The Homeoprotein Alx3 Contains Discrete Functional Domains and Exhibits Cell-specific and Selective Monomeric Binding and Transactivation*

Received for publication, January 23, 2004, and in revised form, June 24, 2004
Published, JBC Papers in Press, June 28, 2004, DOI 10.1074/jbc.M400800200

Beatriz Pérez-Villamil‡, Mercedes Mirasierra§, and Mario Vallecjo§**

From the ‡Reproductive Endocrine Unit, Massachusetts General Hospital, Boston, Massachusetts 02114 and §Instituto de Investigaciones Biomédicas “Alberto Sols,” Consejo Superior de Investigaciones Científicas/Universidad Autónoma de Madrid, Calle Arturo Duperier 4, 28029 Madrid, Spain

Alx3 is a paired class aristaless-like homeoprotein expressed during embryonic development. Transcriptional transactivation by aristaless-like proteins has been associated with cooperative dimerization upon binding to artificially generated DNA consensus sequences known as P3 sites, but natural target sites in genes regulated by Alx3 are unknown. We report the cloning of a cDNA encoding the rat homolog of Alx3, and we characterize the protein domains that are important for transactivation, dimerization, and binding to DNA. Two proline-rich domains located amino-terminal to the homeodomain (Pro1 and Pro2) are necessary for Alx3-dependent transactivation, whereas another one (Pro3) located in the carboxyl terminus is dispensable but contributes to enhance the magnitude of the response. We confirmed that transcriptional activity of Alx3 from a P3 site correlates with cooperative dimerization upon binding to DNA. However, Alx3 was found to bind selectively to non-P3-related TAAT-containing sites present in the promoter of the somatostatin gene in a specific manner that depends on the nuclear protein environment. Cell-specific transactivation elicited by Alx3 from these sites could not be predicted from in vitro DNA-binding experiments by using recombinant Alx3. In addition, transactivation did not depend on cooperative dimerization upon binding to cognate somatostatin DNA sites. Our data indicate that the paradigm according to which Alx3 must act homodiermically via cooperative binding to P3-like sites is insufficient to explain the mechanism of action of this homeoprotein to regulate transcription of natural target genes. Instead, Alx3 undergoes restrictive or permissive interactions with nuclear proteins that determine its binding to and transactivation from TAAT target sites selected in a cell-specific manner.

Homeoprotein domains are a large family of transcription factors that play prominent roles during embryonic development (1, 2). These proteins are characterized by the existence of a common DNA-binding structure, known as the homeodomain, that contains three α-helices spanning 60 amino acids. Comparison of amino acid sequences of known homeodomains from different species shows that residues in the first and third helices are highly conserved. Thus, the relative degree of sequence similarity in the DNA-binding domain defines different classes of homeoprotein domains that relate to protein homologs encoded by Drosophila melanogaster genes (3).

Among the different classes of homeoproteins categorized so far, an important group with major roles in embryonic development is that of paired class proteins, characterized by the presence of a homeodomain homologue to the one encoded by the Drosophila paired gene (4). Additional conserved regions located outside the homeodomain define different subsets of paired class transcription factors.

One of these subsets of paired class homeoproteins is characterized by the presence of a conserved domain located in the carboxyl-terminal region known as OAR or aristaless domain (4, 5). Some of the homeoproteins within this subset are encoded by the so-called group I aristaless-related genes. This group includes the Prx domain-containing proteins Prx1 and Prx2 (6) as well as the homeoproteins Alx3 (7, 8), Alx4 (9), and Cart1 (10) that are highly related both structurally and functionally (5). Aristaless-related genes exhibit overlapping patterns of expression during development in tissues that include the cranial mesenchyme derived from the neural crest, branched arches, body wall mesoderm, and limb buds (11).

Some of the functions of the genes encoding aristaless-like transcription factors have been studied in mice carrying mutant alleles. These studies revealed that group I aristaless-related genes are required for correct skeletal development, since the presence of loss-of-function mutations leads to defects in craniofacial and limb morphogenesis. Specifically, targeted inactivation of Cart1 causes defects in the formation of craniofacial skeleton and in the closure of the neural tube, leading to acrania and meroanencephaly (12), and Alx4 mutant mice exhibit cranial bone defects, polydactyly, and body wall defects (13–15). Alx3 mutant mice have been generated by targeted homologous recombination, but phenotypic defects in these animals have not been reported (16). This finding suggests that there is functional redundancy among these genes, a notion further supported by studies carried out in Alx3/Alx4 and Alx4/Cart1 double mutant mice, which show the existence of overlapping functions of these genes during embryonic development (16, 17).

