IDENTIFICATION, MOLECULAR CLONING, and CHARACTERIZATION OF CEREBROGLYCAN, A CELL SURFACE HEPARAN SULFATE PROTEOGLYCAN of the DEVELOPING RAT BRAIN

by

Christopher S. Stipp

B. S. Biochemistry
Indiana University, 1989

Submitted to the Department of Biology in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in Biology

at the
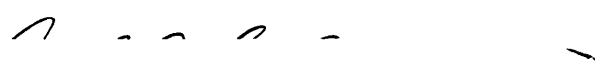Massachusetts Institute of Technology
February, 1996

© 1996, Christopher S. Stipp. All rights reserved

The author hereby grants to MIT permission to reproduce and to
distribute pubilicly paper and electronic copies of this thesis
document in whole or in part.

Signature of Author

Christopher S. Stipp, Department of Biology

Certified by

Arthur D. Lander, Associate Professor, Depts. of Biology
and Brain and Cognitive Sciences, Thesis supervisor

Accepted by

Frank Solomon, Chairman of the Graduate Committee,
Department of Biology

IDENTIFICATION, MOLECULAR CLONING, and CHARACTERIZATION OF CEREBROGLYCAN, A CELL SURFACE HEPARAN SULFATE PROTEOGLYCAN of the DEVELOPING RAT BRAIN

by
Christopher S. Stipp

ABSTRACT

Many of the molecules that are thought to be involved in nervous system development have the property of binding to proteoglycans (PGs). Furthermore, PGs have been shown to modulate the biological activity of many of these molecules in vitro. Taken together, these observations suggest that PGs play roles in nervous system development. Characterization of proteoglycan function and expression in the nervous system is still in the early stages. A recent report revealed that as many as 25 different PG core proteins are found in the rat brain, many with distinct patterns of developmental expression (Herndon and Lander, 1990). The surprising diversity of PG core proteins revealed by this study supports the hypothesis that individual proteoglycans have been adapted to fufill distinct functions in the developing nervous system. To test this hypothesis and to begin to identify potential functions of PGs in the nervous system, we developed a purification strategy to identify the major cell surface heparan sulfate proteoglycans (HSPGs) of the developing rat brain. We report here on studies identifying one of these molecules as a novel cell surface HSPG related to the previously identified cell surface HSPG, glypican. This molecule, which we have named cerebroglycan, is restricted in expression to the developing nervous system. In situ hybridization experiments demonstrated that cerebroglycan is transiently expressed by immature post-mitotic neurons during the

period encompassing axon outgrowth and cell body migration, suggesting a possible role for cerebroglycan in these processes. Immunolocalization experiments revealed that cerebroglycan core protein is strongly localized to the fibre tracts of the developing brain, consistent with a role for cerebroglycan in the growth of axons. Biochemical and structural characterization of cerebroglycan demonstrated that cerebroglycan binds strongly to laminin-1, a molecule thought to be involved in axon guidance. The strong binding may be due, in part, to an interaction of the cerebroglycan core protein with laminin-1. The data described here support the hypothesis that individual PGs fufill distinct functions in the developing nervous system, and are consistent with the possibility that cerebroglycan participates in the growth or guidance of axons.

Thesis superivisor: Arthur D. Lander
Title: Associate Professor
       Departments of Biology and Brain and Cognitive Science

This thesis is dedicated to my teachers and mentors

Daniel Wunderlich
Dennis Peters
and Judith Jaehning

Acknowledgements

Table of Contents

List of Figures and Tables

Chapter 4: High affinity interaction of Cerebroglycan with
    Laminin-1

Chapter 5: Determination of the molecular weight of intact
    Cerebroglycan

CHAPTER 1

PROTEOGLYCAN STRUCTURES AND POTENTIAL ROLES IN NERVOUS SYSTEM DEVELOPMENT

I. Introduction

Proteoglycans (PGs) are simply proteins that have been substituted with one or more glycosaminoglycan (GAG) carbohydrate chains. As will be discussed in the section II below, a GAG is a linear polymer of repeating disaccharides. Several types of GAG exist, each defined by the type of disaccharide of which they are composed. For the familiar GAG heparin (and the related GAG heparan sulfate), the disaccharide subunit is a glucuronic or iduronic acid linked (1-4) to N-acetyl or N-sulfoglucosamine. For chondroitin sulfate (another commonly encountered GAG) the subunit consists of a glucuronate linked (1-3) to N-acetylgalactosamine. Most types of GAG chains undergo a series of modifications in the Golgi apparatus that encompass (1) N and O-sulfations, and (2) epimerizations, resulting in a heterogenous, highly negatively charged structure.

PGs have traditionally been categorized on the basis of the type of GAG chains they bear: a PG core protein that is substituted with a heparan sulfate (HeS) chain is a heparan sulfate proteoglycan (HSPG), and a core protein substituted with chondroitin sulfate (CS) is a CSPG, etc. In recent years, a number cDNA clones for PG core proteins have been isolated, and it has become apparent that there are distinct families of PGs. Table 1 summarizes the families of PGs expressed in the nervous system; many of the PGs in the table will be discussed in section IV of this chapter with regard to their potential ligands and roles in development.

PGs have attracted the interest of developmental biologists as a result of the very large repertoire of molecules with which they can interact. These interactions are typically mediated by the GAG chains, and, have often been found to involve the binding of sulfated residues on the GAG chain to clusters of basic amino acids in the GAG-binding ligand [reviewed in (Cardin and Weintraub, 1989; Jackson et al., 1991)]. The molecules that PGs bind include growth factors such as members of the FGF family; extracellular matrix glycoproteins such as laminin, fibronectin, and thrombospondin; proteases such as thrombin, urokinase, and tissue plasminogen activator; and protease inhibitors such as anti-thrombin III, and protease nexin-1 [reviewed in (Lander, 1989; Lander, 1990; Lander and Calof, 1993)]. Many of these

molecules have been implicated in developmental processes and cell behaviors such as cell proliferation and differentiation, cell adhesion and migration, axon outgrowth and synapse formation, and metastasis and tumor formation. PGs, by the virtue of being able to bind to so many ligands of interest, have become suspected of being "guilty by association" of participating in these same biological processes.

Section III of this chapter will discuss the binding of PGs to their ligands. Section IV will review data on PG expression and evidence that PGs actually do play functional roles in development, with special reference to the nervous system.

II. Glycosaminoglycan Structure: Heterogeneity Resulting from a Complex Biosynthesis

Figure 1 is a schematic comparison of syndecan family and glypican family PGs, and serves to illustrate several features of PG and GAG structure. The biosynthetic steps that lead to these structures are outlined below.

As reviewed in Gallagher et al. (1992), GAG biosynthesis begins in the ER with the synthesis of a PG core protein. Possibly while still in the ER, the hydroxyl group of one or more serine residues on the core protein is substituted with xylose; these serines will ultimately be linked to GAG chains. For glypican family HSPGs, the ER is also the site of lipid attachment. Immediately after translation, a C-terminal cleavage and addition signal is recognized and a glycosyl-phosphatidylinositol (GPI) lipid anchor is attached (Low, 1989). The rules governing GPI attachment have not been completely worked out, but the attachment signal includes a hinge region, rich in prolines and basic amino acids, followed by a hydrophobic region (Kodukula et al., 1993).

The protein sequence requirements for GAG attachment have also not been fully elucidated; however, serines that serve as GAG acceptors are always found adjacent to glycine, and these ser-gly attachment sites are usually preceded and often flanked by acidic amino acid residues. Several more specific motifs have been suggested such as $(SG)_n$-$(X)_{2-5}$-D/E and D/E-XSG-$(X)_n$-D/E (Bernfield

et al., 1992); however no single rule has been found to hold true for all sites.

Two galactose residues and a glucuronate (GlcA) are added to the serine-linked xylose, completing the tetrasaccharide link that will join the GAG chain to the core protein. The GAG is polymerized in the Golgi upon this tetrasaccharide by the successive addition of hexuronate and hexosamine residues which occur as UDP-linked precursors. Heparin and heparan sulfate are polymerized by the alternate addition of GlcA and N-acetylglucosamine (GlcNAc); chondroitin sulfate is polymerized by successive addition of GlcA and N-acetylgalactosamine (GalNAc). Both CS and HeS undergo further modifications in the Golgi. These modifications include N-deacetylation followed by N-sulfation (for HeS), and the epimerization of GlcA to iduronate (IdoA) as well as O-sulfations (for both HeS and CS). Unfortunately, when CS is modified by epimerization of GlcA to IdoA it is identified as a new class of GAG, dermatan sulfate (DS). This designation was in place before it became apparent that CS regions and DS regions often occur on a single GAG chain. Figure 1 shows the structures of some representative disaccharides from an HeS and a CS/DS chain. For HeS, in addition to N-sulfation, O-sulfation occurs at position C2 of IdoA residues, C6 of glucosamine residues, C3 of N-sulfoglucosamine residues (GlcNSO3) and C2 of GlcA residues (although these latter two sulfations are rare compared to the others). CS/DS can be sulfated at C4 or C6 of the GalNAc residue or at C2 of IdoA and GlcA residues. In HeS saccharides are linked (1-4) throughout, while in CS/DS, uronic acid is linked (1-3) to a galactosamine that is linked (1-4) to the next uronic acid, and so on.

Not illustrated in Figure 1 are the structures of two other GAGs, keratan sulfate (KS), and hyaluronic acid (HA). KS is composed of repeating [galactose-GlcNAc] subunits, and has the unusual property of being polymerized on both N-linked and O-linked carbohydrate templates (Hounsell, 1989; Stuhlsatz et al., 1989). HA is composed of [GlcA-GlcNAc] disaccharide subunits, and is found unattached to protein. HA is the only GAG that does not contain sulfate groups [reviewed in (Jackson et al., 1991)].

An interesting feature of HeS structure, illustrated in Figure 1, is the striking domain structure of the HeS chain. HeS chains, which range in the size from 30 to 200 disaccharides (14-100 kDa), are characterized by regions of sparsely modified residues punctuated by islands of highly modified residues (Gallagher et al., 1992). In PGs isolated from human skin fibroblasts, the highly modified regions were on average about 5 disaccharides in length, and these regions were separated by spans of ~18 disaccharides that were very sparsely modified (Gallagher et al., 1992). The sparsely modified regions consisted mostly of disaccharides containing GlcNAc residues, while the GlcNSO$_3$ residues were typically restricted to the highly modified zones. In addition, O-sulfations were found to be restricted to a subset of residues within the N-sulfated regions.

In heparin, such a domain structure is not so apparent. Instead, most disaccharides are found to be highly modified. The domain structure of HeS, and the extensive but incomplete modifications found within the highly modified domains, give HeS a more complex structure than the other GAGs, arguably more complex even than heparin, which, although more extensively modified, does not contain such a striking domain structure.

The domain structue of HeS can be explained by the nature of HeS biosynthesis. The guiding principles of HeS synthesis were first elucidated by studying the synthesis of heparin in a cell-free system derived from a mouse mastocytoma cell line [for review see (Lindahl et al., 1986)]. In this system, sulfation can be controlled by whether or not the sulfate donor 3'-phosphoadenylylsulfate (PAPS) is added, and the products can be labelled to facilitate later analysis by including UDP-[^{14}C]-monosaccharide precursors in the reaction [cf (Lindahl et al., 1973; Kusche et al., 1988)].

Using this system, it was found that heparin biosynthesis occurs in a distinctly stepwise fashion in which the product of each step forms the substrate for the next step. The sequence of the biosynthetic steps has been established: (1) heparin is polymerized as an unmodified polymer of [GlcA-GlcNAc] residues, as described above; (2) the polymer is modified by N-deacetylation, and (3) N-sulfation; (4) disaccharides containing GlcNSO$_3$ are modified by epimerization of

GlcA at C5 to produce IdoA; (5) IdoA is O-sulfated at C2; and (6) GlcNSO$_3$ is O-sulfated at C6. The rare O-sulfations at the C3 position of GlcNSO$_3$ appear to be the final modification for those residues. This rare modification has been shown to be crucial for the formation on heparin and HeS chains of a binding site for antithrombin III (ATIII). The binding of ATIII to HeS and heparin will be discussed further in section III of this chapter. It should be noted that even in heparin, the above modifications do not occur completely along the length of the GAG chain. At each step of the pathway, a fraction of the available substrates are further modified by the next biosynthetic enzyme(s). The result is that complexity is multiplied as biosynthesis proceeds.

The biosynthesis of heparin, as described above, provides a compelling explanation for the domain structure observed for HeS. Since each biosynthetic step requires the modification of the previous step to form its substrate, early modifications control the locations of subsequent modifications along the length of the HeS chain. In heparin, the early steps of N-deacetylation and N-sulfation are extensive, resulting in extensive subsequent modifications. In HeS, the commonly encountered form outside of mast cells, these early modification steps appear to be restricted, resulting in the localization of subsequent modifications to discrete domains. The fact that, as described above, O-sulfations appear to be restricted to regions of N-sulfated residues is also consistent with stepwise HeS synthesis. Another interesting consequence of this scheme of biosynthesis is that the highly modified regions are rich in IdoA whereas poorly modified regions contain mostly GlcA. NMR studies of IdoA groups has demonstrated that IdoA is actually an equilibrium mixture of three different conformations (Casu et al., 1988). This apparent conformational flexibility has important implications in the binding of GAGs to their ligands. Potentially, the added flexibility of these domains could facilitate GAG-protein binding by giving functional groups on the GAG chain freedom to search out binding partners in the protein's GAG binding site.

A qualification of HeS synthesis as described above is that it represents an extrapolation of data for heparin synthesis in a cell free system in vitro. The question of the extent to which the rules worked

out for heparin in vitro will apply to HeS synthesized in vivo has recently been addressed (Bame et al., 1991). In this study the structure of HeS in a Chinese hamster ovary (CHO) cell line deficient in N-sulfotransferase was examined. The mutant CHO cells produced HeS that was 3-5 fold undersulfated at the N-groups of glucosamines, compared to wild type (wt) cells. Although the mutant cells expressed even more O-sulfotransferase activity than wt cells, the HeS of the mutant cells was also significantly undersulfated at all the O-sulfation sites, and the amount of IdoA was also substantially reduced. Thus, as had been observed for heparin, these later modifications appear to depend on the earlier steps of N-deacetylation and N-sulfation, supporting the view that HeS synthesis also proceeds in a stepwise fashion. An important implication of these studies is that regulation of the sites of N-sulfation is likely to be a major mechanism of regulating the overall structure of HeS. The factors responsible for the spacing of N-sulfation in HeS chains are currently unknown.

In addition to the regulation of HeS structure arising from stepwise biosynthesis, there appear to be higher order mechanisms in operation. A variety of studies have reported changes in HeS structure that correlate with developmental stages or the growth state of cells. For example, the size of HeS chains attached to the cell surface HSPG syndecan appears to vary dramatically with developmental age in the mouse lung epithelium, decreasing in size as development proceeds (Brauker et al., 1991). Tissue specific variations have also been noted: syndecan HeS chains in simple epithelia appear to be significantly larger than the HeS chains of syndecan isolated from stratified epithelium (Sanderson and Bernfield, 1988). Monoclonal antibodies that recognize specific HeS-associated epitopes gave distinct staining patterns in developing hamster embryos that varied with developmental age and tissue type, and that only partially overlapped (David et al., 1992). The non-overlapping aspect of the staining could be the result of a core-restricted expression of the GAG epitopes, or a cell type specific expression of the epitopes on multiple cores, but, in any case, it is suggestive of a higher order regulation of HeS structure. Additional strong evidence for such regulation comes from structural

analysis of HeS from transformed cells, which has consistently revealed aberrant patterns of sulfation (Gallagher and Lyon, 1989).

The basis of the developmental, tissue, and growth state specific alterations in HeS structure described above is unknown, but in one recent study, indirect evidence for the existence of a regulatory factor that is probably not one of the biosynthetic enzymes has been obtained (Shworak et al., 1994b). In this study, the effect of PG overexpression on the synthesis of the ATIII binding site on HeS chains was examined. The ATIII binding site, as will be discussed in section III below, is a specific pentasaccharide sequence that contains multiple specific O-sulfations that are absolutely required for high affinity ATIII binding, including a rare 3-O sulfation [reviewed in (Bjork and Lindahl, 1982)]. The binding of heparin or HeS to ATIII greatly potentiates the ability of ATIII to inactivate serine proteases, such as thrombin, in the coagulation cascade. Overexpression of the cell surface HSPG ryudocan (also called syndecan-4) in endothelial cells and in L cells resulted in an overall increase in total HeS synthesized, but a decrease in the absolute amount of ATIII binding HeS (HSact). This result suggests that a factor required for the synthesis of HSact is limiting, and further suggests that the factor contains multiple components since HSact levels actually decline rather than simply plateauing. The additional observation that in all other respects, including the total synthesis of 3-O-sulfated residues, HeS synthesis appears normal in these cells suggests that the limiting factor may be a regulatory element rather than a biosynthetic enzyme (Shworak et al., 1994b). Intriguingly, overexpression of a ryudocan core protein with GAG attachment sites eliminated had a similar effect on the levels of HSact produced in these cells, suggesting that the ryudocan core protein may interact with the putative regulatory control element. Interactions mediated by PG core proteins have been noted before [eg binding of collagen fibrils by the core protein of the CSPG decorin (Vogel et al., 1984), and the binding of TGF-β by the core protein of the HSPG betaglycan (Cheifetz and Massague, 1989)], and data will be presented in Chapters 4 and 5 of this thesis that suggests that PG core interactions with PG ligands may be more widespread than previously appreciated.

As discussed below, subsequent to the discovery of a specific ATIII binding site in heparin and HeS, evidence for specific recognition of particular HeS sequences by other GAG binding molecules has begun to accumulate. Much effort has been focussed on the ligand binding properties of particular PG core proteins on the cell surface and in the extracellular matrix. It is refreshing to note that, in the end, a deeper understanding of the biological processes in which PGs are involved will also require us to learn more about what's going on inside the cell, in the Golgi apparatus, where PG structure is determined.

III. Interactions of PGs with their Ligands

In the discussion below, considerable weight is given to the interactions of HeS and HSPGs with their ligands at the expense of CS and CSPGs. The binding of HeS to GAG binding sites on target molecules typically occurs with higher affinity than the binding of CS, and so is more easily detected and more readily studied. Some notable exceptions include the binding of thrombospondin-1 (TSP-1) and tenascin (TN) to CS, which will be discussed below. Since HeS interactions with proteins have been more intensively studied, they will be considered first, to illustrate some of the underlying principles in GAG-protein interactions.

Table 2 gives a partial list of molecules known to bind GAGs. Many of these molecules were identified as containing heparin-binding regions. In vivo, the more likely ligand for these molecules is heparan sulfate (HeS), which is expressed by virtually all adherent cell types. Heparin in vivo is largely restricted to mast cells. In the discussion below, these sites will often be referred to as HeS binding sites. HeS typically binds to these molecules with a somewhat lower affinity, but affinities can approach those observed for heparin (Binding of brain HeS relative to heparin for a number of molecules has been characterized by Mary Herndon; manuscripts in preparation).

General Features of GAG-Protein Interactions

No one feature of protein structure has been identified that is common to all HeS-binding domains. HeS-binding regions have been identified in a number of protein folding motifs [reviewed in (Lander,

1994)] including fibronectin type III repeats, immunoglobulin domains, and collagen triple helices. In addition, a large number of proteins with unique or uncharacterized protein structures also contain HeS-binding regions. Given that HeS and CS are highly sulfated, polyanionic molecules, it is not surprising that many studies have revealed HeS-binding motifs that are rich in basic amino acids. Early proposals for heparin-binding consensus sequences included BBXB and BBBXXB where B represents a basic amino acid residue (Cardin and Weintraub, 1989). More recently it has become appreciated that HeS-binding domains can be more complex than a contiguous stretch of a few basic residues (Cardin and Weintraub, 1989). HeS-binding elements at distant locations on a polypeptide chain can be brought together to form the complete HeS-binding domain. [cf platelet factor 4 binding to heparin (Stuckey et al., 1992), In addition, some heparin-binding molecules appear to contain heparin-binding sites that are not composed of basic amino acids [e.g. collagen types I, II, and III (San Antonio et al., 1994); and thrombospondin (Adams and Lawler, 1993)].

In addition to the GAG-binding sites on proteins, a complete picture of GAG-protein interactions requires a description of protein-binding sites on GAGs. It has long been appreciated that specific saccharide sequences within GAG chains can form binding sites for particular proteins. The classic example is the ATIII binding site in heparin (see below for more details of ATIII-heparin interactions). This site was elucidated by depolymerization of heparin with nitrous acid and fractionation of the resulting oligosaccharides on an ATIII affinity column [for review see (Bjork and Lindahl, 1982)]. A small fraction of the total oligosaccharides was found to bind to ATIII with high affinity. The surprising result was that, despite the heterogeneity of heparin as a whole, the composition of the oligosaccharides with high affinity for ATIII was remarkably constant. Subsequently, the structure of these oliogosaccharides was analyzed using an arsenal of chemical and enzymatic techniques, with the end result being that a specific pentasaccharide sequence was identified that comprised a binding site for ATIII [reviewed in (Marcum and Rosenberg, 1989)]. The structure of the ATIII binding site is shown in Figure 2 [after

[Jackson et al., 1991)]. A striking feature of the structure is the presence of a rare 3-O-sulfate on the central GlcNSO3 residue. In fact the majority of the 3-O-sulfates present in heparin are now known to be within ATIII binding sites. Removal of this group by a 3-O-sulfatase results in a dramatic decrease in affinity for ATIII.

Recently, the emerging story of HSPG involvement in FGF binding to FGF receptors (see below for details) has sparked interest in determining the structural requirments for HeS binding to FGF. Using techniques similar to those described for ATIII, the HeS binding sites for FGF-1 and FGF-2 have been elucidated. As shown in Figure 2, these sites appear to have less stringent structural requirements than the ATIII binding site. Nevertheless, they are not identical. The preferred binding site for FGF-1 includes a 6-O-sulfate on GlcNSO3 (Mach et al., 1993) whereas the binding site of FGF-2 does not (Turnbull et al., 1992).

Another recent study suggests that specific interactions between GAGs and proteins may be commonplace (San Antonio et al., 1993). In this study, ^{125}I-labelled, low molecular weight heparin (Mr ~ 6000) was fractionated by electrophoresis through agarose gel lanes containing the extracellular matrix molecules laminin-1 (LN-1), fibronectin (FN-1), type I collagen (CI), and thrombospondin (TSP-1). (This technique, called affinity coelectrophoresis or ACE, has been described previously (Lee and Lander, 1991); see also Chapter 4 of this thesis for a description). LN-1, FN-1, and CI were all found to fractionate low molecular weight heparin into high and low affinity fractions. Fractions with high affinity for any one of these molecules were also found to bind with high affinity to the others. However, high affinity binding could not be explained on the basis of net charge, or the presence of ATIII binding sites. Thus there appear to be distinct sequences within heparin that are bound preferentially by extracellular matrix proteins.

To illustrate the complexity and the functional consequences of GAG-protein interactions, two well-characterized examples will be considered next. Additional examples of GAG/PG binding will be discussed in section IV below.

Antithrombin III, a Prototypical Heparin/HeS Binding Molecule

One of the most extensively characterized HeS-binding proteins is ATIII. The studies of ATIII binding to heparin [as reviewed in (Jackson et al., 1991)] provide an excellent example of both the potential complexity of GAG-protein interactions, and the types of techniques that can be brought to bear the problem. One potential heparin binding region in ATIII was identified by studies in which proteolytic fragments of ATIII were isolated that retained some heparin binding capacity. These fragments encompassed a region from residues 124-145 that contains several basic amino acid residues (Rosenfeld and Danishefsky, 1986; Smith and Knauer, 1987). Antibodies raised against peptides based on this region were found to block heparin binding to ATIII, and, interestingly, were also shown to be partially able to substitute for heparin in accelerating the inactivation of thrombin by ATIII (Smith et al., 1990). When the region is modeled as an α-helix, helical wheel presentation reveals that the basic amino acids in this region fall on one face of the helix, and hydrophobic amino acids on the opposite side. Chemical modification of the lysines in this region results in a loss of heparin binding and thrombin inactivation ability (Liu and Chang, 1987; Peterson et al., 1987). Interestingly, pre-binding of heparin protects some of the lysines in the 124-145 region from modification, but results in exposure of other lysines outside this region that had been previously resistant to modification, suggesting a conformation change in ATIII upon heparin binding (Chang, 1989).

Another region of ATIII apparently important for heparin binding was identified on the basis of mutations resulting in disfunction. Thus several hereditary forms of ATIII have been isolated in which Arg-47 has been substituted with another amino acid. These forms exhibit decreased heparin binding and decreased thrombin inactivation (Jackson et al., 1991).

The above findings have led to a view in which the initial binding of heparin to ATIII induces a conformational change that results in the formation of the complete heparin binding domain required for high affinity heparin binding (Jackson et al., 1991). Studies that demonstrate that reduction of disulfide linkages reduces heparin

binding by ATIII, and conversely that heparin binding renders disulfide bonds resistant to reduction, support this view (Sun and Chang, 1989), as do spectroscopic studies that demonstrate an apparent conformational change upon heparin binding [cf (Bjork and Lindahl, 1982; Evans et al., 1992)].

Heparin accelerates the inactivation of thrombin by ATIII by as much as 2000-fold. The contribution of the heparin-induced conformational change to ATIII's activity towards thrombin has been the subject of some debate. Two opposing models were advanced early on, the conformation model and the template model. In the conformation model, the enhancement of ATIII inactivation of thrombin is attributed to the conformational change induced by heparin binding, a change that exposes the active site for more efficient interaction with thrombin (Rosenberg and Damus, 1973). The template model emphasized the fact that heparin also binds to thrombin [cf (Holmer et al., 1979)]. In this model the effect of heparin is to provide a template to which both ATIII and thrombin bind, bringing the molecules into close apposition and thereby accelerating the rate of the reaction. As is often the case for the dichotomies created by scientific debate, evidence has accumulated suggesting that both models are correct. In one recent study using heparin oligosaccharides of defined size, a sharp threshhold for the enhancement of thrombin inactivation by ATIII was observed at the size range of 14-16 disaccharide units. This size requirement favors the template model. However, a similar size requirement was not observed for accleration of ATIII inhibition of the protease, factor Xa, nor for the acceleration of heparin cofactor II inhibition of thrombin, or the potentiation of protein C inhibitor's interaction with several serine proteases (Pratt et al., 1992). The findings emphasize that although patterns may recur, heparin's effect (and that of HeS) on interactions of proteases and their inhibitors will vary from case to case. In the case of the ATIII-thrombin interaction, both a conformational change and the "template effect" appear to be important. In another recent report, the pentasaccharide sequence of heparin that binds ATIII with high affinity caused by itself a conformational change and a partial acceleration of ATIII inactivation

of thrombin. The further addition of low-affinity heparin chains of higher molecular weight potentiated the rate of inactivation to an extent nearly equal to that of intact high affinity heparin chains (Evans et al., 1992).

Binding of PGs to Members of the FGF family: The Coreceptor Hypothesis

A recent incarnation of the debate described above involves the coreceptor hypothesis. In this model, the HeS chains of an HSPG participate in the binding of a ligand to another high affinity, non-PG receptor (Bernfield et al., 1992). Recently, strong evidence in support of the coreceptor hypothesis has come from studies of the binding of FGF-2 (bFGF) to its high affinity receptor. FGF-2 fails to bind to Chinese hamster ovary (CHO) cells deficient in HeS production, even though these cells were transfected to overexpress FGFR-1, a high affinity receptor for FGF-2. The addition of soluble heparin or HeS reconstitutes an FGF-2/FGFR-1 interaction on the surface of these cells (Yayon et al., 1991). In a similar set of experiments treatment of cells with heparitinase (which cleaves HeS chains) or growth of cells in chlorate (which inhibits sulfation of GAG chains) resulted in a loss of FGF-2 binding to the cell surface, blocked the ability of added FGF-2 to stimulate the growth of fibroblasts, and resulted in the terminal differentiation of mouse myoblasts (which is normally suppressed by FGF-2) (Rapraeger et al., 1991).

Several of the experiments aimed at elucidating the molecular basis of the apparent requirement for HeS in the binding of FGFs to tyrosine kinase receptors are reminiscent of experiments done with ATIII and thrombin. Experimental approaches have included mutagenizing basic amino residues thought to be involved in heparin/HeS binding (Heath et al., 1991), and, as described below, examining the effects of using variously sized oligosaccharides in binding assays. Two models that are not mutually exclusive have emerged, a template model, similar to that described for ATIII above, and the "dimerization model." In the dimerization model, the function of HeS is to facilitate the dimerization of FGF molecules that then bind and help to dimerize their high affinity receptors. Dimerization is believed to be a common

step in signalling through tyrosine kinase receptors. The dimerized receptors are thought to reciprocally auto-phosphorylate each other, an initial step that results in a signalling cascade (Ullrich and Schlessinger, 1990).

Evidence in support of the template model has come from a study of the kinetics of the FGF binding to its receptor. The binding of FGF-1 to HeS and FGFR-1 on the cell surface occurred with apparently similar "on" rates, but the "off" rate was higher for FGF bound to HeS than for FGF bound to FGFR-1. However, in solution, FGF-1 bound to FGFR-1 exhibited on off rate similar to that of FGF-1 bound to HeS. This was interpreted to suggest that a cell surface molecule, possibly an HSPG, stabilizes the interaction of FGF-1 with cell surface FGFR-1 by participating in the formation of a ternary complex (Nugent and Edelman, 1992). Additional support for a template effect has come from the identification of an apparently essential heparin binding domain in the FGFR-1 tyrosine kinase receptor (Kan et al., 1993). Antibodies raised to a synthetic peptide from this region as well as point mutations of lysine residues led to a loss of heparin binding and heparin-mediated FGF-1 binding to FGFR-1. In addition, there have been reports of a minimum size of heparin-derived saccharides that will potentiate FGF interactions with FGFR-1 [cf (Ornitz et al., 1992; Guimond et al., 1993)], although the minimum size designated varies from report to report, possibly reflecting differences in the way the saccharides were prepared. Guimond et al. (1993) also demonstrated that heparin has similar effects on the binding of FGFs 1 and 4 to binding to cell surface tyrosine kinase receptors, and that different heparin-derived oligosaccharides have different effects on binding. This indicates there is specificity in the interaction of heparin with the different FGFs and their receptors. Specificity of GAG-protein interactions will be discussed further below.

Ornitz et al. (1992), in a set of experiments designed to demonstrate a requirement for heparin in FGF binding to its receptor in solution, also demonstrated that FGF molecules can be crosslinked into dimers in the presence of added heparin but not in its absence. In a subsequent report, a synthetic heparin analog that binds monovalently to FGF-1 was found to inhibit FGF receptor dimerization

and tyrosine kinase activation (Spivak-Kroizman et al., 1994). Even small di or tri-saccharides, used in relatively high concentrations, can apparently serve to dimerize FGF molecules (presumably by binding to multiple sites on each molecule) and activate the FGF signalling pathway (Ornitz et al., 1995).

None of the above observations rule out a model in which both the template effect and dimerization play a role in FGF binding to FGF receptors. Indeed there has already been one report in which a model was advanced that incorporated elements of both template and dimerization hypotheses (Pantoliano et al., 1994). Most of the studies done to characterize the effects of heparin on the binding of FGFs to their receptors have used heparin as a model for the HeS that is likely to be contributed by HSPGs in vivo. The question of which HSPGs actually serve as the coreceptors for FGFs has only begun to be addressed. A cell surface localization for such HSPGs may seem a priori to be the most likely situation; however it is conceivable that an HSPG in the extracellular matrix (ECM) could serve a similar function. It has long been proposed that ECM HSPGs could act as a reservoir for growth factors such as the FGFs, protecting them from proteolytic degradation [c.f. (Saksela et al., 1988)]. In fact, in a series of recent reports in which distinct PG core proteins were immunopurified from human fetal lung fibroblasts, the ECM HSPG perlecan, and not the cell surface HSPGs glypican (GPN), syndecan-1 (SYN-1), and fibroglycan (syndecan-2, SYN-2), was found to be able to potentiate the binding FGF to CHO cells deficient in HeS. Perlecan also potentiated FGF-2 binding to soluble or immobilized FGFR-1 (Aviezer et al., 1994a; Aviezer et al., 1994b). GPN, SYN-1, and SYN-2 were found to actually *decrease* binding of FGF-2 to immobilized FGFR-1 or to the HeS deficient CHO cells. Whether perlecan will generally be found to serve as an FGF coreceptor remains to be determined, but the above data suggest that it may be an FGF coreceptor for human fetal lung fibroblasts. One caveat in these studies is that the effects of the various HSPGs were measured by adding immunopurified HSPGs to the binding reactions. It is not clear that identical behavior would be observed if the HSPG were in their natural settings (i.e. in the matrix or intercalated in the plasma membrane).

Although FGF binding to FGF receptors is probably the most extensively characterized example of a system implicating HSPGs as coreceptors, support for the coreceptor hypothesis has also come from studies of cell adhesion on fibronectin (FN) and NCAM-NCAM adhesive interactions. CHO cells deficient in HeS synthesis were found to be unable to attach and spread on a heparin-binding fibronectin peptide, and, although they did adhere to intact FN, they did not form focal adhesions as wild type CHO cells did (LeBaron et al., 1988). These results suggest that both cell surface HSPGs and integrins, the classical FN receptors, are required for full adhesive response to FN. The extent to which HSPGs and integrins interact, if at all, during adhesion to FN remains to be determined. NCAM-NCAM interactions have been shown to be sensitive to added heparin, an antibody to the NCAM heparin binding region, or added NCAM proteolytic fragment encompassing the heparin binding region (Cole et al., 1986a; Cole and Glaser, 1986b). Further, a genetic deletion of the heparin binding region results in the loss of NCAM binding to heparin, and a loss of retinal cell adhesion to NCAM coated dishes (Reyes et al., 1990). Taken together, the data suggest that a cell surface HSPG participates in at least some NCAM-NCAM adhesive interactions.

Interactions of CSPGs with their ligands

Comparitively little is known about CS/CSPG interactions with CS-binding proteins. Perhaps because CS affinity for heparin/HeS-binding proteins is typically lower than the affinities of heparin and HeS, less work has been done to characterize CS-protein interactions at the molecular level. However, examples of proteins that bind to CS with affinities comparable to that of HeS exist. Thrombospondin-1 (TSP-1) binds to neonatal rat brain CS with a Kd of 235 nM, very similar to the Kd of 180 nM measured for HeS isolated from the same source (Mary Herndon, unpublished data). This is interesting in light of the observation, mentioned above, that TSP-1 failed to fractionate low molecular weight heparin into low and high affinity fractions (San Antonio et al., 1993). It is possible that TSP-1 contains GAG binding sites that are promiscuous enough to be satisfied by either HeS or CS.

It is also possible that TSP-1 has separate CS and HeS binding sites that each bind to their respective GAGs with similar affinities.

The binding of CS by TSP-1 appears to have functional consequences. In one study, attachment of myoblasts to TSP-1 was not inhibited by heparin or the GRGDS peptide, which interferes with many integrin-mediated interactions, but was inhibited by the addition of chondroitin sulfate (Adams and Lawler, 1994). This suggests that for some cell types, a cell surface CSPG may be important for TSP-1 binding. In another study Chinese hamster ovary cells that were pretreated with chondrotinase showed a decreased ability to bind and internalize TSP-1 (Murphy-Ullrich et al., 1988), supporting the view that cell surface CSPGs contribute to cell interactions with TSP-1.

Another molecule for which CS binding has been noted is tenascin (TN). TN is a large extracellular matrix glycoprotein with six subunits each in the size range of 190,000 - 225,000. TN binds to the CSPGs phosphacan and neurocan in solid phase binding assays. Phosphacan was further shown to inhibit the binding of C6 glioma cells to a TN substrate (Grumet et al., 1994).

IV. Expression and Function of Proteoglycans during Development with Special Reference to the Nervous System

Many of the putative biological roles for PGs have been suggested on the basis of PG expression and the expression of PG-binding molecules during development. This section will review the evidence for PG expression and function in the developing nervous system. The information is organized by developmental process, and the PG-binding molecules that have been implicated in that process. The individual PGs involved (where identified) are discussed.

Regulation of Cell Proliferation

Many molecules that are thought to be involved in the control of cell proliferation bind to heparin and HSPGs. These include heparin binding growth factors and extracellular matrix molecules. Evidence that PGs play a functional role in interacting with these molecules is reviewed below.

Acidic FGF (FGF-1), and basic FGF (FGF-2) are two small (M_r~16 kDa) polypeptides that were first purified from bovine brain on the basis of their ability to stimulate the proliferation of 3T3 fibroblasts (Lobb and Fett, 1984; Thomas et al., 1984). The sequences of these two proteins were subsequently determined directly by protein sequencing and shown to be 55% identical (Gimenez-Gallego et al., 1985). Currently there are seven known members of the FGF family expressed in a wide variety of tissues and developmental stages.

FGF-1 and FGF-2 are expressed at several locations in the developing nervous system, including the ventricular zone and cortical plate of the developing rat cerebral cortex (Gonzalez et al., 1990; Fu et al., 1991; Wilcox and Unnerstall, 1991). In addition, these molecules, have been shown to stimulate proliferation and survival of neural cells in vitro (Walicke et al., 1986; Unsicker et al., 1987; Walicke, 1988; Murphy et al., 1990; Hughes et al., 1993).

As discussed in section III above, HSPGs have been implicated as coreceptors for members of the FGF family, since binding of FGFs to their high affinity tyrosine kinase receptors appears to require HeS. Several studies have reported on findings that support the view that HSPGs can influence FGF activity. As reviewed in Gallagher (1989), these studies have shown that HeS chains and FGFs can form biologically active complexes that have potent mitogenic activity for a variety of cell types. In one of these studies purified HSPGs secreted from bovine capillary endothelial cells were found to bind to FGF-2 and protect it from degradation by the proteinase plasmin, leading to the suggestion that one role of HSPG binding to FGF could be to protect FGF from degradation at sites where proteolytic activity is high, such as zones of neovascularization (Saksela et al., 1988). The view of the extracellular matrix as a reservoir of FGF-2 is supported by a study that showed that FGF-2 can be released from the matrix by treatment with heparitinase or by the addition of soluble heparin (Bashkin et al., 1989). Evidence of the presence of endogenous FGF-2 in an extracellular matrix derived from an in vivo source comes from a study that showed that incubation of Descemet's membranes of bovine corneas with platelets, neutrophils or lymphoma cells resulted in the release of FGF-2 (Ishai-Michaeli et al., 1990). These cells were found

to express a heparanase activity that could be inhibited by carrageenan lambda, a GAG analogue, and carrageenan lambda also inhibited the release of FGF-2. Taken together, the results suggest that some cell types can mobilize matrix-bound FGF-2 by degrading the HeS chains of an extracellular matrix HSPG to which FGF-2 is bound.

To date, the identities of the core proteins that actually interact with members of the FGF family in vivo have not been defined. It has been proposed that syndecan-1 (SYN-1) could act as a coreceptor for FGF-2 (Bernfield et al., 1992). SYN-1 has been shown to bind to FGF-2 via its HeS chains (Bernfield and Hooper, 1991), and expression of SYN-1 in lymphoblastoid cells confers low affinity FGF-2 binding sites on these cells (Kiefer et al., 1990). Furthermore, SYN-1 is transiently expressed on rapidly proliferating, condensing mesenchymal cells in a variety of locations during development, suggesting that one function of SYN-1 could be to bind a growth factor that stimulates mesenchymal cells to divide [reviewed in (Bernfield et al., 1992)]. This possibility is supported by a report that demonstrated a highly similar expression pattern of SYN-1 and epidermal growth factor in developing tooth mesenchyme (Partanen and Thesleff, 1987).

Some authors have reported that a biologically active FGF-2/HSPG complex is released from cell surfaces by the bacterial enzyme phosphoinositol-specific phospholipase C (PIPLC), which cleaves glycosylphosphatidylinositol (GPI) anchors. These studies point to the possibility that members of the glypican (GPN) family of GPI-anchored HSPGs could also serve as coreceptors for members of the FGF family.

However, as described in section III above, an experiment using PG core proteins immunopurified from human fetal lung fibroblasts demonstrated that the large secreted HSPG perlecan, and not GPN or SYN-1, was able to potentiate the binding of FGF-2 to HeS-deficient CHO cells, or to FGFR-1, a tyrosine kinase receptor for FGF (Aviezer et al., 1994a; Aviezer et al., 1994b). In fact, although GPN and SYN-1 bound FGF-2, they were actually found to inhibit the subsequent binding FGF-2 to FGFR-1.

Recently, additional evidence that perlecan family HSPGs may serve as FGF coreceptors has come to light (Nurcombe et al., 1993). In this study, the timing of expression of FGF-1 and FGF-2 mRNAs in the

developing mouse mesencephalic and telencephalic neuroepithelium was examined. FGF-2 mRNA was found on embryonic day 9 (E9) and thereafter, but FGF-1 mRNA was only detected beginning on E11. Interestingly an HSPG [subsequently shown to be a novel member of the perlecan family of HSPGs (Joseph et al., 1995)] isolated from the same tissue on E9 was found to bind with preferentially to FGF-2, whereas the same HSPG isolated from E11 tissue bound preferentially to FGF-1. Thus the switch in the binding preferences of the HSPG coincided with the appearance of FGF-1 in the neuroeptihelium suggesting that this HSPG may be regulated to preferentially interact with FGF-1 when it appears. As discussed above, FGF-1 and FGF-2 have similar but distint binding sites on HeS chains, suggesting one way cells could discriminate between these two growth factors.

The results described above suggest that different HSPGs may exert different effects on FGF interactions with FGF receptors. In particular, the results raise the possibility that cells could regulate their response to FGF-2 by regulating which types of HSPGs they produce. For example, SYN-1 and GPN could act as competitive inhibitors of perlecan potentiation of FGF-2 binding to FGFR-1.

Nevertheless, it appears likely that the individual HSPGs involved in potentiating growth factor action are likely to vary from tissue to tissue. Ratner et al, (1988) have reported on a molecule mitogenic for Schwann cells that was found to be peripherally associated with the surface of dorsal root ganglion neurons (which are themselves mitogenic for Schwann cells). This factor could be extracted from the cell surface with heparin, suggesting it could be held at the cell surface by an HSPG. If this is the case, it suggests that a cell surface HSPG on the neuronal cell surface presents the growth factor to a receptor on the Schwann cell surface.

The β-type transforming growth factor (TGF-β) family of polypeptide growth factors consists of at least five related species, three of which are found in mammals. The structures of TGFβ1, β2, and β3 are similar; each is synthesized as a pre-pro monomeric polypeptide. The pre-pro form is cleaved, but the pro-peptide remains associated, maintaining the latent state until cleaved

monomers dimerize, causing the pro-peptides to dissociate [reviewed in (Lyons and Moses, 1990)]. In in vitro assays, TGFβ family members are mitogenic for some cell types such as cells derived from bone, and growth inhibitory for many other cell types. Members of the TGFβ family are expressed in the developing nervous system. TFGβ2 and β3 are found associated with glia and astrocytes in the embryonic midbrain and ventral spinal cord (Pelton et al., 1991).

The receptor system for the TGFβs is complex. TGFβ receptors type I and II are transmembrane molecules that form a signalling complex when TGFβ is bound (Wrana et al., 1992). The type II receptor contains a serine/threonine kinase in its cytoplasmic domain. The type III receptor for the TGFβs is betaglycan, a heteroglycan with a 100 kDa core protein that can be substituted with CS or HeS chains, and is widely expressed in the developing embryo (López-Casillas et al., 1991). Interestingly, TGFβs bind to betaglycan via the core protein. The functional role of betaglycan's GAG chains remains unknown (Cheifetz and Massague, 1989).

In one recent study, betaglycan expressed in a rat skeletal muscle myoblast cell line (a cell type that does not normally express betaglycan) was found to enhance the growth inhibition response to exogenous TGFβ1, apparently by participating in the formation a ternary complex involving the signalling TGFβ type II receptor (López-Casillas et al., 1993). Thus betaglycan appears to be able to act as a coreceptor, at least for TGFβ1.

The CS/DSPG decorin also binds to TGFβ, and, like betaglycan, the binding is mediated by the core protein. Decorin expression has been detected in the floor plate of the developing mouse spinal cord, a structure believed to be important for the guidance of commisural axons (see below). Decorin expressed in Chinese hamster ovary cells was found to bind endogenous and exogenous TGFβ1 and suppress the growth promoting activity of TGFβ1 for those cells. The mechanism of TGFβ inactivation by decorin remains to be determined (Yamaguchi et al., 1990). Thus betaglycan and decorin both bind to TGFβ via their core proteins but appear to exert opposite effects on TGFβ activity: betaglycan enhances TGFβ activity and decorin neutralizes it.

The large extracellular matrix (ECM) glycoproteins laminin-1 (LN-1), fibronectin (FN), and thrombospondin-1 (TSP-1) have also been shown to exert effects on cell proliferation and differentiation. The structures of these molecules and their tissue distribution will be discussed in greater detail in the next section with regard to their effects on cell migration and axon outgrowth.

LN-1 causes cell proliferation and an increase in ^{3}H-thymidine uptake for several cell types in vitro [reviewed in (Lander and Calof, 1993)]. Interestingly, different LN-1 fragments are active for different cell types. This is possibly a reflection of the large number of cellular receptors for LN-1, including as many as six different combinations of α and β integrin subunits, the cell surface enzyme galactosyl transferase, and, potentially, HSPGs (Lander and Calof, 1993).

In contrast to LN-1, FN has often been found to exert growth inhibitory effects on cells. For example, mitogen-stimulated Schwann cells, in what appears to be an autrocine mechanism of growth control, secrete a metalloprotease that cleaves FN, liberating a 29 kDa, N-terminal, heparin binding fragment that inhibits proliferation (Muir and Manthorpe, 1992). Transformed Schwann cells do not secrete the protease, suggesting that this mechanism may be a primary mode of growth control for Schwann cells.

In the developing rodent cerebral cortex, FN expression, initially found throughout the ventricular zone (a layer where neuronal precursor cells are rapidly dividing), becomes restricted to the preplate overlying the ventricular zone as soon as the preplate appears. (Stewart and Pearlman, 1987). The preplate is formed by the earliest cells to exit the cell cycle and begin differentiation (on E11-E12 in the mouse). Given the activity of FN noted above for Schwann cells it is interesting to speculate that one function of FN in the early cortex could be to act as a brake on proliferation, facilitating the transition from precursor to immature neuron.

The ECM glycoprotein TSP-1 is observed associated with the cell surfaces of neuroepithelial cells throughout developing mouse nervous system at early times (E11 and E13) when cell proliferation is proceeding at a high rate (O'Shea and Dixit, 1988). Outside of the nervous system, TSP expression in the skeletal muscle also correlates

with cell proliferation. In the early embryo, TSP is found on the surface of undifferentiated myoblasts (O'Shea and Dixit, 1988). At adult stages, TSP is upregulated in response to crush injury (Watkins et al., 1990) or in response to amyotrophic lateral sclerosis (Rao et al., 1992), conditions that correspond to renewed myoblast proliferation and migration.

In addition to the binding of intact PGs to ligands that exert an effect of cell proliferation, many authors have reported that free HeS can exert an effect on cell growth by binding to ligands in the nucleus [reviewed in (Hart et al., 1989)] Many of these reports have been met with natural skepticism since HeS chains are synthesized in the Golgi and attached to core proteins destined to be secreted or intercalated in the plasma membrane. Despite the attempts made to demonstrate HeS and CS observed in the nucleus was not a biochemical contaminant or the result of non-specific binding by an immunological probe, a conclusive demonstration of a nuclear localized GAG remained elusive.

A series of more recent reports has provided additional evidence that HeS can be localized to nucleii [reviewed in (Gallagher, 1989)]. In these reports, a GPI-linked HSPG on confluent hepatoma cells was found to express a fraction of HeS chains enriched in 2-O-sulfated-Glc-A, a rare modification. This same modification was found to be enriched on chains isolated from the nuclei of these cells. HeS/HSPGs derived from confluent cells, and containing a mixture of HeS structures, were added to logarithmically growing cells resulting in a selective enrichment of the 2-O-sulfated-GlcA forms in the nucleii, and causing growth arrest at the G1/S boundary of the cell cycle. The fact that the forms isolated in the nuclear pool contained a high percentage of HeS chains with a unique structure uncommon in the other cellular pools argues against non-specific contamination in these experiments (Fedarko and Conrad, 1986; Ishihara et al., 1986; Fedarko et al., 1989).

These reports raise the possibility that a cell surface GPI-anchored HSPG (possibly a member of the glypican family of HSPGs) may be involved in a growth control mechanism at least in these cells. This is

intriguing in light of the recent discovery of a Drosophila homolog of the mammalian glypican family. This molecule, the dally protein, was identified on the basis of a mutant phenotype in which the G2/M transition is blocked at a number of locations in the developing nervous system, such as the lamina furrow of the outer proliferative center, and the morphogenetic furrow of the eye disc (Nakato et al., 1995). The dally protein will be discussed in greater detail in Chapter 6 of this thesis with regard to functions of glypican family members in the developing nervous system.

Cell Migration and Axon Outgrowth

Many of the molecules that that are thought to play roles in cell migration and axon outgrowth are potential ligands for PGs. The section below summarizes the structures of some of these molecules, their expression patterns in the developing nervous system, their biological activities, and their potential interactions with PGs.

The extracellular matrix molecule laminin (LN) is composed of three separate polypeptide chains, the α chain, the β chain, and the γ chain (formerly, the A, B1, and B2 chains), each encoded by a separate gene. The ~440 kDa α chain associates with the β and γ chains, ~ 220 and ~210 kDa respectively, forming a distinctive cross shaped structure, with three short arms and one long arm, that can be visualized by electron microscopy [reviewed in (Beck et al., 1990)]. So far, 4 α chain isoforms, 3 β chain isoforms, and 2 γ chain isoforms have been identified. Of the 24 possible forms of intact LN that could be assembled from these chains, seven have been identified in vivo and are designated LN-1 through LN-7. Classic LN ($\alpha1\beta1\gamma1$), of the kind derived from EHS sarcoma, is LN-1. Other well characterized forms are LN-2 ($\alpha2\beta1\gamma1$; formerly merosin), and LN-3 ($\alpha1\beta2\gamma1$; formerly S-laminin).

LN-1 was originally purified from the conditioned medium of a variety of cell types on the basis of its potent neurite outgrowth promoting activity [cf (Lander et al., 1985)]. Subsequently it has also been demonstrated that LN-1 is actually able to guide neurites in vitro [cf (Gundersen, 1987)]; which is to say, neurites will accurately follow a patterned LN-1 substratum. Axon guidance in vivo by one of the LNs

has yet to be formally demonstrated, but LN expression patterns in the developing nervous system are provocative.

LN-1 is expressed in the avian optic tract during the time when retinal ganglion cells (RGCs) are extending axons through the tract to targets in the optic tectum. Furthermore, responsiveness of purified RGCs to LN-1 in vitro corresponds to the window of axon outgrowth in vivo. RGCs isolated from retina on embryonic day 11, after axon outgrowth is complete, no longer extend neurites on LN-1 in vitro, whereas RGCs from earlier time points do (Cohen et al., 1986; Cohen et al., 1987). Interestingly, the loss of RGC responsiveness to LN-1 correlates with the downregulation of $\alpha 6$ integrin subunit by these cells (deCurtis et al., 1991). LN isoforms are also found associated with the pathways through which spinal cord motoneurons and dorsal root ganglion neurons grow, as well as in the subplate of the developing rat cortex, a region through which cortical and thalamic axons grow (Rogers et al., 1986; Hunter et al., 1992).

LN-1 and LN-2 support the migration of olfactory epithelial (OE) neurons in vitro, whereas fibronectin (FN) and type I collagen type do not. A subset of these neurons is known to migrate from the OE to the hypothalamus during development. The specificity in the preference of OE neurons for LN over other ECM molecules tested suggests that a LN isoform may be part of the migratory substrate for these cells in vivo [reviewed in (Hynes and Lander, 1992).

The PG ligands for LN isoforms in vivo are not well characterized. The HSPG perlecan copurifies with LN-1 [cf (Lander et al., 1985), and perlecan HeS chains bind to LN-1 in in vitro assays. This has led to the suggestion that perlecan may be involved in the proper assembly of LN isoforms in the ECM (Yurchenco and O'Rear, 1993; Colognato-Pyke et al., 1995) In addition, the cell surface HSPGs glypican (GPN) and syndecan-3 (SYN-3) are known to be expressed in the developing rat nervous system (Carey et al., 1992; Gould et al., 1992; Litwack et al., 1994; Gould et al., 1995). The subject of the main part of this thesis is the identification, cloning, and characterization of a novel member of the GPN family of HSPGs. This molecule was named cerebroglycan (CBG), and is also expressed in the developing rat nervous system. Evidence will be presented in Chapters 4 and 5 of this thesis

supporting the possibility that CBG, GPN, and SYN-3 could function as receptors or coreceptors for LN isoforms.

Fibronectin (FN) is another large glycoprotein of the ECM found in the developing nervous system. FN consists of two chains, each ~220 kDa, that are disulfide linked near their C-termini. Although both chains are encoded by the same gene, they are known to be alternately spliced, creating the opportunity for several FN isoforms to be assembled [reviewed in (Hynes, 1990; Reichardt and Tomaselli, 1991)]. FN is a composite of several structural motifs, including FN repeats types I, II, and III. FN type III repeats are also found in other molecules such as tenascin. So far, six different $\alpha\beta$ combinations of integrin subunits have been identified as FN receptors (Hynes, 1992). FN also contains heparin binding domains, one near the N-terminus and one near the C-terminus of each subunit.

In the peripheral nervous system FN is found on associated with the nerve fibres of sensory and sympathetic ganglia, as well as in the matrix through which neural crest cells migrate (Rogers et al., 1989; Hynes and Lander, 1992). In the central nervous system, FN is transiently expressed in the developing cortex. FN is expressed in the early ventricular zone, but becomes restricted to the preplate, a structure comprised of some of the earliest cortical neurons to be born. When the forming cortical plate splits the preplate into the marginal zone (MZ), laterally, and the subplate (SB), medially, FN expression remains associated with both the MZ and the SP. FN expression is also associated with radial glial fibres in the developing cortex; these fibres constitute possible pathways of migration for cortical neurons (Stewart and Pearlman, 1987; Sheppard et al., 1991). At later time points, when axon outgrowth into and out of the cortex has been essentially completed, FN expression disappears from that structure.

Evidence that FN is involved in the migration and axon outgrowth of neuronal cells comes from studies involving a variety of cell types [reviewed in (Rogers et al., 1989; Hynes and Lander, 1992)]. FN supports the migration of neural crest cells in vitro, and, as discussed above, FN is richly expressed in vivo in the extracellular matrix

through which neural crest cells migrate. The RGD peptide, which blocks many of the integrin-FN binding interactions (as well as some integrin interactions with other ECM proteins), inhibits neural crest cell migration when injected into embryos.

PNS neurons respond to a FN substrate in vitro by extending stable neurites. Interestingly, although some CNS neurons extend neurites in vitro on FN, these neurites are not stable and retract within a few days (Rogers et al., 1989). Studies comparing CNS and PNS interactions with FN fragments suggest that CNS neurons may use a cell surface HSPG to interact with FN (Rogers et al., 1983; Rogers et al., 1989). PNS neurons attach and extend neurites on both a 75 kDa fragment of FN containing integrin binding sites alone and a 33 kDa fragment containing the C-terminal heparin binding domain and another integrin binding site. In contrast, CNS neurons (spinal cord motoneurons) attach and extend neurites only on the 33 kDa fragment. Interestingly, the neurites extended by the spinal cord motoneurons in these experiments were stable, persisting for several days in culture, unlike neurites extended on intact FN by a variety of CNS neuronal cell types. One interpretation of these studies is that FN is capable of sending multiple, and sometimes conflicting signals to CNS neurons. The multiple integrin binding sites on the intact FN molecule provide a possible mechanism for such complex interactions. The fact that spinal cord motoneurons responded to the 33 kDa fragment, heparin binding fragment of FN raises the possibility that an HSPG on the cell surface of these cells could be involved in the interaction, although, as noted above, this fragment also contains an integrin binding site adjacent to the heparin binding domain.

A recent study has demonstrated that neuronal cells can utilize a cell surface HSPG to attach and spread on synthetic FN fragments, derived from the 33 kDa fragment, that contain the heparin binding domain without the integrin binding site (Haugen et al., 1992). The adhesion and spreading can be inhibited by heparitinase, or by antibodies raised to a GPI-anchored HSPG in melanoma cells. These same antibodies recognize a core protein of M_r 51 kDa, that shifts upward dramatically upon reduction, on western blots of neuronal PG preparations. The glypican family of cell surface HSPGs, unlike the

syndecan family HSPGs, contains several cysteine residues, and changes in the M_r of the glypican core have been noted before upon reduction (David et al., 1990; Herndon and Lander, 1990). Taken together, the biochemical data suggest that the cell surface PG in these studies is a member of the glypican family, and that it may play a role in the interactions of neuronal cells with FN.

Thrombospondin (TSP) is a large extracellular matrix glycoprotein consisting of 3 chains of chains of M_r~140 kDa, that are linked by disulfide bonds. Each subunit of TSP contains two heparin binding domains and multiple integrin binding sites. TSP also binds to sulphatides (sulfated glycolipids) and to proteases such as plasminogen. There are currently five known isoforms of TSP, TSP-1 through TSP-4, and COMP (cartilage oligomeric matrix protein). COMP differs from the other members of the TSP family in that it is isolated as a pentameric molecule. So far, TSP hybrids that incorporate subunits from different TSP genes into the same molecule have not been identified (Adams and Lawler, 1993).

Vitronectin (VN), in contrast to most other ECM glycoproteins, is a relatively small molecule consisting of a single 78 kDa chain. Its presence in the nervous system is just beginning to be characterized.

TSP-1 and VN are both present in the developing retina (Neugebauer et al., 1991), and both TSP-1 and VN promote attachment and neurite outgrowth by retinal neurons in vitro. Attachment to TSP-1 is inhibited by the addition of soluble heparin (Neugebauer et al., 1991), indicating a possible role for cell surface HSPGs in the interaction of retinal neurons with this molecule.

TSP-1 is also expressed widely throughout other areas of the developing nervous system, including the cerebellum, where it is found associated with granule cell neurons in the premigratory zone of external granule layer and in the molecular layer (O'Shea et al., 1990). Interestingly, an antibody to TSP-1 was found to block granule cell migration from the external granule layer to the molecular layer, suggesting TSP-1 may directly participate in this process (O'Shea et al., 1990). As described in Chapter 3 of this thesis, cerebroglycan (CBG) mRNA in the developing rat cerebellum is also specifically

associated with neurons of the premigratory zone raising the possibility that CBG and TSP-1 interact on the cell surfaces of these neurons. One report has also described extensive colocalization of the cell surface HSPG syndecan-1 with TSP in developing embryos, including in the very early neuroepithelium (embryonic day 8.5 of the mouse) (Corless et al., 1992).

Tenascin (TN) is a multidomain extracellular matrix molecule comprised of six separate chains of ~190-220 kDa each. The chains are linked to each other through a disulfide-rich region near their amino termini. As with FN, the subunits of TN are encoded by a single gene, but can be alternately spliced. TN manifests several different biological activities for neuronal cells in vitro including cell attachment and neurite outgrowth promoting activity for some neuronal cell types, as well as anti-adhesive and neurite repulsive activities for others. The locations of some of these activities have been mapped using TN fragments (Chiquet, 1989; Erickson and Bourdon, 1989; Lander and Calof, 1993).

TN is expressed in the pathways traversed by migrating neural crest cells. In particular, TN is enriched in the rostral portion of somitic sclerotome. This is the same portion that neural crest cells preferentially enter, suggesting that TN may provide a favorable substrate for neural crest cell migration. Attempts to support this notion with in vitro experiments have met with mixed results. TN isolated from chick brain (also called cytotactin) does not support chick neural crest cell migration, and slows down neural crest cell migration on FN when mixed with FN in the substrate. In contrast, TN isolated from chick fibroblasts does support the migration of neural crest cells [reviewed in (Chiquet, 1989)]. The fact that TN can be alternatively spliced provides one possible explanation for this discrepancy.

TN has also been found in the central nervous system, where it appears later and is more widespread than either FN or LN-1. In the developing mouse cortex, TN appears in the subplate and marginal zones as FN is disappearing (see above) and subsequently becomes expressed throughout the maturing cortical plate (Sheppard et al.,

1991). In the somatosensory cortex, TN is expressed at the boundaries of the whisker barrel fields, but not in the center. The barrel fields represent separate functional units within the cortex, and the significance of TN expression at their boundaries is not yet understood. By adulthood TN expression in the cortex has dramatically declined.

TN is also found in the developing mouse cerebellum (Chuong et al., 1987). In this structure, TN expression is associated with the cerebellar granule cell neurons and with the surfaces of the Bergmann glia, along which the granule cells migrate. Antibodies to TN block cerebellar granule cell migration in cerebellar slice cultures. The granule cells migrate into the molecular layer where the Bergmann glia are found, but progress no further.

Although TN can bind to heparin, its interactions with heparin and HSPGs have not been well characterized. However, several experiments have demonstrated interactions of CSPGs with TN. TN isolated from chick brain was found to copurify with a large CSPG termed "cytotactin binding PG" (Hoffman and Edlman, 1987; Hoffman et al., 1988). When this PG was purified away from TN, it was shown to bind purified TN in solid phase assays. More recently, the CSPGs phosphacan, an extracellular variant of RPTP-β, a protein tyrosine phosphatase, and neurocan, a large CSPG related to the aggrecan family, were shown to bind to TN in solid phase assays (Grumet et al., 1994). The TN-binding of these CSPGs was found to be largely mediated by the core proteins. Interestingly, neurocan is also expressed at the boundaries of barrel fields in the somatosensory cortex (Oohira et al., 1994). This, together with the apparent binding of the neurocan core protein to TN in in vitro assays suggest that neurocan may be a ligand for TN in vivo.

Recently, a new class of molecules important for axon guidance was discovered. These molecules, called the netrins, are expressed in the developing spinal cord, and were isolated on the basis of their ability to promote the axon outgrowth of commisural spinal cord axons in vitro (Kennedy et al., 1994; Serafini et al., 1994). Subsequently, netrin-1 was found to be expressed in the floor plate of the spinal

cord, a important structure for commisural axon navigation, and netrin-2 was found to be expressed in the ventral spinal cord above the floor plate. The netrins have homology to the C. elegans protein UNC-6, a laminin-related molecule involved in the circumferential migration of cells and axons in the worm.

The netrins are heparin-binding molecules, and are isolated as peripheral membrane proteins from homogenized brain. This raises the possibility that a cell surface or matrix associated HSPG could serve as a ligand for the netrins. The GPI-anchored cell surface HSPGs glypican (Litwack, 1995), and cerebroglycan [(Litwack, 1995) and Chapter 3, this thesis] have expression associated with the commisural axons of the developing spinal cord, raising the possibility that they could serve as netrin receptors or coreceptors.

Additionally, a recent study has demonstrated that the CS/DSPG decorin is also expressed in the floor plate (Litwack, 1995). Decorin was shown to bind to netrin-1 with an affinity of ~ 90 nM, whereas decorin GAG chains have an affinity for decorin > 1 μM, suggesting that the decorin core protein binds to netrin-1 (Litwack, 1995). Decorin has previously been shown to bind to collagen fibrils and TGFβ via its core protein (Vogel et al., 1984; Cheifetz and Massague, 1989).

In addition to the extracellular matrix molecules discussed above, another class of heparin binding molecules that has been implicated in cell migration is the class of cell adhesion molecules. This family of molecules, typified by neural cell adhesion molecule (NCAM) and L1 features extracellular domains that contain immunoglobin constant region repeats. These molecules mediate cell-cell adhesion via homophilic binding of cell adhesion molecules on the surfaces of the adhering cells. NCAM is widely expressed in the developing brain, and potential functions in cell migration and axon outgrowth were suggested by antibody perterbation experiments in which NCAM antibodies were injected into embryos. In these experiments, the anti-NCAM antibodies caused effects such as perturbation of optic nerve fiber projection to the optic tectum, and changes in the growth patterns of nerves in the developing limb, or nerves undergoing regeneration [reviewed in (Goridis and Brunet, 1992)]. However, the

phenotype of mice in which NCAM is knocked out is rather subtle: the gross morphology of the brain remains intact, there is a modest but notable decrease in the size of the olfactory bulb, and the animals performed below average in the Morris water maze, a test thought to be a measure of spatial learning ability (Cremer et al., 1994)

The related molecule L1 is expressed in migratory and premigratory granule cells in the developing cerebellum and is also found associated with the parallel fibres, axons extended by granule cells as they migrate (Chuong et al., 1987). On the cellular level, electron microscopy shows that L1 is present at the sites of neuron-neuron contact in the developing cerebellum, but not at sites of neuron-glia contact (Persohn and Schachner, 1987), suggesting that L1 may function via neuron-neuron interactions rather than by binding to L1 on the surface Bergmann glia, along which granule cells migrate. Antibodies directed against L1 cause granule cells to cease migrating and pile up at the border of the molecular layer (Chuong et al., 1987), suggestive of a role for L1 in granule cell migration. No PG ligands for L1 have yet been identified, but the large transmembrane CSPG NG2 has been found to inhibit neurite outgrowth of cerebellar granule cells in vitro, and to possess a neurite-repulsive activity for these neurites (Dou and Levine, 1994). Interestingly, NG2 expression is observed in the deep portion of the molecular layer in a region just underlying the parallel fibres of the granule cell neurons. This observation, together with the neurite repulsive activity of NG2 for granule cell neurons has led to the suggestion that NG2 may form a barrier for these axons, constraining them to the axon tracts in the more superficial part of the molecular layer (Dou and Levine, 1994).

Evidence that cell surface HSPGs may be involved in NCAM-NCAM adhesive interactions has been obtained from a series of experiments that demonstrated that heparin, antibodies against cell surface HSPGs, antibodies against the NCAM heparin binding domain, peptides derived from the heparin binding domain, and a genetic deletion of the heparin binding domain all block NCAM-NCAM binding (Lander, 1990). The HSPG(s) involved have not yet been characterized

V. Conclusion

Perhaps the strongest theme emerging from the data reviewed above is that we have a lot to learn about the expression and function of proteolgycans in the nervous system. There are many potential ligands for PGs in the nervous system, but the PGs that bind to them have, in many cases, not been identified, or only tentatively identified. Functional consequences of PG's interactions with their ligands have been inferred from preliminary experiments in vitro, but the task of testing them in vivo is still in the early stages. Some exciting stories have begun to emerge (the apparent requirement of heparan sulfate for FGF function comes to mind), but many more wait to be discovered.

A recent report has demonstrated that as many as 25 distinct core proteins are expressed in the developing and adult rat brain (Herndon and Lander, 1990). At the time of this writing, only three of the core proteins in this study have been positively identified, two as a result of the experiments reported in Chapter 2 below. It is likely that PGs characterized elsewhere will account for several of the remaining 25 cores, but the molecular identity of some probably awaits cloning and characterization of new PG cores. To begin to make sense of which PGs bind to which ligands exerting which effects, a more complete understanding of the PGs involved and the details of their interactions with potential ligands, will be required.

In Chapter 2, we report the cloning and preliminary characterization of a novel HSPG core protein that we have named cerebroglycan. In Chapter 3, cerebroglycan expression in the developing embryo was examined by in situ hybridization and found to be restricted to the developing nervous system, making cerebroglycan the first tissue-specific PG to be reported, and supporting an earlier hypothesis (Herndon and Lander, 1990) that individual PG core proteins are adapted to fulfill specific functional roles during development. Examination of cerebroglycan mRNA expression in several structures strongly supports the view that cerebroglycan is specifically expressed by immature neurons in the process of axon outgrowth and/or cell body migration. Immunohistochemical localization of cerebroglycan core protein further revealed that

cerebroglycan is most strongly associated with axon tracts in the developing nervous system, further supporting the idea that cerebroglycan expression correlates with the motile phases of neuronal development. In Chapter 4, evidence is presented that is consistent with the possibility that cerebroglycan could serve as a receptor or coreceptor for the neurite outgrowth promoting molecule LN-1. Chapter 5 describes experiments in which the molecular weight of intact cerebroglycan is determined and describes additional experiments, that, taken together, suggest a role for the cerebroglycan core protein in the binding of intact CBG to LN-1.

ACKNOWLEDGEMENTS

References

Adams, J. and J. Lawler. 1993. The thrombospondin family. Curr. Biol. 3:188-190.

Adams, J.C. and J. Lawler. 1994. Cell-type specific adhesive interactions of skeletal myoblasts with thrombospondin-1. Molec. Biol. Cell 5:423-437.

Aviezer, D., D. Hecht, M. Safran, M. Eisinger, G. David, and A. Yayon. 1994b. Perlecan, basal lamina proteoglycan, promotes basic fibroblast growth factor-receptor binding, mitogenesis, and angiogenesis. Cell 79:1005-1013.

Aviezer, D., E. Levy, M. Safran, C. Svahn, E. Buddecke, A. Schmidt, G. David, I. Vlodavsky, and A. Yayon. 1994a. Differential structural requirements of heparin and heparan sulfate proteoglycans that promote binding of basic fibroblast growth factor to its receptor. J. Biol. Chem. 269:114-121.

Baciu, P.C., C. Acaster, and P.F. Goetinck. 1994. Molecular cloning and genomic organization of chicken syndecan-4. J. Biol. Chem. 269:669-703.

