Proteoglycans: cell biologists have had a love-hate relationship with these molecules almost since their discovery. Their biochemical properties, dominated by heterogeneous and highly charged glycosaminoglycan (GAG) chains, can make purification challenging and structural analysis painful. Their ability to bind scores of growth factors, growth factor-binding proteins, extracellular proteases, protease inhibitors, extracellular matrix molecules, and other proteins takes the concept of molecular promiscuity to new heights. On top of this, they seem always to be underfoot, showing up on plasma membranes in hundreds of thousands of copies per cell and in extracellular matrices at milligram per milliliter concentrations.

And yet despite these peculiarities (or perhaps because of them) proteoglycans have inspired an extraordinary range of models, theories, and speculation. Since the 1960s, proteoglycans have been credited, in one system or another, with controlling the following: cell division, adhesion, spreading, migration, chemotaxis, axon guidance, matrix assembly, lipoprotein uptake, extracellular proteolysis, and viral entry.

Do proteoglycans do all of these things? Progress on this question during the past decade has been recently kicked into high gear by a flurry of in vivo results (in mice, frogs, flies, and worms), in many cases coming from investigators who never intended to become involved with such difficult molecules. Here, we review some of these findings and discuss how they both confirm old notions of proteoglycan function and suggest new ones. We have chosen to focus exclusively on the heparan sulfate proteoglycans (HSPGs), in part because new data on these molecules have been particularly plentiful, but mostly because those data speak more directly about the functions of GAG chains, the moieties that make proteoglycans unique. However, it should be noted that in vivo studies of proteoglycans that primarily bear chondroitin/dermatan sulfate and keratan sulfate (the other families of GAGs) are also providing important new insights (Fassler et al., 1994; Danielsen et al., 1997; Olsen, 1997; Chakravarti et al., 1998).

The Nature of the Problem

To be sure, answers coming from the genetic manipulation of animals are enriching almost all areas of biology these days. In part, what is different about proteoglycans is that the cell biological tools one can use to study their functions, in particular the functions of their GAG chains, have never been as rich and varied as those for more conventional proteins. For example, GAGs cannot be transfected into cells, nor can one express dominant negative versions of them. Antibodies that bind GAGs have been generated, but ones that block functions have not. Reagents that interfere with GAG–protein interactions exist in the form of GAG-binding peptides and cationic drugs, but their specificity tends to be poor. Free GAGs can also be used to interfere with GAG–protein binding, but their large size and potential to act as protein–protein cross-linkers make it difficult to draw firm conclusions from their biological effects. Finally, there are a few drugs that can be used to interfere with GAG biosynthesis, but each has at least some problems with efficacy, specificity, and/or toxicity.

In spite of these hardships, cell biologists studying proteoglycans have managed to make some extraordinary discoveries. In particular, the finding that fibroblast growth factors (FGFs) require heparan sulfate (HS) for high affinity binding to their receptors stands as a milestone in the elucidation of HSPG function. Yet, even this simple fact was not easily established. Despite the well known high affinity of FGFs for heparin (a rather heavily modified form of HS that, owing its low cost, is widely used for studying protein–HS interactions) and clear evidence that the most abundant cell surface binding sites for FGFs are HSPGs (Gospodarowicz et al., 1984; Moscatelli, 1987), most early studies in this area concluded that HSPGs play no direct role in FGF–receptor interactions and, instead, relegated HS to the job of protecting FGFs from denaturation or proteolysis, or sequestering FGFs in the extracellular matrix (Gospodarowicz and Cheng, 1986; Moscatelli, 1987, 1992). What enabled two groups, working in parallel, to overturn that view, was the judicious use of methods to interfere with GAG expression: in one case, an inhibitor of GAG sulfation and an HS-degrading enzyme were used (Rapraeger et al., 1991); in the other case, it was mutant cell lines that failed to synthesize HS (Y ayon et al., 1991). Since that time, analogous methods have been used by others to extend the notion of HS-dependent growth factor to various other polypeptides, including heparin bind-
ing EGF-like growth factor, hepatocyte growth factor, and Wingless (a Drosophila member of the Wnt family), to name just a few (A viezer and Y ayon, 1994; Zioncheck et al., 1995; Reichsman et al., 1996).

