A proline-rich transcriptional activation domain in murine HOXD-4 (HOX-4.2)

Isabel Rambaldi¹, Erzsebet Nagy Kovacs¹ and Mark S. Featherstone¹,²,*
¹McGill Cancer Centre and ²Departments of Oncology and Medicine, McGill University, 3655 Drummond Street, Montréal, Québec H3G 1Y6, Canada

Received October 4, 1993; Revised and Accepted December 21, 1993

ABSTRACT

The product of the murine Hoxd-4 (Hox-4.2) gene is a transcription factor that acts upon an autoregulatory element in Hoxd-4 upstream sequences (1). Using this activity as an assay in transient transfections of P19 embryonal carcinoma (EC) cells, we performed a mutational analysis to map functional domains in the HOXD-4 protein. The importance of the homeodomain was shown by a single amino acid change in this region that abolished activity. Deletion analysis revealed that many evolutionarily conserved regions outside of the homeodomain were dispensable for activation in our assay. Fusions to the GAL4 DNA-binding domain mapped a transcriptional activation function to the HOXD-4 proline-rich N-terminus. The proline-rich transcription factor AP2 squelched activation by HOXD-4 and by GAL4/HOXD-4 N-terminus fusion proteins. Together, these results suggest that HOXD-4 harbors a transcriptional activation domain of the proline-rich type.

INTRODUCTION

The homeotic genes of insects and the Hox genes of vertebrates are evolutionarily related families represented over a broad range of animal species (2). Homeotic and Hox genes exert their effects through the transcriptional regulation of target genes, few of which have been identified. The best characterized example is the Drosophila homeotic gene Deformed (Dfd). The Dfd protein has been shown to positively regulate the expression of its own gene (3). A minimal Dfd autoregulatory enhancer has four high affinity Dfd binding sites (4).

DNA-binding is mediated by the 61 residue homeodomain encoded by the highly conserved homeobox (5–7) and found in the C-terminal half of the protein. Apart from the homeodomain, other regions of the Hox and homeotic gene products are also conserved. Most notable is the so-called conserved pentapeptide (8, 9), a domain found just N-terminal to the homeodomain of many Antp class homeoproteins. Regions of some conservation are also found immediately C-terminal to the homeodomain and at the extreme N-terminus (2). The extent of homology is greatest amongst members of the same subfamily. Other regions can be identified based on amino acid content or character, rather than percent identity. For example, homeoproteins of the Dfd subfamily have a proline-rich region in the N-terminal half of the molecule (10–12). The C-terminus of the mouse HOXD-4 protein is very serine-rich (10). Determining the roles of these regions is an important approach to understanding differential function among the Hox gene products.

Via transient transfections of P19 embryonal carcinoma (EC) cells, we have recently identified an autoregulatory element of the murine Hoxd-4 gene (1). This activity provided us with an assay by which to assess the functional contributions of regions within the HOXD-4 homeoprotein to transcriptional activation. In a complementary test, HOXD-4 residues were fused to the GAL4 DNA-binding domain to map transcriptional activation functions. Using these approaches, we have shown that many regions of HOXD-4 are dispensable for activation of the Hoxd-4 promoter in our assay. The proline-rich N-terminus of HOXD-4 is a strong transcriptional activation domain.

MATERIALS AND METHODS

Plasmids

We use the recently adopted Hox nomenclature (13). Old Hox names are given in brackets on first use. Two sequencing errors and an error in the Hoxd-4 cDNA were included in the previously published sequence of Hoxd-4 (10). The corrected sequence results in the following changes: G39 to S39, R95 to P95, and P192 to S192 (Fig. 1). Note that with S192 the sequence of the HOXD-4 homeodomain now shares 100% identity with that of human HOXD4 (9). The sequence of Hoxd-4 in the GenBank database has been revised appropriately (accession number J03770).

Hoxd-4 expression vectors were based on the murine phosphoglycerate kinase-1 (pgk-1) promoter and 3' processing sequences (14) which were a kind gift of Dr Michael McBurney. Hoxd-4 coding region sequences were obtained from cDNA and genomic clones (10) and inserted between the pgk-1 elements in pBluescribe (Stratagene). The unique Xho I, Bam HI, and Xba I sites (Fig. 2) were engineered via PCR-mediated mutagenesis and standard recombinant DNA techniques.

