Mechanisms of Development 122 (2005) 887–899

www.elsevier.com/locate/modo

Alteration of the DNA binding domain disrupts distinct functions of the *C. elegans* Pax protein EGL-38

Guojuan Zhang, Sama F. Sleiman, Rong-Jeng Tseng, Vandana Rajakumar, Xiaodong Wang, Helen M. Chamberlin*

Department of Molecular Genetics, Ohio State University, 938 Biological Sciences Building, 484 West 12th Avenue, Columbus, OH 43210, USA

Received 23 November 2004; received in revised form 13 April 2005; accepted 18 April 2005
Available online 11 May 2005

Abstract

The paired-domain-containing Pax transcription factors play an important role in the development of a range of organ, tissue and cell types. Although DNA binding elements and target genes for Pax proteins have been identified, how these proteins identify appropriate DNA elements and regulate different genes in different cellular contexts is not well understood. To investigate the relationship between Pax proteins and their targets, we have studied the in vivo and in vitro properties associated with wild-type and different mutant variants of the *Caenorhabditis elegans* Pax protein EGL-38. Here, we characterize the properties of four mutations that result in an amino acid substitution in the DNA binding domain of EGL-38. We find that animals bearing the different mutant alleles exhibit tissue-preferential defects in *egl-38* function. The mutant proteins are also altered in their activity in an ectopic expression assay and in their in vitro DNA binding properties. Using in vitro selection, we have identified binding sites for EGL-38. However, we show that selected sites function poorly in vivo as EGL-38 response elements, indicating that sequence features in addition to DNA binding determine the efficacy of Pax response elements. The distinction between DNA binding and activity is consistent with the model that other factors commonly play a role in mediating Pax protein target site selection and function in vivo.

Keywords: DNA binding; Transactivation; Transcriptional control; Pax transcription factor

1. Introduction

Pax genes have been shown to play an important role in the development of a range of different cell types, and to promote cell proliferation and cell survival (reviewed by Chi and Epstein, 2002). *Pax* genes encode transcription factors that are identified based on the presence of a conserved DNA binding domain called the paired domain. Although four main classes of Pax proteins have been identified, all Pax proteins share several features common to the paired domain. The paired domain can interact with DNA as a monomer across a relatively large stretch of DNA of more than 20 bp, but it also has a modular structure of two helix-turn-helix domains which can bind to DNA independent of each other (Xu et al., 1995; Jun and Desplan, 1996; Kozmik et al., 1997; Jun et al., 1998; Xu et al., 1999a). Many DNA sequences that bind Pax proteins have been identified. However, how Pax proteins interact with DNA in vivo, and how they recognize appropriate targets from among all sequences that could act as PAX binding sites, remain open questions.

Genetic studies have shown that each Pax protein acts in the development of several different cell and organ types (reviewed by Chi and Epstein, 2002). Different functions are mediated through the regulation of different target genes, suggesting that cellular context plays an important role in Pax protein function. Thus other factors may allow the same Pax protein to regulate one set of target genes in one cell type, and a different set of targets in a different cell type. Pax proteins may function in a combinatorial manner with other proteins acting through separate enhancers (e.g. Brophy et al., 2003). In addition, examples in which Pax proteins form complexes or bind DNA synergistically with other

* Corresponding author. Tel.: +1 614 688 0043; fax: +1 614 292 4466.
E-mail address: chamberlin.27@osu.edu (H.M. Chamberlin).

0925-4773/$ - see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mod.2005.04.005

transcription factors have also been described (Fitzsimmons et al., 1996; Roberts et al., 2001; Jin et al., 2002; Miranda et al., 2002; Sigvardsson et al., 2002; Di Palma et al., 2003). While these studies suggest possible mechanisms for combinatorial control, dissecting how a given Pax factor participates in target gene expression in different cell types in vivo remains an important question.

To better understand how a single Pax transcription factor can regulate different target genes in different cells, we have characterized the functions of the *Caenorhabditis elegans Pax2/5/8* gene *egl-38*. Genetic studies have shown that *egl-38* functions in the development of several different cells and organ types including the egg-laying system, the hindgut, and the development of male mating spicules (Chamberlin et al., 1997). *egl-38* also affects the expression of different genes in the different cell types. For example, expression of the epidermal growth factor gene *lin-3* in the hermaphrodite vulval cells is dependent on *egl-38*, whereas the Ovo zinc finger transcription factor gene *lin-48* is a target for EGL-38 in hindgut cells (Chang et al., 1999; Johnson et al., 2001). *lin-48* is not expressed in the egg-laying system, indicating it is a target in one cell type, but not the other. Although *egl-38* is an essential gene, its different functions have been characterized using non-null, reduction-of-function alleles that allow homozygotes to be viable. These mutant alleles affect *egl-38* in vivo function in different ways, and each allele corresponds to a missense mutation that alters an amino acid in the EGL-38 DNA binding domain (DBD; Chamberlin et al., 1997, 1999; this work). Consequently, to investigate the relationship between Pax proteins and their in vivo targets, we have tested the in vivo and in vitro properties of these mutant EGL-38 variants.

2. Results

2.1. Identification of a new allele of egl-38

Non-null alleles of *egl-38* have been previously identified in genetic screens for egg-laying defective mutants and for mutants with abnormal male development (Trent et al., 1983; Chamberlin et al., 1997, 1999). We carried out a genetic screen for mutations that alter the expression of an EGL-38 target gene *lin-48* using a *lin-48*::*gfp* transgene, and identified a mutation, *gu22*, that affected expression of *lin-48* in hindgut cells. *gu22* mutants also exhibit defective hindgut and male tail development similar to that seen in *egl-38* mutants. We mapped *gu22* to linkage group IV, and showed that it fails to complement other *egl-38* alleles (Table 1 and data not shown). We sequenced the *egl-38* genomic DNA from *gu22* mutants, and identified a single missense mutation affecting the DNA binding domain-coding region of *egl-38* (Fig. 1). Thus, an independent screen for mutants defective in expression of an EGL-38 target gene has identified a fourth *egl-38* allele that alters the DBD of the EGL-38 protein.

Table 1
etl-38 alleles differentially affect expression of *lin-48*

Percent of hindgut cells expressing *lin-48*::*gfp* in trans-heterozygotes						
egl-38 allele	*egl-38* allele					
	s1775	*sy294*	*n578*	*sy287*	*gu22*	+
sy294		0				
n578	73	70	94			
sy287	5	10	79	34		
gu22	12	12	91	38	42	
+						100

2.2. Mutant alleles of egl-38 differentially affect its activity in vivo

To compare the new *egl-38*(*gu22*) allele to previously described alleles, and to extend the analysis of each allele for a range of functions, we examined the phenotypes associated with mutant animals (Table 2). Each allele was tested for its effect on egg-laying ability and male tail morphology, traits associated with *egl-38* mutants that have been described before (Chamberlin et al., 1997). In addition, egg-laying function was characterized at the cellular level by scoring the morphology of the vulF cells in L4 larval stage hermaphrodites, and the excretory system function was assessed by observing the position of the excretory duct cell opening in L3 larval stage animals (Fig. 2). Expression of *lin-48*::*gfp* in hindgut cells was scored in animals homozygous for each allele and in animals bearing different transheterozygous allelic combinations (Table 1; Fig. 3). Although each allele disrupts certain functions of *egl-38*, we have found that they cannot be ordered into an allelic series, going from weakest to strongest allele. Instead, the alleles exhibit cell or tissue preferential defects compared to each other. For example, in *egl-38*(*gu22*) and *egl-38*(*sy287*) mutants, tail functions including male tail development and *lin-48* expression are disrupted, whereas a high level of function in the egg-laying system is retained. In contrast, *egl-38*(*n578*) mutants are defective in the egg-laying system, but retain a high level of activity in the tail.

