Materials and methods

Cloning of p64–Prep1 cDNA

A 50 bp fragment corresponding to the originally obtained 17 amino acid peptide sequence of Prep1 (p64-pep1) was found by RT–PCR using the following procedure: first-strand cDNA was prepared from poly(A)$^+$ RNA from HeLa cells using a cDNA synthesis kit (Pharmacia) with an anchored poly-dT primer, 5′ AACTGGAAGAATTCGCGGCCGCAGGAAT$_{18}$ 3′, according to the manufacture's manual. Two internal degenerate primers, 64pep1–F1 [5′ ATGTGGCT(C/G/T)TT(C/T)CA(A/G)CA(C/T)AT 3′] and 64pep1–R [5′ TT(C/T)TC(A/G)TC(C/T)TC(A/G/T)GT(A/G/T)GG(A/G)TA 3′] derived from the amino acid sequence of p64-pep1 (see Figure 1A) were used for the amplification. cDNA (10 ng) and 100 pmol of each primer were used in a 50-μl reaction with 0.2 mM dNTP, 1unit *Taq* polymerase (Boehringer Mannheim) and supplied buffer. The reaction involved 30 cycles of 94°C denaturation for 30 s, annealing at 55°C for 1 min followed by extension at 72°C for 30 s. The reaction products were separated on a 4% NuSieve (FMC) agarose gel in 0.5×TBE. The expected band of 50 bp was cut out, cloned into pBluescript and sequenced. From the sequence obtained we constructed a nested forward primer, 64pep1–F2 (5′ CAGCACATCGGGCATCCCTA 3′, Figure 1A), that was used, together with a primer containing a sequence equal to the anchor part of the poly(dT) primer, in a 3′-RACE reaction on 10 ng HeLa cDNA as described previously (Frohman *et al.*, 1988). The 180 bp fragment obtained was used as a probe to screen a HeLa dT-primed cDNA library cloned in the λgt11 vector (Clontech) by hybridization (Sambrook *et al.*, 1989). Of 500 000 plaques, four positive ones were isolated and sequenced. The longest clone was 2.1 kb.

The 5′ region of the Pre1 cDNA was found using the Marathon 5′ RACE protocol (Clontech), following the supplier's manual, and a nested primer set, LDR1 and LDR2.

Homology search

All homology searches in data banks were done using the BLAST e-mail server at NCBI at NIH (blast@ncbi.nlm.nih.gov), employing the Blastn, tblastn and blastp programs.

Purification of UEF3 and recombinant Prep1

As a source of purified UEF3, we used a fraction (fraction 12) containing all three UEF3 co-purifying polypeptides. This fraction was taken after the final Mono-S chromatographic step of the UEF3 purification procedure as described previously (Berthelsen *et al.*, 1996).

The cDNA insert of a λ-clone of Prep1 lacking the first 10 codons after the ATG initiation codon (see Figure 1A), was cut out by partial digestion of *Eco*RI and cloned into the *Eco*RI site of pGEX2T (Pharmacia) downstream of, and in frame with, the *GST* gene. The GST–Prep1 fusion protein was expressed in XL-1 blue cells (Stratagene) for 5 h at 37°C by the addition of 0.1 mM IPTG. The cells were washed and lysed by sonication. The fusion protein was purified on Glutathione–Sepharose (Pharmacia) according to the manufacturer's instructions.

Transfections and cell extracts

For expression in eukaryotic cells, we cloned the cDNA of Prep1 into the pEF–BOS expression vector (Mizushimi and Nagata, 1990), between the *Eco*RI and *Xba*I sites (pBOS–Prep1). Contact inhibited NIH 3T3 fibroblast cells were grown in DMEM supplemented with 10% newborn calf serum and antibiotics. Cells were transfected with the pBOSPrep1 plasmid by $CaPO_4$ precipitation in 10 cm dishes. After 48 h the cells were washed with PBS, collected with 100 μl of cold buffer A (10 mM HEPES pH 7.9, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$) into a microfuge tube on ice, and lysed by the addition of Triton X-100 to a final concentration of 0.3%. The nuclei and cell debris were collected by centrifugation. The supernatant was removed to a new tube, 0.11 vol of buffer B added (0.3 M HEPES pH 7.9, 1.5 M KCl, 15 mM $MgCl_2$) and centrifuged. The resulting supernatant was denoted cytoplasmic extract (CE). Nuclear extract was prepared by resuspending the pelleted nuclei in 60 μl of buffer C [20 mM HEPES pH 7.9, 25% glycerol (v/v), 0.42 M NaCl, 1.5 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$] for 30 min on ice. The extract was cleared by centrifugation.

