764–770 *Nucleic Acids Research, 1999, Vol. 27, No. 3*

Transcriptional activation by the homeodomain protein Distal-less 3

Jules A. Feledy, Maria I. Morasso[1], Shyh-Ing Jang[1] and Thomas D. Sargent*

Laboratory of Molecular Genetics, NICHD and [1]Laboratory of Skin Biology, NIAMS, National Institutes of Health, Bethesda, MD 20892, USA

Received October 8, 1998; Revised and Accepted December 4, 1998

ABSTRACT

PCR-based methods and mobility shift competition assays were used to determine the basic biochemical features of the homeodomain transcription factor Distal-less 3 (Dlx3), including an optimal DNA binding site, the binding constant and dissociation rates of this protein. Expression of Dlx3 protein in either HeLa cells or *Xenopus* embryos resulted in strong activation of a model target gene construct containing three tandem copies of the Dlx3 binding site upstream from the TATA element. In addition, deletion analysis revealed that transcriptional activation by Dlx3 depends on two subdomains located on either side of the homeobox: removal of either subdomain resulted in complete loss of Dlx3 function. These observations provide new insight regarding the function of Dlx3 in vertebrate development and tissue differentiation and also suggest a mechanism for the dominant inheritance pattern of a hereditary disease resulting from mutation of the DLX3 gene in human.

INTRODUCTION

Homeodomain proteins comprise a large class of transcription factors that have been shown to be essential regulators of many developmental processes, ranging from organization of the basic body plan to terminal differentiation of individual tissues (1). The homeodomain is a highly conserved 60 amino acid element that is responsible for sequence-specific interactions with DNA. Binding sites for homeodomain proteins center around a core motif, TAAT, with adjacent bases responsible for restricting the interactions between specific homeodomain factors and target genes (2,3). In addition, heteromeric association with other transcription factors and cofactors is thought to further restrict this target specificity.

In *Drosophila*, the Distal-less homeodomain gene (Dll) regulates the formation of ventral appendages (4). In vertebrate genomes, there are six or more genes related to Dll (5,6). In mouse and human these homologs, referred to as Dlx1–6 (DLX1–6), are organized into three pairs of closely linked, convergently transcribed loci (5). Each pair of Dlx genes is located within 1–2 cM of one of the four Hox clusters. Dlx genes are expressed in distinct but overlapping domains, primarily in the forebrain, branchial arches and tissues derived from epithelial/mesenchymal interactions during development (7,8). Dlx3 is transcribed in hair follicles, tooth germ, branchial arch mesenchyme and in the interfollicular epidermis, where it is confined to the suprabasal keratinocyte compartment (7,9,10). Dlx3 differs from the other members of the gene family in that its expression has not been detected in the central nervous system. Also, Dlx3 expression in branchial arches is spatially and temporally distinct from the other members of the Dlx family (8).

The biochemical function of Dlx3 is relevant to important developmental processes. For example, disruption of the DLX3 coding sequence has been shown to be the basis for an inherited human disorder, tricho-dento-osseous (TDO) syndrome. This disorder is characterized by tooth defects, kinky hair at birth and craniofacial bone abnormalities. TDO is inherited as an autosomal dominant trait (11), but the molecular basis for this inheritance pattern is not known. Dlx3 also functions in the differentiation of mammalian epidermis, as shown by ectopic expression of Dlx3 in the basal layer of transgenic mouse skin. This resulted in a number of alterations, including the precocious activation of the terminal differentiation marker profilaggrin. The profilaggrin gene has a binding site for Dlx3 located in the 5′ regulatory region, suggesting the possibility that Dlx3 is a positive regulator of this structural protein (10). However, no data have yet been reported on the transcriptional properties of Dlx3. In this paper we show that Dlx3 is a potent transcriptional activator, map activation domains within the Dlx3 protein sequence and discuss the relevance of these to the function of the Dlx3 homeoprotein *in vivo*.

MATERIALS AND METHODS

Preparation and analysis of *Xenopus* Dlx3 protein

The cDNA encoding the full open reading frame of the *Xenopus laevis* homolog of Dlx3, Xdll2 (12,13; referred to as *Xenopus* Dlx3 in this paper) was cloned into the bacterial expression vector pET 28a (Novagen, Madison, Wl), generating a hexahistidine fusion protein. Following a 1 h induction with 1 mM isopropyl-β-D-thiogalactopyranoside, Dlx3 was purified from lysates by nickel affinity column chromatography according to the manufacturer's instructions. Purification was to near homogeneity, as judged

*To whom correspondence should be addressed. Tel: +1 301 496 0369; Fax: +1 301 496 0243; Email: tsargent@nih.gov

from Coomassie Blue staining of a SDS–polyacrylamide gel. Protein concentrations in column eluates were determined by a trichloroacetic acid precipitation/amido black staining protein assay (Schleicher & Schuell, Keene, NH) to circumvent interference from imidazole in the elution buffer.

