Paraxis Is a Basic Helix-Loop-Helix Protein That Positively Regulates Transcription through Binding to Specific E-box Elements

Members of the Twist subfamily of basic helix-loop-helix transcription factors are important for the specification of mesodermal derivatives during vertebrate embryogenesis. This subfamily includes both transcriptional activators such as scleraxis, Hand2, and Dermo-1 and repressors such as Twist and Hand1. Paraxis is a member of this subfamily, and it has been shown to regulate morphogenetic events during somitogenesis, including the transition of cells from mesenchyme to epithelium and maintaining antagonistic anterior/posterior polarity. Mice deficient in paraxis exhibit a caudal truncation of the axial skeleton and fusion of the vertebrae. Considering the developmental importance of paraxis, it is important for future studies to understand the molecular basis of its activity. Here we demonstrate that paraxis can function as a transcriptional activator when it forms a heterodimer with E12. Paraxis is able to bind to a set of E-boxes that overlaps with the closely related scleraxis. Paraxis expression precedes that of scleraxis in the region of the somite fated to form the axial skeleton and tendons and is able to direct transcription from an E-box found in the scleraxis promoter. Further, in the absence of paraxis, Pax-1 is no longer expressed in the somites and presomitic mesoderm. These results suggest that paraxis may regulate early events during chordogenesis by positively directing transcription of sclerotome-specific genes.

The basic helix-loop-helix (bHLH) superfamily of transcription factors regulates a wide array of developmental processes, including neurogenesis, myogenesis, cardiogenesis, hematopoiesis, and gametogenesis (1–8). The bHLH transcription factors belong to a superfamily of helix-loop-helix transcription factors that contains more than 240 proteins that are expressed in eukaryotic organisms that range from yeast to humans (9). Members of this family possess a highly conserved functional domain containing a stretch of basic amino acids adjacent to two amphipathic α-helices separated by a loop (10). The helices mediate dimerization with a second bHLH factor, bringing the basic domains into close proximity and forming the bipartite DNA binding domain. The bHLH dimers bind specifically to a hexanucleotide sequence (CANNTG), known as an E-box (11). E-boxes are found in the control regions of many lineage-specific genes. In some cases, such as the muscle-specific genes myogenin, myosin light chain, and muscle creatine kinase, the E-box is required for full transcription of the gene (12–16).

The bHLH proteins can be divided into two classes based on tissue distribution and dimerization capabilities. Class A bHLH factors are known as E proteins and include the E2A gene products E12 and E47, HEB, and the Drosophila daughterless gene product (17, 18). These genes are broadly expressed and able to form homo- and heterodimers. Class B bHLH factors exhibit tissue-restricted expression patterns (19, 20). Within this class are both transcriptional activators, such as MyoD and Mash1, and transcriptional repressors, such as Hand1 (21), MyoR (22), and the Enhancer of Split (E(spl)) family of repressors (23). The Twist subfamily of Class B bHLH factors includes Twist (24), Dermo-1 (25), paraxis (26, 27), scleraxis (28), HAND1 (29–31), and HAND2 (2). These genes encode developmentally regulated bHLH factors that are important for the formation of mesodermal derivatives during vertebrate embryogenesis. This subfamily also includes transcription factors that function as transcriptional enhancers (Twist, scleraxis, Dermo-1, and HAND2) (2, 24, 25, 29) and repressors (Twist and HAND1) (32–34).

The classic view of Class B bHLH proteins held that they formed heterodimers with Class A bHLH proteins only. However, recent evidence demonstrates that the HAND proteins can heterodimerize with each other and with other class B bHLH proteins (35–37). Further, Twist can be a repressor or an activator depending on its dimerization partner (38). It is possible that Class B heterodimers expand the potential E-boxes that can be bound or may sequester these proteins in inactive complexes (for review, see Ref. 37). This suggests that an additional level of transcriptional regulation exists based on dimer partners (35–38).

