Biochemistry of the Tale Transcription Factors PREP, MEIS, and PBX in Vertebrates

E. Longobardi, D. Penkov, D. Mateos, G. De Florian, M. Torres, and Francesco Blasi

TALE (three amino acids loop extension) homeodomain transcription factors are required in various steps of embryo development, in many adult physiological functions, and are involved in important pathologies. This review focuses on the PREP, MEIS, and PBX sub-families of TALE factors and aims at giving information on their biochemical properties, i.e., structure, interactors, and interaction surfaces. Members of the three sets of protein form dimers in which the common partner is PBX but they can also directly interact with other proteins forming higher-order complexes, in particular HOX. Finally, recent advances in determining the genome-wide DNA-binding sites of PREP1, MEIS1, and PBX1, and their partial correspondence with the binding sites of some HOX proteins, are reviewed. These studies have generated a few general rules that can be applied to all members of the three gene families. PREP and MEIS recognize slightly different consensus sequences: PREP prefers to bind to promoters and to have PBX as a DNA-binding partner; MEIS prefers HOX as partner, and both PREP and MEIS drive PBX to their own binding sites. This outlines the clear individuality of the PREP and MEIS proteins, the former mostly devoted to basic cellular functions, the latter more to developmental functions. Developmental Dynamics 243:59–75, 2014. © 2013 Wiley Periodicals, Inc.

Key words: PREP, MEIS, and PBX; vertebrates; TALE proteins

Accepted 5 July 2013

INTRODUCTION

TALE proteins share the same basic structure (see below) with a conserved atypical 60-residue-long helix-loop-helix homeodomain (HD), which has a three-amino-acids extension between the first and the second helix with respect to the classical homeobox, a feature that gives this class the name TALE (Three Amino acids Loop Extension) (Bürglin, 1997). The primary goal of this review is to provide readers with practical and tabulary information rather than a comprehensive view of the available information.

In the present review, we will focus on the structure and molecular interactions of the TALE class of transcription factors PREP, MEIS, and PBX, in particular from Homo sapiens, Mus musculus, and Danio rerio. A phylogenetic tree is summarized in Figure 1A showing the evolutionary diversification of the corresponding genes. The extent of conservation in certain regions is very high, but Figure 1 also indicates the diversification of some genes, like the zebrafish Prep1.1, which reflects the duplication of the single mammalian Prep1 gene into Prep1.1 and Prep1.2 in zebrafish. It is interesting to note that the Prep1.2 gene of zebrafish is inducible by retinoic acid (Vaccari et al., 2010). This suggests, therefore, that also in
mammals some of the functions of Prep1 may be induced by retinoic acid; this possibility has, however, not yet been explored.

TALE proteins class includes two families, PBC and MEINOX (a contraction of MEIS and KNOX), the latter further sub-divided in PREP and MEIS sub-families. When we use the term MEINOX, we include both MEIS and PREP. A single MEINOX gene (hth) is present in Drosophila and is frequently thought to be the ancestral MEINOX ortholog. However, other insects like the malaria mosquito, the honey-bee, and the red flour beetle express both Meis and Prep orthologs. Possibly, therefore, a common Prep and Meis precursor was present in a common bilaterian ancestor and the Prep ortholog has been lost in Drosophila (Mukherjee and Bürglin, 2007).

Table 1 summarizes the symbols of the various proteins, their relationship to the single Drosophila melanogaster ortholog, and their correspondence in the various species. Notice that in zebrafish, most genes are duplicated: the single mammalian Prep1, Meis1, and Meis2 genes are split into Prep1.1 and Prep1.2, Meis1 and Meis4.1, and Meis2a and Meis2b. Table 1 also highlights the correspondence of the nomenclature in zebrafish and mammals. For example, the zebrafish equivalent of mammalian Pbx1 is called Lazarus (Lzr) or Pbx4 (zfin.org).

In mammals, the PBC family includes four PBX proteins, whereas the PREP and MEIS subfamilies include two PREP and three MEIS proteins. PBX is the acronym for Pre-B-cell leukemia homeobox since it was isolated as a fusion protein present in a large fraction of human pre-B Acute Lymphocytic Leukemia (Kamps et al., 1990; Nourse et al., 1990); MEIS is the acronym for Myeloid Ectopic Integration Site since it was discovered by viral insertion mutagenesis in the mouse (Moskow et al., 1995; Steelman et al., 1997). PREP stands for PBX Regulatory Protein (Berthelsen et al., 1998a) a name that is widely used instead of the initially proposed PKNOX (Pbx/Knotted 1 homeobox), which, however, is still the official acronym of this protein in the data banks. The originally proposed PKNOX1 acronym for the gene (Chen et al., 1997) stems from the homology with the plants Knotted homeobox family of transcription factors; however, the symbol PREP1 has gained consensus since PKNOX can be confused with a different KNOX family of plants transcription factors (Mukherjee and Bürglin, 2007) while PREP has a distinct and functionally correct connotation. Despite the clear differences underlined in the reports describing their isolation, many authors have considered PREP and MEIS as members of the same gene family. However, the actual analysis shows that they belong to different sub-families of the MEINOX family (Fognani et al., 2002; Mukherjee and Bürglin, 2007).

Although outside the scope of this review, a brief summary of the function of these proteins, as it appears from genetic studies, is necessary. The first TALE gene to be knocked out in mouse was Pbx1, whose deletion gave an embryonic lethal phenotype at E15.5 with hypoplasia or aplasia of several organs and homeotic transformation of elements of the second into elements of the first branchial arch (Selleri et al., 2001). Mice missing the Pbx3 gene died within one day after birth due to central
Developmental Dynamics studies on each other, a hypothesis supported by genes can at least in part compensate et al., 2004). This suggested that KO mice have no phenotype (Selleri appears to be dispensable as the severe phenotype than mice, which showed a much more hypoventilation (Rhee et al., 2004). The Pbx2 gene, on the other hand, appears to be dispensable as the Pbx2 KO mice have no phenotype (Selleri et al., 2004). This suggested that Pbx genes can at least in part compensate each other, a hypothesis supported by studies on Pbx1-Pbx2 double KO mice, which showed a much more severe phenotype than Pbx1 KO alone (Capellini et al., 2006).

Deletion of the Meis1 gene also causes an embryonic lethal phenotype, with death at E14.5, deficiency in definitive hematopoietic stem cells generation, vascular and ocular abnormalities (Hisa et al., 2004; Azcoitia et al., 2005; Carramolino et al., 2010). Importantly, Meis1 overexpression has an important inducing role in the generation of human and murine leukemias and tumors in general (Moskov et al., 1995; Nakamura et al., 1996; Wang et al., 2006; Wong et al., 2007; Crijins et al., 2007; Grubach et al., 2008; Somervaille et al., 2009).

On the other hand, Prep1 null mouse embryos show a very severe phenotype with uterine death before E7.5, lack of gastrulation, and apoptosis of the epiblast (Fernandez-Diaz et al., 2010). An hypomorphic Prep1 mutant (Prep1−/−) embryonic lethal before E7.5, hypomorphic (Prep1+/−, E17.5) and trans-heterozygous (Prep1−/−, E12.5) mutants (Rowan et al., 2010), indicates that Prep1 has multiple critical and essential functions during embryogenesis. In addition, the few adult Prep1+/− mice (25%) that escape embryonic lethality, live a normal-length life but display a variety of phenotypes. Importantly, they develop tumors at high frequency that, together with other properties, defines Prep1 as a haplo-insufficient tumor suppressor (Longobardi et al., 2010). Further studies have shown that Prep1 in fact protects cells from DNA damage (Iotti et al., 2011).

### ANATOMY OF THE TALE PROTEINS

The comprehensive study on TALE genes by Mukherjee and Bürglin (2007) can be used as a reference for genomic structural features, which therefore will not be discussed in this report. However, some data for human, mouse, and zebrafish PREP, MEIS, and PBX genes/proteins are summarized in Table 2. Chromosome position, gene length, number of exons, length of the major transcript, protein length, and length of the 5′ and 3′ untranslated regions (5′-UTR and 3′-UTR), are indicated.

The overall structural organization of the three TALE protein subfamilies PREP, MEIS, and PBX is similar, containing a DNA-binding homeodomain (HD) towards the carboxy-terminus and two protein–protein interaction domains towards the amino terminus (MEIS-A, B in PREP and MEIS, and PBC-A, -B domains in PBX) (Fig. 1B). The HD is conserved in all these proteins whereas the MEIS-A, B domains are conserved only within the MEINOX and the PBC-A, -B domains only within the PBC family (see below). PREP and MEIS proteins belong to different subfamilies. Indeed, sequence similarities are only found within the homeodomain and in two shared domains, MEIS-A and MEIS-B, whereas within each sub-family there is extensive similarity throughout the entire gene/protein (Fognani et al., 2002; Mukherjee and Bürglin, 2007). In fact, both biochemical and genetic studies have highlighted their differential functions (Berthelsen et al., 1998a; Fognani et al., 2002; Penkov et al., 2013, see below). The PBC-A and -B domains of PBX also display a MEIS-A-like box (Bürglin, 1997), but this feature has so far not received attention although it further indicates the common origin of these protein families.

