Cell. Mol. Life Sci. 63 (2006) 1193–1204
1420-682X/06/101193-12
DOI 10.1007/s00018-006-6012-z


Cellular and Molecular Life Sciences

Research Article

# *Caenorhabditis elegans dpy-5* is a cuticle procollagen processed by a proprotein convertase

**C. Thacker[a], J. A. Sheps[b] and A. M. Rose[a, *]**

[a] Department of Medical Genetics, University of British Columbia, 419-2125 East Mall, Vancouver, B. C. V6T 1Z4 (Canada), Fax: +1 604 822 5348, e-mail: arose@gene.nce.ubc.ca
[b] BC Cancer Research Centre, British Columbia Cancer Agency, Vancouver, B. C. V5Z 1L3 (Canada)

Received 13 January 2006; accepted 23 March 2006
Online First 2 May 2006

**Abstract.** Genetic analysis of the nematode *Caenorhabditis elegans* reveals that all *dpy-5* alleles are dominant suppressors of bli-4 blistering. Molecular cloning of *dpy-5* establishes that it encodes a cuticle procollagen, defects in which are responsible for the short-body, dumpy phenotype. The null mutation, *e907* removes the entire coding region, whereas the *dpy-5* reference allele, *e61,* contains a nonsense substitution. RT-PCR analysis and a *dpy-5::gfp* fusion show that *dpy-5* is expressed only in hypodermal cells at all post-embryonic life-cycle stages. Variable expression of *dpy-5* in V lineage-derived seam cells suggests an alternative regulatory mechanism in these cells. The *dpy-5* gene product contains an Arg-X-X-Arg cleavage motif that could be recognized by a proprotein convertase, such as BLI-4. Mutation of this site cause a dominant dumpy phenotype suggesting Dpy-5 procollagen requires processing for normal cuticle production.

**Keywords.** Procollagen, *C. elegans*, suppressor, blisters, proprotein convertase.

## Introduction

The exoskeleton or cuticle of the nematode *Caenorhabditis elegans* is a complex extracellular structure, required for protection, post-embryonic morphology and motility, synthesized at the end of embryogenesis and at each molt. The composition and structure of the L1, dauer, and adult cuticles differ markedly from the L2, L3, and L4 cuticles [1, 2].

The major proteins of the cuticle are small collagens encoded by a family of at least 158 members [3]. During synthesis of a new cuticle, collagens are expressed at discrete times, possibly regulated by heterochronic genes, such as *lin-29* [4], which partly explains the differing cuticles of each developmental stage [5, 6]. Collagens are structural proteins required for a variety of extracellular matrices that form triple helices, either of identical collagen chains (homotrimeric) or a mix of non-identical collagen chains (heterotrimeric). Collagens typically have a repetitive sequence with glycine occurring every third residue, a GXY repeat, with proline or 4-hydroxyproline frequently in the X and Y positions, respectively. The repeats appear involved in the formation and stability of the triple helix. Cuticular collagens in *C. elegans* are composed of three domains, with a GXY repeat region as the middle one of the three. The N-terminal domain contains a predicted signal peptide for secretion. Sequences near the GXY domain are reminiscent of the Arg-X-X-Arg (RXXR) cleavage motif recognized by kex2/subtilisin-like proprotein convertases (PCs) [7, 8]. The GXY repeats are often interrupted by a stretch of residues that may include two cysteines. Studies on mammalian collagens suggest that the non-GXY C terminus is required for initiating triple helix formation, by bringing collagen chains into close proximity and alignment [9], a process called nucleation. Triple helix

* Corresponding author.

formation then takes place by a 'zipper-like' mechanism in the C- to N-terminal direction. The spacing of cysteine residues is conserved between cuticular collagens, suggesting that these aid interactions between collagen molecules by formation of disulfide bridges. This conserved feature provides a convenient means to categorize cuticle collagens into sub-families [3, 6].

Mutations that affect the morphology of *C. elegans* have been identified, most of which result from defects in the cuticle. These are classified into five phenotypic groups; (1) dumpy (Dpy), a shortening of general morphology, normally affecting all developmental stages; (2) blister (Bli), identified by fluid-filled separations of the adult cuticle; (3) roller (Rol), in which the exoskeleton, underlying ectoderm, and musculature is helically twisted along the length of the animal; (4) long (Lon) in which the animal is longer and thinner than wild type, and (5) squat (Sqt), which represents mutations which, when heterozygous are Rol, yet Dpy when homozygous. Examples of each phenotype have been correlated with lesions in cuticle collagen genes, including *dpy-2* [10], *dpy-7* [11], *dpy-10* [10], *dpy-13* [12], *bli-1, bli-2* (J. Kramer, unpublished results), *lon-3* [13], *rol-6* [10], *sqt-1* [14] and *sqt-3* [15]. Most represent substitutions of glycine residues in the GXY repeats. For the most part, these act as null mutations: they inhibit triple helix formation resulting in degradation of abnormal complexes [9]. Normally, the associated decrease in secreted mutant collagen is reflected in a recessive phenotype. However, certain glycine substitutions as well as mutations in the PC-like cleavage motif may result in a severe dominant phenotype. It is easy to speculate that abnormal collagen chains interfere with the function or formation of the triple helical structure. Consistent with this, genetic interactions between collagens have been documented by enhancement or suppression of morphological phenotypes.

*Dpy-5* mutations cause a recessive Dpy phenotype, which becomes progressively more severe. *dpy-5*, *dpy-6* and *dpy-13* act as dominant suppressors of the adult-specific Bli phenotype of *bli-4* [15]. Work on the *bli-4/kpc-4* gene shows it to encode a member of the PC family [16]. Characterization of the blistered *e937* allele, and 13 lethal mutations that arrest in late embryonic development, suggests *bli-4* is required for processing cuticle procollagens [16, 17]. Evidence that cuticle procollagens are processed at RXXR motifs, is provided by *sqt-1* and *rol-6*, where replacements of conserved arginine residues within the PC cleavage site result in dominant phenotypes [10, 18, 19], which can be explained as resulting from insertion of unprocessed procollagens into the cuticle. A similar dominant mutation occurs in *dpy-10* [20]. This study examined how *dpy-5* acts as a dominant suppressor of *bli-4,* and describes the phenotypic range of suppression, including its gender specificity. We describe several lesions in the *dpy-5* gene, which encodes a predicted cuticle procollagen, and provide evidence that DPY-5 proprotein is processed at an RXXR motif.

## Materials and methods

**Nematode strains.** Maintenance and handling of *C. elegans* strains were as described by Brenner [21]. All strains were maintained at 20 °C. The following strains carrying *dpy-5* alleles were used: CB61 *e61,* CB565 *e565,* CB907 *e907,* BC152 *s102,* BC197 *s111* and KR3571 *hEx189 dpy-5(h1964),* KR3809 *hIs19 (dpy-5(h1964),* KR3843 *bli-4(e937) unc-13(e450)*I; *hIs19 dpy-5(h1964)*IV, KR3830 *hIs21* (*dpy-5::gfp*), KR3831 *hIs22* (*dpy-5::gfp*), KR1025 *dpy-5(e61) bli-4(e937),* and KR912 *dpy-5(e61) bli-4(e937) unc-13(e450).* Additional *bli-4* strains were: CB937 *bli-4(e937),* and KR1123 *bli-4(e937) unc-13(e450).*

Strains used for mapping were: CB950 *unc-75(e950)* I, CB120 *unc-4(e120)* II, CB189 *unc-32(e189)* III, CB245 *unc-17(e245)* IV, CB270 *unc-42(e270)* V, and CB55 *unc-2(e55)* X.

**Characterization of Dpy-5 mutant phenotypes.** Viable brood sizes were measured by transferring hermaphrodites individually to fresh plates, daily for 4 days, starting at the L4 stage. Viable progeny were counted 3–4 days later. The development of *dpy-5* homozygotes was measured from hatched L1 larvae to the adult molt. Animals were synchronized by treatment of gravid adults with hypochlorite. Embryos were collected and maintained in M9 buffer without feeding for 24 h to allow embryos to hatch. L1 larvae were placed on NGM plates seeded with *E. coli*. Morphology and size of adult *dpy-5* mutants and wild type worms were measured using a calibrated dissecting scope.

