**INTRODUCTION**

Two homologous receptors, LIN-12 and GLP-1, regulate a variety of cell interactions during *C. elegans* development. Best understood are the roles of LIN-12 in AC/VU lateral signaling (Greenwald et al., 1983) and GLP-1 in germline induction (Austin and Kimble, 1987) and embryonic induction (Priess et al., 1987; Hutter and Schnabel, 1994; Mango et al., 1994; Mello et al., 1994; Moskowitz et al., 1994). In addition to these well-characterized interactions, LIN-12 and GLP-1 control numerous other cell fate decisions during both embryonic and post-embryonic development (Greenwald et al., 1983; Lambie and Kimble, 1991; Newman and Sternberg, 1995; I. Moskowitz and J. Rothman, personal communication). Double mutants lacking the activities of both lin-12 and glp-1 die as L1 larvae with an array of cell transformations that result in the loss of the excretory cell, loss of the rectum and a twisted nose. This collection of double mutant defects is called the Lag phenotype, for *lin-12* and *glp-1* (Lambie and Kimble, 1991). Recent studies indicate that the Lag phenotype results from disruption of specific inductive interactions in the 24- to 87-cell embryo (I. Moskowitz and J. Rothman, personal communication). Therefore, certain cell-fate decisions appear to be controlled specifically by LIN-12 (e.g. AC/VU) or GLP-1 (e.g. germline induction), while others make use of either LIN-12 or GLP-1 (Lag defects).

In an effort to identify additional components of the LIN-12/GLP-1 pathway, a screen for mutants with a phenotype similar to that of the *lin-12 glp-1* double mutants was performed (Lambie and Kimble, 1991). From this screen, two new genes were identified: lag-1 and lag-2. Strong loss-of-function *lag-1* or *lag-2* mutants exhibit a Lag phenotype similar to that of *lin-12 glp-1* double mutant. Because all twenty independent alleles isolated in this screen map to either *lag-1* or *lag-2*, it seems likely that *lag-1* and *lag-2* define central components of the LIN-12 and GLP-1 signal transduction pathways.

Sequence similarities between the LIN-12 and GLP-1 proteins in *C. elegans* and *Drosophila* Notch (Greenwald, 1985; Wharton et al., 1985; Yochem et al., 1988; Austin and Kimble, 1989; Yochem and Greenwald, 1989) suggest that this class of receptors may share a common signaling mechanism. Consistent with this idea, two putative ligands for LIN-12 and GLP-1, called LAG-2 and APX-1, are homologous to the Notch ligands, Delta and Serrate (Henderson et al., 1994; Tax et al., 1994; Mello et al., 1994). The observation that these ligand and receptor families are also present in vertebrates suggests that this pathway may be functionally conserved throughout metazoan phylogeny (reviewed by Artavanis-Tsakonas et al., 1995; Henrique et al., 1995). In this paper, we focus on a molecular analysis of the *lag-1* gene (Lambie and Kimble, 1991). We find that the predicted LAG-1 protein is homologous to a class of DNA-binding proteins in human C Promotor Binding Factor (CBF1) and *Drosophila* Suppressor of Hairless (Su(H)). Furthermore, we show that LAG-1 binds specifically to the DNA sequence RTGGGAA, previously identified as a CBF-1/Su(H)-binding site.

**LAG-1, a gene required for *lin-12* and *glp-1* signaling in *Caenorhabditis elegans*, is homologous to human CBF1 and *Drosophila* Su(H)**

Sioux Christensen, Voula Kodoyianni, Marcus Bosenberg, Lisa Friedman and Judith Kimble

**SUMMARY**

The homologous receptors LIN-12 and GLP-1 mediate diverse cell-signaling events during development of the nematode *Caenorhabditis elegans*. These two receptors appear to be functionally interchangeable and have sequence similarity to *Drosophila* Notch. Here we focus on a molecular analysis of the *lag-1* gene (*lin-12* and *glp-1*), which plays a central role in LIN-12- and GLP-1-mediated signal transduction. We find that the predicted LAG-1 protein is homologous to two DNA-binding proteins: human C Promotor Binding Factor (CBF1) and *Drosophila* Suppressor of Hairless (Su(H)). Furthermore, we show that LAG-1 binds specifically to the DNA sequence RTGGGAA, previously identified as a CBF-1/Su(H)-binding site.

Finally, we report that the 5’ flanking regions and first introns of the *lin-12*, *glp-1* and *lag-1* genes are enriched for potential LAG-1-binding sites. We propose that LAG-1 is a transcriptional regulator that serves as a primary link between the LIN-12 and GLP-1 receptors and downstream target genes in *C. elegans*. In addition, we propose that LAG-1 may be a key component of a positive feedback loop that amplifies activity of the LIN-12/GLP-1 pathway.

Key words: *lag-1*, *lin-12*, *glp-1*, *Caenorhabditis elegans*, *Drosophila*, C Promotor Binding Factor (CBF1), Suppressor of Hairless (Su(H))
proteins that includes human CBF1/KBF2/RBP-Jκ (henceforth called CBF1) and Drosophila Su(H) (Matsunami et al., 1989; Schweisguth and Posakony, 1992). This homology provides important clues about how lag-1 might function in LIN-12/GLP-1 signaling. For example, we show that LAG-1 binds specifically to the DNA sequence RTGGGAA, previously identified as a CBF1/Su(H)-binding site (Tun et al., 1994; Brou et al., 1994). Because lag-1 appears to be required for the majority of cell interactions mediated by LIN-12 and GLP-1 in C. elegans, we propose that the LAG-1 DNA-binding protein may be the primary downstream effector for the LIN-12 and GLP-1 receptors. In addition, we suggest that positive feedback in the LIN-12/GLP-1 signaling pathway may rely, at least in part, on LAG-1 binding to sites located in the genomic sequences of the lin-12, glp-1 and lag-1 genes.

