Null Mutations in the *lin-31* Gene Indicate Two Functions During *Caenorhabditis elegans* Vulval Development

Leilani M. Miller, Heather A. Hess,[1] David B. Doroquez and Noelle M. Andrews

Department of Biology, Santa Clara University, Santa Clara, California 95053

Manuscript received November 16, 1999
Accepted for publication August 3, 2000

ABSTRACT

The *lin-31* gene is required for the proper specification of vulval cell fates in the nematode *Caenorhabditis elegans* and encodes a member of the winged-helix family of transcription factors. Members of this important family have been identified in many organisms and are known to bind specific DNA targets involved in a variety of developmental processes. DNA sequencing of 13 *lin-31* alleles revealed six nonsense mutations and two missense mutations within the DNA-binding domain, plus three deletions, one transposon insertion, and one frameshift mutation that all cause large-scale disruptions in the gene. The missense mutations are amino acid substitutions in the DNA-binding domain and probably disrupt interactions of the LIN-31 transcription factor with its DNA target. In addition, detailed phenotypic analysis of all 19 alleles showed similar penetrances for several characteristics examined. From our analysis we conclude: (1) the null phenotype of *lin-31* is the phenotype displayed by almost all of the existing alleles, (2) the DNA-binding domain plays a critical role in LIN-31 function, and (3) direct screens for multivulva and vulvaless mutants will probably yield only null (or strong) alleles of *lin-31*.

PROTEINS belonging to the winged-helix family of transcription factors appear to play important roles in many developmental processes, including establishment of body axes, cell fate specification, differentiation of tissues, maintenance of cellular differentiation, and tumorigenesis (see Kaufmann and Knöchel 1996 for review of winged-helix proteins). The LIN-31 protein, a member of this winged-helix family, is required for the proper specification of vulval cell fates during *Caenorhabditis elegans* vulval development (Miller *et al.* 1993). During vulval development, six vulval precursor cells (VPCs), which are initially equivalent in developmental potential, differentiate and adopt one of three possible vulval cell fates (for review, see Greenwald 1997). The different fates of the six VPCs (P3.p–P8.p) are determined in response to several extracellular signals and are adopted in the pattern 3°-3°-2°-1°-2°-3°. Cells adopting a 1° or 2° (induced) cell fate give rise to the cells that form the vulva, while cells that adopt a 3° (uninduced) cell fate give rise to cells that fuse with the hypodermal syncytium that surrounds most of the worm. Biochemical and molecular studies of the anchor cell signaling pathway have revealed a conserved receptor tyrosine kinase (RTK)/Ras/mitogen-activated protein (MAP) kinase pathway (for review, see Kornfeld 1997). Mutations in genes belonging to this RTK signaling pathway cause the misspecification of vulval cell fates and typically fall into two broad phenotypic classes, vulvaless (Vul) or multivulva (Muv). In Vul mutants, all six VPCs typically adopt nonvulval (3°) cell fates, due to a less active or completely inactive RTK signaling pathway. In Muv mutants, most or all of the VPCs adopt vulval fates (1° or 2°), due to inappropriate activation of the RTK pathway. In the dissecting microscope, a Muv phenotype is manifested by multiple ectopic pseudovulvae (protrusions) along the ventral side of the animal. Mutations in parallel or modulating pathways may also cause a Vul or Muv phenotype (for review, see Greenwald 1997).

Animals carrying mutations in the *lin-31* gene, however, display an interesting phenotype that is unique among vulval development mutants (Miller *et al.* 1993). In contrast to the Muv and Vul mutations, which affect all six VPCs in the same way, these mutations result in a deregulated cell fate phenotype in which each of the six VPCs can adopt any one of the three possible cell fates. Furthermore, opposite cell fate transformations are often observed in the same *lin-31* mutant animal, in which a cell distal to the anchor cell expresses an induced cell fate (1° or 2°), while a proximal cell simultaneously expresses an uninduced (3°) cell fate. These opposite cell fate transformations indicate that, in *lin-31* mutants, the choice of cell fate is no longer strictly dependent on proximity to the anchor cell. This deregulation of vulval cell fate is manifested by variable Vul and/or Muv phenotypes. Inappropriate expression of an induced fate at P3.p, P4.p, or P8.p results in an ectopic pseudovulva (protrusion), thus causing the ani-

Corresponding author: Leilani M. Miller, Department of Biology, 500 El Camino Real, Santa Clara University, Santa Clara, CA 95053. E-mail: lmiller@scu.edu

[1] *Present address:* Yale University School of Medicine, 333 Cedar St., SHM CE30, P.O. Box 208024, New Haven, CT 06520-8024.

mal to appear Muv when observed in a dissecting microscope. Conversely, inappropriate expression of a 3° nonvulval fate by P5.p, P6.p, or P7.p can result in a Vul animal. Some *lin-31* mutant animals express both Vul and Muv phenotypes simultaneously. The observation that all six VPCs require wild-type *lin-31* activity suggested that *lin-31* has alternative functions (Miller *et al.* 1993). Thus, if the deregulated mutant phenotype does indeed represent the *lin-31* null phenotype, the LIN-31 protein would appear to have two functions: to promote vulval induction (in P5.p, P6.p, and P7.p) and to inhibit it (in P3.p, P4.p, and P8.p). However, because incomplete penetrance is a predictable consequence of a deregulation phenotype, it has been difficult to unequivocally establish the null phenotype for *lin-31* using phenotypic analysis alone.

