1992

Molecular Basis of Loss-of-Function Mutations in the glp-1 Gene of Caenorhabitis elegans

Voula Kodoyianni
University of Wisconsin- Madison

Eleanor M. Maine
Syracuse University

Judith Kimble
University of Wisconsin-Madison

Recommended Citation

Kodoyianni, Voula; Maine, Eleanor M.; and Kimble, Judith, "Molecular Basis of Loss-of-Function Mutations in the glp-1 Gene of Caenorhabitis elegans" (1992). *Biology*. 15.
https://surface.syr.edu/bio/15

Molecular Biology of the Cell
Vol. 3, 1199–1213, November 1992

Molecular Basis of Loss-of-Function Mutations in the *glp-1* Gene of *Caenorhabitis elegans*

Voula Kodoyianni,* Eleanor M. Maine,*† and Judith Kimble*

*Department of Biochemistry and Laboratory of Molecular Biology, University of Wisconsin-Madison, Madison, Wisconsin 53706; and †Department of Biology,
Syracuse University, Syracuse, New York 13244

Submitted July 13, 1992; Accepted August 19, 1992

The *glp-1* gene encodes a membrane protein required for inductive cell interactions during development of the nematode *Caenorhabditis elegans.* Here we report the molecular characterization of 15 loss-of-function (*lf*) mutations of *glp-1*. Two nonsense mutations appear to eliminate *glp-1* activity; both truncate the *glp-1* protein in its extracellular domain and have a strong loss-of-function phenotype. Twelve missense mutations and one in-frame deletion map to sites within the repeated motifs of the *glp-1* protein (10 epidermal growth factor [EGF]-like and 3 LNG repeats extracellularly and 6 *cdc10*/SWI6, or ankyrin, repeats intracellularly). We find that all three types of repeated motifs are critical to *glp-1* function, and two individual EGF-like repeats may have distinct functions. Intriguingly, all four missense mutations in one phenotypic class map to the N-terminal EGF-like repeats and all six missense mutations in a second phenotypic class reside in the intracellular *cdc10*/SWI6 repeats. These two clusters of mutations may identify functional domains within the *glp-1* protein.

INTRODUCTION

Cell interactions regulate growth, differentiation, and pattern formation during the development of multicellular organisms. Yet the mechanisms by which cells communicate to regulate cell fates during development are poorly understood. The application of genetic tools to this question has uncovered numerous genes that mediate regulatory cell interactions (reviewed in Lambie and Kimble, 1991a; Greenwald and Rubin, 1992). Remarkably, similar proteins mediate regulatory cell interactions throughout phylogeny. For example, tyrosine kinase receptors of the epidermal growth factor (EGF) receptor class control vulval induction in *Caenorhabditis elegans* (Aroian *et al.*, 1990), determine the dorsal/ventral axis in *Drosophila* (Price *et al.*, 1989), and have apparently diverse roles during mammalian development (Adamson, 1990).

In this article, we focus on the *glp-1* gene, which regulates several cell interactions during the development of the nematode *C. elegans* (Austin and Kimble, 1987; Priess *et al.*, 1987). Figure 1 diagrams the major developmental defects of *glp-1* loss-of-function (*lf*) mutants. Normally, germline proliferation is induced by a somatic regulatory cell called the distal tip cell (Kimble and White, 1981). If the distal tip cell is ablated, then germ cells that normally are mitotic instead enter meiosis and differentiate. Similarly, animals that lack zygotic *glp-1* fail in germline proliferation; the few germ cells present at hatching enter meiosis prematurely and differentiate (Austin and Kimble, 1987; Priess *et al.*, 1987). In hermaphrodites, this germline defect leads to sterility because only a few sperm and no oocytes are produced; in males, it drastically reduces fertility. Normally, development of the anterior portion of the pharynx is induced during early embryogenesis (Priess and Thomson, 1987). For a few *glp-1* alleles, embryos that lack maternal *glp-1* can undergo morphogenesis normally but fail in pharyngeal induction; these embryos hatch from their eggshell but die as young larvae because they are unable to feed (Priess *et al.*, 1987). Embryos that lack maternal *glp-1* die. Although these embryos generate hundreds of cells, neither pharyngeal induction nor morphogenesis to form an elongated worm occurs (Austin and Kimble, 1987; Priess *et al.*, 1987). Based on these loss-of-function defects, it has been proposed that wild-type *glp-1* activity is essential for germline induction, pharyngeal induction, and embryonic viability and morphogenesis.

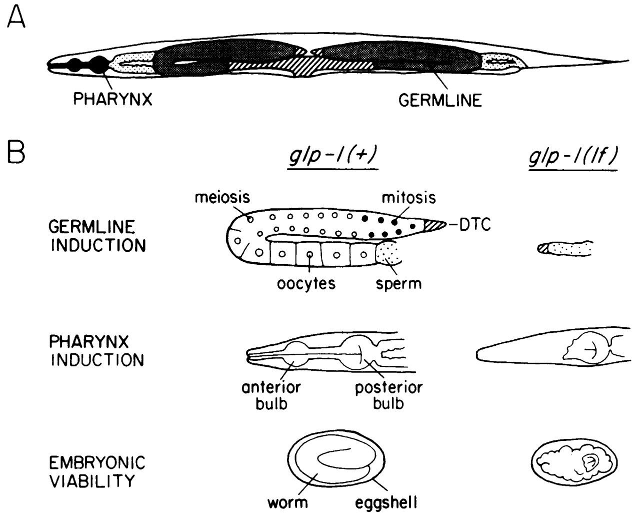

Figure 1. The *glp-1(lf)* mutant phenotypes. (A) Diagram of an adult wild-type hermaphrodite showing positions and relative sizes of germline (grey), pharynx (black), somatic gonad (striped), and intestine (stippled). (B) *glp-1* loss-of-function mutants have three developmental defects. Left, development in wild-type animals; right, defects in *glp-1* loss-of-function mutants. Germline induction, the wild-type germ line is organized into three regions (Kimble and White, 1981). Distally, germline nuclei are mitotic, more proximally, they enter meiosis, and most proximally, they undergo gametogenesis. Both the postembryonic proliferation of the germ line, from 2 to 2500 cells, and its organization into three regions depend on the presence of the distal tip cell (DTC) (Kimble and White, 1981). In *glp-1(lf)* mutants, germline mitoses arrest early during larval development and all germline precursor cells enter meiosis prematurely and differentiate as sperm (Austin and Kimble, 1987). Pharynx induction, the wild-type pharynx possesses anterior and posterior bulbs that are connected by an isthmus (Albertson and Thomsom, 1976). In embryos derived from *glp-1* homozygous mothers, pharyngeal induction fails and the anterior bulb of the pharynx is missing (Priess *et al.*, 1987). Embryonic viability, during the last half of embryogenesis, the wild-type embryo is transformed from a cluster of cells into a little worm. In embryos derived from *glp-1* homozygous mother, the embryo dies as an amorphous cluster of cells and does not hatch (Austin and Kimble, 1987; Priess *et al.*, 1987). Within the dying embryo, the posterior bulb of the pharynx can be detected, but the anterior bulb is missing.

The *glp-1* gene is a member of a small gene family that includes *lin-12* in *C. elegans* (Yochem *et al.*, 1988; Austin and Kimble, 1989; Yochem and Greenwald, 1989); Notch in *Drosophila* (Wharton *et al.*, 1985; Kidd *et al.*, 1986); and homologues in *Xenopus* (Coffman *et al.*, 1990), rat (Weinmaster *et al.*, 1991), and humans (Ellisen *et al.*, 1991). We call this the LNG gene family for its founding members, *lin-12*, Notch, and *glp-1*. The overall organization and amino acid sequences of the LNG proteins are strikingly similar. All are predicted to be membrane proteins that contain three types of repeated motifs (Wharton *et al.*, 1985; Yochem *et al.*, 1988; Austin and Kimble, 1989; Yochem and Greenwald, 1989; Coffman *et al.*, 1990; Ellisen *et al.*, 1991; Weinmaster *et al.*, 1991). Biochemical studies confirm that the Notch product is indeed a membrane protein and that its amino-terminus is extracellular and its carboxy-terminus intracellular (Johansen *et al.*, 1989; Kidd *et al.*, 1989). By analogy, other LNG proteins are likely to be membrane proteins oriented in similar fashion. Extracellularly, the LNG proteins possess a tandem array of EGF-like repeats, first found in the protein precursor of EGF (Gray *et al.*, 1983; Scott *et al.*, 1983). In addition, a cysteine-rich motif, which is unique to LNG proteins, is found in three copies in the region between the EGF-like repeats and the transmembrane domain. We call this cysteine-rich sequence the LNG motif; others have called it the Notch repeat (Wharton *et al.*, 1985), the B-type cysteine-rich repeat (Kidd *et al.*, 1989), or the LNR (*lin-12*/Notch repeat) motif (Yochem and Greenwald, 1989). Intracellularly, the LNG proteins have six copies of the *cdc10*/SWI6 repeat, a motif originally discovered in two yeast genes: *cdc10* is required in *Schizosaccharomyces pombe* for initiation of the cell cycle (Aves *et al.*, 1985) and SW16 is required in *Saccharomyces cerevisae* for HO gene transcription during G_1 (Breeden and Nasmyth, 1987a,b). More recently, *cdc10*/SWI6 repeats have been found in the protein ankyrin (as well as numerous other proteins; see DISCUSSION) and dubbed ankyrin repeats (Lux *et al.*, 1990). The conservation of structural motifs among these proteins suggests that they may function by a similar molecular mechanism.

The *lin-12*, Notch, and *glp-1* genes are all required for cell interactions that control cell fates (Greenwald *et al.*, 1983; Lehmann *et al.*, 1983; Austin and Kimble, 1987; Priess *et al.*, 1987). The *C. elegans* genes, *glp-1* and *lin-12*, control two superficially distinct types of cell–cell interactions: *glp-1*, as described above, mediates inductive interactions (Austin and Kimble, 1987; Priess *et al.*, 1987), whereas *lin-12* regulates interactions between cells of equivalent developmental potential so that they adopt distinct fates (Greenwald *et al.*, 1983). However, recent evidence makes it clear that *glp-1* and *lin-12* are interchangeable and therefore function by a similar mechanism (Lambie and Kimble, 1991b; Mango *et al.*, 1991; Seydoux *et al.*, 1990). The Drosophila gene, Notch, is required for an interaction among neuroepithelial cells that allows some cells to differentiate as epidermal cells and others to develop as neurons (Lehmann *et al.*, 1983). Genetic mosaic analyses of the tissues in which these genes function show that all three are required in the receiving rather than the signaling cell (Hoppe and Greenspan, 1986, 1990; Austin and Kimble, 1987; Seydoux and Greenwald, 1989; Heitzler and Simpson, 1991), suggesting that these presumptive membrane proteins may act as receptors.

In this article, we report the identification of molecular lesions in 15 recessive loss-of-function (*lf*) mutations of the *glp-1* gene. Our characterization of these *glp-1(lf)* mutations provides molecular evidence for the *glp-1* null phenotype and suggests that all three motif types, EGF-like, LNG, and *cdc10*/SWI6, are critical to *glp-1* function. Two phenotypic classes of *glp-1(lf)* mutants carry molecular defects that are clustered to specific regions of the *glp-1* protein. We discuss the implications of these results for identifying functional domains within the *glp-1* protein. In addition to providing insight into the functional domains of the *glp-1* protein, mutants with known molecular lesions provide a battery of reagents that can be used for further genetic and molecular analyses of *glp-1* function. A detailed understanding of how *glp-1* regulates cell fates in *C. elegans* will shed light on how similar proteins regulate cell interactions, not only in worms and flies, but also in higher vertebrates, including humans.