MATERIALS AND METHODS

Strains

C. elegans genetic nomenclature follows guidelines of Horvitz et al. (1979). Most mutations used in this study are described in Hodgkin et al. (1988). Alleles used include: dpy-13(e184); unc-44(e362, e1260); lag-1(q385, q418, q476) (Lambie and Kimble, 1991); smg-3(r929) (provided by B. Cali and P. Anderson); unc-8(e15); deb-1(st385) (provided by R. Barstead and R. Waterston); unc-24(e138). In addition, we used the chromosomal rearrangement, Drn-1, as a balancer for lag-1. The strain unc-44(e1260) lag-1(q385)/deb-1(st385) was used for mutant rescue experiments.

Three-factor mapping and positioning of lag-1 on the physical map

lag-1 was mapped to the right of unc-44 and to the left of unc-8 by standard three factor mapping: 8/38 Unc non-Deb recombinant self-progeny of an unc-44 deb-1/lag-1 hermaphrodite segregated Lag homozygotes, while 30/38 recombinants did not, placing lag-1 to the right of unc-44 and close to unc-44 in the unc-44 to deb-1 interval. 1/24 Unc non-Dpy recombinant self-progeny of a dpy-13 unc-8/lag-1 hermaphrodite segregated Lag animals, placing lag-1 close to and to the left of unc-8 in the dpy-13 to unc-8 interval. Therefore, lag-1 maps on the right arm of linkage group IV between unc-44 and unc-8 (Fig. 1A). Since unc-44 has been cloned (Otsuka et al., 1995), this position provides a left boundary for lag-1 on the physical map; unc-44 coding sequences are present on phage DD#LRF1, which overlaps the left end of YAC Y66E2 (Fig. 1B).

To define a right boundary for lag-1 on the physical map, we mapped lag-1 with respect to the rP8 polymorphism: rP8 is detected by cosmid FS8F9, which spans the right end of Y66E2 (Cali, 1995). From hermaphrodites of genotype unc-44 lag-1(-) rP8 Smg-3 unc-44 lag-1(q385) non-Unc 44 Smg-3 Unc-24 progeny were selected and used to establish lines carrying the recombinant chromosome. The progeny of each recombinant were examined for presence of lag-1(q385) and DNA prepared from each recombinant line was analyzed by Southern blot for the polymorphism rP8 using FS8F9 as probe. Of eleven recombinant lines analyzed, two were homozygous for both lag-1(-) and rP8. The remaining nine recombinant lines were all heterozygous for lag-1(q385), DNA; of these nine lines, four carried rP8, while five did not. Therefore, lag-1 maps to the left of rP8. Together these data place lag-1 in an interval of ~360 kb spanned by a single YAC, Y66E2.

Mutant rescue

Germline transformation was performed using standard techniques (Mello et al., 1991). Cosmids located between DD#LRF1 and FS8F9 were injected (3 μg/ml) together with pRF4 (50-100 μg/ml) as a dominant marker for germline transformation (Mello et al., 1991) into unc-44 lag-1(q385)/deb-1 heterozygotes; the broods of transgenic animals were scored for viable unc-44 progeny. From a pool of four cosmids that rescued lag-1 (Fig. 1B), individual cosmids were assayed for rescue as above. In addition, JK#L78 (described in DNA sequencing and analysis section) and subclones of both JK#L78 and the cosmids F43A12 were tested for rescuing activity as described above.

Isolation of cDNA clones

pJK525 was used to screen approximately 240,000 plaques from each of two C. elegans cDNA libraries (kindly provided by P. Okkema). One full-length cDNA, pJK526, was isolated from the mixed stage library, while one partial cDNA, pJK527, was recovered from the embryonic library.

5’ end determination

Total RNA was prepared from mixed stage wild-type C. elegans and poly(A)+ RNA was selected using the Oligo-tex kit (Qiagen) according to manufacturer instructions. First strand cDNAs were produced using a thermostable reverse transcriptase (Epicenter) at 70°C following manufacturers directions using a lag-1-specific primer, SKC37 (‘5’GGCAA-CATCACTGTTGGCTACT’3’). Second strand synthesis was performed using the RACE-1 primer (‘5’GACTCGAGTCCATCGA(T)7’3’) as described by Frohman et al. (1988). cDNAs were PCR amplified using the RACE-1 primer and a nested lag-1-specific primer, SKC27 reverse (‘5’CGACTGGAGACATTGCCCTTTG’3’). The primary amplification product was reamplified using the RACE-2 primer (‘5’GAGTCCGAAACTG’3’ (Frohman et al., 1988) and the lag-1-specific primer SKC31 (‘5’ATCCAGGTTGGCTGCTGGGTTG’3’). Amplification products were reamplified using a primer complementary to the transspliced leader sequence SL1 (‘5’CTCAAACTTGGGATATTAAACC3’ and SKC31. The resulting PCR product was ligated into the pT7Blue vector (Novagen) and sequenced.