Epistasis experiments placed the *lin-31* gene at or near the end of the RTK signaling pathway specifying vulval cell fate, and sequence analysis showed it to encode a winged-helix transcription factor (Miller *et al.* 1993). Subsequently, immunolocalization experiments revealed that LIN-31 is expressed in the VPCs during the L2 and L3 larval stages and coimmunoprecipitation experiments showed that it forms a heterodimer *in vivo* with the LIN-1 Ets transcription factor that is disrupted upon phosphorylation of LIN-31 by MAP kinase (Tan *et al.* 1998). These results allowed expansion of the model that LIN-31 has two functions: to promote vulval induction (in its phosphorylated, undimerized form) and to inhibit induction (in its unphosphorylated, dimerized form).

The LIN-31 protein contains a winged-helix DNA-binding domain, an acidic region, a serine-rich region, and a proline-rich region (Miller *et al.* 1993). In this study, DNA sequence analysis of 13 *lin-31* mutant alleles revealed many early nonsense mutations and several large-scale disruptions of the gene, unequivocally establishing the null phenotype. DNA sequence analysis also revealed two missense mutations in the winged-helix DNA-binding domain. Although winged-helix proteins are clearly involved in many important developmental regulatory mechanisms, few missense mutations within this DNA-binding domain have been reported, making it difficult to confirm their functions *in vivo*. The identification of these two missense mutations suggests that the winged-helix DNA-binding domain is critical for LIN-31 function *in vivo*.

MATERIALS AND METHODS

General methods and strains: Standard techniques were used for maintenance and handling of *C. elegans* strains (Brenner 1974). Animals were grown at 20° unless otherwise indicated. The animals described as wild type were *C. elegans*, variety Bristol, strain N2 (Brenner 1974). Unless otherwise noted, all *C. elegans* mutations are described in Wood (1988). Mutations and genetic markers were: LGII: *lin-31(e1750, n301, n376, n428, n429, n435, n762, n1048, n1049, n1050, n1053)* (Ferguson and Horvitz 1985), *lin-31(n1282, n1290, n1291)* (Kim and Horvitz 1990), *lin-31(ga9, ga10, ga37)* (Miller *et al.* 1993), *lin-31*(*ga57, ga70)* (gifts from D. Eisenmann); LGIII: *unc-32(e189)*; LGV: *him-5(e1490)*.

Male mating efficiency: To test for male mating efficiency, double mutant strains containing a *lin-31* allele and *him-5(e1490)* were constructed. Each *lin-31* mutant allele was tested in a cross consisting of 12 *lin-31; him-5* L4 males and 6 *unc-32* hermaphrodites. Successful mating was indicated by the presence of non-Unc cross-progeny. Each allele was tested in at least two crosses. Male mating efficiency (ME) ratings were assigned according to Wood (1988), with ME0 indicating no male mating and ME3 indicating wild-type male mating levels.

L2 division phenotype: The presence of L2 divisions in the vulval precursor cell lineage was monitored with Nomarski optics in at least nine animals from each *lin-31* strain scored. Animals in the L2 larval stage were identified on the basis of size using a dissecting microscope. These animals were mounted for Nomarski observation as described in Wood (1988), except that 5 mm NaAzide was used to immobilize the larvae. Larval stage was confirmed by gonad size, and VPCs were spot-checked for inappropriate L2 divisions. To prevent potential confusion between inappropriate L2 VPC divisions and normal early L3 VPC divisions, only animals whose gonad arms had not yet extended past P5.p and P7.p were scored. Although this approach ensures that only animals in the L2 stage are scored, it necessarily underestimates the number of possible inappropriate L2 VPC divisions since the entire L2 stage is not scored in each animal. Mutant strains were scored as positive for the L2 division phenotype if at least one animal displayed L2 divisions.

Cloning of *lin-31* alleles: Genomic DNA for 13 *lin-31* alleles was cloned by a PCR method adapted from Williams *et al.* (1992). For most reactions, the PCR primers used were LM34 (5′-GGGAATTCATGAACAGGAAAATGTGTGGC-3′), and LM35 (5′-CCGAATTCGCAAAACTCTATACATATTCCC-3′). Vent DNA polymerase (0.18 units) was used in each reaction. The LM34 and LM35 primers bracket the 3901-bp genomic coding region of *lin-31* (from 27 nucleotides before the start codon to 70 nucleotides past the stop codon). For some alleles where general location of mutation was known (*e.g.*, *n1049*, *n376*, *ga10*, and *ga37*), primers internal to LM34 and LM35 were used. Gel-purified 3.9-kb fragments from two independent PCR reactions for each mutant allele were cloned using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA).

Sequencing of *lin-31* alleles: For point mutations and frameshift mutation, the entire 3.9-kb insert of each mutant allele clone was sequenced on one strand for two independent clones per mutant allele. When the same mutation was identified in both clones, the other strand was also sequenced in that region to confirm the change. Two independent clones were sequenced in the region of the *n1049* deletion. For each of the other deletion or insertion alleles, one clone was sequenced in the region of the mutation, and the size and position was confirmed by Southern analysis (*n376* and *ga37*; Miller *et al.* 1993) or PCR (*ga10*). Sequencing was performed on an ABI (Foster City, CA) automated DNA sequencer.

