seminars in **CELL & DEVELOPMENTAL BIOLOGY**, Vol. 12, 2001: pp. 65–67
doi:10.1006/scdb.2000.0237, available online at http://www.idealibrary.com on IDEAL®

Cellular functions of proteoglycans—an overview

Norbert Perrimon[a] *and Merton Bernfield*[b]

Introduction

This issue of *Seminars in Cell and Developmental Biology* contains a number of articles reviewing the cellular functions of proteoglycans. Proteoglycans are proteins, substituted with glycosaminoglycans (GAGs), that generally either decorate the cell membrane or fill the extracellular space. GAGs are linear polysaccharides consisting of a repeating disaccharide, generally of an acetylated amino sugar alternating with a uronic acid. Although there are intracellular types, most proteoglycans are destined for the extracellular space. These are very diverse molecules, and various combinations of both different types of proteins [matrix, cell surface transmembrane or covalently linked to membrane glycosylphosphatidylinositol (GPI)] and classes of GAG chains [hyaluronan (HA), chondroitin sulfate (CS), keratan sulfate (KS), dermatan sulfate (DS) and heparan sulfate (HS)] are found in vertebrates. For example, the syndecan transmembrane proteins can be decorated with two HS and one CS chains, while the aggrecan cartilage matrix protein is typically decorated by about 100 CS and 30 KS chains, and decorin, a major proteoglycan of the interstitial matrix, contains a single DS chain. Further, GAGs may decorate only a proportion of a type of protein. The function of the GAG substituent in these so-called 'part-time proteoglycans' is rarely understood.

Proteoglycans can be very abundant. At the surface of epithelial cells, it has been estimated that there may be as many as 1 million syndecan-1 molecules, and cartilage matrix is a several mg per ml composite of aggrecan with HA. HA, the structurally most simple GAG, is not an authentic proteoglycan, as it is not covalently associated with protein. But it is considered here because it is highly abundant in vertebrate embryos and tightly associates with a wide variety of proteins in performing major developmental functions.

The structural diversities of proteoglycans underlie their various functions. For example, via their HS chains, syndecans can bind growth factors, extracellular matrix components, enzymes, protease inhibitors, chemokines among other extracellular constituents, and play an active role in signal transduction. Others, such as aggrecan or versican, are structural components of the extracellular matrix that associate tightly with both HA and proteins.

Extensive biochemical studies of proteoglycans have provided a wealth of information on the structure of these molecules. Much of the early structural work was done by Karl Meyer on HA and by Jorpes and Gardell on HS. The pioneering work of Hascall and Sajdera taught us much about the structure of cartilage proteoglycans, and that of Kraemer showed that every adherent cell contains HS at their surfaces. Dorfman, Silbert, and Lindahl performed now classic studies on GAG biosynthesis. Despite their widespread distribution and this wealth of information, until relatively recently little was known about proteoglycan function beyond their structural role in organizing cartilage and the basal lamina.

In recent years, understanding of proteoglycan functions has been obtained from cell biological and genetic studies. These studies have provided many surprises and have propelled the study of proteoglycans to the forefront of modern biomedical research. In particular, the importance of these molecules has been highlighted by the discovery that a number of human diseases, such as the Simpson-Golabi-Behmel syndrome (SGBS), a rare pre- and post-natal overgrowth, birth defect and tumor susceptibility syndrome associated with mutations in a GPI-linked HS proteoglycan, and multiple hereditary exostoses (EXTs), bony tumors that can undergo malignant transformation, associated with mutations in an HS biosynthetic enzyme. Further, genetic

From the [a] *Department of Genetics, Harvard Medical School, Boston, MA 02115, USA and* [b] *Departments of Pediatrics and Cell Biology, Harvard Medical School, Children's Hospital, Boston, MA 02115, USA.*

1084–9521/01/020065+ 03/$35.00

analyses in the mouse and *Drosophila* model systems have provided novel insights into the function of these molecules during development.

The various chapters illustrate: (1) the role of some proteoglycans in organization of the extracellular matrix (Knudson and Knudson, Toole); (2) the role of proteoglycans in metastasis and invasion of tumor cells (Sanderson) and cell differentiation in the nervous system (Yamaguchi); (3) the role of membrane attached heparan sulfate proteoglycans (HSPGs) in regulation of cell signaling events (Rapraeger, De Cat and David); and (4) Genetic evidence for proteoglycan function in invertebrate models (Selleck).

Proteoglycans and extracellular matrix

Proteoglycans constitute a major component of the extracellular matrix and have been particularly well analyzed in the context of cartilage where the matrix constitutes more than 90% of the dry weight of the tissue. Knudson and Knudson review the role of proteoglycans in organizing cartilage and point out that it is the unique mixing of several proteoglycans together with their organization within the extracellular matrix that gives cartilage its unique physical properties. In particular, they discuss the interactions that illustrate the complex mode of association between a proteoglycan (aggrecan), a GAG (HA), and a protein (link protein) in the cartilage extracellular matrix. HA is a linear GAG composed of alternating residues of glucuronic acid and N-acetylglucosamine which is found at the cell surface, in the extracellular matrix and within cells. It interacts with many proteins (hyaladherins) that modify its structural and physiological properties. Knudson and Knudson also provides a thorough description of other proteoglycans found in cartilage matrix.

