The EMBO Journal Vol. 19 No. 10 pp. 2292–2303, 2000

Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family

Dirk Eberhard, Gerardo Jiménez[1], Barry Heavey and Meinrad Busslinger[2]

Research Institute of Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria and [1]Departamento de Biologia Molecular i Cellular, CID-CSIC, Jordi Girona 18–26, 08034 Barcelona, Spain

[2]Corresponding author
e-mail: Busslinger@nt.imp.univie.ac.at

Pax5 (BSAP) functions as both a transcriptional activator and repressor during midbrain patterning, B-cell development and lymphomagenesis. Here we demonstrate that Pax5 exerts its repression function by recruiting members of the Groucho corepressor family. In a yeast two-hybrid screen, the *groucho*-related gene product Grg4 was identified as a Pax5 partner protein. Both proteins interact cooperatively via two separate domains: the N-terminal Q and central SP regions of Grg4, and the octapeptide motif and C-terminal transactivation domain of Pax5. The phosphorylation state of Grg4 is altered *in vivo* upon Pax5 binding. Moreover, Grg4 efficiently represses the transcriptional activity of Pax5 in an octapeptide-dependent manner. Similar protein interactions resulting in transcriptional repression were also observed between distantly related members of both the Pax2/5/8 and Groucho protein families. In agreement with this evolutionary conservation, the octapeptide motif of Pax proteins functions as a Groucho-dependent repression domain in *Drosophila* embryos. These data indicate that Pax proteins can be converted from transcriptional activators to repressors through interaction with corepressors of the Groucho protein family.

Keywords: Groucho (Grg)/interaction partners/Pax5 (BSAP)/transcriptional repression

Introduction

The Pax proteins constitute a family of paired domain-containing transcription factors that play essential roles in early development from *Drosophila* to man. Based on sequence similarities, these *Pax* genes can be grouped into four subfamilies (Noll, 1993). One of them consists of a single *Drosophila* member, *dPax258* (*spa*) (Czerny *et al*., 1997; Fu and Noll, 1997), and three mammalian genes, *Pax2*, *Pax5* and *Pax8*, which arose by gene duplications at the onset of vertebrate evolution (Pfeffer *et al*., 1998).

Pax5 codes for the transcription factor BSAP, which is essential for brain patterning and B-lymphopoiesis (reviewed by Busslinger and Nutt, 1998). *Pax5* cooperates with *Pax2* in development of the midbrain and cerebellum, in agreement with the overlapping expression patterns of the two genes at the midbrain–hindbrain boundary of the mouse embryo (Urbánek *et al*., 1997). In contrast, *Pax5* is the only member of the *Pax* family that is expressed in B-lymphocytes (Adams *et al*., 1992). As a consequence, B-cell development is arrested at an early pro-B cell stage in mice lacking *Pax5* (Urbánek *et al*., 1994). Surprisingly, the Pax5-deficient pro-B cells are not yet restricted in their lineage fate. Instead, these cells are able to differentiate into various myeloid and lymphoid cell types *in vitro* as well as *in vivo* and thus retain a broad developmental potential characteristic of an uncommitted hematopoietic progenitor cell (Nutt *et al*., 1999; Rolink *et al*., 1999). Pax5-deficient pro-B cells are, however, able to develop along the B-lymphoid lineage once *Pax5* expression has been restored by retroviral transduction. These experiments therefore identified Pax5 as the critical B-lineage commitment factor that restricts the developmental potential of progenitor cells to the B-lymphoid pathway by suppressing alternative cell fates (Nutt *et al*., 1999; Rolink *et al*., 1999).

Insight into the transcriptional role of Pax5 has been provided by the identification of target genes, which was facilitated by the *in vitro* culture of Pax5-deficient pro-B cells and the development of a Pax5-specific induction system (Nutt *et al*., 1998). Pax5 was thus shown to activate *CD19*, *Igα* (*mb-1*), *LEF-1* and N-*myc* expression and simultaneously to repress *PD-1* transcription (Nutt *et al*., 1998). Pax5 therefore fulfills a dual role in early B-cell development, as it functions as both an activator and repressor of gene transcription. The repression function of Pax5 is particularly important for the suppression of alternative lineage fates at B-lineage commitment, which is best illustrated by the regulation of the *M-CSF-R* gene. This myeloid gene is one of several hematopoietic genes that are expressed promiscuously in the uncommitted Pax5-deficient pro-B cell. Upon commitment, Pax5 represses *M-CSF-R* transcription, thus rendering B-cell precursors unresponsive to the myeloid cytokine M-CSF (Nutt *et al*., 1999). Furthermore, Pax5 has been implicated in the repression of the J-chain gene and down-regulation of the activity of immunoglobulin 3′ enhancers during late B-cell differentiation (reviewed by Busslinger and Nutt, 1998).

Structure–function analyses are also consistent with a dual role for Pax5 in transcriptional regulation (see Figure 1A). Pax5 is known to recognize target genes via its N-terminal paired domain and to control transcription through a C-terminal regulatory module consisting of activating and inhibitory sequences (Dörfler and Busslinger, 1996). Furthermore, Pax5 possesses a characteristic octapeptide that was identified originally as a conserved sequence motif found in most Pax proteins (Burri *et al*., 1989; Noll, 1993). The presence of this octapeptide motif was subsequently shown to down-modulate the transcriptional activity of Pax proteins

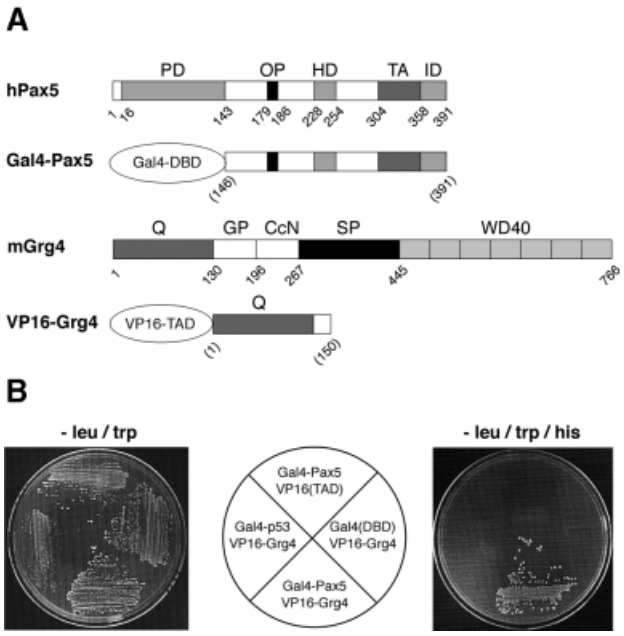


Fig. 1. Interaction of Pax5 with the N-terminal region of Grg4 in yeast. (**A**) Schematic diagram of Pax5, Grg4 and the respective fusion proteins. The different domains of each protein are indicated together with the corresponding amino acid positions. The chimeric Gal4–Pax5 protein was used as a bait to isolate the VP16-Grg4 fusion protein in a yeast two-hybrid screen of a cDNA expression library that was generated by fusing the transactivation domain (TAD) of VP16 to cDNA derived from 9.5/10.5-day-old mouse embryos (Hollenberg *et al*., 1995). See text for description of the Grg4 structure. PD, paired domain; OP, octapeptide; HD, partial homeodomain; TA, transactivation region; ID, inhibitory domain; DBD, DNA-binding domain. (**B**) Specific interaction of Pax5 and Grg4 in yeast. Expression plasmids of the combinations indicated (central panel) were transformed into yeast and selected for by growth on plates lacking leucine and tryptophan (left panel). Activation of a Gal4-dependent *HIS3* gene was examined by growth on selection plates additionally lacking histidine (right panel).

(Lechner and Dressler, 1996). Pax5, like other members of the Pax2/5/8 family, also contains a partial homeodomain that constitutes an interaction surface for both the retinoblastoma (Rb) gene product and the TATA-binding protein (Eberhard and Busslinger, 1999).

As the function of DNA-binding transcription factors is determined by the interaction with cofactors, we have employed the yeast two-hybrid assay to search systematically for Pax5 partner proteins. Here, we describe the identification and characterization of Grg4 as a Pax5 interaction partner. Grg4, which is also known as TLE4, is one of four members of the mammalian Groucho family (Stifani *et al*., 1992; Koop *et al*., 1996). The founding member of this conserved protein family is the *Drosophila* Groucho protein, which is broadly expressed throughout development and plays important roles in diverse processes such as sex determination, segmentation and neurogenesis (Paroush *et al*., 1994). The Groucho proteins consist of several conserved domains including the N-terminal glutamine-rich Q region and C-terminal WD40 repeats (see Figure 1A). Although the Groucho proteins are localized in the nucleus, they lack any recognizable DNA-binding motif. Instead, these proteins interact with various DNA-binding transcription factors and are thus recruited to specific control regions where they function as potent corepressors to inhibit gene transcription (reviewed by Fisher and Caudy, 1998; Parkhurst, 1998). The Groucho proteins form higher order complexes by tetramerizing via the Q domain (Chen *et al*., 1998), bind to the N-terminal tails of histone H3 (Palaparti *et al*., 1997), interact with histone deacetylases (Chen *et al*., 1999; Choi *et al*., 1999) and thus seem to exert their function as part of multiprotein–DNA complexes that locally establish a repressive chromatin structure.

Here we demonstrate by protein binding and co-immunoprecipitation assays that Pax5 is able to interact with the Grg4 protein *in vitro* as well as *in vivo*. Both proteins contact each other via two separate interaction domains. The N-terminal Q domain and central SP region of Grg4 interact cooperatively with the C-terminal transactivation domain and octapeptide motif of Pax5, respectively. Moreover, the interaction with Pax5 induces a specific change in the phosphorylation state of Grg4. As shown by transient transfection assays, Grg4 can repress the transcriptional activity of Pax5 efficiently in an octapeptide-dependent manner. Similar protein interactions resulting in repression were also observed between distantly related members of the Pax2/5/8 and Groucho protein families. As predicted by this evolutionary conservation, the octapeptide motif was shown to function as a Groucho-dependent repression domain in *Drosophila* embryos. Together, these data indicate that Pax proteins can function as active repressors by recruiting corepressors of the Groucho family to selected target genes, thus offering a molecular explanation of how Pax5 represses the transcription of non-B-lymphoid genes at B-lineage commitment.