Immunofluorescence: Rabbit anti-LIN-31 antibodies were a generous gift from P. Tan and S. Kim (Tan *et al.* 1998). Fixed animals (Finney and Ruvkun 1990) were stained with a 1:100 dilution of anti-LIN-31 antibodies as described in Miller and Shakes (1995).

RESULTS

Sequence of *lin-31* mutant alleles: The *lin-31* alleles used in this study are independent and were obtained by several

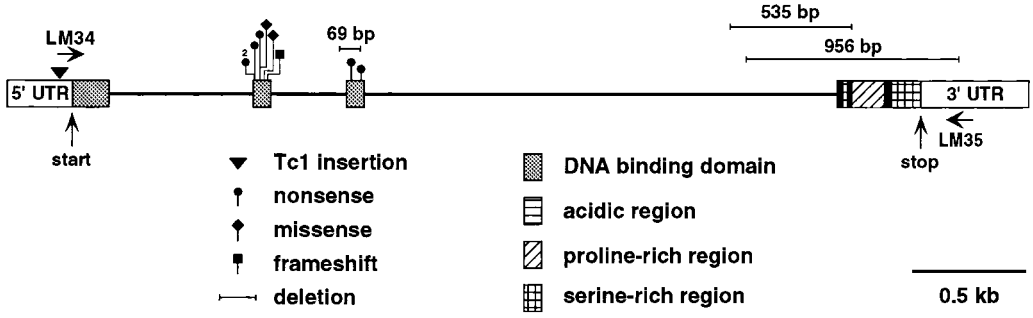

FIGURE 1.—Mutant alleles within the *lin-31* genomic region. Boxes indicate exons, horizontal lines between boxes indicate introns, and open boxes indicate nontranslated regions. The type of mutation is depicted by appropriate symbol. The sizes of deletions are indicated above deletion lines. DNA sequences that encode specific LIN-31 protein domains or regions are indicated by different hatching patterns. Arrows labeled LM34 and LM35 indicate primers used to clone most *lin-31* mutant alleles. The total length of this region is 4593 bp.

different screening methods. *e1750*, *n301*, *n376*, *n428*, *n429*, *n435*, *n762*, *n1048*, *n1049*, *n1050*, and *n1053* were isolated as multivulva animals (Muvs) in a general screen for Muv or Vul animals (FERGUSON and HORVITZ 1985). *ga10* was isolated in a screen for Muv animals (M. HERMAN, personal communication). *n1282*, *n1290*, *n1291*, *ga9*, and *ga37* were isolated as transposon-induced alleles in a mutator background (KIM and HORVITZ 1990). *ga57* and *ga70* were isolated in a screen for animals with protruding vulvae (EISENMANN and KIM 2000).

LIN-31 is a 237-amino-acid protein containing a DNA-binding domain, an acidic region, a serine-rich region, and a proline-rich region (see Figure 1). To elucidate possible functions of these domains, 3.9 kb of genomic DNA spanning the *lin-31* coding region was sequenced in 13 *lin-31* mutant alleles (see MATERIALS AND METHODS). Since Taq polymerase, which was used to clone the mutant alleles, can itself induce mutations, two independent PCR clones from each noninsertion or nondeletion allele were sequenced. The indicated mutation in each case was the only change within the 3.9-kb cloned region exhibited by both clones. The sequence analysis of 13 *lin-31* alleles revealed six nonsense mutations, three deletions, one transposon insertion in the 5′ end of the gene, one frameshift mutation, and two missense mutations. In addition, *n1291* and *n1282*, previously characterized as Tc1 insertions in the 5′ end of the gene, and *n1290*, another mutator allele (MILLER *et al.* 1993), were not sequenced. The results of these experiments are shown in Figure 1 and summarized in Table 1.

Significantly, all six nonsense mutations map to the DNA-binding domain, with two of the alleles (*n1053* and *e1750*) representing identical changes. One deletion (*ga10*) is large, eliminating the carboxyl third of the LIN-31 protein, including all of the acidic, serine-rich, and proline-rich regions. Another deletion (*n376*) effectively removes (or alters) the entire carboxyl third of the protein due to the deletion of the last splice acceptor site. Finally, a small deletion (*n1049*) eliminates the last third of the DNA-binding domain. Three additional alleles (*n1050*, *n428*, and *n429*) are suspected to be large deletions of the entire coding region, as PCR clones were not obtained in several independent reactions using multiple combinations of primers. The Tc1 transposon insertion is located just before the start codon, and the frameshift mutation (*ga9*) mapped to the end of the DNA-binding domain. Finally, the two missense mutations (*ga57* and *n301*) mapped within the DNA-binding domain. Note that all but the two C-terminal deletions affect the DNA-binding domain (see Figure 1), although the sequenced Tc1 insertion is upstream of the DNA-binding domain and may not affect it.

