Products of the \textit{unc-52} gene in \textit{Caenorhabditis elegans} are homologous to the core protein of the mammalian basement membrane heparan sulfate proteoglycan

Teresa M. Rogalski, Benjamin D. Williams, Gregory P. Mullen, and Donald G. Moerman

Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4; \textsuperscript{1}Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110 USA

Mutations in the \textit{unc-52} gene of \textit{Caenorhabditis elegans} affect attachment of the myofilament lattice to the muscle cell membrane. Here, we demonstrate that the \textit{unc-52} gene encodes a nematode homolog of perlecan, the mammalian basement membrane heparan sulfate proteoglycan. The longest potential open reading frame of this gene encodes a 2482-amino-acid protein with a signal peptide and four domains. The first domain is unique to the \textit{unc-52} polypeptide, whereas the three remaining domains contain sequences found in the LDL receptor (domain II) laminin (domain III) and N-CAM (domain IV). We have identified three alternatively spliced transcripts that encode different carboxy-terminal sequences. The two larger transcripts encode proteins containing all or part of domain IV, whereas the smaller transcript encodes a shortened polypeptide that completely lacks domain IV. We have determined that the disorganized muscle phenotype observed in \textit{unc-52(st196)} animals is caused by the insertion of a Tcl transposon into domain IV. Two monoclonal antibodies that recognize an extracellular component of all contractile tissues in \textit{C. elegans} fail to stain embryos homozygous for a lethal \textit{unc-52} allele. We have mapped the epitopes recognized by both monoclonal antibodies to a region of domain IV in the \textit{unc-52}-encoded protein sequence.

[Key Words: \textit{unc-52} gene, \textit{C.elegans}; muscle; basement membrane; myofilament lattice; perlecan]

Received April 22, 1993, revised version accepted June 17, 1993.

Basement membranes are complex structures that are found associated with many different cell types. They consist of thin sheets of extracellular matrix whose major components include collagen type IV, laminin, entactin/nidogen, and heparan sulfate proteoglycan [Martin and Timpl 1987]. In muscle, the basement membrane plays an important role in directing myofilament assembly during differentiation, and in maintaining the ordered structure of the myofilament arrays. In addition, linkage of the contractile proteins to the cell membrane and extracellular matrix is essential for transmitting force during contraction.

In the body wall muscle of \textit{Caenorhabditis elegans} the myofilament lattice lies just beneath the cell surface facing the hypodermis and is anchored to the basement membrane through a series of lateral attachments [Waterston et al. 1980; Francis and Waterston 1985, 1991; Waterston 1988]. The thick filaments of the lattice are held in register by the proteins of the M-line analogs, whereas the thin filaments are anchored by dense bodies and attachment plaques [Waterston et al. 1980]. Dense bodies are finger-shaped structures that project from the sarcolemma, and they perform a function analogous to the Z-line in vertebrate muscle. A transmembrane complex links the dense bodies and M-line analogs to the basement membrane, which is interposed between the muscle cells and the overlying hypodermis [Francis and Waterston 1985]. Several of the proteins involved in these attachments have been identified, including \textit{\alpha}-actinin and vinculin, which are components of the dense bodies [Francis and Waterston 1985; Barstead and Waterston 1989], and \textit{\beta}-integrin, which is part of the transmembrane complex linking the M-lines and the dense bodies to the basement membrane [S. Gettner and L. Reichardt, pers. comm.]. Here, we report that the heparan sulfate proteoglycan component of the extracellular matrix is also required to anchor the myofilament lattice to the muscle cell membrane in \textit{C. elegans}.

Mutations in the \textit{unc-52} gene of \textit{C. elegans} affect myofilament–membrane attachments in the body wall muscle cells. Genetic analyses of this gene have identified both viable and lethal alleles [Brenner 1974, Gilchrist and Moerman 1992; B.D. Williams and R.H. Waterston, unpubl.]. Mutants homozygous for any of the
viable unc-52 alleles are able to move normally as larvae, but adult hermaphrodites are paralyzed, thin, and partially egg-laying defective (Brenner 1974; Gilchrist and Moerman 1992). The paralyzed phenotype is the result of the gradual disruption of the body wall muscle cells posterior to the pharynx that begins at the third or fourth larval stage (MacKenzie et al. 1978; Waterston et al. 1980). The dense bodies in affected cells are fractured, and the myofilaments are no longer anchored to the cell membrane (Waterston et al. 1980). Several unc-52 mutations result in an early larval lethal phenotype. Mutant embryos homozygous for these alleles are paralyzed and fail to elongate past the twofold stage (B.D. Williams and R.H. Waterston, unpubl.). This unc-52 terminal phenotype is similar to that produced by mutations in the muscle-affecting genes, myo-3 (Waterston 1989), unc-45 (Venolia and Waterston 1990), and deb-1 (Barstead and Waterston 1991).

In this paper we present the complete sequence of the unc-52 gene and demonstrate that it is alternatively spliced. The largest predicted unc-52 protein is 720 KD and has three large domains with similarity to repeated sequences found in the LDL receptor, laminin, and cell adhesion molecules. The remarkable similarity of the unc-52 protein sequence to that of perlecan, the mammalian basement membrane heparan sulfate proteoglycan (Noonan et al. 1991; Kallunki and Tryggvason 1992, Murdoch et al. 1992), implies that the unc-52 gene products are proteoglycan components of the extracellular matrix. We have mapped the epitopes recognized by two monoclonal antibodies (mAbs) to a small region of the unc-52-encoded protein sequence. These mAbs recognize an extracellular component of all contractile tissues in C. elegans (Francis and Waterston 1991). The disorganized muscle phenotype seen in unc-52 mutants reveals that nematode perlecan has a role in anchoring the myofilament lattice to the cell membrane.

Results
Isolation of a transposon-induced unc-52 mutant

Mutator strains have been used successfully to isolate Tcl transposon-induced mutations in many C. elegans genes (Greenwald 1985; Moerman et al. 1986; for review see Herman and Shaw 1987). One strain commonly used for transposon tagging is RW7097, which contains the mut-6 IV mutator and has 30 copies of the 1.6-kb Tcl transposable element (Mori et al. 1988). We have obtained a spontaneous unc-52 mutant from this strain after visually screening for rare paralyzed hermaphrodites. The new unc-52 mutation, st196::Tcl, is a typical viable allele of unc-52 (Gilchrist and Moerman 1992). The disorganized body wall muscle structure of an adult st196 mutant hermaphrodite is shown in Figure 1. Although the overall muscle structure in this mutant animal is significantly disorganized, many of the muscle cells retain at least two recognizable sets of A and I bands (the body wall muscle cells in adult, wild-type animals contain 8–10 sets). This mutant strain also exhibits the reduced brood size and disrupted somatic gonad typical of this class of unc-52 mutants (Gilchrist and Moerman 1992).

