Identification of novel chondroitin proteoglycans in Caenorhabditis elegans: embryonic cell division depends on CPG-1 and CPG-2

Sara K. Olson,1,2 Joseph R. Bishop,1 John R. Yates,4 Karen Oegema,1,3 and Jeffrey D. Esko1

1Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, 2Biomedical Sciences Graduate Program, School of Medicine, and 3Ludwig Institute for Cancer Research, University of California, San Diego, La Jolla, CA 92037
4Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037

Vertebrates express at least 25 chondroitin sulfate proteoglycans (CSPGs), distinguished by the primary sequence of their core protein (Olson and Esko, 2004). Major groups include the aggrecan family of large, highly glycosylated proteoglycans present in connective tissues, the small leucine-rich proteoglycans, collagen $\alpha_2IX$, basement membrane proteoglycans (e.g., leprecan and perlecan), membrane bound proteoglycans (e.g., CD44, NG2, and phosphacan), and hybrid proteoglycans containing both chondroitin sulfate and heparan sulfate (syndecans). Secreted CSPGs create a hydrated matrix allowing for tissue expansion, and in cartilage they confer the ability to absorb compressive loading. In other tissues, they bind and help organize fibrillar collagens and can affect growth factor signaling pathways. Mutations affecting the core proteins or the enzymes involved in the assembly of the chondroitin chains result in developmental abnormalities in skin, cartilage, bone, tendon, eyes, brain, and the microvasculature (Iozzo, 1998; Hassell et al., 2002; Schwartz and Domowicz, 2002; Ozerdem and Stallcup, 2002; Schwartz and Domowicz, 2004; Goldberg et al., 2005; Marneros and Olsen, 2005; Niisato et al., 2005).

Vertebrates produce multiple chondroitin sulfate proteoglycans that play important roles in development and tissue mechanics. In the nematode Caenorhabditis elegans, the chondroitin chains lack sulfate but nevertheless play essential roles in embryonic development and vulval morphogenesis. However, assignment of these functions to specific proteoglycans has been limited by the lack of identified core proteins. We used a combination of biochemical purification, Western blotting, and mass spectrometry to identify nine C. elegans chondroitin proteoglycan core proteins, none of which have homologues in vertebrates or other invertebrates such as Drosophila melanogaster or Hydra vulgaris. CPG-1/CEJ-1 and CPG-2 are expressed during embryonic development and bind chitin, suggesting a structural role in the egg.

RNA interference (RNAi) depletion of individual CSPGs had no effect on embryonic viability, but simultaneous depletion of CPG-1/CEJ-1 and CPG-2 resulted in multinucleated single-cell embryos. This embryonic lethality phenocopies RNAi depletion of the SQV-5 chondroitin synthase, suggesting that chondroitin chains on these two proteoglycans are required for cytokinesis.

Vertebrates express at least 25 chondroitin sulfate proteoglycans (CSPGs), distinguished by the primary sequence of their core protein (Olson and Esko, 2004). Major groups include the aggrecan family of large, highly glycosylated proteoglycans present in connective tissues, the small leucine-rich proteoglycans, collagen $\alpha_2IX$, basement membrane proteoglycans (e.g., leprecan and perlecan), membrane bound proteoglycans (e.g., CD44, NG2, and phosphacan), and hybrid proteoglycans containing both chondroitin sulfate and heparan sulfate (syndecans). Secreted CSPGs create a hydrated matrix allowing for tissue expansion, and in cartilage they confer the ability to absorb compressive loading. In other tissues, they bind and help organize fibrillar collagens and can affect growth factor signaling pathways. Mutations affecting the core proteins or the enzymes involved in the assembly of the chondroitin chains result in developmental abnormalities in skin, cartilage, bone, tendon, eyes, brain, and the microvasculature (Iozzo, 1998; Hassell et al., 2002; Schwartz and Domowicz, 2002; Ozerdem and Stallcup, 2002; Schwartz and Domowicz, 2004; Goldberg et al., 2005; Marneros and Olsen, 2005; Niisato et al., 2005).

Each chondroitin proteoglycan (CPG) consists of a protein core and one or more covalently attached chondroitin chains. Assembly occurs in a stepwise manner, starting in the endoplasmic reticulum with translation of core proteins on membrane bound ribosomes and transfer of xylose to specific serine residues. Xylosylation exhibits specificity, occurring only at serine residues that have an adjacent glycine to the COOH-terminal side and one or more flanking acidic residues usually within eight amino acids (Esko and Selleck, 2002). After xylosylation, synthesis of a tetrasaccharide primer ($-\text{GlcA}\beta_3\text{Gal}\beta_3\text{Gal}\beta_4\text{Xyl}\beta-O-\text{Ser}$) takes place in the Golgi, followed by the polymerization of the chain by the alternating addition of N-acetylgalactosamine (GalNAc) and glucuronic acid ($\text{GlcA; [GlcA}\beta\text{GalNAc}\beta_4\text{]_n}$; Sugahara et al., 2003). In vertebrates, the chains undergo further modification by the addition of sulfate at C4 and C6 of GalNAc and C2 of GlcA residues (Habuchi, 2000). Additionally, an epimerase can convert a subset of GlcA residues to l-iduronic acid, which can be subsequently sulfated.