Two putative DNA-binding mutants: DNA sequencing of *lin-31* showed that the N-terminal half of the LIN-31 protein consists of a DNA-binding domain similar to that of the rat transcription factor HNF-3 (MILLER *et al.* 1993). X-ray crystallography studies of the HNF-3γ/DNA cocrystal structure by CLARK *et al.* (1993) subsequently showed that the HNF-3γ DNA-binding domain has a winged-helix motif, which resembles the structure of histone H5. Its three α-helices in the amino-terminal half adopt a compact structure that presents the third (recognition) helix to the major groove, which interacts with its DNA target through both direct and water-mediated major and minor groove base contacts. The carboxyl-terminal half of the motif includes antiparallel β-structure and random coils ("wings") that interact with the minor groove. The DNA-binding domain of the LIN-31 protein displays a striking similarity to the DNA-binding domains of the fork head/HNF-3-related winged-helix family of proteins (Figure 2; MILLER *et al.* 1993; KAUFMANN and KNÖCHEL 1996). Thus, the overall three-dimensional structures of these DNA-binding domains should also be very similar.

From our sequencing studies, we identified two missense alleles of the *lin-31* gene, which are located in the conserved winged-helix DNA-binding domain (Figure 2 and Table 1). One mutant allele, *ga57*, results in the substitution of the amino acid isoleucine for the highly conserved asparagine at the end of the recognition helix (helix 3) within the DNA-binding domain. The HNF-3γ/DNA cocrystal structure revealed that this particular asparagine makes contact with its DNA target via hydrogen bonds with water molecules (CLARK *et al.* 1993). Changing it to a hydrophobic isoleucine presumably disrupts this interaction. The importance of this particu-

TABLE 1

Molecular analysis of *lin-31* mutations

Allele[a]	Mutagen	Type of mutation	DNA change	Amino acid change	Position in protein or gene[b,c]
n301	EMS	Missense	G to A	Arg^{62} to Gln^{62}	Middle of helix 3 in DBD
ga57	EMS	Missense	A to T	Asn^{68} to Ile^{68}	End of helix 3 in DBD
n1053	EMS	Nonsense	G to A	Trp^{57} to amber	Middle of DBD (same as *e1750*)
e1750	EMS	Nonsense	G to A	Trp^{57} to amber	Middle of DBD (same as *n1053*)
n435	EMS	Nonsense	C to T	Gln^{58} to ochre	Middle of DBD
n1048	EMS	Nonsense	C to T	Arg^{62} to amber	Middle of DBD
n762	EMS	Nonsense	G to A	Trp^{87} to amber	Near end of DBD
ga70	EMS	Nonsense	C to T	Arg^{108} to amber	End of DBD
ga37[d]	Mutator	Insertion			Just before start codon (−29 bp)
n1282[d]	Mutator	Insertion			Near 5′ end of gene
n1290[d]	Mutator	Insertion			Unknown
n1291[d,e]	Mutator	Insertion			Near 5′ end of gene
n1049	EMS	Deletion	69 bp		Removes last third of DBD (1216–1285)
n376	EMS	Deletion	535 bp		Removes last splice site (2960–3495)
ga10	EMS	Deletion	956 bp		Deletes all but DBD (3025–3981)
ga9[d]	Mutator	Frameshift	1 bp lost		End of DBD (1292)

DBD, DNA-binding domain.

[a] In addition, *n428*, *n429*, and *n1050* were not amplified by PCR and are likely to be deletions of the entire coding region (see text).

[b] The DNA-binding domain extends from amino acids 13 to 109 (Miller *et al.* 1993).

[c] For insertions and deletions, numbers refer to base pairs, with +1 = A of ATG encoding start codon (Miller *et al.* 1993).

[d] Transposon-induced alleles.

[e] Data from Miller *et al.* (1993).

lar asparagine side chain is underscored by its conservation in all but a few members of known winged-helix proteins (Kaufmann and Knöchel 1996). The other allele, *n301*, also lies within the DNA-binding domain (in the middle of helix 3) and results in the substitution of a glutamine for an extremely highly conserved arginine. Although the X-ray crystallography studies of HNF-3γ did not indicate that this arginine is specifically involved in binding DNA, this substitution may still affect how the mutant protein interacts with the DNA.

FIGURE 2.—Multiple sequence alignment of selected winged-helix DNA-binding domains. LIN-31 (Miller *et al.* 1993) was compared with the founding members of the winged-helix family: *fork head* (FKH), cloned from Drosophila (Weigel *et al.* 1989), and HNF-3α, HNF-3β, and HNF-3γ, cloned from rats (Lai *et al.* 1990, 1991). Also included in this alignment are several proteins from the same winged-helix subclass as LIN-31 (defined in Kaufmann and Knöchel 1996): FD4 from Drosophila (Häcker *et al.* 1992), FKH-4 from mouse (Kaestner *et al.* 1993), and XFD-5 from Xenopus (Lef *et al.* 1996), and several winged-helix DNA-binding domains from *C. elegans*: PHA-4 (Azzaria *et al.* 1996), PES-1 (Hope 1994), and DAF-16 (Lin*et al.*1997; Ogg *et al.* 1997). Lines indicate positions of α-helices (H1–H3), boldface lines indicate positions of β-pleated sheets (S1–S3), and wing-like loop regions are indicated by W1 and W2 (Clark *et al.* 1993). H3 is the DNA recognition helix. Asterisks indicate positions of the two *lin-31* missense alleles, *n301* and *ga57*, respectively. Shading indicates sequence identity to LIN-31. DNA-binding domain percentage identity to LIN-31 is indicated at the beginning of each alignment.