One possible interpretation of the loss-of-function genetics is that there are differences in tissue-specific stability, or subtle differences in protein activity for the different protein variants. To further investigate the differences among the mutants, we assessed the function of EGL-38 proteins when they ectopically expressed in animals under control of a heat-inducible promoter. We find that induction of wild-type EGL-38 results in ectopic expression of the target gene *lin-48* and causes a stringent lethality effect in embryos (Figs. 4,5). We tested the mutant proteins in these ectopic expression assays, and found that the different mutations alter EGL-38 activity to different extents (Table 3, Figs. 4,5). For example, the protein

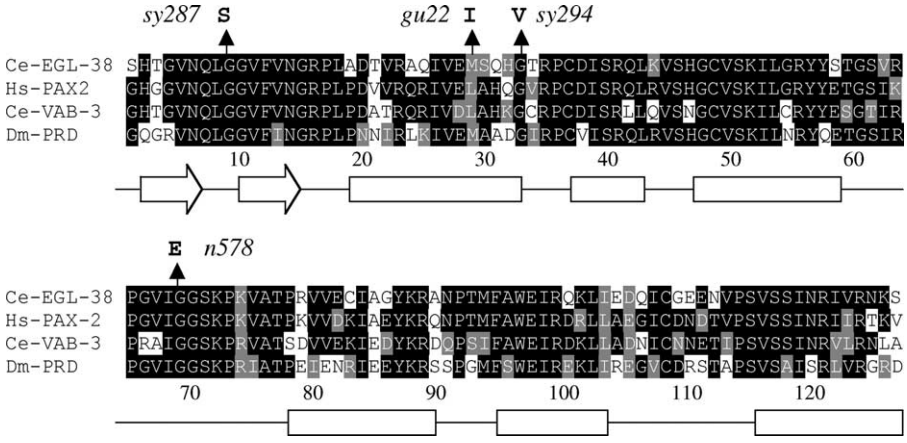

Fig. 1. Alignment of amino acid sequences of the paired domain from different Pax proteins. The paired domain from the *C. elegans* Pax2/5/8 protein EGL-38 (Chamberlin et al., 1997), human PAX2 (Eccles et al., 1992), the *C. elegans* Pax6 protein VAB-3 (Chisholm and Horvitz, 1995), and the Drosophila Paired protein (Bopp et al., 1986) are shown. Identical amino acids are shaded with black, conserved amino acids are gray. Sequences were aligned using ClustalW 1.8 (Jeanmougin et al., 1998), and shaded using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Structural features associated with each domain are indicated below the sequence, with arrows indicating β-sheet and boxes indicating α-helical structure (Xu et al., 1995, 1999b). The amino acid changes associated with each mutant allele are indicated above. Genomic DNA sequence analysis identified that *gu22* corresponds to a missense transition mutation of ATG (methionine) to ATA (isoleucine).

corresponding to the *egl-38*(*n578*) allele (G69E; amino acid position indicates position within the DBD) retains a high level of activity in the *lin-48*::*gfp* transactivation assay but is defective in the viability assay. In contrast, protein corresponding to the *egl-38*(*sy294*) allele (G33V) retains more activity in the viability assay than it does in the *lin-48*::*gfp* transactivation assay. We speculate that the lethality results from inappropriate expression of genes downstream of EGL-38. Although we do not know what gene or genes may be mediating the effect, the functional differences between the different mutant proteins in the two ectopic expression assays suggest it is not mediated through *lin-48*. Taken together, the loss- and gain-of-function genetic assays indicate that different mutations affecting the EGL-38 DNA binding domain can preferentially affect distinct functions of the protein.

2.3. Mutant alleles of egl-38 differentially affect the DNA binding properties of the protein product

Since the different mutant alleles affect the DBD of EGL-38, we asked whether they alter the DNA binding activity of the protein. We expressed the DBD of wild-type and mutant proteins in *E. coli* and tested their DNA binding ability in vitro in an electrophoretic mobility shift assay (EMSA; Fig. 6). As a probe, we used an endogenous DNA response element from the *lin-48* gene (termed lre2) that binds recombinant wild-type EGL-38 with high affinity (Johnson et al., 2001). We found that each mutation indeed affects the DNA binding activity of the protein. In addition, although the EMSA is not quantitative, the DNA binding exhibits some correlation with the ability of each protein to transactivate a *lin-48*::*gfp* transgene in vivo. For example, the protein corresponding to the *egl-38*(*n578*) allele (G69E) retains DNA binding activity in vitro and *lin-48* transactivation activity in vivo, whereas the protein corresponding to the *egl-38*(*sy294*) allele (G33V) is defective in both DNA binding and transactivation. However, the G69E (*n578*) protein and the M29I protein (corresponding to *egl-38*(*gu22*)) exhibit similar DNA binding activity, but very distinct activities in vivo (Tables 1–3; Fig. 5), suggesting that DNA binding to this endogenous element is not the only function that influences a protein's ability to transactivate *lin-48*.

2.4. Selection of DNA binding sequences for wild-type and mutant EGL-38 proteins

To further explore the DNA binding properties associated with wild-type and mutant proteins, we carried out an in vitro DNA selection experiment. We reasoned that selection with the wild-type protein might identify a range of possible binding sites for EGL-38, and allow for a comparison of DNA binding features between EGL-38 and Pax2/5/8 proteins from other species that have also been explored using DNA selection studies (Epstein et al., 1994; Jun and Desplan, 1996). We also selected with mutant

Table 2
egl-38 alleles preferentially affect different functions in vivo

egl-38 allele	DBD change	Percent wild-type			Microns (±SEM)
		Egg-laying	vulF morph.	Male tail	Ex. duct
+	None	100	100	100	13.07 (±1.55)
sy287	G9S	100	90	10	13.96 (±1.29)
gu22	M29I	94	80	2	13.23 (±1.54)
sy294	G33V	27	67	0	27.53 (±3.16)
n578	G69E	0	4	74	13.53 (±1.62)

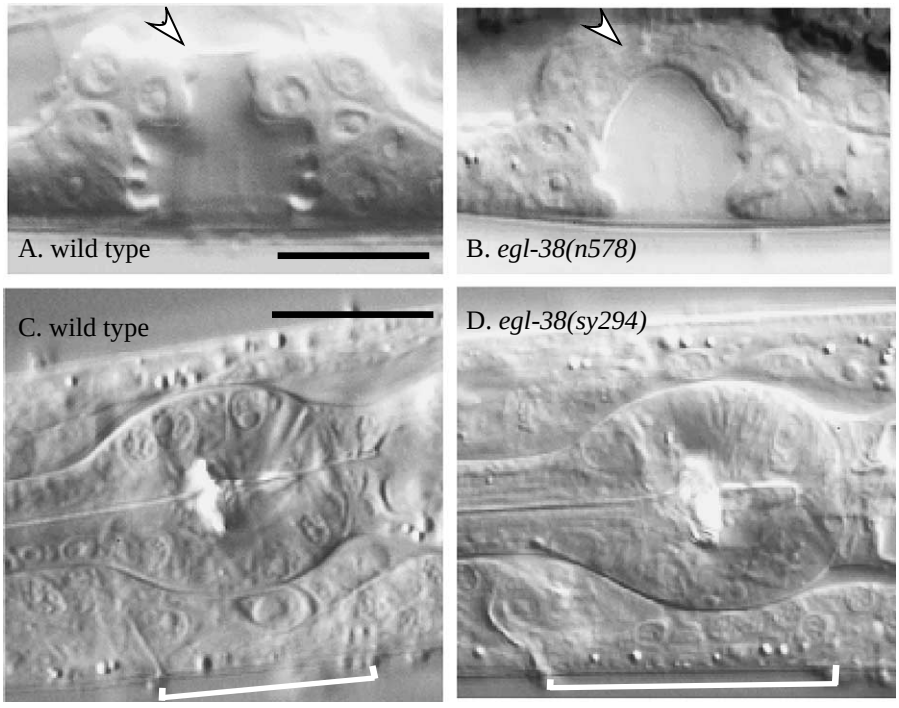

Fig. 2. *egl-38* functions in morphogenesis of the egg-laying and excretory systems. In wild-type L4 larva hermaphrodites (A), the vulva cells separate, leaving a thin laminar process between the vulva and the uterus (arrowhead). In *egl-38* mutants (B), the cells often fail to separate, which contributes to the egg-laying defect of the mutants. This is scored as a defect in vulF morphogenesis (vulF morph.) in Table 2. In wild-type animals (C), the excretory duct connects with the cuticle at a more posterior position compared to some *egl-38* mutants (D). The relative position can be measured as the distance of the pore opening to the base of the posterior pharynx bulb in L3 larvae (white bracket; Table 2). Scale bars are 10 μm.