Paraxis is expressed in the anterior two-thirds of the presomitic mesoderm and throughout the somite in mouse (26, 27), chick (38, 39), and zebrafish embryos (40). As somites compartmentalize, this gene is expressed transiently in the sclerotome while being maintained in the epithelial dermomyotome (27, 38). Transcription of paraxis overlaps with Twist and Dermo-1 in the dermomyotome and transiently with scleraxis in the sclerotome. Mouse embryos deficient in paraxis fail to undergo the normal morphological transition from mesenchyme to epithelium associated with somite formation (41). Paraxis also is required for the maintenance of the anterior/posterior polarity of somites, which prepatterns individual vertebrae and influ-
ences the pattern of peripheral nerves (42). It is also involved in the specification and differentiation of the myogenic and chondrogenic cell lineages that are derived from the somites (41, 43).

Understanding the function of paraxis and its target genes during the regulation of these developmental processes requires determining how paraxis works at the molecular level. In this study, we demonstrate that the paraxis protein is capable of acting as a transcriptional activator and DNA binding is occurring through the bHLH domain. Paraxis was able to bind a subset of E-boxes tested as a heterodimer with E12, but not as a homodimer. Paraxis induced expression of a reporter gene under control of the E-box from the scleraxis promoter that is required for transcription of this gene in vivo. Further, experiments demonstrate that paraxis must be present for normal expression of Pax-1 in the somites and presomitic mesoderm of mouse embryos.

EXPERIMENTAL PROCEDURES

Cells and Transfections—10T1/2 cells and COS-7 cells were grown in Dulbecco’s modified Eagle’s medium (Invitrogen) supplemented with 10% fetal calf serum. 10T1/2 cells were transfected using calcium phosphate as described in Ref. 44. COS-7 cells were transfected using Lipofectamine and the PLUS reagent (Invitrogen), according to the manufacturer.

Plasmids and Construction of Mutants—To generate specific amino acid mutants, a “rolling circle” PCR method was employed (45). For each PCR, 10 ng of template (paraxis-RSET). 1 μl of each primer at 300 nM, and the standard buffer supplied with the Expand Long Template PCR System (Roche Applied Science) was used according to the manufacturer’s protocol. The Q71K/R72K mutant was created by mutating the sequence CACGCGA to AAGAAG, followed by digestion with SnaBI (New England Biolabs). The A76X mutation was created by changing the sequence from CCGGCC to CTAAAC and digestion with HpaI (New England Biolabs). F90P mutation was created by changing the original sequence TTCACC to CCGGTC and digestion with AgeI (New England Biolabs). The truncation of the N terminus (N-) was done by digestion of paraxis-RSET with Ncol and Smal, followed by Klenow fill-in of the ends and religation. The truncation of the C terminus (C-) was done by digestion of paraxis-RSET with MscI followed by religation. The same mutants were subsequently subcloned into a mammalian expression system. Primers were designed with an EcoRI site at the 5′-end and a NolI site at the 3′-end of paraxis, and PCR was done using the RSET constructs as templates. The PCR products were then cloned into pcDNA3.

Electrophoretic Mobility Shift Assays (EMSAs)—Coupled in vitro transcription-translation reactions were performed with 0.5 μg of each plasmid DNA and TNT reticulocyte lysates using T7 polymerase (Promega). Translation products were labeled with Tran35S-label from Promega. Translation products were labeled with [14C]chloramphenicol (PerkinElmer Life Sciences). Translation/translation reactions in the presence of Tran35S-label were tested both as a homodimer and as a heterodimer with E12 proteins generated by coupled transcription and translation using the TNT T7 translation kit (Promega). Translation products were labeled with Tran35S-label. The reticulocyte lysates were diluted in 1 ml of immune TNT buffer (140 mm NaCl, 50 mm Tris, pH 7.5, 1 mM EDTA, 0.1% Triton X-100, 10% glycerol). The diluted lysates were immunoprecipitated using anti-His antibody and 25 μl of protein A/G-Plus agarose (Santa Cruz Biotechnology). Lysates were incubated for 4 h at 4°C with shaking, followed by centrifugation and then washed five times in 1 ml of TNT buffer.