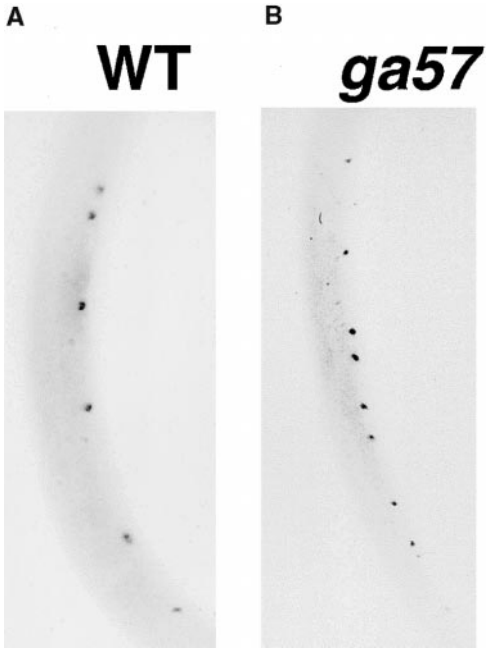

FIGURE 3.—LIN-31 expression patterns in wild-type (N2) and *ga57* animals. (A) Wild-type (N2) and (B) *lin-31(ga57)* mutant animals stained with anti-LIN-31 antibodies during the late L2 or early L3 larval stage. Figure shows midbody region. Dorsal is to the left, ventral to the right, anterior toward the bottom, and posterior toward the top. Figures are negative images of FITC immunofluorescence. Both animals show staining of LIN-31 protein within the nuclei of the VPCs.

This particular arginine is very likely to play an especially critical role in the function of these proteins, as it is completely conserved in all known winged-helix proteins (KAUFMANN and KNÖCHEL 1996).

It is important to determine whether these two mutant proteins are defective because they are no longer able to bind their DNA target or because the amino acid substitutions have destabilized the protein enough to cause unfolding and subsequent degradation. Using immunolocalization studies, we were able to rule out the second possibility by showing that the LIN-31 protein is present in the proper cells at the proper stage of development. Figure 3 shows wild-type and *lin-31(ga57)* animals stained with antibodies to LIN-31 during the late L2 or early L3 stage, when LIN-31 protein is known to be present in the nucleus of each vulval precursor cell (TAN *et al.* 1998). In these animals, the LIN-31 protein is observed in the nuclei of the VPCs. The additional staining cells in the *ga57* animal represent either extra VPCs that are often generated during the L2 larval stage in *lin-31* mutant animals or normal VPC divisions in an early L3 animal. *lin-31(n301)* animals displayed similar results (data not shown). Thus, both *ga57* and *n301* mutant strains express LIN-31 protein at the proper time and place, suggesting that their defects lie in their inability to bind the target DNA, either by disruption of specific protein:DNA interactions or through a folding defect that prevents binding but does not result in significant degradation of the protein.

Phenotype of *lin-31* alleles: Genetic criteria for establishing a null phenotype include ease of mutant isolation, similarity among strong alleles, phenotype *in trans* to a deficiency, and the ability to be suppressed by a nonsense suppressor. Previous studies showed allelic similarity for the vulval cell fate deregulation phenotype and also showed that a strong allele acts like a deficiency *in trans* to another allele (MILLER *et al.* 1993). To determine if *lin-31* alleles display allelic similarity for all known *lin-31* mutant phenotypes, we analyzed three phenotypic characteristics: the deregulation of VPC fates, male mating efficiency, and inappropriate cell division of VPCs during the L2 larval stage.

In *lin-31* deregulated mutants, VPCs that normally choose induced cell fates sometimes choose uninduced fates and VPCs that normally choose uninduced cell fates sometimes choose induced cell fates (MILLER *et al.* 1993). These mistakes in cell fate determination are manifested at the dissecting microscope level as Vul or Muv phenotypes, respectively. Because each vulval precursor cell in a *lin-31* mutant animal appears to make an independent decision, these animals can also sometimes display the Vul and Muv phenotypes simultaneously. All 19 *lin-31* mutant strains produce a high percentage of non-wild-type animals (Muv, Vul, or Muv/Vul), except for *n1291*, which expressed a somewhat weaker deregulation phenotype (see Table 2). In addition, each mutant strain also exibited both Muv and Vul phenotypes (data not shown), indicating that each allele is defective for appropriate adoption of both induced and uninduced cell fates.

In addition to the vulval precursor cell deregulation phenotype in hermaphrodites, *lin-31* mutant males are also mating incompetent (FERGUSON and HORVITZ 1985; MILLER *et al.* 1993), probably due to formation of defective mating spicules (S. BAIRD, personal communication). The 19 *lin-31* mutant strains were tested for the efficiency of male mating and all 19 were found to be completely mating defective, with a mating efficiency rating of "0" (ME0; Table 2).

A third phenotype that is observed in *lin-31* mutant animals is the occasional inappropriate cell division of a VPC during the L2 larval stage to give rise to either an extra VPC or an extra Pn.a-like neuroblast cell (MILLER *et al.* 1993). VPCs in wild-type animals are quiescent during the L2 larval stage. When cell division of a VPC during the L2 results in the creation of an extra VPC in *lin-31* mutant animals, its fate is also deregulated. At least one representative allele from each class of mutation (missense, nonsense, deletion, insertion, and frameshift) was analyzed for this phenotype. L2 divisions of VPCs were observed in all *lin-31* mutant strains tested (Table 2).