Correspondence to Jeffrey D. Esko: jesko52@ucsd.edu

Abbreviations used in this paper: BEMAD, $\beta$-elimination followed by Michael addition with DTT; CPG, chondroitin proteoglycan; CSPG, chondroitin sulfate proteoglycan; DIC, differential interference contrast; dsRNA, double-stranded RNA; GalNAc, N-acetylgalactosamine; GlcA, glucuronic acid; MUDPIT, multidimensional protein identification technology; PH, pleckstrin homology.

The online version of this article contains supplemental material.
as well (dermatan sulfate). The pattern of sulfation varies in different tissues during development and in relation to age.

Interestingly, very little is known about the proteoglycans present in invertebrates, even in the well-studied nematode *Caenorhabditis elegans*. All of the components of the chondroitin biosynthetic machinery are completely conserved between *C. elegans*, mice, and humans, including formation of nucleotide sugar precursors and their import into the Golgi, assembly of the linkage region, and polymerization of the chain (Bulik et al., 2000; Berninsone et al., 2001; Hwang and Horvitz, 2002a,b; Yamada et al., 2002; Hwang et al., 2003a,b; Izumikawa et al., 2004). However, *C. elegans* lacks the sulfotransferases and the epimerase present in vertebrates, so the chains consist of unmodified GalNAc and GlcA residues (Yamada et al., 1999; Toyoda et al., 2000). In spite of its simplicity, genetic experiments demonstrate a crucial role for chondroitin in embryonic cell division and vulval morphogenesis (Herman et al., 1999; Bulik and Robbins, 2002; Hwang and Horvitz, 2002a,b; Hwang et al., 2003a,b; Mizuguchi et al., 2003; Izumikawa et al., 2004).

Although all of the enzymes required for chondroitin synthesis have been identified in *C. elegans*, no protein cores that bear the chondroitin chains have been described. To identify CPGs in *C. elegans*, we conducted searches using BLAST (Basic Local Alignment Search Tool) with all known mammalian CSPG core sequences but did not identify any obvious homologues. We therefore pursued a proteomics-based approach taking advantage of the chemical properties of the long, negatively charged chondroitin chains, a tagging method to modify the serine attachment sites, and mass spectrometry. We report the discovery of nine novel CPG proteins, none of which show homology to vertebrate CSPGs. Simultaneous depletion of two of these proteins, CPG-1/CEJ-1 and CPG-2, by RNAi results in defective cytokinesis during the first embryonic cell division. This phenotype is identical to that observed when the *sqv-5* chondroitin synthase was silenced by RNAi or by a loss-of-function mutation (Herman et al., 1999; Hwang and Horvitz, 2002b; Hwang et al., 2003b; Mizuguchi et al., 2003), suggesting that these two major chondroitin-carrying proteoglycans are essential for embryonic development.

**Results**

* C. elegans* expresses multiple CPGs

Proteoglycans were solubilized from a mixed-stage worm population by a combination of sonication and extraction with guanidine hydrochloride. After dialyzing the crude extract into urea, the proteoglycans were partially purified by stepwise elution from an anion-exchange column, using 0.2 M NaCl with urea to remove contaminating proteins and 1 M NaCl without urea to elute the proteoglycans (see Materials and methods). Treatment of the sample with sodium hydroxide released the chains from their core proteins by β-elimination reaction. The liberated chondroitin chains were generally <20 kD based on their elution by gel filtration chromatography relative to shark cartilage chondroitin sulfate (Fig. 1 A; Yamada et al., 1999), but they exhibited considerable heterogeneity. Treatment of the sample with chondroitinase ABC converted all of the material into nonsulfated disaccharides that eluted near the total volume of the column, confirming that the preparation contains predominantly chondroitin and very little heparan sulfate (Yamada et al., 1999; Toyoda et al., 2000).

To estimate the number of CPGs, the crude *C. elegans* proteoglycans were digested with chondroitinase ABC, separated by SDS-PAGE, and analyzed by Western blotting with the IB5 mAb that recognizes a neoepitope generated by chondroitinase digestion (Fig. 1 B). Several CSPG core proteins ranging in mass from 10 to >200 kD were detected in whole worm (lanes 2 and 3) and embryo extracts (lane 4), whereas none were detected if chondroitinase ABC was omitted (lane 1). The major species had masses of ~118, ~110, and ~60 kD. Treatment of samples with PNGase F had no effect on the mobility of these bands, suggesting that none of the proteins contained Asn-linked glycans (unpublished data). *C. elegans* also synthesizes at least three heparan sulfate–bearing proteoglycans with approximate masses of 45, 28, and 22 kD based on digestion of extracts with heparinase followed by Western blotting with 3G10 mAb, which is specific for a heparan sulfate neoepitope generated by enzyme digestion (Hwang et al., 2003b). In vertebrates, some proteoglycans contain both chondroitin sulfate and heparan sulfate chains (e.g., syndecans 1 and 3). Simultaneous digestion of samples with chondroitinase ABC and heparin lyase II did not alter the intensity of any of the CPG bands or give rise to new ones reactive with mAb IB5 or 3G10 (unpublished data). Thus, hybrid proteoglycans might not exist in *C. elegans*.