Proteins were released from the agarose beads by boiling in 2× Laemmli sample buffer and analyzed on 15% SDS-PAGE.

Chloramphenicol Acetyltransferase (CAT) Assays—Cells were grown in 60-mm dishes and transfected as described above. Cells were harvested 48 h post-transfection, and CAT assays done as described previously (44) with the following modifications. Cells were washed three times in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na2HPO4, 1.4 mM KH2PO4), and then lysates were made by covering cells with 400 μl of reporter lysis buffer (Promega) incubated for 15 min at room temperature with rocking. The plate was then scraped and the lysates transferred to a microcentrifuge tube on ice. Lysates were vortexed for 10 s, heated to 60°C for 10 min, and spun for 2 min at top speed at 4°C in a microcentrifuge. The protein concentration of the lysates was determined by the method of Bradford (37). CAT assays were performed with an equal amount of total protein. CAT assays were done using [3H]chloramphenicol (PerkinElmer Life Sciences) and n-butyl-CoA (Promega) at 37°C for 2 h. The samples were extracted with 500 μl of ethyl acetate and spotted onto a silica gel thin layer chromatography plate and dried. The samples were separated by thin layer chromatography in chloroform:methanol (97:3) in a preequilibrated chamber. The plates were exposed to a PhosphorImage screen overnight and developed and quantified on a Storm PhosphorImager.

In Situ Hybridization—The embryonic pattern of transcription was determined by whole mount in situ hybridization using antisense RNA-labeled digoxigenin-11-UTP probes as described in Ref. 43. Embryonic day (E) 9.5 and E10.5 embryos were probed with antisense RNA specific to Pax-1, and transcripts were visualized using anti-digoxigenin antibodies conjugated to alkaline phosphatase; color reactions were done using BM Purple (Roche Applied Science).

RESULTS

E-boxes (CANTTG) are found in high abundance throughout the genome, due in large part to the low level of complexity. The common E-box and the E-boxes that are flanking sequences that influence the binding preference of individual bHLH dimer pairs (50, 51). The ability of bHLH transcription factors to bind DNA in a sequence-specific manner is determined by specific residues within the basic region of the bHLH domain. For example, in the myogenic bHLH subfamily, conserved alanine and threonine residues in the basic region confer muscle specificity (52–56). Additionally, residues in the junction between the basic and HLH domains can influence E-box specificity without directly binding DNA (57). The bHLH domain of paraxis is highly homologous to that of members of the Twist subfamily (Fig. 1). To examine the ability of paraxis to bind E-boxes in a sequence-specific manner, EMSAs were performed using E-boxes that have been reported to be targets of other Twist subfamily members (13, 28, 46–49). Because Twist can bind DNA as a heterodimer with E proteins, as well as a homodimer (58), the ability of paraxis to bind specific E-boxes was tested both as a homodimer and as a heterodimer with E12.

End-labeled double-stranded E-boxes (Fig. 2A) were incubated with paraxis and/or E12 proteins generated by coupled in vitro transcription and translation reactions in programmed reticulocyte lysates. Among the E-boxes tested was the right E-box of the muscle creatine kinase gene (RMCK), this E-box is bound by E12 homodimers and E12/scleraxis heterodimers (59). Scleraxis will also bind to an E12/paraxis heterodimer from its own promoter as a heterodimer with E12 (59, 60), and rhomboid promoters are bound by E12/Twist heterodimers (58). An E-box from the promoter of the insulin gene (Ins-1) binds to an E12/Beta1 heterodimer (48). As can be seen in Fig. 2B, E12/paraxis heterodimers were able to shift the Scx and Ins-1 E-boxes well, the RMCK and the Da promoter were also shifted. E12/paraxis heterodimers were not able to bind to the E-box from the Rhomboid promoter. Paraxis homodimers demonstrated no detectable binding activity. To confirm protein expression, dupli-
cate Tran35S-labeled in vitro transcription and translation reactions were analyzed by SDS-PAGE (data not shown). Thus, paraxis, a Class B bHLH protein, binds a subset of E-boxes preferentially as a heterodimer with E12, a Class A bHLH factor.

