THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 51, pp. 36285–36301, December 20, 2013

Several Cis-regulatory Elements Control mRNA Stability, Translation Efficiency, and Expression Pattern of Prrxl1 (Paired Related Homeobox Protein-like 1)*[S]

Received for publication, June 6, 2013, and in revised form, November 7, 2013 Published, JBC Papers in Press, November 8, 2013, DOI 10.1074/jbc.M113.491993

Isabel Regadas[‡§], Mariana Raimundo Matos[‡§], Filipe Almeida Monteiro[‡§], José Luis Gómez-Skarmeta[¶], Deolinda Lima[‡§], José Bessa[§¶], Fernando Casares[¶], and Carlos Reguenga[‡§1]

From the [‡]Departamento de Biologia Experimental, Faculdade de Medicina do Porto, Universidade do Porto, Porto 4200-319, Portugal, [§]Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto 4150, Portugal, and [¶]CABD (Consejo Superior de Investigaciones Científicas-UPO-Junta de Andalucía), Seville 41013, Spain

Background: The mechanisms that control the *Prrxl1* expression are poorly understood.
Results: Several regulatory elements present in *Prrxl1* alternative promoters are functionally characterized, including a binding motif for Phox2b required for *Prrxl1* expression in visceral sensory neurons.
Conclusion: We define diverse regulatory modules, which control the spatiotemporal expression of *Prrxl1* in nociceptive neurons.
Significance: A new mechanism involved in the ganglion specific action of *Prrxl1* is described.

The homeodomain transcription factor Prrxl1/DRG11 has emerged as a crucial molecule in the establishment of the pain circuitry, in particular spinal cord targeting of dorsal root ganglia (DRG) axons and differentiation of nociceptive glutamatergic spinal cord neurons. Despite Prrxl1 importance in the establishment of the DRG-spinal nociceptive circuit, the molecular mechanisms that regulate its expression along development remain largely unknown. Here, we show that *Prrxl1* transcription is regulated by three alternative promoters (named P1, P2, and P3), which control the expression of three distinct *Prrxl1* 5′-UTR variants, named 5′-UTR-A, 5′-UTR-B, and 5′-UTR-C. These 5′-UTR sequences confer distinct mRNA stability and translation efficiency to the *Prrxl1* transcript. The most conserved promoter (P3) contains a TATA-box and displays *in vivo* enhancer activity in a pattern that overlaps with the zebrafish Prrxl1 homologue, *drgx*. Regulatory modules present in this sequence were identified and characterized, including a binding site for Phox2b. Concomitantly, we demonstrate that zebrafish Phox2b is required for the expression of *drgx* in the facial, glossopharyngeal, and vagal cranial ganglia.

Sensory perception of peripheral stimuli is primarily mediated by different types of afferent neurons, which are located in the trunk dorsal root ganglia (DRG)[2] and send processes to the periphery and the spinal cord. In the spinal cord, specialized neurons integrate and relay the information to somatosensory centers in the brain where appropriate responses are generated (1, 2). Spinal sensory neurons differentiate from several classes of proliferating progenitor cells, whose establishment requires the expression of proneural genes encoding for basic helix-loop-helix (bHLH) transcription factors (3–5). The various cell lineages are specified according to the combined expression of a set of homedomain transcription factors that confer neural identity to each class (4, 5). One of these factors is Prrxl1 (also known as DRG11). Prrxl1 has emerged as a crucial molecule in the development of the pain-perception circuitry, especially in the establishment of the nociceptive DRG-spinal pathway. Prrxl1 is also expressed in sensory cranial ganglia and their target relay neurons in the hindbrain (6, 7). Nevertheless, the role of Prrxl1 in these tissues is poorly studied.

Prrxl1 null mutant mice present a distorted spinal dorsal horn with scarce superficial nociceptive-responsive neurons (8–10), reduced DRG neuronal population (10), and a marked decrease in nociceptive response capacity in various pain tests (8). Interestingly, although involved in the embryonic differentiation of various subpopulations of superficial dorsal horn excitatory neurons (10), Prrxl1 appears not to be required for the normal development of DRG neurons before birth but rather to be essential for their survival in early postnatal life (9). Although Prrxl1 expression in various cell lineages in DRG and spinal cord is well known, the mechanisms of transcriptional control exerted by different bHLH and homeodomain proteins that modulate *Prrxl1* transcription are still poorly understood.

Recently, a *Prrxl1* alternative spliced variant was identified, and multiple variants of exon 1 in both *Prrxl1* mRNA isoform sequences were discovered, suggesting the existence of various 5′-untranslated regions (5′-UTRs) controlled by distinct promoters (11). Modulation of gene expression through alternative promoter usage is now widely accepted following evidence gathered in the past years (12, 13). According to Baek *et al.* (14),

* This work was supported by the Fundação para a Ciência e a Tecnologia (SFRH/BD/65300/2009 (to I. R.) and PTDC/SAU-OBD/099886/2008), COMPETE: FCOMP-01-0124-FEDER-011262, and Universidade do Porto/Banco Santander Totta (Projectos Pluridisciplinares).
[S] This article contains supplemental Table S1.
[1] To whom correspondence should be addressed: Dept. of Experimental Biology, 4th floor, Faculty of Medicine of Porto, 4200-319 Porto, Portugal. Tel.: 351-220426743; E-mail: cregueng@med.up.pt.
[2] The abbreviations used are: DRG, dorsal root ganglia; bHLH, helix-loop-helix; TSS, transcription start site; HD, homeodomain; TBP, TATA-binding protein; MO, morpholino oligonucleotide; RRA, regulatory region A; RRB, regulatory region B; eGFP, enhanced-GFP; hpf, hours post fertilization; Ngn1, neurogenin1.

about 40–50% of human and mouse genes contain alternative promoters, a condition that seems to be required to initiate transcription in a tissue-specific manner (15–17). The use of multiple promoters, each one controlling at least one transcription start site (TSS), usually originates different 5′-UTRs that might have a role in the control of mRNA stability or translation efficiency (18, 19).

Here, we characterize three *Prrxl1* 5′-UTRs variants and the corresponding promoter regions, which may explain the differential involvement of Prrxl1 in the DRG and spinal cord development. We also present *in vitro* and *in vivo* evidence that the most evolutionarily conserved *Prrxl1* promoter region is sufficient to drive expression to neuronal cells and is regulated by Phox2b specifically in primary afferent neurons.

EXPERIMENTAL PROCEDURES

Animal Care—NMRI mice were bred and housed at the Instituto de Biologia Molecular e Celular, Porto, animal facility under temperature- and light-controlled conditions. The embryonic day 0.5 (E0.5) was considered to be the midday of the vaginal plug. The animals were euthanized (isoflurane anesthesia followed by cervical dislocation), and tissues were collected. Experiments were carried out in compliance with the animal ethics guidelines at Instituto de Biologia Molecular e Celular and approved by the Portuguese Veterinary Ethics Committee.

Wild-type AB/Tuebingen (AB/TU) zebrafish strain were maintained in the breeding colony in CABD, Seville, according to standard procedures. Fertilized eggs were kept at 28 °C in E3 medium with 0.003% 1-phenyl-2-thiourea to prevent pigmentation and were staged according to Kimmel *et al.* (20).

Reverse Transcriptase-PCR—The different 5′-UTR-*Prrxl1* molecules were amplified by reverse transcriptase-PCR (see supplemental Table S1 for primers) from spinal cord total RNA, extracted from mice at different developmental stages (E11.0, E12.5, E14.5, and E16.5) using the Micro-to-midi total RNA purification System (Invitrogen) following the manufacturer's instructions. The first-strand cDNA synthesis was prepared at 42 °C during 1 h from 1 μg of total RNA using 200 units of transcriptase enzyme (Bioline) and 500 ng of oligo(dT)$_{12-18}$ (Bioline). To assess for potential contaminants, a control containing all reagents except the reverse transcriptase enzyme was included for each sample. Normalization was performed by amplification of mouse β-actin using the primers pair listed in supplemental Table S1. The PCR conditions were the following: denaturation at 94 °C for 30 s, annealing at 58 °C for 45 s, and elongation at 72 °C for 45 s. Thirty-two cycles were performed for the amplification of *Prrxl1* 5′-UTR-B and 5′-UTR-C, 29 for 5′-UTR-A and ORF, and 20 cycles for β-actin. The amplification for each gene was in the linear curve (data not shown). Equal amounts of the PCR products were subjected to a 1% agarose gel electrophoresis and visualized by ethidium bromide staining under UV light source. The signals were acquired by a Kodak digital camera DC290, and the densitometric analyses were conducted using the computational program Kodak 1D Image Analysis Software.

In Vitro Transcription-Translation Assay—The different full-length 5′-UTR-*Prrxl1* cDNA sequences (see supplemental Table S1 for primers) were PCR-amplified from mouse E14.5 spinal cord cDNA and cloned by TA overhangs in the pCR2.1 (Invitrogen). The resulting vectors were selected for orientation and used to perform the coupled transcription-translation assay in rabbit reticulocyte lysates using the PROTEINscript® II T7 kit (Ambion). A sequence corresponding to nucleotides −1 to −50 shared by all isoforms was also cloned into the pCR2.1 plasmid and used as a control. The Prrxl1 expression was measured by Western blotting using our homemade rabbit anti-Prrxl1 antibody as described previously (6) and normalized with a mouse anti-tubulin antibody (Sigma).

Expression Vectors—Plasmids used in this work were pRSK-Brn3a (a gift from Dr. Mengqing Xiang), pcDNA3.3-Tlx3, pCAGGS-mPhox2b (a gift from Dr. Christo Goridis), pcDNA3-Islet1 (a gift from Dr. Chunyan Zhou), pcDNA3.1-His-Ngn1 (a gift from Dr. Soyeon Kim) pcDNA3.3-Lmx1b, and pCAGGS-FLAG-Mash1 (a gift from Dr. Diogo S. Castro). The sequences corresponding to the Tlx3 and Lmx1b open reading frame were amplified from mouse E14.5 spinal cord cDNA (for primers see supplemental Table S1) and cloned in the pCDNA3.3-TOPO TA cloning vector (Invitrogen). Protein expression in transfected cells was assessed by Western blotting using the antibodies mouse anti-Brn3a (Santa Cruz Biotechnology), rabbit anti-Tlx3 (Santa Cruz Biotechnology), rabbit anti-Phox2b (a gift from Dr. Qiufu Ma), mouse anti-Islet-1 (40.2D6, Developmental Studies Hybridoma Bank), mouse anti-polyhistidines (Sigma), rabbit anti-Lmx1b (a gift from Dr. Thomas Müller), and mouse anti-FLAG (Sigma).

For the mRNA stability assays, the different 5′-UTR-specific sequences were amplified from E14.5 mouse spinal cord cDNA and cloned by TA overhangs in the pCR2.1 plasmid (Invitrogen). These sequences were then subcloned in the HindIII site of the pGL3-Control vector (Promega) between the SV40 promoter and the luciferase coding region.

To determine the *Prrxl1* alternative promoters, the entire and overlapping fragments of the −1401/−50-bp region upstream of the start codon were PCR-amplified (for primers see supplemental Table S1) from mouse genomic DNA and cloned in the pBlue-TOPO vector (Invitrogen). The sequences were then subcloned in the HindIII site of the promoter-less vector, pGL3-Basic (Promega).

Site-directed mutagenesis of TATA box and HD element were performed from pGL3-REG1 using the QuikChange Site-directed Mutagenesis kit from Stratagene and following the manufacturer's instructions. Primers used are depicted in supplemental Table S1.

Cell Culture—ND7/23, HeLa, HEK293, and PC12 cell lines were maintained and grown in Dulbecco's modified Eagle's minimal essential medium (DMEM; Invitrogen) containing 10% fetal bovine serum (Invitrogen) and 50 units/ml penicillin and streptomycin (Invitrogen). All cells were kept at 37 °C and 5% CO_2 gas phase. Transfection was performed using Lipofectamine2000 agent (Invitrogen), and 24 h later cells were harvested for posterior analysis. Overexpression in differentiated PC12 cells was performed by transfecting PC12 cells, and 6 h later differentiation was induced with 100 ng/ml NGF (Sigma). Luciferase reporter assays were performed 2 days after transfection. To infer the mRNA stability, 24 h after transfection, 100 μg/ml actinomycin D (A9415, Sigma) was added to DMEM

(Invitrogen), and luciferase activity was measured 0, 3, and 6 h later.

DRG Primary Culture and Cell Electroporation—DRG were extracted from newly born mice and, after a 2-h treatment with 10% collagenase (Sigma), were electroporated using the Neon™ Transfection System (Invitrogen) according to the manufacturer's instructions. Afterward, cells were cultured in polyornithine-coated wells in DMEM-F-12 (Invitrogen) containing 10% fetal bovine serum (Invitrogen), 2 mM glutamine (Invitrogen), 4% Ultroser G (Pall), 1× B27 (Sigma), and 10 ng/ml NGF (Sigma). Twenty-four hours later the cells were harvested and processed for luciferase reporter assays.

Luciferase Reporter Assays—Transfected cells from a 96-well plate format were resuspended in 50 μl of lysis buffer (Promega), and the protein extract was cleared by centrifugation. 5 μl of the extract were mixed with the luciferase reagent (Promega), and the signals were measured using a luminometer reader (Tecan). Transfection efficiency was normalized by assessing the β-galactosidase activity using 2-nitrophenyl β-D-galactopyranoside (Sigma) as substrate.