Despite recent advances in our understanding of the develop-
opmental functions of transcription factors encoded by arista-
less-related genes, little is known about their mechanism of
action at the transcriptional level and the target genes that
they regulate. Here we report the cloning of a cDNA encoding
the rat homolog of Alx3 and an initial molecular characteriza-
tion of this transcription factor. We identify novel DNA sites
from which Alx3 can activate transcription, and we show that
a subset of the 319 amino acids 314–319 located near the C
terminal of Alx3 (5-’GCAGTGACACGCAGCTCCGAGCTG
CATCTGGGAAAGG’) is sufficient for protein dimerization and
sequential sequence-specific DNA binding, whereas full transcriptions ability re-
quires the integrity of proline-rich domains located at either
side of the homeodomain.

EXPERIMENTAL PROCEDURES
Cell Lines—Neural RC2.E10 and RH.1C4 cells derive from cortex
and hippocampus, respectively, of rat fetuses of 16 days of gestational
age and were cultured at a temperature of 33 °C as described (18, 19).
HeLa cells, BHK-21 cells (baby hamster kidney fibroblasts) (ATCC
CCL10), and rat pancreatic islet somatostatin-producing RIN-1027-B2
cells (20) were cultured at a temperature of 37 °C in Dulbecco’s modified
Eagle’s medium containing 10% fetal bovine serum in the presence of
penicillin (100 units/ml) and streptomycin (10 μg/ml). RIN-1027-B2 and
RC2.E10 cells were transfected using Lipofectin (Invitrogen) as de-
scribed (21). cDNA in the poly(A) tail of the polyclonal cell line
JM109 were transformed, and plasmid
pBluescript-KS(+) and sequenced after propagation into Stbl2 (Invitrogen).

Reverse Transcription—PCR with Degenerate Oligonucleotides—Total
RNA (5 μg) purified by CsCl gradient centrifugation from RC2.E10 or
RH.1C4 cells was primed with poly(dT)16 (100 ng) and incubated with
avian myeloblastosis virus reverse transcriptase (Roche Applied Sci-
cence) in a total volume of 30 μl to synthesize cDNA. Four μl of this
cDNA preparation were used as template for PCR amplification, using
degenerate oligonucleotides corresponding to the conserved residues
QLDVLE and QVWFKN found in helices 1 and 3, respectively, of
degenerate oligonucleotide amplimers used contain BamHI or
XhoI sites at their 5 ’-end (5’-GAATTTAGGGTACACATAGAAAAGGTGGC
CTTGATGGAAAAG-3’: reverse, 5 ’-TGTAAATACGACTCATAATTGGGTA
GAGGGCTGTTGGAGC-3’). These oligonucleotides incorporated SP6 and T7 polymerase recognition sites, respectively. The PCR prod-
cut was purified, and the riboprobe was generated by using a Promega
system kit in the presence of [32P]CTP, following the instructions
provided by the manufacturer. Hybridization (42 °C) and washing
(65 °C) were carried out following standard protocols.

Plasmids—The full-length Alx3 cDNA was assembled between the
KpnI and BamHI sites of pBluescript-KS(+) by ligating restriction
fragments generated from the partial cDNAs obtained as described
above. A Muta-Gen phagemid kit (Bio-Rad) was used to alter the codon sequence of the Alx3 cDNA by oligonucleotide-directed
DNA synthesis. Partially randomized (5–20 oligonucleotides) oligonucleotides incorporating into the cDNA by oligonucleotide-directed
mutagenesis so that the resulting cDNAs encode versions of Alx3 tran-
crated at the amino or carboxyl termini. To construct cDNAs encoding
Alx3 amino-terminal deletions, the codon encoding the first methionine
was converted into an NcoI restriction site. NcoI restriction sites were
also introduced by modifying codons encoding Leu-57 or Ser-91. A
control sample that lacked RNA was processed in parallel.

Northern Blot—Poly(A) RNA was isolated from total RNA purified from
rat Alx3 cDNA was generated by PCR using the following oligonu-
ucleotides forward, 5 ’-GAATTTAGGGTACACATAGAAAAGGTGGC
CTTGATGGAAAAG-3’: reverse, 5 ’-TGTAAATACGACTCATAATTGGGTA
GAGGGCTGTTGGAGC-3’. These oligonucleotides incorporated SP6 and T7 polymerase recognition sites, respectively. The PCR prod-
cut was purified, and the riboprobe was generated by using a Promega
RNA polymerase; ChIP, chromatin immunoprecipitation; GST, glutathione
S-transferase; EMSA, electrophoretic mobility shift assay; GFAP, glial
fibrillary acidic protein; RACE, rapid amplification of cDNA ends.