The specificity of the DNA binding activity of the E12/paraxis heterodimer detected by EMSA was examined by the addition of an unlabeled competitor probe to the EMSA reaction. Because both E12 and E12/paraxis dimers shifted the Ins-1 E-box it was used as a target for these assays. The use of a 1:1 ratio of E12 and paraxis input plasmids in the coupled transcription/translation reaction resulted in the preferential formation of the heterodimer in the presence of the Ins-1 E-box. However, a 2:1 E12/paraxis plasmid ratio in the transcription/translation reaction generated approximately equal levels of both the E12 homodimer and E12/paraxis heterodimer on this E-box (Fig. 3). Therefore, the 2:1 ratio of input plasmids was used to generate proteins for EMSA to allow for comparison of the effect of competitor on the E12/paraxis heterodimer and the E12 homodimer. As can be seen in Fig. 3, both were inhibited effectively with 20-fold molar excess unlabeled Ins-1 added to the EMSA reaction. In comparison, even 100-fold molar excess of mutated Ins-1 had no observable effect on either dimer (Fig. 3). These observations indicate that paraxis binds in a sequence-specific manner to E-boxes in a heterodimer with E12.

To assess the role of the bHLH domain in mediating the observed DNA binding and dimerization activity, a series of

*Fig. 1. Alignment of the amino acid sequence of the bHLH motifs for paraxis and closely related genes in the Twist subfamily of bHLH transcription factors. The assignment of the basic domain, helix I, and helix II and the numbering of the amino acids is based on the system of Ferre-D’Amare (71). Residues conserved in all members of the family are shaded with a gray box. Paraxis, accession number AAA86825; Dermo-1, Q9D030; scleraxis, AAH62161; Twist, AAH33434; HAND1, AAH50182; and HAND2, Q61039.*

*Fig. 2. Binding site preferences for paraxis and E12 heterodimer complexes. A, oligonucleotide sequences containing E-boxes found in the enhancer region of myosin creatine kinase (RMCK), Ins-1, daughterless (Da), rhomboid, and scleraxis (Scx). The E-box sequence is underlined. B, EMSA using in vitro transcribed and translated paraxis and/or E12 proteins and 32P-labeled E-box oligonucleotide probes. Arrowheads mark the bands generated by the binding of the E12/paraxis heterodimer and E12 homodimer.*
mutations was generated in paraxis. Truncation mutations were made in which amino acids 1–70, which are N-terminal to the bHLH domain (N-), or amino acids 123–196 on the C-terminal side of the bHLH domain (C-), were deleted (Fig. 4A). A third mutation removed all but the bHLH domain of paraxis (N-/C-). Analysis of these mutations will allow us to determine whether residues on either side of the bHLH domain contribute to DNA binding and dimerization. Point mutations within the basic domain and first helix were generated to disrupt these domains (Fig. 4B). X-ray crystallography and mutagenesis studies of the myogenic bHLH factors have demonstrated that specific positions within the basic domain are critical for the sequence specificity of E-box binding (61, 62). This includes basic residues at the first three positions and subfamily specific amino acids at positions 5 and 6. In the Twist subfamily, an alanine at position 5 and an asparagine at position 6 are conserved in all members except HAND1 (Fig. 1). Therefore, Ala was mutated to Asn, resulting in an E12-like Asn-Asn. Another highly conserved residue is the arginine in the third position of the basic domain. Interestingly, in the repressor Hand1 this residue is lysine. This Arg residue is preceded by a conserved glutamine residue, so the Glu-Arg of paraxis was mutated to Lys-Lys. Within the first helix of paraxis there is a phenylalanine residue that is conserved in many bHLH factors. When this residue was mutated to the helix breaker proline in myogenin it blocked dimerization with E12 (52), the Phe → Pro substitution was also constructed in paraxis. All translations were done in duplicate with Tran35S-labeled translations, and all mutants were detected at comparable levels of expression as full-length paraxis (data not shown).
The ability of paraxis and the paraxis mutants to dimerize with E12 was examined by coimmunoprecipitation (Fig. 5). The mutants were constructed with the His6 epitope tag at their N terminus. The mutant/E12 heterodimers were communoprecipitated from Tran35S-labeled reticulocyte lysates after in vitro transcription and translation was done using an anti-His antibody and analyzed on 15% SDS-PAGE. As can be seen in Fig. 5, none of the mutations completely abolished dimerization with E12 because this protein was present in each communoprecipitation. However, truncations of residues that are either N- or C-terminal to the bHLH domain and the Phe → Pro substitution greatly disrupted dimerization, as noted by the greatly decreased intensity of the E12 band in these samples.