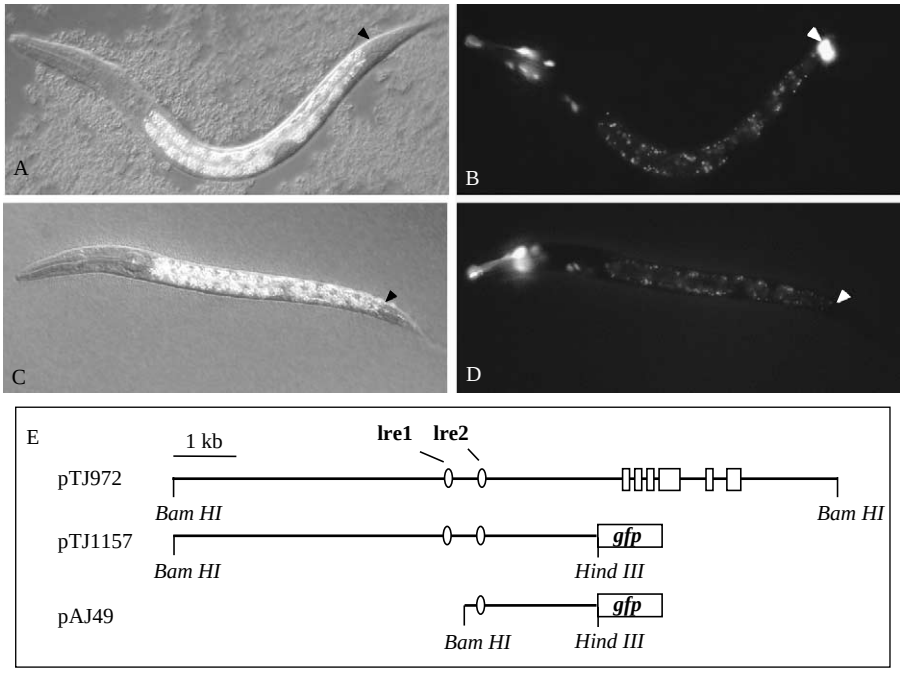

Fig. 3. Expression of *lin-48*::*gfp* is altered in *egl-38* mutants. Nomarski DIC micrographs (A, C) of L1 larvae, and epi-fluorescent micrographs (B, D) of the same animals. Wild-type animals bearing a *lin-48*::*gfp* transgene (pTJ1157) express GFP in hindgut cells (arrowhead, B), whereas *egl-38*(*sy294*) mutants do not (arrowhead, D). E. *lin-48*::*gfp* transgenes used in this study, after Johnson et al. (2001). pTJ972 is a genomic clone that can rescue all *lin-48* defects and is included for comparison to pTJ1157 and pAJ49. Rectangles correspond to *lin-48* exons, and ovals correspond to the two redundant EGL-38 responsive elements, lre1 and lre2. pTJ1157 includes both EGL-38 responsive elements, and mediates a robust response to EGL-38. This clone is included in the *saIs14* transgene used for the experiments summarized in Tables 1 and 3. pAJ49 is a truncated form of pTJ1157 that includes only lre2. This clone was used as the control and the template for the experiments summarized in Fig. 11. Use of the truncated clone for these experiments ensures that the EGL-38 response is mediated through a single response element.

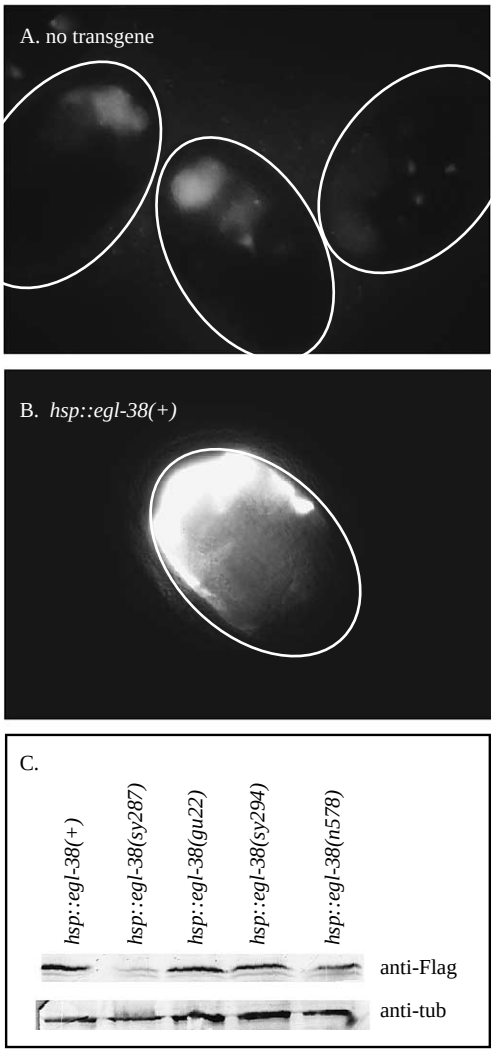

Fig. 4. Ectopic expression of EGL-38 induces ectopic *lin-48*::*gfp* expression. (A) Heat-treated embryos bearing a *lin-48*::*gfp* transgene exhibit a low level of GFP in the intestine. This low-level intestine expression is observed in all *lin-48*::*gfp*-bearing animals under normal culture conditions. (B) Heat-treated embryos bearing both a *hsp*::*egl-38 (+)* and a *lin-48*::*gfp* transgene show a high level of GFP, corresponding to ectopic expression of the *lin-48*::*gfp* transgene. The ectopic expression is commonly restricted to the posterior ectoderm (as in this embryo). This restricted ectopic expression may correspond to cells that contain the appropriate additional factors to promote *lin-48* expression. (C) Western blot of *hsp*::*egl-38* transgene-bearing strains to demonstrate expression of proteins tagged with a FLAG epitope. Although the *hsp*::*egl-38*(*sy287*) strain exhibits lower protein levels than the others, this strain exhibits a high level of *egl-38* activity in functional assays (Fig. 5 and Table 3).

EGL-38 proteins, to allow comparison of the sequences selected by mutant proteins to those selected by the wild-type protein. The rationale for these selections was to test the hypothesis that the mutations preferentially affect EGL-38 function by affecting the ability of the protein to bind to certain target sequences, but not others. One prediction of this hypothesis is that the mutant proteins might select sequences that represent a subset of the sequences selected by the wild-type protein, and that sequences selected by one

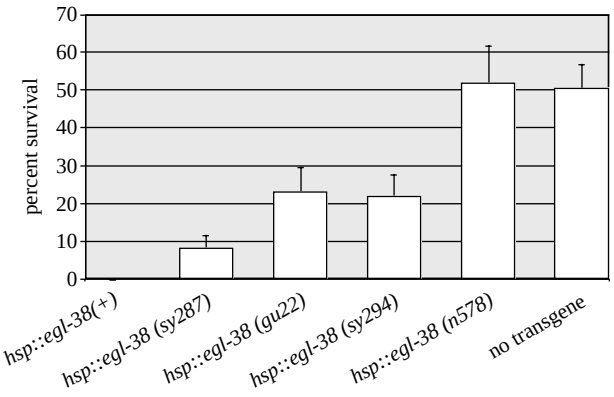

Fig. 5. Ectopic expression of EGL-38 affects embryonic viability. The graph shows the percentage of eggs that hatch following heat-shock treatment. The bars correspond to the average percentage of survivors from three independent trials, with brackets indicating SEM. Induced ectopic expression of wild-type *egl-38* results in embryonic lethality (*hsp*::*egl-38(+)*). This activity is disrupted to different extents by mutations that correspond to the different genetic alleles. Wild-type embryos (no transgene) exhibit some lethality which results from the heat-shock treatment.

mutant protein might not bind to other mutant proteins. Oligo selection from a pool of double-stranded DNAs containing a 25-mer of random nucleotides was carried out for each EGL-38 variant. The resulting selected DNAs were cloned and sequenced. Wild-type, M29I (*gu22*), and G9S (*sy287*) proteins selected multiple sequences from which consensus sequences were derived (Fig. 7). The sequences selected by the wild-type protein are similar to those recovered in similar selections using Pax2 or Pax8 proteins (Epstein et al., 1994; Jun and Desplan, 1996), and similar to that of the endogenous *lin-48* element lre2. Under the same selection conditions G69E (*n578*) and G33V (*sy294*) proteins each selected variants of a distinct single sequence (Fig. 8).