In mammals, the MEINOX family includes three MEIS (Meis1–3) genes and two PREP (Prep1–2) genes. The two highly conserved MEIS-A and

---

**TABLE 1. Correspondence of TALE Gene Symbols Between Danio Rerio, Mus musculus and Homo sapiens**

| Drosophila melanogaster ancestor | Vertebrate family | Danio Rerio proteins | Mus musculus proteins | Homo sapiens proteins |
|---------------------------------|-------------------|----------------------|-----------------------|-----------------------|
| Hth                             | MEINOX            | Prep (pKnox)         | Prep1                 | PREP1                 |
| Hth                             | MEINOX            | Prep (pKnox)         | Prep2                 | PREP2                 |
| Hth                             | Meis              | Meis1Meis4.1a        | Meis1                 | MEIS1                 |
| Hth                             | Meis              | Meis2.1Meis2.2       | Meis2                 | MEIS2                 |
| Hth                             | Meis              | Meis3               | Meis3                 | MEIS3                 |
| Exd                             | PBC               | Pbx                 | Pbx1                  | PBX1                  |
| Exd                             | PBC               | Pbx                 | Pbx2                  | PBX2                  |
| Exd                             | PBC               | Pbx                 | Pbx4 (lazarus)        | PBX3                  |
| Exd                             | PBC               | Pbx                 | Pbx3                  | PBX3                  |
| Exd                             | PBC               | Pbx                 | *                     | PBX4                  |

*No Danio rerio ortholog of mammalian Pbx4 (Wagner et al., 2001) has been identified. In fact lazarus, although called Pbx4 (Popperl et al., 2000), is an ortholog of mammalian Pbx3 and its function during embryonic development is correlated to the function of mammalian Pbx1.*

---
MEIS-B domains are shown in Figure 1B. The sequences of the regions conserved in MEIS and PREP are shown in Figure 2, i.e., MEIS-A, MEIS-B, and the HD domain. At the C-terminus of the HDs, the amino acids sequence diverges in all proteins. Mammalian cells contain four PBX genes (1–4) that code for six proteins, because of two alternatively spliced isoforms of PBX1 and PBX3. The gene products are extremely well conserved throughout the entire amino acids sequence (Fig. 3). PBX1-3 are about 430 residues long, while PBX4 is shorter, missing the first 78 residues including part of the PBC-A domain and a 30-residue stretch in the C-terminal domain. Starting from the N-terminus, all PBX proteins show high conservation including the homeodomain (HD), identifying the 75-residue-long PBC-A domain, PBC-B of 88, and the HD (Fig. 3). The three domains are very conserved (Table 3 and Fig. 3): PBC-A shows 3–9/75 differences in PBX1-4; PBC-B 2–3/88 differences and the HD 0–5/60 (Table 3). The shorter PBC-A of PBX4 allows for restricting the MEIS/PREP binding region to 51 residues since this domain is functional in PBX4 (Wagner et al., 2001). Moreover, in PBC-B also the five serine residues involved in the regulation of PBX1 nuclear export (Kilstrup-Nielsen et al., 2003) are conserved, together with the four not yet characterized threonines (see below).

An important feature of TALE proteins is the formation of dimers between MEINOX and PBC proteins, even in the absence of DNA. The domains participating in this interaction are MEIS-A and -B in MEINOX, and PBC-A and -B in PBC. In addition, PBC proteins can also form dimers with the anterior HOX proteins but for this interaction the region containing the HD is indispensable (Fig. 1B). Since PBC proteins can interact with MEINOX and HOX proteins through separate interaction surfaces, ternary MEINOX-PBC-HOX complexes are formed that are able to bind DNA (see below). Finally, the MEIS sub-family of the MEINOX proteins can also directly interact with the posterior HOX proteins, again using a different interaction surface. These properties and the specific domains involved will be discussed below.

| Table 2. Gene Structure, Chromosomal Location, Transcript, and Protein Length |
|-----------------------------|-----------|----------------|---------------|-------------|----------------|----------------|----------------|
| Gene          | Chromosome | Gene length (kbp) | Transcript length (bp) | Protein length (aa) | 5’UTR (bp) | N of exons | 3’UTR (bp) | N Isoforms |
|----------------|-------------|------------------|-----------------------|---------------------|-------------|-----------|-------------|-----------|
| Hs-PREP1b    | 21          | 59               | 5,003                 | 436                 | 178         | 11        | 5,324       | 1         |
| Mm-PREP1     | 17          | 41               | 2,583                 | 436                 | 145         | 11        | 1,083       | 1         |
| Dr-PREP1.1   | 9           | 17               | 3,193                 | 435                 | 22          | 11        | 3,908       | 3         |
| Dr-PREP1.2   | 1           | 17               | 2,193                 | 431                 | NK          | 10        | NK          | 2         |
| Hs-PREP2     | 11          | 269              | 3,730                 | 472                 | 314         | 13        | 2,057       | 1         |
| Mm-PREP2     | 9           | 256              | 3,488                 | 474                 | 223         | 13        | 1,840       | 1         |
| Dr-PREP2     | 10          | 19               | 2,106                 | 401                 | NK          | 8         | NK          | 1         |
| Hs-MEIS1     | 2           | 137              | 3,180                 | 415                 | 382         | 13        | 1,550       | 1         |
| Mm-MEIS1     | 11          | 138              | 3,178                 | 390                 | 729         | 13        | 1,276       | 2         |
| Dr-MEIS1     | 13          | 79               | 2,258                 | 391                 | 383         | 14        | 1,700       | 2         |
| Hs-MEIS2     | 15          | 209              | 3,687                 | 381                 | 1076        | 13        | 1,465       | 7         |
| Mm-MEIS2     | 2           | 203              | 3,718                 | 401                 | 1026        | 13        | 1,486       | 7         |
| Dr-MEIS2.1   | 20          | 21.3             | 3,104                 | 393                 | 694         | 14        | 1,184       | 1         |
| Hs-MEIS2.2   | 17          | 99               | 3,169                 | 390                 | 544         | 13        | 1,511       | 1         |
| Mm-MEIS3     | 19          | 16               | 2,149                 | 421                 | 447         | 13        | 436         | 1         |
| Dr-MEIS3     | 7           | 11               | 1,781                 | 378                 | 177         | 13        | 467         | 2         |
| Hs-PBX1      | 1           | 292              | 6,635                 | 430                 | 188         | 9         | 5,159       | 2         |
| Mm-PBX1      | 1           | 272              | 1,293                 | 430                 | NK          | 9         | NK          | 2         |
| Dr-PBX1      | 2           | 32               | 1,029                 | 342                 | NK          | 7         | NK          | 3         |
| Hs-PBX2      | 6           | 5.4              | 3,213                 | 430                 | 271         | 9         | 1,649       | 1         |
| Mm-PBX2      | 17          | 5.1              | 3,068                 | 430                 | 292         | 9         | 1,477       | 1         |
| Dr-PBX2      | 16          | 24               | 1,212                 | 403                 | NK          | 8         | NK          | 1         |
| Hs-PBX3      | 9           | 220              | 2,823                 | 434                 | 67          | 9         | 1,450       | 6         |
| Mm-PBX3      | 2           | 200              | 2,393                 | 434                 | 117         | 9         | 972         | 4         |
| Dr-PBX3B     | 8           | 137              | 1,053                 | 350                 | NK          | 7         | NK          | 3         |
| Hs-PBX4      | 19          | 57               | 1,459                 | 374                 | 24          | 8         | 290         | 1         |
| Mm-PBX4      | 8           | 40               | 2,251                 | 378                 | 38          | 8         | 1,076       | 2         |
| Dr-PBX4      | 3           | 50               | 834                   | 278                 | NK          | 5         | NK          | 1         |

aThe length of the transcript refers to the longest transcript present in the data bank, which is also the most abundant form. Details: the data derived from ENSEMBL database (23.11.2007) except where specified. The length of the genes is counted from the starting point of the transcription. The length of the 5’UTR may not be accurate because in most genes the exact starting point of transcription has not been experimentally determined. In many genes also the 3’UTR has not been established. Hs, Homo sapiens; Mm, Mus musculus; Dr, Danio rerio. NK, not known.

bBased on direct determination of the transcription start site (Bernardi et al., 2010).
MEIS genes can form multiple protein isoforms, due to alternative splicing. MEIS1 comes in two major variants of the C-terminus, MEIS1a and MEIS1b, whose differential biochemical function has not been determined (Oulad-Abdelghani et al., 1997). However, MEIS1b has been specifically found to regulate postnatal cardiomyocyte cell cycle exit (Mahmoud et al., 2013). MEIS1 can also produce HD-less isoforms, which have not been characterized (Crist et al., 2011).

Human and murine MEIS2 are expressed in many different isoforms (8 in Homo sapiens) of various lengths containing essentially all the conserved regions except the isoform “8” that is missing part of the MEIS-A and the isoform “5” lacking part of the HD domain. MEIS2 isoform “3” is required for the activity of a PDX1:PBX1b:MEIS2b complex in pancreatic acinar cells involved in the transcriptional activation of the ELA1 enhancer; the complex binds to the enhancer B element and cooperates with the transcription factor 1 complex (PTF1) bound to the enhancer A element (Liu et al., 2001). Probably in complex with PBX1, MEIS2 isoform d is involved in transcriptional activity of KLF4 (Bjerke et al., 2011). MEIS isoform “2b” has been proposed to act in the transcriptional activation of EPHA8 in the developing midbrain (Shim et al., 2007). Moreover, MEIS isoform “2b” may regulate myeloid differentiation (Fujino et al., 2001).