1 The abbreviations used are: CAT, chloramphenicol acetyltrans-
ferase; ChIP, chromatin immunoprecipitation; GST, glutathione
S-transferase; EMSA, electrophoretic mobility shift assay; GFAP, glial
fibrillary acidic protein; RACE, rapid amplification of cDNA ends.
was digested with BamHI and HindIII and cloned into the plasmid pGEX-KG (24). A similar strategy was used to generate GST-Alx3, a fusion protein including full-length Alx3. In this case, the sequence of the forward PCR primer is: 5′-ACCGATCATGGAGCCGAGGCCTG-C′-3′.

To construct the plasmid encoding GST-Alx3-(143–228), a 473-bp fragment was cut with NcoI from the cDNA encoding Alx3-(1–228) in pBluescript-KS(−)I. This fragment was cloned into the NcoI site of pGEX-KG.

**Alx3 Antibodies**—A polyclonal antibody that recognizes the carboxy-terminal region of rat Alx3 (140 residues) was generated in a rabbit inoculated with the GST-Alx3N140C fusion protein expressed in E. coli JM109. Specificity of the antibody was determined by Western immunoblots, using serum presorbed with GST or with GST-Alx3.

**Western Immunoblots**—Nuclear extracts (25) from cells growing in 60-mm dishes were prepared, and proteins were resolved by SDS-PAGE and blotted onto a nitrocellulose membrane. Alx3 immunoreactivity was detected with a rabbit polyclonal primary antiserum (1:5000 dilution) (Bio-Rad). Immunoreactive bands were visualized by using an enhanced chemiluminescence detection system (ECL, Amersham Biosciences).

**DNA-Protein Binding Assays**—Electrophoretic mobility shift assays (EMSA) were carried out with full-length or truncated versions of Alx3 generated by *in vitro* translation by using a rabbit reticulocyte lysate system (Promega) or with nuclear extracts of RC2.E10 cells in the presence of the protease inhibitors pepstatin A (1 mg/ml), leupeptin (10 mg/ml), aprotinin (10 mg/ml), and β-aminobenzamidine (0.1 mM). Synthetic complementary oligonucleotides with 5′-GATC overhangs were annealed and labeled by a fill-in reaction using [α-32P]dATP and Klenow enzyme. Binding reactions were carried out in the presence of 20,000 cpm of radiolabeled probe (1:10,000 dilution) (Bio-Rad). Immunoreactive bands were visualized by using an enhanced chemiluminescence detection system (ECL, Amersham Biosciences).

**Functional Characterization of Alx3**

**Isolation of Rat Alx3 cDNA**—A degenerate reverse transcriptase-PCR-based screen initially aimed at identifying orthodenticle-related homeodomain genes expressed in neural cell lines yielded several homeodomain region-encoding clones whose translated sequence was identical to that of mouse, human, and hamster Alx3 (7, 8, 27). Clones encoding the homodomains of Otx1 and Otx2 were also identified, but no additional Otx-related genes were detected.

The full-length rat Alx3 cDNA isolated using 3′- and 5′-RACE is 1.8 kb in size. However, a number of clones generated by 3′-RACE were found to contain a shorter 3′-untranslated region followed by a poly(A) tail at nucleotide 1724, suggesting that at least two different polyadenylation sites may be used. A putative RNA destabilization signal (AUUUA) (28) was identified within the 3′-untranslated region. Northern blot analysis confirmed the presence of Alx3 mRNA of the expected size in the head and trunk of rat E12 embryos (Fig. 1).

The cloned cDNA encodes a protein of 343 amino acids whose homeodomain is located between residues 153 and 212. The sequence of the rat Alx3 homeodomain is identical to that of the Alx3 homeodomain from human (27), hamster (7), and mouse (8). Outside the homeodomain, rat Alx3 exhibits a high degree of amino acid identity with the known Alx3 orthologs from other species (Fig. 2). These regions are characterized by a relatively high proportion of proline residues distributed into three different domains, two of them in the amino-terminal region, named Pro1 and Pro2, and spanning amino acids 3–79 (25% proline content) and 104–141 (23% proline content), respectively, and one in the carboxy-terminal region named Pro3 and spanning amino acids 223–336 (19% proline content) (Fig. 2).

**Characterization of Transcriptionally Active Domains within Alx3**—To test the functional activity of Alx3, we constructed reporter plasmids bearing three copies of P3- or P5-binding sites, which contain two inverted “TAAT” motifs separated by three or five nucleotides, respectively (22). It is known that aristless-like homeodomain proteins bind preferentially to these sites as dimers (to P3) or monomers (to P5) (17, 29).