Results

Identification of Grg4 as an interaction partner of Pax5

We employed the yeast two-hybrid system to search for potential cofactors of Pax5. To this end, a Gal4–Pax5 fusion protein (Figure 1A) was used as a bait to screen a VP16-tagged cDNA expression library of mouse mid-gestation embryos. One of the isolated cDNA clones encoded a polypeptide that interacted specifically with Pax5 but not with a Gal4–p53 protein or with the Gal4 DNA-binding domain alone (Figure 1B). cDNA sequence analysis revealed that the VP16 transactivation domain of the expression vector was fused in-frame to a 150 amino acid polypeptide that is most highly related to the N-terminal sequences of the *Xenopus* and rat Grg4 (Esp2) proteins, two vertebrate homologs of *Drosophila* Groucho (Schmidt and Sladek, 1993; Roose *et al*., 1998). As only a partial sequence of the mouse Grg4 protein has been characterized thus far (Koop *et al*., 1996), we cloned the full-length murine *Grg4* cDNA by RT–PCR. Like other members of the Groucho family, the mouse Grg4 protein consists of five characteristic domains (Figure 1A): a highly conserved glutamine-rich Q domain at the N-terminus; the GP domain enriched in glycine and proline residues; the conserved CcN domain containing a nuclear localization signal and putative phosphorylation sites for casein kinase II and cdc2 kinase; a serine/proline-

rich region referred to as the SP domain; and a C-terminal region containing seven highly conserved WD40 repeats (Stifani *et al*., 1992; see note added in proof). Interestingly, the VP16-Grg4 protein isolated in the yeast two-hybrid screen contains the entire Q region and part of the GP domain (Figure 1A), indicating that these N-terminal sequences of Grg4 can interact specifically with Pax5 in the yeast cell.

The transactivation domain of Pax5 interacts with the Q domain of Grg4

To verify the Pax5–Grg4 protein interaction in murine cells, we next performed one-hybrid assays in transiently transfected J558L plasmacytoma cells that do not express endogenous Pax5. The expression of limiting amounts of Pax5 protein resulted in a modest increase of the transcriptional activity of the luciferase gene *luc-CD19* (Figure 3A), which is under the control of three high-affinity Pax5-binding sites (Dörfler and Busslinger, 1996). Co-expression of the VP16-Grg4 polypeptide, which was isolated in the yeast two-hybrid screen, enhanced the luciferase activity ~5-fold, whereas the VP16 transactivation domain alone had no effect. Hence, this potent transactivation domain is recruited to the promoter only when linked to the N-terminal Grg4 sequences. Moreover, this recruitment depends on Pax5, as in its absence the VP16-Grg4 protein failed to stimulate the basal promoter activity (Figure 3A). We conclude, therefore, that the Pax5 protein and the N-terminal domain of Grg4 interact with each other in mammalian cells, thus confirming the results of the yeast two-hybrid assay. Importantly, a VP16-Gro protein containing the equivalent N-terminal sequences of *Drosophila* Groucho enhanced the activity of the reporter gene to the same level as VP16-Grg4 (Figure 3A), indicating that the Pax5 interaction domain has been conserved among Groucho proteins.

We next took advantage of the same one-hybrid assay to delineate the domain of Pax5 that is required for interaction with the N-terminal Q domain of Grg4. For this purpose, we analyzed a series of mutant Pax5 proteins that are shown schematically in Figure 2A. The C-terminal sequences of Pax5 are known to harbor a potent transactivation (TA) domain that is negatively regulated by adjacent inhibitory sequences (Dörfler and Busslinger, 1996). Deletion of this inhibitory domain in the mutant protein B4 did not interfere with the Pax5-dependent stimulatory effect of VP16-Grg4 (Figure 3B). In contrast, transcriptional stimulation was abolished by further deletion of the C-terminal transactivation domain in the Pax5 mutants B8 and B9, whereas internal deletion of the conserved octapeptide motif (ΔOP) or the partial homeodomain (ΔHD) of Pax5 did not have any effect (Figure 3B). These results indicate that the C-terminal transactivation domain of Pax5 interacts specifically with the N-terminal Q domain of Grg4.

The interaction of Pax5 with full-length Grg4 depends on the Pax5 octapeptide motif and the Grg4 SP domain

We next used *in vitro* protein-binding assays to study the interaction between Pax5 and Grg4. To this end, we expressed a protein consisting of the Q domain of Grg4 fused to glutathione *S*-transferase (GST) (Figure 4A) for

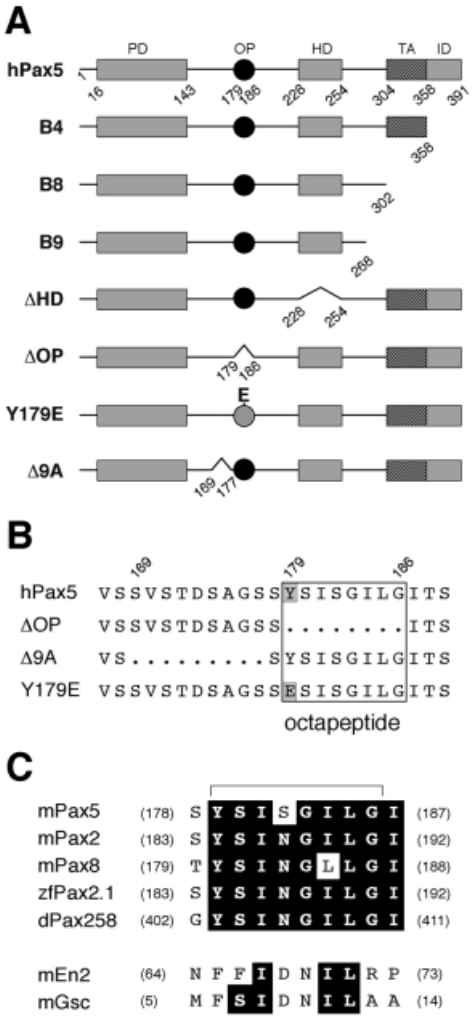


Fig. 2. Schematic diagram of mutant Pax5 proteins. (**A**) Domain structure of Pax5 and extent of amino acid deletion in the different mutant proteins. (**B**) The octapeptide sequence is shown together with the Y179E mutation and the ΔOP and Δ9A deletions. (**C**) Conservation of the octapeptide motif. The octapeptide sequences of mouse (m), zebrafish (zf) and *Drosophila* (d) members of the Pax2/5/8 family are aligned with the corresponding Engrailed homology region 1 (eh1) of the homeodomain transcription factors En2 and Goosecoid (Gsc). Identical amino acids are highlighted by black overlay. For sequences, see Smith and Jaynes (1996), Czerny *et al*. (1997) and Pfeffer *et al*. (1998).

subsequent use in GST pull-down assays. Surprisingly, this Grg4-Q polypeptide, like the control GST protein, failed to bind radiolabeled Pax5 protein (Figure 4B, lanes 2 and 7). In contrast, a GST fusion protein containing the entire Grg4 sequence interacted efficiently with Pax5 (Figure 4B, lane 3), indicating that full-length Grg4 and Pax5 can form a relatively stable complex *in vitro* and that Grg4 sequences other than the N-terminal Q domain are also involved in complex formation. To identify this additional domain(s), we analyzed a series of GST fusion proteins with progressively larger deletions of C-terminal Grg4 sequences (Figure 4A). Elimination of the WD40 repeats in the mutant Grg4-ΔWD40 protein did not significantly affect the interaction with Pax5 (Figure 4B, lane 4), indicating that these protein–protein interaction

motifs (Fisher and Caudy, 1998) are not essential for Pax5 binding. However, further deletion of the SP domain (Grg4-ΔSP) resulted in a dramatic reduction of Pax5 binding, thus identifying this central region of Grg4 as a second Pax5 interaction domain. Interestingly, a GST fusion protein containing only the SP domain was unable to interact with Pax5 (data not shown), suggesting that Pax5 binding depends on cooperative interaction with both the SP and Q domains of Grg4.

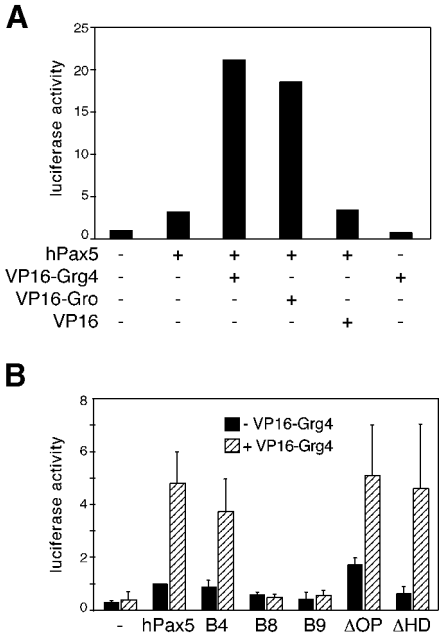


Fig. 3. The N-terminal region of Grg4 interacts with the transactivation domain of Pax5. (**A**) Interaction of Pax5 and VP16-Grg4. The expression plasmids indicated (100 ng) were co-transfected into J558L cells together with the *Renilla* luciferase gene pRL-SV40 (0.4 μg) and firefly luciferase gene *luc-CD19* (5 μg; Dörfler and Busslinger, 1996). After 48 h, the cells were lysed, luciferase activities were measured, and the activity of the firefly luciferase was standardized relative to the control *Renilla* luciferase to normalize for differences in transfection efficiencies. Luciferase values are shown relative to the activity measured with the empty expression vector pKW2T (left bar). (**B**) Interaction of the Q domain of Grg4 with the transactivation domain of Pax5. Both luciferase genes and the Pax5 expression plasmids indicated (100 ng each) were electroporated into J558L cells with (hatched bars) or without (black bars) the VP16-Grg4 expression vector (100 ng). Average values of three experiments are shown.