Bahr, B.A., K. Noremberg, G.A. Rogers, B.W. Hicks, and S.M. Parsons. 1992. Linkage of the acetylcholine transporter-vesamicol receptor to proteoglycan in synaptic vesicles. Biochem. 31:5778-5784.

Bajjalieh, S.M., K. Peterson, R. Shinghal, and R.H. Scheller. 1992. SV2, a brain synaptic vesicle protein homologous to bacterial transporters. Science 257:1271-1273.

Bame, K.J., K. Lidholt, U. Lindahl, and J.D. Esko. 1991. Biosynthesis of heparan sulfate. Coordination of polymer-modification reactions in a chinese hamster ovary cell mutant defective in N-sulfotransferase. J. Biol. Chem. 266:10287-10293.

Bashkin, P., S. Doctrow, M. Klagsbrun, C.M. Svahn, J. Folkman, and I. Vlodavsky. 1989. Basic fibroblast growth factor binds to sub-endotheilial extracellular matrix and is released by heparitinase and heparin-like molecules. Biochemistry 28:1737-1743.

Beck, K., I. Hunter, and J. Engel. 1990. Structure and function of laminin: Anatomy of a multidomain glycoprotein. FASEB J. 4:148-160.

Bernfield, M. and K.C. Hooper. 1991. Possible regulation of FGF activity by syndecan, an integral membrane heparan sulfate proteoglycan. Ann. NY Acad. Sci. 638:182-194.

Bernfield, M., R. Kokenyesi, M. Kato, M.T. Hinkes, J. Spring, R.L. Gallo, and E.J. Lose. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu. Rev. Cell Biol. 8:365-393.

Bjork, I. and U. Lindahl. 1982. Mechanism of the anticoagulant action of heparin. Molec. Cell. Biochem. 48:161-182.

Brauker, J.H., M.S. Trautman, and M. Bernfield. 1991. Syndecan, a cell surface proteoglycan, exhibits a molecular polymorphism during lung development. Dev. Biol. 147:285-292.

Cardin, A.D. and H.J.R. Weintraub. 1989. Molecular modeling of protein-glycosaminoglycan interactions. Arteriosclerosis 9:21-32.

Carey, D., D. Evans, R. Stahl, V. Asundi, K. Conner, P. Garbes, and G. Cizmeci-Smith. 1992. Molecular cloning and characterization of N-syndecan, a novel transmembrane heparan sulfate proteoglycan. J. Cell Biol. 117:191-201.

Carey, D.J., R.C. Stahl, V.K. Asundi, and B. Tucker. 1993. Processing and subcellular distribution of the Schwann cell lipid-anchored heparan sulfate proteoglycan and identification as glypican. Exp. Cell Res. 208:10-18.

Casu, B., M. Petitou, M. Provasoli, and P. Sinay. 1988. Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans. TIBS 13:221-225.

Chang, J.-Y. 1989. Binding of heparin to human antithrombin III acitvates selective chemical modification at lysine 236. LYS-107, LYS-125, and LYS-136 are situated within the heparin-binding site of antithrombin III. J. Biol. Chem. 264:3111-3115.

Cheifetz, S. and J. Massague. 1989. The TGF-β receptor proteoglycan: cell surface expression and ligand binding in the absence of glycosaminoglycan chains. J. Biol. Chem. 264:12025-12028.

Chiquet, M. 1989. Tenascin/J1/Cytotactin: The potential function of hexabrachion proteins in neural development. Dev. Neurosci. 11:266-275.

Chuong, C.M., K.L. Crossin, and G.M. Edelman. 1987. Sequential expression and differential function of multiple adhesion molecules during the formation of cerebellar cortical layers. J. Cell Biol. 104:331-342.

Cohen, J., J.F. Burne, C. McKinlay, and J. Winter. 1987. The role of laminin and the laminin/fibronectin receptor complex in the outgrowth of retinal ganglion cell axons. Dev. Biol. 122:407-418.

Cohen, J., J.F. Burne, J. Winter, and P. Bartlett. 1986. Retinal ganglion cells lose response to laminin with maturation. Nature 322:465-467.

Cole, G.J. and L. Glaser. 1986b. A heparin-binding domain from N-CAM is involved in neural cell-substratum adhesion. J. Cell Biol. 102:403-412.

Cole, G.J., A. Loewy, and L. Glaser. 1986a. Neuronal cell-cell adhesion depends on interactions of N-CAM with heparin-like molecules. Nature 320:445-447.

Colognato-Pyke, H., J.J. O'Rear, Y. Yamada, S. Carbonetto, Y.-S. Cheng, and P.D. Yurchenco. 1995. Mapping of network-forming, heparin-binding, and $\alpha1\beta1$ integrin recognition sites within the α-chain short arm of laminin-1. J. Biol. Chem. 270:9398-9406.

Corless, C.L., A. Mendoza, T. Collins, and J. Lawler. 1992. Colocalization of thrombospondin and syndecan during murine development. Dev. Dynam. 193:346-358.

Cremer, H., R. Lange, A. Christoph, M. Plomann, G. Vopper, J. Roes, R. Brown, S. Baldwin, P. Kraemer, S. Scheff, D. Barthels, K. Rajewsky, and W. Wille. 1994. Inactivation of the N-CAM gene in mice results in size reduction of the olfactory bulb and deficits in spatial learning. Nature 367:455-459.

David, G., X.M. Bai, B.V.D. Schueren, J.J. Cassiman, and H.V.D. Berghe. 1992. Developmental changes in heparan sulfate expression: In situ detection with mAbs. J. Cell Biol. 119:961-975.

David, G., X.M. Bai, B.Y.d. Schueren, P. Marynen, J. Cassiman, and H.V.d. Berghe. 1993. Spatial and temporal changes in the expression of fibroglycan (syndecan-2) during mouse embryonic development. Development 119:841-854.

David, G., V. Lories, B. Decock, P. Marynen, J. Cassiman, and H.V.d. Berghe. 1990. Molecular cloning of a phosphatidylinositol-anchored membrane heparan sulfate proteoglycan from human lung fibroblasts. J. Cell Biol. 111:3165-3176.

deCurtis, I., V. Quaranta, R.N. Tamura, and L.F. Reichardt. 1991. Laminin receptors in the retina: Sequence analysis of the chick integrin $\alpha6$ subunit. Evidence for transcriptional and posttranslational regulation. J. Cell Biol 113:405-416.

Dou, C. and J.M. Levine. 1994. Inhibition of neurite growth by the NG2 chondrotin sulfate proteoglycan. J. Neurosci. 14:7616-7628.

Dou, C.-L. and J.M. Levine. 1994. Inhibition of neurite growth by the NG2 chondroitin sulfate proteoglycan. J. Neurosci. 14:7616-7628.

Erickson, H.P. and M.A. Bourdon. 1989. Tenascin: An extracellular matrix protein prominent in specialized embryonic tissues and tumors. Annu. Rev. Cell Biol. 5:71-92.

Evans, D.L., C.J. Marshall, P.B. Christey, and R.W. Carrell. 1992. Heparin binding site, conformational change and activation of thrombin. Biochemistry 31:12629-12642.

Faissner, A., A. Clement, A. Lochter, A. Streit, C. Mandl, and M. Schachner. 1994. Isolation of a neural chondroitin sulfate proteoglycan with neurite outgrowth promoting properties. J. Cell Biol. 126:783-799.

Feany, M.B., S. Lee, R.H. Edwards, and K.M. Buckley. 1992. The synaptic vesicle protein SV2 is a novel type of transmembrane transporter. Cell 70:861-867.

Fedarko, N.S. and H.E. Conrad. 1986. A unique heparan sulfate in the nuclei of hepatocytes: structural changes with the growth state of cells. J. Cell Biol. 102:587-599.

Fedarko, N.S., M. Ishihara, and H.E. Conrad. 1989. Control of cell division in hepatoma cells by exogenous heparan sulfate proteoglycan. J. Cell. Physiol 139:287-294.

Filmus, J., J.G. Church, and R.N. Buick. 1988. Isolation of a cDNA corresponding to a developmentally regulated transcript in rat intestine. Mol. Cell Biol. 8:4243-4249.

Filmus, J., W. Shi, Z.M. Wong, and M.J. Wong. 1995. Identification of a new membrane-bound heparan sulfate proteoglycan. Biochem. J. 311:561-565.

Fryer, H.J.L., G.M. Kelly, L. Molinaro, and S. Hockfield. 1992. The high molecular weight Cat-301 chondroitin sulfate proteoglycan from brain is related to the large aggregation proteoglycan from cartilage, aggrecan. J. Biol. Chem. 267:9874-9883.

Fu, Y.-M., P. Spirito, Z.-X. Yu, S. Biro, J. Sasse, J. Lei, V.J. Ferrans, S.E. Epstein, and W. Casscells. 1991. Acidic fibroblast growth factor in the developing rat embryo. J. Cell Biol. 114:1261-1273.

Gallagher, J.T. 1989. The extended family of proteoglycans: social residents of the pericellular zone. Curr. Opin. Cell Biol. 1:1201-1218.

Gallagher, J.T. and M. Lyon. (1989). Molecular organisation and functions of heparan sulphate. In Heparin, D. A. Lane and U. Lindahl, editors. (London: Edward Arnold), pp. 135-158.

Gallagher, J.T., J.E. Turnbull, and M. Lyon. 1992. Patterns of sulphation in heparan sulfate: polymorphism based on a common structural theme. Int. J. Biochem. 24:553-560.

Gimenez-Gallego, G., J. Rodkey, C. Bennet, M. Rios-Dandelore, J. DiSalvo, and K. Thomas. 1985. Brain-derived acidic fibroblast growth factor: complete amino acid sequence and homologies. Science 230:1385-1388.

Gonzalez, A., M. Buscaglia, M. Ong, and A. Baird. 1990. Distribution of basic fibroblast growth factor in the 18-day rat fetus: localization in the basement membrane of diverse tissues. J. Cell Biol. 110:753-765.

Goridis, C. and J.F. Brunet. 1992. NCAM: structural diversity, function, and regulation of expression. Semin. Cell Biol. 3:189-197.

Gould, S.E., W.B. Upholt, and R.A. Kosher. 1992. Syndecan 3: A member of the syndecan family of membrane-intercalated proteoglycans that is expressed in high amounts at the onset of chicken limb cartilage differentiation. Proc. Natl. Acad. Sci. USA 89:3271-3275.

Gould, S.E., W.B. Upholt, and R.A. Kosher. 1995. Characterization of chicken syndecan-3 as a heparan sulfate proteoglycan and its expression during embryogenesis. Dev. Biol. 168:438-451.

Grumet, M., P. Milev, T. Sakurai, L. Karthikeyan, M. Bourdon, R.K. Margolis, and R.U. Margolis. 1994. Interactions with tenascin and differential effects on cell adhesion of neurocan and phosphacan, two major chondroitin sulfate proteoglycans of nervous tissue. J. Biol. Chem. 262:12142-12146.

Guimond, S., M. Maccarana, B.B. Olwin, U. Lindahl, and A.C. Rapraeger. 1993. Activating and inhibitory heparin sequences for FGF-2 (Basic FGF). J. Biol. Chem. 268:23906-23914.

Gundersen, R.W. 1987. Response of sensory neurites and growth cones to patterned substrata of laminin and fibronectin in vitro. Dev. Biol. 121:423-431.

Halfter, W. 1993. a heparan sulfate proteoglycan in developing avian axonal tracts. J. Neurosci. 13:2863-2873.

Halfter, W. and B. Schurer. 1994. A new heparan sulfate proteoglycan in the extracellular matrix of the developing chick embryo. Exp. Cell Res. 214:285-296.

Hanemann, C.O., G. Kuhn, A. Lie, C. Gillen, F. Bosse, P. Spreyer, and H.W. Muller. 1993. Expression of decorin mRNA in the nervous system of rat. J. Histochem. Cytochem. 41:1383-1391.

Hart, G.W., R.S. Haltiwanger, G.D. Holt, and W.G. Kelly. 1989. Glycosylation in the nucleus and cytoplasm. Annu. Rev. Biochem. 58:841-874.

Haugen, P.K., P.C. Letourneau, S.L. Drake, L.T. Furcht, and J.B. McCarthy. 1992. A cell surface heparan sulfate proteoglycan mediates neural cell adhesion and spreading on a defined sequence from the C-terminal cell and heparin binding domain of fibronectin, FN-C/H II. J. Neurosci. 12:2597-2608.

Heath, W.F., A.S. Cantrell, N.G. Mayne, and S.R. Jaskunas. 1991. Mutations in the heparin-binding domains of human basic fibroblast growth factor alter its biological acitvity. Biochemistry 30:5608-5615.

Herndon, M.E. and A.D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. Neuron 4:949-961.

Hockfield, S., R.G. Kalb, S. Zaremba, and H.J.L. Fryer. 1990. Expression of neural proteoglycans correlates with the acquisition of mature neruonal properties in the mammalian brain. Cold Spring Harbor Symp. Quant. Biol. 55:505-514.

Hoffman, S., K.L. Crossin, and G.M. Edelman. 1988. Molecular forms, binding properties, and developmental expression pattern of cytotactin and cytotactin-binding proteoglycan, and interactive pair of extracellular matrix molecules. J. Cell Biol. 106:519-532.

Hoffman, S. and G.M. Edlman. 1987. A proteoglycan with HNK-1 antigenic determinants is a neuron-associated ligand for cytotactin. Proc. Natl. Acad. Sci. (USA) 84:2523-2527.

Holmer, E., G. Soderstrom, and L.-O. Andersson. 1979. Eur. J. Biochem. 93:1-5.

Hounsell, E.F. (1989). Structural and conformational analysis of keratan sulphate oligosaccharindes and related carbohydrate structures. In Keratan Sulphate: Chemistry, Biology, Chemical Pathology, H. Greiling and J. E. Scott, editors. (London: The Biochemical Society), pp. 12-15.

51

Hughes, R.A., M. Sendtner, M. Goldfarb, D. Lindholm, and H. Thoenen. 1993. Evidence that fibroblast growth factor 5 is a major muscle-derived survival factor for cultured spinal motoneurons. Neuron 10:369-377.

Hunter, D.D., R. Llinas, M. Ard, J.P. Merlie, and J.R. Sanes. 1992. Expression of S-laminin and laminin in developing rat central nervous system. J. Comp. Neurol. 323:238-251.

Hynes, R.O. (1990). Fibronectins (New York: Springer Verlag).

Hynes, R.O. 1992. Integrins: versatility, modulation, and signaling in cell adhesion. Cell 69:11-25.

Hynes, R.O. and A.D. Lander. 1992. Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. Cell 68:303-322.

Iozzo, R.V., I.R. Cohen, S. Grassel, and A.D. Murdoch. 1994. The biology of perlecan: The multifaceted heparan sulphate proteoglycan of basement membranes and pericellular matrices. Biochem. J. 302:625-639.

Ishai-Michaeli, R., A. Eldor, and I. Vlodavsky. 1990. Heparanase activity expressed by platelets, neurtophils, and lymphoma cells releases active fibroblast growth factor from extracellular matrix. Cell Regulat. 1:833-842.

Ishihara, M., N.S. Fedarko, and H.E. Conrad. 1986. Transport of heparan sulfate into the nuclei of hepatocytes. J. Biol. Chem. 261:13575-13580.

Iwata, M. and S.S. Carlson. 1993. A brain extracellular matrix proteoglycan forms aggregates with hyaluronan. J. Biol. Chem. 268:15061-15069.

Jackson, R.L., S.J. Busch, and A.D. Cardin. 1991. Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. Physiol. Rev. 71:481-539.

Joseph, S.J., M.D. Ford, and V. Nurcombe. 1995. A novel heparan sulfate proteoglycan involved in the regulation of FGFs in mouse embryonic brain is a putative perlecan homologue. Soc. Neurosci. Abstracts 21:1041.

Kan, M., F. Wang, J. Xu, J.W. Crabb, J. Hou, and W.L. McKeehan. 1993. An essential heparin-binding domain in the fibroblast growth factor receptor kinase. Science 259:1918-1921.

Karthikeyan, L., M. Flad, M. Engel, B. Meyer-Puttlitz, R.U. Margolis, and R.K. Margolis. 1994. Immunocytochemical and in situ hybridization studies of the heparan sulfate proteoglycan, glypican, in nervous tissue. J. Cell Sci. 107:3213-3222.

Kennedy, T.E., T. Serafini, and J.R.d.l. Torre. 1994. Netrins are diffusible chemotropic factors for commissural axons in the embyronic spinal cord. Cell 78:425-435.

Kiefer, M.C., J.C. Stephans, K. Crawford, K. Okino, and P.J. Barr. 1990. Ligand-affinity cloning and structure of a cell surface heparan sulfate proteoglycan that binds basic fibroblast growth factor. Proc. Natl. Acad. Sci. USA 87:6985-6989.

Kim, C.W., O.A. Goldberger, R.L. Gallo, and M. Bernfield. 1994. Members of the syndecan family of heparan sulfate proteoglycans are expressed in distinct cell-, tissue-, and development-specific patterns. Mol. Biol. Cell 5:797-805.

Kodukula, K., L.D. Gerber, R. Amthauer, L. Brink, and S. Udenfriend. 1993. Biosynthesis of glycosylphosphatidylinositol (gpi)-anchored membrane proteins in intact cells: Specific amino acid requirements adjacent to the site of cleavage and gpi attachment. J. Cell Biol. 120:657-664.

Kojima, T., N.W. Shworak, and R.D. Rosenberg. 1992. Molecular cloning an expression of two distinct cDNA-encoding heparan sulfate proteoglycan core proteins from a rat endothelial cell line. J. Biol. Chem. 267:4870-4877.

Krueger, N.X. and H. Saito. 1992. A human transmembrane protein-tyrosine phosphatase PTPζ, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases. Proc. Natl. Acad. Sci. USA 89:7417-7421.

Krusius, T. and E. Ruoslahti. 1986. Primary structure of an extracellular matrix proteoglycan core protein deduced form cloned cDNA. Proc. Natl. Acad. Sci. USA 83:7683-7687.

Kusche, M., G. Backstrom, J. Riesenfeld, M. Petitou, J. Choay, and U. Lindahl. 1988. Biosynthesis of heparin. O-sulfation of the antithrombin-binding region. J. Biol. Chem. 263:15474-15484.

Lander, A.D. 1989. Understanding the molecules of neural cell contacts: emerging patterns of structure and function. Trends in Neurosci. 12:189-195.

Lander, A.D. 1990. Mechanisms by which molecules guide axons. Curr. Opin. Cell Biol. 2:907-913.

Lander, A.D. 1994. Targeting the glycosaminoglycan-binding sites on proteins. Chemistry and Biology 1:73-78.

Lander, A.D. and A.L. Calof. (1993). Extracellular matrix in the developing nervous system. In Molecular Genetics of Nervous System Tumors, A. J. Levine and H. H. Schmidek, editors. (New York: Wiley), pp. 341-355.

Lander, A.D., D.K. Fujii, and L.F. Reichardt. 1985. Purification of a factor that promotes neurite outgrowth: isolation of laminin and associated molecules. J. Cell Biol. 101:898-913.

LeBaron, R.G., J.D. Esko, A. Woods, S. Johansson, and M. Höök. 1988. Adhesion of glycosaminoglycan-deficient Chinese hamster ovary cell mutants to fibronectin substrata. J. Cell Biol. 106:945-952.

Lee, M.K. and A.D. Lander. 1991. Analysis of affinity and structural selectivity in the binding of proteins to glycosaminoglycans: development of a sensitive electrophoretic approach. Proc. Natl. Acad. Sci. (USA) 88:2768-2772.

Levine, J.M. 1994. Increased expression of the NG2 chondrotin-sulfate proteoglycan after brain injury. J. Neurosci. 14:4716-4730.

Levine, J.M. and W.B. Stallcup. 1987. Plasticity of developing cerebellar cells in vitro studied with antibodies against the NG2 antigen. J. Neurosci. 7:2721-2731.

Lin, W. 1990. Immunogold localization of basal laminar heparan sulfate proteoglycan in rat brain and retinal capillaries. Brain Res. Bull. 24:533-536.

Lindahl, U., G. Backstrom, L. Jansson, and A. Hallén. 1973. Biosynthesis of heparin II. Formation of sulfamino groups. J. Biol. Chem. 248:7234-7241.

Lindahl, U., D.S. Feingold, and L. Roden. 1986. Biosynthesis of heparin. Trends Biochem. Sci. 11:221-225.

Litwack, E.D. (1995). Expression and function of proteoglycans in the nervous system. Doctoral Dissertation, Massachusetts Institute of Technology.

Litwack, E.D., C.S. Stipp, A. Kumbasar, and A.D. Lander. 1994. Neuronal expression of glypican, a cell-surface

glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan, in the adult rat nervous system. J. Neurosci. 14:3713-3724.

Liu, C.-S. and J.-Y. Chang. 1987. The heparin binding site of human anti-thrombin III. Selective chemical modification at Lys114, Lys125, and Lys287 impairs its heparin cofactor activity. J. Biol. Chem. 262:17356-17361.

Lobb, R.R. and J.W. Fett. 1984. Purification of two distinct growth factors from bovine neural tissue by heparin affinity chromatography. Biochemistry 23:6295-6299.

López-Casillas, F., S. Cheifetz, J. Doody, J.L. Andres, W.S. Lane, and J. Massagué. 1991. Structure and expression of the membrane proteoglycan betaglycan, a componint of the TGF-β receptor system. Cell 67:785-795.

López-Casillas, F., J.L. Wrana, and J. Massagué. 1993. Betaglycan presents ligand to the TGFβ signalling receptor. Cell 73:1435-1444.

Low, M.G. 1989. The glycosyl-phosphatidylinositol anchor of membrane proteins. Biochim. Biophys. Act. 988:427-454.

Lyons, R.M. and H.L. Moses. 1990. Transforming growth factors and the regulation of cell proliferation. Eur. J. Biochem. 187:467-473.

Ma, E., R. Morgan, and E.W. Godfrey. 1994. Distribution of agrin mRNA in the chick embryo nervous system. J. Neurosci. 14:2943-2952.

Mach, H., D.B. Volkin, C.J. Burke, C.R. Middaugh, R.J. Linhardt, J.R. Fromm, D. Loganathan, and L. Mattson. 1993. Nature of the interaction of heparin with acidic fibroblast growth factor. Biochem. 32:5480-5489.

Marcum, J. and R. Rosenberg. 1989. Heparin and Related Polysaccharides. Ann. N. Y. Acad. Sci. 556:81-94.

Maurel, P., U. Rauch, M. Flad, R.K. Margolis, and R.U. Margolis. 1994. Phosphacan, a chondroitin sulfate proteoglycan of brain that interacts with neurons and neural cell adhesion molecules, is an extracellular variant of a receptor-type protein tyrosine phosphatase. Proc. Natl. Acad. Sci. USA 91:2512-2516.

McMahon, U.J. 1990. The agrin hypothesis. Cold Spring Harbor Symp. Quant. Biol. 50:407-418.

Milev, P., D.R. Friedlander, T. Sakurai, L. Karthikeyan, M. Flad, R.K. Margolis, M. Grumet, and R.U. Margolis. 1994. Interactions of the

chondroitin sulfate proteoglycan phosphocan, the extracellular domain of a receptor-type protein tyrosine phosphatase, with neurons, glia, and neural cell adhesion molecules. J. Cell. Biol. 127:1703-1715.

Miller, B., A.M. Sheppard, and A.L. Pearlman. 1992. Expression of two chondroitin sulfate proteoglycan core proteins in the subplate pathway of early cortical afferents. Soc. Neurosci. Abstr. 18:778.

Muir, D. and M. Manthorpe. 1992. Stromelysin generates a fibronectin fragment that inhibits Schwann cell proliferation. J. Cell Biol. 116:177-185.

Murphy, M., J. Drago, and P.F. Bartlett. 1990. Fibroblast growth factor stimulates the proliferation and differentiation of neural precursor cells in vitro. J. Neurosci. Res. 25:463-475.

Murphy-Ullrich, J.E., L.G. Westrick, J.D. Esko, and D.R. Mosher. 1988. Altered metabolism of thrombospondin by Chinese hamster ovary cells defective in glycosaminoglycan synthesis. J. Biol. Chem. 263:6400-6406.

Nakato, H., T.A. Futch, and S.B. Selleck. 1995. The division abnormally delated (dally) gene: A putative integral membrane proteoglycan required for cell division patterning during post-embryonic development of the nervous system in Drosophila. Development 121:3687.

Neugebauer, K.M., C.J. Emmett, K.A. Venstrom, and L.F. Reichardt. 1991. Vitronectin and thrombospondin promote retinal neurite outgrowth: developmental regulation and role of integrins. Neuron 6:345-358.

Nugent, M.A. and E.R. Edelman. 1992. Kinetics of basic fibroblast growth factor binding to its receptor and heparan sulfate proteoglycan: A mechanism for cooperativity. Biochemistry 31:8876-8883.

Nurcombe, V., M.D. Ford, J.A. Wildschut, and P.F. Bartlett. 1993. Developmental regulation of neural response to FGF-1 and FGF-2 by heparan sulfate proteoglycan. Science 260:103-106.

O'Connor, L.T., J.C. Lauterborn, C.M. Gall, and M.A. Smith. 1994. Localization and alternative splicing of agrin mRNA in adult rat brain Transcripts encoding isoforms that aggregate acetylcholine receptors are not restricted to cholinergic regions. J. Neurosci. 14:1141-1152.

O'Shea, K.S. and V.M. Dixit. 1988. Unique distribution of the extracellular matrix component thrombospondin in the developing mouse embryo. J. Cell Biol. 107:2737-2748.

O'Shea, K.S., J.S.T. Rheinheimer, and V.M. Dixit. 1990. Deposition and role of thrombospondin in the histogenesis of the cerebellar cortex. J. Cell Biol. 110:1275-1284.

Oohira, A., F. Matsui, E. Watanabe, Y. Kushima, and N. Maeda. 1994. Developmentally regulated expression of a brain specific species of chondroitin sulfate proteoglycan, neurocan, identified with a monoclonal antibody 1G2 in the rat cerebrum. Neurosci. 60:145-157.

Ornitz, D.M., A.B. Herr, M. Nilsson, J. Westman, C.M. Svahn, and G. Waksman. 1995. FGF binding and FGF receptor activation by synthetic heparan-derived di- and trisaccharides. Science 268:432-436.

Ornitz, D.M., A. Yayon, J.G. Flanagan, C.M. Svahn, E. Levi, and P. Leder. 1992. Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cell. Mol. Cell. Biol. 12:240-247.

Pantoliano, M.W., R.A. Horlick, B.A. Springer, D.E.V. Dyk, T. Tobery, D.R. Wetmore, J.D. Lear, A.T. Nahapetian, J.D. Bradley, and W.P. Sisk. 1994. Multivalent ligand-receptor binding interactions in the fibroblast growth factor system produce a cooperative growth factor and heparin mechanism for receptor dimerization. Biochemistry 33:10229-10248.

Partanen, A. and I. Thesleff. 1987. Localization and quantification of ^{125}I-epidermal growth factor binding in mouse embryonic tooth and other embryonic tissues at different developmental stages. Dev. Biol. 120:186-197.

Pelton, R.W., B. Saxena, M. Jones, H.L. Moses, and L.I. Gold. 1991. Immunohistochemical localization of TGFβ1, TGFβ2, and TGFβ3 in the mouse embryo: Expresson patterns suggest multiple roles during embryonic development. J. Cell. Biol. 115:1091-1105.

Perides, G., F. Rahemtulla, W.S. Lane, R.A. Asher, and A. Bignami. 1992. Isolation of a large, aggregating proteoglycan from human brain. J. Biol. Chem. 267:23883-23887.

Persohn, E. and M. Schachner. 1987. Immunoelectron microscopic localization of the neural cell adhesion molecules L1 and N-CAM during postnatal development of the mouse cerebellum. J. Cell Biol. 105:569-576.

Peterson, C.B., M.T. Morgon, and M.N. Blackburn. 1987. Identification of a lysyl residue in antithrombin which is essential for heparin binding. J. Biol. Chem. 262:8061-8065.

Pratt, C.W., H.C. Whinna, and F.C. Church. 1992. A comparison of three heparin-binding serine proteinase inhibitors. J. Biol. Chem. 267:8795-8801.

Rao, J.S., D. Hantai, and B.W. Festoff. 1992. Thrombospondin, a platelet α-granule and matrix protein, is increased in muscle basement membrane of patients with amyotrophic lateral sclerosis. J. Neuro. Sci. 113:99-107.

Rapraeger, A., A. Krufka, and B.B. Olwin. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science 252:1705-1708.

Rauch, U., P. Gao, A. Janetzko, A. Flaccus, L. Hilgenberg, H. Tekotte, R.K. Margolis, and R.U. Margolis. 1991. Isolation and characterization of developmentally regulated chondroitin sulfate and chondrotin/keratan sulfate proteoglycans of brain identified with monoclonal antibodies. J. Biol Chem. 266:14785-14801.

Rauch, U., L. Karthikeyan, P. Maurel, R.U. Margolis, and R.K. Margolis. 1992. Cloning and primary structure of neurocan, a developmentally regulated, aggregating chondroitin sulfate proteoglycan. J. Biol. Chem. 267:19536-19547.

Reichardt, L.F. and K.J. Tomaselli. 1991. Extracellular matrix molecules and their receptors: Functions in neural development. Annu. Rev. Neurosci. 14:531-570.

Reyes, A.A., R. Akeson, L. Brezina, and G.J. Cole. 1990. Structural requirements for neural cell adhesion molecule-heparin interaction. Cell Regulation 1:567-576.

Rogers, S.L., K.J. Edson, P.C. Letourneau, and S.C. McLoon. 1986. Distribution of laminin in the developing peripheral nervous system of the chick. Dev. Biol. 113:429-435.

Rogers, S.L., P.C. Letourneau, S.L. Palm, J. McCarthy, and L.T. Furcht. 1983. Neurite extension by peripheral and central nervous system neurons in response to substratum-bound fibronectin and laminin. Dev. Biol. 98:212-220.

Rogers, S.L., P.C. Letourneau, and I.V. Pech. 1989. The role of fibronectin in neural development. Dev. Neurosci. 11:248-265.

Rosenberg, R.D. and P.S. Damus. 1973. The purification and mechanism of action of human antithrombin-heparin cofactor. J. Biol. Chem. 248:6490-6505.

Rosenfeld, L. and I. Danishefsky. 1986. A frament of antithrombin that binds both heparin and thrombin. Biochem. J. 237:639-646.

Saksela, O., D. Moscatelli, A. Sommer, and D.B. Rifkin. 1988. Endothelial cell derived heparan sulfate binds basic fibroblast growth factor and protects it from proteolytic degradation. J. Cell Biol. 107:743-751.

San Antonio, J.D., M.J. Karnovsky, S. Gay, R.D. Sanderson, and A.D. Lander. 1994. Interactions of syndecan-1 and heparin with human collagens. Glycobiology 4:327-332.

San Antonio, J.D., J. Slover, J. Lawler, M.J. Karnovsky, and A.D. Lander. 1993. Specificity in the interactions of extracellular matrix proteins with subpopulations of the glycosaminoglycan heparin. Biochem. 32:4746-4755.

Sanderson, R.D. and M. Bernfield. 1988. Molecular polymorphism of a cell surface proteoglycan: distinct structures on simple and stratified epithelium. Proc. Natl. Acad. Sci. USA 85:9562-9566.

Saunders, S., M. Jalkanen, S. O'Farrell, and M. Bernfield. 1989. Molecular cloning of syndecan, an integral membrane proteoglycan. J. Cell Biol. 108:1547-1556.

Scholzen, T., M. Solursh, S. Suzuki, R. Reiter, J.L. Morgan, A.M. Buchberg, L.D. Siracusa, and R.V. Iozzo. 1994. The murine decorin. J. Biol. Chem. 269:28270-28281.

Scranton, T.W., M. Iwata, and S.S. Carlson. 1993. The SV2 protein of synaptic vesicles is a keratan sulfate proteoglycan. J. Neurochem. 61:29-44.

Serafini, T., T.E. Kennedy, M.J. Galko, C. Mirzayan, T.M. Jessell, and M. Tessier-Lavigne. 1994. The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. Cell 78:409-424.

Sheppard, M.M., S.K. Hamilton, and A.L. Pearlman. 1991. Changes in the distribution of extracellular matrix components accompany early morphogenetic events of mammalian cortical development. J. Neurosci. 11:3928-3942.

Shitara, K., H. Yamada, K. Watanabe, M. Shimonaka, and Y. Yamaguchi. 1994. Brain-specific receptor-type protein tyrosine phosphatase RPTPβ-is a chondroitin sulfate proteoglycan in vivo. J. Biol. Chem. 269:20189-20193.

Shworak, N.W., M. Shirakawa, S. Colliec-Jouault, J. Liu, R.C. Mulligan, L.K. Birinyi, and R.D. Rosenberg. 1994b. Pathway-specific regulation of the synthesis of anticoagulantly active heparan sulfate. J. Biol. Chem. 269:24941-24952.

Smith, J.W., N. Dey, and D.J. Knauer. 1990. Heparin binding domain of antithrombin III: Characterization using a synthetic peptide directed polyclonal antibody. Biochemistry 29:8950-8957.

Smith, J.W. and D.J. Knauer. 1987. A heparin-binding site in ATIII. Identification, purification and amino acid sequence. J. Biol. Chem. 262:11964-11973.

Snow, A.D., H. Mar, D. Nochlin, H. Kresse, and T.N. Wight. 1992. Peripheral distribution of dermatan sulfate proteoglycans (decorin) in amyloid-containing plaques and their presence in neurofibrillary tangles of Alzheimer's disease. J. Hisotchem. Cytochem. 40:105-113.

Spivak-Kroizman, T., M.A. Lemmon, I. Dikic, J.E. Ladbury, D. Pinchasi, J. Huang, M. Jaye, G. Crumley, J. Schlessinger, and I. Lax. 1994. Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation. Cell 79:1015-1024.

Spring, J., S. Paine-Saunders, R.O. Hynes, and M. Bernfield. 1994. Drosophila syndecan: Conservation of a cell-surface heparan sulfate proteoglycan. Proc. Natl. Acad. Sci. USA 91:3334-3338.

Stallcup, W.B., L. Beasley, and J. Levine. 1983. Cell surface molecules that characterize different stages in the development of cerebellar interneurons. Cold Spring Harbor Symp. Quant. Biol. 48:761-774.

Steindler, D.A., D. Settles, H.P. Erickson, E.D. Laywell, A. Yoshiki, A. Faissner, and M. Kusakabe. 1995. Tenascin knockout mice: barrels, boundary molceules, and glial scars. J. Neurosci. 15:1971-1983.

Stewart, G.R. and A.L. Pearlman. 1987. Fibronectin-like immunoreactivity in the developing cerebral cortex. J. Neurosci. 7:3325-3333.

Stipp, C.S., E.D. Litwack, and A.D. Lander. 1994. Cerebroglycan: an integral membrane heparan sulfate proteoglycan that is unique to the developing nervous system and expressed specifically during neuronal differentiation. J. Cell Biol. 124:149-160.

Stuckey, J.A., R.S. Charles, and B.F.P. Edwards. 1992. A model of the platelet factor 4 complex with heparin. Proteins 14:277-287.

Stuhlsatz, J.W., R. Keller, G. Becker, M. Oeben, L. Lennarts, D.C. Fisher, and G. Greiling. (1989). Structure of keratan sulphate proteoglycans: core proteins, linkage regions, carbohydrate chains. In Keratan Sulphate: Chemistry, Biology, Chemical Patholgy, H. Greiling and J. E. Scott, editors. (London: The Biochemical Society), pp. 1-11.

Sun, X.-J. and J.-Y. Chang. 1989. Heparin binding domain of human antithrombin III inferred from the sequential reduction of its three disulfide linkages. An efficient method for structural analysis of partially reduced proteins. J. Biol. Chem. 264:11288-11293.

Thomas, K.A., M. Rios-Candelore, and S. Fitzpatrick. 1984. Proc. Natl. Acad. Sci. USA 81:357.

Tsen, G., W. Halfter, S. Kroger, and G.J. Cole. 1995. Agrin is a heparan sulfate proteoglycan. J. Biol. Chem. 270:3392-3399.

Turnbull, J.E., D.G. Fernig, Y. Ke, M.C. Wilkinson, and J.T. Gallagher. 1992. Identification of the basic fibroblast growth factor binding sequence in fibroblast heparan sulfate. J. Biol. Chem. 267:10337-10341.

Ullrich, A. and J. Schlessinger. 1990. Signal transduction by receptors with tyrosine kinase activity. Cell 61:203-212.

Unsicker, K., H. Reichert-Preibsch, R. Schmidt, B. Pettmann, G. Labourdette, and M. Sensenbrenner. 1987. Astroglial and fibroblast growth factors have neurotrophic functions for cultured periperal and central nervous system neurons. Proc. Natl. Acad. Sci. USA 84:5459-5463.

Vogel, K.G., M. Paulsson, and D. Heinegard. 1984. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. Biochem. J. 223:587-597.

Walicke, P. 1988. Basic and acidic fibroblast growth factors have trophic effects on neurons from multiple CNS regions. J. Neurosci. 8:2618-2627.

Walicke, P., W.M. Cowan, N. Ueno, A. Baird, and R. Guillemin. 1986. Fibroblast growth factor promotes survival of dissociated hippocampal neurons and enhances neurite extension. Proc. Natl. Acad. Sci. USA 83:3012-3016.

Watanabe, K., H. Yamada, and Y. Yamaguchi. 1995. K-glypican: a novel gpi-anchored heparan sulfate proteoglycan that is highly expressed in developing brain and kidney. J. Cell Biol. 130:1207-1218.

Watkins, S.C., G.W. Lynck, L.P. Kane, and H.S. Slayter. 1990. Thrombospondin expression in traumatized skeletal muscle. Correlation of appearance with post-trauma regeneration. Cell Tissue Res. 261:73-84.

Wilcox, J.B. and J.R. Unnerstall. 1991. Expression of acidic fibroblast growth factor mRNA in the developing and adult brain. Neuron 6:397-409.

Wrana, J.L., L. Atisano, J. Cárcamo, A. Zentella, J. Doody, M. Laiho, X.-F. Wang, and J. Massagué. 1992. TGF-β signals through a heteromeric protein kinase receptor complex. Cell 71:1003-1014.

Yamada, H., K. Watanabe, M. Shimonaka, and Y. Yamaguchi. 1994. Molecular cloning of brevican, a novel brain proteoglycan of the aggrecan/versican family. J. Biol. Chem. 269:10119-10126.

Yamaguchi, Y., D.M. Mann, and E. Ruoslahti. 1990. Negative regulation of transforming growth factor-β by the proteoglycan decorin. Nature 346:281-284.

Yayon, A., M. Klagsbrun, J.D. Esko, P. Leder, and D.M. Ornitz. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. Cell 64:841-848.

Yurchenco, P.D. and J.J. O'Rear. (1993). . In Molecular and Cellular Aspects of Basement Membranes, D. H. Rohrback and R. Timpl, eds. (New York: Academic Press), pp. 19-47.

Table 1. Proteoglycan Families in the Nervous System

Molecule	GAG Type	Expression and localization[a]	Notes	Refs
Syndecan family PGs				
syndecan-1	HS/CS	(t.m.) early neural plate	(core M_r 33 kD[b]); conserved cytoplasmic domain that interacts with cytoskeleton common to all family members	38,39
syndecan-2	HS	(t.m.) mRNA detected in brain; possibly from meninges	(core M_r 23 kD[b]);	38,40
syndecan-3	HS	(t.m.) widespread in neonatal brain; floorplate of chick neural tube; neurons in culture	(core M_r 43 kD[b])	38,41 42, 45,46
ryudocan	HS/CS	(t.m.) CNS and PNS	(core M_r 22 kD[b]); also called syndecan-4	43,44
Drosophila syndecan	HS	(t.m.) CNS and PNS	(core M_r 39 kD[b])	47
Glypican family PGs				
glypican	HS	(GPI) ventricular zones of early CNS; widespread in CNS later; abundant but restricted in adult; esp. assoc. w/projection neurons; axon tracts	(core M_r 64 kD); lipid (GPI) anchored to plasma membrane; pattern of 14 cys residues conserved in all family members	15, 48-50
OCI-5	HS	(GPI) mRNA detected in rat brain	(core M_r 69 kD)	51-53
k-glypican	HS	(GPI) abundant in developing rat brain, esp. in ventricular zones; less abundant in adult	(core M_r 57.5 kD)	53
cerebro-glycan	HS	(GPI) immature post-mitotic neurons; assoc. w/ axon tracts of developing nervous system	(core M_r 57 kD); restricted to nervous system; (see chapter 3, this thesis)	54
dally	-	(prob. GPI) Dros. nervous system incl. morphogentic furrow and lamina furrow	(core M_r 63 kD); orig. identified in genetic screen for cell cycle progression mutants	55

(Table continued on next page)

Table 1. Proteoglycan Families in the Nervous System (continued)

Molecule	GAG Type	Expression and localization[a]	Notes	Refs
Aggregating CSPGs				
aggrecan	CS/KS	(sec.) found in chick brain	(core M_r 180-370 kD)	1
versican	CS	(sec.) glia of CNS and PNS	(core M_r 290-400 kD)	2
neurocan	CS	(sec.) neurons in cerebellum; cortical subplate; and barrel field boundaries in cortex	(core M_r 220-245 kD)	3,4
Cat-301	CS	(sec.) spinal motoneurons, neurons in visual circuits	(core M_r 580kD); expression highly restricted to subsets of neurons	5,6
brevican	CS	(sec.) brain; glial cells in culture	(core M_r 145kD)	7
pgT1	CS	(sec.) white matter and gray matter throughout nervous system	(core M_r 300 kD)	8
DSD-1	CS/DS	(sec.) barrel field boundaries of somatosensory cortex; glial cells in vitro	(core M_r 350-400 kD)	9,10
Other PGs				
decorin	CS/DS	(sec.) pons; floorplate of spinal cord; amyloid plaques; Schwann cells; neurons	(core M_r 38 kD); binds to netrin, probably via core protein; other decorin-type molecules outside nervous system	11-15
NG2	CS	(t.m.) glial progenitor cells; deep regions of molecular layer of cerebellum	(core M_r 300 kD); neurite repulsive activity in vitro	16-19
phosphacan	CS	(sec./t.m.) glia, including Bergmann glia of cerebellum; spinal cord roof plate; cortical subplate	(core M_r 300-400 kD); alternatively spliced form or RPTPβ, a brain-specif. receptor protein tyrosine phosphatase	7, 20-24
SV2	KS	(ves.) membranes of synaptic vesicles throughout brain	(core M_r 100-250 kD); homology to neurotrans-mitter transporters	25-28
perlecan	HS	(sec.) basement membrane of capillaries within brain; basal lamina of neural tube; endoneurium of peripheral nerves	(core M_r 400-700 kD); a small HSPG of core Mr~ 45kDa and homology to perlecan has been found in emb. mouse neuroepithelium	29-32
agrin	HS	(sec.) neurons and Schwann cells of PNS; ventricular zones of CNS; spinal cord motoneurons; throughout adult brain	(core M_r 250 kD); non-glycanated forms exist; orig. purified on the basis of AChR clustering activity; prob. involved in synaptogenesis	33-37

Table 1. Proteoglycan Families in the Nervous System (continued)

[a] Expression and localization abbreviations are (t.m.), transmembrane molecule; (GPI), glycosyl-phosphatidylinositol- linked protein core; (sec.), secreted proteoglycan; (ves.), molecule that has been localized to the lumen of synaptic vesicles.

[b] Syndecan family core protein values shown here are those predicted from protein sequences (not the much larger values obtained from SDS-PAGE analysis of cores).

The data in this table was assembled from information found in (Litwack, 1995, Chapter 1).

TABLE REFERENCES:

(1) (Krueger and Saito, 1992)
(2) (Perides et al., 1992)
(3) (Rauch et al., 1992)
(4) (Oohira et al., 1994)
(5) (Fryer et al., 1992)
(6) (Hockfield et al., 1990)
(7) (Yamada et al., 1994)
(8) (Iwata and Carlson, 1993)
(9) (Faissner et al., 1994)
(10) (Steindler et al., 1995)
(11) (Krusius and Ruoslahti, 1986)
(12) (Scholzen et al., 1994)
(13) (Snow et al., 1992)
(14) (Hanemann et al., 1993)
(15) (Litwack, 1995)
(16) (Stallcup et al., 1983)
(17) (Levine and Stallcup, 1987)
(18) (Dou and Levine, 1994)
(19) (Levine, 1994)

(20) (Rauch et al., 1991)
(21) (Maurel et al., 1994)
(22) (Shitara et al., 1994)
(23) (Milev et al., 1994)
(24) (Miller et al., 1992)
(25) (Scranton et al., 1993)
(26) (Feany et al., 1992)
(27) (Bajjalieh et al., 1992)
(28) (Bahr et al., 1992)
(29) (Iozzo et al., 1994)
(30) (Lin, 1990)
(31) (Halfter and Schurer, 1994)
(32) (Joseph et al., 1995)
(33) (Tsen et al., 1995)
(34) (McMahon, 1990)
(35) (Ma et al., 1994)
(36) (Halfter, 1993)
(37) (O'Connor et al., 1994)
(38) (Bernfield and Hooper, 1991)

(39) (Saunders et al., 1989)
(40) (David et al., 1993)
(41) (Carey et al., 1992)
(42) (Gould et al., 1992)
(43) (Kojima et al., 1992)
(44) (Baciu et al., 1994)
(45) (Kim et al., 1994)
(46) (Gould et al., 1995)
(47) (Spring et al., 1994)
(48) (Litwack et al., 1994)
(49) (Karthikeyan et al., 1994)
(50) (Carey et al., 1993)
(51) (Filmus et al., 1988)
(52) (Filmus et al., 1995)
(53) (Watanabe et al., 1995)
(54) (Stipp et al., 1994)
(55) (Nakato et al., 1995)

65

Figure 1. Proteoglycan and glycosaminoglycan structure

To illustrate several features of proteoglycan and glycosaminoglycan structure, the structures of the two major families of cell surface HSPGs are compared. **(A)** A representative structure of the syndecan family of HSPGs. The syndecan family is characterized by a conserved intracellular domain that contains potential tyrosine phosphorylation sites and that is also thought to interact with cytoskeletal components. The syndecan family members syndecan-1 and ryudocan (syndecan-4) have been shown to substituted with both heparan sulfate and chondroitin sulfate chains. **(B)** A representative structure of the glypican family of HSPGs. The glypicans are characterized by a lipid anchorage to the outer leaflet of the plasma membrane, and a conserved pattern of cysteine residues that is likely to result in a highly folded tertiary structure, stabilized by disulfide bonds. The heparan sulfate chains on both PGs contains regions of highly modified structure separated by spans of sparsely modified disaccharides. **(C)** Representative structures of sparsely modified heparan sulfate, highly modified heparan sulfate, and chondroitin/dermatan sulfate are shown, as well as the basic structure of the glycosyl-phosphatidylinositol lipid anchor of the glypican family PGs. GlcNAc = N-acetylglucosamine; GlcUA = glucuronic acid; $GlcNSO^{3-}$ = N-sulfoglucosamine; IdoUA = iduronic acid; GalNAc = N-acetyl-galactosamine; 4S, 2S, etc. indicate sites of O-sulfation.

This figure courtesy of Mary Herndon

A

B

C

Heparan sulfate, N-acetlyated domains (▢✕▢◇):

CH$_2$O-SO$_3^-$ COOH   CH$_2$OH   COOH

OH   OH   OH   OH

NH-C-CH$_3$ OH   NH-C-CH$_3$ OH
      ‖O   ‖O

GlcNAc,6S   GlcUA   GlcNAc   GlcUA
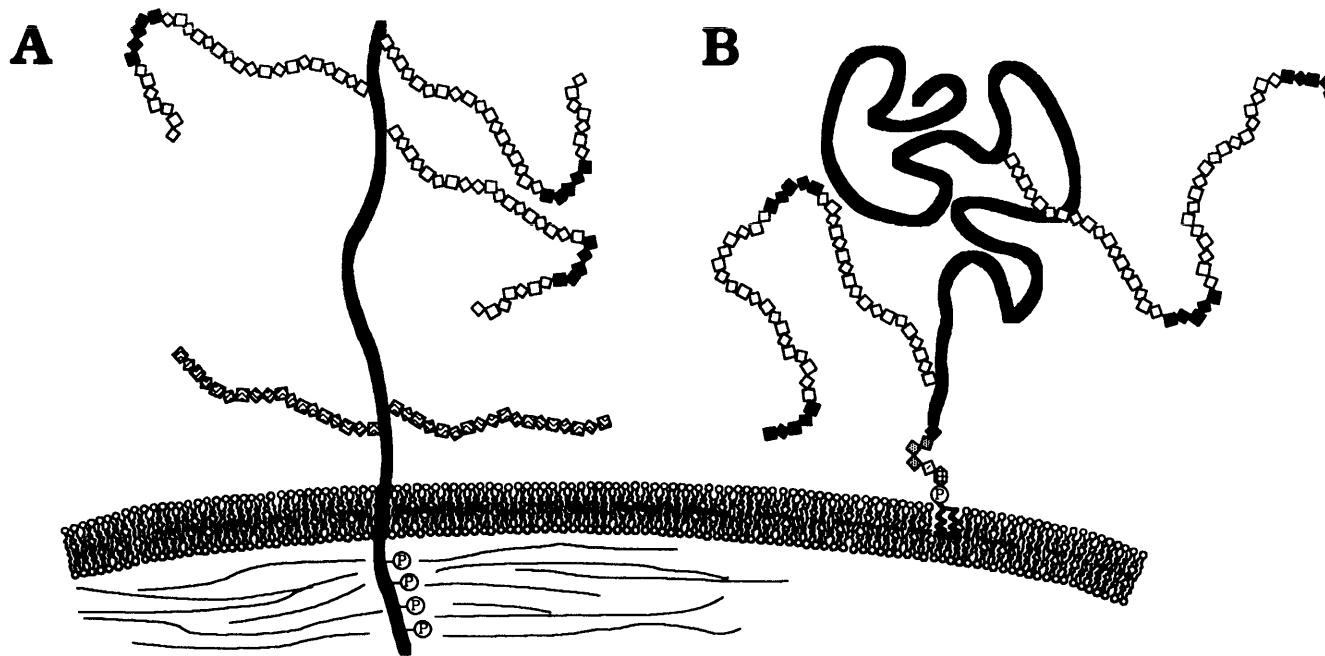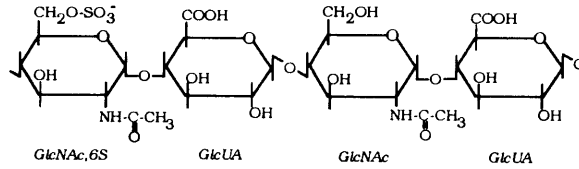
Heparan sulfate, highly sulfated domains (■◆■◆):

CH$_2$O-SO$_3^-$ COOH   CH$_2$O-SO$_3^-$ COOH

OH   OH   O-SO$_3^-$ OH

NH-SO$_3^-$ O-SO$_3^-$ NH-SO$_3^-$ O-SO$_3^-$

GlcNSO$_3^-$,6S   GlcUA,2S   GlcNSO$_3^-$,3,6S   IdoUA,2S
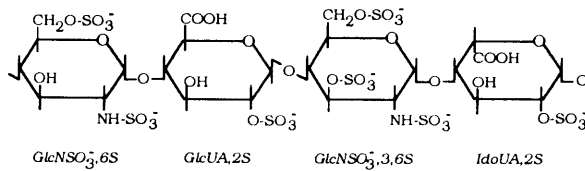
Chondroitin sulfate/Dermatan sulfate (▨◈▨◈):
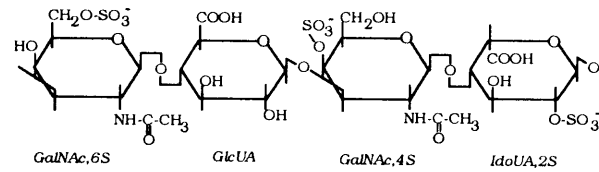
CH$_2$O-SO$_3^-$ COOH   SO$_3^-$ CH$_2$OH

HO   OH   O   COOH   OH

NH-C-CH$_3$ OH   NH-C-CH$_3$ O-SO$_3^-$
      ‖O   ‖O

GalNAc,6S   GlcUA   GalNAc,4S   IdoUA,2S

Potential tyrosine phosphorylation site: -Ⓟ

Glycosylphosphatidylinositol anchor:

◆ — p-ethanolamine
} — mannose residues
— glucosamine
} — phosphoinositol
} — diacylglycerol
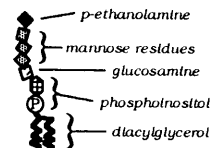
Figure 2. Protein binding sites within heparan sulfate chains
(A) The structure of the antithrombin III binding site. Functional groups required for binding with high affinity to antithrombin III are indicated with asterisks. The dagger indicates the location of a rare 3-O-sulfate that is also required for binding to antithrombin III. The majority of 3-O-sulfates found on heparan sulfate chains are found within antithrombin III binding sites. **(B)** The structure of a disaccharide required for binding to FGF-2. **(C)** The structure of a disaccharide required for binding to FGF-1. The structures in B and C occur in the context of longer oligosaccharides with variable structures. **Abbreviations:** $GlcNSO^{3-}$, N-sulfoglucosamine; GlcUA, glucuronate; IdoUA, iduronate; 2S, 3S, 6S, indicate sites of O-sulfations; R, -H or $-SO^{3-}$ substituent.

This figure courtesy of Mary Herndon

A. Antithrombin III

$CH_2O\text{-}SO_3^-$ *   COOH $CH_2O\text{-}SO_3^-$ $CH_2O\text{-}R$

OH   OH $O\text{-}SO_3^-$ † COO^- * $O\text{-}R$

$NH\text{-}SO_3^-$ OH $NH\text{-}SO_3^-$ * $O\text{-}R$ $NH\text{-}SO_3^-$ *

$GlcNSO_3^-,6S$ $GlcUA$ $GlcNSO_3^-,3,6S$ $IdoUA\pm2S$ $GlcNSO_3^-,\pm3S,\pm6S$

B. FGF-2

CH_2OH

COO^- OH

OH $O\text{-}SO_3^-$ $NH\text{-}SO_3^-$

$IdoUA,2S$ $GlcNSO_3^-$

C. FGF-1

$CH_2O\text{-}SO_3^-$

COO^- OH

OH $O\text{-}SO_3^-$ $NH\text{-}SO_3^-$

$IdoUA,2S$ $GlcNSO_3^-,6S$
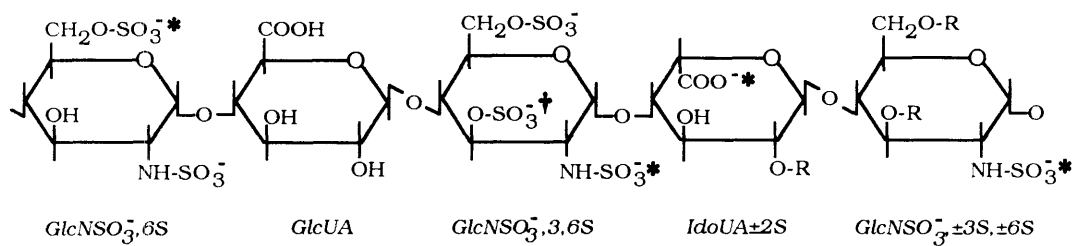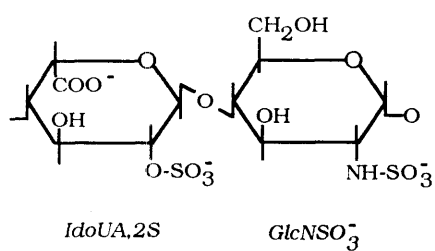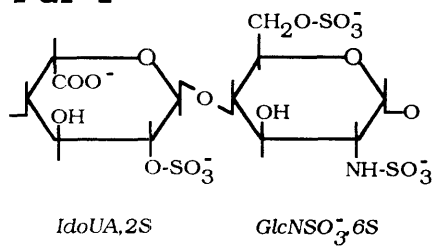
Table 2. Proteins to which Glycosaminoglycans Bind

Extracellular Matrix Molecules

Laminin	Fibronectin	Thrombospondins
tenascin	many Collagens	Vitronectin

Cell Adhesion Molecules

NCAM	L1/NgCAM	PECAM-1
Myelin-associated glycoprotein		

Cytokines

FGFs 1 thru 7	PDGF	VEGF
HB-EGF	TGF-β	IL3
Neural Schwann Cell Mitogen	chemokines (e.g., Il-8, MIP-1β)	

Proteases and Anti-proteases

Antithrombin III	Protease Nexin-1	Heparin Cofactor II
Amyloid β-protein precursor	Urokinase plasminogen activator	Tissue plasminogen activator
Thrombin		

CHAPTER 2

IDENTIFICATION, MOLECULAR CLONING AND PARTIAL GENOMIC CHARACTERIZATION OF CEREBROGLYCAN, A NOVEL GLYPICAN-RELATED HEPARAN SULFATE PROTEOGLYCAN

INTRODUCTION

The development of the nervous system is believed to be guided by an abundant and diverse set of extracellular cues. Growth factors such as FGF-1,2 and 5 can stimulate neural precursor cells to proliferate or differentiate (Walicke et al., 1986; Unsicker et al., 1987; Walicke, 1988; Murphy et al., 1990; Hughes et al., 1993). Migrating neurons and growing axons receive guidance information from large extracellular matrix glycoproteins such as laminin (Gundersen, 1987; Sanes, 1989; Reichardt and Tomaselli, 1991), or fibronectin (Rogers et al., 1989), diffusible factors such as the netrins (Kennedy et al., 1994; Serafini et al., 1994), and adhesion molecules such as NCAM (Neugebauer et al., 1988) and L1 (Lindner et al., 1983; Chuong et al., 1987) on the surfaces of other cells. Even after axons reach their targets, proper synapse formation is still dependent upon extracellular cues from molecules such as S-laminin (laminin-3) (Noakes et al., 1995; Porter et al., 1995) and agrin (Bowe and Fallon, 1995). As divergent as these cues are, many share a common feature: the ability to bind the glycosaminoglycan (GAG) heparan sulfate (HeS), of heparan sulfate proteoglycans (HSPGs).

In most cases, alternative, non-HSPG receptors have been identified for the ligands described above, leaving open the question of what the functional consequence of binding to HeS might be. It has been proposed that cell-surface HSPGs act as "co-receptors" for their ligands, functioning to bring them together with more conventional molecules (e.g. signal-transducing receptors) on the cell surface (Bernfield et al., 1992). Strong evidence for this model comes from studies of the interaction of basic fibroblast growth factor with one of its tyrosine kinase receptors, an interaction that fails to occur in the absence of HeS (Rapraeger et al., 1991; Yayon et al., 1991; Kan et al., 1993). Additional evidence has also come from studies of cell-surface protease inhibition by heparin-binding anti-proteases (Marcum et al., 1987; Cunningham et al., 1992), homophilic cell adhesion mediated by the neural cell adhesion molecule NCAM (Cole et al., 1986a; Reyes et al., 1990), and integrin-mediated cell-matrix interaction (LeBaron et al., 1988; Woods and Couchman, 1992).

The co-receptor model provides a useful framework within which to understand the general functions of HSPGs, but does not address the issue of whether specific functions are associated with individual molecular species of HSPG. Prior to the work described below, five distinct polypeptides had been identified as major carriers of HeS on mammalian cells. Four of these are transmembrane proteins, and form a gene family, the syndecans, on the basis of strong evolutionary conservation in their cytoplasmic and transmembrane domains (Bernfield et al., 1992). The fifth is a glycosylphosphatidylinositol (GPI)-anchored protein known as glypican (David et al., 1990). The expression of these PG core proteins varies dramatically from tissue to tissue, especially during the course of development (Bernfield et al., 1992). Tissue-specific and developmental variation also occurs in the structure of the HeS that is found on HSPG core proteins, and such variation can give rise to substantial differences in the binding properties of intact PGs (Marcum et al., 1987; Sanderson and Bernfield, 1988; Salmivirta et al., 1991; Sanderson et al., 1992a; Nurcombe et al., 1993). Taken together, the data suggest that individual HSPGs may indeed be specialized to carry out particular functions.

Consistent with the hypothesis that distinct HSPG cores have distinct functions in nervous system development, earlier work has shown that as many as nine individual HSPG core proteins are expressed in the developing rat central nervous system (Herndon and Lander, 1990). Furthermore, the expression of many of these HSPGs is developmentally regulated.

Only three of these HSPGs, previously identified as ("M7"), ("M12"), and ("M13"), exhibit detergent partitioning properties indicative of integral membrane proteins (Herndon and Lander, 1990). These three molecules, are thus likely to represent the major cell surface HSPGs of the central nervous system. To begin to assess whether these HSPGs participate in distinct developmental events in the nervous system, we developed a strategy to establish their molecular identities by protein purification and sequencing. We report here on studies that identify HSPG M12 as the rat homologue of glypican, and M13 as a novel, glypican-related HSPG that we have named

cerebroglycan (CBG). The cloning of murine and human CBG cDNA fragments, the partial characterization of mouse CBG genomic structure, and the mapping of CBG in the mouse genome are also described.

Glypican, cerebroglycan, the previously identified molecule OCI-5 (Filmus et al., 1988), and the subsequently indentified HSPGs, K-glypican (Watanabe et al., 1995), and dally, a Drosophila protein (Nakato et al., 1995), define a new family of GPI-anchored cell surface HSPGs. The core proteins in this family exhibit an unexpectedly high degree of structural homology, raising the possibility that these molecules have evolved for more specialized functions than simply to be the passive bearers of GAG chains.

MATERIAL AND METHODS

Materials

Heparitinase was purified from Flavobacterium heparinum by hydroxyapatite and phosphocellulose chromatography (Linker and Hovingh, 1972). The GAPDH cDNA clone was the generous gift of Dr. Timothy Hayes, NIH, Bethesda MD. All restriction enzymes were from New England Biolabs (Beverly, MA). Chondroitinase ABC, TritonX-100, NEM (N-ethylmaleimide), pepstatin A, and PMSF (phenylmethylsulfonyl fluoride), and PVP-40 (polyvinyl pyrrolidone) were obtained from Sigma (St. Louis, MO). Trypsin, CHAPS (3-[(3 Cholamidopropyl) dimethyl-ammonio]-1-propanesulfonate), and yeast RNA were from Boehringer Mannheim (Indianapolis, IN). RNA and DNA ladders, and M-MLV reverse transcriptase were from Gibco BRL (Gaithersburg, MD). RNasin was from Promega. (Nitrocellulose was from Schleicher and Schuell (Keene, NH). DEAE-Sephacel, salmon sperm DNA, and Taq DNA polymerase were from Pharmacia (Piscataway, NJ). Formamide was from Fluka (Buchs, Switzerland). α^{32}P-labelled dCTP and α^{35}S-labelled UTP were from DuPont NEN (Boston, MA). Week 16 human fetal brain was obtained from Dr. Gene Major, NINDS, NIH. Custom primers for PCR and sequencing were synthesized at the MIT Biopolymers Lab. All other reagents were from Mallinckrodt (Paris, KY), except as noted.

Purification and Microsequencing of P0 rat brain HSPGs

Detergent extracts of a crude membrane fraction of neonatal (P0) Sprague-Dawley rat brains were prepared essentially as described previously (Herndon and Lander, 1990). Briefly, whole brains were removed P0 rats and meninges were removed under ice cold saline. Brains were resuspended in 9 volumes of ice cold buffer A (0.3 M sucrose, 4 mM HEPES [pH 7.5], 1mM EDTA, 1μg/ml pepstatin A, 0.25 mg/ml NEM, and 0.4 mM PMSF), and homogenized using a teflon on glass homogenizer. Homogenates were centrifuged at 12,.000 g for 30 minutes at 4°C. The pellet was rehomogenized in buffer A, and centrifuged again. The supernatants from the two 12,000 g spins were pooled and centrifuged at 378,000 g for 30 minutes at 4°C. The membrane containing pellet was recovered and homogenized in buffer

B (50 mM Tris-HCl [pH 8], 0.15 M NaCl, 1% CHAPS, 1 mM EDTA, and 1 µg/ml pepstatin A). After centrifugation at 378,000 g for 60 minutes at 4°C, the supernatant was recovered, and the pellet was rehomongenized and recentrifuged for 40 minutes at 432,000g, 4°C. This supernatant was pooled with the previous buffer B supernatant and stored at -80°C.

Detergent extracts from a total of 34.35 g (wet weight) of brain were thawed, pooled and filtered through a 0.2 µ filter. Total PGs were isolated by anion exchange chromatography on DEAE-Sephacel as described previously (Herndon and Lander, 1990). The detergent extract was loaded at 21 cm/hour to a 254 ml DEAE-Sephacel column pre-equilibrated in buffer C (50 mM Tris-HCl [pH 8], 0.15 M NaCl, 0.1% TritonX-100). After loading was complete, the column was eluted with 4 column volumes of buffer C; 2 column volumes of buffer D (50 mM Tris-HCl [pH 8], 0.25 M NaCl, 0.1% TritonX-100); 2 column volumes of buffer E (50 mM Tris-HCl [pH 8], 0.25 M NaCl, 0.1% TritonX-100, 6 M urea); 5 column volumes of buffer F (50 mM sodium formate [pH 3.5], 0.25 M NaCl, 0.1% TritonX-100, 6 M urea); 2.5 column volumes of 100 mM Tris-HCl [pH 8], 0.1% TritonX-100; and 1 column volume of 50 mM Tris-HCl, 0.1% TritonX-100. PGs were eluted with 3 column volumes of buffer G (50 mM Tris-HCl [pH 8], 0.75 M NaCl, 0.1% TritonX-100, and collected in ~6 ml fractions. Fractions were analyzed by passing through an aliquot of each through a nitrocellulose filter pre-wet with 6% trichloroacetic acid, staining with amido black (0.25% in 45% methanol, 5% acetic acid), and destaining with 2% acetic acid in 90% methanol. On the basis of the staining, 27 fractions containing a total of 165 ml were pooled. At this point, PGs previously isolated from the membrane fraction of an additional 12.0 grams (wet weight) of P0 rat brains were added to the pool along with 0.3 µCi of ^{125}I-labelled, membrane associated PGs, purified and radiolabelled as previously described (Herndon and Lander, 1990). The addition of the radiolabelled PGs allowed subsequent purification steps to be followed.

Centriprep-10 tubes (Amicon; Beverly, MA) were used to concentrate the PG pool and exchange the elution buffer for 50 mM Tris-HCl (pH 8.0), 0.15 M NaCl, 0.1% Triton X-100. The sample was

then treated with Chondroitinase ABC (0.09 units/ml) in the presence of 1mM EDTA, 1μg/ml pepstatin, and 0.4 mM PMSF for 2 hr at 37°C, and loaded immediately onto a DEAE-Sephacel column (0.2 ml packed volume), equilibrated in buffer C. The column was washed successively with 4 ml 0.2 M NaCl, 100 mM sodium acetate (pH 3.5), 0.1% TritonX-100; 2 ml 0.1 M Tris-HCl (pH 8.0), 0.1% TritonX-100; 3 ml 50 mM Tris-HCl (pH 8.0), 0.1% TritonX-100; and 0.8 ml 25 mM ammonium acetate (pH 7.0), 0.1% TritonX-100. An HSPG-enriched fraction was eluted with a 20 ml gradient of NaCl from 0.15 M to 0.75 M, in 25 mM ammonium acetate (pH 7.0) and 0.1% TritonX-100. All significantly radioactive fractions eluted by ≥ 0.3 M NaCl were pooled. In this manner, approximately 100 μg of HSPG were recovered (as determined by amino acid analysis).

A Centricon-10 (Amicon) tube was used to concentrate the pooled sample and exchange the buffer to 50 mM NaCl, 25 mM ammonium acetate, 0.1% Triton X-100. The sample was made 25 mM in Tris-HCl (pH 7.1 at 37°C), and protease inhibitors (as with the chondroitnase digest above) added along with heparitinase to a final concentration of 9 μg/ml. The HSPG sample was digested for 3 h at 37°C, concentrated in a Centricon-10 tube and loaded to a 6% SDS-PAGE gel, polymerized over a strip of 9% acrylamide (see Figure 2). A control digest of identical composition but with no HSPGs was loaded in an adjacent lane of the gel. After electrophoresis to resolve HSPG M7, HSPGs M12 and M13, captured in the strip of 9% acrylamide, were excised (along with the enzyme blank) and recast in a second 9% acrylamide gel. After further electrophoresis to resolve M12 and M13, both the 6% gel (containing M7) and the 9% gel were electroblotted overnight at 25 V and 4°C to nitrocellulose in 25mM Tris-HCl (pH 8.3), 20% (v/v) ethanol, 0.005% SDS. Bands were visualized by staining with amido black (Schaffner and Weissman, 1973), and the bands corresponding to HSPGs M7, M12, and M13 were excised, as were corresponding regions from the enzyme blank lanes.

The nitrocellulose-bound HSPGs and their blanks were blocked with PVP-40 (0.5% in 100 mM acetic acid at 37°C for 30 min) and then digested with trypsin for 18 h at 37°C in 100 mM ammonium carbonate/acetonitrile (95:5 by volume) (Tempst et al., 1990).

Released digestion products were analyzed by reverse phase HPLC on a Hewlett Packard model 1090L HPLC equipped with a Vydac 218TP52 C18 column (Vydac, Hesperia, CA), and multiple peaks of absorbance at 210 nm were collected and subjected to microsequencing by automated Edman degradation.