Neural RC2.E10 cells were cotransfected with either 3xP3TK-Luc or 3xP5TK-Luc reporters and pcDNA-Alx3, an expression vector encoding full-length rat Alx3. We found that Alx3 increased the luciferase activity elicited by 3xP3TK-Luc by 2–3-fold but did not increase the activity elicited by 3xP5TK-Luc, which was low and close to background levels (data not shown).

To avoid interferences with endogenous Alx3 expressed in RC2.E10 cells, we used BHK-21 cells, since they do not express Alx3 as shown by Western immunoblot (see Fig. 6A). Fig. 3A shows that cotransfection of the Alx3 expression vector in BHK-21 fibroblasts results in higher luciferase activity elicited from the P3-bearing reporter plasmid but not from the P5-bearing reporter plasmid.

To identify regions of Alx3 corresponding to functional domains required for transactivation, we introduced sequential deletions in the amino terminus to residues 57, 91, or 143,
without affecting the integrity of the homeodomain (Fig. 3B). We found that deletion of residues 1–57 (Alx3-(57)) reduced 3xP3TK-Luc activity elicited by Alx3 to 60–70% that obtained with the full-length protein, and that a further deletion to residue 91 (Alx3-(91)) reduced Alx3 transactivational activity to less than 50% (Fig. 3C). Deletion of residues 1–143 (Alx3-(143)), which eliminated the entire amino-terminal region to a position close to the homeodomain, resulted in levels of luciferase activity similar to those obtained with cotransfection of control empty pcDNA3 vector. A similar lack of activation was observed when we used Alx3-(143–228), a truncated Alx3 protein in which both the amino- and the carboxyl-terminal regions were deleted, leaving the homeodomain intact (Fig. 3C). These experiments indicate that regions located in the amino terminus are required for transactivation by Alx3.

However, further experiments demonstrated that regions located within the carboxyl-terminal domain of Alx3 are important although not essential for full transcriptional transactivation activity. Thus, a deletion of the last 64 residues (Alx3-(1–279)) reduced 3xP3TK-Luc luciferase activity to 70% of that induced by full-length Alx3, whereas a deletion of the last 115 residues (Alx3-(1–228)), spanning almost entirely the Pro domain, reduced Alx3 transcriptional transactivation activity by 50% (Fig. 3C).

The Homeodomain of Alx3 Is Sufficient for Dimeric and Monomeric Binding to DNA—To investigate whether the deletion of the proline-rich domains affects binding to DNA, we carried out EMSA by using [35S]Met-labeled P3 and P5 oligonucleotide probes and truncated versions of Alx3 generated in vitro with a reticulocyte lysate system. Fig. 4A shows that Alx3-(143) forms two complexes on the P3 probe, likely to correspond to monomeric and dimeric forms, respectively. As predicted, only Alx3-(143) binds to P3 in a DNA-independent manner (31). To test whether Alx3 dimerization can take place in the absence of a P3-like DNA-binding site, we carried out GST-pull-down experiments. As shown in Fig. 4C, [35S]Met-labeled full-length Alx3 is able to interact specifically with the GST-Alx3 fusion protein but not with control GST-Alx3N140C, which lacks the amino-terminal region and the homeodomain, or with GST-C/EBPβ. Truncated versions of [35S]Met-labeled Alx3 generated by deletion of residues spanning either the carboxyl or the amino termini of Alx3, but leaving intact the homeodomain, did not affect dimerization with the GST-Alx3 fusion protein (Fig. 4D).

Dimerization was also observed by using labeled Alx3-(143–228) (Fig. 4E), but not labeled Alx3-(175–343), which lacks part of the homeodomain (not shown). In addition none of the labeled versions of Alx3 was observed to bind to control GST or to GST-Alx3N140C. Finally, we observed that Alx3-(143–228) can dimerize with GST-Alx3-(143–228) (Fig. 4F). Thus, these experiments indicate that the segment of Alx3 corresponding to the homeodomain is necessary and sufficient for homodimerization and that the presence of DNA corresponding to a cognate P3-like site is not required for Alx3 homodimer formation.

Alx3 Binds Selectively to Specific TAAT-containing DNA Sites—A number of studies have characterized the binding properties of paired-like and aristless-like homeoproteins to P3 and related DNA elements (13, 17, 22, 29, 31). However, it is important to bear in mind that P3-like elements appear not to be common in nature and that target genes regulated by Alx3 and related transcription factors are unknown. To investigate whether Alx3 can regulate transcription in the absence of P3-dependent dimerization, we took advantage of the circumstance that Alx3-expressing RC2.E10 cells also express the genes encoding the neuropeptide hormone somatostatin and the glial fibrillary acidic protein (GFAP) (18, 19), both of which are regulated by promoters that operate in a cell-specific manner (18, 32, 33).

