The EMBO Journal vol.11 no.13 pp.4969-4979, 1992

Regulation of the *mec-3* gene by the *C.elegans* homeoproteins UNC-86 and MEC-3

Ding Xue, Michael Finney[1], Gary Ruvkun[1] and Martin Chalfie

Department of Biological Sciences, 1012 Fairchild, Columbia University, New York, NY 10027 and [1]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA

Communicated by J.Hodgkin

The *mec-3* gene encodes a homeodomain protein with LIM repeats that is required for the specification of touch cell fate in *Caenorhabditis elegans*. Previous experiments suggested that *mec-3* expression requires the product of the *unc-86* gene, a POU-type homeoprotein, and *mec-3* itself. We have analyzed the control of *mec-3* expression by identifying potential *cis* regulatory elements in the *mec-3* gene (by conservation in a related nematode and by DNase I footprinting using *unc-86* and *mec-3* proteins) and testing their importance by transforming *C.elegans* with *mec-3lacZ* fusions in which these sites have been mutagenized *in vitro*. Both *unc-86* and *mec-3* proteins bind specifically to the promoter of the *mec-3* gene, suggesting that both proteins may be directly involved in the regulation of the *mec-3* gene. In addition, the footprint pattern with *mec-3* protein is altered in the presence of *unc-86* protein. *In vivo* transformation experiments reveal that some of the binding regions of the two proteins are needed for general positive control and maintenance of *mec-3* expression while others have no detectable, unique function. Interestingly, the *unc-86* gene appears to be required not only to initiate *mec-3* expression but also to maintain it.

Key words: Caenorhabditis elegans/DNase I footprinting/*mec-3*/promoter/*unc-86*

Introduction

Initiation and maintenance of cell differentiation are usually regulated by a hierarchy of transcriptional events. The activation of cell-type-specific regulators to initiate cell differentiation and the autoregulation of such regulators to maintain the differentiated state is quite a common theme in mammalian cell differentiation (Serfling, 1989; Karin, 1990; Emerson, 1990). We have been studying the regulation of cell-type specification in the development of a set of six touch receptor neurons in the nematode *Caenorhabditis elegans*. Previous genetic and molecular studies have led us to propose a simple genetic and developmental pathway involving genes needed to generate touch cell precursors, specify touch cell differentiation, maintain the differentiated state, and act as differentiation targets of this regulation (Chalfie and Au, 1989; Way and Chalfie, 1988, 1989). The *unc-86* gene is needed to generate the touch cells; the fates of precursors in the touch cell lineages are altered in *unc-86* mutants so that no touch cells arise (Chalfie *et al.*, 1981). The *unc-86* gene encodes a POU-type homeodomain protein (Finney *et al.*, 1988), so it is likely to affect transcription in the cells in which it is expressed. Interestingly, while the *unc-86* product is found in touch cell precursors, it is also found in cells produced by these lineages, including the touch cells (Finney and Ruvkun, 1990). This late expression suggests that *unc-86* may have a direct role in touch receptor differentiation.

Mutations in the *mec-3* gene do not affect touch cell lineages, but do affect touch cell differentiation (Chalfie and Sulston, 1981; Chalfie and Au, 1989; Way and Chalfie, 1988, 1989). Because the cells that would normally develop as touch receptors appear to be transformed into other types of neurons in *mec-3* mutants, we believe that *mec-3*, which encodes a homeoprotein with cysteine-rich LIM repeats (Freyd *et al.*, 1990; Karlsson *et al.*, 1990), is important in specifying touch cell fate. However, expression of *mec-3* alone cannot determine touch cell fate because the gene is expressed and functions in two pairs of neurons, the PVD and FLP cells, in addition to the six touch cells (Way and Chalfie, 1989). The expression of *unc-86* protein in the touch receptor neurons and the requirement of *unc-86* for *mec-3* expression suggest that *mec-3* may be directly regulated by *unc-86* (Way and Chalfie, 1989; Finney and Ruvkun, 1990).

Two genes, *mec-17* and *mec-3*, appear to be needed to maintain touch receptor differentiation. *mec-17* larvae are initially touch sensitive, but become touch insensitive as they mature (Chalfie and Au, 1989). This insensitivity is accompanied by the reduction of expression of a *mec-3lacZ* fusion in the touch cells (but not in the PVD or FLP cells; Way and Chalfie, 1989). The *mec-3* gene appears to maintain its own expression, since *mec-3lacZ* fusion expression is only seen in early, not late, larvae in *mec-3* mutants (Way and Chalfie, 1989).

Several genes, including *mec-7* and *mec-4*, appear to be target genes in the differentiation of the touch cells (Chalfie and Sulston, 1981; Chalfie and Au, 1989). *mec-7* encodes a β-tubulin that is needed for the cell-specific microtubules found in the touch receptor neurons (Savage *et al.*, 1989); this protein is found at high levels only in these cells (S.Mitani, H.P.Du, D.H.Hall, M.Driscoll and M.Chalfie, in preparation). *mec-4* encodes a putative membrane protein that can be mutated to cause the cell-specific degeneration of the touch cells (Driscoll and Chalfie, 1991); expression of a *mec-4lacZ* fusion is only detected in the six touch receptor neurons (S.Mitani, H.P.Du, D.H.Hall, M.Driscoll and M.Chalfie, in preparation). These genes are potential targets of *mec-3* action.

These observations suggest a central role for *mec-3* in the differentiation of the touch cells. In the experiments described in this paper, we have identified potential *cis*-regulatory elements controlling *mec-3* expression by comparing the promoter sequence of the *mec-3* gene in two nematode species and by using DNA binding assays with *mec-3* and *unc-86* proteins. We have used *in vitro* mutagenesis and

germline transformation to test the importance of these elements. We find that both *unc-86* and *mec-3* proteins bind specifically to *mec-3* DNA, suggesting a direct role for both gene products in *mec-3* expression. *mec-3lacZ* expression studies suggest that the binding of *unc-86* protein (designated as UNC-86) to two sites is required for *mec-3* expression. The binding of *mec-3* protein (MEC-3) to at least one site appears to be important for the maintenance of *mec-3* expression. In addition, we find that one UNC-86 binding site is also required to maintain *mec-3* expression and that UNC-86 binding affects MEC-3 binding.

Results

Isolation of mec-3 cDNAs

The previously identified *mec-3* cDNA (Way and Chalfie, 1988) had only a partial coding sequence. We obtained seven additional full-length *mec-3* cDNAs from a size-selected wild-type *C.elegans* cDNA library (see Materials and methods). These cDNAs encode the same polypeptide of 321 amino acids (see Figure 1a). The predicted protein has two cysteine-rich LIM repeats (the *mec-3* sequence was used to define the LIM motif; Freyd *et al.*, 1990) as well as the previously noted homeodomain and acid-rich region (a putative transcription activation domain; Way and Chalfie, 1988). These three structural features (albeit with different putative activator domains) are also found in the products of two other homeobox genes, *lin-11* (Freyd *et al.*, 1990) and *Isl-1* (Karlsson *et al.*, 1990). Since all seven *mec-3* cDNAs result from the same pattern of splicing, the expression of alternatively spliced transcripts to generate cell-type specificity is unlikely.

5′ flanking regions of the mec-3 gene in C.elegans and C.briggsae

To identify important *cis* elements in putative regulatory sequences of *mec-3* and conserved domains within the coding sequence, we cloned and sequenced 3.8 kb of genomic DNA for the *mec-3* homolog from another nematode species, *Caenorhabditis briggsae* (see Materials and methods). This