MATERIALS AND METHODS

Worm Strains

All *C. elegans* strains described in this paper were derived from the wild-type parent *C. elegans* var. Bristol strain N2 (Brenner, 1974). The mutations used are LGI, *smg-1(r861)*, LGIII, *glp-1* (*q46, q50, q158, q172, q175, q224, q231, q415, bn18, e2072, e2141, e2142, e2144, oz25* and *sy56*), *dpy-17* (*el64*), *dpy-19* (*el259*), *unc-32* (*el89*), *unc-69* (*e587*), and LGV, *him-5* (*el490*). All *glp-1* alleles (except *qDf2*) were obtained after ethyl methane sulfonate mutagenesis (EMS). *qDf2* is a γ-ray induced allele (Austin and Kimble, 1989). Most *glp-1* mutations are described in Austin and Kimble (1987) and Priess *et al.* (1987). *glp-1* (*q415*) is a new allele isolated by P. Balandyk in our lab, *bn18* was provided by S. Strome (Indiana University), *oz25* by T. Schedl (Washington University), and *sy56* by W. Boorstein and P. Sternberg (Cal. Tech.). Other mutations are described in Hodgkin *et al.* (1988). In addition, we used the translocation *eTI* (*III; V*) (Rosenbluth and Baillie, 1981), the free duplication *qDp3* (*III; f*) (Austin and Kimble, 1987), and the lethal mutation *qDf2* (Austin and Kimble, 1989).

Cloning of Mutant glp-1 Genes

Genomic DNA was prepared by the method of Emmons and Yesner (1984). For the six *glp-1* (*ts*) alleles, genomic DNA was isolated from homozygous stocks grown at permissive temperature. For alleles that could not be grown as homozygous strains, genomic DNA was prepared from hand-picked *glp-1(x)* homozygotes identified by one of the closely linked markers *unc-32* or *dpy-19*.

For *glp-1* alleles, *q224, q231, bn18, q158, q46, q172, e2142, qDf2*, genomic DNA was digested with *Xho* I and cloned into the *Sal* I site of EMBL3 λ phage vector. Recombinant phage clones containing the 9.3 kilobase (kb) *glp-1* insert were identified by screening the resulting libraries with ^{32}P-labeled *glp-1* genomic probes. This 9.3-kb *Xho* I *glp-1* genomic insert includes 1.7 kb of 5′ flanking sequence, the 3′ UTR (360 bases), and 110 bases of 3′ flanking sequence. The 5′ end of the *glp-1* coding region is known by primer extension; *glp-1* mRNA is trans-spliced at a splice acceptor site (TTCCAG) 72 nucleotides upstream of the translation initiation codon (Kodoyianni, unpublished data). In addition, the *glp-1* polyadenylation site and 3′ untranslated region (3′UTR) were defined by cDNA clones (Kodoyianni, unpublished data; Yochem and Greenwald, 1989).

For *glp-1* alleles *q50, q175, q415*, and *e2072*, four overlapping DNA fragments containing the entire *glp-1* genomic region (as described above but with only 60 bases of 3′ flanking sequence) were generated by polymerase chain reaction (PCR) (Saiki *et al.*, 1988) using Taq DNA polymerase (Perkin Elmer-Cetus, Norwalk, CT). For *glp-1* alleles *oz25, e2141, e2144*, and *sy56*, only specific regions of the *glp-1* gene were cloned by PCR (see below). The conditions used for amplification were as follows: 1–3 μg of genomic DNA was amplified for 35 cycles of 1 min at 94°C, 40 s at 55°C, and 3.5 min at 72°C for each cycle. The pairs of 5′ and 3′ primers used and their nucleotide coordinates (from Yochem and Greenwald, 1989) are shown below. Each pair of primers flanks restriction sites that were used to clone the PCR fragments.

Primer	Sequence	Nucleotide
VK22	5′ GGCATTAGGACCTTATAAGGT	505–525
VK21	5′ TAGAGGGAAATATGGACAGTG	2850–2830
EM7	5′ CACCAAGAGCTGCTCTAACA	2456–2475
VK3	5′ CAGGTTCACAGACACAG	5586–5570
VK4	5′ GGATACTGTGCCCATGA	5381–5397
VK1	5′ GTGACAACAGCTTGCCG	7230–7214
EM4	5′ GTGGTTTTGACGGTGGAG	6935–6952
EM1	5′ TCGGATCGAAATGAGGAG	9986–9969

Identification of glp-1(lf) Mutations

For most *glp-1* alleles, the approximate site of the lesion was first identified by one of two methods for mismatch detection, using either PCR amplified or cloned *glp-1* genomic DNA (see above); the exact base change was then determined by sequencing. For a few alleles, the lesion was identified directly by sequencing.

Mismatch Detection by Chemical Modification. Mutations in *glp-1* (*q50, q158, q172, q175, q415, e2072*, and *e2142*) were examined by the chemical modification method of Cotton *et al.* (1988) with minor modifications. Cloned wild-type *glp-1* genomic regions were used as probes: probe A, *Xho* I (804) to *Hin*dIII (2781); probe B, *Hin*dIII (2781) to *Eco*RI (5481); probe C, *Eco*RI (5481) to *Eco*RI (7021); and probe D, *Eco*RI (7021) to *Eco*RI (9964) (nucleotide numbering from Yochem and Greenwald, 1989). The following nucleotide numbers are given for reference: *glp-1* trans-splice acceptor site TTTCA<u>G</u>, 2440; translation initiation codon <u>A</u>UG 2513; translation termination codon <u>T</u>AA 9530 or 9539; polyadenylation site 9899. Fragments to be used as probes were first gel purified and then digested with restriction enzyme(s) to yield a ladder of fragments with sizes between 100 and 1000 base pair (bp) and finally labeled with T4 kinase and ^{32}P-ATP. End-labeled wild-type probe was then hybridized to a 100- to 200-fold excess of unlabeled mutant DNA, which had also been digested with the same restriction enzyme(s). The mutant DNA was either prepared from a recombinant phage or a PCR fragment (see above). The DNA hybrids (wild-type probe + mutant) were treated first with hydroxylamine (5M, pH 6.3), which modifies all mismatched cytosines, and then with piperidine, which cleaves at the modified cytosines. Analysis of the products on a 4% polyacrylamide gel revealed extra bands when the hybrids contained mismatched cytosines. Negative controls (i.e., wild-type probe hybridized to wild-type DNA) were run with every set of reactions. To rule out PCR artifacts, two independently amplified PCR fragments were examined in each experiment. To locate the site of a specific base substitution more accurately, the mismatch was usually repeated twice for each region, each time using different restriction enzymes to generate the fragment ladder.

Detection by DNA:RNA Mismatch. Mutations in *glp-1* (*q46, q172, bn18, q224, q231*, and *qDf2*) were mapped by the DNA:RNA mismatch detection method of Myers *et al.* (1985) with the following modifications. For this procedure, we used genomic DNA cloned in λ phage EMBL3 as described above. One hundred to 200 ng of DNA in 1 ml of Tris (pH 8.0), EDTA (TE) was mixed with 1 ml of 200 mM NaOH and incubated at 37° for 10–15 min. The DNA solution was neutralized by addition of 30 μl of hybridization solution, and the ^{32}P-labeled RNA probe was immediately added. Hybridizations were carried out at 47°C for 1 h. RNase A was typically used at a concentration of 1–4 μg/μl.

RNA probes were synthesized (Melton *et al.*, 1984) from a series of clones containing DNA spanning the wild-type *glp-1* gene (probes A to D above except that two smaller probes were used to subdivide the D region: probe D-1, *Eco*RI [7021] to *Pst* I [8288] and probe D-2, *Pst* I [8288] to *Eco*RI [9964]). A wild-type DNA control was included in each experiment; in addition, each probe (without prior DNA hybridization) was treated with RNase A to ensure that digestion was complete and did not produce artifactual bands.

Detection by Direct Sequencing. Mutations in *glp-1* (*q50, oz25, e2141, e2144,* and *sy56*) were identified directly by DNA sequencing of specific genomic regions.

1) *glp-1* (*q50*). For *q50*, the chemical modification mismatch method, which detects G → A,C,T and C → A,G,T nucleotide changes, yielded no aberrant fragments. A "reversed labeling" mismatch procedure (i.e., ^{32}P-labeled mutant DNA used as probe and wild-type DNA used as target DNA) that should detect A → C,G and T → C,G transitions, similarly yielded no changes. To find the T → A change in *glp-1*(*q50*), the *glp-1* coding sequence plus flanking regions were subcloned in duplicate from PCR fragments and the entire coding region sequenced.

2) *glp-1* (*oz25*). The *oz25* mutation was identified within the 1.5-kb *Eco*RI fragment (5481–7021) containing EGFL-7 to 10 and the 3 LNG repeats. This fragment was cloned from two separate PCR reactions and sequenced in duplicate.

3) *glp-1* (*e2141, e2144, sy56*). Because we had mapped three *ts* alleles to the *cdc10*/SWI6 repeats, we subcloned the 2.9-kb *Eco*RI fragment (7021–9964) containing these repeats from each of the three remaining *ts* alleles. Direct sequencing of clones from duplicate PCR reactions revealed the lesion sites.

Nucleic Acid Manipulations

Standard procedures were used as described (Sambrook *et al.*, 1989). PCR fragments were cut with restriction enzymes to generate cloning sites, purified on 0.8% agarose gels (Geneclean, BIO101 Inc., La Jolla, CA), and subcloned into pGEM-7z(+) (Promega, Madison, WI) or pBSKII(+) (Stratagene, La Jolla, CA). Plasmid DNA was isolated by alkaline lysis and purified either on a Qiagen column or by cesium chloride density gradient centrifugation. Templates were sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977) with Sequenase 2.0 (United States Biochemical, Cleveland, OH).

Phenotypic Characterization

We have characterized the phenotypes of several *glp-1* homozygotes (*glp-1(x)/glp-1(x)*; Table 1) and transheterozygotes (*glp-1(x)/glp-1(y)*; Table 3). Homozygotes characterized include four new alleles, *oz25* (isolated by T. Schedl), *sy56* (isolated by W. Boorstein), *bn18* (isolated by S. Strome), *q415* (isolated by P. Balandyk), and four previously reported alleles (*e2072, e2141, e2142,* and *e2144*). The new alleles were obtained after EMS mutagenesis, backcrossed, mapped to chromosome III, and shown to fail to complement *glp-1(q46)* (Schedl, Boorstein, and Strome, personal communication; Kodoyianni, Maine, and Kimble, unpublished data).

The effect of each mutant on germline proliferation was assayed in two ways. First, when germline proliferation fails, hermaphrodites are sterile. Therefore, fertility was scored by counting the number of embryos produced. Second, when germline proliferation fails, the number of germ cells is drastically reduced. Therefore, the total number of germline cells produced was determined by one of two methods. For some mutants, we first counted sperm in 4′6-diamidino-2-phenylindole·2 HCl (DAPI)-stained young adult homozygotes raised at 25°C and then calculated the number of germ cells produced as the number of sperm divided by four; for other mutants, we used Nomarski microscopy to count total germline nuclei in L3 animals raised at 25°C; the latter method was applied only to mutants in which all germline nuclei were primary spermatocytes by the L3 stage.