To identify the CPG core proteins detected by Western blotting, we searched the *C. elegans* genome for sequences related to any of the 25 known mammalian CSPG core proteins (Olson and Esko, 2004). No obvious homologues were found, with the exception of ORF Y47D3A.26, which resembled bamacan/SMC3 (Wu and Couchman, 1997). However, the coding sequence for this gene, as well as vertebrate bamacan, lacks
a signal peptide present in all known proteoglycans, consistent with the observation that most of bamacan is found in the nucleus in a nonglycosylated form (Ghiselli and Iozzo, 2000). Thus, a new method was needed to identify the various CPGs in *C. elegans*.

Fig. 2 describes the scheme used to identify *C. elegans* CPGs. After the DEAE chromatography step, samples were reduced, alkylated, and digested with trypsin. Peptides containing chondroitin chains (glycopeptides) were recovered by a second round of anion-exchange chromatography, which removed much of the contaminating protein from the sample. In some experiments, we used gel filtration to separate glycopeptides of large hydrodynamic size from contaminating peptides rich in acidic amino acids, which coeluted under the low salt conditions used to prepare the samples. Next, the samples were treated with sodium hydroxide, which caused β-elimination of the chains and the formation of dehydroalanine from serine residues. The dehydroalanine residues were then reacted with DTT, which adds across the double bond (Wells et al., 2002). In some experiments, the DTT-tagged peptides were further purified with a thiol column. Samples were then analyzed by mass spectrometry using multidimensional protein identification technology (MUDPIT) to identify peptides containing DTT (+167 D) or dehydroalanine (−18 D; Washburn et al., 2001; Wolters et al., 2001). The corresponding full-length proteins were identified by searching a *C. elegans* proteome database, parsing the data using the following criteria: (1) putative CPGs had to contain at least one glycosaminoglycan attachment site, defined as a modified serine residue flanked on the COOH-terminal side by glycine with one or more aspartate or glutamate residues near the site of chondroitin addition (Esko and Zhang, 1996), and (2) the proteins had to have a hydrophobic signal peptide, which would direct the protein into the secretory pathway, where chondroitin synthesis occurs. Nine tentative CPG core proteins were identified by this procedure and designated CPG-1 through -9 (Fig. 3).

**Characteristics of the *C. elegans* CPGs**

CPG-1 (C07G2.1a) and CPG-2 (B0280.5) have predicted masses of 62 and 54 kD, respectively (Fig. 3). CPG-1 was identified previously as CEJ-1 based on its reactivity with an antibody to a mammalian tight junction protein (Siddiqui, S.S., personal communication). It contains five putative chondroitin attachment sites based on amino acid sequence. One of these sites was confirmed by mass spectrometry based on presence of the DTT tag. In contrast, CPG-2 has 34 potential chondroitin attachment sites (four confirmed by mass spectrometry; Fig. 3). CPG-1 and -2 also have in common multiple peritrophin-A chitin binding motifs, defined by the arrangement of six cysteine residues within a six-cysteine loop.

**Figure 2. Scheme for identifying CPGs in *C. elegans*.**

**Figure 3. *C. elegans* CPGs identified by BEMAD/MUDPIT.** CPGs were purified from worm extracts and identified by mass spectrometry analysis. Nine independent purifications resulted in identification of nine CPGs, most of which were identified in multiple runs. The predicted masses are based on amino acid sequence, and the apparent masses are based on SDS-PAGE migration of recombinant proteins expressed in COS-7 cells after digestion with chondroitinase ABC and reduction. Putative glycosylation sites (short vertical lines) consist of Ser-Gly dipeptides flanked by one or more acidic amino acids. Identified sites (long vertical lines) represent serine residues modified with DTT by the BEMAD method (see Materials and methods). Genes enriched for germline expression were identified by Reinke et al. (2000). Schematic drawings of each CPG are shown. Gray circles indicate signal peptides, black ovals are peritrophin-A chitin binding domains, diagonally hatched boxes identify C-type lectin domains, and dotted lines below the protein indicate peptide coverage discovered by mass spectrometry.
residues that can form three disulfide bridges in a characteristic pattern (Wright et al., 1991; Beintema, 1994; Venegas et al., 1996; Merzendorfer and Zimoch, 2003). In CPG-1, two peritrophin-A domains lie in the NH2-terminal half of the protein and one toward the COOH-terminal end, whereas in CPG-2 the six peritrophin-A domains lie between sites predicted to be glycosylated (Fig. 3). Additionally, the COOH-terminal half of CPG-1 contains many Thr, Val, and Pro residues, resembling a segment of Muc-2, a membrane mucin present in vertebrates.

CPG-3 (R06C7.4) was not originally identified by the DTT-tagging method or by the presence of dehydroalanine (Wells et al., 2002) but rather by 5'-RACE using PCR primers designed against the repetitive SG region of cpg-2. Reevaluation of the mass spectral data revealed that peptides from CPG-3 were present in nearly all datasets (Fig. 3). Based on primary sequence, CPG-3 contains 15 putative chondroitin attachment sites and shows 21% sequence identity to CPG-2, mainly in the putative glycosylated region. The lack of mass spectral data for peptides that contain putative attachment sites in CPG-3 may have been caused by the failure of trypsin to gain access to and cleave the protein in heavily glycosylated regions.