DNA sequence analysis supports the conclusion that

TABLE 2

Phenotypic analysis of *lin-31* mutant alleles

Allele	% deregulated[a,b]	ME	L2 div.[c]
Wild type	0	3	0/30
e1750	95 (n = 484)	0	ND
ga10	87 (n = 814)	0	3/15
ga37[d]	87 (n = 906)	0	2/24
ga57	80 (n = 648)	0	8/30
ga70	84 (n = 438)	0	ND
ga9[d]	98 (n = 1294)	0	2/21
n1048	96 (n = 510)	0	ND
n1049	96 (n = 1167)	0	4/18
n1050	94 (n = 1521)	0	ND
n1053	94 (n = 934)	0	24/37[e]
n1282[d]	89 (n = 1195)	0	ND
n1290[d]	95 (n = 1684)	0	ND
n1291[d]	54 (n = 676)	0	3/15
n301	91 (n = 1098)	0	5/21[e]
n376	93 (n = 1149)	0	1/9
n428	95 (n = 1011)	0	ND
n429	95 (n = 1162)	0	ND
n435	82 (n = 989)	0	ND
n762	91 (n = 525)	0	ND

ME, male mating efficiency; ND, not determined.

[a] % deregulated, Muv + Vul + Muv/Vul animals.

[b] Numbers from *e1750*, *ga9*, *n1049*, *n1050*, *n1282*, *n301*, *n428*, *n429*, and *n435* are from Miller *et al.* (1993) and were included in the table for completeness.

[c] Numbers indicate number of animals in which any L2 VPC divisions were observed over the total number of animals scored. Note: given the method of scoring (see materials and methods), this number is variably underestimated and should not be used to compare relative strength of alleles.

[d] Transposon-induced alleles.

[e] Some data published in Miller *et al.* (1993, 1996).

at least some of the *lin-31* mutant alleles represent the null phenotype. Since all 13 of these sequenced mutant alleles, as well as two previously mapped Tc1 insertions and the three deletions that likely delete the entire coding region, show similar mutant phenotypes, this suggests that LIN-31 protein function is similarly defective in 18 of the 19 mutant strains. The *n1291* strain was the only one to show a weaker phenotype in any of the phenotypic analyses. Its 54% vulval cell fate deregulation phenotype fell significantly below the next highest allele of 80% (see Table 2), indicating that it may represent a somewhat weaker allele. It was, however, indistinguishable from other alleles in the remaining two phenotypic analyses (male mating and inappropriate VPC division during the L2 larval stage). Earlier experiments characterized *n1291* as a Tc1 insertion in the 5′ end of the gene, resulting in a smaller mRNA (1.4 *vs.* 1.5 kb observed in wild-type animals; Miller *et al.* 1993). Thus, with the possible exception of *n1291*, all existing *lin-31* alleles display the null phenotype and are likely to be null mutations.

DISCUSSION

In this article, we report the molecular and phenotypic characterization of 19 mutant alleles of the gene that encodes the winged-helix transcription factor LIN-31. All 19 alleles displayed similar phenotypes, and DNA sequencing results indicated that most of these alleles must be null alleles. In addition, we show that two missense alleles in the DNA-binding domain are likely to disrupt the interaction of the transcription factor with its DNA target, revealing the critical importance of the DNA-binding domain to the function of the LIN-31 protein.

The null phenotype of *lin-31*: Due to the nature of the *lin-31* mutant phenotype (presence of both Muv and Vul animals and less than 100% penetrance for deregulation of VPCs), it has been difficult to unequivocally establish the null phenotype by phenotypic characteristics alone. In this article, allelic similarity was demonstrated for *three* phenotypic characteristics, whereas previous analyses had relied on only one phenotype. In addition, six early nonsense mutations, four insertions in the 5′ end of the gene, one early frameshift mutation, and at least three deletions were identified. Furthermore, P. Tan (personal communication) has shown, using antibody studies, that there is no detectable LIN-31 protein for at least two alleles, *n1053* and *ga37*.

With the establishment of the null phenotype, the transformation of both induced and uninduced vulval cell fates in *lin-31* null mutants indicates that LIN-31 plays an important role in the specification of both types of cell fates. This is consistent with models proposing that LIN-31 functions both to promote and to inhibit vulval induction (Miller *et al.* 1993; Tan *et al.* 1998). Thus, a cell lacking all *lin-31* activity would be incapable of correctly specifying either an induced or uninduced cell fate. This would require the cell to adopt a VPC fate in a stochastic manner, resulting in the deregulated pattern of cell fate specification observed in *lin-31* mutants. The adoption of a specific induced or uninduced cell fate in the absence of LIN-31 rather than a default or intermediate cell fate can be explained by a model in which target genes of the LIN-31 transcription factor are autoregulated, and that stochastic perturbances can cause one of the target proteins to rise above a threshold level, triggering autoactivation and subsequent deactivation of other target genes, thus "locking in" a particular cell fate (Miller *et al.* 1993; Tan *et al.* 1998).