The effect of each mutant on embryogenesis was assayed by scoring the percentage of embryos that hatched; hermaphrodites were transferred every day to fresh plates through their reproductive period of adulthood. For each plate, embryos were counted first when the hermaphrodite was transferred and then again after 24 h (25°C) or 48 h (15°C). Embryos that fail to hatch were defective for both pharyngeal induction and embryonic morphogenesis. Embryos that hatched might either die, due to a failure in pharyngeal induction, or they might survive. The percentage of embryos with normal pharynx development was determined by counting how many embryos hatched and grew to adulthood; embryos that hatched but then died as young larvae were scored as defective in pharyngeal induction.

When our phenotypic characterization differed from those previously reported by Priess et al. (1987), we repeated it with new strains sent by J. Priess, with the same allele after repeated backcrossing, or both. We have no explanation for the differences that remain.

Construction and Analysis of glp-1 Transheterozygotes

General. All transheterozygotes reported in Table 3 were generated by the matings as described below. For all constructions, hermaphrodites were purged of their own sperm by allowing them to lay eggs at 20°C for several days before being mated. All transheterozygotes (except for those containing *q50*; see below) were homozygous for *unc-32*, a marker closely linked to *glp-1* on LGIII.

***glp-1* (+, *q415*, or *e2072*)/*glp-1*(*q175*).** Hermaphrodites of genotype *unc-32 glp-1*(+, *q415* or *e2072*)/*dpy-19 unc-69* were mated with *unc-32 glp-1* (*q175*)/*dpy-19 unc-69; him-5* males at the indicated temperature (Table 3). Unc non-Dpy F1 progeny, presumably of genotype *unc-32 glp-1(x)/unc-32 glp-1* (*q175*) were picked onto separate plates as L4s and scored for fertility and embryonic viability 24 or 48 h later. For controls, homozygous *unc-32 glp-1* hermaphrodites, derived as self progeny from heterozygous mothers, were scored concurrently.

***unc-36 glp-1* (*q50*)/*unc-32 glp-1*(*q175*).** *glp-1*(*q50*) arose on the translocation chromosome *eT1* (*unc-36*[*e873*]). Hermaphrodites of genotype *unc-36 glp-1*(*q50*)/*dpy-19 unc-69* were mated with *unc-32 glp-1* (*q175*)/*dpy-19 unc-69; him-5* males. All non-Dpy non-Unc F1 hermaphrodites were picked onto separate plates as L4s and scored 24 h later. *unc-36 glp-1* (*q50*) homozygotes were picked as F1 progeny of heterozygous mothers and were scored concurrently.

***unc-32 glp-1*(*e2142*)/*unc-32 glp-1* (*q175*).** Hermaphrodites of genotype *unc-32 glp-1* (*e2142*)/+ were raised at 15°C, purged of endogenous sperm, and mated either with *unc-32 glp-1* (*q175*)/*dpy-19 unc-69; him-5* or *unc-32 glp-1* (*e2142*)/+ males at either 15 or 25°C. Unc F1 hermaphrodite progeny from these crosses, presumed to be *unc-32 glp-1*(*e2142*)/*unc-32 glp-1* (*q175*) or, for the control, *unc-32 glp-1* (*e2142*), were transferred to individual plates as L4s and scored concurrently.

Examination of smg-1; glp-1(0) Double Mutants

Construction of *smg-1*(*r861*); *glp-1*(*q46*). *smg-1; him-5* males were mated with *glp-1*(*q46*)/*eT1* hermaphrodites. Male F1 progeny, presumably of genotype *smg-1/+; q46/+* or *eT1/+; him-5/+*, were crossed to *dpy-19 unc-69* hermaphrodites. Cross progeny of genotype *smg-1/+; q46/dpy-19 unc-69* were identified by progeny testing individual L4 hermaphrodites. Among self progeny of this double heterozygote, *smg-1; q46/dpy-19 unc-69* were identified by the genitalia defect typical of homozygous *smg-1* animals (Hodgkin *et al.*, 1989) and by progeny testing.

Construction of *smg-1*(*r861*); *unc-32*(*e189*) *glp-1* (*q175*). *smg-1; glp-1*(*q46*)/*dpy-19 unc-69* hermaphrodites were purged of endogenous sperm at 20°C and were mated to *unc-32 glp-1*(*q175*)/*dpy-19 unc-69; him-5* males. Non-Dpy non-Unc hermaphrodites were transferred onto individual plates as L4 larvae and individuals of genotype *smg-1/+; unc-32 glp-1*(*q175*)/*dpy-19 unc-69* identified by progeny testing. Among self progeny of this double heterozygote, *smg-1; q175/dpy-19 unc-69* were identified as described above.

Scoring *smg; glp-1*(*0*) double mutants. To look for *smg-1* effects on *q46* and *q175*, complete broods of three *smg-1; glp-1*(*q46*)/*dpy-19 unc-69* hermaphrodites and three *smg-1; unc-32 glp-1*(*q175*)/*dpy-19 unc69* hermaphrodites were scored for brood size and segregation of phenotypes at the expected ratios. In addition, heterozygous mothers were examined for vulval aberrations and fertility.

RESULTS

glp-1(lf) Mutations Fall into Three Phenotypic Classes

The 15 recessive *glp-1(lf)* alleles that we have examined can be grouped into three phenotypic classes.

1) Five *glp-1* alleles (*q46, q158, q172, q175,* and *oz25*) are nonconditional, fully penetrant, and eliminate germline induction (Table 1). In homozygotes, the two germline precursor cells that are present at hatching undergo one or two mitotic divisions, enter meiosis prematurely, and differentiate. Therefore, hermaphrodites are sterile and produce no embryos. These five alleles have the strongest *glp-1(lf)* phenotype known and are likely to be null mutations.

2) Six *glp-1* alleles (*q224, q231, e2141, e2144, bn18,* and *sy56*) are temperature sensitive (*ts*) in both germline and embryo (Table 1). When raised at permissive temperature (15°C), *glp-1(ts)* mutants appear wild-type, but when newly hatched larvae are shifted to restrictive temperature (25°C), germline induction fails and hermaphrodites are sterile. The germline effect of the *glp-1(ts)* mutations mimics that of putative *glp-1* null alleles. When hermaphrodites are raised at 15°C and shifted to 25°C as young adults, they are fertile, but their embryos die at restrictive temperature.

3) Four *glp-1* alleles (*q415, q50, e2072,* and *e2142*) have a more severe effect on embryogenesis than on germline induction (Table 1). These mutations range in strength. The weakest allele, *e2142*, is virtually wild-type for germline induction and temperature sensitive for embryonic lethality. The strongest allele, *q415*, is temperature sensitive for germline induction and nonconditional for embryonic lethality. Two alleles, *q50* and *e2072*, are intermediate in strength; germline induction fails in some but not all animals, but embryos always die (Table 1). Among those *q50* and *e2072* homozygotes that produce embryos, 80% have two wild-type ovotestes and 20% have one wild-type and one Glp-sterile ovotestis.

Mapping the Molecular Defects of glp-1 Mutations

To determine their approximate locations, 11 of the 15 *glp-1(lf)* mutations were examined by one of two methods for mismatch detection: RNAse A cleavage of RNA: DNA hybrids (Myers *et al.*, 1985) or chemical modification of DNA:DNA hybrids followed by cleavage (Cotton *et al.*, 1988). Probes that cover the *glp-1* transcription unit plus 1.7 kb of 5′ flanking sequence and 60 nucleotides of 3′ flanking sequence were used to roughly localize each mutation (see MATERIALS AND METHODS). Once a region was identified as distinct from wild-type, the exact base change was determined by directly sequencing. In this way, we identified the sequence changes in 10 *glp-1* mutants but not in *glp-1(q50)*. We therefore sequenced the entire *q50* coding region and found a single sequence change. In addition, the nucleotide changes of four other alleles were identified by sequencing a single region of *glp-1* without the aid of mismatch detection (see MATERIALS AND METHODS). Finally, we identified the breakpoint of a lethal deletion mutant, *qDf2*, which breaks within the *glp-1* gene (see MATERIALS AND METHODS).

The positions of the *glp-1(lf)* mutations within the *glp-1* protein are diagrammed in Figure 2 and the molecular changes are summarized in Table 2. During the course of these experiments, we confirmed the *glp-1* sequence reported previously (Yochem and Greenwald, 1989). In the discussion that follows, each amino acid motif is abbreviated (EGFL for EGF-like, LNG for *lin-12*/Notch/*glp-1*, SWI6 for *cdc10*/SWI6) and each repeat is numbered, starting from the amino-terminus. Hence, the second EGF-like repeat from the amino-terminus is called EGFL-2 and the fourth *cdc10*/SWI6 repeat from the amino-terminus is called SWI6-4. In addition, *glp-1* alleles are often referred to by allele number, e.g., *q46* rather than *glp-1(q46)*.

Molecular Basis of the Putative Null glp-1(lf) Mutations

The putative null alleles of *glp-1* include nonsense, missense, and deletion mutations (Table 2). The reference allele, *q46*, is an ochre mutation in the middle of LNG-2 (Table 2, Figure 3), and *q175* is an opal mutation at the end of EGFL-3 (Table 2; Figure 5). The *q46* and *q175* alleles are predicted to encode fragments of protein consisting of the amino-terminal 536 or 190 amino acids, respectively. These nonsense fragments should lack both the membrane spanning and cytoplasmic domains of the *glp-1* protein. Two other severe *glp-1(lf)* alleles, *q158* and *oz25*, are missense mutations; *q158* substitutes tyrosine for a conserved cysteine in LNG-2 (Figure 3), whereas *oz25* makes the same amino acid substitution in EGFL-7 (Figure 5A). Both of these cysteines are conserved residues among all members of the LNG protein family. The remaining strong *glp-1(lf)* allele, *q172*, was previously identified as a small deletion (Austin and Kimble, 1989). We find that *glp-1(q172)* is an in-frame deletion of 261 nucleotides (Table 2, Figure 2). The *q172* protein is predicted to lack the last 11 amino acids of LNG-1, which includes one of the conserved cysteines plus all of LNG-2 and LNG-3 (Figure 3).

The two nonsense mutations, *glp-1(q46)* and *glp-1(q175)*, are likely to represent complete loss of *glp-1* activity. The truncated proteins encoded by these two mutants should not be membrane associated and therefore should not function in signal transduction. Yet the *q175* fragment contains three complete EGF-like repeats and the *q46* fragment has all EGF-like repeats and one intact LNG motif. Either of these nonsense fragments might retain some function. One potential difficulty in detecting the activity of a nonsense fragment is that, in

Table 1. Phenotypes of *glp-1* loss-of-function mutants

Allele	Temperature (°C)[a]	Germline induction (%)[b]	Embryos hatch (%)[c]	Pharynx induction (%)[d]
Wild-type				
glp-1(+)[e]	15, 25	100	100	100
Putative null				
glp-1(q46, q158, q172, q175, oz25)[f]	15, 25	0	NA	NA
Temperature sensitive				
glp-1(q224, q231, bn18, e2141, e2144, sy56)[g]	15	100	100	100
	25	0	NA	NA
	15 → 25	NA	0	0
Partial loss-of-function				
glp-1(q415)[h]	15	100	0	0
	25	0	NA	NA
	15 → 25	NA	0	0
glp-1(q50)[i]	25	30	0	0
glp-1(e2072)[j]	25	23	80	0
glp-1(e2142)[k]	15	100	100	100
	25	100	90	0

NA, not applicable.

[a] Temperature at which mutant was raised from early embryogenesis through adulthood. To examine the embryonic phenotype of certain alleles, fourth larval stage hermaphrodites were shifted from 15 to 25°C (15 → 25); in these cases, oogenesis, fertilization, and embryogenesis all took place at 25°C.