To test the function associated with the selected sequences, we focused on three representative elements: an 'optimal' sequence derived from the sequences selected by wild-type EGL-38, the sequence selected by G33V (*sy294*), and the sequence selected by G69E (*n578*). These sequences were chosen due to the different in vivo functions associated with each selecting protein. Specifically, wild-type protein retains all EGL-38 activities, the G69E (*n578*) protein retains a high level of *lin-48* transactivation activity, and the G33V (*sy294*) protein is defective in *lin-48* transactivation (Tables 1 and 3). The selected sequences

Table 3
Ectopic expression of EGL-38 induces ectopic expression of *lin-48*

Transgene	Ectopic *lin-48*::*gfp* (%)	*N*
none	0	16
hsp::*egl-38(+)*	73	26
hsp::*egl-38(sy287)*	80	25
hsp::*egl-38(sy294)*	7	14
hsp::*egl-38(n578)*	60	20

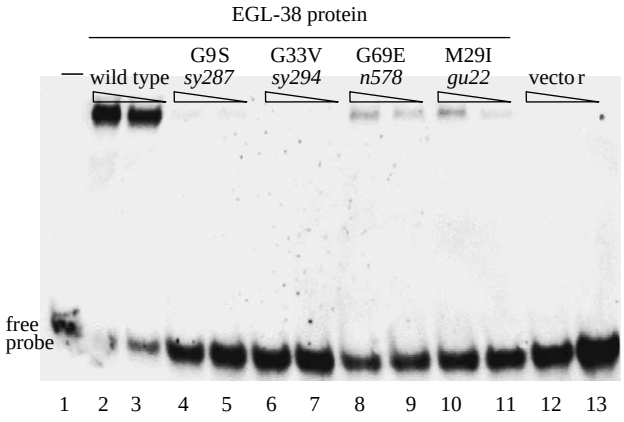

Fig. 6. DNA binding ability is affected in mutant EGL-38 proteins. Two dilutions (1:20 and 1:40) of wild-type and mutant EGL-38 DNA binding domain cell lysates were incubated with a probe corresponding to the endogenous response element from *lin-48*, lre2, and separated on a 6% polyacrylamide gel. Vector lanes (12, 13) contain lysates from cells expressing the pET-32 vector without *egl-38* cDNA. Lane 1 (–) includes no lysate (probe only).

were tested by altering lre2 oligos as illustrated in Fig. 8. In an EMSA, we found that the 'optimal' sequence bound wild-type protein better than mutant protein, whereas the sequences selected by G33V (*sy294*) and G69E (*n578*) show similar binding to all three proteins (Fig. 9). Thus, although the selection experiment indeed selected for sequences that bind mutant proteins, at least some of the recovered sequences do not preferentially bind one protein over another. Nevertheless, from the in vitro selection experiment we conclude that EGL-38 can select sequences similar to those selected by other Pax proteins, and that there are sequences that the mutant proteins can bind to as effectively as do wild-type proteins.

2.5. Features in addition to binding affinity play a critical role in the in vivo function of an EGL-38 response element

Although, we did not identify DNA sequences that preferentially bind one mutant protein over another in the selection experiment, we did identify a sequence that preferentially binds wild-type protein and thus behaves like the endogenous lre2, as well as two sequences that can bind wild-type and mutant proteins. To better understand the features that contribute to an effective EGL-38 response element in vivo, we compared the properties of these sequences selected in vitro to the endogenous EGL-38 response element lre2 in in vitro and in vivo assays. To compare DNA binding affinity, we carried out an EMSA competition assay (Fig. 10). We observed that in vitro, each selected sequence competes at least as well as lre2 for binding to wild-type EGL-38.

To test the function of the selected sequences in vivo, we incorporated each element into a reporter transgene by altering the endogenous lre2 site in a *lin-48*::*gfp* reporter clone (pAJ49; see Fig. 3). This clone is different from the one used in Table 1, in that it includes only one EGL-38 response element (lre2), rather than two EGL-38 response elements (lre1 and lre2). Thus, the effectiveness of the response element in the lre2 position can be assayed, as the clone lacks the redundant element that can mediate the response to EGL-38. However, with only one functional response element, the transgenes are not as robust, and

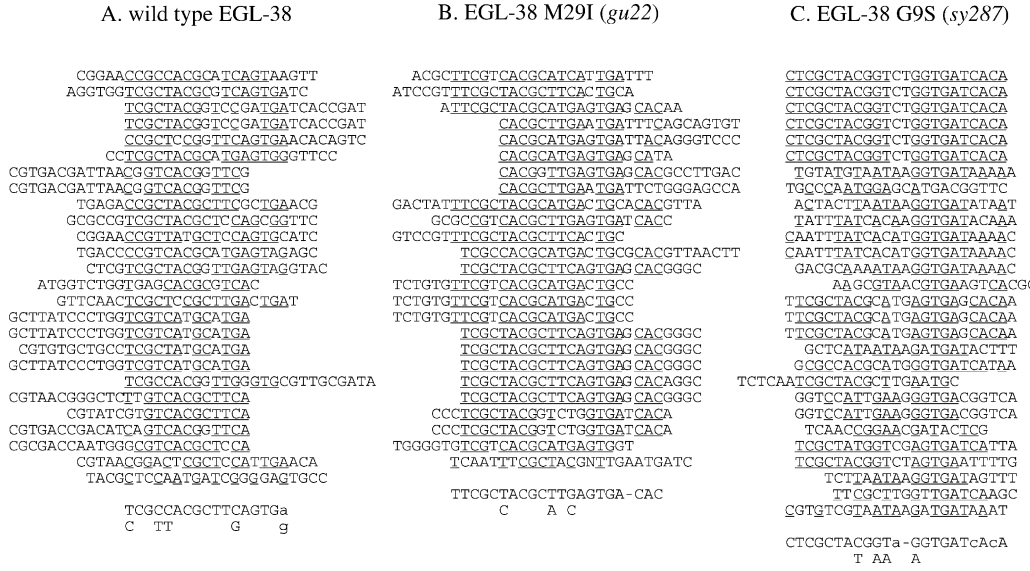

Fig. 7. Different EGL-38 variants select distinct sequences in vitro. The sequences obtained after successive rounds of selection with the DBD from wild-type EGL-38, EGL-38 M29I (*gu22*) and EGL-38 G9S (*sy287*). Each sequenced clone is listed, with a summary sequence below each column. In the summary sequence, upper case letters indicate non-random consensus $P<0.01$, lower case letters indicate non-random consensus, $0.01<P<0.05$. Nucleotides in each clone that match the consensus are underlined. Selections with EGL-38 G33V (*sy294*) and EGL-38 G69E (*n578*) identified essentially identical sequence from 7 to 8 sequenced clones. The sequences selected by these two proteins is listed in Fig. 8.

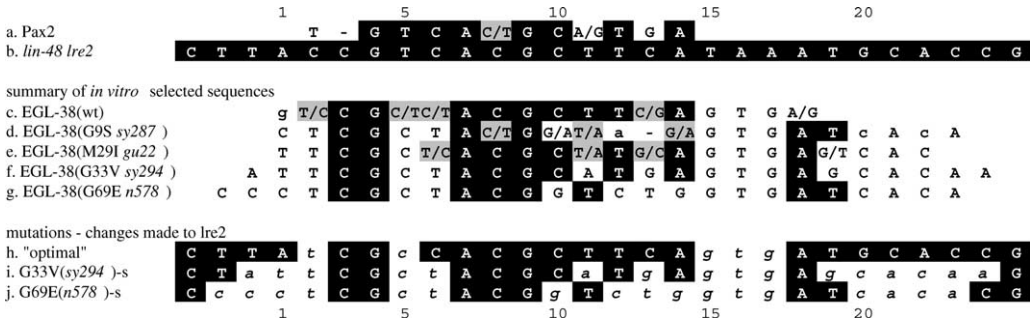

Fig. 8. Summary of Pax binding sequences. Nucleotide position corresponds to position from Xu et al., (1999a). Based on the structure of Pax6 protein bound to DNA, EGL-38::DNA contacts are predicted to extend from positions 2 to 22, with the more amino-terminal portion of the paired domain associated with sequences to the left, and the carboxyl-terminal portion associated with sequences to the right. Sequence is shaded according to its match with the EGL-38 response element lre2. Black indicates a sequence match to lre2, and gray indicates one of two common nucleotides matches. (A) Composite sequence selected by Pax2 (Epstein et al., 1994). (B) Endogenous sequence lre2 from the EGL-38 target gene *lin-48* (Johnson et al., 2001). (C–G) Summary of EGL-38 selected sequences. (H–J) Sequence alterations tested in vitro and introduced into the lre2 site of *lin-48* (Figs. 9–11). 'Optimal' indicates a consensus from sequences selected by wild-type EGL-38 protein. G33V (*sy294*)-s and G69 (*n578*)-s indicate sequence selected by the protein from *egl-38*(*sy294*) and *egl-38*(*n578*) mutants, respectively. Each variant incorporates selected nucleotides into the context of an lre2 oligo or transgene. Changes from the lre2 element are not shaded, and are also indicated in lower-case italics.

exhibit altered response in mutant *egl-38* backgrounds compared to full length transgenes containing both response elements (Johnson et al., 2001). Nevertheless, altering the lre2 element in this reporter provides an in vivo context to compare the function of an endogenous DNA binding element to those selected in vitro. Surprisingly, we found that selected sequences all function poorly when tested in wild-type animals. The 'optimal' sequence failed to mediate an EGL-38 response, whereas the sequence elements selected by G33V (*sy294*) and G69E (*n578*) mediated a reduced level of response (Fig. 11). In general, each selected element functions less efficiently as an EGL-38 response element than does the endogenous lre2.