**Determination of *bli-4* penetrance.** The *dpy-5* allele *e61* is a dominant suppressor of blistering in heteroallelic combination with *bli-4(e937*) [16]. Suppression of blistering was measured by crossing *bli-4(e937)* homozygous L4 males with L4 homozygous *dpy-5(x) bli-4(e937) unc-13(e450)* hermaphrodites, where x indicates the alleles *e61, e565, e907, s102,* or *s111.* Penetrance was determined by scoring non-Dpy non-Unc F1 adult hermaphrodites and males.

**Rescue of the Dpy-5 mutant phenotype.** A 3.3-kb *Nco* I fragment containing a predicted cuticle collagen gene was isolated from the cosmid F27C1, and cloned into the *Nco* I site of pGEM-5 to generate the clone pCeh361. This plasmid was injected into *dpy-5* mutant adult hermaphrodites at 1 µg/ml, 20 µg/ml and 100 µg/ml as described [22]. Total injected DNA concentration was brought to 100 µg/ml by adding pGEM-5. Successful rescue was scored by the presence of wild type F1 progeny, which

gave rise to stable transgenic lines. pCeh361 rescued the *dpy-5* phenotype at all concentrations, although lines segregating extrachromosomal arrays were more easily established from animals injected with more rescuing plasmid. The presence of pCeh361 in transformed worms was verified by PCR.

**Sequencing of *dpy-5* mutations.** Genomic DNA was amplified from single *dpy-5* homozygotes using primers KRp249 and KRp250 as described previously [17]. These primers generate products that encompass the entire *dpy-5* coding region: KRp249 anneals 918 bp upstream of the predicted start codon, while KRp250 anneals 1362 bp downstream of the stop codon. These primers, on wild type genomic DNA, *e61*, *s102*, and *s111,* yielded the expected product of 3135 bp. Amplification from *e565* and *e907* yielded products of approximately 2700 bp and 2100 bp, respectively, which suggests they contain deletions. Amplification products were resolved by agarose gel electrophoresis, purified using Qiaquick Spin Columns (Qiagen) and directly sequenced. Sequence was obtained from both strands using gene-specific oligonucleotide primers.

**RT-PCR analysis.** Total RNA (1 µg) from each developmental stage was converted to first-strand cDNA using Superscript reverse transcriptase according to the manufacturer's conditions (Gibco Life Sciences). From each reaction mix, 1 µl cDNA template was amplified by PCR with an SL-1 leader-specific primer (KRp 60 5′-ATAAGAATGCGGCCGCGGTTTAATTACCCAAGTTTG-3′) and a dpy-5 antisense primer, KRp156 (5′-GTCCGAAGATTCCACGAACG-3′). The expected product was 277 bp. Primers KRp12 and KRp14, which amplify a 509 bp product specific for S-adenosyl homocysteine hydrolase (AHH) [18], were also included as a control for PCR amplification. AHH is expressed during all developmental stages (C. Thacker, unpublished results).

**GFP reporter gene construction and generation of transgenic animals.** The tissue specificity of *dpy-5* was determined using a translational fusion of *dpy-5* with green fluorescent protein (GFP). The *dpy-5::gfp* reporter construct pCeh358 was generated by insertion of a 750 bp *Sph* I fragment from pCeh361 into the *Sph* I site of the *gfp* expression vector pPD95.69 (kindly provided by A. Fire). This fragment contains 5′ sequences from an *Sph* I site in the polylinker of pCeh361 to a site 30 bp downstream from the predicted DPY-5 initiator methionine, resulting in an in-frame fusion of the first 12 codons of *dpy-5* with *gfp*. Transgenic animals were generated by microinjection of pCeh358 (5 ng/µl) and pBluescript KS (100 ng/µl), or in combination with 50 ng/µl pCes1943, which carries a dominant *rol-6* mutation [*rol-6(su1006)*] used as a morphological marker for successful transformation. Arrays were generated with and without *rol-6* to avoid any possible interference from the *rol-6* collagen gene. Integration of extrachromosomal arrays was by irradiation of L4 transgenic larvae using a γ-ray source (3800 rad) as described [22]. F3 animals were isolated, which gave 100% transmission of the roller phenotype and/or GFP expression. Worms used to study the expression of *dpy-5::gfp* carried the *gfp*-marked gene either as an extrachromosomal array, as in *hEx212* (pCeh358 and pBluescript), and *hEx214* (pCeh358, pBluescript and pCes1943) or as an integrated copy, in *hIs20* and *hIs2*. Multiple animals from all transgenic lines were examined and each exhibited similar GFP expression patterns.

**Site-directed mutagenesis of the putative proprotein convertase cleavage site.** Complementary oligonucleotides encompassing the nucleotide sequence of the cleavage motif were synthesized, incorporating a missense mutation that substitutes Arg-80 with a serine residue. The sequences of the oligonucleotides were: KRp281 5′-CGGACGTCACAAG**A**GTTCCAACAGCC-3′ and KRp282 5′-GGCTGTTGGAAC**T**CT-TGTGACGTCCG-3′, (the nucleotide substitution is indicated in bold). Site-directed mutagenesis of a 196 bp *Sph* I–*Aat* II subclone, which includes the cleavage motif, was performed using the QuickChange kit, according to the manufacturers conditions (Stratagene). The fragment was sequenced after mutagenesis to confirm that the only nucleotide change was the one desired. The *Sph* I–*Aat* II fragment was then reinserted into pCeh361 using standard techniques to generate plasmid pCeh384. This construct was injected into wild type and *dpy-5(e61)* adults as described above. Extrachromosomal arrays of pCeh384 were integrated into the genome of wild type worms by γ-irradiation as described. Three integrated lines were isolated and out-crossed six times with wild type males and all three showed identical characteristics. One line, KR3809, carrying the integrated array *hIs19* was selected for subsequent analyses. Chromosomal sites of integration for arrays were determined by establishing linkage with selected uncoordinated (Unc) mutants. For each chromosomal *unc* marker, heterozygous males were mated with hermaphrodites homozygous for the array. We then looked for the presence or absence of Dpy Unc progeny in the F2 generation. The absence of such progeny establishes linkage of the array to that particular chromosome. The markers used included *unc-75(e950) I*, *unc-4(e120) II*, *unc-32(e189) III*, *unc-17(e245) IV*, *unc-42(e270) V*, and *unc-2(e55) X*.

**Genetic analysis of the dominant *dpy-5 h1964* mutation.** The effect of the *h1964* mutation on body length, brood size, development and suppression of blistering was examined. To measure suppression of blistering, we constructed a strain with the genotype *bli-4(e937) unc-13(e450)*; *hIs19*. *bli-4(e937) unc-13(e450)* hermaphrodites were mated

with *hIs19* homozygous males, and F1 heterozygous hermaphrodites allowed to self-fertilize. Dpy, Unc hermaphrodites were selected from the F2 progeny and plated individually, and a line that did not segregate non-Dpy animals (signifying that *hIs19* was homozygous) established. Since the Dpy, Unc animals did not blister, the presence of the *bli-4(e937)* mutation was verified by PCR using the primers KRp70 (5′-AGTTCTCTCACGCGTTCATC-3′) and KRp248 (5′-TAACTCTTACCCTACTCCTC-3′) which amplify an 850 bp product specific for the deletion. To examine whether the integrated transgene, *h1964,* could dominantly suppress the Bli phenotype, we mated *bli-4(e937) unc-13(e450)*; *hIs19* hermaphrodites with homozygous *bli-4(e937)* males. In this way, all out-crossed F1 progeny would be homozygous for the recessive *bli-4* mutation, yet contain a single copy of *hIs19*. Penetrance of blistering was measured as a percentage of Bli, Dpy, non-Unc male and hermaphrodite animals in the F1 generation.