The developmental roles of HA are further discussed by the paper of Toole that reviews the function of this GAG at the surface of mesenchymal cells. Toole describes the motifs found in the multiple hyaladherins that interact with HA, and then describes how changes in the hydration of the pericellular matrix influence the physical nature of HA, and regulate signal transduction events as well as cell behaviors such as mitosis and cell migration. Altogether, the reviews of Knudson and Kundson and Toole illustrate how changes in the physical characteristics of a GAG can influence a variety of cellular events.

Cellular function of proteoglycans

All cellular processes that involve molecular interactions at the cell surface, such as cell–matrix, cell–cell and ligand–receptor interactions, likely involve proteoglycans because these molecules avidly bind proteins and are quite abundant at this site. This diversity of interactions is well illustrated during the development of the nervous system and concomitant with tumor invasion and metastasis, instances where cell movements and cell extensions are the major cellular behaviors.

Yamaguchi describes the role of HSPGs during three critical phases of the development of the mammalian nervous system. He reviews their roles in the generation and differentiation of neurons from stem cells, in axonal guidance and in synapse development. In particular, the roles of cell surface HSPGs and growth factors in these processes are described.

The critical aspects of tumor cell behavior that influence the health of an organism are local invasion and distant metastasis. These behaviors depend on cell adhesion, motility and growth, each readily affected by HSPGs. Sanderson reviews these effects of proteoglycans, and emphasizes that HSPGs can either promote or inhibit these processes depending on the tissue type, the pathophysiological state of the tumor, and the step within the metastatic cascade that is affected.

Proteoglycans and signal transduction

Two major families of cell surface HSPGs, syndecans and glypicans, have been identified. These bind a multitude of growth factors and extracellular matrix molecules, and have been implicated in several signal transduction pathways that regulate cell proliferation and cell shape. While syndecans are transmembrane proteins with an intracellular cytoplasmic domain, glypicans are attached to discrete cell membrane regions by a GPI lipid anchor.

Rapraeger details the function and various binding partners of the four known mammalian syndecans. The syndecan core proteins are structurally quite similar but have distinct regions and cellular distributions, but share conserved cytoplasmic, juxtamembrane and transmembrane domains. Both the extracellular domain and a region within the cytoplasmic domain are divergent underlying the

diverse functions of these molecules. Rapraeger also describes both the conserved and divergent partner proteins associated with syndecan core proteins and relates these to the cellular and developmental functions of these proteoglycans.

DeCat and David review the structure and function of the glypicans. Recent studies in *Drosophila*, humans and mice have implicated these cell surface molecules in the control of cell growth and differentiation. In particular they have been implicated in regulating the signaling pathways of molecules such as FGFs, BMPs, Wnts, Hhs and IGFs. Although the detailed molecular mechanisms by which these molecules operate is not yet understood, one attractive possibility is that they activate or determine the activity-ranges of morphogens and growth factors.

Proteoglycan function in model organisms

Much of the enhanced current interest in proteoglycans results from genetic studies of the fruit fly, *Drosophila melanogaster*, and the nematode worm, *Caenorhabditis elegans*. Identification of a syndecan homolog in the fly was initially a surprise, but subsequent work, reviewed by Selleck, has revealed multiple physiological roles for cell surface HSPGs in these organisms. Despite major differences in physiology and development, the functions of the two *Drosophila* glypicans and single syndecan are often mimicked in mammals by their six glypicans and four syndecans, providing a clear example of conservation of function during the two genome duplications that transpired between the times when these organisms emerged.

Concluding remarks: whither proteoglycans?

Proteoglycans are ubiquitous, highly abundant, chemically complex and functionally disparate. They differ from proteins and other glycoproteins by their GAG chain substituents. While in some instances their functions reflect the unique chemical properties of the linear anionic polysaccharide GAGs (e.g. binding of water by aggrecan to provide cartilage with the ability to resist compression), in other instances the rationale for Nature employing a GAG chain is not clear (e.g. the co-receptor role of cell surface HSPGs when several cell surface proteins similarly serve as co-receptors). Further, the evolutionary advantage for synthesizing GAG chains, which appear to require specific sequences to interact selectively with proteins, by an enzymatic, non-template mechanism is quite unclear. The resulting sequence complexity of GAG chains has all the characteristics of an informational code, one which we cannot as yet read, but which is used by the very large variety of extracellular proteins that bind to GAG chains. Importantly, except for HA, no GAG-containing proteoglycans exist in prokaryotes or in yeast, and it appears that proteoglycans arose evolutionarily with the emergence of multicellularity in the metazoans. A critical role for proteoglycans in survival is inferred from the retention of so many proteoglycan encoding genes in higher vertebrates.

These considerations suggest that the key cellular role of these molecules is to mediate cell interactions, whether direct, via soluble mediators or via the insoluble extracellular matrix. Once we understand the rationale for these molecules and decipher the code, the reasons for their ubiquity and abundance should become apparent and we should be able to modulate their interactions for therapeutic purposes. With so much more to learn, the next several years will be exciting for proteoglycan research.

Acknowledgement

We thank Charlotte McManus for organizing and compiling these essays.