Calculation of Yields

The inclusion of radioiodinated PGs in the purification allowed yields for some purification steps to be estimated. For the first concentration step (prior to chondroitinase digestion), the yield was calculated by (cpm, ret.)/[(cpm, tot.) - (cpm, filt.)], where "cpm, ret." was the cpm retained and recovered from the Centricon concentrator, "cpm, tot." was the total number of cpm in the pool prior to concentration, and "cpm, filt." was the cpm that passed through the Centricon filter. This latter value represents primarily unincorporated ^{125}I, and so needs to be subtracted from total cpm to get an accurate indication of protein yield.

In the second DEAE step (after chondroitinase treatment), CSPG core proteins either flowed through, or were removed with a stringent wash. Earlier experiments showed that the PGs remaining bound to DEAE were essentially depleted of CSPG cores (not shown). To estimate the yield of HSPGs eluted from the second DEAE column, the following expression was used:

yield = (cpm, gradient)/[(cpm, tot.)-(cpm, f.t.)-(cpm, s.w.)];
"cpm, gradient" is the total number of counts eluted by the salt gradient (primarily HSPGs), "cpm, tot." is the total number of counts applied to the column, "cpm, f.t." is the number of counts that flowed through the column, and "cpm, s.w." is the number of counts that were eluted by the stringent wash.

The yield for the second concentration step, after the second DEAE column was estimated simply by (cpm, recovered)/(cpm, applied) since there was little if any unincorporated ^{125}I remaining by this step.

PCR Protocols
Amplification of an HSPG-M13 cDNA

Degenerate PCR primers were based on the sequence of peptide M13-22 (see Results and Table 1). The sequences of the primers were 22L: 5'-GGCCTCTAGA(T/C)ATGCA(T/C)GA(T/C)GCNGA-3' and 22R: 5'-GCGCGGGCCCCT(G/A)TCNGC(G/A)TA(C/T)TG(C/T)TG-3'. PCR reactions were performed using 25 pmol of each PCR primer, 2 µg of P0 rat brain cDNA as template, 80 µM of each dNTP, and 2 U of Taq polymerase. The reaction profile was denature at 95°C for 60 sec., anneal at 50°C for 90 sec., and extend for 90 sec. at 72°C. The resulting 80 bp product was gel purified and subcloned into pCR1000 (Invitrogen; San Diego, CA) for sequencing.

Amplification of a murine cerebroglycan homolog

Total RNA was isolated from 1.2 grams (wet weight) of embryonic day 17 mouse brain by guanidinium isothiocyanate extraction (Chomczynski and Sacchi, 1987). Reverse transcriptase reactions were set up in 1X Taq polymerase buffer with (1) 0.4-3.8 µg RNA (heated to 90°C for 5 minutes and placed on ice), (2) 100 pmol of random hexamers or 100 pmol of the primer OM13-165 (see below), based on the sequence of rat cerebroglycan, (3)1 mM dNTPs, (4) 40 units of RNasin, and (5) 200 units of M-MLV reverse transcriptase. Reactions were incubated at room temperature for 10 minutes and at 42°C for 60 minutes. Reactions were terminated by boiling for 8 minutes and stored at -20°C until PCR.

PCR reactions were performed by adding 80 µl of 1X Taq polymerase buffer containing 2.5 units of Taq polymerase, 50 pmol of primer OCBG-137, and 50 pmol or 20 pmol of primer OM13-165 (20 pmol was used for reactions were OM13-165 had been used in the reverse transcription step). The reaction profile was 95°C for 30 seconds, 48°C for 1 minute, and 72°C for 1 minute for a total of 30 cycles. The major ~400 bp band visible in all reactions after gel purification was subcloned directly into the pCR1000 vector (Invitrogen; San Diego, CA) for sequencing. The sequence of the primer OM13-165 was 5'-GAGAGACTCGAGTGAGGCGTCCCACATTCC-3'. The underlined residues are a XhoI site and clamp sequences. The remaining are complementary to nucleotides 1284-1265 of the CBG cDNA sense strand. The sequence of the primer OCBG-137 was

5'-CACTGGTGGCAGCCCGG-3'. This primer is identical to nucleotides 894-910 in the CBG cDNA sense strand.

Amplification of a human cerebroglycan homolog

Total RNA was isolated by guanidinium isothiocyanate extraction (Chomczynski and Sacchi, 1987) from week 16 human fetal brain. cDNA was synthesized in a reaction containing 13 μg of total RNA, 2.8 μg of primer 22R (see above), 1 mM dNTPs, 1 unit/μl of RNasin, and 20 units/μl of M-MLV reverse transcriptase for 10 minutes at room temperature and 60 minutes at 42°C. A nested set of degenerate primers was used. The outer set was gdoL1 (5'-CCICA[A/G]G[A/G]-ITA[T/C]ACITG[T/C]TG-3') and gdoR2 (5'-CC[A/G]CAI[T/C]C[T/C]TG-IA[T/A]IAC[T/C]TT-3'). The inner set was gdoL2 (5'-TT[T/C]TGG-GCI[C/A][A/G]I[T/C]TI[T/C]TIGA-3') and gdoR3 (5'-IGCIGCIACIA[G/A]-IGCIC[G/T]IG-3'). See Figure 10 for the relationship of these primers to sequences of cerebroglycan, glypican, and K-glypican.

The first round of PCR used 100 pmol of primers gdoL1 and gdoR2, 100μM of dNTPs, 2.5 units of Taq polymerase, and 0.4μg of cDNA, prepared as described above, in a 100 μl reaction. The reaction profile was 95°C for 1 minute, 53°C for 1 minute, and 72°C for 1.5 minutes for a total of 24 cycles. A diffuse band of 800-1000 bp was gel purified and used in a second round of reactions. The second round of PCR was performed essentially as above, yielding a products of ~180 bp and ~168 bp. The 168 bp product was gel purified and subcloned into pCR1000 (Invitrogen; San Diego, CA) for sequencing.

Alternatively, a second round of PCR using modified gdoL2 and gdoR3 primers was performed. These primers had restriction sites added to their 5' ends (extra sequence in italics, restriction sites underlined):
BamHIL2 (5'-*CG*GGATCCCGTT[T/C]TGGGCI[C/A][A/G]I[T/C]TI[T/C]-TIGA-3') and EcoRIR3 (5'-*G*GAATTCCIGCIGCIACIA[G/A]IGCIC[G/T]-IG3'). PCR with these primers was performed with a different profile: (1) 95°C for 1 minute, (2) 37°C for 1 minute, (3) slope= +0.7°C/3sec., (4) 72°C for 1 minute, (5) repeat (1)-(4) 2 more times, (6) 95°C for 1 minute, (7) 53°C for 1 minute, (8) 72°C for 45 seconds, (9) repeat (6)-(8) 19 times.

This reaction profile also resulted in two major products at ~168 and 180 bp. The 180 bp product was subcloned into the BamHI/EcoRI restriction sites in pBluescript (Stratagene) for sequencing.

Amplification of a mouse genomic fragment for SSCP analysis

PCR reactions were performed with 25 ng of genomic DNA in a 40 μl reaction that was 250 μM dNTPs, 0.05 units/μl Taq polymerase, and 25 nM for each primer. The primers were muCBG-L (5'-GGTCC-AGGGTCTGGAGAC-3'), and muCBG-R (5'-GCCTGTCTGCAGCCTTCC-3') (see Figure 9). The reaction profile was 1 minute at 94°C, 1 minute at 55°C, and 1 minute 10 seconds at 72°C, for a total of 35 cycles. For SSCP analysis, 10μCi of 3000 Ci/mmol α32P-dCTP was included in each reaction. To verify authenticity, the major product just above 615 bp was gel purified and subcloned into pCR1000 (Invitrogen; San Diego, CA) for sequencing.

Library Construction and Screening

P0 rat brain polyA-enriched RNA was prepared by fractionation on oligo dT cellulose using a Fast Track mRNA isolation kit (Invitrogen; San Diego, CA). cDNA was synthesized from 5 μg of polyA$^+$ RNA using an oligo dT primer with the ZAP cDNA synthesis kit (Stratagene; La Jolla CA). The resulting cDNA was ligated into EcoRI/XhoI digested and dephosporylated Uni-ZAP XR vector arms (Stratagene) and packaged for plating (Giga Pack Gold II, Stratagene). The library was amplified once before plating on XL1-Blue host cells. The library was screened with a probe generated by a PCR reaction (as described above), that included α^{32}P-dATP.

A primer-extended P0 rat brain cDNA library was synthesized from 5μg of polyA-enriched RNA (prepared as described above) and 120 pmol of primer OM13-165 (based on the sequence of cDNA M13-A, see Results, and above) using a Timesaver cDNA synthesis kit (Pharmacia). The resulting cDNA, ligated to EcoRI adaptors, was cloned into EcoRI-digested, dephosphorylated λgt10 arms (Stratagene), packaged (Giga Pack Gold II, Stratagene), and plated on NM514 host cells for screening with a probe generated from the 100

bp Bgl II fragment of M13-A (see Results) by random-primer labelling (70200 Random Primed DNA Labelling Kit, United States Biochemical; Cleveland, OH).

A mouse genomic library (129/Sv) (gift of the lab of Susumu Tonegawa, MIT) was plated on LE392 host cells for screening with a probe generated by random-primer labelling (70200 Random Primed DNA Labelling Kit, United States Biochemical; Cleveland, OH) of the 215 bp Aat II/Sma I restriction fragment of the rat CBG cDNA. This fragment begins 95 bp upstream from the start ATG. Two CBG genomic clones were identified, plaque purified, and subcloned into pBluescript (Stratagene) for sequencing and restriction mapping. The library was rescreened with a probe made by random-primer labelling of the 818 bp Aha II/AflI restriction fragment of the rat CBG cDNA. Five overlapping clones were isolated, and subcloned for further characterization.

DNA Sequencing

Gel purified fragments were subcloned into pBluescript SK⁻ (Stratagene), pSL301 (Invitrogen), or pCR1000 (Invitrogen). cDNA inserts were sequenced by the dideoxy chain termination method using a modified T7 DNA polymerase (Sequenase kit 2.0, United States Biochemical). T3, T7, and M13 (-40) primers, as well as custom made primers, were used in reactions with denatured, supercoiled plasmids. All rat cerebroglycan cDNA sequences were determined with both dGTP and dITP; some of the remaining sequencing reactions were performed with dGTP only.

Southern Blot Analysis

Rat genomic DNA was prepared from Sprague-Dawley rat thymocytes according to established methods (Sambrook et al., 1989), and stored at 4°C under ethanol. 10μg of DNA, digested to completion with HindIII, KpnI, or XhoI, was electrophoresed and blotted to nitrocellulose (Sambrook et al., 1989). Prehybridization, hybridization, and washing, were done as previously described (Church and Gilbert, 1984). The hybridization step and the subsequent washes

were performed at 68°C (high stringency). The filter was autoradiographed with XAR-5 film (Kodak; Rochester, NY).

Northern Blot Analysis

Total RNA was isolated by guanidinium isothiocyanate extraction (Chomczynski and Sacchi, 1987). 20µg of total RNA from embryonic day 18 (E18), P0, and adult rat brain was used for the Northern analysis, performed according to Sambrook et al. (Sambrook et al., 1989), except that yeast RNA (200 µg/ml) and salmon sperm DNA (100 µg/ml) were included in the hybridization buffer. Filters were exposed to Kodak XAR-5 film or imaged with a phosphorimager (Molecular Dynamics; Sunnyvale, CA). The blot was re-probed for glyceraldehyde-phosphate dehydrogenase (GAPDH) mRNA to normalize for different loading and transfer efficiencies between the lanes.

Analysis of single stranded conformational polymorphisms (SSCP)

PCR reactions were performed as described above with genomic DNA samples from the test cross NCI-Frederick #77-122 (B6 X spretus) X B6. B6 refers to C57B16 (Mus musculus) and spretus refers to Mus spretus. 9 volumes of 95% formamide, 10 mM NaOH, 0.02% bromphenol blue and 0.02% xylene cyanole was added to an aliquot of each reaction. After heating at 95°C for 5 minutes, samples were placed on ice, spun briefly and loaded to a 0.4 mm thick, 20 cm long gel of 0.5X MDE acrylamide gel (AT Biochem Inc., Malvern, PA). Gels were run at 220 volts for 18 h with circulating water to cool. After electrophoresis, gels were dried down and autoradiographed or exposed in a phosphorimaging cassette (Molecular Dynamics). The SSCP was scored for each sample as heterozygous (the B6 X Spretus pattern) or homozygous (the B6 pattern). The tabulated results were analyzed by Dr. Mark Daly at the Whitehead Institute for Biomedical Research, MIT, using the program Mapmaker.

RESULTS

Preparation of a partially purified HSPG fraction
from neonatal rat brain

As a first step toward establishing the molecular identity of the HSPGs M7, M12, and M13, a neonatal rat brain fraction enriched for HSPGs and depleted of CSPGs was prepared. PGs were prepared by DEAE-chromatography of a detergent extract of neonatal rat brain membranes as previously described (Herndon and Lander, 1990). A trace quantity (0.3 μCi) of previously isolated, radioiodinated neonatal membrane PGs was added to the PG pool to allow the subsequent purification steps to be followed.

CSPGs were removed by treatment with chondroitinase ABC and re-fractionation on DEAE-Sephacel. Figure 1 shows the elution profile for the second DEAE purification. After chondroitinase digestion, a substantial fraction of the CSPG core proteins flow through a DEAE column when applied in 0.15 M NaCl (peak I in Figure 1). In preliminary experiments, SDS-PAGE analysis of the material in peak I demonstrated that it is comprised primarily of the CSPG core proteins identified previously (Herndon and Lander, 1990) as M1, M6, and M8 (not shown). A small but significant amount of the M1 protein core does bind DEAE in 0.15 M NaCl even after chondroitinase digestion. This material can be largely removed by a stringent wash (0.2 M NaCl, 0.1 M sodium acetate pH 3.5)(peak II in Figure 1). After pH was restored, HSPGs were eluted with a salt gradient (peak III). The bulk of the HSPG material eluted at ~0.45 M NaCl.

To obtain an indication of the amount of material in the HSPG pool, an aliquot was subjected to amino acid analysis. The results of the analysis indicated that approximately 100 μg of total protein had been recovered from 46.4 g (wet weight) of brain.

Electrophoretic separation and protein sequencing of neonatal rat
brain HSPG core proteins.

To obtain amino acid sequence from individual HSPG cores, the HSPG pool was digested with heparitinase, concentrated, and separated on an SDS-PAGE minigel to resolve the core proteins. The

proteins were electroblotted to nitrocellulose, visualized by amido black staining, and excised. After digestion in situ with trypsin, liberated peptides were separated by HPLC and subjected to Edman degradative sequencing.

One obstacle, identified in preliminary experiments, was the incompatibility of M7 with M12 and M13 in electrophoresis and electroblotting. Control experiments demonstrated that an acrylamide gel of at least 9% total acrylamide was required to adequately resolve M12 from M13. However, M7 did not transfer efficiently from gels higher than 7% acrylamide. To overcome this obstacle, a two stage electrophoresis step was used. As Figure 2 illustrates, M7 was separated on a 6% acrylamide gel polymerized over a strip of 9% acrylamide. M12 and M13, captured in the strip, were excised and recast into a second acrylamide gel of 9% total acrylamide, where they were resolved by further electroporesis.

A second obstacle stemmed from the requisite use of heparitinase in the preparation of the HSPG cores for electrophoresis. To minimize the possibility of heparitinase contamination in the final sequencing step, an enzyme blank was processed in parallel to the HSPGs. As an example, Figure 3 shows the HPLC chromatograms of HSPG M12 and its enyzme blank. Peaks present in the blanks were disqualified from further analysis.

As shown in Table 1, sequence data were obtained for several M12 and M13 peptides. No peptides were obtained for HSPG M7 (see Discussion). Peptides derived from HSPG M12 were 74-100% identical to regions of the GPI-linked HSPG glypican (not shown), cloned from human lung fibroblasts (David et al., 1990). This suggested that HSPG M12 is the rat homologue of glypican and subsequent cloning of rat glypican cDNA verified this [see (Karthikeyan et al., 1994; Litwack et al., 1994)]. Peptides derived from HSPG M13 could also be aligned with human glypican, but had an overall identity of only 58%, suggesting that HSPG M13 might be a novel glypican-related HSPG. Because of its unique pattern of expression in the developing rat nervous system (see Chapter 3, this thesis), HSPG M13 was subsequently named cerebroglycan (CBG), and will be referred to as CBG hereafter.

cDNA cloning and sequencing of cerebroglycan

To facilitate further characterization of CBG, a cDNA was obtained from neonatal rat brain by PCR. Using primers 22L and 22R, which were based on unambiguous sequence at either end of peptide M13-22 (Table 1), a major PCR product of ~80 bp was obtained. This product was subcloned and two independent clones were isolated and sequenced. Both contained the sequences of primers 22L and 22R, and the intervening sequences of both gave a deduced protein sequence that was an exact match for peptide M13-22 (data not shown), verifying that a genuine CBG cDNA had been amplified.

To obtain a complete cDNA, a neonatal rat brain cDNA library was screened with a probe generated by PCR using α^{32}P-labelled dATP and primers 22L and 22R, with the gel-purified 80 bp CBG cDNA fragment as the template. Three positive plaques were identified out of 900,000 screened in the amplified library. All three were plaque purified, and all contained inserts of ~1600 bp. Preliminary sequence data supported the conclusion that all three inserts were identical products of a single cloning event. Additional sequencing of this cDNA, referred to hereafter as M13-A, showed that it contained an 862 bp open reading frame followed by 659 bp of 3' untranslated sequence, and was unlikely to encode full length CBG. Therefore, a primer extended library was constructed using neonatal rat brain RNA and primer OM13-165 based on sequence 165 bp from the 5' end of M13-A. The library was screened with a random-primer labelled probe generated from the 100 bp Bgl II fragment of M13-A (see Figure 4), and 52 positives were identified out of 9 X 10^4 plaques. Two of these, M13PX-1 and M13PX-2, were plaque purified and found to contain inserts of ~1200 and ~1100 bp respectively. The combined length of M13-A and the overlapping M13PX-1 is an excellent match for the size of the CBG mRNA observed in Northern analyses (see below), indicating that together, M13-A and M13PX-1 constitute an essentially full length CBG cDNA. Figure 4 illustrates how the complete CBG cDNA sequence was obtained from M13-A, M13PX-2, and M13PX-1 by a combination of subcloning and the use of custom sequencing primers.

Sequence Analysis

The combined sequences of M13-A, M13PX-2, and M13PX-1 comprise a cDNA of 2631 bp, shown in Figure 5. The first ATG codon (at position 70), occurs in a poor context for initiation of translation (Kozak, 1991), and the open reading frame that follows terminates after 240 bp. The second ATG codon (at position 236) occurs in a favorable context, with a G in the -3 position. Beginning with this codon, an open reading frame extends 1740 bases before terminating at position 1975. An AATAAA cleavage and polyadenylation signal is found at position 2582, followed 13 bases later by a poly-A tract. The deduced amino acid sequence of the long open reading frame correctly predicts the sequences of all four CBG tryptic peptides (underlined in Figure 5), including the three that were not used in screening the cDNA library.

The deduced amino acid sequence features a stretch of hydrophobic amino acids at both the amino and carboxy termini, that could serve as a signal sequence and a GPI-attachment signal respectively. CBG was previously shown to be GPI-linked on the basis of its ability to be released from Triton X-114 micelles by phosphatidylinositol specific-phospholipase C (Herndon and Lander, 1990). With the removal of these signal sequences, the mature protein would have a predicted molecular weight of 58.6 kDa, in good agreement with the size of the CBG core protein, as observed on reducing SDS-PAGE gels [57 kDa; (Herndon and Lander, 1990)]. The protein contains seven "ser-gly" sequences, five of which (boxed in Figure 5) are potential GAG attachment sites (see Discussion). No "N-X-S" or "N-X-T" N-linked glycosylation consensus sequences are present.

As Figure 6 (panel A) illustrates, the CBG protein sequence shows significant homology to four other molecules, rat glypican (Karthikeyan et al., 1994; Litwack et al., 1994), the deduced protein sequence of OCI-5, a developmentally regulated transcript cloned from a primitive intestinal epithelial rat cell line (Filmus et al., 1988), K-glypican, cloned from mouse kidney cDNA (Watanabe et al., 1995), and the Drosophila protein, dally (Nakato et al., 1995). The homology is

moderate but extensive, with islands of highly conserved sequence surrounded by divergent sequences. Most striking is the conservation of cysteine residues (black and underlined in Figure 6A): Of the 14 cysteines in the predicted sequence of the mature CBG protein, all are conserved in the other molecules. Glypican, K-glypican, and OCI-5 have and additional pair of cysteines near the carboxy terminus that is absent from the other molecules. In addition, all of the proteins have at least one putative ser-gly GAG attachment site (Figure 6A, on the left hand side of the second set from the bottom), in a cluster near the carboxy terminus. Figure 6, panel B shows the percent identity of each of the five molecules to the other four. Among the mammalian proteins CBG, glypican, and K-glypican are more closely related to each other than to OCI-5. The dally protein has a similar level of identity to each of the mammalian proteins.

Southern hybridization

To assess whether CBG is present as a single copy in the genome, 10 μg of rat thymus DNA was digested to completion with HindIII, KpnI, or XhoI, separated by agarose gel electrophoresis, and transferred to nitrocellulose. The filter was probed with a random primer-labelled probe made from the 100 bp BglII frament of the M13-A cDNA. The resulting autoradiogram is shown in Figure 7. In each lane a single band appears indicating that the gene encoding CBG is present as a single copy in the rat genome. The high level of stringency at which the hybridization was performed would not have detected the presence of CBG-related proteins such as glypican, OCI-5, K-glypican, or other as yet unidentified molecules.

Northern hybridization

The CBG core protein is observed in the membrane fraction of embryonic day 18 (E18) and newborn (P0) rat brain, but not in the adult (Herndon and Lander, 1990). To examine the expression of CBG mRNA, Northern blot analysis was performed using 20 μg of rat brain total RNA from each of these three developmental stages, and a random-primer labelled probe generated using cDNA M13-A as the template (Figure 8). A single transcript of about 2.7 kb was seen that

is relatively abundant at E18, less so at PO, and undetectable in the adult sample. The filter was re-probed with a glyceraldehyde-phosphate dehydrogenase (GAPDH) probe (Figure 8, lower), demonstrating that there was not an appreciable difference in the amount of RNA loaded in each lane. Thus, the developmental expression of CBG mRNA agrees with that previously observed for CBG protein.

Isolation of homologous CBG cDNAs from mouse and human

To assess the evolutionary conservation of CBG cDNA and protein sequence, and to generate probes for future experiments, murine and human homologues of CBG were isolated by PCR. To obtain a murine cDNA, primers 0M13-165 and OCBG-137 (see Materials and Methods) were used in an RT-PCR reaction with embryonic day 17 mouse brain total RNA. A major product of the predicted 400 bp size was obtained and subcloned. Two independent subclones muCBG-1 and muCBG-2 were isolated, and preliminary sequencing indicated that they were identical. The complete sequence of muCBG-1 was obtained by a second round of sequencing; however, the complete sequence of *both strands* has not been determined. Comparison of muCBG-1 with rat CBG (Figure 9) revealed 94% identity at the nucleic acid level, and 99% identity at the protein level. The one substitution that was observed was conservative -- valine for alanine.

A cDNA likel;y to be human CBG was obtained using a nested set of degenerate primers, based on conserved sequences in rat glypican, mouse K-glypican, and rat CBG (see Figure 10, top). The outer set of primers was used in a PCR reaction with week 16 human human fetal brain cDNA. To enrich for CBG-like sequences, the cDNA was sythesized from total RNA using the degenerate primer 22R, based on rat CBG protein sequence (Table 1). The products of this round of PCR, separated on an agarose gel and stained with ethidium bromide, contained, in addition to several non-specific products, a diffuse band just below 1 kb, close to the predicted size of 954 bp. This material was purified and used in a second round of PCR with the inner set of primers. Two major products near the predicted size of 180 bp were obtained. Subcloning and sequencing revealed that the larger of these

two products is likely to be a human CBG cDNA fragment. The second product, about 12 bp smaller, is likely to be a human K-glypican cDNA fragment.

Figure 10 (bottom) shows the alignment of these cDNAs with rat CBG and mouse K-glypican respectively. The putative human CBG fragment is 81% identical with rat sequence at the nucleic acid level with 87% identy at the protein level. The putative human K-glypican fragment is 81% identical to mouse K-glypican nucleic acid, and 78% identical at the protein level.

CBG genomic structure and mapping

To begin to characterize CBG genomic structure a mouse genomic library was screened with a probe based on the rat CBG cDNA sequence. To bias toward isolating clones containing sequence from the first translated exon, a 215 bp Aat II/Sma I restriction fragment, beginning 94 bp ustream of the start ATG, was used as a template to generate α ^{32}P-dCTP labelled probe. Hybridization of the probe to the mouse genomic library identified two positive plaques out of a total of 760,000 screened. Overlapping clones were isolated by re-screening the library with the 818 bp Aha II/Afl I restriction fragment of the rat CBG cDNA. Figure 11 shows a partial restriction map of the CBG genomic locus.

Sequencing of the CBG genomic clones has revealed the intron/exon boundaries of the first two translated exons. Additionally, two CBG cDNA clones containing a partially unspliced cDNA were fortuitously isolated. Sequencing of these clones (Figure 12) revealed the locations of additional intron/exon boundaries. The locations of all intron/exon boundaries identified so far are summarized in Figure 13. In Figure 13, the assumption is made that mouse and rat CBG have the same genomic structure, an assumption that has been validated for two intron/exon boundaries so far (see below).

To map CBG in the mouse genome, it was necessary to find a polymorphism linked to the CBG locus. This was accomplished by identifying a single stranded conformational polymorphism (SSCP) in the CBG loci of C57B16 (Mus musculus) and Mus spretus mice. To identify the SSCP, two primers (muCBG-L and muCBG-R) were

designed based on the sequence of the muCBG-1 PCR fragment (see Figure 9). These primers were chosen to flank the site of an intron identified in the same region of two partially unspliced rat CBG cDNAs (Figure 12). Because intron sequences can be highly divergent, they are the best candidates when searching for a SSCP. PCR with the primers muCBG-L and muCBG-R and Mus spretus or Mus musculus genomic DNA yielded a major product of about 615 bp, very close to the 613 bp predicted by the size of the intron in the rat clones. Subcloning and partial sequencing of the Mus musculus product demonstrated that it contained the sequences of the primers, and coding sequence identical to muCBG-1, interrupted by sequence highly similar to the sequence of the rat intron (not shown). This demonstrated that an authentic mouse genomic CBG fragment had been amplified.

To test the genomic fragment for a SSCP, α^{32}P-dCTP was included in PCR reactions to generate labelled genomic CBG fragments from Mus musculus and Mus spretus genomic DNA. The gel purified fragments were re-run on a strand separating gel (see Materials and Methods), revealing the presence of a SSCP (Figure 14). This SSCP was then scored by PCR with genomic DNAs from the test cross [NCI-Frederick #77-122 (Mus musculus X Mus spretus) X Mus musculus]. The results, summarized in Table 2, indicated that the CBG locus in the mouse genome maps to the distal end of the long arm of chromosome 5 (distal to Epo, a marker ~6 cM from the telomere). No previously identified mouse mutants (e.g. reeler) correspond to this location in the mouse genome (see Discussion).

DISCUSSION

Purification of the major HSPGs of neonatal rat brain.

Amino acid analysis of the HSPG pool obtained after chondroitinase treatment and re-purification on DEAE-Sephacel indicated that 100 μg of protein had been isolated.

The table at right gives the calculated or predicted yields for the steps prior to amino acid analysis. The yield for the first DEAE step is predicted to be ~40% based on the results of add back experiments in an earlier

PurificationStep	Estimated Yield
1st DEAE column	40%
1st concentration	89%
2nd DEAE column	54%
2nd concentration	63%
total yield	12%

study (Herndon and Lander, 1990) in which radioiodinated PGs were added back to a crude homogenate and followed through a subsequent DEAE purification step. Calculated yields for the second DEAE step and the concentration steps were made possible by the inclusion of trace amounts of radioiodinated membrane PGs in the present purification. Although the true yield of the first DEAE step is unknown, the apparently slightly higher yield in the second DEAE purification step may be a result of gradient elution rather than step elution. An earlier study has shown that as much as 9% of the recoverable PG material is lost in the "tail" of a single step elution (Herndon and Lander, 1990). Acceleration of the tail's elution by a gradient would be expected to improve yield. Of the apparent 46% of the HSPGs not recovered in the second DEAE step, a substantial fraction was probably lost due to irreversible binding to DEAE. Previous studies showed that as much as 44% of bound PGs may remain associated with DEAE, even after elution with high salt (Herndon and Lander, 1990).

The total yield implies that roughly 0.83 mg of HSPG core protein was present in the crude membrane starting material derived from 46.35 grams (wet weight) of neonatal rat brain. This results in a value of 0.02 mg HSPG/gram of neonatal rat brain (or 2.7 mg HSPG per gram of brain membrane protein). This value is somewhat higher than previously reported values (Herndon and Lander, 1990), which may be a reflection of the fact that these earlier estimates were based on

binding of the anionic dye amido black, which binds less as efficiently to anionic PGs than to the protein standard (bovine serum albumin) used in the assay. In the present study, the amount of HSPG was measured by amino acid analysis, a more reliable method. If CBG accounts for ~20% of the total HSPG at this age, it would constitute ~0.05% of the total brain membrane protein.

Preliminary experiments with radioiodinated neonatal membrane PGs indicated that the material in the HSPG pool eluted from the second DEAE column is virtually depleted of CSPG core proteins. The amount of labelled PG that flowed through or was eluted by the stringent wash in the second DEAE step, taken as a fraction of the labelled PG applied, indicates that CSPGs account for ~50% of the material in the neonatal membrane PG fraction. Assuming HSPG cores and CSPG cores are radioiodinated to a similar specific activity, this implies that as much as 0.04 mg total PG/gram brain (wet weight) may be present in crude membrane fraction of neonatal rat brain.

Molecular identities of HSPGs M7, M12, and M13

The protein sequence of peptides derived from HSPG M12 suggested that M12 is the rat homolog of glypican. The sequences of all four M12-derived peptides have subsequently been identified within predicted amino acid sequence of two independently isolated rat glypican cDNAs (Karthikeyan et al., 1994; Litwack et al., 1994). Rat brain M12, like human glypican cloned from lung fibroblasts, is a glycosylphosphatidylinositol (GPI)-anchored HSPG with a core protein of M_r 65 kDa (David et al., 1990; Herndon and Lander, 1990). Antibodies raised against a rat glypican fusion protein specifically precipitate M12 from a crude mixture of brain PGs and recognize a band that comigrates with the M12 core protein in Western blots (Litwack et al., 1994).

No microsequence was obtained for HSPG M7. Possible reasons for the failure to obtain sequence include poor transfer of M7 to nitrocellulose after electrophoresis, and poor recovery of M7 tryptic peptides after in situ digestion. Autoradiography of the minigel from which M7 was transferred revealed that a significant portion of the M7 band remained in the gel. M7 aggregates with other species in PG

fractions as demonstrated by gel filtration (Herndon and Lander, 1990). This behavior may have caused M7 to partially precipitate in the gel and become resistant to electroblotting. The M7 core protein size and pattern of developmental expression in the rat brain are indistiguishable from those of the HSPG syndecan-3 (Carey et al., 1992; Gould et al., 1992), and an antibody to syndecan-3 has recently been shown to immunoprecipitate M7 from growth cone particle preparations (Ivins et al., 1995) The predicted protein sequence of syndecan-3 contains an extensive serine-threonine-proline rich region similar to the highly glycosylated regions of several mucin-like proteins (Gould et al., 1992). If M7 -- which is known to be highly sialylated (Herndon and Lander, 1990)-- is rat syndecan-3, the presence of this domain may have rendered it protease-resistant, decreasing the yield of tryptic peptides.

The sequences of the four M13-derived peptides suggested that M13 is a novel glypican related HSPG. To further characterize M13, an M13 cDNA clone was isolated from a newborn rat brain cDNA library on the basis of the sequence of one of the M13 peptides. Several lines of evidence confirm that the cDNA obtained in this study encodes the GPI-anchored HSPG core protein previously designated as M13 (Herndon and Lander, 1990): The cDNA encodes all four of the peptides that were obtained from the tryptic digest of M13; it encodes a protein of the correct size, that contains likely GPI-anchorage and GAG-attachment sequences (see below); and it hybridizes to an mRNA that is present in embryonic and neonatal rat brain, but not in adult brain [matching the pattern previously observed for M13 protein (Herndon and Lander, 1990)]. As described in Chapter 3 of this thesis, expression of HSPG M13 is limited to the nervous system. This result is surprising because no other proteoglycan is known to be restricted in expression to a single tissue. To reflect the nervous system-specific expression of this proteoglycan, we have chosen to name PG M13 "cerebroglycan."

Cerebroglycan Sequence Analysis

The deduced amino acid sequence of cerebroglycan (CBG) contains seven serine-glycine pairs, representing possible glycosaminoglycan

(GAG) attachment sites. Of these, ser[155], ser[498] and ser[500] all match the sequence SGXG, and all occur within regions of protein sequence that are rich in acidic residues. Both features have been proposed to define a consensus for GAG attachment (Bourdon et al., 1987; Bernfield et al., 1992). In addition, two other serine-glycine pairs (ser[55] and ser[92]) occur either just before or just after an acidic residue, and might also represent GAG attachment sites (Bernfield et al., 1992). Interestingly, ser[498] and ser[500] were sequenceable residues of peptide M13-22. Since substitution with GAG would be expected to result in a blank round in Edman degradative sequencing, it would appear the majority of the CBG population is substituted at only one or perhaps neither of these sites. Amino acid yields of these serines in the Edman degradation reaction were about 30% of the yields obtained for the neighboring glycines. This is the typical yield ratio observed for a labile amino acid such as serine compared to a non-labile one such as glycine (Richard Cook, MIT Biopolymer Lab, personal communication). Thus amino acid sequencing provided no evidence of substitution at either of these sites. See Chapter Five of this thesis for further discussion of CBG GAG attachment.

The cerebroglycan protein sequence has a stretch of hydrophobic amino acids at both the amino and carboxy termini, as indicated in Figure 5. These sequences could serve as the signal sequence and a GPI-attachment signal respectively, being cleaved off in the mature protein. The precise cleavage sites are unknown, and those indicated in Figure 5 are predictions based upon the -1,-3 rule (von Heijne, 1990), for the signal sequence, and the ω, ω + 2 rule (Kodukula et al., 1993), for the GPI-attachment signal. Of the two potential GPI-attachment sites indicated, ser[557] has a slightly higher probability than asn[553], based on amino acid preferences at the ω and ω+2 sites of other GPI-anchored proteins. Following the candidate GPI-attachment sites are a potential "hinge" region (rich in proline and charged amino acids), and a stretch of hydrophobic amino acids; both the hinge and hydrophobic stretch are features of GPI attachment signals (Kodukula et al., 1993).

Cerebroglycan Belongs to a Family of GPI-Anchored PGs

The deduced amino acid sequence of cerebroglycan shares significant homology with four other molecules: (1) the HSPG glypican (David et al., 1990); (2) OCI-5, originally identified as a developmentally regulated transcript cloned from a rat intestinal epithelial cell line, and recently shown to be expressed as an HSPG (Filmus et al., 1988; Filmus et al., 1995); (3) K-glypican, cloned from a mouse kidney cDNA library (Watanabe et al., 1995); and (4) the product of the Drosophila gene dally (Nakato et al., 1995) (see Figure 6). Thus it is now apparent that in addition to the syndecan family of transmembrane cell surface HSPGs, there is also an entire family of molecules related to glypican. All members of the glypican family are intercalated in the membrane by a GPI lipid anchor. Glypican family members also all have at least one putative GAG attachment site in a conserved cluster near the carboxy terminus.

The homology between glypican family members is moderate but extensive, being characterized by blocks of identity or conservative substitutions separated by stretches of divergent sequence. A pattern of 14 cysteine residues is found in all family members so far identified. The presence of these cysteines in the Drosophila dally protein, which appears to be related to the ancestor of the mammalian glypicans, indicates that this well conserved structural pattern is of an ancient origin.

The conserved pattern of cysteines also suggests that these PGs may have similar, compact tertiary structures, stabilized by disulfide bonds. This view is supported by the fact that both glypican and cerebroglycan show a dramatic increase in apparent molecular weight in SDS-PAGE gels upon reduction (Herndon and Lander, 1990). In addition, a proteolytic fragment obtained in some preparations of K-glypican can be only be observed in SDS-PAGE gels upon reduction (Watanabe et al., 1995).

It is also of interest to note the high level of conservation among orthologues of the mammalian glypican family members. The mature polypeptides of human glypican and rat glypican are 91% identical at the amino acid level, and the predictied protein sequences of the human cDNA fragments of cerebroglycan and K-glypican obtained in

this study are 87% and 78% identical to their respective rat and mouse counterparts. The ectodomains of the human and mouse syndecan-1 and human and rat syndecan-4 are only 70% and 76% identical respectively. In the syndecan family ectodomains, only a protease susceptible domain near the plasma membrane and the regions surrounding the GAG attachment sites are strongly conserved (Bernfield et al., 1992)., suggesting the primary function of the syndecan family ectodomains may be to present GAG chains at the cell surface. In contrast, the protein cores of the glypican family members may serve a more complex function that has constrained their evolution. For example, these core proteins might bind as yet unknown ligands. Examples of ligand-binding PG core proteins include the core proteins of betaglycan [which binds TGF-β (Cheifetz and Massague, 1989)] and decorin [which binds to collagen fibrils (Vogel et al., 1984; Krusius and Ruoslahti, 1986)].

Alternatively, the conserved protein structure of the glypican family HSPGs may play a regulatory role in post-translational processing. For example, the protein might interact with components of the Golgi apparatus in order to influence the type of GAG that it receives (e.g. HeS versus chondrotin sulfate), or even the sequence of GAG modifications (e.g. sulfation) that is subsequently generated. Recent studies underscore the importance of differences in GAG modification on the biological properties of HSPGs (Sanderson et al., 1992a; Nurcombe et al., 1993; San Antonio et al., 1993).

Finally, among the mammalian glypican family members, glypican, K-glypican, and cerebroglycan are all more closely related to each other than to OCI-5. This suggests that there may be a subfamilies within the glypican family. It will be interesting to see if, in the future, additional glypican family members are identified.

Genomic structure of Cerebroglycan

As a result of the partial characterization of CBG genomic structure, three CBG exons have been identified (Figure 13). The first exon (exon I, Figure 13) codes for the first 55 amino acids, including the signal sequence, and contains one of the fourteen conserved cysteine residues. Interestingly, this exon ends with the serine of the first ser-

gly putative GAG attachment site. Given the tendency for individual exons to contain intact functional domains (Darnell et al., 1986), the fact that the serine and glycine of this ser-gly pair are on different exons seems to disfavor the possibility that GAG is attached at this site. In syndecan-1 and syndecan-4, none of the serine-glycine pairs that are thought to attach GAG are split to separate exons (Vihinen et al., 1993; Baciu et al., 1994).

The second exon (exon II, Figure 13) encodes the next 53 amino acids, including three of the conserved cysteine residues and one putative GAG attachment site. This ser-gly (ser92) fits the minimal description of one type of GAG attachment site observed in the syndecan family, namely, a single ser-gly that is preceded and often flanked by acidic amino acids (Bernfield et al., 1992) -- D/E-X-S-G-(X)n-D/E -- although it has no flanking acidic amino acid residue.

Although the 3' boundary of the third exon has not been defined, another downstream exon (exon Q, Figure 13) was defined by the fortuitous cloning of a partially spliced cDNA (Figure 12). This small exon, which begins at Ile^{217} in CBG, encodes only 27 amino acids and contains no putative GAG attachment sites or conserved cysteine residues. However, the level of homology among glypican, K-glypican, and cerebroglycan is unusually high in this region. Glypican and K-glypican, with an overall identity of 41%, are 70% identical in the region of the exon. CBG and K-glypican, 36% identical overall, are 52% identical here, and CBG and glypican, which share a 37% identity overall, are 48% identical in the region of the exon. In addition, many of the non-identical residues represent conservative substitutions. The high level of conservation in the region of this exon suggests that it fulfills some important structural or functional role, at least in the most closely related mammalian family members, glypican, K-glypican, and CBG.

The 3' splice acceptor sequence of one intron in Figure 13 shows substantial deviation from the typical consensus sequence for splice acceptors (Mount, 1982). Typically, the invariant AG at the 3' intron: exon boundary is preceded by a distinctive polypyrimidine tract. In this case of this intron, many of the pyrimidines in the tract have been substituted with purines. This raises the possibility that CBG mRNA

expression is regulated at the level of splicing. It has been established that the efficient expression of many mammalian genes requires the presence of at least one intron. One recent study has shown that the level of increased expression conferred on an intronless minigene by the inclusion of an intron varies dramatically with the efficiency with which the intron is spliced out (Korb et al., 1993). Another study demonstrated that muscle-specific expression of a Ca++/ATPase splice variant is likely the result of trans-acting factors that activate an otherwise inefficient splicing process (Van den Bosch et al., 1994). Given the exquisite pattern of spatial and temporal CBG mRNA expression (as will be described in Chapter 3), the possibility of regulation at the level of splicing is intriguing. However, preliminary experiments have shown that the ~4 kb Xba fragment, upstream from exon I in the CBG genomic clone, is sufficient to confer restricted, CBG-like expression to a lacZ reporter construct in transgenic mice (John Fesenko, unpublished observations).

Localization of the CBG gene in the mouse genome

The segregation of a CBG-specific single stranded conformational polymorphism (SSCP) in a test cross of a B6/Spretus mouse with a B6 mouse resulted in the assignment of the CBG gene to the distal end of the long arm of chromsome 5. The known markers in this region are shown below:

```
                                            FLK2
        GNB-2                    ACT-1      FLT1
        ZP3      EPO     PDGFA              GSH-2
         ↓        ↓        ↓        ↓       ↓↓
         ▼        ▼                 ▼       ▼▼  END OF
        (1.3)    (4.1)            (1.1)  (0.5) CHROMOSOME 5

         |◄──────────────────────►|
                    CBG
```
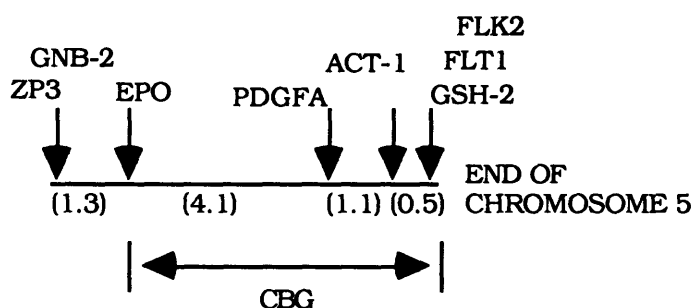
Cerebroglycan maps somewhere between EPO and the end of the chromosome. The numbers refer to distances in centimorgans between the arrows. The markers shown above and their location in the human genome are:

Marker	Name	hum. locat.
ZP3	zonapellucida protein 3	ND
GNB-2	guanine nucleotide binding protein B-2	7q21
EPO	erythropoetin	7q21
PDGFA	platelet derived growth factor, alpha	7pter-q21
ACT-1	actin-1	7pter-q22
GSH-2	genomic screening homeobox-2	ND
FLT-1	FMS-like tyrosine kinase-1	13p12
FLK-2	fetal liver tyrosine kinase-2	ND

Because of the nervous system-specific expression of CBG mRNA and protein (see Chapter 3), it was of interest to determine whether any known behavioral or neurological mutants colocalized with CBG in the mouse genome. No such markers are known to be present in the distal portion of chromosome 5 (Wayne Frankel, Jackson Labs, personal communication). The markers that map close to CBG are found on chromosomes 7 and 13 in the human genome, making these two chromosomes possible candidates for the location of the human CBG gene.

In summary, two of the three major cell surface HSPGs of the developing rat brain have been identified. These molecules are rat glypican, and a novel glypican-related protein that we have named cerebroglycan (CBG), and for which we have obtained a full length cDNA. Glypican and cerebroglycan belong to a family of HSPGs characterized by GPI-lipid anchorage to the cell surface, and conserved protein structure, including a pattern of 14 cysteine residues that is absolutely conserved. The developmental expression

of CBG mRNA and protein is restricted: no CBG is detected in adult brain. The genomic structure of CBG has been partially elucidated and its location in the mouse genome has been determined. In addition, mouse and human CBG cDNA fragments have been obtained. The reagents obtained in this study lay the foundation for further characterization of CBG function and expression in the developing rat nervous system.

ACKNOWLEDGEMENTS

References

Baciu, P.C., C. Acaster, and P.F. Goetinck. 1994. Molecular cloning and genomic organization of chicken syndecan-4. J. Biol. Chem. 269:669-703.

Bernfield, M., R. Kokenyesi, M. Kato, M.T. Hinkes, J. Spring, R.L. Gallo, and E.J. Lose. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu. Rev. Cell Biol. 8:365-393.

Bourdon, M.A., T. Krusius, S. Campbell, N.B. Schwartz, and E. Ruoslahti. 1987. Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. Proc. Natl. Acad. Sci. USA 84:3194-3198.

Bowe, M.A. and J.R. Fallon. 1995. The role of agrin in synapse formation. Annu. Rev. Neurosci. 18:443-462.

Carey, D., D. Evans, R. Stahl, V. Asundi, K. Conner, P. Garbes, and G. Cizmeci-Smith. 1992. Molecular cloning and characterization of N-syndecan, a novel transmembrane heparan sulfate proteoglycan. J. Cell Biol. 117:191-201.

Cheifetz, S. and J. Massague. 1989. The TGF-β receptor proteoglycan: cell surface expression and ligand binding in the absence of glycosaminoglycan chains. J. Biol. Chem. 264:12025-12028.

Chomczynski, P. and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate phenol-chloroform extraction. Anal. Biochem. 162:156-159.

Chuong, C.M., K.L. Crossin, and G.M. Edelman. 1987. Sequential expression and differential function of multiple adhesion molecules during the formation of cerebellar cortical layers. J. Cell Biol. 104:331-342.

Church, G.M. and W. Gilbert. 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81:1991-1995.

Cole, G.J., A. Loewy, and L. Glaser. 1986a. Neuronal cell-cell adhesion depends on interactions of N-CAM with heparin-like molecules. Nature 320:445-447.

Cunningham, D.D., S.L. Wagner, and D.H. Farrell. (1992). Regulation of protease nexin-1 activity by heparin and heparan sulfate. In Heparin and Related Polysaccharides, D. A. Lane, I. Bjork, and U. Lindahl, editors. (New York: Plenum), pp. 297-306.

Darnell, J., H. Lodish, and D. Baltimore. (1986). Molecular Cell Biology (New York: Scientific American Books, Inc.).

David, G., V. Lories, B. Decock, P. Marynen, J. Cassiman, and H.V.d. Berghe. 1990. Molecular cloning of a phosphatidylinositol-anchored membrane heparan sulfate proteoglycan from human lung fibroblasts. J. Cell Biol. 111:3165-3176.

Filmus, J., J.G. Church, and R.N. Buick. 1988. Isolation of a cDNA corresponding to a developmentally regulated transcript in rat intestine. Mol. Cell Biol. 8:4243-4249.

Filmus, J., W. Shi, Z.M. Wong, and M.J. Wong. 1995. Identification of a new membrane-bound heparan sulfate proteoglycan. Biochem. J. 311:561-565.

Gould, S.E., W.B. Upholt, and R.A. Kosher. 1992. Syndecan 3: A member of the syndecan family of membrane-intercalated proteoglycans that is expressed in high amounts at the onset of chicken limb cartilage differentiation. Proc. Natl. Acad. Sci. USA 89:3271-3275.

Gundersen, R.W. 1987. Response of sensory neurites and growth cones to patterned substrata of laminin and fibronectin in vitro. Dev. Biol. 121:423-431.

Herndon, M.E. and A.D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. Neuron 4:949-961.

Hughes, R.A., M. Sendtner, M. Goldfarb, D. Lindholm, and H. Thoenen. 1993. Evidence that fibroblast growth factor 5 is a major muscle-derived survival factor for cultured spinal motoneurons. Neuron 10:369-377.

Ivins, J.K., E.D. Litwack, A. Kumbasar, C.S. Stipp, A. Yang, and A.D. Lander. 1995. Cerebroglycan and glypican are cell surface heparan sulfate proteoglycans expressed on axons and growth cones in the developing rat nervous system. Soc. Neurosci. Abstracts 21:795.

Kan, M., F. Wang, J. Xu, J.W. Crabb, J. Hou, and W.L. McKeehan. 1993. An essential heparin-binding domain in the fibroblast growth factor receptor kinase. Science 259:1918-1921.

Karthikeyan, L., M. Flad, M. Engel, B. Meyer-Puttlitz, R.U. Margolis, and R.K. Margolis. 1994. Immunocytochemical and in situ hybridization studies of the heparan sulfate proteoglycan, glypican, in nervous tissue. J. Cell Sci. 107:3213-3222.

Kennedy, T.E., T. Serafini, and J.R.d.l. Torre. 1994. Netrins are diffusible chemotropic factors for commissural axons in the embyronic spinal cord. Cell 78:425-435.

Kodukula, K., L.D. Gerber, R. Amthauer, L. Brink, and S. Udenfriend. 1993. Biosynthesis of glycosylphosphatidylinositol (gpi)-anchored membrane proteins in intact cells: Specific amino acid requirements adjacent to the site of cleavage and gpi attachment. J. Cell Biol. 120:657-664.

Korb, M., Y. Ke, and L.F. Johnson. 1993. Stimulation of gene expression by introns: conversion of an inhibitory intron to a stimulatory intron by alteration of the splice donor sequence. Nuc. Acids Res. 21:5901-5908.

Kozak, M. 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266:19867-19870.

Krusius, T. and E. Ruoslahti. 1986. Primary structure of an extracellular matrix proteoglycan core protein deduced form cloned cDNA. Proc. Natl. Acad. Sci. USA 83:7683-7687.

LeBaron, R.G., J.D. Esko, A. Woods, S. Johansson, and M. Höök. 1988. Adhesion of glycosaminoglycan-deficient Chinese hamster ovary cell mutants to fibronectin substrata. J. Cell Biol. 106:945-952.

Lindner, J., F.G. Rathjen, and M. Schachner. 1983. L1 mono- and polyclonal antibodies modify cell migration in early postnatal mouse cerebellum. Nature 305:427-430.

Linker, A. and P. Hovingh. 1972. Heparinase and heparitinase from flavobacteria. Meth. Enzymol. 28:902-910.

Litwack, E.D., C.S. Stipp, A. Kumbasar, and A.D. Lander. 1994. Neuronal expression of glypican, a cell-surface glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan, in the adult rat nervous system. J. Neurosci. 14:3713-3724.

Marcum, J.A., C.F. Reilly, and R.D. Rosenberg. (1987). Heparan sulfate species and blood vessel wall function. In Biology of Proteoglycans, T. N. Wight and R. P. Mecham, editors. (Orlando: Academic Press Inc.), pp. 301-338.

Mount, S.M. 1982. Nuc. Acids Res. 10:459.

Murphy, M., J. Drago, and P.F. Bartlett. 1990. Fibroblast growth factor stimulates the proliferation and differentiation of neural precursor cells in vitro. J. Neurosci. Res. 25:463-475.

Nakato, H., T.A. Futch, and S.B. Selleck. 1995. The division abnormally delated (dally) gene: A putative integral membrane proteoglycan required for cell division patterning during post-embryonic development of the nervous system in Drosophila. Development 121:3687.

Neugebauer, K.M., K.J. Tomaselli, J. Lilien, and L.F. Reichardt. 1988. N-cadherin, NCAM, and integrins promote retinal neurite outgrowth on astrocytes in vitro. J. Cell Biol. 107.

Noakes, P.G., M. Gautam, J. Mudd, J.R. Sanes, and J.P. Merlie. 1995. Aberrant differentiation of neuromuscular junctions in mice lacking s-laminin/laminin beta 2. Nature 6519:258-262.

Nurcombe, V., M.D. Ford, J.A. Wildschut, and P.F. Bartlett. 1993. Developmental regulation of neural response to FGF-1 and FGF-2 by heparan sulfate proteoglycan. Science 260:103-106.

Porter, B.E., J. Weis, and J.R. Sanes. 1995. A motoneuron-selective stop signal in the synaptic protein S-laminin. Neuron 3:549-559.

Rapraeger, A., A. Krufka, and B.B. Olwin. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science 252:1705-1708.

Reichardt, L.F. and K.J. Tomaselli. 1991. Extracellular matrix molecules and their receptors: Functions in neural development. Annu. Rev. Neurosci. 14:531-570.

Reyes, A.A., R. Akeson, L. Brezina, and G.J. Cole. 1990. Structural requirements for neural cell adhesion molecule-heparin interaction. Cell Regulation 1:567-576.

Rogers, S.L., P.C. Letourneau, and I.V. Pech. 1989. The role of fibronectin in neural development. Dev. Neurosci. 11:248-265.

Salmivirta, M., K. Elenius, S. Vainio, U. Hofer, R. Chiquet-Ehrismann, I. Thesleff, and M. Jalkanen. 1991. Syndecan from embryonic tooth mesenchyme binds tenascin. J. Biol. Chem. 266:7733-7739.

Sambrook, J., E.F. Fritsch, and T. Maniatis. (1989). Molecular Cloning, A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

San Antonio, J.D., J. Slover, J. Lawler, M.J. Karnovsky, and A.D. Lander. 1993. Specificity in the interactions of extracellular matrix proteins with subpopulations of the glycosaminoglycan heparin. Biochem. 32:4746-4755.

Sanderson, R.D. and M. Bernfield. 1988. Molecular polymorphism of a cell surface proteoglycan: distinct structures on simple and stratified epithelium. Proc. Natl. Acad. Sci. USA 85:9562-9566.

Sanderson, R.D., T.B. Sneed, L.A. Young, G.L. Sullivan, and A.D. Lander. 1992a. Adhesion of B lymphoid (MPC-11) cells to type I collagen is mediated by the integral membrane proteoglycan, syndecan. J. Immunol. 148:3902-3911.

Sanes, J.R. 1989. Extracellular matrix molecules that influence neural development. Annu. Rev. Neurosci. 12:491-516.

Schaffner, W. and C. Weissman. 1973. A rapid, sensitive, and specific method for the determination of protein in dilute solution. Anal. Biochem. 56:502-514.

Serafini, T., T.E. Kennedy, M.J. Galko, C. Mirzayan, T.M. Jessell, and M. Tessier-Lavigne. 1994. The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. Cell 78:409-424.

Tempst, P., A.J. Link, L.R. Riviere, M. Fleming, and C. Elicone. 1990. Internal sequence analysis of proteins separated on polyacrylamide gels at the submicrogram level: Improved methods, applications and gene cloning strategies. Electrophoresis 11:537-553.

Unsicker, K., H. Reichert-Preibsch, R. Schmidt, B. Pettmann, G. Labourdette, and M. Sensenbrenner. 1987. Astroglial and fibroblast growth factors have neurotrophic functions for cultured periperal and central nervous system neurons. Proc. Natl. Acad. Sci. USA 84:5459-5463.

Van den Bosch, L., J. Eggermont, H.D. Smedt, L. Mertens, F. Wuytack, and R. Casteels. 1994. Regulation of splicing is responsible for the expression of the muscle-specific 2a isoform of the sarco/endoplasmic-reticulum Ca^{2+}ATPase. Biochem. J. 302:559-566.

Vihinen, T., P. Auvinen, L. Alanen-Kurki, and M. Jalkanen. 1993. Structural organization and genomic sequence of mouse syndecan-1 gene. J. Biol. Chem. 268:17261-17269.

Vogel, K.G., M. Paulsson, and D. Heinegard. 1984. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. Biochem. J. 223:587-597.

von Heijne, G. 1990. The signal peptide. J. Membr. Biol. 115:195-201.

Walicke, P. 1988. Basic and acidic fibroblast growth factors have trophic effects on neurons from multiple CNS regions. J. Neurosci. 8:2618-2627.

Walicke, P., W.M. Cowan, N. Ueno, A. Baird, and R. Guillemin. 1986. Fibroblast growth factor promotes survival of dissociated hippocampal neurons and enhances neurite extension. Proc. Natl. Acad. Sci. USA 83:3012-3016.

Watanabe, K., H. Yamada, and Y. Yamaguchi. 1995. K-glypican: a novel gpi-anchored heparan sulfate proteoglycan that is highly expressed in developing brain and kidney. J. Cell Biol. 130:1207-1218.

Woods, A. and J.R. Couchman. (1992). Heparan sulphate proteoglycans and signalling in cell adhesion. In Heparin and Related Polysaccharides, I. Bjork, D. A. Lane, and U. Lindahl, editors. (New York: Plenum), pp. 87-96.

Yayon, A., M. Klagsbrun, J.D. Esko, P. Leder, and D.M. Ornitz. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. Cell 64:841-848.

Figure 1. Purification of neonatal rat brain membrane HSPGs

Total PGs, isolated from a neonatal (P0) rat brain membrane fraction by DEAE chromatography, were digested with chondroitnase ABC and fractionated again on DEAE-Sephacel. The addition of a trace amount of previously isolated, radioiodinated P0 membrane PGs, prior to chondroitinase digestion, allowed the purification to be monitored. The cpm of ^{125}I in each fraction was measured in a γ-counter. Most of the CSPG core proteins [e.g. CSPGs "M1," "M6," and "M8" as identified in Herndon and Lander (1990)] flowed through when applied in 0.15 M NaCl (Peak I). A small but significant amount of the M1 protein core was retained, but could be eluted with a stringent wash: 0.1 M sodium acetate [pH 3.5], 0.2 M NaCl (peak II). The HSPGs were subsequently eluted with a gradient of 0.15-0.75 M NaCl (peak III). Earlier experiments demonstrated that peak III is centered around 0.45 M NaCl (not shown). See Materials and Methods for details.
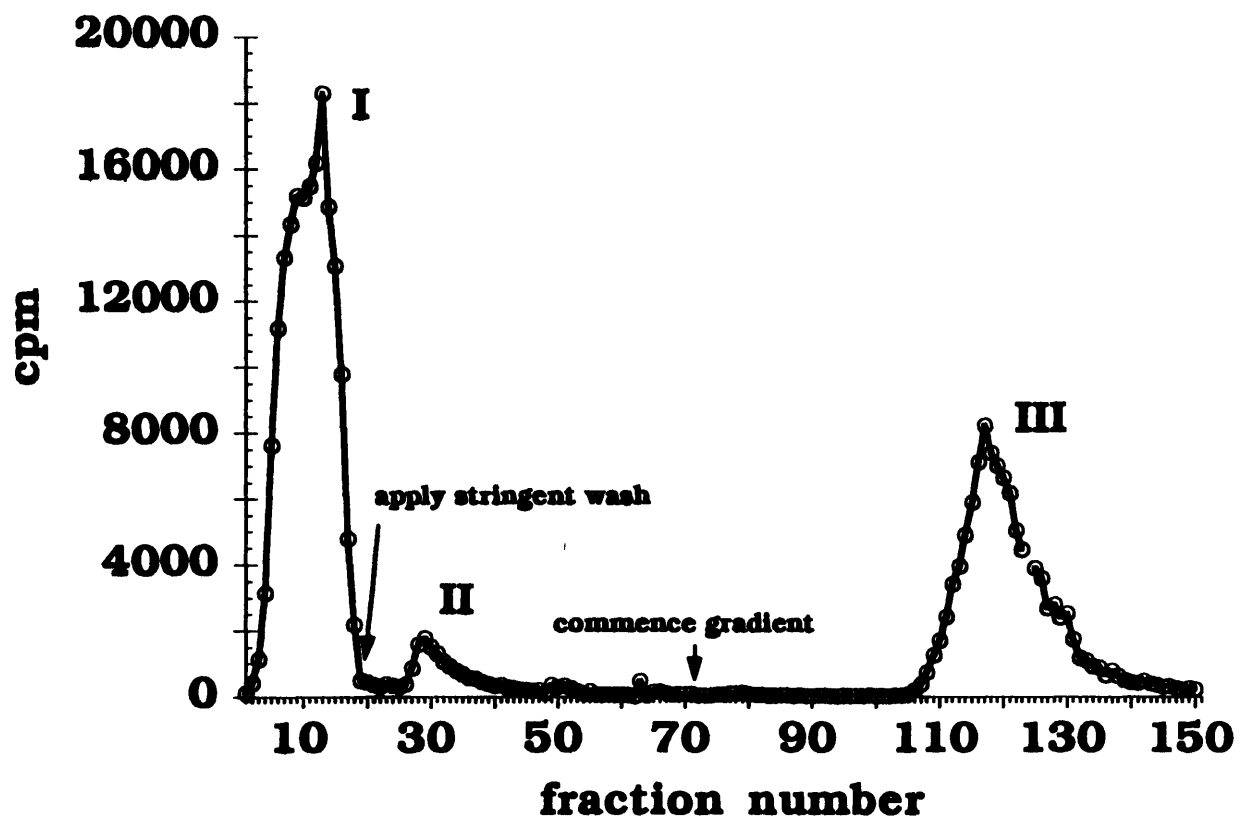
Figure 2. Electrophoretic separation of HSPGs M7, M12, and M13
To resolve the HSPG core proteins M7, M12, and M13, a two stage electrophoresis step was used.   (I) The heparitinase-digested, concentrated HSPG pool was loaded onto a 6% T acrylamide, SDS minigel, polymerized over a strip of 9% T acrylamide (lane 1 of the gel).   An enzyme blank and pre-stained molecular weight markers were loaded in lanes 2 and 3.   (II) After electrophoresis to resolve the M7 core protein, the M12 and M13 cores (in the 9% T acrylamide strip) were excised and re-cast in a second 9% T acrylamide gel.   (III) Further electrophoresis resolved the M12 and M13 cores.   The locations of the molecular weight markers are indicated as are the rough locations of the HSPG core proteins as revealed by subsequent electoblotting and amido black staining.   See Materials and Methods for details.
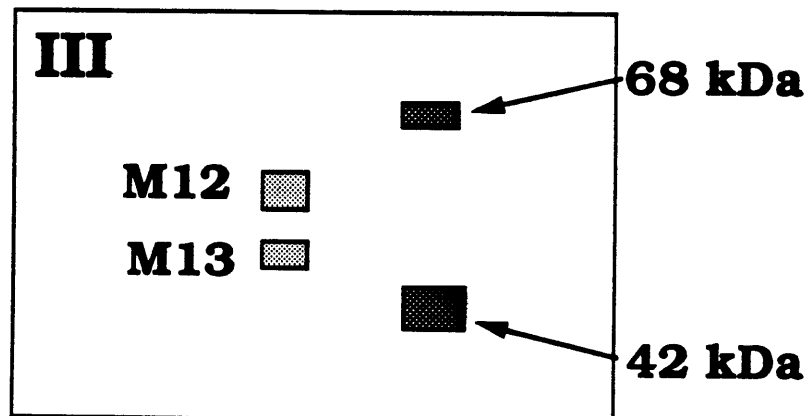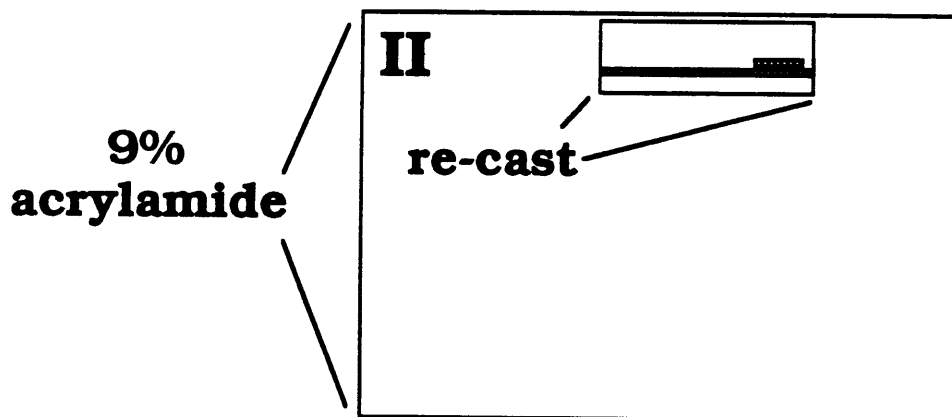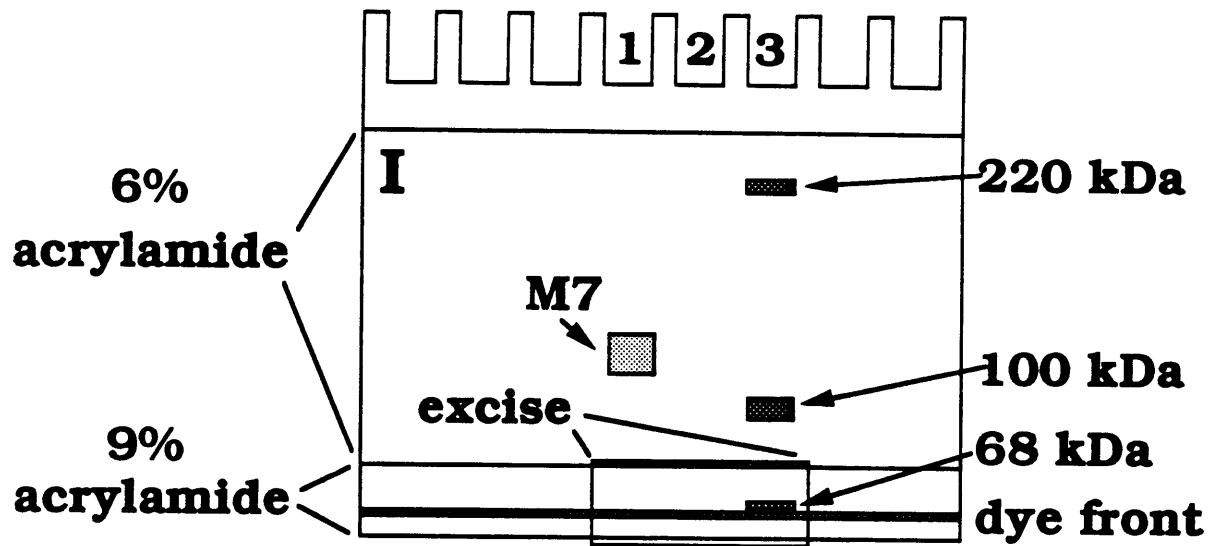
Figure 3. HPLC chromatography of M12-derived peptides
The HSPG M12 core protein was digested in situ with trypsin after electroblotting to nitrocellulose. The peptides liberated by digestion were separated by HPLC as described in Materials and Methods. The chromatograms show milli-absorbance units (mAU) versus elution time in minutes. Purple line is absorbance at 210 nm; green line is absorbance at 277 nm. (A) The chromatogram of the M12-derived tryptic peptides. The numbered peaks contained peptides whose sequences were determined (see Table 1 below). (B) The chromatogram of the enzyme blank. The blank was obtained by trypsin digestion performed on a congruent section of the nitrocellulose excised from the same level of the blot as the M12 band in an adjacent lane. The blank was composed of heparitinase at identical concentration in an identical buffer to the M12 sample. Peaks in (A) that eluted at identical times to peaks in (B) were disqualified from further analysis.

Peptide / Sequence

<table>
<tr><th>Peptide</th><th>Sequence</th></tr>
</table>

M13-9

   C   A TN
R S S A E E R P P T T A A G T N L X R
K W S   R G K Y L   H

M13-10

RALVAAR (with KSS below RAL)

M13-21

  A A
K T S E G L MHLQENSVK
R S C L G
   W

M13-22

   C L
KA A A GQDL**DMHDADEDA**SGSGGG**QQYADD**
RI W G
                22L   22R

M12-15a

K/R MELETALHDSSR

M12-15

K/R XAEALRPFGDAPR

M12-18

K/R TPLTHALPGLSEQEGQK

M12-26

K/R ALQATLATQLHGIDDHFQ

M12-29

K/R XL **PEVMGD**/FAITNN GL A/T NQGXNPEVD

Table 1. Peptide Sequence from HSPGs M12 and M13

Selected M12 and M13 tryptic peptides were subjected to automated Edman degradation after separation by reverse-phase HPLC. Multiple letters indicate ambiguities in the sequence, bold letters indicate the sequence which was later verified by from the deduced amino acid sequence of M 12 and M13 cDNAs. 22L and 22R: bars indicate the unambiguous amino acid sequence upon which the PCR primers 22L and 22R were based.

Figure 4. Cerebroglycan cDNA sequencing strategy
The restriction sites used in subcloning and sequencing are indicated
by letters: A= ApaI, B= BglII, E= EagI, K= KpnI, M= MspI, P= PstI,
R= EcoRV, S= SmaI, Sa= Sau3AI, Sc= SacI, Sp= SphI, and T= TaqI.
Arrows indicate individual sequencing reactions.   The custom
oligonucleotides used in sequencing are numbered.   Oligonucleotide
three was OM13-165, the primer used in the construction of the
primer extended library.   The three CBG cDNAs are indicated by the
heavy lines below the restriction map.