Electrophoretic Mobility Shift Assay—Recombinant TATA-binding protein (sc-4000, Santa Cruz Biotechnology, Inc.) was incubated with 50 fmol of the double-stranded oligonucleotide TATA containing the TATAbox promoter sequence (5′-TTATGCGTGAGAT**TATA**AAGGCGAGTGCTGAGCGGCGGCGCGCGCTG-3′) and end-labeled with a dyomic dye DY682 (Thermo Scientific) in a buffer containing 12 mM Tris-HCl, pH 8.0, 0.15 mM EDTA, 6 mM $MgCl_2$, 90 mM KCl, 1 mM DTT, 10% glycerol, 0.5 mg/ml BSA, and 0.4 μg/μl poly(dI/dC) (adapted from Riquet *et al.* 21). Competition experiments were performed using a non-labeled oligonucleotide with the same sequence. For supershift experiments, 200 ng of the anti-TBP antibody (sc-273, Santa Cruz Biotechnology, Inc.) were used. All samples were run in a 5% PAGE with 10 mM $MgCl_2$.

Nuclear proteins were extracted from ND7/23 cells previously transfected with pcDNA3.3, pcDNA3.3-Tlx3, pCAGGS-mPhox2b, or pRSK-Brn3a using first a low salt lysis buffer (30 mM Tris-HCl, pH 7.8, 20 mM NaCl, 1 mM EDTA, 1 mM DTT, 0,1% Triton X-100, and proteases and phosphatases inhibitors cocktails). After nuclear fractionation, proteins were resuspended in 10 mM Tris-HCl, pH 7.5, 60 mM KCl, 200 mM NaCl, 10% glycerol, 5 mM $MgCl_2$, and 0.1% Triton X-100, cleared by centrifugation, and incubated with 50 fmol of the double-stranded oligonucleotide HD (5′-CTGGAAA**TAAT**CAG**ATTA**AGGC-3′) end-labeled with dyomic dye DY682. The samples were run in a 5% polyacrylamide electrophoresis get. The fluorescent signals were detected using the Odyssey Infrared Imaging System (LI-COR Biosciences).

Chromatin Immunoprecipitation Assays—For TATA-binding protein (TBP) chromatin immunoprecipitation (ChIP) assays, dorsal spinal cords from E14.5 mouse embryos were dissected and fixed with 2 mM di(*N*-succinimidyl) glutarate (Sigma) in phosphate buffer saline (PBS) for 45 min followed by 1% formaldehyde in PBS for 10 min and lysed in 50 mM Tris-HCl, pH 8.0, 1% SDS, and 10 mM EDTA. Chromatin shearing was performed using Bioruptor (Diagenode) at high power settings for 60 cycles (30 s on/30 s off). ChIP assays with or without (mock control) mouse monoclonal anti-TBP antibody (Mab-002–100, Diagenode) were performed using 80 μg of chromatin/assay in ChIP buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 5 mg/ml BSA), and protease inhibitor mixture (Roche Applied Science). Immunoprecipitates were retrieved with 50 μl of Protein G Dynabeads (Invitrogen) per assay and washed once with wash buffer I (20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), once with wash buffer II (20 mM Tris, pH 8.0, 250 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), twice with wash buffer III (10 mM Tris pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% Nonidet P-40, 1% sodium deoxycholate), and once with TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA) and eluted with lysis buffer at 65 °C for 10 min. Eluted and input chromatin were subjected to proteinase K (Roche Applied Science) treatment for 2 h at 42 °C and reverse-cross-linked at 65 °C overnight. Immunoprecipitated and input DNA samples were purified by phenol-chloroform extractions followed by isopropyl alcohol precipitation. DNA sequences were quantified by real-time PCR (primers are listed in supplemental Table S1) using a StepOnePlus Real Time PCR system (Applied Biosystems) and a SYBR Green chemistry for quantitative PCR (Maxima master mix, Fermentas). Quantities of immunoprecipitated DNA were calculated by comparison with a standard curve generated by serial dilutions of input DNA. Data were plotted as the means of at least two independent ChIP assays and three independent amplifications; *error bars* represent S.E.

Phox2b ChIP assays were performed essentially as described above, with the following modifications: (i) chromatin samples were extracted from dorsal medulla oblongata of E14.5 mouse embryos; (ii) mouse monoclonal anti-Phox2b antibody (sc-376997, Santa Cruz Biotechnology) was used; (iii) NaCl concentration of ChIP buffer was reduced to 20 mM; (iv) immunoprecipitates were washed 6 times with wash buffer III containing only 0.7% sodium deoxycholate.

Production of Transgenic Zebrafish—A sequence (−751/−584 bp) that includes the P3 core promoter was PCR amplified from mouse genomic DNA, cloned in the pCR/GW/TOPO vector (Invitrogen), and then recombined by the Gateway *in vitro* recombination technology using the Gateway LR Clonase II Enzyme mix (Invitrogen) into a Tol2 vector (22) containing an *iroquois* enhancer with midbrain activity (Z48; Refs. 23 and 24) and the enhanced GFP reporter gene. This vector was assembled by cloning a SalI/NotI fragment of pCS2eGFP (25) containing the CMV promoter, the enhanced GFP reporter gene, and the poly(A) of SV40 into SalI/NotI restriction sites of a modified pminiTol2/MCS vector (26) that has a fragment of the pUC19 polylinker that goes from EcoRI to HindIII. Z48 *iroquois* enhancer was isolated from the Z48 TOPO vector (23) by cutting with EcoRI and cloned into NotI restriction site after blunting with Klenow. Finally, a gateway cassette (Invitrogen) was cloned in blunt between the SalI/BamHI restriction sites, replacing the CMV promoter. This vector lacks a promoter and does not drive GFP reporter expression. Therefore, it is useful to test for promoter activity of selected DNA sequences. The following primers were used for amplification: 5′-TAAGGCCCAATAGACCTATC-3′ and 5′-CAGGACCAGAGAAGTGACTG-3′). About 5 nl of the reaction mix containing 50 ng/μl transposase mRNA, 50 ng/μl phenol/chloroform purified vec-

tor, and 0.05% phenol red were injected in the cell of one-stage zebrafish embryos. The GFP expression was then documented from the next 24 to 72 h.

The fragment containing the Ebox and the HD element was PCR-amplified from human genomic DNA, cloned in the pCR/GW/TOPO plasmid (Invitrogen), and then recombined to the ZED vector (23) by the Gateway *in vitro* recombination technology described above. The following primers were used for amplification: 5′-GTGGTGGTTGTATCGTTCTC-3′ and 5′-GCATAATTGGCCTTAATCTG-3′. The injections were performed as described above. Positive transgenic embryos strongly expressing red fluorescent protein were selected 72 h later. Those F0 embryos were raised, and the F1 generation of embryos expressing GFP was analyzed.

RNA Probe Synthesis and Whole-mount in Situ Hybridization—The *drgx* and *tlx3b* sequences were amplified from zebrafish cDNA using the following primers: 5′-ATGTTTTACTTTCACTGTCCTCCA-3′ and 5′-CATTTCTTATCCGGACCCTC-3′ for *drgx* and 5′-TTCGGGTGGTGAGGATGGAC-3′ and 5′-GATTTTGGGATGCAACAGCA-3′ for *tlx3b*. PCR products were cloned in the pGEM-T Easy vector (Promega), and a phenol-chloroform purification was performed after linearization with NsiI (for drgx) or NcoI (for tlx3b). Each vector was used as a template for the *in vitro* synthesis of a DIG-labeled RNA probe for zebrafish *drgx* and *tlx3b*.

Wild-type embryos at 48 and 72 h post fertilization were fixed overnight with 4% paraformaldehyde at 4 °C. After brief washes with PBS, 0.1% Tween 20 (PBST), they were treated with 10 μg/ml proteinase K and fixed for 20 min with 4% paraformaldehyde at room temperature. After a 20-min wash with PBST, embryos were incubated in hybridization buffer for 1 h at 70 °C. The respective RNA probe was then added to the hybridization buffer (50% formamide, 2× SSC, and 0.1% Tween 20) to a final concentration of 1 ng/μl, and embryos were incubated overnight at 70 °C. The next day embryos were sequentially washed at 70 °C in solutions containing different concentrations (75, 50, 25, and 0%) of hybridization buffer diluted in 2× SSC (75 mM NaCl and 7.5 mM sodium citrate, pH 7.0). After a further wash with 0.05× SSC for at least 1 h at 70 °C, embryos were finally washed with PBST for 10 min at room temperature. Embryos were blocked in 2% normal goat serum/PBST for at least 1 h at room temperature. Anti-digoxigenin coupled with alkaline phosphatase was added in fresh 2% normal goat serum, PBST (1:5000) for 2 h and then allowed to wash overnight in PBST. The next day embryos were washed with AP reaction buffer without $MgCl_2$ (100 mM Tris, pH 9.5, 100 mM NaCl, and 0.1% Tween 20). Detection was performed with 3.5 μl of nitro blue tetrazolium (50 mg/ml) and 1 μl of 5-bromo-4-chloro-3-indolyl phosphate (50 mg/ml) per 1 ml of complete AP-reaction buffer (100 mM Tris, pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl, and 0.1% Tween 20). The signal was allowed to develop for 3–5 h.

Antisense Morpholino Oligonucleotide (MO) Analysis—Antisense MO targeted to the translation initiation site of Phox2b (CATTGAAAAGGCTCAGTGGAGAAGG) was obtained from Gene Tools, LLC, diluted to a working concentration in Milli-Q water (0.4 ng/nl) with 0.05% phenol red, and about 5 nl were injected into 1- to 2-cell-stage embryos.

Statistical Analysis—In the present study, all the data presented (except for ChIP experiments; see above) were derived from at least three independent experiments with three replicates. When necessary, a two-tailed *t* test was performed. Both mean and S.D. values were calculated and included in the figures.

RESULTS

Prrxl1 5′-UTR Variants Present Distinct mRNA Stability and Translation Efficiency—Prrxl1 is a transcription factor first identified by Saito (27) in a subtractive hybridization screening with rat DRG. The *Prrxl1* mRNA sequence described by these authors contained the start codon in exon 2, whereas exon 1 corresponded to a 5′-UTR. More recently, by the use of 5′-rapid amplification of cDNA ends assays with mice spinal cord RNA extracts, we described two novel variants of *Prrxl1* containing alternative exon 1 that gives rise to distinct 5′-UTRs (11). BLAST searches of GenBank™ database led us to identify sequences that correspond to all the three 5′-UTRs that contained both the first and the last coding exons of the annotated *Prrxl1*. We named these *Prrxl1* 5′-UTR variants 5′-UTR-A, 5′-UTR-B, and 5′-UTR-C, which encompass, respectively, nucleotides −622 to −485, −484 to −298, and −148 to −85 relative to the start codon (Fig. 1*A*).

To assess for the presence of the different *Prrxl1* 5′-UTR variants, we performed reverse transcriptase-PCR experiments using a common reverse primer mapping within exon 7 (containing the stop codon), shared by all isoforms, and isoform-specific forward primers mapping within alternative exon 1 (*arrows* in Fig. 1*A*). With this primer design, we wanted to make sure that the amplicons contained the entire coding region and to exclude from our analysis other putative non-annotated *Prrxl1* splicing variants. The three transcripts are detected in the spinal cord at developmental stages where Prrxl1 expression has been previously reported (6). The levels of amplification are suggestive that 5′-UTR-A variant is the most abundant transcript, being detectable three PCR cycles earlier than the other two 5′-UTR transcripts (Fig. 1*B*). The semiquantitative reverse transcriptase-PCR analysis with spinal cord extracts from E11.0 to E16.5, a period that includes the early-born (E10.5-E12.0) and late-born (E12.5-E14.5) neurogenesis waves, showed that 5′-UTR-A reaches the maximum level at E14.5, a profile similar to the ORF (Fig. 1*B*). On the other hand, the highest expression levels of 5′-UTR-B and 5′-UTR-C are reached at E12.5 (Fig. 1*B*), which is suggestive of a more preponderant role of these two isoforms during early neurogenesis.

Prrxl1 5′-UTR variants result from alternative processing of non-coding exon 1 and, therefore, have no consequences in the *Prrxl1* coding region. Because mRNA-untranslated regions have been associated to post-transcriptional regulation mechanisms (18, 19), we wondered if 5′-UTR sequences confer distinct mRNA stability or translation efficiency to *Prrxl1* transcripts. To address this question, we performed coupled transcription-translation *in vitro* assays using different vectors containing the *Prrxl1* ORF associated to each 5′-UTR under the control of the bacterial T7 RNA polymerase. The amount of Prrxl1 translated was determined by Western blotting (Fig. 1*C*). *Prrxl1* transcript containing 5′-UTR-A was about three times