Labeled random oligonucleotide pool for selection

A random oligonucleotide pool was generated for selection of ligands in the SELEX procedure (14,15). Individual DNA sequences containing a central 14 bp random sequence between two 15 bp primer sequences incorporating *Eco*RI and *Bam*HI restriction sites [GACGAATTCACGTG(N)$_{14}$GTACGGATCCA-TGCG] were synthesized (BioServe Biotechnologies, Laurel, MD). Single-stranded DNA oligonucleotides (10 ng) were made double-stranded by 14 cycles of PCR using 600 ng each of primers corresponding to the non-random flanking regions [5′-GACGAATTCACGTG and 3′-CGCATGGATCCGTAC] using *Taq* DNA polymerase, 50 µM unlabeled deoxyribonucleotide triphosphates, 2 mM $MgCl_2$ and 1× PCR buffer (Perkin Elmer, Norwalk, CT) in a final volume of 100 µl. Reaction times and temperatures for each cycle were 92°C 40 s, 46°C 40 s, 72°C 20 s. The product was loaded on a 6% polyacrylamide gel in non-denaturing buffer and electrophoresed. The 44 bp band was excised, crushed and eluted overnight in 0.5 M ammonium acetate, 1 mM $MgCl_2$ and 0.1 mM EDTA. The eluted DNA was precipitated with ethanol and resuspended in 0.1 M NaCl, 0.1 mM EDTA, 10 mM Tris, pH 8.0.

SELEX

For each round of enrichment, electrophoretic mobility shifting was used to select oligonucleotides that bound to recombinant Dlx3. Each mobility shift was performed by incubating 12 ng Dlx3 recombinant protein with purified, radiolabeled DNA in a 20 µl reaction containing 7 mM Tris, pH 7.5, 81 mM NaCl, 112.5 mM imidazole (from the Dlx3 elution buffer), 2.75 mM dithiothreitol, 5 mM $MgCl_2$, 0.05% Nonidet P40, 1 mg/ml bovine serum albumin, 25 µg/ml poly(dl–dC)-poly(dl–dC) (Pharmacia Biotech, Piscataway, NJ), 10% glycerol, at 0°C for 30 min. The protein–DNA complexes were electrophoresed on a 6% acrylamide, 0.16% bis-acrylamide gel with 0.5× Tris–borate/EDTA buffer at 10 V/cm. For a positive control a radiolabeled 20 bp oligonucleotide containing a Dlx3 binding site from the human profilaggrin gene (10) was also incubated with Dlx3 and electrophoresed on the same gel. Gels were autoradiographed and the appropriate region was excised, eluted and subsequently amplified by PCR. A total of five cycles of binding and amplification were carried out.

Determination of consensus sequences

Gel-purified oligonucleotides were cloned in the plasmid vector pCRII (Invitrogen, San Diego, CA). Individual clones were sequenced using the Sequenase kit (US Biochemical Inc., Cleveland, OH). The Pileup and Consensus programs from the Genetics Computer Group (GCG) v.8 software package (GCG, Madison, Wl) were used to tabulate consensus sequences.

Gel mobility shift competition assays

Mobility shift assays were performed with various binding sites obtained from the SELEX procedure, as well as with potential sites selected from genomic DNA sequences based on similarity to the Dlx3 consensus binding site. Each site was synthesized as complementary 20 bp oligonucleotides, in which the 14 (SELEX sites) or 12 bp (genomic sites) determined by the individual selected site was bracketed by four or six G residues, respectively, for thermal stability. These oligonucleotides were annealed and used as unlabeled, double-stranded competitors to determine the relative affinity of each sequence compared with a probe containing the restricted consensus sequence. The binding reactions were carried out at increasing concentrations of unlabeled competitor (0, 5, 20, 50 and 100 ng) with 12 ng (17 nM) Dlx3 protein and 1 ng (4 nM) labeled consensus DNA, in a 20 µl reaction as described above. The gels were dried and autoradiographed.

Determination of the equilibrium constant (K_d)

A constant amount (1 ng) of radiolabeled probe including the experimentally determined restricted consensus sequence (GGGG-GGATAATT GCTGGGGGG) was incubated with increasing amounts of Dlx3 protein (0.5 pg–70 ng) for 30 min at 0°C under standard mobility shift conditions (above). After electrophoresis, gels were dried and autoradiographed. Quantification of the free and complexed protein–DNA was performed with a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). The K_d was calculated by Scatchard analysis. In brief, a graph of $\ln x/1 - x$ versus $\ln[C_r - xC_o]$ was made, in which the value of x for each point represented the fraction of the probe bound by Dlx3 protein, C_r was the molar concentration of protein and C_o was the molar concentration of DNA. The K_d was calculated from the inverse log of the y-intercept (16).

Determination of the Dlx3 dissociation rate

Radiolabeled probe (1 ng) including the restricted consensus was incubated with 12 ng Dlx3 protein for 30 min under standard mobility shift conditions. An excess of unlabeled restricted consensus competitor (100 ng) was added. Aliquots were sampled at selected time points and directly loaded onto running gels to minimize dissociation prior to entry into the polyacrylamide. After electrophoresis, the gels were dried and exposed. Quantification of the free and complexed protein–DNA complexes was performed by scanning densitometry. The relative intensity of the autoradiographed complexes was plotted against time.