The promoter region of the somatostatin gene contains at least four homeodomain-binding DNA cis-regulatory elements with TAAT core motifs, known as SMS-UE-B, SMS-TAAT1, SMS-TAAT2, and SMS-TAAT3, the function of which has been studied in detail in pancreatic and neural cells (18, 33–35). In the case of the GFAP gene, studies on the regulation of its expression by homeodomain proteins have not been reported. However, inspection of the sequence of the rat GFAP gene promoter revealed the presence of three different sites containing a TAAT motif. We named them GFAPT1, GFAPT2, and GFAPT3, respectively. GFAPT1 is located 1485 bp upstream from the transcription start site defined by Condorelli et al. (36), and GFAPT2 and GFAPT3 are located about 796 and 450 nucleotides from the transcription initiation site, respectively.

As none of these somatostatin and GFAP elements resemble P3 or P5 sequences, we first used EMSA to test whether Alx3 can indiscriminately bind to all of them or whether its binding is specific only for some selected TAAT-containing sequences. Incubation of recombinant Alx3 with [35S]P-labeled synthetic oligonucleotides corresponding to each of the above-mentioned somatostatin or GFAP sites indicated that Alx3 can only bind to SMS-TAAT2 and GFAPT3 but not to any of the other TAAT probes (Fig. 5A). Deletions of amino- or carboxyl-terminal segments of Alx3 did not alter binding to these sites, and, similar to what we found using the P3 and P5 oligonucleotides, we determined that the region of Alx3 corresponding to the
homeodomain and adjacent amino acids (Alx3-(143–228)) is sufficient for binding to SMS-TAAT2 (Fig. 5B) and to GFAPT3 (not shown). In addition, we confirmed that Alx3 binds to SMS-TAAT2 and GFAPT3 as a monomer by mixing Alx3-(143–228) with other truncated versions of Alx3 in the binding reaction. Contrary to what was observed in the case of the P3 oligonucleotide, this did not result in the formation of additional intermediate bands corresponding to dimeric complexes (Fig. 5C).

Nuclear Proteins Alter the DNA Binding Specificity of Alx3—To verify whether endogenous Alx3 is able to bind selectively to these elements in the same sequence-specific manner, we carried out EMSA by using nuclear extracts of RC2.E10 cells in the presence of a specific anti-Alx3 antiserum. For comparison, we also used nuclear extracts of pancreatic islet-derived somatostatin producing RIN-1027-B2 cells. We have previously shown that nuclear proteins from both RC2.E10 and RIN-1027-B2 cells bind to somatostatin TAAT-containing elements forming several sequence-specific protein-DNA complexes (18, 33, 35). Expression of Alx3 in RC2.E10 and RIN-1027-B2 cells was confirmed by Western blot (Fig. 6A).

In the experiments reported above, we found that recombinant Alx3 does not bind SMS-UE-B. Surprisingly, however, we found that addition of the Alx3 antiserum to RC2.E10 nuclear extracts caused the disappearance of the upper band detected with both probes (Fig. 6B). Western immunoblot of nuclear extracts of transfected cells confirming the expression of truncated versions of Alx3. The Alx3 homeodomain (Alx3-(143–228)) was run on a different gel (20% polyacrylamide) due to its relative small size (not shown). Values for luciferase activities represent the mean ± S.E. of at least three independent experiments carried out in duplicate.
tein, respectively (37, 38), we confirmed IDX1 binding to SMS-TAAT1 and SMS-TAAT2 in nuclear extracts of RIN-1027-B2 cells but not in nuclear extracts of RC2.E10 cells (Fig. 6B). Most interestingly, the lower band affected by the IDX1 antisera in the case of the SMS-TAAT1 probe coincides with one of the bands affected by the Alx3 antisera, suggesting that both proteins may interact as part of the same complex bound to SMS-TAAT1 in pancreatic cells.

To gain information about the binding of Alx3 to native chromatin in vivo, we used a ChIP assay to address whether this transcription factor binds to its cognate elements on the somatostatin promoter in the context of the endogenous gene. We found that the anti-Alx3 antisera, but not control serum, immunoprecipitates a fragment of formaldehyde cross-linked chromatin from RIN-1027-B2 cells that contains the SMS-TAAT elements (Fig. 6D). Therefore, it is possible that the somatostatin gene is a target for regulation by Alx3 in these cells.