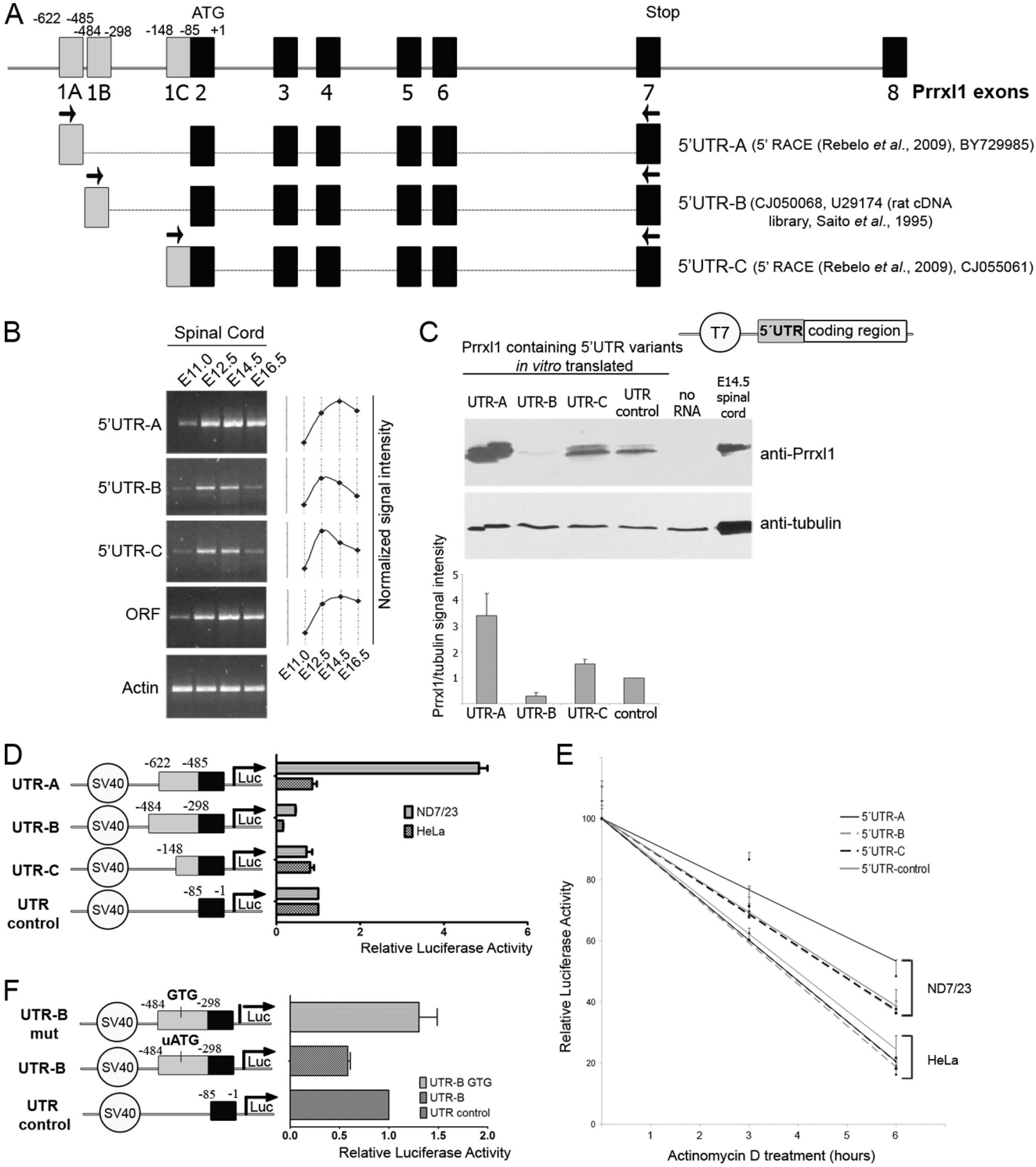

FIGURE 1. **Differential expression and stability of *Prrxl1* 5′-UTR variants.** *A*, scheme of the *Prrxl1* gene. The gene gives rise to three alternative transcripts that are composed by eight exons and differ on their untranslated first exon (*1A*, *1B*, or *1C*). Those untranslated regions were named 5′-UTR-A, 5′-UTR-B, and 5′-UTR-C. *B*, expression studies of *Prrxl1* transcripts containing different 5′-UTRs (5′-UTR-A, 5′-UTR-B, and 5′-UTR-C) and *Prrxl1* ORF by reverse transcriptase-PCR and gel electrophoresis analysis using mouse spinal cord at different developmental ages (E11 to E16.5). The graph illustrates a typical expression profile, normalized with β-actin signal intensity, from three independent experiments. *C*, Western blot analysis of Prrxl1 *in vitro* translated from mRNA containing distinct 5′-UTR and transcribed by T7 RNA Polymerase. The graph represents the mean of signal intensity (normalized with β-tubulin) from three independent experiments. *D*, luciferase reporter assays in ND7/23 and HeLa cells transfected with vectors containing the different 5′-UTR fused to the luciferase encoding gene. *E*, analysis of the stability of the luciferase mRNA molecule conferred by the distinct 5′-UTR. The constructs used in *D* were used to transfect the ND7/23 and HeLa cell lines. Transcription was halted using actinomycin D and luciferase activity measured at different time points (0, 3, and 6 h of treatment). *F*, comparison of luciferase expression regulated by different versions of 5′-UTR-B. 5′-UTR-B contains an upstream ATG (*uATG*), which was replaced by a GTG sequence (*5′-UTR-B mut*). In *C–E*, 5′-UTR control represents a fragment encompassing nucleotides −1 to −85 and shared by all isoforms.

more efficiently translated than the UTR control (a fragment immediately upstream of the start codon encompassing nucleotides −1 to −85 and shared by all isoforms). On the contrary, 5′-UTR-B led to a loss of translational rate, whereas 5′-UTR-C was as expressed as the 5′-UTR control. Similar results were obtained in luciferase reporter assays using the ND7/23 cell line transfected with each 5′-UTR cloned between the SV40 promoter and the firefly luciferase coding region (Fig. 1*D*). The ND7/23 cells are an appropriate *in vitro* model for the study of Prrxl1-associated mechanisms as they endogenously express this transcription factor and display a phenotype characteristic of nociceptive neurons (28). Interestingly, the increased luciferase activity of the 5′-UTR-A construct observed in ND7/23 cells was not observed in the non-neuronal HeLa cell line. To test if the neuron-specific activity induced by 5′-UTR-A could be due to an increase in the transcript stability, mRNA decay associated to each 5′-UTR variant was inferred by measuring the luciferase activity at different time points upon treatment with actinomycin D in both cell lines (Fig. 1*E*). All the different 5′-UTRs promoted a similar mRNA decreasing rate in HeLa cells. In ND7/23 cells, 5′-UTR-B- and 5′-UTR-C-containing transcripts displayed decay similar to the control, whereas the 5′-UTR-A was more stable. This result suggested that the 5′-UTR-A sole effect, both in the mRNA translation rate and stability, is conferred by a neuronal context likely mediated by specific RNA-binding proteins.

On the contrary, the 5′-UTR-B variant reduced the *in vitro* translation efficiency (Fig. 1*C*) and protein expression both in ND7/23 and in HeLa cells (Fig. 1*D*) without interfering with mRNA stability (Fig. 1*E*). A careful analysis of the 5′-UTR-B sequence led us to the identification of an ATG located upstream (*uATG*) of the *Prrxl1* main ATG that could be mistaken as an alternative start codon, modifying the reading frame and thereby explaining the feature of this variant (Fig. 1*F*). Upstream ORFs are widely recognized as cis-regulatory elements that can affect mRNA translation and thus are the molecular base of severe disorders (29, 30). By changing this upstream ATG to GTG, the luciferase activity of the mutated 5′-UTR-B increased to values similar to the control sequence (Fig. 1*F*), suggesting that 5′-UTR-B sequence works as a negative modulator of Prrxl1 expression levels. 5′-UTR-C did not confer any particular trait to the mRNA molecule.

Identification of Prrxl1 Alternative Promoter Regions and Evolutionarily Conserved Regulatory Elements—The existence of TSSs specific to each *Prrxl* 5′-UTR suggested that *Prrxl1* expression is controlled by a mechanism of alternative promoter usage. To identify these promoter regions and further dissect the mechanisms of *Prrxl1* regulation, we selected a region of 1351 bp (−1401/−50) upstream of the *Prrxl1* translation initiation site (+1) based on the high degree of conservation observed in a genomic alignment of homologous region in various species (human, chick, *Xenopus tropicalis*, and zebrafish) (Fig. 2). This sequence (named REG-1) was amplified and cloned into the promoter-less pGL3-basic vector. Upon transient transfection into mouse DRG primary cell culture and neuronal derived ND7/23 cells, the luciferase reporter gene expression was activated indicating the presence of promoter activity in the cloned region (Fig. 2). We then evaluated the luciferase activity of shorter overlapping sequences (named REG-2 to REG-16) in ND7/23 cells. When the entire fragment was divided in two (REG-2 and REG-11), we verified that both sequences were able to drive the transcription of the luciferase gene. Because we used a promoter-less vector, this result indicated that each fragment harbored at least one promoter. The reduction of the most distal fragment (REG-2) from its 5′ end (REG-3 to REG-7) led to the identification of a minimum sequence displaying transcriptional activity (REG-7, −772/−584). This sequence is adjacent to 5′-UTR-A TSS and, therefore, was considered to be a promoter region. The same analysis was performed for the fragment most proximal to the Prrxl1 start codon (REG-11), and two minimal fragments eliciting luciferase activity (REG-13, −622/-481 and REG-14, −157/−50), located in the vicinity of the TSS of 5′-UTR-B and 5′-UTR-C, were considered as promoter regions. By this analysis, we identified three alternative promoters, named P1 (−85/−157), P2 (−485/−604), and P3 (−622/−772) that likely control, respectively, the transcription of Prrxl1 5′-UTR-C, 5′-UTR-B, and 5′-UTR-A variants.

Some fragments, even though containing promoter regions, did not display any transcriptional activity (REG-6 and REG-9) (Fig. 2). When these fragments were shortened (REG-7 and REG-10), the luciferase activity increased to values closer to fragments of similar size, which is suggestive of the existence, in the region encompassing nucleotides −811/−772, of a regulatory element (termed regulatory region A (RRA)) with the capacity to strongly suppress the transcription of the three alternative promoters. Interestingly, by the 5′ expansion (REG-4 and REG-8) of the fragments containing the RRA, the luciferase activity increased, again revealing the presence of a new regulatory element in the region −891/−922 (termed Regulatory Region B, RRB), with the potential to inhibit the action of the repressive motif RRA and consequently to activate *Prrxl1* transcription.

It is also important to note the presence of transcriptional regulatory elements located in the sequence between −1401 and −958 bp. The abrogation of this region, which originates REG-8 construct (−958/−50), resulted in a strong decrease (about 75%) in the luciferase activity when compared with the full sequence (REG-1). Given that the deletion of the region between −1401 and −958 bp did not significantly alter the reporter activity of fragments containing P3 as sole promoter (compare the activity of REG-2 with REG-3 and REG-4 in Fig. 2), we concluded that these elements are required for transcription driven by promoters P1 and P2 rather than for the P3 promoter activity.

Because Prrxl1 expression has only been detected in neuronal tissues, namely all sensory ganglia and second order relay sensory neurons (6), we found it pertinent to investigate whether the promoter regions here identified could also drive transcription in non-neuronal cell models. Thus, we performed similar luciferase reporter assays using HeLa and HEK293 cells and compared the results with those obtained for the ND7/23 cells (Fig. 3*A*). The longer sequence in analysis (REG-1) exhibited the capacity to drive the luciferase transcription 10 (HeLa) to 15 (HEK293) times lower than when transcription was promoted in the neuronal-derived cells ND7/23 (Fig. 3*A*). An

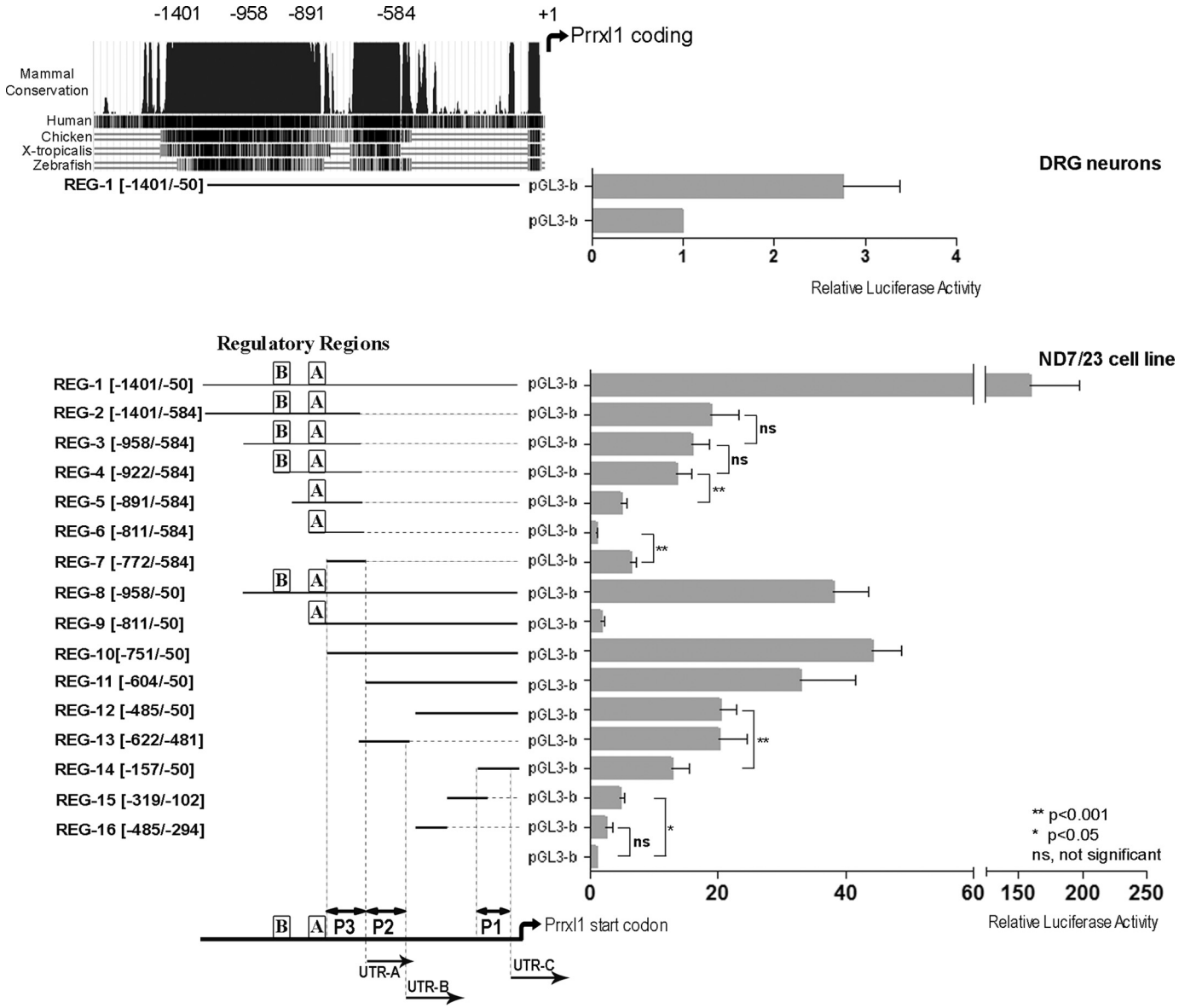

FIGURE 2. **Identification of regulatory cis-elements involved in the modulation of *Prrxl1* transcription.** Genomic alignment using the UCSC Genome Browser of the 1401-bp sequence upstream of Prrxl1 coding region from evolutionarily distant species (human, chick, *X. tropicalis* and zebrafish) is shown. *Black peaks* correspond to mammal conservation. Luciferase reporter assays using the −1401/−50-bp region were performed in mouse E15.5 DRG primary cultures and ND7/23 cells to test its transcriptional capacity. By successive deletion analysis from the longer region, three regions displaying promoter activity (termed P1, P2, and P3) were identified as well as two regulatory regions (*A* and *B*). Each promoter is located in the vicinity of the transcription start site of Prrxl1 5′-UTR.

individual analysis of each promoter and its correspondent luciferase activity demonstrated a neuron-specific activity for promoters P1 (REG-14) and P3 (REG-7), as no relevant activity in the HeLa and HEK293 cells was detected. On the other hand, promoter P2 still led to some expression of the reporter enzyme in non neuronal cells. We assume that the potential displayed by promoter P2 to drive transcription in these cells is probably responsible for the transcriptional activity exhibited by the whole fragment (REG-1) due to the presence of constitutive regulatory elements (see below).