To determine how the mutations affected the ability of E12/paraxis heterodimers to bind DNA, EMSA was performed using the Ins-1 E-box as a target. Initially, the translation efficiency and stability of the different mutant paraxis proteins were determined by analysis of Tran35S-labeled proteins from coupled transcription/translation reactions on SDS-PAGE. The mutants were translated at nearly equivalent levels (data not shown). Each mutant paraxis protein and E12 were cotranslated in unlabeled reticulocyte lysates to determine the effect of these mutants on DNA binding in EMSA. Truncation of the amino acids N-terminal to the bHLH domain (N-) did not affect DNA binding; this was surprising because this mutant demonstrated the least dimerization ability (Fig. 5), but the E12/N-heterodimer bound the scleraxis E-box similarly (data not shown). This may indicate that the N-mutant only forms stable heterodimers in the presence of DNA. Truncation of the amino acids C-terminal to the bHLH domain (C-) abolished DNA binding activity (Fig. 6). The bHLH domain alone (N/-C-) bound weakly to the Ins-1 E-box with E12. Taken together this indicates that the C-mutant may be unstable structurally or may not fold correctly, and this abrogates DNA binding. Of the point mutants in the basic domain and helix 1 of paraxis, only the Phe → Pro substitution mutant did not bind DNA (Fig. 6). This is consistent with data from other bHLH proteins (2, 56, 61).

Like full-length paraxis, none of the mutants bound DNA as a homodimer (data not shown).

The transactivation function of paraxis was tested using a reporter gene constructed with four multimerized scleraxis E-boxes controlling expression of the CAT gene in the pCAT basic vector (4×ScxCAT). The scleraxis E-box was chosen because this demonstrated the highest affinity of binding in EMSA (Fig. 5).
Paraxis Is a Transcriptional Activator

Fig. 7. Paraxis acts as a transcriptional activator. A, paraxis was tested for its ability to enhance transcription from an E-box. Paraxis and/or E12 expression vectors were cotransfected into 10T1/2 cells with 4×ScaCAT. Transcription was measured as the fold increase of CAT activity over the vector only control. E12/paraxis heterodimers induced the 4×ScaCAT reporter 5-fold over background. B, plasmids encoding GAL4 DNA binding domain fused to paraxis, N- or C-mutants, or with the strong viral activator VP16, or the parent vector were cotransfected with GAL4E1b.CAT, and transcriptional activation was measured by CAT assay. The results are expressed as fold activation over the vector only control. E12/paraxis heterodimers induced the 4×ScaCAT reporter 5-fold over background. When paraxis and E12 were cotransfected it resulted in 5-fold activation of the CAT reporter gene (Fig. 7A). These data indicate that paraxis can bind to E-boxes and transactivate target genes and is thus a transcriptional activator; however, the level of transcription was not robust.