Cell culture and co-transfections

To test for the effect of Dlx3 binding on transcription, a synthetic oligonucleotide containing three tandem copies of the restricted consensus binding site (AAG CTT GCG ATA ATT GCG GCG ATA ATT GCG GCG ATA ATT GCG AAA GCTT) was inserted into the *Hin*dIII site of the reporter plasmid ΔTKCAT (17). This reporter construct comprises the HSV thymidine kinase proximal promoter region from –50 to +51 driving a CAT reporter cassette (18). As a control, an identical construct was prepared in which the central TAATT of each Dlx3 site was changed to TGGCC. This mutated site does not bind to Dlx3 protein *in vitro* (data not shown). HeLa cells (ATCC, Rockville, MD) were grown following recommended procedures. Transient transfections were performed according to Jang *et al.* (19). The transfections were done in duplicate using Lipofectin (Life Technologies Inc., Gaithersburg, MD), following the manufacturer's instructions. Normally, 2–3 × 10^5 cells were plated in 6-well culture plates

16–20 h prior to transfection. A constant amount (1 µg) of each reporter construct was mixed with variable amounts (0.0, 0.01, 0.025, 0.05 and 0.1 µg) of pRSV-Dlx3, which is an expression plasmid comprising the Dlx3 coding sequence driven by the RSV LTR (Invitrogen, San Diego, CA). The DNA mixtures were combined with Lipofectin and incubated at room temperature for 20 min. The Lipofectin/DNA mixture was then added to the cells and incubated for 4 h. At the end of the transfection period the mixture was replaced by medium and the cells were harvested after 50–60 h of culture at 37°C. Cellular extracts were prepared according to Pothier *et al*. (20). Aliquots were used for CAT assays (21) and total protein quantification. Cellular extracts of untransfected cells were used as negative controls. The values of CAT activity were normalized to protein content in the extracts and expressed as d.p.m./protein. The relative CAT values are averages of at least two independent experiments, each with duplicate samples.

Deletion analysis of Dlx3

The open reading frame of Dlx3 was subdivided into regularly spaced regions and 30mer primers for PCR were synthesized including an *Eco*RI restriction site on the upstream primer and a *Bam*HI site on the downstream primer. Specific primers were as follows: (i) GAG GAG GAA TTC ATG AGT GGC CCC TAT GAG AAG; (ii) GAG GAG GAA TTC TCT ACT GGG CAG CAC GAC TTC; (iii) GAG GAG GAA TTC GGG TAT CGG CCG TTT GTC CAT; (iv) GAG GAG GAA TTC CGG AAG CCA AGA ACC ATC TAC; (v) GAG GAG GGA TCC ACT GTG CTC CAT ATC AGG ACC; (vi) GAG GAG GGA TCC GGT TCG GCT TCC ACT ATT GTC; (vii) GAG GAG GGA TCC ATG TTG CCC AGA TTG GTT CTG; (viii) GAG GAG GGA TCC GCG ATA CAC TGT ATC GGG AGG AGG. Individual constructs (Fig. 6) were prepared using the following primer pairs: ORF1–277, a+h; ΔN43–277, b+h; ΔN93–277, c+h; ΔN130–277, d+h; ΔC1–261, a+g; ΔC1–225, a+f; ΔC1–200, a+e. Twenty-five cycles of PCR using 100 ng of each primer corresponding to the selected region of Dlx3 were amplified with Deep Vent polymerase (New England BioLabs, Beverly, MA), 50 µM unlabeled deoxyribonucleotide triphosphates and 1× ThermoPol Buffer in a final volume of 50 µl. Reaction times and temperatures for each cycle were 94°C 30 s, 56°C 60 s, 72°C 45 s. The DNA product was electrophoresed on an agarose gel, purified with Geneclean silica matrix (Bio101, La Jolla, CA) and digested with *Eco*RI and *Bam*HI.

The Dlx3 fragments were cloned into a modified version of the plasmid vector pCS2+ (22) which provided a common translational initiation site, as well as a Myc epitope tag at the N-terminus. These plasmids were linearized with *Asp*718 and transcribed *in vitro* using a mMessage Machine kit from Ambion Inc. (Austin, TX). RNA concentrations were determined spectrophotometrically and confirmed using denaturing agarose gel electrophoresis. All constructs were verified by DNA sequencing.

Eggs were stripped from female *Xenopus* primed the previous day with 300–600 IU of human chorionic gonadotrophin (US Biochemical, Cleveland, OH). These were fertilized with macerated testes and dejellied by treatment with 2% cysteine, pH 7.8, for 3–4 min. A mixture of either 15 or 150 pg of synthetic Dlx3 mRNA, 500 pg of the Dlx3-CAT plasmid (above) and 50 pg of a TK-luciferase control plasmid (23) in a total of 10 nl of water was injected into the animal pole region of one to two cell stage embryos.

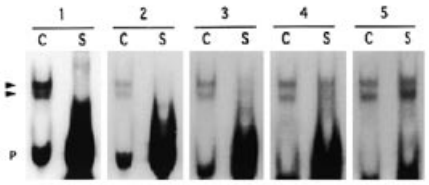

Figure 1. Mobility shift analysis of oligonucleotides binding to recombinant Dlx3. Five cycles of purification are shown (1–5). C, control Dlx3 binding site probe (profilaggrin gene promoter); S, SELEX product; P, unbound probe. The specific DNA–protein complexes are indicated by arrows.

These were allowed to develop to the early gastrula stage (10.5) and then homogenized in Reporter Lysis Buffer (Pt no. E397A; Promega Inc., Madison, WI). Homogenates were centrifuged to remove yolk platelets and the supernatants extracted once with 1,1,2-trichlorotrifluoroethane (Freon) to remove lipid. The clarified supernatants were assayed for CAT and luciferase activity (21,23).