M13PX-2

M13PX-1

M13-A

——100 bp

Custom oligo
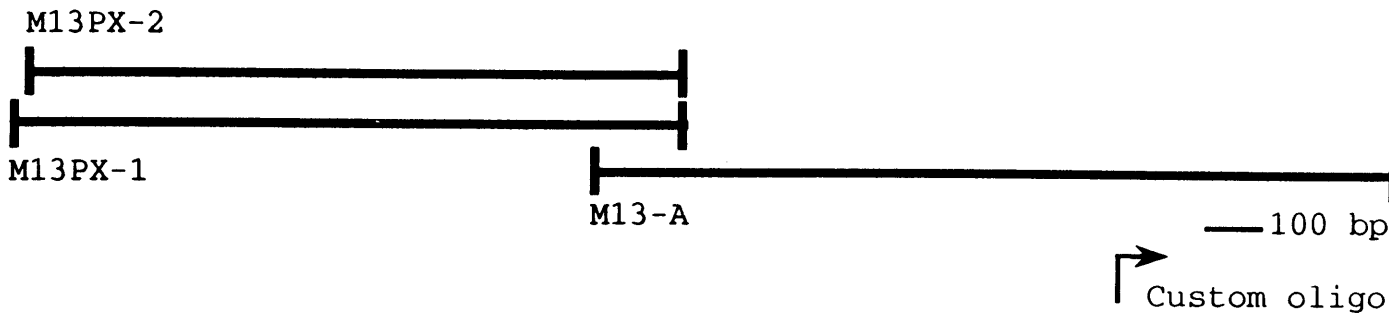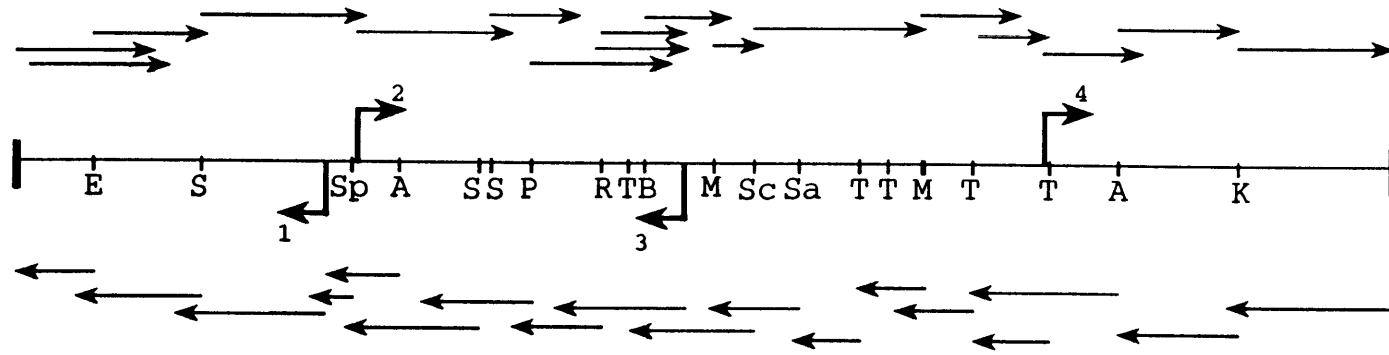
Figure 5. Nucleotide sequence of cerebroglycan cDNA and translation
The merged sequences of cDNAs M13-A, M13PX-1, and M13PX-2 are given along with the deduced amino acid sequence of M13. Potential glycosaminoglycan attachment sites are boxed. Cysteine residues are circled. Arrowheads indicate potential cleavage sites of the signal sequence (after Gly^{21}) and the glycosylphosphatidylinositol attachment signal sequence (after Asn^{553} or Ser^{557}). (See text for discussion.) Sequences that correspond to M13 peptides (Table 1) are underlined. The cerebroglycan cDNA sequence has been deposited in the GenBank/EMBL database (accession number L20468).

```
        1
       GATCATTGGCGGGCTCCAGCTTCGGGTCCCGCCCCCTGTCCCGCTCCCCGGCTCCCATTGTCTCCGCAGATGCCTCCTAGTCCAGCTACCGTGCTGGGGCTGTTCAGTTTTGGGG   115
  116  GGGACGCTCGTATTCTTTTCTTCGAGACGTCCCGCGGCCGAGAACAGGATTTTAGAGTTCCGTGGACTCTTTTTTCCTCCTTAGGACTCACTTTGCCCCGTCCCAGGAAGGAGAGCAGCT   235
  236  ATGTCCGCGGTGCGACCTCTCCTGCTTTTGCTGCTCCCTCTGTGTCCCGGTCCAGGACCGGGACATGGGAGTGAAGCAAAGGTCGTCCGGAGTTGTGCAGAGACTCGGCAGGTGCTGGGG   355
        M  S  A  V  R  P  L  L  L  L  L  L  L  P  L  C  P  G  P  G  P  G  H  G  S  E  A  K  V  V  R  S  Ⓒ  A  E  T  R  Q  V  L  G
        1                                              21                          ▲                                           40

  356  GCCCGGGGATATAGCTTAAACCTTATCCCTCCCTCCCTCATCTCAGGTGAGCACCTTCAGATCTGTCCTCAGGAGTACACCTGCTGTTCCAGTGAGACAGAGCAGAAGTTGATCCGGGAT   475
        A  R  G  Y  S  L  N  L  I  P  P  S  L  I  [S  G] E  E  L  Q  I  Ⓒ  P  Q  E  Y  T  Ⓒ Ⓒ  S  S  E  T  E  Q  K  L  I  R  D
        41                               61                          Ⓒ                                                        80

  476  GCTGAGGTCACCTTCCGTGGCCTGGTGGAAGACAGTGGCTCCTTCCTGATTCACACGCTGGCTGCCCGCCACAGAAAGTTTAACGAGTTTTTTCGGGAGATGCTGTCTATATCCCAGCAT   595
        A  E  V  T  F  R  G  L  V  E  D  [S  G] S  F  L  I  H  T  L  A  A  R  H  R  K  F  N  E  F  F  R  E  M  L  S  I  S  Q  H
        81                          101                                                                                       120

  596  TCTTTGGCCCAGCTCTTCTCGCATTCCTATGGTCGCCTGTATTCCCAGCATGCCGTCATCTTCAATAGCCTGTTCTCCGGCCTGCGGGACTACTATGAGAAGTCTGGTGAGGGGTTAGAT   715
        S  L  A  Q  L  F  S  H  S  Y  G  R  L  Y  S  Q  E  A  V  I  F  N  S  L  F  S  G  L  R  D  Y  Y  E  K  [S  G] E  G  L  D
        121                        141                                                                  [S  G]              160

  716  GACACCTTGGCAGACTTCTGGGCCCAGCTCCTGGAGAGAGCCTTCCCTTTGCTGCACCCACAATACAGTTTCCCTCCTGATTTCCTGCTCTGCCTTACTCGGCTCACTTCAACTGCTGAT   835
        D  T  L  A  D  F  W  A  Q  L  L  E  R  A  F  P  L  L  H  P  Q  Y  S  F  P  P  D  F  L  L  Ⓒ  L  T  R  L  T  S  T  A  D
        161                        181                                              Ⓒ                                        200

  836  GGCTCTCTCCAGCCCTTCGGGGACTCGCCCCGCCGCCTCCGCCTACAGATAACCCGGGCACTGGTGGCAGCCCGGGCCTTGGTCCAGGGTCTGGAGACCGGAAGAAATGTGGTCAGCGAA   955
        G  S  L  Q  P  F  G  D  S  P  R  R  L  R  L  Q  I  T  R  A  L  V  A  A  R  A  L  V  Q  G  L  E  T  G  R  N  V  V  S  E
        201                                             221                                                                 240

  956  GCACTTAAGGTACCCATGTTGGAAGGCTGCAGACAGGCCCTCATGCGTCTGATTGGCCCGCTGCGTGGGGTACCCTCGCTTATGCCCTGCCGGGGCTTCTGCCTCAATGTGGCC   1075
        A  L  K  V  P  M  L  E  G  Ⓒ  R  Q  A  L  M  R  L  I  G  Ⓒ  P  L  Ⓒ  R  G  V  P  S  L  M  P  Ⓒ  R  G  F  Ⓒ  L  N  V  A
        241                      Ⓒ                   261                                                                    280

 1076  CACGGCTGTCTCAGCAGCAGGGGATTGGAGCCTGAATGGGGCGGATATCTGGATGGTCTCCTGCTGCTGAGAAACTCCAGGGACCCTTTTCCTTCGAGCTGGCTGCTGAGTCCATA   1195
        H  G  Ⓒ  L  S  S  R  G  L  E  P  E  W  G  G  Y  L  D  G  L  L  L  L  L  A  E  K  L  Q  G  P  F  S  F  E  L  A  A  E  S  I
        281                                             301                                                                 320

 1196  GGGGTGAAGATCTCAGAAGGTTTGATGCACTTGCAGGAAAAACAGTGTAAAGGTGTCAGCCAAGGTGTTTCAGGAATGTGGGACGCCTCACCCGGTGCAATCCCGTAACCGGCGAGCACCA   1315
        G  V  K  I  S  E  G  L  M  H  L  Q  E  N  S  V  K  V  S  A  K  V  F  Q  E  Ⓒ  G  T  P  E  P  V  Q  S  R  N  R  R  A  P
        321                                             341                                                                 360

 1316  GCACCCCGGGAAGAGACTAGCCGGTCTTGGAGGTCGTCAGCTGAGGAAGAACGCCCCACAACAGCGGCAGGCACTAATCTGCACCGTCTGGTTTGGGAGCTCCGAGAGAGGCTTAGCCGA   1435
        A  P  R  E  E  T  S  R  S  W  R  S  S  A  E  E  E  R  P  T  T  A  A  G  T  N  L  H  R  L  V  W  E  L  R  E  R  L  S  R
        361                                             381                                                                 400

 1436  GTGCGGGGCTTCTGGGCTGGACTTCCCGTTACGGTGTGTGGGGACTCCCGCATGCGGCAGATCTCTCCCAGGAGGCGGCACCCTGCTGGACTGGGGTTGGGAGAGGCCGGTATATGTCG   1555
        V  R  G  F  W  A  G  L  P  V  T  V  Ⓒ  G  D  S  R  M  A  A  D  L  S  Q  E  A  A  P  Ⓒ  W  T  G  V  G  R  G  R  Y  M  S
        401                                             421                                                                 440

 1556  CCTGTGGTCGTGGGCTCCTTGAACGAGCAGCTCCATAACCCGGAGTTGGATACTTCGAGTCCCGATGTCCCGACACGTCGGCGGAGGCTCCATCTCCGAGCAGCTACTGCTCGAATGAAG   1675
        P  V  V  V  G  S  L  N  E  Q  L  H  N  P  E  L  D  T  S  S  P  D  V  P  T  R  R  R  R  L  H  L  R  A  A  T  A  R  M  K
        441                                             461                                                                 480

 1676  GCGGCTGCGCTGGGCCAGGACCTTGACATGCATGATGCGGATGAAGACGCCAGTGGCTCCGGAGGGGGACAGCAGTACGCAGACGACTGGAAAGCTGGGGCAGCGCCTGTGGTTCCCCCA   1795
        A  A  A  L  G  Q  D  L  D  M  H  D  A  D  E  D  A  [S  G  S  G] G  G  Q  Q  Y  A  D  D  W  K  A  G  A  A  P  V  V  P  P
        481                                             501                                                                 520

 1796  GCCCAGGCCTCCAAGGCCACCTCGTCCTCCTCGAAGGGATGGTCTTGGAGTGAGACGGAGGAAGTGGCCAGATACAACCAGGGCCGAAGCAGGAATCTGGGGTCCTCTGTTGGCCTC   1915
        A  R  P  P  R  P  P  R  P  P  R  R  D  G  L  G  V  R  G  S  G  S  A  R  Y  N  Q  G  R  S  R  N  L  G  S  S  V  G  L
        521                                             541                                                          ▲    ▲    560

 1915  CATGCCCCACGTGTCTTCATTCTCCTCCCCTCAGCCCTGACCCTGCTTGGACTTCGATAACAAGCGAGGGGCGCCCTAGCGTCAGGAGGCTTAGTGGCCTTTTCCTTCCACCTCCCTCAG   2035
        H  A  P  R  V  F  I  L  L  P  S  A  L  T  L  L  G  L  R  *
        561

 2036  CTGGGCATGAGGAAGAAATGGAAGGGGGCTACCGGGAGGGCAGAAAGGGACTTTGTAGGTAAATGGCTGGGGCCCAAATCCAGGAGTTTCCTATTAGTGGGTTGGGAGGGGAGGGTAGATG   2155
 2156  GGTGTTCATAATATTTATTTTTTCATTTGGGAACTGGGGATATTTATTTAGGAAGGGATTGTGGTTCAGCTCTTTTAAGGCTTGGGTCTCATTGATGCAAAAGGCAGGGGGTGGGGTTGG   2275
 2276  GATCGGGTTGAAGGAACTTGGAAGTGGGGAGGGGTGGCACTATTGGCATGGGGGTACCTGATATTGAAAATGAACAAATAAACTAGCTGAGAGTGATGTCACTTGCCTGTAAACCATTAT   2395
 2396  TTTGGAGGCTGAGGCAGGAGGATTGAGCGTTGGAGACCCATAAAATTAGGTGCTGATAGTGGAACAGCCGGGGTGGGAGGTAGGACCTGAGTTTTATTGACTTTTAAAATCATCTGTTTC   2515
 2516  TTAAATATGGGGGAAACCCATCTTGTGACCTTTGCTGTCTAATGTTTCAGTGTTCACATCGGTGTTTATTGGTAAATAAACAATTTGACCATGAAAAAAAAAAAAAAAAAAAAAAAAA   2631
```

Figure 6. Cerebroglycan is a member of a family of glypican-related HSPGs. (A) The protein deduced sequence of the CBG cDNA (rCEREBRO) is shown aligned with the sequences of four other molecules: the Drosophila protein, dally (dDLY); rat OCI-5 (rOCI-5); mouse K-glypican (mK-GLYP); and rat glypican (rGLYP). The top line represents the sequence of human glypican (hGLYP). The sequence of human glypican is identical to the sequence of rat glypican except where noted. All five molecules share a pattern of 14 conserved cysteines (black, underlined). Conservative substitutions are indicated with the use of colors, as indicated at the bottom of the figure. (B) The percent identity of each molecule for the others is shown.

A

```
hGLYP                                                                                      G
rGLYP      MELRARGWWLLCAAAALVACARGD--PASK---------------------------------------SRSCSEVR
mK-GLYP    MARLGLLALLCTLAALSASLLA----AELK---------------------------------------SKSCSEVR
rCEREBRO   MSAVRPLLLLLLPLCPGPGPGHGS--EAKV---------------------------------------VRSCAETR
rOCI-5     MAGTVRTACLLVAMLLGLGCLG A PPPPP----------------------------------------DATCK VR
dDLY       MAARSVRLA LLLFTLLCGFVGLSAAKHLDLDGIKKH HHLHSATTHRRRL RDSRAKDAVGGSTH CDAVK


                                             R HA     R     V     M        RSF
           IYGARGFSLSDVP AEISGEHLRIC-P GYTCCISEMEE LA HSRMELETALHDSSRAL ATLAT LHGID
           RLYVSKGF K DAPLYEI GDHLKIC-P DYTCCS EMEEKYSL SKDDFKTVVSE C HL AIFASRYKKFD
           VLGARGYSL LIPPSLISGEHL IC-P EYTCCSSETE KLIRDAEVTFRGLVEDSGSFLIHTLAARHRKF
           SFF RL PGLKWVPETPVPGSDL VCLPKGPTCCSRKMEEKY LTARL ME LL- SASMELKFLII  AAV-
           SYF-----ESIDIKSSGTYSEKGAIC---GG CC ATELELRDKAAGMFE LLHHFTSSLRGVLET AK F


                     K         AT      E        A         S
           DHF RLL DSERTL DAFPGAFGDLYT  TRAFRDLYAEL-----------------RLYYRGA LHLEETL
           EFFKELLE AEKSL DMFVKTYGHLYM  SELFKDLFVEL-----------------KRYYVAG V LEEML
           EFFREMLSIS HSLA LFSHSYGRLYS HAVIF SLFSGL-----------------RDYYEKSGEGLDDTL
           --F EAFEIVVRHAK YT AMFK  YPSLTP AFEFVGEFFTDV-------------SLYILGSDI VDDMV
           SHVLELA ISE MTHSLFSKVYTRMVPSSRMMIH LYTEIM HLIYTS YT S G LGRRGIGSV S LEEAV


                                                              E
           AEFWARLLERLFK LHP L-----LLPDDYLDCLGK ----AEALRPFGDAPRELRLRATRAFVAARSFV GL
           DFWARLLERMFRLV S Y-----HFTDEYLECVSKY----TE LKPFGDVPRKLKL VTRAFVAARTFA GL
           ADFWA LLERAFPLLHP Y-----SFPPDFLLCLTRLTSTADGSL PFGDSPRRLRL ITRALVAARALV GL
           ELFDSLFPVIYT MM PG---LPESVLDI ECLRGA----RRDLKVFGSFPKLIMT VSKSL VTRIFL AL
           HHFFV LFPVAYH MVHLSK  LGDLHEDYV CL H ----FDEMHPFGDIP  V S LGKSVHMS VFM AL


                            G                L
           GVASDVVRKVA-- VPLAPE-CSRAVMKLVYCAHCRG---------VPGARPCPDYCR VLKGCLA  A-DLD
           AVARDVVSKVS--VV PTA -CTHALLKMIYCSHCRG---------LVTVKPCY YCS IMRGCLA  G-DLD
           ETGR VVSEAL--KVPMLEG-CR ALMRLIGCPLCRG---------VPSLMPCRGFCL VAHGCLSSHG--LE
           LGIEVI TTD--HLKFSKD-CGRMLTRMWYCSYC G---------LMMVKPCGGYC VVM GCMAGVV-EID
           L AAEVLSEADALYGE LTDTCKLHLLKMHYCP C GHHSSSRSET----KLCYGYCK VMRGCSAEYAGLLD


                        T    V S        T           R                        P
           AEWR LLDSMVLITDKFWGP-SGAE VIGSVHMWLA--EAI AL D KDTLTAKVI GCG PKV---- PHGS
           FEW  FIDAMLMVAERLEGP-F IESVMDPIDVKIS--DAIM M D SV  S KVF GCGPPKP----LPAGR
           PEWGGYLDGLLLLAEKL GPFSF-ELAAESIGVKIS--EGLMHL E SVKVSAKVF ECGTPHPV SR RRAP
           KYWR---EYILSLEELV GMYRIYDME VLLGLFSTIHDSI YV K GGKLTTTIGKLCAHS--- R YRSA
           SPWSGVVDSL LVTTHILSDTGII VIKHL TYFS--EAIMAAMH GPELEKKVKKTCGTP----SLTPYSS


                             PR R PPS                V                  L T
           GPEEKKRGK--LAL EK--SSTGT---LEKLVSEAKA LRDI D----YWISLPGTLCSEKMAMSPASDD--
           ISRSISESAF--SARFRPYHPE RPTTAAGTSLDRLVTDVKEKLK AKKFWSSLPSTVC DERMAAG E ED-
           APREETSRSWRSSAEEER--PTTAAGT LHRLVWELRERLSRVRG----FWAGLPVTVCGDSRMAADLS EAA
           YYPEDLFIDKKVLKVARVEHEETLS---SRR--RELI KLKSFIS----FYSALPGYICSHSPVAE DTL---
           GEPDARPPPHK VKWATDPDPGMV------LFLSTIDKSKEFYT-------TIVD FCDE  HSRDDHS---


                    MA R                              S
           RCW GIS-KGRYLPEVMGDGLA  I  PEVEVDITKPDMTIR  IM LKIMT RLRGAYGG----- DVDF D
           DCW GKG-KSRYLFAVTG GLA  G  PEV VDTSKPDILILR IMALRVMTSKMK AY G----- DVDFFD
           PCWTGVG-RGRYMSPVVVGSL E LH PEL--DTSSPDVPTRRRRLHLRAATARMKAAALG----- DLDMHD
           -CW G ELVERYS KAAR GMK  F LHELKMKGPEPVVS IIDKLKHI LLRTMSVPKG-----KVVDKSL
           -CWSGDH-FGDYT LLI PGTDS RY PEVPF A-KA TGKL ELVDKLFKIRKSIGAAAPS SI ATHDI D


                     D L    L  K  R                                          SC
           ASDDGSGSGSGGGCPDD--ACGRRVSKKSSSSRTPLTHALPGLSE EG-------------- KTSAATR--
           ISDESSGEGSGSGCEY -- CPSEFEY ATD------HSGKSA EKAD----------------SAGGA--
           ADEDASGSGGG  YADDWKAGAAPVVPPARPPRPPPRPPPRRDGLGVRGG--SGSARY  GRSR LGSSVGL--
           -DEEGLESGD---CGDDEDECIGS--SGDGMMKVK  LRFLAELAYDLDVDDAPG K HG  KD EITTSHS
           M GEGSGGGEG I-GDDEEEYGGAHGSGDGSGDGPHTPIEESEGTTT EV---------ESRDSGKTSGS--


                      PT L                  PL LF A TV
           --------PEPHYFF------------LLFLFTLVLAAARPRWR*
           --------HAETKPY-----------LLAALCILFLAV GEWR*
           --------HAPRVFI------------LLPSALTLL-----GLR*
           VG MPSPLKILISVAIYVACFFSWCTDLPCPCLCCPAAPCGPPT*
           -------- PLEGTA------------TWMLLTLVTMLFSSCS*
```

R,K,H are green; E,D are red; F,W,Y,I,L,V,M are dark blue; S,T are cyan; Q,N are
C is black and <u>underlined</u>; G,A,P are black

B

	DALLY	OCI-5	GLYPICAN	CEREBROGLYCAN
OCI-5	21%			
GLYPICAN	22%	21%		
CEREBROGLYCAN	21%	19%	37%	
K-GLYPICAN	20%	20%	41%	36%

Figure 7. Southern analysis

10 µg of rat genomic DNA was digested to completion with one of three restriction enzymes, separated by electrophoresis, blotted to nitrocellulose, and probed with an α^{32}P-dCTP labelled CBG probe. Hybridization and subsequent washes were performed at high stringency. The sizes of DNA standards are indicated. The restriction enzymes used were HindIII (H), KpnI (K), and XhoI (X).
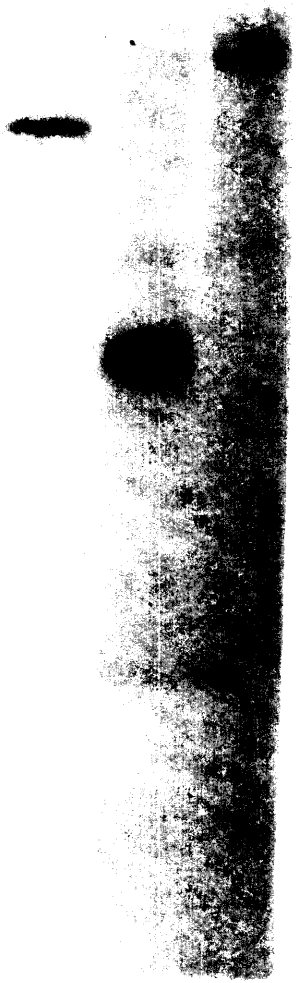
Figure 8. Northern analysis

20 μg of rat brain total RNA, isolated from embryonic day 18 (E18), post-natal (P0), and adult (A) rats, was subjected to electrophoresis through formaldehyde agarose and transferred to nitrocellulose. Arrow: the 2.7 kb transcript detected with an α^{32}P-dCTP labelled CBG probe. Lower panel: reprobing of the blot with a GAPDH probe demonstrated that a similar amount of RNA was present in each lane. The positions of molecular weight markers and ribosomal RNA bands are indicated.

E18 P0 Ad

9.5 —
7.5 —

4.4 —   ▬ 28S

2.4 —

                                          ▬ 18S

1.4 —

0.24 —

GAPDH
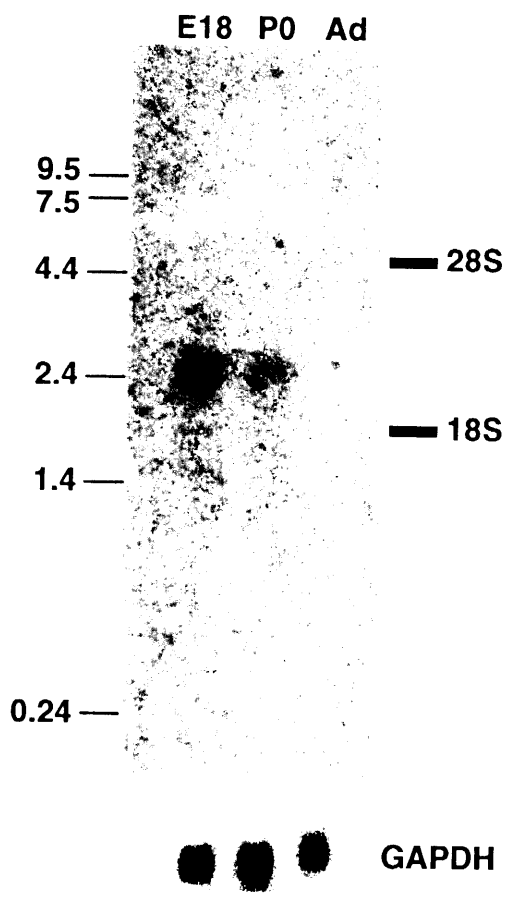
Figure 9. Comparison of mouse and rat cerebroglycan sequence
A mouse CBG cDNA fragment was isolated by PCR with the primers OM13-165 and OCBG-137 and embryonic mouse brain cDNA. (A) The sequence of the mouse CBG cDNA fragment (Mu) is shown compared to the rat CBG (Rat) sequence. The numbers correspond to the location of nucleotides in the rat CBG cDNA sequence shown in Figure 5. The primer sites of OM13-165 and OCBG-137 are indicated. (Δ): the location of an intron in rat genomic CBG, discovered by the fortuitous cloning of cDNAs from partially spliced CBG mRNA (see Figure 12, and Results below). The primer sites of primers used to amplify a mouse genomic CBG fragment for mapping CBG in the mouse genome are enclosed by lines (see Results below). The primer site upstream of the intron corresponds to primer muCBG-L, and the primer site downstream of the intron corresponds to muCBG-R. (B) The sequence of the deduced protein sequence of mouse CBG (Mu) is shown compared to the corresponding sequence from rat (Rat). The single amino acid substitution between the two sequences is shown bold and underlined.

A

```
Mu      ACTGGTGGCAGCCCGGGCCTTGGTCCAGGGTCTGGAGACTGGAAG
        ocbg-137       ||||||||||||||||||||||| |||||
Rat(911)               GCCTTGGTCCAGGGTCTGGAGACCGGAAG

        AAATGTGGTCAGCGAAGCACTTAAGGTACCAGTGTTGGAAGGCTG
        ||||||||||||||||||||||||||||||||||| |||||||||||
940     AAATGTGGTCAGCGAAGCACTTAAGGTACCCATGTTGGAAGGCTG
                                         Δ
        CAGACAGGCCCTGATGCGTCTGATTGGCTGCCCTCTTTGTCGGGG
        |||||||||||| |||||||||| ||||||| |||||||||||
985     CAGACAGGCCCTCATGCGTCTGATCGGCTGCCCACTTTGTCGGGG

        AGTTCCTTCACTTATGCCCTGCCGGGGCTTCTGCCTCAATGTAGC
        ||| || || ||||||||||||||||||||||||||||||||| ||
1030    AGTACCCTCGCTTATGCCCTGCCGGGGCTTCTGCCTCAATGTGGC

        CCATGGCTGCCTCAGCAGTAGGGGACTGGAGCCTGAATGGGGCGG
        ||| ||||| ||||||||| |||||| ||||||||||||||||||
1075    CCACGGCTGTCTCAGCAGCAGGGGATTGGAGCCTGAATGGGGCGG

        ATATCTGGATGGTCTCCTGCTGCTGGCTGAGAAACTCCAAGGACC
        |||||||||||||||||||||||||||||||||||||| |||||
1120    ATATCTGGATGGTCTCCTGCTGCTGGCTGAGAAACTCCAGGGACC

        CTTTTCCTTTGAACTGGCTGCTGAGTCCATTGGGGTGAAGATCTC
        |||||||| || ||||||||||||||||| |||||||||||||||
1165    CTTTTCCTTCGAGCTGGCTGCTGAGTCCATAGGGGTGAAGATCTC

        AGAAGGTTTGATGCACTTACAGGAAAACAGTGTAAAGGTGTCAGC
        |||||||||||||||||| |||||||||||||||||||||||||||
1210    AGAAGGTTTGATGCACTTGCAGGAAAACAGTGTAAAGGTGTCAGC

        GAAGGTGTTTCAGGAATGTGGGACGCCTCA   Mu
        |||||||||||              om13-165
1255    CAAGGTGTTTCA                   Rat
```

B

```
Mu:       ALVQGLETGRNVVSEALKVPVLEGCRQALMRLIGCPLCRGVPSLM
Rat:  226 ALVQGLETGRNVVSEALKAPVLEGCRQALMRLIGCPLCRGVPSLM

          PCRGFCLNVAHGCLSSRGLEPEWGGYLDGLLLLAEKLQGPFSFEL
      271 PCRGFCLNVAHGCLSSRGLEPEWGGYLDGLLLLAEKLQGPFSFEL

          AAESIGVKISEGLMHLQENSVKVSAKVF
      316 AAESIGVKISEGLMHLQENSVKVSAKVF
```

Figure 10. Isolation of human cerebroglycan and K-glypican cDNA fragments. Putative human CBG and K-glypican cDNA fragments were isolated by PCR using a nested set of primers. The four primers gdoL1, gdoL2, gdoR2, and gdoR3, were based on highly conserved protein sequences in CBG, glypican, and K-glypican. (A) The protein sequences upon which PCR primers were based. The sequences represent the consensus sequence of CBG, glypican, and K-glypican; numbered amino acids refer to the location of the amino acid in the rat CBG protein sequence as shown in Figure 5. The arrows indicate the directionality of the primers. See Materials and Methods for details of the PCR reactions. The cDNA and deduced protein sequences of human and rat CBG and human and rat K-glypican are compared in (B) and (C) respectively. Both human cDNA sequences are 81% identical to their rodent counterparts. The predicted protein sequence of the human K-glypican fragment is 78% identical to the mouse protein sequence, and the predicted protein sequence of the human cerebroglycan fragment is 87% identical to the rat protein sequence.

A

PQ(G/D/E)YTCC⁶⁹→ . . . FWA(Q/R)LLE¹⁷²→ . . .
  gdoL1   gdoL2

. . . ← ²¹⁸TRA(F/L)VAAR . . . ← ³⁴¹KV(I/F)Q(G/E)CG
           gdoR3   gdoR2

Note: The superscript numbers above are sequence position markers in the original figure.

B

```
          173R  A  F  P  L  L  H  P  Q  Y  S  F  P  P  D
rCBG   752AGAGCCTTCCCTTTGCTGCACCCACAATACAGTTTCCCTCCTGAT
          ||||  ||||| |||||||||||||| ||||| ||||| |||| |
huCBG     AGAGTGTTCCCGCTGCTGCACCCACAGTACAGCTTCCCCCCTGAC
           R  V  F  P  L  L  H  P  Q  Y  S  F  P  P  D


             E  L  L  C  L  T  R  L  T  S  T  A  D  G  S
rCBG      TTCCTGCTCTGCCTTACTCGGCTCACTTCAACTGCTGATGGCTCT
          | ||||||||||| | || | | ||| || | || ||||||
huCBG     TACCTGCTCTGCCTCTCACGCTTGGCCTCATCTACCGACGGCTCT
             Y  L  L  C  L  S  R  L  A  S  S  T  D  G  S


             L  Q  P  F  G  D  S  P  R  R  L  R  L  Q  I217
rCBG      CTCCAGCCCTTCGGGGACTCGCCCCGCCGCCTCCGCCTACAGATA
          || |||||||| |||||||| ||||||||||||||||| ||||||
huCBG     CTGCAGCCCTTTGGGGACTCACCCCGCCGCCTCCGCCTGCAGATA
             L  Q  P  F  G  D  S  P  R  R  L  R  L  Q  I
```

C

```
           171R  M  F  R  L  V  N  S  Q  Y  H  F  T  D  E
mKglyp  980CGCATGTTTCGCCTGGTGAACTCCCAGTACCATTTTACAGATGAG
           |||||||||| ||| | ||| | |||| || || | || |||
hKglyp     CGGATGTTTCAGCTGATAAACCCTCAGTTACACTTCAGTGAAGAC
             R  M  F  Q  L  I  N  P  Q  L  H  F  S  E  D


             Y  L  E  C  V  S  K  Y  T  E  Q  L  K  P  F
mKglyp    TACTTGGAATGTGTGAGCAAATATACAGAGCAGCTGAAGCCTTTT
          ||| |||||||||||||||||||||| || ||||||||| ||| |||
hKglyp    TACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTT
             Y  L  E  C  V  S  K  Y  T  D  Q  L  K  P  F


             G  D  V  P  R  K  L  K  L  Q  V  T  R  A  F215
mKglyp    GGAGATGTCCCTCGGAAACTGAAGCTCCAGGTTACCCGCGCATTT
          ||||| || || ||||||||||||||| | |||||||||||||||  |
hKglyp    GGAGACGTGCCCCGGAAACTGAAGATTCAGGTTACCCGCGCCCTC
             G  D  V  P  R  K  L  K  I  Q  V  T  R  A  L
```

Figure 11. Partial restriction map of the cerebroglycan genomic locus
Overlapping CBG genomic clones were isolated by screening a mouse genomic DNA library with fragments of the rat CBG cDNA, as described in the text. A partial restriction map of the clones is shown. The locations of the intron/exon boundaries, determined by sequencing the clones, are indicated. *This figure was courtesy of Scott Saunders.*

ATG

Exon 1   Exon 2

Native Cerebroglycan Gene

Xba I
Bgl II

Bgl II

EcoR I

Hind III

Eag I
Xba I

EcoR I

Xba I
EcoR I

Bgl II

1 kb

131
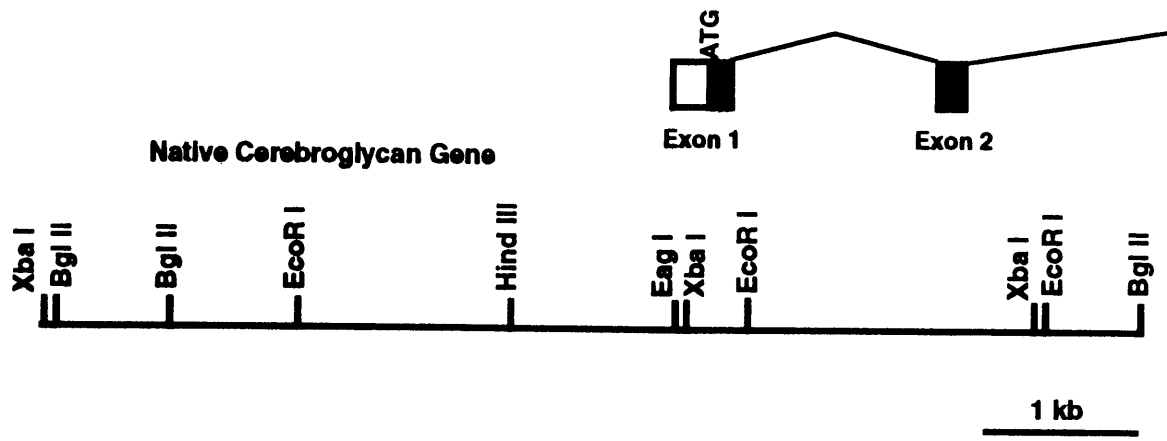
Figure 12. Sequence of cDNA from partially unspliced CBG mRNAs
Two of the cDNA clones isolated from the primer extended rat brain
cDNA library (see Materials and Methods) were eventually shown to be
complementary to partially unspliced CBG mRNAs. (A) The combined
sequence of the clones. Splice donors and splice acceptors are shown
in bold italics. Coding sequence is underlined, and the translation is
shown. The numbers refer to the position of the amino acids in the
CBG protein sequence as shown in Figure 5. (B) Comparison of the 3'
splice acceptor upstream from val244 to the 3' splice acceptor
consensus sequence (Mount, 1982). The invariant AG nucleotides are
bold. Of the 12 pyrimidines in the consensus sequence, 8 have been
substituted with purines in the endogenous splice acceptor.

A

GGTAATCCCACAAACGCAACCTCCACTCCAGTCCTGACTAACTGATAATCA
TTTTTGTGGGTTCGTTTTGTTTTGTTTTGTTTTTCTTTCTCAT*CTTCTTTC*
*CCTCCAG*ATAACCCGGGCACTGGTGGCAGCCCGGGCCTTGGTCCAGGGTCT
     217I T R A L V A A R A L V Q G .

<u>GGAGACCGGAAGAAATGTGGTCAGCGAAGCACTT*AAGGTTAGA*GGGGTCCT</u>
 L E T G R N V V S E A L K243

GAATATGGACAGAGCCCAGGAAGTAAAGAAGACAGACACCAGCTGGCAGGA
GAGGCCTAGAGCTGTCTCTTACTTGTTCCTGGGCTCTGGTGCTCCAGAGGA
CATCTGGAAGGGAGGCCAAAGTCAGGCCTCCACGGGGTGTCTGACATGATC
TCCTGGGATTTCTGGTCTTCCCTCTGGTGACGCACATGCCTTCTGCTAGTC
TTTCCATAGCAGTTCAAGATTGTAGTGAACCCCTCAGTGCCCCTCCCTCCT
CCCAAAAGGCCTCTCTTCATCACTGGTCCTGTCCCTAGCCTCACCTCTCAC
CTCCTTCCTCTCTCCTCAGCCTGGACTCACTCTTGCCCCTGAATCTTTCCT
CACTACATCCTCCTCAGTAACACCCTAGTCAAACCCCTCTCATGAAGCAGA
ATCAGCAACCTTTTCTCTAGGTACCCTCCTTCTCAGTCCTGTATCACCGGG
TGCCACATGGGGTCACCAGCCACAGTTGCTCATATCCAGCTTCCAAC*AGAT*
AGCCTGGGGAG<u>GTACCCATGTTGGAAGGCTGCAGACAGGCCCTCATGCGTC</u>
    244V P M L E G C R Q A L M R

<u>TGATCGGCTGCCCACTTTGTCGGGGAGTACCCTCGCTTATGCCCTGCCGGG</u>
 L I G C P L R G V P S L M P C R

<u>GCTTCTGCCTCAATGTGGCCCACGGCTGTCTCAGCAGCAGGGGATTGGAGC</u>
 G F C L N V A H G C L S S R G L E

<u>CTGAATGGGGCGGATATCTGGATGGTCTCCTGCTGCTGGCTGAGAAACTCC</u>
 P E W G G Y L D G L L L L A E K L

<u>AGGGACCCTTTTCCTTCGAGCTGGCTGCTGAGTCCATAGGGGTGAAGATCT</u>
 Q G P F S F E L A A E S I G V K I

<u>CAGAAGGTTTGATGCACTTGCAGGAAAACAGTGTAAAGGTGTCAGCCAAGG</u>
 S E G L M H L Q E N S V K S A K V

<u>TGTTTCAGGAATGTGGGACGCCTCA</u>....
 F Q E C G T P H350...

B

TTTTTTTTTTTT-T**AG**   CONSENSUS 3'
CCCCCCCCCCCC-C**AG**   SPLICE ACCEPTOR

AGATAGCCTGGGG**AG**   ENDOGENOUS 3'
                   SPLICE ACCEPTOR

Figure 13. Partial exon organization of cerebroglycan

The location and relative size of the CBG exons identified so far are shown schematically. Exons I and II were identified by sequencing of CBG mouse genomic clones, and exon q was identified by sequencing of fortuitously isolated, partially unspliced CBG rat cDNA fragments (see Figure 12). The exon structure of the remaining regions, shown by broken line, has not been determined. Structural features: shaded regions indicate putative signal sequence (in exon I) and GPI cleavage and addition signal (near carboxy terminus); arrows show the location of conserved cysteine residues (C); "lollipops" indicate the location of potential glycosaminoglycan (GAG) attachment sites. One potential ser- gly GAG attachment site is split between exons I and II and is not shown.

I   II   q
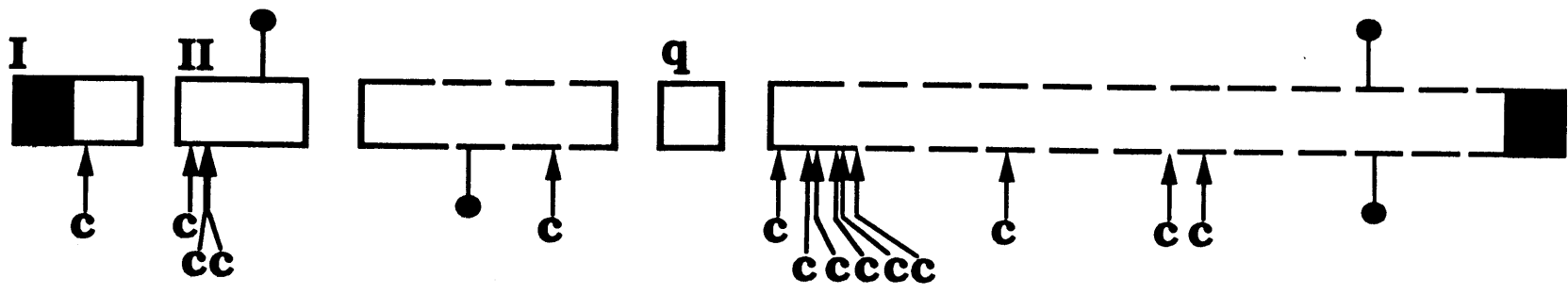
—— 30 amino acids

Figure 14. Single stranded conformational polymorphism in B6 and spretus mouse genomic CBG loci. Genomic CBG DNA fragments from B6 (C57B16, Mus musculus) and spretus (Mus spretus) mice were isolated by PCR (see text for details). Electrophoresis of the ^{32}P-dCTP-labelled fragments on a strand separating gel identified a single stranded conformational polymorphism (SSCP) in the B6 and spretus genomic CBG loci. The figure shows an example of a gel scoring the SSCP in genomic DNA samples from the test cross NCI-Frederick #77-122 (B6 X spretus) X B6. Numbers refer to DNA samples from different animals from the test cross; "**B**" is the B6 pattern, and "**S**" is the spretus pattern. Samples were scored as "**b**", the B6 pattern; "**h**" the heterozygous B6 X spretus pattern; or "**x**" ambiguous. Samples with a low yield in the PCR reaction were typically scored as ambiguous since the presence or absence of a unique band in the (B6 X spretus) pattern was the basis of the SSCP, and this band was difficult to discern in low yield samples. Subsequent rounds of SSCP analysis often resolved these amibiguities.
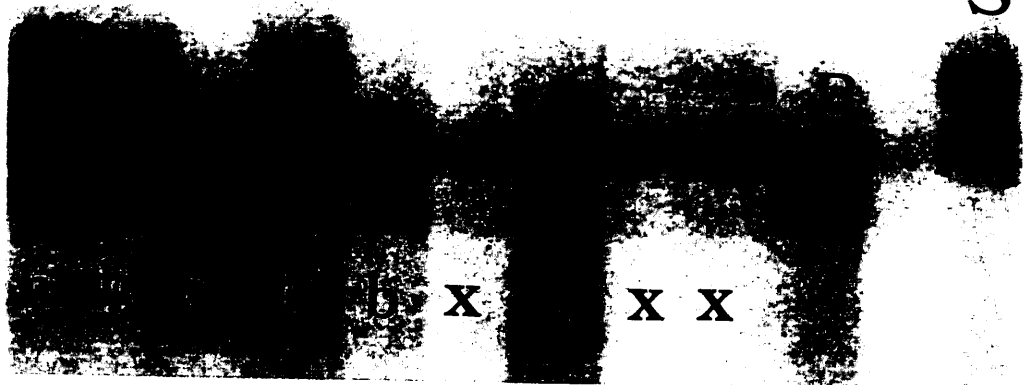
SAMPLE	SSCP	SAMPLE	SSCP
77	het.	100	B6
78	het.	101	het.
79	B6	102	n.c.
80	het.	103	het.
81	B6	104	het.
82	n.c.	105	B6
83	B6	106	het.
84	het.	107	n.c.
85	het.	108	n.c.
86	B6	109	n.c.
87	B6	110	n.c.
88	B6	111	n.c.
89	het.	112	B6
90	B6	113	het.
91	B6	114	B6
92	het.	115	B6
93	het.	116	het.
94	het.	117	het.
95	B6	118	B6
96	het.	119	het.
97	het.	120	het.
98	B6	121	B6
99	het.	122	B6

Table 2 Genomic Mapping of Cerebroglycan
A single stranded conformational polymorphism (SSCP) in the cerebroglycan loci of B6 and Spretus mice was used to map cerebroglycan in the mouse genome. PCR reactions were performed with genomic DNA of animals from the NCI-Frederick test cross (#77-122), a (B6 X Spretus) X B6 test cross. α^{32}P-dCTP labelled PCR products were analyzed on a strand separating gel, and the SSCP was scored either as B6 or heterozygous (het.); (n.c.) = no call possible, ambiguous. The data shown above was analyzed with the program, Mapmaker, resulting in the assignment of cerebroglycan to the distal end of mouse chromosome 5 (see text for details).

CHAPTER 3

EXPRESSION OF CEREBROGLYCAN mRNA AND PROTEIN IN THE DEVELOPING NERVOUS SYSTEM

INTRODUCTION

The major stages of neural development -- neurogenesis, cell body migration, axon outgrowth, and synaptogenesis -- all have the potential to be regulated by cell surface heparan sulfate proteoglycans (HSPGs). For example, FGF-1 and FGF-2 are expressed at several locations in the developing nervous system, including the ventricular zone and cortical plate of the developing rat cerebral cortex (Gonzalez et al., 1990; Fu et al., 1991; Wilcox and Unnerstall, 1991). The activity of these molecules, which have been shown to stimulate proliferation and survival of neural cells in vitro (Walicke et al., 1986; Unsicker et al., 1987; Walicke, 1988; Murphy et al., 1990; Hughes et al., 1993), appears to require or at least to be greatly potentiated by the presence of cell surface heparan sulfate (Rapraeger et al., 1991; Yayon et al., 1991; Olwin and Rapraeger, 1992; Ornitz and Leder, 1992; Kan et al., 1993). Immature cortical neurons migrate from the ventricular zone to the cortical plate along radial glial fibers. The extracellular matrix molecule fibronectin (FN), which is found closely associated with these fibers, (Sheppard et al., 1991) is known to bind to heparin and HSPGs, and a recent study has shown that neural cells can use a cell surface HSPG to attach and spread on a peptide from the FN heparin-binding domain (Haugen et al., 1992). Many cortical neurons subsequently extend axons through the underlying layers containing FN, laminin-1 (LN-1), and laminin-3/S-laminin (LN-3) (Sheppard et al., 1991; Hunter et al., 1992), all of which have heparin-binding domains, and LN-3 is present at the synaptic cleft of neuromuscular junctions (Sanes et al., 1990), and appears to be important for formation and maintenance of proper synaptic structure (Noakes et al., 1995; Porter et al., 1995). Thus there are many potential ligands and many potential roles for cell surface HSPGs in the development of the nervous system.

Evidence that individual HSPG core proteins can participate in distinct developmental events has come from studies of expression of the syndecan family of HSPGs in non-neural tissues. For example, in several embryonic epithelial structures, syndecan-1 expression is lost during morphogenesis and re-expressed in the mature, stable epithelium. Conversely, the mesenchymal cells underlying the

embryonic epithelia are initially negative for syndecan-1, but become syndecan-positive when they proliferate and condense around epithelium undergoing morphogenesis, only to lose expression again upon terminal differentiation (Trautman et al., 1991; Bernfield et al., 1992) Thus syndecan-1 may play two distinct roles during epithelial morphogenesis: its expression in epithelial cells is consistent with a role in the maintenance of a stable epithelial structure, while its expression in mesenchymal cells correlates with proliferation and differentiation. Studies demonstrating that syndecan-1 acts as a receptor for extracellular matrix molecules and, potentially, for growth factors such as FGF-2, point to possible mechanisms of syndecan function during development (Saunders and Bernfield, 1988; Kiefer et al., 1990; Bernfield and Hooper, 1991; Sanderson et al., 1992a).

Comparison of the patterns of syndecan-1 and syndecan-3 expression during development suggests an additional level of specificity in the roles of the syndecans during embryogenesis. Both molecules are expressed in distinct temporal and spatial patterns, that are often reciprocal. For example, syndecan-1 is expressed in the early neurectoderm (Sutherland et al., 1991), but then disappears, while syndecan-3, is expressed later in the embryonic central nervous system, including in the floor plate of the developing spinal cord (Gould et al., 1995), and becomes highly expressed in the brain around the day of birth, decreasing again in the adult brain (Carey et al., 1992). Thus the different members of the syndecan family may fulfill distinct roles during the development of a single tissue type.

Recently, the glycosylphosphatidylinositol (GPI)-anchored HSPG glypican has also been shown to be expressed in a restricted pattern that varies during development. In the developing bone, for example, glypican appears to be expressed in differentiated osteoblasts, the cells that deposit the bone matrix replacing the cartilaginous model during osteogenesis. However, chondrocytes in the developing bone do not appear to express glypican (Litwack, 1995).

In the early rat nervous system, glypican is initially highly expressed in the ventricular zones containing proliferating neuronal precursor cells (Litwack, 1995). At later stages, glypican expression is

widespread in the developing nervous system, but by adulthood, glypican expression has become restricted in the brain (Litwack et al., 1994). Many adult rat brain structures that retain glypican expression are populated by projection neurons (e.g. the dorsal thalamus, and layers II, III, V, and VI of the cerebral cortex) (Litwack et al., 1994).

Preliminary experiments with whole brain extracts indicated that cerebroglycan (CBG) protein and mRNA expression are limited to development. To determine whether CBG's expression pattern correlates with any of the known events in nervous system development, we performed CBG in situ hybridization and immunolocalization experiments on rat embryos of several developmental stages. Surprisingly, CBG expression was found to be restricted to the nervous system, making it the first tissue specific PG to be identified. Close examination of CBG mRNA expression in several neural structures strongly suggests that CBG is transiently expressed by immature neurons, appearing around the time of their final mitosis and vanishing after cell body migration and axon outgrowth have been completed. Immunolocalization experiments revealed that the CBG protein is particularly abundant in the axon tracts of the developing rat brain.

MATERIALS AND METHODS

In Situ Hybridization Protocols

In situ hybridization to rat embryos and neonates

Freshly dissected rat embryos were quick-frozen in isopentane on dry ice and stored at -80°C. P0 rats were anaesthetized on ice and decapitated; the heads were quick-frozen in isopentane on dry ice and stored at -80°C. P7 rats were anaesthetized with metofane and decapitated. The brains were dissected out and quick-frozen as described above. Prior to sectioning, samples were equilibrated to -20°C and embedded in OCT medium. 12-20 μm cryostat sections were collected on gelatin-subbed slides or on Probe-On Plus slides (Fisher Biotech; Pittsburgh, PA) and stored at -80°C with dessicant. Sections were equilibrated to room temperature, fixed in 4% paraformaldehyde in PBS at room temperature for 5 min, and washed twice in 2X SSC.

Riboprobes were synthesized using an RNA Transcription Kit (Stratagene, Cat. # 200340). T7 or T3 RNA polymerase was used in reactions with ^{35}S-UTP (DuPont, NEN), and linearized CBG templates. The templates consisted of pBluescript subclones of the 301 bp SacI fragment or the 223 bp ApaI/KpnI fragment of the M13-A CBG cDNA. Probes made with the SacI fragment (from CBG coding region) and the ApaI/KpnI fragment (from the 3' untranslated region) gave identical results.

Prehybridization treatments, hybridization, and post-hybridization washes were performed according to Simmons et al. (Simmons et al., 1989), except that the proteinase K step was omitted. After hybridization, sections were exposed to hyperfilm βmax (Amersham; Arlington Heights, IL) for 5 to 7 days at -80°C. Selected sections were dipped in Kodak NTB-2 liquid emulsion that had been diluted 1:1 in ddH$_2$O, and exposed for 12-21 days at -80°C. Developed sections were stained in Hoechst 33258 (bisbenzimide) at 1 μg/ml in ddH$_2$O or cresyl violet (0.5% in ddH$_2$O), and examined by fluorescence, bright field, and dark field illumination on a Zeiss Axiophot microscope, or a Nikon Diaphot microscope.

In situ hybridization to mouse olfactory epithelium(OE) and OE explants

Freshly dissected embryonic day 15 mouse embryos were decapitated, and the heads were fixed overnight at 4°C in freshly prepared 4% paraformaldehyde in phosphate buffered saline (PBS). To cryoprotect, the fixed heads were transferred to a fresh tube with 5% sucrose in PBS and stored overnight at 4°C. The heads were then transferred to 15% sucrose in PBS and stored overnight at 4°C. Finally, the heads were transferred to fresh 15% sucrose in PBS and stored at 4°C until sectioning.

Cryoprotected heads were gradually frozen in OCT mounting medium and horizontal cryostat sections 20 μ thick were prepared and collected on ProbeOn Plus slides (Fisher Biotech). Sections were stored at -85°C with dessicant until in situ hybridization.

Sections were thawed in an EZ Breeze air dryer (Hoeffer) without heat until dry, and post-fixed for 2 hours in fresh 4% paraformaldehyde at 4°C. After three 5 minute washes in PBS with 0.1% Tween-20 (PBT), sections were bleached for 90 minutes at room temperature in 5:1 (V/V) methanol (MeOH): 30% H_2O_2, and stored overnight at -20°C in absolute methanol.

Sections were rehydrated by washing 5 minutes each in (1) 75% MeOH/25% PBT, (2) 50% MeOH/50% PBT, (3) 25% MeOH/75% PBT, (4) PBT, and (5) PBT. Sections were permeabilized by treating with proteinase K (Boehringer Mannheim) at 10 μg/ml in PBT for 5 minutes at room temperature. Sections were next fixed for 5 minutes at room temperature with 4% paraformaldehyde, 0.2% glutaraldehyde in PBS, rinsed twice with PBT, and treated with 0.1% fresh sodium borohydride in PBT for 20 minutes at room temperature.

Sections were pre-hybridized in hybridization buffer (HB): 50% formamide, 5X SSC pH 4.5, 50 μg/ml yeast RNA, 1% SDS, 50 μg/ml heparin, for 90 minutes at 60°C. Digoxigenin (DIG)-labelled riboprobes were applied in HB at a concentration of 1 μg/ml, and sections were coverslipped and incubated overnight at 60°C in a humid oven. Probes were synthesized from linearized templates consisting of the muCBG-1 mouse CBG cDNA fragment (Figure 9 Chapter 2), subcloned into pCR1000 (Invitrogen) or pBluescript

144

(Stratagene). The Genius RNA Labeling Kit (Boehringer Mannheim) was used to synthesize the probes and estimate probe concentrations.

After hybridization, non-specific probe was removed by the following treatments: (1) two 15 minute washes at 60°C in 50% formamide, 5X SSC pH 4.5, 1% SDS; (2) two 5 minute washes in 0.5 M NaCl, 10 mM Tris-HCl pH 7.5, 0.1% Tween-20; (3) a 40 minute treatment 100 μg/ml RNase A in the same buffer as (2) above at 37°C; (4) a 5 minute wash in Tris buffered saline (pH 7.4) with 0.5% Tween-20 (TBST) at room temperature, and a 60 minute wash in 50% formamide, 2X SSC pH 4.5 at 60°C.

Sections were prepared for DIG immunohistochemistry by three 5 minute washes in TBST, followed by a 60 minute blocking step in TBST with 10% heat-inactivated normal goat serum (HI-NGS). Alkaline phosphatase conjugated, goat anti-DIG antibody Fab fragments (Boehringer Mannheim) were pre-adsorbed with 10 mg/ml mouse brain acetone powder (Sigma) in 1% HI-NGS, 2 mM levamisole (Sigma), in TBST. The pre-adsorbed Fab fragments were applied to the sections at a final concentration of 0.3 μg/ml in TBST, 1% HI-NGS, with 2 mM levamisole and allowed to bind overnight at 4°C.

Hybridized probe was detected by (1) washing five times, 10 minutes each in TBST with 2 mM levamisole; (2) washing three times, 10 minutes each in 0.1 M NaCl, 0.1 M Tris-HCl (pH 9.5), 50 mM MgCl$_2$, 1% Tween-20, and 2 mM levamisole (NTMT); (3) incubating with NBT/BCIP (Nonradioactive Nucleic Acid Detection Kit, Boehringer Mannheim) in NTMT with 2 mM levamisole until color had developed.

For postnatal day 15 mouse, heads were dissected, the lower jaws were removed, and the remainder was fixed as described above. Muscle and bone overlying the OE were partially removed by further dissection, and the heads were cryoprotected and processed for in situ hybridization essentially as described above.

For OE explants, OE was isolated as described below from embryonic day 14/15 mice, and plated on merosin coated coverslips. At set times after plating, coverslips were fixed and processed for in situ hybridization essentially as described above. For NCAM riboprobes, a 391 bp EcoRI/Bcl I fragment from the coding region of

the mouse NCAM cDNA was subcloned into the EcoRI and BamHI cloning sites of pGEM. DIG-labelled NCAM riboprobes were synthesized from linearized templates as described above, using SP6 or T7 RNA polymerase.

Tissue Culture Protocols

Olfactory epithelium explant culture

OE was dissected from E14/E15 mouse animals and purified as previously described (Calof and Chikaraishi, 1989). Finely chopped pieces of OE were cultured on laminin or merosin coated coverslips in serum-free, low calcium (0.1 mM Ca^{2+}) Dulbecco's modified Eagle's medium supplemented with 5 mg/ml BSA, 10 μg/ml insulin, 10 μg/ml transferrin, 20 nM progesterone, 100 μM putrescine, and 30 nM selenium, as described (Calof and Chikaraishi, 1989; Calof et al., 1991a).

Establishment of Ψ2, 3T3 cell and PC12 cell CBG expression lines

Ψ2 cells were maintained at 8% CO_2 in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% iron-supplemented calf serum (FeSCS), glutamine, and penicillin/streptomycin. A CBG expression vector, pLNCX-CBGΔ3', was constructed by subcloning the 1835 bp Eag I/Nar I restriction fragment of the full length CBG cDNA into the cloning site of the pLNCX retroviral expression vector (Miller and Rosman, 1989)(see Figure 9 for the cloning strategy). This fragment is missing most of the 3' untranslated region of the CBG cDNA. pLNCX uses a CMV promoter to drive expression of the transgene, and the 5' LTR to drive expression of a neomycin resistance gene.

Ψ2 producer cell lines were made by transfecting Ψ2 cells with pLNCX-CBGΔ3' using the lipofectamine reagent (Gibco BRL) according to manufacturers directions. Conditioned medium from the transfected Ψ2 cells was used to infect Ψ2 cells treated with 200 ng/ml tunicamycin and 8 μg/ml polybrene, creating stable, virus producing, infected cell lines. Individual lines were selected by growth in 1mg/ml G418 (Gibco BRL), and isolation of individual colonies with cloning rings.

3T3 cells were maintained identically to Ψ2 cells. 3T3 cell lines expressing CBG were generated by infecting 3T3 cells with conditioned medium, containing 8 μg/ml polybrene, from Ψ2 producer cell lines. Infected cells were selected by growth in 1 mg/ml G418 and maintained as an uncloned population by growth in 0.5 mg/ml G418. This population of cells was designated 3T3inf$_u$CBGΔ3'

PC12 cells were grown at 8% CO$_2$ in RPMI with 10% heat-inactivated horse serum (JRH), 5% fetal calf serum, glutamine, and penicillin/streptomycin. PC12 cells were grown on Vitrogen (Celltrix), coated plates (1:100 dilution of Vitrogen in PBS to coat plates for ≥1 hour at 37°C). Cells expressing CBG were generated as with 3T3 cells, and maintained as an uncloned population (PC12inf$_u$CBGΔ3')

CBG Immunohistochemistry and Immunofluorescence

A rabbit polyclonal antiserum was raised against a peptide (CRPPRPPRPPRRDGL) based on amino acids 521-535 of the CBG protein sequence as shown in Figure 5, Chapter 2. The peptide (designated "521") was coupled to keyhole limpet hemocyanin (KLH) (Pierce) with the cross-linker sulfo-SMCC (Pierce). Rabbits were injected intradermally with 2.5 mg of the KLH-521 in complete Freund's adjuvant and boosted two times intramuscularly with an additional 2.5 mg in Freund's incomplete adjuvant. Anti-521 antibodies were purified from serum using an affinity column of peptide 521 coupled to Sulfo-link gel (Pierce). This antibody was designated 521-2.

E13.5 mouse embryos, E13 rat embryos, and E19 rat heads were processed for cryosectioning essentially as described above for OE in situ hybridization experiments. For immunohistochemistry, sections were incubated in block (2% BSA, 100 mM Tris-HCl pH 7.4, 0.15 M NaCl, 0.3% TritonX-100), and then washed twice in Tris-buffered saline (TBS), twice for 30 minutes in 0.3% H$_2$O$_2$, and once more in TBS. 521-2 was applied at 2.5 μg/ml in block and allowed to bind overnight at 4°C. After washing in TBS, a biotin-conjugated goat-anti rabbit secondary antibody (Vector) was applied at 1:300 in block and

allowed to bind 4 hours to overnight at 4°C. Sections were washed in TBS and incubated in ABC reagent (Vector) diluted in block, washed again and then developed in 50 μg/ml DAB, 0.3% H_2O_2, 50 mM Tris pH 7.4. Sections were washed in H_2O, dehydrated through xylenes, and coverslipped in Permount. Some sections were counterstained in methyl green.

3T3inf$_u$CBGΔ3' cells and PC12inf$_u$Δ3' cells were grown on LN or FN-coated coverslips (proteins were coated at 20 μg/ml in PBS overnight at 4°C.). Cells were rinsed once in pre-warmed DMEM or RPMI and then fixed for 10 minutes at room temperature in 4% paraformaldehyde. After 20 minutes in block (as above), 521-2 was applied at 5 μg/ml in block and allowed to bind for 60 minutes at room temperature. Coverslips were rinsed 3 times with TBS, and a Cy-3 conjugated goat-anti-rabbit secondary antibody (Jackson Immunoresearch) was applied in block at a 1:100 dilution. Hoechst 33258 was added with the secondary antibody at 2.5 μg/ml to stain nuclei. After a 40 minute incubation at room temperature, coverslips were washed 3 times in TBS and mounted in Aquamount.

Isolation of 3T3 Cell PG Fractions

Freshly confluent plates of 3T3 cells and 3T3inf$_u$CBGΔ3' cells were harvested by scraping into homogenization buffer A (see Chapter 2 Materials and Methods, Purification and protein sequencing of P0 rat brain HSPGs). Membrane-associated PGs were isolated essentially as described in Chapter 2. PGs were re-bound to DEAE-Spectragel (Pharmacia) and iodinated on the beads as previously described (Herndon and Lander, 1990). ^{125}I-labelled PGs were digested with chondroitinase ABC and/or heparitinase as described (Herndon and Lander, 1990) and analyzed on SDS-PAGE minigels.

For immunoprecipitation experiments, 3.9×10^6 cpm aliquots of ^{125}I-labelled PGs from 3T3 cells and 3T3inf$_u$CBGΔ3' cells were either treated with chondroitinase ABC and heparitinase, or were treated with neither enzyme in a mock digest. The PGs, in a volume of 75 μl, were diluted to 500 μl in TBS with 0.5% CHAPS, along with 20 μl (packed volume) of protein A sepharose beads (Sigma). The PGs were precleared for 40 minutes with rotation at 4°C. After centrifugation

to pellet the beads, the supernatants were transferred to a fresh tube along with 6.75 µg of affinity purified 521-2 anti-CBG antibody. The immune complexes were collected with another 20 µl of protein A sepharose beads after an overnight binding step at 4°C. The beads were washed with (1) TBS, 0.5% CHAPS for 30 minutes at room temperature with rotation, (2) 0.5M NaCl, 50 mM Tris-HCl (pH 7.4), 0.5% CHAPS for 20 minutes at room temperature, and (3) 50 mM Tris-HCl (pH7.4), 0.5% CHAPS briefly at room temperature. The beads were counted in a γ-counter, demonstrating that while an average of 16,000 cpm had been collected in immunoprecipitates from parental 3T3 cells, an average of 375,000 cpm had been collected immunoprecipitates from 3T3inf$_U$CBGΔ3' cells. The protein A-sepharose beads from the 3T3inf$_U$CBGΔ3' immunoprecipitations were boiled in SDS-PAGE loading buffer and the supernatants analyzed on SDS-PAGE minigels along with the diagnostic digests of labelled PGs from the two cell lines.

RESULTS

Developmental expression of cerebroglycan

The CBG core protein is observed in the membrane fraction of embryonic day 18 (E18) and newborn (P0) rat brain, but not in the adult (Herndon and Lander, 1990). As shown in Chapter 2, the CBG mRNA is expressed in a similar pattern. To localize CBG expression in the developing animal, *in situ* hybridization was performed. The 306 bp SacI or the 223 bp ApaI-KpnI pBluescript subclones of the CBG cDNA, M13-A (Figure 4, Chapter 2), were used to generate α^{35}S-UTP labelled RNA probes. Figure 1 shows the results obtained when these probes were hybridized to fresh-frozen sections of embryonic day 14 (E14), 16 (E16), and 19 (E19), and newborn (P0) animals. The results are presented as "reverse contrast" images, printed directly from autoradiograms of the sections.

In E14 and E16 animals (Panels B and D of Figure 1), strong CBG expression occurs along the entire neuraxis, including the brain, spinal cord, dorsal root ganglia, and cranial nerve ganglia. No expression in any non-neural tissue could be detected. By E19 (Figure 1, panel F), expression in the spinal cord and dorsal root ganglia is dramatically lower than in the many regions of the brain that still express high levels of CBG message, including the cerebral cortex, hippocampus, corpus striatum, inferior colliculus, and pontine nuclei. As at earlier stages, hybridization to non-neural tissues was not above background. In the P0 head (Figure 1, panel I), expression persists in the cerebral cortex, hippocampus, striatum, inferior colliculus, and cerebellar cortex, but only weak expression if any is seen elsewhere in the brain. Again, expression was not observed outside the nervous system. Panel G shows an example of the results obtained when the opposite strand control probe was hybridized to a nearly adjacent section (in this case, in the E19 animal). No hybridization above background was observed in any tissue.

As Figure 1 illustrates, CBG is globally expressed in the nervous system at early stages. As neural development proceeds, expression is selectively lost from certain structures while remaining in others. A closer examination of the pattern of expression suggested that CBG mRNA remains in structures for several days after neurogenesis--the

production of post-mitotic neurons--has ended (see Discussion). This raised the possibility that CBG expression might be associated with neuronal precursors and/or immature neurons.

To more accurately examine the relationship between CBG expression and neuronal development, *in situ* hybridization was carried out using sections of postnatal day 7 (P7) brain. At this age, the generation and early development of neurons has finished in most of the brain, but persists in three late-developing areas: the external granule layer of the cerebellum, the dentate gyrus of the hippocampus, and interneurons of the olfactory bulb (Jacobson, 1991). As shown in Figure 2, CBG is expressed at this stage in precisely these three areas: the folia of the cerebellum, the dentate gyrus of the hippocampus, and the olfactory bulb. As with the earlier experiments, no hybridization with the opposite-strand control probe was detected (not shown).

A higher magnification view of the cerebellum (Figure 3A) reveals that CBG-expressing cells are found in the external granule layer (EGL), the zone in which granule cells (cerebellar interneurons) are generated and undergo intitial stages (~28 hours) of development (Fujita, 1967). In contrast, the internal granule layer (IGL), which contains mature granule cells, was essentially devoid of hybridization. A higher magnification view of the olfactory bulb (not shown) revealed that CBG-expressing cells are localized to the subventricular zone, the site of postnatal production of olfactory bulb interneurons [granule cells and periglomerular cells (Luskin, 1993)]. In contrast, layers containing mature olfactory bulb neurons were devoid of CBG hybridization (these layers lie anterior to the indicated region of hybridization in Figure 2, and were visualized by Hoechst 33258 fluorescence [not shown]).

These results suggest that, at least in these tissues, expression of CBG is restricted to an early stage in the development of neurons. An indication of how early CBG appears can be obtained by examining the distribution of cells within the external granule layer of the cerebellum. As shown in figure 3A, CBG-expressing cells are strongly concentrated within the deeper half of the external granule layer, the so-called premigratory zone (Fujita, 1967). This zone consists of newly generated, post-mitotic neurons, whereas the superficial half of

the external granule layer contains the proliferating cells that give rise to those neurons. This result indicates that CBG expression begins in neurons at or about the time they become post-mitotic.

To further test this hypothesis, *in situ* hybridization was used to examine CBG expression in the embryonic day 12 (E12) spinal cord. At this stage, the spinal cord consists of a large central proliferative zone (the ventricular zone), from which newly post-mitotic neurons migrate laterally to form the nascent ventral and dorsal horns (Altman and Bayer, 1984). As Figure 3B illustrates, CBG is strongly expressed by regions of the E12 spinal cord that contain post-mitotic neurons, but weakly expressed, if at all, by the still-proliferating neural precursors of the ventricular zone. In addition, expression can be seen in the dorsal root ganglia (DRG) located on either side of the spinal cord. These structures also contain recently post-mitotic neurons (Altman and Bayer, 1984), the progeny, in this case, of neural crest cells.