CPG-4 (C10F3.1) has the largest predicted mass of the C. elegans CPGs at 84 kD. The majority of its 35 predicted glycosaminoglycan attachment sites reside in the COOH-terminal half of the protein, of which four sites were confirmed by DTT modification (Fig. 3).

CPG-5 (C25A1.8, CLEC-87) and CPG-6 (K10B2.3a, CLEC-88) are highly homologous, showing 67% identity and 90% similarity. CPG-5 and -6 contain only one putative chondroitin site toward the NH2 terminus, which was identified in both proteins by DTT modification. Both proteins contain a C-type lectin domain in the COOH-terminal half of the protein, like the aggrecan family of vertebrate CSPGs (Zelensky and Gready, 2003). This motif therefore classifies cpg-5 and -6 as members of the clec (C-type lectin) gene class. Proteins containing C-type lectins are common in C. elegans (88 members have been found thus far), but their glycan binding specificity has not been determined.

CPG-7 (K09E4.6) and CPG-8 (K03B4.7a) have predicted masses of only ~12 kD (Fig. 3). CPG-7 contains 11 putative chondroitin attachment sites, whereas CPG-8 contains 6 sites. In both cases, five of the sites were confirmed by mass spectrometry. Of the multiple peptides identified in CPG-8 samples, some contained a single modified site at either residue 61 or 63 (the tandem mass spectrometry data did not discriminate which site actually carried the chain). Interestingly, other peptides were modified at both serine residues. A similar phenomenon was observed at residues 84 and 88. CPG-9 (Y67D8C.8) has a predicted mass of only 7 kD, and mass spectrometry confirmed two of four attachment sites.

Interestingly, none of the CPGs identified using the scheme in Fig. 2 showed homology to vertebrate CSPGs. Furthermore, no homologues were found in searches of the D. melanogaster genome or Hydra vulgaris EST and genomic databases, suggesting that the core proteins have evolved independently in different organisms. In contrast, homologues for each C. elegans CPG were present in Caenorhabditis briggsae, another free-living nematode.

The glycosylation sites of many C. elegans CPGs appeared to follow an “SGX” motif, where X was usually G, A, T, or S (Fig. 4). BLAST searches using this sequence motif identified CPG-2, -3, -4, and -7 but no other proteins in C. elegans or vertebrates. D. melanogaster contains one protein containing this motif (CG6048-PA), but it has not yet been shown to contain a chondroitin chain. The aggrecan family of CSPGs in vertebrates also contains repetitive SG sequences within domains densely substituted with chondroitin chains, but the intervening sequences differ from those found in C. elegans CPGs.

C. elegans core proteins carry chondroitin sulfate chains after expression in mammalian cells

To confirm that the C. elegans CPGs identified by mass spectrometry could serve as a scaffold for chondroitin synthesis, the cDNA sequences were subcloned into a mammalian expression vector with a COOH-terminal Myc tag and expressed in COS-7 cells.
Conditioned media was collected 48 h after transfection, purified by anion-exchange chromatography, treated with chondroitinase ABC, and Western blotted with an antibody to Myc (Fig. 5). CPG1-6 yielded reactive protein bands that did not appear in the absence of chondroitinase ABC digestion. Attempts to express CPGT-9 in COS cells have thus far been unsuccessful, possibly because of the small size of the protein core. None of the proteoglycans migrated at the molecular mass predicted by their primary sequence. Furthermore, a comparison of the bands in Fig. 1 B to the calculated masses in Fig. 3 did not demonstrate obvious correspondence. However, the migration pattern of CPG-2 expressed in COS cells appears similar to the 120-kD core in the crude proteoglycan preparation (Fig. 1 B). Likewise, CPG-3 migrates at 55–60 kD, the same position as a prevalent core in the crude proteoglycan preparation (Fig. 1 B). These observations suggest the presence of higher order structures in the core proteins or that the chondroitin stub oligosaccharides remaining after chondroitinase digestion altered their migration. Mammalian CSPGs also chondroitin stub oligosaccharides remaining after chondroitinase digestion altered their migration. Mammalian CSPGs also

Simultaneous RNAi of cpg-1 and -2 leads to penetrant embryonic lethality

Expression of cpg-1, -2, -3, -5, and -6 is enriched in the C. elegans germline, suggesting possible activity in germline development and progression, spermatogenesis, oogenesis, or embryogenesis (Reinke et al., 2000; Fig. 3). However, genomewide RNAi screens have not reported phenotypes for any of the cpg genes, suggesting either functional redundancy or subtle phenotypes (www.wormbase.org). Lee and Schedl (2001) reported that GLD-1 acts as a translational repressor of several mRNA targets in the germline, including cej-1 and B0280.5, which were identified in our proteomic screen as cpg-1 and -2, respectively. They also reported that RNAi of each gene alone had no phenotype, whereas simultaneous depletion of both genes resulted in embryonic lethality. Because this phenotype is reminiscent of the embryonic phenotype resulting from inhibiting the SQV-5 chondroitin synthase, we decided to analyze CPG-1 and -2 in greater detail.