* To whom correspondence should be addressed at: McGill Cancer Center, Room 714, 3655 Drummond Street, Montréal, Quebec H3G 1Y6, Canada
Sequences naturally flanking the ATG of the third codon conform to a Kozak consensus (15) whereas those for the first codon do not (10), suggesting that translation of Hoxd-4 transcripts begins with the methionine at position 3 (Fig. 1). Creation of the Bam HI site on the methionine restriction at a Bal I site and removed approximately 300bp of leader sequence and the upstream initiation codon (M1) (Fig. 1). The Kozak consensus at the initiation codon for M3 was lost in this process. A Kozak consensus sequence was restored by site-directed mutagenesis to give the expression vector p4.2 (Table 1). p4.2-transfected cells synthesize a product recognized by polyclonal antibodies directed against the HOXD-4 N- and C-termini (Fig. 3). The apparent size of this protein on Western blots is larger than expected, also noted for the products of the chicken Hoxd-4 and Hoxa-4 homologs which are also proline-rich (16).

Derivatives of p4.2 mutated in the Hoxd-4 coding region were generated by oligonucleotide-directed mutagenesis on single-stranded DNA (deletions and substitutions) and by standard recombinant DNA techniques (Table 1). The procedure of Kunkel et al. (17) was used for site-directed mutagenesis. All mutants were verified by sequencing of double-stranded DNA by the chain termination method (18). Mutant vectors gave products of appropriate size in Western blot analysis.

The construction of the luciferase reporter psXAP2 under the control of the Hoxd-4 promoter/autoregulatory element is described elsewhere (1). An expression vector for the GAL4 DNA binding domain (residues 1 to 148) was made from the vector pSG5Gal-VF16 (19) (kind gift of Dr John White) by enzymatic removal of VP16 coding sequences. Hoxd-4 coding sequences were introduced in frame with the Gal4 coding region. The GAL4-responsive luciferase reporter pXPAL7 was made by PCR amplification of the promoter region of the vector pAL7 (20) (kind gift of Dr John White) and transfer to the luciferase vector pXP1 (21). Luciferase reporter p34a was made by H.Popperl (unpublished). All constructs were verified by sequence analysis. The AP2 expression vector sp(RSV)AP2 was a kind gift of Dr Trevor Williams (22).

Cell culture, transfections and enzymatic assays
P19 EC cells were cultured and calcium phosphate transfection assays performed as described previously (1). Transfections of differentiating cells were done 48 hours after retinoic acid (RA) addition to 3×10⁻⁷M. For assays on the Hoxd-4 promoter, each dish received 1µg of expression vector, 4µg of luciferase reporter psXAP2, and 1µg of pRSVlacZ as an internal standard. For assays with GAL4 fusions, the amount of expression vector was reduced to 100ng. Precipitates for each 100mm dish were brought to 10µg total DNA with supercoiled pBluescribe (Stratagene). All luciferase values were normalized to those for β-galactosidase.

Antibody production and Western analysis
Fusions of the N-terminal half or C-terminal half of HOXD-4 to the amino terminus of trpE were accomplished through the pATH2 and pATH3 vectors, respectively (23). Immunizations were according to Harlow and Lane (24). Approximately 125µg