The GST pull-down experiments also pointed to the existence of an additional Grg4-binding region in Pax5. To identify this domain, we analyzed mutant Pax5 proteins (Figure 2A) for their ability to bind full-length Grg4 in the GST pull-down assay (Figure 4C). The deletion mutant B8 bound Grg4 with an affinity similar to the full-length Pax5 protein, indicating that this assay failed to detect an interaction between the C-terminal transactivation domain of Pax5 and the N-terminal Q domain of Grg4, in agreement with the data shown in Figure 4B (lane 7). Grg4 binding was, however, dramatically reduced by deletion of the octapeptide motif (ΔOP) in Pax5, whereas elimination of nine amino acids (Δ9A) adjacent to the octapeptide sequence did not influence protein binding (Figure 4C). The octapeptide motif of Pax2/5/8 proteins is closely related to a short sequence present in the repression domain of the transcription factors Engrailed (En) and Goosecoid (Gsc; Figure 2C). This short domain is known as En homology region 1 (eh1) or Gsc-En homology (GEH) element (Smith and Jaynes, 1996) and was previously shown to mediate interaction with the *Drosophila* Groucho protein (Jiménez *et al*., 1997, 1999). Groucho binding was, however, abolished by mutating a conserved phenylalanine of the eh1/GEH sequence to glutamic acid (Jiménez *et al*., 1999). The octapeptide sequence of all known Pax2/5/8 proteins contains a tyrosine residue at the corresponding position (Figure 2C), which could, however, be substituted by phenylalanine without affecting Grg4 binding (data not shown). In contrast, the glutamic acid substitution Y179E completely abrogated the interaction of Pax5 with Grg4 (Figure 4C), further emphasizing the importance of the octapeptide motif as an interaction domain for Grg4.

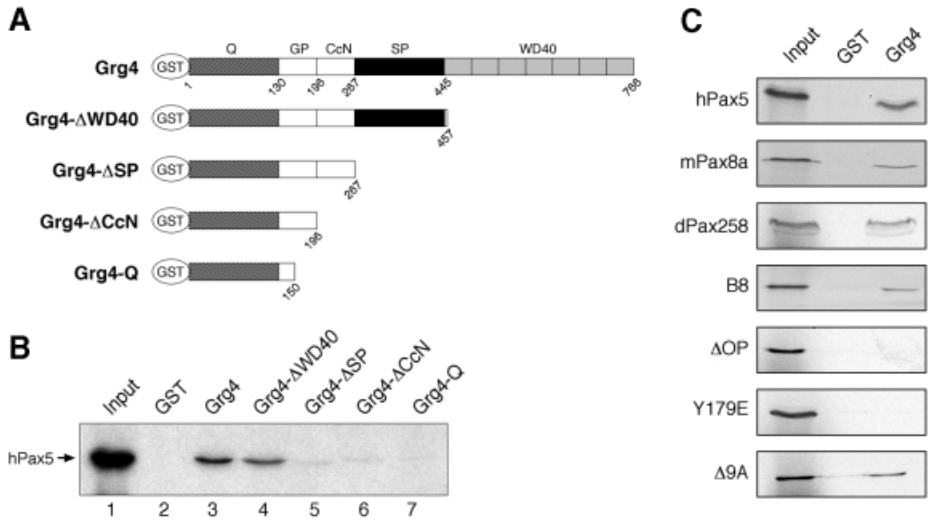


Fig. 4. *In vitro* binding of Pax5 to Grg4. (**A**) C-terminal deletions of GST–Grg4 proteins. (**B**) The SP domain of Grg4 interacts with Pax5. GST pull-down assays were used to study the interaction between *in vitro* translated, ^{35}S-labeled Pax5 protein and GST (lane 2) or GST–Grg4 proteins (lanes 3–7) bound to glutathione–Sepharose. Lane 1 contained 10% of the Pax5 protein input. (**C**) The octapeptide of Pax5 mediates binding of Grg4. The ^{35}S-labeled Pax proteins indicated were analyzed for binding to GST or GST–Grg4. The input lane contained ~10% of the total Pax protein used in each assay.

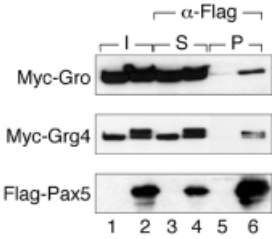


Fig. 5. *In vivo* interaction of Groucho and Pax5. Myc-tagged Grg4 or Groucho proteins were expressed in transiently transfected COP-8 fibroblasts either alone (lanes 1, 3 and 5) or with a Flag-tagged Pax5 protein (lanes 2, 4 and 6). After 48 h, whole-cell lysates were prepared and immunoprecipitated with the anti-Flag M2 antibody. Immunocomplexes were analyzed by Western blotting first with a polyclonal anti-Myc antibody and then with a polyclonal antiserum directed against the Pax5 paired domain (Adams *et al*., 1992). Lanes 1–4 contained ~7.5% of the total input (I) protein or supernatant (S), respectively. P, precipitated protein.

In vivo interaction of Grg4 and Pax5

To investigate complex formation *in vivo*, we expressed Myc epitope-tagged Grg4 or Groucho either alone or together with Flag-tagged Pax5 in transiently transfected COP-8 fibroblasts. Pax5 was precipitated subsequently from the cell lysate with a monoclonal anti-Flag antibody, and the immunoprecipitate was analyzed for the presence of Groucho and Pax5 proteins by Western blotting. A low but significant amount of Myc-tagged Grg4 or Groucho was only detected in the immunoprecipitate of COP-8 cells co-expressing Pax5 (Figure 5, compare lanes 5 and 6). These data therefore indicate that the mouse Grg4 and *Drosophila* Groucho proteins can interact with Pax5 *in vivo*. However, we failed to co-immunoprecipitate endogenous Pax5 and Grg proteins from nuclear extract of B cells, which may reflect a low abundance or stability of the Pax5–Grg complex (see Discussion).

Induced phosphorylation of Grg4 upon interaction with Pax5

The *Drosophila* Groucho and mammalian Grg proteins are expressed predominantly as a constitutively phosphorylated polypeptide that can be detected as a single band of ~90 kDa by Western blot analysis (Husain *et al*., 1996). Additional phosphorylation of Grg proteins was observed upon neural differentiation of P19 embryonal carcinoma cells, which resulted in the appearance of a doublet of 90–93 kDa (Husain *et al*., 1996). Similar Grg isoforms were also generated in COP-8 cells co-expressing Grg4 and Pax5 (Figure 5, lanes 2 and 4). However, a single Grg4 species was detected in the absence of Pax5 (Figure 5, lanes 1 and 3), suggesting that the interaction with Pax5 leads to further modification of the Grg4 protein. Moreover, both Grg4 isoforms could be co-immunoprecipitated together with Pax5 from COP-8 cell extracts (Figure 5, lane 6). Incubation of the immunoprecipitate with λ protein phosphatase resulted in a single Grg4 species that migrated slightly faster on SDS–PAGE than both isoforms prior to phosphatase treatment (Figure 6A). Hence, we conclude that the two Grg4 isoforms expressed in COP-8 cells differ in their phosphorylation state.

The Pax5-dependent phosphorylation of Grg4 suggested that binding of Grg4 to Pax5 is a prerequisite for this modification to occur. To test this hypothesis, we co-

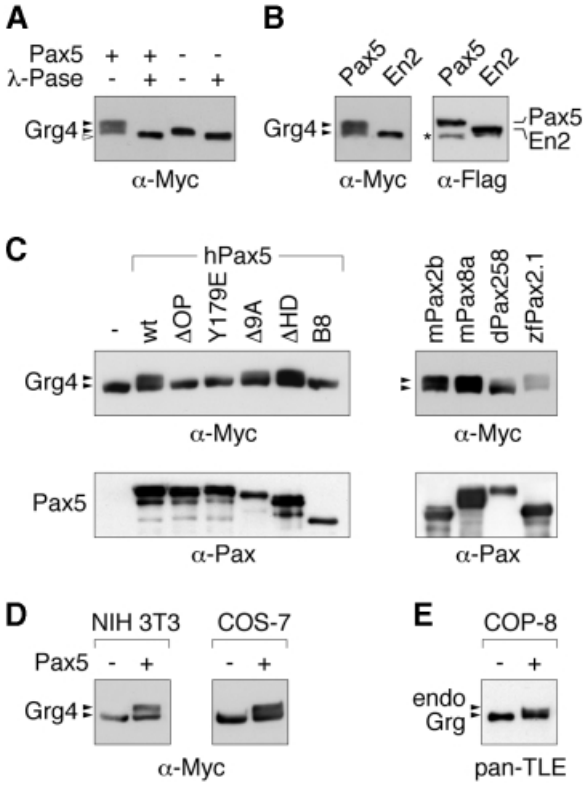


Fig. 6. Induced phosphorylation of Grg4 upon interaction with Pax5. (**A**) Pax5-dependent phosphorylation of Grg4. Pax5 (where indicated) and Myc-tagged Grg4 were transiently expressed in COP-8 cells for 48 h. Grg4 was then precipitated from whole-cell lysates with a monoclonal anti-Myc antibody, incubated with λ protein phosphatase (λ-Pase) and then analyzed by Western blotting with a polyclonal anti-Myc antibody. (**B**) Specificity of the Grg4 phosphorylation. Myc-tagged Grg4 and Flag-tagged Pax5 or En2 proteins were co-expressed in COP-8 cells followed by Western blot analysis with anti-Myc and anti-Flag antibodies, respectively. The asterisk denotes a cross-reacting protein. (**C**) The phosphorylation of Grg4 depends on direct interaction with Pax2/5/8 proteins. The Pax proteins indicated were co-expressed with Myc-tagged Grg4 in COP-8 cells followed by Western blotting with polyclonal anti-Myc and anti-paired domain antibodies. (**D**) Phosphorylation of Myc-tagged Grg4 in NIH 3T3 and COS-7 cells co-expressing Pax5. (**E**) Endogenous (endo) Grg proteins are phosphorylated in Pax5-expressing COP-8 cells, as shown by Western blot analysis with a pan-TLE (Grg) antibody (Stifani *et al*., 1992).