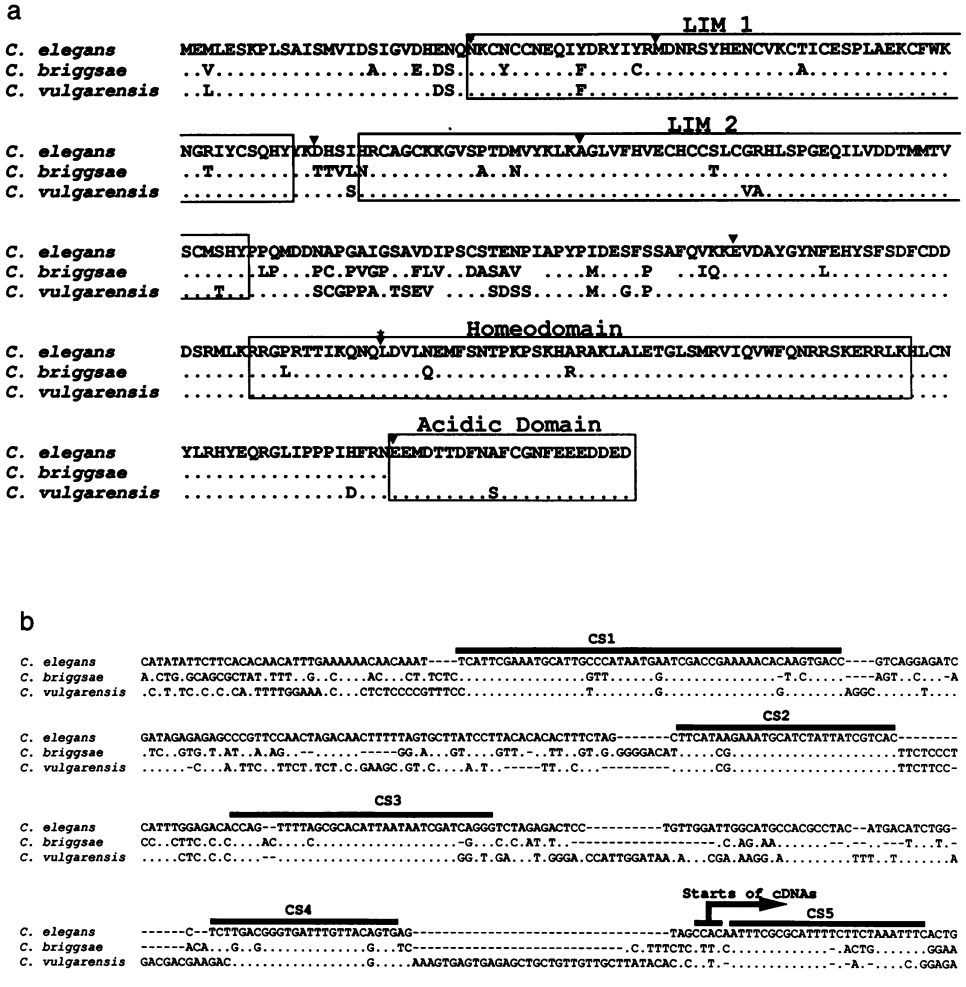

Fig. 1. Alignment of protein sequence and the promoter region of nematode *mec-3* genes. (**a**) Alignment of protein sequence. Amino acids identical to that of the *C.elegans* gene are indicated by dots. The positions of the exon junction are indicated with filled triangles. The *C.briggsae* gene lacks the intron (*) between exons six and seven of the *C.elegans* gene. The positions of the LIM repeats, homeodomains and acidic domains are shown. The 3′ end of the *C.briggsae* gene has not yet been cloned. (**b**) Alignment of promoter region. Nucleotides identical to those in the *C.elegans* gene are indicated by dots. Extensive conserved sequences (CS1–CS5) between *C.elegans* and *C.briggsae* are indicated with black lines above the sequences (these are also found in *C.vulgarensis*). The 5′ ends of the *mec-3* cDNA are shown. The *C.vulgarensis* sequences were taken from Way *et al.* (1991). Computer alignment was performed using the algorithm of Pearson and Lipman (1988).

DNA contains ~1300 bp 5′ of the putative start of transcription (as estimated from the *C.elegans* cDNAs) and DNA for all but the last exon (the one encoding the acid-rich region). The two genes are very similar, giving further support that the coding sequence deduced from the cDNAs is correct. The LIM domains and homeodomains are 89 and 95% identical, respectively, suggesting that these domains are functionally important (Figure 1a).

Upstream of the initial methionine codon are five extensive regions that are conserved in both species (Figure 1b). These regions are 51, 29, 33, 25 and 26 bp long (CS1−CS5, respectively). All but CS5 are found 5′ of the beginning of the cDNA sequences; CS5 is located at the beginning of the cDNA sequences. As this manuscript was being prepared, Way *et al.* (1991) reported a similar comparison of *mec-3* genomic sequences from *C.elegans* and *Caenorhabditis vulgarensis*. They identified four conserved regions 71, 29, 28 and 24 bp long (regions I−IV) upstream of the *mec-3* coding sequence which correspond to our CS1−CS4, respectively. Although Way *et al.* (1991) did not identify the fifth conserved region, they did note that a 13 bp sequence that we recognize as a subset of CS5 was found in both the *mec-3* genes they examined. Interestingly, we find that the CS5 regions in all three *mec-3* genes contain the sequence ATTTC at least twice (Figure 1b).

We have generated *mec-3lacZ* fusions with the *mec-3* genes of both *C.elegans* and *C.briggsae* and transformed wild-type *C.elegans* with both fusions to determine whether the genes have the same expression pattern. In the *C.briggsae mec-3lacZ* construct (TU#49), 2.47 kb of *mec-3* DNA, which contains part of the third exon and all of the upstream region, was fused into the *lacZ* coding sequence in pPD16.51 vector (Materials and methods). The *C.elegans mec-3lacZ* construct pTU28 (Way and Chalfie, 1989) has the *lacZ* coding sequence inserted in-frame into the homeobox region of the whole *mec-3* gene. As found previously (Way and Chalfie, 1989), the *C.elegans* gene construct pTU28 is reproducibly expressed in the six touch receptor neurons and the FLP and PVD neuron pairs (Figure 2A and Table I). The *C.briggsae* construct is expressed in the same 10 cells, indicating that the 2.47 kb *mec-3* DNA fragment of *C.briggsae* is sufficient for the correct cell-specific expression of *mec-3* and that the region after the third exon including the 3′ untranslated region is not required for this cell-specific expression of the *mec-3* gene (Figure 2B and Table I).

The *mec-3* promoter appears to be quite small. A deletion construct of pTU28 (Δ1) that has only 311 bp 5′ upstream of the first cDNA start and ends just 5′ of CS1 produced the same staining pattern as pTU28 (Figure 3a and Table I), suggesting that this region is sufficient for normal *mec-3* expression. Another deletion construct (Δ2) which deletes a 180 bp region containing CS1 and CS2 has no β-galactosidase staining, suggesting that CS3, CS4 and CS5 alone are insufficient to direct any *mec-3* expression (Figure 3a and Table I).

An examination of the 5′ flanking region for putative transcription factor binding sites reveals several possible regulatory domains, but, as indicated below, such analysis does not reveal all of the sites. Both CS1 and CS2 contain a sequence (AAATGCAT) that is similar to the binding sites for several POU proteins (Garcia-Blanco *et al.*, 1989; Ruvkun and Finney, 1991; Rosenfeld, 1991), suggesting that these sites are potential UNC-86 binding sites. There are also several potential MEC-3 binding sites. CS1, CS2 and CS3 all have at least one TAAT motif, a sequence suggested

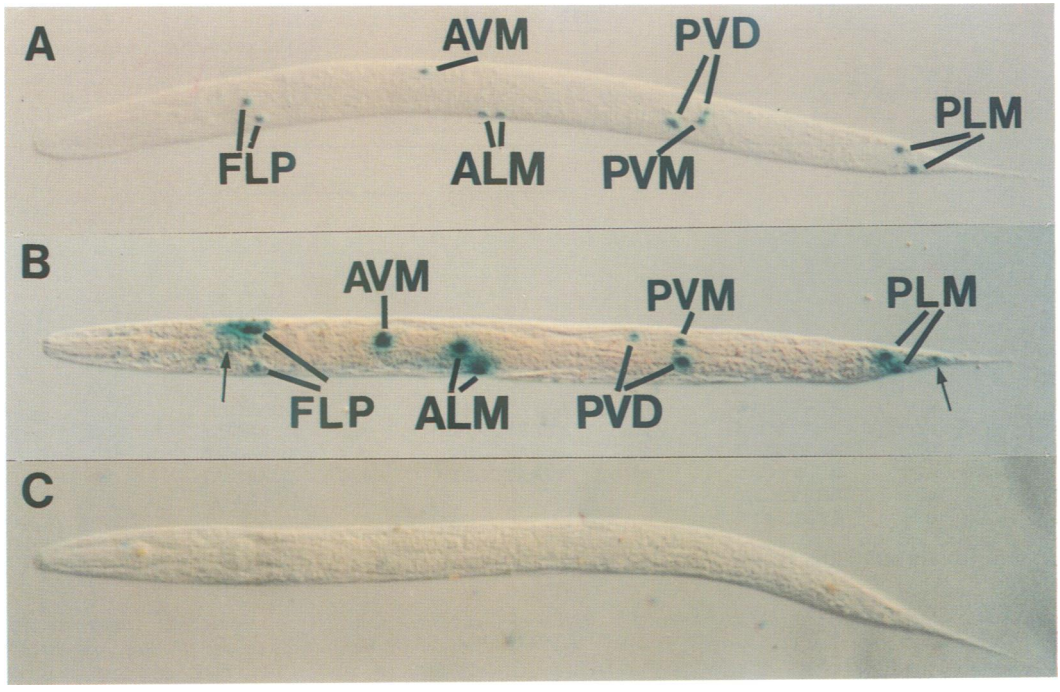

Fig. 2. β-galactosidase expression from wild-type and mutant *mec-3lacZ* fusions. Wild-type *C.elegans* contained (**A**) the wild-type *C.elegans* fusion pTU28, (**B**) the wild-type *C.briggsae* fusion TU#49, (**C**) the u3m3(−) mutant fusion. All animals shown are third stage larvae. Compared with the pTU28 construct, some differences in expression are seen in the TU#49 construct: (i) the percentage of animals staining for a given cell is somewhat lower for the *C.briggsae* DNA than for the *C.elegans* DNA (Table I); (ii) the *C.briggsae* DNA results in more intense staining than that of *C.elegans*; and (iii) often the FLP, ALM and PLM cell processes in animals with the *C.briggsae* construct are also stained (arrows).