To test whether the altered response elements simply reduce the response of the reporter transgene to endogenous EGL-38, we introduced the two altered transgenes that retained some activity into different *egl-38* genetic mutant backgrounds. The endogenous lre2 element exhibits differential responses, with *lin-48*::*gfp* expression still observed in *egl-38*(*n578*) mutants, but an absence of expression in *egl-38*(*sy294*) mutants (Fig. 11). This reporter gene response is similar to the level of activity that each allele retains with respect to hindgut cell development (Johnson et al., 2001). In contrast, the transgenes modified to contain the selected sequences exhibited a similar, low level of activity in each *egl-38* genetic background. This result shows that the mutant proteins (e.g. the G33V protein in the *egl-38*(*sy294*) mutant that normally cannot transactivate *lin-*

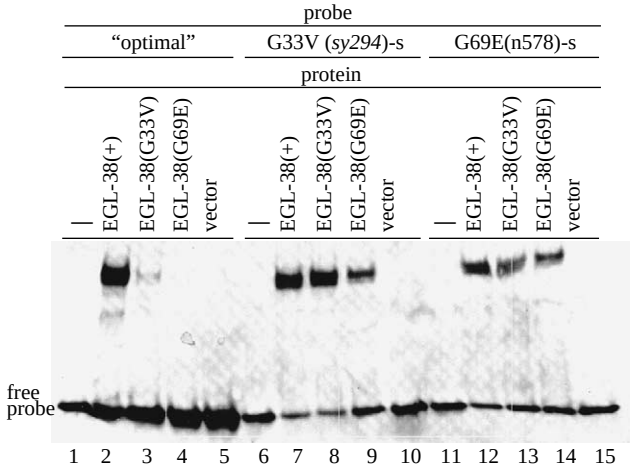

Fig. 9. EGL-38 proteins bind selected sequences in vitro. A 1:30 dilution of protein lysates including wild-type and mutant EGL-38 proteins were incubated with probes corresponding to different selected sequences ('optimal' G33V (*sy294*)-s, and G69E (*n578*)-s selected sequences are shown in Fig. 8). Vector lanes contain lysates from cells expressing the pET-32 vector without *egl-38* cDNA. –, indicates oligo only (no lysate).

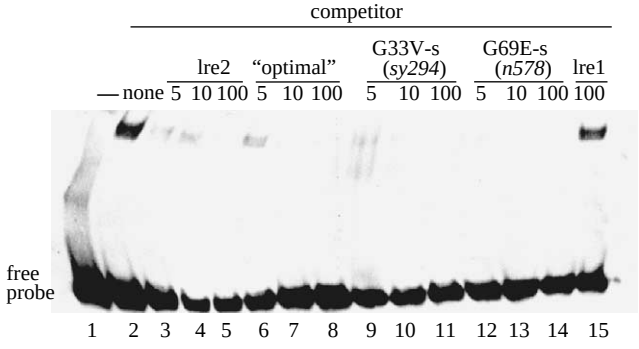

Fig. 10. Comparison of EGL-38 binding to endogenous and in vitro selected sequences. A 1:30 dilution of wild-type EGL-38 DBD cell lysates were incubated with the lre2 probe and increasing-fold unlabeled competitor oligo. Sequences corresponding to lre2, 'optimal' G33V(*sy294*)-s, and G69E(*n578*)-s are shown in Fig. 8. lre1 is an element from the *lin-48* gene that can mediate the response to EGL-38, but does not bind the protein in vitro (Johnson et al., 2001), and serves as a negative control. –, indicates oligo only (no lysate).

	egl-38 genotype	EGL-38 response element in *lin-48::gfp* transgene	Expression in hindgut (percent of total cells)	Expression in hindgut (av. cells/animal +/– s.e.m.)	N
			0 100		
a.	wild type	lre2 (+)		**2.75** +/– 0.21	56
b.	wild type	"optimal"		**0.00** +/– 0.00	54
c.	wild type	G33V(*sy294*)-s		**1.10** +/– 0.16	50
d.	wild type	G69E(*n578*)-s		**0.89** +/– 0.16	46
e.	*egl-38(n578)*	lre2 (+)		**0.88** +/– 0.15	69
f.	*egl-38(n578)*	G33V(*sy294*)-s		**0.63** +/– 0.14	67
g.	*egl-38(n578)*	G69E(*n578*)-s		**0.71** +/– 0.15	67
h.	*egl-38(sy294)*	lre2 (+)		**0.00** +/– 0.00	42
i.	*egl-38(sy294)*	G33V(*sy294*)-s		**1.30** +/– 0.15	47
j.	*egl-38(sy294)*	G69E(*n578*)-s		**0.42** +/– 0.12	48

Fig. 11. The endogenous lre2 element is a more effective EGL-38 response element than are in vitro selected sites. *lin-48*::*gfp* transgenes bearing the endogenous (lre2) or modified EGL-38 response element were tested for their ability to promote GFP expression in the four hindgut cells that normally can express the transgene. The data are presented in two ways. Graphically, with the percentage of total cells expressing GFP (black bar) or not expressing GFP (white bar). Numerically, with the average number of cells expressing per animal, ±SEM. The percent expression for lre2-based transgenes in this figure is lower than those reported in Table 1 because this assay uses extrachromosomal transgenes with only one EGL-38 response element (lre2), whereas the transgene in Table 1 is integrated into a chromosome and contains two EGL-38 response elements (lre1 and lre2; see Fig. 3). *N*, number of animals scored (four cells are scored for each animal).

48) are capable of acting in vivo when challenged with a response element to which they can bind. Each protein (wild-type and mutant) can act through the selected sequences equally well, but with reduced efficiency compared to how wild-type protein acts through the endogenous lre2 element. Overall, the selected sequences reduce the efficacy of the response, but also change how the target gene responds to the different EGL-38 proteins in a cellular context.

3. Discussion

3.1. Alterations to a Pax protein DNA binding domain can affect its function in a tissue-preferential manner

DNA binding affinity is a key component of transcription factor function. Our results with EGL-38 indicate that mutations in the DNA binding domain of Pax proteins can alter their DNA binding and their in vivo activity. Many mutations that affect the DNA binding domain of Pax proteins have been characterized previously (Tang et al., 1997; DeStefano et al., 1998; Vilain et al., 2001; Jumlongras et al., 2004). However, since we have characterized several alleles and functions of *egl-38*, we have been able to demonstrate that the alleles that affect the DBD of EGL-38 can preferentially affect its activity in certain functions, but not others. Our results with EGL-38 suggest one of two possible models to explain the tissue-preferential effects. One possibility is that the different mutant alleles affect the ability of EGL-38 to bind to certain sequences, but not others. This model predicts that either there are one or a few key targets in each cell type, or that there are specific response element sequences that are associated with all the targets for EGL-38 in a particular cell type. An alternative model is that, although the mutations affect DNA binding, their tissue preferential effects result from their impact on EGL-38 function in combination with other proteins that influence target sequence recognition in vivo. This model predicts that there would be cell-specific co-factors that function with EGL-38 and participate in identifying which sequences correspond to an appropriate DNA binding site or that enhance EGL-38 transactivation activity. It also suggests that the interaction between the co-factor and EGL-38 would be mediated, or otherwise affected, by the EGL-38 DBD. Although we favor the second model, clearly distinguishing between the two will require identification of proteins that act with EGL-38 and a large collection of cell-specific EGL-38 target genes.