<table>
<thead>
<tr>
<th>Vector</th>
<th>Manipulation</th>
<th>Nucleotide sequence across site of mutation</th>
<th>Protein sequence across junction</th>
<th>Residues altered</th>
</tr>
</thead>
<tbody>
<tr>
<td>p4.2</td>
<td>deletion, insertion</td>
<td>J: GACCTGCAAGATCCA/CCATGATTCTAT</td>
<td>T: MPPSYM 7</td>
<td>1 and 2</td>
</tr>
<tr>
<td>pd1B</td>
<td>deletion</td>
<td>O: TCTGATATAGTTGGAATGTCAGGGTGTCAT</td>
<td>M3: MPPSYM/LQGGY 28</td>
<td>10 to 23</td>
</tr>
<tr>
<td>pd1A</td>
<td>deletion</td>
<td>O: ACAAGGCTCAATGCATGTCAGGGTGCTAC</td>
<td>M3: MPPSYM/LQGGY 28</td>
<td>4 to 23</td>
</tr>
<tr>
<td>pd1D</td>
<td>Acc U/Eco RI deletion and adaptor insertion</td>
<td>O: CT/TC/CACGTT</td>
<td>128 PAVV/HVNS 142</td>
<td>132 to 138</td>
</tr>
<tr>
<td>pd1G</td>
<td>deletion</td>
<td>O: AATTACCAGT/ACA</td>
<td>209 KDHK/GRSS 221</td>
<td>213 to 217</td>
</tr>
<tr>
<td>pd1GH</td>
<td>deletion</td>
<td>O: AAAAAAAGACCAAA/GGCAGGTCCTTCTCT</td>
<td>209 KDHK/GRSS</td>
<td>213 to 250</td>
</tr>
<tr>
<td>pdbK201</td>
<td>substitution</td>
<td>O: CAAGATATGTGTT/GAAGAACCCGGAGGA</td>
<td>197 KIFW/ANRM 205</td>
<td>201</td>
</tr>
<tr>
<td>psbY199</td>
<td>substitution</td>
<td>J: GATCAGATCAT/AT/TCACAGACCG</td>
<td>195 KIWI/y/QNOR 203</td>
<td>199</td>
</tr>
<tr>
<td>p4.2FS</td>
<td>Bgl II, blunt, ligate</td>
<td>J: CATAGCATAG/GATCTGGTT</td>
<td>195 KIWI/dl...</td>
<td>199 onward</td>
</tr>
</tbody>
</table>

*a*, for junction, indicates the sequence across the site of mutation. 'O' designates the actual oligonucleotide used in the mutagenesis. Slashes indicate boundaries between regions juxtaposed or created as a result of the manipulations.

bSmall letters indicate amino acids not found in the wild-type HOXD-4 protein. 'x' indicates the presence of a termination codon in the corresponding transcript. Numbering as in figure 1.

cIncludes deletions and substitutions. Numbering of altered residues is inclusive.
of each induced fusion protein were electrophoresed from SDS/polyacrylamide gels, lyophilized, resuspended in 500μl water and emulsified with 500μl of complete Freund’s adjuvant. Rabbits were injected subcutaneously at 4 sites each, 250μl per site. Boosts were given at 4 week intervals using incomplete Freund’s adjuvant. Serum was obtained at the same time by ear bleeding (25ml) or by cardiac puncture 4 weeks after the third boost. Western blots were performed with Biorad minigel and minitransfer modules and a Biorad immunoblotting kit. A 1:100 dilution of the primary antibody was used with alkaline phosphatase-conjugated goat antirabbit second antibody. The availability of antisera independently directed against the HOXD-4 N- and C-termini allowed us to assess the levels of proteins mutated in one domain by using antisera directed against the other, thus avoiding detection problems due to epitope loss.

RESULTS
The Hoxd-4 coding region
The HOXD-4 protein can be divided into 8 regions based on homologies within and beyond the Dfd subfamily of proteins, and on the basis of amino acid composition (Fig. 1). We used these to guide a mutational analysis of HOXD-4 structure and function (Table 1). To assay HOXD-4 activity, we relied on its ability to activate transcription through an autoregulatory element in the Hoxd-4 promoter in transient transfections of P19 EC cells (1). The wild-type expression vector for HOXD-4 (p4.2) and the HOXD-4-responsive luciferase reporter under the control of the Hoxd-4 promoter (pSXΔP2) are diagrammed in Figure 2.

After transfection into pluripotent and RA-treated P19 cells, p4.2 directed the synthesis of a product reactive on Western blots with antibodies against HOXD-4 (Fig. 3). Mutated derivatives of p4.2 were also examined for the production of immunoreactive species which were detected in all but four cases (Fig. 4). One of these four was p4.2FS which had a frame shift mutation in the homeobox in the region coding for the recognition helix (Table 1; Fig. 2). Transcriptional activation by p4.2 and its mutated derivatives was expressed as the ratio of standardized luciferase activity over that obtained with p4.2FS.