Additionally, both regulatory regions RRA and RRB appeared to act exclusively in the neuronal model, as little differences in the transcriptional activity of the fragments REG-8 (−958/−50) and REG-9 (−811/−50) in comparison to REG-10 (−751/−50) were observed in HeLa and HEK293 cells (Fig. 3*A*).

Interestingly, a careful search for chromatin modifications in the human REG-1 sequence (chr10:50,603,551–50,604,877; GRCh37/hg19 assembly) using the UCSC Genome Browser and ENCODE annotations (31) of ChIP-seq assays revealed trimethylation in the lysine 4 of the histone H3 (H3K4me3) in all the three promoter regions (Fig. 3*B*). This is a chromatin signature of promoters actively transcribing protein-coding genes (32, 33) and was detected in human embryonic stem cells (H7/H1-hESC) and in different types of neuronal cells, such as a neuroblastoma cell line (SK-N-SH) and neurons derived from embryonic stem cells (H1-neurons) but not in HeLa cells (Fig. 3*B*). Moreover, binding events for RNA Polymerase II, TBP, and the enhancer-associated protein P300 (34, 35) are also detected in the *Prrxl1* promoter regions in H1-neurons, H1-hESC, and SK-N-SH cells but not in HeLa cells. Note that

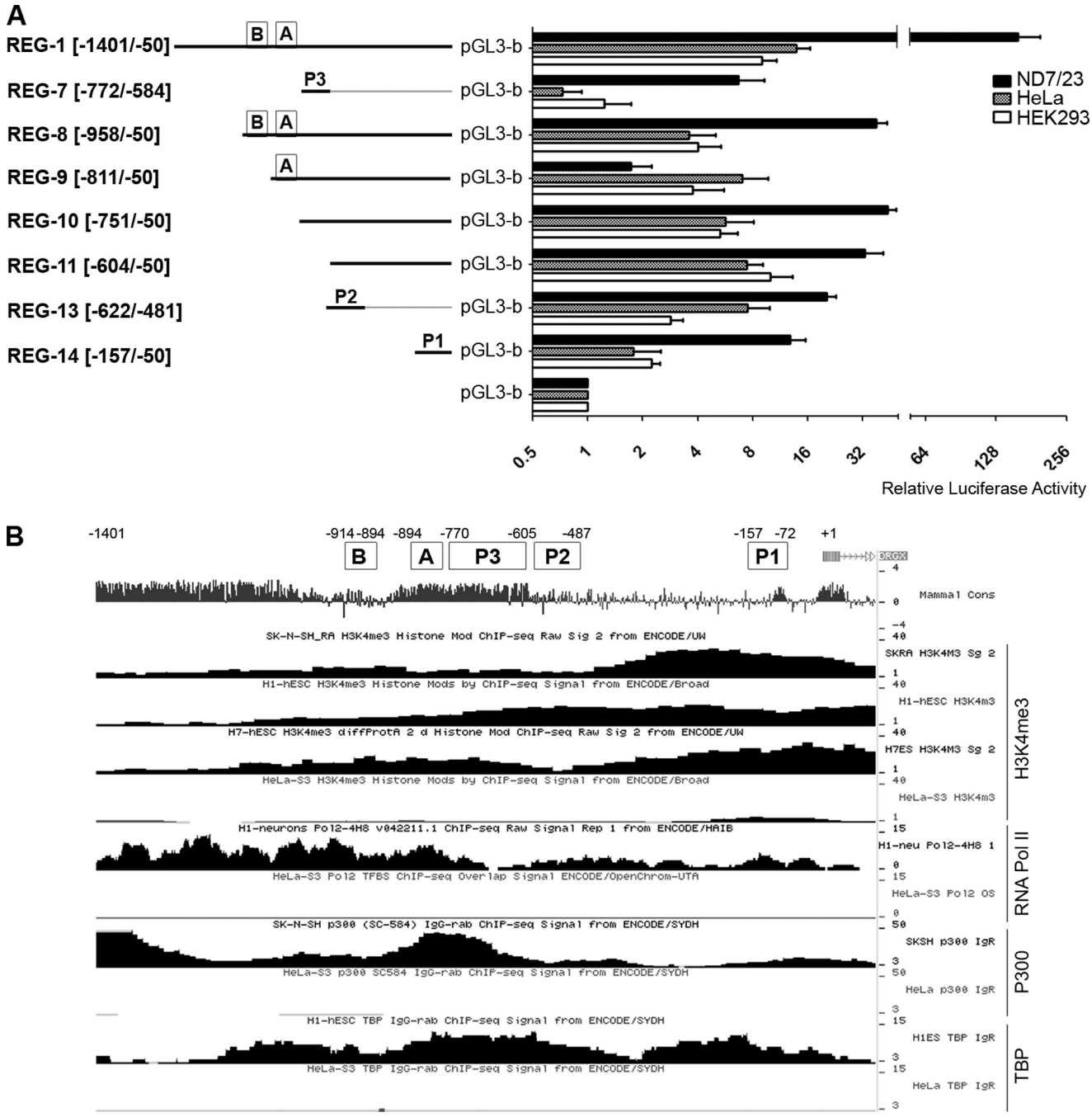

FIGURE 3. **Dependence on neuronal context for the transcriptional activity of *Prrxl1* alternative promoters.** *A*, the luciferase activity induced by some selected fragments was compared in ND7/23 and the non-neuronal HeLa and HEK293 cells. P1 and P3 only promote significant transcription in neuronal-derived cells, whereas P2 exhibits activity in the three cell lines. The repressive or activator transcriptional effect induced by the regulatory elements A and B in ND7/23 cells was no longer observed in HeLa and HEK293 cells. *B*, the presence of binding peaks of H3K4me3, RNA Polymerase II, P300, and TBP on the *Prrxl1* promoters was specifically detected in neuronal cells, using the UCSC genome Browser and ENCODE annotations of ChIP-seq assays.

the three-peak pattern observed with RNA Polymerase II and P300 overlaps with the three alternative promoter regions (Fig. 3*B*). These data reinforce our previous observation that *Prrxl1* promoters displayed neuron-specific activity.

The Promoter P3 Displays Neuron-specific Activity—Nucleotide alignment of *Prrxl1* alternative promoter sequences from different species revealed that the region comprising promoter P3 presents a high degree of conservation, from zebrafish to human, whereas promoters P1 and P2 appeared to be specific to mammals (Fig. 4). We screened these sequences to identify conserved DNA binding elements for transcription factors using the bioinformatics prediction tool MatInspector from Genomatix. Some putative motifs known to be important for transcriptional regulation during embryonic development were identified.

Promoter P1 has no evident conserved motifs, whereas promoter P2 contains a GC box element (Fig. 4), known to be a binding site for Sp1, Sp3, and Sp4 transcription factors (36). CpG islands are elements often associated with the transcription of genes whose expression is ubiquitous and feature >50%

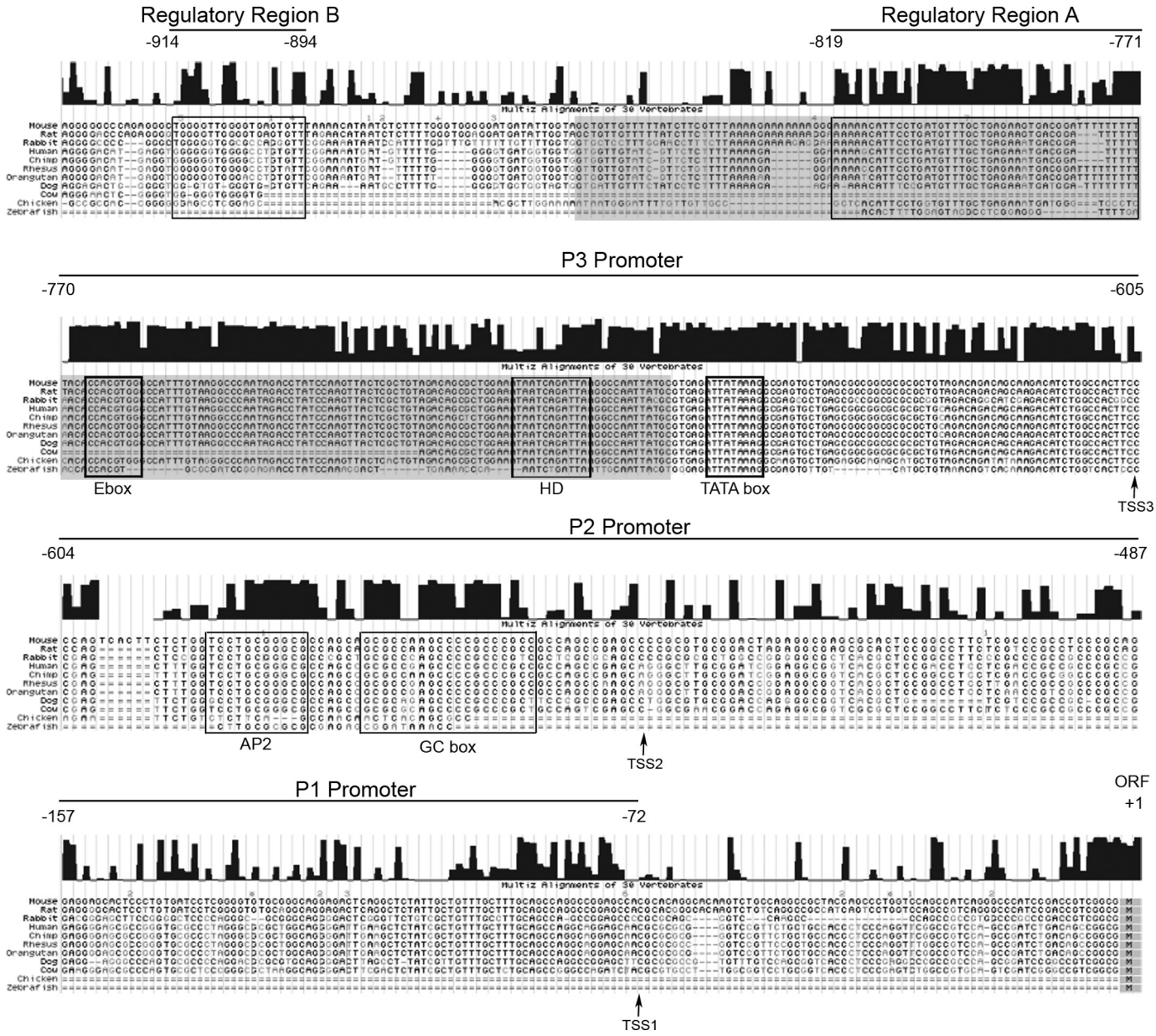

FIGURE 4. **Nucleotide alignment of the regulatory regions A and B and the *Prrxl1* alternative promoters.** The nucleotide alignment was performed using the UCSC Genome Browser and sequences from several mammals, chick, and zebrafish. The *boxes* that delimit the regulatory regions A and B are depicted. P3 is the most conserved promoter and contains binding sites for the TFIID complex (*TATA box*), bHLH (*Ebox*), and homeodomain transcription factors (*HD*). P2 and P1 are only present in mammals. P2 contains putative motifs for the binding of SP1-like family transcription factors (*GC box*) and for the activation protein 2 (*AP2*). The TSSs of each 5′-UTR are also represented. TSS1 corresponds to 5′-UTR-C, TSS2 to 5′-UTR-B, and TSS3 to 5′-UTR-A. Binding sites were predicted by the bioinformatics tool MatInspector (Genomatix). *Black peaks* represent mammal conservation. The *gray box* highlights the human sequence used to produce the zebrafish transgenic line (see Fig. 6). Not to scale.

of vertebrate regulatory sequences (37). The presence of this GC-box may explain the activity displayed by the promoter P2 in HeLa cells (Fig. 3*A*). Moreover, a conserved binding site for the activator protein 2 family was also predicted on the P2 promoter (Fig. 4). Activator protein 2 transcription factors are general regulators of vertebrate development, controlling the balance between proliferation and differentiation during embryogenesis (38, 39).