The 5’ end of the lag-1 cDNA was amplified from an independent C. elegans cDNA library (kindly provided by R. Barstead) using the vector-derived primer MB9 (‘5’GGAGCGATTGGCAGCATTTG’3’) and a lag-1-specific primer SKC32 (‘5’GGTGTTGGCGAAGAATCT’3’). The primary PCR product was reamplified with a nested, vector-derived primer, MB10 (‘5’GGCATCGCGTAGAGGTGGTTGTC’3’) and a second lag-1-specific primer SKC27(reverse (see the 5’ end). The PCR product was cloned into pT7Blue and sequenced to determine the 5’ end.

Northern blot

Northern blots were performed as described (Sambrook et al., 1989) using approximately 4 μg poly (A)+ mRNA per lane. The northern blot was probed with a full-length cDNA insert excised from pJK526 and labeled using the Prime-a-Gene kit (Promega) according to manufacturers directions.

PCR walking

To clone genomic sequences between the end of the genomic phase clone JK#L78 and the first exon of lag-1, total genomic DNA was amplified using an oligo complementary to sequences in the first exon of lag-1, SKC8 (‘5’GAATGCTCCTGCTGCCTAC’3’) and an oligo complementary to sequences within the first intron of lag-1, SKC-Breverse (‘5’CAGGAGCGCATCATTGTTG’3’), located near the 5’ end of JK#L78. The resulting PCR fragment was cloned into pT7Blue and sequenced.

To isolate genomic sequence upstream of the first lag-1 exon, total genomic DNA was digested overnight with Clal and cloned into Blue-script (Stratagene). The ligation mixture was PCR amplified using the T7 primer (Promega) and the lag-1-specific primer SKC39 (‘5’GAAACTTGTGCAGCATCT’3’). The resulting PCR products were gel purified, cloned into pT7Blue and sequenced.
DNA sequencing and analysis

All sequencing reactions were performed using Sequenase Version 2.0 (USB) according to manufacturers’ directions. The cDNA inserts from phage clones pJK526 and pJK527 were subcloned into Bluescript vectors and both strands of each insert were sequenced. Genomic fragments containing the lag-1 locus, as identified by Southern blot analysis (data not shown), were subcloned from cosmids F43A12 (pJK525) or phage Jk#L78 (pJK550, pJK551, pJK552) (see Fig. 1C) and sequenced as described for cDNAs except that intron sequences were determined by sequencing a single DNA strand. There is perfect agreement between the proposed coding sequences of the cDNA and genomic clones. Approximately 99.5% of the genomic sequence has been unambiguously determined. Genomic and cDNA sequences were aligned using the Align program (DNASTAR) to establish the exon-intron structure of lag-1. LAG-1, Su(H) and CBF1 sequences were compared using the MegAlign program of DNASTAR.

Identification of lag-1 mutations

DNA was prepared from 5-10 lag-1 homozygous mutant larvae and fragments corresponding to lag-1 exons were generated by PCR, cloned into pT7 Blue and sequenced using standard procedures (Sambrook et al., 1989). For each allele, PCR products covering all coding sequences, the majority of the 3’ UTR and the splice junctions of flanking introns were sequenced. Two or more clones were sequenced for each region; exons containing nucleotide changes were confirmed by sequencing at least two additional independent clones. Oligonucleotide primers used to amplify each exon are listed below; numbers in parentheses indicate the position of the 5’ end of each primer in the genomic sequence, with position 1 corresponding to the first nucleotide of the cDNA clone pJK526.

Exon 1: SKC E0.1- 5’CGATTTCCTCTGTTTTCTACGTG3’ (7037); SKC E1.2- 5’CGGATTTACGATTCGAAACGTG3’ (134).

Exons 2 & 3: SKC E1.1- 5’GCTCCCTTTTTGACACGTCC3’ (7037); SKC E1.2- 5’CCCTTTCCTGGGCAACATGAC3’ (7389).

Exons 4 & 5: SKC E2.1- 5’CCGAGACCTGTAAATTCTC3’ (8250); SKC E2.2- 5’CAGTCTGGAACACATTTGC3’ (8756).

Exons 6 & 7: SKC E3.1- 5’GAAATTCCAGCGTCAAGTAG3’ (11683); SKC E4.2- 5’CTGCTGCTGCTGCAAAGAAGAC3’ (12211).

Exon 8: SKC E5.1- 5’CGACATATGTTGCTGAACAG3’ (13170); SKC E5.2- TCCGCGGCCACAAAATTCTGCTG3’ (13574); SKC E5B.2- 5’TGTGCGCAATCTGGAATTC3’ (14210).

Exon 11: SKC E6.1- 5’GTGATGTGTGCAATGGA3’ (14551); SKC E6.2- 5’CTGAAATTCAGAACACTCTTCC3’ (15164).

Exon 12: SKC E7.2- 5’GGTATTTCGAGTCAAATTC3’ (15591); SKC-12- 5’AGGAGAGATACCCACGTTGA3’ (15993).

