dpy-18 Encodes an α-Subunit of Prolyl-4-Hydroxylase in *Caenorhabditis elegans*

Katherine L. Hill,*,† Brian D. Harfe,† Carey A. Dobbins† and Steven W. L'Hernault*,†

**Program in Genetics and Molecular Biology, Graduate Division of Biological and Biomedical Sciences and †Department of Biology, Emory University, Atlanta, Georgia 30322*

Manuscript received September 21, 1999
Accepted for publication March 27, 2000

ABSTRACT

Collagen is an extracellular matrix (ECM) component encoded by a large multigene family in multicellular animals. Procollagen is post-translationally modified by prolyl-4-hydroxylase (EC 1.14.11.2) before secretion and participation in ECM formation. Therefore, collagen processing and regulation can be studied by examining this required interaction of prolyl-4-hydroxylase with procollagen. High-resolution polymorphism mapping was used to place the *Caenorhabditis elegans dpy-18* gene on the physical map, and we show that it encodes a prolyl-4-hydroxylase α catalytic subunit. The Dpy phenotype of *dpy-18(e364)* amber mutants is more severe when this mutation is *in trans* to the noncomplementing deficiency *tDf7*, while the *dpy-18*(*e499*) deletion mutant exhibits the same phenotype as *dpy-18(e499)/tDf7*. Furthermore, *dpy-18* RNA interference (RNAi) in wild-type worms results in Dpy progeny, while *dpy-18* (RNAi) in *dpy-18(e499*) mutants does not alter the Dpy phenotype of their progeny. These observations suggest that the *dpy-18* null phenotype is Dpy. A *dpy-18*::*gfp* promoter fusion construct is expressed throughout the hypodermis within the cells that abundantly produce the cuticle collagens, as well as in certain head and posterior neurons. While prolyl-4-hydroxylase has been studied extensively by biochemical techniques, this is the first report of a mutationally defined prolyl-4-hydroxylase in any animal.

COLLAGENS are an important structural component of the extracellular matrix (ECM) in all multicellular animals. About 30 genes encode collagen in mammals, and ~25% of total body protein is some form of collagen (reviewed by van der Rest and Garrone 1991). The nematode *Caenorhabditis elegans* has >100 genes that encode collagen (Cox *et al.* 1984), and at least 60 of these genes are known to be expressed (Politz and Edgar 1984). Two of the *C. elegans* collagen genes, *let-2* and *emb-9*, encode type IV basement membrane collagens (Guo and Kramer 1989), while the remaining genes encode cuticle collagens. *C. elegans* expresses collagen throughout its life cycle, but the amount and type of collagen expressed vary greatly according to developmental stage. In general, expression of cuticle collagens increases during the four larval molts, when a new cuticle is synthesized (Cox *et al.* 1981a; Cox and Hirsh 1985). In addition, several collagen-encoding genes are expressed only by dauer larvae, which are a stress-resistant developmental stage (Cox *et al.* 1981b; Kramer *et al.* 1985).

A structural feature common to all collagens is the presence of at least one Gly-X-Y tripeptide repeat that interacts with the Gly-X-Y domains of two other collagen polypeptides to form a triple helix. The X and Y positions within the tripeptide repeat are often occupied by proline. Collagen proproteins undergo co- and post-translational modification within the endoplasmic reticulum (ER) before they are secreted from the cell to be incorporated into the ECM (for a review of collagen post-translational processing, see Kivirikko *et al.* 1992). In particular, proline residues in the Y position are often modified to 4-hydroxyproline by the enzyme prolyl-4-hydroxylase. This modification allows collagen polypeptides to hydrogen bond to each other during triple-helix formation and is essential for triple-helix stability at physiological temperatures (reviewed by Brodsky and Ramshaw 1997). Moreover, an incompletely hydroxylated procollagen polypeptide remains bound to prolyl-4-hydroxylase, is retained within the ER lumen, and is unable to contribute to the ECM (Walmsley *et al.* 1999).

Some aspects of collagen biochemistry are similar in *C. elegans* and in vertebrates, but the distribution of collagen subtypes found in *C. elegans* is significantly different from that found in vertebrates. For instance, most vertebrate collagens are fibrillar (types I, II, and III), but *C. elegans* collagens have the highest sequence and predicted structural homology to two classes of vertebrate nonfibrillar collagen. The *let-2* and *emb-9* genes encode network-forming type IV basement membrane collagens, while the cuticle is composed principally of collagens that are most similar to the nonfibrillar vertebrate fibril-associated collagen with interrupted triple helix (FACIT; Gordon and Olsen 1990; Shaw and Olsen 1991; Kramer 1994a,b). Vertebrate FACIT collagens are hypothesized to control collagen fibril geome-

Corresponding author: Steven W. L'Hernault, Department of Biology, Emory University, 1510 Clifton Rd., Atlanta, GA 30322.
E-mail: bioslh@biology.emory.edu

try, fibrillar collagen interactions with other ECM proteins, or the flexibility/compressibility of the ECM. Although their precise role(s) is uncertain, vertebrate FACITs appear to function in association with fibrillar collagens.

In vertebrates, collagen triple-helix stability requires post-translational modifications that include formation of hydroxyproline. *C. elegans* collagen contains ~12% hydroxyproline (Kramer 1994a,b), suggesting that this collagen post-translational modification is important for collagen function in this invertebrate. *C. elegans* lives at cooler temperatures than the warm-blooded vertebrates used for most collagen biochemical studies, and it is known that the requirement for hydroxyproline during collagen assembly can be affected by temperature. For instance, vertebrate fibrillar collagen that lacks hydroxyproline can be assembled into triple-helical filaments *in vitro* at low, nonphysiological temperatures, but these filaments fall apart at the normal body temperature. However, one vertebrate collagen type that is similar to that found in the *C. elegans* cuticle, FACIT type XII collagen, requires proline hydroxylation for assembly of the triple helix at all assayed temperatures (Mazzorana *et al.* 1993). This suggests that proline hydroxylation should be required to assemble the *C. elegans* cuticle at the temperatures (16°–25°) normally employed to grow this organism in the laboratory.

Prolyl-4-hydroxylase consists of catalytically active α-subunits and protein disulfide isomerase (PDI) β-subunits (EC 5.3.4.1; Koivu *et al.* 1987; Pihlajaniemi *et al.* 1987) that form an α2β2 tetramer in most animals (reviewed by Kivirikko and Pihlajaniemi 1998). The biochemistry of prolyl-4-hydroxylase, and in particular the α-subunit, has been studied extensively because of its essential role in collagen biosynthesis. A cDNA for a *C. elegans* α-subunit was shown to encode a protein that could form a catalytically active dimer with a human β-subunit (Veijola *et al.* 1994). This *in vitro* result was initially surprising because this enzyme is usually active only as a tetramer. However, it was later shown that the *C. elegans* active enzyme is also an αβ-dimer *in vitro* (Veijola *et al.* 1996).

Here we report that this previously identified *C. elegans* α-subunit of prolyl-4-hydroxylase (Veijola *et al.* 1996) is encoded by *dpy-18*, and this is the first mutationally defined prolyl-4-hydroxylase in any animal. We have sequenced two *dpy-18* mutations and show that the null phenotype for *dpy-18* is Dpy, demonstrating that this prolyl-4-hydroxylase α-subunit plays an important role during morphogenesis. The *dpy-18* promoter is active in the hypodermis, consistent with the production of cuticle collagens in this tissue throughout the *C. elegans* life cycle. *C. elegans* produces many different collagens that are probably all post-translationally modified by prolyl-4-hydroxylase. Therefore, DPY-18 might provide both new insights into the regulation of collagen biosynthesis and a target for antinematode pharmaceuticals.