## Results

**Differences in size and growth curves for *dpy-5* alleles.** *dpy-5* results in a Dpy phenotype, which appears in L2 and continues into the adult stage. We examined body sizes of the *dpy-5* alleles, *e61, e565, e907, s102* and *s111*. All adults showed a similar reduction in size (Table 1) except *s111,* which exhibited an intermediate size compared with the reference allele *e61* and wild type (Fig. 1). The effect of *dpy-5* on development and reproduction were also measured (Table 1). Brood sizes were reduced compared with wild type, and development from L1 stage to the adult molt was retarded in all *dpy-5* mutants by as much as 12 h. We and others [23] have observed that the *e61* allele exhibits a weak dominant phenotype, such that *e61*/+ animals are slightly Dpy. This effect, which was often difficult to score, was observed for all other *dpy-5* heterozygous animals, to varying degrees. However, *e61*/+, *s102*/+ and *s111*/+ individuals consistently displayed a weak Dpy phenotype. The behavior of these three alleles suggests that they are antimorphs; they produce an aberrant gene product that interferes with the normal product.

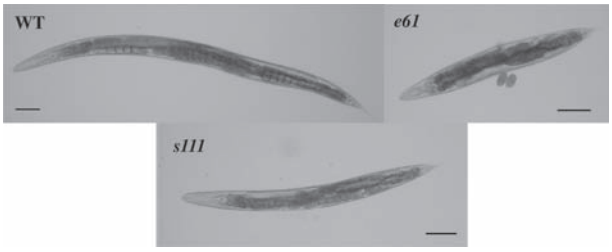



**Figure 1.** The Dpy-5 phenotype. Photomicrographs of a wild type (WT) animal compared with the reference *dpy-5* mutant *e61* and a representative homozygous *s111* mutant, which typically are larger than other *dpy-5* mutants. Scale bars, 100 μm.

***Dpy-5* suppression of the *bli-4* adult-specific blistered phenotype is reduced in males.** We had previously shown that *dpy-5*(*e61)* allele dominantly suppressed the adult-specific blistering by *bli-4(e937)* [15]. Penetrance of blistering was examined in adult F1 progeny of *bli-4(e937)* homozygous males and *dpy-5(*x*) bli-4(e937) unc-13*(*e450*) hermaphrodites (Table 2). Complete suppression of blistering was observed in all non-Dpy non-Unc hermaphrodite progeny. Male progeny of most strains showed only approximately 50% blistering (Table 2). Our results for penetrance of blistering in *dpy-5*(*e61*) *bli-4*(*e937*)/ + *bli-4*(*e937*) males differ from those reported in our previous work [16]. The difference may be attributed to the use of homozygous *bli-4* males in this study. Most affected males blistered at the same location, on the dorsal surface close to the fan (data not shown). Therefore, *dpy-5* dominant suppression of the *bli-4* phenotype differs between the sexes, being incomplete in males.

**Positional cloning of the *dpy-5* gene and identification of molecular lesions.** The mechanism of dominant suppression was not immediately obvious, as many *dpy-5*

**Table 1.** Effects of *dpy-5* mutations on body size, development and reproduction.

| Genotype | Mean length (mm) [a] | Viable brood size [b] | Development time (h) [c] |
|---|---|---|---|
| Wild type | 1.30 (0.040) | 321 ± 25 (*n* = 10) | 58 |
| *e61* | 0.49 (0.051) | 280 ± 32 (*n* = 10) | 69 |
| *s102* | 0.48 (0.044) | 278 ± 40 (n = 10) | 68 |
| *e565* | 0.45 (0.040) | 269 ± 45 (*n* = 9) | 72 |
| *e907* | 0.49 (0.036) | 292 ± 28 (*n* = 10) | 68 |
| *s111* | 0.75 (0.050) | 282 ± 35 (*n* = 10) | 65 |
| *h1964* | 0.44 (0.040) | 222 ± 28 (n = 10) | 74 |
| *h1964/+* | 0.51 (0.090) | 201 ± 49 (*n* = 10) | 81 |

[a] Average length of 20 adult hermaphrodites. Standard error is presented in parentheses.
[b] Mean ± standard deviation.
[c] Time from hatched/L1–L4/adult molt. Values are medians from five animals.

**Table 2.** *bli-4(e937)* penetrance in *dpy-5* heterozygote backgrounds.

| *dpy-5* allele | Percent penetrance [a] | |
|---|---|---|
| | Hermaphrodites | Males |
| *e61* | 0 (559) | 50 (570) |
| *s102* | 0 (524) | 52 (536) |
| *e565* | 0 (520) | 54 (549) |
| *e907* | 0 (644) | 53 (624) |
| *s111* | 0 (589) | 48 (578) |
| *h1964* | 0 (608) | 0 (540) |

[a] Numbers in parentheses indicate total adult animals counted.

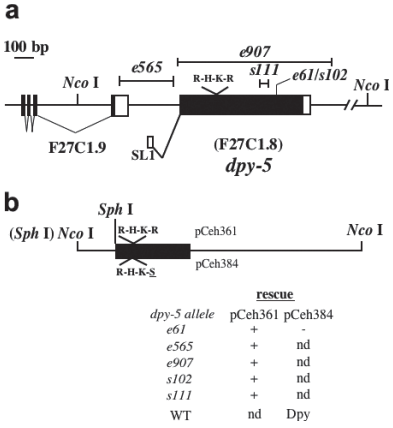
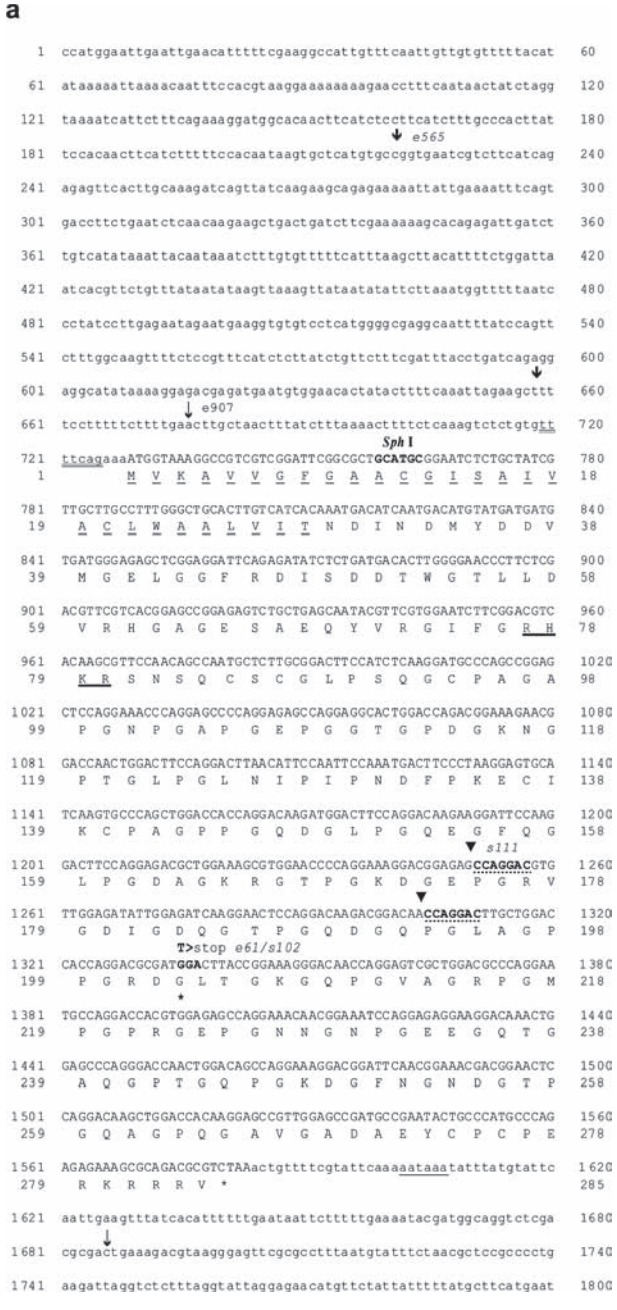