3.2. The relationship between DNA binding and Pax target gene transactivation

Our results with sequences selected by EGL-38 wild-type and mutant proteins indicate that DNA binding affinity is only one feature that influences whether a sequence acts as a response element for Pax proteins. Indeed, altering the endogenous lre2 response element to a sequence that binds EGL-38 protein in vitro with high affinity both reduces the overall effectiveness of the element, and also alters the nature of how the element responds in different *egl-38* genetic backgrounds. Whereas the endogenous sequence responds to the different mutant proteins in a manner similar to the apparent cell-specific in vivo function associated with each allele, the elements selected in vitro respond similarly in each genetic background, suggesting that the functional properties of the in vitro assay are not the same as those in vivo. This may be either because the DNA is presented to the protein differently in vivo compared to in vitro, EGL-38 protein is modified in vivo in a manner different from in vitro, or because other proteins influence how EGL-38 binds DNA in a cellular context. However, we have not ruled out an alternative possibility that each of the altered sequences changes the normal EGL-38 binding element to one that now binds a different protein that can compete with EGL-38 in vivo.

Previous studies have shown that although Pax proteins can bind DNA as monomers, they can also form complexes with other transcription factors. For example, Pax5 and Ets transcription factors form a complex on the *mb-1* promoter that results in improved DNA binding affinity compared to that seen with either protein alone (Fitzsimmons et al., 1996). Complex formation between Pax5 and Ets is mediated through the β-hairpin domain of the EGL-38 DBD (Wheat et al., 1999). In vitro, EGL-38 and *C. elegans* Ets proteins can also form ternary complexes on DNA (Fitzsimmons et al., 2001). Although genetic experiments have not identified a role for *C. elegans* Ets genes in regulating *lin-48* (H.M.C., unpublished), these studies demonstrate that the DBD can play roles in both DNA binding and protein interaction. We find in our current studies that the DNA binding ability of the different EGL-38 mutant proteins is affected. However, the tissue-preferential features may also reflect disruption of specific protein interactions mediated by the DBD. In the current experiments, we have only tested one response element in one gene. However, many endogenous response elements identified for other Pax proteins also exhibit lower in vitro DNA binding affinity compared to 'optimal' sequences (Czerny et al., 1993). We speculate that the sub-optimal features of endogenous Pax binding sites might indicate a common role for other factors in the recognition of in vivo Pax response elements. This combinatorial control would allow for the recognition of Pax targets in both a cell-specific and sequence-specific manner.

4. Materials and methods

4.1. Caenorhabditis elegans strains

Caenorhabditis elegans strains were cultured according to standard techniques (Sulston and Hodgkin, 1988). Mutations are described in wormbase (http://www.wormbase.org/). Mutations used: *unc-119*(*e2498*) *III*; *unc-22*(*s7*), *dpy-20*(*e1282*), *egl-38*(*n578*), *egl-38*(*sy294*), *egl-38*(*sy287*), *egl-38*(*gu22*), *egl-38*(*s1775*) *IV*; *him-5*(*e1490*) *V*. Integrated transgene: *saIs14* (*lin-48*::*gfp*; Sewell et al., 2003).

4.2. Identification and characterization of egl-38(gu22)

gu22 was identified in a genetic screen for mutations that result in reduced expression of a *lin-48*::*gfp* reporter transgene in hindgut cells. Fourth larval stage P0 hermaphrodites from the strain CM3 (*unc-119*(*e2498*); *him-5*(*e1490*); *saIs14*) or CM117 (*unc-119*(*e2498*); *saIs14*) were mutagenized with 50 mM EMS for 4 h (Sulston and Hodgkin, 1988), placed individually on agar plates, and allowed to produce self-progeny for two generations. F2 animals were screened for reduced expression of *lin-48*::*gfp* in the hindgut, a trait that phenocopies the expression pattern in *egl-38* mutants. From a screen of about 18,000 mutagenized gametes, several mutations were recovered, one of which was a new allele of *egl-38* (*gu22*). *gu22* was mapped to LGIV by crossing with mapping strains containing markers from each linkage group: MT3751 (*dpy-5*(*e61*) *I; rol-6*(*e187*) *II; unc-32*(*e189*) *III*) and MT464 (*unc-5*(*e53*) *IV; dpy-11*(*e224*) *IV; lon-2*(*e678*) *X*). *gu22* exhibited linkage with *unc-5*(*e53*), but not the other markers. Failure to complement *egl-38* was tested by crossing *gu22* heterozygous males with *dpy-20*(*e1282*) *egl-38*(*sy294*) hermaphrodites, and observing *egl-38*-phenotypes among the non-Dpy cross progeny.

4.3. Phenotypic analysis

Quantification of egg-laying ability, male tail morphology, *lin-48*::*gfp* expression, vulval morphology, and excretory duct cell position was carried out as previously described (Chamberlin et al., 1997; Johnson et al., 2001), and as detailed below. Each phenotype was analyzed in at least 25 animals for each genotype.

For *lin-48*::*gfp* (Table 1; Fig. 3), the transgene *saIs14* was used. This transgene contains the pTJ1157 reporter gene with the upstream regulatory sequences from *lin-48* that include both identified EGL-38-responsive elements (lre1 and lre2; Fig. 3E). In wild-type animals, these elements exhibit functional redundancy, but transgenes containing a single response element exhibit enhanced sensitivity in different *egl-38* genetic backgrounds (Johnson et al., 2001). Thus, the most robust EGL-38 response is mediated through both elements. Trans-heterozygous animals were created by crossing *saIs14*-bearing males with *unc-22*(*s7*) *egl-38*(*x*)

hermaphrodites. Male offspring (*unc-22*(*s7*) *egl-38*(*x*)/++) were crossed to *dpy-20*(*e1282*) *egl-38*(*y*)/*DnT1* hermaphrodites. Twi (Unc-22 heterozygotes) non-Unc (non-DnT1) animals were scored. *egl-38*(*s1775*) is a lethal allele that behaves as a genetic null for *egl-38* (Chamberlin et al., 1997).

To quantify vulF cell morphology, L4 animals were observed under Nomarski optics (Table 2, Fig. 2). When the vulval nuclei anterior and posterior to the anchor cell separate, a thin laminar process is formed between the vulva and uterus. In *egl-38* mutants the vulval nuclei separate to a lesser extent, and the tissue between the vulva and uterus remains thick (Chamberlin et al., 1997).

Quantification of the positioning of the excretory duct opening was done by measuring the distance from the excretory pore to the base of the posterior pharynx bulb in L3 animals as previously described (Table 2, Fig. 2; Wang and Chamberlin, 2002).

4.4. Ectopic expression of EGL-38 proteins in worms

A cDNA coding for full-length wild-type EGL-38 protein was cloned downstream of the *hsp16–41* promoter in the vector pPD49.83 (a gift from A. Fire). Sequences corresponding to the FLAG epitope were introduced in-frame at the 3′ end of the cDNA to allow detection of the expressed protein. This construct represents *hsp::egl-38 (+)* and also served as the template to produce *egl-38* mutant variants using the QuikChange mutagenesis method (Stratagene). Mutant clones were sequenced to confirm the clone and ensure that additional mutations were not incorporated during the mutagenesis process. Transgenes for the heat-inducible constructs were produced by microinjection of plasmid DNA into the mitotic germline of hermaphrodites according to the method of Mello et al. (1991). For all experiments, 40–50 ng/μl experimental clone was co-injected with 15 ng/μl pDP#MM016 (*unc-119(+)* plasmid; Maduro and Pilgrim, 1995) into *unc-119*(*e2498*) animals. The resultant extrachromosomal transgenes were integrated into the chromosome by treating transgene-bearing animals with 3.0–3.5 kRad gamma irradiation, and selecting for stable transgene transmission in subsequent generations following the protocol of Inoue et al. (2002).

Integrated transgenes were back-crossed several times, and tested for protein expression level following heat induction of the transgene using western blotting (Fig. 4). Protein expression was visualized by subjecting mixed-stage animals to a 35C heat-shock for 30 min. After 30 min recovery at room temperature, worms were washed 2× in M9 buffer, and frozen in liquid nitrogen. Thawed pellets were sonicated in SDS-PAGE buffer, boiled for 5 min, and separated by SDS-PAGE. Proteins were transferred to membrane, blocked overnight with 5% non-fat dried milk in PBS plus 0.1% tween, then washed 1× with PBS tween. Proteins were detected with anti-FLAG and anti-α-tubulin antibodies (Sigma) diluted 1:100.