expressed mutant Pax5 proteins with Myc-tagged Grg4 in transiently transfected COP-8 cells, followed by Western blot analysis. Indeed, deletion of the Pax5 transactivation domain (B8) as well as mutation of the octapeptide motif (ΔOP, Y179E) prevented further phosphorylation of the Grg4 protein (Fgure 6C). In contrast, mutation of other Pax5 domains (Δ9A or ΔHD) had no effect on Grg4 phosphorylation (Figure 6C), indicating that the integrity of the two Grg4 interaction domains in Pax5 is essential for this additional modification to occur. Moreover, the Pax5-induced phosphorylation of Grg4 was also observed in transfected NIH 3T3 fibroblasts as well as in COS-7 kidney cells, and thus represents a cell type-independent phenomenon (Figure 6D). Importantly, even endogenous Grg proteins underwent phosphorylation in COP-8 cells overexpressing Pax5 (Figure 6E). En2, another Groucho-interacting transcription factor (Jiménez *et al*., 1997), was, however, unable to induce additional phosphorylation of

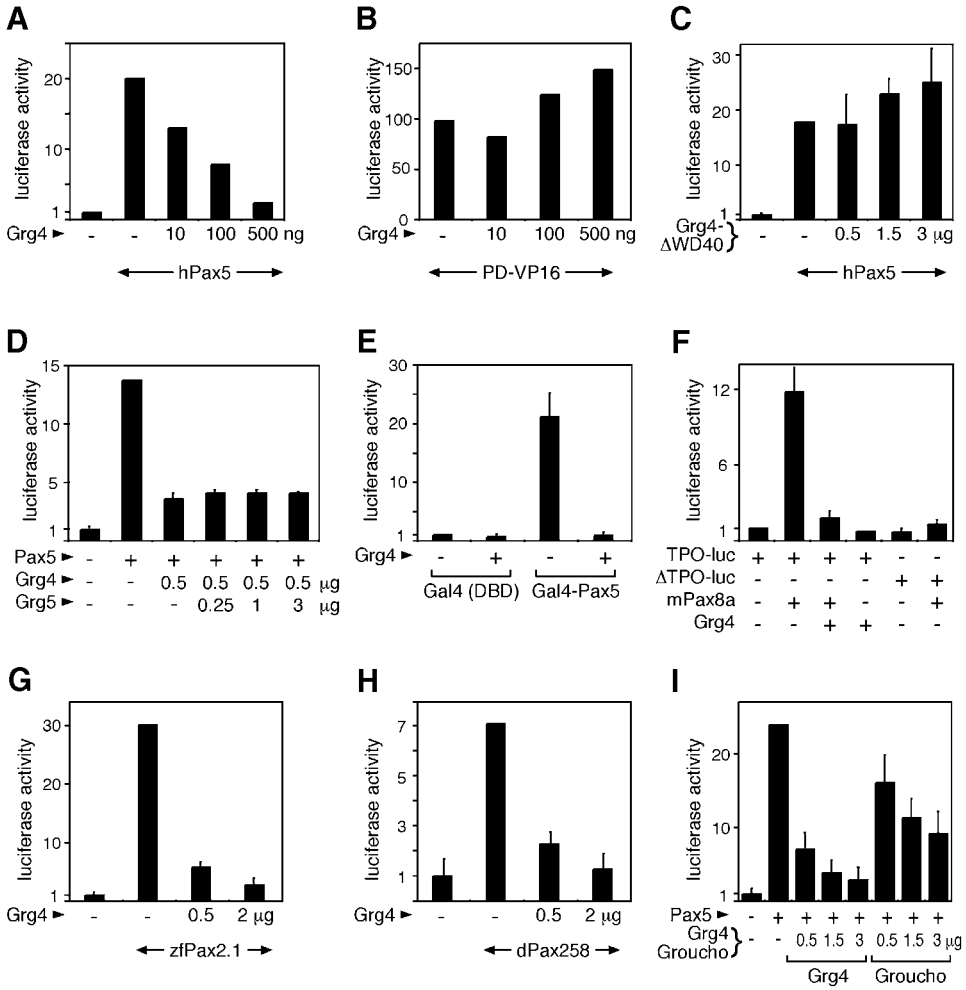


Fig. 7. Grg4 represses the transcriptional activity of Pax5. The Pax expression vector (0.5 μg), luciferase genes *luc-CD19* (5 μg) and pRL-SV40 (0.4 μg), and increasing amounts of the indicated Grg expression vector were used for transient transfection of SP2/0 cells (A–D and G–I). The amount of expression plasmid was equalized by the addition of pKW2T, and all data were evaluated as described in the legend to Figure 3A. Normalized luciferase values of one representative experiment (A and B) or three independent transfections (C–I) are shown relative to the luciferase activity measured in the absence of the Pax protein. (**A**) Repression of the Pax5 transactivation function by Grg4. (**B**) Failure of Grg4 to repress the function of a paired domain (PD)-VP16 protein. (**C**) Inability of the Grg4-ΔWD40 protein to repress the transcriptional activity of Pax5. (**D**) Grg5 fails to antagonize Grg4-mediated repression of Pax5 activity. (**E**) Repression of a chromatinized reporter gene by Grg4. Expression vectors (1 μg) encoding Gal4–Pax5 or the Gal4 DNA-binding domain (DBD) were transfected together with a Grg4 expression plasmid (0.5 μg) into U2-OS cells containing an integrated luciferase gene, followed by luciferase analysis as described (Alkema *et al*., 1997). (**F**) Grg4-mediated repression of the rat *TPO* promoter. HeLa cells were transfected with mPax8a (0.1 μg) and Grg4 (0.5 μg) expression plasmids and the luciferase gene *TPO-luc* (2.5 μg) and pRL-SV40 (10 ng). Δ*TPO-luc* contains a mutated Pax8-binding site (Zannini *et al*., 1992). (**G** and **H**) Grg4-mediated repression of the transcriptional activity of zebrafish zfPax2.1 and *Drosophila* dPax258. (**I**) Repression of the Pax5 transactivation function by *Drosophila* Groucho.

Grg4 (Figure 6B). Hence, this phosphorylation reaction cannot be promoted by all Groucho-binding proteins.

Grg4 represses the transcriptional activity of Pax5

As Groucho proteins are known to function as transcriptional corepressors (reviewed by Fisher and Caudy, 1998; Parkhurst, 1998), we investigated whether full-length Grg4 is able to modulate the transcriptional activity of Pax5. We thus studied the effect of Grg4 on the Pax5-dependent transcription of the reporter gene *CD19-luc* in transiently transfected plasmacytoma cells. As shown in Figure 7A, the Grg4 protein repressed Pax5-mediated activation of this reporter gene in a concentration-dependent manner. This repression was specific, as the transcriptional activity of a chimeric protein consisting of the VP16 transactivation domain fused to the Pax5 paired domain was not affected by Grg4 (Figure 7B). Interestingly, a Grg4 protein lacking all WD40 repeats failed to repress the activity of Pax5 (Figure 7C) despite the fact that this mutant Grg4 protein binds Pax5 with an efficiency similar to full-length Grg4 (Figure 4B). A fifth member of the *Grg* family, *Grg5*, codes for a short 197-amino-acid protein consisting only of the Q and SP domains (Mallo *et al*., 1993). This small protein is able to antagonize Grg4-dependent repression by the transcription factors TCF and Blimp-1 (Roose *et al*., 1998; Ren *et al*., 1999). However, the Grg5 protein could not reverse Grg4-mediated repression of the Pax5 transactivation function (Figure 7D). In conclusion, Grg4 can repress the transcriptional activity of Pax5 efficiently by a mechanism that depends on the integrity of the WD40 repeats and is insensitive to the dominant-negative action of Grg5.

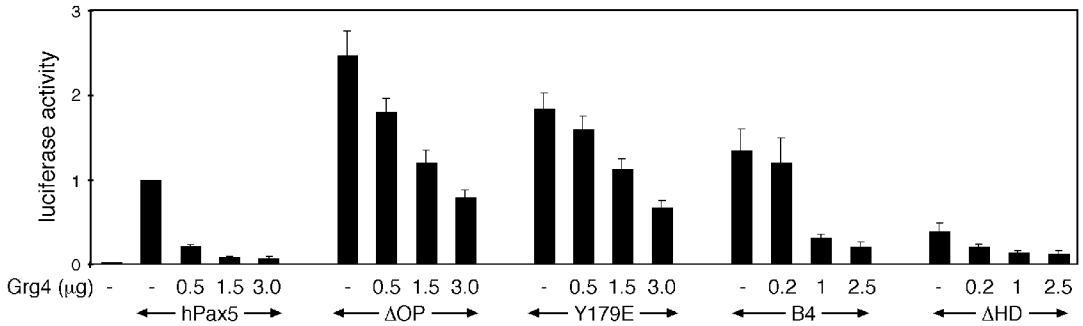


Fig. 8. The octapeptide motif of Pax5 mediates repression by Grg4. Transient transfection experiments with the expression plasmids indicated were performed in SP2/0 cells and subsequently evaluated as described in the legend to Figure 7.

We next examined whether Grg4 can also repress Pax5-mediated activation of a reporter gene that is stably inserted in the genome. For this experiment, we used a human osteosarcoma cell line that contains an integrated luciferase gene under the control of a thymidine kinase promoter and five upstream Gal4-binding sites (Alkema *et al*., 1997). Luciferase expression was stimulated in these cells ~20-fold by a Gal4–Pax5 fusion protein relative to the Gal4 DNA-binding domain alone (Figure 7E). Co-expression of Grg4 efficiently inhibited transactivation by Gal4–Pax5 (Figure 7E), indicating that packaging of a reporter gene into chromatin still results in Grg4-mediated repression of Pax5 activity.