MATERIALS AND METHODS

Strains, culture conditions, nomenclature, and genetic markers: Bristol N2 is the reference wild-type strain used in all *C. elegans* experiments (Brenner 1974). The strains used in this study, with the exception of CB4856, were all derived from N2. The *C. elegans* strain CB4856 was isolated in 1972 from a pineapple field in Hawaii (Hodgkin and Doniach 1997) and was provided by T. Schedl. All genetic manipulations and culturing of *C. elegans* were performed as previously described (Brenner 1974), and standard nomenclature was used (Horvitz *et al.* 1979). The following genes and mutations were used in this study: LGIII: *vab-7(e1562)* (Hodgkin 1983), *fer-2(hc2ts)* (Ward *et al.* 1981), *spe-16(hc54ts, eb35ts)* (Shakes 1988; K. L. Hill and S. W. L'Hernault, unpublished data), *dpy-18(e499, e364)* (Brenner 1974); LGIV: *him-3(e1147)* (Hodgkin *et al.* 1979); LGV: *him-5(e1490)* (Hodgkin *et al.* 1979). The balanced deficiency strain *unc-32(e189) tDf7/qC1 dpy-19(e1259) glp-1(q339); him-3(e1147)* was provided by H. and R. Schnabel.

Body length measurement of *dpy-18* homo- and hemizygotes: *dpy-18(e499)* hemizygotes were obtained by crossing homozygous *dpy-18(e499)* L4 hermaphrodites to heterozygous males carrying the noncomplementing deficiency *tDf7*. The P_0 adults were allowed to lay eggs on plates for 1 day and they were transferred individually to fresh plates on each successive day. After ascertaining the presence of outcrossed progeny on the first-day mate plates, L4 Dpy *dpy-18(e499)/tDf7* hermaphrodites were picked from the second-day mate plates to fresh plates, incubated at 16° overnight, and measured the following day. Age-matched homozygous *dpy-18(e499)* control hermaphrodites were also picked and incubated at 16° at the same time. *dpy-18(e364)* hemizygotes and homozygous controls were obtained in a similar manner, using *spe-16(hc54ts) dpy-18(e364)*.

Hemi- and homozygous animals were immobilized in 2.5 mm levamisole and mounted on 2% agar pads under coverslips. Worms were observed by differential interference contrast (DIC) at ×10 on a Zeiss Axiophot compound microscope with a DAGE CCD300T-RC camera (Dage-MTI Inc., Michigan City, IN). Resulting images were captured with a Scion LG3 Frame capture board (Scion Corp., Frederick, MD) and manipulated with NIH Image software (Wayne Rasband, NIH). Images were subsequently displayed on a monitor with Adobe Photoshop PC v.5.0 (Adobe Systems, San Jose, CA). A length of dental floss was stretched from the tip of the pharynx to the tip of the tail of each worm, and the length in centimeters was recorded from the dental floss on a sheet of acetate overhead projector film. An image of a stage micrometer slide captured at the same magnification and manipulated in the same fashion allowed measurements of worms in centimeters to be converted to microns. Data analysis was performed using the Microsoft Excel 98 program (Microsoft, Redmond, WA); all *t*-tests are two-tailed, assuming unequal variance.

DNA nomenclature and sequencing: As described below, *dpy-18* encodes an α-subunit of prolyl-4-hydroxylase, and a cDNA sequence for this gene has been published without knowledge of its association with *dpy-18* (Veijola *et al.* 1994). Additionally, the genomic sequence of the chromosome III α-subunit of prolyl-4-hydroxylase and several corresponding expressed sequence tags (ESTs) have been described; the predicted gene is Y47D3B.10 (Figure 1A). Y47D3B.10 is encoded by the yeast artificial chromosome (YAC) Y47D3 and is partially present on the cosmid T28D6. In this article, the first A of the start AUG (as proposed by Veijola *et al.* 1994 and for Y47D3B.10) is numbered 1 and the last A of the UAA stop

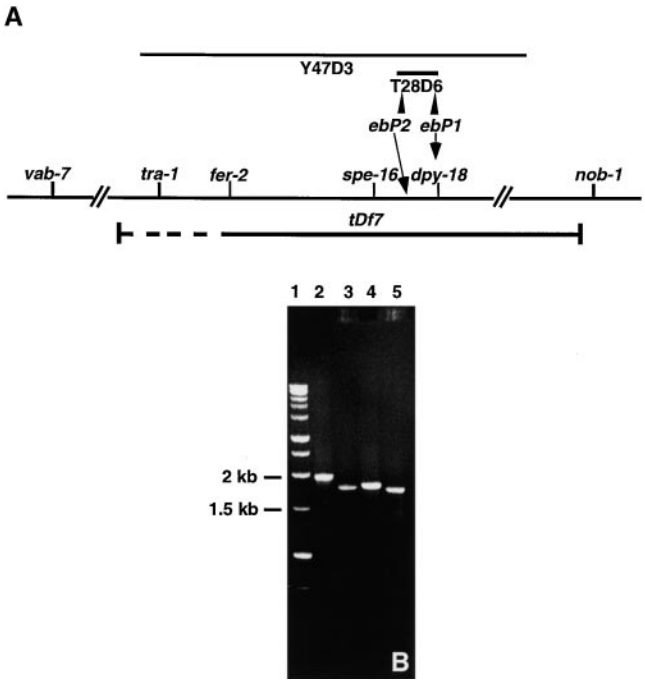

Figure 1.—(A) *dpy-18* region of chromosome IIIR. The deficiency *tDf7* fails to complement *dpy-18*, *spe-16*, and *fer-2*, but complements *vab-7* and *nob-1*. The ~7 map units between the cloned genes nearest to *dpy-18*, *tra-1* (left), and *nob-1* (right) correspond to ~645 kb. The prolyl-4-hydroxylase α-subunit gene encoded by *dpy-18* is wholly within the YAC Y47D3 and is partially present on the cosmid T28D6, and the 2-map-unit interval between *tra-1* and *dpy-18* is ~270 kb. *ebP1* and *ebP2* denote PCR product-size polymorphisms that lie within and left of *dpy-18*, respectively. (B) Agarose gel electrophoresis of DNA size standards (lane 1) and PCR products (lanes 2–5). The primer pair 5 and 6 (see Table 1) was used to detect *ebP1*, which is ~1.9 kb in Bristol N2 control genomic DNA (lane 2) and 1.7 kb in CB4856 (lane 3). The primer pair 9 and 10 (see Table 1) was used to detect *ebP2*, which is ~1.8 kb in Bristol N2 control genomic DNA (lane 4) and 1.65 kb in CB4856 (lane 5). The *dpy-18* gene is transcribed from right to left.

codon is numbered 4934 (based on the genomic sequence that encodes Y47D3B.10). This numbering scheme includes all introns, and positive numbering continues past the stop codon into the 3′ untranslated region (UTR). Nucleotides that are 5′ to the start AUG are indicated by negative numbers.

All DNA sequencing in this study was performed by the DNA sequencing facility at Iowa State University, using standard ABI (Perkin-Elmer Corp., Foster City, CA) automated fluorescent sequencing methods for polymerase chain reaction (PCR) products. The chromosome III prolyl-4-hydroxylase α-subunit gene was PCR-amplified from genomic DNA derived from *dpy-18(e364)* and *dpy-18(e499)*. The PCR products from four independent reactions for each genotype were pooled, purified using phenol-chloroform extraction or the Qiaquick PCR purification kit (QIAGEN, Valencia, CA), and sequenced to identify the mutant lesions. Sequence traces from the *dpy-18* mutants were compared with the corresponding control sequence traces from N2 and also to the sequence available from the Sanger Centre (Hinxton, England).