To confirm that CPG-1 and -2 contain chondroitin chains in vivo, worms were fed bacteria expressing double-stranded RNA (dsRNA) directed against cpg-1 or -2, and samples were analyzed by Western blotting. No differences were initially observed in CPG-1-depleted extracts (Fig. 6 A), but CPG-2–depleted extracts showed a dramatic reduction of the major band at ~120 kD, as well as the bands at ~80 and ~45 kD (Fig. 6 A, arrowheads), suggesting that all three were related to each other (the predicted mass of the protein was 54 kD; Fig. 3). The simultaneous reduction of these three bands was confirmed by using a second construct targeting a different region of cpg-2 (unpublished data). The RNAi effect was specific, as the other major core protein band at ~60 kD remained unaltered.

CPG-1 and -2 have in common multiple peritrophin-A chitin binding domains (Fig. 3). To test the functionality of these domains, crude worm extract was incubated with chitin beads. Four CPG bands were able to bind chitin, including those at 120, ~80, and 45 kD (Fig. 6 B, filled arrowheads), as well as a new band at ~150 kD (Fig. 6 B, open arrowhead). Chitin binding appeared to be selective, as the major band at 60 kD did not interact with the resin. cpg-2 RNAi treatment diminished the three bands at 120, ~80, and 45 kD, confirming the presence and functionality of one or more peritrophin-A chitin binding domains in these proteoglycans (Fig. 6 C, filled arrowheads). The band migrating at ~150 kD was present in vector and CPG-2–depleted extracts but absent in extracts depleted of CPG-1 and CPG-1/CPG-2 (Fig. 6 C, open arrowhead). Recombinant CPG-1 expressed in COS-7 cells had a similar mass (Fig. 5), suggesting the 150-kD band represents CPG-1. Thus, both CPG-1 and -2 behave as CSPGs in vivo and have functional chitin binding domains.

To study the function of CPG-1 and -2, we compared the effect of silencing their expression by RNAi to depletion of SQV-5, which encodes the chondroitin synthase. Brood sizes of cpg-1(RNAi), cpg-2(RNAi), and cpg-1/cpg-2(RNAi) animals were comparable to uninjected worms or worms injected with

Figure 5. CPG-1 through -6 behave as CSPGs in COS-7 cells. CPG-1 through -6 were expressed as Myc-tagged recombinant proteins in COS-7 cells. Proteoglycans from conditioned media were purified by anion-exchange chromatography (see Materials and Methods). Samples were digested with chondroitinase ABC (+) or left untreated (−), separated by SDS-PAGE, and Western blotted with an anti-Myc mAb.

Figure 6. CPG-1 and -2 bind chitin and exist as CSPGs in vivo. (A) Proteoglycan extracts of vector and RNAi-treated animals were digested with chondroitinase ABC, analyzed by SDS-PAGE, and Western blotted with the 1B5 mAb that recognizes the chondroitin stub remaining after enzyme digestion. cpg-2(RNAi) reduced the bands at 120, ~80, and ~45 kD (arrowheads). (B) 20 μg of total worm extract (input) was incubated with chitin beads. Four bands specifically bound chitin (arrowheads), but the 60 kD band did not. (C) Proteoglycan extracts from RNAi-treated animals were enriched by affinity chromatography on chitin beads. cpg-1(RNAi) reduced the band at ~150 kD (open arrowhead), whereas cpg-2(RNAi) reduced the bands at 120, ~80, and 45 kD (filled arrowheads).
buffer alone (Table I; P > 0.05 by one-way analysis of variance). Depletion of CPG-1 or -2 alone also had no effect on viability, as 96–99% of the embryos hatched into healthy larvae that grew into normal fertile adults (Table I). However, simultaneous depletion of CPG-1 and -2 had a synergistic effect, resulting in penetrant embryonic lethality. Whether this synergy is due to the functional redundancy of CPG-1 and -2, the partially penetrant nature of RNAi, other functions of the proteoglycans besides carrying chondroitin chain, or the cross-reactivity of cpg-2 dsRNA with other targets remains to be determined. sqv-5(RNAi) treatment also had no effect on egg production or laying but had a hatching phenotype similar to cpg-1/cpg-2(RNAi) (95% of progeny failed to hatch). To look for genetic interactions between sqv-5 and cpg-1 or -2, a heterozygous sqv-5(n3611) mutation was compounded with single cpg-1(RNAi) or cpg-2(RNAi). However, no combination affected brood sizes or embryonic viability (unpublished data), possibly because the heterozygous mutation did not reduce chondroitin levels sufficiently to see the effect of depleting a single CPG.