Spontaneous germ-line revertants of the original unc-52(st196) strain were obtained at a frequency of 2 x 10⁻³. Removal of the mutator genetic background from this strain by outcrossing to the N2 strain of C. elegans greatly reduced the reversion frequency (<10⁻⁶) as expected for a Tcl transposon-induced mutation (Mori et al. 1988; Moerman and Waterston 1989).

Cloning the unc-52 gene

The Tcl transposable element associated with the st196 mutation was identified by probing DNA from an outcrossed st196 mutant strain and a coisogenic st196 revertant strain with Tcl sequences. A 6.4-kb Tcl containing BamHI fragment present in the mutant, but absent in the revertant, was identified (data not shown) and cloned. A 4.8-kb EcoRV–BamHI DNA fragment adjacent to the Tcl element was subcloned and used to probe Southern blots of BamHI-digested DNA from the mutant strain and five independent st196 revertant strains. The probe hybridized to a 4.8-kb BamHI fragment in all five st196 revertant strains and to the expected 6.4-kb fragment in the mutant strain (data not shown). Thus, reversion of the st196 mutant phenotype was absolutely correlated with the excision of this particular Tcl element. The Tcl insertion event always segregated with

---

**Figure 1.** Disorganized body wall muscle in unc-52 (st196) mutants. Polarized light micrographs of body wall muscle cells of adult hermaphrodites are shown. (a) Wild type; (b) unc-52(st196). The arrowhead in a points to a dense body. The arrowhead in b points to the edge of a muscle cell. Bar, 5 μm.
unc-52 encodes a basement membrane proteoglycan

The unc-52 mutant phenotype in genetic recombination experiments demonstrating that it was linked to unc-52.

To obtain the complete wild-type unc-52 gene a λ 1059 genomic library (Karn et al. 1980; Moerman et al. 1986) was screened with the unique 4.8-kb EcoRV–BamHI fragment, and seven overlapping clones covering a region of ~30 kb were isolated. Three of the clones were sent to A. Coulson [MRC Laboratory, Cambridge, England] for correlation with the physical map of the C. elegans genome (Coulson et al. 1986), and they were assigned to a “contig” ~1.5 Mb in length. More than 50 cDNAs were obtained after screening a λZAP cDNA library (Barstead and Waterston 1989) with fragments from various regions of the unc-52 gene. Our analysis of these clones identified 26 different cDNAs ranging in size from <1.0 to 4.7 kb. Six of these cDNAs were completely sequenced and another eight were partially sequenced. cDNAs covering the 5′ end of unc-52 were generated by coupling reverse transcription of RNA with PCR. These fragments were also sequenced but were not cloned.

The unc-52 gene is large and produces several transcripts

The unc-52 gene spans almost 15 kb of genomic DNA and consists of 26 exons and 25 introns (Fig. 2). A total of 16.5 kb of genomic DNA was sequenced, and 26 exons were identified by searching for open reading frames bounded by consensus intron–exon junctions (Emmons 1988) and by sequencing unc-52 cDNA clones. The first exon contains a putative initiator methionine and signal peptide. The presence of a signal peptide and the absence of open reading frames in the 1.8 kb of sequence upstream of exon 1 suggests that all of the unc-52 gene has been identified. The predicted amino acid sequence is presented in Figure 3.

It was apparent from the cDNA sequence data that the unc-52 gene produces more than one transcript. cDNA clones corresponding to three separate but overlapping transcripts were identified [DM#156, DM#150, and DM#135 in Fig. 2]. DM#156 is a partial cDNA 2.6 kb in length with a 360-bp 3′-untranslated region ending at a poly(A) addition signal located between exons 10 and 11 [Fig. 2]. The two other cDNAs used a different poly(A) addition signal located downstream of exon 26 [Fig. 2]. DM#150 was the largest cDNA obtained and included exons 21 and 22, but exons 16 and 17 had been removed. Because exon 22 encodes an in-frame stop codon, DM#150 has a 1.28-kb 3′-untranslated region. The DM#135 cDNA does not contain exons 21 and 22 and, therefore, must represent a third transcript. Northern hybridizations using probes from various regions of these cDNAs identified at least three mRNA species (Fig. 4). A unique probe from the 3′ end of DM#156 hybridized to a 4.0-kb transcript [data not shown], whereas a probe from the 3′ region common to both DM#150 and DM#135 hybridized to larger transcripts ~6.3 and 8.0 kb in size [Fig. 4, lane B]. As expected, a probe from the 5′ end of DM#156 identified all three bands [Fig. 4, lane A]. These data suggest that the unc-52 gene produces at least three polypeptides, one of which is much shorter than the others. Also, the shorter gene product appears to be less abundant than the larger gene products.

The 8.0-kb transcript detected by Northern hybridization is large enough to include the longest open reading frame of the unc-52 gene (all exons except 21 and 22), which encodes a 2464-amino-acid protein [minus the signal sequence] of ~270 kD (Fig. 3). A structural model of

---

**Figure 2.** Physical map of the unc-52 gene. A diagram of the unc-52 gene. The top line represents 16.5-kb of sequenced genomic DNA and shows the positions of 26 exons [solid boxes]. Partial cDNAs representing alternatively spliced transcripts are shown below. The hatched boxes indicate 3′-untranslated regions. The DM#135, DM#150, and DM#156 cDNA clones were obtained from a λZAP cDNA library, whereas the two unlabeled cDNAs were generated by reverse transcription and PCR. All of the cDNAs shown here have been sequenced. The complete nucleotide sequence of the unc-52 gene can be found in the GenBank data base under accession number L13458.
Figure 3. Predicted amino acid sequence of the unc-52 encoded polypeptides. The longest open reading frame of the unc-52 gene includes all exons except 10b, 21, and 22 and encodes the 2482 amino acids shown here. Exon 10b and exons 21 and 22 encode alternative carboxy-terminal sequences, which are shown below the main sequence. The putative signal sequence (amino acids 1-18) is indicated by a double underline. Potential amino-linked oligosaccharide attachment sites are underlined and marked with an asterisk (*) and the two RGD sites are underlined.

1474 GENES & DEVELOPMENT
components of the extracellular matrix. This has been confirmed by antibody staining [see below].

The putative start methionine of the unc-52 protein is followed by a signal peptide consisting of a 15-amino-acid hydrophobic core [Fig. 3]. A likely cleavage site for the removal of the leader peptide is between amino acids 18 and 19, which would make threonine the first amino acid of the mature protein [von Heijne 1986]. Domain I of the unc-52 protein is a short, 28-amino-acid region rich in aspartic acid residues. There is no similarity between the amino-terminal domain I of the mammalian perlecans and the first domain of the unc-52 protein, except that they are all rich in acidic amino acids and lack cysteine residues. In perlecan, domain I is much larger and contains three possible SGD attachment sites for the heparan sulfate side chains. These sites are not present in the unc-52 sequence; however, there are two serine residues in domain I that may be sites of attachment.