Polymorphism mapping: Previously, *dpy-18* was mapped ~2 map units right of *tra-1* on chromosome III (Hodgkin 1997). We narrowed the DNA region that contains *dpy-18* by mapping relative to sequence polymorphisms (Jakubowski and Kornfeld 1999) because this region is characterized by unstable recombinant DNA clones (Hodgkin 1993; K. L. Hill and S. W. L'Hernault, unpublished data) and there are no cloned genes close to *dpy-18*. This approach requires locating suitable sequence polymorphisms, and our new data, together with previously determined sequence data (see above), guided design of single-stranded DNA primer pairs. Each primer was designed to a predicted exon of a predicted open reading frame, and a primer pair was designed for every ~10 kb of genomic sequence. Each primer pair allows amplification of an ~1–2-kb DNA fragment that includes at least one predicted intron, because noncoding regions usually have more frequent sequence alterations than exons. PCR fragments were amplified from genomic DNA prepared from Bristol N2 and the Hawaiian *C. elegans* isolate CB4856 by standard methods (Barstead and Waterston 1991). Some primer pairs generated PCR products that showed a size difference between the two different *C. elegans* strains, and one such polymorphism, *ebP1*, was used to identify a right limit on the position of *dpy-18*. Primer pair 5 and 6 reveals fragment size polymorphism *ebP1* while primer pair 9 and 10 reveals an additional fragment size polymorphism, *ebP2*, when PCR products are fractionated by agarose gel electrophoresis (Table 1; Figure 1B).

CB4856 males were crossed to *vab-7(e1562) fer-2(hc2ts) spe-16(hc54ts) dpy-18(e364)* self-sterile hermaphrodites (raised restrictively at 25°); F_1 outcross heterozygous hermaphrodites were raised at 20°, and their F_2 progeny were screened for Vab non-Dpy recombinants. We obtained 21 Vab non-Dpy recombinants, and homozygous recombinant lines were used to prepare genomic template DNA (Barstead and Waterston 1991) for polymorphism analysis. PCR products from each recombinant template were typed by agarose gel electrophoresis to map *ebP1* with respect to *dpy-18*.

RNA interference: Three EST clones (yk339d8, yk276b3, and yk247d7) corresponding to the *C. elegans* prolyl-4-hydroxylase α-subunit gene on chromosome III were obtained from Yuji Kohara (Genome Biology Lab, Center for Genetic Resource Information, National Institute of Genetics, Mishima, Japan). The cDNA yk247d7, which contains a *dpy-18* insert that lacks only the first 19 nucleotides of protein-coding sequence, was used for these studies. RNA interference (RNAi; Fire *et al.* 1998), performed using dsRNA generated from this cDNA, produced an embryonic lethal phenotype substantially different from that of any known *dpy-18* alleles. The yk247d7 cDNA contains several regions of high sequence homology to a second prolyl-4-hydroxylase α-subunit present in the *C. elegans* genome (F35G2.4, CAA93466–CAA93470, F35G2 cosmid accession no. Z69637); therefore the dsRNA generated from yk247d7 may have targeted both prolyl-4-hydroxylase α-subunits. To prevent RNAi crosstalk from affecting expression of both genes, a specific template derived from yk247d7 was prepared. yk247d7 was truncated with the restriction enzyme *Kpn*I (Figure 2A), and the resulting 337-bp fragment of 5′ *dpy-18* sequence does not show significant nucleotide homology to F35G2.4. The resulting cloned construct, pKH19, was the template for *in vitro* synthesis of a double-stranded RNA that was injected into either N2 or *dpy-18(e499)* hermaphrodites by standard techniques (Fire *et al.* 1998). The injected animals were incubated at 25°, transferred to fresh plates at 24-hr intervals for 2 days following injection, and the resulting progeny were examined after an additional 24–48 hr of incubation. Embryos were scored as arrested if they had not hatched 24 hr after removal of the P_0. The age-matched sibs of injected *dpy-18(e499)* hermaphrodites were picked and transferred in parallel to provide control progeny for comparison. Offspring of both injected and uninjected *dpy-18(e499)* animals fell into four categories: Dumpy (Dpy), Lumpy Dumpy (LpyDpy; me-

TABLE 1

Single-stranded DNA primers used for PCR

Primer	5′ to 3′ sequence	Position in *dpy-18* sequence
1	**ACATGCATGC**TTTTGGCTCTCCTAAGTTTCAGC	5′ untranslated sequence, sense primer
2	CTTTCATTGCAGAGATATGCGTG	Intron 4, antisense primer
3	TACGCGGATCAGGGAAATTACAC	Exon 3, sense primer
4	**ATAGTTTAGCGGCCGC**CTAGGGAATTGTCGGCTGCTTTC	3′ untranslated region, antisense primer
5	GGCGTTTGCACTGTACAAGCAG	Exon 5, sense primer
6	CTTTGGCAAGATTTGGGGACACG	Exon 8, antisense primer
7	CTCGTCGCCGAGTCTTCTCTG	Exon 4, antisense primer
8	**TTTTCTGCAG**CATTCTGAAAAATTGAGA	Intron 1 and exon 2, antisense primer
9	TCTACACCGGCGACAGTTATG	NA; sense primer left of *dpy-18*
10	CACAGACGATGGTGATCAACG	NA; antisense primer left of *dpy-18*
11	GGAATTCGGTTTAATTACCCAAGTTTGAG	NA; SL1 *trans*-spliced leader, sense primer
12	GGAATTCGGTTTTAACCCAGTTACTCA	NA; SL2 *trans*-spliced leader, sense primer

Primer sequences that are in boldface letters are linkers that encode restriction sites used for DNA cloning; they are not encoded by the *dpy-18* region.

dium to strong Dpy with mild swellings and/or constrictions), severe Lumpy Dumpy (strong Dpy with gross abnormalities in body shape), and dead eggs. Because there was no qualitative difference in phenotype between *dpy-18(e499)* and *dpy-18(e499); (dpy-18(RNAi))*, the percentage of total progeny falling into each class was compared between injected and uninjected animals, using Student's *t*-test to determine if there was a quantitative difference in phenotype (see above). Selected progeny in M-9 buffer were mounted on 2% agar pads under coverslips, examined by DIC with a Zeiss Axiophot compound microscope, and digital images were prepared as described above to document each phenotypic class.

Reverse transcription PCR: Total RNA was prepared from *him-5(e1490)* using the QIAGEN (Chatsworth, CA) RNEasy kit, treated with RQ1 RNase-free DNase (Promega Corp., Madison, WI), and repurified using the RNEasy kit. Reverse transcription (RT)-PCR was performed using Ready-to-Go You-Prime first-strand synthesis beads (Pharmacia Biotech, Piscataway, NJ) and an oligo(dT) primer (Promega). PCR on the resulting cDNA was carried out using Ready-to-Go PCR beads (Pharmacia). Primer pair 5 and 6 (Table 1) was used as a positive control to confirm the presence of *dpy-18* cDNA in the first-strand RT pools. Primer 7 was used with either primer 11 or 12 (Table 1) to test for the presence of a *trans*-spliced exon at the 5′ end of *dpy-18* mRNA.