RNAi of cpg-1 and -2 results in a cytokinesis defect

To determine the underlying cause for embryonic lethality, we looked more closely at the first cell division, where defects have been reported in sqv-5 chondroitin synthase mutants (Hwang et al., 2003b). To better visualize the plasma membrane, a strain expressing GFP fused to a pleckstrin homology (PH) domain derived from mammalian phospholipase-C was injected with dsRNA against cpg-1, cpg-2, cpg-1/cpg-2, or sqv-5. The PLC161 PH domain binds to phosphatidylinositol 4,5 bisphosphate located specifically on the plasma membrane and is an excellent probe for monitoring the changes in cell shape that accompany cytokinesis (Audhya et al., 2005). The depleted embryos were simultaneously imaged by differential interference contrast (DIC; Fig. 7, top) and spinning-disc confocal microscopy (Fig. 7, bottom). Embryos were imaged in utero because depletion of either SQV-5 or both CPG-1 and -2 resulted in embryos that were fragile and osmotically sensitive. Depletion of CPG-1 or -2 alone had no effect on fertilization, membrane ruffling, pseudocleavage, pronuclear meeting and rotation, karyokinesis, and cytokinesis (compare Video 2 to buffer-injected worms in Video 1, available at http://www.jcb.org/cgi/content/full/jcb.200603003/DC1). Initiation of the cleavage furrow occurred normally (Fig. 7 D, bottom), and two separate daughter cells routinely appeared (Fig. 7 E). A four-celled embryo was distinctly seen at the end of filming, demonstrating the fidelity of the second round of cell division (Fig. 7 G). In contrast, embryonic development was severely perturbed in embryos simultaneously depleted of CPG-1 and -2 (Video 3). In wild-type embryos, fertilization triggers the oocyte pronucleus to undergo two rounds of meiotic segregation, which produces two polar bodies that are extruded from the embryo by small cytokinesis-like events. In CPG-1/CPG-2–depleted embryos, fertilization occurred normally but polar body extrusion failed, resulting in extra nuclear material that remained in the embryo (Fig. 7, I and J). The pronuclei migrated and the complex of the pronuclei and centrosomes rotated onto the long axis of the cell as in wild type, but the membrane ruffling that normally precedes and accompanies these events was absent, as was the space between the embryonic plasma membrane and the eggshell that is normally present during cell division (compare Fig. 7, C [bracket] and J [solid white lines]). Spindle formation appeared normal and the chromosomes appeared to segregate, but the cleavage furrow failed to form and ingress to introduce a new cell surface between the segregated chromosomes (Fig. 7, K and L). Without the barrier of a new cell membrane, the daughter nuclei fused (Fig. 7 M) and repeatedly attempted cell division without cytokinesis (Fig. 7 N). An essentially identical phenotype was observed after depletion of SQV-5 (Fig. 7, O–U; and Video 4). Since depleting CPG-1 and -2 mimics the defect in embryonic cytokinesis by depletion of the SQV-5 chondroitin synthase, we conclude that CPG-1 and -2 are two functionally important CPGs required at this early developmental stage.

Discussion

A proteomic approach for identifying novel proteoglycans

We report here the identification of a group of novel CPGs in C. elegans. The identification scheme used a combination of conventional methods to purify proteoglycans and glycopeptides based on the polyanionic character and size of the chondroitin chains coupled with mass spectrometry. Additional selection criteria ensured that the majority of identified proteins had the properties of a proteoglycan. These included chondroitin attachment sites modified with DTT, the presence of a serine residue followed by glycine and one or more nearby acidic residues, and a signal peptide. Although DTT addition was not stoichiometric, the β-elimination step introduced an alternate method to identify the protein via the characteristic mass difference between dehydroalanine and serine (Wells et al., 2002). Attempts to optimize the DTT-addition reaction have not yet yielded higher efficiency, and other methods are needed to improve this step to better assess the stoichiometry of chondroitin attachment. Similarly, adjustment of the proteolysis conditions might improve peptide coverage, as the absence of trypsin

Table I. Brood size and viability in RNAi-treated worms

<table>
<thead>
<tr>
<th>RNAi treatment</th>
<th>Brood size*</th>
<th>Range</th>
<th>Viability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninjected</td>
<td>133 ± 43</td>
<td>75–197</td>
<td>98.9</td>
</tr>
<tr>
<td>Buffer</td>
<td>120 ± 45</td>
<td>65–178</td>
<td>99.6</td>
</tr>
<tr>
<td>cpg-1</td>
<td>95 ± 34</td>
<td>64–184</td>
<td>95.9</td>
</tr>
<tr>
<td>cpg-2</td>
<td>144 ± 42</td>
<td>87–202</td>
<td>99.2</td>
</tr>
<tr>
<td>cpg-1/cpg-2</td>
<td>129 ± 56</td>
<td>62–203</td>
<td>0.0</td>
</tr>
<tr>
<td>sqv-5</td>
<td>77 ± 22</td>
<td>22–114</td>
<td>4.8</td>
</tr>
</tbody>
</table>

The number of embryos and hatched L1 larvae were assessed 24 h after placing control worms or worms injected with dsRNA on individual plates. Viability was determined >24 h after the initial brood size was counted and was calculated as the number of hatched larvae divided by the total number of embryos laid. n = 6–11.

*One-way analysis of variance showed no significant difference between brood sizes. P > 0.05.
cleavage sites and the potential inhibitory effects of closely spaced chondroitin chains could interfere with the analysis. Regardless of its limitations, the current protocol led to the identification of nine proteins that had the hallmarks of CPGs, and six served as substrates for chondroitin sulfate assembly when expressed in mammalian cells. The simplicity of the identification strategy suggests that it should be generally applicable to other organisms, as well as tissues, cells, and secretions from vertebrates for which little information about the proteoglycan composition exists.