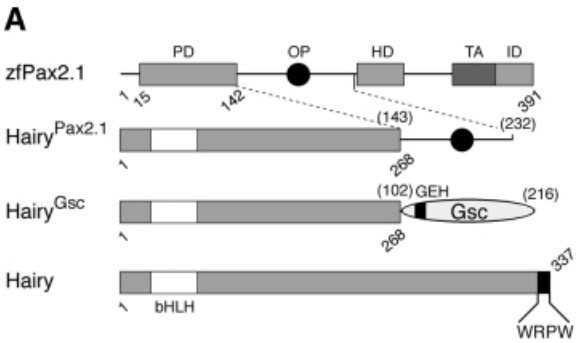
Although several B cell-specific target genes of Pax5 are known, they all contain TATA-less promoters that are almost inactive in transient transfection assays (reviewed by Busslinger and Nutt, 1998). To study Grg4-mediated repression of a naturally occurring Pax target gene, we analyzed the promoter of the rat thyroperoxidase (*TPO*) gene, which contains a Pax8-binding site immediately upstream of a TATA-box (Zannini *et al*., 1992). A luciferase gene under the control of the *TPO* promoter was strongly activated by Pax8 in transiently transfected HeLa cells, whereas a promoter mutation inactivating the Pax8-binding site (Δ) prevented transcriptional stimulation (Figure 7F) (Zannini *et al*., 1992). Co-expression of Grg4 efficiently repressed the Pax8-dependent activation of the *TPO* promoter, whereas Grg4 expression on its own did not affect basal promoter activity (Figure 7F). In summary, these data demonstrate that Pax5 and Pax8 can efficiently recruit Grg4 to artificial or naturally occurring promoters, which results in repression of gene transcription.

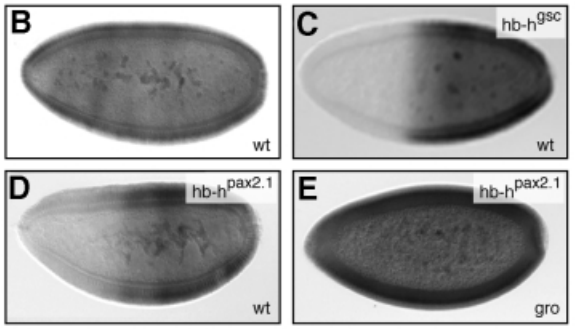
The octapeptide of Pax5 mediates repression by Grg4

To identify the domains of Pax5 that confer Grg4-mediated repression, we analyzed the transcriptional activity of several mutant Pax5 proteins in transiently transfected plasmacytoma cells. In the absence of exogenous Grg4 protein, the mutant proteins already differed in their transactivation potential (Figure 8). A Pax5 protein lacking the partial homeodomain (ΔHD) was 2- to 3-fold less active than the wild-type protein, whereas the transcriptional activity of the octapeptide deletion mutant (ΔOP) was increased consistently by a factor of 2–3 (Figure 8). Western blot analysis indicated, however, that these differences in transcriptional activity were not caused by different expression levels of the mutant Pax5 proteins (data not shown). The lower activity of the Pax5-ΔHD protein could be explained by our recent finding that the partial homeodomain is an interaction motif for the TATA-binding protein and may thus facilitate recruitment of the basal transcription factor complex TFIID to the promoter (Eberhard and Busslinger, 1999). Moreover, RNase protection analyses revealed that the known mouse *Grg(1/3a/4)* genes are constitutively expressed at all stages of B-cell development as well as in the plasmacytoma cells analyzed (data not shown). Hence, endogenous Grg proteins may interact with Pax5 and reduce its activity in transfected cells, whereas the increased transactivation potential of the Pax5-ΔOP protein could reflect the loss of Grg binding.

Consistent with this hypothesis, expression of exogenous Grg4 protein could repress the transcriptional activity of the Pax5-ΔOP protein only 2-fold, in contrast to the wild-type Pax5 protein, whose transactivation function was completely inhibited (Figure 8). Even at the highest Grg4 concentration, the Pax5-ΔOP protein was transcriptionally as active as the wild-type Pax5 protein in the absence of any exogenous Grg4 protein. Moreover, the single amino acid substitution Y179E prevented Grg4-mediated repression of Pax5 activity to the same extent as deletion of the entire octapeptide sequence (Figure 8), further demonstrating that the octapeptide motif of Pax5 is essential for *in vivo* binding and thus recruitment of the corepressor Grg4. The residual 2-fold repression of the octapeptide mutants by Grg4 suggests that the two proteins can still bind to each other weakly *in vivo* through the second, still intact interaction involving the transactivation domain of Pax5 and the Q domain of Grg4 (Figure 3B).

The C-terminal sequences of Pax5 contain an inhibitory domain that negatively regulates the adjacent transactivation region of Pax5 in most cell lines including the J558L cells, but not in SP2/0 cells (Dörfler and Busslinger, 1996). A mutant Pax5 protein (B4) lacking this inhibitory domain was still repressed efficiently by Grg4 in both SP2/0 and J558L cells (Figure 8; data not shown). Furthermore, Grg4 was also able to repress the low transcriptional activity of the homeodomain deletion mutant Pax5-ΔHD (Figure 8A). These data therefore indicate that neither the C-terminal inhibitory sequence nor the partial homeodomain of Pax5





F

construct	females	males
hb-h^{gsc}	0	59
hb-h$^{pax2.1}$	61	54

Fig. 9. The octapeptide motif functions as a repression domain in *Drosophila* embryos. (**A**) Schematic diagram of the Hairy derivatives used in the *Sxl* repression assay. The Hairy$^{Pax2.1}$ protein was generated by fusing the Hairy sequences at amino acid 268 to a 90-amino-acid sequence (143–232) of the zebrafish Pax2.1 protein encompassing the conserved octapeptide motif (Pfeffer *et al*., 1998). The Hairy^{Gsc} protein containing the eh1/GEH motif of Goosecoid (Gsc) was described previously (Jiménez *et al*., 1999). (**B**–**E**) *Sxl* repression by chimeric Hairy proteins in female blastoderm embryos of *Drosophila*. The Hairy^{Gsc} (C) and Hairy$^{Pax2.1}$ (D and E) proteins were expressed under the control of the *hunchback* (*hb*) promoter in wild-type (wt, C and D) or *gro*$^{E48mat-}$ mutant embryos (E). Sxl expression was analyzed in transgenic (C–E) and control (B) female embryos by whole-mount staining with a monoclonal antibody detecting the female-specific form of Sxl (Bopp *et al*., 1991). (**F**) Effects of Hairy^{Gsc} and Hairy$^{Pax2.1}$ expression on female viability. Wild-type females were crossed with *hb* transgenic males, and the number of viable progeny is indicated.

are essential for Grg4-mediated repression of the Pax5 transactivation function.

Evolutionary conservation of the interaction between Groucho and Pax proteins

Three *groucho*-related genes coding for full-length Grg proteins (Grg1, 3a and 4) have been identified to date in the mouse genome (Koop *et al*., 1996; Leon and Lobe, 1997). Using transient transfection assays, we have shown that all three murine Grg proteins are phosphorylated in a Pax5-dependent manner and can repress the transcriptional activity of Pax5 efficiently (data not shown). Even the distantly related Groucho protein of *Drosophila* was able to interact with Pax5 (Figures 3A and 5) and to down-modulate the activity of this transcription factor in heterologous mammalian cells (Figure 7I).

Furthermore, GST pull-down assays demonstrated that the mouse Pax8 and *Drosophila* Pax258 proteins can bind full-length Grg4 with an affinity similar to that of human Pax5 (Figure 4C). Moreover, the transcriptional activity of the mouse Pax8, zebrafish Pax2.1 and *Drosophila* Pax258 proteins could be repressed efficiently by Grg4 in transfected plasmacytoma cells (Figure 7F, G and H). These different Pax proteins were also able to promote additional phosphorylation of Grg4 in transfected COP-8 fibroblasts (Figure 6C). Collectively, these data demonstrate, therefore, that the interaction between distantly related members of the Pax2/5/8 and Groucho protein families has been conserved in evolution.

The octapeptide motif functions as a Groucho-dependent repression domain in Drosophila embryos

Inspired by the high evolutionary conservation of the Groucho–Pax2/5/8 protein interaction, we next investigated whether the octapeptide motif can function *in vivo* as a repression domain during *Drosophila* development. For this, we took advantage of a repression assay that is based on the transcriptional regulation of the *Sex lethal* (*Sxl*) gene in *Drosophila* embryos (Parkhurst *et al*., 1990). Sxl is a key regulator of sex determination and dosage compensation, whose transcription is initiated only in female blastoderm embryos. In male embryos, *Sxl* expression is prevented by the transcriptional repressor Deadpan (Dpn), which is a member of the Hairy-related basic helix–loop–helix (bHLH) protein family. The negative effect of Dpn can be mimicked in female embryos by ectopic expression of the related Hairy protein at the time of sex determination (Parkhurst *et al*., 1990). Premature Hairy expression under the control of the *hunchback* (*hb*) promoter represses *Sxl* transcription in the anterior part of female embryos, which leads to female-specific lethality (Parkhurst *et al*., 1990). Repression of *Sxl* by Hairy depends on the interaction of its C-terminal WRPW motif with Groucho (Paroush *et al*., 1994) and, consequently, does not occur in embryos deprived of maternal Groucho function (Jiménez *et al*., 1997). Moreover, substitution of the C-terminal Hairy sequences by a heterologous repression domain still leads to down-regulation of *Sxl* expression, thus providing a convenient assay for studying Groucho-dependent repression domains *in vivo* (Jiménez *et al*., 1999).