Transgenic rescue of *dpy-18*: Two overlapping PCR fragments were amplified from Bristol N2 genomic DNA using primer pairs 1 and 2, and 3 and 4 (Table 1). Resulting 5′ and 3′ *dpy-18*-containing fragments were cloned, cut with compatible enzymes, and the inserts were ligated into Bluescript (Stratagene, La Jolla, CA) to create pKH9. pKH9 is a 7140-bp construct that contains 2008 bp of sequence 5′ to the predicted start AUG, the entire *dpy-18* gene sequence, and 198 bp 3′ to the stop codon. pKH9 was coinjected with the dominantly expressed fluorescent marker plasmid pPD118.33 [a *myo-2* pharyngeal promoter driving green fluorescent protein (GFP; Chalfie *et al.* 1994; A. Fire, S. Xu, J. Ahnn and G. Seydoux, personal communication)] into either *dpy-18(e499)* or *unc-49(e382) spe-16(eb35ts) dpy-18(e364)/qC1* [*dpy-19(e1259) glp-1(q339)*] young adult hermaphrodites following standard techniques (Fire 1986; Mello *et al.* 1991). The mean body length of GFP(+) F_1 progeny from injected animals was compared to the mean length of their GFP(−) sibs to assess rescue of *dpy-18(e499)* (see above). *dpy-18(e364)* rescue was inferred from the presence of GFP(+) Unc non-Dpy progeny in broods of injected *unc-49(e382) spe-16(eb35ts) dpy-18(e364)/qC1* [*dpy-19(e1259) glp-1(q339)*] hermaphrodites. This strain does not normally segregate Unc non-Dpy animals because the balancer *qC1* effectively suppresses recombination between *unc-49* and *dpy-18* (Edgley *et al.* 1995).

***dpy-18::gfp* expression construct:** PCR was used to amplify a 2011-bp genomic fragment from N2 DNA with primer pair 1 and 8 (Table 1). This PCR fragment, which contains 2008 bp of sequence 5′ to the predicted start AUG and has a restriction enzyme linker at each end, was restriction digested and ligated in frame to the promoterless *gfp* expression vector pPD95.69 (Chalfie *et al.* 1994; A. Fire, S. Xu, J. Ahnn and G. Seydoux, personal communication) to create pBH*dpy-18::gfp2* (Figure 2C). This construct was restriction mapped and the *dpy-18*-pPD95.69 junction was sequenced to verify the clone construction. The expression construct was coinjected with the dominant transformation marker pRF4 [*rol-6(su1006)*; Mello *et al.* 1991] by standard methods (Fire 1986; Mello *et al.* 1991), and stable transgenic lines were established before assessing the pattern of GFP expression in transgenic animals. Adult Rol animals were anesthetized and images were acquired as described above.

RESULTS

Polymorphism mapping of *dpy-18*: The *dpy-18* gene is located on the right arm of *C. elegans* chromosome III, between the cloned genes *tra-1* (to the left) and *nob-1* (to the right; Y75B8A.2; see Figure 1). This ~7-map-unit interval contains ~645 kb of DNA, based on the sequence information available from the *C. elegans* Genomic Center (Sanger Centre). Conventional genetic analyses could not be used to further refine the *dpy-18* map position because this region has few mutationally defined or cloned genes. Instead, polymorphism mapping was used to narrow the interval in which *dpy-18* must reside. This technique allows direct correlation of the physical and genetic maps by ordering sequence polymorphisms with respect to visible genetic markers in mapping crosses (Cox *et al.* 1985; Ruvkun *et al.* 1989; Jakubowski and Kornfeld 1999). The

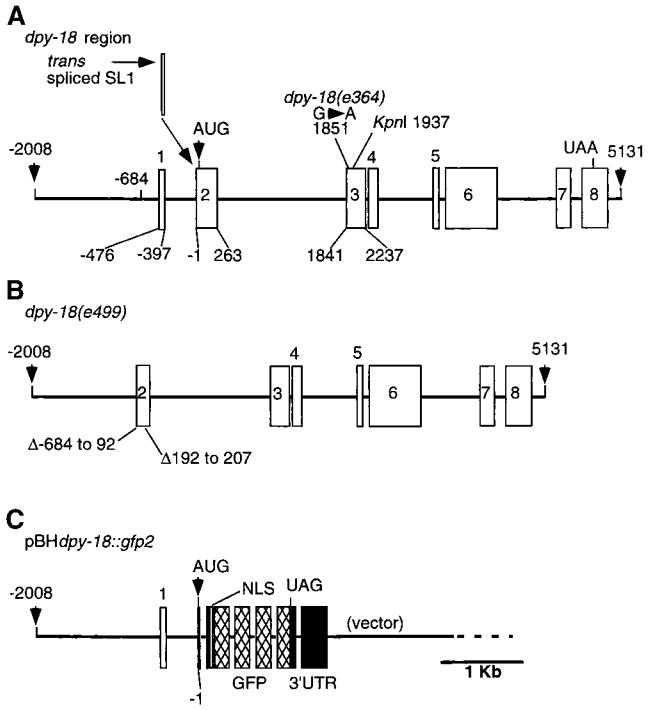

Figure 2.—*dpy-18* rescuing and expression constructs. Thin horizontal lines represent 5′ UTR, introns, and 3′ UTR; numbered boxes represent exons. (A) The 7140-bp chromosome III region cloned in pKH9 that rescues *dpy-18* mutants. The AUG indicated in exon 2 is the proposed start codon under normal conditions, and the A of this codon is numbered 1. The numbered positions of the 5′ and 3′ ends of exons 1, 2, and 3 are indicated below the line. There are at least two known variants for exon 1 and both are spliced to position −1 of exon 2. The first is *cis* spliced and has been described previously (Veijola *et al.* 1994), while the second is a *trans*-spliced SL1 exon. The positions of the *dpy-18(e364)* nonsense amber mutation in exon 3 and the UAA stop codon in exon 8 are indicated above the line. (B) The 6348-bp *dpy-18(e499)* region, using the same numbered coordinates as for wild type in A. The coordinates of the two regions deleted (Δ) by the *dpy-18(e499)* compound deletion are indicated below the line. The first deletion removes 5′ UTR sequence, the *cis* exon 1, and part of exon 2, including the proposed start AUG codon while the second deletion removes a second region of exon 2. (C) The structure of plasmid pBH*dpy-18::gfp2*. The wild-type region of *dpy-18* from −2008 through +3 was ligated to the *Pst*I site of plasmid pPD95.69, placing *dpy-18* sequence 5′ to the SV40 nuclear localization signal (NLS), followed by GFP coding sequence and an *unc-54* 3′ UTR. For additional information on plasmid construction, see materials and methods.

CB4856 *C. elegans* isolate from Hawaii (Hodgkin and Doniach 1997) was used for this purpose because it has been shown to contain a sequence polymorphism (*vs.* Bristol N2 sequence) every ∼1.5–2 kb (see http://genome.wustl.edu/gsc/CEpolymorph/snp.shtml).

The nearly complete genome sequence of *C. elegans* Bristol N2 facilitated our search for CB4856 polymorphisms within the *dpy-18* region. Consistent with prior data (see Web site in the previous paragraph), some CB4856 polymorphisms discovered in the *dpy-18* region were single-base changes (data not shown). However, several CB4856 polymorphisms, including *ebP1* and *ebP2*, changed the sizes of PCR products from those obtained with a Bristol N2 template, which allowed recombinant typing by agarose gel electrophoresis (Figure 1B). A total of 21 Vab non-Dpy recombinants, including 5 within the 0.3-map-unit *spe-16-dpy-18* interval, had the CB4856 allele of *ebP1* and did not have the *dpy-18(e364)* mutation, indicating that this polymorphism lies to the right of, or within, the *dpy-18* gene. This right-hand limit for the position of *dpy-18* is ∼270 kb to the right of *tra-1*. Many previously studied *dpy* genes affect collagen (Kramer 1997), so this genomic sequence interval was examined for genes involved in collagen biosynthesis. This interval contains a prolyl-4-hydroxylase α-subunit gene, and functional prolyl-4-hydroxylase is required for collagen assembly and stability (Mazzorana *et al.* 1993; Kivirikko and Pihlajaniemi 1998). Since the *ebP1* polymorphism was located within intron 6 of this prolyl-4-hydroxylase α-subunit gene, this gene was a plausible *dpy-18* candidate (Figure 2).