[b] Percent homozygotes able to produce embryos (see MATERIALS AND METHODS).

[c] All embryos are self-progeny of homozygous *glp-1* hermaphrodites (see MATERIALS AND METHODS).

[d] For *e2072* and *e2142*, the embryos that hatch are missing the anterior pharynx and die as L1 larvae (see MATERIALS AND METHODS).

[e] Each wild-type hermaphrodite makes ~2500 germline descendants and produces ~300 embryos by self-fertilization.

[f] Putative null *glp-1* homozygotes have few germline nuclei (4–8) and produce no embryos (Austin and Kimble, 1987; this article). See Austin and Kimble (1987) for phenotypes of *glp-1(q46, q158, q172,* and *q175)*. For *glp-1(oz25)*, *oz25* homozygotes were Unc progeny of *unc-32 glp-1(oz25)/dpy-19 unc-69* mothers; all were sterile (n = 132, 25°C; n = 120, 15°C); germ cells totalled three to eight in homozygotes raised at 25°C (n = 10).

[g] Temperature sensitive *glp-1* homozygotes have a reduced number of germline nuclei (6–56) when raised at 25°C from hatching, but are virtually wild-type when raised at 15°C (Austin and Kimble, 1987; Preiss et al., 1987; this article). See Austin and Kimble (1987) for phenotypes of *glp-1(q224* and *q231)* and see Priess et al. (1987) for *glp-1(e2141* and *e2144)*. Our data for *glp-1(bn18, e2141, e2144, sy56)* homozygotes raised at 25°C are as follows: for *glp-1(bn18)*, germ cells totalled 20–56 (n = 12). For *glp-1(e2141* and *e2144)*, germ cells totalled 2–12 (n = 31) and 9–14 (n = 10), respectively. For *glp-1(sy56)*, germ cells totalled 60–80 (n = 10).

[h] For *glp-1(q415)*, germ cells in homozygotes raised at 25°C totalled 10–22 (n = 10); germ cells in homozygotes raised at 15°C totalled ~800. Older adults raised at 15°C were shown to continue germline mitoses by DAPI.

[i] Austin and Kimble (1987); among those *q50* homozygous hermaphrodites that produce embryos, 73% have two wild-type ovotestes and 27% have one wild-type and one Glp-sterile ovotestis (n = 22, examined with Nomarski optics) (this article).

[j] Priess et al. (1987); this article: *glp-1(e2072)*, homozygotes raised at 25°C fell into two distinct classes: 77% had few germline nuclei (the Glp phenotype) and 23% had nearly a wild-type number (n = 172); among those animals producing embryos, 80% had two ovotestes with a nearly wild-type number of germ cells and 20% had one nearly wild-type and one Glp-sterile ovotestis (n = 36, examined with Nomarski optics). The total number of germ cells in mutants with few germ cells was 6–11 (n = 16); the total number of germ cells in mutants with two nearly wild-type ovotestes was ~600–1000 (n = 4).

[k] Priess et al. (1987); this article: germ cells in *glp-1(el2142)* homozygotes raised at 25°C totalled 700–1200 (n = 10); older 25°C adults were examined after DAPI staining to confirm that mitosis continues as in wild-type (n = 20).

general, they are synthesized at a low level: mRNAs carrying nonsense mutations are much less abundant than wild-type mRNAs (Losson and Lacroute, 1979; Daar and Maquat, 1988; Pulak and Anderson, unpublished data). In *C. elegans,* mutations in any of six *smg* loci allow nonsense mRNAs to accumulate at wild-type levels (Hodgkin *et al.,* 1989; Pulak and Anderson, personal communication). We previously showed that the loss-of-function phenotype of another *glp-1* nonsense allele, *q35*, is suppressed by a mutation in *smg-1* (Mango *et al.,* 1991). Therefore, this suppression is likely to reflect an increase in the levels of both *glp-1(q35)* mRNA and protein.

To ask whether the nonsense fragments of *q46* and *q175* might have activity, we examined the phenotypes of *glp-1(q46* or *q175); smg-1* and *glp-1(q46* or *q175)/+; smg-1* animals. We found no effect of *smg-1* on the phenotype of either mutant. Among the self progeny of *glp-1(q46* or *q175)/+; smg-1* hermaphrodites, one-quarter were sterile as expected. Furthermore, we saw no

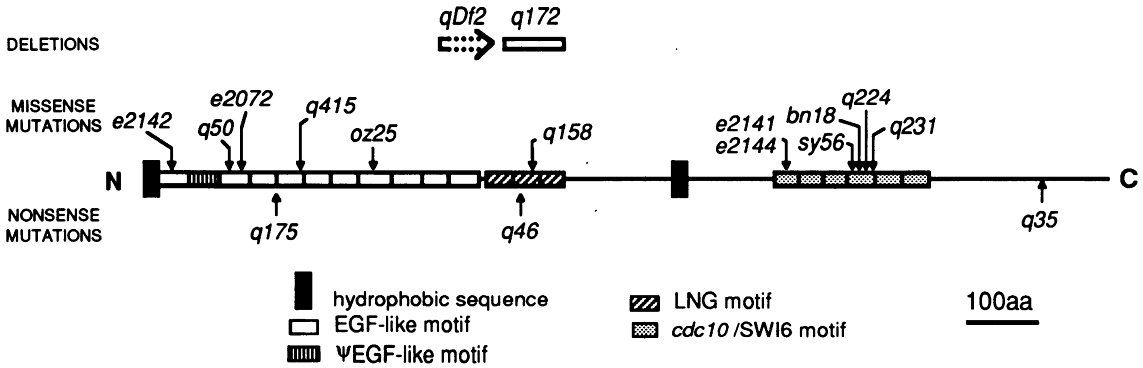

Figure 2. Molecular basis of *glp-1*(*lf*) mutations. The positions of 15 recessive loss-of-function *glp-1* mutations are shown on a schematic representation of the *glp-1* protein (1295 amino acids; see Table 1 for the associated mutant phenotypes and Table 2 for a list of codon changes). The hydrophobic sequence at the N-terminus is likely to be a signal sequence and that found in the middle of the protein has been predicted to be a transmembrane domain. In this diagram, the extracellular domain lies to the left of the transmembrane domain and contains two sets of cysteine-rich amino acid motifs: 10 copies of an "EGF-like" motif followed by three copies of the LNG motif. The first and second EGF-like repeats are separated by a pseudo-EGF-like repeat (ψ-EGFL, previously called T + Y; Yochem and Greenwald, 1989), which contains seven cysteines. The putative cytoplasmic domain contains six copies of the *cdc10*/SWI6 motif. Arrows above the line indicate the positions of missense mutations; arrows below the line mark the sites of nonsense mutations and include *glp-1*(*q35*) (Mango *et al.*, 1991). The deleted region in *q172* is shown as an open box. Also shown is *qDf2*, a γ-ray induced deficiency that breaks within EGFL9 and removes all of *glp-1* from that position extending 3′-ward in the gene. The orientation of the *glp-1* protein in this diagram is opposite that of the standard genetic map; therefore, *qDf2* extends to the left from *glp-1* on chromosome III.

effect on viability or morphology among *glp-1*(*q46* or *q175*); smg-1 or *glp-1*(*q46* or *q175*)/+; smg-1 animals. One interpretation of these data is that the extracellular portion of the *glp-1* protein, when produced as a nonsense fragment, is inactive and does not interfere with wild-type *glp-1*. Alternatively, the nonsense *glp-1* fragment may be unstable. In either case, the mutant phenotype of *glp-1*(*q46* and *q175*) is likely to reflect a complete lack of *glp-1* activity.

Six glp-1(ts) Alleles are Missense Mutations in the cdc10/SWI6 Repeats

Six *glp-1*(*ts*) alleles are missense mutations in the *cdc10*/SWI6 repeats (Table 2, Figures 2 and 4). All six of these *glp-1*(*ts*) mutants are virtually wild-type at permissive temperature but are sterile at restrictive temperature (Austin and Kimble, 1987; Priess *et al.*, 1987; this article). Four *glp-1*(*ts*) mutations map to SWI6-4 and change amino acids that are conserved between *glp-1* and *lin-*

Table 2. Molecular lesions in *glp-1*(*lf*) mutants

Allele	Nucleotide affected[a]	Base change	Codon change	Amino acid change
glp-1(*0*)				
q46	6745	C → T	CAA → UAA	Gln 537 → ochre
q158	6815	G → A	UGU → UAU	Cys 560 → Tyr
q172	6709–6969		In-Frame Deletion	Asp 524 to Thr 611 deleted
q175	5102	C → T	CGA → UGA	Arg 191 → opal
oz25	5538	G → A	UGU → UAU	Cys 336 → Tyr
glp-1(*ts*)				
q224	8721	G → A	GGA → GAA	Gly 1043 → Glu
q231	8763	G → A	GGA → GAA	Gly 1057 → Glu
bn18	8693	G → A	GCA → ACA	Ala 1034 → Thr
e2141	8378	C → T	CUU → UUU	Leu 929 → Phe
e2144	8378	C → T	CUU → UUU	Leu 929 → Phe
sy56	8678	C → T	CGG → UGG	Arg 1029 → Trp
glp-1(*lf*)				
q415	5208	G → A	GGA → GAA	Gly 226 → Glu
q50	4907	T → A	UGC → AGC	Cys 126 → Ser
e2072	4941	G → A	GGA → GAA	Gly 137 → Glu
e2142	2609	G → A	GGA → AGA	Gly 33 → Arg

[a] Nucleotide coordinates from Yochem and Greenwald (1989).

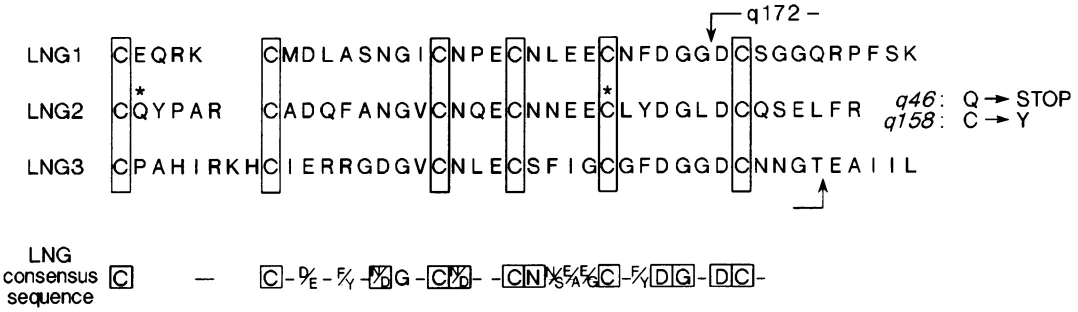

Figure 3. *glp-1* mutations mapping to the LNG repeats. The three copies of the LNG motif found in the *glp-1* protein are aligned with respect to their conserved cysteine residues (boxed). The breakpoints of the inframe deletion in *q172* are marked by arrows. The cysteine that is changed to a tyrosine by the missense mutation *glp-1*(*q158*) and the glutamine altered in the nonsense mutation *glp-1*(*q46*) are marked with asterisks. The LNG consensus is based on the LNG repeats of *glp-1* (Yochem and Greenwald, 1989), *lin-12* (Yochem *et al.*, 1988), Notch (Wharton *et al.*, 1985; Kidd *et al.*, 1986), Xotch (Coffman *et al.*, 1990), rat Notch (Weinmaster *et al.*, 1991) and TAN-1 (Ellisen *et al.*, 1991). Within the consensus, boxed residues are conserved in at least 17 of the 18 LNG repeats in these proteins; remaining consensus amino acids are found in at least 11 of the 18 repeats.