Domain II is composed of three repeated units of ~40 residues, each similar to a cysteine-rich sequence motif found in the LDL receptor family of proteins [Fig. 6; Yamamoto et al. 1984, Hobbs et al. 1990]. This region of the LDL receptor is required for ligand binding [Russell et al. 1989]. The amino acid identity between domain II in the unc-52 protein and the last three repeats of domain II in murine perlecan is 49% [Fig. 6].

Domain III is similar to the amino-terminal, short arms of the laminin A and B chains [Sasaki and Yamada 1987; Sasaki et al. 1987,1988]. This 733-amino-acid domain can be divided into two similar subdomains [Fig. 7]. The first consists of a cysteine-rich segment of 41 amino acids, followed by a 182-amino-acids cysteine-free globular region and then another cysteine-rich segment. The second subdomain is similar to the first subdomain except that it has three additional copies of the cysteine-rich motif after the globular region [Fig. 7]. Domain III in mammalian perlecan has three laminin-like subdomains, and the ratio of cysteine-rich repeats relative to globular domains in these subdomains is somewhat different from that observed in the unc-52 molecule [Noonan et al. 1991; Kallunki and Tryggvason 1992, Murdoch et al. 1992]. There is extensive similarity [43% amino acid identity] overall between the second subdomain of the unc-52 protein and the third subdomain of mouse perlecan. The cysteine-rich regions are more highly conserved between the mouse and the nematode proteins [53% amino acid identity] than the globular domains.

The carboxy-terminal domain IV of the unc-52 protein is comprised of a series of 14 repeated segments that occur in members of the immunoglobulin superfamily of proteins [Williams and Barclay 1988]. There are two additional immunoglobulin repeats flanking domain II in the amino-terminal half of the unc-52 protein for a total of 16 [Fig. 8]. These repeated elements vary in length between 82 and 100 amino acids and are characterized by two completely conserved cysteines, a completely conserved tyrosine, and several other highly conserved amino acids. Additional cysteine residues are found in immunoglobulin R2 [IgR2], IgR4, and IgR14b. The 16
Figure 5. Structure of the unc-52 polypeptides. (A) Diagrammatic representation of the largest predicted unc-52 protein showing the organization of the different domains. Domain I is a short, acidic region located at the amino terminus (solid area). Domain II contains three copies of a cysteine-rich sequence that occurs in the amino-terminal ligand-binding domain of the LDL receptor protein (open areas). Domain III contains sequences similar to the amino-terminal region of laminin. This domain contains two copies of a larger cysteine-free globular sequence (solid circles) and seven copies of a cysteine-rich element (horizontal stripes). Domain IV consists of 14 immunoglobulin C2-like repeats (shown hatched and numbered). There are also two of these repeats flanking domain II. (B) Diagrammatic representations of the partial polypeptides encoded by the DM#150 and DM#156 cDNAs [see Fig. 2]. The protein encoded by DM#156 does not contain domain IV, and the protein encoded by DM#150 has fewer immunoglobulin repeats in domain IV than the predicted full-length protein. The regions of the unc-52 proteins expressed as fusions in clones DM#178 and DM#180 [see Fig. 10] are indicated by the lines above the partial polypeptides.

immunoglobulin repeats found here are most similar to members of the C2 subfamily, which includes the neural cell adhesion molecule (N-CAM) [Barthels et al. 1987; Cunningham et al. 1987].

The mammalian perlecan core proteins contain 14 (mouse) and 21 (human) immunoglobulin repeats in domain IV but only one in the amino-terminal portion of the molecule between domains II and III. They lack the single immunoglobulin repeat between domain I and II that is present in the nematode protein [Noonan et al. 1991; Kallunki and Tryggvason 1992; Murdoch et al. 1992]. The amino acid identity between the domain IV repeats in the mouse and nematode proteins is ~23% overall, which is typical for members of this superfam-

Figure 6. The LDL receptor-like repeats of domain II. Alignment of the unc-52 domain II cysteine-rich repeats with three of four domain II repeats from mouse perlecan [Noonan et al. 1991] and three of seven repeats from the ligand-binding domain of the LDL receptor protein [Yamamoto et al. 1984]. The conserved amino acids are boxed and shaded.
unc-52 encodes a basement membrane proteoglycan

Figure 7. The laminin-like repeats of domain III. Alignment of the two domain III subdomains of the unc-52-encoded polypeptides with each other, with the third subdomain of domain III in mouse perlecan [Noonan et al. 1991], and with the laminin A chain domain IVb and the flanking cysteine repeats [Sasaki et al. 1988]. Each subdomain consists of one cysteine-rich repeat, a globular domain, and one, three, or four cysteine-rich repeats. The conserved amino acids are boxed and shaded.

illy. Both mammalian core proteins contain a fifth, carboxy-terminal domain that is similar to the laminin A chain domain G. This domain is not present in nematode perlecan.

Alternative splicing leads to different carboxy-terminal peptides

The three alternatively spliced cDNA clones identified here encode proteins with different carboxy-terminal amino acid sequences (Fig. 2 and 3). The polypeptide deduced from the sequence of DM#150 has 10 carboxy-terminal immunoglobulin repeats in domain IV [Fig. 5B]. It does not contain IgR8 and IgR9, which are encoded by exons 16 and 17, or IgR15 and IgR16, which are downstream of the stop codon in exon 22. The second half of the last immunoglobulin repeat (IgR14b) and the carboxy-terminal 15 amino acids are encoded by exons 21 and 22. The DM#135 cDNA does not contain exons 21 and 22 and would produce a protein that includes IgR15 and IgR16, which are downstream of the stop codon in exon 22. The second half of the last immunoglobulin repeat (IgR14b) and the carboxy-terminal 15 amino acids are encoded by exons 21 and 22. The DM#135 cDNA does not contain exons 21 and 22 and would produce a protein that includes IgR15 and IgR16 as well as the carboxy-terminal amino acids encoded by exon 26. The DM#156-encoded protein lacks the 14 carboxy-terminal immunoglobulin repeats of domain IV and contains 32 carboxy-terminal residues not found in the larger gene product [Fig. 3 and 5B]. Our analysis of PCR-generated cDNAs has revealed that there is also variability in the amino-terminal region of the unc-52 polypeptides. One of the transcripts does not contain exon 6, which encodes the first laminin-like subdomain of domain III (Fig. 2 and 5). The protein encoded by this transcript would have a shortened domain III.

Identification of the st196 insertion site

The Tc1 insertion site associated with the st196 mutation was identified by sequencing the cloned BamHI fragment from the st196::Tc1 mutant strain. The Tc1 element inserted between a TA dinucleotide in exon 18 [nucleotides 11395 and 11396 in the genomic sequence] 68 bp upstream of the splice donor site. This 279-bp exon encodes a single immunoglobulin repeat (IgR10) in domain IV of the unc-52 proteins. Thus, the st196 mutation would be expected to affect the two large unc-52 gene products but not the smaller one that lacks domain IV.