Identification of *dpy-18* as an α-subunit of prolyl-4-hydroxylase: Two overlapping genomic DNA-derived PCR products, which together contained the chromosome III prolyl-4-hydroxylase α-subunit gene and flanking sequence, were cloned to create the ∼7.1-kb insert-containing pKH9 plasmid (Figure 2A). Homozygous *dpy-18(e499)* worms that were transgenic for pKH9 had significantly longer bodies than their nontransformed Dpy sibs and were not significantly shorter than wild-type controls (Table 2); similar results were obtained with *dpy-18(e364)* (not shown; see materials and methods). The rescue of the Dpy phenotype of both *dpy-18(e499)* and *dpy-18(e364)* suggested that the chromosome III α-subunit of prolyl-4-hydroxylase is the *dpy-18* gene.

The sequences of two *dpy-18* mutant alleles were determined to confirm that this locus encodes the chromosome III α-subunit of prolyl-4-hydroxylase. Prior work showed that *dpy-18(e364)* was suppressible by amber-suppressing tRNA mutations such as *sup-5* and *sup-7* (Waterston and Brenner 1978; Waterston 1981). Since amber mutations usually alter coding sequence, PCR fragments containing the exon sequence of *dpy-18(e364)* and *dpy-18(e499)* were sequenced and compared to N2. This analysis showed that the *dpy-18(e364)* mutation is a G→A transition at position 1851 (the fourth codon of exon 3) that changes a UGG tryptophan codon to a UAG amber stop codon (Figure 2A). One PCR product that included the 5′ end of the prolyl-4-hydroxylase α-subunit gene from *dpy-18(e499)* was significantly smaller than the comparable wild-type product (data not shown). The sequence of this region revealed that *dpy-18(e499)* contains a compound mutation that is two deletions that both affect the coding sequence (Figure 2C). The first deletion removes 776 bp starting at −684, extending through the start AUG

TABLE 2

***dpy-18* body length measurements**

Genotype	*n*	Body length (μm) ± SE	$P(T \leq t)$[a] vs. N2	$P(T \leq t)$[a] vs. *dpy-18*[b]
N2	15	865 ± 14	NA[c]	NA
dpy-18(e499)/*dpy-18(e499)*	39	569 ± 17	$1.7\ e^{-17}$	NA
dpy-18(e499)/*tDf7*	38	59 ± 10	NA	0.15
dpy-18(e499); pKH9	10	834 ± 29	0.37	$7.8\ e^{-7}$
dpy-18(e364)/*dpy-18(e364)*[d]	14	756 ± 10	$2.02\ e^{-6}$	NA
dpy-18(e364)[d]/*tDf7*	15	691 ± 12	NA	$3.6\ e^{-4}$

[a] See materials and methods for calculation of *P* values.
[b] Hemizygous animals were compared to homozygotes for the same *dpy-18* allele.
[c] NA, not applicable.
[d] Full genotype of the *dpy-18(e364)* chromosome is *spe-16*(*hc54ts*) *dpy-18*(*e364*); *spe-16* has no effect on body length (data not shown).

and ending at +92. The second deletion removes 18 bp of exon 1, starting at 192 and ending at 207. These transformation rescue data and the presence of sequence alterations in two independent *dpy-18* alleles confirm that *dpy-18* encodes this α-subunit of prolyl-4-hydroxylase.

A cDNA corresponding to the *C. elegans dpy-18*-encoded α-subunit of prolyl-4-hydroxylase has been described previously (accession no. U12762; Veijola *et al.* 1994). Our genomic sequence results, and those of the *C. elegans* Genome Project (T28D6.1, accession no. Z81134, or Y47D3B.10, accession no. AL031635) differ from this previously published sequence. The first difference is the 5′ six nucleotides of the published cDNA (Veijola *et al.* 1994) that do not correspond to genomic sequence and do not appear to be due to splicing. The second region of difference occurs within exon 6, starting at position 3260 and extending to position 3307: the encoded amino acid sequence is changed from VSRRHLRLYCYYLAGPSFL (Veijola *et al.* 1994) to VSQKDISRLYCYYKRDRPFL. A cartoon illustrating the exon and intron structure of *dpy-18* is shown in Figure 2A.

Characterization of residual *dpy-18* function in mutant alleles and determination of the *dpy-18* null phenotype: Double-stranded RNA (Fire *et al.* 1998) was created from the EST clone yk247d7-derived plasmid pKH19, which contains 337 bp of the 5′ coding sequence corresponding to the chromosome III prolyl-4-hydroxylase α-subunit gene (see materials and methods). This RNA was microinjected into N2 hermaphrodites, and resulting progeny were examined for phenotypic defects relative to uninjected N2 wild type. Progeny of the injected animals showed a primarily Dpy phenotype, similar to that of *dpy-18(e364)* or *dpy-18(e499)* homozygotes (data not shown), suggesting that the *dpy-18(null)* phenotype is Dpy.

The mutations found in *dpy-18(e364)* and *dpy-18(e499)* suggest that each should be null for the encoded α-subunit of prolyl-4-hydroxylase. To determine whether the alleles *dpy-18(e364)* and *dpy-18(e499)* are null for DPY-18 function, each was placed *in trans* to the noncomplementing deficiency *tDf7* (Figure 1). The lengths of the resulting hemizygous animals were measured to assess whether the mutant *dpy-18* gene products possess any residual catalytic activity. Worms with the genotype *spe-16(hc54ts) dpy-18(e364)*/*tDf7* were significantly shorter than those homozygous for *spe-16(hc54ts) dpy-18(e364)*, suggesting that the *dpy-18(e364)* mutation does not entirely abolish DPY-18 function (Table 2). While *dpy-18(e499)* homozygotes were shorter than wild type or *dpy-18(e364)*, *dpy-18(e499)*/*tDf7* hemizygous animals were not significantly more Dpy than *dpy-18(e499)* homozygotes (Table 2), suggesting that the compound deletion *dpy-18(e499)* is null for DPY-18 activity.

To confirm that *dpy-18(e499)* is null, homozygous *dpy-18(e499)* hermaphrodites were injected with the pool of pKH19-derived dsRNA that produced a Dpy phenotype when injected into N2 (see above). Progeny of *dpy-18(e499)* homozygotes are mostly Dpy, but some show more extreme phenotypes: Lumpy Dumpy, severe Lumpy Dpy, or embryonic lethal. When the broods of *dpy-18(e499)* animals injected with *dpy-18* dsRNA were compared to the broods of uninjected sibs, the same four phenotypic classes were observed, demonstrating no qualitative difference in phenotype. To determine whether a quantitative difference in phenotype existed, the percentage of offspring falling into each phenotypic category from injected and uninjected animals was compared using Student's *t*-test. This analysis found no statistical difference in the distribution of phenotypes between the two populations ($P = 0.06$–0.6), except that there were extra dead eggs. However, the physical trauma associated with injection of foreign material into the *C. elegans* germline can cause some embryonic lethality, so the presence of dead eggs cannot be attributed solely to the *dpy-18(RNAi)*. Therefore, *dpy-18* RNAi in a *dpy-18(e499)* background did not noticeably change the phenotype of resulting progeny, and we conclude that *dpy-18(e499)* is null.