A winged-helix DNA-binding domain is essential for function: While the winged-helix family of DNA-binding proteins is clearly an important one in developmental biology, most of its eukaryotic members have been identified simply by sequence homology, and very few missense mutations within the DNA-binding domain have been identified through genetic screens. Thus, definitive proof of this domain's *in vivo* importance in eukaryotes is limited. For example, site-directed mutation stud-

ies in HNF3α have identified specific residues in the DNA-binding domain important for its stability (Stevens *et al.* 2000), but those studies were limited to transcription activation assays in tissue culture and *in vitro* binding experiments. Other studies have shown that targeted null mutations in HNF3α and HNF3β are lethal *in vivo*, but those experiments deleted almost the entire protein-coding region (Ang and Rossant 1994; Weinstein *et al.* 1994) or at least the entire DNA-binding domain (Kaestner *et al.* 1999). Furthermore, some transcription factors, such as *fushi tarazu*, are capable of regulating transcription without a DNA-binding domain (Fitzpatrick *et al.* 1992). In this study, we show that the winged-helix domain is indispensable to the function of the LIN-31 protein.

One indication that the LIN-31 DNA-binding domain is critical for function is the number of mutations that lie within it (Figure 1). Eleven of the 16 alleles with known positions are located within the DNA-binding domain, including two missense mutations and a small deletion. The most compelling argument for the importance of the DNA-binding domain is the existence of the two missense mutations within this domain that leave the rest of the protein intact. Since these mutations lie in or near the recognition helix and since LIN-31 protein is not degraded due to unfolding caused by the amino acid substitutions in these two mutant strains, it is likely that the only defect these mutant proteins display is an inability to bind the correct target DNA. Furthermore, at least one of these missense alleles represents a change in an amino acid that, in a similar winged-helix homolog, is known to make contact with the target DNA. The null phenotype of this allele strongly suggests that the ability to bind target DNA is a critical function of the LIN-31 protein. Finally, since the phenotypes of all *lin-31* alleles affecting the DNA-binding domain are equivalent to the *lin-31* null phenotype, it is clear that this winged-helix domain is essential for protein function. The importance of this domain in LIN-31 is consistent with a study showing that a missense mutation in helix 1 of the DNA-binding domain of a human winged-helix protein (TTF-2) results in impaired DNA binding and loss of transcriptional function (Clifton-Bligh *et al.* 1998).

Standard genetic screens yield only null alleles: All six existing *lin-31* nonsense mutations map to the DNA-binding domain, located in the N-terminal half of the protein. Interestingly, no late nonsense mutations were observed, indicating that a LIN-31 protein truncated late in the protein may not result in a phenotype detectable by standard Muv and Vul screens. While there are more codons in the N-terminal half of the LIN-31 protein that are capable of being mutated to a stop codon in a single step (37 *vs.* 27 in the C-terminal half), this difference does not adequately explain the complete absence of late nonsense mutations. The most likely explanation for the lack of nonsense mutations in the latter part of the protein is that a mutation leaving more than half of LIN-31 intact results in a protein with some residual function. These weak alleles would not have been detected in a standard screen for Muvs or Vuls, since the penetrance for either phenotype would probably be well below the 31–76% Muv penetrance or the 15–41% Vul penetrance (L. Miller, unpublished results) exhibited by null alleles. Thus, the isolation of weak or other nonnull *lin-31* alleles awaits the development of nonstandard mutant screening procedures or *in vitro* mutagenesis experiments. Such studies (in progress) may reveal the functional domains responsible for the two different activities of LIN-31 postulated by the model discussed above.

We thank L. Corbett and D. Schatz for help with DNA sequencing; P. Tan and S. Kim for the generous gift of LIN-31 antibodies; S. Kim, D. Eisenmann, and M. Herman for *lin-31* alleles; K. Warren for identification of the *ga70* deletion; J. Holzman, K. Warren, R. Childers, L. Treat, and K. Donnelly for help with male mating experiments; members of the Kim lab for antibody advice; and P. Schatz, D. Preuss, D. Eisenmann, P. Tan, J. Grainger, C. Stephens, and B. Meyer for comments on the manuscript. This work was supported by National Institutes of Health (NIH) research grant 1 R15 GM52636-01, National Science Foundation CAREER grant IBN-9600864, and the Clare Boothe Luce Foundation. Some nematode strains were supplied by the Caenorhabditis Genetics Center, which is supported by a contract between the NIH National Center for Research Resources and the University of Minnesota.

LITERATURE CITED

Ang, S.-L., and J. Rossant, 1994 *HNF-3*β is essential for node and notochord formation in mouse development. Cell **78:** 561–574.

Azzaria, M., B. Goszczynski, M. A. Chung, J. M. Kalb and J. D. McGhee, 1996 A fork head/HNF-3 homolog expressed in the pharynx and intestine of the *Caenorhabditis elegans* embryo. Dev. Biol. **178:** 289–303.

Brenner, S., 1974 The genetics of *Caenorhabditis elegans.* Genetics **77:** 71–94.

Clark, K. L., E. D. Halay, E. Lai and S. K. Burley, 1993 Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature **364:** 412–420.