In the case of the EMSA carried out with the TAAT-containing sites of the GFAP gene, we observed that incubation of RC2.E10 nuclear extracts with each one of the corresponding probes resulted in the generation of several DNA-protein complexes. Competition with homologous and heterologous oligonucleotides demonstrated that binding is sequence-specific (data not shown). Addition of anti-Alx3 or anti-IDX1 antisera did not alter the banding pattern observed with GFAP-T1 or GFAP-T2 (Fig. 6C). However, when GFAP-T3 was used, addition of antisera against Alx3 resulted in the disappearance of the slowest migrating band, an effect that was not observed with control normal rabbit serum or with an IDX1-specific antisera (Fig. 6C). Thus, in the presence of nuclear proteins, Alx3 shows a pattern of sequence-specific binding to TAAT-containing elements.
GFAP promoter sites similar to the one observed with recombinant Alx3.

Alx3 Activates Transcription from Somatostatin TAAT-containing Elements in a Cell-specific Manner

To determine whether Alx3 binding to somatostatin TAAT-containing DNA sites correlates with transcriptional activity, we carried out transient transfection experiments using RC2.E10 and RIN-1027-B2 cells. For this purpose, we used reporter plasmids constructed by placing the SMS-UE-B, SMS-TAAT1, or SMS-TAAT2 elements at position /H11002 65 of the somatostatin promoter in the plasmid SMS65-CAT, which contains the smallest 5/-deletion fragment of the rat somatostatin promoter that retains activity in both RIN-1027-B2 and RC2.E10 cells (18, 33). Co-transfection of the Alx3 expression plasmid with these reporter plasmids in neural RC2.E10 cells did not result in a significant increase in CAT activity (Fig. 7A). In pancreatic RIN-1027-B2 cells, Alx3 did not increase the activity elicited by UE-B-SMS65 but produced a 2–3-fold increase in the activity elicited by TAAT1-SMS65 or TAAT2-SMS65 (Fig. 7A).

In addition to neural and pancreatic cells that express endogenous Alx3, we tested two other heterologous cell lines. In HeLa cells, Alx3 stimulated the activity of UE-B-SMS65 by ~2-fold but induced a robust enhancement of the activity elicited by TAAT1-SMS65 or TAAT2-SMS65 (Fig. 7A). In contrast, Alx3 did not enhance the CAT activity elicited by any of these reporters when cotransfected in BHK-21 cells (Fig. 7A).

It is generally accepted that transcriptional transactivation by aristaless-like homeoproteins requires their binding to DNA in the form of cooperative dimers (17, 29). However, our data generated by EMSA indicate that Alx3 does not bind to somatostatin TAAT-containing elements as a homodimer. Because IDX1 binds to these elements in RIN-1027-B2 cells and interactions between different types of homeodomain transcription factors are not uncommon, we tested whether Alx3 and IDX1 may cooperate to activate transcription from these sites. We found that cotransfection of Alx3 and IDX1 expression vectors with the SMS-UE-B reporter plasmid in pancreatic RIN-1027-B2 cells did not result in an increase in CAT activity (Fig. 7B). Also, the increase in transcriptional activity elicited by Alx3 on the SMS-TAAT1 and SMS-TAAT2 reporters was not enhanced in the presence of IDX1, which in fact caused a decrease in Alx3-induced CAT activity in the case of SMS-TAAT2 (Fig. 7B). Additional experiments indicated that IDX1 does not enhance Alx3-dependent transcriptional activity elicited from somatostatin reporters cotransfected in HeLa or BHK-21 cells. These experiments indicate that there are no functional interactions between Alx3 and IDX1 to transactivate somatostatin gene regulatory elements in the cell lines tested.
DISCUSSION

Proline-rich Domains Confer Transcriptional Transactivation Functions to Alx3—In the present study, we show that Alx3 and the homeoprotein IDX1 present in nuclear extracts of neural RC2.E10 cells bind to the somatostatin SMS-UE-B element (left) but not to the SMS-TAA1 or SMS-TAA2 elements. However, Alx3 and IDX1 present in nuclear extracts of pancreatic RIN-1027-B2 bind to SMS-TAA1 and SMS-TAA2. Arrows indicate the complexes disrupted by the anti-Alx3 and anti-IDX1 antisera. C, Alx3 present in nuclear extracts of neural RC2.E10 cells binds selectively to the GFAPT3 element (right) but not to the GFAPT1 (left) or GFAPT2 (center) elements. Arrow indicates the band disrupted by the anti-Alx3 antiserum. Note that IDX1 did not bind any of the GFAP TAAT-containing elements. Control normal rabbit serum (NRS), an antiserum against Alx3, or one of two antisera to IDX1 (251 or 253) were added to the binding reactions shown in B and C 15 min before the addition of the probes. D, chromatin immunoprecipitation assay carried out in the presence of anti-Alx3 antiserum or normal rabbit serum (NRS). Shown are PCR products corresponding to the region of the somatostatin gene promoter that contains the SMS-TAAT sites recognized by Alx3.