We used this assay to examine the *in vivo* function of the octapeptide motif by replacing the C-terminal region of Hairy with a sequence encompassing the 90 amino acids located between the paired domain and partial homeodomain of zfPax2.1 (Figure 9A). The octapeptide motif is the only conserved element that is shared between this zebrafish Pax2.1 sequence and the corresponding region of the *Drosophila* Pax258 protein (Czerny *et al*., 1997). Expression of the chimeric Hairy$^{Pax2.1}$ protein under the control of the *hb* promoter resulted in significant reduction of Sxl expression in the anterior half of transgenic female embryos (Figure 9D) compared with the uniform Sxl staining of wild-type embryos (Figure 9B). Moreover, the repression of *Sxl* by Hairy$^{Pax2.1}$ was dependent on Groucho, as it was not observed in embryos lacking

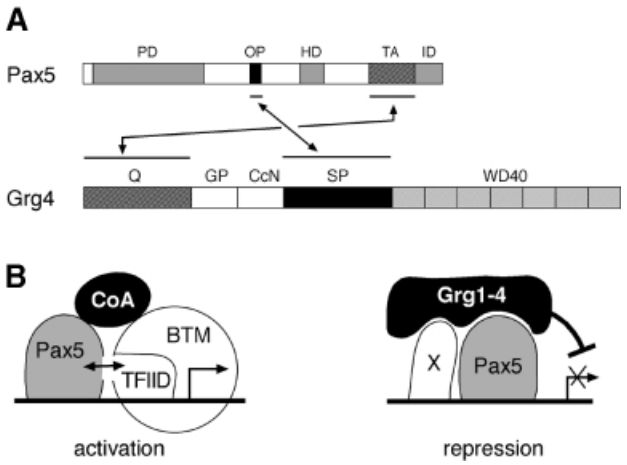


Fig. 10. Model of Grg4 recruitment by Pax5 (BSAP). (**A**) Summary of the identified interactions between Pax5 and Grg4. (**B**) Model of Grg-mediated gene repression by Pax5. As different members of the Grg family are co-expressed with Pax5 throughout midbrain and B-cell development (Koop *et al*., 1996; D.Eberhard, unpublished data), we hypothesize that Pax5 can stably recruit Grg proteins to a specific promoter only in collaboration with a second Grg-binding transcription factor (X). CoA, coactivator protein(s); BTM, basal transcription machinery.

maternal *gro* function (Figure 9E; Materials and methods). However, the Hairy$^{Pax2.1}$ protein was clearly less active in repressing the *Sxl* gene than a Hairy^{Gsc} protein (Figure 9C) containing the GEH motif of Goosecoid (Gsc) as a potent repression domain (Jiménez *et al*., 1999). This difference in repression activity is also reflected by the fact that ectopic expression of Hairy^{Gsc} caused female lethality, whereas Hairy$^{Pax2.1}$ did not significantly affect female viability (Figure 9F). These data indicate that the octapeptide motif of the zebrafish Pax2.1 protein can function as a weak Groucho-dependent repression domain in *Drosophila* embryos.

Discussion

Commitment to the B-lymphoid lineage critically depends on the repression of lineage-inappropriate genes by the transcription factor Pax5 (BSAP) (Nutt *et al*., 1999). By identifying Grg4 as a corepressor of Pax5, we have now elucidated a molecular mechanism by which Pax5 can act as an active repressor of gene transcription. Grg4 was shown to interact with Pax5 *in vitro* in GST pull-down assays as well as *in vivo* in yeast and various mammalian cell types. Grg4 and Pax5 both contain two separate interaction domains that cooperate together in protein binding (Figure 10A). As a consequence, Grg4 could repress the transcriptional activity of Pax5 efficiently in transiently transfected cells. Similar protein interactions resulting in repression were also observed between distantly related members of both the Pax2/5/8 and Groucho protein families. Moreover, Grg4 also interacts with Pax1, Pax3 and Pax6 (D.Eberhard, unpublished data), which are representative members of the other three Pax subfamilies. Hence, Groucho proteins appear to act as corepressors of all Pax transcription factors. It has recently been suggested that Pax3 mediates gene repression by recruiting the corepressors HIRA or Daxx, respectively (Magnaghi *et al*., 1998; Hollenbach *et al*., 1999). In addition, we have shown that the Rb protein can interact with the partial homeodomain of Pax5 (Eberhard and Busslinger, 1999). However, none of these proteins (HIRA, Daxx or Rb) was able to repress the transcriptional activity of Pax5 (data not shown), suggesting that the Pax transcription factors exert their repression function primarily by recruiting corepressors of the Groucho family.

Two distinct interactions contribute to the formation of the Grg4–Pax5 complex

Several transcription factors are known to interact with Groucho proteins and, wherever studied, a single domain in each protein was responsible for this interaction (reviewed by Fisher and Caudy, 1998; Parkhurst, 1998). Pax5 is thus the first example of a transcription factor that relies on two separate interactions for Groucho binding (Figure 10A). One of these interactions is mediated by the N-terminal Q domain of Grg4 and the C-terminal transactivation domain of Pax5. Detailed *in vitro* mutagenesis indicated that the integrity of the entire Q domain is essential for Pax5 binding (data not shown). The highly conserved Q domain contains two leucine zipper-like motifs that are both required for tetramerization of Groucho proteins (Chen *et al*., 1998). In addition, the Q domain is sufficient to mediate binding of the Blimp-1 (PRDI-BF1) (Ren *et al*., 1999) and TCF proteins (Roose *et al*., 1998). Consequently, the short Grg5 protein consisting only of the Q and SP domains (Mallo *et al*., 1993) can act as a dominant-negative protein to reverse Grg4-mediated repression by these two transcription factors (Roose *et al*., 1998; Ren *et al*., 1999). In contrast, the interaction between the Q domain of Grg4 and the transactivation domain of Pax5 is too weak to be detected in *in vitro* binding assays. In addition, the Grg5 protein is unable to antagonize Grg4-mediated repression of the Pax5 transcriptional activity. Both findings are consistent with the notion that two separate but cooperative interactions are required for efficient formation of the Grg4–Pax5 complex.

The second interaction that contributes to complex stability involves the SP domain of Grg4 and the octapeptide motif of Pax5 (Figure 10). The SP domain, which was previously implicated in the interaction between Groucho and Hairy-related bHLH proteins (Paroush *et al*., 1994; Jiménez *et al*., 1997), was again insufficient to mediate Pax5 binding, and even three copies of the octapeptide motif linked to a Gal4 protein failed to interact with Grg4 in *in vitro* binding assays (data not shown). The octapeptide sequence was identified 10 years ago as a conserved motif of most Pax proteins (Burri *et al*., 1989). Here, we show that this short amino acid sequence constitutes a protein interaction domain, is required in mammalian cells for Grg4-mediated repression of the Pax5 transactivation function and acts as a weak Groucho-dependent repression domain in *Drosophila* embryos. The octapeptide motif is related in sequence to the eh1/GEH region that is found in several homeodomain transcription factors (Noll, 1993; Smith and Jaynes, 1996). Consistent with this sequence similarity, the Y179E substitution in the octapeptide motif abrogated Groucho protein binding (Figure 4) in analogy to the equivalent F-to-E substitution in the eh1/GEH region (Jiménez *et al*., 1999). Apart from

these similarities, the octapeptide differs in several aspects from the eh1/GEH region. First, the eh1/GEH sequence is both necessary and sufficient for Groucho binding and transcriptional repression (Jiménez *et al*., 1999), in contrast to the octapeptide motif (this study). Secondly, the eh1/GEH region functions as a more potent Groucho-dependent repression domain in *Drosophila* embryos than the octapeptide sequence (Jiménez *et al*., 1997, 1999). Thirdly, the WD40 repeats of Groucho are essential for binding of the eh1/GEH region (Jiménez *et al*., 1997), whereas the SP domain of Grg4 seems to mediate interaction with the Pax5 octapeptide motif. However, the WD40 repeats of Grg4 are still required for repression of the Pax5 activity, suggesting that these conserved protein interaction motifs recruit additional factors into the Grg4–Pax5 complex.

Pax5-dependent phosphorylation of Grg proteins

Groucho proteins are known to be constitutively phosphorylated on serine and threonine residues (Husain *et al*., 1996). Hence, we were surprised to see that the Grg proteins undergo further phosphorylation upon Pax5 binding. This phosphorylation reaction is strictly dependent on the interaction with Pax5 and appears to be catalytic, as the majority of the Grg proteins become phosphorylated in cells overexpressing Pax5 (Figure 6) despite the fact that only a small proportion of these proteins could be co-precipitated with Pax5, possibly due to the relatively low stability of the Grg4–Pax5 complex (Figure 5). In contrast, the En2 protein, which interacts efficiently with Grg4 *in vivo* (data not shown), was unable to induce this additional phosphorylation, thus revealing a remarkable selectivity with regard to the Groucho-binding transcription factors involved in this phenomenon. Two different hypotheses could account for these observations. Pax5, in contrast to En2, binds Grg4 via two different interactions that may induce a conformational change in Grg4 and thus render cryptic phosphorylation sites accessible to a constitutively active kinase. Alternatively, Pax5 may specifically recruit a kinase into the Groucho complex in analogy to the homeodomain protein NK-3, which interacts simultaneously with Groucho and the nuclear kinase HIPK2 (Choi *et al*., 1999). Interestingly, signaling of the Torso receptor via the MAP kinase pathway has been implicated in antagonizing Groucho-mediated gene repression at the terminal pole regions of the *Drosophila* embryo (Paroush *et al*., 1997). Torso signaling does not, however, interfere with all Groucho functions, as it specifically affects only certain Groucho-containing complexes (Paroush *et al*., 1997), in analogy to the transcription factor selectivity observed for the induced Grg phosphorylation described in this study.

Grg-dependent transcriptional repression by Pax5

The different Groucho-binding transcription factors can be grouped into three classes according to their mode of Grg protein recruitment. The first class consists of active repressors comprising the Hairy-related bHLH proteins (Paroush *et al*., 1994), the homeodomain proteins En, Gsc and NK-3 (Jiménez *et al*., 1997, 1999; Choi *et al*., 1999) and the zinc finger proteins Hkb and Blimp-1 (Goldstein *et al*., 1999; Ren *et al*., 1999). All of these negative regulators utilize short sequence motifs to recruit Groucho proteins stably and thus act as constitutive repressors of transcription. In contrast, the HMG-box proteins of the TCF family bind Grg proteins in a manner that is regulated by Wnt signaling. These TCF proteins interact with Groucho proteins and thus function as transcriptional repressors in the absence of a Wnt signal. Upon signal transduction, the TCF proteins are converted into transcriptional activators by displacement of the Grg protein with the coactivator β-catenin (Roose *et al*., 1998). The third class consists of intrinsic transcriptional activators, as exemplified by the Dorsal and Runt proteins (Aronson *et al*., 1997; Dubnicoff *et al*., 1997). These transcription factors on their own are unable to recruit Groucho proteins to their target genes *in vivo*. Instead, they require the assistance of other Groucho-interacting transcription factors to recruit Grg proteins, and thus function as repressors only in a regulatory context-dependent manner (Valentine *et al*., 1998).