4.5. Functional assays of ectopically expressed EGL-38 proteins

The activity of ectopically expressed EGL-38 proteins was assessed by characterizing its impact on *lin-48* transactivation and on embryonic viability. For *lin-48* transactivation, 50 ng/μl of the *lin-48::gfp* clone pTJ1157 (Fig. 3; Johnson et al., 2001) was co-injected with 180 ng/μl pRF4 (*rol-6*(*d*); Mello et al., 1991) into animals bearing the integrated *hsp::egl-38* transgenes. Expression of *lin-48::gfp* in the expected pattern (Fig. 3B) was confirmed in the resulting transgenic lines. Transgene-bearing hermaphrodites were allowed to lay eggs for 5–12 h, and then removed from the plate. The eggs were subjected to a 35C heat shock for 30 min, followed by incubation at 15C for 10–11 h, and analysis for expression of GFP. The data of Table 3 are all from embryos prior to elongation, but with clear differentiation and alignment of intestinal cells.

For viability, transgene-bearing hermaphrodites were allowed to lay eggs for 5 h, removed from the plate, and the eggs were counted. The eggs were then subjected to a 35 °C heat-shock for 30 min. On the second day, the eggs and resulting larvae were subjected to a second heat-shock treatment. On the fourth day, the number of larvae at different developmental stages was counted, and the number of dead eggs was inferred by comparing resulting larva to number of eggs counted on the first day. The data in Fig. 5 result from three independent trials. Each trial included 50 or more eggs for a given genotype. We have found that all embryos are sensitive to heat shock. However, heat treatment for 30 min at 35 °C induces transgene expression yet permits significantly greater embryonic survival than other heat treatment protocols such as those used for larvae (32–33 °C for 1–2 h). Nevertheless, about half of control embryos do not survive heat-shock treatment.

4.6. EGL-38 protein expression and EMSA

Production of recombinant EGL-38 DNA binding domain (EGL-38 DBD) in *E. coli* was carried out as described previously (Johnson et al., 2001). Briefly, cDNA coding for amino acids 29–156 was amplified using RT-PCR and cloned into vector pET-32 (Novagen). This construct also served as the template to produce *egl-38* mutant cDNA using the QuikChange method (Stratagene). Mutant clones were sequenced to confirm the clone and ensure that additional mutations were not incorporated during the mutagenesis process. Recombinant EGL-38 DBDs were expressed in *E. coli* strain BL21 (DE3) (Novagen). Cells were lysed with sonication in buffer Z (Fitzsimmons et al., 2001). Recombinant protein expression was confirmed with SDS-PAGE, and expressed protein amount was estimated by comparing dilutions to known quantities of BSA. For the experiments, lysates were utilized that included similar levels of expressed recombinant protein. Cell lysates, rather than purified protein, were

used for DNA binding assays since previous experiments have found that purification of Pax proteins significantly reduces their activity (Fitzsimmons et al., 2001). In all cases, lysates derived from cells expressing the pET-32 vector with no cDNA insert was used as a negative control.

All the oligonucleotides used are illustrated in Fig. 8, except for lre1 which is described in Johnson et al. (2001). Probes were purified on a polyacrylamide gel, and labeled at the 3′ end with digoxigenin (Roche). 0.8 ng of each labeled oligonucleotide probe was used in EMSA. The binding reactions were carried out in a 20 μl volume including 1× binding buffer (10 mM Hepes, pH 7.9; 50 mM NaCl; 0.1 mM DTT; 1 mM EDTA; 5% glycerol; and 0.25% BSA), 2 μg poly[d(I-C)], 0.2 μg L-lysine, and 1 μl diluted 1:20, 1:30 or 1:40 cell lysate. These dilutions were selected to better observe the relatively weak binding of the EGL-38 mutant proteins. For the competition assay (Fig. 10), labeled and unlabeled oligo were mixed together prior to adding the cell lysate to the reaction. An equal amount of labeled oligo was used for each reaction. The reactions were incubated at room temperature for 30 min, then loaded on a 6% polyacrylamide gel in 1X TGE buffer for electrophoresis. The oligonucleotides were transferred to membrane by electroblotting. Digoxigenin label was detected using chemiluminescence (Roche).

4.7. In vitro DNA binding site selection

An 83-bp oligonucleotide containing 25 degenerate bases was used as described in Epstein et al. (1994). One hundred nanograms of double-stranded oligomer were used for each round of selection. The selection reactions were carried out in a 20 μl volume containing 1× binding buffer (above), 2 μg poly [d(I-C)], 0.2 μg poly L-lysine, and 4 μl cell lysate. Selection with lysates derived from cells expressing the pET-32 vector with no insert was used as a negative control. These control lysates selected DNAs unrelated to those selected by the EGL-38 protein variants (data not shown). The reactions were incubated at room temperature for 30 min with gentle mixing by flicking the tubes every 5 min. Next, 20 μl Ni-NTA Magnetic Agarose Beads (QIAGEN) and 20 mM imidazole were added. The reactions were incubated on ice for 1 h with gentle flicking to allow the DNA-protein complex to bind to the beads. The beads were separated from the supernatant using a Magnetic Particle Concentrator (Dynal A.S., Oslo, Norway). The beads were washed once with 250 μl 1× binding buffer containing 20 mM imidazole, and then twice with 500 μl washing buffer (50 mM NaH2PO4, 50 mM NaCl, 20 mM imidazole, pH 8.0). The washing buffer was removed with the concentrator. Selected oligomer was eluted with 25 μl elution buffer (50 mM NaH2PO4, 50 mM NaCl, and 250 mM imidazole, pH8.0). The oligomer was amplified by PCR using primers A and B (Epstein et al., 1994). The PCR products were purified on a 15% polyacrylamide gel, then used for the next round of selection. After five rounds of selection, the selected sequences were used as the probe in an EMSA assay, but shifted bands were not detected. Consequently, four additional rounds of selection were carried out. Selected sequences were cloned into pBluscript SK (+).

4.8. Analysis of selected sequences

Clones were sequenced at the OSU Cancer Center Genotyping and Sequencing Unit. Sequences were aligned using SeqApp (http://www.iubio.bio.indiana.edu), and analyzed using Gibbs Recursive Sampler (Thompson et al., 2003) with manual editing to identify summary sequences (Fig. 7). For each position in the summary sequence, nucleotide distributions that were non-random by Chi-square analysis ($P<0.01$ for upper case letter, $0.01<P<0.05$ for lower case letter) were included in the summary (Figs. 7,8).

4.9. Construction and analysis of modified *lin-48*::*gfp* transgenes

Alterations to the EGL-38 response element lre2 were made in the clone pAJ49 (Fig. 3E). This clone incorporates regulatory sequences from *lin-48* upstream of the coding sequences for green fluorescent protein (GFP). It is a truncated form of pTJ1157 that lacks the redundant element lre1, allowing lre2 to be the only EGL-38 response element influencing GFP expression in hindgut cells (Johnson et al., 2001). Although the most robust response is obtained from transgenes with both lre1 and lre2, it is necessary to use this compromised transgene for this assay so that the redundancy between lre1 and lre2 does not interfere with our ability to assess the function associated with the altered response elements. The nucleotide changes corresponding to each selected sequence variant were incorporated to precisely alter lre2 using either the QuikChange method (Stratagene) or a two-step, PCR-based method (Ausubel et al., 2000). Altered clones were sequenced to confirm that the desired change was incorporated. Transgenes for the reporter constructs were produced by microinjection of plasmid DNA into the mitotic germline of as described above. For all experiments that test expression of the original or altered *lin-48*::*gfp*, 40–50 ng/μl reporter clone was co-injected with 15 ng/μl pDP#MM016 (*unc-119(+)* plasmid; Maduro and Pilgrim, 1995) into *unc-119*(*e2498*); *him-5*(*e1490*) animals. GFP expression levels were assessed in L1 and L2 animals as previously described (Johnson et al., 2001). As *egl-38* genotype and the lre2 response element do not affect the expression of *lin-48* in other cell types, each animal was confirmed to carry the transgene based on expression of GFP in head cells prior to scoring the four hindgut cells that are normally capable of expressing *lin-48*::*gfp*.