RESULTS

Cloning lag-1

The lag-1 locus was initially positioned on the genetic and physical maps to an interval defined by unc-44 and the polymorphism rP8 on chromosome IV (see Materials and Methods) (Fig. 1A,B). This region is spanned by a single YAC, Y66E2, of approximately 360 kb (Fig. 1B). To locate lag-1 within Y66E2, cosmids underlying the YAC were tested for rescue of lag-1(q385) homoyzogotes by germline transformation (see Materials and Methods). Whereas lag-1(q385) homoyzogotes die as first stage larvae (Lambie and Kimble, 1991), transgenic lag-1(q385) homoyzogotes carrying either of two overlapping cosmids, M02G2 and F43A12, survived to adulthood (Fig. 1B). Although the somas of rescued animals have a wild-type morphology, the germlines fail to proliferate, a phenotype similar to that of weak lag-1 and glp-1 null mutants. Since C. elegans, transgenes are often poorly expressed in the germline (Mello and Fire, 1995), the inability of transgenic lag-1 to rescue the germline defect may indicate that lag-1 expression is required in the germline for continued mitotic proliferation. Alternatively, both cosmids that rescue the somatic lag-1 defects may lack regulatory elements required for germline rescue.

Restriction mapping and Southern blot analysis of M02G2 and F43A12 indicated that the largest genomic fragment common to both cosmids was a 7 kb HindIII fragment. However, this fragment was not sufficient to rescue the lag-1 mutant phenotype. To identify lag-1 within F43A12, the cloned 7 kb HindIII fragment, pJK525 (Fig. 1C), was used as a probe to isolate phage from a genomic library. One phage, JK#L78, which contained the 7 kb HindIII fragment and flanking sequences, rescued lag-1 somatic defects in a transgenic assay. Subclones of JK#L78, (pJK550, pJK551 and pJK552, Fig. 1C), were generated and tested for rescuing activity; none of these subclones was able to rescue lag-1(q385) homoyzogotes. We therefore used pJK525 to isolate cDNAs. Two cDNAs were isolated and sequenced: pJK526 is 2550 nt long and is likely to correspond to the full-length transcript.

Exon/intron structure of lag-1

To deduce the exon/intron structure of lag-1, the genomic lag-1 region was sequenced (see Materials and Methods). We first sequenced ∼ 14.2 kb of genomic sequence derived from genomic subclones pJK525, pJK550, pJK551 and pJK552 (Fig. 1C). Comparison of the genomic and cDNA sequences revealed at least 12 exons (11 within pJK#L78 and one at the 5’ end found in the cDNA, but not in the phage) (Fig. 1C); introns ranged in size from 48 nt to 6117 nt. All exon/intron boundaries are flanked by conserved C. elegans splice junction sequences (Emmons, 1988). Because the phage clone JK#L78 did not contain the first exon or 5’ flanking sequences, a PCR walk was used to obtain the 5’ portion of lag-1. In this way, genomic DNA corresponding to the first exon plus approxi-
mately 1.2 kb of its 5′ flanking region was cloned and sequenced (Fig. 1C).

The genomic lag-1 sequence contains a consensus C. elegans splice acceptor sequence (TTTCAG) immediately upstream of the first nucleotide of the longer cDNA clone, pJK526. This may indicate either that the lag-1 message is transspliced in vivo, or that pJK526 represents a truncated transcript. To distinguish between these possibilities, the 5′ ends of two additional lag-1 cDNA clones were examined. First, PCR was used to amplify the 5′-end of lag-1 from a random primed C. elegans cDNA library; such libraries do not require transcription of the full-length cDNA and thus are more likely to contain complete 5′ cDNA sequences. Second, the product of a lag-1-specific RACE reaction was amplified with a primer complementary to the C. elegans transspliced leader sequence SL1. For both of the resulting clones, pJK533 (random primed library) and pJK534 (RACE reaction), the 5′ end corresponded to that identified in pJK526. Furthermore, pJK534 contained SL1 sequences immediately upstream of the first nucleotide predicted by pJK526, suggesting that the lag-1 message can be transspliced in vivo.

Sequence analysis of the predicted LAG-1 protein

Conceptual translation of the full-length lag-1 cDNA predicts a protein of 673 amino acids with an $M_r \approx 74 \times 10^3$. Searches of protein databases revealed a striking homology between LAG-1 and the mammalian CBFI and Drosophila Su(H) proteins (Fig. 4A) (Matsunami et al., 1989; Furukawa et al., 1991; Schweiguth and Posakony, 1992); we suggest that proteins of this class be called CSL proteins, for C. elegans, Drosophila and the mammalian CBFI and Su(H) proteins. Over a region of 412 amino acids, LAG-1 is 60% identical (71% similar) to human CBFI and 62% identical (72% similar) to Drosophila Su(H) (Fig. 4B). Two potential nuclear localization sequences (Chelsky et al., 1989) are found within the predicted LAG-1 protein and a consensus MAP kinase site (P X S/T P) (Clark-Lewis et al., 1991) is conserved among the CSL proteins (Ser$^{560}$ of LAG-1) (Fig. 2). Within the conserved region, four intron splice sites are identical between lag-1 and the corresponding human and mouse genes (Fig. 2, filled arrowheads) (Amakawa et al., 1993). The N terminus and C terminus of the C. elegans, Drosophila and vertebrate proteins are variable in length and are not conserved (Fig. 4B).