**Figure 2.** *dpy-5* gene structure and location of mutations. (*a*) Rectangles represent exons of genes predicted by GENEFINDER, including F27C1.9 and *dpy-5* (F27C1.8). The *dpy-5* transcript is *trans*-spliced to the SL1-splice leader. The approximate position of the putative RXXR (RHKR in DPY-5) cleavage motif recognized by the PC family is shown. The deletions found in alleles *e565, e907,* and *s111,* are indicated by bars. (*b*) Plasmids used for transformation of nematodes. pCeh361 contains a 3.3-kb *Nco* I fragment that was cloned from cosmid F27C1. *Nco* I sites within the *dpy-5* locus are shown in (*a*). pCeh361 was injected directly into *dpy-5* animals and wild type F1 progeny were scored. Rescue was considered successful (+) if stable transgenic lines, verified by PCR, were established from wild type progeny. pCeh384 is identical to pCeh361 except for a mutation changing the putative PC cleavage site to RHKS, designated *h1964.* Introducing pCeh384 into *dpy-5(e61)* animals failed to rescue the DPY phenotype. Injection of wild type animals resulted in dominant dumpy (Dpy) progeny. The *Sph* I sites demarcate the 750 bp fragment that was cloned into the GFP expression vector pPD95.69. The bracketed site is located within the polylinker of pCeh361.

heterozygotes have no detectable phenotype. To gain a better understanding of the interaction, we identified the molecular nature of *dpy-5*. We took advantage of phenotypic rescue by the cosmid B0342 [24]. An overlapping cosmid F27C1, contained a group 1 cuticle collagen gene designated F27C1.8 [3]. We cloned a 3.3-kb *Nco* I fragment from F27C1, completely encompassing F27C1.8, and tested it for rescue of the *dpy-5* phenotype. Microinjection of all five *dpy-5* strains resulted in complete rescue of the Dpy phenotype (Fig. 2), confirming that F27C1.8 encodes *dpy-5*. The gene structure for *dpy-5* appears quite simple. The coding sequence is a single exon of 857 bp (Figs 2, 3). Immediately upstream of *dpy-5* is a putative gene identified by GENEFINDER and annotated as F27C1.9. The distance between the 3′ end of F27C1.9 and the 5′ splice acceptor of *dpy-5* is 517 bp. RT-PCR using a *dpy-5*-specific primer and a primer specific for the *trans*-spliced leaders SL1 or SL2 showed that *dpy-5* is *trans*-spliced to SL1 (data not shown). The *Nco* I rescuing fragment included only a portion of the last predicted exon of F27C1.9. These data suggest that *dpy-5* is not part of an operon with the upstream gene. The single 857 bp exon contains an open reading frame that begins 3 bp downstream of the SL1 splice acceptor and encodes a predicted cuticle collagen chain composed of 284 amino acids. A putative polyadenylation site, AATAAA, is found 19 bp 3′ of the termination codon (Fig. 3a).

We identified the molecular lesions responsible for the individual *dpy-5* alleles using PCR followed by sequencing. Alleles *e565*, *e907* and *s111* result from deletions in the *dpy-5* locus. The ethylmethane sulfonate (EMS)-induced allele *e565* is a 436 bp deletion upstream of the predicted *dpy-5* coding region (Figs 2, 3a). This presumably removes elements, essential for expression, contained in the relatively short 517 bp 5′ upstream region. Essential regulatory elements are found in close proximity to the initiator methionine in other cuticle collagen genes such as *dpy-7*, *dpy-12*, and *col-19* [4, 21, 25]. The *e907* mutation is a 1009 bp deletion, removing the entire *dpy-5* coding region. This allele was isolated from worms treated with $^{32}P$ [26]. The *s111* mutation is a 54 bp in-frame deletion in the coding region, which removes 16 amino acids from the second GXY repeat domain. The *s111* allele was isolated using formaldehyde, a mutagen shown to cause deletions [27]. In this case, the first seven nucleotides at the right breakpoint of the 54 bp deletion (CCAGGAC; dotted underline in Fig. 3a) are repeated in the wild type sequence. This repeat suggests a possible mechanism for the deletion involving a looping-out of intervening DNA and its subsequent loss in the germline. The EMS-induced *e61* and *s102* mutations are independent isolates of identical nucleotide substitutions, in which Gly 203 is replaced by a stop codon (GGA→TGA), resulting in a predicted truncation of the collagen. We conclude that *dpy-5* encodes a cuticle collagen essential for wild type appearance in *C. elegans*.

**DPY-5 comparison to other cuticle collagens using conserved cysteine spacing.** The structure of cuticle collagen chains is generally conserved starting with a signal peptide, and an N-terminal GXY domain. These features are conserved in DPY-5 (Fig. 3b), including the presence of a predicted PC cleavage motif, Arg-His-Lys-Arg (RHKR) at positions 77–80. The protein contains two GXY repeat domains. The first domain is preceded by a



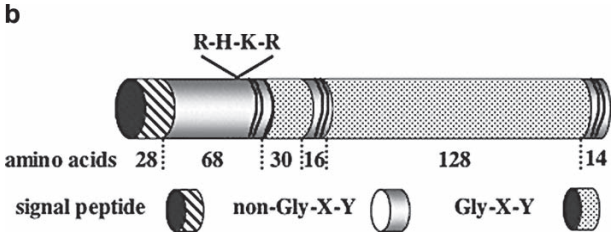




**Figure 3.** DNA and predicted protein sequence of *dpy-5*. (*a*) Sequence starts from 5′ *Nco* I site of the rescuing clone pCeh361. Sequences removed by the *e565* deletion are between bold arrows (↓). Breakpoints for the *e907* deletion mutant are demarcated by downward arrows (↓); those for the *s111* 54 bp in-frame deletion by solid arrowheads (▼). A seven bp repeat at the 5′ end of the *s111*-deleted region and after the right breakpoint are underlined (…). The nucleotide change that results in a stop codon (*) in *e61* and *s102* is in bold. The acceptor for *trans*-splicing the SL1-splice leader is double underlined. The predicted *dpy-5* polyadenylation site in the 3′ untranslated region is underlined. A predicted N-terminal signal peptide sequence is indicated (dashed-underlined), as is the putative furin-like cleavage motif (RHKR) at amino acids 77–80 (bold underlined). Also highlighted is the *Sph* I restriction site used to construct the *dpy-5::gfp* fusion gene plasmid pCeh358. (*b*) Schematic presentation of the overall DPY-5 structure, showing the signal peptide, the N-terminal non-GXY region, the GXY domains, and the C-terminal non-GXY domain. Numbers indicate the length of regions in amino acids. Also shown are the positions of the putative PC cleavage site and cysteine residues (solid lines).

stretch of 16 amino acids after the putative PC cleavage site that includes three cysteines. Between the first and second GXY domains are 16 amino acids that include two cysteines. The non-GXY C terminus is 14 amino acids with two additional cysteines. Spacing of the cysteines may be important for establishing disulfide bonds with collagens and/or non-collagen molecules during cuticle formation. The conserved spacing of the cysteines has been used to order the 158 known cuticle collagens into six groups [3, 6]. On this basis, the *dpy-5* gene product has been assigned to Group I, which contains 68 members [3]. A FASTA search of the *C. elegans* genome identified the Group I cuticular collagens (Fig. 4). The *dpy-5* gene is one of only five genes in this group that produce a visible phenotype, and the only one to produce a Dpy phenotype.

***dpy-5* expression in hypodermal cells during larval development.** RT-PCR was used to determine the temporal expression pattern of *dpy-5*. RNA from each developmental stage was converted to cDNA and amplified using primers for the SL1 splice-leader and *dpy-5*. We detected *dpy-5* transcripts in all larval stages, adults and dauer larvae (Fig. 5). The *dpy-5*-specific product (277 bp) was not amplified from embryo cDNA, suggesting that DPY-5 is not a component of the first larval cuticle.