Immunolocalization of CBG protein in the developing nervous system

CBG in situ hybridization experiments in the P7 cerebellum showed that, in that structure, CBG mRNA expression colocalizes with a population of neurons that are undergoing axon outgrowth and cell body migration. In the P0 cortical plate, which is also CBG positive, neuronal cell bodies are, as a population, in their final locations. However, the process of axon outgrowth by these cells is still incomplete at this stage. Thus the question of whether CBG expression correlates best with cell body migration, axon outgrowth, or both, remains open. To gain a more complete picture of CBG expression in the developing nervous system, the CBG protein was localized in sections from embryos of several stages using an affinity purified, polyclonal antibody, "521-2," raised against a CBG peptide (see Materials and Methods; this antibody has previously been shown to specifically recognize the CBG core protein in Western blots of a crude rat brain membrane preparation (Litwack, 1995).) Figure 4 shows some examples of these experiments.

As was seen by in situ hybridization, CBG immunolocalization demonstrates that the CBG protein is nervous system-specific. In

Figure 4A, a saggital section of an E19 rat head, only the brain is stained with the 521-2 anti-CBG antibody. Additionally, immunolocalization revealed that CBG expression is largely associated with axon tracts. In Figure 4A, the most strongly stained structure is the intermediate zone, a region containing cortical and thalamic axons. Strong CBG expression is also seen associated with the axons of the optic tract. In a coronal section from a P0 rat brain (Figure 4B), CBG expression is observed in the internal capsule, which contains cortical and thalamic axons, and in the tract containing the hippocampal commisural fibers. In the E13 rat spinal cord (Figure 4C), CBG localizes to the commisural axon tracts on the lateral surfaces of the spinal cord, and in the ventral root, which contains the axons of the spinal cord motoneurons.

Cerebroglycan expression in the olfactory epithelium

The data from *in situ* hybridization strongly suggest that CBG mRNA appears in neurons around the time of their final mitosis, raising the possibility that CBG may be involved in facilitating the transition from proliferating neuronal precursor to differentiating neuron, and/or in the early steps of neuronal differentiation. Immunolocalization of the CBG protein is consistent with a role for CBG in the growth of axons.

One structure where the process of neurogenesis and neuronal differentiation has been well characterized is the olfactory epithelium (OE). The OE has the unusual feature of being a neural structure in which neurogenesis and neuronal differentiation continue even in adult animals (Graziadei and Graziadei, 1978). To gain additional insight into the timing of CBG expression relative to neurogenesis, and into potential roles for CBG in neuronal differentiation, CBG expression in the OE was examined.

As a first step in characterizing CBG expression in the OE, a CBG *in situ* hybridization was performed on sections of developing and mature OE. At embryonic day 15 (E15) in the mouse, the OE is relatively unstructured, and neurogenesis is occurring across the width of the entire epithelium (Smart, 1971). The CBG mRNA is expressed throughout the E15 OE, although not as highly as in some other neural structures such as the retinal neuroepithelium (Figure 5).

By postnatal day 15, the OE has achieved an essentially adult organization (Figure 6A). In the mature OE, NCAM positive olfactory receptor neurons lie over a thin layer of keratin positive, NCAM negative basal, epithelial cells (Calof and Chikaraishi, 1989). The NCAM positive neurons can be further subdivided into immature neurons (in the process of axon outgrowth), and mature neurons (in contact with targets in the olfactory bulb). GAP-43, a neuron-specific marker whose expression correlates with neuronal development (Benowitz and Routtenberg, 1987), is present on immature but not mature neurons, and OMP (olfactor marker protein) is only present on mature olfactory neurons (Verhaagen et al., 1989). CBG mRNA localizes to near the basal surface of the OE in the postnatal day 15 mouse (Figure 6B). This region corresponds to the area where olfactory receptor neurons are born and begin differentiation. Thus, CBG expression in the OE appears to be consistent with the expression pattern that had been observed elsewhere (e.g. the P7 cerebellum) in that CBG mRNA was found in a zone that contains immature neurons in the early stages of differentiation.

An in vitro model of OE development is available that recapitulates many of the essential features observed in the OE in vivo (Calof and Chikaraishi, 1989; Calof et al., 1991a). In the model, pieces of purified E14/15 mouse olfactory epithelium are cultured as explants on laminin (LN) or merosin (MN) coated coverslips (see Figure 7). Immature neuronal precursors (INPs) sort out from the explant, and, under proper conditions, a substantial fraction migrate away from the explant onto the LN or MN substrate. Typically, the INPs divide once to give rise to two NCAM-positive, immature neurons (Calof and Chikaraishi, 1989). Subsequently, these neurons extend neurites, and depending on the culture conditions, may even express markers indicative of mature olfactory receptor neurons (Calof et al., 1993).

To further characterize the timing of CBG expression with respect to neurogenesis, CBG in situ hybridizations were performed on OE explants. For comparison, hybridization with a probe for mouse NCAM was also performed.

By six hours in culture, substantial cell migration has occurred in the explant cultures. The NCAM probe detected several NCAM-

positive cells on the substrate away from the explant (Figure 8, A and B), indicating that, as had been observed previously, many of the migrating cells were INPs, which divide to give rise to NCAM-positive, immature neurons. The NCAM probe also detected substantial numbers of positive cells at the perimeter of the explants; this is the normal consequence of the fact that, in any given culture, as many as 50% of the INPs remain near the explant boundary after sorting out. When these explant-associated INPs are dissociated with a mild trypsin treatment and plated on fresh coverslips, they differentiate in a manner indistinguishable from that of the INPs that migrate (Anne Calof, unpublished observations).

In contrast, the CBG probe never detected CBG-positive cells on the substrate away from the explant (Figure 8, C and D). The opposite strand control probe gave little background hybridization (Figure 8, E and F). Significant CBG expression was, however, detected associated with the explant. As shown in Figure 8 (G and H), large numbers of the cells that sorted out from the basal epithelial cell ball, but that remained at the lateral surfaces of the explant, showed substantial hybridization to the CBG probe. The negative control probe showed only a low level of background hybridization to cells at the explant boundary (data not shown). After 26 hours, strong NCAM expression is still detected throughout the culture, but CBG expression has largely disappeared (not shown). Thus, although the INPs around the explant and the INPs on the substrate both differentiate into NCAM-positive neurons, only those associated with the explant were observed to express CBG at detectable levels (See Discussion).

Cerebroglycan expression in cell lines infected with a CBG retroviral expression vector

Several questions about CBG expression and structure might best be addressed by cell culture experiments. For example, although several PG core proteins (e.g. syndecan-1 (Rapraeger et al., 1985), syndecan-4 (Shworak et al., 1994a), and serglycin (Seldin et al., 1985)) are known to attach both heparan sulfate (HeS) and chondroitin sulfate GAG chains, members of the glypican family of PGs appear to be substituted only with HeS. The (presumably) structural features of the glypican

family core proteins that are responsible for the specifity for HeS attachment have yet to worked out.

Additionally, in several instances, glycosylphosphatidylinositol (GPI)-anchored molecules have been found to be sorted to distinct domains of the cell membrane (Dodd et al., 1988; Dotti et al., 1991; Faivre-Sarrailh et al., 1992), including caveolae, protein microdomains that are involved in cellular uptake of small molecules (Anderson et al., 1992), and that contain several cytoplasmically oriented signalling molecules including members of the src family of protein tyrosine kinases (Sargiacomo et al., 1993; Lisanti et al., 1994). It would be of interest to determine if CBG, as a GPI-linked HSPG, is sorted preferentially to a particular membrane domain.

To address the question of whether the CBG cDNA obtained in Chapter 2 is in fact expressed as an HSPG in cell lines, NIH 3T3 cells and PC12 cells were infected with a CBG retroviral expression vector. This vector, LCNX-CBGΔ3' (see Figure 9), contains the CBG cDNA minus most of its 3' untranslated region, inserted downstream of the cytomegalovirus (CMV) promoter. LNCX-CBGΔ3' also contains a neomycin resistance gene under the control of the retroviral 5' LTR promoter.

The infected 3T3 and PC12 cells, maintained as an uncloned, neomycin-resistant population, were tested for CBG expression with the 521-2 affinity purified, anti-CBG polyclonal antibody. Binding of 521-2 was detected with a Cy-3 conjugated fluorescent secondary antibody. High levels of expression were observed in both cell types, as shown in Figure 10 (A and C). Controls omitting the 521-2 antibody, and parental cell lines gave a low background of non-specific binding (Figure 10, B and D).

To obtain a preliminary assessment of CBG localization within CBG-expressing cell lines, cells were examined at high power on a confocal microscope with the aperture nearly closed to produce the thinnest possible optical sections. For CBG-expressing 3T3 cells (3T3inf$_{U}$CBGΔ3') plated on laminin (LN), strong CBG signal, typically with a punctate quality, was often detected at the level of cellular attachment to the substrate (Figure 11A). The middle section of the cells showed reduced signal (Figure 11B), and high levels of punctate

signal were again detected at the dorsal cell surface (Figure 11C). The relatively higher levels of signal at the levels of the dorsal and ventral cells surfaces was suggestive of a cell surface localization of CBG in these cells.

3T3inf$_u$CBGΔ3' cells plated on fibronectin (FN), were much more spread, and appeared to be much flatter when examined on the confocal microscope. Consequently, it was not possible to estimate CBG localization in these cells. Interestingly, these cells, as an overall population, appeared to be more uniformly positive for CBG (not shown), and signal in individual cells was detected throughout the entire cell, even in the space occupied by the nucleus (as determined by Hoechst fluorescence) (Figure 11, D-F).

CBG-expressing PC12 cells (PC12inf$_u$CBGΔ3') plated on LN showed comparable signal at the substrate level, and an intermediate level, but little signal at the dorsal surface, over the nuclei. (Figure 11, G-I). It was difficult to discern whether the signal at the substrate level in these cells was truly associated with the ventral cell surface, or was merely bleed-through from the intermediate levels above. However, the 521-2 antibody did detect CBG expression in the numerous short processes and micro-extensions of PCinf$_u$CBGΔ3' cells (see for example, Figure 11G, and Figure 10C).

To determine whether the CBG cDNA is expressed as an HSPG, PGs were prepared from the membrane fraction of 3T3inf$_u$CBGΔ3' cells by the method described in Chapter 2. The PG profiles of these two cell lines were not identical; however it is possible to discern an HSPG core protein, similar in size to the CBG core, that is present in 3T3inf$_u$CBGΔ3' cells but not in the parental cell line (Figure 12). Immunoprecipitations of CBG from 3T3inf$_u$CBG3' cells, although contaminated with some other PGs, contain a prominent PG core protein of appropriate molecular weight that comigrates with the novel HSPG core in the preparations of 3T3inf$_u$CBGΔ3' cell PGs, and that is absent from PGs prepared from parental 3T3 cells (Figure 12).

157

DISCUSSION

Cerebroglycan expression marks a distinct stage in neuronal development

Examination of patterns of cerebroglycan message expression by *in situ* hybridization suggests that cerebroglycan mRNA *1)* is produced by neurons, *2)* appears at around the time of a neuron's terminal mitosis, and *3)* disappears by the time that neuronal migration and axonal extension have been completed.

Strong support for this interpretation comes from examination of the developing cerebellum, in which neuronal birth, migration and maturation occur in distinct anatomical layers (Sidman and Rakic, 1973). In that structure, cells that proliferate in the superficial part of the external granule layer give rise to post-mitotic neurons that reside in the deep part of the same layer, and these in turn migrate rapidly through the molecular layer to reach their final positions in the internal granule layer. The concentration of cerebroglycan mRNA in the deep half of the external granule layer (Figure 3A) implies that cerebroglycan begins to be expressed around the final mitosis, while its absence from the internal granule layer implies that cerebroglycan message disappears by the time cell body migration has finished. From studies of the timing of cerebellar neuron migration (Fujita, 1967), it can be concluded that the window of cerebroglycan mRNA expression in granule cells is ~36 hours or less.

The view that cerebroglycan is first expressed around the time neurons become post-mitotic is supported by results in the early spinal cord, in which mRNA is found only in regions containing newly generated neurons (Figure 8B), and by the fact that all sites of cerebroglycan mRNA expression throughout the nervous system contain neurons that have been generated within the preceding 2-3 days (e.g. E12-E16 spinal cord and DRG [Fig. 1,3 (Altman and Bayer, 1984)]; E14-P0 cerebral cortex [Fig. 1; (Bayer and Altman, 1987)]; E19-P0 inferior colliculus [Fig. 1; (Altman and Bayer, 1981)]; E19 pontine nuclei [Fig. 1; (Altman and Bayer, 1978)]; E15 mouse retina [Fig. 5; (Young, 1985)]; and P7 hippocampus and olfactory bulb [Fig. 2; (Bayer, 1980; Bayer, 1983)]). In contrast, little if any cerebroglycan mRNA is found in the proliferating neural precursors that form the

ventricular zone, both in the E12 spinal cord (Figure 3B) and throughout the developing brain (not shown).

The time of disappearence of cerebroglycan message in at least one structure other than the cerebellum also correlates with completion of neuronal migration: Cells expressing cerebroglycan are found in the subventricular zone pathway along which olfactory bulb interneuron precursors migrate (Luskin, 1993), but not in the layers of the olfactory bulb that contain interneurons that have completed their migration. In other parts of the nervous system, however, a simple correlation between cerebroglycan expression and neuronal cell body migration cannot be drawn. The ventral horn of the spinal cord, the dorsal root ganglia, the E19 thalamus, the superficial layers of the E19 and P0 cerebral cortex, and the dentate gyrus of the hippocampus (Figures 1-3), all primarily consist of neurons that have reached their final positions, yet express cerebroglycan. Interestingly, although the cell bodies of these neurons are not migrating at these times, their axons are generally still growing. If the same cells are examined after their axons have reached target areas, cerebroglycan message is undetectable [e.g. after E20 for thalamus (Erzurumlu and Jhaveri, 1992); E18 for dorsal root ganglia (Killackey et al., 1990); and P2-P5 for cerebral cortex (O'Leary et al., 1990)]. Since axon growth and cell migration are both forms of cell motility (Wilcox and Unnerstall, 1991; Cypher and Letourneau, 1992), the data suggest a general correlation between cerebroglycan expression and periods of time during which neurons engage in motile behaviors.

Although the data presented cannot rule out the possibility that some glial cells transiently express cerebroglycan, neither the timing nor pattern of expression of cerebroglycan match the timing or sites of glial production in the brain, which is largely postnatal (Jacobson, 1991). In addition, several of the structures in the developing brain that highly express cerebroglycan mRNA contain few if any glial cell bodies, e.g. the external granule layer of the cerebellum (Figure 3) and the E15 mouse retina (Figure 5). Interestingly, recent evidence indicates that, in the adult rat brain, glypican is also expressed largely, if not exclusively, by neurons (Litwack et al., 1994).

Cerebroglycan is expressed in axon tracts of the developing nervous system

Immunohistochemistry with an affinity purified anti-CBG antibody confirmed that the CBG core protein, like the CBG mRNA, is restricted to the nervous system. Additionally, the antibody revealed that CBG protein is primarily associated with axon tracts (Figure 4). The localization of CBG to axon tracts, together with the fact that the CBG mRNA appears to be expressed during axon outgrowth in many neuronal populations, suggests that CBG may have a role in guidance of elongating axons. Consistent with this possibility, CBG -- as well as glypican -- has recently been shown to be enriched in biochemical preparations of growth cone particles from neonatal rat brain (Ivins et al., 1995).

A recent study has shown that glypican is also expressed in axon tracts of the developing rat central nervous system (Litwack, 1995). One possible function of the GPI lipid anchor in glypican and CBG could be to target these molecules to axons. Several other GPI-linked molecules are known to be expressed on axons (Dodd et al., 1988; Furley et al., 1990; Xue et al., 1990; Dotti et al., 1991; Faivre-Sarrailh et al., 1992), and it has been demonstrated for some polarized epithelial cells in culture that GPI-linked molecules are preferentially sorted to the apical cell surface, which in neurons may be analogous to the axonal cell surface (Olwin and Rapraeger, 1992). For glypican, it has not been possible to address unambiguously whether expression is actually restricted to axons. However, one study recently suggested that CBG is in fact expressed on the axons but not the dendrites of a particular cell type, the granule cell neuron of the dentate gyrus (Litwack, 1995).

CBG expression in the olfactory epithelium: evidence of regulation by extrinsic factors

The pattern of CBG expression in the olfactory epithelium in vivo is consistent with what has been observed in other parts of the developing nervous system. In the E15 OE, the ordered layers present in the adult OE have not yet arisen (Smart, 1971). Consequently, neurogenesis and early neuronal differentiation are not

localized within the OE at this stage. Likewise, CBG expression in the E15 OE is detected throughout the entire width of the structure.

In contrast, by P15, the adult structure of the OE has essentially been established. In the P15 OE, neurogenesis is restricted to near the basal surface of the epithelium, and neuronal differentiation is restricted to the bottom one third to one half of the epithelium, as marked by staining with GAP-43, a marker for neurons in the process of neurite outgrowth [(Verhaagen et al., 1989), and Anne Calof, unpublished observations]. CBG mRNA was found to be highly localized to near the basal surface of the OE (Figure 6B). It is possible, on the basis of CBG mRNA expression, that CBG may be even more restricted than GAP-43 in the P15 OE. In the future it will be of interest to determine the relationship of GAP-43 expression and CBG expression in the OE using co-immunolocalization experiments.

In OE explant cultures, CBG was not expressed in the expected pattern. It is known that, in these cultures, a large number of immediate neuronal precursors (INPs) migrate away from the explant and divide once to give rise to two NCAM-positive, immature neurons. Based on CBG expression in the OE and elsewhere in vivo, we predicted that CBG expression would also appear in the daughter neurons of the INP divisions. However, no CBG-positive cells were detected on the substrate away from the explant in any of the cultures. Many CBG positive cells were detected, but these were always associated with the perimeter of the explant, in clusters near the keratin-positive ball of basal epithelium.

One reason, a priori, for the failure to see CBG-positive cells on the substrate away from the explant would be if the window of CBG expression in these cells is so short that finding a CBG expressing cell would be a very rare event. However, the large numbers of CBG-positive cells around the explant suggest that an extremely short window of expression is not a fundamental feature of CBG expression in these cells, and points to the possibility that other factors may be involved.

One possibility is that cell-cell contact is important for the normal expression of CBG. In experiments with cerebellar granule cells, culturing cells as cellular reaggregates, as apposed to dispersed cells,

caused a ten fold increase in DNA synthesis. The addition of astroglial cells or cell membranes to the reaggregates arrested DNA synthesis in a dose dependent manner (Gao et al., 1991). This experiment points out the possible importance of homotypic interactions among proliferating neuronal precursors in regulating proliferation and early differentiation. In the OE explant cultures, cells that migrate away from the explant must complete proliferation and differentiation without the benefit of the cell-cell contacts that cells remaining with the explant, and cells in vivo, normally enjoy.

Another possibility is that a factor supplied by basal epithelial cells is required for normal CBG expression. In an experiment in which dissociated OE cells were plated with cells from the olfactory bulb (the normal target of olfactory receptor neurons), a fraction of the OE cells were found to express carnosine, a putative neurotransmitter for olfactory receptor neurons. These carnosine-positive neurons were always found associated with clusters of keratin-positive basal epithelial cells, suggesting association with basal epithelial cells may influence the differentiation of OE receptor neurons (Perroteau et al., 1994).

Nevertheless, INPs that migrate off the explant can, under the proper culturing conditions, differentiate into neurons that express markers of mature olfactory receptor neurons such as OMP (olfactory marker protein) (Calof et al., 1993). The fact that these cells were not observed to express CBG raises the possibility that CBG is not required for the neurite outgrowth by of these cells. However, unlike INPs in culture, which extend neurites at random, immature neurons in the OE in vivo must extend axons to a defined target, the olfactory bulb. Thus if CBG is involved in axon outgrowth in vivo, it may function in axon guidance rather than the growth of axons per se. The enrichment of CBG in preparations of growth cone particles (Ivins et al., 1995) is consistent with a potential role for CBG in axon guidance (See Chapter 6 for a further discussion of potential functions for CBG in the nervous system).

The CBG cDNA is expressed as a membrane-associated HSPG in cell lines

PC12 and NIH3T3 cell lines expressing CBG were obtained by infection with a CBG retroviral expression vector. Preliminary experiments demonstrated CBG expression in these cells by immunofluorescence, using an affinity purified anti-CBG antibody and a fluorescent secondary antibody. 3T3infᵤCBGΔ3' cells, the 3T3 cells expressing CBG, were plated on both laminin (LN) and fibronectin (FN). When plated on laminin (LN), 3T3infᵤCBGΔ3' cells were less fully spread than when plated on FN. Differences in adhesiveness to LN and FN has been reported before for fibroblasts (Botti et al., 1987; Codogno et al., 1988). In one study with mouse B82L fibroblasts, the parental cell line adhered to well FN but not at all to LN. However, individual clones were isolated that had upregulated expression of α6 integrin subunit and become adherent and migratory on LN. Adhesion and motility of the transfected clones to LN could be inhibited by an antibody to the α6 integrin subunit (Lin and Bertics, 1995).

Interestingly, OE neurons also adhere poorly to LN and well to FN, but are migratory on LN but not on FN. As with the 3T3 cell in the study mentioned above, migration of OE neurons on LN can be inhibited by an antibody to the α6 integrin subunit (Calof and Lander, 1991b; Calof et al., 1994). Thus it is possible that the differences in cell shape of 3T3infᵤCBGΔ3' cells on LN and FN is the result of LN and FN binding to distinct integrin receptors that convey different signals to the cells (e.g., stimulation of migration vs. attachment and spreading).

The differences in cell shape of 3T3infᵤCBGΔ3' cells on LN and FN were accompanied by differences in CBG staining. On LN, several cells showed substantial, punctate CBG staining near the ventral and dorsal cell surfaces (Figure 11, A-C). At intermediate levels, less staining was observed, suggestive of a cell surface localization for CBG in these cells. The abundant amount of CBG found near the substrate level in these cells raises the possibility that CBG was concentrated at the basal surface by interacting with LN on the substrate.

LN is a good candidate ligand for CBG in vivo because it is expressed in the developing nervous system in several areas now

known to be traversed by axons in CBG-positive axon tracts. For example, CBG expression is associated with axons of the optic tract, axons in the intermediate zone of the cerebral cortex, and the axons of spinal cord motoneurons (Figure 4), all of which are known to navigate in areas containing LN (Rogers et al., 1986; Cohen et al., 1987; McLoon et al., 1988; Hunter et al., 1992).

On FN, 3T3inf$_u$CBGΔ3' cells exhibited a flat, fully spread morphology that made any estimation of CBG levels in cytoplasm versus on the cell membrane impossible (Figure 11, D-F). Surprisingly, even the space occupied by the nucleus stained positive for CBG. Whether it is an artifact of fixation and permeabilization, or whether the epitope was actually imported into the nucleus by the cells, the significance of this observation remains to be determined. No CBG signal was detected in the nuclei of 3T3 and PC12 cell lines plated on LN, indicating that the unusual staining is not a typical consequence of the staining protocol employed. Examination of the 3T3inf$_u$CBGΔ3' cells as a population suggested that they were more uniformly positive for CBG when plated on FN (not shown). It is possible that, on FN, for reasons not understood, CBG turnover is slower, and the high level of CBG expression by the constitutive CMV promoter results in "overflow" of CBG, or CBG breakdown products, into other cellular compartments. There have been some reports of heparan sulfate (HeS), biochemically demonstrated to be in the nucleus of some cell types (Fedarko and Conrad, 1986; Ishihara and Conrad, 1989), although the findings in some cases may be artifacts of the way the nuclei were prepared (Hiscock et al., 1994). However, one set of experiments has shown HeS with a rare modification is apparently selectively enriched in the nuclei of confluent hepatocytes, a finding that argues against non-specific contamination (Fedarko and Conrad, 1986; Ishihara et al., 1986; Fedarko et al., 1989). In one instance, an antibody to HeS was found to stain a small number of neuronal cell nuclei in Alzheimer's brain, whereas no nuclear staining was seen in normal brain (Su et al., 1992), suggesting that, in some cases, the presence of HeS in cell nuclei is a pathological response.

For PC12inf$_u$CBGΔ3' cells plated on LN, high levels of CBG expression were typically detected at the level of the substrate, and in

the intracellular space around the nucleii (Figure 11, G and H).
However, unlike 3T3inf$_u$CBGΔ3' cells, PC12inf$_u$CBGΔ3' cells did not
show strong CBG staining on their dorsal cell surfaces (Figure 11 I).
The staining often illuminated many small micro-extensions on ventral
and lateral surfaces of these cells. In lightly stained cells, viewed at
lower magnification, this fringe of stained processes was often the
predominant feature of CBG staining (Figure 10C). In the future, it
would be of interest to determine whether these micro-extensions
also contain actin or other submembranous cytoskeletal components
such as ezrin or moesin, indicative of microspikes or filopodia.

Analysis of the PGs in the membrane fraction of 3T3inf$_u$CBGΔ3'
cells revealed that an HSPG core protein of a M_r ~53 kDa is present
that is not found in the membrane fraction PGs of parental 3T3 cells.
This same core protein is the most prominent band in
immunoprecipitations of 3T3inf$_u$CBGΔ3' cell membrane PGs using the
affinity-purified 521-2 anti-CBG antibody (Figure 12). Taken together,
the data strongly suggest that the CBG cDNA obtained in Chapter 2
encodes a membrane associated HSPG. In the future it will be
important to demonstrate that CBG can be released from these cells by
treatment with PIPLC (phosphoinositol-specific phospholipase C) an
enzyme that cleaves GPI anchors.

In summary, in situ hybridization and immunolocalization
experiments have demonstrated that CBG expression is restricted to
the developing nervous system. CBG mRNA appears in immature
neurons around the time of their final mitosis and disappears shortly
after cell body migration and axon outgrowth have been completed,
raising the possibility that CBG plays a role in these motile behaviors.
The fact that the CBG core protein is localized to axon tracts is also
suggestive of a role for CBG in axon outgrowth.

Expression of CBG in olfactory epithelium (OE) in vivo is consistent
with patterns observed elsewhere in the developing nervous system:
CBG expression colocalizes with populations of newly post-mitotic
immature neurons. However, CBG expression in OE explants in vitro
suggests that extrinsic factors, possibly homotypic cell-cell contact of

neurons or precursors within the OE, or a factor(s) derived from the basal epithelium, are required for proper CBG expression.

Expression of CBG in cell lines strongly suggests that the CBG cDNA obtained in Chapter 2 encodes a membrane associated HSPG core protein. These cell lines or others like them could be used to address questions of CBG structure such as where glycosaminoglycans are attached. Although many cell lines naturally express the related HSPG, glypican, the use of CBG in these experiments would have the advantage that modified versions of CBG could be expressed and isolated without the use of an epitope tag since no cell line tested, including PC12, 3T3, C6 glioma, and NBA1 neuroblastoma, has been found to express a detectable level of CBG (not shown). Further, it may prove interesting to compare the structure of CBG with that of endogenous glypican produced by the same cell type.

The apparent concentration of CBG at the substrate level observed in CBG-infected 3T3 cells plated on LN suggests CBG on the surface of these cells may bind to LN on the substrate. This is of potential importance because LN is an excellent candidate ligand for CBG in vivo. As discussed above, many examples exist of CBG expression associated with axons that navigate through LN rich regions. The fact that CBG appears to be expressed specifically during periods of axon outgrowth in the nervous system serves to underscore the possibility that these two molecules interact in vivo.

References

Altman, J. and S.A. Bayer. 1978. Prenatal development of the cerebellar system in the rat. II. Cytogenesis and histogenesis of the inferior olive, pontine gray, and the precerebellar reticular nuclei. J. Comp Neurol. 179:49-76.

Altman, J. and S.A. Bayer. 1981. Time of origin of neurons of the rat inferior colliculus and the relations between cytogenesis and tonotopic order in the auditory pathway. Exp. Brain Res. 42:411-423.

Altman, J. and S.A. Bayer. (1984). The development of the rat spinal cord. In Advances in Anatomy, Embryology, and Cell Biology, F. Beck, W. Hild, J. van Limborgh, R. Ortmann, J. E. Pauly, and T. H. Schiebler, editors. (New York: Springer-Verlag), pp. 1-150.

Anderson, R.G.W., B.A. Kamen, K.G. Rothberg, and S.W. Lacey. 1992. Potocytosis: sequestration and transport of small molecules by caveolae. Science (Wash. DC) 255:410-411.

Bayer, S.A. 1980. Development of the hippocampal region in the rat. I. Neurogenesis examined with ^{3}H-thymidine autoradiography. J. Comp. Neurol 183:89-106.

Bayer, S.A. 1983. Thymidine radiographic studies of neurogenesis in the rat olfactory bulb. Exp. Brain Res 50:329-340.

Bayer, S.A. and J. Altman. 1987. Directions in neurogenetic gradients and patterns of anatomical connections in the telencephalon. Prog. Neurobiol. 29:57-106.

Benowitz, L.I. and A. Routtenberg. 1987. A membrane phosphoprotein associated with neural development, axonal regeneration, phospholipid metabolism, and synaptic plasticity. Trends Neurosci. 10:527-532.

Bernfield, M. and K.C. Hooper. 1991. Possible regulation of FGF activity by syndecan, an integral membrane heparan sulfate proteoglycan. Ann. NY Acad. Sci. 638:182-194.

Bernfield, M., R. Kokenyesi, M. Kato, M.T. Hinkes, J. Spring, R.L. Gallo, and E.J. Lose. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu. Rev. Cell Biol. 8:365-393.

Botti, J., M.A. Doyennett-Moyne, E. Gouet, M. Aubrey, and P. Codogno. 1987. Alterations of fibroblast interactions with fibronectin and laminin during chick embryo development. Cell Biol. Int. Rep. 11:849-859.

Calof, A.L., M.D. Adusumalli, M.K. DeHamer, J.L. Guevara, J.S. Mumm, S.J. Whitehead, and A.D. Lander. (1993). Generation, differentiation, and maturation of olfactory receptor neurons in vitro. In Olfaction and Taste XI: proceedings of the 11th International Symposium on Olfaction and Taste, K. Kurihara, N. Suzuki and H. Ogawa, editors. (Tokyo; New York: Springer-Verlag), pp. 36-40.

Calof, A.L., M.R. Campanero, J.J. O'Rear, P.D. Yurchenco, and A.D. Lander. 1994. Domain-specific activation of neuronal migration and neurite outgrowth-promoting activities of laminin. Neuron 13:117-130.

Calof, A.L. and D.M. Chikaraishi. 1989. Analysis of neurogenesis in a mammalian neuroepithelium: proliferation and differentiation of an olfactory neuron precursor in vitro. Neuron 3:115-127.

Calof, A.L. and A.D. Lander. 1991b. Relationship between neuronal migration and cell-substratum adhesion: Laminin and merosin promote neuronal migration but are anti-adhesive. J. Cell Biol. 115:779-794.

Calof, A.L., A.D. Lander, and D.M. Chikaraishi. (1991a). Regulation of neurogenesis and neuronal differentiation in primary and immortalized cells from mouse olfactory epithelium. In Regeneration of Vertebrate Sensory Receptro Cells: Ciba Found. Symp.), pp. 249-276.

Carey, D., D. Evans, R. Stahl, V. Asundi, K. Conner, P. Garbes, and G. Cizmeci-Smith. 1992. Molecular cloning and characterization of N-syndecan, a novel transmembrane heparan sulfate proteoglycan. J. Cell Biol. 117:191-201.

Codogno, P., M.A. Doyennette-Moyne, J. Botti, and M. Aubry. 1988. Concanavilin A-induced impairment of fibroblast spreading on laminin but not on fibronectin. J. Cell Physiol. 136:463-470.

Cohen, J., J.F. Burne, C. McKinlay, and J. Winter. 1987. The role of laminin and the laminin/fibronectin receptor complex in the outgrowth of retinal ganglion cell axons. Dev. Biol. 122:407-418.

Cypher, C. and P.C. Letourneau. 1992. Growth cone motility. Curr. Opin. Cell Biol. 4:4-7.

Dodd, J., S.B. Morton, D. Karagogeos, M. Yamamoto, and T.M. Jessell. 1988. Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. Neuron 1:105-116.

Dotti, C.G., R.G. Parton, and K. Simmons. 1991. Polarized sorting of glypiated proteins in hippocampal neurons. Nature 349:158-160.

Erzurumlu, R.S. and S. Jhaveri. 1992. Emergence of connectivity in the embryonic rat parietal cortex. Cerebral Cortex 2:336-352.

Faivre-Sarrailh, C., G. Gennarini, C. Goridis, and G. Rougon. 1992. F3/F11 cell surface molecule expression in the developing mouse cerebellum is polarized at synaptic sites and with granule cells. J. Neurosci. 12:257-267.

Fedarko, N.S. and H.E. Conrad. 1986. A unique heparan sulfate in the nuclei of hepatocytes: structural changes with the growth state of cells. J. Cell Biol. 102:587-599.

Fedarko, N.S., M. Ishihara, and H.E. Conrad. 1989. Control of cell division in hepatoma cells by exogenous heparan sulfate proteoglycan. J. Cell. Physiol 139:287-294.

Fu, Y.-M., P. Spirito, Z.-X. Yu, S. Biro, J. Sasse, J. Lei, V.J. Ferrans, S.E. Epstein, and W. Casscells. 1991. Acidic fibroblast growth factor in the developing rat embryo. J. Cell Biol. 114:1261-1273.

Fujita, S. 1967. Quantitative analysis of ell proliferation and differentiation in the cortex of the postnatal mouse cerebellum. J. Cell Biol. 32:277-287.

Furley, A.J., S.B. Morton, D. Manalo, D. Karagogeos, J. Dodd, and T.M. Jessell. 1990. The axonal glycoprotein TAG-1 is an immunoglobulin superfamily member with neurite outgrowth-promoting activity. Cell 61:157-170.

Gao, W.O., N. Heintz, and M.E. Hatten. 1991. Cerebellar granule cell neurogenesis is regulated by cell-cell interactions in vitro. Neuron 6:705-715.

Gonzalez, A., M. Buscaglia, M. Ong, and A. Baird. 1990. Distribution of basic fibroblast growth factor in the 18-day rat fetus: localization in the basement membrane of diverse tissues. J. Cell Biol. 110:753-765.

Gould, S.E., W.B. Upholt, and R.A. Kosher. 1995. Characterization of chicken syndecan-3 as a heparan sulfate proteoglycan and its expression during embryogenesis. Dev. Biol. 168:438-451.

Graziadei, P.P.C. and G.A.M. Graziadei. (1978). Continuous nerve cell renewal in the olfactory system. In Handbook of Sensory Physiology, Volume IX: Development of Sensory Systems, M. Jacobson, editor. (New York: Springer Verlag), pp. 55-83.

Haugen, P.K., P.C. Letourneau, S.L. Drake, L.T. Furcht, and J.B. McCarthy. 1992. A cell surface heparan sulfate proteoglycan mediates neural cell adhesion and spreading on a defined sequence from the C-terminal cell and heparin binding domain of fibronectin, FN-C/H II. J. Neurosci. 12:2597-2608.

Herndon, M.E. and A.D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. Neuron 4:949-961.

Hiscock, D.R., M. Yanagishita, and V.C. Hascall. 1994. Nuclear localization of glycosaminoglycans in rat ovarian granulosa cells. J. Biol. Chem. 269:4539-4546.

Hughes, R.A., M. Sendtner, M. Goldfarb, D. Lindholm, and H. Thoenen. 1993. Evidence that fibroblast growth factor 5 is a major muscle-derived survival factor for cultured spinal motoneurons. Neuron 10:369-377.

Hunter, D.D., R. Llinas, M. Ard, J.P. Merlie, and J.R. Sanes. 1992. Expression of S-laminin and laminin in developing rat central nervous system. J. Comp. Neurol. 323:238-251.

Ishihara, M. and H.E. Conrad. 1989. Correlations between heparan sulfate metabolism and hepatoma growth. J. Cell Physiol. 138:467-476.

Ishihara, M., N.S. Fedarko, and H.E. Conrad. 1986. Transport of heparan sulfate into the nuclei of hepatocytes. J. Biol. Chem. 261:13575-13580.

Ivins, J.K., E.D. Litwack, A. Kumbasar, C.S. Stipp, A. Yang, and A.D. Lander. 1995. Cerebroglycan and glypican are cell surface heparan sulfate proteoglycans expressed on axons and growth cones in the developing rat nervous system. Soc. Neurosci. Abstracts 21:795.

Jacobson, M. (1991). Time and sequence of origin of neurons and glial cells in the central nervous system. In Developmental Neurobiology, (New York: Plenum), pp. 58-64.

Kan, M., F. Wang, J. Xu, J.W. Crabb, J. Hou, and W.L. McKeehan. 1993. An essential heparin-binding domain in the fibroblast growth factor receptor kinase. Science 259:1918-1921.

Kiefer, M.C., J.C. Stephans, K. Crawford, K. Okino, and P.J. Barr. 1990. Ligand-affinity cloning and structure of a cell surface heparan sulfate proteoglycan that binds basic fibroblast growth factor. Proc. Natl. Acad. Sci. USA 87:6985-6989.

Killackey, H.P., M.F. Jacquin, and R.W. Rhoades. (1990). Development of somatosensory structures. In Development of Sensory Systems In Mammals, J. R. Coleman, editor. (New York: J. Wiley and Sons), pp. 403-429.

Lin, M.L. and P.J. Bertics. 1995. Laminin responsiveness is associated with changes in fibroblast morphology, motility, and anchorage-independent growth: A cell system for examining the interaction between laminin and EGF signaling pathways. J. Cell Physiol. 164:593-604.

Lisanti, M.P., P.E. Scherer, J. Vidugiriene, Z.L. Tang, A. Hermanowski-Vosatka, Y.-H. Tu, R.F. Cook, and M. Sargiacomo. 1994. Characterization of caveolin-rich membrane domains isolated from an endothelial-rich source: implications for human disease. J. Cell Biol. 126:111-126.

Litwack, E.D. (1995). Expression and function of proteoglycans in the nervous system. Doctoral Dissertation. Massachusetts Institute of Technology.

Litwack, E.D., C.S. Stipp, A. Kumbasar, and A.D. Lander. 1994. Neuronal expression of glypican, a cell-surface glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan, in the adult rat nervous system. J. Neurosci. 14:3713-3724.

Luskin, M.B. 1993. Restricted proliferation and migration of postnatally generated neurons derived from the forebrain subventricular zone. Neuron 11:173-189.

McLoon, S.C., L.K. McLoon, S.L. Palm, and L.T. Furcht. 1988. Transient expression of laminin in the optic nerve of the developing rat. J. Neurosci. 8:1981-1990.

Miller, A.D. and G.J. Rosman. 1989. Improved retroviral vectors for gene transfer and expression. Biotechniques 7:980-990.

Murphy, M., J. Drago, and P.F. Bartlett. 1990. Fibroblast growth factor stimulates the proliferation and differentiation of neural precursor cells in vitro. J. Neurosci. Res. 25:463-475.

Noakes, P.G., M. Gautam, J. Mudd, J.R. Sanes, and J.P. Merlie. 1995. Aberrant differentiation of neuromuscular junctions in mice lacking s-laminin/laminin beta 2. Nature 6519:258-262.

O'Leary, D.D.M., A.R. Bicknese, J.A.D. Carlos, C.D. Heffner, S.E. Koester, L.J. Kutka, and T. Terashima. 1990. Target selection by cortical axons: alternative mechenisms to establish axonal connections in the developing brain. Cold Spring Harbor. Symp. Quant. Biol. 55:453-468.

Olwin, B.B. and A. Rapraeger. 1992. Repression of myogenic differentiation by aFGF, bFGF, and K-FGF is dependent on cellular heparan sulfate. J. Cell Biol. 118:631-639.

Ornitz, D.M. and P. Leder. 1992. Ligand specificity and heparin dependence of fibroblast growth factors 1 and 3. J. Biol. Chem. 267.

Perroteau, I., S. Biffo, E. Tolosano, G. Tarozzo, P. Bovolin, H. Vaudry, and A. Fasolo. 1994. In vitro study of olfactory receptor neurones expressing the dipeptide carnosine. Neuroreport 5:569-572.

Porter, B.E., J. Weis, and J.R. Sanes. 1995. A motoneuron-selective stop signal in the synaptic protein S-laminin. Neuron 3:549-559.

Rapraeger, A., M. Jalkenen, and M. Bernfield. 1985. The cell surface proteoglycan from mouse mammary epithelial cells bears chondroitin sulfate and heparan sulfate glycosaminoglycans. J. Biol. Chem. 260:11046-11052.

Rapraeger, A., A. Krufka, and B.B. Olwin. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science 252:1705-1708.

Rogers, S.L., K.J. Edson, P.C. Letourneau, and S.C. McLoon. 1986. Distribution of laminin in the developing peripheral nervous system of the chick. Dev. Biol. 113:429-435.

Sanderson, R.D., T.B. Sneed, L.A. Young, G.L. Sullivan, and A.D. Lander. 1992a. Adhesion of B lymphoid (MPC-11) cells to type I collagen is mediated by the integral membrane proteoglycan, syndecan. J. Immunol. 148:3902-3911.

Sanes, J.R., E. Engvall, R. Butkowski, and D.D. Hunter. 1990. Molecular heterogeneity of basal laminae: Isoforms of laminin and collagen IV at the neuromuscular junction and elsewhere. J. Cell Biol. 111:1685-1699.

Sargiacomo, M., M. Sudol, Z.L. Tang, and M.P. Lisanti. 1993. Signal transducing molecules and GPI-linked proteins form a caveolin-rich insoluble complex in MDCK cells. .

Saunders, S. and M. Bernfield. 1988. Cell surface proteoglycan binds mouse mammary epithelial cells to fibronectin and behaves as a receptor for interstitial matrix. J. Cell Biol. 106:423-430.

Seldin, D.C., K.F. Austen, and R.L. Stevens. 1985. Purification and characterization of protease-resistant secretory granule proteoglycans containing chondroitin sulfate di-B and heparin-like

glycosaminoglycans from rat basophilic leukemia cells. J. Biol. Chem. 260:11131-11139.

Sheppard, M.M., S.K. Hamilton, and A.L. Pearlman. 1991. Changes in the distribution of extracellular matrix components accompany early morphogenetic events of mammalian cortical development. J. Neurosci. 11:3928-3942.

Shworak, N.W., M. Shirakawa, R.C. Mulligan, and R.D. Rosenberg. 1994a. Characterization of ryudocan glycosaminoglycan acceptor sites. J. Biol. Chem. 269:21204-21214.

Sidman, R.L. and P. Rakic. 1973. Neuronal migration with special reference to developing human brain: a review. Brain Res. 62:1-35.

Simmons, D., J. Arriza, and Swanson. 1989. A complete protocol for in situ hybridization of messenger RNAs in brain and other tissues with radiolabelled single-stranded RNA probes. J. Histotechnology 12:169-181.

Smart, I.H.M. 1971. Location and orientation of mitotic figures in the developing mouse olfactory epithelium. J. Anat. 109:243-251.

Su, J.H., B.J. Cummings, and C.W. Cotman. 1992. Localization of heparan sulfate glycosaminoglycan and proteoglycan core protein in aged brain and Alzheimer's disease. Neuroscience 51:801-813.

Sutherland, A.E., R.D. Sanderson, M. Mayes, M. Seibert, P.G. Calarco, M. Bernfield, and C.H. Damsky. 1991. Expression of syndecan, a putative low affinity fibroblast growth factor receptor, in the early mouse embryo. Development 113:339-351.

Trautman, M.S., J. Kimelman, and M. Bernfield. 1991. Developmental expression of syndecan, an integral membrane proteoglycan, correlates with cell differentiation. Development 111:213-220.

Unsicker, K., H. Reichert-Preibsch, R. Schmidt, B. Pettmann, G. Labourdette, and M. Sensenbrenner. 1987. Astroglial and fibroblast growth factors have neurotrophic functions for cultured periperal and central nervous system neurons. Proc. Natl. Acad. Sci. USA 84:5459-5463.

Verhaagen, J., A.B. Oestreicher, W.H. Gispen, and F.L. Margolis. 1989. The expression of the growth associated protein B50/GAP43 in the olfactory system of neonatal and adult rats. J. Neurosci. 9:683-691.

Walicke, P. 1988. Basic and acidic fibroblast growth factors have trophic effects on neurons from multiple CNS regions. J. Neurosci. 8:2618-2627.

Walicke, P., W.M. Cowan, N. Ueno, A. Baird, and R. Guillemin. 1986. Fibroblast growth factor promotes survival of dissociated hippocampal neurons and enhances neurite extension. Proc. Natl. Acad. Sci. USA 83:3012-3016.

Wilcox, J.B. and J.R. Unnerstall. 1991. Expression of acidic fibroblast growth factor mRNA in the developing and adult brain. Neuron 6:397-409.

Xue, G.P., R.A. Calvert, and R.J. Morris. 1990. Expression of the neuronal surface glycoprotein Thy-1 is under post-transcriptional control, and is spatially regulated, in the developing olfactory nervous system. Development 109:851-864.

Yayon, A., M. Klagsbrun, J.D. Esko, P. Leder, and D.M. Ornitz. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. Cell 64:841-848.

Young, R.W. 1985. Cell proliferation during postnatal development of the retina in the mouse. Dev. Brain Res. 21:229-239.

Figure 1. CBG expression in the developing nervous system
A, C, E, and H: cresyl violet stained sections. B, D, F, G, I: reverse contrast images printed directly from autoradiograms of the same sections hybridized to α^{35}S-UTP labelled, *in vitro* transcribed, CBG RNA probe. A and B: embryonic day 14 whole embryo; C and D: embryonic day 16 whole embryo; E, F, and G: embryonic day 19 whole embryo; H and I: neonatal head. (s)--spinal cord; (te)--telencephalon; (cx)--cerebral cortex; (cb) cerebellar cortex; (d)--dorsal root ganglia; (p)--pontine nuclei; (i)--inferior colliculus, (ob)--olfactory bulb; (hi)--hippocampus; (st)--striatum, (th)--thalamus. Panel G shows the result when a sense orientation (negative control) probe was hybridized to an adjacent section of the embryonic day 19 animal; the image was printed at high contrast to allow visualization of the faint background. Bar = 5 mm.

A

B
te
d

C

D
te
tg
d
s

E
hi
st
th

F
cx
i
ob
cb
p

G

H
hi
st
th

I
cb
i
cx
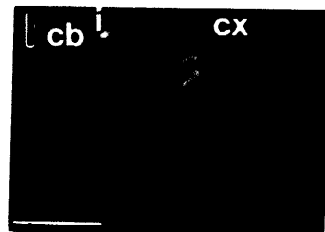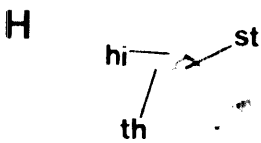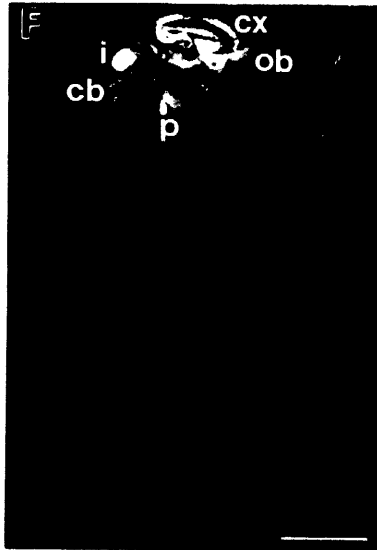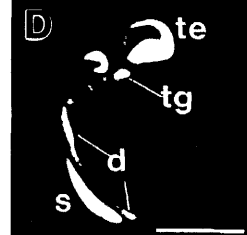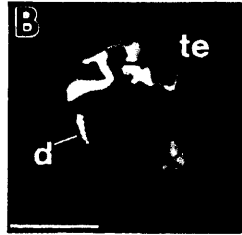
Figure 2. Expression of CBG in postnatal rat brain
Darkfield photomicrograph of a postnatal day 7 rat brain saggital section after *in situ* hybridization with an α^{35}S-UTP labelled CBG RNA probe. Hybridization is detected in the cerebellum (cb), the dentate gyrus of the hippocampus (dg), and in migrating granule cell precursors of the olfactory bulb (arrowhead). No hybridization was detected with the sense orientation (negative control) probe (not shown). Bar = 1 mm. Top is dorsal; right is rostral.
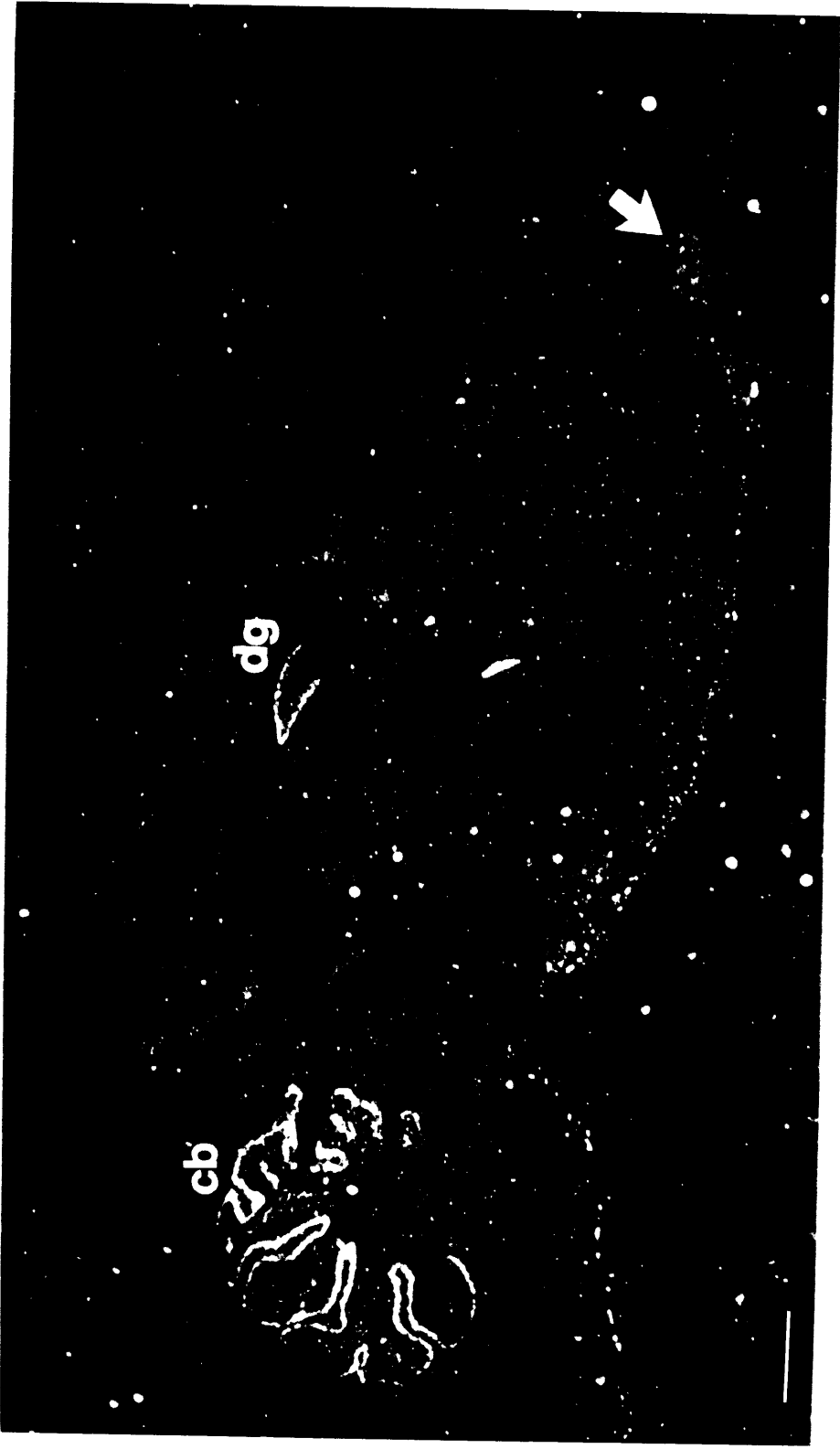
Figure 3. CBG is expressed by post-mitotic neurons
Simultaneous visualization of CBG mRNA (darkfield optics) and cell
nuclei (Hoechst blue fluorescence). (A) P7 cerebellum, sagittal
section. Hybridization is localized to the external granule layer, and is
concentrated in the premigratory zone, which contains newly
postmitotic granule neurons. EGL, external granule layer; ML,
molecular layer; IGL, internal granule layer; pz, proliferative zone;
pmz, premigratory zone. (B) E12 spinal cord, transverse section.
Hybridization is detected in the forming dorsal horns (DH), and
ventral horns (VH) of the spinal cord, and the dorsal root ganglia
(arrowheads), but not the ventricular zone (VZ). The cleft in the
ventricular zone of the spinal cord was an artifact of sectioning. No
signal was detected in adjacent sections hybridized with sense
orientation (negative control) probe (not shown). Bar = 100 μm.

A

EGL
ML
IGL
pz
pmz

B

DH
VZ
VH
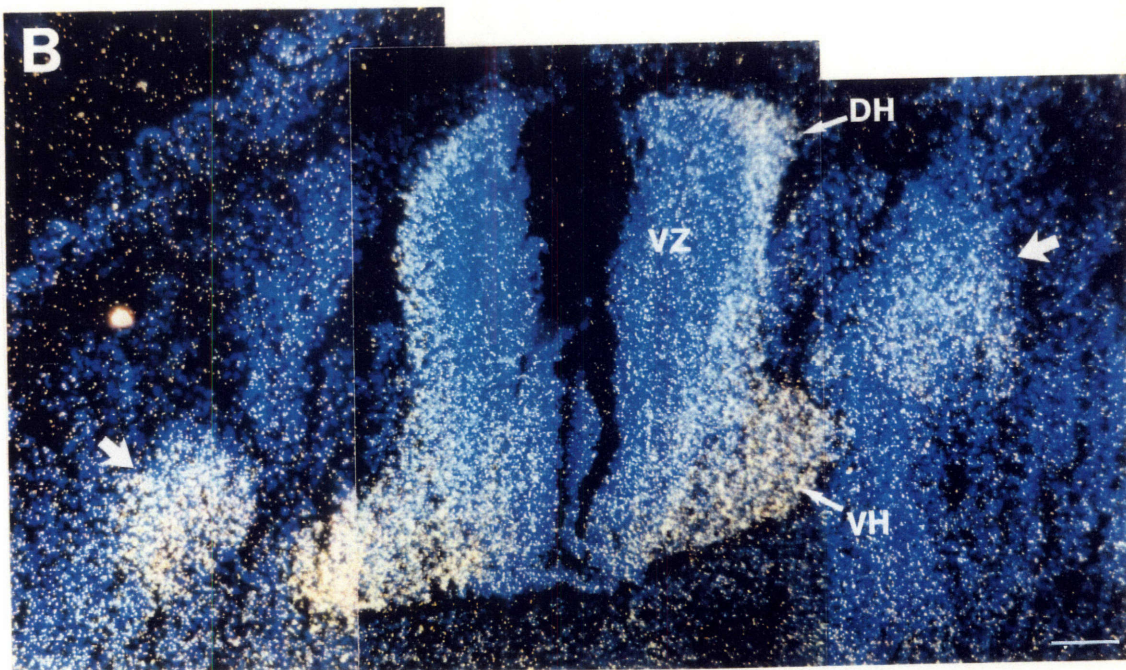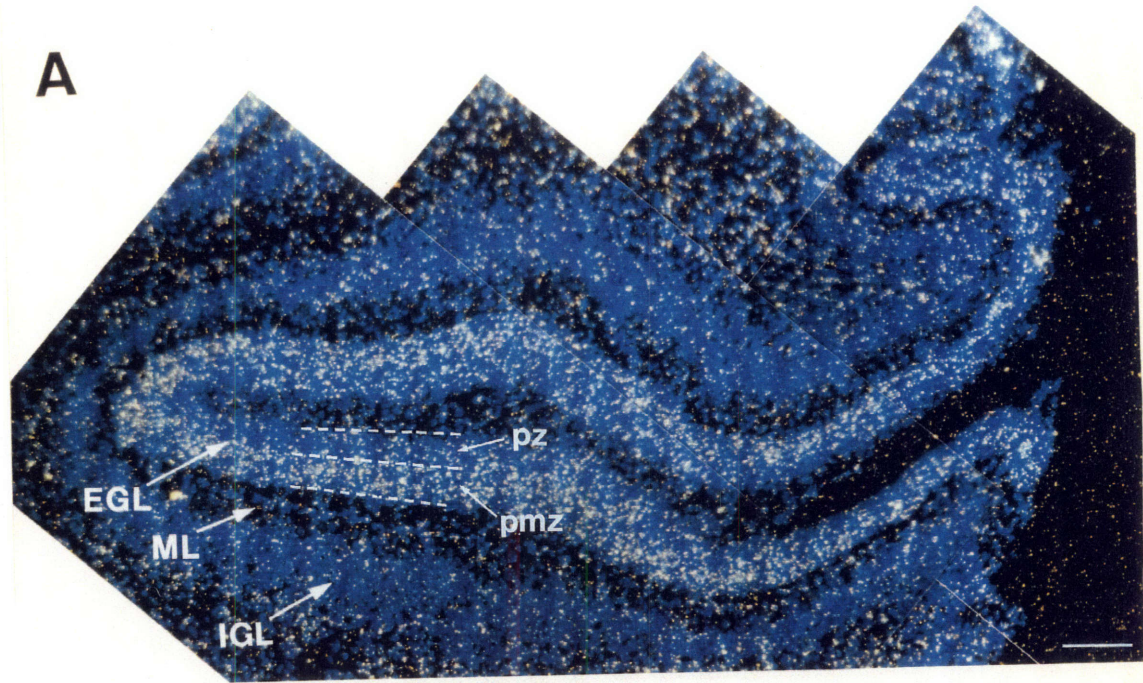
Figure 4. The CBG protein core is localized to fiber tracts
(A) Sagittal section of an E19 rat head stained with the 521-2 affinity purified anti-CBG antibody and developed with a horseradish peroxidase conjugated secondary antibody. Top is dorsal; left is rostral. (B) Coronal section of a P0 rat brain stained for CBG with 521-2. Top is dorsal; left is lateral. (C) Transverse section of an E13 rat spinal cord stained with 521-2. Top is dorsal. *Open arrows* points to spinal cord commisural axons. Abbreviations: c, hippocampal commisural fibers; d, dorsal root ganglion; h, hippocampus; i, internal capsule; iz, intermediate zone; m, motoneurons; o, optic tract; vr, ventral root. Scale bars: (A) and (B), 200 μm; (C), 50 μm.

This Figure courtesy of E. David Litwack (Litwack, 1995).
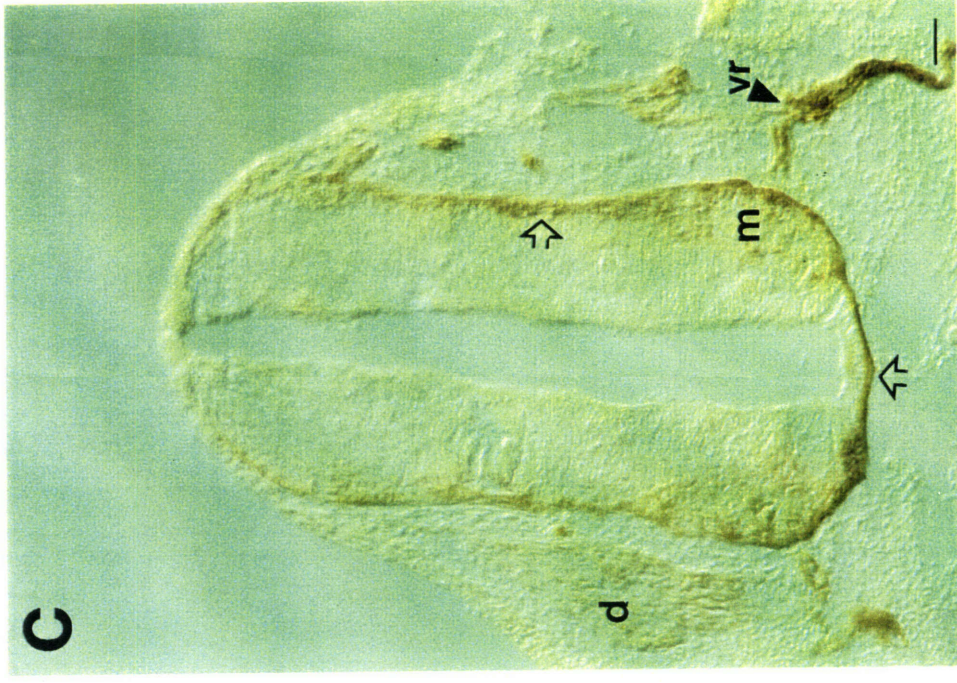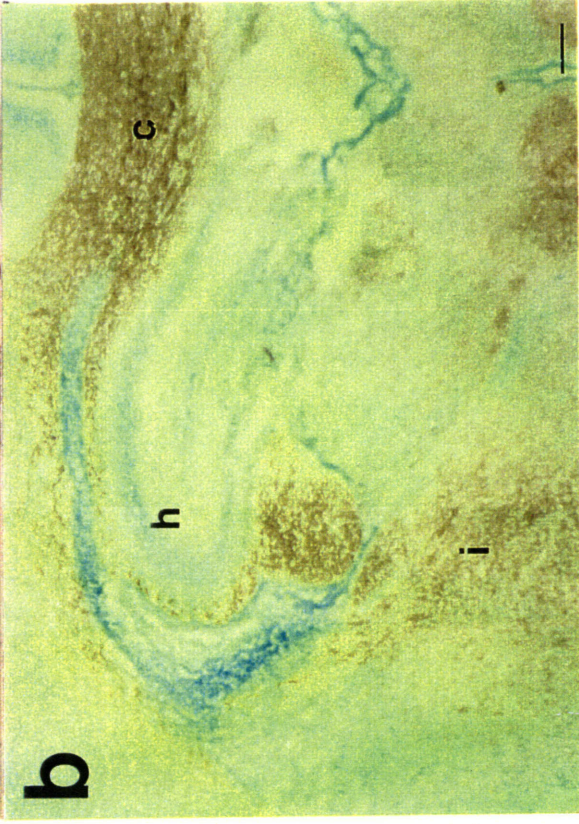
Figure 5. CBG expression in E15 mouse olfactory epithelium
Three nearly adjacent, horizontal sections from an embryonic day 15 mouse head are shown. (A) A cresyl violet stained section. OE, olfactory epithelium; R, retinal neuroepithelium. (B) A section hybridized with a digoxigenin-labelled CBG riboprobe reveals strong CBG expression in the retinal neuroepithelium, as well as expression throughout the olfactory epithelium. (C) The opposite strand control probe showed no background hybridization.
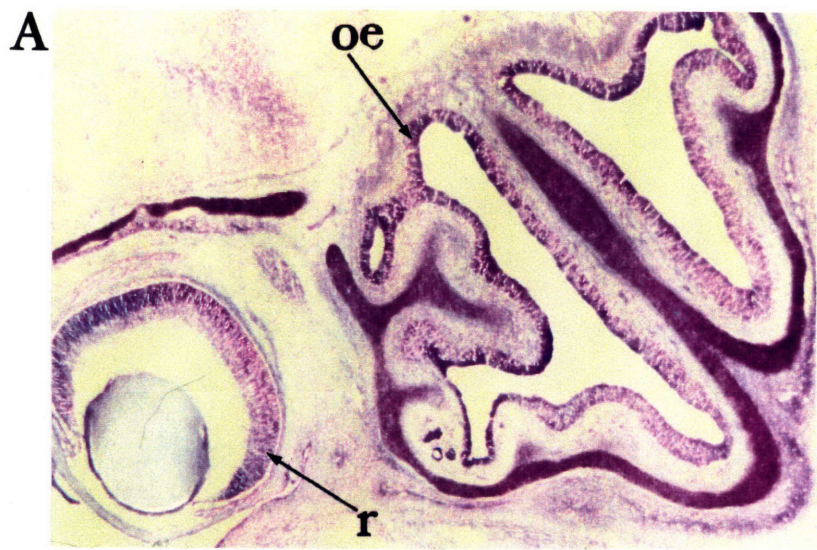
Figure 6. Expression of CBG in the P15 mouse olfactory epithelium

(A)   A schematic showing the organization of the mature olfactory epithelium (OE).   **B**, keratin positive basal epithelial cells;   **GB**, globose basal cells are morphologically distinguishable from other basal cells, and are believed to be the neuronal precursor cells in the OE;   **ORN**, immature olfactory receptor neurons in the process of axon outgrowth are $NCAM^+$, $GAP43^+$, but olfactory marker protein $(OMP)^-$;   **MORN**, mature olfactory receptor neurons have made contact with their target, the olfactory bulb in the brain, and are $NCAM^+$, OMP^+, $GAP43^-$;   **S**, sustentacular cells are a glial cell type in the OE.

(B)   High magnification view of a postnatal day 15 mouse OE, hybridized with a digoxigenin CBG riboprobe.   **m**, the mesenchyme underlying the olfactory epithelium;   **oe**, the olfactory epithelium;   **lu**, the lumen of the airway at the apical surface of the OE.   CBG expression is localized to near the basal surface of the OE.

A

Apical Surface

dendrites

S

$\dfrac{\text{MORN}}{\text{NCAM}^+}$
OMP$^+$

axon

$\dfrac{\text{ORN}}{\text{NCAM}^+}$
GAP43$^+$

Keratin$^+$

G   B

B

m   /   oe   /   lu
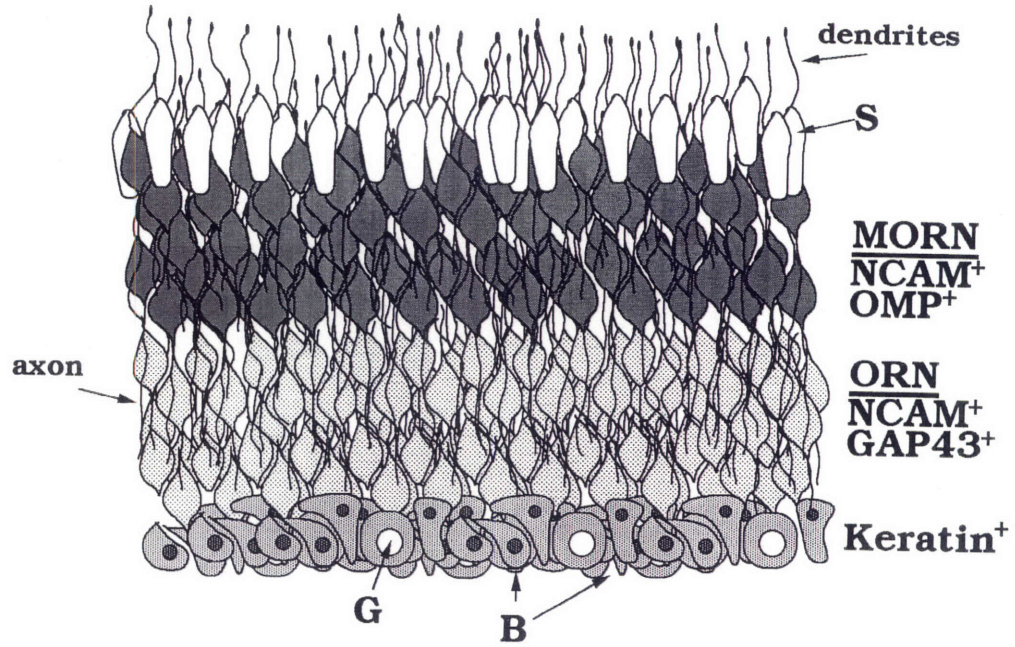
Figure 7. Proliferation and differentiation of olfactory neurons in vitro
Figure 7 is after Calof et. al. (1993). The culture of purified olfactory epithelium explants from embryonic day 14 or 15 mice is represented schematically. (A) 30 minutes after plating to laminin or merosin coated coverslips. (B) 6 hours after plating, immediate neuronal precursors have sorted out from the ball of basal epithelial cells, and many have migrated onto the substrate. Several of the neuronal precursors have already divided to give rise to NCAM+ immature neurons. (C) 24 hours after plating, most of the neuronal precursors have divided to give rise to immature neurons, and many GAP43+ neurons are extending neurites. Abbreviations: INP, immediate neuronal precursors; ORN, olfactory receptor neurons.