Interestingly, none of the *C. elegans* CPGs identified in this study had predicted membrane spanning domains or consensus sites for glycosphatidylinositol-anchor attachment. *C. elegans* expresses homologues of the membrane heparan sulfate proteoglycan syndecan, which has a single membrane spanning segment, and the glycosphatidylinositol-anchored heparan sulfate proteoglycan, glypicanc. Because vertebrates express multiple membrane CSPGs (e.g., NG2, CD44, and phosphacan; Olson and Esko, 2004), our failure to identify membrane CPGs in *C. elegans* could indicate either their absence or that they represent only minor components of the proteoglycan population. Heparan sulfate proteoglycans also were not detected, most likely because *C. elegans* expresses ~250 times less heparan sulfate than chondroitin (Yamada et al., 1999; Toyoda et al., 2000). Thus, it may be necessary to modify the procedure to selectively enrich membrane proteins and increase sample size to detect rare proteoglycans.

The lack of homology between *C. elegans* CPGs and vertebrate CSPGs and the observation that none of the CPGs have homologues in two other invertebrates (*D. melanogaster* and *H. vulgaris*) raises interesting questions about the evolution of these molecules. All of the core proteins identified in this study have sequence motifs required for initiation of glycosaminoglycan biosynthesis similar to those found in vertebrates (Esko and Zhang, 1996). Furthermore, *C. elegans* expresses orthologues of all of the vertebrate enzymes required for assembly of the linkage region tetrasaccharide (xylosyltransferase, galactosyltransferases I and II, and glucuronyltransferase I; Bulik et al., 2000; Hwang et al., 2003a), as well as the chondroitin polymerizing system (Hwang et al., 2003b; Mizuguchi et al., 2003), including the newly discovered chondroitin polymerizing factor (Izumikawa et al., 2004). These findings indicate that the mechanism of chain initiation and polymerization evolved early in metazoans and has been maintained by strong selection. In contrast, the core proteins on which the chains assemble have continued to evolve, presumably to serve specialized functions. For example, CPG-1 and -2 have functional chitin binding domains that could interact with the nematode eggshell, whereas no vertebrate proteoglycans contain this motif and vertebrates lack chitin as a structural component. The presence of chitin in insect exoskeletons and peritrophic matrices that line the gut suggests that orthologues of these CPGs might exist in *D. melanogaster* (Merzendorfer and Zimoch, 2003), but BLAST searches have not yet uncovered them. Additional studies using techniques like the one reported here are needed to identify the proteoglycans present in *D. melanogaster* and other organisms, which in turn might help define the evolution of proteoglycans during the metazoan expansion.
CPG-1/CEJ-1 and CPG-2 play essential roles in embryonic cell division in *C. elegans*

Our preliminary analysis of RNAi-depletion experiments indicated that silencing *cpg-1* or *-2* had no effect on embryogenesis or morphogenesis. However, depletion of both genes resulted in a strong embryonic phenotype characterized by failure of polar body extrusion after fertilization, loss of membrane ruffling preceding pronuclear fusion, and failure to initiate the cleavage furrow before cytokinesis. The necessity of depleting both CPG-1 and -2 to uncover a phenotype suggests functional redundancy of the proteins, though the only commonality is the presence of chitin binding domains and at least one chondroitin chain (Fig. 3). An alternative explanation is that depletion of each protein by RNAi may not have been fully penetrant, and the phenotype after double depletion could be the result of more extensive depletion of chondroitin. However, this explanation seems unlikely because CPG-1 is a minor component and had to be enriched by chitin affinity chromatography to detect its presence. Instead, the requirement for simultaneous depletion of both proteoglycans could be due to other functional aspects of the CPGs (e.g., their ability to bind chitin or the presence of other glycans). Generation of deletion mutants will help resolve this issue.

The strong embryonic lethal phenotype seen with *cpg-1/-cpg-2(RNAi)* resembles that of the *sqv-5* chondroitin synthase mutant described by Hwang et al. (2003b). How the CPGs mediate these effects is unknown. One of the roles of CPGs may be to fill the space between the eggshell and the embryo, and the presence of a high concentration of polyanions and their counterions might cause sufficient hydrostatic pressure to aid in formation of the extraembryonic space and subsequent ingress of the cleavage furrow. Thus, the loss of CPGs in the perivitelline space surrounding the embryo could result in collapse of the extracellular space and apposition of the plasma membrane to the eggshell (Fig. 7). Interestingly, chondroitin has been shown to be present on both the embryonic cell surface and the eggshell, suggesting that it may function at both locations (Sugahara et al., 2003). This leads to an alternative explanation: CPG-1 and -2 could act as structural elements of the eggshell or might bridge chitin polymers in the eggshell with other components of the embryonic plasma membrane that result in transmembrane signaling to cytoskeletal components involved in cytokinesis. Regardless of which hypothesis is correct, we have shown that CPG-1 and -2 play a crucial role in cytokinesis, as their simultaneous depletion results in multinucleated single cell embryos.