***dpy-18* expression pattern:** The *dpy-18* expression con-

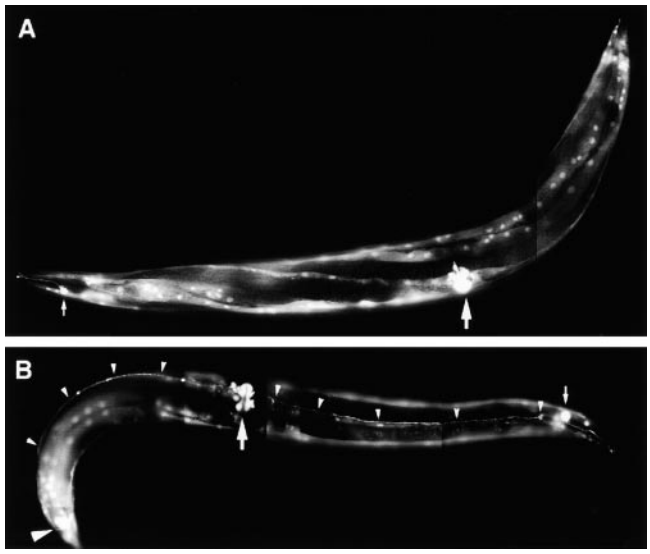

Figure 3.—(A and B) Composite images of pBH*dpy-18::gfp2* expression in the hypodermis of adult hermaphrodites. Helical twisting of the hypodermis is due to the presence of the dominant transgenic marker *rol-6 (su1006).* Large arrows indicate the labeled hypodermal cells of the vulva. Small arrows indicate head neurons. Arrowheads in B indicate a posterior neuron and its process extending anteriorly along the ventral nerve cord; large arrowhead, cell body; small arrowheads, process.

struct pBH*dpy-18::gfp2* was made by ligating 2008 bp of wild-type upstream sequence and the proposed start AUG in frame to GFP (Figure 2C). N2 animals transformed to GFP(+) after injection of this construct together with the dominant *rol-6* plasmid pRF4 (Fire 1986; Mello *et al.* 1991) were examined by DIC and fluorescence microscopy to determine where the *dpy-18* promoter is active. GFP expression was first observed at the threefold embryonic stage (not shown) and was continuously expressed thereafter throughout embryonic and postembryonic development and into adulthood. The GFP signal was observed in the hypodermis (Figure 3), two pairs of bilaterally symmetrical head neurons that extend processes to the tip of the pharynx (small arrows in Figure 3, A and B), and one posteriorly located neuron that sends a process along the ventral nerve cord to the head (arrowheads, Figure 3B).

***dpy-18* is SL1 *trans* spliced:** Comparison of the 5′ sequence of the *dpy-18* cDNA published by Veijola *et al.* (1994) with the *dpy-18* genomic sequence revealed a splicing event immediately upstream of the putative start AUG, indicating that the *dpy-18* mRNA contains at least one exon 5′ to the one containing the proposed start AUG. To investigate the splicing of *dpy-18* mRNA, we performed RT-PCR using SL1 (Table 1, primer 11) and SL2 (Table 1, primer 12) spliced leader forward primers, and a reverse primer within exon 4 (Table 1, primer 7). RT-PCR using the SL1 forward primer generated abundant product, but no product was obtained using the SL2 forward primer, demonstrating that *dpy-18* is SL1 *trans* spliced (data not shown). Sequencing of the SL1 RT-PCR product revealed that the SL1 spliced leader sequence is spliced to the *dpy-18* transcript 1 bp 5′ to the proposed start AUG codon (data not shown). This transcript does not contain the upstream sequence present in the published cDNA (Veijola *et al.* 1994), suggesting that alternative splicing may occur.

DISCUSSION

Two lines of evidence indicate that *dpy-18* encodes the chromosome III-encoded α-subunit of prolyl-4-hydroxylase in *C. elegans.* First, a 7140-bp genomic construct containing the α-subunit coding region and 5′ and 3′ flanking sequences restores a wild-type body length to two *dpy-18* mutants (Figure 2, Table 2, and materials and methods). Consequently, this sequence encodes *dpy-18* and contains sufficient promoter sequence to allow *dpy-18* expression. Second, we sequenced the chromosome III prolyl-4-hydroxylase α-subunit coding region of *dpy-18(e499)* and *dpy-18(e364)* mutants and found that each has a mutation in this gene (Figure 2).

The *dpy-18* gene has eight exons, and several of its introns are large by *C. elegans* standards (Figure 2). In particular, the second intron is >1 kb in length, suggesting that regulatory sequences might be present within this intron (*e.g.*, Okkema *et al.* 1993). While pBH*dpy-18::gfp2*, which contains only 2008 bp of upstream sequence and the putative start AUG codon, showed robust expression (Figure 3), other *dpy-18::gfp* fusions that included intron 2 (and other *dpy-18* sequences) were expressed inconsistently and at low levels (data not shown). This suggests that *dpy-18* intron 2 might contain elements that negatively regulate expression from the *dpy-18* promoter. Despite inconsistent expression, several intron 2-containing *dpy-18::gfp* fusions caused transgenic animals to occasionally produce dead embryos and Dpy progeny. This suggests that these *dpy-18::gfp* fusion constructs may have caused an RNAi effect, as has been seen for other genes (*e.g.*, Fire *et al.* 1990).

Previous work showed that *dpy-18(e364)* is an amber nonsense mutation, because it is suppressible by amber-suppressing tRNAs (Waterston and Brenner 1978; Waterston 1981). While amber mutations are frequently null, *dpy-18(e364)/tDf7* are more Dpy than *dpy-18(e364)* homozygotes, showing that this allele is hypomorphic (Table 2). The hypomorphic nature of *dpy-18(e364)* indicates that some catalytically active protein is synthesized; this is likely due to occasional readthrough of the stop codon, which has also been seen, for instance, in the nonnull *spe-9(eb23)* ochre nonsense mutant (Singson *et al.* 1998).

The *dpy-18(e499)* mutant contains two deletions that remove upstream, exon 1 and exon 2 sequence, including the putative start AUG codon and the putative ER signal sequence (Figure 2B; see Veijola *et al.* 1994 for discussion of ER signal sequence). The absence of the

start AUG should abolish translation or cause an N-terminal deletion in the resulting protein. Unlike *dpy-18(e364)*, the mean body length of *dpy-18(e499)*/*dpy-18(e499)* is the same as that of *dpy-18(e499)*/*tDf7*. This result implies that *dpy-18(e499)* is null, and is confirmed by *dpy-18(RNAi)* results.

Double-stranded RNAi is well-established in *C. elegans* and other organisms as a way to block the function of the gene that encodes that RNA and mimic its null phenotype (Fire *et al.* 1998; reviewed by Montgomery and Fire 1998; Fire 1999; Sharp 1999). We found that *dpy-18(RNAi)* produced a Dpy phenotype in N2 animals, and had no obvious effect on *dpy-18(e499)* animals, demonstrating that the null phenotype for *dpy-18* is Dpy, and *dpy-18(e499)* is null. However, *dpy-18(e499)* homozygotes sometimes arrest as embryos or have a severely disrupted body morphology. RNAi experiments performed using dsRNA with homology to both *dpy-18* and the second prolyl-4-hydroxylase α-subunit gene found on chromosome IV (F35G2.4) produced extensive embryonic lethality (not shown; see materials and methods), demonstrating the essential nature of prolyl-4-hydroxylase function. These RNAi-arrested embryos show phenotypes that are very similar to the occasional *dpy-18(e499)* dead embryo. This suggests that prolyl-4-hydroxylase activity is essential, and that these genes are partially redundant with respect to function during embryogenesis. Activity of F35G2.4 is usually sufficient to allow survival of *dpy-18* embryos, but both F35G2.4 and *dpy-18* must function if the embryonic lethality associated with insufficient prolyl hydroxylase α-subunit is to be reliably avoided.