Pax5 is known simultaneously to activate B cell-specific genes (such as *CD19*) and to repress lineage-inappropriate genes (such as *M-CSF-R*) in the same early B-lymphoid progenitor cell (Nutt *et al*., 1997, 1998, 1999). Hence, Pax5 appears to belong to the class of transcription factors that are converted from activators to repressors by co-recruitment of Groucho proteins in a context-dependent manner (Figure 10B). The transcriptional activity of Pax5 was, however, repressed by Grg proteins in transient transfection assays, even though the artificial promoter of the reporter gene contained only Pax-binding sites. This apparent contradiction could be explained by the fact that the high expression levels attained in transiently transfected cells force the Grg4 protein to interact with Pax5. In addition, the Pax5 proteins bound to the multimerized sites in the reporter gene may cooperate in Grg4 binding, thus mimicking a co-recruitment paradigm. Pax5 is known to negatively regulate the activity of the 3′α enhancer present in the immunoglobulin heavy chain locus (reviewed by Busslinger and Nutt, 1998). Interestingly, the Pax5-dependent down-regulation of this enhancer not only depends on the presence of a Pax5 recognition sequence, but also requires the integrity of adjacent transcription factor-binding sites (Singh and Birshtein, 1996). By reconstituting the 3′α enhancer activity in heterologous cells, we have recently been able to demonstrate that Pax5 requires cooperation with other transcription factors to repress this enhancer in a Grg4-dependent manner (Y.Linderson, D.Eberhard, S.Pettersson and M.Busslinger, unpublished data). Hence, context-dependent recruitment of Groucho proteins may be a general mechanism that converts Pax proteins from activators to repressors of gene transcription.

Materials and methods

DNA constructs

All cDNAs were cloned into the eukaryotic expression vector pKW2T (Dörfler and Busslinger, 1996). The mouse *Grg4* cDNA was amplified by RT–PCR from RNA of 9.5/10.5-day-old embryos with the primers 5′-CCCAAGCTTACC**ATG**GTTCCGCAGACGCGC-3′ and 5′-GCTCTAGATGCTATGAGGAGGAGTCCAG-3′. Likewise, the mouse *Grg5* cDNA was amplified with the primers 5′-CCCAAGCTTACC**ATG**ATGTTTCCGCAAAGC-3′ and 5′-CAGCCAGAACCAGGACTG-3′. The *Drosophila groucho* cDNA was isolated by RT–PCR from embryonic RNA using the primers 5′-ACATGACC**ATG**GTTCCCT-

CACCGGTGCGCC-3′ and 5′-ACCCAAGCTTGGATCCTTTTGTTT-TACTGCCGATGCT-3′. The sequence context surrounding the start codon of the *Drosophila Pax258* cDNA (Czerny *et al*., 1997) was optimized by inserting the oligonucleotide 5′-GCGGAATTCCA-CC**ATG**GGCAGTATTTCGGGTGATGGTCATGGAGGCGTTAATC-AAC-3′ by PCR. Myc or Flag epitope tags were added by PCR at the N-terminus of the expression constructs. The Pax5 mutants ΔOP, Δ9A and Y179E were generated by PCR-based mutagenesis in addition to the previously described mutants B4, B8, B9 and ΔHD (Dörfler and Busslinger, 1996; Eberhard and Busslinger, 1999). A PCR fragment encoding amino acids 146–391 of Pax5 was inserted into the expression vector pPuroGal4 (Alkema *et al*., 1997) to obtain plasmid pPuroGal4-Pax5, which was used for transfection of U2-OS cells. VP16 and VP16-Grg4 expression plasmids were generated by inserting the *Hin*dIII–*Eco*RI fragment of the yeast vector pVP16 or the respective two-hybrid clone into pKW2T. The VP16-Gro construct contained a PCR fragment coding for amino acids 1–158 of *Drosophila* Groucho in the *Not*I site of pVP16. All GST–Grg4 constructs were obtained by insertion of the respective PCR fragments into pGEX vectors (Amersham Pharmacia Biotech).

Cell lines

The murine plasmacytoma cells SP2/0 and J558L, murine COP-8 and NIH 3T3 fibroblasts, monkey COS-7 kidney cells, and human U2-OS osteosarcoma and HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum. These cell lines do not express endogenous Pax5 (Dörfler and Busslinger, 1996).

Yeast two-hybrid screen

The Matchmaker™ two-hybrid system (Clontech) was used to screen for interaction partners of Pax5. Briefly, a PCR fragment of human Pax5 (amino acids 146–391) was inserted into the *Bam*HI site of the Gal4 DNA-binding domain plasmid pGBT9. This construct was transformed into the yeast strain HF7c (Clontech). Single colonies grown on synthetic medium lacking tryptophan were transformed with a mouse cDNA-VP16 fusion library prepared from RNA of 9.5/10.5-day-old embryos (Hollenberg *et al*., 1995). One colony out of 3.2×10^6 transformants activated the *HIS3* gene in a Pax5 bait-dependent manner, as it contained a VP16-Grg4 cDNA fragment.

GST pull-down assay

Purified GST fusion proteins (2–5 μg) immobilized on glutathione–Sepharose beads were incubated for 2 h at 4°C with *in vitro* synthesized ^{35}S-labeled protein (5 μl) in buffer BC100 (200 μl) supplemented with 2 mg/ml bovine serum albumin (BSA), 0.2% NP-40 and 100 μg/ml ethidium bromide, then washed extensively, eluted and analyzed by SDS–PAGE as previously described (Eberhard and Busslinger, 1999).

Cell transfection assay

J558L and SP2/0 cells were transiently transfected by electroporation with the firefly luciferase reporter gene *luc-CD19*, the *Renilla* luciferase plasmid pRL-SV40 (Promega) and Pax5 and Grg4 expression plasmids, as described (Dörfler and Busslinger, 1996). The COP-8 and U2-OS cells were transfected with LipofectAMINE PLUS™ Reagent (Gibco-BRL) and HeLa cells with the FuGENE™ 6 Transfection Reagent (Boehringer Mannheim) according to the supplier's instructions. Luciferase activities were measured by the Dual-Luciferase™ reporter assay (Promega) in a Lumat LB 9507 bioluminescence counter (EG&G Berthold, Bad Wildbad, Germany).

Co-immunoprecipitation analysis

Transiently transfected COP-8 cells were lysed in buffer A [20 mM Tris–HCl pH 7.9, 120 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 0.2% NP-40, 1 mM dithiothreitol (DTT), 10% glycerol supplemented with 2 mM benzamidine hydrochloride, 0.1 mg/ml Pefabloc, 5 μg/ml each of pepstatin, leupeptin and aprotinin, 2 μg/ml each of antipain and chymostatin]. Lysates were incubated for 30 min on ice, cleared from cellular debris by centrifugation and subsequently mixed with 10 μl of anti-Flag M2 affinity beads (Sigma) for 2 h at 4°C under constant rotation. After extensive washing of the beads, the precipitated proteins were analyzed by SDS–PAGE and Western blotting using polyclonal anti-Myc and anti-Pax5 antibodies.

Generation of transgenic flies and analysis of Sxl expression

The *hb-h$^{pax2.1}$* transgene was constructed and injected into *y w* embryos as described (Jiménez *et al*., 1997). The analysis of several transformant lines yielded equivalent results. For the experiments shown in Figure 9, males carrying an insertion on the X chromosome were crossed with wild-type or mosaic *gro* females (see below), so that all female embryos inherited the *hb-h$^{pax2.1}$* transgene. Embryos deprived of maternal *gro* function were obtained using the *gro^{E48}* allele and the *ovo^{D}-FLP-FRT* system as described (Jiménez *et al*., 1999). This system generates homozygous mutant clones in the germline of heterozygous females, thus circumventing the lethality of homozygous *gro* females. Sxl staining was performed with a monoclonal antibody specific for the active form of the protein (Bopp *et al*., 1991).

Accession number

The mouse *Grg4* cDNA has been submitted to DBBJ/EMBL/GenBank (accession No. AF229633).

Acknowledgements

We thank S.Stifani for anti-TLE antibodies, S.M.Hollenberg for the cDNA library, H.Clevers and R.Di Lauro for plasmids, M.van Lohuizen for U2-OS cells, and D.Ish-Horowicz for his support. D.E. was a recipient of a Marie Curie fellowship. This work was supported by the I.M.P. Vienna and, in part, by the Austrian Industrial Research Promotion Fund.

References

Adams,B., Dörfler,P., Aguzzi,A., Kozmik,Z., Urbánek,P., Maurer-Fogy,I. and Busslinger,M. (1992) *Pax-5* encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS and adult testis. *Genes Dev.*, **6**, 1589–1607.

Alkema,M.J., Jacobs,J., Voncken,J.W., Jenkins,N.A., Copeland,N.G., Satijn,D.P.E., Otte,A.P., Berns,A. and van Lohuizen,M. (1997) *MPc2*, a new murine homolog of the *Drosophila* Polycomb protein is a member of the mouse Polycomb transcriptional repressor complex. *J. Mol. Biol.*, **273**, 993–1003.

Aronson,B.D., Fisher,A.L., Blechman,K., Caudy,M. and Gergen,J.P. (1997) Groucho-dependent and -independent repression activities of Runt domain proteins. *Mol. Cell. Biol.*, **17**, 5581–5587.

Bopp,D., Bell,L.R., Cline,T.W. and Schedl,P. (1991) Developmental distribution of female-specific Sex-lethal proteins in *Drosophila melanogaster*. *Genes Dev.*, **5**, 403–415.

Burri,M., Tromvoukis,Y., Bopp,D., Frigerio,G. and Noll,M. (1989) Conservation of the paired domain in metazoans and its structure in three isolated human genes. *EMBO J.*, **8**, 1183–1190.

Busslinger,M. and Nutt,S.L. (1998) Role of the transcription factor BSAP (Pax-5) in B-cell development. In Monroe,J.G. and Rothenberg,E.V. (eds), *Molecular Biology of B-Cell and T-Cell Development*. Humana Press Inc., New York, NY, pp. 83–110.