The observation that the regulatory protein GLD-1 may translationally regulate *cpg-1* and *-2* transcripts indicates a sophisticated level of control over proteoglycan expression during early embryogenesis. GLD-1 is localized to the germine cytoplasm (Jones et al., 1996) and translationally regulates the expression of a large number of genes by interacting with mRNA 3'UTR sequences (Lee and Schedl, 2001). It also protects some transcripts from nonsense-mediated mRNA decay by binding to 5'UTR sequences (Lee and Schedl, 2004). One example is *gna-2*, which encodes the enzyme glucosamine 6-phosphate N-acetytransferase. This enzyme plays an essential role in the biosynthesis of the nucleotide sugar UDP-N-acetylgalcosamine (UDP-GlcNAc). The biosynthesis of a variety of glycans depends on UDP-GlcNAC, including chitin (a polymer of GlcNAC) and chondroitin by way of UDP-GalNAC. Interestingly, a null allele of *gna-2(qa705)* exhibits maternal effect lethality and multinucleated embryos (Johnston, W., and Dennis, J., personal communication). The similar phenotype of *cpg-1/cpg-2(RNAi)*, *sqv-5(RNAi)*, and *gna-2* mutants may indicate a common mechanism of action through formation of CPGs. Posttranscriptional regulation of CPG assembly in *C. elegans* via GLD-1 is reminiscent of the regulation of enzymes involved in heparan sulfate formation in vertebrates (Grobe and Esko, 2002) and may indicate the general importance of translational control over proteoglycan assembly during development.
a tandem mass spectrometer. The identities of the DTT-tagged peptides were determined by searching the tandem mass spectra against a C. elegans proteome database using SEQUEST software and a computer array. The unique mass signature imported by DTT (+167 D) and unmodified dehydroalanine residues (~18 D) were used to determine the sites of chondroitin addition.

Recombinant protein expression
C. elegans cDNA was prepared from total RNA with the SuperScript III First-Strand kit (Invitrogen). cDNA was prepared from total RNA with the SuperScript III Reverse transcriptase (Invitrogen) following the manufacturer’s instructions. Media was harvested 48 h later and purified over DEAE-Sephalac as described previously (Bame and Esko, 1989). COS-7 cells were cultured in Dulbecco’s modified Eagle’s medium (CellGro) supplemented with 10% fetal bovine serum, 100 U/ml penicillin G, and 100 μg/ml streptomycin sulfate.

RNAi
C. elegans cDNA was amplified by PCR with primers engineered to contain T7 (forward primer) or T3 (reverse primer) bacterial promoter sequence and base pairs 1–649 for cpg-1 and 1–775, 754–1572, and 1190–1572 for cpg-2. dsRNA was generated with the Megascript T7 and T3 transcription kits (Ambion) according to the manufacturer’s instructions. ODS5 L4 or young adult worms were injected with dsRNA and allowed to recover at 16°C or 20°C. To count brood sizes, worms were transferred 24 h after injection to individual plates, and 1 d later the number of eggs and hatched L1 larvae were counted (Maddox et al., 2005). Viability was measured 24–36 h later by counting the number of hatched larvae and un-hatched embryos. Percentage of viability was calculated as number of hatched progeny divided by the total number of eggs laid.

Early embryonic cell division was assessed 24 h after injection. Embryos were filtered in utero because cpg-1/cpg-2(RNAi) embryos were fragile and osmotically sensitive. Injected animals were anesthetized with 1 mM levamisole in M9 buffer, mounted on an agarose pad, and filmed as described previously (Maddox et al., 2005). Images were acquired with a DeltaVision deconvolution microscope (Applied Precision) equipped with a charge-coupled device camera (CoolSnap; Roper Scientific) at 20×. DIC and GFP images were acquired using a 100× Oil-Planapo objective (Olympus) with a 2×2 binning and a 480×480 pixel area. DIC and GFP images were acquired at 10-s intervals by sequentially rotating the analyzer and GFP filter set into the light path. Illumination was attenuated with a 10% neutral density filter. Images were analyzed with MetaMorph software (Universal Imaging Corp.).


References


Published June 19, 2006

Online supplemental material
The videos show the first two rounds of cell division in C. elegans embryos exposed to the RNAi treatments described in Fig. 7. Video 1 shows normal cell division in buffer-injected embryos. Video 2 shows normal cell division in cpg-2(RNAi) embryos. Video 3 shows that cpg-1/cpg-2(RNAi) embryos fail to complete the first cell division. Video 4 shows that sqv–5(RNAi) produces the same phenotypic defects as cpg-1/cpg-2(RNAi). Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.200603003/DC1.

We would like to thank lan McLeod (The Scripps Research Institute) for helpful discussion regarding mass spectrometry; Lindsay Lewellyn, Amy Maddox, and Nathan Porter for assistance with worm imaging studies; and Kafi Aroian for many helpful and inspiring discussions.

This work was supported by National Institutes of Health grants GM33063 (to J. Esko) and RR11823 (to J. Yates) and Genetics Training Grant GM018666 (to S. Olson). K. Oegema is a Pew Scholar in the Biomedical Sciences and is supported by funding from the Ludwig Institute for Cancer Research.

Submitted: 1 March 2006
Accepted: 11 May 2006