Acknowledgements

We thank A. Fire for vectors, A. Kauffman of the OSU Reactor Lab for use of, and assistance with, the gamma irradiation source, and Chamberlin lab members for critical reading of the manuscript. Some of the nematode strains used in this project were provided by the Caenorhabditis Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR). This work was supported by the NIH (GM62336).

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 2000. Current Protocols in Molecular Biology. Wiley, New York, NY.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., Noll, M., 1986. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. Cell 47, 1033–1040.

Brophy, P.D., Lang, K.M., Dressler, G.R., 2003. The secreted frizzled related protein 2 (SFRP2) gene is a target of the Pax2 transcription factor. J. Biol. Chem. 278, 52401–52405.

Chamberlin, H.M., Palmer, R.E., Newman, A.P., Sternberg, P.W., Baillie, D.L., Thomas, J.H., 1997. The PAX gene egl-38 mediates developmental patterning in *Caenorhabditis elegans*. Development 124, 3919–3928.

Chamberlin, H.M., Brown, K.B., Sternberg, P.W., Thomas, J.H., 1999. Characterization of seven genes affecting *Caenorhabditis elegans* hindgut development. Genetics 153, 731–742.

Chang, C., Newman, A.P., Sternberg, P.W., 1999. Reciprocal EGF signaling back to the uterus from the induced *C. elegans* vulva coordinates morphogenesis of epithelia. Curr. Biol. 9, 237–246.

Chi, N., Epstein, J.A., 2002. Getting your Pax straight: Pax proteins in development and disease. Trends. Genet. 18, 41–47.

Chisholm, A.D., Horvitz, H.R., 1995. Patterning of the *Caenorhabditis elegans* head region by the Pax-6 family member vab-3. Nature 377, 52–55.

Czerny, T., Schaffner, G., Busslinger, M., 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. Genes. Dev. 7, 2048–2061.

DeStefano, A.L., Cupples, L.A., Arnos, K.S., Asher Jr.., J.H., Baldwin, C.T., Blanton, S., et al., 1998. Correlation between Waardenburg syndrome phenotype and genotype in a population of individuals with identified PAX3 mutations. Human Genet. 102, 499–506.

Di Palma, T., Nitsch, R., Mascia, A., Nitsch, L., Di Lauro, R., Zannini, M., 2003. The paired domain-containing factor Pax8 and the homeodomain-containing factor TTF-1 directly interact and synergistically activate transcription. J. Biol. Chem. 278, 3395–3402.

Eccles, M.R., Wallis, L.J., Fidler, A.E., Spurr, N.K., Goodfellow, P.J., Reeve, A.E., 1992. Expression of the PAX2 gene in human fetal kidney and Wilms' tumor. Cell Growth. Differentiat. 3, 279–289.

Epstein, J., Cai, J., Glaser, T., Jepeal, L., Maas, R., 1994. Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. J. Biol. Chem. 269, 8355–8361.

Fitzsimmons, D., Hodsdon, W., Wheat, W., Maira, S.M., Wasylyk, B., Hagman, J., 1996. Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. Genes. Dev. 10, 2198–2211.

Fitzsimmons, D., Lutz, R., Wheat, W., Chamberlin, H.M., Hagman, J., 2001. Highly conserved amino acids in Pax and Ets proteins are required for DNA binding and ternary complex assembly. Nucl. Acids Res. 29, 4154–4165.

Inoue, T., Sherwood, D.R., Aspock, G., Butler, J.A., Gupta, B.P., Kirouac, M., et al., 2002. Gene expression markers for *Caenorhabditis elegans* vulval cells. Gene Express. Patterns 2, 235–241.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., Gibson, T.J., 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. 23, 403–405.

Jin, Z.X., Kishi, H., Wei, X.C., Matsuda, T., Saito, S., Muraguchi, A., 2002. Lymphoid enhancer-binding factor-1 binds and activates the recombination-activating gene-2 promoter together with c-Myb and Pax-5 in immature B cells. J. Immunol. 169, 3783–3792.

Johnson, A.D., Fitzsimmons, D., Hagman, J., Chamberlin, H.M., 2001. EGL-38 Pax regulates the *ovo*-related gene *lin-48* during *Caenorhabditis elegans* organ development. Development 128, 2857–2865.

Jumlongras, D., Lin, J.Y., Chapra, A., Seidman, C.E., Seidman, J.G., Maas, R.L., Olsen, B.R., 2004. A novel missense mutation in the paired domain of PAX9 causes non-syndromic oligodontia. Hum. Genet. 114, 242–249.

Jun, S., Desplan, C., 1996. Cooperative interactions between paired domain and homeodomain. Development 122, 2639–2650.

Jun, S., Wallen, R.V., Goriely, A., Kalionis, B., Desplan, C., 1998. Lune/eye gone, a Pax-like protein, uses a partial paired domain and a homeodomain for DNA recognition. Proceedings of the National Academy of Sciences of the United States of America 95, 13720–13725.

Kozmik, Z., Czerny, T., Busslinger, M., 1997. Alternatively spliced insertions in the paired domain restrict the DNA sequence specificity of Pax6 and Pax8. Eur. Mol. Bio. Organ. J. 16, 6793–6803.

Maduro, M., Pilgrim, D., 1995. Identification and cloning of unc-119, a gene expressed in the *Caenorhabditis elegans* nervous system. Genetics 141, 977–988.

Mello, C.C., Kramer, J.M., Stinchcomb, D., Ambros, V., 1991. Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. Eur. Mol. Bio. Organ. J. 10, 3959–3970.

Miranda, G.A., Villalvazo, M., Galic, Z., Alva, J., Abrines, R., Yates, Y., et al., 2002. Combinatorial regulation of the murine RAG-2 promoter by Sp1 and distinct lymphocyte-specific transcription factors. Mol. Immun. 38, 1151–1159.

Roberts, E.C., Deed, R.W., Inoue, T., Norton, J.D., Sharrocks, A.D., 2001. Id helix-loop-helix proteins antagonize pax transcription factor activity by inhibiting DNA binding. Mol. Cell. Biol. 21, 524–533.

Sewell, S.T., Zhang, G., Uttam, A., Chamberlin, H.M., 2003. Developmental patterning in the *Caenorhabditis elegans* hindgut. Dev. Biol. 262, 88–93.

Sigvardsson, M., Clark, D.R., Fitzsimmons, D., Doyle, M., Akerblad, P., Breslin, T., et al., 2002. Early B-cell factor, E2A, and Pax-5 cooperate to activate the early B cell-specific mb-1 promoter. Mol. Cell. Biol. 22, 8539–8551.

Sulston, J.E., Hodgkin, J., 1988. In: Wood, W.B. (Ed.), The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 587–606.

Tang, H.K., Chao, L.Y., Saunders, G.F., 1997. Functional analysis of paired box missense mutations in the PAX6 gene. Hum. Mole. Gene. 6, 381–386.

Thompson, W., Rouchka, E.C., Lawrence, C.E., 2003. Gibbs Recursive Sampler: finding transcription factor binding sites. Nucl. Acids Res. 31, 3580–3585.

Trent, C., Tsuing, N., Horvitz, H.R., 1983. Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. Genetics 104, 619–647.

Vilain, C., Rydlewski, C., Duprez, L., Heinrichs, C., Abramowicz, M., Malvaux, P., et al., 2001. Autosomal dominant transmission of congenital thyroid hypoplasia due to loss-of-function mutation of PAX8. J. Clin. Endocrinol. Metabol. 86, 234–238.

Wang, X., Chamberlin, H.M., 2002. Multiple regulatory changes contribute to the evolution of the Caenorhabditis *lin-48 ovo* gene. Genes. Deve. 16, 2345–2349.

Wheat, W., Fitzsimmons, D., Lennox, H., Krautkramer, S.R., Gentile, L.N., McIntosh, L.P., Hagman, J., 1999. The highly conserved beta-hairpin of the paired DNA-binding domain is required for assembly of Pax-Ets ternary complexes. Mol. Cell. Biol. 19, 2231–2241.

Xu, W., Rould, M.A., Jun, S., Desplan, C., Pabo, C.O., 1995. Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations. Cell 80, 639–650.

Xu, H.E., Rould, M.A., Xu, W., Epstein, J.A., Maas, R.L., Pabo, C.O., 1999a. Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes. Dev. 13, 1263–1275.

Xu, P.X., Zhang, X., Heaney, S., Yoon, A., Michelson, A.M., Maas, R.L., 1999b. Regulation of Pax6 expression is conserved between mice and flies. Development 126, 383–395.