to be a binding site for homeodomain proteins which, like MEC-3, have a glutamine at the 9th amino acid of helix 3 of the homeodomain (Hanes and Brent, 1989; Treisman *et al.*, 1989; Kissinger *et al.*, 1990). Moreover, CS3 contains a sequence (TTAATAATC) that is a binding site for the LIM homeoprotein produced from the *Isl-1* gene (Karlsson *et al.*, 1990).

MEC-3 and UNC-86 footprints

To test whether UNC-86 and MEC-3 can bind to any of these potential sites, we have performed DNase I footprinting on a 477 bp DNA fragment that includes the minimal 311 bp promoter region of *mec-3* using full-length *unc-86* protein (plus 16 extra N-terminal amino acids) and full-length *mec-3* protein, both generated in *Escherichia coli*. Proteins from

Table I. Expression of various *mec-3lacZ* constructs in wild-type animals

mec-3lacZ construct[a]	Strain	Number of animals[b]	% cells stained[c]								
			FLPL/R	AVM	ALML/R	PVDL/R	PVM	PV?	PLML/R	Embryo	Other
*WT-*C.e.*	TU1514	126	74	76	84	22	22	41	90	100	−
*WT-*C.e.*	TU1516	114	82	77	89	28	36	43	92	100	+
WT-*C.e.*	TU1419	237	74	58	71	39	38	32	76	91	+
WT-*C.e.*	TU1390	163	64	48	71	22	23	37	77	100	+
WT-*C.b.*	TU1420	147	86	28	53	27	23	40	67	88	−
WT-*C.b.*	TU1418	93	80	16	30	23	13	53	49	82	−
*Δ1	TU1511	120	81	69	88	24	23	33	88	64	−
*Δ1	TU1524	125	50	62	83	12	18	28	89	92	−
*Δ2	TU1531	92	0	0	0	0	0	0	0	0	−
Δ2	TU1410	36	0	0	0	0	0	0	2	0	−

[a]The asterisk indicates strains in which mutagenized DNA and *rol-6* were injected at a 1:10 ratio, others were injected at a 1:1 ratio. WT, wild-type; *C.e.*, *C.elegans*; *C.b.*, *C.briggsae*. The Δ1 and Δ2 deletions are diagrammed in Figure 4(a).
[b]Animals ages range from L2 to adult. No differences in staining with age were detected.
[c]Numbers listed are the percentage of cells stained. Most cell types are found in pairs except AVM and PVM which are single cells, and PV?, which could be a single PVM or PVD cell. The numbers under Embryo indicate the percentage of adults that contained at least one stained embryo. Additional stained cells (Other) were found near the positions of the FLP cells (in 10, 20 and 9% of the TU1516, TU1419 and TU1390 animals, respectively) and in the posterior ventral cord (in 7 and 2% of TU1419 and TU1390 animals, respectively). As many a six ventral cord cells stained. This ectopic staining was only seen in older larvae (L3−L4) and adults. In addition, 7% of TU1419 animals had one or two stained cells between the two pharyngeal bulbs.

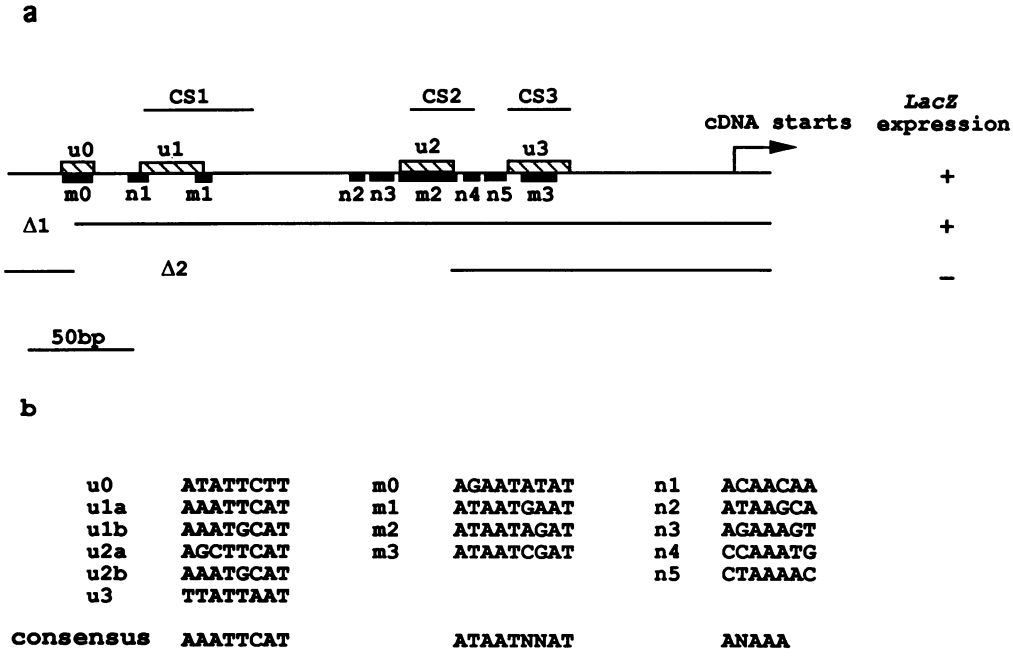

Fig. 3. Binding of MEC-3 and UNC-86 to the *mec-3* promoter. (**a**) Location of MEC-3 and UNC-86 binding sites. The central line shows the *mec-3* promoter. The arrow indicates the *mec-3* cDNA starts. The three short black lines above the central line show the three conserved sequences CS1−CS3. Hatched boxes indicate regions protected by UNC-86 in the DNase I footprint, whereas filled boxes indicate regions protected by MEC-3. The extent and expression of two *mec-3lacZ* deletions are also indicated. Δ1 contains 311 bp upstream of the cDNA starts and is sufficient to direct wild-type *mec-3lacZ* staining. Δ2 has a 180 bp internal deletion and is unable to produce any *mec-3lacZ* staining. (**b**) Alignment of the UNC-86 and MEC-3 binding sites. u0, u1a, u1b, u2a, u2b, m1, m3 and n1 sites are found on the upper strand of the gene and the other sites are found on the opposite strand. Alignment of UNC-86 binding sites (u) reveals the consensus binding motif AAATTCAT. Sites u1 and u2 each contain two copies of this motif, though the u1a and u2a sites are only partially conserved in *C.briggsae* and *C.vulgarensis*. Alignment of MEC-3 binding sites (m) that overlap with UNC-86 binding sites reveals the consensus sequence ATAATNNAT. Alignment of MEC-3 binding sites (n) outside of the UNC-86 binding regions reveals the consensus sequence ANAAA.

both genes bind to this DNA (Figure 4). Both UNC-86 and MEC-3 bind to four similar areas of the DNA; three within CS1, CS2 and CS3 (u1−u3 for UNC-86; m1−m3 for MEC-3) and one upstream of the conserved regions (u0 and m0). All UNC-86 binding regions overlap with MEC-3 binding sites, though MEC-3 binds to several other regions of the DNA to which UNC-86 does not bind (Figures 3 and 4). The UNC-86 binding sites are either positioned 5′ of the MEC-3 binding sites (CS1 and CS2) or co-centered with them (u0 and CS3). The four sequences footprinted by UNC-86 share a consensus core of AAATTCAT, which is similar to the POU binding motif. CS1 and CS2 each have two copies of this motif; the two copies in both cases are separated by 12 bp. This extended structure is characteristic of several functional POU protein binding sites (Ruvkun and Finney, 1991). UNC-86, however, appears to bind with higher affinity and with a large protected area to CS3, with its single copy of the consensus core, than it does to CS1 and CS2 (Figure 4). Since the Δ1 and *C.briggsae mec-3lacZ* constructs, both of which lack u0 and m0, resulted in essentially wild-type expression (Table I), these sites do not appear to be important for *mec-3* expression.