Molecular identification of lag-1 mutants

To confirm the molecular identification of lag-1 and to begin exploring the functional domains of the LAG-1 protein, we sequenced the coding regions of three strong loss-of-function lag-1 alleles. All three contain single nucleotide changes within the lag-1 coding region, confirming identity of the locus as lag-1 (Fig. 4B). Two mutations introduce premature stop codons: lag-1(q418) is predicted to truncate LAG-1 N-terminal to the conserved region, while lag-1(q385) removes only the C-terminal 48 residues. In C. elegans, mRNAs containing premature stop codons are often unstable; such mRNAs can be

![Genetic map:](attachment://fig1a.png)

**Fig. 1.** Genetic and molecular identification of lag-1. (A) Genetic map of the lag-1 region on the right arm of linkage group IV showing the position of lag-1 with respect to markers used in three-factor mapping of the lag-1 locus. The expanded region below shows the position of flanking genes and the polymorphism rP8 used to correlate the genetic and physical maps in the lag-1 region. (B) Physical map between unc-44 and rP8. DD#LRF1 is a genomic phage clone containing part of unc-44; the cosmid clone F58F9, detects the RFLP polymorphism rP8. The results of mutant rescue experiments using various cosmids and phage are shown to the left of the relevant DNAs. (C) The lag-1 gene. Above, solid line indicates the sequenced portion of JK#L78, including its breakpoint in the first lag-1 intron. pJK525 is the 7 kb subclone used to screen genomic and cDNA lambda libraries (see text). The pJK525, pJK550, pJK551 and pJK552 subclones cover most of the genomic lag-1 locus (see Results) H, HindIII; B, BstXI; P, PstI. Below, exon-intron structure of lag-1 based on comparison of genomic and cDNA sequence. Filled boxes represent exons; lines represent introns.
LAG-1 and LIN-12/GLP-1 signaling

Fig. 2. The lag-1 cDNA sequence and predicted protein. Nucleotide sequence of the lag-1 cDNA and the predicted amino acid sequence of LAG-1 protein. cDNA numbering begins at the first nucleotide of pJK526, which is likely to be nearly full length, missing only its transspliced leader (see text). Amino acid numbering begins at the first nucleotide of pJK526, which is likely to be nearly full length, missing only its transspliced leader (see text). Amino acid numbering begins at the first methionine. Lower case letters represent predicted 5' and 3' untranslated regions; upper case letters represent predicted coding sequence. Oligonucleotide primers described in the text are shown below the corresponding sequence. Open arrowheads indicate positions of the eleven introns; filled arrowheads mark those splice sites conserved between lag-1 and its single mutant alleles are circled (see Fig. 4A and text). The asterisk indicates the predicted termination codon. The proposed cleavage and polyadenylation signal is boxed. On both cDNA clones, a stretch of poly(A) begins immediately after the last nucleotide shown.

stabilized in a smg mutant background (Hodgkin et al., 1989). We therefore examined the phenotype of a smg-1; lag-1(q385) double mutant, but found no change from the lag-1(q385) single mutant phenotype. Therefore, the strong lag-1(q385) phenotype is likely to result from loss of LAG-1 protein activity rather than from message instability. In addition to these two nonsense
mutants, which establish the lag-1 null phenotype by molecular criteria, we identified one missense mutation, lag-1(q476), in which a semi-conserved aspartic acid at position 303 is converted to histidine. The mutant phenotype of lag-1(q476) is indistinguishable from that of the two nonsense mutants, suggesting that this amino acid replacement eliminates LAG-1 activity.

DNA binding by LAG-1

The homology of LAG-1 to CBF1 and Su(H) provides important clues to its function. CBF1 can activate transcription in combination with EBNA2 (Waltzer et al., 1994; Henkel et al., 1994; Grossman et al., 1994; Zimmer-Strobl et al., 1994) and can repress transcription in certain cases (Hsieh and Hayward, 1995; Dou et al., 1994). Both CBF1 and Su(H) bind a defined DNA sequence, GTGGGAA, in vitro (Tun et al., 1994; Brou et al., 1994). To determine whether LAG-1 may have a similar activity, we tested the ability of LAG-1 to bind DNA containing the CBF1-binding site GTGGGAA as well as the related sequence ATGGGAA. Mobility shift assays were performed with in vitro translated LAG-1 and MB22, a radioactively labeled double-stranded oligonucleotide derived from the...
Identification of LAG-1-binding sites in the genomic sequences of lag-1, lin-12 and glp-1

The distribution of potential LAG-1-binding sites in the genomic sequences of lin-12, glp-1 and lag-1 reveals an intriguing pattern (Fig. 6). If distributed randomly, one RTGGGAA site would be expected in 4 kb on average. Within 15 kb spanning glp-1, 9 sites were found; within 18 kb spanning lin-12, 20 sites were found; and within 16 kb spanning lag-1, 18 sites were observed (Fig. 6; Table 1). Furthermore, in lin-12, glp-1 and lag-1, these sites are preferentially located in 5’ flanking regions and in large introns at the 5’ end of the genes. Particularly striking is the presence of 12 potential LAG-1-binding sites within 2.4 kb region in the first intron of lag-1. 6 of these sites occur as tandem repeats with similar spacing, orientation and conserved linker sequences (Table 1). This region of the first intron is present in the rescuing phage, pJK#L78, and may provide regulatory sequences required for lag-1 activity. In contrast, upon examination of the genomic sequences of lag-2, ced-4, tra-2, and xol-1, sites were either not present (lag-2 and ced-4) or found at a random frequency (tra-2 and xol-1). In the latter case, potential LAG-1-binding sites were not preferentially located in predicted regulatory regions. A search of all available C. elegans genomic sequence using FindPatterns (Devereux et al., 1984) reveals potential LAG-1 sites distributed at the expected frequency of approximately 1 site every 4 kb. Furthermore, no unique sequence was found with a cluster of LAG-1-binding sites as dense (12 sites in 2.4 kb) as that observed in the first lag-1 intron.