The transcriptional function of paraxis was tested further in the context of GAL4 fusion proteins. Full-length paraxis and the N- and C-mutants were cloned in-frame with the DNA binding domain (DBD) of GAL4 and cotransfected into 10T1/2 cells with a CAT reporter gene under control of GAL4 (GAL4E1b.CAT). The GAL4DBD-paraxis fusion protein induced 11.2-fold activation of the GAL4E1b.CAT reporter. Deletion of either the N or C terminus of paraxis resulted in the loss of transcriptional activation in this context. This indicates that amino acid residues that are necessary for transactivation are found in both the N- and C-terminal regions of paraxis. Alternatively, this could reflect structural instability of the proteins (Fig. 6). When fused to the GAL4 DBD, the strong viral transcriptional activator VP16 induced more than 100-fold activation of this reporter (Fig. 7B).

The skeletal deficiency in the paraxis1−/− neonates and its ability to direct transcription from the scleraxis E-box predict that paraxis is required for normal transcription of sclerotome-specific genes. The transcription factor, Pax-1, is necessary for normal sclerotome formation during embryogenesis (63). Pax-1-deficient mice have defects of the axial skeleton and ribs that are similar to defects in the paraxis mutant mice (63). Pax-1 expression was analyzed by whole mount in situ hybridization using E9.5 and E10.5 paraxis1−/− and wild type embryos. In paraxis1−/− embryos, Pax-1 transcripts were absent in the sclerotome (Fig. 8). Pax-1 transcription was unaltered in the nonsomatic tissues, such as the pharyngeal pouches and the anterior proximal aspect of the forelimb (63–65).

Paraxis has been found to be an important regulator of morphological events during somitogenesis. In the absence of this gene, somites fail to undergo the mesenchymal to epithelial transition associated with somite formation (41). Further, individual cells within the somite fail to maintain their positional identity along the anterior/posterior axis (42). Paraxis is also required for the proper differentiation of the hypaxial skeletal muscle and axial skeleton that is derived from the somites (41, 43). Despite our understanding of the developmental role of paraxis, its downstream target genes and cofactors are unknown. As an important first step, we have examined the DNA binding activity, dimerization preference, and transactivation activity of paraxis.

Paraxis was initially isolated through a low stringency hybridization screen of a mouse embryonic cDNA library using the bHLH domain of scleraxis as a probe (27). Scleraxis is required for mesoderm formation during gastrulation (59). Later during embryogenesis, scleraxis is expressed in mesenchymal precursors of the axial and appendicular skeleton and tendons (29, 66, 67). Scleraxis has been shown to act as a transcriptional activator by binding to E-box sequences as a heterodimer with E proteins. In osteoblastic ROS17/2.8 cells, scleraxis has been shown to induce the transcription of aggrecan through two E-boxes in the promoter of this gene (68). Based on the high degree of homology within the bHLH domain (7 of 9 amino acids in the basic domain and 40 of 46 amino acids in HLH domain are conserved), we anticipated that paraxis would be a transcriptional activator and have E-box binding preferences similar to those of scleraxis.

Initially, the ability of paraxis to shift E-boxes favored by other Twist family members was examined. In EMSAs (Fig. 2B) paraxis bound to the RMCK and Sca E-boxes, similar to scleraxis, but it also bound to the Ins-1 and Rhomboid E-boxes that scleraxis was not able to bind (59). Further, paraxis did not bind DNA as a homodimer, only as a heterodimer with E12.

DISCUSSION

Paraxis has been found to be an important regulator of morphological events during somitogenesis. In the absence of this gene, somites fail to undergo the mesenchymal to epithelial transition associated with somite formation (41). Further, individual cells within the somite fail to maintain their positional identity along the anterior/posterior axis (42). Paraxis is also required for the proper differentiation of the hypaxial skeletal muscle and axial skeleton that is derived from the somites (41, 43). Despite our understanding of the developmental role of paraxis, its downstream target genes and cofactors are unknown. As an important first step, we have examined the DNA binding activity, dimerization preference, and transactivation activity of paraxis.