Analysis of the P3 sequence highlighted a putative TATA motif, a conserved element of some basal promoters (Fig. 4). To test if this sequence could bind a TBP, the binding of a recombinant TBP to an oligonucleotide spanning the P3 TATA motif present in the promoter P3 was tested by EMSA (Fig. 5*A*). A gel shift, which was impaired by the use of competitor oligonucleotides, was observed, demonstrating the specificity of the binding to this region (Fig. 5*A*). The incubation with an antibody directed to TBP resulted in a supershift (*arrow* in Fig. 5*A*), strongly indicating that P3 is a TATA box-containing promoter. The *in vivo* association of the TBP to the TATA motif of the P3 sequence was tested in mouse embryonic spinal cord by performing ChIP assays combined with real-time PCR (Fig. 5*B*). In these experiments a clear enrichment in the binding of the TBP was observed in the P3 region comprising the TATA motif compared with the

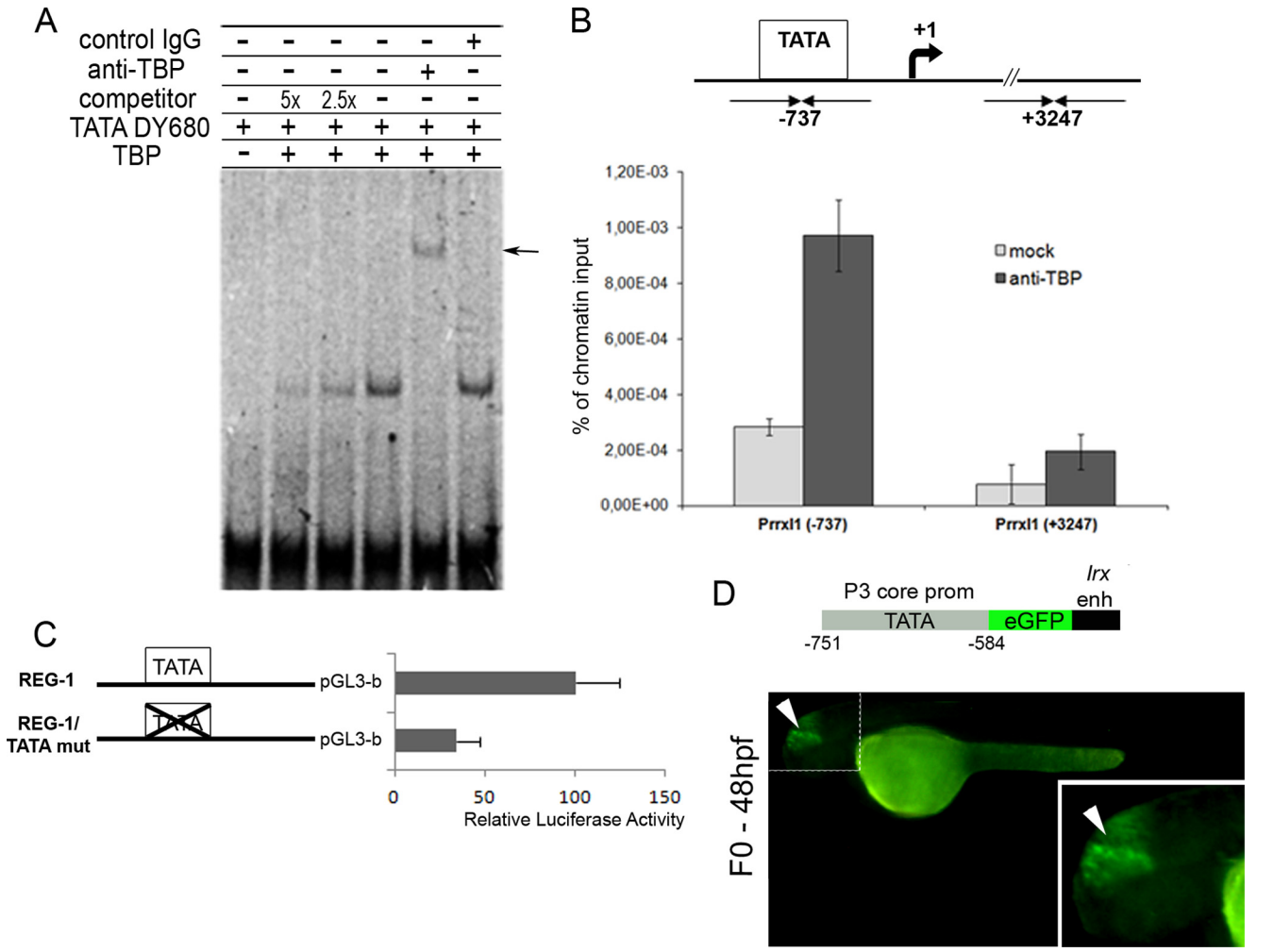

FIGURE 5. **Validation of P3 as a TATA-containing promoter.** *A*, electrophoretic mobility shift assay using recombinant TBP and a fluorescent probe (*TATA DY680*) corresponding to a region of P3 with the TATA box motif. A shift, corresponding to a complex formed between TBP and the probe, was detected, and decreases of intensity when different amounts (2.5× and 5×) of non-labeled competitor were mixed. The supershift (*arrow*) represents the binding of anti-TBP to the protein-DNA complex. *B*, chromatin immunoprecipitation-quantitative PCR assays with (*anti-TBP*) or without (*mock*) an anti-TBP antibody were performed using dorsal spinal cord chromatins from E14.5 mouse embryos followed by quantitative PCR using primers targeting P3 region (Prrxl1 −737) and a downstream region (Prrxl1 + 3247). Enrichment is observed in the region where the TATA box is present (−737). A region corresponding to *Prrxl1* intron was used as control (+3247). *Mock* represents the condition without antibody. *C*, the evaluation of the functional importance of the TATA box for the transcriptional potential of the fragment REG-1 was assessed by luciferase reporter assays in ND7/23 cells, comparing the REG-1 fragment with a sequence containing a mutated TATA motif (REG-1/TATA mut). *D*, to evaluate the ability of the P3 core promoter to activate transcription in zebrafish, a region encompassing nucleotides −584/−751 bp was cloned upstream to an *Irx* (*Iroquois*) enhancer, which drives expression to the zebrafish midbrain (*white arrowheads*). The vector was injected in zebrafish embryos at one-cell stage, and 48 hpf GFP signal was recorded.

downstream site (Fig. 5*B*, compare position −737 to + 3247) and the control ChIP (*mock*).

Moreover, to evaluate the importance of the TATA box in the transcriptional activity displayed by the fragment REG-1 in the ND7/23 cells, we compared the luciferase activity of this fragment with a fragment containing a mutated TATA motif (Fig. 5*C*). This mutation resulted in a 3-fold decrease of luciferase expression, indicating that this element is required for correct transcription. The remaining luciferase expression observed for the fragment REG-1/TATAmut is probably due to the activity of the alternative promoters P1 and P2, which was unaltered.

Taking into account that the P3 sequence is well conserved among species, we used the zebrafish model to test the potential of this TATA promoter to drive transcription *in vivo*. One- to two-cell- stage embryos were injected with a vector containing the enhanced-GFP (eGFP) reporter gene under the control of the P3 minimum region combined with a zebrafish *iroquois* enhancer that could direct the expression of eGFP to the midbrain (24). Twenty-four hours post fertilization (hpf), eGFP was already observed in the midbrain of the embryos (Fig. 5*D*), an expression that lasted at least until 72 hpf. This result indicated that the P3 region exhibits promoter activity *in vivo* in this transgenic zebrafish assay.

Adjacent to the TATA motif, the P3 region contains evolutionarily conserved motifs potentially bound by bHLH (Ebox motif) and homeodomain (HD motif) transcription factors (Fig. 4). Transcription factors belonging to these families play important roles in the determination of neuronal fates from undifferentiated progenitor cells (for review, see Refs. 5 and 40). To evaluate their functional relevance, a region of 172 bp containing these sites but excluding the TATA motif was cloned in a Tol2 vector, specifically designed to analyze cis-regulatory elements in zebrafish (23) and carrying the eGFP reporter gene

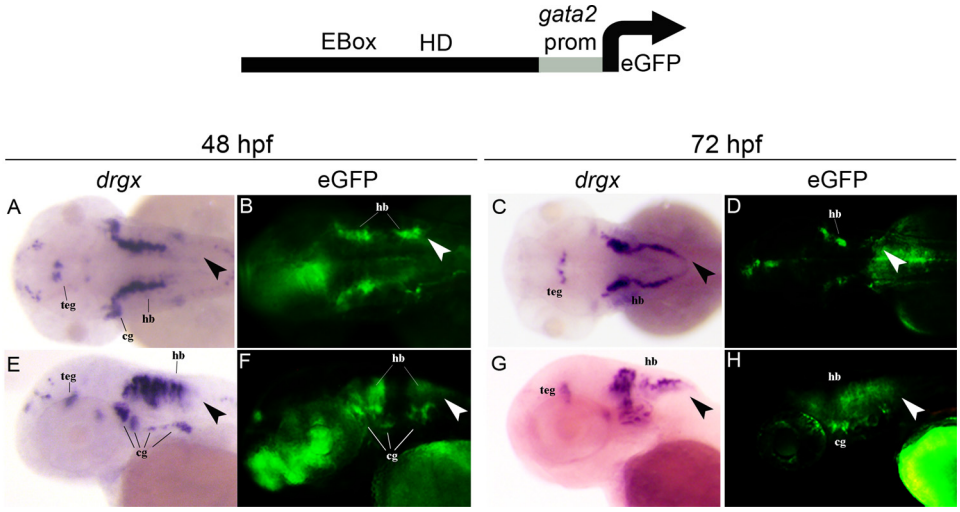

FIGURE 6. **Expression patterns of endogenous *drgx* and eGFP in the head region of a P3 promoter-eGFP zebrafish stable line.** *A*, *C*, *E*, and *G*, analysis of endogenous *drgx* expression by *in situ* hybridization. *drgx* expression is detected in embryos at 48 hpf in the anterior region of the hindbrain (*hb*), in the tegmentum (*teg*) (see dorsal view (*A*) and lateral view (*E*)), and in cranial ganglia (*cg*) (see *E*). At 72 hpf, *drgx* expression is more widespread in the hindbrain, and it still maintains its expression in the cranial ganglions and in the tegmentum (see dorsal view (*C*) and lateral view (G)). *B*, *D*, *F*, and *H*, analysis of eGFP expression in transgenic zebrafish embryos. eGFP expression is regulated by a module that encloses the *gata2* minimal promoter and a region of the *Prrxl1* P3 promoter that contains the Ebox and HD elements, excluding the TATA motif (see the *scheme*). In 48 hpf embryos, eGFP expression is distinguished in the hindbrain and in cranial ganglia (see dorsal view (*B*), lateral view (*F*)). The *arrowhead* indicates a region where eGFP is apparently activated prematurely in the transgenic embryos at 48 hpf (*B* and *F*) but whose *drgx* expression is only detected at 72 hpf (*C* and *G*). Although eGFP expression is reduced at 72 hpf, it is still detected in a small region of the anterior hindbrain and in the cranial ganglia (see dorsal view (*D*) and lateral view (*H*)).

under the control of the *gata2a* minimal promoter (Fig. 6). This vector was used to generate a zebrafish stable transgenic line. eGFP signal was recorded in F1 embryos at different developmental ages and compared with endogenous expression of *drgx*, the *Prrxl1* zebrafish homologue. The expression of *drgx* was previously reported (41) as restricted to dorsal spinal cord, DRG, and, in the developing brain, to sensory neuron populations of the midbrain and hindbrain, cranial sensory ganglia, and the habenula. From 48 hpf, a strong eGFP signal was detected in the developing hindbrain and cranial ganglia (Fig. 6, *B* and *F*) in a pattern that only partially overlaps the expression of endogenous *drgx*, as revealed by *in situ* hybridization (Fig. 6, *A* and *E*). Indeed, eGFP expression was prematurely observed in a posterior region of the hindbrain (marked by an *arrowhead* in Fig. 6, *B* and *F*) that only expresses *drgx* later at 72 hpf (compare the *arrowheads* in Fig. 6, *C* and *G*, with the *arrowheads* in Fig. 6, *B* and *F*). Due to the very small length of this fragment (172 bp), it is conceivable that repressive elements may be missing. Expression of eGFP in cranial ganglia displayed a spatiotemporal pattern that perfectly matched with *drgx* staining (figure 6*E*-H). At 72hpf, a decrease in the eGFP staining was observed (compare Fig. 6, *D* and *H*, with Fig. 6, *panels C* and *G*). Albeit the expression of *drgx* in DRG and spinal cord is well reported in zebrafish (41), the 172-bp sequence does not drive eGFP transcription to these tissues (data not shown). Together, these experiments pointed out that the region of 172 bp containing the conserved Ebox and HD elements was sufficient to drive expression of the reporter gene in *drgx*-expressing neurons.