12 (Figure 4B). Two of these, *sy56* and *bn18*, reside in a region with high similarity among all SWI6 repeats (see consensus in Figure 4A), whereas the other two, *q224* and *q231*, are found in the less conserved half of SWI6-4 (Figure 4, A and B). Two other *glp-1*(*ts*) mutations, *e2141* and *e2144*, though apparently isolated independently (Priess *et al.*, 1987), carry the same nucleotide change, resulting in a leucine to phenylalanine substitution in SWI6-1 (Table 2, Figure 4A). This position is occupied by isoleucine in *lin-12*. The clustering of lesions that result in *glp-1*(*ts*) mutations points to the SWI6-4 repeat as playing a critical role in *glp-1* function.

glp-1 Alleles that Affect Embryos More Than the Germ Line are Missense Mutations in the Amino-Terminal EGF-Like Repeats

Four *glp-1* alleles (*q50*, *q415*, *e2072*, and *e2142*) are missense mutations in the amino-terminal EGF-like repeats (Figures 2 and 5); all four have a more severe effect on embryos than on the germ line (Table 1). Three of the four mutations affect residues that are highly conserved among all EGF-like repeats and one affects a glycine immediately adjacent to such a conserved residue (Figure 5, A and B). Assuming that such conservation reflects an underlying structural significance, these amino acid substitutions are likely to disrupt the structure of the EGF-like repeat in which they occur (see DISCUSSION).

The differential effect of these four alleles in embryos and the germline can be explained in two distinct ways. One possibility is that these mutations alter domains with embryo-specific functions. Alternatively, the amino acid substitutions may simply reduce *glp-1* activity. If embryos are more sensitive than the germline to reduced levels of *glp-1* activity, weak *glp-1* alleles would affect embryos more severely than the germline. To distinguish between these explanations, we compared the phenotype of *glp-1*(*x*)/*glp-1*(*x*) homozygotes with that of *glp-1*(*x*)/*glp-1*(*0*) heterozygotes (where *glp-1*(*x*) represents each of the four *glp-1* alleles that affect embryogenesis more than the germ line). We selected *glp-1*(*q175*) for *glp-1*(*0*), because this nonsense mutant truncates the *glp-1* protein close to the amino-terminus (see above) and because the *glp-1*(*q175*) protein could not be detected using antibodies (Crittenden, Troemel, and Kimble, unpublished data). Although a *glp-1* deficiency is preferable for this test, the only deficiency in this region, *qDf2*, leaves part of *glp-1* intact (Figure 1). For the arguments presented below, we assume that *q175* produces no functional *glp-1* product. If *glp-1*(*x*) disrupts an embryo-specific domain, a reduction in its copy number is predicted to enhance the embryo-specific defect preferentially without affecting other tissues (i.e., the germline). However, if the embryo were simply more sensitive to gene dose, a reduction in copy number is predicted to enhance the phenotype in all affected tissues.

We find that, for all four alleles, the germ line is more defective in *glp-1*(*x*)/*glp-1*(*0*) than in *glp-1*(*x*)/*glp-1*(*x*) animals (Table 3). For three alleles, (*q50*, *q415*, and *e2072*), the effect is dramatic, and for one allele, *e2142*, the effect is slight. The most probable explanation is that at least three of the four alleles are weak or partial loss-of-function *glp-1* mutations; *e2142* could either be a partial loss-of-function mutation or interfere with an embryo-specific function. We also examined the embryonic phenotypes of these animals. As might be expected, viability was reduced in *glp-1*(*e2142*)/*glp-1*(*0*) compared with *glp-1*(*e2142*)/*glp-1*(*e2142*) embryos (Table 3). Unexpectedly, *e2072*/*q175* embryos were more viable than *e2072*/*e2072* embryos. This might be explained by interference of the *e2072* mutant protein, either with itself or another protein, which is alleviated when the copy number is reduced. In summary, we suggest that three of the four missense mutations in the N-terminal EGF-like repeats are partial loss-of-function mutations and that the embryo is more sensitive to the

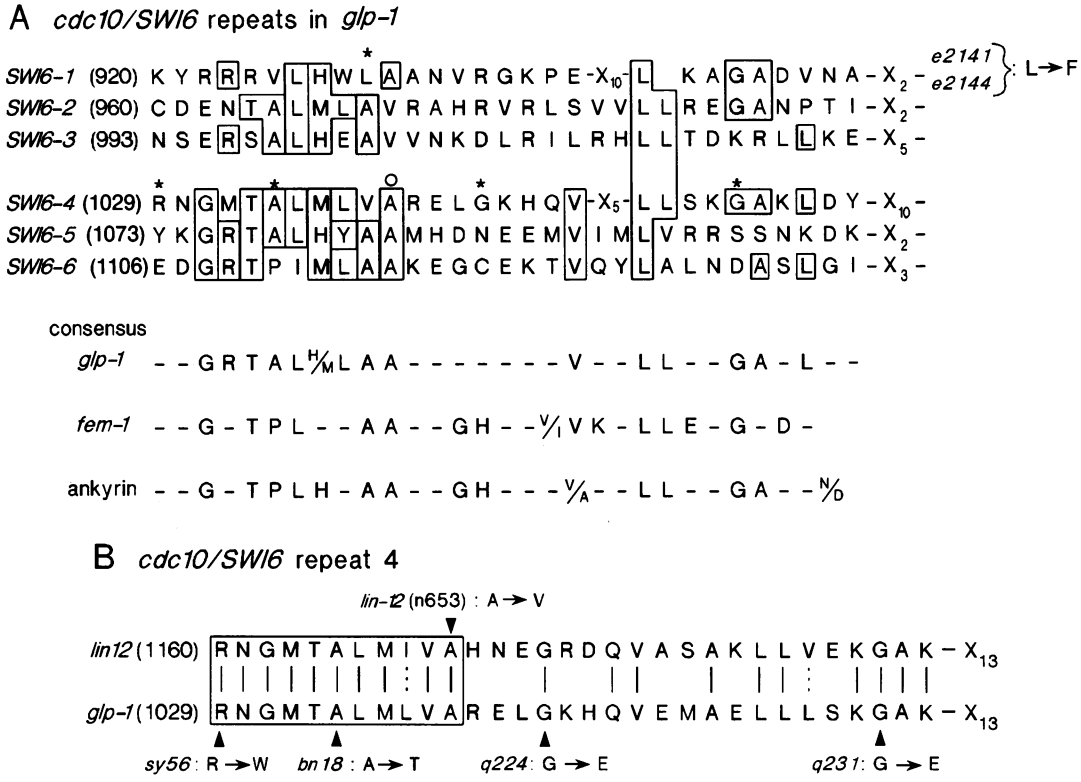

Figure 4. *glp-1* mutations mapping to the *cdc10*/SWI6 repeats. (A) The six *cdc10*/SWI6 repeats of *glp-1* are aligned to maximize homology. The number of the first amino acid of each repeat, shown in parentheses, is based on the numbering of Yochem and Greenwald (1989) (X, amino acid). Asterisks mark sites of amino acid substitutions in *glp-1* mutants and a circle indicates the amino acid substitution of *lin-12(n653)* (Greenwald and Seydoux, 1990). Residues identical in at least three of the six repeats are boxed and appear in the *glp-1* consensus. The consensus sequences derived from the *cdc10*/SWI6 repeats of *fem-1* (Spence *et al.*, 1990) and ankyrin (Lux *et al.*, 1990) are shown for comparison. (B) Alignment of the SWI6-4 repeats of *glp-1* and *lin-12*, with emphasis on the region with mutations. Identical amino acids are connected with vertical lines and similar amino acids with broken lines. The region of highest homology is boxed. The changes in *glp-1* mutations are shown below the repeat and that associated with *lin-12(n653)* (Greenwald and Seydoux, 1990) is above the repeat.

dose of *glp-1* than the germ line. Furthermore, we suggest that *glp-1(e2142)* may have an embryo-specific defect.

DISCUSSION

All glp-1(lf) Mutations Map Within the EGF-like, LNG, and cdc10/SWI6 Repeats

glp-1 is required for inductive cell interactions, during both embryonic and postembryonic development of *C. elegans* (Austin and Kimble, 1987; Priess *et al.*, 1987). The *glp-1* protein belongs to a small family of proteins, the LNG family (see INTRODUCTION), and is likely to function as a receptor (Austin and Kimble, 1987, 1989; Yochem and Greenwald, 1989). Repetitive motifs (EGF-like, LNG, and *cdc10*/SWI6) make up ~60% of the *glp-1* protein. In addition to the evolutionary conservation of these motifs among all LNG proteins, molecular analyses of mutations in Notch and *lin-12* indicate that certain repeats are important for the function and regulation of individual LNG proteins (Hartley *et al.*, 1987; Kelley *et al.*, 1987; Kidd *et al.*, 1989; Greenwald and Seydoux, 1990).

In this article, we report the characterization of 15 *glp-1(lf)* alleles. Two are nonsense mutations, 12 are missense mutations, and 1 is an in-frame deletion. All of these mutations map within the repetitive motifs. Therefore, the integrity of the EGF-like, LNG, and *cdc10*/SWI6 repeats must be crucial to *glp-1* function. By analogy, these repeats will surely be essential to the function of all LNG proteins. In the following discussion, we present the major conclusions that can be drawn from our analysis and provide ideas for how the various repeated motifs of the *glp-1* protein function to mediate cell interactions.

Nonsense Mutants Define the glp-1 Null Phenotype

The most severe phenotype of *glp-1(lf)* mutations is the failure of germline induction. Two mutations with this severe *glp-1(lf)* phenotype, *glp-1(q175* and *q46*), are nonsense mutants: one truncates the *glp-1* protein just after EGFL-3 and the other within LNG-2 (Figure 1).