Clifton-Bligh, R. J., J. M. Wentworth, P. Heinz, M. S. Crisp, R. John *et al.*, 1998 Mutation of the gene encoding human TTF-2 associated with thyroid agenesis, cleft palate and choanal atresia. Nat. Genet. **4:** 399–401.

Eisenmann, D. M., and S. K. Kim, 2000 Protruding vulva mutants identify novel loci and Wnt signaling factors that function during *C. elegans* vulva development. Genetics **156:** 1097–1116.

Ferguson, E. L., and H. R. Horvitz, 1985 Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans.* Genetics **110:** 17–72.

Finney, M., and G. Ruvkun, 1990 The *unc-86* gene product couples cell lineage and cell identity in *C. elegans.* Cell **63:** 895–905.

Fitzpatrick, V. D., A. Percival-Smith, C. J. Ingles and H. M. Krause, 1992 Homeodomain-independent activity of the fushi tarazu polypeptide in Drosophila embryos. Nature **356:** 610–612.

Greenwald, I., 1997 Development of the vulva, pp. 519–541 in *C. elegans II*, edited by D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Häcker, U., U. Grossniklaus, W. J. Gehring and H. Jäckle, 1992 Developmentally regulated *Drosophila* gene family encoding the fork head domain. Proc. Natl. Acad. Sci. USA **89:** 8754–8758.

Hope, I. A., 1994 PES-1 is expressed during early embryogenesis in *Caenorhabditis elegans* and has homology to the fork head family of transcription factors. Development **120:** 505–514.

Kaestner, K. H., K. H. Lee, J. Schlöndorff, H. Hiemisch, A. P. Monaghan *et al.*, 1993 Six members of the mouse forkhead gene family are developmentally regulated. Proc. Natl. Acad. Sci. USA **90:** 7628–7631.

Kaestner, K. H., J. Katz, Y. Liu, D. J. Drucker and G. Schütz, 1999 Inactivation of the winged helix transcription factor *HNF3* α affects glucose homeostasis and islet glucagon gene expression in vivo. Genes Dev. **13:** 495–504.

Kaufmann, E., and W. Knöchel, 1996 Five years on the wings of fork head. Mech. Dev. **57:** 3–20.

Kim, S. K., and H. R. Horvitz, 1990 The *Caenorhabditis elegans* gene *lin-10* is broadly expressed while required specifically for the determination of vulval cell fates. Genes Dev. **4:** 357–371.

Kornfeld, K., 1997 Vulval development in *Caenorhabditis elegans*. Trends Genet. **13:** 55–61.

Lai, E., V. R. Prezioso, E. Smith, O. Litvin, R. H. Costa *et al.*, 1990 HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. Genes Dev. **4:** 1427–1436.

Lai, E., V. R. Prezioso, W. F. Tao, W. S. Chen and J. E. Darnell, Jr., 1991 Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the Drosophila homeotic gene fork head. Genes Dev. **5:** 416–427.

Lef, J., P. Dege, M. Scheucher, V. Forsbach-Birk, J. H. Clement *et al.*, 1996 A fork head related multigene family is transcribed in *Xenopus laevis* embryos. Int. J. Dev. Biol. **40:** 245–253.

Lin, K., J. B. Dorman, A. Rodan and C. Kenyon, 1997 *daf-16*: an HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans* [see comments]. Science **278:** 1319–1322.

Miller, D. M., and D. C. Shakes, 1995 Immunofluorescence microscopy, pp. 365–394 in *Methods in Cell Biology, Volume 48: Modern Biological Analysis of an Organism*, edited by H. F. Epstein and D. C. Shakes. Academic Press, San Diego.

Miller, L. M., M. E. Gallegos, B. A. Morisseau and S. K. Kim, 1993 *lin-31*, a *Caenorhabditis elegans* HNF-3/fork head transcription factor homolog, specifies three alternative cell fates in vulval development. Genes Dev. **7:** 933–947.

Miller, L. M., D. A. Waring and S. K. Kim, 1996 Mosaic analysis using a *ncl-1(+)* extrachromosomal array reveals that *lin-31* acts in the Pn.p cells during *Caenorhabditis elegans* vulval development. Genetics **143:** 1181–1191.

Ogg, S., S. Paradis, S. Gottlieb, G. I. Patterson, L. Lee *et al.*, 1997 The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. Nature **389:** 994–999.

Stevens, K., L. Cirillo and K. S. Zaret, 2000 Creating temperature-sensitive winged helix transcription factors: amino acids that stabilize the DNA binding domain of HNF3. J. Biol. Chem. **275:** 30471–30477.

Tan, P. B., M. R. Lackner and S. K. Kim, 1998 MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN-31 WH transcription factor complex during *C. elegans* vulval induction. Cell **93:** 569–580.

Weigel, D., G. Jürgens, F. Kuttner, E. Seifert and H. Jäckle, 1989 The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. Cell **57:** 645–658.

Weinstein, D. C., A. R. Altaba, W. S. Chen, P. Hoodless, V. R. Prezioso *et al.*, 1994 The winged-helix transcription factor *HNF-3*β is required for notochord development in the mouse embryo. Cell **78:** 575–588.

Williams, B. D., B. Schrank, C. Huynh, R. Shownkeen and R. H. Waterston, 1992 A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. Genetics **131:** 609–624.

Wood, W. B., 1988 *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Communicating editor: R. K. Herman