RESULTS

Identification of Dlx3 binding sites

The SELEX procedure (14,15) was used to identify sequences capable of binding to recombinant *Xenopus* Dlx3 protein (Fig. 1). A previously identified Dlx3 binding site in a human profilaggrin gene (10) was used as a positive control. The presence of a doublet band in the mobility shifts was not observed equally in all preparations (data not shown) and probably reflects partial protein degradation. Five cycles of binding and amplification resulted in the generation of a highly enriched subpopulation of DNA binding sites, which were cloned and sequenced.

A total of 30 candidate binding sites were identified by this procedure (Fig. 2A), containing the core element TAAT that has been shown by previous studies to be generally required for homeodomain binding (2,3). A small number of oligonucleotides lacking this core failed to bind Dlx3 protein (data not shown) and were assumed to be contaminants. The core TAAT element was used to align individual sites for comparison. In several cases a TAAT was present on both strands and one of the cloned oligonucleotides contained two TAAT elements on one strand, which generated ambiguity in the alignment. In these cases all possible alignments were included and are indicated in Figure 2A. A subset of sites was compiled, in which the only clones included were those in which there was no ambiguity, i.e. only a single TAAT core was present. This smaller group of 12 sites yielded an 8 base ‘restricted’ consensus of (A/C/G)TAATT(G/A)(C/G) (Fig. 2B). This was essentially identical to the consensus derived from the unfiltered set of candidate sites and was used to design the probe for subsequent mobility shift assays.

DNA binding assays

To confirm the validity of the Dlx3 binding site established by SELEX analysis and determine relative binding affinities, mobility shift assays were carried out using individual sites from the compilation as unlabeled competitors. An oligonucleotide conforming to the consensus shown in Figure 2B (ATAATTGCT), flanked with four G residues for thermal stability, was radiolabeled and used as probe. A selection of

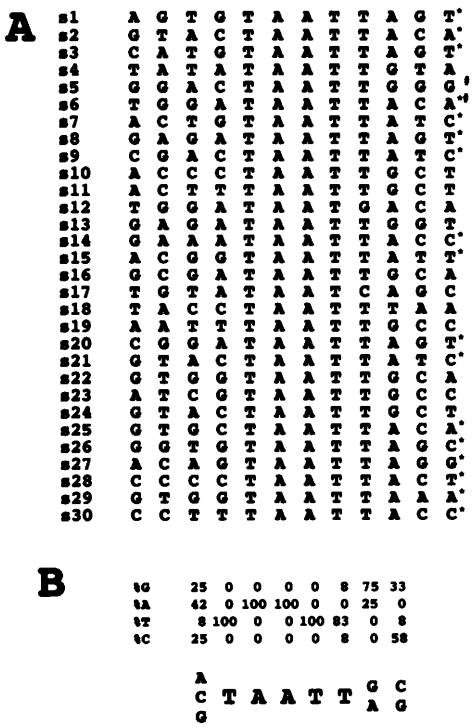

Figure 2. Compilation of TAAT-containing sites from individual SELEX clones. (**A**) A total of 30 potential sites identified (s1–s30). *, sites from clones with TAAT core elements on both strands; #, sites from clones containing two inserts. Only 12 of the 14 degenerate positions are shown, since the outer bases exhibited no significant preferences. (**B**) Consensus derived from clones in which only one TAAT core was present.

results from these competition binding experiments is shown in Figure 3. SELEX sites 5, 6, 8, 1, 14 and 22 bound approximately as well as the consensus, as judged by the elimination of most radiolabeled probe in the shifted complex at a competitor:probe ratio of 20:1 (third lanes). Other sites bound with significantly lower affinity, for example, S9 (CGACTAATTATC) and S4 (TATATAATTGTA). At the lower end of the scale were sites such as S27 (ACAGTAATTAGG), which competed poorly compared with the other sites. Figure 3 also shows the results of competition experiments using the consensus binding site for the related Msx homeodomain factors (3) and the previously reported element in the profilaggrin gene (10). Both of these sites bound well to the probe oligonucleotide, compared with the optimized consensus site (Fig. 3, first panel).

The results of these assays can be summarized as follows. A TAATT core is essential for binding. The flanking bases are degenerate, but there is a strong preference for a G residue at position –2 and either an A or C at –1 (numbers relative to the central TAATT box). There is also a strong preference for a purine at position +1 and either a G or C at +2. The presence of a high affinity Dlx3 binding site (G-A/C-TAATT-A/G-G/C) near a gene that is expressed in cells where Dlx3 is also expressed is thus a valid criterion for a potential regulatory association with this homeoprotein. The recombinant Dlx3 protein used in this experiment appears to bind with similar affinity to the SELEX consensus site and the published optimal binding site for Msx1, suggesting significant overlap in the target genes for these two homeoproteins *in vivo*. Another important aspect of these data is the degeneracy of the consensus site. A simple calculation that factors the redundant and invariant positions suggests that there should be on the order of 100 000 potential binding sites for Dlx3 in a typical mammalian genome, implying that other factors are likely to play an important role in determining which genes are actually regulated by Dlx3 *in vivo*.

Determination of binding constants

The equilibrium dissociation constant (K_d) was determined by binding saturation experiments, in which the amount of restricted consensus binding site DNA was increased in the presence of a fixed amount of protein. DNA–protein complex formation was monitored by mobility shift and data were evaluated by Scatchard analysis. As calculated from the data shown in Figure 4A, Dlx3 binds to the target DNA with a K_d of 7.8×10^{-8} M. In addition, the slope of the Scatchard plot, 1.15, suggests that Dlx3 binds to DNA as a monomer under these conditions, i.e. *in vitro* using purified recombinant protein. The half-life of the complex between Dlx3 and the target DNA was very short, with 50% dissociation taking place within 15 s from the addition of excess unlabeled competitor DNA (Fig. 4B).