Transcriptional activation by mutant HOXD-4 proteins
Site-specific binding by the homeodomain is in part mediated by the third a helix (recognition helix) which makes contacts in the major groove (25—32). A key determinant of specificity is the residue at position 50 of the homeodomain, within the recognition helix (33—35). Substitution for lysine converts the specificity of binding to that of the bicoid homeodomain (30). As for all Hox and homeotic gene products, HOXD-4 has a glutamine at this position (Q201, Fig. 2). Conversion of this glutamine to lysine abrogated activation of the Hoxd-4 promoter (sbK201, Fig. 4) and abolished binding to a HOXD-4 binding site in an electrophoretic mobility shift assay (data not shown). Nonetheless, this protein gave a band intensity comparable to wild-type on Western blots. This result underscores the importance of the homeodomain in transcriptional activation by HOXD-4. The tryptophan at position 199 of HOXD-4 is highly conserved and has been implicated in homeodomain function by genetic and structural observations. Conversion to a tyrosine residue (sbY199, Fig. 4) resulted in the loss of a detectable product on Western blots and a severe reduction in transcriptional activation.

Deletion of conserved N-terminal regions B, or A plus B (dB and dBAB, Fig. 1 and 4), did not greatly affect HOXD-4 function, though the activity of dBAB was somewhat elevated. Products bearing deletions of much of the proline-rich region C or the
Figure 4. Effect of coding region mutations on transcriptional activation by HOXD-4 in pluripotent (−RA) and RA-treated (+RA) P19 EC cells. At top, representation of the wild-type HOXD-4 product of the p4.2 expression vector, divided into regions and lettered according to Figure 2. The homeodomain (region F) is stippled. Along left, names of mutants indicating deletion (dl) or substitution (ab) or other alteration. 'A-4' designates the C-terminus of the HOXA-4 product. Single amino acid substitutions within the homeodomain at positions 199 and 201 are indicated. HOXD-4FS denotes the frame-shifted product of the p4.2FS expression vector. 'FOLD ACTIVATION' gives the ratio of luciferase activity in co-transfections with vectors encoding given HOXD-4 proteins over that obtained with the control vector p4.2FS. Activation by HOXD-4 was set to 14 (average over multiple experiments) and other values normalized accordingly. Standard deviations are shown. The number of independent experiments for a given construct is given in brackets. '+' under 'WESTERN' indicates that a given protein could be detected at wild-type levels on Western blots of transfected cell extracts using rabbit polyclonal serum directed against the HOXD-4 N- or C-terminus (see Materials and Methods). 'n.d.' not determined.

entire N-terminal half of HOXD-4 could not be detected in Western analysis. A six amino acid deletion which removed four of the five amino acids of the conserved pentapeptide (8, 9) (dID, Fig. 1 and 4) was 2.6 fold more active for transcriptional activation in pluripotent P19 cells. A five amino acid deletion of the conserved region G immediately C-terminal to the homeodomain (dIG, Fig. 1 and 4) also caused a modest increase in HOXD-4 activity. Full deletion of the C-terminus, including the serine-rich region H (dGH, Fig. 4), impaired the ability of HOXD-4 to activate the Hoxd-4 promoter. Replacement with the C-terminus of the paralogous Hoxa-4 (Hax-1.4) product (sbHA-4, Fig. 4) restored activation to wild-type levels. Results were similar in both pluripotent and RA-treated P19 cells.

To ensure that transcriptional activation by HOXD-4 derivatives was mediated by the autoregulatory enhancer in the Hoxd-4 promoter, we tested three HOXD-4 vectors on two different promoters. One was a minimal adenovirus major late promoter, while the other had, in addition, the Hoxd-4 autoregulatory enhancer (1). As shown in Table 2, wild-type HOXD-4 as well as derivatives dID and dIG were considerably more active on the promoter containing the autoregulatory element (pNBML). Consistent with results on the Hoxd-4 promoter (Fig. 4), dID is more active than wild-type. We conclude that the effects of HOXD-4 derivatives are mediated through the Hoxd-4 autoregulatory enhancer.