Phox2b Controls Prrxl1 Expression by Binding the HD Motif in P3 Promoter—To unravel the trans-acting factors that may be responsible for the modulation of *Prrxl1* transcription via Ebox or HD element, we tested the overexpression effect of a set of transcription factors previously implicated in the control of Prrxl1 expression, as described in diverse epistatic studies employing mutant mice. In *Tlx3* and *Lmx1b* null mutant mice, Prrxl1 expression is affected in the spinal cord but not in the DRG (42, 43), whereas studies performed with DRG of *islet1* inducible conditional knock-out mice suggested that Prrxl1 expression is regulated by islet1 only at an early stage of neurogenesis (44). Moreover, Prrxl1 is activated by Brn3a in the DRG and trigeminal ganglion (45) and repressed by Phox2b in the facial, glossopharyngeal, and vagal cranial ganglia (7). Mash1 and Neurogenin1 (Ngn1) are proneuronal genes implicated in the specification of progenitors cells from which Prrxl1-expressing neurons arise either in the DRG or the neural tube (3, 46–48). Ngn1 and Mash1 belong to the bHLH transcription factor family and bind the Ebox consensus motif CANNTG (49), whereas the remaining proteins belong to homeodomain family and recognize the HD bipartite element TAATNNNATTA (50, 51). Thus, we assessed the expression of luciferase under the control of the REG-1 fragment in ND7/23 cells overexpressing Brn3a, Tlx3, Phox2b, Islet1, Ngn1, Lmx1b, and Mash1 (Fig. 7*A*). Three transcription factors strongly induced the transcriptional activity of the fragment REG-1. Those were Brn3a, Tlx3, and Phox2b. As homeodomain proteins, they are candidates to bind the HD element present in the P3 promoter. Therefore, we evaluated by luciferase reporter assays the effect of the overexpression of these transcription factors on the REG-1 construct containing the mutated HD motif compared with wild-type REG-1 (Fig. 7*B*). Tlx3 and Brn3a overexpression displayed the same luciferase activity in both constructs, whereas the induction of Phox2b decreased about 50% when the HD motif is mutated. To ascertain if Phox2b binds directly to the HD element present in *Prrxl1* P3 promoter, EMSA and ChIP-PCR were performed (Fig. 7, *C* and *D*). By EMSA, a DNA-protein shift was detected after the incubation of an oligonucleotide comprising the HD motif sequence and nuclear protein extracts from ND7/23 cells overexpressing Phox2b. On the

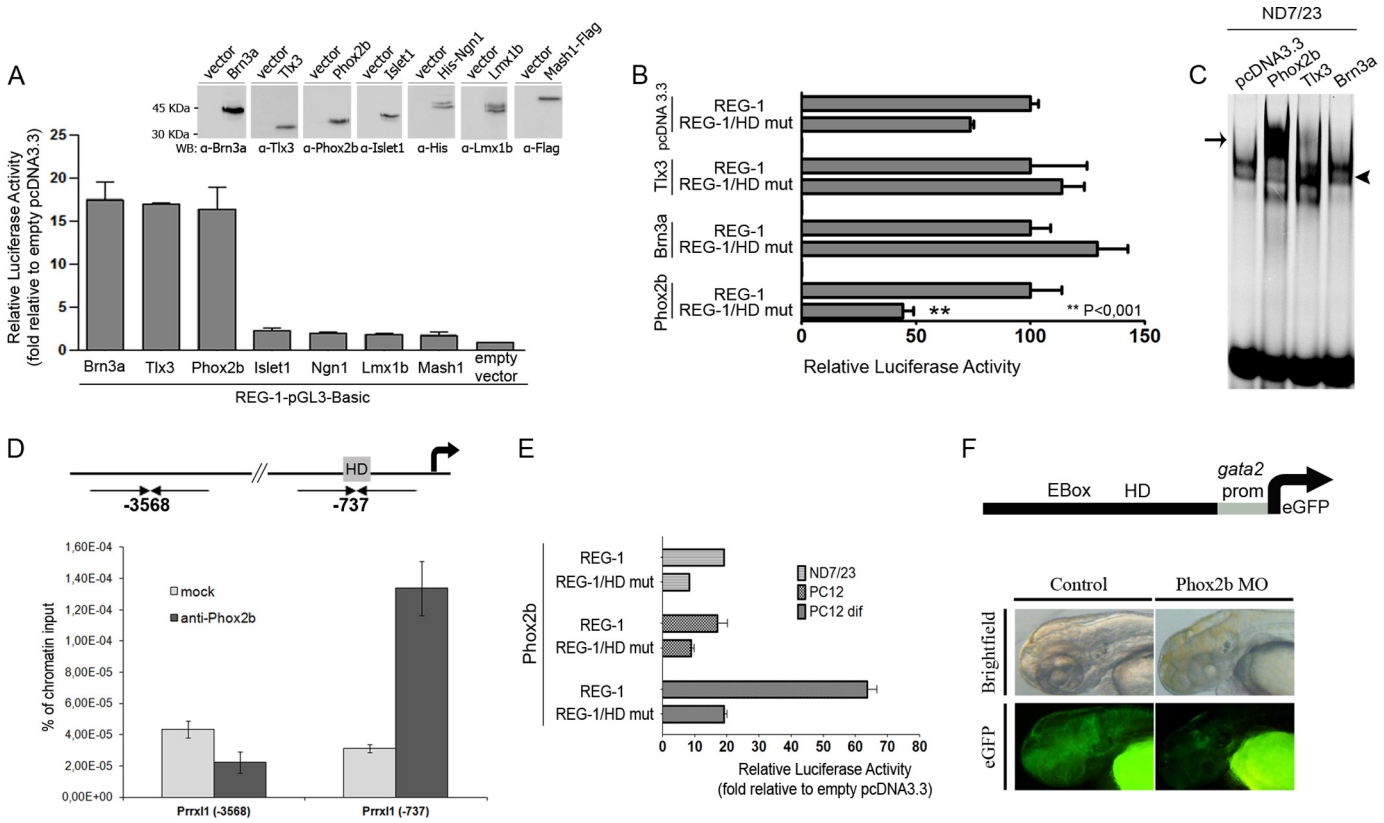

FIGURE 7. **Phox2b modulates Prrxl1 transcription by binding to the HD element in the P3 promoter.** *A*, the effect on REG-1 transcriptional activity by overexpressing the bHLH transcription factors Mash1 and Ngn1 and the homeodomain proteins Lmx1b, Islet1, Phox2b, Tlx3, and Brn3a in ND7/23 cells was assessed by luciferase reporter assays. The immunoblot confirms the protein expression and corresponds to one representative experiment. *B*, luciferase activity induced by REG-1 and REG-1/HD mut was measured in ND7/23 cells overexpressing Tlx3, Brn3a, and Phox2b or transfected with an empty pcDNA3.3. *C*, electrophoretic mobility shift assay using nuclear extracts from ND7/23 cells overexpressing Brn3a, Tlx3, or Phox2b or transfected with an empty pcDNA3.3 and a fluorescent probe (HD DY680) that comprises the HD motif present in the P3 promoter. The shift (*arrow*) is only observed in ND7/23 cells that overexpressed Phox2b. The *arrowhead* indicates a basal shift detected in all samples. *D*, chromatin immunoprecipitation assays with (*anti-Phox2b*) or without (*mock*) an anti-Phox2b antibody were performed using dorsal medulla oblongata chromatins from E14.5 mouse embryos followed by quantitative PCR using primers targeting regions, which comprises the HD motif (*Prrxl1 −737*) and an upstream control sequence (*Prrxl1 −3568*). An enrichment was only observed in the region where the HD element was present (*−737*). *E*, luciferase activity induced by REG-1 and REG-1/HD mut was measured in ND7/23, PC12, and differentiated PC12 cells overexpressing Phox2b. *F*, analysis of eGFP expression in transgenic zebrafish embryos at 72 hpf, containing the module that comprises the Phox2b-binding site (HD motif) present in the P3 promoter. Expression of eGFP was reduced in Phox2b MO-injected embryos when compared with controls. The fluorescence acquisition settings were exactly the same in both images.

contrary, overexpression of Tlx3 or Brn3a only led to a weak basal shift in the gel, which is also observed in extracts containing Phox2b overexpressed (marked by an *arrowhead* in Fig. 7*C*). We assume that this interaction represents a binding with another homeodomain protein that remains to be identified. The binding of Phox2b on *Prrxl1* promoter was further validated by ChIP-PCR assays (Fig. 7*D*). A chromatin enrichment (relative to the no antibody control, *mock*) was only observed with primers that amplified the region comprising the HD element (Fig. 7*D*, position −737).

Co-expression of Phox2b and Prrxl1 or its zebrafish orthologue, *drgx*, was only reported in some sensory ganglia, namely in the facial, glossopharyngeal, and vagal cranial ganglia, and their target relay neurons in the hindbrain (6, 7). As Phox2b is not detected in the ND7/23 cell line, we performed the same analysis using non-differentiated and differentiated PC12 cells, which endogenously expressed Phox2b and Prrxl1 (52). Again, overexpression of Phox2b increased the luciferase activity of REG-1 fragment, which is impaired by mutating the HD motif (Fig. 7*E*).

To better investigate the regulation of Prrxl1 expression by Phox2b *in vivo*, we used a validated specific antisense MO to interfere with the translation of zebrafish Phox2b protein. According to what was previously reported by Elworthy *et al.* (53), our zebrafish Phox2b morphants also presented scarce Hu-positive cells in the hindgut at 5 dpf (53%, 8 in 15 embryos). We injected Phox2b MO in the transgenic zebrafish line described above (Fig. 6), which contains eGFP under the control of the 172-bp sequence comprising the Phox2b binding element (HD motif). A decrease in the eGFP expression was observed (Fig. 7*F*), supporting our hypothesis that Phox2b binds to the HD element present in the P3 promoter region of *Prrxl1* and is required for this promoter activity.

Phox2b Is Required for drgx Expression in the Glossopharyngeal, Vagal, and Facial Ganglia—To further address the role of Phox2b in the control of *drgx* expression, we injected Phox2b MO in wild-type zebrafish embryos at one- to two-cell stage. Expression analysis of *drgx* by *in situ* hybridization showed that in 48 hpf morphant animals, *drgx* is still maintained in the hindbrain, in the tegmentum, and in the trigeminal ganglia (Fig. 8,

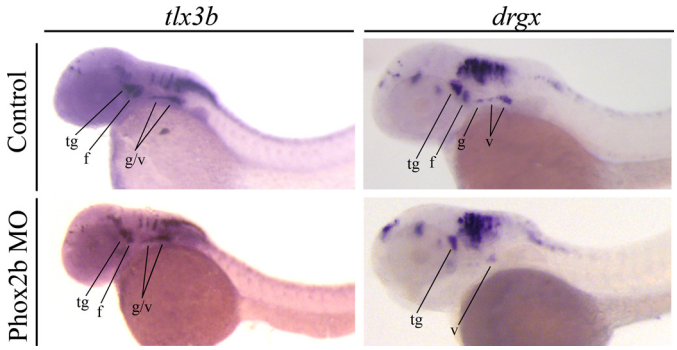

FIGURE 8. **Phox2b is required for the expression of *drgx* in the glossopharyngeal, vagal, and facial ganglia of zebrafish embryos.** Analysis of the *tlx3b* and *drgx* expression by *in situ* hybridization in control or Phox2b MO-injected embryos is shown. The *upper panels* show the control embryos, and the *lower panels* show representative images of morphant embryos. *tlx3b* and *drgx* expression are maintained in the hindbrain and trigeminal ganglion (*tg*) of both control and Phox2b MO-injected animals. Expression of *drgx* in the facial (*f*), glossopharyngeal (*g*), and vagal (*v*) ganglia is greatly reduced in Phox2b MO-injected animals, whereas *tlx3b* expression was still observed in all these ganglia.

compare *drgx* expression in control with Phox2b MO). Nevertheless, in 44% of embryos, *drgx* expression is much weaker or absent in the glossopharyngeal, vagal, and facial ganglia (Fig. 8), suggesting that, as predicted by *in vitro* assays, Phox2b is required for the transcriptional control of *drgx*.

To discard the hypothesis that the absence of *drgx* expression was due to a failure in the development of these ganglia induced by Phox2b knockdown, *in situ* hybridizations were performed using a *tlx3b* probe. Tlx3 expression has been shown to extensively overlap with Prrxl1 (10, 42) and thus was used here as a reliable marker of neuronal differentiation. As expected, *tlx3b* expression is maintained in Phox2b morphants (Fig. 8, *Phox2b MO tlx3b*).

DISCUSSION

Prrxl1 is a transcription factor with an important role in the establishment and maintenance of the nociceptive DRG-spinal cord neuronal circuit. The function of Prrxl1 in this process has been characterized in detail, but the molecular determinants causing its activation or repression are not yet understood. To shed some light on the molecular mechanisms regulating *Prrxl1* gene expression, we isolated and characterized the alternative promoters that control the expression of three distinct *Prrxl1* 5′-UTR variants, named 5′-UTR-A, 5′-UTR-B, and 5′-UTR-C. These variants are originated from different exon 1 during *Prrxl1* splicing and do not have any consequence in the protein reading frame as the AUG start codon is present in exon 2.

The alternative use of different exon 1 has been recognized as another mechanism of control of gene expression. Within the human and mouse genome, >3000 genes with multiple first exons have been identified (54). 5′-UTR-mediated regulation were shown to modulate gene expression through mechanisms that influence post transcriptional modification of RNA (secondary structure and mRNA stability) and translational efficiency (18). This proved here to be also the case for *Prrxl1* 5′-UTR variants. The 5′-UTR-A displayed a neuron-specific effect, increasing both the rate of Prrxl1 protein translation and the stability of the mRNA molecule. This observation implies a contribution of neuronal specific RNA-binding proteins. Taking into account that Prrxl1 has an important role in the development of the neuronal circuit connecting the DRG to the spinal dorsal horn (8, 9), possible candidates are Hu RNA-binding proteins (human homologues of *Drosophila* ELAV), namely the neuronal specific HuC and HuD isoforms, which are long used as markers of neuronal differentiation. Overexpression of HuC or HuD in PC12 cells increased the rate of neuronal differentiation, whereas down-regulation resulted in an impairment of neurite growth (55, 56). Likewise, a decrease in neuronal differentiation was observed in HuD null mutant mice (57). Although HuC/D proteins are strong candidates to modulate the neuron-specific activity of *Prrxl1* 5′-UTR-A, their involvement remains to be demonstrated.