The temporal and spatial expression pattern was investigated in transgenic animals carrying a *dpy-5::gfp* fusion gene. We analyzed several extrachromosomal and integrated reporter arrays. Although the reporter construct contained a nuclear localization signal, fluorescence was seen in both the nuclei and cytoplasm of expressing cells, which helped define cell boundaries. For all arrays examined, the reporter gene was expressed exclusively in hypodermal cells (Fig. 6), as expected since these cells synthesize and secrete components of the cuticle. The nematode hypodermis consists of individual cells and multinucleate syncytia, the largest being hyp7, which envelops the main body region. During development, hyp7 receives additional nuclei by fusions with the progeny of the lateral seam cells (from the H1, H2 and

V lineages) and the ventral P cells [28, 29]. GFP expression was first observed in mid to late L1 larvae (Fig. 6A), becoming more abundant and consistent throughout the L2–L4 larval stages before disappearing in gravid adults, consistent with the RT-PCR results. GFP was abundant in all larval stages within the hypodermal cells in the head region (hyp3, hyp4, hyp5 and hyp6), the major syncytial hyp7 cell, those nuclei contributed from the P cells, and hypodermal cells in the tail (the hyp8, hyp9, hyp10, hyp11 and hyp12). A typical expression pattern for the L3 larval hermaphrodite is shown in Fig. 6B and C. In contrast, GFP was absent or of low abundance in most larval hypodermal seam cells of the V lineage, seen as a line of dark shadows in Fig. 6A, C and E, although occasional expression was seen in the more anterior seam cells descended from H2 and V1 (filled arrow in Fig. 6C). GFP expression decreases after the L4 molt, and gradually disappears as synthesis of the adult cuticle is completed. Fluorescence diminishes within

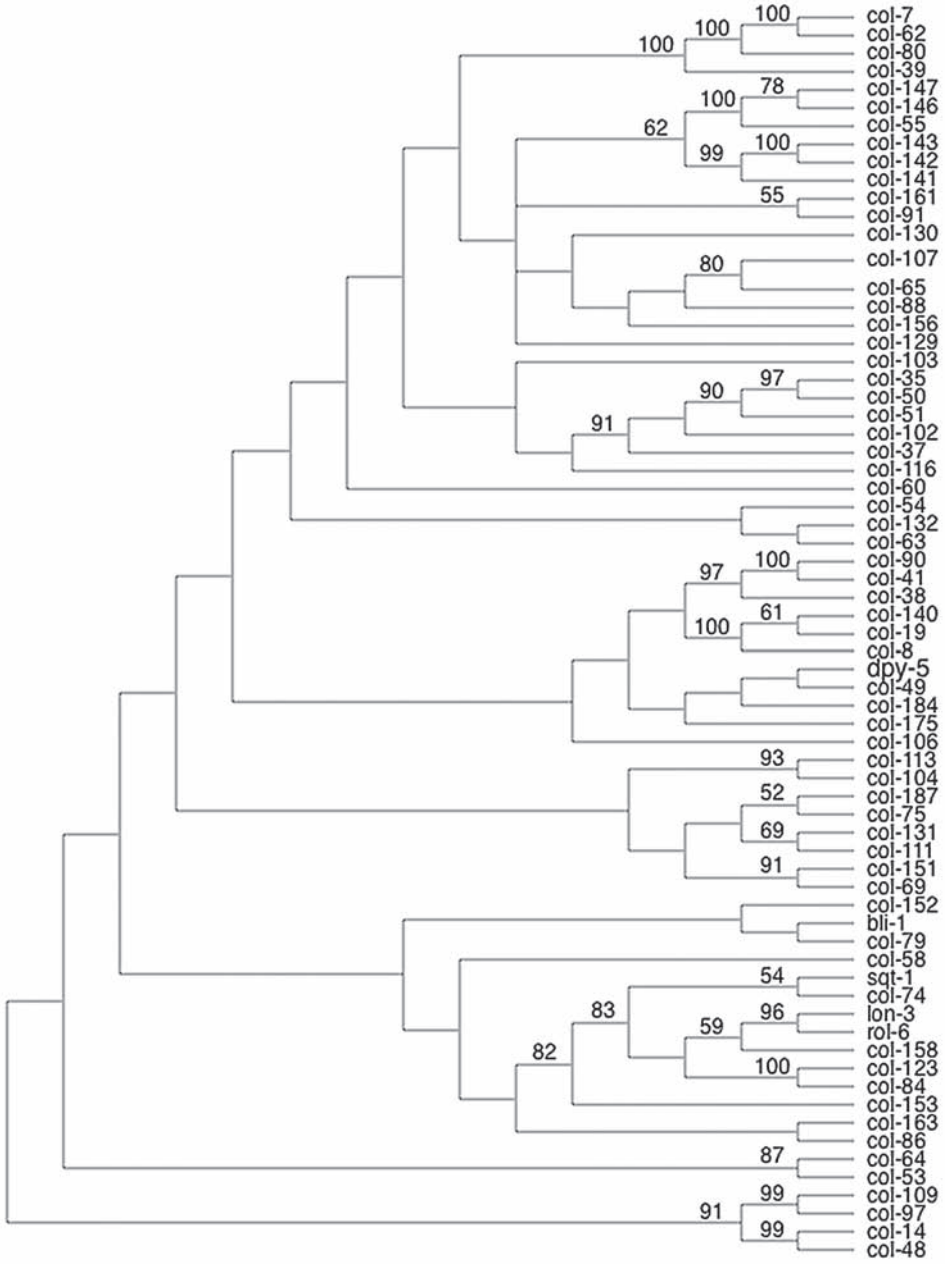



**Figure 4.** *dpy-5* is a Group 1 cuticle collagen. A FASTA search identified all *C. elegans* cuticle collagens with significant amino acid similarity to DPY-5. Of the 158 genes, 68 Group 1 collagens [3, 6] were aligned using ClustalX 1.83 [42]. Shown is the strict consensus of the two most parsimonious trees generated using 'random addition sequence' and 'TBR branch-swapping' in PAUP*4.0b10 [43] using several hundred random starting trees. The shortest tree using 'closest addition sequence' and 'TBR branch swapping' was only one step longer. Bootstrap values [44] (1000 pseudo-replicates) over 50% using the latter procedure are shown.

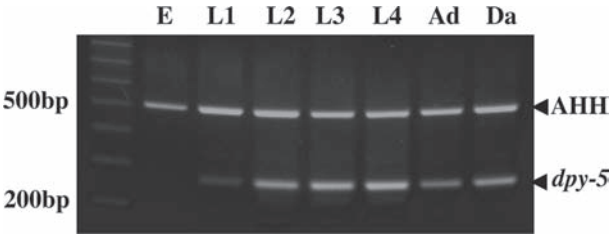



**Figure 5.** Temporal expression of *dpy-5*. An SL-1 primer and antisense *Dpy-5* primer were used to amplify *dpy-5* cDNA from RNA representing each developmental stage, including embryos (E), larvae (L1–L4), adults (Ad), and dauer larvae (Da). *dpy-5* transcripts were amplified from all but embryonic stages as indicated by the expected product of 277 bp. Primers for S-adenosyl homocysteine hydrolase (AHH) were included as a control for the PCR. The AHH primers amplify a product of 509 bp. Numbers on the left indicate a 100 bp DNA ladder.

the head and tail hypodermal cells, the P cells and hyp7 in early adults, although abundant expression continues in the V lineage seam cells (Fig. 6D).

In males, expression of *dpy-5::gfp* generally resembles that of hermaphrodites (Fig. 6E, F). Fluorescence is observed within the head and tail hypodermal cells, the P cells and hyp7 (Fig. 6E; arrowheads), and is absent or of low abundance within the seam cells (open arrows in Fig. 6E, F). This is more easily seen in a higher magnification of the tail region (boxed in Fig. 6E) in which GFP is notably absent in the V5-derived seam cell (open arrow), yet abundant in its sister set cell and the V6- and T-derived specialized hypodermal cells (arrowheads in Fig. 6F) that envelope the sensory rays used in copulation. In summary, expression of the *dpy-5::gfp* reporter gene in the hypodermal cells begins in L1, and continues throughout larval development, ending in adulthood. In seam cells, expression is variable, suggesting that it may be regulated differently than in the hypodermal cells.