Co-transfections

To determine the effects *in vivo* of Dlx3 association with a proximal promoter region via a high affinity binding site, co-transfection experiments were performed. HeLa cells were transfected with a mixture of two plasmids, one an expression construct in which the Dlx3 coding sequence was driven by an

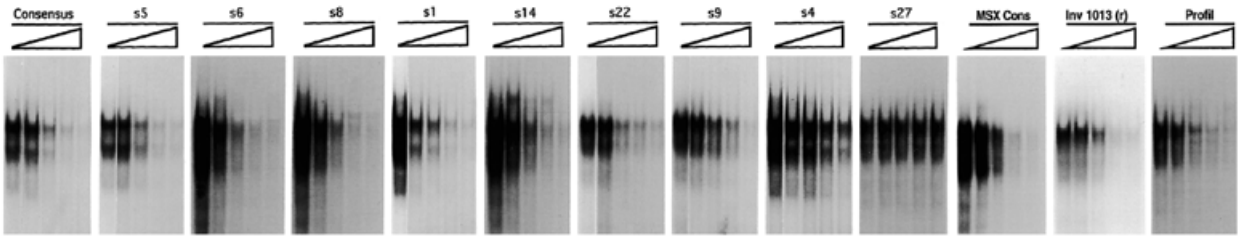

Figure 3. Competition assays. Mobility shifts using a probe corresponding to the restricted consensus binding site for Dlx3 protein and competitor oligonucleotides. For each experiment, 0, 5, 20, 50 and 100 ng competitors were used, loaded from left to right in each panel.

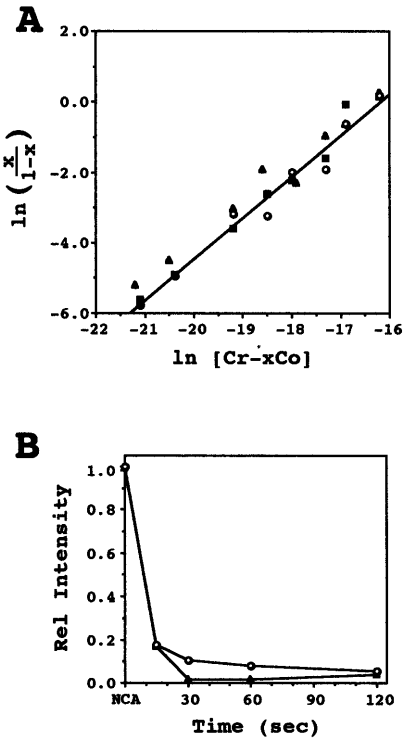

Figure 4. DNA binding properties of Dlx3. (**A**) Scatchard analysis. Mobility shift assays of protein–DNA interaction between recombinant Dlx3 and the consensus binding site, using increasing ratios of protein to DNA. Data from three independent experiments are shown (triangles, circles and squares). The line is a simple fit (Cricket Graph; Cricket Software, Malvern, PA) and the calculated *y*-intercept, slope and R value are 18.9, 1.15 and 0.981, respectively. (**B**) Off-rate determination. Recombinant Dlx3 complexed with labeled consensus probe was mixed with excess unlabeled probe and assayed by mobility shift (Materials and Methods). Percentage remaining bound is plotted versus time in seconds after addition of competitor. NCA, no competitor added. Data from two independent experiments are shown (triangles and circles).

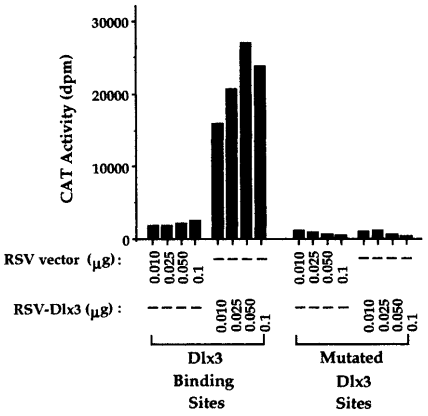

Figure 5. Co-transfection experiments. Dlx3-CAT or Mutated-CAT plasmids (1.0 μg) were transfected into HeLa cells with the indicated amount of RSV-Dlx3 or with RSV vector DNA. Total DNA for each transfection was equalized by adding an appropriate amount of pCAT-Basic plasmid (Promega Inc). CAT activities are presented as d.p.m. of ^{3}H normalized by protein concentration in the extracts. Data are an average of two independent experiments, each with duplicate samples.

RSV promoter (RSV-Dlx3) and the other a reporter, the CAT coding sequence driven by a minimal promoter derived from the herpes simplex virus thymidine kinase gene (18), in which three tandem Dlx3 binding sites had been inserted 50 bp upstream from the start of transcription (Dlx3-CAT; Materials and Methods). Variable ratios of RSV-Dlx3 plasmid to reporter plasmid were used. In addition, control experiments were carried out in parallel in which the Dlx3 expression plasmid was replaced by the RSV vector and the Dlx3 binding sites were replaced with mutated, non-binding sites in which the core TAATT was replaced by TGGCC.