Antibodies and immunoblotting

Anti-Prep1 rabbit antiserum was produced by immunizing a rabbit with purified GST–Prep1 recombinant protein. GST–Prep1 (2 mg) was used to immunize a rabbit over a period of 4 months, after which the antiserum was taken. IgG was isolated by Protein A chromatography, dialysed and suspended in PBS to a concentration of 10 mg/ml.

Antibodies against PBX proteins were obtained from Santa Cruz Biotechnology. The antibodies αPbx1, αPbx2 and αPbx3 recognize a peptide from the N-terminal of the respective proteins and are type-specific. αPbx1 and αPbx3 will, however, recognize both splice variants of their type. The antibody αPbx1/2/3 recognizes a common C-terminal peptide in all of the 50 kDa splice variants of the PBX proteins.

Cell extracts or purified UEF3 were separated by a 10% SDS–PAGE and the proteins were blotted onto PVDF membrane in a semidry blotting apparatus. Immunoblotting with Prep1 antisera (1:10 000 dilution) or PBX antibodies (1:1000 dilution) was carried out using the BM Chemoluminoscent Kit (Boehringer Mannheim) according to the supplier's manual.

Immunoprecipitation

Proteins produced *in vitro* were immunoprecipitated by resuspending 3 μl of translation lysate in 200 μl of NET buffer (50 mM Tris–Cl

pH 7.5, 150 mM NaCl, 2% Gelatine, 0.1% NP-40), 20 µg of anti-Prep1 IgG or 5 µg αPbx1 antibodies were added and the mixture was incubated on ice for 1 h. Immunocomplexes were precipitated by the addition of 5 µl (bed volume) of protein A–Sepharose (Pharmacia Biosystems), and rotating for 1 h. The resin was recovered by brief centrifugation, and washed three times with 200 µl of NET600 (NET buffer containing 600 mM NaCl) for 20 min each wash and once in TENP (10 mM Tris–Cl, pH 7.5, 1 mM EDTA, 0.1% NP-40). The resin was finally resuspended in 10 µl of 2× SDS loading buffer (100 mM Tris–Cl pH 6.8, 100 mM DTT, 4% SDS, 20% glycerol) and boiled for 5 min. A quantity (10 µl) was separated using 10% SDS–PAGE, after which the gel was fixed (10% HAc, 10% methanol), soaked in Amplify (Amersham), dried and exposed to Kodak X-OMAT at –80°C.

EMSA

EMSAs were carried out essentially as described by Berthelsen *et al.* (1996). Briefly, 10 µg of nuclear extract, 30 µg of cytoplasmic extract or 2 µl of *in vitro* translation lysate were incubated with 2 µg of poly(dIdC) in 9 µl of H2K150 [25 mM HEPES pH 7.9, 20% glycerol (v/v), 150 mM KCl, 0.2 mM EDTA, 1 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$] on ice for 10 min. We then added 20 000 c.p.m. ^{32}P-labelled oligonucleotide together with the eventual cold competitor oligonucleotide or antibody and the mixture was left at room temperature for 20 min. The reactions were separated by a 5% PAGE in 0.5×TBE. The TGACAG-sequence containing oligonucleotide o-1 CACCTGAGAG**TGACAG**AAGGAAGGCAGGGAG and the mutated o-1m CACCTGAGAT**T**C**ACAG**AAGGAAGGCAGGGAG were used as binding sites (Berthelsen *et al.*, 1996).