Mouse Prep1 also produces an alternatively spliced isoform (Fernandez-Diaz and Blasi, unpublished results) but again its function has not been investigated. Both mouse and human PREP1 come with 3'-UTR of two different lengths (unpublished data). Mouse Prep2 was cloned from an E8.5 embryo cDNA library and shown to have three different isoforms by Northern analysis (Haller et al., 2002). Specific antibodies against the C-terminal and the N-terminal part of Prep2 showed five different bands by immunoblotting. Interestingly, the shorter mRNA form of mouse Prep2 produces an HD-less protein (Prep2ΔHD). Although the different Prep2 isoforms have a different subcellular localization, there is no clue about their biological roles (Haller et al., 2004). The production of alternatively spliced isoforms by human PREP2, cloned from HeLa cells RNA, has not been investigated (Fognani et al., 2002).

Also some PBX genes present alternatively spliced protein isoforms, in particular PBX1 and PBX3. PBX1a and 1b, and PBX3a and 3b derive from the same genes but express a protein different in the C-terminus, i.e., after the homeodomain (Monica et al., 1991) (Fig. 4). In H. sapiens, the sequence of PBX1b is 100% identical to PBX1a up to residue 333, lacks the subsequent 92 residues, and then contains 18 additional residues in the C-terminal tail, bringing the total length of PBX1b to 348 residues. The same form is also found in the mouse, where PBX1b includes residues 1–333 of PBX1 with 23 additional residues in the C-terminal region, bringing the overall size again to 347 residues. PBX3, tumor-specific, isoforms have been described (Milech et al., 2001).

Functional differences between alternatively spliced forms has been very little addressed in vertebrates. However, a deeper study is available for the Hth gene of Drosophila melanogaster (Noro et al., 2006). However, since further, still unpublished,
studies are now in progress, we have chosen not to review this issue at this time.

FUNCTIONAL DOMAINS

Both PREP and MEIS proteins (like Hth in *D. melanogaster*) (Johnson et al., 1995) have been shown to interact with PBX, and the interaction is essential for many of their functions. In fact, PREP1 from HeLa cells is purified as a complex with PBX1 and PBX2 (Berthelsen et al., 1998a). This type of interaction requires the N-terminal moiety of PREP1 or MEIS1. Likewise, MEIS can also form similar complexes with PBX (Jacobs et al., 1999).

The sequences required for the interaction with PBX are the MEIS-A and MEIS-B domains that are extremely conserved in all MEIS and PREP sub-family members. The sequence and length of the MEIS-A domain is similar in PREP and MEIS proteins but MEIS3 contains an insert of six residues in the carboxy-terminal half of the MEIS-A domain whose functional relevance is unknown (Fig. 2).

Table 4 summarizes the extent of identity of the three conserved domains of the MEINOX proteins in *H. sapiens*, *M. musculus*, and *D. rerio*. The extent of identity is very high, with the unique exception of the zebrafish Prep1.1, which is slightly less conserved.

MEIS-A and MEIS-B domains are required for the interaction of PREP/MEIS (and Hth) with PBX (Exd) proteins (Knoepfler et al., 1997; Ryoo et al., 1999; Jacobs et al., 1997; Berthelsen et al., 1998b). MEIS-A and MEIS-B domains from PREP or MEIS are thought to interact with the PBC-A and PBC-B domains of PBX; however, structural studies are still not available. While the requirement for MEIS-A seems absolute for the formation of a PREP1-PBX1 complex, MEIS-B, though structurally similar to MEIS-A, has not yet a totally clear function and may be (in part) dispensable for Pbx binding. In fact, deletion of the MEIS-A domain of PREP1 essentially abolishes the interaction (Berthelsen et al., 1998b).

Formation of a MEIS/PREP-PBX complex does not interfere with DNA binding, in fact it stimulates it and increases the selectivity (Knoepfler et al., 1997; Ryoo et al., 1999; Jacobs et al., 1999; Berthelsen et al., 1998b). When a MEIS/PREP-PBX complex binds DNA, both HD are required as the mutation in a single HD is sufficient to prevent binding. Thus, both PREP and MEIS HDs may have to contact DNA.

In PREP1, a binding site for PBX (and MYBBP1A, see below) has been localized to the sequence 64LFPLLALL71 of the MEIS-A domain (Fig. 2) (Diaz et al., 2007) while the overlapping sequence 60YRHPLFPL67 is required for binding 4EHP (see below) (Villaescusa et al., 2009). Mutations in residues Y60, L64, L67, and L68 strongly impair not only PBX1 but also MYBBP1a and 4EHP binding (Diaz et al., 2007; Villaescusa et al., 2009). Binding to MYBBP1a is also a feature of MEIS1a (Dardaei and Blasi, unpublished data). Whether MEIS2, MEIS3, and PREP2
also bind MYBBP1a and 4EHP remains to be investigated.

In addition to MEIS-A and MEIS-B, the MEIS subfamily has three additional internally conserved regions of unknown function: MEIS-C (about 30 aa-long) and MEIS-D (20 residues) immediately upstream and downstream of the HD, respectively, and MEIS-N located at the N-terminus (Mukherjee and Bürglin, 2007).

NUCLEAR LOCALIZATION

As transcription factors, TALE proteins must act in the nucleus; however, in some cells/tissues they have a cytoplasmic localization. The nuclear localization of PBX proteins has been mainly studied in Drosophila. Both Exd and Hth have a nuclear localization signal (NLS), but apparently the one on Hth is dispensable (Rieckhoff et al., 1997; Ryoo et al., 1999). Exd and its homologs have two putative NLSs (NLS1 and NLS2) within the homeodomain (Abu-Shaar et al., 1999; Saleh et al., 2000a) (Fig. 1B). The weaker NLS1 is located in the N-terminal arm (amino acids 234–239: RRKRR) while the stronger NLS2 is in helix 3 (residues 285–294: KRIR-YKKN), and is less conserved. Notice that Figure 1B also indicates the NLS for MEIS. This, however, has been marked because of its correspondence with a canonical NLS sequence (KK/R-X-K/R), but has not been experimentally proven. No clear NLS has been found in PREP1 (Berthelsen et al., 1999).

Fig. 4. The a and b isoforms of human and mouse Pbx1. The “a” forms are 430, the “b,” forms 358 (H. sapiens) and 357 (M. musculus) residues long.

| TABLE 3. Sequence Characterization of Human Pbx Proteinsa |
|----------------------------------------------------------|
| Length of protein or protein domain | PBX1a | PBX2 | PBX3a | PBX4 |
|-----------------------------------|-------|------|-------|------|
| Protein                           | 430   | 430  | 434   | 330  |
| PBC-A                             | 75    | 75   | 75    | 51   |
| PBC-B                             | 88    | 88   | 88    | 88   |
| HD                                | 60    | 60   | 60    | 60   |
| % Identity with PBX1              |       |      |       |      |
| Pbx2                              | 69/75 | 80/88| 59/60 |      |
| Pbx3a                             | 72/75 | 86/88| 60/60 |      |
| Pbx4                              | 66/75 | 75/88| 55/60 |      |

aThe size of the full-length proteins indicates the longest isoform (like PBX1a and PBX3a) or the most frequent of the known isoforms (in the case of PBX2 and PBX4).
import and export, which is mediated by the NLSs and the NES and the nuclear import and export pathway (Stevens and Mann, 2007; Abu-Shaar et al., 1999; Berthelsen et al., 1999; Saleh et al., 2000a). The PBC-A domain of PBX contains a NES-like sequence that, however, is not directly interacting with the exportin but inhibits nuclear localization indirectly, by binding intramolecularly to its own homeodomain, masking the NLSs (Saleh et al., 2000a). Moreover, phosphorylation of serine residues in the PBC-B domain of PBX1 regulates Pbx nuclear localization (Kilstup-Nielsen et al., 2003). The NES sequence of Exd/PBX (IHKKFSSIQM) is highly conserved among Drosophila, M. musculus, and D. rerio.

The nuclear localization of PBX is also dependent on its dimerization status. Indeed, in the absence of Hth, Exd remains in the cytoplasm (Abu-Shaar et al., 1999; Jaw et al., 2000). Moreover, even though PREP1 has no functional NLS, mammalian PBX1 requires PREP1 to be stably localized in the nucleus as PREP1 appears to mask the NES from the exportin (Berthelsen et al., 1999).

Also PREP1 nuclear localization requires dimerization with a PBX protein. Interestingly, in the mouse thymus (which expresses almost uniquely PREP1 and PBX2 among the TALE proteins), expression of a mutant PBX1 missing the NLS but still capable of binding PREP1 is capable to completely deplete PREP1 from the nucleus and localize it in the cytoplasm, with functional consequences similar to the absence of PREP1 (Penkov et al., 2005).