**Introduction of a missense mutation in the PC cleavage motif causes a dominant phenotype.** To investigate the function of the predicted PC cleavage site, we mutated *dpy-5*, converting the putative cleavage site from RHKR to RHKS. This substitution was designated *h1964*. We pre-

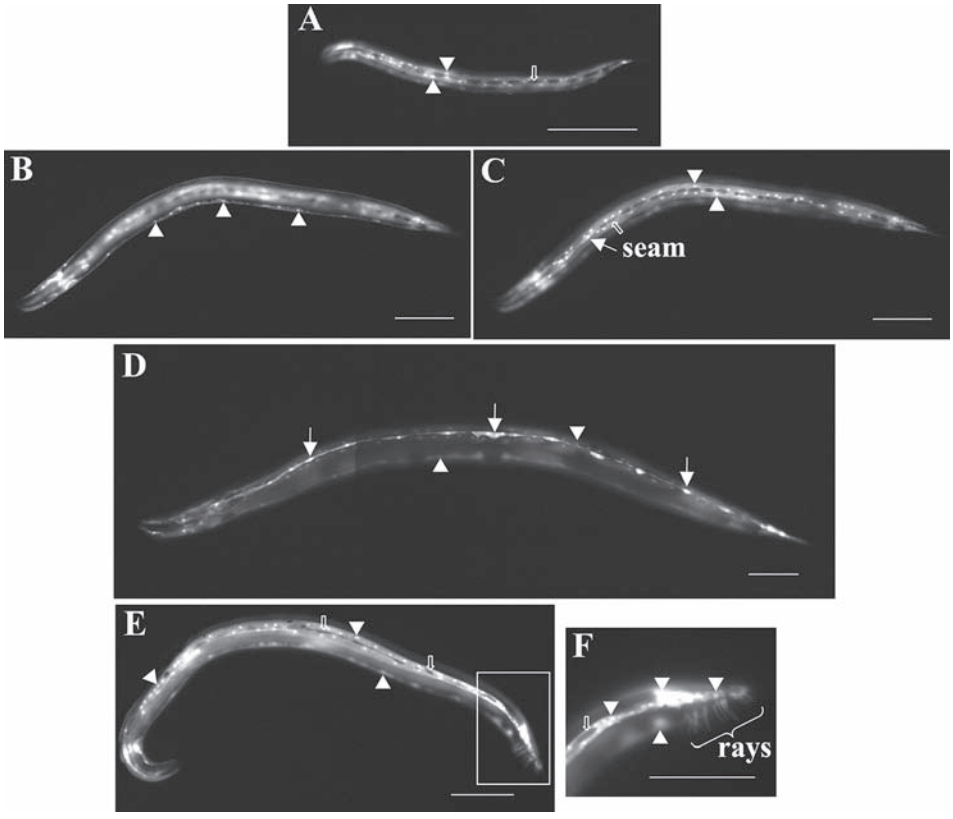



**Figure 6.** Expression pattern of the *dpy-5::gfp* fusion gene. Photomicrographs showing GFP expression in transgenic animals carrying an integrated *dpy-5::gfp* array (*hIs20*). Expression first occurs during mid to late L1 in hypodermal cells. (*A*) Lateral view of a late L1 larva: GFP is localized in the head and tail hypodermal cells, and hyp7 nuclei (filled arrowheads), but not in seam cells (open arrow). This pattern is maintained in later larval stages as shown by a representative L3 worm (*B*, *C*). A mid-body view shows GFP expression in the hypodermal P cell nuclei within the ventral cord (*B*, arrowheads). A lateral plane of focus (*C*) highlights GFP expression in head and tail hypodermal cells, and hyp7 nuclei (arrowheads). Expression is either absent (open arrow) or low (filled arrow) in seam cells. (*D*) Fluorescence diminishes in the majority of hypodermal cells in early adults (arrowheads), but abundant GFP is observed in the seam cells (arrows). (*E*) Expression in an early adult male. Arrowheads indicate the syncytial hypodermal cells. Note the absence of fluorescence in the seam cells (open arrows). (*F*) Higher magnification of boxed area in (*E*) showing absence of GFP in the V5-derived seam cell (open arrow), but within the tail hypodermal cells and the specialized epidermal cells from the V6 and T lineages (arrowheads). Scale bars, 100 μm.

dicted that the mutated site would no longer be recognized for cleavage by a member of the PC family, based on studies of the vertebrate enzymes [19] and the cuticle procollagens SQT-1 and ROL-6 [20]. Injection of pCeh384, carrying the *h1964* mutation, failed to rescue the phenotype of *dpy-5(e61)* animals. As expected, wild type worms that received pCeh384 segregated Dpy progeny, some of which gave rise to stable transgenic lines, *hIs19*. The dominant Dpy phenotype appeared more severe than other alleles of *dpy-5* (Fig. 7). Transgenic worms showed a variable 'blunt' head phenotype not seen for other *dpy-5* alleles. These results suggest that, like SQT-1 and ROL-6, the *dpy-5* gene product is synthesized as a precursor that requires processing at the RHKR site.

We characterized these worms with regard to body length, brood size, development and suppression of blistering. All the progeny of *hIs19* showed a dominant Dpy phenotype. In most respects *hIs19* was similar to other *dpy-5* alleles, although brood sizes were slightly reduced (Table 1). A *bli-4 unc-13 I*; *hIs19*[*dpy-5(h1964)*] *IV* strain was constructed and hermaphrodites failed to exhibit any cuticle blistering. *bli-4(e937)* homozygotes that were heterozygous for *h1964* were generated by crossing this strain to *bli-4(e937)* males. Two interesting observations were made: first, that both hermaphrodite and male progeny from the crosses displayed hypodermal defects (not shown) similar to those displayed by *h1964* homozygotes. The defects were not fully penetrant and thus were difficult to score. These defects were not a result of *bli-4* because a similar effect was found when wild type males were crossed to a non-*bli-4* strain carrying *h1964* (data not shown). Animals carrying one copy of the *h1964* transgene show lower brood sizes and a developmental period that is more retarded than wild type or other *dpy-5* alleles (Table 1). The hypodermal defects may simply reflect mosaic expression of the integrated array. However, we believe this is due to competition between endogenous DPY-5 and that expressed from the *h1964* locus. Secondly, unlike other *dpy-5* alleles, none of the *bli-4(e937)*; *hIs19*[*dpy-5(h1964)*]/+ males were blistered (Table 2). Although the *h1964* dominant epistasis of *bli-4* was complete, we are cautious about interpreting this result because we have not estimated the copy number of the integrated plasmid. It may be that increased expression from the integrated array causes a decrease in penetrance of blistering in the heterozygotes. Disruption of the putative PC cleavage site results in a severe Dpy phenotype. The dominance of the phenotype may be explained by expression of unprocessed collagen interfering with cuticle formation.


**Figure 7.** Expression of a *dpy-5* transgene with a mutation of the putative proprotein cleavage site causes a dominant Dpy phenotype. Photomicrograph of a transgenic animal in which the RHKR site has been mutated to RHKS (*h1964)* shows a dumpy morphology. Note that *h1964* animals show a definitively more rounded head region compared with the typical *e61 dpy-5* morphology (arrow). A wild type animal is shown for comparison. Scale bar, 100 μm.