In vitro transcription/translation

For use in *in vitro* transcription/translation reactions, the Prep1 cDNA was inserted into the *Eco*RI site of pBluescript SK, downstream of the T7 promoter. The Pbx cDNA constructs (pSG–Pbx) were under the control of the T7 promoter of the pSG5 plasmid (Di Rocco *et al.*, 1997). The Meis1 cDNA (kindly provided by Dr N.Copeland) was inserted into pBluescript SK as a *Bam*HI–*Not*1 fragment, downstream of the T7 promoter. The cDNA plasmid constructs were transcribed and translated *in vitro* using the TNT coupled transcription–translation reticulocyte lysate (Promega) according to the manufacturer's instructions, using T7 polymerase and ^{35}S-methionine (Amersham). Plasmids were added either alone, or co-translated by mixing equimolar amounts of the two plasmids. Translation products were always analysed by SDS–PAGE followed by exposure to a Kodak X-OMAT film.

Accession number

Prep1 DDBJ/EMBL/GenBank database accession number: Y13613.

Acknowledgements

The authors wish to thank Dr S.Tronick (Santa Cruz Biotechnology, Inc.) for the generous gift of Pbx antibodies. We are grateful to Dr Neal Copeland for providing the Meis1a expression plasmid. We are grateful to Telethon Italy (grant No. E485) and AIRC (Associazione Italiana Ricerche sul Cancro) for funding this work. J.B. is supported by a fellowship from the Danish Research Academy.

References

Basu,A., Dong,B., Krainer,A.R. and Howe,C.C. (1997) The intracisternal A-particle proximal enhancer-binding protein activates transcription and is identical to the RNA- and DNA-binding protein p54nrb/NonO. *Mol. Cell. Biol.*, **17**, 677–686.

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.

Bertolino,E., Reimund,B., Wildt-Perinic,D. and Clerc,R.G. (1995) A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *J. Biol. Chem.*, **270**, 31178–31188.

Besser,D., Verde,P., Nagamine,Y. and Blasi,F. (1996) Signal transduction and the uPA/uPAR system. *Fibrinolysis*, **10**, 215–237.

Burglin,T.R. (1994) A comprehensive classification of homeobox genes. In Duboule,D. (ed.), *Guidebook to the Homeobox Genes*. Oxford University Press, Oxford, UK, pp. 27–41.

Burglin,T.R. (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, IROQUIS, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res.*, **25**, 4173–4180.

Burglin,T.R. and Ruvkun,G. (1992) New motif in PBX genes. *Nature Genet.*, **1**, 319–320.

Chan,S.-K. and Mann,R.S. (1996) A structural model for a homeotic protein–extradenticle–DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc. Natl Acad. Sci. USA*, **93**, 5225–5228.

Chan,S.-K., Popperl,H., Krumlauf,R. and Mann,R.S. (1996) An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.*, **15**, 2476–2487.

Chang,C.P., Shen,W.F., Rozenfeld,S., Lawrence,H.J., Largman,C. and Cleary,M.L. (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.*, **9**, 663–674.

Chang,C.P., Brocchieri,L., Shen,W.F., Largman,C. and Cleary,M.L. (1996) Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. *Mol. Cell. Biol.*, **16**, 1734–1745.

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.

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.

De Cesare,D., Vallone,D., Caracciolo,A., Sassone-Corsi,P., Nerlov,C. and Verde,P. (1995) Heterodimerization of c-Jun with ATF-2 and c-Fos is required for positive and negative regulation of the human urokinase enhancer. *Oncogene*, **11**, 365–376.

De Cesare,D., Palazzolo,M. and Blasi,F. (1996) Functional characterization of COM, a DNA region required for cooperation between AP-1 sites in urokinase gene transcription. *Oncogene*, **13**, 2551–2562.

De Cesare,D., Palazzolo,M., Berthelsen,J. and Blasi,F. (1997) Characterization of UEF-4, a DNA-binding protein required for transcriptional synergism between two AP-1 sites in the human urokinase promoter. *J. Biol. Chem.*, **272**, 23921–23929.

van Dijk,M.A., Peltenburg,L.T. and Murre,C. (1995) Hox gene products modulate the DNA binding activity of Pbx1 and Pbx2. *Mech. Devel.*, **52**, 99–108.

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.

Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) Rapid amplification of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.

Johnson,F.B., Parker,E. and Krasnow,M.A. (1995) Extradenticle protein is a selective cofactor for the *Drosophila* homeotics: Role of the homeodomain and the YPWM amino acid motif in the interaction. *Proc. Natl Acad. Sci. USA*, **92**, 739–743.