### PREP AND MEIS INTERACTORS

**TALE transcription factors interact** with each other and with proteins of other families. The most studied interactions are between MEINOX and PBC, PBC and HOX, and MEIS and HOX proteins. However, other interactors have been discovered. For some of them, the interaction has been validated in vivo and specific phenotypes have been identified.

MEINOX proteins are major interactors of PBX. This interaction involves a well-defined sequence, the MEIS-A domain, whose tertiary structure is however not known yet. Mutations in the sequence outlined in Figure 2B abolish or strongly decrease the binding, as well as prevent the nuclear localization of both MEINOX and PBC proteins (Diaz et al., 2007; Villaescusa et al., 2009). In the case of PREP1, this same sequence has been shown to bind also to MYBBP1A (Myb-binding protein 1A) and 4EHP (eukaryotic translation initiation factor 4E homolog protein) (see below). These proteins in fact compete for binding to PREP1 and affect its transcriptional activity (Diaz et al., 2007).

Dimerization of PBX with PREP or MEIS affects its stability in vivo. While the detailed mechanism is not known, it has been observed that in the absence of PREP1 the stability of PBX1, PBX2 and MYBBP1a is decreased (Longobardi and Blasi, 2003; Diaz et al., 2007; Oriente et al., 2008).

MYBBP1A is a non-DNA-binding transcriptional regulator of PGC1A (Fan et al., 2004). The binding of PREP1 to MYBBP1A increases protein stability, and indeed its concentration in mouse muscle is PREP1-dose-dependent, whereas its mRNA level is not. Hence, in the absence of PREP1, MYBBP1A fails to inhibit the transcriptional activity of PGC1A in muscle cells, increasing insulin sensitivity (Oriente et al., 2008). MYBBP1A acts as a tumor suppressor and its down-regulation in HeLa cells leads to drastic mitotic abnormalities, a block in G2/M, and increased sensitivity to oncogenic transformation (Mori et al., 2012).

4EHP is a cytoplasmic protein that in mouse oocytes interacts with the MEIS-A domain of cytoplasmic PREP1 (Villaescusa et al., 2009). The interaction of 4EHP with homebox-containing proteins is not unique to PREP1 and might be conserved from Drosophila to man. Indeed, in Drosophila the homeodomain factor Bicoid (Bcd) interacts with 4EHP to repress translation of Caudal (cdx) mRNA and to drive Drosophila embryo development (Cho et al., 2005). In agreement with the presence of a Bcd-like 4EHP-binding sequence in PREP1, cytoplasmic PREP1 binds both 4EHP and the 3′-UTR of at least two Hox mRNAs (HoxB4 and HoxB8), inhibiting their translation, a molecular function that may be important in the maturation of oocytes (Villaescusa et al., 2009).

Other PREP1 interactors include (Table 5) SMAD2-4 (Bailey et al., 2004) and PAX6 (Mikkola et al., 2001), the ribosomal protein S3a/Fte, the splicing cofactors PSF and p54/NRB/NonO, and the cytoskeletal proteins beta-actin and myosin NMMHCIIa (Diaz et al., 2007; Ferrai

### TABLE 4. Amino Acids Sequence Identity in Human, Mouse, and Dr Meinox Protein Domains

| Protein domain | Human/mouse | D. rerio/mouse |
|----------------|-------------|---------------|
| Prep1 MEIS-A   | 31/32       | 26/32         |
| Prep1 MEIS-B   | 45/45       | 38/45         |
| Prep1 HD       | 60/60       | 57/60         |
| Prep2 MEIS-A   | 32/32       | 30/32         |
| Prep2 MEIS-B   | 45/45       | 44/45         |
| Prep2 HD       | 60/60       | 60/60         |
| Meis1 MEIS-A   | 40/40       | 39/40         |
| Meis1 MEIS-B   | 45/45       | 45/45         |
| Meis1 HD       | 60/60       | 59/60         |
| Meis2 MEIS-A   | 40/40       | 39/40         |
| Meis2 MEIS-B   | 45/45       | 44/45         |
| Meis2 HD       | 60/60       | 60/60         |
| Meis3 MEIS-A   | 45/46       | 32/51         |
| Meis3 MEIS-B   | 44/45       | 44/45         |
| Meis3 HD       | 60/60       | 59/60         |

*aFor this comparison, the sequence of D. rerio Meis2.1 has been used.*

*bThe data refer to D. rerio Prep1.*
et al., 2009). Although no structural information is available, some of these factors may interact with PREP1 through the PBX moiety of a PREP1-PBX1 complex (Ferrai et al., 2009; Naum-Ongania et al., unpublished data). The interaction of PREP1 with beta-actin is essential for the retinoic-acid induction of HOX genes expression in human N-TERA2 terato-carcinoma cells. In this context, beta-actin is part of a larger complex containing RNA polymerase II, PREP1, PBX1, the actin polymerizing agent N-WASP, and the splicing cofactors PSF and p54/NRB/NonO. This complex, in fact, allows the polymerization of nuclear actin and its binding to the enhancer of at least the HOXB2 gene, and is required for the initiation of colinear transcription of the HOXB genes (Ferrai et al., 2009).

The molecular details for the requirement of the polymerized nuclear actin in HOX gene expression are, in fact, not known.

Also Oct1 can interact with Prep1 in pull-down experiments in vitro. In this case, it appears that the HD domain is essential for the interaction (Table 5). This finding is rather curious as Oct1 appears to be an interactor of Pbx1 in vivo (Rave-Harel et al., 2004) (see Table 6).

The carboxy-termini of MEIS and PREP are totally divergent. In the case of MEIS, the carboxy terminus is the site for further direct interactions and was previously called transactivation region, i.e., a site for interaction with proteins that will induce activation or repression of target gene expression. For example, MEIS1 C-terminus is required to directly interact with posterior HOX11-13 proteins, a property unique to the MEIS members of the MEINOX family. For this interaction, MEIS uses carboxy-terminal sequences (18 residues in MEIS1 and 93 residues in MEIS2) (Williams et al., 2005). Moreover, MEIS can also bind PBX-HOX complexes (in particular those involving HOXD4 and HOXD5) without directly contacting DNA, again a feature requiring MEIS carboxy-terminal sequences (Shanmugan et al., 1999).

MEIS1 C-terminus harbors transcriptional activation domains that respond to chromatin structure and signaling pathways and that regulate transactivation. Indeed, four regions of the carboxyterminal 56 residues, required for transcriptional activation, are responsive to the HDAC inhibitor TSA and to CBP-dependent Protein Kinase-A. Mutations in all four sites eliminated the response to TSA and to protein kinase-A. C-terminal deletions impair transactivation but do not disturb DNA binding or MEIS-A-mediated formation of HOX or PBX complexes (Huang et al., 2005). Likewise, it has been shown that the 49 C-terminal residues of MEIS1 contain two domains responsible for leukemia induction and HOX A1 gene activation (Mamo et al., 2006).

Not much is known of the carboxy-terminal domains of PREP1 and PREP2 that are not highly conserved. Indeed, addition of the carboxy-terminal domain of MEIS1 to the full-length PREP1 protein transforms it from a tumor suppressor into an oncogene; intriguingly, the same result is

---

**TABLE 5. PREP1 and MEIS1 Interactors**

| PREP1 interactor | Nuclear/cytoplasmic | Interacting site | PREP1 motif^a^ | Interactor motif | Reference |
|------------------|---------------------|-----------------|----------------|-----------------|----------|
| PBX              | N                   | 63LFPLLALL70     | PBCA           | Ferhlesten et al., 1998b |
| Smads 2,3,4      | N                   |                 |                | Diaz et al., 2007  |
| Pax6             | N                   |                 |                | Bailey et al., 2004 |
| p160MBP          | N                   | 63LFPLLALL70     | LXXLL domains, N terminus | Mikkola et al., 2001 |
| 4EHP             | N                   | 58YRHPLFPLL66    |                | Diaz et al., 2007  |
| RFX3             | (2HS)               |                 |                | Ferrai et al., 2009 |
| ZBTB3            | (2HS)               |                 |                | Diaz et al., 2007  |
| TLX1             | (2HS)               |                 |                | Ravasi et al., 2010 |
| βACTIN^b^        | N                   |                 |                | Diaz et al., 2007  |
| NMMHCIIA^b^      | N                   |                 |                | Diaz et al., 2007  |
| Oct1             | N                   | HD              |                | Rave-Harel et al., 2004 |
| p54/Nrb/NonO^b^  | N                   |                 |                | Ferrai et al., 2009 |
| PSF^b^           | N                   |                 |                | Diaz et al., 2007  |
| MEIS1 interactor | Nuclear/cytoplasmic | Interacting site | MEIS1 motif | Interactor motif | Reference |
| PBX              | N                   | aa 64-202       | PBCA           | Chang et al., 1996 |
| HOX9-13          | N                   | C terminal      |                | Williams et al., 2005 |
| PKA              | C                   | C terminal      |                | Huang et al., 2005  |
| TLX1             | (2HS)^c^            |                 |                | Ravasi et al., 2010 |
| TLX2             | (2HS)^c^            |                 |                | Ravasi et al., 2010 |
| EMX2             | (2HS)^c^            |                 |                | Ravasi et al., 2010 |

^aIn bold are the residues whose mutation into Ala drastically decrease binding (Diaz et al., 2007; Villaescusa et al., 2009).