Two monoclonal antibodies recognize an unc-52-encoded antigen

Francis and Waterston [1991] described two mAbs, MH2 and MH3, that recognize a common antigen localized to the basement membrane of contractile tissues in C. elegans. In adult hermaphrodites, MH2 and MH3 show a general staining of the basement membrane adjacent to the body wall muscle cells, with more intense staining occurring at the positions of the muscle cell dense bodies and M-lines [Francis and Waterston 1991]. The antigen recognized by these mAbs is detected in wild-type embryos; however, it is not detected in embryos homozygous for st549, a lethal allele of unc-52 [Fig. 9].

The lethal unc-52 mutations produce a Pat (paralyzed, arrested elongation at twofold) terminal phenotype [B.D. Williams and R.H. Waterston, unpubl.]. Mutant embryos do not exhibit any of the movements that normally oc-
### Caenorhabditis unc-52 Immunoglobulin Repeats

| Domain | Repeat | Sequence |
|--------|--------|----------|
| 1      | TVF    | SADYS    |
| 2      | TVD    | LSGYN    |
| 3      | IMK    | NTEPV    |
| 4      | FYDE   | QTSEGGY  |
| 5      | FKDR   | TRQLOSGD |
| 6      | GOO    | SNPGNYD  |
| 7      | P O    | YC R2     |
| 8      | PI     | DGRV     |
| 9      | RG     | LPQV         |
| 10     | KV     | A SM         |
| 11     | AV     | V Q N         |
| 12     | KV     | I T    |
| 13     | VKL    | Q V H       |
| 14     | V     | P S         |
| 15     | H     | Q V G        |
| 16     | V     | Q V G        |

### Vertebrate Immunoglobulin Repeats

| Domain | Repeat | Sequence |
|--------|--------|----------|
| P1     | VYTP   | ASRSQETV |
| P2     | VQ     | YCTVCATVP | TPI | IMRN  |
| P3     | VRMT   | TSEGGQ  |
| P15    | VF     | EDQAX    |
| N1     | VI     | G K         |
| N5     | KG     | E Y         |

**Figure 8.** The immunoglobulin-like repeats of domain IV. Alignment of the 16 unc-52-encoded immunoglobulin superfamily repeats with each other. The most highly conserved amino acids are boxed and shaded, and a consensus sequence is indicated. Also shown is an alignment with the first (P1) and fifteenth (P15) immunoglobulin-like repeats from mouse perlecans (Noonan et al. 1991), and the first (N1) and fifth (N5) repeats from the mouse N-CAM sequence (Barthels et al. 1987).
unc-52 encodes a basement membrane proteoglycan

cle cells in the wild-type embryo (Fig. 9c) had well-organized A-bands that are seen most clearly in the dorsal quadrant (each embryonic quadrant has a double row of muscle cells). In contrast, the muscle cells in the mutant were obviously disorganized (Fig. 9d). This result suggests a failure in the earliest stages of myofilament lattice organization within the body wall muscle cells of st549 mutant embryos.

Although the st549 mutant embryos stained with MH27 and DM5.6, they failed to stain with MH3 (Fig. 9b) or with MH2 (data not shown). Thus, this mutant does not accumulate the MH2/MH3 antigens. MH3 staining was observed in the wild-type embryo at positions corresponding to the body wall muscle quadrants: Two of the four quadrants are visible in the plane of focus, one ventral and one dorsal (Fig. 9a). The failure of st549 mutant embryos to stain with MH2 and MH3 suggests that these mAbs recognize an unc-52-encoded gene product.

The region of the unc-52 protein sequence recognized by MH2 and MH3 has been identified by expressing specific unc-52 peptides as fusions with glutathione S-transferase (GST). Restriction fragments from unc-52 cDNAs were subcloned into the appropriate pGEX vector (Smith and Johnson 1988), and two constructs producing inducible fusion products were obtained (Fig. 10). The DM#178 clone contained a 1.47-kb EcoRI fragment encoding 490 amino acids from domain IV and directed the synthesis of an 80-kD fusion protein (Figs. 5B and 10). DM#180 contained a 1.0-kb EcoRI fragment encoding most of the second subdomain from domain III (Fig. 5B). The protein synthesized by this clone was ~65 kD (Fig. 10). The results of Western hybridizations with MH3 and an anti-GST antibody are shown in Fig. 10. The anti-GST sera identified both fusion products as well as the 26-kD GST protein produced by the pGEX vector. In contrast, MH3 specifically recognized the unc-52-encoded portion during embryogenesis and fail to elongate past the “twofold” length. They hatch as misshapen L1 larvae but die soon after. Figure 9 shows the results obtained when wild-type and unc-52(st549) mutant embryos were stained with the mAbs MH3, DM5.6, and MH27. No antibody staining with MH3 was detected in the st549 mutant embryos, whereas staining with MH27 and DM5.6 was detected. MH27 stains belt desmosomes at hypodermal cell junctions [Waterston 1988], and was used as a control for antibody staining and to define the shape of the embryos. The shape and arrangement of the hypodermal cells in the mutant embryos were normal [Fig. 9b,d]. The DM5.6 mAb recognizes myosin isoform A, which is specific to the body wall muscle cells [Miller et al. 1983]. The staining pattern of this mAb in the unc-52 mutant embryo revealed defects in the organization of the developing muscle cells. The body wall muscle cells in the wild-type embryo (Fig. 9c) had well-organized A-bands that are seen most clearly in the dorsal quadrant (each embryonic quadrant has a double row of muscle cells). In contrast, the muscle cells in the mutant were obviously disorganized (Fig. 9d). This result suggests a failure in the earliest stages of myofilament lattice organization within the body wall muscle cells of st549 mutant embryos.

Although the st549 mutant embryos stained with MH27 and DM5.6, they failed to stain with MH3 (Fig. 9b) or with MH2 (data not shown). Thus, this mutant does not accumulate the MH2/MH3 antigens. MH3 staining was observed in the wild-type embryo at positions corresponding to the body wall muscle quadrants: Two of the four quadrants are visible in the plane of focus, one ventral and one dorsal (Fig. 9a). The failure of st549 mutant embryos to stain with MH2 and MH3 suggests that these mAbs recognize an unc-52-encoded gene product.
Rogalski et al.

of the DM#178 fusion product. A similar result was obtained with MH2 (data not shown). Thus, these mAbs recognize epitopes in domain IV of at least one of the unc-52-encoded polypeptides.