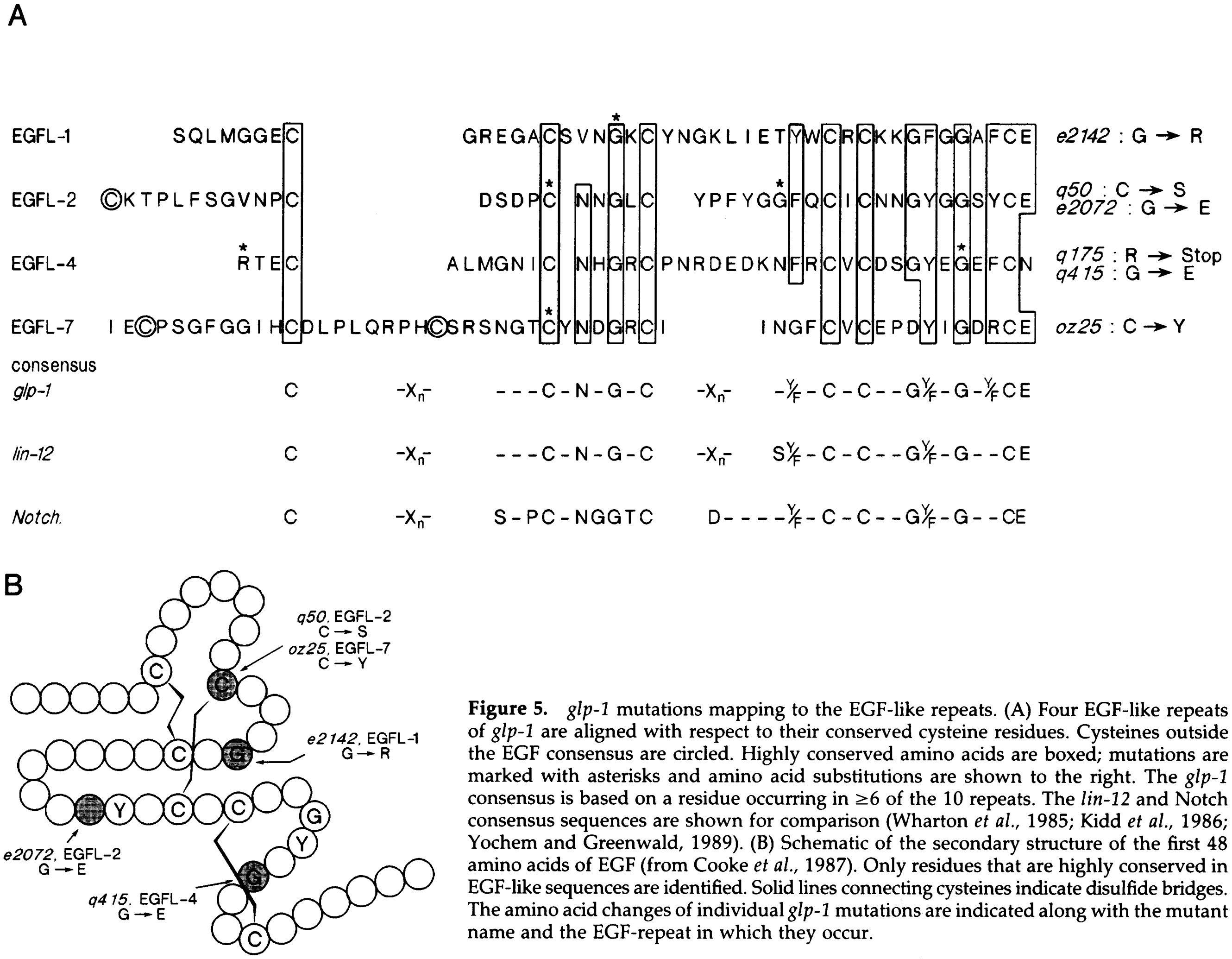

Figure 5. *glp-1* mutations mapping to the EGF-like repeats. (A) Four EGF-like repeats of *glp-1* are aligned with respect to their conserved cysteine residues. Cysteines outside the EGF consensus are circled. Highly conserved amino acids are boxed; mutations are marked with asterisks and amino acid substitutions are shown to the right. The *glp-1* consensus is based on a residue occurring in ≥6 of the 10 repeats. The *lin-12* and Notch consensus sequences are shown for comparison (Wharton *et al.*, 1985; Kidd *et al.*, 1986; Yochem and Greenwald, 1989). (B) Schematic of the secondary structure of the first 48 amino acids of EGF (from Cooke *et al.*, 1987). Only residues that are highly conserved in EGF-like sequences are identified. Solid lines connecting cysteines indicate disulfide bridges. The amino acid changes of individual *glp-1* mutations are indicated along with the mutant name and the EGF-repeat in which they occur.

Previously, we identified a different nonsense mutant, *glp-1*(*q35*), that truncates the *glp-1* protein by 122 amino acids from the carboxy-terminus (Mango *et al.*, 1991). The *glp-1*(*q35*) nonsense mutant has a weak loss-of-function phenotype that is suppressed by a mutation in *smg-1* and a weak gain-of-function phenotype that is enhanced by the same mutation (Mango *et al.*, 1991). The *smg* mutations allow nonsense mRNAs to accumulate at a wild-type level (Pulak and Anderson, personal communication). The influence of the *smg-1* mutation on the *glp-1*(*q35*) phenotype implies that the truncated *glp-1*(*q35*) protein is functional when produced in sufficient quantity. By contrast, we show in this article that the amino-terminal nonsense mutants, *glp-1*(*q46* and *q175*), are not affected by *smg-1*. A more standard genetic test for a null allele depends on the use of a deficiency; however, the only deficiency in this region, *qDf2*, only deletes a portion of *glp-1* and its products have not yet been analyzed. Therefore, the standard tests cannot be done. Recent evidence with antibodies to the *glp-1* protein supports the identification of *glp-1*(*q175*) as a protein null: no *glp-1* protein is detected in *glp-1*(*q175*) homozygotes that nonetheless have a proliferative germ line due to the presence of other mutations that are not in *glp-1* (Crittenden, Troemel, and Kimble, unpublished data). Therefore, *glp-1*(*q46* and *q175*) lack *glp-1* activity by both genetic and molecular criteria and are likely to be null mutants.

Missense Mutations and an In-Frame Deletion may Identify Functional Domains in the glp-1 Protein

Thirteen mutants are predicted to produce *glp-1* proteins that are either full-length or nearly full-length. These 13 mutations fall into three phenotypic classes. The first class includes three alleles with a strong nonconditional phenotype, which is identical to that of the two nonsense mutants. Among these, *oz25* is a missense mutation in EGFL-7, *q158* is a missense mutation in LNG-2, and *q172* is an in-frame deletion that removes most of the LNG repeats. In the second class are six *glp-1*(*ts*) alleles with a strong loss-of-function phenotype at re-

Table 3. Genetic analyis of *glp-1* mutations with a more severe effect on embryogenesis than germline development

Genotype[a]	Temperature (°C)	% fertile (n)[b,c]	% hatching (n)[b,d]
+/+	15	100 (24)	100 (564)
	25	100 (20)	100 (574)
+/*q175*	15	100 (17)	98 (605)
	25	100 (68)	99 (914)
q415/q415	15	97 (60)	0 (738)
q415/q175	15	3 (31)	0 (4)[e]
q50/q50	25	33 (48)	0 (520)
q50/q175	25	0 (40)	NA
e2072/e2072	15	65 (194)	41 (3438)
e2072/q175	15	42 (161)	70 (3517)
e2072/e2072	25	22 (82)	74 (1176)
e2072/q175	25	7 (162)	80 (436)
e2142/e2142	15	100 (45)	97 (1666)
e2142/q175	15	95 (66)	0 (737)
e2142/e2142	25	100 (87)	21 (2219)
e2142/q175	25	100 (72)	0 (1285)

NA, not applicable.

[a] Transheterozygotes were generated from matings (see MATERIALS AND METHODS). All animals (except those containing *q50*) were homozygous for *unc-32*, a closely linked marker on chromosome III (see MATERIALS AND METHODS).

[b] All cross-progeny of at least three matings were scored.

[c] n, number of animals scored.

[d] n, number of embryos scored.

[e] Only one hermaphrodite produced four eggs, none of which hatched.

strictive temperature and a virtually wild-type phenotype at permissive temperature. All six are missense mutations in the *cdc10*/SWI6 repeats. Finally, the third class includes four alleles with a stronger effect on embryos than the germline. All four are missense mutations in the amino-terminal EGF-like repeats. Therefore, mutations in the latter two phenotypic classes correspond to clustered changes within the *glp-1* protein.

Our finding that 12 missense mutations and 1 in-frame deletion of *glp-1* alter the repeated motifs of the *glp-1* protein suggests that the repeated motifs are critical to *glp-1* function. Although we cannot rule out an effect on protein stability, it seems unlikely that all of these mutations yield unstable products. Therefore, in the discussion that follows, we discuss specific amino acid substitutions with the idea that they perturb protein function rather than stability.

Individual EGF-Like Repeats May Have Distinct Functions

Five missense mutations in *glp-1* alter the EGF-like repeats of the *glp-1* protein (Figure 5). Many proteins other than LNG proteins possess arrays of EGF-like repeats, but the function of these tandem arrays of EGF-like repeats is not understood. Some participate in protein–protein interactions (EGF, Cohen *et al.*, 1980; thrombomodulin, Kurosawa *et al.*, 1988; urokinase, Appella *et al.*, 1987; Notch, Rebay *et al.*, 1991), whereas those in the LDL receptor have been implicated in endocytosis (Davis *et al.*, 1987). Characterization of mutations in the *glp-1* EGF-like repeats may therefore provide insight into the function of this array in the *glp-1* protein.

An EGF-like repeat is defined by its similarity to EGF. All EGF-like repeats contain at least six cysteines, three glycines, and three tyrosines/phenylanines in a conserved pattern (Figure 5). In EGF, the six cysteines establish disulfide bridges that are essential to its secondary structure (Cooke *et al.*, 1987; Montelione *et al.*, 1987) (Figure 5B). The conserved amino acids in the EGF-like repeats may therefore be required for the repeat to assume a structure similar to that of EGF. Among the *glp-1* missense mutations, two alter a conserved cysteine, two alter a conserved glycine, and one alters a glycine that lies adjacent to a conserved phenylalanine. One mutation, *e2142*, makes a change analogous to that found in Factor IX in two independent cases of Hemophilia B (Denton *et al.*, 1988). The conserved nature of the amino acids altered in these mutants suggests that each mutation may change the secondary structure of its EGF-like domain.

Four missense mutations in the EGF-like repeats affect embryos more severely than the germline. The differential effect of these alleles on embryos could be attributed either to an embryo-specific function of the mutated EGF-like repeats or to a more stringent requirement for *glp-1* activity in embryos than in the germline. We found that the germline defect was more pronounced in animals with only one dose, although the effect of one mutation, *e2142*, was only slightly enhanced. These results suggest that mutations in EGFL-1, EGFL-2, and EGFL-4 reduce *glp-1* activity partially and that the embryo therefore has a lower tolerance for reduced *glp-1* activity than the germline. Furthermore, the mutation in EGFL-1, *e2142*, may interfere with an embryo-specific function.

One missense mutation in the EGF-like repeats, *glp-1(oz25)*, results in a null phenotype. This mutation substitutes tyrosine for a conserved cysteine in EGFL-7 (Figure 5). The sequence of EGFL-7 has two features that are unique among the *glp-1* EGF-like repeats. First, EGFL-7 possesses eight rather than six cysteines. Second, EGFL-7 contains the sequence CDPGYIGSR, which is also found in laminin (Figure 6). Domain III of the B1 subunit of laminin has eight EGF-like repeats, each with eight rather than six cysteines; its sequence is conserved in *Drosophila* (Montell and Goodman, 1988), mouse (Sasaki *et al.*, 1987), and human (Pikkarainen *et al.*, 1987). The fourth EGF-like repeat of domain III of laminin from all three species includes the recognition site for the laminin receptor, CDPGYIGSR; the binding of laminin to laminin receptor promotes cell attachment and migration (Graf *et al.*, 1987). Therefore, EGFL-7 of the *glp-1* protein shares two characteristics of the laminin-subclass of EGF-like repeats. It is intriguing that the only missense mutation in the EGF-like repeats that appears to eliminate *glp-1* function is located in EGFL-7. One interpretation is that EGFL-7 is a functionally unique domain and that its function is essential for *glp-1* activity. Indeed, this domain might mediate interactions between the *glp-1* protein and the extracellular matrix (Maine and Kimble, 1989).

glp-1 EGFL-7 (aa 347-359)	C	V	C	E	P	D	Y	I	G	D	R	C	E
human Laminin B1 (aa 944-956)	C	V	C	D	P	G	Y	I	G	S	R	C	D
mouse Laminin B1 (aa 944-956)	C	V	C	D	P	G	Y	I	G	S	R	C	D
drosophila Laminin B1 (aa 937-949)	C	H	C	Q	E	G	Y	S	G	S	R	C	E

Figure 6. Similarity in sequences of EGFL-7 of *glp-1* and the EGF-like repeats of domain III of the B1 subunit of laminin from human (Pikkarainen *et al.*, 1987), mouse (Sasaki *et al.*, 1987), and *Drosophila* (Montell and Goodman, 1988). Identical or similar (Asp [E] and Glu [D]) residues are boxed. EGF-like repeat 10 of *lin-12*, which also has eight cysteines, does not contain the sequence found in laminin.