The reduced transcriptional activity of the truncated versions of Alx3 are not likely due to changes in their ability to bind DNA, because even the smallest fragment used (Alx3-(143–228)) contains two nuclear localization signals similar to those characterized in the related homeoprotein Cart1 (31) and binds the P3 site efficiently. Rather, it is possible that proline-rich domains provide a three-dimensional docking structure for transcriptional coupling proteins or coactivators (9, 19, 32), which would not be recruited by truncated Alx3 due to absence of an interacting surface, as it is known that proline residues are critical for many types of protein-protein interactions (39). However, it is also possible that the function of proline-rich residues is linked to local changes in protein structure (40) so that the absence of a specific segment of the protein would alter the conformation of an interacting surface located at a distance (41, 42).

The relatively high proline content of Alx3 may also explain the discrepancy between its predicted molecular mass (36.9 kDa) and the apparent molecular mass indicated by its electrophoretic migration (see Fig. 4C and 6A), because it is known that proline-rich proteins migrate anomalously on SDS-polyacrylamide gels (43). The difference between the predicted and the observed molecular mass is smaller in the case of the related transcription factor Alx4, which contains a significantly smaller number of proline residues (9, 14, 44).

Functional domains outside the homeodomains of Alx4 and
Alx3, because deletion of the region that contains the aristaless inhibitory function (45). This does not seem to be the case for though these proteins exhibit redundant or overlapping func-

tions including Alx4 can also bind to a TAATGG half-site (22),

binding to P5 sites used in this study and fits the consensus half-site preferred by paired Q50 homeodomain pro-

teins (22). Transcriptional activity of aristaless-like proteins bound to this motif requires cooperative dimerization on the DNA, which is made possible by the existence of a palindromic TAAT motif in P3 sites. Therefore, GFAP3 resembles a P5-like site in the sense that binding of an Alx3 monomer is not capable of eliciting transcriptional transactivation, as we observed in our transfection experiments using GFAP-based luciferase reporters (not shown).

The TAATGG motif present in SMS-TAA22 corresponds to a preferred consensus binding site for paired transcription factors with a histidine (His-50) or isoleucine (Ile-50) residue at position 50 of the homeodomain (47). Most interestingly, we found that recombinant Alx3 and Alx3 present in pancreatic RIN-1027-B2 binds to SMS-TAA2, but Alx3 present in nuclear extracts of neural RC2.E10 cells was prevented from binding to this site. Although the presence of competing nuclear proteins that recognize the same site cannot be excluded (49), a more likely possibility is that Alx3 interacts with other cell-specific proteins to form complexes with different sequence

Specificities in neural cells.

The TAATGG motif of the SMS-UE-B oligonucleotide is a preferred consensus binding site for Fushi tarazu (47), which is not a paired-like homeoprotein but contains a glutamine at position 50 within its homeodomain. Paired Q50 homeoproteins including Alx4 can also bind to a TAATGG half-site (22), and therefore, it is not entirely surprising that although recombinant Alx3 does not recognize the SMS-UE-B, Alx3 expressed in RC2.E10 cells binds this site, perhaps favored by neural Alx3-interacting proteins. An analogous situation may explain the occupation of the SMS-TAAT1 site by Alx3 present in neural RC2.E10 cells (which also express GFAP) and in pancreatic RIN-1027-B2 cells to explore the binding preferences of Alx3 in two different types of nuclear protein environments and to compare them to those exhibited by recombinant Alx3 in isolation. Coexpression of Alx3 with somatostatin in neural RC2.E10 cells (which also express GFAP) and in pancreatic RIN-1027-B2 cells does not necessarily imply that these genes are regulated by Alx3 in vivo, but this possibility cannot be formally excluded because our ChIP experiments show that Alx3 occupies the promoter of the somatostatin gene, and it is expressed in neural and pancreatic islet cells of embryonic and adult rats.2