Paraxis was initially isolated through a low stringency hybridization screen of a mouse embryonic cDNA library using the bHLH domain of scleraxis as a probe (27). Scleraxis is required for mesoderm formation during gastrulation (59). Later during embryogenesis, scleraxis is expressed in mesenchymal precursors of the axial and appendicular skeleton and tendons (29, 66, 67). Scleraxis has been shown to act as a transcriptional activator by binding to E-box sequences as a heterodimer with E proteins. In osteoblastic ROS17/2.8 cells, scleraxis has been shown to induce the transcription of aggrecan through two E-boxes in the promoter of this gene (68). Based on the high degree of homology within the bHLH domain (7 of 9 amino acids in the basic domain and 40 of 46 amino acids in HLH domain are conserved), we anticipated that paraxis would be a transcriptional activator and have E-box binding preferences similar to those of scleraxis.

Initially, the ability of paraxis to shift E-boxes favored by other Twist family members was examined. In EMSAs (Fig. 2B) paraxis bound to the RMCK and Sca E-boxes, similar to scleraxis, but it also bound to the Ins-1 and Rhomboid E-boxes that scleraxis was not able to bind (59). Further, paraxis did not bind DNA as a homodimer, only as a heterodimer with E12.
Paraxis Is a Transcriptional Activator

Paraxis is required for transcription of Pax-1 in the sclerotome. Whole mount in situ hybridizations were performed using a digoxigenin-labeled antisense Pax-1 riboprobe to examine Pax-1 transcription in paraxis+/+ (A and C) and paraxis−/− (B and D) embryos. In wild type E9.5 embryos (A), Pax-1 is transcribed in the pharyngeal pouches (P1, P2, and P3) and the sclerotome (s). At E10.5 (C), Pax-1 is also transcribed in the proximal anterior domain of the forelimb (fl); black arrow). The sclerotomal staining is absent in paraxis−/− embryos.

The ability to form heterodimers with a Class A bHLH protein, such as E12, is a hallmark of the Class B bHLH proteins. E12/paraxis heterodimer E-box binding could be specifically inhibited by the addition of an unlabeled competitor, but not by a random sequence, indicating that this is sequence-specific DNA binding (Fig. 3). Flanking sequences have been known to influence E-box preference. The Rhomboid and Scr E-boxes demonstrate this well for paraxis; these are identical E-boxes with completely different flanking sequences and paraxis bound to the Scr but not the Rhomboid E-box.

The grouping of paraxis into the Twist subfamily of the HLH transcription factor family is based on a high level of sequence homology in the bHLH domains at the amino acid level. Interestingly, members of the Twist family can be either transcriptional activators such as scleraxis and HAND2 (69, 70), or repressors like HAND1 (32, 72), and mammalian Twist, which has been reported to be both an activator and a repressor (34). When the amino acid sequences of the bHLH domain of paraxis and the other Twist family members are compared (Fig. 1), there are several conserved residues in Twist, paraxis, scleraxis, and HAND2 that are replaced in HAND1. These residues are arginine in position 2 substituted for lysine, and alanine and asparagine at positions 5 and 6 instead of glycine and proline. This proline residue is also found in the basic domains of the hairy/E(spl) family of bHLH repressors (23). Using an amino acid substitution frequency table, based on the study of 34 protein superfamilies grouped into 71 evolutionary trees (72), a change of Pro for Asn is very rare, it occurs in only 1.2% of all substitutions for Asn. Both paraxis and scleraxis have another substitution of threonine for a highly conserved arginine in the loop domain. Paraxis is highly homologous to other Twist family members and lacks the Pro substitution; overall the amino acid sequence of the paraxis bHLH domain predicts that it is a transactivator.