In Search of Genetic Insights

With HS-deficient cell lines proving to be among the more useful tools for studying HSPG function in vitro, did it not escape the notice of researchers that mutations affecting proteoglycans in intact animals could also be extremely informative, especially with regard to cellular functions that might be difficult to study in vitro. Yet by mid-1995, although a few mutations had been described that directly or indirectly affect the major chondroitin/keratan sulfate proteoglycan of cartilage (Li et al., 1993; Hästbacka et al., 1994; Watanabe et al., 1994), a general lack of mutations involving HSPGs—even in the more genetically tractable organisms, such as C. elegans and Drosophila—was becoming disturbingly clear. Are HSPGs simply not very important? Or are they so important that mutations are difficult to recover in vivo? Or are the genes that control their synthesis highly redundant? Fortunately, the wait for answers to these questions was not long, with the first informative studies beginning in late 1995.

Before reviewing these studies, it is useful to distinguish between two classes of mutations that affect HSPG expression. The first involves the enzymes and transporters required for HS biosynthesis. As illustrated in Fig. 1, at least 14 biochemical steps contribute to the synthesis of HS chains. A block at some steps would eliminate HS expression altogether, whereas at others only subtle changes in the structure of HS chains would be seen. A second class of mutations involves the core proteins on which HS is synthesized. We now know that most cell surface HS is carried by syndecans, transmembrane proteins of which four exist in vertebrates, and glypicans, GPI-anchored proteins of which at least six occur in vertebrates (Bernfield et al., 1999). On some cells, HS may also be contributed by other integral membrane proteins, such as CD44 and betaglycan, which have HS-independent functions as well. In the extracellular matrix, the major carriers of HS appear to be perlecan (Iozzo et al., 1994) and agrin (Cole and Halfter, 1996).

Glypicans Pop Up Twice

A first glimpse at HSPG function in vivo came in 1995 with the analysis of a Drosophila mutant known as division abnormally delayed, or dally (Nakato et al., 1995). Mutations in dally were identified on the basis of cell division patterning defects in the eye and larval brain. In particular, certain sets of neuronal precursors show a disruption in progression from the G2 to M phases of the cell cycle. In the brain, this failure of orderly cell division disrupts a subsequent division cycle that is triggered by photoreceptor axons arriving from the developing eye. Sequencing of the dally gene revealed that it encodes a protein belonging to the glycan family of cell surface HSPGs (Nakato et al., 1995). Biochemical studies have since shown that Dally bears all the features of vertebrate glypicans including glycosylphosphoinositol (GPI)–anchorage and selective glycosylation with HS (Jackson et al., 1997; Tsuda et al., 1999).

The functions of dally are not limited to the nervous system. Aanalysis of a collection of mutant alleles showed that it is required for the proper morphogenesis of other tissues, including the wing, antenna, and genitalia. Furthermore, difficulties in isolating complete loss-of-function dally alleles suggest that dally is a haplolethal locus (Lin and Perrimon, 1999; Tsuda et al., 1999), i.e., the level of dally expression is so critical that a reduction to 50% of the wild-type is lethal.

The initial characterization of dally mutants was rapidly followed by a report that an X-linked human disorder known as Simpson-Golabi-Behmel syndrome (SGBS) is caused by deletions and mutations in the gene encoding glypican-3 (Pilia et al., 1996). SGBS is characterized by prenatal overgrowth of multiple tissues and organs, together with an increased susceptibility to the formation of certain tumors.