We found that from a total of seven TAAT-containing sites present in the promoter regions of the somatostatin and GFAP genes, recombinant Alx3 only binds to the GFAP3 and SMS-TAA2 elements, which contain the core sequence motifs TA-ATTG and TAATTA, respectively. None of these two motifs is present in any of the other somatostatin or GFAP oligonucleotides tested. The TAATTG motif contained in GFAP3 is also present in the P3 and P5 sites used in this study and fits the consensus half-site preferred by paired Q50 homeodomain proteins (22). Transcriptional activity of aristaless-like proteins bound to this motif requires cooperative dimerization on the DNA, which is made possible by the existence of a palindromic TAAT motif in P3 sites. Therefore, GFAP3 resembles a P5-like site in the sense that binding of an Alx3 monomer is not capable of eliciting transcriptional transactivation, as we observed in our transfection experiments using GFAP-based luciferase reporters (not shown).

The TAATGG motif present in SMS-TAA22 corresponds to a preferred consensus binding site for paired transcription factors with a histidine (His-50) or isoleucine (Ile-50) residue at position 50 of the homeodomain (47). Most interestingly, we found that recombinant Alx3 and Alx3 present in pancreatic RIN-1027-B2 binds to SMS-TAA2, but Alx3 present in nuclear extracts of neural RC2.E10 cells was prevented from binding to this site. Although the presence of competing nuclear proteins that recognize the same site cannot be excluded (49), a more likely possibility is that Alx3 associates with other cell-specific proteins to form complexes with different sequence

Specificities in neural cells.

The TAATGG motif of the SMS-UE-B oligonucleotide is a preferred consensus binding site for Fushi tarazu (47), which is not a paired-like homeoprotein but contains a glutamine at position 50 within its homeodomain. Paired Q50 homeoproteins including Alx4 can also bind to a TAATGG half-site (22), and therefore, it is not entirely surprising that although recombinant Alx3 does not recognize the SMS-UE-B, Alx3 expressed in RC2.E10 cells binds this site, perhaps favored by neural Alx3-interacting proteins. An analogous situation may explain the occupation of the SMS-TAAT1 site by Alx3 present in nuclear extracts of RIN-1027-B2 cells, even though recombinant Alx3 does not recognize this site.

2 M. Mirasierra and M. Vallejo, unpublished observations.
Thus, recombinant Alx3 used in isolation from other nuclear proteins selectively recognizes different versions of TAAT sites that differ on their flanking nucleotides. In turn, this binding selectivity can be altered in the presence of different nuclear protein environments. These data support the notion that Alx3 shows structural plasticity for the recognition of distinct target sites that is modulated in a cell-specific manner via interactions with other nuclear proteins.

Alx3 Transactivates Transcription from Target DNA Regulatory Elements in a Cell-specific Manner—Transactions in different cell lines confirmed the notion that the nuclear protein environment may modulate binding and transcriptional activity of Alx3 from selected somatostatin regulatory elements in a cell-specific manner. Thus, in neural RC2-E10 cells Alx3 is unable to transactivate from any of the somatostatin regulatory elements tested, despite the fact that it was found by EMSA to occupy at least the SMS-UE-B site in nuclear extracts prepared from these cells. In pancreatic RIN-1027-B2 cells Alx3 was found to increase transcription from the SMS-TAAT2 site, consistent with the binding data obtained with recombinant and nuclear Alx3, and from the SMS-TAAT1 site, it was consistent with the observed cell-specific occupation of this site detected with nuclear extracts. Finally, in two heterologous cell lines, opposite effects were observed. On the one hand, in HeLa cells Alx3 was shown to transactivate from all the somatostatin reporters tested, but in BHK-21 cells, on the other hand, it was unable to transactivate from any of them.

It is generally accepted that aristless-related homeoproteins are unable to activate transcription acting in monomeric form. Thus, transactivation from SMS-TAAT2 and SMS-TAAT1, where Alx3 binds as a monomer, may require the interaction with other homeodomain or nonhomeodomain proteins bound in its vicinity. Three-dimensional studies show that a TAATTAAT motif similar to the one in SMS-TAAT2 could accommodate the binding of two homeodomain proteins located at opposite sides of the DNA helix (50, 51). As for nonhomeodomain-related transcription factors, Alx3 can interact with basic helix-loop-helix proteins, and Alx4 has been shown to establish functional interactions with LEF-1 to regulate the activity of the N-CAM promoter (52). Further studies are required to identify functionally active Alx3-interacting proteins.

Acknowledgments—We thank Dr. Petra Schwartz for early discussions and contributions, Dr. Jorge Ferrer (Hospital Clinic, Barcelona) for advice on ChIP assays, and Dr. Joel F. Habener (Massachusetts General Hospital, Boston) for anti-IDX1 antisera.