Discussion

The unc-52 gene of *C. elegans* encodes a nematode homolog of perlecan, the most abundant mammalian basement membrane heparan sulfate proteoglycan [Noonan et al. 1991; Kallunki and Tryggvason 1992; Murdoch et al. 1992]. Perlecan is present in all basement membranes and plays a major role in their assembly [Martin and Timpl 1987]. It binds to laminin, collagen type IV, and fibronectin [Laurie et al. 1986; Heremans et al. 1990] and has been implicated in cell attachment through interactions with cell-surface receptors [Singer et al. 1987; Hayashi et al. 1992]. The longest predicted open reading frame of the unc-52 gene encodes a 2482-amino-acid protein with a signal peptide and four recognizable domains. The first domain is very short, only 28 amino acids, and is unique to this protein. However, the three remaining domains contain sequences that are found in other proteins, including the LDL receptor (domain II), laminin (domain III), and N-cam (domain IV). The similarity with the mouse and human perlecan core proteins occurs throughout domains II, III, and IV, and it is possible to align all but the first 126 amino acids of the nematode protein with the mouse proteoglycan by introducing a gap in domain II to accommodate the extra laminin-like repeats in perlecan. Although the amino acid sequences of the nematode and mammalian proteins are remarkably similar, they differ in their amino- and carboxy-terminal regions. Most notably, mammalian perlecan contains a fifth carboxy-terminal domain similar to the laminin A chain domain G that is absent in the nematode protein [Noonan et al. 1991; Kallunki and Tryggvason 1992; Murdoch et al. 1992].

The molecular analysis described here has revealed the physical organization of the unc-52 gene, which spans almost 15-kb of DNA. We have identified three transcripts that encode proteins with different carboxy-terminal sequences. These alternative transcripts are generated by the utilization of two different poly(A) addition sites located 8.5-kb apart and by the differential splicing of exons at the 3' end of the gene. Interestingly, all of the exons that encode domain IV are located between the two poly(A) addition sites. The two larger transcripts encode proteins containing all or part of this domain, whereas the smaller transcript encodes a shortened polypeptide that completely lacks domain IV. It will be necessary to determine the distribution of these proteins in vivo before we can fully understand their function.

The exon–intron boundaries in the unc-52 gene are well placed for generating multiple domain combinations by alternative splicing. Exons 1–4, 5–8, and 13–18 are interrupted such that any donor site in either region can be spliced to any acceptor site within the same region and maintain the reading frame. Our results demonstrate that unc-52-encoded proteins with variable immunoglobulin and laminin-like repeats can be generated by alternative splicing. The mouse and human perlecan structures were deduced from sequencing cDNAs, and these proteins show immunoglobulin repeat variability in domain IV. It has been suggested that this variability is the result of alternative splicing [Kallunki and Tryggvason 1992; Noonan and Hassell 1993].

Little is known about the biochemistry of the unc-52 gene products. Like perlecan, the unc-52 proteins have several potential amino-linked glycosylation sites that occur mainly in domain IV. However, the potential heparan sulfate attachment sites that have been identified in the mammalian molecules [Noonan et al. 1991; Kallunki and Tryggvason 1992; Murdoch et al. 1992] are not present in the nematode sequence. Thus, if the unc-52 products contain glycosaminoglycan side chains, they must use different attachment sites. The presence of a signal peptide at the amino terminus and the absence of a transmembrane domain imply that the unc-52 gene products are extracellular proteins. Our finding that the MH2 and MH3 mAbs [Francis and Waterston 1991] recognize an unc-52-encoded antigen confirms that at least one of the unc-52 proteins is a basement membrane component. The epitope recognized by MH2 and MH3 is in a region of domain IV that is present in one, and possibly both, of the larger unc-52-encoded proteins. This protein (or proteins) is synthesized early in embryogenesis and is present in the basement membranes of all contractile tissues in adult hermaphrodites [Francis and Waterston 1991]. MH2 and MH3 detect a 200-kD polypeptide on Western blots of protein fractions from wild-type nematodes [Francis and Waterston 1991]. Because we do not know the exact amino acid composition of any of the unc-52 gene products, it is not clear whether this 200-kD polypeptide is an intact unc-52 protein or a proteolytic fragment.

The disorganized body wall muscle phenotypes produced by the lethal and viable unc-52 alleles imply a role for this gene in myofilament assembly and/or attachment of the myofilament lattice to the cell membrane. The st549 allele of unc-52 appears to be a complete loss-of-function mutation for at least one of the unc-52 gene products. As observed in st549 mutant embryos, lack of unc-52 gene function leads to defects in the earliest stages of myofilament lattice formation in the developing body wall muscle cells. Because the viable alleles result in a milder phenotype than st549, they cannot eliminate unc-52 gene function completely. Animals homozygous for these mutations exhibit defects in myofilament assembly and sarcomere growth but not until the third or fourth larval stage [MacKenzie et al. 1978; Waterston et al. 1980]. These mutants also have abnormal somatic gonads, indicating that unc-52 gene function is not limited to the body wall muscle cells [Gilchrist and Moerman 1992]. This is consistent with the antibody results of Francis and Waterston [1991] that reveal the presence of an Unc-52 protein in the basement membrane of the myoepithelial sheath of the gonad.

The paralyzed phenotype produced by the st196 mutation is the result of a Tc1 transposon insertion in the...
exon encoding IgR10 of Domain IV. This mutation should only affect the two larger gene products and may result in the production of partially functional truncated proteins. Two additional mutations, e444 and e998, which also impart a paralyzed phenotype, have been localized to the same immunoglobulin repeat as st196 [E.J. Gilchrist and D.G. Moerman, unpubl.]. These data suggest that the disruption of the body wall muscles observed in the viable unc-52 mutants may be attributable to defects in domain IV of one, or both, of the larger polypeptides. If this is the case, then the lethal phenotype might result only when the smaller protein is also affected. It will be necessary to identify the sequence alterations in the lethal alleles to test this hypothesis.

Recent findings by Hayashi et al. [1992] and Rescan et al. [1993] suggest that perlecan may interact with integrin receptors. Integrins have been shown to play a pivotal role in the terminal stages of myogenic differentiation [Menko and Boettiger 1987], and studies with lethal myosporid mutants in Drosophila have determined that β-integrin is essential for the formation of stable Z-bands in muscle [Volk et al. 1990]. In the nematode, perlecan is concentrated over muscle dense bodies and M-lines [Francis and Waterston 1991], which are associated with β-integrin [Francis and Waterston 1985; S. Gettner and L. Reichardt, pers. commun.]. Thus, perlecan may be an extracellular anchor for integrin receptors in muscle.

**Materials and methods**

**Nematode strains and genetic procedures**

Stock maintenance and strain growth for protein and nucleic acid procedures were as described in Brenner [1974] and Sulston and Brenner [1974]. Nematode strains used in this study were derived from the N2 strain of C. elegans var. Bristol [Brenner 1974], or RW7097 [Mori et al. 1988].

