Abstract
The unc-52 gene encodes the nematode homologue of mammalian perlecan, the major heparan sulphate proteoglycan of the extracellular matrix. This is a large complex protein with regions similar to low-density lipoprotein receptors, laminin and neural cell-adhesion molecules. Three major classes of UNC-52/perlecan isoforms are produced through alternative splicing, and these distinct proteins exhibit complex spatial and temporal expression patterns throughout development. The unc-52 gene plays an essential role in myofilament assembly in body-wall muscle during embryonic development.

Introduction
The unc-52 gene in Caenorhabditis elegans produces several large proteins that are homologous with perlecan [1], the mammalian basement-membrane heparan sulphate proteoglycan (Figure 1) [2–4]. The perlecan core protein has a signal sequence and five distinct domains. The first domain is very short, and is unique to perlecan.
sequences with similarity to the low-density lipoprotein (LDL) receptor (domain II), laminin (domains III and V) and the neural cell-adhesion molecule (NCAM; domain IV). Mammalian perlecan is the major heparan sulphate proteoglycan of the extracellular matrix (ECM), and has been

**Figure 1**

Structural comparison of the human, mouse and nematode perlecan core proteins

Adapted from lozzo et al. [1] with permission [(1994) © Biochemical Society].

**Figure 2**

Diagram showing the proteins associated with dense bodies in C. elegans

perlecan (*unc-52*)

ECM

plasma membrane

alpha and beta integrin

(*pat-2*, *pat-3*)

**UNC-112**

**ILK**

**PINCH (unc-97)**

(t *pat-4*)

talin

 vinculint

 ( *deb-1*)

actin filament

(*act-123*, *pat-10*, *lev-11* etc)
implicated in a number of biological processes, including glomular filtration [5], mitogenesis and angiogenesis [6], and cell adhesion through interactions with focal-adhesion complexes [7,8]. Biochemical studies indicate that perlecan binds both itself and other ECM components including laminin, collagen, nidogen and fibronectin [9–13].

In the nematode C. elegans, UNC-52/perlecan is found in the basement membrane between the body-wall muscle cells and the hypodermis [1,14,15]. The absence of UNC-52/perlecan prevents myofilament lattice assembly during late embryogenesis, resulting in a paralysed, arrested elongation at two-fold (Pat) terminal phenotype [16]. Body-wall muscle cells in C. elegans are arranged in four quadrants, two dorsal and two ventral, beneath the hypodermis (reviewed in [17]). Each quadrant runs the length of the animal and consists of a double row of spindle-shaped cells. Within each muscle cell, the thin and thick filaments of the myofilament lattice lie just beneath the plasma membrane facing the hypodermis. The myofilament lattice is anchored to the muscle-cell membrane and adjacent basement membrane by integrin-containing structures called dense bodies and M-lines, which, in turn, are linked to intermediate filament arrays that extend throughout the hypodermis and attach to the cuticle [16–18]. The body-wall muscle in unc-52-null mutants lacks organized A- and I-bands, and morphological studies reveal that even the earliest stages of myofilament lattice assembly are defective [1,16].

UNC-52/perlecan is concentrated at the muscle-cell dense bodies and M-lines [1,14,15]. These structures contain many of the components found in focal-adhesion plaques, the cell–ECM contacts of tissue-culture cells (Figure 2) [16,19–25]. The interaction of UNC-52/perlecan and αPAT-2/βPAT-3/integrin, either directly or indirectly, is a key early event in the assembly of dense bodies and M-lines [1,15,16,18]. Careful analyses of wild-type and mutant animals with antibodies to muscle components have revealed

**Figure 3**
Diagram showing the exon structure of the unc-52 gene and the location of the st549, st560, e444 and gk3 mutations

Also shown are the three major protein isoforms encoded by this gene, and the regions of these proteins that are recognized by the GM1, GM3 and GM9 polyclonal antibodies.
that body-wall muscle assembly is initiated by events occurring at the muscle-cell membrane [1,16,18,23,26,27]. Early in embryonic body-wall muscle development, muscle proteins accumulate at membranes where adjacent muscle cells contact each other and the hypodermis [18]. This has been termed muscle-cell polarization, and at this stage, muscle, basement membrane and hypodermal components are all co-localized [18]. In the later stages of body-wall muscle development, UNC-52/perlecan and zPAT-2/βPAT-3/integrin become organized along the basal cell membrane in structures resembling cell-matrix adhesion complexes. Dense body and M-line components then assemble at these sites. The thick filaments of the lattice assemble with the M-line and the thin filaments with the dense bodies [1,16,18,26-28]. When UNC-52/perlecan is not present in the basement membrane, the zPAT-2/βPAT-3 integrin heterodimers do not localize to the basal surface of the muscle cell membrane and dense bodies and M-lines do not assemble [1,18].