^bIn these proteins it is not known whether they interact with Prep/Meis or Pbx.

^cMammalian 2-Hybrid system.
| Interactor                                      | Notes                                                   | Reacting site in PBX | Reacting site in the interactor | Reference                                |
|------------------------------------------------|---------------------------------------------------------|----------------------|----------------------------------|------------------------------------------|
| **Hox 1–10**                                    | Clustered homeodomain transcription factors             | HD, C terminal       | Hexapeptide (usually YPWMX),     | Piper et al., 1999                      |
| **Prep1**                                       | Pbx regulating protein-1 (pKnox1)                       | PBC-A, N terminal    | 63LFPLALL70                      | Berthelsen et al., 1998b                |
| **Prep2**                                       | Pbx regulating protein-2 (pKnox2)                       | PBC-A, N terminal    | HRI, N terminal                  | Fognani et al., 2002                    |
| **Meis1**                                       | Myeloid Ecotropic Insertion site-1                      | N terminal (aa 1–88) | N terminal (aa 30–60)            | Chang et al., 1996                      |
| **MyoD and bHLH proteins (Myf5, E2a, Myogenin, Meis1)** | Necessary for induction of muscle differentiation.     | HD, C terminal       | CL-X-W motif                     | Knoepfler et al., 1997; Berkes et al., 2004; Maves et al., 2007 |
| **Pdx1**                                        | Forms ternary complex with Pbx and Prep1/Meis1.         | HD, C terminal       | Hexapeptide (FPWMK), N terminal  | Peers et al., 1995; Kim et al., 2002    |
| **NMHCA (non-muscle heavy chain Myosin IIa)**   | Involved in Pbx1 retention in the cytoplasm             | PBC-B                | Coiled-coil tail (aa 161–177)     | Huang et al., 2003                      |
| **Hox11**                                       | Same interaction as for clustered Hox                   | HD, C terminal       | Hexapeptide (FPWME), N terminal  | Allen et al., 2000; Brendolan et al., 2004 |
| **PbxIP1 (Pbx1-interacting Protein1)**          | Inhibits the formation of the Pbx-Hox complexes on DNA | HD (aa 160–230)      |                                   | Abramovich et al., 2000                 |
| **ZF-Pip (Zinc Finger Pbx1 Interacting Protein)** | Prevents HOXA9-Pbx binding to DNA                      | aa 215–241           |                                  | Laurent et al., 2007                    |
| **Rnx**                                         | Forms a DNA binding complex with Pbx3                   | HD, C terminal       |                                  | Rhee et al., 2004                       |
| **Protein Kinase A**                            | Phosphorylate-specific serines in the PBC-B domain      | PBC-B, N terminal    |                                  | Kilstrup-Nielsen et al., 2003           |
| **HDAC1 (histone deacetylase 1)**               | HDACs recruits to Pbx a corepressor complex containing N-CoR/SMRT and mSIN3B | N terminal (aa 89–172) |                                  | Saleh et al., 2000b                     |
| **Oct-1**                                       | Coactivator of Pbx1b                                   | C terminal           |                                  | Rave-Harel et al., 2004                 |
| **FoxC1**                                       | Transcription factor important for neural crest-derived tissue and mesenchymal mesoderm cells. Inhibits Pbx1a transactivation. | N terminal |                                  | Berry et al., 2005                      |
| **P54/Nrb (p54/NonO)**                          | Accessory splicing factor playing a role in actin-dependent transcription-regulating role of Hox | N terminal |                                  | Ferrai et al., 2009; Ongania et al. (unpublished data) |
| **NR3C1 (glucocorticoid receptor)**             | 2HSb                                                   |                      |                                  | Ravasi et al., 2010                     |
not obtained by substitution of the PREP1 C-terminus with that of Meis1 (Bisaillon et al., 2011).

Other uncharacterized interactors of MEIS are LOBE (EPOLM), KLF4, and TLX1 (HOX11). In *Drosophila*, Hth is involved in defining the boundary between the eye and the head cuticle on the ventral margin. In this function, Hth interaction with Lobe, homolog of human EPOLM (epilepsy, occipitotemporal lobe, and migraine with aura) through the MEIS-A domain is required, while the HD is not (Singh et al., 2011). Kruppel-like factor KLF4 is implicated in tumorigenesis and maintaining stem cell pluripotency. PBX1 and MEIS2 homeodomain proteins interact with KLF4 and are recruited to DNA elements comprising a KLF4 site or GC box, with adjacent MEIS and PBX sites (Bjerke et al., 2011). The interaction details are not known. Aberrant expression of the TLX1 (aka HOX11) proto-oncogene is associated with a significant subset of T-cell acute lymphoblastic leukemias (T-ALL). TLX1 and MEIS proteins both interact and are co-expressed in T-ALL (Milech et al., 2010). The details of the interaction are not known but it has also been confirmed by mammalian 2Hybrid analysis, in which both MEIS1 and PREP1 were identified as TLX1 interactors (Ravasi et al., 2010). This interaction is particularly interesting since PBX1 is required in mouse spleen formation where it is involved, among others, in the expression of TLX1, together with a MEINOX protein (Brendolan et al., 2005). The MEINOX protein was suggested to be Prep1, but maybe the newly discovered interaction with Meis1 will bring to light subtle regulatory features, i.e., the possibility that both PREP1-PBX1-TLX1 and MEIS1-PBX1-TLX1 complexes regulate TLX gene expression, possibly in different directions or in different cells.

A recent analysis using mammalian 2Hybrid System has identified, in addition to the classical ones, further novel interactions (Table 5) of PREP and MEIS proteins (Ravasi et al., 2010).

**PBX INTERACTORS**

In the presence of a HOX-PBX-responsive sequences, PBX and most
anterior HOX1-10 proteins form complexes that strongly increase the DNA-binding activity and selectivity of HOX proteins (Chan et al., 1994; Johnson et al., 1995; Merabet et al., 2009). The X-ray structure of DNA-bound PBX1-HOXB1 and PBX1-HOXA9 dimeric homeodomains (Gehringer et al., 1994; Passner et al., 1999; Piper et al., 1999; LaRonde-LeBlanc and Wolberger, 2003) has revealed that each HD binds one half of an octameric DNA sequence, and that the third helix of one half of an octameric DNA double helix. HOX proteins contribute to stabilize this interaction in most cases through short amino acids sequence (the tryptophan-containing hexapeptide) located N-terminally of the HD; in some Drosophila Hox proteins and C-terminally of the HD, a UbdA motif is also involved in the interaction with Exd. PBX proteins, on the other hand, provide HOX-interacting motifs located likewise N-terminally (Chang et al., 1996) and C-terminally (Piper et al., 1999; LaRonde-LeBlanc and Wolberger, 2003). The DNA-bound complex hence forms a structure that embraces the DNA double helix. The DNA-sequence specificity of the PBX-HOX complex is given by the binding of the PBX and HOX moieties to the 5′ and 3′ tetranucleotide of an octanucleotidic sequence of the TGATTXXT type (Gehringer et al., 1994; Piper et al., 1999). The sequence divergence of the HOX proteins HDs and of their target sequence explains how they have evolved to perform more recent unique fusions while keeping common ancestral functions (Sambrani et al., 2013; Hudry et al., 2012; Saadaoui et al., 2011; Slattery et al., 2011a; Lelli et al., 2011).

Full-length PBX1, or any fragment of it, is unable on its own to modulate transcription. However, the 39–232 residues region specifically represses transcription induced by the transcription domain of the SP1 transcription factor, but not of VP16 or p53. C-terminal sequences present in the PBX1a (but not PBX1b) isofrom (see above) block this repressor function (repression domain), the core of which may be a sequence of nine contiguous alanine residues. Interestingly, repression does not require the HD, implying that it is exerted via other proteins and not by competing for target DNA (Knoepfler et al. 1996; Lu and Kamps, 1997).

In complex with HOXB1, PBX1 undergoes a switch from repressor to activator of transcription, upon inhibition of histone deacetylases. Indeed, a region in the amino-terminus of PBX1 (residues 89–172) recruits the repressing HDAC-1 or -3-mSin3B-N-CoR/SMRT complex, while the HOX moiety recruits the CREB co-activator protein (CBP) (Saleh et al., 2000b). Interestingly, this region overlaps also parts of the PBC-A and PBC-B domains. Whether this entails an effect on the interaction with Prep or Meis proteins is not known.

PBX can form ternary complexes with both HOX and PREP/MEIS (Jacobs et al., 1999; Ryoo and Mann, 1999; Ferretti et al., 2000) because different interaction surfaces of PBX are required (the PBC-A domain for PREP/MEIS and the HD for HOX). Not much is known of the structure of the ternary complexes. In order to bind as a trimer, PREP/MEIS, PBX, and HOX select a DNA sequence that can accommodate all three HDs. In vitro experiments have shown the requirement for both an octanucleotide of the HOX-PBX-responsive element-type, and of a hexanucleotide like TGACAG. These sequences are required and functional in responding to the correct cues in vivo (Jacobs et al., 1999; Ryoo et al., 1999; Ferretti et al., 2000, 2005). However, in the case of MEIS, the TGACAG sequence might not be always necessary since MEIS binding to PBX may stabilize the complex without directly contacting the DNA (Shanmugan et al., 1999).