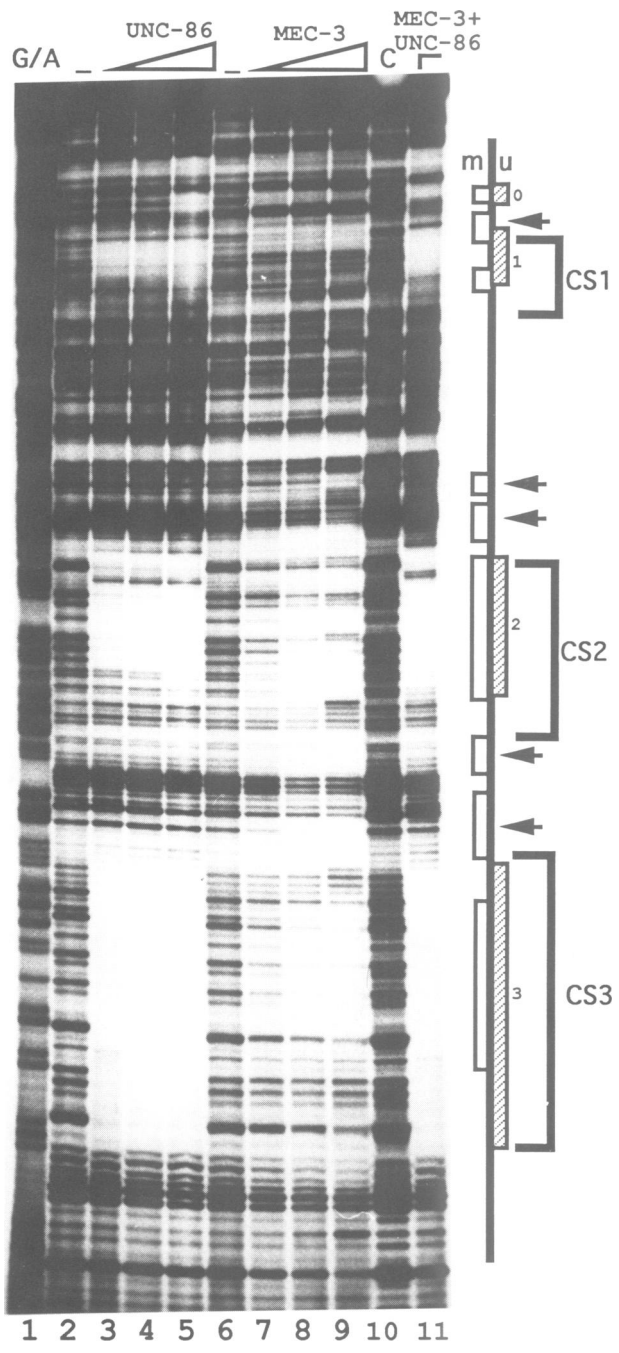

Fig. 4. DNaseI footprints of MEC-3 and UNC-86 to the *mec-3* promoter. Empty boxes and hatched boxes represent regions protected by MEC-3 and UNC-86, respectively. Three conserved sequences (CS1−CS3) are bracketed. Samples in lanes 3, 4 and 5 had 30, 120 and 360 ng of UNC-86, respectively. Those in lanes 7, 8 and 9 had 0.5, 2 and 6 μg of MEC-3, respectively. The sample in lane 11 contained 2 μg of MEC-3 and 120 ng of UNC-86. Samples in lanes 2 and 6 contained no proteins, and that in lane 10 had 6 μg of protein extract from induced cells containing the expression vector pET-3a with no insert. Lane 1 is a G/A DNA ladder. The MEC-3 binding regions which disappear in lane 11 are indicated by arrows. The fragment for footprinting was generated by PCR using oligonucleotide corresponding to positions −30 to −6 and −483 to −463; the former primer was end-labeled at position −6.

The MEC-3 binding sites have AT-rich sequences; the core consensus for those overlapping with UNC-86 binding sites (m sites) is AATAATNNAT. The other sites (n sites) appear to share no sequence homology except that they are all A-rich (ANAAA). Most of the n sites in the *C.elegans mec-3* gene are not conserved in the *C.briggsae* gene, yet the *mec-3lacZ* expression pattern from the *C.briggsae* construct is essentially similar to that from the *C.elegans* construct. Thus, these sites do not appear to be essential for *mec-3* expression and may be non-specific MEC-3 binding sites.

Because UNC-86 and MEC-3 bind to overlapping regions in the *mec-3* promoter, we tested whether the presence of UNC-86 affected MEC-3 DNA binding specificity. Interestingly, the footprint produced when both UNC-86 and MEC-3 are present differs from the ones produced by the same amount of each protein alone (lanes 4, 8 and 11 in Figure 4). All the MEC-3 binding sites that do not overlap with UNC-86 binding regions (the n sites) disappear. The disappearance of these MEC-3 binding sites does not seem to result from the general inhibition of MEC-3 binding by UNC-86, since under the conditions tested, UNC-86 does not detectably protect the u0 region and part of the u2 region, yet these regions are still protected by MEC-3. Thus, UNC-86 appears to cause the preferential binding of MEC-3 to these overlapping regions, and, as a result, increases the specificity of MEC-3 binding to DNA.

Of the five conserved non-coding regions, the combined UNC-86 and MEC-3 footprints cover all of CS3 and most of CS1 and CS2. The remaining conserved regions may bind other factors required for *mec-3* expression or maintenance.

In vivo significance of the MEC-3 and UNC-86 binding sites

The sequences conserved in the 5′ flanking regions of the *C.elegans* and *C.briggsae mec-3* genes and the footprint sites for UNC-86 and MEC-3 suggest several regions that may be functionally important for *mec-3* expression. We tested the importance of these regions by mutating the sites (Figure 5a) in the wild-type *mec-3lacZ* construct pTU28 (Way and Chalfie, 1989) *in vitro* and transforming wild-type animals with the mutagenized DNA (Table II). [In the following, mutations are labeled according to the mutated site, e.g. u1(−) is a set of point mutations that abolishes UNC-86 binding at CS1.] As described below, we found that CS2