The results of these experiments are shown in Figure 5. The combination of co-transfected Dlx3 with the reporter plasmid carrying high affinity Dlx3 binding sites (Dlx3CAT) increased expression dramatically. The extent of stimulation depended on the relative amounts of the two plasmids, with a peak observed at a ratio of 0.05 μg RSV-Dlx3 to 1.0 μg CAT reporter construct. Co-transfection of RSV-Dlx3 and Dlx3-CAT yielded up to a 12.3-fold higher signal than RSV vector combined with Dlx3-CAT reporter plasmid (27 000 versus 2190 d.p.m. ^{14}C/unit total protein). Comparing Dlx3-CAT with mutated-CAT, both co-transfected with RSV-Dlx3, gives a larger stimulation, 44-fold at the maximum (27 000 versus 614 d.p.m./protein). The latter comparison may be more valid, since the Dlx3 binding site has the potential to interact with other homeodomain proteins that might be expressed in HeLa cells, some of which could be transcriptional activators. The mutated site is unlikely to bind to any homeodomain factors, although unexpected interaction with other activator or repressor proteins cannot be excluded. In any case, it is clear that in this context, Dlx3 acts as a potent activator of transcription.

Dlx3 RNA is not detectable by northern blot in HeLa cells (M.I.Morasso, unpublished results) and thus it is possible that different results might be obtained in a more physiologically meaningful cell type. Dlx3 is expressed in *Xenopus* development beginning at gastrulation, in the presumptive epidermis. Injection of plasmid DNA into *Xenopus* embryos has been used to identify regulatory elements (24–26) and provides a convenient method for evaluating the function of transcription factors in a living vertebrate embryo, where Dlx3 is naturally expressed. We used this approach to map the region of Dlx3 responsible for target gene activation.

As shown in Figure 6, removal of the sequence encoding the first 43 residues of Dlx3 (ΔN43–277) had only a modest effect on *trans*-activation, suggesting this region is largely dispensable for this function. However, further deletion to residue 93 (ΔN93–277) resulted in reduction of activation to a level apparently even lower than the controls (β-galactosidase or no RNA). Likewise, deletion of the C-terminal 16 residues (ΔC1–261) had little effect, whereas removal of an additional 36 amino acids (ΔC1–225) significantly reduced the activation level. Deletion to near the homeodomain (ΔC1–200) resulted in a loss of activation to below that of controls, similar to the ΔN130–277 deletion. None of these effects were due to differences in protein expression, since all truncated proteins accumulated *in vivo* to levels falling within an ~5-fold range as revealed by western blot (data not shown) and essentially identical results were obtained with both 15 and 150 pg Dlx3 RNA doses. These results suggest that transcriptional activation by Dlx3 depends on a pair of domains separated by the homeobox region. Removal of either of these subdomains essentially

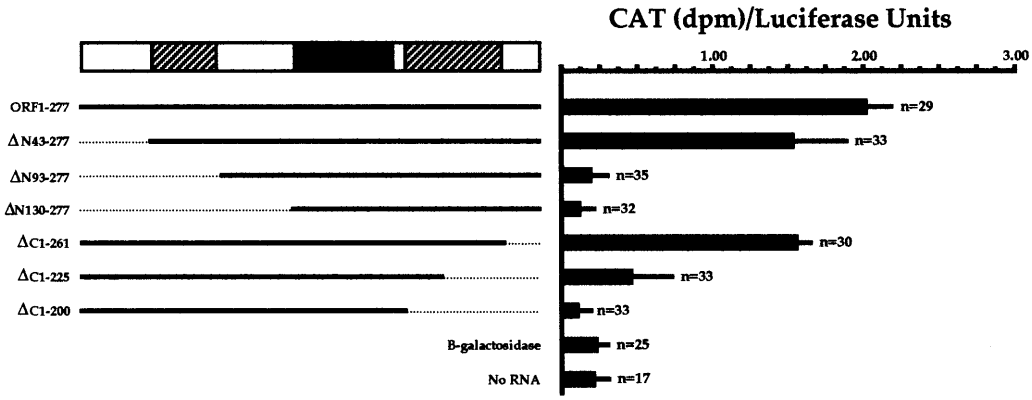

Figure 6. *Xenopus* embryo expression assays. Six different deletions of Dlx3 were prepared, with sequences removed from the N- or C-terminus as described in Materials and Methods. The amino acid numbers shown on the left indicate the portion of the protein remaining, which is also represented by the solid bars. The filled bars on the right indicate the ratio, in arbitrary units, of CAT activity (from the Dlx3-CAT plasmid) to luciferase activity (from the control, co-injected TK-luciferase plasmid). Each set of data was derived from at least two independent experiments, using both 15 and 150 pg of Dlx3 RNA, which gave similar results. The total number (*n*) of embryos in each assay is shown to the right of each bar. The extended bars indicate half of the range in values from individual experiments. The results at the bottom of the figure correspond to embryos injected with β-galactosidase RNA instead of Dlx3 RNA or with the two plasmid DNAs alone, without any co-injected RNA. The diagram at upper left indicates the position of the homeodomain (gray) and two activation subdomains (hash marks) in the Dlx3 protein. The N-terminal activation domain corresponds to residues 43–93 and the C-terminal domain to residues 200–261.

inactivates the Dlx3 protein and may actually result in transcriptional repression for the shortest variants.