Expression of a *dpy-18::gfp* fusion construct in the hypodermis is consistent with cuticle collagen expression in that tissue (Kramer 1994a,b). However, type IV collagens are produced primarily in muscle (Graham *et al.* 1997), and no *dpy-18::gfp* expression was seen in muscle at any time during the life cycle. One possible explanation is that the pBH*dpy-18::gfp2* fusion construct lacks regulatory elements required for expression in muscle cells, or expression in muscle might be below the level required for detection of GFP fluorescence. Alternatively, the F35G2.4 prolyl-4-hydroxylase α-subunit (discussed above) might be the major form employed within muscle cells. Although the *dpy-18* promoter fusion is expressed in head and posterior neurons, no neural defects are obvious in *dpy-18* mutants, so the significance of this expression pattern is unclear. Perhaps *dpy-18* mutants have subtle behavioral defects that would be revealed by specific assays (*e.g.*, Bargmann and Mori 1997).

There are at least two different transcripts expressed by the *dpy-18* gene and both contain sequence 5′ to the proposed start AUG. The previously published cDNA sequence contains a 79-bp noncoding exon that is spliced one nucleotide 5′ to the proposed initiating AUG codon (Veijola *et al.* 1994). Our experiments indicate that SL1 *trans* splicing to this same splice acceptor can also occur. Therefore, while both cDNA sequences suggest that the same AUG is utilized to initiate translation, they do not agree with regard to the 5′ end of the *dpy-18* transcript. The potential roles of these sequences 5′ to the proposed start AUG and whether they are a consequence of tissue- or stage-specific regulation remain to be established.

Collagen mutations have been studied extensively in *C. elegans*, and they are associated frequently with either altered body morphology or lethality (reviewed by Kramer 1994a,b). Mutation of cuticle collagen genes leads to a Dumpy body in the case of *col-2* or *dpy-13*, a Long body in the case of *sqt-1(sc101)*, helical twisting of the long body axis in *sqt-1* or *rol-6*, and occasional larval lethality in the case of restrictively raised *sqt-3(e2117ts)* (Kramer *et al.* 1982, 1988, 1990; von Mende *et al.* 1988; van der Keyl *et al.* 1994). Mutations in the *emb-9*- or *let-2*-encoded type IV collagens, which are the principal components of basement membranes, cause failure of elongation, hypodermal rupture, and embryonic lethality (Guo *et al.* 1991). Prior to the work described in this article, none of the enzymes involved in the post-translational modification of collagen had been mutationally defined in *C. elegans*.

Human syndromes caused by disruption of collagen biosynthesis have also been extensively studied, and they cause phenotypes ranging from mild skin abnormalities and joint hyperflexibility to pre- or perinatal lethality (reviewed by Prockop 1992; Sauberlich 1994). These disorders are most frequently caused by mutations in genes that encode collagen polypeptides, leading to abnormal, reduced, or absent collagen helix formation. An exception is Ehlers-Danlos syndrome type VI, which is usually due to impaired lysyl hydroxylase activity (reviewed by Burrows 1999). Prolyl-4-hydroxylase has not been implicated in any inherited human disorder, but it is known that proper function of this enzyme is required in humans. For instance, insufficient dietary intake of the essential cofactor ascorbate (vitamin C) reduces prolyl-4-hydroxylase activity and causes the disease scurvy. Ascorbate is required for formation of hydroxyproline and, in its absence, procollagen remains bound to prolyl-4-hydroxylase and is retained within the ER (Walmsley *et al.* 1999), leading to deterioration of the ECM; chronic cases of scurvy can be lethal (reviewed by Sauberlich 1994).

The *C. elegans dpy-18* gene is the first mutationally defined prolyl-4-hydroxylase α-subunit described in any animal. Because of its unique ability to be mutated without causing lethality, *dpy-18* may provide a powerful, natural *in vivo* system for analysis of prolyl-4-hydroxylase function and regulation, which may lead to novel treatments for collagen disorders in humans. In addition, the lethality in *C. elegans* associated with absence of prolyl-4-hydroxylase function may provide a target for

pharmaceutical control of nematodes through disruption of cuticle formation.

We thank Alan Coulson, Andy Fire, Yuji Kohara, Tim Schedl, Sam Ward, and Heinke and Ralf Schnabel for providing *C. elegans* strains, DNA clones, or PCR primers. We thank Tim Schedl for suggesting the polymorphism mapping procedure utilized in this article and Shawn Ahmed for useful discussions. We also thank two anonymous reviewers for numerous suggestions that resulted in a much-improved manuscript. The *Caenorhabditis* Genetic Center provided some nematode strains, and it is funded by the National Institutes of Health National Center for Research Resources. This work was supported by U.S. Public Health Service grant GM RO1 GM40697 and National Science Foundation grant IBN-9631102 to S.W.L.

LITERATURE CITED

Bargmann, C. I., and I. Mori, 1997 Chemotaxis and thermotaxis, pp. 717–737 in *C. elegans II*, edited by D. Riddle, T. Blumenthal, B. J. Meyer and J. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Barstead, R., and R. H. Waterston, 1991 Vinculin is essential for muscle function in the nematode. J. Cell Biol. **114:** 715–724.

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

Brodsky, B., and J. A. Ramshaw, 1997 The collagen triple-helix structure. Matrix Biol. **15:** 545–554.

Burrows, N. P., 1999 The molecular genetics of the Ehlers-Danlos syndrome. Exp. Dermatol. **24:** 99–106.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, 1994 Green fluorescent protein as a marker for gene expression. Science **263:** 802–805.

Cox, G. N., and D. Hirsh, 1985 Stage-specific patterns of collagen gene expression during development of *C. elegans.* Mol. Cell. Biol. **5:** 363–372.

Cox, G. N., M. Kusch, K. Denevi and R. S. Edgar, 1981a Temporal regulation of cuticle synthesis during development of *Caenorhabditis elegans.* Dev. Biol. **84:** 277–285.

Cox, G. N., S. Strapans and R. S. Edgar, 1981b The cuticle of *Caenorhabditis elegans* II. Stage-specific changes in ultrastructure and protein composition during postembryonic development. Dev. Biol. **86:** 456–470.

Cox, G. N., J. M. Kramer and D. Hirsh, 1984 Number and organization of collagen genes in *C. elegans.* Mol. Cell. Biol. **4:** 2389–2395.

Cox, G. N., S. Carr, J. M. Kramer and D. Hirsh, 1985 Genetic mapping of *C. elegans* collagen genes using DNA polymorphisms as phenotypic markers. Genetics **109:** 513–528.

Edgley, M., D. L. Baillie, D. L. Riddle and A. M. Rose, 1995 Genetic balancers, pp. 148–184 in *Caenorhabditis elegans: Modern Biological Analysis of an Organism*, edited by H. F. Epstein and D. C. Shakes. Academic Press, San Diego.

Fire, A., 1986 Integrative transformation of *Caenorhabditis elegans.* EMBO J. **5:** 2673–2680.

Fire, A., 1999 RNA-triggered gene silencing. Trends Genet. **15:** 358–363.

Fire, A., S. W. Harrison and D. Dixon, 1990 A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans.* Gene **93:** 189–198.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver *et al.*, 1998 Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans.* Nature **391:** 806–811.

Gordon, M. K., and B. R. Olsen, 1990 The contribution of collagenous proteins to tissue-specific matrix assemblies. Curr. Opin. Cell Biol. **2:** 833–838.

Graham, P. L., J. J. Johnson, S. Wang, M. H. Sibley, M. C. Gupta *et al.*, 1997 Type IV collagen is detectable in most, but not all, basement membranes of *Caenorhabditis elegans* and assembles on tissues that do not express it. J. Cell Biol. **137:** 1171–1183.

Guo, X., and J. M. Kramer, 1989 The two *Caenorhabditis elegans* basement membrane (type IV) collagen genes are located on separate chromosomes. J. Biol. Chem. **264:** 17574–17582.