**DISCUSSION**

LAG-1 is a conserved central component in signaling by both LIN-12 and GLP-1

Only two genes, lag-1 and lag-2, were identified in extensive mutageneses designed to find loci that participate in both lin-12- and glp-1-mediated signaling (Lambie and Kimble, 1991). Therefore, lag-1 and lag-2 are believed to encode core components of the LIN-12/GLP-1 pathway. This prediction has been confirmed for lag-2, which encodes a homolog of Delta and is believed to function as a ligand for GLP-1 and LIN-12 (Henderson et al., 1994; Tax et al., 1994). In this paper, we...
report that this prediction also holds true for lag-1, which appears to encode the primary downstream effector for LIN-12/GLP-1-mediated interactions. Furthermore, LAG-1 bears a striking sequence similarity to vertebrate CBF1 and Drosophila Su(H), which suggests that at least one downstream effector of the LIN-12/Notch/GLP-1 signaling pathway is conserved.

### Table 1. Potential LAG-1 binding sites in lin-12, glp-1 and lag-1

<table>
<thead>
<tr>
<th>Site</th>
<th>Location*</th>
<th>Direction</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>lin-12</td>
<td>1</td>
<td>-5421</td>
<td>→</td>
<td>TCGTGGGAAAT</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-5208</td>
<td>→</td>
<td>TTATTGGGAAT</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-4283</td>
<td>←</td>
<td>TAATGGGAAAC</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-3664</td>
<td>←</td>
<td>TAATGGGAAAT</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2854</td>
<td>←</td>
<td>CGATGGGAAATG</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>-2091</td>
<td>←</td>
<td>TGATGGGAAATT</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>-2060</td>
<td>←</td>
<td>ATATGGGAAAT</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>-1666</td>
<td>←</td>
<td>GCAGTGGGAAAC</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>-822</td>
<td>←</td>
<td>AGAGTGGGAAAC</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>733</td>
<td>←</td>
<td>AGATGGGAAATG</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>892</td>
<td>←</td>
<td>GTATTGGGAAT</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>1107</td>
<td>←</td>
<td>GCATGGGAAAC</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1831</td>
<td>←</td>
<td>GCATGGGAAATG</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>1894</td>
<td>←</td>
<td>CACTTGGGAAAT</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>3039</td>
<td>←</td>
<td>GACATGGGAAAC</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>3712</td>
<td>←</td>
<td>GTATGGGAAAC</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>4400</td>
<td>←</td>
<td>CACTTGGGAAAT</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>5198</td>
<td>←</td>
<td>TGCTATGGGAACT</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>6785</td>
<td>←</td>
<td>TCATGGGAAATG</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>7085</td>
<td>←</td>
<td>AAAATGGGAAAAG</td>
</tr>
</tbody>
</table>

| glp-1 | 1 | -5517 | ← | GAAATGGGAAATG | promoter |
|      | 2 | -3940 | ← | GCTATTGGGAAT | promoter |
|      | 3 | -3203 | ← | AAGCATGGGAAAT | promoter |
|      | 4 | -3196 | ← | GAAATGGGAAAC | promoter |
|      | 5 | -2329 | ← | GTATGGGAAATC | promoter |
|      | 6 | 1581 | ← | TGATGGGAAATT | promoter |
|      | 7 | 242 | ← | GTTCTGGGAAACA | intron 1 |
|      | 8 | 368 | ← | AGAGTGGGAAAC | intron 1 |
|      | 9 | 2607 | ← | TTATGGGAAATC | exon 4 |

| lag-1 | 1 | 674 | ← | AGAGTGGGAAATG | intron 1 |
|      | 2 | 689 | ← | CGCTGGGGAAAT | intron 1 |
|      | 3 | 781 | ← | GAATGGGAAAC | intron 1 |
|      | 4 | 1632 | ← | AGATGGGAAAT | intron 1 |
|      | 5 | 1765 | ← | AGATGGGAAATG | intron 1 |
|      | 6 | 1780 | ← | GTATGGGAAAAA | intron 1 |
|      | 7 | 2079 | ← | CGATGGGAAAAA | intron 1 |
|      | 8 | 2096 | ← | ATATGGGAAAAA | intron 1 |
|      | 9 | 2297 | ← | TCTGGGAAATAC | intron 1 |
|      | 10 | 2405 | ← | GTTCTGGGAAAT | intron 1 |
|      | 11 | 2917 | ← | TCATGGGAAATC | intron 1 |
|      | 12 | 3078 | ← | ATATGGGAACT | intron 1 |
|      | 13 | 5802 | ← | GTATGGGAAATG | intron 5 |
|      | 14 | 7840 | ← | GACATGGGAAATC | intron 5 |
|      | 15 | 8092 | ← | GTATGGGAAATC | intron 5 |
|      | 16 | 9001 | ← | GCGATGGGAAAC | intron 5 |
|      | 17 | 9119 | ← | TCGATGGGAAAT | intron 5 |
|      | 18 | 11371 | ← | AGATGGGAAAT | intron 7 |

| lag-1 tandem sites | a | 689, 674 | ← | CGCTGGGGAAAAACTAGGATG | |
|                   | b | 1780,1765 | ← | GTTCTGGGAAATACGAGGATG | |
|                   | c | 2096, 2079 | ← | ATCTGGGGAAATGGCGCATG | |