Besides the PBX-MEINOX interactions, many other proteins have been shown to interact with PBX, but their interaction surface is much less defined (Table 6).

**DNA SEQUENCE SELECTIVITY AND TARGET GENES SELECTION IN VIVO**

While a fair amount of knowledge has been accumulated on the genetics and biochemistry of the TALE proteins, the complete picture of their in vivo interactions is still far from complete. Due also to the numerous and redundant members of the families, the definition in vivo of the protein–DNA interactions, the proteins distribution on the genome, its variation in different cell lineages, and the functional outcome of these interactions, are still unanswered questions. Other questions are: is there a preference between PREP and MEIS in forming dimers with PBX and do these recognize the same genes and DNA sequences? While in some cells only (or mostly) one member of a family/subfamily is expressed (for example Prep1 and Pbx2 in mouse thymocytes), in other cells many and even all genes of the families are expressed. Since the sequence of the HD is totally conserved within a subfamily, which of the PREP, MEIS, or PBX proteins is bound to a specific gene? The answer to this question is important since, for example, in the analysis of a specific KO mouse we do not know whether the binding sites on DNA remain free or are occupied by other orthologs. Recent data have allowed a step forward in understanding the complex in vivo biochemistry of this network of transcription factors. Finally, an in-depth analysis of the differences in binding sites between Drosophila in which a single MEINOX gene is present (hth) and mammals, would likely contribute to put the mammalian data into a simpler but functionally significant frame.

**The Target Genes and Recognition Sites**

ChiPseq analysis (Chromatin Immunoprecipitation followed by parallel sequencing of the precipitated DNA) for Prep, Meis, and Pbx1 proteins has recently been performed on the whole trunk tissues of E10.5 mouse embryos (Penkov et al., 2013) identifying several thousand genomic sites (peaks) for each transcription factor. Only a fraction of the sites were bound by more than one factor, and in those cases they represented DNA sequences bound by one specific dimer. This analysis has allowed us to draw some general rules that are summarized in Table 7. Since the antibodies used can recognize both Meis1 and Meis2, these rules apply to the binding of...
both proteins. Likewise, the anti-
Prep1 antibody recognizes in part also Prep2 and hence the same rules may also apply to Prep2.

Prep1-Pbx1 and Meis1/2-Pbx1 DNA-binding peaks frequently overlap whereas binding sites uniquely bound by Meis1/2 or Prep1 very rarely overlap. Alignment of the peaks with the mouse genome sequence showed that Prep1 largely prefers promoters while Meis1 binds preferentially to intra- and intergenic sites, which indicates a different mechanism of action. Moreover, Prep1 appears to be the main partner of Pbx1 as the number of Prep1-Pbx1 peaks is about three-fold more than those for Meis1/2-Pbx1.

ChIPseq analysis allows the identification of the DNA sequences bound by a specific transcription factor. This has revealed that in the mouse embryo the peaks found exclusively by Pbx1 do not have a very strict DNA-binding consensus, whereas those bound by dimers with Prep1 or Meis1 do (see below) (Penkov et al., 2013). Since Pbx1 may have additional DNA-binding interactors, the absence of a binding site consensus may also depend on dimerization with others partners that, as Prep1 and Meis1, may direct Pbx1 to their specific binding site, a possibility that has not yet been explored.

Prep1 and Meis1-2 select the same specific DNA sequences both when binding alone and when in combination with Pbx1 (Penkov et al., 2013). Prep1 and Prep1-Pbx1 bind preferentially a decameric consensus with a general structure TGAXTGACAG, as expected (Knoepfler et al., 1997), whereas Meis1-Pbx1 binds mostly both an octameric TGATTGXX and an hexameric TGA-
CAG sequence. Hence, the DNA sequence selectivity resides in Prep1 or Meis1, may direct Pbx1 to their specific binding site, a possibility that has not yet been explored.

Prep1 and Meis1-2 target preferentially embryonic development genes whereas Prep1 targets mainly basic cellular functions. Despite the low level of overlap and the clearly different and separate functional categories, Prep and Meis still show a low level coordination, for example, in the regulation of expression of Hox genes (Penkov et al., 2013).

The binding sites identified by the above study indeed confirm previously recognized regulatory regions. Meis1/2, Pbx1, and Prep1 bind several times in the Hox clusters often together, suggesting an important degree of cross-talk. Meis peaks are the most abundant, being most frequent in the HoxA and least frequent in the HoxC cluster. Pbx1 shows less binding sites and Prep1 the lowest number of binding sites among the three factors. In all cases Pbx1, Meis and Prep1 peaks coincide. Interestingly, all peaks are concentrated in the 1 to 9 paralog region, subdividing the Hox clusters in Meis/Pbx/Prep-interactive and Meis/Pbx/Prep-non-interactive regions. Previous biochemical and genetic studies had identified 6 TALE protein-bound regions as important regulatory sites in the Hox clusters, often in cooperation with HoxB1 and functional in auto- and cross-regulatory interactions (Gould et al., 1997; Jacobs et al., 1999; Lampe et al., 2008; Manzanares et al., 2001; Popperl et al., 1995; Tumpel et al., 2007). Despite that those studies were carried out at a different developmental stage and in specific tissues, 4 out of the 6 regions were found to coincide with the described regions (Penkov et al., 2013).

One other important novel finding obtained in such a study relates to the
octameric TGGATGXX sequence that is present in a large fraction of the peaks bound by Meis1 (exclusively or in combination with Pbx1). Interestingly, these peaks identify a consensus sequence previously thought to bind Pbx-Hox (Piper et al., 1999). This suggests that Hox are major partners of Meis1 in addition to Pbx1. Indeed, analysis of the recently published ChIP seq analysis for Hoxa2 and Hoxc9 (Jung et al., 2010; Donaldson et al., 2012) reveals that a rather large percent of these binding sites overlaps with Meis1 peaks, and that this is more frequent than with Prep1 or Pbx1 (Penkov et al., 2013). Therefore, a large fraction of the Meis peaks corresponds in fact to Hox target genes. The fact that Hox peaks’ identification was carried out in different tissues does not affect this conclusion. In agreement with the above finding, cytological analysis of genome-wide chromosomal binding sites of Drosophila Hth (Cohen and Salzberg, 2008), and ChIP-chip analysis of Ubx and Hth-bound regions in the haltere and T3 leg imaginal discs (Slattery et al., 2011b) has revealed a remarkable amount of tissue-specificity of binding and of tissue-specific overlap. Unfortunately, data on the co-binding of Exd are still outstanding.

The amino acids sequence of the HDs within the various sub-families is nearly identical; hence, one might expect that in the absence of a specific TALE protein another one of the same sub-family would take its place. For example, in the absence of Pbx1 a different Pbx isoform might be bound to the Pbx1 sites. Indeed, a ChIP-on-chip analysis of mouse thymocytes and T3 leg-bound regions in the haltere and T3 leg imaginal discs (Slattery et al., 2011b) has revealed a remarkable amount of tissue-specificity of binding and of tissue-specific overlap. Unfortunately, data on the co-binding of Exd are still outstanding.

The amino acids sequence of the HDs within the various sub-families is nearly identical; hence, one might expect that in the absence of a specific TALE protein another one of the same sub-family would take its place. For example, in the absence of Pbx1 a different Pbx isoform might be bound to the Pbx1 sites. Indeed, a ChIP-on-chip analysis of mouse thymocytes and T3 leg-bound regions in the haltere and T3 leg imaginal discs (Slattery et al., 2011b) has revealed a remarkable amount of tissue-specificity of binding and of tissue-specific overlap. Unfortunately, data on the co-binding of Exd are still outstanding.

The amino acids sequence of the HDs within the various sub-families is nearly identical; hence, one might expect that in the absence of a specific TALE protein another one of the same sub-family would take its place. For example, in the absence of Pbx1 a different Pbx isoform might be bound to the Pbx1 sites. Indeed, a ChIP-on-chip analysis of mouse thymocytes and T3 leg-bound regions in the haltere and T3 leg imaginal discs (Slattery et al., 2011b) has revealed a remarkable amount of tissue-specificity of binding and of tissue-specific overlap. Unfortunately, data on the co-binding of Exd are still outstanding.

The amino acids sequence of the HDs within the various sub-families is nearly identical; hence, one might expect that in the absence of a specific TALE protein another one of the same sub-family would take its place. For example, in the absence of Pbx1 a different Pbx isoform might be bound to the Pbx1 sites. Indeed, a ChIP-on-chip analysis of mouse thymocytes and T3 leg-bound regions in the haltere and T3 leg imaginal discs (Slattery et al., 2011b) has revealed a remarkable amount of tissue-specificity of binding and of tissue-specific overlap. Unfortunately, data on the co-binding of Exd are still outstanding.