Mutants of paraxis were examined for their ability to dimerize with E12 by coimmunoprecipitation and for their ability to bind to the Insa-1 E-box by EMSA (Figs. 5 and 6). Although several mutations demonstrated inhibition of dimerization, as evidenced by a reduced E12 band, none of the mutations completely blocked the ability of paraxis to interact with E12. The Phe→Pro mutant was surprising as a similar mutant in myogenin abolished dimerization. Interestingly, the C- and N-C truncations and the Phe→Pro substitution mutant that demonstrated reduced dimerization with E12 also demonstrated no DNA binding. Either the low level of E12 observed in the coimmunoprecipitations was not enough to produce a shift in EMSA or the dimers were not stable. Neither paraxis (Figs. 2 and 3) nor any mutants of paraxis could shift DNA as a homodimer (data not shown).

To determine whether paraxis was a transcriptional activator we examined the ability of paraxis to induce transcription of a reporter gene (Fig. 7, A and B). In transfected COS-7 cells, a GAL4-paraxis fusion protein induced transcription of a GAL4ElbCAT reporter. GAL4-paraxis was a potent transactivator of this reporter, indicating that paraxis contained a transactivation domain. Further, a multimerized scleraxis E-box reporter gene was constructed, and paraxis induced transcription of this reporter gene in the presence of E12; however, activation of this reporter gene was not as robust as the GAL4CAT gene (Fig. 7, A and B). Currently there are no known target genes for paraxis, so it is possible that none of the E-boxes used in EMSA would be targets in vivo. Alternatively, there is another cofactor that is not present in these cells which is required for high levels of transactivation by paraxis. For example, the myogenic bHLH factors form complexes with Mef2 proteins and synergistically activate muscle-specific genes through E-boxes or Mef2 sites (45).

The ability of paraxis to bind to an E-box from the scleraxis promoter that is required for transcription of this gene during embryogenesis (60) suggests that paraxis may play a role in chondrogenesis. Further, paraxis is expressed in the region of the somite fated to become the chondrogenic lineage (41, 42), and this led us to examine the expression of Pax-1, another sclerotomal transcriptional regulator. Embryos that lack paraxis demonstrated a loss of Pax-1 transcription in the somites (Fig. 8). This gene is a sclerotomal marker whose expression precedes scleraxis, and embryos deficient for Pax-1 have defects of the axial skeleton and ribs (63). This phenotype is consistent with the paraxis-null mutants (42). This may reflect an overlapping set of target genes during sclerotome formation and unique gene targets for paraxis and scleraxis during myogenesis and bone and tendon formation, respectively.

Acknowledgments—We thank Dr. Brian Black for the paraxis cDNA and Drs. Doris Brown and Eric Olson for the scleraxis-pCITE plasmid.

REFERENCES

1. Olson, E. N., and Klein, W. H. (1994) Gene Dev. 8, 1–8
2. Srivastava, D., Cserjesi, P., and Olson, E. N. (1995) Science 270, 1995–1999
3. Yun, K., and Wold, B. (1996) Cell 87, 777–889
4. Guillemot, F. (1999) Exp. Cell Res. 253, 357–364
5. Massari, M. E., and Murre, C. (2000) Mol. Cell. Biol. 20, 429–440
6. Morrison, S. J. (2001) Curr. Opin. Cell Biol. 13, 666–672
7. Staal, F. J., Weerkamp, F., Langerak, A. W., Hendriks, R. W., and Clevers, H. C. (2001) Stem Cells 19, 165–179
8. Dean, J. (2002) J. Reprod. Immunol. 53, 171–180
9. Atchley, W., and Fitch, W. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 5172–5176
10. Ellenberger, T., Pass, D., Arnaud, M., and Harrison, S. (1994) Genes Dev. 8, 970–980
11. Ephrussi, A., Church, G. M., Tonegawa, S., and Gilbert, W. (1985) Science 227, 144–149
12. Buskin, J. N., and Hauschka, S. D. (1989) Mol. Cell. Biol. 9, 2627–2640
13. Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and Weintraub, H. (1998) Cell 98, 823–831
14. Cheng, T. C., Wallace, M. C., Merlie, J. P., and Olson, E. N. (1993) Science 261, 215–218
15. Yee, S. P., and Rigby, P. W. (1993) Genes Dev. 7, 1277–1289