DISCUSSION

The expression pattern, transcriptional regulation and to a lesser, but significant, extent the protein sequence of Dlx3 are conserved in vertebrate phylogeny (7,12,27). This conservation suggests that Dlx3 plays an important role in the development of higher organisms, just as the single Dll gene is critical in *Drosophila* ontogeny. There are convincing experimental data supporting this hypothesis: loss of Dlx3 function by mutation results in an inherited disease, TDO, in humans (11) and targeted disruption of the mouse Dlx3 gene has an early embryonic lethal phenotype due to placental failure (28). Furthermore, ectopic expression of Dlx3 in murine epidermis leads to several major defects, including termination of mitotic activity and precocious expression of differentiation markers, resulting in perinatal lethality (10). Characterizing the molecular and biochemical properties of Dlx3 is an essential prerequisite for understanding the role this protein plays in regulating tissue differentiation.

The finding that Dlx3 is a transcriptional activator implies the existence of target genes that are up-regulated in response to the presence of Dlx3. In the experiments reported here, the *Xenopus* protein has been used, but it is likely that Dlx3 cognates in other species will behave similarly to *Xenopus* Dlx3. The N- and C-terminal activation domains have a slightly acidic bias, both in *Xenopus* and in the corresponding regions of the mouse Dlx3 sequence and there is significant conservation of protein sequence in these regions. However, the *Xenopus* protein has a poly-glutamine tract in the C-terminal domain that is not present in the mouse gene. Certain highly conserved elements, in addition to the homeodomain, have been noted in the Dlx family (7). Some of these conserved subdomains appear to be dispensable for *trans*-activation; for example, there is a highly conserved peptide located downstream from residue 266 which can be deleted without affecting activation. It may be that some of these conserved domains are important in protein–protein interactions that modulate Dlx3 function, but are not responsible for target gene activation *per se*.

The functional map of Dlx3 provides some useful insights into the molecular basis of the TDO syndrome. A genetic lesion responsible for this disease, which is inherited as an autosomal dominant trait, has been recently identified as a 4 nt deletion immediately downstream from the DLX3 homeodomain (11). Assuming the human and *Xenopus* Dlx3 proteins have similar properties, our results predict that the truncated Dlx3 protein in TDO patients has lost the ability to activate target gene expression. In addition, if important interactions between Dlx3 and other regulatory factors are mediated by the N-terminal portion of the protein, the TDO deletion polypeptide might also result in the formation of heteromeric complexes that are unable to properly activate target gene transcription. The dominant nature of the known TDO alleles would thus be due to a molecular interference mechanism, as opposed to a simple haploinsufficiency of Dlx3 gene product in heterozygous individuals with the disease. This interpretation is also consistent with the observation that targeted deletion of the murine Dlx3, which removes most of the protein coding sequences, does not lead to a detectable phenotype in the heterozygous state (28).

The homeodomain gene family most closely related to the Dlx genes encodes the Msx homeoproteins, which are vertebrate homologs of the *Drosophila* gene Msh (29,30). The optimal binding site for Msx1 is very similar to the Dlx3 binding consensus (3) and, as shown in Figure 3, Dlx3 is able to bind quite well to a site conforming to the Msx1 binding consensus. The transcriptional effects of Msx1 and Dlx3 are very different. Whereas Dlx3 functions as an activator protein, Msx1 acts as a transcriptional repressor (31–33). This negative effect does not require binding of Msx1 to DNA but rather seems to involve a direct interaction between the Msx1 homeodomain and the TATA binding protein (31,32). Other members of the Dlx family (Dlx1, 2 and 5) have been shown to interact directly with Msx proteins, both *in vitro* and *in vivo*, and to have antagonistic effects on target gene transcription (33). These results, taken together with those

presented here, suggest that the entire Dlx family may act as positive transcription factors, in opposition to the repressive effects of members of the Msx family. When members of each family are co-expressed in individual tissues, the balance of Dlx and Msx protein levels could determine the cellular phenotype both by direct protein–protein interaction and by competition for regulatory target sites in the genome.

The activation of a model target gene, Dlx3-CAT, in co-transfection assays provides an explanation for the effect observed on the profilaggrin gene in basal keratinocytes when Dlx3 is ectopically expressed in these cells in transgenic mice (10). In these animals, profilaggrin transcripts accumulate in basal cells, much earlier than the normal time of activation of this gene, which occurs late in suprabasal keratinocyte differentiation. There is a potential binding site for Dlx3 in the proximal promoter region of the profilaggrin gene (10), suggesting the possibility that Dlx3 interacts directly with the regulatory region of this gene and activates transcription. This would represent an unusually direct relationship between a homeoprotein and a tissue-specific structural gene.

As an activator, Dlx3 appears to be quite potent, resulting in up to 40-fold stimulation of the reporter construct in HeLa cells, with a similar effect in *Xenopus* embryos which express Dlx3 normally, starting at about the time when the injected embryos were harvested for analysis. The DNA binding data presented here are consistent with the interpretation that Dlx3 engages DNA as a monomer, but this is with purified protein *in vitro* and does not preclude the formation of homodimeric or heterodimeric forms *in vivo*. In fact, Zhang *et al*. (33) have shown that Dlx2 and Dlx5 can both form homodimers *in vitro*, as well as heterodimers within the Dlx/Msx families.