On the contrary, the 5′-UTR-B reduced mRNA half-life and suppressed mRNA translation, resulting in a decrease in the luciferase expression both in the neuronal-derived ND7/23 and HeLa cells. This was due to the presence of an AUG in the 5′-UTR-B, as shown by luciferase reporter gene assays and site-directed mutagenesis. This AUG is used as an alternative start codon, modifying, therefore, the protein reading frame. AUG codons upstream of the main open reading frame are present in ~10% of all mRNAs (58). Although the functional impact of this mechanism has not been investigated in detail, a recent study on *male-specific lethal-2* mRNA suggested an important role in negative translational control by increasing initiation of scanning ribosomes at the upstream open reading frame and blocking downstream translation (58). The cis-regulatory upstream ORF present in the *Prrxl1* 5′ UTR-B exerts a negative influence on translational efficacy, as seen in Fig. 1*C*, and is probably responsible for controlling the amount of Prrxl1 protein during early neurogenesis, the stage where 5′-UTR-B is most expressed.

By studying deletion derivatives of the *Prrxl1* 5′-flanking region, we demonstrated that transcription of each *Prrxl1* 5′-UTR variants is controlled by specific promoters. The here-named promoter P1, P2, and P3 controlled, respectively, the expression of 5′-UTR-C, 5′-UTR-B, and 5′-UTR-A. Recent genome wide analyses indicated that alternative promoter usage is a common event that occurs at least with the same order of frequency as alternative splicing, affecting about 52% of human genes (37). On average, there are 3.1 alternative promoters per gene, with the composition of one CpG-island-containing promoter per 2.6 CpG-less promoters. Moreover, it was demonstrated that genes that undergo complex transcriptional regulation often include at least one CpG island-containing promoter, expressed ubiquitously, and are accompanied by other promoters used for tissue-specific or signal-dependent expression (37). Our results on the *Prrxl1* promoters are in accordance with these observations. Indeed, a GC-rich region is contained within promoter P2, whereas P3 is a TATA promoter. P2 displayed activity in all cell lines tested, whereas P3 and P1 were regulated depending on the cellular context being only active in neurons (for a summary of these findings see the model in Fig. 9).

Among the three *Prrxl1* alternative promoters, P3 is the most conserved and displays *in vivo* activity in the zebrafish. Reverse

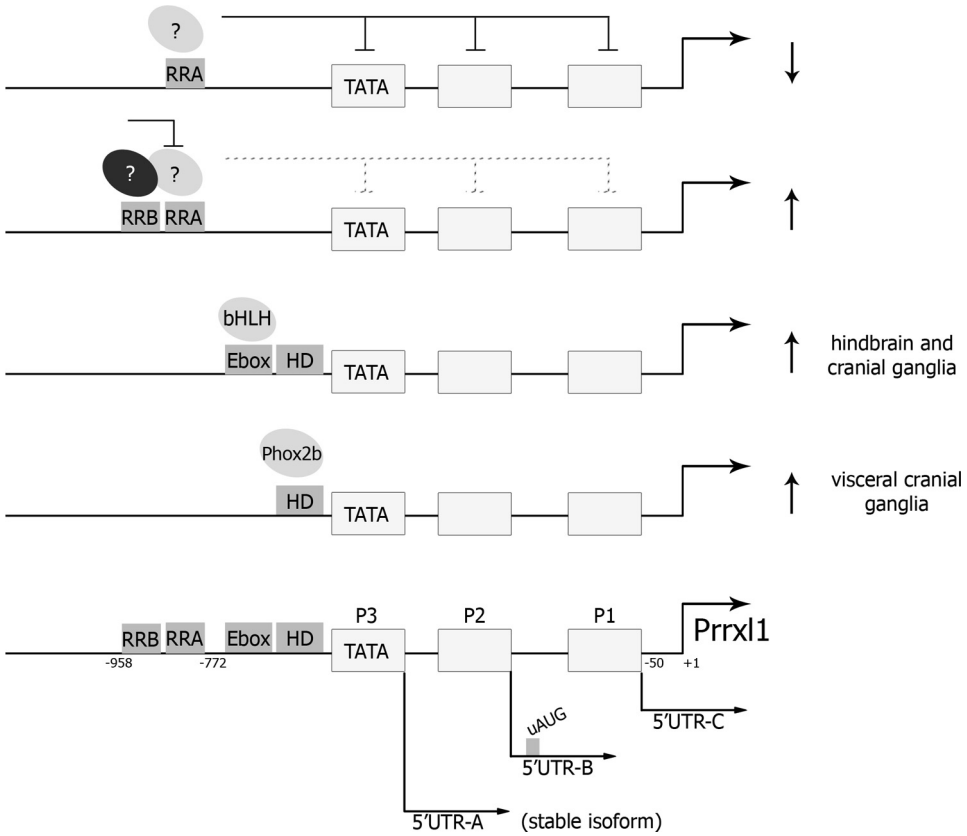

FIGURE 9. **Schematic representation of the regulatory elements that control the expression of *Prrxl1* 5′-UTR variants.** Promoter regions and other regulatory elements were identified upstream of *Prrxl1* translation start site (+1). Transcription of the three *Prrxl1* mRNA variants 5′-UTR-C, 5′-UTR-B, and 5′-UTR-A is controlled by distinct promoters, termed *P1*, *P2*, and *P3*, respectively. 5′-UTR-A is the most stable *Prrxl1* transcript and is enriched at E14.5 spinal cord, a developmental age associated with late neurogenesis. Located between −958 and −772 bp, two preponderant elements were identified: RRA is able to totally suppress the transcriptional activity regulated by the three promoters, and RRB inhibits the repressive effect of RRA resulting in an increase of transcription. In addition, the Ebox and HD motifs located upstream of the TATA box drive the transcription of *Prrxl1* to hindbrain and cranial ganglia. Although the bHLH transcription factor that binds the Ebox motif remains to be identified, the homeodomain protein Phox2b binds the HD element and is required for Prrxl1 expression in visceral cranial ganglia. *uAUG*, AUG codons upstream.

transcriptase-PCR experiments also showed that the P3-derived transcript, *Prrxl1* 5′-UTR-A, is more abundant than the other two variants. These two observations led us to conclude that promoter P3 may play a more prominent role in *Prrxl1* transcriptional regulation and prompted us to focus our study on transcriptional mechanisms modulating P3 activity. Core promoters were shown to comprise DNA sequence motifs, such as the TATA box, the Initiator (Inr), and the downstream promoter element, located within −30 to +30 nucleotides relative to the TSS and to mediate the recognition and recruitment of the RNA polymerase II to the transcriptional apparatus (59, 60). *Prrxl1* P3 core promoter contained a TATA motif located at 35 bp upstream of 5′-UTR-A TSS. The functional validation of this motif was attained by *in vitro* and *in vivo* evidence, such as (i) TBP interacted with the TATA motif as detected by EMSA and ChIP using mouse spinal cord samples, (ii) mutation of the TATA motif greatly reduced promoter activity, and (iii) P3 core promoter, which contains the TATA element, is sufficient to drive the reporter gene expression in the zebrafish.

Located upstream of the P3 promoter, a highly conserved region (−1401/−958) was identified as an important modulator of *Prrxl1* transcription. This region contains binding sites for actively transcribing protein such as RNA Polymerase II, TBP, and P300. By sequence deletion experiments, the presence of still uncharacterized regulatory modules was predicted. These elements appeared to be mainly associated with P1 and P2 promoters and thereby to the transcription of 5′-UTR-B and 5′-UTR-C. Because these Prrxl1 mRNA variants are enriched in early-born neurons of the developing spinal cord, we hypothesize that the modules included in the −1401/−958 bp region may be responsible for enhancing temporal-specific transcription controlled by promoters P1 and/or P2 promoters.

We also defined two neuron-specific elements, RRA (−811/−772) and RRB (−891/−922), which exhibit the opposite effect on *Prrxl1* transcription. RRA has the potential to strongly suppress the transcription of the three alternative promoters, whereas RRB counteracts the action of the RRA repressive motif and consequently induces *Prrxl1* transcription (see the model in Fig. 9). Such tight regulation could be understood as a way to modulate Prrxl1 expression in different neuronal types, namely in glutamatergic (Prrxl1-positive) over GABAergic (Prrxl1-negative) neurons during the developing spinal cord. Because these two neuronal populations derived from the same progenitor domain and are mutually exclusive, molecular interrepression between inhibitory and excitatory interneurons has been suggested. For instance, Tlx3 acts as an inhibitor of Lbx1 expression inducing the glutamatergic transmitter phenotype (61), whereas Ptf1a represses Tlx3 specifying a GABAergic cell

fate (62). A similar mechanism could be envisaged for Prrxl1. In GABAergic neurons, Prrxl1 expression could be prevented by a specific transcription factor acting on the RRA element. This repressive effect could be suppressed in glutamatergic neurons through the RRB element. It would be interesting to assess if one possible RRB binding candidate could be Tlx3 as this transcription factor induces *Prrxl1* promoter activity (Fig. 7*A*) and highly co-localizes with Prrxl1 (10).

Furthermore, the 172-bp 5′ region adjacent to the P3 core promoter was sufficient to drive specific neuronal activity in zebrafish. GFP expression under the control of putative regulatory elements present in this 172-bp region was only observed in the area in which the transcript of the zebrafish *Prrxl1* ortologue, *drgx*, is detected. Such specificity suggested that important cis-regulatory modules are present in this region, and further analysis of this sequence revealed the presence of a highly conserved putative binding site for homeodomain transcription factors, proteins that are expected to control the expression of Prrxl1 as a function of developmental age and neuronal context. Among the tested candidates, Tlx3, Brn3a, and Phox2b induced *Prrxl1* promoter activity, but only Phox2b was able to bind to this HD motif. Co-expression of Phox2b and Prrxl1 or its zebrafish ortologue, *drgx,* was only detected in the visceral sensory pathway, namely in the facial, glossopharyngeal, and vagal cranial ganglia and their target relay neurons in the hindbrain (6, 7, 63).We, therefore, presume that the P3 activity observed in the transgenic line (Fig. 6) may be controlled, at least to some extent, by Phox2b. This positive regulation is further supported by silencing experiments with Phox2b morphants (Fig. 7*F*).

Nonetheless, in the mouse, visceral sensory neurons switch to a somatic fate in the absence of Phox2b acquiring a molecular profile similar to that of somatic sensory neurons at later developmental stages, with higher expression of Prrxl1 and Brn3a (7). This increase in Prrxl1 expression was suggested to be mediated by the increase of Brn3a, which, in turn, is directly repressed by Phox2b (7). On the contrary, Prrxl1 expression in the nuclear TS progenitor domains of the hindbrain is not affected by Phox2b inactivation, suggesting that the repression of Prrxl1 by Phox2b only occurs in the visceral sensory ganglia (7). In accordance, Phox2b is not required for *drgx* expression in the zebrafish hindbrain. However, contrary to the mouse, Phox2b down-regulation induced a loss of *drgx* expression in the facial, glossopharyngeal, and vagal cranial ganglia. Binding of Phox2b on the HD motif increased *Prrxl1* promoter activity in ND7/23 and PC12 cells. Our data strongly suggest that Phox2b has the potential to work as a direct Prrxl1 activator (see model in Fig. 9).

Although Phox2b is a determinant of visceral fate, its mode of action varies with neuronal types (64, 65) likely due to cell type-specific combinations of transcription factor complexes. On the other hand, Prrxl1 is transiently expressed at early stages of the development of the facial-glossopharyngeal ganglion and the distal part of the vagal ganglion, as observed in mice (7) and zebrafish (41). During this particular developmental window, co-expression of Prrxl1, Phox2b, Tlx3, and Islet1 is observed (7). Recently, studies performed in Islet1 inducible conditional knock-out mice (44) suggested that Prrxl1 expression in the DRG is regulated by Islet1 only at an early stage of neurogenesis. Thus, it is conceivable that Phox2b, in combination with Islet1, binds to the HD motif on the P3 *Prrxl1* alternative promoter to activate *Prrxl1* expression at early stages of sensory ganglia development, whereas later (from E13.5 on), as part of a distinct transcriptional machinery, Phox2b shuts down Prrxl1 expression likely through the repression of Brn3a. Understanding how Brn3a acts on *Prrxl1* promoters will add new insight on this mechanism. The present data thus support that the P3 alternative promoter is involved in the ganglion specific action of Prrxl1 (see the model in Fig. 9), which appear to be controlled by Phox2b in the case of visceral sensory neurons.

Acknowledgments—We thank Diogo S. Castro (Instituto Gulbenkian de Ciência, Portugal) for the transcription factor binding sites analysis prediction and Solangel Rivero-Gil (CABD, Seville) for all the help with care and maintenance of zebrafish transgenic lines.