Is there a common basis for the very different fly and human phenotypes that result from glypican mutations? At a very basic level, both phenotypes suggest a derangement of cellular growth control, an interpretation that fits with the currently popular model that HSPGs are components of growth factor signaling pathways. To address this issue, dally was tested for its ability to affect signaling mediated by two known HS-binding growth factors in Drosophila, Decapentaplegic (Dpp), a TGF–b/bone morphogenetic protein–related protein, and Wingless (Wg), a Wnt family member. The results strongly suggest that Dally potentiates cellular responses to both molecules, but with surprising tissue specificity. In the embryo, Dally influ-

![Figure 1. Pathways in heparan sulfate proteoglycan biosynthesis.](image-url)
ences Wg but not Dpp-directed events, whereas in the developing genitalia daily promotes Dpp signaling yet is antagonistic to Wg (Tsuda et al., 1999).

In the case of SGBS, initial attempts to understand the phenotype focused on the possibility that glypican-3 might affect signaling by insulin-like growth factors (IGFs). This idea was suggested by the phenotypic overlap between SGBS and another overgrowth syndrome, Beckwith-Wiedemann syndrome. The latter is a complex genetic disorder that causes the loss of imprinting, and resultant overexpression, of the gene encoding IGFI-2 (Reik and Maher, 1997). The hypothesis was put forth that loss of glypican-3 also leads to increased IGF-I expression or activity, i.e., that glypican-3 is an endogenous inhibitor of the growth-promoting effects of IGFs. Indeed, Pillia et al. (1996) initially asserted that glypican-3 binds directly, via its core protein, to IGFI-II, but subsequent studies contradict that finding (Song et al., 1997). Instead, the possibility has been raised that glypican-3 acts more at the level of stimulation of apoptotic cell death, rather than restraint of cell proliferation. For example, cell culture data show that transfection of exogenous glypican-3 into certain cell lines can trigger apoptosis (Gonzalez et al., 1998). More recently, Cano-Gauci et al. (1999) have generated a glypican-3–null mouse, which replicates some of the features of SGBS. In these animals alterations in systemic and tissue IGF-I levels are not observed.

**A Flurry of Mutants in Biosynthetic Enzymes**

Whereas studies on daily and glypican-3 showed that mutations in HSPG core protein genes can have dramatic phenotypes, recent work on other HSPG-s (e.g., syndecan-1, glypican-2), suggests that loss-of-function phenotypes can also be subtle or undetectable, most likely because of compensation or functional redundancy among HSPGs (Bernfield et al., 1999; Saunders, S., and A.D. Lander, unpublished data). One way to circumvent this potential problem is to study the effects of mutations that alter glucosaminoglycan biosynthesis, since these should presumably affect all HSPGs.

The first descriptions of phenotypes stemming from mutations in glucosaminoglycan biosynthesis came from what was, at the time, an unexpected quarter: Drosophila laboratories searching for new components of the Wg and Dpp signaling pathways. Three groups simultaneously reported that sugarless (sgl, also known as supenkasper or kiwi), a gene encoding a protein homologous to vertebrate UDP glucose dehydrogenase (UDPGDH), is required for Wg-directed patterning of the early embryo (Binari et al., 1997; Haerry et al., 1997; Hacker et al., 1997; Saedler et al., 1997). Mutations in sgl also suppressed the effects of expressing activated Dpp receptors in the wing (O’Connor and Haerry, 1999), suggesting a role for sgl in Dpp signaling.

UDPGDH is the enzyme that generates UDP-glucuronic acid, which donates the glucurionate that is one of the two sugars in the repeating disaccharide backbones of both HS and chondroitin sulfate (Fig. 1). Indirect evidence that sgl protein indeed has UDPGDH activity comes from the observation that GAG modification of both D-syndecan and Daily is disrupted in sgl mutants (Haerry et al., 1997; Tsuda et al., 1999). Interestingly, sgl mutants not only have defects in Wg signaling, but also show genetic interactions with daily, suggesting that glycosaminoglycan modification of Daily is critical for its activity in Wg-directed events (Tsuda et al., 1999).