Guo, X., J. J. Johnson and J. M. Kramer, 1991 Embryonic lethality caused by mutations in basement membrane collagen of *C. elegans.* Nature **349:** 707–709.

Hodgkin, J., 1983 Male phenotypes and mating efficiency in *C. elegans.* Genetics **103:** 43–64.

Hodgkin, J., 1993 Molecular cloning and duplication of the nematode sex-determining gene *tra-1.* Genetics **133:** 543–560.

Hodgkin, J., 1997 Appendix 1, Genetics, pp. 881–1047 in *C. elegans II*, edited by D. Riddle, T. Blumenthal, B. J. Meyer and J. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hodgkin, J., and T. Doniach, 1997 Natural variation and copulatory plug formation in *Caenorhabditis elegans.* Genetics **146:** 149–164.

Hodgkin, J., H. R. Horvitz and S. Brenner, 1979 Nondisjunction mutants of the nematode *C. elegans.* Genetics **91:** 67–94.

Horvitz, H. R., S. Brenner, J. Hodgkin and R. K. Herman, 1979 A uniform genetic nomenclature for the nematode *C. elegans.* Mol. Gen. Genet. **175:** 129–133.

Jakubowski, J., and K. Kornfeld, 1999 A local, high-density, single nucleotide polymorphism map used to clone *Caenorhabditis cdf-1.* Genetics **153:** 743–752.

Kivirikko, K., and T. Pihlajaniemi, 1998 Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. Adv. Enzyme Mol. Biol. **72:** 325–398.

Kivirikko, K., R. Myllyla and T. Pihlajaniemi, 1992 Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins, pp. 1–51 in *Post-Translational Modification of Proteins*, edited by J. J. Harding and M. J. C. Crabbe. CRC Press, Boca Raton, FL.

Koivu, J., R. Myllyla, T. Helaakoski, T. Pihlajaniemi, K. Tasanen *et al.*, 1987 A single polypeptide acts both as the β subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase. J. Biol. Chem. **262:** 6447–6449.

Kramer, J. M., 1994a Structures and functions of collagens in *Caenorhabditis elegans.* FASEB J. **8:** 329–336.

Kramer, J. M., 1994b Genetic analysis of extracellular matrix in *C. elegans.* Annu. Rev. Genet. **28:** 95–116.

Kramer, J. M., 1997 Extracellular matrix, pp. 471–500 in *C. elegans II*, edited by D. Riddle, T. Blumenthal, B. J. Meyer and J. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Kramer, J. M., G. N. Cox and D. Hirsh, 1982 Comparisons of the complete sequences of two collagen genes from *Caenorhabditis elegans.* Cell **30:** 599–606.

Kramer, J. M., G. N. Cox and D. Hirsh, 1985 Expression of the *C. elegans* collagen genes *col-1* and *col-2* is developmentally regulated. J. Biol. Chem. **260:** 1945–1951.

Kramer, J. M., J. J. Johnson, R. S. Edgar, C. Basch and S. Roberts, 1988 The *sqt-1* gene of *C. elegans* encodes a collagen critical for organismal morphogenesis. Cell **55:** 555–565.

Kramer, J. M., R. P. French, E.-C. Park and J. J. Johnson, 1990 The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. Mol. Cell. Biol. **10:** 2081–2089.

Mazzorana, M., H. Gruffat, A. Sergeant and M. van der Rest, 1993 Mechanisms of collagen trimer formation. Construction and expression of a recombinant minigene in HeLa cells reveals a direct effect of prolyl hydroxylation on chain assembly of type XII collagen. J. Biol. Chem. **268:** 3029–3032.

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

Montgomery, M. K., and A. Fire, 1998 Double-stranded RNA as a mediator in sequence-specific genetic silencing and co-suppression. Trends Genet. **14:** 255–258.

Okkema, P. G., S. W. Harrison, V. Plunger, A. Aryana and A. Fire, 1993 Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans.* Genetics **135:** 385–404.

Pihlajaniemi, T., T. Helaakoski, K. Tasanen, R. Myllyla, M. L. Huhtala *et al.*, 1987 Molecular cloning of the β subunit of human prolyl 4-hydroxylase. This subunit and protein disulfide isomerase are products of the same gene. EMBO J. **6:** 643–649.

Politz, S., and R. S. Edgar, 1984 Overlapping stage-specific sets of numerous small collagenous polypeptides are translated *in vitro* from *Caenorhabditis elegans* RNA. Cell **37:** 853–860.

Prockop, D. J., 1992 Mutations in collagen genes as a cause of connective-tissue diseases. N. Engl. J. Med. **326:** 540–546.

Ruvkun, G., V. Ambros, A. Coulson, R. Waterston, J. Sulston *et al.*, 1989 Molecular genetics of the *Caenorhabditis elegans* heterochronic gene *lin-14*. Genetics **121:** 501–516.

Sauberlich, H. E., 1994 Pharmacology of vitamin C. Annu. Rev. Nutr. **14:** 371–391.

Shakes, D. C., 1988 A genetic and pharmacological analysis of spermatogenesis in the nematode *Caenorhabditis elegans*. Ph.D. Thesis, Johns Hopkins University, Baltimore.

Sharp, P. A., 1999 RNAi and double-strand RNA. Genes Dev. **13:** 139–141.

Shaw, L. M., and B. R. Olsen, 1991 FACIT collagens: diverse molecular bridges in extracellular matrices. Trends Biochem. Sci. **16:** 191–194.

Singson, A., K. B. Mercer and S. W. L'Hernault, 1998 The *C. elegans spe-9* gene encodes a sperm transmembrane protein that contains EGF-like repeats and is required for fertilization. Cell **93:** 71–79.

van der Keyl, H., H. Kim, R. Espey, C. V. Oke and M. K. Edwards, 1994 *Caenorhabditis elegans sqt-3* mutants have mutations in the *col-1* collagen gene. Dev. Dyn. **201:** 86–94.

van der Rest, M., and R. Garrone, 1991 Collagen family of proteins. FASEB J. **5:** 2814–2823.

Veijola, J., P. Koivunen, P. Annunen, T. Pihlajaniemi and K. I. Kivirikko, 1994 Cloning, baculovirus expression and characterization of the α subunit of prolyl 4-hydroxylase from the nematode *Caenorhabditis elegans*. J. Biol. Chem. **269:** 26746–26753.

Veijola, J., P. Annunen, P. Koivunen, A. P. Page, T. Pihlajaniemi *et al.*, 1996 Baculovirus expression of two protein disulfide isomerase isoforms from *Caenorhabditis elegans* and characterization of prolyl 4-hydroxylases containing one of these polypeptides as their β subunit. Biochem. J. **317:** 721–729.

von Mende, N., D. M. Bird, P. S. Albert and D. L. Riddle, 1988 *dpy-13*: a nematode collagen gene that affects body shape. Cell **55:** 567–576.

Walmsley, A. R., M. R. Batten, U. Lad and N. J. Bulleid, 1999 Intracellular retention of procollagen within the endoplasmic reticulum is mediated by prolyl 4-hydroxylase. J. Biol. Chem. **274:** 14884–14892.

Ward, S., Y. Argon and G. A. Nelson, 1981 Sperm morphogenesis in wild-type and fertilization-defective mutants of *C. elegans*. J. Cell Biol. **91:** 26–44.

Waterston, R. H., 1981 A second informational suppressor, *sup-7 X*, in *Caenorhabditis elegans*. Genetics **97:** 307–325.

Waterston, R. H., and S. Brenner, 1978 A suppressor mutation in the nematode acting on specific alleles of many genes. Nature **275:** 715–719.

Communicating editor: R. K. Herman