A
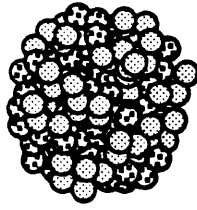
B
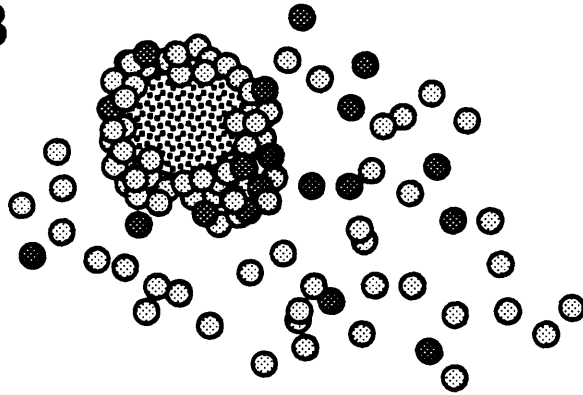
C
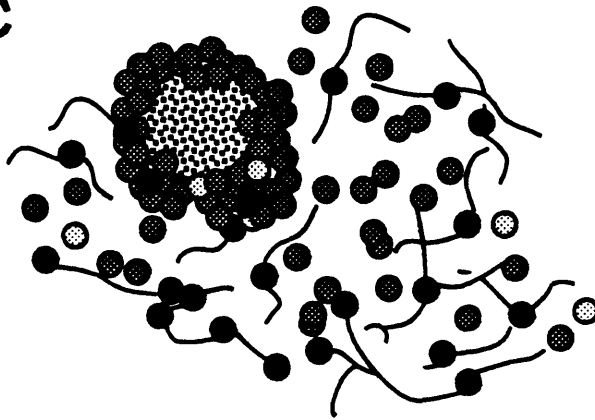
INP, NCAM(-)
ORN, NCAM(+)
ORN, NCAM(+), GAP43(+)
Basal epithelium

Figure 8. CBG expression in olfactory epithelium explants
(A), (C), (E), (G): differential interference contrast (DIC) views of olfactory epithelium (OE) explant cultures. (B), (D), (F), and (H): bright field views of the same fields as in (A), (C), (E), and (G), to allow visualization with truer color. In (A) and (B), a 6 hour explant has been hybridized with a digoxigenin (DIG)-labelled NCAM riboprobe. In (C) and (D), a 6 hour explant has been hybridized with a DIG-labelled CBG riboprobe. (E) and (F): hybridization with the opposite strand negative control probe. (G) and (H): A view at a higher focal plane of a 6 hour explant hybridized with a CBG riboprobe, revealing CBG-positive cells at the top lateral surface of the explant. Negative control probes showed little hybridization to cells at this level (not shown). Cells on the substrate below are out of the plane of focus. Abbreviations: be, basal epithelium cells.
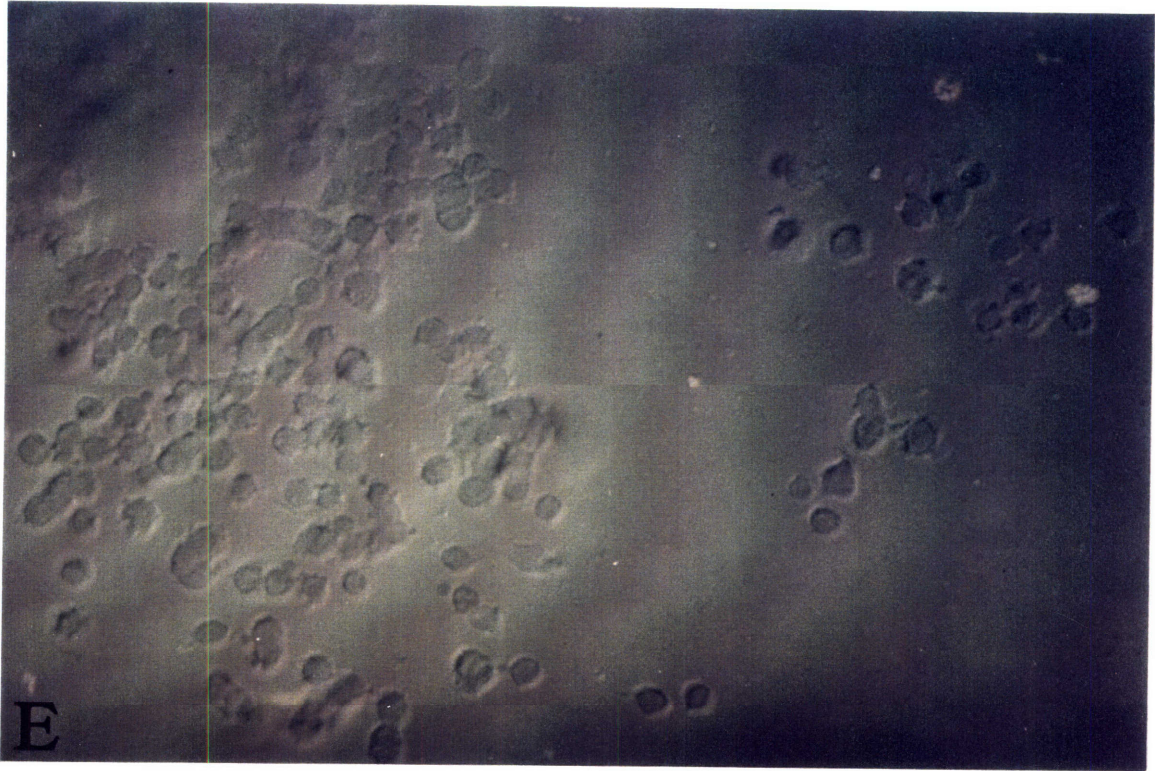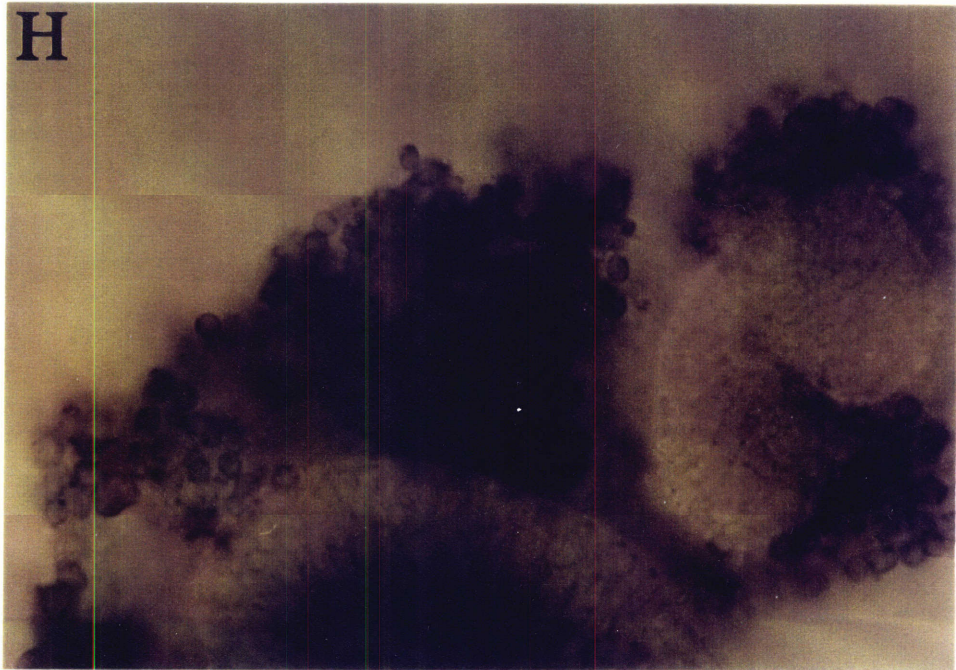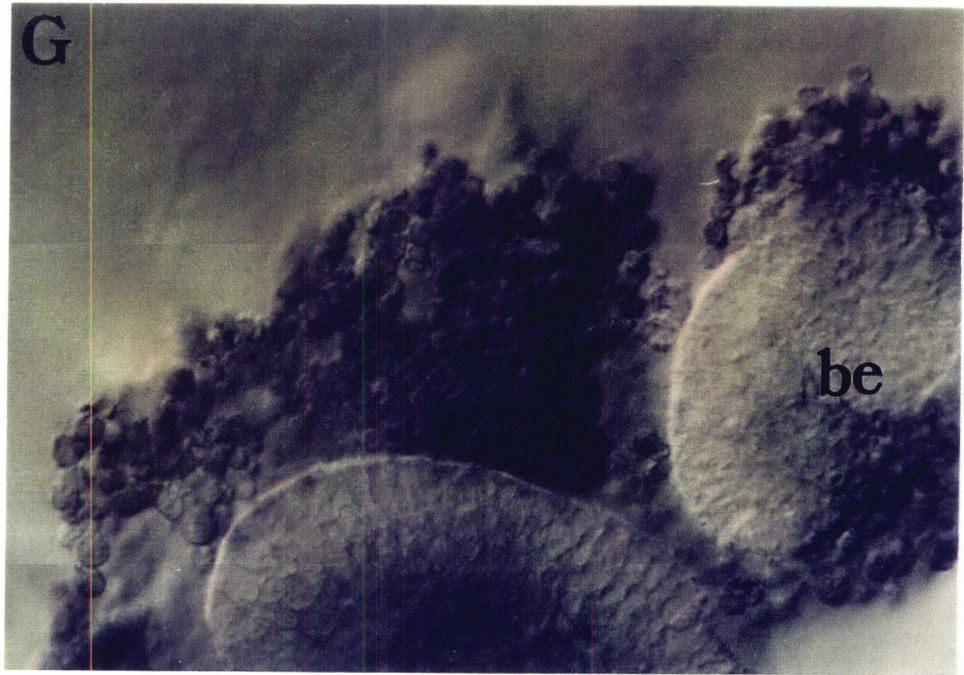
E

F
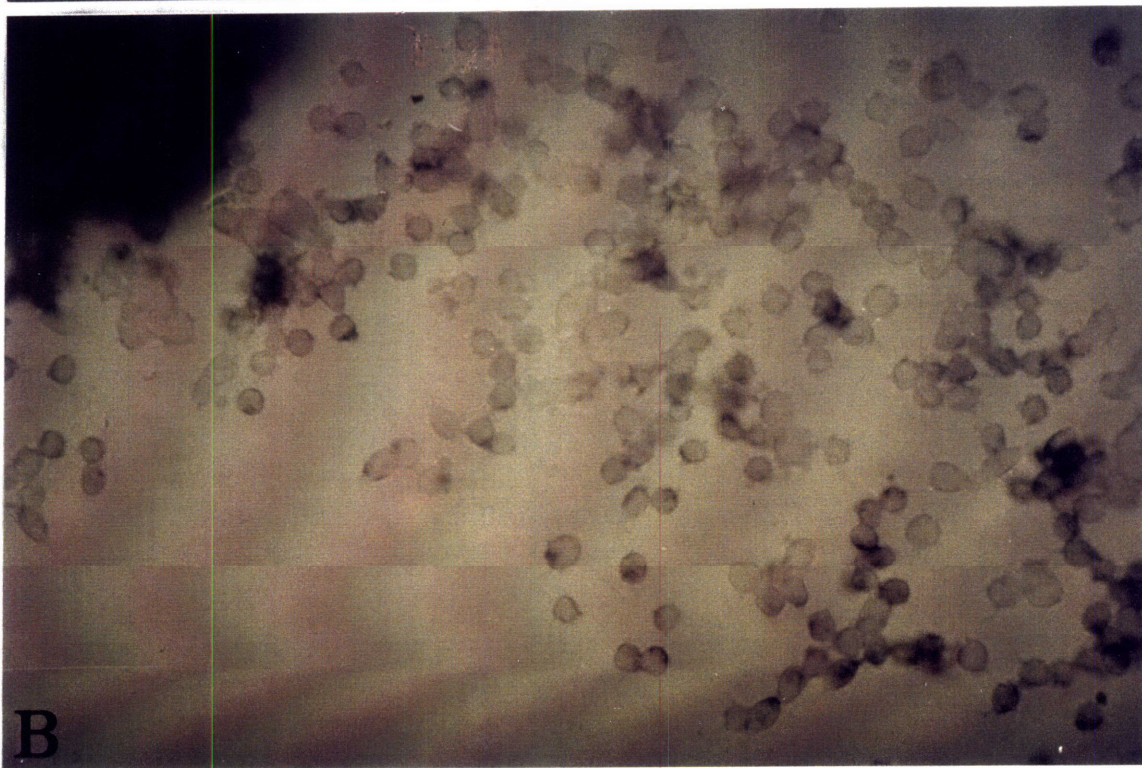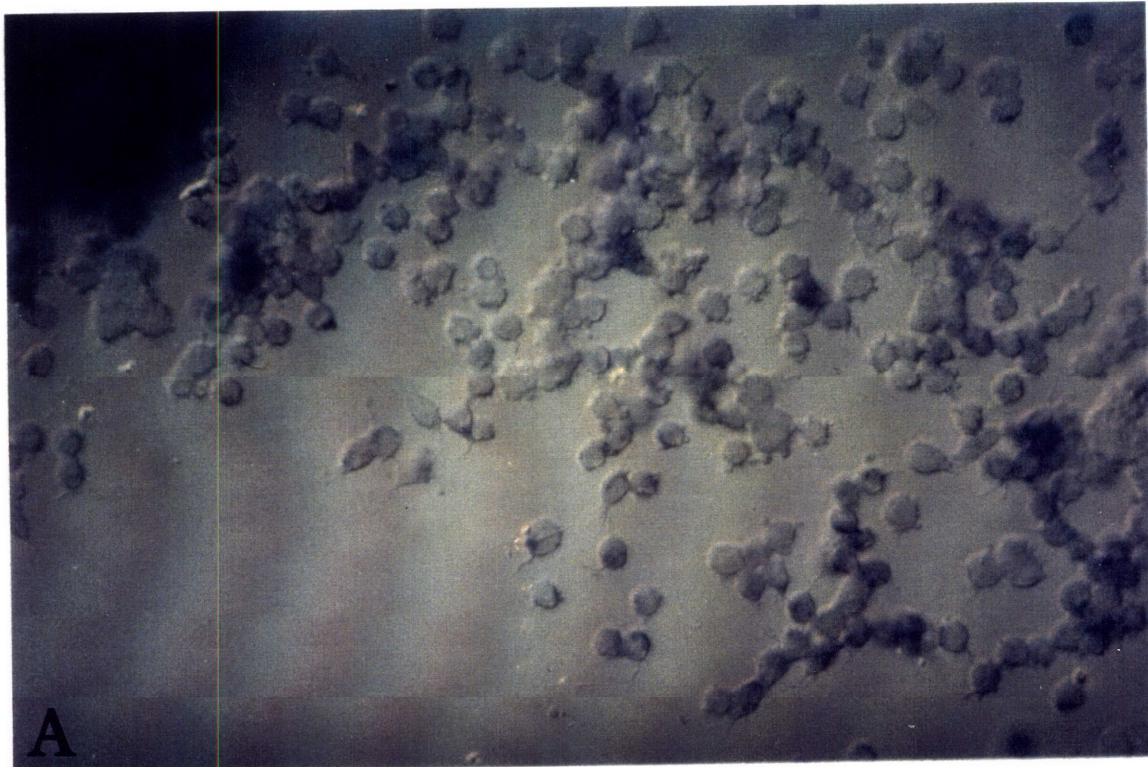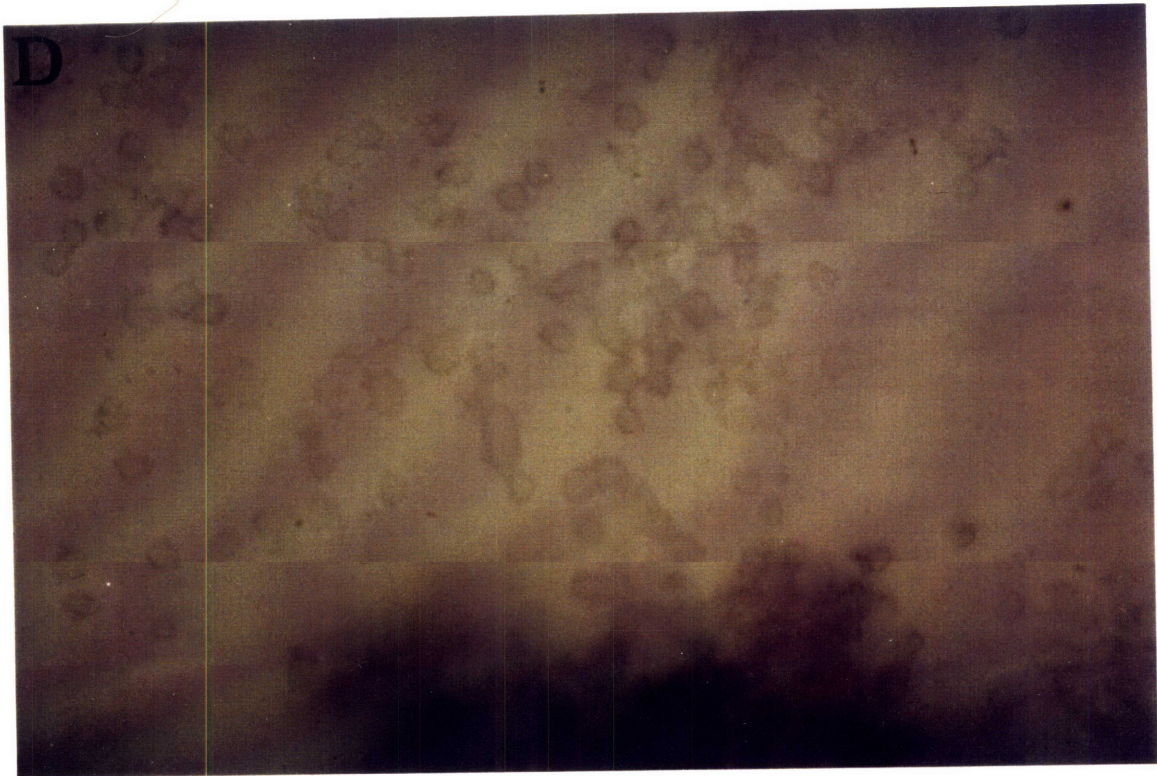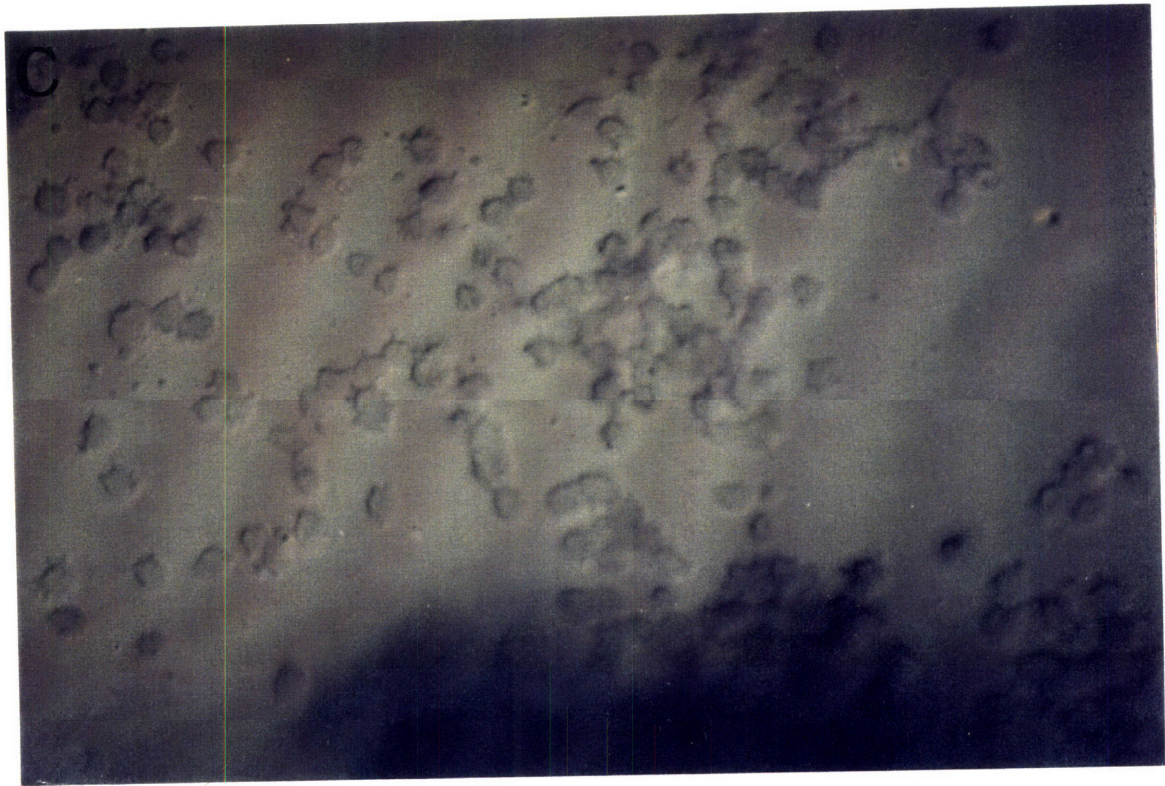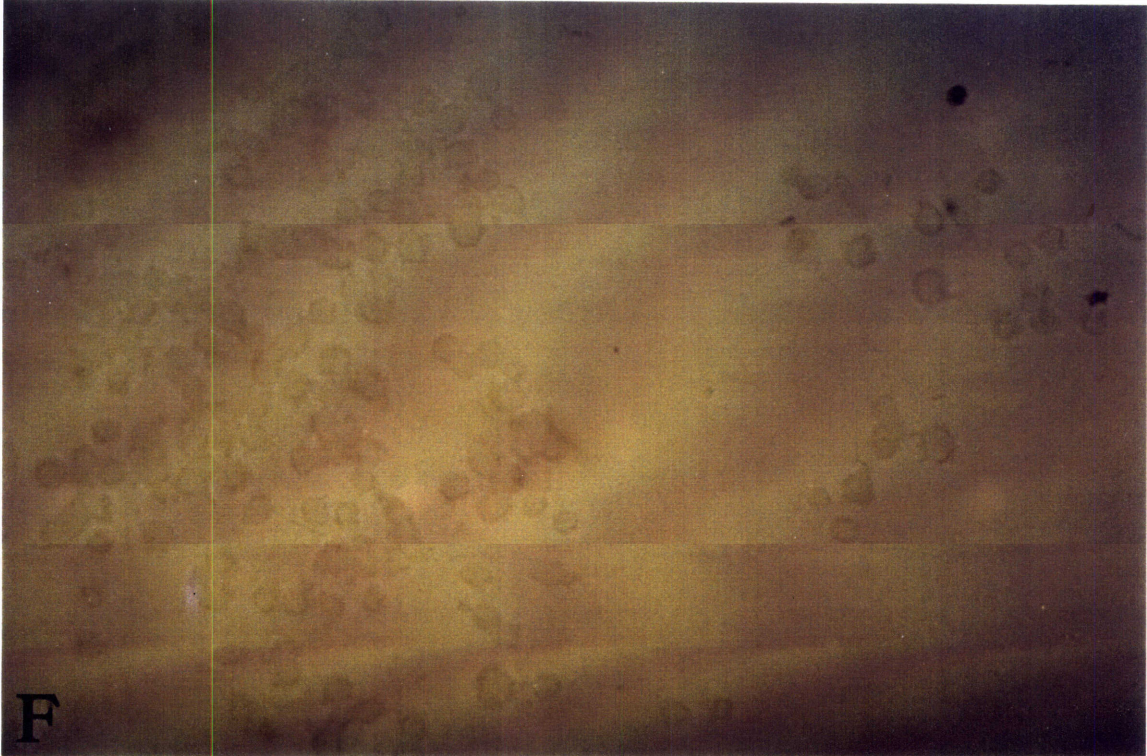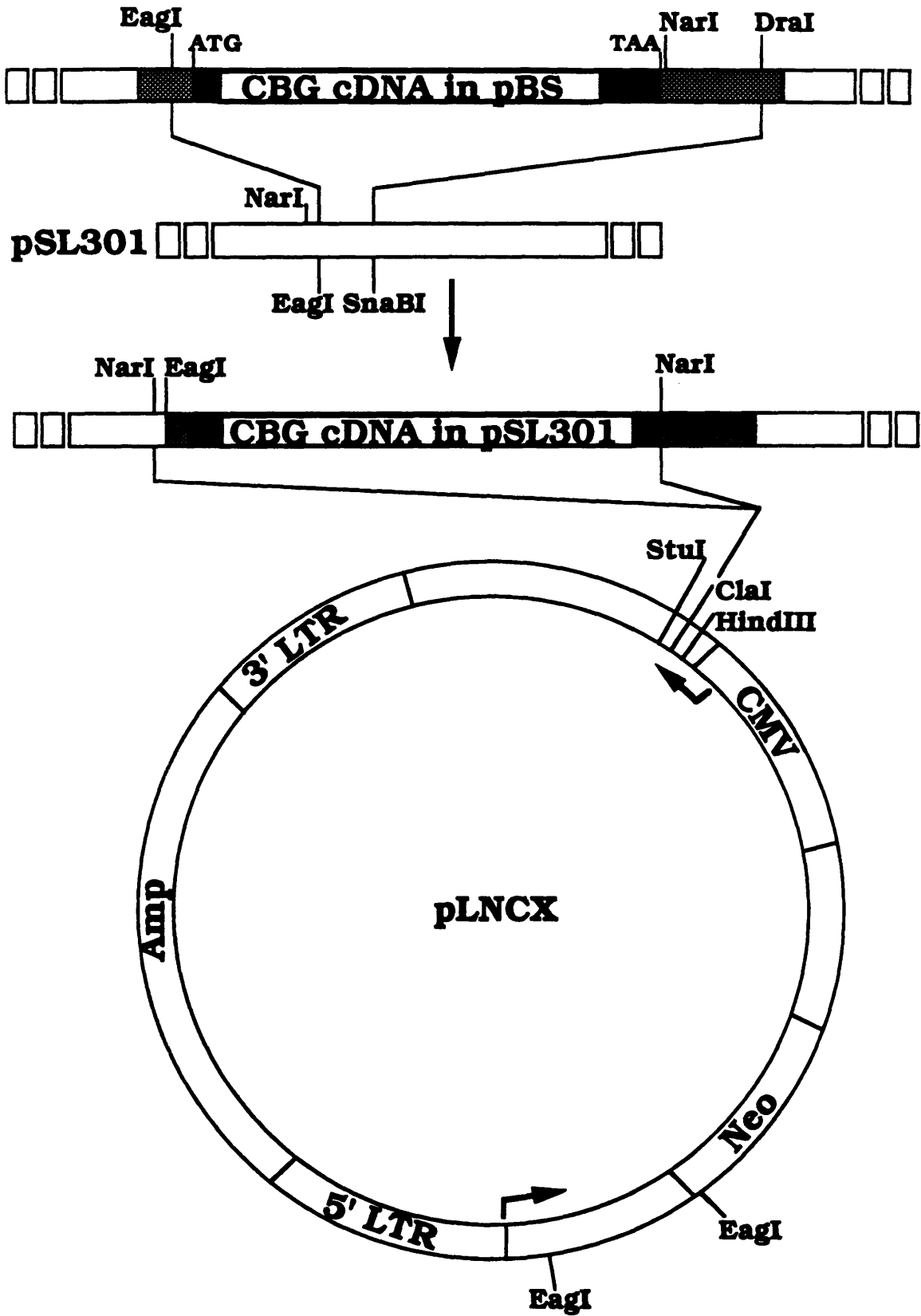
Figure 9. Construction of a CBG expression vector
The cloning strategy and structural features of the CBG expression vector LNCX-CBGΔ3' are shown schematically. Vector sequences are shown in white, CBG cDNA is shaded, and the CBG coding region is black. The start and stop codons for CBG coding sequence are indicated. Abbreviations: pBS, pBluescript plasmid (Stratagene); pSL301, Super-linker plasmid (Invitrogen); LTR, long terminal repeat; Amp, ampicillin resistance selectable marker; Neo, neomycin resistance selectable marker; CMV, cytomegalovirus promoter. Arrows indicate direction of transcription. The orientation of the CBG cDNA in LNCX-CBGΔ3' was verified by restriction digest using Eag I, and by sequencing with primers based on vector sequence on either side of the insert.

EagI ATG ... TAA NarI DraI

CBG cDNA in pBS

NarI

pSL301

EagI SnaBI

NarI EagI NarI

CBG cDNA in pSL301

StuI

ClaI
/HindIII

3' LTR

CMV

Amp

pLNCX

Neo

5' LTR

EagI

EagI

**Figure 10. CBG expression in 3T3 and PC12 cell lines**
NIH3T3 cells and PC12 cells were infected with the CBG expression
vector LNCX-CBGΔ3' to create the cell lines 3T3inf$_u$CBGΔ3' and
PC12inf$_u$CBGΔ3', which were maintained as uncloned populations in
selective medium. (A) and (B): 3T3inf$_u$CBGΔ3' cells plated on
fibronectin, and viewed with the 40X objective of a confocal
microscope. (A) Cells stained with the 521-2 anti-CBG antibody; (B)
No first antibody negative control. (C) and (D) PC12inf$_u$CBGΔ3' cells
plated on laminin and viewed with the 40X objective of a confocal
microscope. (C) PC12inf$_u$CBGΔ3' cells stained with 521-2; (D)
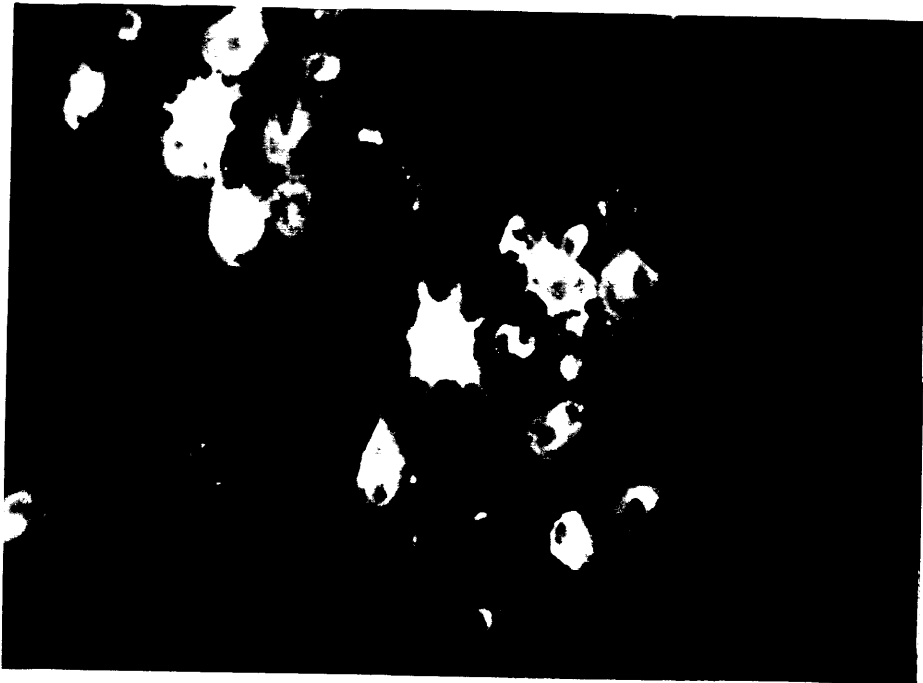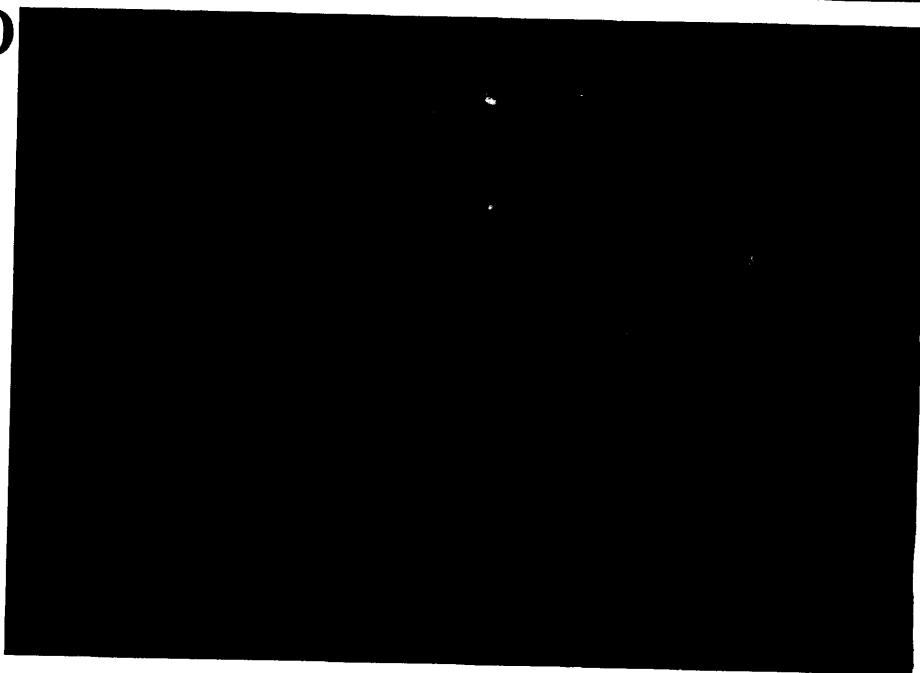parental PC12 cell line stained with 521-2. Bar = 50 μm.

C

D

198
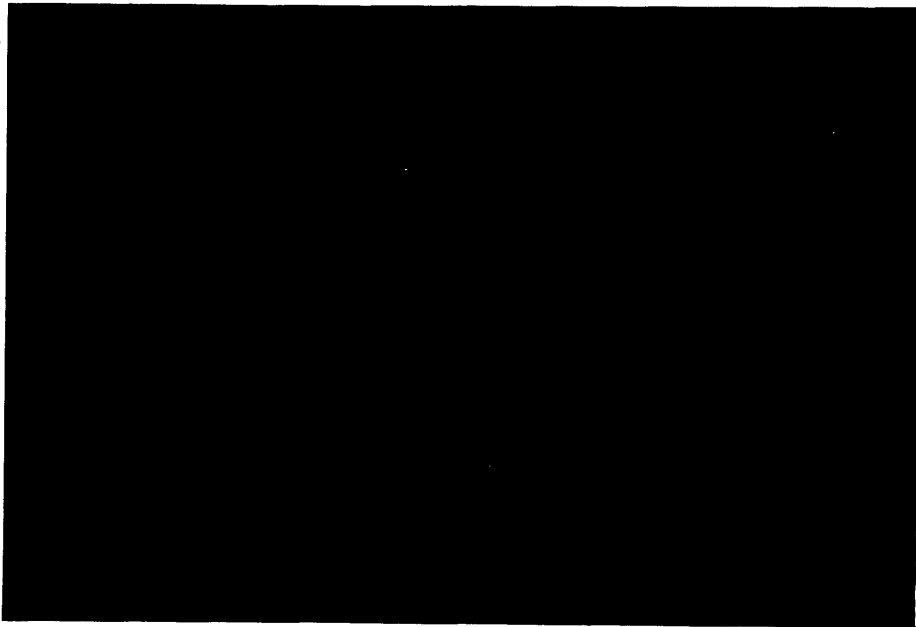
Figure 11. Distribution of CBG protein in cells of CBG expression lines

(A-C) 3T3inf_uCBGΔ3' cells were plated on laminin, and CBG protein was visualized by staining with the 521-2 anti-CBG antibody and a Cy-3 conjugated secondary antibody. Cells were viewed with a 100X objective, 2X zoom, with the aperture nearly closed to allow for the thinnest possible optical sections. **(A)** View at the substrate level. **(B)** View at an intermediate level. **(C)** View at the dorsal cell surface. Bar = 10 μm.

(D-F) 3T3inf_uCBGΔ3' cells plated on fibronectin, stained with 521-2, and viewed with the 63X objective of a confocal microscope. **(D and E)** Staining was observed throughout the entirety of these cells, even in the nucleus (compare to A and B above, where the nucleus is readily visualized by lack of staining; also, compare Figure 10A to Figure 10C). **(F)** The same cell as in (E) but at a slightly different focal plane and higher contrast to illustrate the nuclear staining in this cell. Bar = 25 μm.

(G-I) PC12inf_uCBGΔ3' cells plated on laminin, stained with 521-2, and visualized with the 100X objective of a confocal microscope, with the aperture nearly closed. **(G)** View at substrate level. **(H)** View at an intermediate level. **(I)** View at the dorsal cell surface. Bar = 25 μm.
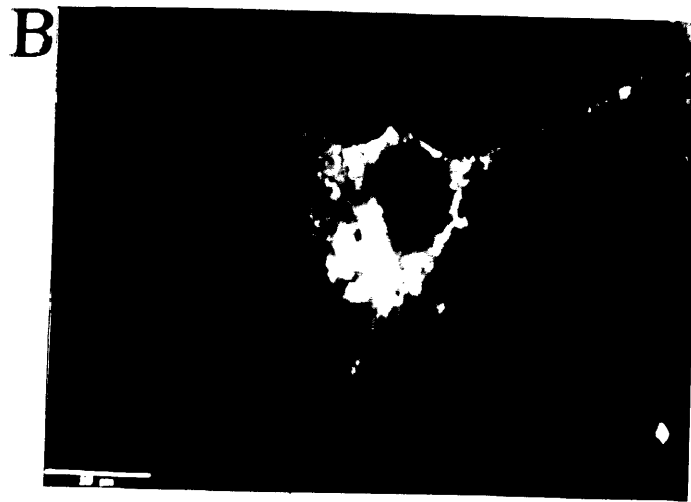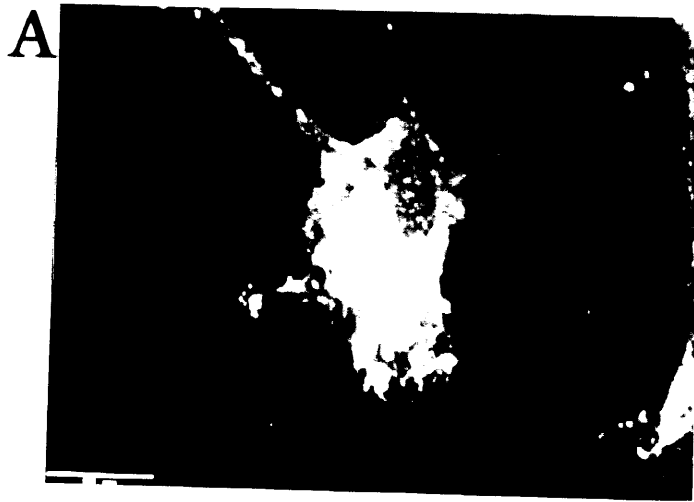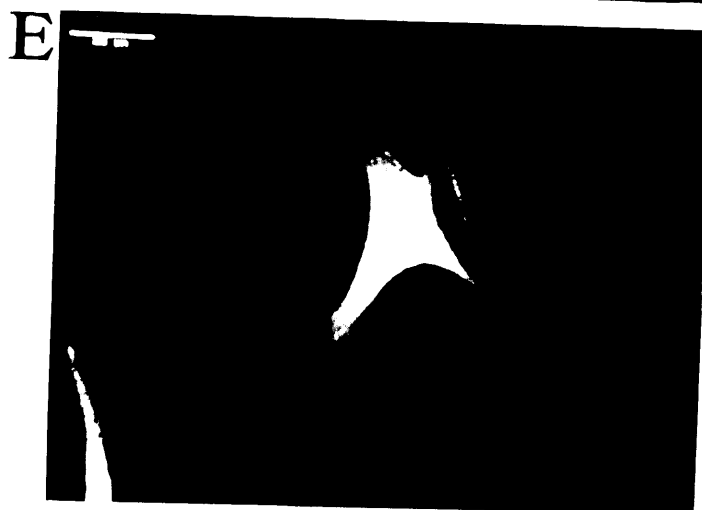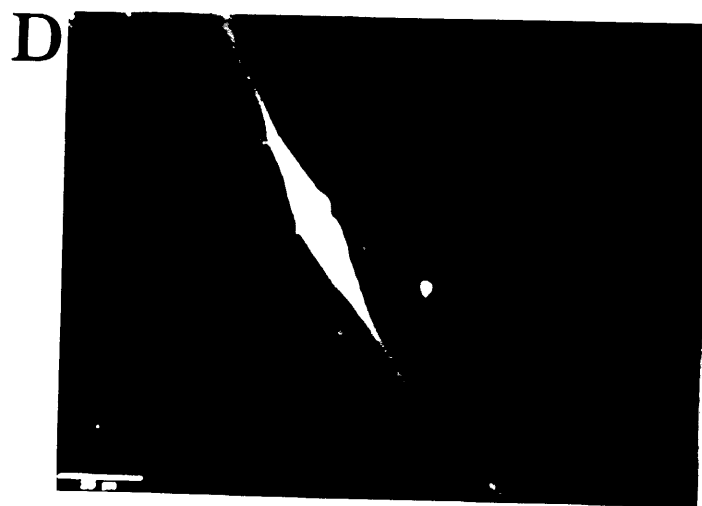
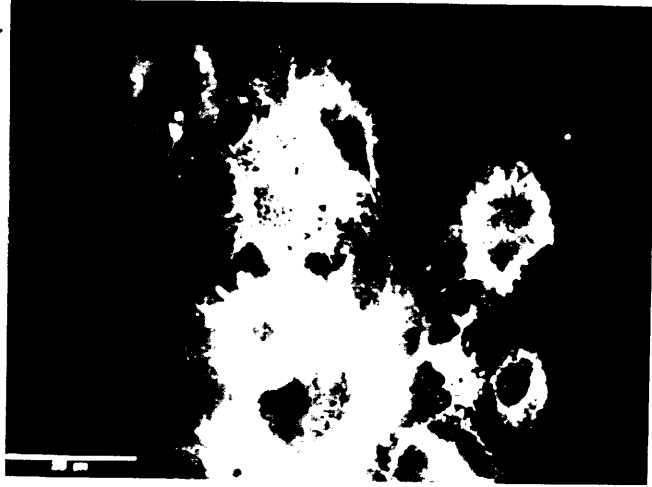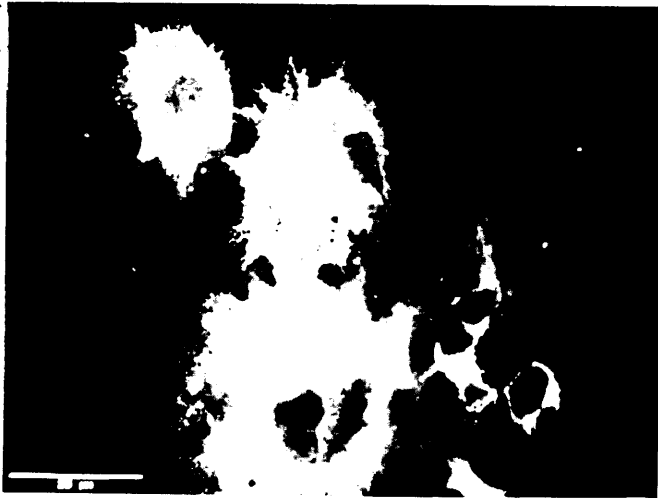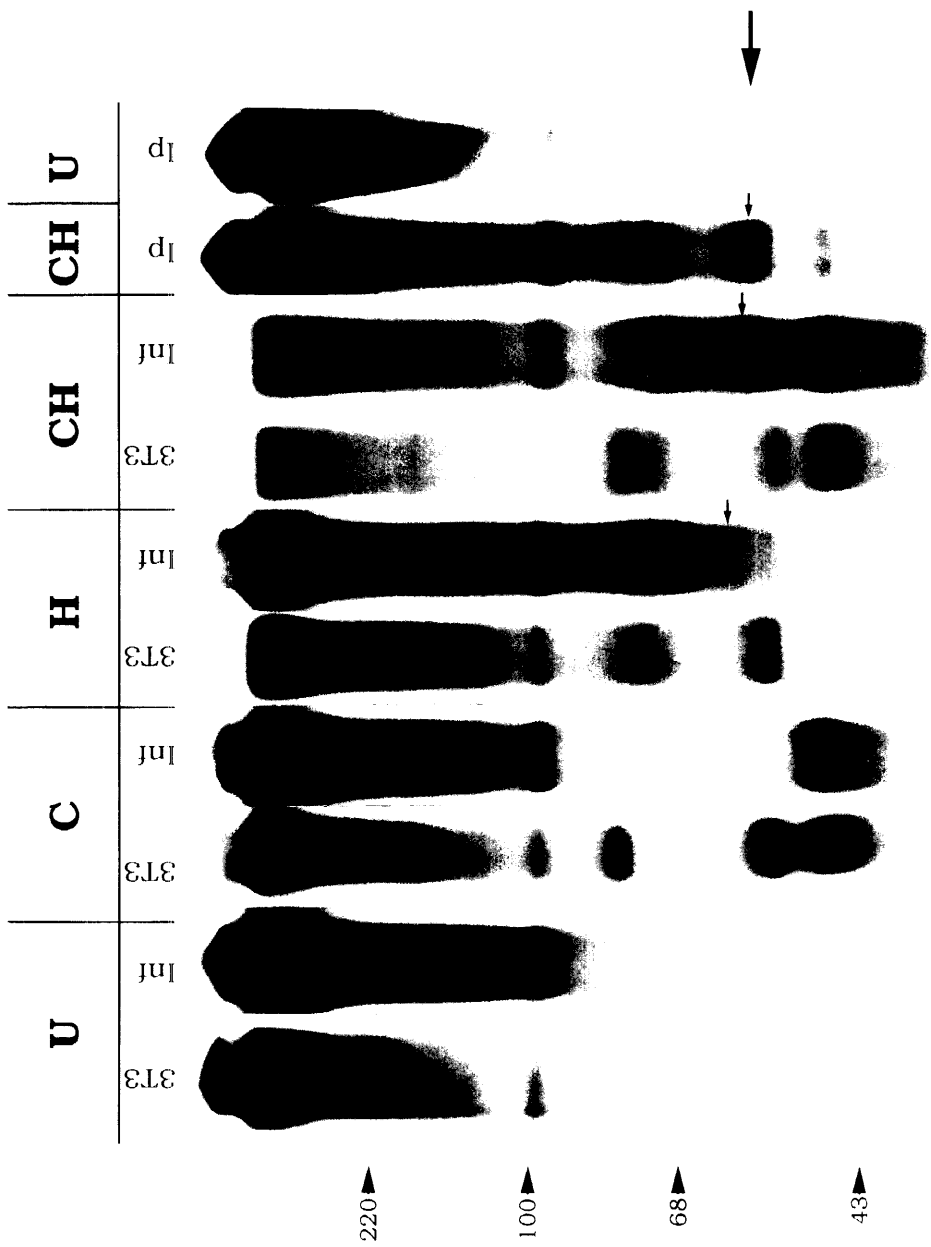**Figure 12. The CBG cDNA is expressed as a membrane associated HSPG in 3T3 cells**

PGs were isolated (as in Chapter 2) from the membrane fraction of 3T3 cells and the 3T3 cell line 3T3inf$_u$CBGΔ3', that had been infected with the CBG retroviral expression vector. PGs were labelled with ^{125}I as previously described (Herndon and Lander, 1990) and analyzed by SDS-PAGE after digestion with heparitinase or chondroitinase, to visualize PG core proteins. Additionally, undigested and digested PGs from 3T3inf$_u$CBGΔ3' cell membrane fraction were immuno-precipitated with the 521-2 anti-CBG antibody and the immuno-precipitates were analyzed on the same SDS-PAGE minigel. Abbreviations: **U**, undigested PGs; **C**, chondroitinase ABC digested PGs; **H**, heparitinase digested PGs; **CH**, PGs digested with both heparitinase and chondroitinase; **3T3**, PGs from parental 3T3 cell membrane fraction; **Inf**, PGs from the 3T3inf$_u$CBGΔ3' cell membrane fraction; **Ip**, PGs from 3T3inf$_u$CBGΔ3' cell membrane fraction, immunoprecipitated with 521-2. The large arrow at the side and the small arrows within the gel indicate the CBG core protein bands. Molecular weight markers are indicated. The table below shows the cpm recovered from immunoprecipitations with the 521-2 antibody of membrane fraction PGs from 3T3inf$_u$CBGΔ3' cells and parental 3T3 cells.

cell line	cpm input	cpm recovered
3T3	3.9×10^6	16,000
3T3inf$_u$CBGΔ3'	3.9×10^6	375,000

CHAPTER 4

HIGH AFFINITY INTERACTION
OF CEREBROGLYCAN WITH LAMININ-1

205

INTRODUCTION

The previous chapters have described the identification, cloning, and developmental expression of a novel cell surface heparan sulfate proteoglycan (HSPG) that we have named cerebroglycan (CBG). CBG is expressed only in the nervous system, and only during development. CBG expression appears in neurons around the time of their final mitoses, and disappears shortly after cell body migration and axon outgrowth have been completed. Immunolocalization of the CBG protein core has revealed that CBG is highly expressed in axon tracts throughout the developing nervous system, and, in one favorable case --the granule cell neurons of the dentate gyrus -- it has been possible to demonstrate that CBG may actually be polarized to axons (i.e., expression was found associated with axons and not cell bodies or dendrites) (Litwack, 1995).

The available data are consistent with the possibility that CBG plays a role in axon outgrowth during neuronal development. CBG expression is associated with axons during the period when outgrowth is occurring, but is not present later, when outgrowth has been completed. Biochemical data have also demonstrated that CBG (as well as glypican and syndecan-3) is enriched in growth cone particles (GCPs) (Ivins et al., 1995). These particles morphologically resemble growth cones when visualized by electron microscopy (Pfenninger et al., 1983) and are enriched for markers such as GAP-43 that are known to be localized to growth cones (Skene et al., 1986). The apparent abundance of CBG protein in growth cones, the sensory structures responsible for axon guidance, is further evidence that CBG may be involved in the growth of axons.

As a cell surface HSPG, CBG has the potential to interact with several extracellular cues in the developing nervous system that may influence axon outgrowth. The heparin-binding molecules fibronectin, netrins 1 and 2, thrombospondin, vitronectin, and laminins 1 and 3 have all been implicated in axon outgrowth and are all expressed in the developing nervous system at places and times that overlap with CBG expression (Cohen et al., 1987; McLoon et al., 1988; Rogers et al., 1989; Neugebauer et al., 1991; Sheppard et al., 1991; Hunter et al., 1992; Kennedy et al., 1994; Serafini et al., 1994). To begin to assess

the potential of CBG to serve as a receptor for such molecules, we have chosen to focus on laminin-1 (LN-1).

LN-1 is a good candidate ligand for CBG for several reasons. LN-1 contains multiple heparin-binding domains, potentially capable of interacting with a cell surface HSPG (Sung et al., 1993; Colognato-Pyke et al., 1995). The basement membrane HSPG perlecan copurifies with LN-1 (Lander et al., 1985), and perlecan heparan sulfate (HeS) chains bind to LN-1 (Yurchenco and O'Rear, 1993) suggesting that at least one of these domains actually binds HSPGs in vivo. LN-1 has potent neurite outgrowth promoting and guiding activity in vitro (Rogers et al., 1983; Gundersen, 1987) and is expressed in vivo in CBG-positive axon tracts, such as the subplate of the developing cerebral cortex, and the optic tract [(Cohen et al., 1987; McLoon et al., 1988; Hunter et al., 1992); Chapter 3, this thesis].

Evidence that LN-1 is actually used as a substrate for axon outgrowth in vivo has come from studies of the avian optic tract. LN-1 is expressed in the avian optic tract during the time when retinal ganglion cells (RGCs) are extending axons through the tract to targets in the optic tectum. At embryonic day 6 (E6), RGC axon outgrowth in vivo is occurring, and E6 RGCs respond to a LN-1 substrate in vitro by extending neurites. By E11, RGC axon outgrowth in vivo is essentially complete, and E11 RGCs cultured in vitro no longer respond to a LN-1 substrate (Cohen et al., 1986; Cohen et al., 1987). Interestingly, the loss of RGC responsiveness to LN-1 correlates with the downregulation of $\alpha6$ integrin subunit by these cells (deCurtis et al., 1991). This same integrin subunit has been implicated in motile responses of olfactory receptor neurons (Calof et al., 1994) and fibroblasts (Lin and Bertics, 1995) to LN-1.

Despite the potential for cell surface HSPG/LN-1 interactions in the developing nervous system, little is known about cell surface HSPG binding to LN-1. We report here on studies that demonstrate that CBG binds to LN-1 with sub-nanomolar affinity, around four hundred-fold better than the affinity of heparin for LN-1. This high affinity binding to LN-1 requires an intact CBG core since trypsin digestion of CBG produced fragments that bound with affinities similar to that of

heparin. In addition, two other neuronal HSPGs, glypican and syndecan-3/N-syndecan, were also found to bind LN-1 with high affinity. These results are consistent with the view that cell surface HSPGs could serve as LN-1 receptors or coreceptors in the developing nervous system.

MATERIALS AND METHODS

Materials

Laminin-1 was purified from Engelbreth-Holm-Swarm (EHS) sarcoma according to published methods (Kleinman et al., 1982; Timpl et al., 1982). The recombinant G domain (rG) of LN-1 was the gift of Dr. Peter Yurchenco, (Dept. of Pathology, UMDNJ-Robert Wood Johnson Medical School, Piscataway , NJ.) FGF-2 (bFGF) was purified from bovine brain (Lobb and Fett, 1984). Immunopurified ^{125}I-labelled syndecan-3 from postnatal day 0 (P0) rat brain growth cone particles (see below) was kindly supplied by John Ivins (UCI, Irvine, CA). Syndecan-3 was precipitated using the mAb MSE-3 (Kim et al., 1994), supplied by the lab of Merton Bernfield (Children's Hospital/Harvard Medical School, Boston, MA). Metabolically labelled heparan sulfate (HeS) chains were supplied by Mary Herndon (Dept. of Pathology, Harvard Medical School, and were prepared essentially as described in the Materials and Methods section of Chapter 5, this thesis). Thrombospondin-1 was supplied by Jack Lawler, Dept. of Pathology, Harvard Medical School). Porcine intestinal heparin was from Sigma. Trypsin was from Worthington Biochemical Co. Chondroitinase, heparitinase, pepstatin A, PMSF (phenylmethyl-sulfonyl fluoride), DEAE-Sephacel, and CHAPS (3-[(3 Cholamidopropyl) dimethyl-ammonio]-1-propanesulfonate) were obtained as described in Chapter 2. All other chemicals were obtained from Mallinkrodt except where noted.

Preparation of ^{125}I-Labelled, Low Molecular Weight Heparin

Heparin was substituted with tyramine as described in San Antonio, et al. (1993). 3-10 μg of heparin in 50 μl of 0.25 M NaPO$_4$, pH 7.5 and 2-5 mCi of Na^{125}I (NEN) were mixed in a 13 x 75 mm glass tube, that had previously been coated with 20 μg of Iodogen (Pierce) (50 μl of 0.4 mg/ml Iodogen in dichloromethane). ^{125}I-labelled GAGs were separated from free iodine by filtration over a 2 ml G-25 (Pharmacia) column that had been pre-blocked with 200 μg of crystalline BSA. For isolation of low molecular weight heparin (LMW-heparin), ^{125}I-labelled heparin was further fractionated over a G-100 (Pharmacia) gel filtration column: the low molecular weight pool

(Mr<6000) was isolated as the last 11% of the heparin peak to elute (San Antonio et al., 1993). ^{125}I-labelled LMW-heparin was radioprotected with the addition of 0.25 mg/ml crystalline BSA, aliquotted, and stored at -80°C. LMW-heparin is the heparin used in all of the binding assays described below.

Preparation of Anti-Cerebroglycan and Anti-Glypican Affinity Matrices

Affinity purified anti-CBG (521-2) and anti-glypican (343-1) polyclonal antibodies were isolated as described [Chapter 3, this thesis; (Litwack, 1995). 521-2 and 343-1 antibodies were bound to protein-A sepharose and crosslinked with dimethyl-pimelimidate (DMP) according to instructions, using an Immunopure Protein A IgG Orientation Kit (Pierce). The resulting affinity matrices contained 0.6 mg/ml (packed volume) of 521-2 antibody, and 0.3 mg/ml (packed volume) of 343-1 antibody.

Proteoglycan Isolation, Purification, and Analysis

PGs were isolated from the membrane fraction of embryonic day 16 (E16) rat brain using a protocol described previously [(Herndon and Lander, 1990); Chapter 2, this thesis]. PGs were labelled with ^{125}I according to Herndon and Lander (1990). ^{125}I-labelled PGs from the membrane fraction of the 3T3 cell line 3T3inf$_{u}$CBGΔ3', that has been infected with a CBG retroviral expression vector, were prepared as described in Chapter 3.

Growth cone particles (GCPs) were isolated from P0 rat brains essentially as described previously (Pfenninger et al., 1983). Briefly, P0 rat brains were homogenized in 0.32 M sucrose, 1 mM MgCl$_2$, 1 mM TES ([N-Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid) (pH7.3) in a teflon-on-glass homogenizer. The homogenate was centrifuged at 1660 g for 15 minutes, and the supernatant was loaded onto a discontinuous sucrose gradient and centrifuged for 40 minutes at 242,000 g. The GCPs were collected from the interface between the sample load and the first step (0.75 M sucrose) in the gradient. This fraction is substantially enriched for GCPs, but is also heavily contaminated with soluble protein (Pfenninger et al., 1983).

To isolate PGs from GCPs, GCPs at ~1 mg/ml total protein in 1 mM TES, 1 mM $MgCl_2$, ~0.65 M sucrose, were made 50 mM in Tris-HCl [pH 7.4], 0.15 M in NaCl, and 1.0% in CHAPS, and vortexed for 1 minute. After centrifuging for 3 minutes at 12,000 g in a microfuge, the supernatant was recovered and pepstatin A and PMSF were added at 1 μg/ml and 0.4 mM respectively. The supernatant was then loaded to a DEAE-Sephacel column, and PGs were isolated as described previously (Herndon and Lander, 1990). After elution, PGs were radiolabelled with ^{125}I as previously described (Herndon and Lander, 1990).

Immunopurified cerebroglycan and glypican were prepared using the affinity matrices described above. ^{125}I-labelled PGs from E16 rat brain or from P0 rat brain growth cone particles, both in 0.75 M NaCl, 50 mM Tris-HCl [pH7.4], 0.5% CHAPS, were diluted five fold with 50 mM Tris-HCl, 0.5% CHAPS, and crystalline BSA (ICN) was added to a final concentration of 0.1 mg/ml; 0.05 to 0.1 volumes (packed beads) of affinity matrix were added, and PGs were allowed to bind overnight at 4°C. Beads were washed for 20-60 minutes with rotation at room temperature with 1 ml each of (1) 50 mM Tris-HCl, 0.15 M NaCl, 0.5% CHAPS and (2) 0.5M NaCl, 50 mM Tris-HCl, 0.5% CHAPS, and then briefly with 10 mM Tris-HCl, 0.5% CHAPS. Bound PGs were eluted with 100 mM NaH_2PO_4/H_3PO_4 [pH 2.5], 0.5% CHAPS for 10 minutes and immediately neutralized with 0.2 volumes of 0.5 M Na_2HPO_4.

For reduction and alkylation, 10 μl of 1M Tris-HCl [pH 8.5], 50 μl of 8M urea, and 10 μl of 0.2 M dithiothreitol were added to 30 μl (~90,000 cpm) of immunopurified CBG, prepared as described above from the E16 rat brain membrane fraction. After 30 minutes at 37°C, 10μl of 0.5M iodoacetamide was added and the sample was incubated for an additional 45 minutes at 37°C. The sample was diluted with 1.9 ml of 50 mM Tris-HCl [pH 7.4], 0.5% CHAPS, and loaded to a Centricon-30 tube (Amicon) that had been pre-blocked with 0.5 mg/ml crystalline BSA in Tris/CHAPS. The buffer was exchanged to 50 mM Tris-HCl, 0.5% CHAPS, and the sample was concentrated by succesive rounds of dilution and centrifugation.

For trypsin digestion, 6 μl (36 μg) of freshly prepared trypsin (TRTPCK grade, Worthington Biochemical Co.) in phosphate buffered saline (PBS) was added to 30 μl (~90,000 cpm) of immunopurified CBG, prepared as described above, and the sample was incubated for 2 hours at 37°C. 12 μl (72 μg) of soybean trypsin inhibitor in PBS was added, and the sample was incubated an additional 1 hour at 37°C. Aliquots of the digest were analyzed on an 18% SDS-PAGE minigel, with or without prior digestion with heparitinase [as previously described, (Herndon and Lander, 1990)]. At least one ^{125}I-labelled, glycanated peptide was obtained which migrated as a smear with a Mr of ~40 kDa prior to digestion, and as a ~7.5 kDa band after heparitinase digestion.

Affinity coelectrophoresis (ACE) and Analysis

Affinity co-electrophoresis (ACE) and its analysis was carried out essentially as described (Lee and Lander, 1991; Lim et al., 1991) (See also, Figure 2). ACE buffer was 50 mM MOPSO [3-(N-morpholino)-2-hydroxypropane-sulfonate], 125 mM NaAcetate pH 7.0. Dried ACE gels were exposed in a phosphorimager cassette (Molecular Dynamics, Sunnyvale, CA), analyzed using ImageQuant Software (Molecular Dynamics) and quantified manually or as described in San Antonio et al. (1993), using Microsoft Excel. Kaleidagraph (Synergy Software) was used to fit electrophoretic mobility data to equilibrium equations. Briefly, R values (retardation coefficient, the electrophoretic mobility of bound PG or heparin in a given lane normalized to the mobility of unbound PG or heparin) were graphed with respect to the protein concentration for each lane and the resulting curve was fit to the equation:

$$R = \frac{R_{\infty}}{1 + (K_d/[P_{tot}])}$$

where R_{∞} = R value when the PG or heparin is maximally shifted, K_d = Equilibrium dissociation constant, [G][P]/[G-P], for the reaction: G (PG or heparin) + P (test protein) ⇔ G-P (complex of PG/heparin with test protein); [P_{tot}] = Protein concentration in each lane of the ACE gel (at equilibrium, this value should actually be ([P_{tot}] - [G-P]),

but because PG concentrations are so low under normal ACE conditions ($<0.1K_d$), the G-P term is negligible); n = Number of protein molecules binding to each PG or heparin molecule, i.e. order of binding mechanism (usually taken to be 1).

Overruns (cases where the PG or heparin migrates out of the protein-containing zone of a particular lane) were compensated for using methods also outlined in Lim et al. (1991).

RESULTS

Cerebroglycan binds to laminin-1 with high affinity

To begin to assess the potential for CBG-laminin-1 (LN-1) interactions, immunopurified CBG was prepared from embryonic day 16 (E16) rat brain for use in binding studies. ^{125}I-labelled PGs isolated from E16 rat brain (a rich source of CBG; data not shown) were immunoprecipitated using a 521-2 anti-CBG affinity matrix, as described in Materials and Methods. As shown in Figure 1 this step resulted in a substantial purification of CBG. The eluate from the affinity matrix contained a single HSPG core protein of Mr ~53 kDa, similar to the size previously observed for the CBG core protein (Herndon and Lander, 1990). Several non-PG proteins including one that comigrates with the CBG core protein -- compare the "undigested" lanes in starting material and CBG samples in Figure 1 -- were substantially removed by the immunopurification step.

Immunopurified CBG was tested for binding to LN-1 using affinity coelectrophoresis (ACE). In ACE, a horizontal, 1% agarose gel is cast with 9 lanes arranged orthogonally to a thin sample well (see Figure 2). A test protein at nine different concentrations is mixed with agarose and embedded in the 9 lanes. A PG or glycosaminoglycan (GAG) that has been radioiodinated or metabolically labelled is loaded into the sample well and driven by electrophoresis through the lanes containing the test protein. If a binding interaction occurs, the PG or GAG will be retarded to an extent that depends on the dissociation constant (K_d) for the interaction (Lee and Lander, 1991; Lim et al., 1991).

Figure 3A shows the result when ^{125}I-labelled CBG is tested for binding to LN-1 by ACE; CBG is found to bind LN-1 with a high affinity. The CBG sample is near maximally retarded by concentrations of LN-1 as low as 5 nM, but ^{125}I-heparin run in parallel is not significantly retarded by concentrations of LN-1 lower than 50 nM, and is not maximally retarded even at 200 nM, the highest concentration tested (Figure 3B). The affinity patterns in Figure 3 (A and B) were converted to the binding isotherms, as described in Materials and Methods. As shown in Figure 3C, CBG bound to LN-1 with an apparent K_d of 0.18 nM while a K_d of 203 nM (similar to values reported by others (San

Antonio et al., 1993)) was obtained for heparin -- a 1000-fold difference in affinities.

The high affinity interaction of CBG with LN-1 requires an intact CBG core protein.

To further characterize the interaction of CBG with LN-1, immunopurified ^{125}I-labelled CBG was digested exhaustively with trypsin. The resulting material contained at least one glycanated, ^{125}I-labelled peptide with a Mr of ~7.5 kDa, as determined by SDS-PAGE, with or without prior heparitinase digestion (not shown). Tryptic CBG fragments bind LN-1 with an affinity similar to that of free heparin, much lower than that of intact CBG run in parallel (Figure 4), indicating that an intact CBG core protein is required for high affinity binding to LN-1. Interestingly, trypsinized CBG and intact CBG bind to FGF-2 (bFGF) with very similar affinities (Figure 4). This indicates that an intact core protein is not a requirement for maximal binding of CBG to all GAG-binding proteins, and suggests that trypsin treatment did not non-specifically decrease the affinity of CBG for GAG-binding molecules (e.g. by damaging the HeS chains).

Further characterization of LN-1 binding to cell surface HSPGs

To more completely assess the potential for LN-1 binding to cell surface HSPGs, LN-1 affinity for several additional molecules was tested. The results of these experiments are presented in Table 1. In addition to CBG, two other neuronal HSPGs were found to bind LN-1 with high affinities. Glypican and syndecan-3 were immuno-precipitated from P0 rat brain growth cone particles using the polyclonal 343-1 anti-glypican antiserum (Litwack, 1995) and the monoclonal antibody MSE-3 (Kim et al., 1994). Immunopurified glypican and syndecan-3, with K$_d$s of 2.2 and 2.7 nM respectively, both bound to LN-1 with much higher affinities than does heparin. CBG isolated from other sources also binds to LN-1 with high affinity, including CBG immunopurified from 3T3 cells infected with a CBG retroviral expression vector. Taken together, these results suggest that high affinity binding may be a common feature of LN-1 interactions with PGs.

In further experiments aimed at elucidating the basis of the high affinity interaction of LN-1 with CBG, CBG that had been reduced and alkylated was tested for affinity to LN-1. While the diffuse migration of reduced CBG made it difficult to accurately fit a single binding isotherm to the resulting ACE binding pattern, it was clear that reduction and alkylation had substantially reduced the over-all affinity of CBG for LN-1, (Figure 5 and Table 1), suggesting that tertiary structure is important for high affinity CBG binding to LN-1.

To explore the LN-1 structural requirements for a high affinity interaction with CBG, CBG affinity for the recombinant G (rG) domain of LN-1 was tested. The G domain of LN-1 is at the carboxy terminus of the LN-1 $\alpha 1$ chain (formerly the A chain), and contains two heparin binding sites (although one of these sites may be masked in the intact LN-1 molecule) (Sung et al., 1993). The rG domain is the G domain of LN-1 expressed in a baculovirus system in Sf9 cells, and has been shown to possess heparin binding properties similar to that of the G domain isolated proteolytically from murine LN-1 (Sung et al., 1993). ACE analysis revealed that, although heparin binds to rG and LN-1 with similar affinity, CBG's affinity for rG is much lower than its affinity for intact LN-1 (Table 1). CBG binds to rG with a K_d of 29 nM, only about nine-fold better than the affinity of heparin for rG. This suggests that structural elements in LN-1 outside of the G domain are required for high affinity binding by CBG.

For comparison, the binding of CBG and syndecan-3 to thrombospondin-1 (TSP-1) was also examined. TSP-1 is another example of a large, multidomain, extracellular matrix glycoprotein that contains multiple heparin binding sites (Adams and Lawler, 1993). As shown in Table 1, both CBG and syndecan-3 bind TSP-1 with an affinity higher than that of metabolically labelled HeS chains isolated from neonatal rat brain material; however, the the affinity of the intact molecules is only around four to five-fold higher than the affinity of the free chains, much less dramatic than the difference seen with LN-1.

Finally, free HeS chains isolated from neonatal rat brain bound LN-1 with a lower affinity than the ^{125}I-labelled heparin used as a control in most of the CBG binding experiments described above. This result

suggests that the difference between the binding of free HeS and intact HSPGs to LN-1 is even greater than first estimated.

DISCUSSION

High affinity interactions between cell surface HSPGs and laminin-1

In the present study, the binding of three neuronal cell surface HSPGs to LN-1 has been measured. Glypican, syndecan-3, and CBG were all found to bind LN-1 with high affinity -- K_ds were 2.2, 2.7, and 0.5-0.7 nM for glypican, syndecan-3, and CBG respectively. These affinities are several hundred fold better than the affinity of heparin or brain heparan sulfate for LN-1. This result demonstrates that the binding of heparin or HeS to a ligand cannot be taken to be an accurate indicator of the affinity of an HSPG for the same ligand. In some cases (e.g. FGF-2, and TSP-1; Table 1) the affinities for HSPGs and HeS are similar, whereas for other molecules such as LN-1, the affinities are dramatically different. An accurate picture of HSPG binding to various ligands will require a case by case investigation, and will probably depend not only on the particular molecules involved, but the source from which they are derived.

For example, in one recent report, a glycosylphosphatidylinositol-anchored cell surface HSPG from Schwann cells (subsequently identified as glypican) was found to bind LN-1 with an affinity of only ~ 1 µM (Carey et al., 1990). It is possible that the lower affinity is due to experimental differences -- binding of the Schwann cell PG to LN-1 was measured by equilibrium gel filtration with LN-1 present throughout the elution buffer, which was 0.1 M Tris-HCl, pH 7.5. With time, solutions of LN-1 at low ionic strength can experience substantial non-specific losses (data not shown); such losses would result in an underestimation of LN-1 affinity. However, it is also possible that the discrepancy is simply the result of the different source of the glypican used in the binding experiments.

A more clear-cut example of a case where the same HSPG from two different sources exhibits two different affinities for the same ligand has also been recently reported (Sanderson et al., 1994). Syndecan-1 derived from MPC-11 myeloma cells binds to type I collagen with an affinity 20-fold better than does syndecan-1 derived from P3 cells, another myeloma cell line. This difference is observed even though the core proteins from the two sources are indistinguishable, and the HeS chains are of a similar size *and* net charge. Interestingly, the

difference in the affinities appears to have a tangible effect at the level of whole cells: MPC-11 cells adhere to type I collagen while P3 cells do not (Sanderson et al., 1994).