Whereas UDPGDH generates one of the building blocks of HS, the enzyme that actually assembles the HS polymer is the HS copolymerase, so named because biochemical studies suggest that a single protein adds both glucuronate and 4-acetylglucosamine to growing HS polymer (Lind et al., 1993). Until recently, the molecular identity of this enzyme was a mystery, but in the last few years, both its identity and mutant phenotypes in flies and man have emerged.

An early clue to the identity of the copolymerase came from somatic cell, rather than animal, genetics. The discovery by Spear and colleagues that the binding of herpes viruses to cell-surface HS is a key initial step in viral infection (Shieh et al., 1992) implied that selection of cultured cells for viral resistance might allow one to isolate mutants in HS biosynthesis. Tufaro and colleagues did just this (Banfield et al., 1995), and, subsequently, used an expression cloning strategy to identify a cDNA that restored virus sensitivity to a severely HS-deficient cell line (McCormick et al., 1998). Subsequent biochemical studies showed that the protein encoded by the cDNA was likely to be all or part of the HS copolymerase (Lind et al., 1998).

The cDNA isolated by these studies corresponds to EXT2, one of three tumor suppressor genes that had been identified by linkage analysis as responsible for the autosomal dominant disorder hereditary multiple exostoses. A parent loss-of-function mutations in either EXT1 or EXT2 predispose to the development of benign cartilage-capped tumors from the growth plates of bones, tumors that occasionally progress to malignancy (Wufts et al., 1998). Like EXT2, EXT1 also can restore HS copolymerase activity in deficient cell lines (Lind et al., 1998).

Taken together with the linkage of SGBS to glypican-3 (see above), the data on the EXT genes provide a second clear connection between disruption of HSPG function, loss of growth control, and tumor development. For a researcher arguing that HSPGs act as coreceptors for mammalian growth factors, these results were reassuring, but also disquieting in that in vivo data were now pointing toward a role of HSPGs in the restraint, rather than promotion, of cell growth. Whether this reflects a major role for HS-dependent growth factors in growth inhibition in vivo, or just an incomplete understanding of how HSPGs affect growth factors, is unclear.

Helping in sorting out this puzzle may have come from the Drosophila field. At the same time that the connection between mammalian EXT genes and HS biosynthesis was made, Bellaiche et al. (1998) described the Drosophila gene tout-velu (ttv). ttv encodes an EXT-1 homologue and, nicely confirming the data on vertebrate EXT genes, ttv mutants exhibit a great reduction in HS (but not chondroitin sulfate) levels (Toyoda et al., 2000). Mutations in ttv were picked up in a screen for maternally acting genes affecting early patterning of the embryo (tout-velu means all hair, a reference to the appearance of the cuticles of affected embryos). Embryos lacking ttv function show segment polarity defects also found in wingless and hedgehog mutants, although a closer study suggested that ttv prima-
rily disrupts hedgehog signaling. Hedgehogs are secreted factors critical for patterning both in flies and vertebrates.

The nature of the defect in tv mutants is intriguing. Evidence is accumulating that hedgehog protein is made and secreted, and cells are able to respond to it, but the movement of the protein through tissue is abnormal. For example, in the developing wing, hedgehog protein normally travels and acts at a distance of 8–10-cell dimensions from the site of its production. When clones of cells that are mutant for tv are generated in the field of cells that normally respond to hedgehog, only those cells directly adjacent to hedgehog-producing cells receive the signal. Staining for the hedgehog protein suggests it is not being transported through the mutant cells and, indeed, one can observe that wild-type cells on the other side of the mutant clone are also deprived of their hedgehog signal (Bellaiche et al., 1998).