## Discussion

To investigate the basis for *dpy-5* suppression of the phenotype of *bli-4* and other Bli mutants (Rose and Thacker, unpublished results), we identified the molecular product of *dpy-5*. The gene encodes a cuticle procollagen, mutations in which reduce body size and retard development. The coding region is a single exon; completely deleted in the null mutation *e907*. Expression is likely controlled by the 517 bp upstream sequence, some of which is removed by the *e565* deletion. This region directs expression of a *gfp* reporter in hypodermal cells from the mid-L1 stage to early adults. Variable GFP expression was observed in the seam cells originating from V cell lineages. The compact regulatory region of *dpy-5* and variability of expression in seam cells is unclear, although DPY-5 may be required for the specialized adult cuticle structures, called alae, synthesized by these cells. Recent evidence suggests that *nhr-25*, a member of the nuclear receptor family of transcriptional regulators, could contribute to differential expression in seam cells and hypodermal syncytium [31]. Since our characterization of *dpy-5*, two reports provide evidence that it is required for proper synthesis of the cuticle. Scanning electron micrographs of *dpy-5* worms showed abnormal contraction of the cuticle above the lateral hypodermal seam cells [32] and narrowing of the annuli (circumferential ridges on the surface of the cuticle) [33]. These irregularities result in a dumpy morphology. Furthermore, *dpy-5* results in mislocalization of COL-19 [32] and DPY-7 [33]. Peak *dpy-5* transcription correlates with peak *dpy-13* levels, approximately 2 h before secretion of each new cuticle. The similar expression pattern and structural abnormalities of *dpy-5* and *dpy-13* led to the proposal that the two are obligate partners in the same substructure in the annulus, and function to define its breadth [33].

The types of mutations identified, and the likelihood that DPY-5 is involved in a number of physical interactions, can explain many of the genetic behaviors of *dpy-5* alleles. The reference allele, *e61*, shows a variable, weak dominant effect causing heterozygotes to become slightly Dpy. This is also true for *s111*, an in-frame deletion. Both mutations may behave similarly to mutant vertebrate collagens that contain substitutions of glycine in the GXY domains. Lesions of this type act like null mutations in that the proteins they encode are not incorporated into extracellular structures, but accumulate in the cytoplasm [34]. These aberrant proteins are often degraded along with any proteins with which they physically associate. Consequently, incomplete triple helices including mutant DPY-5 may degrade, leading to a weak dominant phenotype. This may also explain the genetic interaction between *e61* and *smg* mutants. The seven *smg* genes control the selective degradation of mRNAs containing premature stop codons [30]. Mutations in *smg* genes suppress the *e61* Dpy phenotype, *i.e. e61*/*e61*; *smg(-)* animals have a body length similar to wild type. This suggests that the truncated *e61* product is incorporated into an apparently functional cuticle, compatible with the prediction that DPY-5 forms homotrimers. If DPY-5 functioned as a heterotrimer, one might expect an abnormal body morphology in a *smg(-)* background due to interference of the truncated chain with formation of triple helices by other collagen species. This may also explain why *smg* mutations enhance the weak dominance of *e61* since *e61*/+; *smg(-)* worms are always dumpy [35]. In this case, the *e61* product probably interferes with the production of homotrimers with normal DPY-5, perhaps leading to their accumulation and subsequent degradation.

Vertebrate fibrillar collagen molecule assembly begins with association of the carboxyl non-GXY domains and requires the cysteines in this region for disulfide bonding [9]. However, directed mutagenesis of SQT-1 shows that the carboxyl domain cysteines are not required for assembly into triple helices [14]. Our interpretation that truncated DPY-5 is partially functional also suggests that this domain is not required for triple helix formation, and that cuticle collagen molecules assemble without needing to form disulfide bonds. This feature of collagen assembly appears conserved with that of nonfibrillar vertebrate type IX and XII FACIT collagens [34–36].

We have demonstrated that the putative PC cleavage site, arginine at position 80, is essential for normal DPY-5 function. The simplest interpretation is that this site is recognized and processed by a member of the PC family. The presence of a PC-like cleavage site in most, if not all, cuticle collagens was described as 'homology box A' [6], and studies on SQT-1 and ROL-6 [20, 37] and a similar mutation in *dpy-10* [23] strongly suggest that cuticle collagens are synthesized as precursor proteins. For cuticle procollagens to assembly correctly, it appears that they need to be cleaved to remove N-terminal sequences. Cleavage of procollagens by these enzymes is not unique to *C. elegans*. It has recently been shown that a PC, furin-like, activity processes human procollagen types V [38] and XVII [39]. However, in these cases, processing occurs in the C-terminal region rather than the N terminus. The dominant phenotype of animals expressing the *h1964* transgene probably reflects an interaction between unprocessed *h1964* product and wild type DPY-5. Assembly of collagen fibers composed of mutant and normal DPY-5 could perturb or inhibit generation of higher order structures. Reducing the *dpy-5(h1964)* array copy number was more deleterious to the morphology and development of the animals, causing a slower growth rate as well as hypodermal defects. Some of these characteristics could be attributed to a mosaic expression of wild type cuticle by the hypodermal cells or perhaps by necrosis of hypodermal cells as aberrant collagens accumulate. There is evidence that SQT-1 procollagen can be incorporated into cuticles [37]. Unprocessed SQT-1 has been identified in cuticle preparations by Western analysis, with the majority having mobility consistent with being in a dimeric form, albeit at reduced levels compared with wild type SQT-1.

Reduction of wild type DPY-5 or expression of aberrant forms suppresses blistering of the adult cuticle. The adult cuticle is composed of an outer cortical layer and an inner basal layer separated by a fluid-filled space containing columnar structures called struts [40]. Blister formation is thought to occur by separation of the outer and inner layers caused by compromised or unstable struts. It could be that DPY-5 is located at one or both surfaces of the cortical and basal layers, and when missing or abnormal, these layers fuse, reducing the formation of blisters. Examination of adult *dpy-5* cuticles by electron microscopy demonstrates that this may be the case [32, 33, 41]. Work from our laboratory has shown that two other mutants, *dpy-6* and *dpy-13*, also show dominant suppression of blistering. At present, the molecular nature of *dpy-6* has not been determined. DPY-13 is a different class of cuticle collagen to DPY-5, although recent evidence suggests these proteins function within the same cuticle substructure [33]; consequently mutations in either would probably similarly affect the structure or integrity of the cuticle. The *bli-1* and *bli-2* genes encode cuticle collagens that are reportedly components of the struts (J. Kramer, unpublished results). One explanation why *bli-4(e937)* homozygotes exhibit a phenotype might be incomplete processing of proBLI-1 or/and proBLI-2, since both contain PC-like cleavage sites. Compromised strut formation or stability would then lead to blistering.

One reason for *dpy-5* alleles' inability to suppress blistering in males is that the cuticle composition may differ between hypodermal cells expressing different repertoires of cuticle collagens. Different collagen genes are expressed in different subsets of hypodermal cells [6]. In

addition, the morphology of the male tail is highly specialized for its role in copulation. It may be that the male tail is subject to more mechanical stress, which might lead it to blister more readily when the structure of the cuticle is perturbed. Clearly, the combination of temporal, spatial and quantitative differences in cuticle collagen expression; combined with the large numbers of collagen genes, produces a high level of diversity in the complexion of the cuticle.

*Acknowledgements.* Several strains were provided by the *Caenorhabditis elegans* Genetics Center. Thanks to Jennifer Crew and Jim Kramer for sharing unpublished results regarding *bli-1* and *bli-2*, Andy Fire for the *gfp* vector pPD95.69, and Alan Coulson for cosmids. We also thank Erin Gilchrist, Alex Parker, and David Baillie for critical comments on the manuscript. This work supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