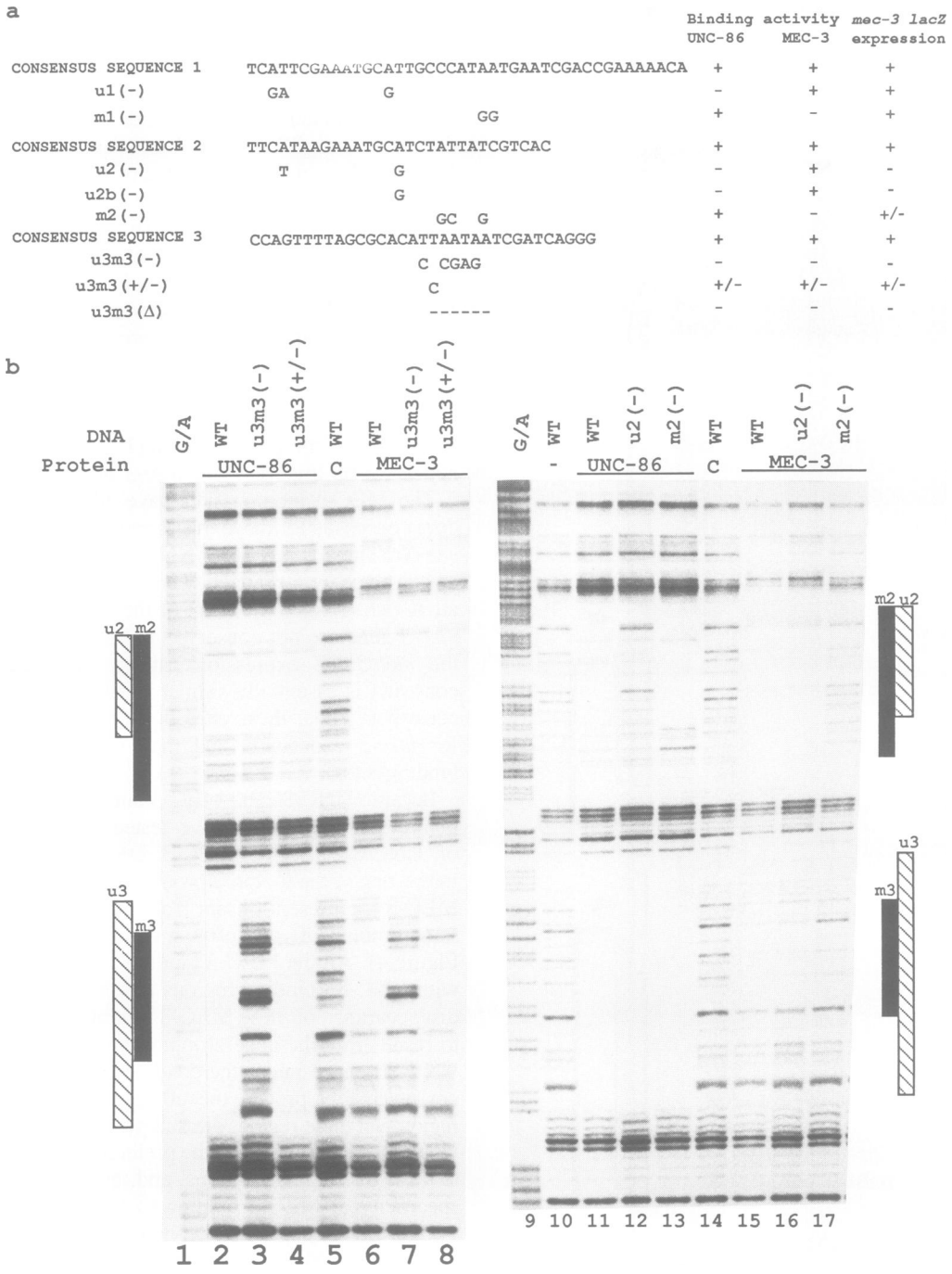

Fig. 5. Mutations in the *mec-3* promoter and their effects on protein binding. (**a**) Mutations in the *mec-3* promoter. Only the altered nucleotides are shown in the mutated sequences (− denotes the loss of a nucleotide). Wild-type binding is indicated by +, loss of binding by −, and partial binding by +/−. For *mec-3lacZ* expression, wild-type expression is indicated by +, loss of expression by −, and partial loss of expression by +/−. (**b**) DNase I footprints of MEC-3 and UNC-86 on wild-type and mutated *mec-3* DNA. Hatched boxes and filled boxes represent regions protected by UNC-86 and MEC-3 in wild-type sequences, respectively. 240 ng of UNC-86 and 4 μg of MEC-3 were used in each reaction. −, no protein added; C, 4 μg protein extract from induced cells containing pET-3a without insert; G/A, G/A DNA ladder. DNA fragments were generated by PCR using oligonucleotides corresponding to positions −30 to −6 and −324 to −304; the former primer was end-labeled at position −6.

and CS3 are important for the initial and maintained expression of *mec-3*.

As with most *C.elegans* transformants, our transformants, which contain presumed extrachromosomal arrays with multiple copies of the injected DNA, are unstable. In our initial experiments utilizing an equimolar concentration of plasmid pRF4, which results in a dominant Roller phenotype, and the *lacZ* constructs, we often saw ectopic staining (usually of up to six ventral cord cells and two cells near the FLP cells) in some strains (Tables I and II). This ectopic staining, which increases with the age of the animals, seems to result from the high copy number of the *lacZ* construct in the extrachromosomal array, since this staining was not seen in subsequent experiments in which transformants were constructed using a pRF4 to *lacZ* construct ratio of 10:1 (see Materials and methods). Presumably, the ectopic staining arises because the high copy number of construct DNA in cells of animals injected at a 1:1 ratio titrates out the limiting

amount of negative factors that normally would silence *mec-3* expression. [Interestingly, this ectopic expression does not require *unc-86* (D.Xue and M.Chalfie, unpublished data).] Similarly, a large number of copies of the construct DNA could also titrate out positive factors. Consistent with this suggestion, ~20–30% of the animals resulting from the 1:1 injections were touch-insensitive either at the head or at the tail. Where these two sets of transformants have different expression patterns, we feel that the latter set more closely reflects the *in vivo* situation because of the presumed lower copy number.

Mutation of two regions, u3m3 and u2, results in the virtual loss of *mec-3* expression. A five nucleotide change at the MEC-3 and UNC-86 binding region (TTAATA to CTCGAG) of CS3 abolishes the binding of both proteins to this region as verified by footprinting (lanes 2, 3 and 7, Figure 5b). Animals with this mutant *lacZ* construct, u3m3(−), usually show almost no β-galactosidase-expressing cells in embryos, larvae and adults. Occasionally animals with one or two expressing cells are seen; these cells are in the usual position of *mec-3*-expressing cells and stain with the same intensity as with the wild-type construct (Figure 2C and Table II).

Even a single point mutation in CS3, that only slightly reduces the binding of both UNC-86 and MEC-3 (lanes 4 and 8 in Figure 5b) causes a significant reduction of *lacZ* expression in animals carrying this mutant construct, u3m3(+/−) (Table II). These data strongly suggest that CS3 is critical for *mec-3* expression. Since we have not been able to mutate separately the UNC-86 and MEC-3 binding in this region (data not shown), we cannot identify specific effects of UNC-86 and MEC-3 at this site. In contrast to our results, Way *et al.* (1991) reported that a 6 bp deletion in CS3 (at the center of the regions footprinted by UNC-86 and MEC-3) had only a weak effect on the maintenance of *mec-3* expression. In addition, these authors found that this same construct also caused the ectopic expression of *mec-3* at two additional cells at the tail of the animals, a result which we did not observe in any of our CS3 mutant constructs. In order to resolve this apparent discrepancy, we made the same 6 bp deletion in CS3 [u3m3(Δ); Figure 5a]. This mutation abolished both MEC-3 and UNC-86 binding at CS3 (data not shown). Consistent with our results with the u3m3(−) and u3m3(+/−) mutations, we found essentially no β-galactosidase-expressing cells in animals carrying the u3m3(Δ) construct (Table II).

A two nucleotide change at position u2 greatly reduced the UNC-86 binding at u2 but did not detectably affect MEC-3 binding at CS2 (Figure 5a; lanes 10, 12 and 16, Figure 5b). Similar to the case with u3m3(−), β-galactosidase expression was virtually abolished in animals with this mutant *lacZ* construct, u2(−) (Table II). Thus, both CS3 and u2 are needed for general *mec-3* expression.

Although u2(−) does not affect MEC-3 binding at CS2, wild-type animals containing this mutant construct show no *lacZ* expression at later stages, i.e. when wild-type animals still express the endogenous *mec-3* gene. This result indicates that the maintenance of *mec-3* expression requires a functional u2 sequence, suggesting a possible role of *unc-86* in maintaining *mec-3* expression.

A strain obtained after equimolar injection of u2(−) and pRF4 displayed a slightly different pattern of expression: young (L2) larvae had normal β-galactosidase activity in PLML/R, PVD and PVM, greatly reduced activity in FLPL/R, and essentially no activity in ALML/R and AVM. These data suggest that a functional u2 site is essential for *mec-3* expression in ALML/R and AVM, but is not essential in the other cells under conditions of presumed high copy number (this strain, TU1534, has the ectopic staining described above). At the later stages (L3 to adult), β-galactosidase expression in the usual *mec-3*-expressing cells gradually disappeared, but appeared in the ectopic cells often seen in these presumably high copy-number transformants (Table II). The loss of *lacZ* expression again indicates that a functional u2 site is required for the maintenance of *mec-3* expression.

Similar staining patterns were seen when even a single nucleotide in the second UNC-86 consensus core in CS2 (u2b) was mutated. Such a change abolished UNC-86 binding at CS2 but did not affect MEC-3 binding at CS2 (data not shown). Of the two strains that carried this mutant construct [u2b(−)], the one that did not have ventral cord staining (TU1301) had the same staining pattern as the u2(−) strains TU1523 and TU1529. The other strain (TU1290), which had ventral cord staining, looked very similar to strain TU1534 (Table II).

A three nucleotide change at CS2 greatly reduced the binding of MEC-3 to CS2 yet had no detectable effect on UNC-86 binding at u2 (Figure 5a; lanes 13 and 17 in Figure 5b). Four independent strains which contain this mutant construct, m2(−), were examined. In three of the strains the frequency of *mec-3lacZ* expression in embryos and at all stages examined was greatly reduced. In the remaining strain the staining frequency was almost at wild-type levels (Table II). These results suggest that *mec-3* may be needed directly for its own maintenance.