The amino acids sequence of the HDs within the various sub-families is nearly identical; hence, one might expect that in the absence of a specific TALE protein another one of the same sub-family would take its place. For example, in the absence of Pbx1 a different Pbx isoform might be bound to the Pbx1 sites. Indeed, a ChIP-on-chip analysis of mouse thymocytes and T3 leg-bound regions in the haltere and T3 leg imaginal discs (Slattery et al., 2011b) has revealed a remarkable amount of tissue-specificity of binding and of tissue-specific overlap. Unfortunately, data on the co-binding of Exd are still outstanding.
modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. Mol Cell Biol 16:1734–1745.

Chen H, Varela C, Nakamura Y, Lyon A, Chakravarti A, Antonarakis S. 1997. Cloning of a novel homebox-containing gene, PKNX1, and mapping to chromosome 21q22.3. Genomics 41:193–200.

Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD. 2005. A new paradigm for translational control: inhibition via 59–39 mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell 121:411–423.

Cohen L, Salzberg A. 2008. Chromosomal binding sites of the homeotic cofactor Homothorax. Mol Genet Genom 280:73–81.

Crijns AP, de Graaff P, Geerts D, Ten Hoor KA, Hollema H, van der Sluis T, Hofstra RM, de Bock GH, de Jong S, van der Zee AG, de Vries EG. 2007. MEIS and PBX homebox proteins in ovarian cancer. Eur J Cancer 43:2495–2505.

Crist RC, Roth JJ, Handschin C, Donaldson IJ, Amin S, Hensman JJ, Di Rosa P, Villaescusa JC, Longobardi E, Crist RC, Roth JJ, Waldman SA, Crippa MP, Scita G, Blasi F. 2009. Longobardi E, Disanza A, Diaz VM, Jenkins NA, Fiorenza MT, Copeland S, Turco M, Longobardi E, Iotti G, Ferretti E, Cambronero F, Tumpel S, Weidemann L, Blasi F, Krumlauf R. 2005. The absence of Hoxb1 enhancer and control of rhombomere 4 expression: Complex interplay between PREP1-PBX1-HOXB1 binding sites. Mol Cell Biol 25:8541–8552.

Ferretti E, Villaescusa JC, Di Rosa P, Fernandez-Diaz LC, Longobardi E, Mazzieri R, Micco A, Miceli N, Selleri L, Ferrari G, Blasi F. 2006. Human PREP-2, a novel interactor of PBX proto-oncogene, defines a novel subfamily of TALE homeodomain transcription factors. Nucleic Acid Res 30:2043–2051.

Fognani C, Kilstrup-Nielsen C, Ferretti E, Zappavigna V, Blasi F. 2002. Human PREP, MEIS AND PBX BIOCHEMISTRY 73

Huang H, Palouras M, Rambaldi I, Lasko P, Featherstone M. 2003. Non-muscle myosin promotes cytoplasmic localization of PBX. Mol Cell Biol 23:3652–3665.

Huang H, Rastegar M, Bodner C, Goh SL, Rambaldi I, Featherstone M. 2005. MEIS C termini harbor transcriptional activation domains that respond to cell signaling. J Biol Chem 280:10119–10127.

Hudry B, Remacle S, Delfini MC, Rezsohazy R, Graba Y. 2012. Hox proteins display a common and ancestral ability to diversify their interaction mode with the PBC class cofactors. PLoS Biol 10:e1001351.

Iotti G, Longobardi E, Masella S, Dardaei L, De Santis F, Miceli N, Blasi F. 2011. The homeodomain transcription factor PREP1 is required to maintain genomic stability. Proc Natl Acad Sci USA 108: E314–E322.

Jegers W, Schnabel CA, Cleary ML. 1999. Trimeric association of Hox and TALE homeodomain proteins mediates Hoxb2 hindbrain enhancer activity. Mol Cell Biol 19:134–142.

Jaw Tj, You LR, Knoepfner PS, Yao LC, Pai CY, Tang CY, Chang LP, Berthelsen J, Blasi F, Kamps MP, Sun YH. 2000. Direct interaction of two homeoproteins, homothorax and extradenticle, is essential for EXD nuclear localization and function. Mech Dev 91:279–291.

Johnson FB, Parker E, Krasnow MA. 1995. Extradenticle protein is a selective cofactor for the Drosophila homeotic molar phenotype, role of the homeodomain and YYPWM amino acid motif in the interaction. Proc Natl Acad Sci USA 92:739–743.

Jung H, Lacombe J, Mazzoni EO, Liem Jr, Grinstein J, Mahony S, Mukhopadhyay D, Gifford DK, Young RA, Anderson K. 2010. Global control of motor neuron topography mediated by the repressive actions of a single hox gene. Neuron 67:781–796.

Kamps MP, Murre C, Sun XH, Baltimore D. 1990. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. Cell 60:547–555.

Kilstrup-Nielsen C, Alessio M, Zappavigna V. 2003. PBX1 nuclear export is regulated independently of PBX-MEINOX interaction by PKA phosphorylation of the PBC-B domain. EMBO J 22:89–99.

Kim SK, Selleri L, Lee JS, Zhang AY, Gu X, Jacobs Y, Cleary ML. 2002. Pbx1 inactivation disrupts pancreas development and in Ifp-deficient mice promotes diabetes mellitus. Nat Genet 30:430–435.
MesI1 and pKnoX1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. Proc Natl Acad Sci USA 94: 8795–8799.

Lampe X, Samad OA, Guiguean A, Matis C, Remacle S, Picard JJ, RiMI, Rokzehazay R. 2008. An ultrasconserved Hox-Pbx responsive element resides in the coding sequence of Hoxa2 and is active in the VegT regulon. Nucleic Acids Res 36:3214–3225.

Laronde-LeBlanc NA, Wolberger C. 2003. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. Genes Dev 17:2060–2072.

Laurent A, Bihan R, Deschamps S, Nuciforo P, Ponzoni M, Doglioni C, Remacle S, Picard JJ, Rijli FM. 2006. Independent regulation of initiation anterior to posterior. Genes Dev 134:2060–2072.

Lelli KM, Noro B, Mann RS. 2011. Variations of the Pbx homedomain-less isoforms encoded by dna-independent recruitment of different homeodomain proteins. J Biol Chem 286:39235–39241.

Lelli KM, Noro B, Mann RS. 2011. Variational motif utilization in homeotic selector (Hox)-cofactor complex formation controls specificity. Proc Natl Acad Sci USA 108:21122–21127.

Lu Y, MacDonald RJ, Swift GH. 2001. DNA binding and transcriptional activation by a PDX1.PBX1b. MEIS2.b trimer and cooperation with a pancreas-specific basic helix-loop-helix complex. J Biol Chem 276:17985–17993.

Longobardi E, Blasi F. 2003. Overexpression of PREP-1 in F9 teratocarcinoma cells leads to a functionally relevant increase of PDX-2 by preventing its degradation. J Biol Chem 3:38235–39241.

Longobardi E, Iotti G, Di Rosa P, Miettta S, Nicotaro P, Ponzoni M, Doglioni C, Caniatti P, Bianchi F, Di Fiore PP, Blasi F. 2010. The homeodomain transcription factor Prep1 gene (pKnoX1) is a haploinsufficient tumor suppressor in man and mice. Mol Oncol 4:226–234.

Lu Q, Kamps MP. 1997. Heterodimerization of Hox proteins with Pbx1 and oncoprotein E2a-Pbx1 generates unique DNA-binding specificities at nucleotides predicted to contact the N-terminal arm of the Hox homeodomain: demonstration of Hox-dependent targeting of E2a-Pbx1 in vivo. Oncogene 14:75–85.

Mahmoud AI, Kocbas F, Mundalihar SA, Kimura W, Koura AS, Tset S, Porrello ER, Sadek HA. 2013. MesI1 regulates postnatal cardiomyocyte cell cycle arrest. Nature 497:249–253.

Mario A, KroI J, Kroas E, Bijl J, Thoorel A, Mayotte N, Girard S, Bisiaux R, Besiu N, Featherstone M, Sauvageau G. 2006. Molecular dissection of MesI1 reveals 2 domains required for leukemia induction and a key role for Hoxa gene activation. Blood 108:622–629.

Mananares M, Bel-Vialar S, Ariza-MeNaughton L, Ferretti E, Marshall H, Macconechie MM, Blasi F, Krumlauf R. 2001. Independent regulation of initiation and maintenance phases of Hoxa3 expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms. Development 128:3595–3607.

Maves L, Waskiewicz AJ, Paul B, Cao Y, Tyler A, Moens CB, Tapsestoff AJ. 2007. Pbx homeodomain proteins direct Myod activity to promote fast-muscle differentiation. Development 134:3371–3382.

Merabet S, Hadry B, Saadassui M, Graba Y. 2005. Characterization of sequence signatures: a guide to Hox protein function. Bioessays 31:500–511.

Mikkola I, Bruun JA, Holm T, Johansen T. 2001. Superactivation of Pax6-mediated transactivation from paired domain-binding sites by DNA-binding specificities at nucleotides. Mol Cell 7:2060–2072.

Milech N, Kees UR, Watt FM. 2001. Novel alternative PBX3 isoforms in leukemia cells with distinct interaction specificities. Genes Chromosomes Cancer 32:275–280.

Milech N, Gottardo NG, Ford J, D’Souza D, Greene WK, Kees UR. Watt FM. 2010. MEIS proteins as partners of the TLX1/HOX11 oncoprotein. Leuk Res 34:358–363.