The unc-52 gene encodes three distinct protein isoforms

The unc-52 gene is a large, complex locus consisting of 37 exons (Figure 3) [15]. Three different polyadenylation sites, located downstream of exons 10, 26 and 37, are used to generate mRNAs encoding three distinct UNC-52/perlecan polypeptides (Figure 3). The short (S) isoform is an 1160 amino acid protein containing three of the five encoded protein domains (I-III), the medium (M) isoform is a 2482 amino acid protein containing four of the five encoded protein domains (I-IV) and the long (L) isoform is a 3375 amino acid protein containing all five domains. In addition to the generation of these three major variants of UNC-52/perlecan, alternative splicing of exons 6, 16, 17, 18, 21 and 22 generates diversity within domains III and IV. There could be as many as 50 UNC-52/perlecan isoforms if all combinatorial possibilities between S, M and L variants are realized.

M isoforms of unc-52/perlecan are sufficient for myofilament lattice assembly in the body-wall muscles of C. elegans

The st549 mutation results in the complete elimination of all UNC-52/perlecan isoforms and defines the null phenotype of this gene (see Figure 3) [1,15,29]. Other mutations affecting different regions of the unc-52 gene have proven useful in elucidating the functions of the three major protein isoforms. For example, the gk3 mutation results in the elimination of the L isoform of UNC-52/perlecan but does not affect the M or S polypeptides [15]. Animals homozygous for this mutation are essentially wild type, indicating that domain V is not essential for integrin anchorage and myofilament assembly. The st560 mutation results in the elimination of the M and L UNC-52/perlecan polypeptides but does not affect the S isoform [15]. Since embryos homozygous for this mutation exhibit a Pat terminal phenotype and are defective in myofilament lattice assembly, we can conclude that the S isoform of UNC-52/perlecan is not sufficient for integrin anchorage and myofilament assembly during embryogenesis [15]. Taken together, these results reveal that the M UNC-52/perlecan polypeptide is sufficient for myofilament lattice assembly. We do not know whether the L isoform of UNC-52 can substitute for this function because we do not have a mutation that specifically eliminates the M isoform while retaining the L isoform.

Spatial and temporal differences in the localization of UNC-52 isoforms

Polyclonal antibodies have been raised against different domains of the UNC-52/perlecan protein sequence to determine the expression patterns of the various isoforms [15]. The polyclonal serum GM1 recognizes a region of domain III present in all UNC-52/perlecan isoforms, whereas the polyclonal serum GM3 recognizes a conserved region of domain IV (Figure 3). The domain V-specific serum GM9 does not detect the L isoform and has not proven useful for determining the localization of this protein isoform. Staining with either GM1 or GM3 has revealed that UNC-52/perlecan is specifically associated with contractile tissues in C. elegans and is not found in the basement membranes lining the pseudocoelom or surrounding the intestine [15]. Thus, UNC-52 is not a general basement-membrane component in C. elegans. UNC-52/perlecan expression is first observed in comma-stage embryos at regions of contact between adjacent muscle cells, and then spreads over the basal face of each muscle cell where the basement membrane is located [15]. Both GM1 and GM3 stain the body-wall muscles in embryos, and both antisera also stain the basement membranes associated with the body wall, pharyngeal, anal and sex-specific muscles in
larvae and adult animals [15]. The body-wall, pharyngeal and anal muscles are required for movement, eating and defecation, respectively. The only difference in staining between the two antisera occurs in embryos where GMI stains the pharyngeal and anal muscles in addition to the body-wall muscles [15]. The fact that GM3 does not stain these tissues reveals that the M and/or L isoforms of UNC-52/perlecan are restricted to the body-wall muscles during embryogenesis. However, these isoforms are expressed in the pharyngeal and anal muscles at later stages of development.