REFERENCES

1. Goulding, M., Lanuza, G., Sapir, T., and Narayan, S. (2002) The formation of sensorimotor circuits. *Curr. Opin Neurobiol.* **12,** 508–515
2. McGlone, F., and Reilly, D. (2010) The cutaneous sensory system. *Neurosci. Biobehav. Rev.* **34,** 148–159
3. Gowan, K., Helms, A. W., Hunsaker, T. L., Collisson, T., Ebert, P. J., Odom, R., and Johnson, J. E. (2001) Cross-inhibitory activities of Ngn1 and Math1 allow specification of distinct dorsal interneurons. *Neuron* **31,** 219–232
4. Caspary, T., and Anderson, K. V. (2003) Patterning cell types in the dorsal spinal cord. What the mouse mutants say. *Nat. Rev. Neurosci.* **4,** 289–297
5. Helms, A. W., and Johnson, J. E. (2003) Specification of dorsal spinal cord interneurons. *Curr. Opin. Neurobiol.* **13,** 42–49
6. Rebelo, S., Reguenga, C., Osório, L., Pereira, C., Lopes, C., and Lima, D. (2007) DRG11 immunohistochemical expression during embryonic development in the mouse. *Dev. Dyn.* **236,** 2653–2660
7. D'Autréaux, F., Coppola, E., Hirsch, M. R., Birchmeier, C., and Brunet, J. F. (2011) Homeoprotein Phox2b commands a somatic-to-visceral switch in cranial sensory pathways. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 20018–20023
8. Chen, Z. F., Rebelo, S., White, F., Malmberg, A. B., Baba, H., Lima, D., Woolf, C. J., Basbaum, A. I., and Anderson, D. J. (2001) The paired homeodomain protein DRG11 is required for the projection of cutaneous sensory afferent fibers to the dorsal spinal cord. *Neuron* **31,** 59–73
9. Rebelo, S., Chen, Z. F., Anderson, D. J., and Lima, D. (2006) Involvement of DRG11 in the development of the primary afferent nociceptive system. *Mol. Cell. Neurosci.* **33,** 236–246
10. Rebelo, S., Reguenga, C., Lopes, C., and Lima, D. (2010) Prrxl1 is required for the generation of a subset of nociceptive glutamatergic superficial spinal dorsal horn neurons. *Dev. Dyn* **239,** 1684–1694
11. Rebelo, S., Lopes, C., Lima, D., and Reguenga, C. (2009) Expression of a Prrxl1 alternative splice variant during the development of the mouse nociceptive system. *Int. J. Dev. Biol.* **53,** 1089–1095
12. Ayoubi, T. A., and Van De Ven, W. J. (1996) Regulation of gene expression by alternative promoters. *FASEB J.* **10,** 453–460
13. Grandien, K., Berkenstam, A., and Gustafsson, J. A. (1997) The estrogen receptor gene. Promoter organization and expression. *Int. J. Biochem. Cell Biol.* **29,** 1343–1369
14. Baek, D., Davis, C., Ewing, B., Gordon, D., and Green, P. (2007) Characterization and predictive discovery of evolutionarily conserved mammalian alternative promoters. *Genome Res.* **17,** 145–155
15. Nicholas, S. B., Yang, W., Lee, S. L., Zhu, H., Philipson, K. D., and Lytton, J. (1998) Alternative promoters and cardiac muscle cell-specific expression of the Na^+/Ca^{2+} exchanger gene. *Am. J. Physiol.* **274,** H217–H232
16. Kim, J. D., Kim, C. H., and Kwon, B. S. (2011) Regulation of mouse 4–1BB expression. Multiple promoter usages and a splice variant. *Mol. Cells* **31,** 141–149
17. Banday, A. R., Azim, S., and Tabish, M. (2011) Alternative promoter usage

and differential expression of multiple transcripts of mouse Prkar1a gene. *Mol. Cell Biochem.* **357,** 263–274
18. Wang, G., Guo, X., and Floros, J. (2005) Differences in the translation efficiency and mRNA stability mediated by 5'-UTR splice variants of human SP-A1 and SP-A2 genes. *Am. J. Physiol. Lung Cell Mol. Physiol.* **289,** L497–L508
19. Gauss, K. A., Bunger, P. L., Crawford, M. A., McDermott, B. E., Swearingen, R., Nelson-Overton, L. K., Siemsen, D. W., Kobayashi, S. D., Deleo, F. R., and Quinn, M. T. (2006) Variants of the 5'-untranslated region of human NCF2. Expression and translational efficiency. *Gene* **366,** 169–179
20. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203,** 253–310
21. Riquet, F. B., Tan, L., Choy, B. K., Osaki, M., Karsenty, G., Osborne, T. F., Auron, P. E., and Goldring, M. B. (2001) YY1 is a positive regulator of transcription of the Col1a1 gene. *J. Biol. Chem.* **276,** 38665–38672
22. Kawakami, K. (2007) Tol2. A versatile gene transfer vector in vertebrates. *Genome Biol.* **8,** S7
23. Bessa, J., Tena, J. J., de la Calle-Mustienes, E., Fernández-Miñán, A., Naranjo, S., Fernández, A., Montoliu, L., Akalin, A., Lenhard, B., Casares, F., and Gómez-Skarmeta, J. L. (2009) Zebrafish enhancer detection (ZED) vector. A new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev. Dyn.* **238,** 2409–2417
24. de la Calle-Mustienes, E., Feijóo, C. G., Manzanares, M., Tena, J. J., Rodríguez-Seguel, E., Letizia, A., Allende, M. L., and Gómez-Skarmeta, J. L. (2005) A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. *Genome Res.* **15,** 1061–1072
25. Bessa, J., Tavares, M. J., Santos, J., Kikuta, H., Laplante, M., Becker, T. S., Gómez-Skarmeta, J. L., and Casares, F. (2008) meis1 regulates cyclin D1 and c-myc expression, and controls the proliferation of the multipotent cells in the early developing zebrafish eye. *Development* **135,** 799–803
26. Balciunas, D., Wangensteen, K. J., Wilber, A., Bell, J., Geurts, A., Sivasubbu, S., Wang, X., Hackett, P. B., Largaespada, D. A., McIvor, R. S., and Ekker, S. C. (2006) Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet.* **2,** e169
27. Saito, T., Greenwood, A., Sun, Q., and Anderson, D. J. (1995) Identification by differential RT-PCR of a novel paired homeodomain protein specifically expressed in sensory neurons and a subset of their CNS targets. *Mol. Cell. Neurosci.* **6,** 280–292
28. Wood, J. N., Bevan, S. J., Coote, P. R., Dunn, P. M., Harmar, A., Hogan, P., Latchman, D. S., Morrison, C., Rougon, G., and Theveniau, M. (1990) Novel cell lines display properties of nociceptive sensory neurons. *Proc. Biol. Sci.* **241,** 187–194
29. Rogozin, I. B., Kochetov, A. V., Kondrashov, F. A., Koonin, E. V., and Milanesi, L. (2001) Presence of ATG triplets in 5' untranslated regions of eukaryotic cDNAs correlates with a "weak" context of the start codon. *Bioinformatics* **17,** 890–900
30. Geballe, A. P., and Morris, D. R. (1994) Initiation codons within 5'-leaders of mRNAs as regulators of translation. *Trends Biochem. Sci.* **19,** 159–164
31. ENCODE Project Consortium (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* **9,** e1001046
32. Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D., and Ren, B. (2005) A high-resolution map of active promoters in the human genome. *Nature* **436,** 876–880
33. Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., Wang, W., Weng, Z., Green, R. D., Crawford, G. E., and Ren, B. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39,** 311–318
34. Cho, H., Orphanides, G., Sun, X., Yang, X. J., Ogryzko, V., Lees, E., Nakatani, Y., and Reinberg, D. (1998) A human RNA polymerase II complex containing factors that modify chromatin structure. *Mol. Cell. Biol.* **18,** 5355–5363
35. Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Ren, B., Rubin, E. M., and Pennacchio, L. A. (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457,** 854–858
36. Hagen, G., Müller, S., Beato, M., and Suske, G. (1992) Cloning by recognition site screening of two novel GT box binding proteins. A family of Sp1 related genes. *Nucleic Acids Res.* **20,** 5519–5525
37. Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S. (2006) Diversification of transcriptional modulation. Large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res.* **16,** 55–65
38. Knight, R. D., Javidan, Y., Zhang, T., Nelson, S., and Schilling, T. F. (2005) AP2-dependent signals from the ectoderm regulate craniofacial development in the zebrafish embryo. *Development* **132,** 3127–3138
39. Mitchell, P. J., Timmons, P. M., Hébert, J. M., Rigby, P. W., and Tjian, R. (1991) Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes Dev.* **5,** 105–119
40. Guillemot, F. (1999) Vertebrate bHLH genes and the determination of neuronal fates. *Exp. Cell Res.* **253,** 357–364
41. McCormick, L. J., Hutt, J. A., Hazan, J., Houart, C., and Cohen, J. (2007) The homeodomain transcription factor drg11 is expressed in primary sensory neurons and their putative CNS targets during embryonic development of the zebrafish. *Gene Expr. Patterns* **7,** 289–296
42. Qian, Y., Shirasawa, S., Chen, C. L., Cheng, L., and Ma, Q. (2002) Proper development of relay somatic sensory neurons and D2/D4 interneurons requires homeobox genes Rnx/Tlx-3 and Tlx-1. *Genes Dev.* **16,** 1220–1233
43. Ding, Y. Q., Yin, J., Kania, A., Zhao, Z. Q., Johnson, R. L., and Chen, Z. F. (2004) Lmx1b controls the differentiation and migration of the superficial dorsal horn neurons of the spinal cord. *Development* **131,** 3693–3703
44. Sun, Y., Dykes, I. M., Liang, X., Eng, S. R., Evans, S. M., and Turner, E. E. (2008) A central role for Islet1 in sensory neuron development linking sensory and spinal gene regulatory programs. *Nat. Neurosci.* **11,** 1283–1293
45. Dykes, I. M., Tempest, L., Lee, S. I., and Turner, E. E. (2011) Brn3a and Islet1 act epistatically to regulate the gene expression program of sensory differentiation. *J. Neurosci.* **31,** 9789–9799
46. Ma, Q., Fode, C., Guillemot, F., and Anderson, D. J. (1999) Neurogenin1 and neurogenin2 control two distinct waves of neurogenesis in developing dorsal root ganglia. *Genes Dev.* **13,** 1717–1728
47. Wildner, H., Müller, T., Cho, S. H., Bröhl, D., Cepko, C. L., Guillemot, F., and Birchmeier, C. (2006) dILA neurons in the dorsal spinal cord are the product of terminal and non-terminal asymmetric progenitor cell divisions and require Mash1 for their development. *Development* **133,** 2105–2113
48. Ma, Q., Sommer, L., Cserjesi, P., and Anderson, D. J. (1997) Mash1 and neurogenin1 expression patterns define complementary domains of neuroepithelium in the developing CNS and are correlated with regions expressing notch ligands. *J. Neurosci.* **17,** 3644–3652
49. Johnson, J. E., Birren, S. J., Saito, T., and Anderson, D. J. (1992) DNA binding and transcriptional regulatory activity of mammalian achaete-scute homologous (MASH) proteins revealed by interaction with a muscle-specific enhancer. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 3596–3600
50. Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., and Desplan, C. (1993) Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev.* **7,** 2120–2134
51. Wilson, D. S., Guenther, B., Desplan, C., and Kuriyan, J. (1995) High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA. *Cell* **82,** 709–719
52. Yang, C., Kim, H. S., Seo, H., Kim, C. H., Brunet, J. F., and Kim, K. S. (1998) Paired-like homeodomain proteins, Phox2a and Phox2b, are responsible for noradrenergic cell-specific transcription of the dopamine β-hydroxylase gene. *J. Neurochem.* **71,** 1813–1826
53. Elworthy, S., Pinto, J. P., Pettifer, A., Cancela, M. L., and Kelsh, R. N. (2005) Phox2b function in the enteric nervous system is conserved in zebrafish and is sox10-dependent. *Mech. Dev.* **122,** 659–669
54. Zhang, T., Haws, P., and Wu, Q. (2004) Multiple variable first exons. A

mechanism for cell- and tissue-specific gene regulation. *Genome Res.* **14,** 79–89

55. Akamatsu, W., Okano, H. J., Osumi, N., Inoue, T., Nakamura, S., Sakakibara, S., Miura, M., Matsuo, N., Darnell, R. B., and Okano, H. (1999) Mammalian ELAV-like neuronal RNA-binding proteins HuB and HuC promote neuronal development in both the central and the peripheral nervous systems. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 9885–9890
56. Mobarak, C. D., Anderson, K. D., Morin, M., Beckel-Mitchener, A., Rogers, S. L., Furneaux, H., King, P., and Perrone-Bizzozero, N. I. (2000) The RNA-binding protein HuD is required for GAP-43 mRNA stability, GAP-43 gene expression, and PKC-dependent neurite outgrowth in PC12 cells. *Mol. Biol. Cell* **11,** 3191–3203
57. Akamatsu, W., Fujihara, H., Mitsuhashi, T., Yano, M., Shibata, S., Hayakawa, Y., Okano, H. J., Sakakibara, S., Takano, H., Takano, T., Takahashi, T., Noda, T., and Okano, H. (2005) The RNA-binding protein HuD regulates neuronal cell identity and maturation. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 4625–4630
58. Medenbach, J., Seiler, M., and Hentze, M. W. (2011) Translational control via protein-regulated upstream open reading frames. *Cell* **145,** 902–913
59. Ponjavic, J., Lenhard, B., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., and Sandelin, A. (2006) Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. *Genome Biol.* **7,** R78
60. Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A., Taylor, M. S., Engström, P. G., Frith, M. C., Forrest, A. R., Alkema, W. B., Tan, S. L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S. M., Wells, C. A., Orlando, V., Wahlestedt, C., Liu, E. T., Harbers, M., Kawai, J., Bajic, V. B., Hume, D. A., and Hayashizaki, Y. (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nat. Genet.* **38,** 626–635
61. Cheng, L., Samad, O. A., Xu, Y., Mizuguchi, R., Luo, P., Shirasawa, S., Goulding, M., and Ma, Q. (2005) Lbx1 and Tlx3 are opposing switches in determining GABAergic *versus* glutamatergic transmitter phenotypes. *Nat. Neurosci.* **8,** 1510–1515
62. Glasgow, S. M., Henke, R. M., Macdonald, R. J., Wright, C. V., and Johnson, J. E. (2005) Ptf1a determines GABAergic over glutamatergic neuronal cell fate in the spinal cord dorsal horn. *Development* **132,** 5461–5469
63. Pattyn, A., Morin, X., Cremer, H., Goridis, C., and Brunet, J. F. (1997) Expression and interactions of the two closely related homeobox genes Phox2a and Phox2b during neurogenesis. *Development* **124,** 4065–4075
64. Pattyn, A., Morin, X., Cremer, H., Goridis, C., and Brunet, J. F. (1999) The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives. *Nature* **399,** 366–370
65. Pattyn, A., Vallstedt, A., Dias, J. M., Samad, O. A., Krumlauf, R., Rijli, F. M., Brunet, J. F., and Ericson, J. (2003) Coordinated temporal and spatial control of motor neuron and serotonergic neuron generation from a common pool of CNS progenitors. *Genes Dev.* **17,** 729–737