Mutations that abolish UNC-86 binding at CS1 [u1(−)] or MEC-3 binding at CS1 [m1(−)] (data not shown) have no apparent effect on *mec-3lacZ* expression (Figure 5a and Table II). As described above, deletion of the upstream region of CS1 which includes u0 and m0 sites (Δ1) does not affect *mec-3lacZ* expression either (Figure 3a and Table I). Either u1, m1, u0 or m0 are not important for *mec-3* expression *in vivo* or their *in vivo* importance is obscured because we are examining animals with many copies of the *mec-3lacZ* DNA.

Discussion

Previous experiments suggested that both *unc-86* and *mec-3* were needed, perhaps in a direct fashion, for the proper expression of *mec-3* (Way and Chalfie, 1988, 1989; Chalfie and Au, 1989; Finney and Ruvkun, 1990). The expression of both *unc-86* (Finney and Ruvkun, 1990) and *mec-3* (Way and Chalfie, 1989) in the touch cells and our present demonstration of UNC-86 and MEC-3 binding to *mec-3* DNA and the importance of these sites for *mec-3* gene expression further support the hypothesis of direct involvement of these proteins.

The finding that the UNC-86 binding sites u3 and u2 are required for general *mec-3* expression further supports the hypothesis that *unc-86* is required for the initiation of *mec-3* expression by directly interacting with the *mec-3* gene. Interestingly, u2 also appears to be needed for *mec-3* maintenance, since late expression of mutant genes defective at this site is not seen. All of our *in vitro* mutants were tested

Table II. Expression of the *mec-3lacZ* mutants in wild-type animals

Defect[a]	Strain	Stage[b]	Number	% cells stained[c]						
				Embryo	FLP[d]	AVM	ALM	PVD/PVM	PLM	VC[d]
*Wild-type *C.elegans*	TU1514	L2	18	–	44	83	75	13	83	
		L3–L4	59	–	68	69	85	35	88	–
		A	49	100	92(2)	80	86	45	96	
Wild-type *C.elegans*	TU1419	L2	59	–	56	69	72	47	81	0
		L3–L4	115	–	74(15)	59	70	47	75	1
		A	63	91	91(20)	46	72	52	73	25
*u3m3(−)	TU1532	L2	19	–	8	0	3	2	3	
		L3–L4	52	–	3	0	0	2	6	–
		A	23	ND	20	0	0	4	15	
u3m3(−)	TU1397	L2	17	–	0	0	0	0	3	
		L3–L4	50	–	2	0	0	1	1	–
		A	23	0	0	0	0	0	0	
u3m3(−)	TU1396	L2	12	–	0	0	4	11	25	0
		L3–L4	55	–	5	0	1	5	7	2
		A	36	3	18(13)	0	0	5	4	31
*u3m3(+/−)	TU1522	L2	20	–	3	0	0	1	15	
		L3–L4	48	–	2	4	4	3	20	–
		A	44	25	13	4	4	10	26	
u3m3(+/−)	TU1279	L2	15	–	0	0	0	0	0	0
		L3–L4	58	–	11(3)	2	1	10	9	5
		A	36	36	42(20)	0	0	13	14	20
u3m3(Δ)	TU1663	L2	11	–	9	0	0	6	9	0
		L3–L4	44	–	9	0	0	2	5	0
		A	21	0	26(10)	0	0	0	0	19
u3m3(Δ)	TU1664	L2	17	–	3	0	0	0	3	
		L3–L4	37	–	1	0	0	5	1	–
		A	24	0	13	0	2	3	4	
*u2(−)	TU1529[e]	L2	24	–	0	0	0	0	0	
		L3–L4	56	–	0	0	0	1	0	–
		A	48	4	0	0	0	0	0	
*u2(−)	TU1523[e]	L2	19	–	11	0	0	21	5	
		L3–L4	59	–	4	0	0	7	1	–
		A	25	ND	6	0	0	5	0	
u2(−)	TU1534	L2	29	–	14	20	14	55	86	0
		L3–L4	73	–	25?	7	3	68	75	64
		YA	21	–	90?	0	0	37	50	100
		A	23	87	89?	0	0	6	13	100
u2b(−)	TU1290	L2	16	–	0	6	0	17	31	0
		L3–L4	28	–	20?	0	0	25	18	18
		A	11	9	73?	0	0	0	5	91
u2b(−)	TU1301	L2	16	–	3	0	0	6	3	
		L3–L4	46	–	0	2	0	7	0	–
		A	24	8	4	0	0	0	0	
*m2(−)	TU1528	L2	18	–	61	50	86	33	64	
		L3–L4	58	–	74	52	67	19	29	–
		A	37	71	82	49	66	41	38	
*m2(−)	TU1530	L2	18	–	33	39	56	15	39	
		L3–L4	62	–	19	16	18	6	11	–
		A	33	36	24	18	26	13	20	
m2(−)	TU1402	L2	10	–	0	0	0	0	0	
		L3–L4	33	–	2	0	1	0	0	–
		A	36	67	4	0	0	0	2	
m2(−)	TU1409	L2	18	–	14	17	14	11	22	
		L3–L4	96	–	10	3	4	6	2	–
		A	72	51	9	3	6	2	0	
*u1(−)	TU1501	L2	24	–	75	58	90	33	90	
		L3–L4	70	–	79	73	88	34	86	–
		A	63	100	87	41	84	44	90	
u1(−)	TU1533	L2	33	–	33	67	56	58	82	
		L3–L4	95	–	34	29	53	45	69	–
		A	56	100	55	25	46	23	45	
*m1(−)	TU1535	L2	9	–	67	77	94	44	83	
		L3–L4	37	–	88	78	88	37	82	–
		A	24	83	91	87	96	58	90	

m1(−)	TU1416	L2	34	−	38	59	81	34	91	
		L3−L4	83	−	54	53	81	31	78	−
		A	64	100	74	53	83	39	81	

[a]The asterisk indicates strains in which mutagenized DNA and *rol-6* were injected at a 1:10 ratio, others were injected at a 1:1 ratio.
[b]A indicates both young (non-egg-bearing) and gravid adults unless specified separately.
[c]Numbers listed are the percentages of cells stained. The categories are the same as in Table I except that PVM/PVD gives the percentage of these three cells that stained.
[d]Several strains injected with *mec-3* and *rol-6* DNAs at a ratio of 1:1 had more than two cells stained in the region of FLP cells. The percentage of animals with more than two cells in this region is given in parentheses. These same strains also had stained cells in the posterior ventral cord (VC). Because of this ectopic staining, it is difficult to identify the cells that stained in the FLP region in strains TU1534 and TU1290. Of the cells that normally stain in wild-type, all but those in the FLP region were seen with decreasing frequency as the animals matured. These cells may not be FLP cells and are thus labeled with a question mark.
[e]In TU1534 and TU1290, in addition to the other abnormalities in these strains (see note d), 28 and 10% of animals (L2 and L3 respectively) had at least one additional cell that stained in the tail.

in wild-type animals. If *unc-86* binding to this site were only required to initiate *mec-3* expression, we would expect endogenous *mec-3* protein (and other factors) to activate transcription. We envisage several ways in which *unc-86* may exert its effects: (i) UNC-86, as an activator, could be required directly and independently of MEC-3 for the maintenance of *mec-3* expression; (ii) UNC-86 and MEC-3 may synergistically activate *mec-3* expression; (iii) UNC-86 may act as a coupling factor for MEC-3 to target MEC-3 binding to specific sites and stabilize its binding to those sites; and (iv) the binding of UNC-86 could be permissive for *mec-3* maintenance, e.g. it could lead to changes in chromatin structure that permit MEC-3 binding. The results of UNC-86 and MEC-3 co-footprint provide support for the third possibility.

The involvement of the *unc-86* gene in the maintenance of *mec-3* expression and the presence of UNC-86 throughout the life of the animals (M.Finney and G.Ruvkun, unpublished data) raises the question of why *mec-3* maintenance occurs. Possibly, some other factor or factors needed to initiate *mec-3* expression along with UNC-86 are not maintained subsequently in the *mec-3*-expressing cells. The function of these other factors could be replaced by *mec-3* and, possibly, *mec-17* (see Introduction).

Among the MEC-3 binding sites, m2 seems to be required for maintenance of *mec-3* expression, since the disruption of this site affects *mec-3lacZ* expression. Interestingly, one strain obtained after transformation with a 1:10 ratio of m2(−) and pRF4 DNA had almost wild-type staining. We do not know the reason for this discrepancy although it may result from the different copy number of m2(−) DNA in the extrachromosomal arrays. For example, if several sites contribute to *mec-3* maintenance, dilution of factors by high copy number expression may reveal the importance of individual sites, such as m2. Nevertheless, our results suggest that MEC-3 binding at m2 may be important for the maintenance of *mec-3* expression. Consistent with our results, Way *et al.* (1991) reported that a 38 bp insertion at m2 causes the loss of the maintenance of *mec-3* expression.