**Cloning procedures**

Gel electrophoresis and Southern blotting and hybridization conditions followed standard procedures (Sambrook et al. 1989). To clone the Tcl-containing fragment, BamHI-digested DNA was separated on a preparative agarose gel and the 6.4-kb region was purified and cloned into λ 2001 [Karn et al. 1984]. This mini-library was screened with a Tcl probe following the method of Benton and Davis [1977], and a clone containing the appropriate fragment was isolated. The 6.4-kb insert was then subcloned into Bluescribe M13 + to become plasmid RW#DM11. The transposon was removed by cutting with EcoRV, leaving a 4.8-kb EcoRV-BamHI fragment to use as a unique probe [RW#DM12].

This 4.8-kb BamHI fragment was used to obtain several overlapping clones from a λ 1059 N2 genomic library [Karn et al. 1980; Moerman et al. 1986] and two clones [B. Barstead, unpub.] from a AZAP mixed stage RNA cDNA library [Barstead and Waterston 1989]. Purified fragments from the two cDNA clones, DM#101 and DM#102, were used singly or in combination to isolate an additional 50 cDNA clones [see Results]. Methods for phage growth, library screening, and DNA isolation from phage and plasmid clones have been described [Karn et al. 1980, Short et al. 1988, Sambrook et al. 1989].

**Sequence analysis**

Phage DNA prepared from RW#DM42 and RW#DM44 was digested with BamHI, EcoRI, and HindIII and size fractionated on LMP agarose gels. Fourteen overlapping restriction fragments were extracted from the gels, subcloned into Bluescribe M13 – and transformed into Escherichia coli strain XL1-Blue [Bullock et al. 1987] for sequencing. The Tc1 insertion site was identified by sequencing the appropriate region of RW#DM12. This clone was obtained by digesting RW#DM11 with EcoRV to remove most of the Tc1 insert and then religating the plasmid. Recombinant φBluescript SK(–) plasmids were excised from the λZAP phage using the protocol supplied by Stratagene (La Jolla, CA) and propagated in XL1-Blue cells. Six cDNA clones were completely sequenced and another eight were partially sequenced. Double-stranded DNA templates were sequenced by the dideoxy chain termination method [Sanger et al. 1977] using either the T7 Sequencing kit (Pharmacia) or Sequenase version 2.0 (U.S. Biochemical Corp.). Universal primers were used to obtain initial sequence data, and oligonucleotides corresponding to previously determined sequence were then synthesized to use as primers. Primers were either made by us on an Applied Biosystems PCR-MATE or by the University of British Columbia Nucleic Acid Facility. Using this strategy, both strands of DNA were sequenced except for some of the larger introns, where only one strand was sequenced. Sequence data were assembled and analyzed using the Staden [1987] programs [Amersham], and searches of the EMBL/ SWISS-PROT data base were performed with the FASTA program of Pearson and Lipman [1988].

**Reverse transcription and PCR amplification**

Reverse transcription was performed with 5 μg of total RNA, 25 pmoles of the appropriate downstream primer, 1 mm dNTPs (Perkin Elmer), 10 mm DTT, 20 units of RNasin (Promega) and 5 units of reverse transcriptase (Promega) in 10 μl of 1 x PCR buffer [Promega; 15 mm MgCl2]. The reactions were incubated for 60 min at 42°C and then for 15 min at 99°C. To amplify the cDNA products, 25 pmoles of the forward primer, 1.5 μl of 10 x PCR buffer, 2.5 units of Taq polymerase [Promega], and dH2O were added to the reverse transcription reactions for a final volume of 25 μl. The mixture was amplified using a Perkin-Elmer Cetus DNA thermocycler. Amplification conditions were 30 sec at 95°C, 60 sec at 55°C, and 90 sec at 72°C for 35 cycles. This was followed by 5 min at 72°C. The PCR fragments were sequenced directly using the BRL dsDNA Cycle Sequencing System.

**Northern blot analysis**

RNA for Northern blot analysis was isolated from a mixed population of N2 hermaphrodites. Nematodes were disrupted by grinding in liquid nitrogen in three volumes of 5 m guanidine isothiocyanate, 25 mm Tris-HCl [pH 7.5], 0.5% N-lauroylsarcosine, and 0.1% β-mercaptoethanol. The lysate was extracted three times with phenol-chloroform [1 : 1], after which the RNA was precipitated with one volume of ethanol. The pellet was redissolved in 20 mm EDTA and RNA was reprecipitated by adding one volume of 5 m LiCl and storing on ice for 2 hr. After a final ethanol precipitation, the RNA was dissolved in water and 15- to 20-μg samples were resolved on a 1.2% agarose gel containing 6% formaldehyde, transferred by blotting to a nylon membrane [Nytran, Schleicher & Schuell], fixed by baking at 80°C, and hybridized at 42°C in 50% formamide with a 32P-
labeled probe. Stringent washes were carried out in 0.1 x SSPE at 42°C.

Immunofluorescence techniques and polarized light microscopy

Immunofluorescence staining was performed on populations of embryos from the N2 (wild-type) and RW6010 strains. RW6010 hermaphrodites are homozygous for the unc-52(st549) lethal mutation and carry mnDp34 [Herman et al. 1979], a free duplication with a wild-type copy of the unc-52 gene. Hermaphrodites with this genotype segregate wild-type progeny. Populations of embryos were fixed and stained using methods described previously [Barstead and Waterston 1991]. Briefly, embryos were released from adult hermaphrodites by treatment with a mixture of hypochlorite and KOH, fixed in 3% formaldehyde followed by 100% methanol, and stained with mixtures of the appropriate mAbs. The secondary antibody was DTAF-labeled goat anti-mouse IgG [Chemicon International, Inc. Temecula, CA]. Embryos positive for MH2 and negative for MH3 were observed only in the RW6010 strain and never in the N2 strain. Similarly, the disorganized myosin A staining pattern characteristic of unc-52(st549) homozygotes was never seen in populations of wild-type embryos. Specimens were viewed with a Zeiss Universal microscope [Carl Zeiss, Inc., Thornwood, NY] equipped with epifluorescent illumination. Photomicrographs were prepared as described in Francis and Waterston [1985]. Liveworms were viewed under polarized light microscopy following the procedure described in Waterston et al. [1980]. Photomicrographs were prepared using Kodak TMAX400 film on a Zeiss Axioshot Photomicroscope [Carl Zeiss D-7082 Oberkochen].

Fusion protein analysis

EcoRI restriction fragments from the DM#101 and DM#150 cDNAs were extracted from low-melting-point agarose gels and subcloned into the appropriate pGex expression vector [Smith and Johnson 1988]. The various clones were transformed into E. coli strain DH5α, and synthesis of the GST–unc-52 fusion proteins was induced with IPTG (Sigma). Fusion proteins were affinity purified using the mass screening protocol described by Smith and Johnson [1988]. Purified fusions were separated by denaturing gel electrophoresis and transferred to 0.45-μm-pore Hybond-ECL nitrocellulose filters [Amersham] by the electrophoretic method of Towbin et al. [1979]. Transfer was for 15–20 min at 15 V in a Trans-Blot SD Electroblot Transfer Cell [Bio-Rad Laboratories]. Filters were blocked overnight at 4°C in 5% milk powder/TBS-T [TBS-T: 10 mM Tris-HCl [pH 7.4], 0.9% NaCl, 0.1% Tween 20] and then incubated with the primary antibody [MH2, MH3, or anti-GST]. The secondary antibody was either horseradish peroxidase-conjugated goat anti-mouse IgG or goat anti-rabbit IgG [Amersham]. Each incubation was in 0.5% milk powder/TBS-T, followed by several washes in TBS-T and TBS. The final wash was followed by incubation for 1 min with Enhanced chemiluminescence detection reagents [Amersham].