Table 2. Transcriptional activation by HOXD-4 derivatives is dependent on the autoregulatory element in the Hoxd-4 promoter

<table>
<thead>
<tr>
<th>HOXD-4 derivative</th>
<th>Fold activation</th>
<th>Specific increase mediated by autoregulatory enhancer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RA</td>
<td>WESTERN</td>
</tr>
<tr>
<td>dID</td>
<td>14 (10)</td>
<td>14 (11) +</td>
</tr>
<tr>
<td>dIG</td>
<td>15 : 2 (3)</td>
<td>15 : 3 (5) +</td>
</tr>
<tr>
<td>dGH</td>
<td>24 : 7 (3)</td>
<td>24 : 5 (3) +</td>
</tr>
<tr>
<td>dID</td>
<td>37 : 5 (5)</td>
<td>37 : 7 (9) +</td>
</tr>
<tr>
<td>dIG</td>
<td>24 : 1 (2)</td>
<td>24 : 5 (5) +</td>
</tr>
<tr>
<td>dGH</td>
<td>7 : 2 (5)</td>
<td>7 : 2 (3) +</td>
</tr>
<tr>
<td>sbHA-4</td>
<td>120 : 2 (2)</td>
<td>120 : 1 (2) n.d.</td>
</tr>
<tr>
<td>sbX201</td>
<td>2 : 1 (6)</td>
<td>2 : 0 (3) +</td>
</tr>
<tr>
<td>sbY109</td>
<td>3 : 2 (3)</td>
<td>3 : 0 (2) +</td>
</tr>
<tr>
<td></td>
<td>1 : 1 (10)</td>
<td>1 : 0 (11) -</td>
</tr>
</tbody>
</table>

*HOXD-4 and two derivatives, dID and dIG, were tested for specific activation through the autoregulatory enhancer in the Hoxd-4 promoter in RA-treated P19 cells. Appropriate expression vectors were tested at 1 μg/100mm plate in cotransfections with luciferase reporters pML and pNBML. pML contains the adenovirus major late promoter. In pNBML, the 217bp NruI/BsmHI fragment containing the Hoxd-4 autoregulatory element has been inserted just upstream of the adenoviral TATA box (1). The fold increase over basal promoter activity mediated by HOXD-4 expression vectors is shown.

The ratio of fold activation obtained with pNBML over pML.

Figure 5. Mapping of transcriptional activation domains in HOXD-4 by fusion to the GAL4 DNA binding domain. A. Schematic of luciferase reporter used to measure transcriptional activation by GAL4 derivatives. The promoter (20) consists of the adenovirus-2 major late TATA box and transcriptional start site (substened arrow) downstream of five copies of the 17-mer GAL4 dimer binding site (boxed open arrows). B. Transcriptional activation with fusions of the GAL4 DNA binding domain to N- and C-terminal regions of HOXD-4. GAL4 is stippled. HOXD-4 regions (open rectangles) are lettered as in Figure 4. Numbers below indicate amino acid residues as in Figure 1. C is a subregion of C as indicated. HA-4 is the HOXA-4 C-terminus. 'FOLD ACTIVATION' gives the increase in luciferase activity relative to that obtained with the GAL4 DNA binding domain (GAL4). Standard deviations and the numbers of independent experiments are shown. '+' under 'WESTERN' indicates that the protein could be readily detected in extracts from transient transfections using polyclonal antibodies directed against the HOXD-4 N-terminus. n.d. not determined.

A transcriptional activation domain in the HOXD-4 N-terminus

While the intact yeast GAL4 protein is a strong transcriptional activator of responsive promoters in eukaryotic cells, the isolated GAL4 DNA-binding domain is not. Fusion of the GAL4 DNA-binding domain to heterologous polypeptides thus provides an assay for transcriptional activation domains distinct from those involved in DNA binding (36). HOXD-4 N-terminal regions A through D were fused to the GAL4 DNA-binding domain and tested for their ability to activate transcription from a GAL4-responsive reporter in differentiating P19 cells. As shown in Figure 5, this fusion product (GAL4ABCD) was 250 times more active than GAL4 which has the GAL4 DNA-binding domain alone (507 fold vs 2 fold). This result maps one or more transcriptional activation domains to the proline-rich HOXD-4 N-terminus.
Deletion of regions A and B from HOXD-4 had no effect on transcriptional activation (Fig. 4), indicating a role for regions C or D in N-terminal function. Fusion of the GAL4 DNA-binding domain to just a portion of the proline-rich region C (GAL/C'; residues 24 to 109) also increased activation GAL/ (Fig. 5). This suggests that region C spans a transcriptional activation domain of the proline-rich variety (37).