*The numbering of the lag-1 sequence is described in Materials and Methods. For lin-12 and glp-1, the A in the initiation codon is denoted as base pair 1. LAG-1 binding sites were numbered according to the position of the R in the RTGGGAA sequence element.*

LAG-1 controls most signaling interactions mediated by LIN-12 and GLP-1

The lag-1 gene is required for numerous cell interactions mediated by LIN-12 and GLP-1. The larval lethality of lag-1 null alleles demonstrates that LAG-1 is required for those embryonic cell-fate decisions relying on expression of either lin-12 or glp-1; furthermore, the defect in germline proliferation of weak lag-1 mutants indicates that LAG-1 is required for germline induction (Lambie and Kimble, 1991). However, no vulval phenotype has been observed in weak lag-1 alleles (Lambie and Kimble, 1991), even though lin-12 regulates lateral signaling in two types of cell interactions required for vulval development (AC/VU and VPCs) (Greenwald et al., 1983). The lack of vulval defects in lag-1 mutants may reflect a lower threshold requirement for lag-1 in these cell interactions, or it may indicate that lag-1 does not participate in lin-12-mediated cell-fate decisions affecting vulval development.

A maternal requirement for lag-1 is suggested by the embryonic lethality of progeny derived from animals homozygous for either of two weak alleles, lag-1(p416) (Christensen, 1995) and lag-1(m13) (Qiao et al., 1995). The effects of stronger lag-1(lf) mutations on the early embryo cannot be examined, because homozygotes either die as larvae or become sterile adults. The finding that lag-1 antisense RNA injections in the adult hermaphrodite gonad results in embryonic lethality (V. Kodoyianni, unpublished) is consistent with the idea that lag-1 is required maternally for embryogenesis.

Although the full extent of lag-1 function remains to be determined, LAG-1 activity is clearly crucial for numerous cell interactions during C. elegans development and is therefore one of the key players in both LIN-12 and GLP-1 signaling. We suggest that the two LAG-1 homologs, vertebrate CBF1 and Drosophila Su(H), may also play a broad role in signaling by their counterpart receptors. Indeed, recent evidence suggests that Su(H) has a broader role than previously suspected in Drosophila development (Lecourtois and Schweisguth, 1995) and that CBF1 may participate in signaling by the vertebrate Notch pathway (Oka et al., 1995).

CSL proteins link receptor activity to target genes during signal transduction

Two major conclusions have emerged during the past year concerning the function and regulation of CSL proteins (for CBF1, Su(H) and LAG-1). First, they appear to be transcriptional regulators, either activators (Grossman et al., 1994; Henkel et al., 1994; Waltzer et al., 1994; Zimerl-Strobl et al., 1994; Bailey and Posokony, 1995; Lecourtois and Schweisguth, 1995) or repressors (Dou et al., 1994; Hsieh and Hayward, 1995). Second, they appear to link the Notch receptor with downstream target genes (Fortini and Artavanis-Tsakonas, 1994; Bailey and Posokony, 1995; Jarriault et al., 1995; Lecourtois and Schweisguth, 1995). Su(H) protein binds the intracellular domain of Notch in yeast two-hybrid experiments and Notch activation appears to regulate the translocation of Su(H) from cytoplasm to nucleus (Tamura et al., 1995; Fortini and Artavanis-Tsakonas, 1994). In addition, Su(H) transcriptionally regulates the expression of target genes in the Enhancer of Split cluster (E(Sp)-C), in response to Notch activation (Bailey and Posokony, 1995; Lecourtois and Schweisguth, 1995). Similarly, CBF1 mediates activation of expression of HES-1, the mammalian homolog of E(Spl), in the presence of activated
Notch1 (Jarriault et al., 1995). Therefore, although the molecular details of how activation of Notch leads to activation of Su(H) are not yet understood, CBF1 and Su(H) clearly link Notch and Notch-related homologs with transcriptional activation of E(Spl)-C-related genes.

Several lines of evidence suggest that the function and regulation of LAG-1 may be similar to that of CBF1 and Su(H). First, LAG-1 binds to the consensus CBF1/Su(H)-binding site RTGGGAA (this paper). Second, LAG-1 appears to act downstream of the LIN-12/GLP-1 receptors: genetically, a weak lag-

1 mutation is epistatic to a glp-1 gain-of-function mutation (L. W. Berry and T. Schedl, personal communication) and molecularly, lag-1 mRNA is present in the mitotic region of the germline, the receiving tissue of a glp-1-mediated interaction (V. Kodoyianni, unpublished data). However, downstream target genes of LIN-12/GLP-1 signaling have not yet been identified in C. elegans.

LAG-1 may mediate positive feedback in the LIN-12/GLP-1 signaling pathways

Positive feedback has been suggested to regulate both in GLP-1 inductive signaling (Kodoyianni et al., 1992) and LIN-12 lateral signaling (Seydoux and Greenwald, 1989; Wilkinson et al., 1994). For induction processes, feedback was hypothesized because of an all-or-none phenotype in mutants carrying either of two unusual glp-1 alleles: within a single worm, one germline arm may exhibit no proliferation, while the other may be indistinguishable from wild type. One interpretation of this all-or-none phenotype relies on positive feedback in the GLP-1 signaling pathway: partially active GLP-1 may reach a threshold sufficient to establish positive feedback, which then ensures continuous glp-1 expression or activity (Kodoyianni et al., 1992). For lateral signaling events, positive feedback was hypothesized from a bias in cell fate choice that was observed in lin-12 mosaics but not wild-type animals (Seydoux and Greenwald, 1989). In Drosophila, a positive feedback mechanism has also been hypothesized to play a role in the Notch signaling pathway (Heitzer and Simpson, 1991).