Of the other MEC-3 binding sites, m0 and m1 are not required for the maintenance of *mec-3* expression under the conditions we have tested. Mutation of m1 alone has no effect on the expression pattern, and disruption of both the m1 and m2 sites has a similar expression pattern as that of m2(−) alone (data not shown). We are not certain whether the m3 site is important for this function due to our inability to separate the MEC-3 binding from UNC-86 binding at this site.

Our conclusions about the nature of the *mec-3* promoter, while agreeing in some respects with those of Way *et al.* (1991), differ in others from those reported by these workers. Both groups have found that (i) the promoter contains four regions that are conserved in several nematode species and are thus likely to be functionally important (although the size of these conserved regions is slightly different and we have identified a fifth conserved region that lies near the putative start of transcription); (ii) the POU binding site in CS1 [region I in Way *et al.* (1991)] is not essential for *mec-3* expression; (iii) the CS2 region (region II) is important for the maintenance of *mec-3* expression. Our results differ, however, with respect to the initiation of *mec-3* expression, the maintenance of this expression, and restriction of this expression to a set of 10 cells.

With regard to initiation, Way *et al.* (1991) concluded that the first conserved region but not its POU binding site was necessary and sufficient for the initiation of *mec-3* expression. Coupled with our results on UNC-86 binding, this would imply that UNC-86 is not needed for the initiation of *mec-3* transcription. Their conclusion was based on the observation that the first conserved region fused to a heterologous heat shock promoter could mediate transient *lacZ* expression in *mec-3*-expressing cells. But this result does not imply that this region is sufficient for the initiation of *mec-3* expression from the wild-type *mec-3* promoter. The UNC-86 binding site in CS1 may cooperate with the heat shock promoter or simply have a stronger effect because it is much closer to the start of transcription. Ip *et al.* (1992) have found that a non-functional activator binding site became functional when brought near a basal promoter. In contrast to the results of Way *et al.* (1991), we find that a combination of UNC-86 binding sites (u2 and CS3) are required for initiation. This result not only corresponds with but also provides direct evidence for the hypothesis from genetic analysis that the *unc-86* gene is required for the initiation of *mec-3* expression. We cannot exclude, however, a role for elements that do not bind UNC-86 within CS1 in *mec-3* initiation.

One important difference with regard to initiation is the role of the CS3 region. Way *et al.* (1991) found a weak maintenance defect when part of CS3 was deleted, whereas we find that the region is needed for initiation (we could not test whether this region is needed for maintenance). As indicated in Results, when we used this same deletion in our

transformation experiments, we did not get the results that Way *et al.* (1991) reported with their single transformed line (although the two groups used different marker DNAs). Since in our hands multiple independent strains containing two different mutations and the deletion all resulted in the loss of *mec-3lacZ* expression, we believe that CS3 is needed for the initiation of *mec-3* transcription *in vivo*.

Another difference with regard to the maintenance of *mec-3* expression is that Way *et al.* (1991) found that the 3′ end of CS2 is required, whereas we find that both the 5′ (UNC-86 binding) and 3′ (MEC-3 binding) regions are important.

Finally, Way *et al.* (1991) often observed ectopic staining in the tails of their transformants which we did not see in any of our 1:10 injected animals (14 strains) as well as in the CS3 deletion, whereas we often observed ectopic staining in the posterior ventral cord and FLP region in animals carrying high copy number fusions, which they did not report. At this time it is unclear whether in either case, the ectopic staining identifies functionally important, negatively acting elements in the *mec-3* promoter. Further experiments are needed to exclude the possibilities that this ectopic staining is a consequence of either the particular co-injected DNAs used in the experiments [Way *et al.* (1991) used a different marker DNA] or the copy number of fusion constructs or changes in chromosomal structure such as deletions which may alter the relative spacing of transcriptional machinery.

Several conserved regions in the *mec-3* promoter that bind neither UNC-86 nor MEC-3, may be required for the correct expression of this gene. These sites include CS4, CS5 and regions in CS1 and CS2. Perhaps these regions are important in restricting *mec-3* expression to 10 of the 57 neurons that express UNC-86 or in determining the level of *mec-3* expression. Further analysis of the *mec-3* promoter will help to elucidate the function of these sites.

Materials and methods

Strain maintenance

Caenorhabditis elegans strains were maintained at 25°C according to Brenner (1974) as modified by Way and Chalfie (1988).

Cloning of C.elegans mec-3 cDNAs and C.briggsae mec-3 genomic DNA

A *ClaI*−*SpeI* fragment from *mec-3* genomic DNA (this contains exons 2−7 of *mec-3*) was used to probe a 1−2 kb size-selected wild-type *C.elegans* cDNA library (C.Martin and M.Chalfie, unpublished data) constructed in SHLX2 vector (Palazzolo *et al.*, 1990). Seven *mec-3* cDNAs, all ~1.26 kb in length, were found among 3.5×10^6 plaques. These seven cDNAs were sequenced by the dideoxy method (Sanger, 1978) using oligo primers and Sequenase (United States Biochemical). We believe that the entire coding sequence is contained in the cDNAs since the first ATG codon corresponds to the eukaryotic translation initiation site consensus (A 3 bp before the ATG, G immediately after the ATG; Kozak, 1989) and is preceded, in-frame, by a stop codon. This ATG codon is also found in the *mec-3* gene from *C.briggsae* but the ATG at codon position 3 is not (see Figure 1a). The seven cDNAs have 5′ untranslated regions of 64−68 nucleotides and 3′ untranslated regions of 231−244 nucleotides. The sequences are listed under GenBank accession number L02877.

A 4.6 kb *Eco*RI fragment, which contains the entire *C.elegans mec-3* gene, from pTU23 (Way and Chalfie, 1988) was used as probe to screen a Charon 4A library of *C.briggsae* genomic DNA provided by D.Baillie (University of British Columbia). Four positive clones were isolated. All contain a 3.7 kb DNA fragment that hybridizes with *mec-3* on Southern blots. This fragment was subcloned from one of these clones into pBluescript KS(−). A nested set of deletions in one orientation was constructed using exonuclease III, and the insert was sequenced as above. The 5′ untranslated region was sequenced in both directions. The *C.briggsae mec-3* sequence has GenBank accession number L02878.

Expression of mec-3 and unc-86 proteins

Site-directed mutagenesis (Kunkel, 1985) was performed to create an *NdeI* site in the *mec-3* cDNA at a position corresponding to the initial methionine codon. A *NdeI*−*Hin*dIII fragment, which contains the whole ORF of *mec-3*, was cloned into the T7 expression vector pET-3a (Rosenberg *et al.*, 1987), resulting in plasmid TU#47. Expression was induced in the cell line BL21(DE3) (Studier and Moffat, 1986). Cells containing pET-3a or TU#47 were grown in parallel at room temperature in 250 ml of 2 × YT medium, 0.4% glucose, 200 μg/ml ampicillin to an OD_{600} of ~0.6, then induced with 0.8 mM IPTG, and grown for another 2 h. The cells were harvested and lysed by treatment of lysozyme and sonication in buffer Z (Hoey and Levine, 1988) supplemented with 50 μM $ZnCl_2$ and 1 mM β-mercapto-ethanol (buffer ZZM). The lysates were centrifuged at 22 000 *g* for 10 min. The pellets were washed with 1% NP-40, solubilized with 8 ml of 6 M urea in the same buffer and dialyzed against 1 l of 1 M urea in ZZM overnight and then with 1 l of ZZM alone twice each for 6 h. The soluble fractions of this step were used for DNase I footprint analysis. [Li *et al.* (1991) have suggested that LIM proteins might also bind iron; we have made MEC-3 in the presence of both zinc and iron and although this material gave similar results in gel shift assays, it produced poorer DNase I footprints as compared with the material made with zinc (D.Xue and M.Chalfie, unpublished).]

UNC-86 was expressed from the pET-3a vector as described (Finney and Ruvkun, 1990). Inclusion bodies were washed in 50 mM Tris (pH 8.0), 10 mM EDTA and 0.5% Triton X-100, pelleted and dissolved by boiling in 50 mM Tris (pH 8.0), 10 mM EDTA and 1% SDS. The protein was >95% pure by SDS−PAGE. Protein used for binding experiments was dialyzed extensively against distilled water and then against 0.05 × PBS. The denaturation and renaturation of UNC-86 were similar to that of MEC-3 except that buffer Z was used.