Deletion and substitution mutants suggested a role in transcriptional activation for the HOXD-4 C-terminus and the equivalent region of the Hoxa-4 product (Fig. 4). However, fusion of the GAL4 DNA-binding domain to the HOXD-4 C-terminus (GAL/H) had no effect on transcriptional activation (Fig. 5). A similar result was obtained with the HOXA-4 C-terminus (GAL/HA-4). Coexpression of either of these products with GAL/ABCD did not result in any synergistic activation of reporter gene expression (not shown). The role of these C-terminal sequences thus remains to be determined.

The transcription factor AP2 has been shown to harbor a transcriptional activation domain of the proline-rich variety (22) and would be expected to squelch the activity of HOXD-4, if HOXD-4 also contains a proline-rich activation domain. Cotransfections with an expression vector for the AP2 transcription factor revealed an inhibition of HOXD-4 activity on both artificial and Hox-d-4 promoters (Table 3). By contrast, a GAL/VP16 expression vector did not squelch the activity of HOXD-4 on the Hox-d-4 promoter (Table 3). In fact, we note a two fold increase in luciferase activity in this case. This is not due to an effect of GAL/VP16 on the Hox-d-4 promoter (data not shown). Cotransfections with a luciferase reporter under the control of the pgk-1 promoter showed a 10 and 50% repression by 0.4 and 1.0 μg of the AP2 expression vector, or 1.0 μg of the GAL/VP16 expression vector was cotransfected with fixed amounts of the HOXD-4 expression vector (pgk-7) (not shown). We conclude that the activation of the Hox-d-4 promoter by HOXD-4 requires the homeodomain, unlike some homeodomain-independent activities noted for other proteins (38, 39).

A pentapeptide motif (consensus VYPWM) (8, 9) is located N-terminal to the homeodomain in Atp class homeoproteins and in the labial subfamily. It is not present in Abd-B, nor in Abd-B-like vertebrate proteins described to date (40). In our assay we did not observe a dramatic change in activity upon deletion of the pentapeptide motif (Fig. 4). Though a possible explanation is that P19 EC cells lack cofactors which could modulate homeoprotein activity through the pentapeptide motif, we obtained similar results in transfections of differentiated P19 cells (Fig. 4). We were also unable to detect activity for conserved regions A and B at the extreme N-terminus of HOXD-4. Conservation suggests function, therefore we conclude that transient transfections of P19 cells cannot reveal all the biochemical activities of homeoprotein molecules.

C-termini have been implicated in quantitative roles in homeoprotein function (41). Deletion of residues C-terminal to the HOXD-4 homeodomain (d1GH, Fig. 4) reduced activity on the Hox-d-4 promoter, and this could be rescued by replacement with the HOXA-4 C-terminus. However, neither was able to confer activity on the GAL4 DNA-binding domain (GAL/H and GAL/HA-4, Fig. 5). Nor did they cooperate with the HOXD-4 N-terminus in cotransfections with GAL/ABCD (not shown). Others report a mild (4 fold) increase in HOXD-4 activity upon deletion of regions G and H (42). However, these earlier studies made use of high levels of transfected Hox-d-4 expression vectors (5μg/plate) in HeLa cells. We have also noted 4 to 6 fold increases over wild-type in the activity of HOXD-4 C-terminal deletions at ≥4 μg/plate of transfected expression vector (unpublished observations). The significance of these increases at such high vector concentrations is not clear.