Acknowledgments

We thank Bob Waterston, Bob Barstead, and Linda Matsuuchi for invaluable comments on the manuscript; Bob Barstead for the gift of the cDNA library and for isolating DM#101 and DM#102; David Miller for mAb DM5.6; Ross Francis for mAbs MH2, MH3, and MH27; Dennis Dixon for providing the anti-GST sera and for advice on ECL detection; Steve Rioux for assistance in identifying the initial st196 Tc1 dimorphism; Terry Snutch for assistance in making oligonucleotides; Mary Gilbert for advice on DNA sequencing; Erin Gilchrist for permission to cite unpublished results and for photographic advice and assistance; Alistair Blachford and the Biosciences Data Center for computer facilities and advice; and David Baillie and Terry Starr for assistance with the data base searches. D.G.M. gratefully acknowledges Bob Waterston, in whose laboratory this project was initiated, for his constant support and encouragement.

Some nematode strains used in this work were provided by the Caenorhabditis Genetics Center, which is funded by the National Institutes of Health National Center for Research Resources. This work was supported by grants from the Medical Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and the Muscular Dystrophy Association of Canada to D.G.M.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Barstead, R.J. and R.H. Waterston. 1989. The basal component of the nematode dense-body is vinculin. J. Biol. Chem. 264: 10177–10185.
  ———. 1991. Vinculin is essential for muscle function in the nematode. J. Cell Biol. 114: 715–724.
Barthels, D., M.J. Santoni, W. Wille, C. Ruppert, J.C. Chaix, M.R. Hirsch, J.C. Fontecilla-Camps, and C. Goridis. 1987. Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000 polypeptide without a membrane-spanning region. EMBO J. 6: 907–914.
Benton, W.D. and R.W. Davis. 1977. Screening lambda gt recombinant clones by hybridization to single plaques in situ. Science 196: 180–182.
Brenner, S. 1974. The genetics of Caenorhabditis elegans. Genetics 77: 71–94.
Bullock, W.O., J.M. Fernandez, and J.M. Short. 1987. XLI-Blue: A high efficiency plasmid transforming reca Escherichia coli strain with beta-galactosidase selection. BioTechniques 5: 376–379.
Coulson, A., J. Sulston, S. Brenner, and J. Karn. 1986. Towards a physical map of the genome of the nematode Caenorhabditis elegans. Proc. Natl. Acad. Sci. 83: 7821–7825.
Cunningham, B.A., J.J. Hemperly, B.A. Murray, E.A. Prediger, B. Brackenbury, and G.M. Edelman. 1987. Neural cell adhesion molecule structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. Science 236: 799–806.
Emmons, S.W. 1988. The genome. In The nematode Caenorhabditis elegans [ed. W.B. Wood], pp. 47–80. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
Francis, R. and R.H. Waterston. 1985. Muscle organization in Caenorhabditis elegans: Localization of proteins implicated in thin filament attachment and 1 band organization. J. Cell Biol. 101: 1532–1549.
  ———. 1991. Muscle cell attachment in Caenorhabditis elegans. J. Cell Biol. 114: 465–479.
Gilchrist, E.J. and D.G. Moerman. 1992. Mutations in the sup-38 gene of Caenorhabditis elegans suppress muscle-attachment defects in unc-52 mutants. Genetics 132: 431–442.
Greenwald, I. 1985. *lin-12*, a nematode homeotic gene, is homolo-
gous to a set of mammalian proteins that includes epi-
dermal growth factor. *Cell* 43: 583–590.

Hayashi, K., J. Madri, and P. Yurchenco. 1992. Endothelial cells
interact with the core protein of basement membrane perle-
can through B1 and B3 integrins: An adhesion modulated by
glycosaminoglycan. *J. Biol. Chem.* 267: 945–959.

Herman, A., B. DeCock, J.J. Cassiman, H. Van den Berghe, and
G. David. 1990. The core protein of the matrix-associated hepatan
sulfate proteoglycan binds to fibronectin. *J. Biol. Chem.* 265: 8716–8724.

Herman, R.K. and J. Shaw. 1987. The transposable genetic ele-
ment Tcl in the nematode *Caenorhabditis elegans*. *Trends Genet.*
3: 222–225.

Herman, R.K., J.E. Madl, and C.K. Kari. 1979. Duplications in
*Caenorhabditis elegans*. *Genetics* 92: 419–435.

Hobbs, H., D.W. Russell, M.S. Brown, and J.L. Goldstein. 1990.
The LDL receptor locus in familial hypercholesterolemia: Mu-
tational analysis of a membrane protein. *Annu. Rev. Genet.*
24: 133–170.

Kallunki, P. and K. Tryggvason. 1992. Human basement mem-
brane heparan sulfate proteoglycan core protein: A 467-kD
protein containing multiple domains resembling elements of
the low density lipoprotein receptor, laminin, neural cell ad-
hesion molecules and epidermal growth factor. *J. Cell Biol.*
116: 559–571.

Karn, J., S. Brenner, L. Barnett, and G. Cesareni. 1980. Novel
bacteriophage lambda cloning vector. *Proc. Natl. Acad. Sci.*
77: 5172–5176.

Karn, J., H.W.D. Matthes, M.J. Gait, and S. Brenner. 1984. A new
selective phage cloning vector, lambda 2001, with sites for
XbaI, BamHI, HindIII, EcoRI, SstI and XhoI.

Laurie, G.W., J.T. Bing, H.K. Kleinman, J.R. Hassell, M. Aumail-
ley, G.R. Martin, and R.J. Feldmann. 1986. Localization of
binding sites for laminin, heparan sulfate proteoglycan and
fibronectin on basement membrane (type IV) collagen. *J. Mol.
Biol.* 189: 205–216.

MacKenzie, J.M., R.L. Garcea, J.M. Zengel, and H.F. Epstein.
1978. Muscle development in *Caenorhabditis elegans*: Mut-
tants exhibiting retarded sarcomere construction. *Cell* 15:
751–763.

Martin, G.R. and R. Timpl. 1987. Laminin and other basement
membrane components. *Annu. Rev. Cell Biol.* 3: 57–85.

Menko, A.S. and D. Boettiger. 1987. Occupation of the extra-
cellular matrix receptor, integrin, as a control point for myo-
genic differentiation. *Cell* 51: 51–57.