We have also determined that the S isoform of UNC-52/perlecan is restricted to the pharyngeal and anal muscles during embryonic development by staining embryos homozygous for the st560 mutation with GMI [15]. This mutation does not affect the S UNC-52/perlecan isoform but does eliminate the M and L isoforms. Thus, the only UNC-52/perlecan protein present in these mutants should be the S isoform. Staining with GMI revealed that the S isoform of UNC-52/perlecan is associated with the pharynx and the anal sphincter and depressor muscles during embryogenesis. The lack of body-wall muscle staining with GMI leads us to conclude that this particular UNC-52/perlecan polypeptide is not present in the body-wall muscle during assembly of the myofilament lattice. The role of this isoform in pharyngeal muscle development is not clear since there is no evidence of pharyngeal disruption in unc-52-null mutants.

Evidence for a temporal shift between early and late UNC-52 isoforms

Alternative splicing of exons 16, 17 and 18 gives rise to isoforms that vary in the number of NCAM-like Ig repeats within domain IV [1,29]. Null mutations in these three exons eliminate some, but not all, of the M and L UNC-52/perlecan isoforms. Animals that are homozygous for any one of these mutations develop normally as young larvae but become progressively paralysed as they mature. This paralysis is caused by the gradual disruption of the myofilament lattice in the body-wall muscle cells posterior to the head [30,31]. We have shown that these mutations do not disrupt the accumulation of UNC-52 during early development, but instead have dramatic effects on the accumulation of UNC-52 isoforms associated with the body-wall muscles in older animals [15,29]. In homozygous e444 hermaphrodites, for example, staining of the body wall muscles appears to be normal until the late L4 stage but is greatly reduced relative to wild-type animals in adults [15]. Within the body-wall muscles, myosin becomes highly disorganized and forms large aggregates [31]. Curiously, staining of the body-wall muscles in the head is not affected, even in older animals, and staining of the pharynx and uterine muscles also appears normal. Thus, the e444 mutation affects an adult-specific subset of UNC-52 isoforms associated with most of the body-wall muscles. Other null alleles in exons 16, 17 and 18 behave in a similar manner [15].

Several NCAM repeats within domain IV are dispensable for muscle assembly

Alternative splicing between exons 15 and 19 regulates the number of NCAM-like Ig repeats within domain IV of UNC-52/perlecan, and is associated with both spatial and temporal differences in isoform expression [15,29]. This expression pattern is regulated by an RNA-binding protein encoded by the mec-8 gene which appears to mask certain splice-site choices during late embryogenesis [32]. Considering this complexity of UNC-52/perlecan isoform expression, it is curious that a deletion eliminating all variants except the 15–19 splice product does not noticeably affect either muscle development or function. This observation suggests that alternative splicing within this region is associated with fine modulation of function rather than large-scale changes in biophysical properties. In contrast, alternative splicing events that give rise to the three major groups of isoforms are likely to significantly change the properties of these proteins, including their ability to interact with transmembrane receptors such as integrin.

Conclusion

The localization of UNC-52 over the body-wall muscles and the effects of unc-52-null mutations on myofilament assembly suggest that UNC-52 anchors the dense bodies and M-lines, perhaps through interactions with integrin. Whether UNC-52 plays an instructional role or simply an attachment role in the assembly of integrin complexes at the muscle-cell membrane is not clear. However, without a stable attachment structure at the muscle-cell membrane, sarcomere units within muscle cells cannot be properly organized [16]. Several observations suggest that UNC-
52/perlecan interacts directly with βPAT-3/integrin. First, the distribution of UNC-52 in the basement membrane overlaps with that of PAT-3 at the plasma membrane [14,16,18]. Second, UNC-52 and Pat-3 mutants both exhibit a severe Pat phenotype and have similar defects in myofilament lattice assembly [16]. Third, βPAT-3/integrin anchorage and myofilament lattice assembly occurs in mutants, supporting the idea that UNC-52/perlecan is required to anchor integrin [18].

Of particular importance in our study of UNC-52/perlecan function is the demonstration that the M isoform is sufficient for myofilament organization and attachment in the body-wall muscle of C. elegans [15]. Since this polypeptide does not contain domain V, it is clearly dispensable for myofilament lattice assembly [15]. However, this domain most likely has a role in other biological developmental processes. If UNC-52/perlecan serves as an adhesive substrate for anchoring the myofilament lattice, then these results help to identify regions that are critical for integrin anchorage and myofilament lattice assembly.

This work was supported by grants from the Medical Research Council of Canada and the Natural Sciences and Engineering Research Council of Canada to D.G.M.

References
26 Waterston, R. H. (1989) EMBO J. 8, 3429-3436

Received 6 November 2000