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.

Lu,Q. and Kamps,M.P. (1996) Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: Proposal for a model of a Pbx–Hox–DNA complex. *Mol. Cell. Biol.*, **16**, 1632–1640.

Lu,Q., Knoepfler,P.S., Scheele,J., Wright,D.D. and Kamps,M.P. (1995) Both Pbx1 and E2A–Pbx1 bind the DNA motif ATCAATCAA cooperatively with the products of multiple murine Hox genes, some of which are themselves oncogenes. *Mol. Cell. Biol.*, **15**, 3786–3795.

Mann,R.S. and Chan,S.-K. (1996) Extra specificity from *extradenticle*: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet.*, **12**, 258–262.

Mizushimi,S. and Nagata,S. (1990) pEF–BOS, a powerful mammalian expression vector. *Nucleic Acids Res.*, **18**, 5322.

Moskow,J.J., Bullrich,F., Huebner,K., Daar,I.O. and Buchberg,A.M. (1995) *Meis1*, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Mol. Cell. Biol.*, **15**, 5434–5443.

Nerlov,C., Rørth,P., Blasi,F. and Johnsen,M. (1991) Essential AP-1 and PEA3 binding elements in the human urokinase enhancer display cell type specific activity. *Oncogene*, **6**, 1583–1592.

Nerlov,C., Cesare,D.D., Pergola,F., Caracciolo,A., Blasi,F., Johnsen,M. and Verde,P. (1992) A regulatory element that mediates co-operation between a PEA3/AP-1 element and an AP-1 site is required for phorbol ester induction of urokinase enhancer activity in HepG2 hepatoma cells. *EMBO J.*, **11**, 4573–4582.

Neuteboom,S.T., Peltenburg,L.T., van Dijk,M.A. and Murre,C. (1995) The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins. *Proc. Natl Acad. Sci. USA*, **92**, 9166–9170.

Nomiyama,H., Hieshima,K., Hirokawa,K., Hattori,T., Takatsuki,K. and Miura,R. (1993) Characterization of cytokine LD78 gene promoters: positive and negative transcriptional factors bind to a negative regulatory element common to LD78, interleukin 3 and granulocyte-macrophage colony stimulating factor gene promoters. *Mol. Cell. Biol.*, **13**, 2787–2801.

Peltenburg,L.T.C. and Murre,C. (1996) Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. *EMBO J.*, **15**, 3385–3393.

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.

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 *extradenticle*-related homeoprotein. *Cell*, **91**, 171–183.

Rørth,P., Nerlov,C., Blasi,F. and Johnsen,M. (1990) Transcription factor PEA3 participates in the induction of urokinase plasminogen activator transcription in murine keratinocytes stimulated with epidermal growth factor or phorbol-ester. *Nucleic Acids Res.*, **18**, 5009–5017.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Stacey,K.J., Fowles,L.F, Colman,M.S., Ostrowski,M.C. and Hume,D.A. (1995) Regulation of urokinase-type plasminogen activator gene transcription by macrophage-colony stimulating factor. *Mol. Cell. Biol.*, **15**, 3430–3441.

Teesalu,T., Blasi,F. and Talarico,D. (1996) Embryo implantation in mouse: fetomaternal coordinationin the pattern of expression of *uPA*, *uPAR*, *PAI-1* and α_2*MR/LRP* genes. *Mech. Devel.*, **56**, 103–116.

Treismann,J., Gönczy,P., Vashishtha,M., Harris,E. and Desplan,C. (1989) A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell*, **59**, 553–562.

Verde,P., Boast,S., Franze,A., Robbiati,F. and Blasi,F. (1988) An upstream enhancer and a negative element in the 5 flanking region of the human urokinase plasminogen activator gene. *Nucleic Acids Res.*, **16**, 10699–10716.

Vershon,A.K. (1996) Protein interactions of homeodomain proteins. *Curr. Opin. Biotech.*, **7**, 392–396.

Received October 23, 1997; revised and accepted December 23, 1997

Note added in proof

A recent paper identifies the HR1 and HR2 domains also in the *C.elegans* CEH-25 homeoprotein by sequence comparison with Meis-proteins (Burglin, 1997).