The binding of syndecan-1 to LN-1 has also been investigated. Several authors have reported on experiments in which syndecan-1 failed to bind to LN-1 in solid phase binding assays [c.f. (Saunders and Bernfield, 1988; Sanderson et al., 1992a)], a result that further underscores the importance of testing HSPGs individually for binding to a given ligand. However, solid phase assays have also indicated that although syndecan-1 binds type 1 collagen, it fails to bind to type IV collagen [for review see (Bernfield et al., 1992)], whereas affinity coelectrophoresis (ACE) has demonstrated that syndecan I binds type IV and type I collagens equally well (San Antonio et al., 1994). This serves to point out an advantage of the ACE technique: ligands can be tested for PG binding without adsorption to plastic or chemical crosslinking, either of which might significantly alter the biochemical properties of the ligand.

The molecular basis of CBG binding to LN-1

The binding of CBG to LN-1 occurs with an affinity that is ~300 fold higher than heparin affinity for LN-1, and ~1500 fold higher than the affinity of HeS from P0 rat brain membrane associated PGs for LN-1 (Table 1). These data indicate that a single HeS chain binding by itself to LN-1 is unlikely to account for the measured affinity of intact CBG for LN-1. Two possible reasons for the high affinity binding of CBG to LN-1 are (1) the HeS chain(s) of CBG bind simultaneously to multiple sites in LN-1, and (2) the CBG core protein contributes to the overall affinity by interacting with LN-1. While the experiments described above do not distinguish between these possibilities, several of the findings do bear on this issue.

For example, TSP-1 is a large multidomain glycoprotein of the extracellular matrix composed of three subunits, each of which contains two known heparin binding domains (Adams and Lawler, 1993). However, CBG, heparin, and HeS all bind to TSP-1 with similar affinities (Table 1). This indicates that the mere presence of multiple heparin-binding sites within a molecule is not sufficient for high affinity binding (relative to heparin) by CBG. If high affinity binding by

CBG is the result of multivalent HeS interactions, the arrangement of the HeS binding sites within the ligand would have to be important since both LN-1 and TSP-1 contain multiple HeS binding sites but only LN-1 is bound with high affinity. A similar argument can be made on the basis of CBG binding to the recombinant G (rG) domain of LN-1; rG also contains multiple heparin binding sites (Sung et al., 1993), but binds CBG only nine fold better than it binds heparin (Table 1).

Many factors could render the arrangement of heparin binding sites important. For example, HeS structure is known to be characterized by clusters of highly sulfated residues separated by spans of largely unmodified residues (Gallagher et al., 1992). The number of sites on any given HeS chain capable of interacting with HeS binding sites in a ligand molecule is therefore likely to be limited. If each chain contains an average of only one ligand binding site, multiple chains would be required to take advantage of multiple HeS binding sites on the ligand. If the HeS binding sites were spaced too close together, electrostatic repulsion between HeS chains could prevent both sites from being utilized simultaneously.

Even if multiple ligand binding sites were present on any one HeS chain, the spacing of the sites might preclude more than one site per chain from binding to a ligand. Given the fact that HeS and HeS binding molecules both have distinct domain structures, the unique geometry of each HeS-binding ligand seems likely to have a profound effect on its affinity for various HSPGs.

The fact that trypsinized CBG shows such a dramatically reduced affinity for LN-1 (Table 1) suggests that, if multiple interactions of HeS with LN-1 are indeed the basis of the high affinity binding of CBG to LN-1, then these interactions would have to be mediated by multiple HeS chains: if a single HeS chain were capable of binding to LN-1 with the same high affinity as intact CBG, then trypsin digestion of CBG would be unlikely to have such a dramatic effect on the K_d. HeS chains isolated from P0 rat brain also bind LN-1 with a much lower affinity than any of the intact HSPGs tested. This further supports the idea that CBG would need multiple HeS chains for high affinity binding to LN-1 if the high affinity were strictly due to multiple HeS chain interactions with LN-1.

An alternative basis for the high affinity of CBG for LN-1 is that an interaction between the CBG core protein and LN-1 contributes to the overall binding. The fact that reduction and alkylation of the CBG core has a marked effect on binding to LN-1 would seem to support this hypothesis. However, the issue of geometry raised above also applies in this case. It is possible that the tertiary structure of CBG is required for maximal LN-1 binding not because of a protein-protein interaction, but because the conformation of native CBG is necessary to present the HeS chains in the optimum geometry for LN-1 binding.

If a protein-protein interaction is responsible for the high affinity binding of CBG to LN-1, it appears to be relatively specific to LN-1 since the binding of CBG to rG and TSP-1 is only marginally better, if at all, than the binding of heparin or HeS, and the affinities of CBG and trypsinized CBG for FGF-2 are essentially identical. Examples of PGs that bind their ligands via the core protein have been reported elsewhere [e.g. the binding of TGF-β by the PG betaglycan, and the binding of collagen fibrils by decorin (Vogel et al., 1984; Krusius and Ruoslahti, 1986; Cheifetz and Massague, 1989).]

Summary

CBG, glypican, and syndecan-3 all bind to LN-1 much more strongly than would be predicted on the basis of heparin's affinity for LN-1. This result is consistent with the possibility that these neuronal cell surface HSPGs could serve as receptors or coreceptors for LN-1 during nervous system development. In the case of CBG, the high affinity binding requires an intact, native core protein suggesting that the conformation of CBG is important either for a core protein binding to LN-1, or for proper presentation of HeS chains for optimal LN-1 binding. High affinity binding (relative to heparin) is not a general feature of CBG interactions with all potential ligands, since CBG's affinity for TSP-1, FGF-2, and the rG domain of LN-1 is similar in each case to that of heparin. Regardless of whether LN-1 is an authentic ligand for CBG in vivo, these findings illustrate the importance of testing intact HSPGs for binding to individual ligands rather than relying on heparin affinity as an indicator of the potential for an interaction.

References

Adams, J. and J. Lawler. 1993. The thrombospondin family. Curr. Biol. 3:188-190.

Bernfield, M., R. Kokenyesi, M. Kato, M.T. Hinkes, J. Spring, R.L. Gallo, and E.J. Lose. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu. Rev. Cell Biol. 8:365-393.

Calof, A.L., M.R. Campanero, J.J. O'Rear, P.D. Yurchenco, and A.D. Lander. 1994. Domain-specific activation of neuronal migration and neurite outgrowth-promoting activities of laminin. Neuron 13:117-130.

Carey, D.J., D.M. Crumbling, R.C. Stahl, and D.M. Evans. 1990. Association of cell surface heparan sulfate proteoglycans of Schwann cells with extracellular matrix proteins. J. Biol. Chem. 265:20627-20633.

Cheifetz, S. and J. Massague. 1989. The TGF-β receptor proteoglycan: cell surface expression and ligand binding in the absence of glycosaminoglycan chains. J. Biol. Chem. 264:12025-12028.

Cohen, J., J.F. Burne, C. McKinlay, and J. Winter. 1987. The role of laminin and the laminin/fibronectin receptor complex in the outgrowth of retinal ganglion cell axons. Dev. Biol. 122:407-418.

Cohen, J., J.F. Burne, J. Winter, and P. Bartlett. 1986. Retinal ganglion cells lose response to laminin with maturation. Nature 322:465-467.

Colognato-Pyke, H., J.J. O'Rear, Y. Yamada, S. Carbonetto, Y.-S. Cheng, and P.D. Yurchenco. 1995. Mapping of network-forming, heparin-binding, and α1β1 integrin recognition sites within the α-chain short arm of laminin-1. J. Biol. Chem. 270:9398-9406.

deCurtis, I., V. Quaranta, R.N. Tamura, and L.F. Reichardt. 1991. Laminin receptors in the retina: Sequence analysis of the chick integrin α6 subunit. Evidence for transcriptional and posttranslational regulation. J. Cell Biol 113:405-416.

Gallagher, J.T., J.E. Turnbull, and M. Lyon. 1992. Patterns of sulphation in heparan sulfate: polymorphism based on a common structural theme. Int. J. Biochem. 24:553-560.

Gundersen, R.W. 1987. Response of sensory neurites and growth cones to patterned substrata of laminin and fibronectin in vitro. Dev. Biol. 121:423-431.

Herndon, M.E. and A.D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. Neuron 4:949-961.

Hunter, D.D., R. Llinas, M. Ard, J.P. Merlie, and J.R. Sanes. 1992. Expression of S-laminin and laminin in developing rat central nervous system. J. Comp. Neurol. 323:238-251.

Ivins, J.K., E.D. Litwack, A. Kumbasar, C.S. Stipp, A. Yang, and A.D. Lander. 1995. Cerebroglycan and glypican are cell surface heparan sulfate proteoglycans expressed on axons and growth cones in the developing rat nervous system. Soc. Neurosci. Abstracts 21:795.

Kennedy, T.E., T. Serafini, and J.R.d.l. Torre. 1994. Netrins are diffusible chemotropic factors for commissural axons in the embyronic spinal cord. Cell 78:425-435.

Kim, C.W., O.A. Goldberger, R.L. Gallo, and M. Bernfield. 1994. Members of the syndecan family of heparan sulfate proteoglycans are expressed in distinct cell-, tissue-, and development-specific patterns. Mol. Biol. Cell 5:797-805.

Kleinman, H.K., M.L. McGarvey, L.A. Liotta, P.G. Robey, K. Tryggvason, and G.R. Martin. 1982. Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan form the EHS sarcoma. Biochemistry 21:6188-6193.

Krusius, T. and E. Ruoslahti. 1986. Primary structure of an extracellular matrix proteoglycan core protein deduced form cloned cDNA. Proc. Natl. Acad. Sci. USA 83:7683-7687.

Lander, A.D., D.K. Fujii, and L.F. Reichardt. 1985. Purification of a factor that promotes neurite outgrowth: isolation of laminin and associated molecules. J. Cell Biol. 101:898-913.

Lee, M.K. and A.D. Lander. 1991. Analysis of affinity and structural selectivity in the binding of proteins to glycosaminoglycans: development of a sensitive electrophoretic approach. Proc. Natl. Acad. Sci. (USA) 88:2768-2772.

Lim, W.A., R.T. Sauer, and A.D. Lander. 1991. Analysis of DNA-protein interactions by affinity coelectrophoresis. Meth. Enzymol. 208:196-210.

Lin, M.L. and P.J. Bertics. 1995. Laminin responsiveness is associated with changes in fibroblast morphology, motility, and anchorage-independent growth: A cell system for examining the interaction between laminin and EGF signaling pathways. J. Cell Physiol. 164:593-604.

Litwack, E.D. (1995). Expression and function of proteoglycans in the nervous system: Massachusetts Institute of Technology).

Lobb, R.R. and J.W. Fett. 1984. Purification of two distinct growth factors from bovine neural tissue by heparin affinity chromatography. Biochemistry 23:6295-6299.

McLoon, S.C., L.K. McLoon, S.L. Palm, and L.T. Furcht. 1988. Transient expression of laminin in the optic nerve of the developing rat. J. Neurosci. 8:1981-1990.

Neugebauer, K.M., C.J. Emmett, K.A. Venstrom, and L.F. Reichardt. 1991. Vitronectin and thrombospondin promote retinal neurite outgrowth: developmental regulation and role of integrins. Neuron 6:345-358.

Pfenninger, K.H., L. Ellis, M.P. Johnson, L.B. Friedman, and S. Somlo. 1983. Nerve growth cones isolated from fetal rat brain: Subcellular fractionation and characterization. Cell 35:573-584.

Rogers, S.L., P.C. Letourneau, S.L. Palm, J. McCarthy, and L.T. Furcht. 1983. Neurite extension by peripheral and central nervous system neurons in response to substratum-bound fibronectin and laminin. Dev. Biol. 98:212-220.

Rogers, S.L., P.C. Letourneau, and I.V. Pech. 1989. The role of fibronectin in neural development. Dev. Neurosci. 11:248-265.

San Antonio, J.D., M.J. Karnovsky, S. Gay, R.D. Sanderson, and A.D. Lander. 1994. Interactions of syndecan-1 and heparin with human collagens. Glycobiology 4:327-332.

San Antonio, J.D., J. Slover, J. Lawler, M.J. Karnovsky, and A.D. Lander. 1993. Specificity in the interactions of extracellular matrix proteins with subpopulations of the glycosaminoglycan heparin. Biochem. 32:4746-4755.

Sanderson, R.D., T.B. Sneed, L.A. Young, G.L. Sullivan, and A.D. Lander. 1992a. Adhesion of B lymphoid (MPC-11) cells to type I collagen is mediated by the integral membrane proteoglycan, syndecan. J. Immunol. 148:3902-3911.

Sanderson, R.D., J.E. Turnbull, J.T. Gallagher, and A.D. Lander. 1994. Fine structure of heparan sulfate regulates syndecan-1 function and cell behavior. J. Biol. Chem. 269:13100-13106.

Saunders, S. and M. Bernfield. 1988. Cell surface proteoglycan binds mouse mammary epithelial cells to fibronectin and behaves as a receptor for interstitial matrix. J. Cell Biol. 106:423-430.

Serafini, T., T.E. Kennedy, M.J. Galko, C. Mirzayan, T.M. Jessell, and M. Tessier-Lavigne. 1994. The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. Cell 78:409-424.

Sheppard, M.M., S.K. Hamilton, and A.L. Pearlman. 1991. Changes in the distribution of extracellular matrix components accompany early morphogenetic events of mammalian cortical development. J. Neurosci. 11:3928-3942.

Skene, J.H.P., R.B. Jacobson, G.J. Snipes, C.B. McGuire, J.J. Norden, and J.A. Freeman. 1986. A protein induced during nerve growth (GAP-43) is a major component of growth-cone membranes. Science 233:783-786.

Sung, U., J.J. O'Rear, and P.D. Yurchenco. 1993. Cell and heparin binding in the distal long arm of laminin: Identification of active and cryptic sites with recombinant and hybrid glycoprotein. J. Cell. Biol. 123:1255-1268.

Timpl, R., H. Rohde, L. Risteli, U. Ott, P.G. Robey, and G.R. Martin. 1982. Laminin. Meth. Enzymol. 82:831-838.

Vogel, K.G., M. Paulsson, and D. Heinegard. 1984. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. Biochem. J. 223:587-597.

Yurchenco, P.D. and J.J. O'Rear. (1993). . In Molecular and Cellular Aspects of Basement Membranes, D. H. Rohrback and R. Timpl, editors. (New York: Academic Press), pp. 19-47.

Figure 1. Immunopurification of CBG from embryonic day 16 rat brain membrane fraction. ^{125}I-labelled PGs from E16 rat brain membrane fraction were immunoprecipitated with the 521-2 affinity matrix as described in Materials and Methods. The eluate from the immunoprecipitate was digested with chondroitinase ABC, heparitinase, or both enzymes and compared to total PGs from the E16 membrane fraction that had been similarly treated. Abbreviations: **mem.**, E16 membrane fraction PGs; **Ip**, the eluate from the 521-2 affinity matrix; **u**, undigested; **c**, chondroitinase digested; **h**, heparitinase digested; **d**, double digest. The position of the molecular weight markers is indicated. The open arrowhead indicates the CBG core protein, which migrated with an M_r of 53 kDa, similar to what had been observed previously (Herndon and Lander, 1990).
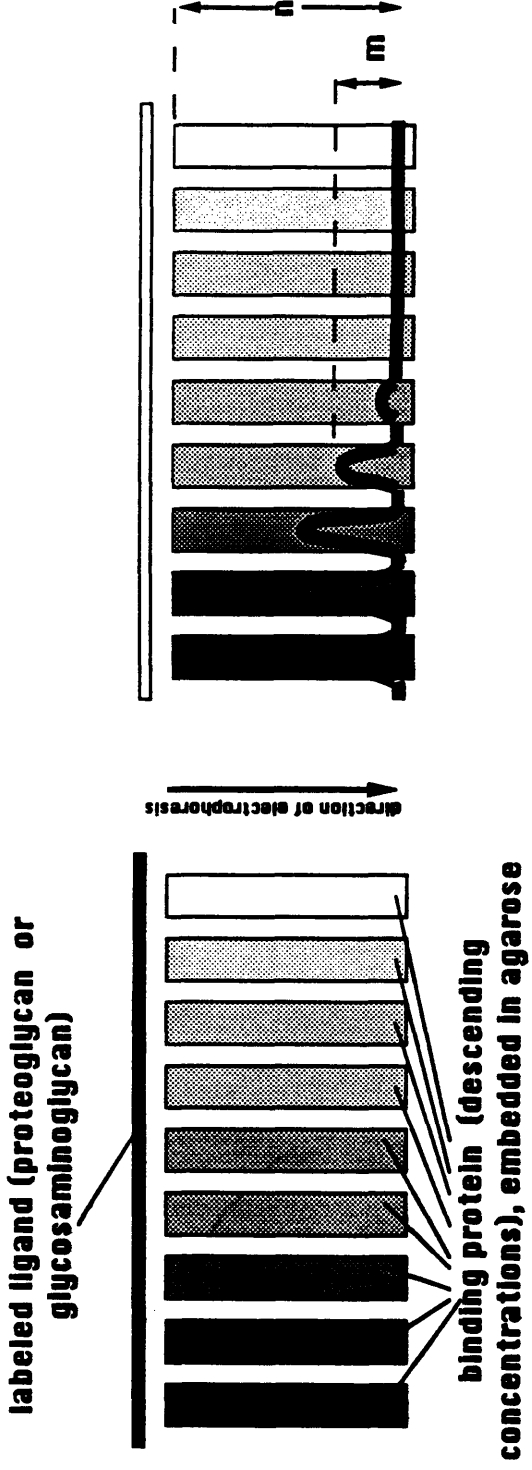
mem.   Ip

u c h d u c h d

220→
100→
68→
43→
27→

Figure 2. Affinity coelectrophoresis

A schematic of an affinity coelectrophoresis (ACE) gel is shown as viewed from above. Labelled PGs and GAGs migrate through lanes containing test ligand at various concentrations, embedded in 1% low melting point agarose. Under these conditions, there is little or no sieving of the PG or GAG by the agarose. Due to a high net negative charge, PGs and GAGs typically migrate with a high mobility relative to the test ligand. An interaction with the ligand results in the formation of a lower mobility complex. The extent to which the PG or GAG migration is retarded, as a function of ligand concentration, can be used to calculate a K_d for the interaction.

Before

labeled ligand (proteoglycan or glycosaminoglycan)

binding protein (descending concentrations), embedded in agarose

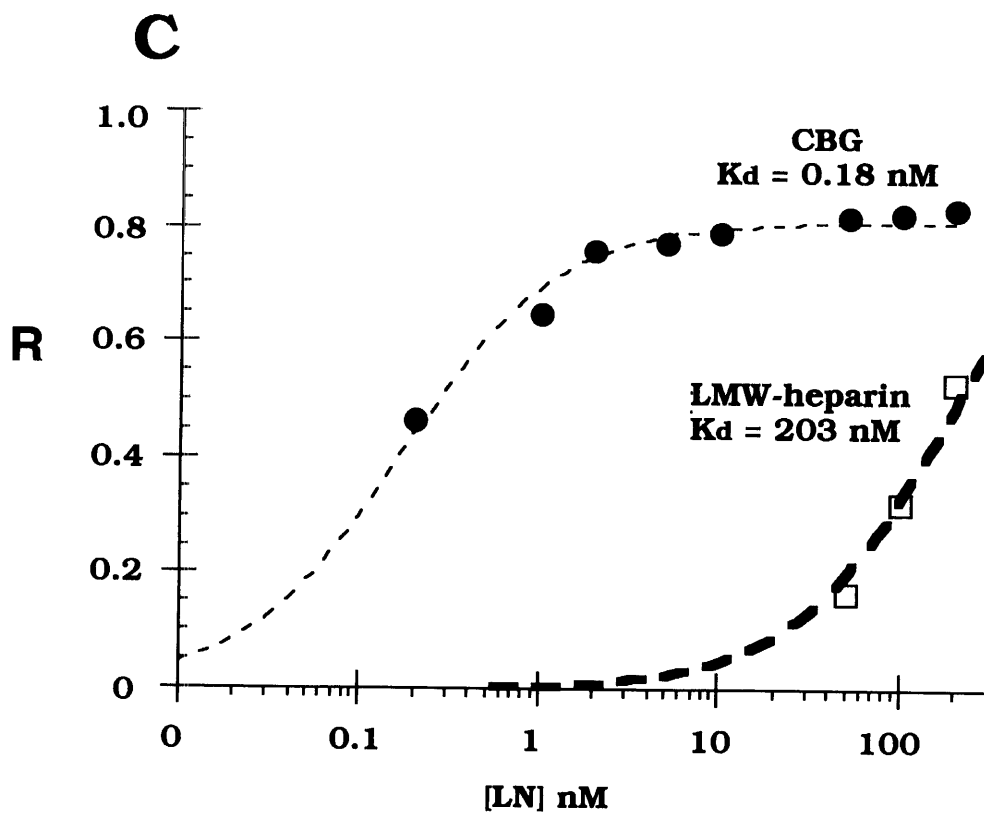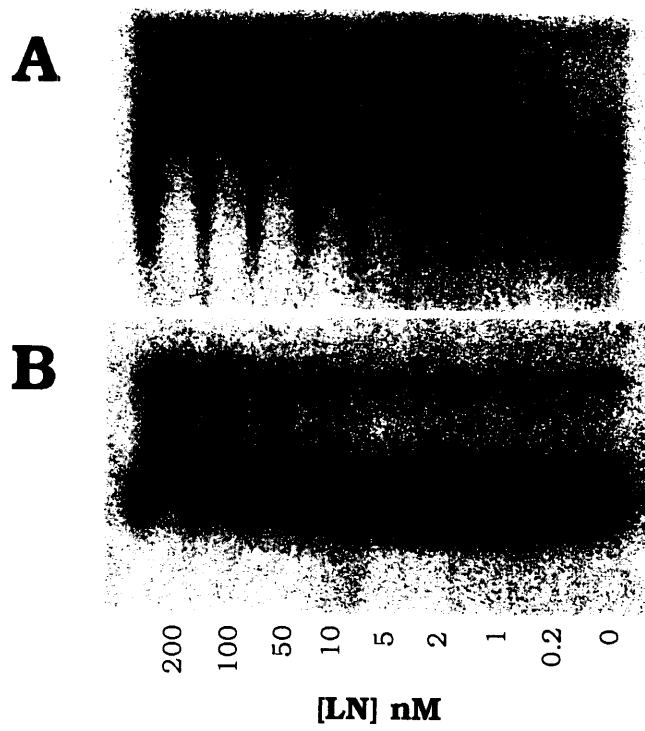direction of electrophoresis

After

If the concentration of labeled ligand is $\ll K_d$, the value of K_d may be calculated from:

$$R = R_\infty \Big/ \left(1 + \frac{K_d}{[\text{Protein}]}\right)$$

where R is the retardation coefficient, i.e. the mobility shift, 'm' divided by the mobility of free glycosaminoglycan 'n'. R_∞ is the value of R approached at saturating protein concentrations.

Figure 3. CBG binds to LN-1 with high affinity
^{125}I-labelled, immunopurified CBG from E16 rat brain was compared
with ^{125}I-labelled low molecular weight (LMW-heparin) for binding to
LN-1 using affinity coelectrophoresis (ACE). (A) The affinity pattern
obtained for CBG. (B) The affinity pattern obtained for LMW-heparin.
The concentrations of LN-1 in each lane are indicated. (C) The
binding isotherms derived from the affinity patterns in (A) and (B).
CBG bound to LN-1 with an apparent affinity of 0.18 nM, while the
measured affinity of LMW-heparin was 203 nM. For LMW-heparin, R_∞,
the R value for maximally retarded PG or GAG, was constrained to 1.0
during curve fitting, based on values obtained from subfractions of
heparin with higher affinity for LN-1 that acheive maximal retardation
within the range of LN-1 concentrations tested (San Antonio et al.,
1993). R_∞ for CBG was not constrained during curve fitting, and was
0.92 for the binding isotherms in this experiment.

A

B

200 100 50 10 5 2 1 0.2 0

[LN] nM

C

R

1.0
0.8
0.6
0.4
0.2
0

CBG
Kd = 0.18 nM

LMW-heparin
Kd = 203 nM

0   0.1   1   10   100

[LN] nM

Figure 4. Binding of CBG and trypsin-digested CBG to LN-1 and FGF-2
Immunopurified, ^{125}I-labelled CBG from embryonic day 16 rat brain
was digested exhaustively with trypsin as described in Materials and
Methods. The affinities of trypsinized CBG (A and C) and intact CBG (B
and D) for LN-1 (A and B) and FGF-2 (C and D) were compared using
affinity coelectrophoresis. Trypsin treatment reduced CBG affinity for
LN-1 by a factor of 400, but had no effect on CBG affinity for FGF-2.
See Table 1 for K_ds, which were calculated as described in Materials
and Methods.

A

B

C

D

[FGF-2] nM

0
1
2
5
10
20
40
60
80

[LN] nM

0
0.1
0.5
1
2
5
10
50
100
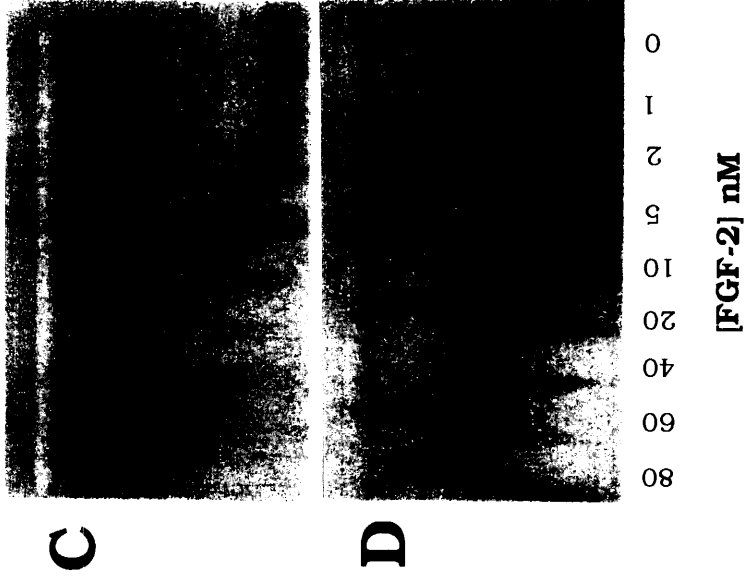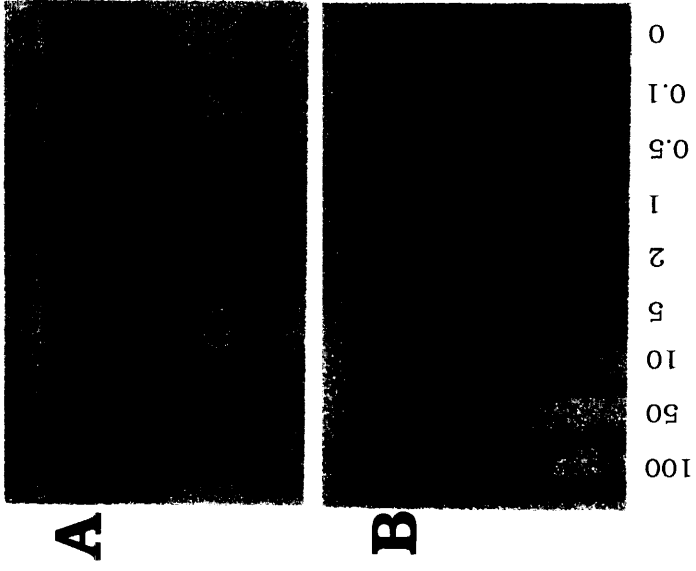
Table 1. Interactions of neuronal cell surface HSPGs and HeS with LN-1, FGF-2, and TSP-1

| PG/GAG	source	LN-1 K_d	LN-1 $(K_d hep)/(K_d)$	rG K_d	rG $(K_d hep)/(K_d)$	FGF-2 K_d	FGF-2 $(K_d hep)/(K_d)$	TSP-1 K_d	TSP-1 $(K_d hep)/(K_d)$
CBG	E16 mem.	0.49± 0.21 nM [a]	400	29 nM	9	50 nM	0.2	33 nM [c]	1.2
CBG	P0 GCP	0.7 nM [b]	280	-	-	-	-	-	-
CBG	3T3 cells	0.11 nM	1800	-	-	-	-	-	-
CBG red.	E16 mem.	~10-100 nM	2-20	-	-	-	-	-	-
CBG tryp.	E16 mem.	215 nM	0.9	-	-	20 nM	0.5	-	-
GPN	P0 GCP	2.2 nM	90	-	-	-	-	-	-
SYN-3	P0 GCP	2.7 nM	70	-	-	-	-	48 nM [c]	0.9
heparin	porcine intest.	194 nM [b]	-	250 nM	-	9 nM [c]	-	41 nM [c]	-
HeS	P0 mem.	886 nM [c]	0.2	-	-	47 nM [c]	0.2	180 nM [c]	0.2

Table 1. continued...

The table summarizes the affinities of CBG, glypican (GPN), and syndecan-3 (SYN-3) for LN-1, the recombinant G domain (rG) of LN-1, FGF-2, and TSP-1, as measured using affinity coelectrophoresis (ACE) in several separate experiments. Abbreviations: **(red.)**, reduced and alkylated; **(tryp.)**, trypsin digested; **(HeS)**, metabolically labelled heparan sulfate; **(E16 mem.)**, embryonic day 16 rat brain membrane fraction; **(P0 GCP)**, neonatal rat brain growth cone particles; **(P0 mem.)**, neonatal rat brain membrane fraction; **(3T3 cells)**, membrane fraction of the CBG-expressing 3T3 cell line 3T3inf$_u$CBGΔ3'.

For comparison, the affinity of low molecular weight (LMW)-heparin **(heparin)** for was also measured in many cases. The ratio of the K_d of heparin to the K_d of a particular PG or GAG appears in the columns labelled $(K_d hep)/(K_d)$. Numbers in this column >1.0 indicate that the PG binds stronger than LMW-heparin to a particular ligand, and numbers <1.0 indicate weaker binding than LMW-heparin.

Footnotes: (a) the average of 3 separate determinations. (b) the average of 2 separate determinations. (c) these values courtesy of Mary Herndon.

Figure 5. Reduction and alkylation of the CBG core protein results in decrease in affinity for LN-1. ^{125}I-labelled, immunopurified CBG was reduced and alkylated as described in Materials and Methods, and its affinity for LN-1 was compared to the affinity of untreated CBG using affinity coelectrophoresis. (A) Affinity pattern of reduced and alkylated CBG. (B) Affinity pattern of untreated CBG. The concentrations of LN-1 in each lane are indicated. The affinity pattern of reduced and alkylated CBG could not be accurately converted into a single binding isotherm, but was estimated to be in the range of 10-100 nM (see also, Table 1).

A

B
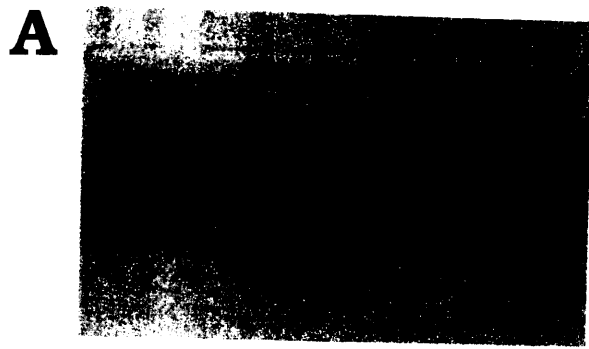
100 50 10 5 2 1 0.5 0.1 0

[LN] nM

CHAPTER 5

DETERMINATION OF THE MOLECULAR WEIGHT OF INTACT CEREBROGLYCAN

INTRODUCTION

Heparan sulfate proteoglycans (HSPGs) have the potential to interact with a diverse set of molecules including growth factors, extracellular matrix glycoproteins, cell adhesion molecules, proteases and antiproteases. As a result, HSPGs may play roles in a wide variety of cell behaviors including cell adhesion, migration, axon outgrowth, cell proliferation and cancer [for reviews, see (Kjellén and Lindahl, 1991; Bernfield et al., 1992; Wight et al., 1992)]. Where investigated, interactions of HSPGs with their ligands have often been found to be mediated by the heparan sulfate (HeS) chains of HSPGs (Jackson et al., 1991). This raises the possibility that one way cells regulate their interactions with their environment is by altering the number or composition of the glycosaminoglycan (GAG) chains presented by PGs at the cell surface.

Support for this idea has come from studies of the cell surface PG syndecan-1, which can be substituted with both HeS and chondroitin sulfate (CS) chains. Several examples have been documented in which the HeS/CS ratio, the size of the GAG chains, and potentially, the average number of chains attached per core have been found to change during development, from tissue to tissue, or in response to growth factor exposure in vitro (Rasmussen and Rapraeger, 1988; Sanderson and Bernfield, 1988; Sanderson et al., 1989; Brauker et al., 1991; Sanderson et al., 1992b). For example, syndecan-1 from simple epithelia (e.g. gastric, cutaneous) and syndecan from stratified epithelia (e.g. uterus) have been shown to have different structures. The syndecan in simple epithelia has larger HeS and CS chains than syndecan from stratified epithelia, and intact syndecan from simple epthelia migrates with an apparent molecular weight (M_r) in SDS-PAGE gels that is significantly higher than that of syndecan from stratified epithelia (Sanderson and Bernfield, 1988).

One limitation in the structural characterization of PGs arises from the anomalous electrophoretic behavior of intact PGs in SDS-PAGE gels. It has long been established that glycosylation can result in an inaccurate estimate --typically an overestimation -- of molecular weight by SDS-PAGE [c.f. (Leach et al., 1980)]. Without an accurate estimate of the molecular weight of the intact PG, guesses about the

average number of chains attached per PG core protein are likely to be incorrect.

In the previous chapter, we demonstrated that the neuronal cell surface HSPGs cerebroglycan (CBG), glypican (GPN), and syndecan-3 all bind to laminin-1 (LN-1) with unexpectedly high affinity (on the order of 0.5-2 nM). The basis of CBG's high affinity for LN-1 (which was 400-fold higher than the affinity of heparin for LN-1) was explored further and found to require an intact, native CBG core protein. Two possible reasons for this requirement are (1) the high affinity binding is the result of multiple HeS chains binding to multiple sites on LN-1, and an intact core is required to tether the chains and possibly to hold them in the optimal conformation for binding, and/or (2) an intact core protein is required for a protein-protein interaction with LN-1 that contributes to the observed high affinity binding.

To distinguish between these possibilities, it would be helpful to know how many HeS chains on average are attached to each CBG core protein. We report here on studies in which the molecular weight of intact CBG and GPN are determined by a combination of gel filtration chromatography and sucrose density centrifugation in H_2O and deuterium oxide (D_2O). These results, combined with an independent estimate of HeS chain size by gel filtration, are consistent with a model where each CBG and GPN core protein is substituted with, on average, a single HeS chain, although substitution with two HeS chains per core could not be ruled out. This finding favors the hypothesis that it is an interaction of the CBG core protein with LN-1 that improves the affinity of CBG for LN-1 from values of ~200 to 800 nM (seen for heparin, trypsin-digested CBG, and free HeS) to the value of 0.5 nM, seen for intact CBG. This view is further supported by studies of LN-1 binding by recombinant ryudocan/syndecan-4 isoforms expressed in L cells. A form of ryudocan with two HeS chains attached was found to bind to LN-1 no better than a form with a single chain attached. Interestingly, ryudocan binding to LN-1 occurred with a substantially lower affinity than that observed for the other cell surface HSPGs tested so far, suggesting a level of specificity in the interactions of LN-1 with HSPG core proteins.

MATERIALS AND METHODS

Preparation of Cerebroglycan and Glypican for Molecular Weight Determination

Growth cone particle PGs were isolated and radiolabelled as described in Chapter 4. To eliminate the complication of detergent binding in the molecular weight determinations, the glycosylphosphatidylinositol (GPI) lipid anchors of CBG and GPN were removed by digestion with phosphoinositol-specific phospholipase C (PIPLC, Boehringer Mannheim). 50 μl of ^{125}I-labelled PGs (containing 2 X 10^7 cpm) in 50 mM TrisHCl, 0.75 M NaCl, 0.5% CHAPS were diluted with 3910 μl of 50 mM Tris-phosphate pH 7.5, 0.1% TritonX-114 (TX-114, Pierce), with 1 mM EDTA, and 1μg/ml pepstatin A, and 40 μl PIPLC (200 U) was added. After a 1 hour digestion at 37°C, another 200 U of PIPLC was added, and incubation was continued for an additional hour. The reaction was stopped by the addition of 240 μl of 5 M NaCl, and the reaction was extracted with TX-114 to remove any uncleaved CBG and GPN: 800 μl of 10% TX-114 was added, followed by a 5 minute incubation at 37°C. Under these conditions, TX-114 resolves into a separate phase, and integral membrane proteins are selectively enriched in the detergent phase. After centrifugation to clarify the detergent and aqueous phases, the aqueous phase was removed and subjected to a second round of TX-114 extraction. The second aqueous phase was loaded into a Centricon-30 microconcentrator (Amicon) that had been pre-blocked with 1 mg/ml BSA and rinsed. The PGs were concentrated from a volume of ~4 ml to a volume of 290 μl. To remove residual TX-114, the PGs were passed over a 200 μl column of Extractigel D (Pierce) that had been pre-equilibrated in TBS (50 mM TrisHCl pH 7.4, 0.15 M NaCl) with1 mg/ml crystalline BSA. By monitoring the absorbance at 275 nm before and after the Extractigel D step, it was determined that the concentration of TX-114 had been reduced from 0.08% to 0.01%, below the critical micelle concentration of ~0.02% (Neugebauer, 1990).

Gel Filtration of Cerebroglycan and Glypican

6×10^6 cpm (114 µl) of the PGs prepared as described above were loaded to a Sepharose CL-6B (Pharmacia) column along with 32 µg of E. Coli β-galactosidase (β–gal, Sigma) and 1 µCi of $Na_2{}^{35}SO_4$ (to measure the included volume, V_t) in a buffer with 5% glycerol and 0.015% phenol red in a final volume of 150 µl. The dimensions of the column were 28.5 cm $\times$ 0.7 cm dia. The column was eluted with TBS at a rate of 3.9 cm/hour. In subsequent runs, the column was calibrated with bovine thyroid gland thyroglobulin (Sigma), bovine liver catalase (Worthington), and rabbit muscle glyceraldehyde phosphate dehydrogenase (GAPD, Worthington)(see below for enzyme assays). β-gal and $^{35}SO_4$ were included in each run to allow all runs to be normalized. The void volume (V_o) was measured by the leading edge of Blue Dextran 2000 (Fluka) and was found to correspond with a sharp peak of ^{125}I-labelled PGs that eluted ahead of CBG and GPN and that contains HSPGs with core proteins of 120 kDa and higher (not shown). CBG and GPN were located in the elution profile by heparitinase digestion and SDS-PAGE, performed as described in previous chapters. The CBG and GPN bands were quantified using a phosphorimager (Molecular Dynamics).

Sucrose Density Gradient Centrifugation

4.7 ml linear gradients of $5 \rightarrow 20\%$ sucrose in TBS were prepared in H_2O or D_2O (ICN) using a Model J5 Gradient Former (Jule Inc., New Haven, CT) in 5 ml ultracentrifuge tubes. The tubes were covered with parafilm, and the gradients were allowed to age for 3 hours at 4°C. Samples containing 1×10^6 cpm of labelled P0 GCP PGs, prepared as described above, in a 100 µl volume were overlaid onto the gradients. Bovine pancreatic trypsin inhibitor (BPTI, Worthington), aggregate-free bovine pancreatic ribonuclease A (RNAse A, Worthington), human erythrocyte carbonic anhydrase B (CAH, Sigma), pig heart mitochondrial malate dehydrogenase (MDH, Worthington), and calf intestinal alkaline phosphatase (AP, Sigma) were included in the samples as protein standards (see below for enzyme assays). The gradients were centrifuged for 20 hours at 45,000 rpm, 4°C in a Sorvall AH650 swinging bucket ultracentrifuge

rotor ($\omega^2 t$ = 1.60 X 10^{12} s^{-1}). Fractions were collected from the bottom of the tubes using a fraction collector (Hoeffer). The locations of CBG and GPN in the gradient were determined by SDS-PAGE after heparitinase digestion, as described for the gel filtration experiments above. A separate run of GCP PGs in a 5 → 20% sucrose gradient in TBS/H_2O performed on a different day gave an essentially identical sedimentation profile as did parallel runs of CL-6B purified CBG and GPN performed at the same time as the run described above. The run described above was used as the basis for molecular weight calculations because the smaller number of cpm in the CL-6B purified pool made the quantification of the CBG and GPN bands by phosphorimager more difficult.

Enzyme Assays

BPTI was assayed by measuring the inhibition of trypsin (Worthington) activity using Chromozyme TH (Boehringer Mannheim) as a substrate and measuring absorbance at 405 nm (A_{405}) of timed reactions. RNase A was measured by monitoring the digestion of poly-cytidylic acid (Fluka cat.#81306) as described by Zimmerman and Sandeen (1965). This assay proved to be extremely sensitive, linear and highly reproducible. CAH and MDH were assayed densitometrically after SDS-PAGE and silver staining. AP was assayed with p-nitrophenyl phosphate (Sigma) by monitoring A_{410} of timed reactions. Thyroglobulin was assayed by measuring A_{278} of aliquots of selected column fractions. β-gal was assayed with o-nitrophenyl-β-D-galactopyranoside (Sigma) by monitoring A_{405} of timed reactions. Catalase was assayed with 0.2% H_2O_2 by monitoring A_{240} of timed reactions, and GAPD was assayed with DL-glyceraldehyde-3-phosphate and β-nicotinamide adenine dinucleotide (NAD^+) by monitoring A_{340} of timed reactions. The bases of the enzyme assays and composition of the buffers are described in the Worthington Enzyme Manual (Worthington, 1993) which can be obtained upon request from Worthington Biochemical Corp. Table 1 lists the protein standards used in this study along with the relavent hydrodynamic constants.

Molecular Weight Calculations

The basis of the molecular weight determinations is the Svedberg equation which relates the hydrodynamic properties of a molecule to its molecular weight: $M = sRT/D(1-\bar{V}\rho)$; where M is molecular weight (in grams/mol); s is the sedimentation coefficient in svedberg units (1 svedberg is 10^{-13} sec); R is 8.31×10^7 erg mol^{-1} deg^{-1}; T is temperature in degrees Kelvin; D is the diffusion coefficient in cm^2 sec^{-1}; $\bar{V}$ is the partial specific volume in cm^3 g^{-1}; and ρ is the density of the solution in g cm^{-3}. As described by Clarke (1975), an adaptation of the Svedberg equation is $M = (s_{app, Ravg})(6\pi)(\eta_{app, Ravg})(a)(N)/(1 - \bar{V}\rho_{Ravg})$. In this equation, $s_{app,Ravg}$ is the apparent sedimentation coefficient for the molecule at **Ravg**, *the position of half distance of travel of the molecule in the gradient*, $[s_{app,Ravg} = ((r - r_0)/t/\omega^2 Ravg)$; where r = the distance of the peak at time t, (the end of the run), r_0 is the distance of the applied band at time zero, and Ravg $= (r + r_0)/2]$. $\eta_{app,Ravg}$ is the apparent viscosity of the solution at Ravg; ρ_{Ravg} is the density of the solution at Ravg; "**a**" is the Stokes radius, and **N** is Avagadro's number. In the present study, the Stokes radius was calculated by gel filtration chromatography as described above using protein standards of known Stokes radii to calibrate the column.

Although the reader is referred to Clarke, (1975) for a more complete discussion, some additional features of the calculations are described below. Values of $\eta_{app,Ravg}$ were determined for CBG and GPN by including in each run protein standards with a known $s_{20,w}$ (sedimentation coefficient at 20°C in H$_2$O) and a known $\bar{V}$. The value of $s_{app,Ravg}$ measured for each standard, along with the $s_{20,w}$ and $\bar{V}$ values, was used to calculate values of $\eta_{app,Ravg}$ for each protein standard. The values of $\eta_{app,Ravg}$ for CBG and GPN were determined by interpolating between the values of $\eta_{app,Ravg}$ calculated for the protein standards. In this way the apparent viscosity, as a function of the position in the gradient, becomes a general correction factor which gives the system self-consistency (Clarke, 1975).

The values of $\bar{V}$ for CBG and GPN were determined on the basis of their differing rates of sedimentation in D$_2$O and H$_2$O: the measured values of $s_{app,Ravg}$, $\eta_{app,Ravg}$, and ρ_{Ravg} for CBG and GPN in D$_2$O and H$_2$O were used to calculate $\bar{V}$, using equations from Clarke, (1975).

In Clarke, (1975), values relating to the amount of detergent bound by the molecules enter into these equations; these values are equal to zero in the present study since sedimentation and gel filtration were performed in the absence of detergent.

Preparation of Labelled Heparan Sulfate from a Neonatal Rat Brain Membrane Fraction

Dissections of neonatal (P0) rat brains and removal of meninges was accomplished as outlined in Chapter 2. Meninges-free whole brains were cut into 1mm x 350 μm prisms with a McIlwain Tissue Chopper, then settled through Hanks buffered salt solution (HBSS, Gibco). Serum free, sulfate free Dulbecco's modified Eagle's medium (SSF-DMEM) was formulated to GIBCO/BRL specifications for DMEM - but without sulfate - using culture supplements and 10X Earle's Balanced Salt Solution from GIBCO. SSF-DMEM was supplemented ("labelling medium") to 50 μM with unlabeled Na_2SO_4, 10 μg/ml insulin, 5 mg/ml crystalline BSA, 10 μg/ml transferrin, 20 nM progesterone, 100 μM putrescine, 30 nM selenium, and 250 mCi/ml $Na_2^{35}SO_4$ (ICN). Labelling medium was added to tissue prisms at 20 ml per ml tissue (settled volume) and prisms were cultured 16-20 hours in 5% CO_2 with gentle gyratory rocking (Nutator).

PGs were isolated from the membrane fraction of labelled prisms as previously described [(Herndon and Lander, 1990); Chapter 2 this thesis]. To the PG pool, in a volume of 1.7 ml, 32 μl of 5 M NaOH and 16 μl of fresh Na_2BH_4 in 10 mM NaOH were added to liberate glycosaminoglycan (GAG) chains by β-elimination. After 90 minutes at 45°C, 3 μl of 0.5% phenol red were added to monitor neutralization with ~32 μl of 5 M HCl. The reaction was chilled on ice and centrifuged for 10 minutes at 12,000 g to remove insoluble, denatured protein. The supernatant was transferred to a fresh tube and precipitated on ice for 30 minutes with the addition of 10 μl of 10 mg/ml hyaluronate and 5.9 ml of ice cold 100% ethanol. After centrifugation for 20 minutes at 4°C, 12,000 g, the pellet was resuspended in 450 μl of H_2O, and reprecipitated with 1440 μl of ethanol. The final pellet was resuspended in TBS and stored at -80°C until use. Labelled GAGs were treated with chondroitnase ABC to

eliminate chondroitin sulfate contamination, in a reaction performed essentially as previously described (Herndon and Lander, 1990), overnight at 37°C.

Gel Filtration of $^{35}SO_4$-Labelled Heparan Sulfate Chains

Labelled heparan sulfate (HeS) chains, prepared as described above, were sized by gel filtration chromatography using a Sephadex-G100 column (Pharmacia). The dimensions of the column were 18 cm by 0.5 cm dia. The column was eluted at a flow rate of 4 cm/hour; the running buffer was TBS. HeS chains were loaded in a volume of 100 μl in a buffer containing 5% glycerol, 0.015% phenol red, and 1 μCi of free $^{35}SO_4{}^{2-}$, which was included to ensure a peak at the included volume, Vt. The column had been previously calibrated using fluorescein isothiocyanate (FITC)-conjugated dextrans of known modal molecular weight (Sigma), the elution of the FITC-dextrans was monitored by reading A_{493} of portions of selected fractions in a spectrophotometer. The void volume of the column was measured as the elution volume of the leading edge of Blue Dextran 2000 (Fluka). To verify that labelled HeS was being sized, an aliquot of the material was digested overnight at 37°C with heparitinase, essentially as described (Herndon and Lander, 1990), and run on the G100 column. Column fractions from both runs were assayed by counting aliquots of each fraction in a scintillation counter (LKB-Wallach).

Preparation and LN-1 Affinity Coelectrophoresis of Ryudocan Isoforms

Growth and metabolic labelling of the mouse L cell lines, L215 and L218 were performed essentially as previously described (Shworak et al., 1994a). These cell lines express epitope-tagged ryudocan isoforms with two or one GAG acceptor site respectively. Immunopurification of ^{35}S-labelled ryudocan isoforms was performed as described (Shworak et al., 1994a) except that immunoprecipitates were eluted with a low pH step, as described for CBG and GPN in Chapter 4. The affinity of the ryudocan isoforms for LN-1 was measured by affinity coelectrophoresis, as described in Chapter 4.

RESULTS

Cerebroglycan and glypican molecular weight determination

In a previous study, cerebroglycan (CBG) and glypican (GPN) were found to bind to laminin-1 (LN-1) with high affinity (0.5 and 2 nM respectively)(Chapter 4). The molecular basis of these high affinity interactions is unknown. To obtain a more accurate picture of the structure of intact CBG and GPN, the molecular weights of these two related molecules were determined. Because of the aberrant electrophoretic behavior of glycosylated molecules in SDS-PAGE (Leach et al., 1980), an accurate measure of the molecular weight (Mw) of intact PGs is unlikely to be obtained using this method. We have chosen instead a method based on the Svedberg equation which relates Mw to the hydrodynamic behavior of molecules. The Svedberg equation allows a determination of Mw that is independent of the details of molecular shape (Cantor and Schimmel, 1980).

The hydrodynamic parameters that must be measured to make use of the Svedberg equation are (1) Stokes radius, (2) the sedimentation coefficient, and (3) the partial specific volume. The Stokes radii of CBG and GPN [isolated from neonatal (P0) rat brain growth cone particles (GCPs) (Pfenninger et al., 1983) and labelled with ^{125}I] were measured by gel filtration on Sepharose CL6B. Heparitinase digestion followed by SDS-PAGE analysis of the column fractions allowed the fractions containing CBG and GPN to be identified (Figure 1). CBG and GPN core proteins were easily distinguishable in this assay as had been previously observed (Herndon and Lander, 1990). The identities of these bands were confirmed by SDS-PAGE with immunopurified CBG and GPN in subsequent experiments (not shown). Figure 2A shows the elution profile of CBG and GPN, as determined by densitometric quantification of the CBG and GPN core protein bands. Also shown are the elution positions of the protein standards that were used to calibrate the column. Comparison of the CBG and GPN elution profiles to the calibration curve generated by the standards yielded values of Stokes radii of 6.51 and 6.57 nm respectively (Figure 2B).

To determine the sedimentation coefficients (s) and the partial specific volumes ($\bar{V}$) of CBG and GPN, sucrose density gradient centrifugation in H_2O and D_2O was performed (Clarke, 1975). To

calibrate the gradients, protein standards of known $s_{20,w}$ (the s value at 20°C in water) and $\bar{V}$ were included in the sedimentation runs. Figure 3 (A and B) shows the sedimentation profiles of the standards. Plots of s versus fraction number derived from these profiles were linear, demonstrating that the gradients were effectively linear (Figure 3, C and D). The apparent s values of the standards in the sedimentation runs, and their previously determined values of $s_{20,w}$ and $\bar{V}$ were used to determine apparent viscosity as a function of position in the gradients (Clarke, 1975), one of the parameters required for the subsequent calculations of s and $\bar{V}$. The sedimentation profiles of CBG and GPN were determined by SDS-PAGE after heparitinase digestion as described for the analysis of the CL6B fractions above (Figure 4). The apparent s values of CBG and GPN, their $s_{20,w}$ values, and their $\bar{V}$ values were calculated from the CBG and GPN sedimentation profiles as described in Clarke, (1975). These values and the molecular weights calculated from them are summarized in Table 2. The calculated Mw for CBG was 100 kDa, while the Mw of GPN was calculated at 103 kDa. With the core protein contribution subtracted, these Mw values imply that both CBG and GPN carry an average of ~42 kDa of HeS per core protein.

Estimation of the average size of HeS from the neonatal rat brain membrane fraction

To obtain an estimate of the average number of HeS chains per core protein for CBG and GPN, the size of HeS isolated from P0 rat brain membrane fraction was measured. To prepare labelled HeS for these experiments, finely chopped prisms of P0 rat brain were cultured in vitro in a defined medium containing $^{35}SO_4^{2-}$. HeS was isolated from the membrane fraction of the labelled prisms as described in Materials and Methods. Labelled HeS was analyzed by gel filtration chromatography with or without prior digestion with heparitinase (Figure 5). Comparison of the elution profile of intact HeS with the elution volumes of fluoresceinated dextrans of known molecular weight allowed the size of the P0 HeS chains to be estimated (Figure 5). The bulk of the P0 HeS eluted as a relatively broad peak, centered at ~40 kDa, suggesting that the entire amount of HeS found on average

on each CBG and GPN core protein could be accounted for by a single HeS chain (See discussion).

The binding of one and two chain ryudocan isoforms to LN-1

Ryudocan is a member of the syndecan family of cell surface HSPGs and contains three potential glycosaminoglycan (GAG) attachment sites. A recent study has characterized these sites, demonstrating that all three are utilized and all three can be substituted with both HeS and chondroitin sulfate (CS) (Shworak et al., 1994a). This was accomplished by generating ryudocan isoforms in which the three GAG attachment sites had been systematically eliminated by changing the serine acceptor residues to threonines, and expressing them in mouse L cells.

The molecular weight calculations presented above suggest that CBG and GPN may be, on average, substituted with a single HeS chain per protein core. Since free HeS binds to LN-1 with a much lower affinity than intact CBG and GPN (Chapter 4 this thesis), it is unlikely that a single HeS chain binding to LN-1 by itself could account for the observed affinities of CBG and GPN for LN-1. This raises the possibility that CBG and GPN core proteins bind to LN-1, contributing to the overall affinity observed for the intact PGs. If the strong binding of CBG and GPN observed in Chapter 4 results from an extra contribution to binding by the core protein, it raises the possibility that the number of GAG chains per core protein may not be an important determinant of a PG's affinity for LN-1. To test this possibility the binding of ryudocan isoforms with one or two GAG chains to LN-1 was measured.

Exponentially growing L cell lines were used for metabolic labelling and isolation of the ryudocan isoforms, as previously described (Shworak et al., 1994a). Under conditions of exponential growth, ~ 90% of the GAG chains attached to ryudocan are HeS chains (Nicholas Shworak, personal communication) minimizing any potential effect of CS chain substitution. (In a separate experiment using total L cell PGs, pretreatment with chondroitinase ABC resulted in no difference in the binding pattern or electrophoretic migration in ACE.) The LN-1 affinity of immunopurified ryudocan with a single HeS chain was compared to the affinity of ryudocan with two HeS chains using affinity

coelectrophoresis as described in Chapter 4. Both ryudocan isoforms were found to bind LN-1 with similar affinities (K_d of ~120-140 nM)(Figure 6). This result provides further support for the idea that factors other than multivalent HeS chain interactions are responsible for the high affinity binding of CBG and GPN to LN-1.

DISCUSSION

Determination of the molecular weights of cerebroglycan and glypican

An estimate for the molecular weights (Mw) of intact CBG and GPN has been obtained based on the measurement of hydrodynamic properties. The basis of the calculations was the Svedberg equation that permits a calculation of Mw that is independent of the details of molecular shape, provided that the various hydrodynamic measurements were made under conditions as similar as possible (Cantor and Schimmel, 1980). The sedimentation coefficients and the partial specific volumes of CBG and GPN were measured by sedimentation on a $5 \rightarrow 20\%$ sucrose gradients in H_2O and D_2O. In this technique, the positions of CBG and GPN in the gradient are determined at only one time point -- the end of the sedimentation run. A more rigorous method would require monitoring the positions optically at all times during the run (Cantor and Schimmel, 1980). Nevertheless, the uncertainty in measurements by the sucrose gradient method have previously been estimated to be $\pm 8\%$, and much of this was due to uncertainties about amounts of detergent bound to the proteins (Hartshorne et al., 1980). Since we took steps to minimize the contribution of bound detergent, the uncertainty in our measurements may be lower still.

The Mw values that were obtained for CBG and GPN were 100 and 103 kDa respectively. These values are substantially lower than the estimated ~200,000 kDa apparent molecular weight of the undigested molecules by the SDS-PAGE method (for example, see Figure 1, Chapter 4). This result is not surprising since the overestimation of the molecular weights of glycosylated proteins by SDS-PAGE has been reported before (Leach et al., 1980). The aberrant electrophoretic behavior of PGs may be related to an unusually large hydrodynamic radius. In the gel filtration experiments described above, CBG and GPN Stokes radii were measured at 6.51 and 6.57 nm respectively, about as large as that of β-galactosidase, a protein of 540 kDa. These results demonstrate that estimations of PG molecular weight based soley on SDS-PAGE or gel filtration are unlikely to be accurate.

Heparan sulfate chain number for CBG and GPN

When the sizes of the core proteins are substracted from the total molecular weights determined for CBG and GPN in this study, a value of ~42 kDa is obtained for the amount of HeS attached on average to each core protein. An independent measurement of HeS chain size from P0 rat brain membrane fraction resulted in a broad peak centered at ~40 kDa. Thus it is possible that the entire mass of HeS attached on average to CBG and GPN core proteins could be accounted for by a single HeS chain, although the peak was sufficiently broad to permit the possibility of substitution with two smaller chains. This result is different from an earlier assessment for GPN from human lung fibroblasts (David et al). Although a similar size for the HeS chains was reported, the estimation of the M_W of intact GPN was based on SDS-PAGE, and was set at ~200,000, leading to the conclusion that several HeS chains must be attached to each GPN protein core ((Lories et al., 1989; David et al., 1990)). As discussed above, SDS-PAGE is likely to result in an overestimate of the size of intact PGs.

Another source of disagreement is in the estimation of HeS chain size. One earlier study reported a value for brain HeS of ~14-15 kDa, significantly lower than the value of ~40 kDa reported here (Ripellino and Margolis, 1989). A possible source for this discrepancy is developmental age. In the earlier study, HeS chains were isolated from adult rat brain. In the present study HeS was isolated from neonatal rat brain. Variation of HeS chain size with developmental stage has been reported before. For example, in the mouse lung epithelium, the average size of HeS chains is 51 kDa on embryonic day 14, but has decreased to only 28 kDa by embryonic day 18 (Brauker et al., 1991). In this study, HeS chains were prepared by an in vitro labelling protocol similar to the one employed in our experiments. Another study reported that when mouse keratinocytes stratify in vitro, the average size of HeS chains decreases from 34 to 21 kDa (Sanderson et al., 1992b). It is possible that a similar decrease in average size of HeS chains occurs during brain development. The value of ~40 kDa obtained in the present study is similar to the value of 40 kDa obtained from HeS isolated from the retinal neuroepithelium

of embryonic chick (Cole and Burg, 1989); thus the value we report is not an unprecedented size for neuronal HeS, and demonstrates that, at a minimum, HeS chains in the size range of 40 kDa are present as a substantial fraction of material derived from P0 rat brain.

Determination of the size of HeS chains attached to immunopurified CBG and GPN will be necessary to resolve this issue. The elution profile shown in Figure 5 is sufficiently broad that chain sizes of ~20 kDa could be present in significant quantities. If for some reason HeS from CBG and GPN was smaller on average than total HeS, it would be possible for more than one HeS chain to be attached to each core protein. It will also be important to include glycosaminoglycan (HeS or CS) chains of defined size in future calibrations since the dextrans used in the present study may not be the most suitable standards because of differences in structure (i.e. sulfation) between dextrans and GAGs.

Thus, the results of the present study cannot conclusively distinguish between a model where each CBG or GPN core protein is substituted with one HeS chain or a model in which each core is substituted with two HeS chains. However, they do establish that HeS chains derived from neonatal rat brain contain a broad range of sizes, including a substantial fraction that appears to be 40 kDa or larger.

Implications of CBG and GPN structure on binding to laminin

In the previous chapter, we demonstrated that CBG and GPN bind to LN-1 with high affinity. Experiments aimed at determining the basis of the high affinity interaction demonstrated that an intact CBG core protein was required. Since trypsin treatment liberated glycanated CBG peptides that bound to LN-1 with much lower affinity, is appears unlikely that high affinity binding of CBG to LN-1 is mediated solely by a single HeS chain. One possibility is that multiple chains attached to a single CBG core protein interact with multiple binding sites on LN-1. However, as discussed above, it is possible that at least some CBG and GPN cores may bear only one HeS chain.

An alternative hypothesis is that a core protein interaction with LN-1 contributes to the overall affinities of CBG and GPN for LN-1. The observation that reduction and alkylation appear to significantly reduce

CBG affinity for LN-1 (Chapter 4), is consistent with this hypothesis. Further support for this view comes from the binding of ryudocan isoforms with defined numbers of HeS chains to LN-1. Although the HeS chains on these isoforms are just as large or larger (51 kDa, (Shworak et al., 1994a)), than those on GPN or CBG, the presence of two chains on a single ryudocan core protein did not improve ryudocan's affinity for LN-1.

It is interesting to note that the contribution of a putative core protein interaction to overall affinity needn't be that large in terms of binding energy to make a substantial difference in the value of the K_d. This is because of the exponential relationship between K_d and the standard free energy of binding ($K_d = e(\Delta G^\circ/RT)$. For example, for a K_d of 500 nM (in the range of heparan sulfate binding to LN-1), the value of ΔG° is -4.63 kcal/mole. For a Kd of 500 pM (in the range of CBG binding to LN-1), ΔG° is -6.82 kcal/mole. The difference between these two values, -2.18 kcal/mole, is well within the range of the binding energy that could be provided by single hydrogen bond. However, it is important to emphasize that the contribution to binding energy from a putative core protein interaction could be significantly higher or lower than this value. In the absence of an independent measurement of the K_d for core protein binding to LN-1 it is impossible to estimate any contribution of the core protein to the binding energy. The above discussion serves only to illustrate that such a contribution needn't be large in order to have a substantial impact on K_d. In fact, the contribution of the core protein to binding could be so small that it would be impossible to measure by ACE, since the concentrations of LN-1 required would become prohibitive.

In conclusion, the available data are consistent with a model where CBG both with its HeS chain(s) and its protein core, resulting in a high affinity interaction. Although the putative contribution of the protein core in binding to LN-1 needn't be large to have an effect on the K_d of intact CBG for LN-1, it appears to be relatively specific: the affinity of CBG for FGF-2, TSP-1, and the recombinant G domain of LN-1 was only marginally better, if at all, than the affinity of heparin for the same three molecules (Chapter 4). In addition, ryudocan binding to LN-1 is of substantially lower affinity than CBG binding (Figure 6 above), which

suggests that LN-1 may exhibit some selectivity in which core proteins it binds.

The vast majority of putative ligands for HSPGs have been identified on the basis of binding to heparin or HeS chains. The results reported here suggest that protein-protein interactions involving PG core proteins could be more widespread than previously appreciated, and emphasize the importance of taking the core protein into account when analyzing the structures and functions of proteoglycans.

ACKNOWLEDGEMENTS

References

Bernfield, M., R. Kokenyesi, M. Kato, M.T. Hinkes, J. Spring, R.L. Gallo, and E.J. Lose. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu. Rev. Cell Biol. 8:365-393.

Brauker, J.H., M.S. Trautman, and M. Bernfield. 1991. Syndecan, a cell surface proteoglycan, exhibits a molecular polymorphism during lung development. Dev. Biol. 147:285-292.

Cantor, C.R. and P.R. Schimmel. (1980). Analysis of Sedimentation Measurements. In Biophysical Chemistry Part II: Techniques for the study of biological structure and function, A. C. Bartlett, P. C. Vapnek and L. W. McCombs, editors. (San Francisco: W. H. Freeman and Co.), pp. 605-623.

Clarke, S. 1975. The size and detergent binding of membrane proteins. J. Biol. Chem. 14:5459-5469.

Cole, G.J. and M. Burg. 1989. Characterization of a heparan sulfate proteolglycan that copurifies with the neural cell adhesion molecule. Exp. Cell Res. 182:44-60.

David, G., V. Lories, B. Decock, P. Marynen, J. Cassiman, and H.V.d. Berghe. 1990. Molecular cloning of a phosphatidylinositol-anchored membrane heparan sulfate proteoglycan from human lung fibroblasts. J. Cell Biol. 111:3165-3176.

Hartshorne, R.P., J. Coppersmith, and W.A. Catterall. 1980. Size characteristics of the solubilized saxitoxin receptor of the voltage-sensitive sodium channel from rat brain. J. Biol. Chem. 255:10572-10575.

Herndon, M.E. and A.D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. Neuron 4:949-961.

Jackson, R.L., S.J. Busch, and A.D. Cardin. 1991. Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. Physiol. Rev. 71:481-539.

Kjellén, L. and U. Lindahl. 1991. Proteoglycans: Structures and interactions. Annu. Rev. Biochem. 60:443-475.

Leach, B.S., J.F.C. Jr., and W.W. Fish. 1980. Behavior of glycopolypeptides with empirical molecular weight estimation methods. 1. In sodium dodecyl sulfate. Biochemistry 19:5734-5741.

Lories, V., J.J. Cassiman, H. Van den Berghe, and G. David. 1989. Multiple distinct membrane heparan sulfate proteoglycans in human lung fibroblasts. J. Biol. Chem. 264:7009-7016.

Neugebauer, J. (1990). A Guide to the Properties and Uses of Detergents in Biology and Biochemistry (Potsdam, New York: CALBIOCHEM Corporation).

Pfenninger, K.H., L. Ellis, M.P. Johnson, L.B. Friedman, and S. Somlo. 1983. Nerve growth cones isolated from fetal rat brain: Subcellular fractionation and characterization. Cell 35:573-584.

Rasmussen, S. and A. Rapraeger. 1988. Altered structure of the hybrid cell surface proteoglycan of mammary epithelial cells in response to transforming growth factor-β. J. Cell Biol. 107:1959-1967.

Ripellino, J.A. and R.U. Margolis. 1989. Structural properties of the heparan sulfate proteoglycans of brain. J. Neurochem. 52:807-812.

Sanderson, R.D. and M. Bernfield. 1988. Molecular polymorphism of a cell surface proteoglycan: distinct structures on simple and stratified epithelium. Proc. Natl. Acad. Sci. USA 85:9562-9566.

Sanderson, R.D., M.T. Hinkes, and M. Bernfield. 1992b. Syndecan, a cell-surface proteoglycan, changes in size and abundance when keratinocytes stratify. J. Invest. Dermatol. 99:390-396.

Sanderson, R.D., P. Lalor, and M. Bernfield. 1989. B lymphocytes express and lose syndecan at specific stages of differentiation. Cell Regulat. 1:27-35.

Shworak, N.W., M. Shirakawa, R.C. Mulligan, and R.D. Rosenberg. 1994a. Characterization of ryudocan glycosaminoglycan acceptor sites. J. Biol. Chem. 269:21204-21214.

Wight, T.N., M.G. Kinsella, and E.E. Qwarnstrom. 1992. The role of proteoglycans in cell adhesion, migration and proliferation. Curr. Opin. Cell Biol. 4:793-801.

Worthington. (1993). Worthington Enzyme Manual, enzymes and related biochemicals, V. Worthington, editor. (Freehold, New Jersey: Worthington Biochemical Corporation).

| Molecule	Source	Stokes radius (nm)	$S_{20,w}$ (S)	$\bar{V}(cm^3/g)$	M_w (kDa)	refs.
Trypsin inhibitor	bovine pancreas	1.66	1.0 S	0.718	6.67	a
RNaseA	bovine pancreas	2.0	1.78	0.703	13.6	a
Carbonic anhydrase B	human erythrocyte	2.0	3.23	0.729	27.0	a
Malate Dehydrog.	pig heart	3.72	4.53	0.742	73.9	a
Alkaline Phosphatase	calf intestine	-	6.45	0.756	140	b
Glycerald. phosphate dehydrog.	rabbit muscle	3.92	-	-	140	c
catalase	bovine liver	5.22	-	-	232	d
beta-galactosidase	E. Coli	6.9	-	-	540	c
Thyroglobulin	bovine thyroid	8.5	-	-	669	d

Table 1. Hydrodynamic constants for protein standards
The hydrodynamic data for the protein standards used in the experiments described below are shown. $S_{20,w}$ is the Svedberg coefficient at 20°C in water and is reported in svedbergs (1 S =10^{-13} sec). $\bar{V}$ is the partial specific volume. M_w is the molecular weight. References for the values reported in this table were (a) Creighton, (1984), (page 268); (b) Fosset et al, (1974); (c) compiled in Clarke et al, (1975), from earlier references; and (d) In Gel Filtration, Principles and Methods, 6th edition, (page 103), published by Pharmacia Biotech, and available upon request from Pharmacia.