Positive Regulation of glp-1 may be Mediated by the EGF-Like Repeats

The two mutations located in EGFL-2 have a unique "all or none" effect on germline proliferation. In *glp-1*(*q50* or *e2072*) homozygotes, germline proliferation is either severely defective or nearly wild-type. Even within an individual animal, one ovotestis can be severely defective and the other virtually wild-type. This phenotype is rare or undetectable (e.g., <1%) among other *glp-1* mutants (Austin and Kimble, 1987; Kodoyianni, unpublished data). The finding that one gonadal arm can be phenotypically Glp, whereas the other is phenotypically wild-type suggests that *glp-1* is regulated independently in the two arms.

An all or none effect on germline proliferation can be explained by invoking a positive feedback control that maintains *glp-1* in an active state. The altered *glp-1* protein in *glp-1*(*q50* or *e2072*) may provide sufficient activity in some ovotestes to initiate the positive feedback, whereas in other ovotestes, the defective protein may have too little activity to trigger this control. The positive control might be autoregulatory or it might act through a positive feedback loop.

Why is the all or none effect on germline proliferation observed in the two EGFL-2 mutants but not in other *glp-1* mutants? One model is that EGFL-2 itself might be required for initiating the positive regulation. This interpretation is consistent with the finding that individual EGF-like repeats of Notch have distinct functions (Kelley *et al.*, 1987; Rebay *et al.*, 1991). An alternative model is that the EGF-like repeats act together to mediate the positive regulation of *glp-1* activity in the germline. Consistent with this idea is the finding that the five missense mutations in the EGFL-repeats can be placed loosely into an allelic series, with the relative strength of the mutations corresponding to their distance from the amino-terminus. Thus, the weakest allele is *glp-1*(*e2142*) in EGFL-1 and the strongest is *glp-1*(*oz25*) in EGFL-7 (Table 1, Figure 5A). Mutations in EGFL-2 (*q50* and *e2072*) are intermediate in strength and exhibit the all or none character described. Mutations in EGFL-4 (*q415*) are temperature sensitive for germline activity. A mutation that inactivates EGFL-1 would leave the majority of repeats intact, whereas mutations located more internally, e.g., EGFL-4 or EGFL-7; might severely disrupt the array and inactivate it. The position of EGFL-2 in the array would be poised so that some proteins with EGFL-2 defects remain functional, whereas others do not. Therefore, in some ovotestes, the level of *glp-1* activity may reach a threshold level required to promote wild-type germline proliferation, and in other ovotestes, that threshold level may not be obtained. It is possible that elements of each model are correct. For example, the EGF-like repeats may act together to achieve positive regulation, but EGFL-2 may have evolved to play a more critical role in this function than the other EGF-like repeats.

LNG Repeats are Required for glp-1 Function

Two mutations that affect the LNG repeats of the *glp-1* protein lead to a severe loss-of-function phenotype that is indistinguishable from that of null mutants. In *glp-1*(*q172*), 87 amino acids within the LNG repeats are deleted; in *glp-1*(*q158*), a conserved cysteine in LNG-2 is replaced by a tyrosine (Figure 3). The only other known mutations in LNG repeats are single amino acid substitutions in LNG-2 and LNG-3 of *lin-12*, which elevate rather than reduce *lin-12* activity (Greenwald *et al.*, 1983; Greenwald and Seydoux, 1990). One of the *lin-12*(*gf*) mutants, *lin-12*(*n302*), occurs within a stretch of eight amino acids in LNG-3 that is identical in *glp-1* and *lin-12* and may be a common site of regulation. The finding that mutations in the LNG repeats can lead to both loss- and gain-of-function suggests that the LNG region has a pivotal role in the function of LNG proteins.

cdc10/SWI6 Repeats of glp-1 may bind Proteins that Control Cell Fate

Six *glp-1* mutations alter the *cdc10*/SWI6 or ankyrin repeats in the intracellular domain (Figures 2 and 4). The germline phenotype of these *glp-1*(*ts*) mutants, when raised at restrictive temperature, is similar to that of *glp-1*(*0*) mutants, as defined above. Therefore, it is likely that the *glp-1*(*ts*) proteins are not active at restrictive temperature. This inactivity may reflect either degradation of the *glp-1* protein or inability of that protein to function. In either case, the embryonic defect of the *glp-1*(*ts*) mutations, which cannot be observed in *glp-1*(*0*) mutants, is likely to represent the effect of a lack of *glp-1* activity in the early embryo.

Remarkably, four of the *glp-1(ts)* mutations cluster in SWI6-4 (Figure 2). Furthermore, a *lin-12(ts)* mutation maps to the same repeat (Greenwald and Seydoux, 1990). These findings suggest that SWI6-4 is essential either to the structure of the intracellular domain or its function. Comparison of the sequence of SWI6-4 to the sequences of the other *cdc10*/SWI6 repeats reveals no obvious unique domains within SWI6-4. Therefore, the significance of SWI6-4 may lie in its position within the group of repeats or on some unique feature of SWI6-4 that we cannot recognize.

The *cdc10*/SWI6 repeats occur in diverse proteins and have been implicated in mediating protein–protein interactions. Three mammalian proteins exemplify this function. First, four *cdc10*/SWI6 repeats in the $\beta 1$ subunit of the GA binding protein (GABP) mediate a stable interaction between the $\beta 1$ and α subunits of GABP (LaMarco *et al.*, 1991; Thompson *et al.*, 1991). Second, five *cdc10*/SWI6 repeats are present in IkB, a negative regulator of the transcription factor NF-kB (Haskill *et al.*, 1991). NF-kB is maintained in an inactive cytoplasmic form by binding to IkB (Nolan *et al.*, 1991). Third, the 22 *cdc10*/SWI6 repeats in ankyrin appear to tether membrane proteins (Lux *et al.*, 1990). By analogy, it is likely that the six *cdc10*/SWI6 repeats of LNG proteins are important for protein–protein interactions. The temperature sensitivity of the mutations mapping to the *cdc10*/SWI6 repeats of *glp-1* is consistent with this hypothesis. Many *cdc10*/SWI6 repeat-containing proteins bind and/or regulate a putative transcription factor: SWI4 and SWI6 of *S. cerevisae* (Andrews and Herskowitz, 1989a,b), *fem-1* of *C. elegans* (Spence *et al.*, 1990), cactus of *Drosophila* (Nusslein-Volhard, personal communication), IkB (Haskill *et al.*, 1991) and the $\beta 1$ subunit of GABP from mammals (Thompson *et al.*, 1991). The exceptions are ankyrin and α-latrotoxin, which are also exceptional in their high number of repeats (Kiyatkin *et al.*, 1990; Lux *et al.*, 1990). An intriguing possibility is that the *cdc10*/SWI6 repeats in the LNG proteins may bind a transcriptional regulator. If true, they may regulate entry of a protein into the nucleus by sequestering it to the cytoplasm. In the germ line of *C. elegans*, the bound factor might be predicted to instruct entry into meiosis or to negatively regulate mitosis.

CONCLUSION

A knowledge of the amino acid changes in mutant *glp-1(lf)* proteins, when coupled with a description of their phenotypic effects, provides insight into how the *glp-1* protein functions. In this article, we provide molecular evidence for the *glp-1* null phenotype and identify two mutants that are likely to be protein nulls. Furthermore, we find clusters of amino acid changes that may identify functional domains in the *glp-1* protein. In addition, we present ideas about molecular mechanisms that may underlie certain aspects of *glp-1* regulation and function. First, we speculate that the unusual "all or none" effect of EGFL-2 mutants on germline proliferation may reflect a role of the EGF-like repeats in a positive feedback control that regulates *glp-1* activity. Positive regulation of *glp-1* could serve to maintain *glp-1* in an active state and commit cells carrying *glp-1* to a certain fate. Second, we speculate that the EGFL-7 repeat may play a modular role in *glp-1* function. This repeat is unique among the EGF-like repeats of *glp-1*, but it shares features with the laminin-subclass of EGF-like repeats. EGFL-7 may therefore mediate interactions of the *glp-1* protein with the extracellular matrix or receptors of the extracellular matrix. Third, we speculate that SWI6-4 may be central to the function of the *glp-1* intracellular domain. We further postulate that the *cdc10*/SWI6 repeats of the LNG proteins may function by sequestering a regulatory protein, perhaps a transcription factor, to the membrane and by preventing its transport into the nucleus.

ACKNOWLEDGMENTS

We are grateful to Phil Balandyk for isolation of *glp-1(q415)* and to the following people for their generosity in sending *glp-1* alleles: S. Strome (Indiana University), (*bn18*), W. Boorstein (Cal. Tech.), (*sy56*), T. Schedl (Washington University), (*oz25*), and J. Priess (*e2072*, *e2141*, *e2142*, and *e2144*). Sarah Crittenden, Tom Evans, Betsy Goodwin, and Steve McKnight have provided stimulating conversations throughout the course of this work. Technical illustrations were done with the ever-patient assistance of Leanne Olds. Finally, we express special appreciation to Joel Rothman and members of the Kimble laboratory for critical reading of the manuscript.

This research was support by NIH grant GM-31816 to J.K. V.K., and E.M. were supported by NIH postdoctoral grants GM-14186 and GM-11569, respectively.

REFERENCES

Adamson, E.D. (1990). EGF receptor activities in mammalian development. Mol. Reprod. Dev. *27*, 16–22.

Albertson, D.G., and Thomson, J.N. (1976). The pharynx of *C. elegans*. Philos. Trans. Roy. Soc. (Lond.) B Biol. Sci. *275*, 299–325.

Andrews, B.J., and Herskowitz, I. (1989a). The yeast SWI4 protein contains a motif present in developmental regulators and is part of a complex involved in cell-cycle dependent transcription. Nature *342*, 830–833.

Andrews, B.J., and Herskowitz, I. (1989b). Identification of a DNA binding factor involved in cell-cycle control of the yeast HO gene. Cell *57*, 21–29.

Appella, E., Robinson, E.A., Ullrich, S.J., Stoppelli, M.P., Corti, A., Cassani, G., and Blasi, F. (1987). The receptor-binding sequence of urokinase. J. Biol. Chem. *262*, 4437–4440.

Aroian, R.V., Koga, M., Mendel, J.E., Ohshima, Y., and Sternberg, P.W. (1990). *let-23*, a gene necessary for *C. elegans* vulval induction, encodes a tyrosine kinase of the EGF receptor subfamily. Nature *348*, 693–699.

Austin, J., and Kimble, J. (1987). *glp-1* is required in the germ line for regulation of the decision between mitosis and meiosis in *C. elegans*. Cell *51*, 589–599.

Austin, J., and Kimble, J. (1989). Transcript analysis of *glp-1* and *lin-12*, homologous genes required for cell interactions during development of *C. elegans*. Cell *58*, 565–571.

Aves, S.J., Durckacz, B.W., Carr, A., and Nurse, P. (1985). Cloning, sequencing and transcriptional control of the *Schizosaccharomyces pombe cdc10* "start" gene. EMBO J. *4*, 457–463.

Breeden, L., and Nasmyth, K. (1987a). Cell cycle control of the yeast HO gene: cis- and trans-acting regulators. Cell *48*, 389–397.

Breeden, L., and Nasmyth, K. (1987b). Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of *Drosophila*. Nature *329*, 651–654.