DNA binding assays

DNase I footprint analysis was performed essentially as described by Hoey and Levine (1988). The DNA fragments were generated by polymerase chain reaction (PCR) by labeling one of the primers. The PCR is performed in 32 cycles at 94°C for 1 min, 58°C for 1 min and 72°C for 30 s. UNC-86 or MEC-3 was incubated with ~10 000 c.p.m. of labeled DNA fragment in 50 mM KCl, 25 mM HEPES (pH 7.9), 1 mM DTT, 0.05% NP-40, 5% glycerol, 5 μg BSA and 1 μg poly(dG−dC) (total volume: 50 μl) for 1 h at 4°C, then 30 min at room temperature. The samples were digested with DNase I for 40 s at room temperature by addition of 50 μl of 10 mM $MgCl_2$, 5 mM $CaCl_2$ and 9.0 ng of DNase I (Worthington) and stopped by the addition of 90 μl 10 mM EDTA (pH 8.0), 1% SDS and 10 μg yeast tRNA. The samples were extracted twice with 1:1 phenol:chloroform, ethanol precipitated, dissolved in 10 μl of loading buffer and loaded onto 8% polyacrylamide gels with 7.5 M urea.

Construction and transformation of mec-3lacZ mutants

The *Eco*RV−*Bam*HI fragment of pTU23, which contains the entire *mec-3* promoter region (Way and Chalfie, 1988), was subcloned into pBluescript KS(−). *In vitro* mutagenesis was performed on the resulting plasmid TU#48 by the method of Kunkel (1985). The presence of the desired mutations was confirmed by DNA sequencing. A 3.15 kb *Bgl*II−*Bam*HI fragment which contains the desired mutation then replaced the corresponding wild-type fragment in pTU28 (Way and Chalfie, 1989). Δ1 was made by creating a *Bgl*II site in TU#48 corresponding to position 1640 in the *mec-3* gene (Way and Chalfie, 1988) and replacing the wild-type *Bgl*II−*Bam*HI fragment in pTU28 with 1.51 kb *Bgl*II−*Bam*HI fragment from this mutated plasmid. Δ2 was made by creating *Nhe*I sites in TU#48 at positions 1640 and 1819 in *mec-3* (Way and Chalfie, 1988), digesting the construct with *Nhe*I and religating. The 2.97 kb *Bgl*II−*Bam*HI fragment from the resulting plasmid was used to replace the corresponding wild-type fragment in pTU28. The *C.briggsae mec-3lacZ* construct was made by subcloning the 2.47 kb *Pst*I fragment of the *C.briggsae mec-3* gene into pPD16.51 (Fire *et al.*, 1990) kindly provided by Andy Fire. The *lacZ* coding sequence is followed by the 3′ untranslated region of the *C.elegans unc-54* gene in this vector. Wild-type *C.elegans* were transformed by injecting DNAs into adult gonads as described by Mello *et al.* (1991). The *mec-3lacZ* mutant plasmids were co-injected with plasmid pRF4, which includes *rol-6(su1006)*, a dominant allele that causes animals to roll and facilitates identification of transformed progeny. For *mec-3lacZ/rol-6* injected at a ratio of 1:10, the concentrations of *mec-3laZ* and *rol-6* were 5 and 50 μg/ml, respectively. For *mec-3laZ/rol-6* injected at a ratio of 1:1, the concentration of *mec-3lacZ* and *rol-6* were both 50 μg/ml. None of the injected DNAs integrated into the *C.elegans*

genome, so transformed strains were maintained by picking Rol animals onto new plates. Only Rol animals were stained for β-galactosidase activity following the procedure of A.Fire (personal communication). In brief, Rol animals were fixed in 50% acetone for at least 20 min, washed twice with 50 mM sodium phosphate (pH 7.5) and 1 mM $MgCl_2$ and twice with staining mix buffer (50 mM sodium phosphate, pH 7.5, 1 mM $MgCl_2$, 10 mM potassium ferrocyanide, 10 mM potassium ferricyanide and 0.004% SDS), and then stained in staining mix buffer with 0.12% of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) for 48 h. Animals were also stained with 2.5 μg/ml diamidinophenylindole (DAPI) to assist in identifying cells and characterizing the developmental stage of the animals (Li and Chalfie, 1990). For most of the initial lines (i.e. those injected at a 1:1 ratio), observations were made on animals on more than one occasion. These different batches always gave virtually the same results. Results from the batch with the largest number of animals are given in the tables.

Acknowledgements

We thank Maud Kinnell and Yuan Tu for technical help, Jing Jiang for advice on DNase I footprint analysis, and Andy Fire for kindly providing us with the *lacZ* vectors. We are grateful to Ron Prywes, Stephen Mount, Tulle Hazelrigg, and the members of our laboratories for advice and comments on the manuscript. This work was supported by US Public Health Service grant GM30997 and a McKnight Development Award to MC. and by a grant from Hoeschst AG and a McKnight Development Award to G.R.

References

Brenner,S. (1974) *Genetics*, **7**, 71–94.
Chalfie,M. and Au,M. (1989) *Science*, **243**, 1027–1033.
Chalfie,M. and Sulston,J.E. (1981) *Dev. Biol.*, **82**, 358–370.
Chalfie,M., Horvitz,H.R. and Sulston,J.E. (1981) *Cell*, **24**, 59–69.
Driscoll,M. and Chalfie,M. (1991) *Nature*, **349**, 588–593.
Emerson,C.P. (1990) *Curr. Opin. Cell Biol.*, **2**, 1065–1075.
Finney,M. and Ruvkun,G. (1990) *Cell*, **63**, 895–905.
Finney,M., Ruvkun,G. and Horvitz,H.R. (1988) *Cell*, **55**, 757–769.
Fire,A., Harrison,S.W. and Dixon,D. (1990) *Gene*, **93**, 189–198.
Freyd,G., Kim,S.K. and Horvitz,H.R. (1990) *Nature*, **334**, 876–879.
Garcia-Blanco,M.A., Clerc,R. and Sharp,P.A. (1989) *Genes Dev.*, **3**, 739–745.
Hanes,S.D. and Brent,R. (1989) *Cell*, **57**, 1275–1283.
Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858–861.
Ip,Y.T., Park,R.E., Kosman,D., Yazdanbakhsh,K. and Levine,M. (1992) *Genes Dev.*, **6**, 1518–1530.
Karin,M. (1990) *Curr. Opin. Cell Biol.*, **2**, 996–1002.
Karlsson,O., Thor,S., Norberg,T., Ohlsson,H. and Edlund,T. (1990) *Nature*, **344**, 879–882.
Kissinger,C.R., Liu,B., Martin-Blanco,E., Kornberg,T.B. and Pabo,C. (1990) *Cell*, **63**, 579–590.
Kozak,M. (1989) *J. Cell Biol.*, **108**, 229–241.
Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.
Li,C. and Chalfie,M. (1990) *Neuron*, **4**, 681–695.
Li,P.M., Reichert,J., Freyd,G., Horvitz,H.R. and Walsh,C.T. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 9210–9213.
Mello,C.M., Kramer,J.M., Stinchcomb,D. and Ambros,V. (1991) *EMBO J.*, **10**, 3959–3970.
Palazzolo,M.J., Hamilton,B.A., Ding,D., Martin,C.H., Mead,D.A., Mierendorf,R.C., Raghavan,K.V., Meyerowitz,E.M. and Lipshitz,H.D. (1990) *Gene*, **88**, 25–36.
Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2444–2448.
Rosenfeld,M.G. (1991) *Genes Dev.*, **5**, 897–907.
Rosenberg,A.H., Lade,B.N., Chui,D.S., Lin,S.W., Dunn,J.J. and Studier,F.W. (1987) *Gene*, **56**, 125–135.
Ruvkun,G. and Finney,M. (1991) *Cell*, **64**, 475–478.
Sanger,F., Nicklen,S. and Coulson,A.R. (1978) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Savage,C., Hamelin,M., Culotti,J.G., Coulson,A., Albertson,D.G. and Chalfie,M. (1989) *Genes Dev.*, **3**, 870–881.
Serfling,E. (1989) *Trends Genet.*, **5**, 131–133.
Studier,F.W. and Moffat,B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.
Treisman,J., Gonczy,P., Vashishtha,M., Harris,E. and Desplan,C. (1989) *Cell*, **59**, 553–562.
Way,J.C. and Chalfie,M. (1988) *Cell*, **54**, 5–16.
Way,J.C. and Chalfie,M. (1989) *Genes Dev.*, **3**, 1823–1833.
Way,J.C., Wang,L.L., Run,J.Q. and Wang,A. (1991) *Genes Dev.*, **5**, 2199–2211.

Received on July 21, 1992; revised on September 17, 1992