Figure 1. Gel filtration of cerebroglycan and glypican

^{125}I-labelled, PIPLC-treated P0 rat brain PGs isolated from growth cone particles were analyzed by gel filtration chromatography on CL6B Sepharose. To determine the elution profiles, portions of selected column fraction were digested with heparitinase and separated on SDS-PAGE minigels; an example of one such gel is shown. The numbers refer to CL6B column fractions. For comparison an aliquot of fraction 34 that had not been treated with heparitinase was included. The positions of the CBG and GPN bands, which appear in response to heparitinase treatment, are indicated by open arrowheads. The molecular weight markers are also shown. The relative abundance of GPN compared to CBG is typical for this stage in development, a time when CBG mRNA and protein are beginning to disappear from the rat brain [Chapter 2, this thesis; (Herndon and Lander, 1990)].
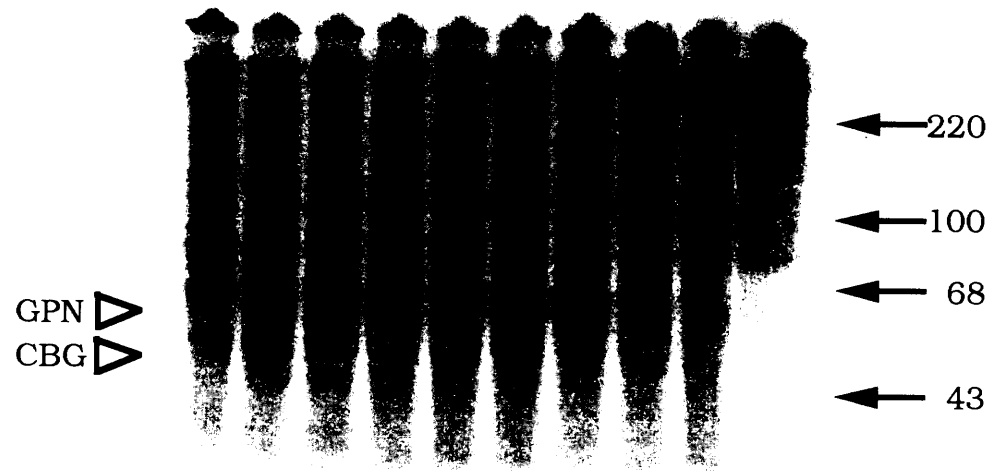
Figure 2. Determination of the Stokes radii of CBG and GPN
The elution volumes of CBG and GPN, analyzed by gel filtration on Sepharose CL6B, were determined by densitometry of the gel shown in Figure 1. (A) The elution profiles of CBG and GPN. The positions of the void volume (V_o) and the included volume (V_t) are indicated, as are the positions of the peaks of the protein standards used to calibrate the column. The precise points of the CBG and GPN peaks were determined by constructing a symmetrical curve based on the peak fraction and the two fractions adjacent to the peak fraction. (B) Comparison of CBG and GPN elution to the calibration curve derived from the standards. Calculated values of Stokes radii for CBG and GPN were 6.51 and 6.57 nm respectively. Abbreviations: TGB, thyroglobulin; β-gal, beta-galactosidase; CSE, catalase; GAPD, glyceraldehyde-phosphate dehydrogenase.

A
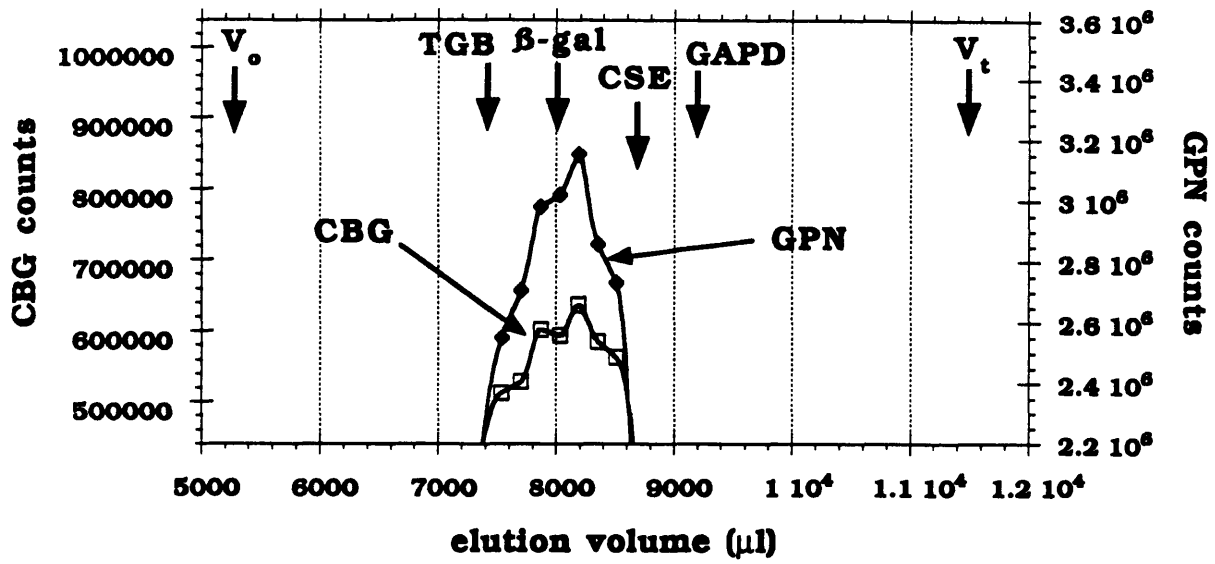
B
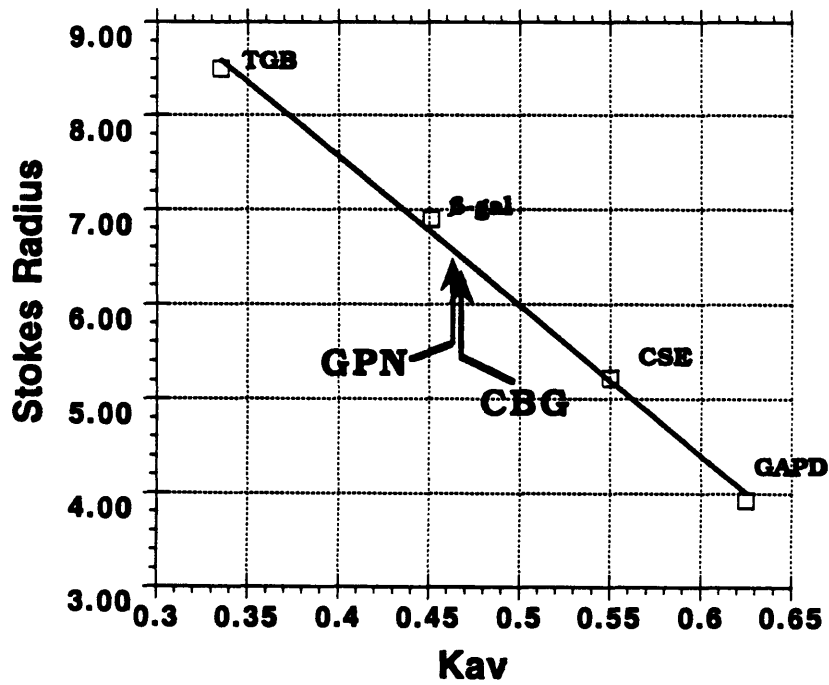
Figure 3. Calibration of sucrose density gradients
To determine the sedimentation coefficients and the partial specific
volumes of CBG and GPN, sucrose density gradient centrifugation was
performed in H_2O and D_2O. (A) The H_2O sedimentation profiles of
the standards used to calibrate the gradients is shown. The protein
standards were AP, alkaline phosphatase; MDH, malate
dehydrogenase; CAH, carbonic anhydrase; RNase, RNase A; BPTI,
bovine pancreatic trypsin inhibitor. The bottom and top of the
gradient were oriented as indicated. (C) A plot of the Svedberg
coefficient versus fraction number demonstrates that the gradient was
effectively linear. (B and D) As with (A) and (C) but in D_2O rather
than H_2O.
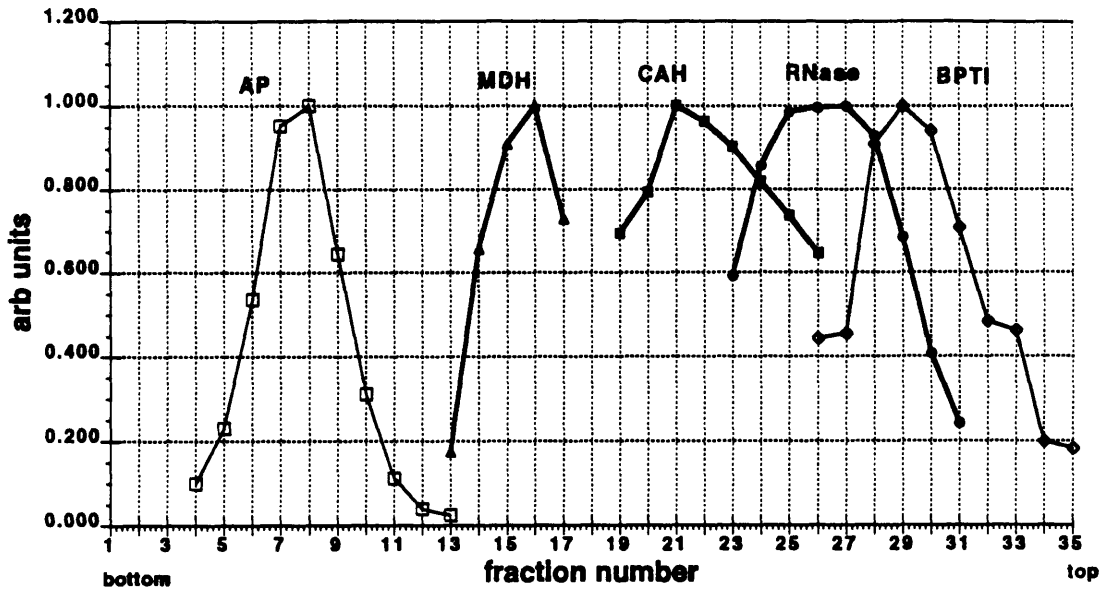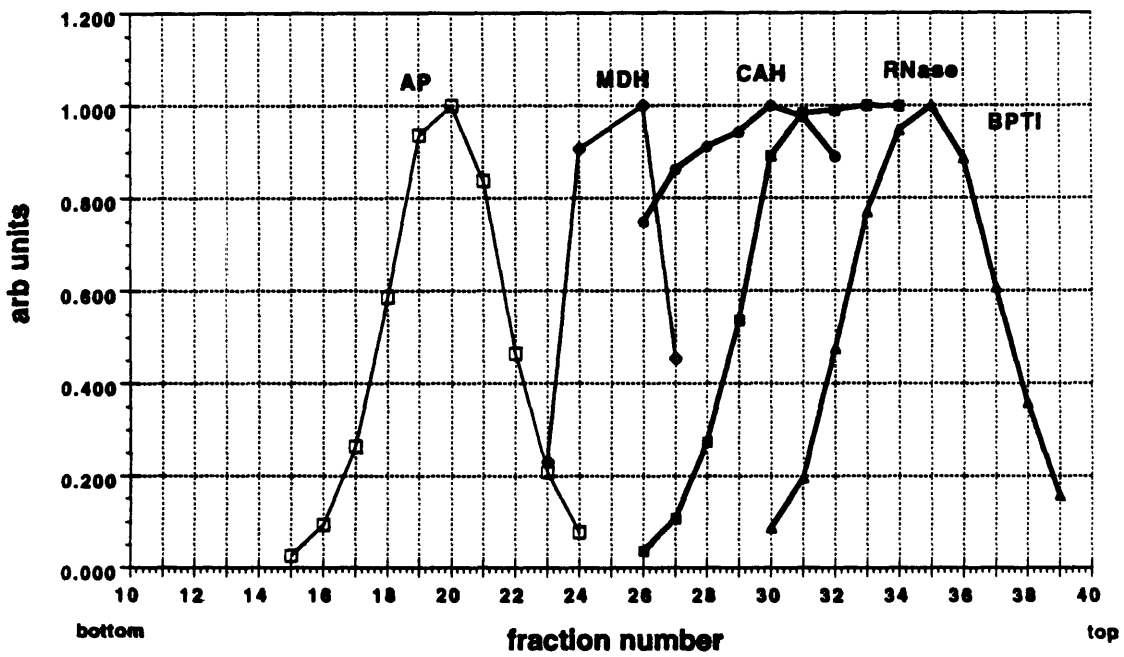
C
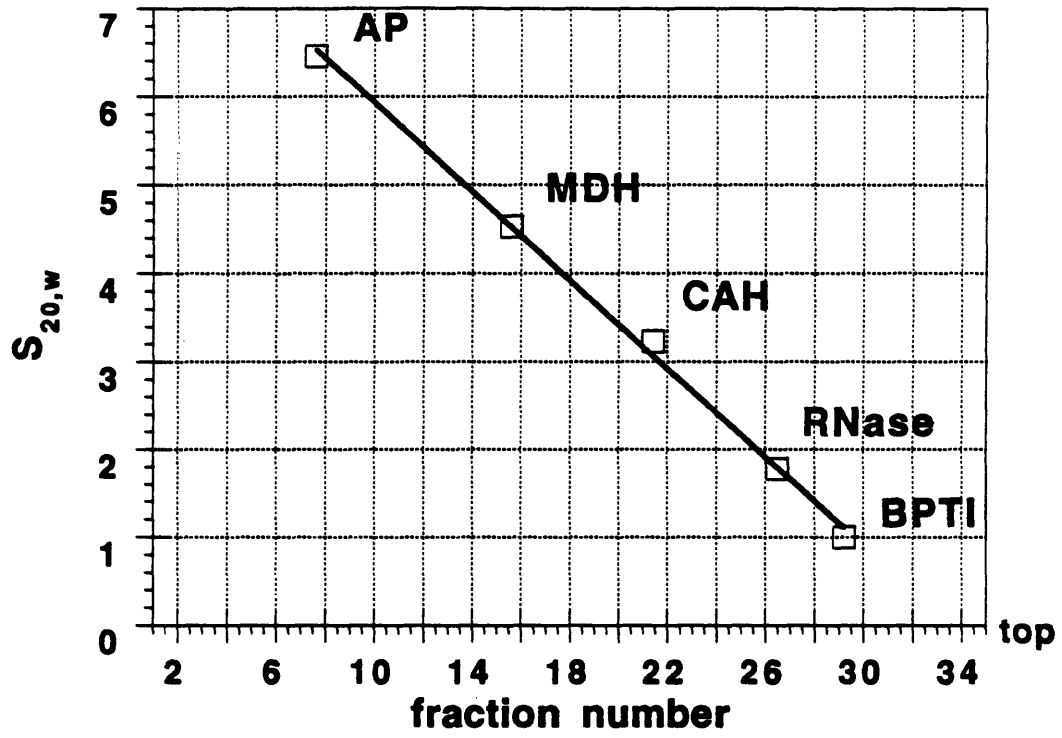
D
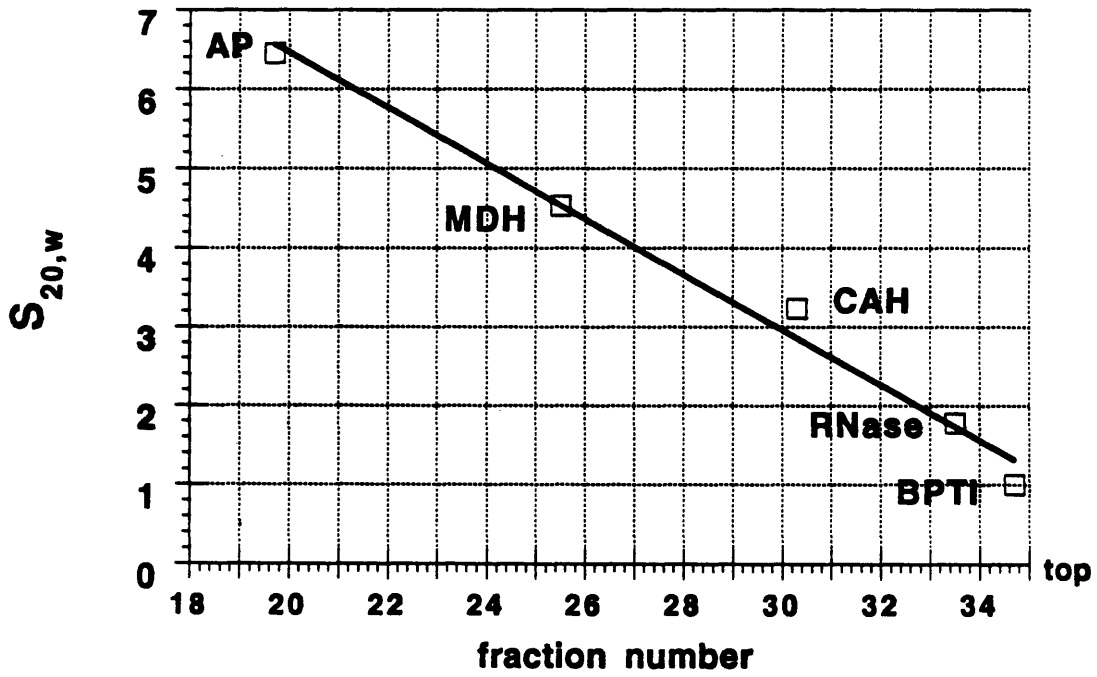
Figure 4. Sedimentation profiles of CBG and GPN in H2O and D2O sucrose density gradients. The sedimentation profiles of CBG and GPN were determined by densitometry of CBG and GPN bands on SDS-PAGE minigels, as described in Figure 1. (A) H2O sedimentation profiles. (B) D2O sedimentation profiles. The locations of the protein standards in the gradients is also indicated: BPTI, bovine pancreatic trypsin inhibitor; RNase, RNase A; CAH, carbonic anhydrase; MDH, malate dehydrogenase; AP, alkaline phosphatase. (top and bottom): indicates the orientation of the tops and bottoms of the gradients. These profiles were used to calculate the values of $S_{20,w}$, $\bar{V}$, and M_W, reported in Table 2, using equations from Clarke, (1975).

A

BPTI RNase   CAH   MDH   AP

1.00

CBG   GPN

0.80

arb units

0.60

0.40

0.20   H₂O

top 31   28   25   22   19   16   13   10   7
                                                              bottom
                    fraction number

B

BPTI RNase CAH   MDH   AP

1.00   GPN

CBG

0.80

arb units

0.60

0.40   D₂O

35   32   29   26   23   20   17   14   11   8
top   bottom
                    fraction number
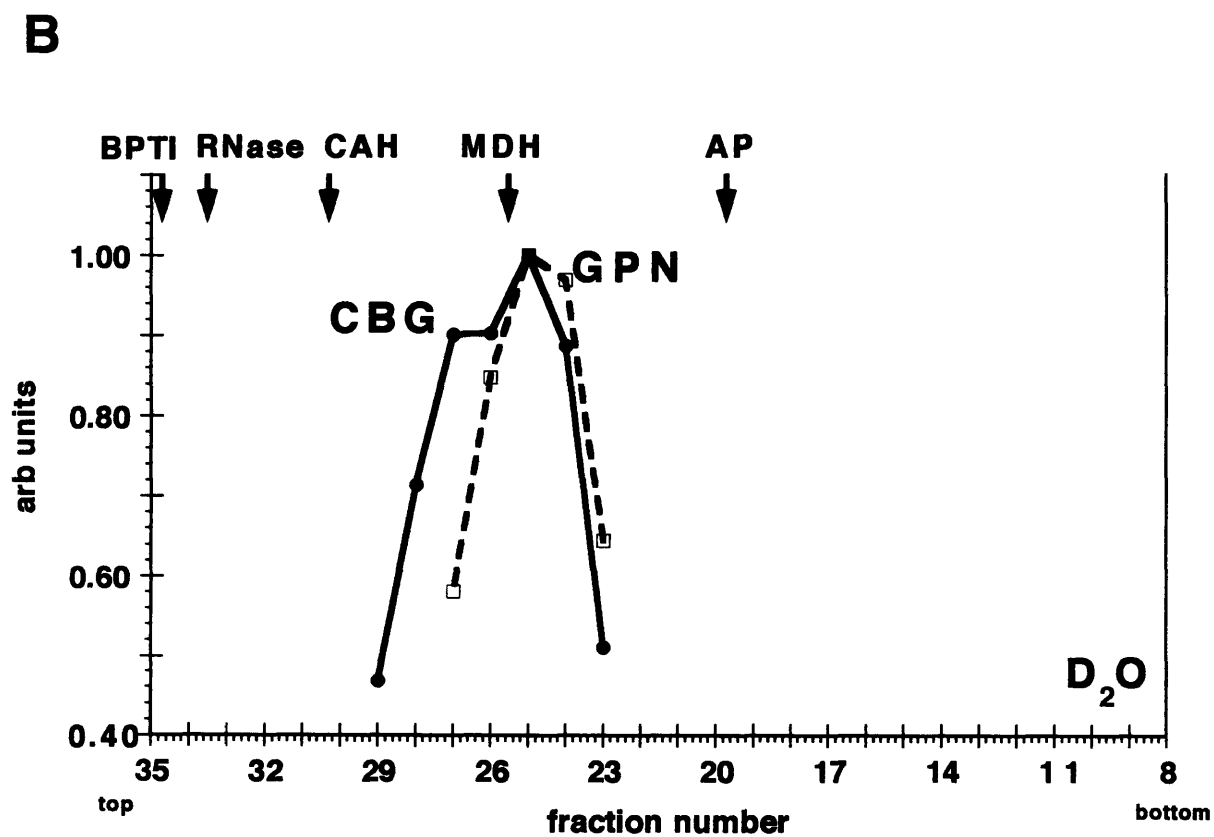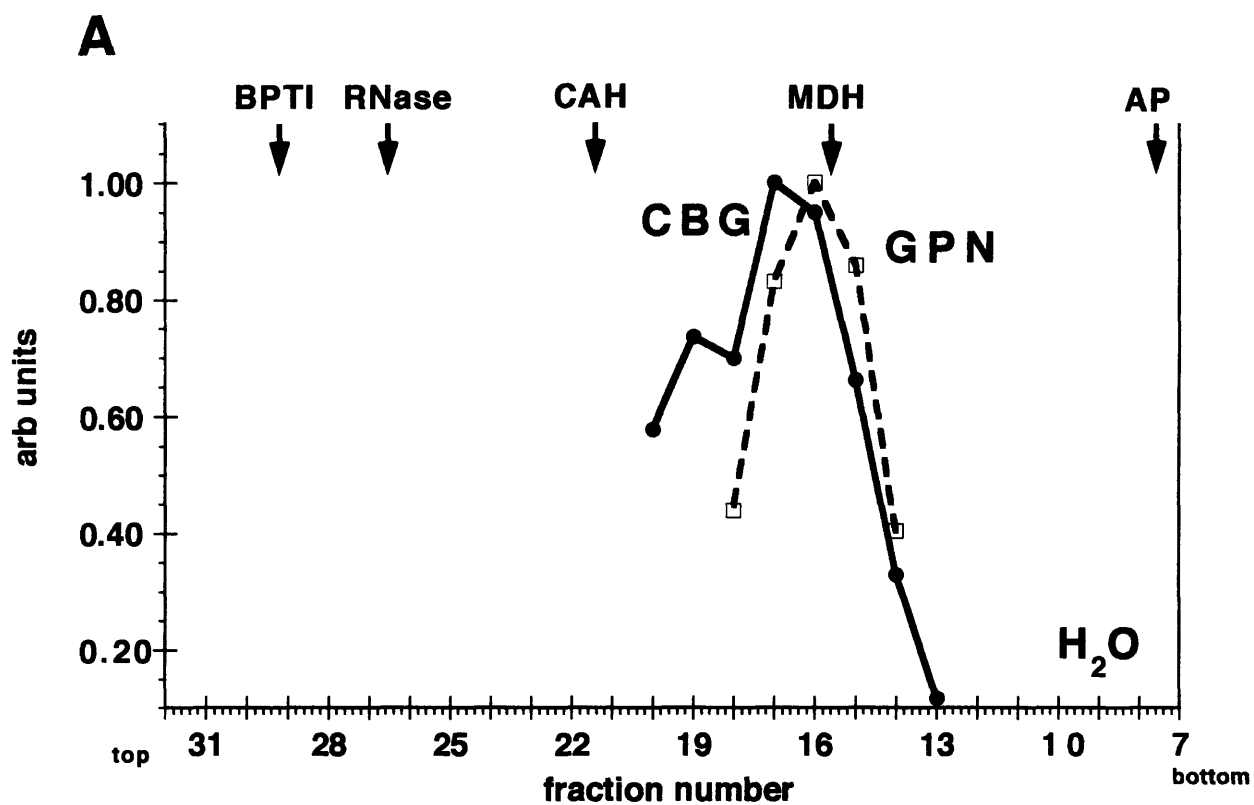
	Cerebroglycan	Glypican
$S_{20,w}$	4.15×10^{-13} sec	4.28×10^{-13} sec
$\bar{V}$	0.693 cm^3 g^{-1}	0.690 cm^3 g^{-1}
Stokes radius	6.51 nm	6.57 nm
Calculated Molecular Weight of intact PG	100 kDa	103 kDa
— Core protein	—58 kDa	—56 kDa
N-linked oligosac-— charides (predicted)	— 0 kDa	— 3-7 kDa
Total Predicted Molecular Weight of GAG	42 kDa	40-44 kDa

Table 2. Hydrodynamic data and molecular weight calculation
The hydrodynamic data obtained for cerebroglycan (CBG) and glypican (GPN) is summarized. Stokes radius was obtained by gel filtration; $S_{20,w}$, the Svedberg coefficients, and $\bar{V}$, the partial specific volumes were obtained by sucrose density gradient centrifugation in H2O and D2O as described in Materials and Methods. The molecular weights were calculated using the Svedberg equation. The total amount of glycosaminoglycan (GAG) attached per protein core was estimated by substracting the molecular weights of the core proteins. In the case of glypican, this number also includes an unknown amout of N-linked oligosaccharide. There are two potential N-glycosylation sites in GPN, and the number above is an estimate based on typical N-linked oligosaccharide structure (as described in Darnell et al, 1986), assuming either one or both sites are utilized. It has been previously reported that the GPN core protein shifts from Mr ~64 kDa to ~56 kDa in SDS-PAGE gels after treatment with N-glycanase (DeBoeck et al, 1987).

Figure 5. Gel filtration of ^{35}S-labelled HeS from P0 rat brain membrane fraction. Finely chopped prisms of P0 rat brains were cultured overnight in defined medium with Na$_2$^{35}SO$_4$. ^{35}S labelled HeS was isolated from the membrane fraction of metabolically labelled prisms as described in Materials and Methods. The diamonds show the elution profile of labelled HeS analyzed on a Sephadex G100 column; undigested HeS eluted in a broad peak centered at ~40 kDa. The squares show the elution profile of this same material after heparitinase digestion. The locations of the void volume (V_o), the included volume (V_t), and the FITC-labelled dextrans used to calibrate the column are indicated. For the FITC-dextrans, molecular weight values as determined by low angle scattering (Sigma technical services) are given in kDa. The elution profile of the undigested HeS is not shown beyond 3800 µl because free ^{35}SO$_4$, included in the run to ensure an accurate determination of Vt, obscured the HeS elution profile beyond this point. The material in the digested sample which elutes ahead of Vt likely consists of oligosaccharides remaining after digestion with heparitinase. Heparitinase does not cleave within highly sulfated domains of heparan sulfate, and even after heparitinase digestion, some oligosaccharides remain that are large enough to be precipitated with ethanol (Mary Herndon, unpublished observations). In addition, the reaction products of the chondroitinase ABC digestion were also still present in these samples.

PO HeS
PO HeS, heparitinase-treated
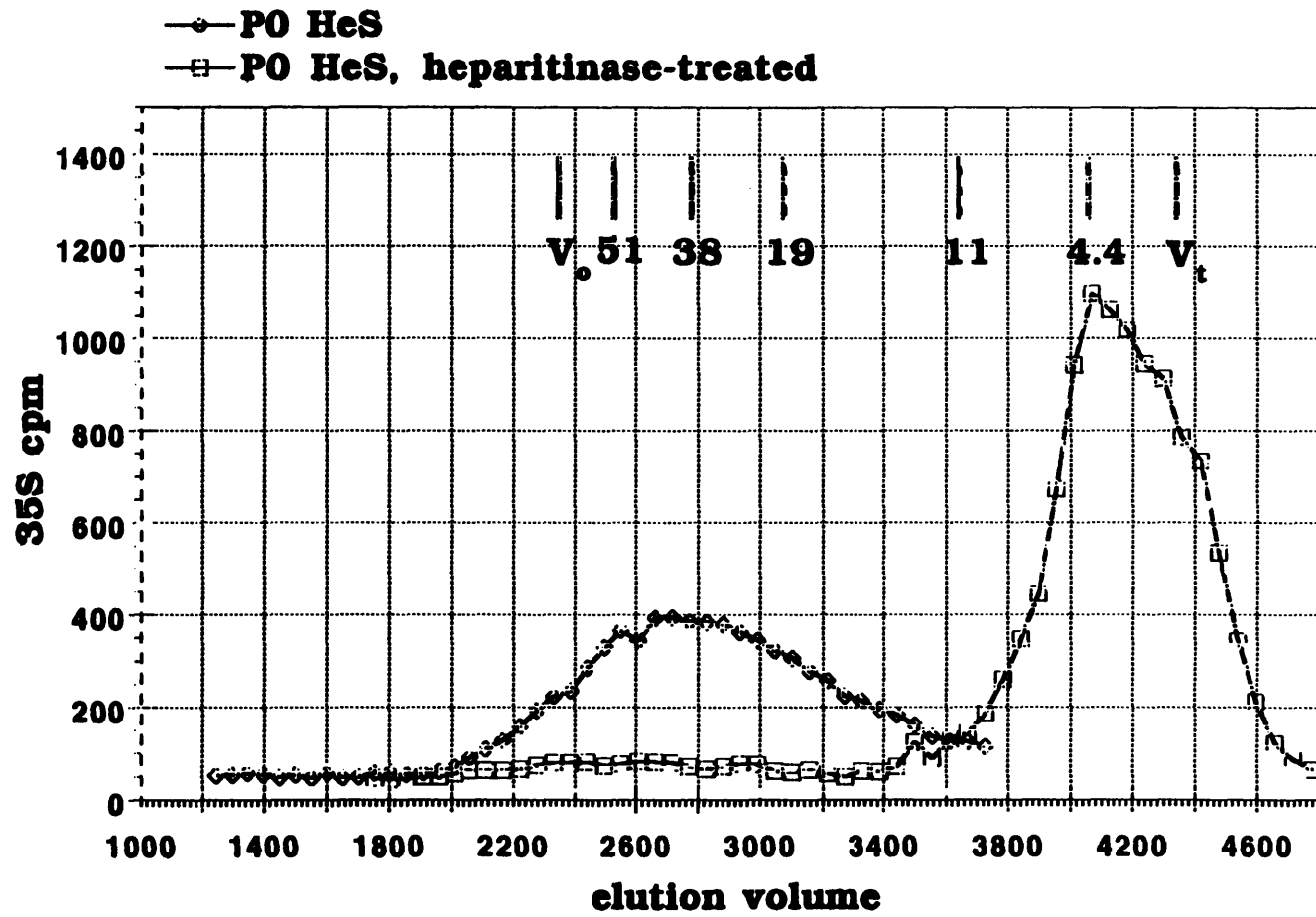
Figure 6. LN-1 affinity coelectrophoresis with ryudocan isoforms containing one or two HeS chains. ^{35}S-labelled, immunopurified ryudocan isoforms with one or two HeS chains were tested for binding to LN-1 using affinity coelectrophoresis, as described in Chapter 4. (A) Ryudocan with one HeS chain. (B) Ryudocan with two HeS chains. (C) ^{125}I-labelled low molecular weight heparin. The concentrations of LN-1 in each lane are indicated as well as the dissociation constants (K_d) calculated from the affinity patterns.

A

143 nM

B

118 nM

C

152 nM

50 30 20 10 5 2 1 0.5 0

[LN] nM

CHAPTER 6

CEREBROGLYCAN EXPRESSION IN THE DEVELOPING NERVOUS SYSTEM: CONCLUSIONS AND FUTURE DIRECTIONS

Cerebroglycan Structure, Expression, and Possible Functions in the Developing Nervous System: A Comparison to other Gypican Family HSPGs

This study reports on the identification, cloning, expression pattern, and biochemical characterization of a novel cell surface HSPG, cerebroglycan (CBG). CBG was found to be related to the glycosylphosphatidylinositol (GPI)-linked cell surface HSPG glypican. While this work was in progress, data on neuronal expression of glypican, as well as other newly discovered glypican family members, were reported. The section below summarizes what has been learned about CBG structure and expression in the developing nervous system by way of comparison to other members of the glypican family.

Structure

The HSPG glypican, cloned from human lung fibroblasts, was the first member of the glypican family to be characterized at the protein level (David et al., 1990). Glypican was found to be anchored to the cell surface by a glycosylphosphatidylinositol (GPI) lipid anchor. Another family member, OCI-5, had been previously identified as a developmentally regulated transcript in a rat primitive intestinal epithelial cell line (Filmus et al., 1988). However, the relationship between glypican and OCI-5 was not immediately recognized. With the cloning from neonatal rat brain of cerebroglycan (CBG), the third glypican family member (Chapter 2, this thesis) and, more recently, K-glypican from mouse kidney (Watanabe et al., 1995), it has become possible to identify the common structural features of the glypican family members. All members of the family have, in addition to an amino-terminal signal sequence, a stretch of hydrophobic amino acids near the carboxy terminus that can serve as a cleavage and addition signal for the GPI-anchor. In addition, glypican family members all have at least one conserved, putative GAG attachment site in a cluster near the carboxy terminus.

Elsewhere, the homology between glypican family members is moderate but extensive, being characterized by blocks of identity or conservative substitutions separated by stretches of divergent sequence (See Figure 6 of Chapter 2). Most striking is the absolute

conservation of a pattern of 14 cysteine residues, found in all family members so far identified. This conserved pattern of cysteines suggests that these PGs may have similar, compact tertiary structures, stabilized by disulfide bonds. This view is supported by the fact that both glypican and cerebroglycan show a dramatic increase in apparent molecular weight in SDS-PAGE gels upon reduction (Herndon and Lander, 1990). In addition, a proteolytic fragment obtained in some preparations of K-glypican can be only be observed in SDS-PAGE gels upon reduction (Watanabe et al., 1995).

More recently, the product of the Drosophila gene *dally* was found to encode a protein related to the ancestor of the glypican family PGs (Nakato et al., 1995). The pattern of 14 cysteine residues found in the mammalian glypicans is present in the dally protein as well, indicating that this well conserved structural motif is of an ancient origin. It is also of interest to note that the mature polypeptides of human glypican and rat glypican are 91% identical at the amino acid level (Litwack et al., 1994). In contrast, the ectodomains of the human and mouse syndecan-1 are only 70% identical; only a protease susceptible domain near the plasma membrane and the regions surrounding the GAG attachment sites are strongly conserved (Bernfield et al., 1992). This suggests that while a primary function of the syndecan family ectodomains may be to present GAG chains at the cell surface, the protein cores of the glypican family members may serve a more complex function that has constrained their evolution.

Among the mammalian glypican family members, glypican, K-glypican, and cerebroglycan are all more closely related to each other than to OCI-5. This suggests that there may be a subfamilies within the glypican family. It will be interesting to see if, in the future, additional glypican family members are identified.

Expression

At a gross level, northern blot analyses have shown that all four mammalian glypican family members are expressed in the rat brain [Chapter 2, Figure 8, this thesis; (Lisanti et al., 1994; Watanabe et al., 1995)]. Glypican mRNA levels are roughly constant across

development. OCI-5 and K-glypican mRNAs, abundant in embryonic rat brain, are present at lower levels in neonates, and adults, and CBG mRNA is only detected in embryonic and early postnatal rat brain.

In situ hybridization and immunolocalization experiments have revealed a complex pattern of glypican mRNA and protein expression. In early embryos (e.g. embryonic day 14 rat), glypican mRNA expression is primarily associated with the nervous system (Karthikeyan et al., 1994; Litwack, 1995). In particular, high levels of expression are found in the ventricular zones where proliferating neuronal and glial precursors are found (Litwack, 1995). Much lower levels of expression are observed outside the nervous system at this stage (Litwack, 1995). By embryonic day 18, glypican mRNA expression is widespread throughout the nervous system in a pattern that generally correlates with the number of cell bodies (Litwack, 1995).

Widespread non-neuronal expression of the glypican mRNA is also apparent at this stage. For example, in the developing bones of embryonic day 18 rats, glypican mRNA is detected in a pattern consistent with expression by differentiated osteoblasts, the cell type responsible for deposition of the bone matrix during bone formation (Litwack, 1995). Other cell types in the developing bones, such as chondrocytes, appear to be glypican-negative.

In the post-natal stages, neuronal glypican mRNA expression begins to be restricted, disappearing from some structures such as the corpus striatum and the cerebellar granule cell layer, until the adult pattern of expression is achieved (Litwack, 1995).

In the adult rat nervous system, the highest levels of glypican expression are found in the hippocampus, the dorsal thalamus, and motor nerve nuclei (Karthikeyan et al., 1994; Litwack et al., 1994). Many brain regions express no glypican mRNA above background. A common feature of many of the glypican positive brain structures in the adult is that they contain projection neurons, neurons with long axons projecting to distant targets, rather than local circuit neurons whose shorter axons contact close neighbors (Litwack et al., 1994). In structures where neurons and glia are morphologically distinguishable (such as hippocampus and dorsal root ganglia), glypican mRNA

expression is associated with neurons rather than glia (Litwack et al., 1994). However, glial expression of glypican cannot be ruled out, and there is one report that Schwann cells express glypican in vitro (Carey et al., 1993).

The observation that many of the glypican-positive brain structures contain projection neurons is in good agreement with glypican immunostaining which has revealed glypican protein in axon tracts of the developing and adult nervous system (Karthikeyan et al., 1994; Litwack, 1995). Several neuronal GPI-anchored proteins have been found to be localized preferentially (or polarized) to axons, suggesting one possible function for GPI attachment [examples include TAG-1, Thy-1, and F3/F11 (Dodd et al., 1988; Dotti et al., 1991; Faivre-Sarrailh et al., 1992)]. Glypican immunostaining in the pyramidal layer of the hippocampus and in dorsal root ganglia is consistent with axonal polarization, but the data are not conclusive (Litwack, 1995). In vitro cultured hippocampal neurons express glypican uniformly over the entire cell surface (Asli Kumbasar, unpublished data), and there has been one report of glypican expression on dendrites in vivo (Karthikeyan et al., 1994), so the question of glypican localization in neurons at the cellular level remains open.

Non-neuronal glypican staining in the mouse embryo demonstrates widespread, but restricted patterns of expression (David Litwack, unpublished data). For example, glypican staining is observed in striated and smooth muscle, but not in cardiac muscle, and in epidermal but not endothelial cells. Expression was also detected in the periosteum of bones -- one region where osteoblasts are found -- consistent with in situ hybridization experiments that indicated osteoblasts express the glypican mRNA.

In contrast to glypican, cerebroglycan mRNA and protein are found in a simple, striking, regional and developmental pattern. Cerebroglycan (CBG) mRNA is completely restricted to the nervous system throughout development (See Figure 1, Chapter 3). Within the rat nervous system, the CBG message is globally expressed in early embryos, but by embryonic day 19 begins to disappear from structures, such as the spinal cord, where neuronal development is advanced. By

postnatal day 7, only late developing structures such as the cerebellum, the dentate gyrus, and the olfactory bulb still retain CBG mRNA expression (Figure 2, Chapter 3). No CBG mRNA is detected in the adult rat brain.

Closer examination of CBG mRNA expression at several stages has shown that CBG expression is associated with immature neurons in early, post-mitotic stages. Such neurons are, as a group, undergoing processes such as cell body migration and axon outgrowth. Little or no expression is detected in regions containing proliferative neuroepithelium. For example, in the postnatal day 7 cerebellum, CBG expression is associated with the deep half of the external granule layer (the so called premigratory zone), a region containing neurons that have ceased to divide and have begun axon outgrowth and cell body migration (Figure 3, Chapter 3). The internal granule layer, containing neurons that have completed migration and axon outgrowth, is negative. These data are consistent with the view that CBG expression is tightly correlated with the motile phases of neuronal life.

As with glypican, immunolocalization of cerebroglycan is consistent with the pattern of CBG mRNA expression. CBG protein is found only in the developing nervous system, and, in particular is expressed by post-mitotic neurons and not proliferating neuroepithelium (Litwack, 1995). The CBG protein is also found in axon tracts in the developing nervous system, consistent with the observation that CBG mRNA is expressed by populations of neurons undergoing axon outgrowth (Figure 4, Chapter 3). In addition, immunolocalization of CBG in the developing dentate gyrus strongly suggests that, at least in that structure, CBG protein is polarized to axons (Litwack, 1995).

Recently, the first data for neuronal expression of the other glypican family members have become available. K-glypican mRNA is highly expressed in the ventricular zones of the embryonic day 13 mouse nervous system (Watanabe et al., 1995). In this regard, the expression pattern of K-glypican appears to be similar to that of glypican at a similar stage of development. Currently, no immunolocalization data for K-glypican are available. Although, as

mentioned above, OCI-5 mRNA has been detected in the rat brain, in the absence of in situ hybrization data, it is not possible to determine whether OCI-5 is expressed by neurons, glia, or non-parenchymal cells (e.g. vascular tissue, meningeal tissue, blood cells, or choroid plexus).

The *dally* gene product in Drosophila is also expressed in the nervous system (Nakato et al., 1995). Studies with enhancer trap lines in which a ß-galactosidase (ß-gal) reporter is inserted into the *dally* locus have revealed that dally mRNA is expressed in the lamina precursor cells (LPCs) adjacent to the anterior outer proliferative center of the fly brain. The LPCs differentiate into lamina neurons that receive input from axons of the photoreceptors in the eye, and, in turn project to other targets in the Drosophila nervous system. The LPCs are situated along the lamina furrow, which marks the progress of differentiation in the lamina, and LPCs at distinct stages of their final two rounds of cell division are positioned at distinct locations along the lamina furrow (Selleck et al., 1992). LPC's complete the first of these final two divisions and then remain in G1 phase of the second round until they receive a signal from innervating photoreceptor axons that stimulates them to complete their final division and differentiate into lamina neurons (Selleck et al., 1992).

Expression of the dally enhancer trap ß-gal reporter RNA is restricted to LPCs in G2/M phase of their first round of division. Colocalization of ß-gal protein and cyclin-B (a marker for late G2/early M phase) demonstrated that the dally enhancer trap insertion is expressed in a cell cycle-dependent fashion at several locations in the developing fly nervous system, although ß-gal positive cells that did not overlap with regions of cyclin-B expression were also observed (Nakato et al., 1995).

Possible Functions

Although it is unclear whether they are true nulls or merely hypomorphs, the phenotypes of dally mutant flies provide the first direct evidence for in vivo function of a member of the glypican family of HSPGs. In flies homozygous for the most severe alleles of *dally* analyzed thus far, the epithelium containing the anterior outer proliferative center and the LPCs is disordered, as is the eye disc, and

the flies show a severe defecit in phototaxis (Nakato et al., 1995). Several other, nonneural structures such as antennae and genitalia also show abnormal morphology. To further characterize the *dally* phenotype, *dally* flies were stained with propidium iodide to visualize mitotic figures, and with antibodies to cyclin-B to mark cells in G2, or pulsed with bromodeoxyuridine to mark cells in S phase. These experiments demonstrated that in the LPCs of *dally* flies, there is a complete absence of cells in their second S, G2, or M phases. In addition, the region containing cells in their first G2 phase is enlarged, and the M phase cells of the first division are displaced. Taken together, these observations suggest the loss of functional dally protein causes a delayed entry into the first M phase in the lamina furrow. This phenotype is the source of the name *dally* which stands for *division abnormally delayed*. In the developing eye disc, where the morphogenetic furrow is an analogous structure to the lamina furrow, a similar phenotype is seen in *dally* flies. The second round of division in the morphogenetic furrow does occur, but, as with the lamina furrow, the zone of cells in their first G2 phase is enlarged, and cells in their first M phase are displaced from their normal location relative to the furrow (Nakato et al., 1995).

The apparent involvement of the dally protein in regulating the transition of proliferating precursors to differentiating neurons is especially intriguing in light of the expression patterns of glypican and cerebroglycan in the early rat nervous system. As discussed above, in early embryos glypican is highly expressed by cells in proliferative neuroepithelia whereas cerebroglycan expression is seen in the newly post-mitotic neurons that overlie these proliferative zones. Close comparison of adjacent sections has shown that the expression patterns of glypican and cerebroglycan are mutually exclusive, or nearly so, in these early neural structures (Litwack, 1995). Thus the expression of glypican family members appears to be tightly regulated around the transition from proliferative neuronal precursor to immature neuron, suggesting that individual members may play distinct roles in facilitating this transition. In the future it will be important to determine which if any of the mammalian glypican family members is able to complement the *dally* phenotype. As Nakato et al

have pointed out (Nakato et al., 1995), the dally protein, as a putative integral membrane HSPG, is an excellent candidate receptor for any of several extracellular HeS-binding factors that may influence neuronal differentiation.

The expression of glypican and cerebroglycan on axons, and the presence of K-glypican, glypican, and OCI-5 in the adult nervous system suggest that whatever role, if any, glypican family HSPGs may have in regulating precursor proliferation, they may also be adapted for additional functions in the developing and adult nervous system. Potential clues concerning these functions come from studies of PG function and GPI-anchored molecules in other systems. A survey of the literature suggests the following possibilities.

Glypican family members may act as coreceptors for HeS binding growth factors.

Several HeS-binding growth factors are found in the developing nervous system (Gonzalez et al., 1990; Merenmies and Rauvala, 1990; Fu et al., 1991; Hutchins and Jefferson, 1992; Reddy and Pleasure, 1992; Chang et al., 1994; Ho et al., 1995). Perhaps the best characterized of these is bFGF or FGF-2. Studies with fibroblasts and Chinese hamster ovary cells deficient in HeS synthesis have demonstrated that, at least under certain conditions, cell surface HeS appears to be required for FGF-2 binding to its high affinity receptor (Yayon et al., 1991; Ornitz et al., 1992), as well as for an FGF-2 induced mitogenic response (Rapraeger et al., 1991; Ornitz et al., 1992). Studies such as these have led to the concept of PGs as coreceptors. In the coreceptor hypothesis, cell surface HSPGs function to facilitate the binding of an extracellular ligand to its classical, high affinity receptor (Bernfield et al., 1992). Some authors have reported that a biologically active FGF-2/HSPG complex is released from cell surfaces by the bacterial enzyme phosphoinositol-specific phospholipase C (PIPLC), which cleaves GPI anchors. These studies point to the possibility that glypican family HSPGs in particular serve as coreceptors for members of the FGF family (Bashkin et al., 1989; Brunner et al., 1991).

Glypican family members may act as receptors for extracellular matrix molecules.

Besides soluble factors, other potential ligands for glypican family members include extracellular matrix molecules such as laminins (LNs), fibronectin (FN), thrombospondins (TSPs), and vitronectin (VN). Both LNs and FN have been shown to promote neurite outgrowth of various neuronal cell types, and LN-1 can actually guide neurites in vitro (Gundersen, 1987; Rogers et al., 1989; Sanes, 1989; Reichardt and Tomaselli, 1991). Both have the ability to bind to HSPGs, and both are present in the developing nervous system. FN, LN-1 and LN-3 are found transiently in the subplate (SP) of the developing rat cortex at developmental time points that coincide with the outgrowth through the SP of CBG positive axons [Chapter 3, this thesis; (Sheppard et al., 1991; Hunter et al., 1992)]. Similarly, axons of developing retinal neurons, neurons of the dorsal root ganglia, and spinal cord motoneurons all navigate through regions that contain LN-1 (Rogers et al., 1986; Cohen et al., 1987; McLoon et al., 1988). Immunolocalization of glypican and CBG has revealed that many of these axons may be expressing CBG and/or glypican [(Karthikeyan et al., 1994; Litwack, 1995), Chapter 3, this thesis]. The observation that both CBG and glypican bind to LN-1 with affinities in the nanomolar to subnanomolar range further supports the idea that they could serve as LN-1 receptors in vivo (see Chapter 4).

A recent study has demonstrated that neuronal cells can utilize a cell surface HSPG to attach and spread on FN fragments that contain a heparin-binding domain (Haugen et al., 1992). The adhesion and spreading can be inhibited by heparitinase, or by antibodies raised to a GPI-anchored HSPG in melanoma cells. These same antibodies recognize a core protein of Mr 51 kDa, that shifts upward dramatically upon reduction on western blots of neuronal PG preparations. Taken together, the biochemical data suggest that the cell surface PG in these studies is a member of the glypican family, and that it plays a role in the interactions of neuronal cells with FN.

The heparin binding molecules TSP-1 and VN are both present in the developing retina (Neugebauer et al., 1991), a structure where high levels of CBG mRNA were found to be expressed. Both TSP-1 and

VN promote attachment and neurite outgrowth by retinal neurons in vitro, and attachment to TSP-1 is inhibited by the addition of soluble heparin (Neugebauer et al., 1991), indicating a possible role for cell surface HSPGs in the interaction of retinal neurons with this molecule.

TSP-1 is also expressed widely throughout other areas of the developing nervous system, including the cerebellum, where it is found associated with granule cell neurons in the premigratory zone of external granule layer and in the molecular layer (O'Shea et al., 1990). Interestingly, an antibody to TSP-1 was found to block granule cell migration from the external granule layer to the molecular layer, suggesting TSP-1 may directly participate in this process (O'Shea et al., 1990). As described in Chapter 3 above, CBG mRNA in the developing rat cerebellum is also specifically associated with neurons of the premigratory zone raising the possibility that CBG and TSP-1 interact on the cell surfaces of these neurons.

Glypican family HSPGs as signalling molecules?

Antibody crosslinking of GPI-anchored proteins on the surfaces of T cells results in proliferation and interleukin-2 secretion, suggesting that GPI-anchored molecules may be involved in signalling (Presky et al., 1990). Several reports have been published in which GPI-anchored proteins are found to be associated with src family tyrosine kinases (Stefanova et al., 1991; Thomas and Samelson, 1992; Draberova and Draber, 1993; Stefanova et al., 1993), suggesting one way signalling through a GPI-anchored receptor could occur. Others have suggested that the GPI moiety itself could be involved. Although inherently controversial because of the localization of the GPI anchor to the outer leaflet of the cell membrane, this idea is not without some experimental support: PIPLC treatment alone can cause cellular "activation" (Depper et al., 1984; Ransom and Cambier, 1986), and exogenously added GPI or phosphoinositol stimulates cytokine secretion by monocytes, and simulates the effects of insulin on adipocytes (Saltiel and Sorbara-Cazan, 1987; Schofield and Hackett, 1993).

Biochemical evidence that seems to support the notion that GPI anchored molecules can associate somehow with intracellular src

family kinases comes from studies of caveolae. Caveolae are caveolin-rich membrane microdomains that are thought to be involved in the uptake of small extracellular molecules by a process distinct from coated-pit endocytosis (Anderson et al., 1992). Biochemical characterization of caveolae-like subcellular particles has revealed that they are enriched for GPI-anchored proteins and contain signal transducing molecules, including GTP-binding proteins and src family protein tyrosine kinases (Sargiacomo et al., 1993; Lisanti et al., 1994). Although caveolin mRNA has not been detected in the brain, it is possible that similar domains, in which GPI-anchored molecules are clustered together with signal transducing proteins, exist on the surfaces of neural cells.

Future Directions

Cerebroglycan-null transgenic mice.

The genomic clones obtained in Chapter 1 lay the groundwork for a project to generate cerebroglycan null transgenic mice. The tissue specific and developmental stage specific expression of cerebroglycan makes it an attractive candidate for targeted disruption: no cerebroglycan expression has been detected outside of the nervous system, and within the nervous system, cerebroglycan is expressed by post-mitotic, immature neurons. Neurogenesis occurs largely during the second half of embryogenesis. For example, in the developing rat cortex, the earliest neurons are born beginning around embryonic day 11-12. In addition, embryos with severe neurological defects are often brought to term. Unless cerebroglycan expression is different in the very early embryo than at later stages (a possibility that cannot be dismissed), the phenotype of a cerebroglycan homozygous null animal (if there is an obvious one) should be relatively easy to score. The fact that the loss or reduction of function phenotype of the Drosophila homologue, dally, is readily detected would seem to favor the chances that targeted disruption of cerebroglycan would be informative.

Identification of cerebroglycan transcriptional control elements.

As indicated in Chapter 2, a considerable portion of the mouse genomic locus upstream of the start site of transcription has been

285

obtained. In addition, some of the intron/exon boundaries have been identified. This information should facilitate the identification of the transcriptional control elements that confer on cerebroglyan its striking pattern of tissue specific, temporally restricted expression. The necessary cis-acting elements may all be contained in the promoter, or additional elements within the introns may be required. One approach would be to generate expression vectors containing a reporter gene such as β-galactosidase downstream from various fragments of the cerebroglycan genomic sequence from the region 5' to the transcriptional start site. If necessary, intronic sequence could also be incorporated. These constructs could then be injected into the male pronuclei of in vitro fertilized oocytes to generate chimeric embryos in which the expression of the reporter gene would be analyzed. Because the extent of incorporation of the transgene will vary from embryo to embryo, a number of embryos for each construct will have to be analyzed to score expression patterns with confidence.

Using this approach, a 4 kb fragment upstream from the cerebroglycan transcriptional start site has been tentatively identified as containing the cis-acting elements required for the cerebroglycan expression pattern (John Fesenko, unpublished data). Once identified, these elements might prove useful for further experiments. For example, a line of transgenic mice could be generated in which all the immature, motile neurons of the central nervous system were labelled by the expression of an easily assayble reporter gene. These mice would be useful for identifying extrinsic factors that influence the processes of neuronal differentiation (see below for one example).

Identification of extrinsic factors controlling cerebroglycan expression

Studies of cerebroglycan expression in olfactory epithelial (OE) explant cultures suggested that extrinsic factors may influence cerebroglycan expression. In vivo, cerebroglycan begins to be expressed in neurons around the time of the last mitosis, and continues to be expressed during the period encompassing axon outgrowth and cell body migration (See Figure 4, Chapter 3). In OE explant cultures, however, this was not seen to be the case. A large number of neuronal precursors migrate away from the explants in this

system before dividing to give rise to NCAM-positive immature neurons. These immature neurons should also be cerebroglycan-positive based on expectations from expression patterns in vivo, but no cerebroglycan expression was seen among these cells. Instead, cerebroglycan expression was only detected in cells that had remained associated with the explants, suggesting that extrinsic factors supplied either by the explant, or by the neuronal precursors clustered around the explant, were responsible either for initiating cerebroglycan expression, or for maintaining cerebroglycan expression.

One potential strategy for identifying these putative extrinsic factors would involve a biochemical approach. Homogenates of olfactory epithelium could be tested for the ability to restore cerebroglycan expression to dispersed OE neuronal cells in vitro. Such an approach would be facilitated by the availability of the cerebroglycan-reporter gene line of transgenic mice described above. After preliminary tests to show that the reporter construct gave an expression pattern in OE cultures identical to that described in Chapter 3, OE cells isolated from these transgenic mice could be used in assays to search for extrinsic factors affecting cerebroglycan expresson. A parallel approach of screening already identified factors for the ability to restore CBG expression in these cells would also be a part of this strategy. This parallel approach might yield results first, especially if the factors being searched for are present at very low concentrations in the OE. The use of whole brain homogenates rather than OE homogenates would also simplify this approach although this runs the risk of isolating a different factor than the one in the OE (which might be interesting, nontheless).

Analysis of glycosaminoglycan attachment to cerebroglycan

The molecular weight data in Chapter 5 suggest raise the possibility that at least a fraction of cerebroglycan core proteins may be substituted with only one HeS chain. This model could be tested by construction of cerebroglycan isoforms in which the five potential glycanation sites are systematically eliminated. Such an approach has already been successfully utilized to determine that glycosaminoglycan

chains are attached to all three acceptor sites of ryudocan (syndecan-4) (Shworak et al., 1994a).

This approach would also yield cerebroglycan for which it is known for certain that only one HeS chain is attached. This isoform could then be used to confirm that the high affinity binding of cerebroglycan to LN-1 observed in Chapter 4 is not the result of multivalent HeS chains on a single cerebroglycan core protein binding to multiple HeS binding sites on LN-1.

Overexpression of unglycanated cerebroglycan core protein in this same system could also provide material for an attempt at measuring a potentially rather weak interaction between the cerebroglycan core protein and LN-1. The results of Chapters 4 and 5 point to the possibility that such an interaction occurs in the binding of intact cerebroglycan with LN-1. The affinity of this interaction could be impossible to measure by affinity coelectrophoresis since the LN-1 concentrations required might be impossible to attain. The availability of large amounts of unglycanated cerebroglycan core protein could provide one way to circumvent this problem. For example, limiting amounts of labelled LN-1 could be used in binding reactions with excess cerebroglycan core, and LN-1/cerebroglycan core protein complexes could be collected using an anti- cerebroglycan affinity matrix. The observation that cerebroglycan synthesized in 3T3 cells binds to LN-1 with the same high affinity observed for cerebroglycan derived from brain suggests that high affinity binding to LN-1 is not the result of some sort of special brain-specific modification to the cerebroglycan core or chains, and supports use of a cell line approach in further elucidating the mechanism of cerebroglycan-LN-1 binding.

Glypican family members have, without exception, been found to be homoglycans, attaching HeS chains only. In contrast, syndecan-1 and ryudocan are heteroglycans, attaching both HeS and CS. A recent study has shown that all three glycosaminoglycan attachment sites in the ryudocan core protein can attach either HeS or CS. The factors controlling whether a site will be an HeS attachment site, a CS attachment site, or both have not been determined. One way to approach this problem might be to make limited chimeras of ryudocan and cerebroglycan or glypican. As a first step, glycosaminoglycan

attachment sites from the two molecules could be substituted for each other. A result in which a homoglycanated cerebroglycan site becomes a heteroglycanated site in the context of the ryudocan core protein, or vice versa would argue for a model where structural elements outside of the attachment site regulate the type of chain attached. The converse result (e.g. a homoglycanated cerebroglycan site remains homoglycanated in ryudocan) would argue that structural elements within the attachment site control glycanation.

Conclusion

The identification and characterization of cerebroglycan have (1) contributed to the realization that a whole family of glypican-related HSPGs exists; (2) demonstrated that PG core proteins can be expressed in a tissue-specific fashion; (3) provided additional evidence that cell surface HSPGs may be involved in neuronal cell migration and axon outgrowth; (4) demonstrated that intact PGs can bind to their ligands with much higher affinity than that of free heparin or heparan sulfate; and (5) provided evidence suggesting that interactions of HSPG core proteins with HSPG ligands may be more widespread than previously appreciated. Future studies promise to reveal more, not only about the cell biology of cerebroglycan, but also about the processes underlying the development of the nervous system.

References

Anderson, R.G.W., B.A. Kamen, K.G. Rothberg, and S.W. Lacey. 1992. Potocytosis: sequestration and transport of small molecules by caveolae. Science (Wash. DC) 255:410-411.

Bashkin, P., S. Doctrow, M. Klagsbrun, C.M. Svahn, J. Folkman, and I. Vlodavsky. 1989. Basic fibroblast growth factor binds to sub-endothelial extracellular matrix and is released by heparitinase and heparin-like molecules. Biochemistry 28:1737-1743.

Bernfield, M., R. Kokenyesi, M. Kato, M.T. Hinkes, J. Spring, R.L. Gallo, and E.J. Lose. 1992. Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu. Rev. Cell Biol. 8:365-393.

Brunner, G., J. Gabrilove, D.B. Rifkin, and E.L. Wilson. 1991. Phospholipase C release of basic fibroblast growth factor from human bone marrow cultures as a biologically active complex with a phosphatidylinositol-anchored heparan sulfate proteoglycan. J. Cell Biol. 114:1275-1283.

Carey, D.J., R.C. Stahl, V.K. Asundi, and B. Tucker. 1993. Processing and subcellular distribution of the Schwann cell lipid-anchored heparan sulfate proteoglycan and identification as glypican. Exp. Cell Res. 208:10-18.

Chang, D., D. Wen, M. Schwartz, and Y. Yarden. 1994. Brain neurons and glial cells express Neu differentiation factor/heregulin: a survival factor for astrocytes. Proc. Natl. Acad. Sci. (USA) 91:9387-9391.

Cohen, J., J.F. Burne, C. McKinlay, and J. Winter. 1987. The role of laminin and the laminin/fibronectin receptor complex in the outgrowth of retinal ganglion cell axons. Dev. Biol. 122:407-418.

David, G., V. Lories, B. Decock, P. Marynen, J. Cassiman, and H.V.d. Berghe. 1990. Molecular cloning of a phosphatidylinositol-anchored membrane heparan sulfate proteoglycan from human lung fibroblasts. J. Cell Biol. 111:3165-3176.

Depper, J.M., W.J. Leonard, M. Kronke, P.D. Noguchi, R.E. Cunningham, T.A. Waldmann, and W.C. Greene. 1984. Regulation of interleukin 2 receptor expression: effects of phorbol diester, phospholipase C, and reexposure to lectin or antigen. J. Immunol. 133:3054-3061.

Dodd, J., S.B. Morton, D. Karagogeos, M. Yamamoto, and T.M. Jessell. 1988. Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. Neuron 1:105-116.

Dotti, C.G., R.G. Parton, and K. Simmons. 1991. Polarized sorting of glypiated proteins in hippocampal neurons. Nature 349:158-160.

Draberova, L. and P. Draber. 1993. Thy-1 glycoprotein and src-like protein tyrosine kinase p53/56 lyn are associated in large detergent-resistant complexes in rat basophilic leukemia cells. Proc. Natl. Acad. Sci. USA 90:3611-3615.

Faivre-Sarrailh, C., G. Gennarini, C. Goridis, and G. Rougon. 1992. F3/F11 cell surface molecule expression in the developing mouse cerebellum is polarized at synaptic sites and with granule cells. J. Neurosci. 12:257-267.

Filmus, J., J.G. Church, and R.N. Buick. 1988. Isolation of a cDNA corresponding to a developmentally regulated transcript in rat intestine. Mol. Cell Biol. 8:4243-4249.

Fu, Y.-M., P. Spirito, Z.-X. Yu, S. Biro, J. Sasse, J. Lei, V.J. Ferrans, S.E. Epstein, and W. Casscells. 1991. Acidic fibroblast growth factor in the developing rat embryo. J. Cell Biol. 114:1261-1273.

Gonzalez, A., M. Buscaglia, M. Ong, and A. Baird. 1990. Distribution of basic fibroblast growth factor in the 18-day rat fetus: localization in the basement membrane of diverse tissues. J. Cell Biol. 110:753-765.

Gundersen, R.W. 1987. Response of sensory neurites and growth cones to patterned substrata of laminin and fibronectin in vitro. Dev. Biol. 121:423-431.

Haugen, P.K., P.C. Letourneau, S.L. Drake, L.T. Furcht, and J.B. McCarthy. 1992. A cell surface heparan sulfate proteoglycan mediates neural cell adhesion and spreading on a defined sequence from the C-terminal cell and heparin binding domain of fibronectin, FN-C/H II. J. Neurosci. 12:2597-2608.

Herndon, M.E. and A.D. Lander. 1990. A diverse set of developmentally regulated proteoglycans is expressed in the rat central nervous system. Neuron 4:949-961.

Ho, W.H., M.P. Armanini, A. Nuigens, H.S. Phillips, and P.L. Osheroff. 1995. Sensory and motor neuron-derived factor. A novel heregulin variant highly expressed in sensory and motor neurons. J. Biol. Chem. 270:14523-14532.

Hunter, D.D., R. Llinas, M. Ard, J.P. Merlie, and J.R. Sanes. 1992. Expression of S-laminin and laminin in developing rat central nervous system. J. Comp. Neurol. 323:238-251.

Hutchins, J.B. and V.E. Jefferson. 1992. Developmental distribution of platelet-derived growth factor in the mouse central nervous system. Brain Res. Dev. Brain Res. 67:121-135.

Karthikeyan, L., M. Flad, M. Engel, B. Meyer-Puttlitz, R.U. Margolis, and R.K. Margolis. 1994. Immunocytochemical and in situ hybridization studies of the heparan sulfate proteoglycan, glypican, in nervous tissue. J. Cell Sci. 107:3213-3222.

Lisanti, M.P., P.E. Scherer, J. Vidugiriene, Z.L. Tang, A. Hermanowski-Vosatka, Y.-H. Tu, R.F. Cook, and M. Sargiacomo. 1994. Characterization of caveolin-rich membrane domains isolated from an endothelial-rich source: implications for human disease. J. Cell Biol. 126:111-126.

Litwack, E.D. (1995). Expression and function of proteoglycans in the nervous system. Doctoral Dissertation. Massachusetts Institute of Technology.

Litwack, E.D., C.S. Stipp, A. Kumbasar, and A.D. Lander. 1994. Neuronal expression of glypican, a cell-surface glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan, in the adult rat nervous system. J. Neurosci. 14:3713-3724.

McLoon, S.C., L.K. McLoon, S.L. Palm, and L.T. Furcht. 1988. Transient expression of laminin in the optic nerve of the developing rat. J. Neurosci. 8:1981-1990.

Merenmies, J. and H. Rauvala. 1990. Molecular cloning of the 18-kDa growth associated protein of developing brain. J. Biol. Chem. 265:16721-16724.

Nakato, H., T.A. Futch, and S.B. Selleck. 1995. The division abnormally delated (dally) gene: A putative integral membrane proteoglycan required for cell division patterning during post-embryonic development of the nervous system in Drosophila. Development 121:3687.

Neugebauer, K.M., C.J. Emmett, K.A. Venstrom, and L.F. Reichardt. 1991. Vitronectin and thrombospondin promote retinal neurite outgrowth: developmental regulation and role of integrins. Neuron 6:345-358.

O'Shea, K.S., J.S.T. Rheinheimer, and V.M. Dixit. 1990. Deposition and role of thrombospondin in the histogenesis of the cerebellar cortex. J. Cell Biol. 110:1275-1284.

Ornitz, D.M., A. Yayon, J.G. Flanagan, C.M. Svahn, E. Levi, and P. Leder. 1992. Heparin is required for cell-free binding of basic fibroblast

growth factor to a soluble receptor and for mitogenesis in whole cell. Mol. Cell. Biol. 12:240-247.

Presky, D.H., M.G. Low, and E.M. Shevack. 1990. Role of phosphatidylinositol-anchored proteins in T cell activation. J. Immunol. 144:860-868.

Ransom, J.T. and J.C. Cambier. 1986. B cell activation. VII. Independent and synergistic effects of mobilized calcium and diacylglycerol on membrane potential and I-A expression. J. Immunol. 136:66-72.

Rapraeger, A., A. Krufka, and B.B. Olwin. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science 252:1705-1708.

Reddy, U.R. and D. Pleasure. 1992. Expression of platelet-derived growth factor (PDGF) and PDGF receptor genes in the developing rat brain. J. Neurosci. Res. 31:670-677.

Reichardt, L.F. and K.J. Tomaselli. 1991. Extracellular matrix molecules and their receptors: Functions in neural development. Annu. Rev. Neurosci. 14:531-570.

Rogers, S.L., K.J. Edson, P.C. Letourneau, and S.C. McLoon. 1986. Distribution of laminin in the developing peripheral nervous system of the chick. Dev. Biol. 113:429-435.

Rogers, S.L., P.C. Letourneau, and I.V. Pech. 1989. The role of fibronectin in neural development. Dev. Neurosci. 11:248-265.

Saltiel, A.R. and L.R. Sorbara-Cazan. 1987. Inositol glycan mimics the action of insulin on glucose utilization in rat adipocytes. Biochem. Biophys. Res. Commun. 149:1084-.

Sanes, J.R. 1989. Extracellular matrix molecules that influence neural development. Annu. Rev. Neurosci. 12:491-516.

Sargiacomo, M., M. Sudol, Z.L. Tang, and M.P. Lisanti. 1993. Signal transducing molecules and GPI-linked proteins form a caveolin-rich insoluble complex in MDCK cells.   .

Schofield, L. and F. Hackett. 1993. Signal transduction in host cells by a glycosylphophatidylinositol toxin of malaria parasites. J. Exp. Med. 177:145-.

Selleck, S.B., C. Gonzalez, D.M. Glover, and K. White. 1992. Regulation of the G1-S transition in post-embryonic neuronal precursors by axon ingrowth. Nature 355:253-255.

Sheppard, M.M., S.K. Hamilton, and A.L. Pearlman. 1991. Changes in the distribution of extracellular matrix components accompany early morphogenetic events of mammalian cortical development. J. Neurosci. 11:3928-3942.

Shworak, N.W., M. Shirakawa, R.C. Mulligan, and R.D. Rosenberg. 1994a. Characterization of ryudocan glycosaminoglycan acceptor sites. J. Biol. Chem. 269:21204-21214.

Stefanova, I.M.L., E.M. Corcoran, L. Horak, M. Wahl, J.B. Bolen, and I.D. Horak. 1993. Lipopolysaccharide induces activation of CD-14-associated protein tyosine kinase p53/56 lyn. J. Biol. Chem. 28:20725-20728.

Stefanova, I.V., I. Horejsi, J. Ansotegui, W. Knapp, and H. Stockinger. 1991. GPI-anchored cell surface molecules complexed to protein tyrosine kinases. Science 254:1016-1019.

Thomas, P.M. and L.E. Samelson. 1992. The glycosylphosphatidylinositol-anchored Thy-1 molecule interacts with the p60fyn protein tyrosine kinase in T cells. J. Biol. Chem. 267:12317-12322.

Watanabe, K., H. Yamada, and Y. Yamaguchi. 1995. K-glypican: a novel gpi-anchored heparan sulfate proteoglycan that is highly expressed in developing brain and kidney. J. Cell Biol. 130:1207-1218.

Yayon, A., M. Klagsbrun, J.D. Esko, P. Leder, and D.M. Ornitz. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. Cell 64:841-848.