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

Coffman, C., Harris, W., and Kintner, C. (1990). Xotch, the *Xenopus* homolog of *Drosophila* Notch. Science *249*, 1438–1441.

Cohen, S., Carpenter, G., and King, L., Jr. (1980). Epidermal growth factor-receptor-protein kinase interactions: co-purification of receptor and epidermal growth factor-enhanced phosphorylation activity. J. Biol. Chem. *255*, 4834–4842.

Cooke, R.M., Wilkinson, A.J., Baran, M., Pastore, A., Tappin, M.J., Campbell, D., Gregory, H., and Sheard, B. (1987). The solution structure of human epidermal growth factor. Nature *327*, 339–341.

Cotton, R.G.H., Rodrigues, N.R., and Campbell, R.D. (1988). Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc. Natl. Acad. Sci. USA *85*, 4397–4401.

Daar, I.O., and Maquat, L.E. (1988). Premature translation termination mediates triosephosphate isomerase mRNA degradation. Mol. Cell. Biol. *8*, 802–813.

Davis, C.G., Goldstein, J.L., Südhof, T.C., Anderson, R.G.W., Russel, D.W., and Brown, M.S. (1987). Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region. Nature *326*, 760–765.

Denton, P.H., Fowlkes, D.M., Lord, S.T., and Reisner, H.M. (1988). Hemophilia B. Durham: a mutation in the first EGF-like domain of Factor IX that is characterized by polymerase chain reaction. Blood *72*, 1407–1411.

Ellisen, L.W., Bird, J., West, D.C., Soreng, A.L., Reynolds, T.C., Smith, S.D., and Sklar, J. (1991). TAN-1, the human homolog of the Drosophila *Notch* gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. Cell *66*, 649–661.

Emmons, S.W., and Yesner, L. (1984). High frequency excision of transposable element Tc1 in the nematode *C. elegans* is limited to somatic cells. Cell *32*, 55–65.

Graf, J., Iwamoto, Y., Sasaki, M., Martin, G.R., Kleinman, H.K., Robey, F.A., and Yamada, Y. (1987) Identification of an amino acid sequence in laminin mediating cell attachment, chemotaxis and receptor binding. Cell *48*, 989–996.

Gray, A., Dull, T.J., and Ulrich, A. (1983). Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein precursor. Nature *303*, 722–725.

Greenwald, I., and Seydoux, G. (1990). Analysis of gain-of-function mutations of the *lin-12* gene of *Caenorhabditis elegans*. Nature *346*; 197–199.

Greenwald, I.S., and Rubin, G. (1992). Making a difference: the role of cell-cell interactions in establishing separate identities for equivalent cells. Cell *68*, 271–281.

Greenwald, I.S., Sternberg, P.W., and Horvitz, H.R. (1983). The *lin-12* locus specifies cell fates in *C. elegans*. Cell *34*, 435–444.

Hartley, D.A., Xu, T., and Artavanis-Tsakonas, S. (1987). The embryonic expression of the Notch locus of *Drosophila melanogaster* and the implications of point mutations in the extracellular EGF-like domain of the predicted protein. EMBO J. *6*, 3407–3417.

Haskill, S., Berg, A.A., Tompkins, S.M., Morris, J.S., Yurochko, A.D., Sampson-Johannes, A., Mondal, K., Ralph, P., and Baldwin, A.S., Jr. (1991). Characterization of an immediate-early gene induced in adherent monocytes that encodes the IkB-like activity. Cell *65*, 1281–1289.

Heitzler, P., and Simpson, P. (1991). The choice of cell fate in the epidermis of *Drosophila*. Cell *64*, 1083–1092.

Hodgkin, J., Edgley, M., Riddle, D.L., and Albertson, D.G. (1988). List of mapped genes and mutant phenotypes. In: Biology of *Caenorhabditis elegans*, ed. W.B. Wood, Cold Spring Harbor, NY: Cold Spring Harbor Press, 502–559.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode *Caenorhabditis elegans*. Genetics *123*, 301–313.

Hoppe, P.E., and Greenspan, P.J. (1986). Local function of the Notch gene for embryonic ectodermal pathway choice. Cell *46*, 773–783.

Hoppe, P.E., and Greenspan, P.J. (1990). The Notch locus of *Drosophila* is required in epidermal cells for epidermal development. Development *109*, 875–885.

Johansen, K.M., Fehon, R.G., and Artavanis-Tsakonas, S. (1989). The Notch gene product is a glycoprotein expressed on the cell surface of both epidermal and neuronal precursor cells during *Drosophila* development. J. Cell Biol. *109*, 2427–2440.

Kelley, M.R., Kidd, S., Deutsch, W.A., and Young, M.W. (1987). Mutations altering the structure of epidermal growth factor-like coding sequences at the *Drosophila* Notch locus. Cell *51*, 539–548.

Kidd, S., Baylies, M.K., Gasic, G.P., and Young, M.W. (1989). Structure and distribution of the Notch protein in developing *Drosophila*. Genes Dev. *3*, 1113–1129.

Kidd, S., Kelley, M.R., and Young, M.W. (1986). Sequence of the *Notch* locus of *Drosophila:* relationship of the encoded protein to mammalian clotting and growth factors. Mol. Cell. Biol. *6*, 3094–3109.

Kimble, J., and White, J.G. (1981). On the control of germ cell development in *Caenorhabditis elegans*. Dev. Biol. *81*, 208–219.

Kiyatkin, N., Dulubova, I., Chekhovskaya, I., and Grishin, E. (1990). Cloning and structure of cDNA encoding α-latrotoxin from black widow spider venom. FEBS Lett. *270*, 127–131.

Kurosawa, S., Stearns, D.J., Jackson, K.W., and Esmont, C.T. (1988). A 10-kDa cyanogen bromide fragment from the epidermal growth factor homology domain of rabbit thrombomodulin contains the primary binding site. J. Biol. Chem. *263*, 5993–5996.

LaMarco, K., Thomson, C.C., Byers, B.P., Walton, E.M., and McKnight, S.L. (1991). Identification of Ets- and Notch-related subunits in GA binding protein. Science *253*, 789–792.

Lambie, E.J., and Kimble, J. (1991a). Genetic control of cell interactions in nematode development. Annu. Rev. Genet. *25*, 411–436.

Lambie, E.J., and Kimble, J. (1991b). Two homologous regulatory genes, *lin-12* and *glp-1*, have overlapping functions. Development *112*, 231–240.

Lehmann, R., Jiminez, F., Dietrich, U., and Campos-Ortega, J.A. (1983). On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. Wilhelm Roux's Arch. Dev. Biol. *192*, 62–74.

Losson, R., and Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic mRNA stability. Proc. Natl. Acad. Sci. USA *76*, 5134–5137.

Lux, S.E., John, K.M., and Bennett, V. (1990). Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell cycle control proteins. Nature *344*, 36–41.

Maine, E., and Kimble, J. (1989). Identification of genes that interact with *glp-1*, a gene required for inductive cell interactions in *C. elegans.* Development *106*, 133–143.

Mango, S.E., Maine, E.M., and Kimble, J. (1991). A carboxy-terminal truncation activates the *glp-1* protein to specify vulval fates in *C. elegans.* Nature *352*, 811–815.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage Sp6 promoter. Nucleic Acids Res. *12*, 7035–7056.

Montelione, G.T., Wuthrich, K., Nice, E.C., Burgess, A.W., and Sheraga, N.A. (1987). Solution structure of murine epidermal growth factor: determination of the polypeptide backbone chain-fold by nuclear magnetic resonance and distance geometry. Proc. Natl. Acad. Sci. USA *84*, 5226–5230.

Montell, D.J., and Goodman, C.S. (1988) *Drosophila* substrate adhesion molecule: sequence of laminin B1 chain reveals domains of homology with mouse. Cell *53*, 463–473.

Myers, R.M., Larin, Z., and Maniatis, T. (1985). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA: DNA duplexes. Science *230*, 1242–1246.

Nolan, G.P., Ghosh, S., Liou, H.C., Tempst, P., and Baltimore, D. (1991). DNA binding and IkB inhibition of the cloned p65 subunit of NF-kB, a rel-related polypeptide. Cell *64*, 961–969.

Pikkarainen, T., Eddy, R., Fukushima, Y., Byers, M., Shows, T., Pihlajaniemi, T., Saraste, M., and Tryggvason, K. (1987). Human laminin B1 chain. J. Biol. Chem. *262*, 10454–10462.

Price, J.V., Clifford, R.J., and Schüpbach, T. (1989). The maternal ventralizing locus *torpedo* is allelic to *faint little ball*, an embryonic lethal, and encodes the Drosophila EGF receptor homolog. Cell *56*, 1085–1092.

Priess, J.R., and Thomson, J.N. (1987). Cellular interactions in early *C. elegans* embryos. Cell *48*, 241–250.

Priess, J.R., Schnabel, H., and Schnabel, R. (1987). The *glp-1* locus and cellular interactions in early *C. elegans* embryos. Cell *51*, 601–611.

Rebay, I., Fleming, R.J., Fehon, R.G., Cherbas, L., Cherbas, P., and Artavanis-Tsakonas, S. (1991). Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. Cell *67*, 687–699.

Rosenbluth, R.E., and Baillie, K.L. (1981). The genetic analysis of a reciprocal translocation, eT1 (III;V) in C. elegans. Genetics *99*, 415–428.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1988). Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science *230*, 1350–1354.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, R., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Sasaki, M., Kato, S., Kohno, K., Martin, G.R., and Yamada, Y. (1987). Sequence of the cDNA encoding the laminin B1 chain reveals a multidomain protein containing cysteine-rich repeats. Proc. Natl. Acad. Sci. USA *84*, 935–939.

Scott, J., Urdea, M., Quiroga, M., Sanchez-Pescador, R., Fong, N., Selby, M., Rutter, W.J., and Bell, G.I. (1983). Structure of a mouse submaxillary messenger RNA encoding epidermal growth factor and seven related proteins. Science *221*, 236–240.

Seydoux, G., and Greenwald, I. (1989). Cell autonomy of *lin-12* function in cell fate decision in *C. elegans.* Cell *57*, 1237–1245.

Seydoux, G., Schedl, T., and Greenwald, I. (1990) Cell–cell interactions prevent a potential inductive interaction between soma and germline in *C. elegans.* Cell *61*, 939–951.

Spence, A.W., Coulson, A., and Hodgkin, J. (1990). The product of *fem-1*, a nematode sex-determining gene, contains a repeated motif found in cell cycle control proteins and receptors for cell–cell interactions. Cell *60*, 981–990.

Thompson, C.C., Brown, T.A., and McKnight, S.L. (1991). Convergence of Ets- and Notch-related structural motifs in heteromeric DNA binding complex. Science *253*, 762–768.

Weinmaster, G., Roberts, V.J., and Lemke, G. (1991). A homolog of *Drosophila Notch* expressed during mammalian development. Development *113*, 199–205.

Wharton, K.A., Johansen, K.M., Xu, T., and Artavanis-Tsakonas, S. (1985). Nucleotide sequence from the neurogenic locus Notch implies a gene product which shares homology with proteins containing EGF-like repeats. Cell *43*, 567–581.

Yochem, J., and Greenwald, I. (1989). *glp-1* and *lin-12*, genes implicated in distinct cell–cell interactions in *C. elegans*, encode similar transmembrane proteins. Cell *58*, 553–563.

Yochem, J., Weston, K., and Greenwald, I. (1988). *C. elegans lin-12* encodes a transmembrane protein similar to *Drosophila Notch* and yeast cell cycle products. Nature *335*, 547–550.