Miller, D.M., I. Ortiz, G.C. Berliner, and H.F. Epstein. 1983.
Differential localization of two myosins within nematode
thick filaments. *Cell* 34: 477–490.

Moerman, D.G. and R.H. Waterston. 1989. Mobile elements in
*Caenorhabditis elegans* and other nematodes. In *Mobile
DNA* (ed. D.E. Berg and M.M. Howe), pp. 537–556. Ameri-
can Society for Microbiology, Washington, D.C.

Moerman, D.G., G.M. Benian, and R.H. Waterston. 1986. Mo-
lecular cloning of the muscle gene unc-22 in *Caenorhabditis
elegans* by Tcl transposon tagging. *Proc. Natl. Acad. Sci.*
83: 2579–2583.

Mori, I., D.G. Moerman, and R.H. Waterston. 1988. Analysis of
a mutator activity necessary for germline transposition and
exclusion of Tcl transposable elements in *Caenorhabditis elegans*. *Genetics* 120: 397–407.

Murdoch, A.D., G.R. Dodge, I. Cohen, R.S. Tuan, and R.V.
Iozzo. 1992. Primary structure of the human heparan sulfate
proteoglycan from basement membrane [HSPG2/perlecan].
*J. Biol. Chem.* 267: 8544–8557.

Noonan, D.M. and J.R. Hassell. 1993. Perlecan, the large low
density proteoglycan of basement membranes: Structure and
variant forms. *Kidney Int.* 43: 53–60.

Noonan, D.M., A. Fulle, P. Valente, S. Cai, E. Horigan, M.
Sasaki, Y. Yamada, and J.R. Hassell. 1991. The complete
sequence of perlecan, a basement membrane heparan sulfate
proteoglycan, reveals extensive similarity with laminin A
chain, low density lipoprotein-receptor, and the neural cell
adhesion molecule. *J. Biol. Chem.* 266: 22939–22947.

Pearson, W.R. and D.J. Lipman. 1988. Improved tools for bi-
ological sequence comparison. *Proc. Natl. Acad. Sci.* 85:
2444–2448.

Rescan, P.-Y., O. Lorcal, J.R. Hassell, Y. Yamada, A. Guillouz-
ou, and B. Clement. 1993. Distribution and origin of the base-
ment membrane component perlecan in rat liver and pri-
mary hepatocyte culture. *Am. J. Pathol.* 142: 199–208.

Russell, D.W., M.S. Brown, and J.L. Goldstein. 1989. Different
combinations of cysteine-rich repeats mediate binding of
low density lipoprotein receptor to two different proteins. *J.
Biol. Chem.* 264: 1682–1688.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular
cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor
Press, Cold Spring Harbor, New York.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequenc-
ing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*
74: 5463–5467.

Sasaki, M. and Y. Yamada. 1987. The laminin B2 chain has a
multidomain structure homologous to the B1 chain. *J. Biol.
Chem.* 262: 17111–17117.

Sasaki, M., S. Kato, K. Kohno, G.R. Martin, and Y. Yamada.
1987. Sequence of the cDNA encoding the laminin B1 chain
reveals a multidomain protein containing cysteine-rich re-
peats. *Proc. Natl. Acad. Sci.* 84: 935–939.

Sasaki, M., H. Kleinman, H. Huber, R. Deutzmann, and Y. Ya-
mada. 1988. Laminin, a multidomain protein. The A chain
has a unique globular domain and homology with the base-
ment membrane proteoglycan and the laminin B chains. *J.
Biol. Chem.* 263: 16536–16544.

Short, J.M., J.M. Fernandez, J.A. Sorge, and W.D. Huse. 1988.
Lambda ZAP: A bacteriophage lambda expression vector
with in vivo excision properties. *Nucleic Acids Res.* 16:
7583–7600.

Singer, L.I., S. Solomon, D.W. Kawka, and J.R. Hassell. 1987.
Extracellular matrix fibers containing fibronectin and base-
ment membrane heparan sulfate proteoglycan coalign with
focal contacts and microfilament bundles in stationary fi-
broblasts. *Exp. Cell Res.* 173: 558–571.

Smith, D.B. and K.S. Johnson. 1988. Single-step purification of
polypeptides expressed in *Escherichia coli* as fusions with
glutathione S-transferase. *Gene* 67: 31–40.

Staden, R. 1987. Computer handling of DNA sequencing pro-
jects. In *Nucleic acid and protein sequence analysis: A prac-
tical approach* (ed. M.J. Bishop and C.J. Rawlings), pp. 173–
217. IRL Press, Oxford, UK.

Sulston, J.E. and S. Brenner. 1974. The DNA of *Caenorhabditis
elegans*. *Genetics* 77: 95–104.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic
transfer of proteins from polyacrylamide gels to nitrocel-
lose sheets: Procedure and some applications. *Proc. Natl.
Acad. Sci.* 76: 4350–4354.

Venolia, L. and R.H. Waterston. 1990. The unc-45 gene of *Ca-
enorhabditis elegans* is an essential muscle-affecting gene
with maternal expression. *Genetics* 126: 345–353.

Volk, T., L.I. Fessler, and J.H. Fessler. 1990. A role for integrin
in the formation of sarcomeric cytoarchitecture. *Cell* 63:
525–536.
von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acid Res.* 14: 4683–4690.
Waterston, R.H. 1988. Muscle. In *The nematode Caenorhabditis elegans* (ed. W.B. Wood), pp. 281–335. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
———. 1989. The minor myosin heavy chain, mhcA, of *Caenorhabditis elegans* is necessary for the initiation of thick filament assembly. *EMBO J.* 8: 3429–3436.
Waterston, R.H., J.N. Thomson, and S. Brenner. 1980. Mutants with altered muscle structure in *Caenorhabditis elegans*. *Dev. Biol.* 77: 279–302.
Williams, A.F. and A.N. Barclay. 1988. The immunoglobulin super-family: Domains for cell surface recognition. *Annu. Rev. Immunol.* 6: 381–405.
Yamamoto, T., C.G. Davis, M.S. Brown, W.J. Schneider, M.L. Casey, J. Goldstein, and D. Russell. 1984. The human LDL-receptor: A cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* 39: 27–38.
Products of the unc-52 gene in Caenorhabditis elegans are homologous to the core protein of the mammalian basement membrane heparan sulfate proteoglycan.

T M Rogalski, B D Williams, G P Mullen, et al.

*Genes Dev.* 1993, 7:
Access the most recent version at doi:10.1101/gad.7.8.1471

References

This article cites 49 articles, 27 of which can be accessed free at: [http://genesdev.cshlp.org/content/7/8/1471.full.html#ref-list-1](http://genesdev.cshlp.org/content/7/8/1471.full.html#ref-list-1)

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](http://genesdev.cshlp.org/content/7/8/1471.full.html#ref-list-1).