Monica K, Galili N, Nourse J, Saltman D, Cleary ML. 1991. PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-oncogene PBX1. Mol Cell Biol 11:6149–6157.

Mori S, Bernardi R, Laurent A, Resnati M, Crippa A, Gabrieli A, Keough R, Gonda TJ, Blasi F. 2012. Myb-binding protein 1a (MYBBP1A) is developmentally essential, acts as a tumor suppressor and is required in the G2/M phase of the cell cycle. PLOS ONE 7:e39723.

Moskov JJ, Bullrich F, Huebner K, Daar IO, Buechber AM. 1995. MesI1, a Pbx1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. Mol Cell Biol 5:5434–5443.

Mukherjee K, Batista R, Kiehl D, Greene WK, Kees UR, Watt FM. 2010. MEIS proteins as partners of the TLX1/HOX11 oncoprotein. Leuk Res 34:358–363.

Nakamura T, Largesaedpa DA, Shaunghnessy JD Jr, Jenkins NA, Copeland, NG. 1996. Cooperative activation of Hoxa and Pbx1-related genes in murine myeloid leukemia. Nat Genet 12:149–153.

Noro B, Culi J, McKay DJ, Zhang W, Mann RS. 2006. Distinct functions of Hox dependent targeting of E2a-Pbx1 in vivo. Oncogene 14:75–85.

Nourore J, Melltentin JD, Galili N, Wilkinson J, Stanbridge E, Smith SD, Cleary ML. 1990. Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell 60:535–545.

Oriente F, Fernandez Diaz LC, Rosello CA, Torroja T. 2003. Superactivation of Pax6-homeodomain-less isoforms encoded by dna-independent recruitment of different homeodomain proteins. J Biol Chem 276:4109–4118.

Oriente F, Fernandez Diaz LC, Rosello CA, Torroja T. 2003. Superactivation of Pax6-homeodomain-less isoforms encoded by dna-independent recruitment of different homeodomain proteins. J Biol Chem 276:4109–4118.

Peers B, Sharma S, Johnson T, Kamps M, Montminy M. 1995. The pancreatic islet factor STF-1 binds cooperatively with Pbx to a regulatory element in the somatostatin promoter: importance of the FPWMK motif and of the homeodomain. Mol Cell Biol 15:7091–7097.

Penkov D, Di Rosa P, Fernandez Diaz L, Basso V, Ferretti E, Grassi F, Mondino A, Blasi F. 2005. Involvement of Prep-1 in the abTCR-T-lymphocytic potential of hematopoietic precursors. Mol Cell Biol 25:7689–7681.

Penkov D, Mataos San Martin D, Fernandez-Diaz LC, Rosello CA, Torroja C, Sanchez-Cabo F, Warnatz HJ, Sultan M, Yaspo ML, Gabrieli A, Tkachuk V, Brendolan A, Blasi F, Torres M. 2013. Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. Cell Rep 25:1321–1333.

Piper DE, Batchelor AH, Chang CP, Cleary ML, 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.

Popperl H, Bienen M, Stuven M, Chan SK, Aparicio S, Brambilla F, Mansour MS, Krumlauf R. 1995. Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx. Cell 81:1031–1042.

Popperl H, Rikhof H, Chang H, Hafer H, Kimmel CB, Moens CB. 2000. Lazarus is a novel pbx gene that globally mediates hox gene function in Zebrafish. Mol Cell 6:255–267.

Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, Tan K, Akalin A, Schmeier S, Kanamori-Katayama M, Bertin N, Carninci P, Daub CO, Forrest AR, Gough J, Grimmond S, Han JH, Hashimoto T, Hide W, Hofmann O, Kamburov A, Kaur M, Kawaji H, Kauppi MA, Langmann T, van Nimwegen E, MacPherson CR, Ogawa C, Radovanovic A, Schwartz A, Teasdale RD, Tegner J, Lenhard B, Teichmann SA, Arakawa T, Ninomiya N, Murakami K, Tazaki S, Fukuda S, Imanura K, Rie C, Ishihara H, Kitazume Y, Kawai J, Hume DA, Ideker T, Hayashizaki Y. 2010. An atlas of combinatorial transcriptional regulation in mouse and man. Cell 140:744–752.
PREP, MEIS AND PBX BIOCHEMISTRY 75

Developmental Dynamics

Rieckhoff G., Casares F, Ryoo HD, Abu-Rhee IW, Arata A, Selleri L, Jacobs Y, Rave-Harel N, Givens ML, Nelson SB, Sambrani N, Hudry B, Maurel-Zaffran C, Saleh M, Rambaldi I, Yang XY, Saleh M, Huang H, Green NC, Saadaoui M, Merabet S, Litim-Mecheri I, Ryoo HD, Marty T, Casares F, Affolter M, Pax6 via enhancer binding site affinity. Cell 91:171–183.

extradenticle-related homeodomain protein. Cell 91:171–183.

ectodermal change in PBX1A is necessary for its nuclear localization. Exp Cell Res 260:105–115

adult stem cells. Cell Stem Cell 4:129–140.

expression independently and via interactions with Oct-1. J Biol Chem 279:30287–30297

Rhee IW, Arata A, Selleri L, Jacobs Y, Arata S, Onimaru H, Cleary ML. 2004. The control of trunk Hox specificity and activity by Extradoenticle. Genes Dev 18:1704–1716.

Rowan S, Siggers T, Lachke SA, Yue Y, Bulyk ML, Maas RL. 2010. Precise temporal control of the eye regulatory gene Pox6 via enhancer binding site affinity. Genes and Dev 24:980–985

Ryoo HD, Mann RS. 1999. The control of trunk Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. Development 126:5137–5148.

Saadaoui M, Merabet S, Litim-Mecheri I, Arbeille E, Sambrani N, Damen W, Brena C, Pradel J, Graba Y. 2011. Selection of distinct Hox-Extradenticle interaction modes fine-tunes Hox protein activity. Proc Natl Acad Sci USA 108:2276–2281

Saleh M, Huang H, Green NC, Featherstone MS. 2000a. A conformational change in PBX1A is necessary for its nuclear localization. Exp Cell Res 260:105–115

Saleh M, Rambaldi I, Yang XY, Featherstone MS. 2000b. Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone desacytlases and histone acetyltransferases. Mol Cell Biol 20:8623–8633.

Sambrani N, Hudry B, Maurel-Zaffran C, Zouaz A, Mishra R, Merabet S, Graba, Y. 2013. Distinct molecular strategies for Hox-mediated limb suppression in Drosophila: from cooperativity to dispensability/antagonism in TALE partnerships. PLoS Genet 9:e1003307.

Selleri L, Depew MJ, Jacobs Y, Chanda SK, Tsang KY, Cheah KS, Rubenstein JL, O’Gorman S, Cleary ML. 2001. Requirement for Pbx1 in skeletal patterning and programming chondrocyte proliferation and differentiation. Development 128:3543–3557.

Selleri L, DiMartino J, Van Deursen J, Brendolan A, Sanyal M, Boon E, Capellini T, Smith KS, Rhee J, Popperl H, Grosveld G, Cleary ML. 2004. The TALE homeodomain protein Pbx2 is not essential for development and long-term survival. Mol Cell Biol 24:5324–5331.

Shanmugam K, Green NC, Rambaldi I, Saragovi HU, and Featherstone MS. 1999. PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. Mol Cell Biol 19:7577–7588.

Shim S, Kim Y, Shin J, Kim J, Park S. 2007. Regulation of Epha8 gene expression by TALE homeobox transcription factors during development of the mesencephalon. Mol Cell Biol 27:1614–1630.

Singh A, Tare M, Kango-Singh M, Son SK, Kim Y, Shin J, Park S. 2007. Expression of Hoxa2 in rhombomere 4 is regulated by a conserved cross-regulatory mechanism dependent upon Hoxb1. Dev Biol 302:646–660.

Vaccari E, Delforian G, Bernardi E, Pauls S, Tso N, Bortolus M, Argenton F. 2010. prep1.2 and aldh1a2 participate to a positive loop required for branchial arches development in zebrafish. Dev Biol 343:94–103.

Villaescusa JC, Buratti C, Penkov D, Mathiasen L, Planagumà J, Ferretti E, Blasi F. 2009. Cytoplasmic Prep1 interacts with 4EHP inhibiting HoxB4 translation. PLoS One 4:e5213.

Wagner K, Mincheva A, Korn B, Lichter P, Popperl H. 2001. Pbx4, a new Pbx family member on mouse chromosome 8, is expressed during spermatogenesis. Mech Dev 103:127–131.

Wang GG, Pasillas MP, Kamps MP. 2006. Persistent transactivation by meis1 replaces hox function in myeloid leukemia-mogenesis models: evidence for co-occupancy of meis1-pbx and hox-pbx complexes on promoters of leukemias-associated genes. Mol Cell Biol 26:3902–3916.

Williams TM, Williams ME, Innis JW. 2005. Range of HOX/TALE superclass associations and protein domain requirements for HOX13:MEIS interaction. Dev Biol 277:457–471.

Wong P, Iwasaki M, Somerville TC, So CW, Cleary ML. 2007. Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. Genes Dev 21:2762–2774.