In Dfd, the region N-terminal to the homeodomain is 365 residues long, whereas in HOXD-4 it is 149 residues. Within
these regions, there is little sequence conservation apart from the extreme amino terminus and the region of the pentapeptide. Moreover, we have no direct evidence for transcriptional activation domains in either of these conserved regions. N-terminal to the homeodomain, Dfd has regions up to 10% in proline, but has domains richer in histidine, serine, glycine and acidic residues (43). By contrast, the HOXD-4 N-terminal half is proline-rich and displayed strong transcriptional activation function when fused to the GAL4 DNA-binding domain (GAL/ABCD, Fig. 5). Region C of HOXD-4 is 21% in proline. A 74 residue stretch of region C (residues 48 to 121) is 30% proline. Most of region C was fused to the GAL4 DNA-binding domain and conferred an 11 fold activation above background (GAL/C', Fig. 5). This is less than with the whole HOXD-4 N-terminus (371 fold), but comparable to that seen with the CTF/NF-1 proline-rich region fused to the SP1 DNA-binding domain (5 fold)(44).

The decreased activity of GAL/C' suggests additional activity in regions A, B or D. However, deletion of these regions did not impair HOXD-4 function (Fig. 4). It should be noted that 7 proline residues between positions 109 and 141 are missing from GAL/C', and the absence of this stretch may have been detrimental. Moreover, using extracts of transiently transfected cells, GAL/C' gave a less intense retarded band than GAL/ABCD in electrophoretic mobility shift assays with a GAL4 binding site (not shown). It is thus not clear for the moment whether the reduced activity of GAL/C' relative to GAL/ABCD is due to the loss of complementing residues. A HOXD-4 protein containing a duplication of residues 6 to 81 appears to be much more active than the wild-type protein (42). This stretch is proline- and glycine-rich and straddles much of the region we have called C', supporting a role for these residues in transcriptional activation.

Squelching (45) is a phenomenon whereby one transcription factor can inhibit the activity of another independent of both its ability to bind DNA and to directly interact with the inhibited protein. AP2 has a proline-rich transcriptional activation domain (22) and is able to squelch HOXD-4 activity, but not that of the acidic transcriptional activator GAL/VP16 (Table 3). Neither is GAL/VP16 able to squelch HOXD-4. These data support the presence of functionally similar proline-rich transcriptional activation domains in AP2 and HOXD-4 that are distinct from that of GAL/VP16. AP2 expression overlaps with that of many Hox genes both spatially and temporally in development. Tissues expressing both include neural crest, limb bud mesenchyme and the meso- and metanephros (46). Overlapping expression and similar proline-rich transcriptional activation domains may suggest cooperation between HOX proteins and AP2 in embryonic patterning.

Our results strongly suggest that a transcriptional activation domain of the proline-rich type (37) contributes to the activity of the HOXD-4 N-terminus, and that proline-rich domains of other homeoproteins may function in a similar capacity. HOX proteins with proline-rich N-terminal regions (≥20% over at least 60 residues) include the HOXD-4 paralogs HOXA-4 (11), HOXB-4 (12) and HOXC-4 (47). Others have suggested that the proline-rich N-terminus of chicken HOXA-4 functions as a transcriptional activator, though their results are difficult to interpret due to competitive effects on the high endogenous activity of a Jun-responsive reporter (48). By contrast, the background in our GAL-4 assay is low.

Dfd itself does not have proline-rich domains. Nonetheless, human HOXD4 can substitute for Dfd for the activation of the Dfd transcription unit in flies (49). In addition, the Dfd autoregulatory enhancer may respond to Dfd subfamily Hox products in mice (50). These data suggest that for some of the genetic pathways regulated by Dfd subfamily members, transcriptional activation can be accomplished by disparate domains. We also note that many Hox genes do not encode proline-rich regions, suggesting that these proteins use functionally distinct transcriptional activation domains. In the embryo, where the products of multiple Hox genes may bind responsive enhancers in various combinations, the specific array of such regulatory domains could be an important determinant of the transcriptional response.

ACKNOWLEDGEMENTS

We thank Wendy Hauck, Sylvie Mader, Christopher Mueller, Alaka Mullick, Heike Pöpperl and John White for useful discussions, Alain Ben David for technical help, and André Veillette and Dominique Davidson for advice with the trpE overexpression system. M.S.F. is the recipient of a Medical Research Council of Canada Scholarship. This work was supported by grants to M.S.F. from the National Cancer Institute of Canada, the Medical Research Council of Canada, and the Fonds de la Recherche en Santé du Québec.

REFERENCES