1 Cox G. N., Kusch M. and Edgar R. S. (1981) Cuticle of *Caenorhabditis elegans:* its isolation and partial characterization. J. Cell Biol. **90:** 7–17
2 Cox G. N., Staprans S. and Edgar R. S. (1981) The cuticle of *Caenorhabditis elegans*. II. Stage-specific changes in ultrastructure and protein composition during postembryonic development. Dev. Biol. **86:** 456–470
3 Johnstone I. L. (2000) Cuticle collagen genes. Expression in *Caenorhabditis elegans.* Trends Genet. **16:** 21–27
4 Liu Z., Kirch S. and Ambros V. (1995) The *Caenorhabditis elegans* heterochronic gene pathway controls stage- specific transcription of collagen genes. Dev. Biol. **121:** 2471–2478
5 Johnstone I. L. and Barry J. D. (1996) Temporal reiteration of a precise gene expression pattern during nematode development. EMBO J. **15:** 3633–3639
6 Kramer J. M. (1997) Extracellular matrix. In: *C. elegans* II, pp. 471–500, Riddle D. L., Blumenthal T., Meyer B. J. and Priess J. R. (eds.), Cold Spring Harbor Laboratory Press, New York
7 Seidah N. G. and Chretien M. (1997) Eukaryotic protein processing: endoproteolysis of precursor proteins. Curr. Opin. Biotechnol. **8:** 602–607
8 Steiner D. F. (1998) The proprotein convertases. Curr. Opin. Chem. Biol. **2:** 31–39
9 Engel J. and Prockop D. J. (1991) The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. Annu. Rev. Biophys. Biophys. Chem. **20:** 137–152
10 Kramer J. M. and Johnson J. J. (1993) Analysis of mutations in the *sqt-1* and *rol-6* collagen genes of *Caenorhabditis elegans.* Genetics **135:** 1035–1045
11 Johnstone I. L., Shafi Y. and Barry J. D. (1992) Molecular analysis of mutations in the *Caenorhabditis elegans* collagen gene *dpy-7.* EMBO J. **11:** 3857–3863
12 Von Mende N., Bird D. M., Albert P. S. and Riddle D. L. (1988) *dpy-13:* a nematode collagen gene that affects body shape. Cell **55:** 567–576
13 Nyström J., Shen Z. -Z., Aili M., Fleming A. J., Leroi A. and Tuck S. (2002) Increased or decreased levels of *Caenorhabditis elegans lon-3*, a gene encoding a collagen, cause reciprocal changes in body length. Genetics **161:** 83–97
14 Kramer J. M., Johnson J. J., Edgar R. S., Basch C. and Roberts S. (1988) The *sqt-1* gene of *C. elegans* encodes a collagen critical for organismal morphogenesis. Cell **55:** 555–565
15 Van Der Keyl H., Kim H., Espey R., Oke C. V. and Edwards M. K. (1994) *Caenorhabditis elegans sqt-3* mutants have mutations in the *col-1* collagen gene. Dev. Dyn. **201:** 86–94
16 Peters K., McDowall J. and Rose A. M. (1991) Mutations in the bli-4 (I) locus of *Caenorhabditis elegans* disrupt both adult cuticle and early larval development. Genetics **129:** 95–102
17 Thacker C., Peters K., Srayko M. and Rose A. M. (1995) The *bli-4* locus of *Caenorhabditis elegans* encodes structurally distinct kex2/subtilisin-like endoproteases essential for early development and adult morphology. Genes Dev. **9:** 956–971
18 Thacker C., Marra M. A., Jones A., Baillie D. L. and Rose A. M. (1999) Functional genomics in *Caenorhabditis elegans:* An approach involving comparisons of sequences from related nematodes. Genome Res. **9:** 348–359
19 Molloy S. S., Anderson E. D., Jean F. and Thomas G. (1999) Bicycling the furin pathway: from TGN localization to pathogen activation and embryogenesis. Trends Cell Biol. **9:** 28–35
20 Yang J. and Kramer J. M. (1994) *In vitro* mutagenesis of *Caenorhabditis elegans* cuticle collagens identifies a potential subtilisin-like protease cleavage site and demonstrates that carboxyl domain disulfide bonding is required for normal function but not assembly. Mol. Cell. Biol. **14:** 2722–2730
21 Brenner S. (1974) The genetics of *Caenorhabditis elegans*. Genetics **77:** 71–94
22 Mello C. and Fire A. (1995) DNA transformation. In: *C. elegans:* modern biological analysis of an organism, pp. 452–482, Epstein H. F. and Shakes D. C. (eds.), Academic press, New York
23 Levy A. D., Yang J. and Kramer J. M. (1993) Molecular and genetic analyses of the *Caenorhabditis elegans* dpy-2 and dpy-10 collagen genes: a variety of molecular alterations affect organismal morphology. Mol. Biol. Cell **4:** 803–817
24 Browning H., Berkowitz L., Madej C., Paulsen J. E., Zolan M. E. and Strome S. (1996) Macrorestriction analysis of *Caenorhabditis elegans* genomic DNA. Genetics **144:** 609–619
25 Gilleard J. S., Barry J. D. and Johnstone I. L. (1997) cis regulatory requirements for hypodermal cell-specific expression of the *Caenorhabditis elegans* cuticle collagen gene dpy-7. Mol. Cell. Biol. **17:** 2301–2311
26 Babu P. (1974) Biochemical genetics of *Caenorhabditis elegans.* Mol. Gen. Genet. **135:** 39–44
27 Johnsen R. C. and Baillie D. L. (1988) Formaldehyde mutagenesis of the eT1 balanced region in *Caenorhabditis elegans:* dose-response curve and the analysis of mutational events. Mutat. Res. **1:** 137–147
28 Sulston J. E. and Horvitz H. R. (1977) Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans.* Dev. Biol. **56:** 110–156
29 Sulston J. E., Schierenberg E., White J. G. and Thomson J. N. (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans.* Dev. Biol. **100:** 64–119
30 Hodgkin J., Papp A., Pulak R., Ambros V. and Anderson P. (1989) A new kind of informational suppression in the nematode *Caenorhabditis elegans.* Genetics **123:** 301–313
31 Gissendanner C. R. and Sluder A. E. (2000) nhr-25, the *Caenorhabditis elegans* ortholog of the ftz-f1, is required for epidermal and somatic gonad development. Dev. Biol. **221:** 259–272
32 Thein M. C., McCormack G., Winter A. D., Johnstone I. L., Shoemaker C. B. and Page A. P. (2003) *Caenorhabditis elegans* exoskeleton collagen COL-19: an adult-specific marker for collagen modification and assembly and the analysis of organismal morphology. Dev. Dyn. **226:** 523–539
33 McMahon L., Muriel J. M., Roberts B., Quinn M. and Johnstone I. L. (2003) Two sets of interacting collagens form functionally distinct substructures within the *Caenorhabditis elegans* extracellular matrix. Mol. Biol. Cell **14:** 1366 –1378
34 Myllyharju J. and Kivirikko K. I. (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet. **20:** 33–43
35 Labourdette L. and Van Der Rest M. (1993) Analysis of the role of the COL1 domain and its adjacent cysteine- containing sequence in the chain assembly of type IX collagen. FEBS Lett. **320:** 211–214
36 Mazzorana M., Gruffat H., Sergeant A. and Van Der Rest M. (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

37 Yang J. and Kramer J. M. (1999) Proteolytic processing of *Caenorhabditis elegans* SQT-1 cuticle collagen is inhibited in right roller mutants whereas cross-linking is inhibited in left roller mutants. J. Biol. Chem. **274:** 32744–32749

38 Imamura Y., Steiglitz B. M. and Greenspan D. S. (1998) Bone morphogenetic protein-1 processes the $NH_2$-terminal propeptide and a furin-like proprotein convertase processes the COOH-terminal propeptide of pro-alpha1(V) collagen. J. Biol. Chem. **273:** 27511–27517

39 Schacke H., Schumann H., Hammanmi-Hauasli N., Raghunath M. and Bruckner-Tuderman L. (1998) Two forms of collagen XVII in keratinocytes. A full-length transmembrane protein and a soluble ectodomain. J. Biol. Chem. **273:** 25937–25943

40 White J. (1988) The anatomy. In: The nematode *Caenorhabditis elegans*, pp. 81–122, Wood W. B. (ed.), Cold Spring Harbor Laboratory Press, New York

41 Ouazana R., Garrone R. and Godet J. (1985) Characterization of morphological and biochemical defects in the cuticle of a dumpy mutant of *Caenorhabditis elegans.* Comp. Biochem. Physiol **80B:** 481–484

42 Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F. and Higgins D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:** 4876–4882

43 Swofford D. L. (2001) *Paup* (Phylogenetic analysis using parsimony and other methods)*. 4.0b10 beta version, Sinauer Associates, Sunderland

44 Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783–791