Positive feedback appears to occur at the transcriptional level in the case of lin-12 (Wilkinson et al., 1994). Specifically, a 67 bp region in the 5' flanking region of lin-12 is required to maintain lin-12 expression in the VU precursor cell during the AC/VU decision. This 67 bp region, LCS1 (lin-12 conserved sequence) was identified by conservation with C. briggsae lin-12 and contains a LAG-1-binding site. Deletion of LCS1 from a lin-12 transgene results in a reduction of its ability to regulate the AC/VU decision, but not other lin-12-mediated processes. Therefore, the LCS1 element appears to specifically control positive feedback for the AC/VU decision.

Examination of the lin-12, glp-1 and lag-1 genomic sequences reveals a greater number of potential LAG-1-binding sites than would be predicted on a random basis (Fig. 6). Furthermore, the LAG-1-binding sites are clustered in the 5' flanking regions and large first introns of these genes, including one in LCS1. Indeed, MB20, one of the oligonucleotides used to compete LAG-1 DNA-binding, is derived from LCS1. Based on the non-random distribution of potential LAG-1-binding sites in glp-1, lin-12 and lag-1, we suggest that these genes may be transcriptionally regulated by LAG-1 in a positive feedback loop (Fig. 7). However, the importance of LAG-1-binding sites for positive feedback in the LIN-12/GLP-1 pathway remains to be tested.

Signaling cell Receiving cell

Fig. 7. Model for positive feedback regulation by the LAG-1 DNA-binding protein. Activation of GLP-1/LIN-12 may affect the ability of LAG-1 to transcriptionally up regulate its own synthesis, as well as that of the GLP-1/LIN-12 receptors.

Specificity in the LIN-12/GLP-1 signaling pathway

The central role of LAG-1 in signaling by both LIN-12 and GLP-1 suggests that LAG-1 itself does not confer specificity in the cellular responses to activation of these receptors. Similarly, specificity does not appear to depend on which receptor or ligand is used: LIN-12 and GLP-1 are functionally interchangeable (Mango et al., 1991; Lambie and Kimble, 1991; Fitzgerald et al., 1993) as are their predicted ligands LAG-2 and APX-1 (Gao and Kimble, 1995; Fitzgerald and Greenwald, 1995). Thus, specificity is likely to be generated by cell- or tissue-specific proteins that may interact with the core components of the pathway. For example, the CSL proteins may direct such proteins to specific promoters. Consistent with this idea, the LCS1 region is relatively large (67 nt) and appears to be specific for the AC/VU decision (Wilkinson et al., 1994). Therefore, as yet unidentified factors may work together with LAG-1 at the LCS1 regulatory site to regulate the lin-12 promoter during lateral signaling between the AC and VU cells. As an extension of this idea, we suggest that the number of potential LAG-1-binding sites in the glp-1, lin-12 and lag-1 genomic sequences may reflect the complexity of transcriptional regulation required to express these genes in specific cells and at specific times during development. Molecular analyses of individual sites, initiated with LCS1, will clarify the importance of these sites and their individual functions during cell signaling.

Extent of conservation in Notch-related pathways

LAG-1 is likely to provide the primary link between the LIN-12 and GLP-1 receptors and downstream genes in C. elegans, and a similar role may be played by Su(H) in Drosophila and by CBF1 in vertebrates (Fortini et al., 1994; Bailey and Posakony, 1995; Jarriault et al., 1995; Lecourtois and Schweiguth, 1995). Therefore, the Delta-like ligands, Notch-like receptors and CSL proteins are core components of the LIN-12/Notch/GLP-1 pathway that have been conserved throughout phylogeny. In Drosophila and vertebrates, regulation of the E(Spl) complex by the CSL proteins is also conserved (Bailey and Posakony, 1995; Jarriault et al., 1995; Lecourtois and Schweiguth, 1995). In nematodes, no connection has yet been made between LAG-1 and a nematode homolog of E(Spl), and therefore the extent to which the pathway is conserved downstream of lag-1 is not yet clear.

We are grateful to A. Coulson and J. Sulston for providing cosmids; C. White for the genomic library and P. Okkema and B. Barstead for cDNA libraries; B. Cali for providing the RFLP mapping strain and...
collaboration in recombinant analysis. We thank L. Berry and T. Schedl, and I. Moskowitz and J. Rothman for communicating unpublished results, Leanne Olds for her expertise with illustration, all the members of the Kimble laboratory for encouragement and discussions, and Lisa Kadyk for comments on the manuscript. This work was supported in part by grants from the NIH and the Council for Tobacco Research; M. B. is a Jane Coffin Childs Postdoctoral Fellow. J. K. is an investigator with the Howard Hughes Medical Institute.

REFERENCES


Disruption of the mouse RBP-Jk gene results in early embryonic death. Development 121, 3291-3301.


(accepted 5 February 1996)

Note added in proof:
The nucleotide sequences reported in this paper have GenBank accession numbers U49794 and U49795.