Chen,G., Nguyen,R.H. and Courey,A.L. (1998) A role for Groucho tetramerization in transcriptional repression. *Mol. Cell. Biol.*, **18**, 7259–7268.

Chen,G., Fernandez,J., Mische,S. and Courey,A.J. (1999) A functional interaction between the histone deacetylase Rpd3 and the corepressor Groucho in *Drosophila* development. *Genes Dev.*, **13**, 2218–2230.

Choi,C.Y., Kim,Y.H., Kwon,H.J. and Kim,Y. (1999) The homeodomain protein NK-3 recruits Groucho and a histone deacetylase complex to repress transcription. *J. Biol. Chem.*, **274**, 33194–33197.

Czerny,T., Bouchard,M., Kozmik,Z. and Busslinger,M. (1997) The characterization of novel *Pax* genes of the sea urchin and *Drosophila* reveal an ancient evolutionary origin of the *Pax2/5/8* family. *Mech. Dev.*, **67**, 179–192.

Dörfler,P. and Busslinger,M. (1996) C-terminal activating and inhibitory domains determine the transactivation potential of BSAP (Pax-5), Pax-2 and Pax-8. *EMBO J.*, **15**, 1971–1982.

Dubnicoff,T., Valentine,S.A., Chen,G., Shi,T., Lengyel,J.A., Paroush,Z. and Courey,A.J. (1997) Conversion of Dorsal from an activator to a repressor by the global corepressor Groucho. *Genes Dev.*, **11**, 2952–2957.

Eberhard,D. and Busslinger,M. (1999) The partial homeodomain of the transcription factor Pax-5 (BSAP) is an interaction motif for the retinoblastoma and TATA-binding proteins. *Cancer Res.*, **59**, 1716s–1724s.

Fisher,A.L. and Caudy,M. (1998) Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. *Genes Dev.*, **12**, 1931–1940.

Fu,W. and Noll,M. (1997) The *Pax2* homolog *sparkling* is required for

development of cone and pigment cells in the *Drosophila* eye. *Genes Dev.*, **11**, 2066–2078.
Goldstein,R.E., Jiménez,G., Cook,O., Gur,D. and Paroush,Z. (1999) Huckebein repressor activity in *Drosophila* terminal patterning is mediated by Groucho. *Development*, **126**, 3747–3755.
Hollenbach,A.D., Sublett,J.E., McPherson,C.J. and Grosveld,G. (1999) The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repessor hDaxx. *EMBO J.*, **18**, 3702–3711.
Hollenberg,S.M., Sternglanz,R., Cheng,P.F. and Weintraub,H. (1995) Identification of a new family of tissue-specific basic helix–loop–helix proteins with a two-hybrid system. *Mol. Cell. Biol.*, **15**, 3813–3822.
Husain,J., Lo,R., Grbavec,D. and Stefani,S. (1996) Affinity for the nuclear compartment and expression during cell differentiation implicate phosphorylated Groucho/TLE1 forms of higher molecular mass in nuclear functions. *Biochem. J.*, **317**, 523–531.
Jiménez,G., Paroush,Z. and Ish-Horowicz,D. (1997) Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. *Genes Dev.*, **11**, 3072–3082.
Jiménez,G., Verrijzer,C.P. and Ish-Horowicz,D. (1999) A conserved motif in Goosecoid mediates Groucho-dependent repression in *Drosophila* embryos. *Mol. Cell. Biol.*, **19**, 2080–2087.
Koop,K.E., MacDonald,L.M. and Lobe,C.G. (1996) Transcripts of *Grg4*, a murine *groucho*-related gene, are detected in adjacent tissues to other murine neurogenic gene homologues during embryonic development. *Mech. Dev.*, **59**, 73–87.
Lechner,M.S. and Dressler,G.R. (1996) Mapping of Pax-2 transcription activation domains. *J. Biol. Chem.*, **271**, 21088–21093.
Leon,C. and Lobe,C.G. (1997) *Grg3*, a murine *groucho*-related gene, is expressed in the developing nervous system and in mesenchyme-induced epithelial structures. *Dev. Dynam.*, **208**, 11–24.
Magnaghi,P., Roberts,C., Lorain,S., Lipinski,M. and Scambler,P.J. (1998) HIRA, a mammalian homologue of *Saccharomyces cerevisiae* transcriptional co-repressors, interacts with Pax3. *Nature Genet.*, **20**, 74–77.
Mallo,M., Franco del Amo,F. and Gridley,T. (1993) Cloning and developmental expression of *Grg*, a mouse gene related to the *groucho* transcript of the *Drosophila Enhancer of split* complex. *Mech. Dev.*, **42**, 67–76.
Noll,M. (1993) Evolution and role of *Pax* genes. *Curr. Opin. Genet. Dev.*, **3**, 595–605.
Nutt,S.L., Urbánek,P., Rolink,A. and Busslinger,M. (1997) Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced *V*-to-*DJ* recombination at the *IgH* locus. *Genes Dev.*, **11**, 476–491.
Nutt,S.L., Morrison,A.M., Dörfler,P., Rolink,A. and Busslinger,M. (1998) Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. *EMBO J.*, **17**, 2319–2333.
Nutt,S.L., Heavey,B., Rolink,A.G. and Busslinger,M. (1999) Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*, **401**, 556–562.
Palaparti,A., Baratz,A. and Stifani,S. (1997) The Groucho/transducin-like enhancer of split transcriptional repressors interact with the genetically defined amino-terminal silencing domain of histone H3. *J. Biol. Chem.*, **272**, 26604–26610.
Parkhurst,S.M. (1998) Groucho: making its Marx as a transcriptional co-repressor. *Trends Genet.*, **14**, 130–132.
Parkhurst,S.M., Bopp,D. and Ish-Horowicz,D. (1990) X:A ratio, the primary sex-determining signal in *Drosophila*, is transduced by helix–loop–helix proteins. *Cell*, **63**, 1179–1191.
Paroush,Z., Finley,R.L., Kidd,T., Wainwright,S.M., Ingham,P.W., Brent,R. and Ish-Horowicz,D. (1994) Groucho is required for *Drosophila* neurogenesis, segmentation and sex determination and interacts directly with Hairy-related bHLH proteins. *Cell*, **79**, 805–815.
Paroush,Z., Wainwright,S.M. and Ish-Horowicz,D. (1997) Torso signalling regulates terminal patterning in *Drosophila* by antagonising Groucho-mediated repression. *Development*, **124**, 3827–3834.
Pfeffer,P.L., Gerster,T., Lun,K., Brand,M. and Busslinger,M. (1998) Characterization of three novel members of the zebrafish *Pax2/5/8* family: dependency of *Pax5* and *Pax8* expression on the *Pax2.1* (*noi*) function. *Development*, **125**, 3063–3074.
Ren,B., Chee,K.J., Kim,T.H. and Maniatis,T. (1999) PRDI-BF1/Blimp-1 repression is mediated by corepressors of the Groucho family of proteins. *Genes Dev.*, **13**, 125–137.
Rolink,A.G., Nutt,S.L., Melchers,F. and Busslinger,M. (1999) Long-term *in vivo* reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature*, **401**, 603–606.
Roose,J., Molenaar,M., Peterson,J., Hurenkamp,J., Brantjes,H., Moerer,P., van de Wetering,M., Destrée,O. and Clevers,H. (1998) The *Xenopus* Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature*, **395**, 608–612.
Schmidt,C.J. and Sladek,T.E. (1993) A rat homolog of the *Drosophila* Enhancer of split (*groucho*) locus lacking WD-40 repeats. *J. Biol. Chem.*, **268**, 25681–25686.
Singh,M. and Birshtein,B.K. (1996) Concerted repression of an immunoglobulin heavy-chain enhancer, 3′αE (hs1,2). *Proc. Natl Acad. Sci. USA*, **93**, 4392–4397.
Smith,S.T. and Jaynes,J.B. (1996) A conserved region of engrailed, shared among all en-, gsc-, Nk1-, Nk2- and msh-class homeoproteins, mediates active transcriptional repression *in vivo*. *Development*, **122**, 3141–3150.
Stifani,S., Blaumueller,C.M., Redhead,N.J., Hill,R.E. and Artavanis-Tsakonas,S. (1992) Human homologs of a *Drosophila Enhancer of split* gene product define a novel family of nuclear proteins. *Nature Genet.*, **2**, 119–126.
Urbánek,P., Wang,Z.-Q., Fetka,I., Wagner,E.F. and Busslinger,M. (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell*, **79**, 901–912.
Urbánek,P., Fetka,I., Meisler,M.H. and Busslinger,M. (1997) Cooperation of *Pax2* and *Pax5* in midbrain and cerebellum development. *Proc. Natl Acad. Sci. USA*, **94**, 5703–5708.
Valentine,S.A., Chen,G., Shandala,T., Fernandez,J., Mische,S., Saint,R. and Courey,A.J. (1998) Dorsal-mediated repression requires the formation of a multiprotein repression complex at the ventral silencer. *Mol. Cell. Biol.*, **18**, 6584–6594.
Zannini,M., Francis-Lang,H., Plachov,D. and Di Lauro,R. (1992) Pax-8, a paired domain-containing protein, binds to a sequence overlapping the recognition site of a homeodomain and activates transcription from two thyroid-specific promoters. *Mol. Cell. Biol.*, **12**, 4230–4241.

Received December 21, 1999; revised and accepted March 16, 2000

Note added in proof

Sequence analysis of the mouse Grg4 protein with hidden Markov models derived from two different alignments of WD40 repeats [present in PFAM (Bateman *et al*., 2000) and REP (Andrade *et al*., 2000)] revealed five common hits with both alignments. The two models additionally located a sixth repeat at non-overlapping positions within Grg4. Hence, Grg proteins also contain seven WD40 repeats like other proteins containing WD40 motifs (Frank Eisenhaber, unpublished data).

Andrade,M.A., Ponting,C.P., Gibson,T.J. and Bork,P. (2000) Homology-based method for identification of protein repeats using statistical significance estimates. *J. Mol. Biol.*, in press.
Bateman,A., Birney,E., Durbin,R., Eddy,S.R., Howe,K.L. and Sonnhammer,E.L. (2000) The Pfam protein families database. *Nucleic Acids Res.*, **28**, 263–266.