Although hedgehogs are known to bind heparin (Bruno et al., 1995), this implication of HS in hedgehog function was entirely novel. Interestingly, one of the three vertebrate hedgehogs, Indian hedgehog, is primarily known as a regulator of bone development, where it acts to limit chondrocyte differentiation (Vortkamp et al., 1996). It is tempting to speculate that the cartilaginous overgrowth associated with the EXT mutations in man is explained, at least in part, by a disruption of Indian hedgehog function or localization. Whatever the mechanism, the notion that HS plays a major role in cartilage development is becoming increasingly apparent, as was most recently—and unexpectedly—driven home by the finding that abnormalities in cartilage growth are among the most obvious phenotypes resulting from targeted deletion of the mouse gene for the HS perlecain (Costell et al., 1999).

**The Finer Points of Fine Structure**

Enzymes like UDPGDH and the HS copolymerase are required to generate HS chains altogether. Many enzymes in the HS biosynthetic pathway, however, function downstream of these to modify HS, generating patterns of sulfation and sugar isomerization known as fine structure. There is evidence that these modifying enzymes act sequentially, with early steps required for later ones to proceed (Lindahl et al., 1998). The earliest such modification transforms blocks of N-acetylgalactosamine residues into N-sulfoglucosamine. At least three different, but homologous deacetylase/sulfotransferase enzymes, can carry out this function in mammalian cells. However, in Drosophila, a single homologue is known. Like the UDPGDH homologue sgl, it was recovered from a screen for mutations that disrupt wingless signaling, and it has been named sulfatless (sfl) (Lin and Perrimon, 1999).

Curiously, in sfl mutants Dally molecules carry reduced amounts of HS, rather than just HS that is less sulfated (Lin and Perrimon, 1999), a result that suggests some sort of feedback regulation of HS polymerization. Regardless, the dally, sgl, and sfl phenotypes all strongly point to a critical role for HS modifications downstream of N-sulfation. So far the only step at which mutant phenotypes have emerged is the 2-O-sulfation of uronic acid residues (Fig. 1). And once again, flies and mammals have taken center stage almost simultaneously.

On the mammalian side, Bullock et al. (1998) recently described the phenotype of a gene-trap mutation in mice that disrupts the only known HS 2-O-sulfotransferase. These animals exhibit absent kidneys, abnormalities of the skeleton and eye, and perinatal death. Kidney development arrests at a relatively early stage, in which mesenchyme condenses around the ureteric bud and branching morphogenesis begins. Although the factors involved in kidney morphogenesis are numerous, they include multiple HS-binding growth factors, especially Wnts (Kispert et al., 1998).

On the Drosophila side, Sen et al. (1998) reported that the product of the pipe gene is a homologue of the mammalian HS 2-O-sulfotransferase. Pipe is a gene involved in setting up dorsal-ventral polarity in the Drosophila embryo. It is expressed in the ventral follicle cells of the ovary, where it is required for the proteolytic activation of the secreted protein, Spätzle, which occurs after egg deposition. A putative Spätzle engages a receptor in the plasma membrane of the early embryo, establishing a nuclear dorsal-ventral gradient of the NF-kB/Rel-like transcription factor, Dorsal. Pipe is not only necessary for directing a ventral fate, but also sufficient, since misexpression of Pipe in dorsal follicle cells produces a completely ventralized embryo.

Assuming Pipe does encode an HS 2-O-sulfotransferase (which has yet to be established), these findings suggest that a spatially localized HS controls the proteolytic activation of a growth factor. Such a mechanism recalls the well-studied role of heparin and HS in the regulation of proteolysis by thrombin (although, in the case of thrombin the role of HS is to accelerate inhibition, rather than activation, of the protease [Olson and Björk, 1992]).

**New Questions, New Models**

The genetic experiments described above have provided a bountiful harvest of exciting information. Clearly HS modifications are key players in development. Clearly they influence cell–cell signaling and morphogenesis. Yet each answer provokes new questions and raises new cautions.