The mechanisms by which homeodomain proteins regulate specific subsets of target genes is an important problem in biology. The small and degenerate binding sites associated with homeoproteins cannot provide sufficient specificity to explain the regulatory properties of these genes in development, and interactions between homeoproteins and between homeoproteins and other transcription factors probably supply this additional level of specificity. It will be interesting to determine how Dlx3 interacts *in vivo* with other factors, such as Msx homeoproteins, to mediate regulatory functions in development and differentiation.

ACKNOWLEDGEMENTS

We thank Dr Nedialka Markova for her help in designing and preparing the constructs for co-transfection experiments and Dr Jeff Franklin for helpful suggestions and advice on the initial phase of this work.

REFERENCES

1 McGinnis,W. and Krumlauf,R. (1992) *Cell*, **68**, 283–302.
2 Gehring,W.J., Qian,Y.Q., Billeter,M., Furokubo-Todunaga,K., Schier,A.F., Resendez-Perez,D., Affolter,M., Otting,G. and Wuthrich,K. (1994) *Cell*, **78**, 212–223.
3 Catron,K.M., Iler,N. and Abate,C. (1993) *Mol. Cell. Biol.*, **13**, 2354–2365.
4 Gorfinkel,N., Morata,G. and Guerrero,I. (1997) *Genes Dev.*, **11**, 2259–2271.
5 Simeone,A., Acampora,D., Pannese,M., D'Esposito,M., Stornaiuolo,A., Gulisano,M., Mallamaci,A., Kastury,K., Druck,T., Huebner,K. and Boncinelli,E. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 2250–2254.
6 Stock,D.W., Ellies,D.L., Zhao,Z., Ekker,M., Ruddle,F.H. and Weiss,K.M. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 10858–10863.
7 Robinson,G.W. and Mahon,K. (1994) *Mech. Dev.*, **48**, 199–215.
8 Qiu,M., Bulfone,A., Ghattas,I., Meneses,J.J., Christensen,L., Sharpe,P.T., Presley,R., Pederson,R.A. and Rubenstein,J.L. (1997) *Dev. Biol.*, **185**, 165–184.
9 Morasso,M.I., Jamrich,M. and Sargent,T.D. (1994) *Dev. Biol.*, **162**, 267–276.
10 Morasso,M.I., Markova,N.G. and Sargent,T.D. (1996) *J. Cell Biol.*, **135**, 1879–1887.
11 Price,J.A., Bowden,D.W., Wright,J.T., Pettenati,M.J. and Hart,T.C. (1998) *Hum. Mol. Genet.*, **7**, 563–569.
12 Dirksen,M.L., Morasso,M.I., Sargent,T.D. and Jamrich,M. (1994) *Mech. Dev.*, **46**, 63–70.
13 Papalopulu,N. and Kintner,C. (1994) *Development*, **117**, 961–975.
14 Tuerk,C. and Gold,L. (1990) *Science*, **249**, 505–510.
15 Cui,Y., Wang,Q., Stormo,G. and Calvo,J. (1995) *J. Bacteriol.*, **177**, 4872–4880.
16 Chadwick,P., Pirrotta,V., Steinberg,R., Hopkins,N. and Ptashne,M. (1970) *Cold Spring Harbor Symp. Quant. Biol.*, **35**, 283–294.
17 Lee,J.-H., Jang,S., Yang,J.-M., Markova,N.G. and Steinert,P.M. (1996) *J. Biol. Chem.*, **271**, 4561–4568.
18 Wagner,M.J, Sharp,J.A. and Summers,W.C. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 1441–1445.
19 Jang,S., Steinert,P.M. and Markova,N.G. (1996) *J. Biol. Chem.*, **271**, 24105–24114.
20 Pothier,F., Ouellet,M., Julien,J.P. and Guerin,S.L. (1992) *DNA Cell Biol.*, **11**, 83–90.
21 Neumann,J.R., Morency,C.A. and Russian,K.O. (1987) *Biotechniques*, **5**, 444–447.
22 Turner,D.L. and Weintraub,H. (1994) *Genes Dev.*, **8**, 1434–1447.
23 Minucci,S., Zand,D.J., Dey,A., Marks,M.S., Nagata,T., Grippoj,J.F. and Ozato,K. (1994) *Mol. Cell. Biol.*, **14**, 360–372.
24 Krieg,P.A. and Melton,D.A. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 2331–2335.
25 Jonas,E.A., Snape,A.M. and Sargent,T.D. (1989) *Development*, **106**, 399–405.
26 Rebbert,M.L. and Dawid,I.B. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 9717–9722.
27 Morasso,M.I., Mahon,K.A. and Sargent,T.D. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 3968–3972.
28 Morasso,M.I., Grinberg,A., Robinson,G., Sargent,T.D. and Mahon,K.A. (1999) *Proc. Natl Acad. Sci. USA*, **96**, 162–167.
29 Robert,B., Sassoon,D., Jacq,B., Gehring,W. and Buckingham,M. (1989) *EMBO J.*, **8**, 91–100.
30 Davison,D. (1995) *Trends Genet.*, **11**, 405–411.
31 Zhang,H., Catron,K.M. and Abate-Shen,C. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 1764–1769.
32 Catron,K.M., Wang,H., Hu,G., Shen,M.M. and Abate-Shen,C. (1996) *Mech. Dev.*, **55**, 185–199.
33 Zhang,H., Hu,G., Wang,H., Sciavolino,P., Iler,N., Shen,M.M. and Abate-Shen,C. (1997) *Mol. Cell. Biol.*, **17**, 2920–2932.