Perhaps the most important question right now concerns the mechanisms of action of HS modifications. In vitro studies have suggested that some growth factors require HS modifications as coreceptors to bind to, or signal at, their receptors. Drosophila Wg appears to fall into this category (Riechmann et al., 1996), which is likely to explain the effects of dally, sgl, and sfl mutations on Wg function in vivo. FGFs are thought of as the quintessential HS-dependent growth factors, and it is indeed comforting to know that further study of sgl and sfl mutants has shown clear defects in FGF receptor signaling (Lin and Perrimon, 1999).

In contrast, the evidence that HS modifications potentiate Dpp signaling, obtained from the study of dally and sgl mutants, comes as a surprise to those laboring in vitro, whose work has suggested that interactions with HS, if anything, inhibit the functions of BMP2, a mammalian orthologue of Dpp (Ruppert et al., 1996). Researchers are also at a loss for an HS-stimulated growth-inhibitory signaling pathway onto which to pin the abnormalities of SGBS. Whereas it is certainly possible that HS dependence in the activities of certain growth factors has been missed until now, the
rather sharp tissue specificity of several HS mutant phenotypes (e.g., mammalian 2-O-sulfotransferase), together with the curiously tissue-specific effects of dally (see above) suggest that HS dependence is a phenomenon that may be context dependent. To accommodate this idea, simple models such as promotion of growth factor dimerization, or increasing local growth factor concentration by HS, may need to be extended and/or modified (Lander, 1999).

Similarly, the stunning and unexpected defects in Hedgehog transport in tvt mutants also suggest that new thinking about mechanism is called for. Until recently, much has been said about the ability of PGs to arrest or slow the diffusion of ligands, with only a few suggestions that they might accelerate it (Lander, 1998). One possibility is that in the developing fly wing, HSPG s are released from cells in soluble form and, in binding Hedgehog, inhibit its interaction with other cell-surface or extracellular binding sites, thereby allowing it to diffuse more freely. That HSPGs are efficiently shed from mammalian cells has been well established (Bernfield et al., 1999). A nother possibility is that HSPGs attract Hedgehog transport indirectly, by affecting Hedgehog proteolytic processing or cholesterol modification in a manner as yet unknown (but perhaps analogous to their role in processing of Spätzle). A third possibility is that movement of Hedgehog through tissues is not mediated by diffusion at all but by a vesicular transcellular transport, such as has been suggested for Wg (Gumbiner, 1998). To this end, it is interesting that HS HSPG s have been observed to mediate internalization of proteins through highly specific mechanisms (Rieland and Rapraeger, 1993). A ternatively, the fact that one family of HSPG s is GPI-anchored raises the additional possibility of movements of HSPG-attached Hedgehog by direct partitioning of GPI-anchored proteins from one plasma membrane into another (Kooyman et al., 1995).

**Coming Attractions**

The in vivo data of the last few years are having a dramatic impact on the way proteoglycans are viewed by biologists. In response, we are likely to see renewed and expanded efforts by biochemists and cell biologists to address the critical mechanistic questions raised above. In addition, it is important to remember that we have only seen the first wave of animal data on proteoglycans. For genes that have essential functions throughout life, early genetic studies invariably highlight developmental roles since developmental phenotypes often prevent adulthood from being reached. Pinning down the adult functions of HSPG s through genetic means will require more effort, but is feasible. Given the intriguing data over the years that have linked HSPG s to cell adhesion (LeBaron et al., 1988), wound healing (Bernfield et al., 1999), microbial invasion (van Putten and Paul, 1995), viral infection (Shieh et al., 1992), lipoprotein metabolism (Ji et al., 1997), and cancer (Kleef et al., 1998), it is likely that such an effort would be handsomely rewarded.

Submitted: 30 Aug 1999
Revised: 8 Dec 1999
Acepted: 9 Dec 1999

**References**


