Interactions between heparan sulfate and proteins: the concept of specificity

Johan Kreuger,1 Dorothe Spillmann,2 Jin-ping Li,2 and Ulf Lindahl2

Proteoglycan (PG) coreceptors carry heparan sulfate (HS) chains that mediate interactions with growth factors, morphogens, and receptors. Thus, PGs modulate fundamental processes such as cell survival, division, adhesion, migration, and differentiation. This review summarizes recent biochemical and genetic information that sheds new light on the nature of HS–protein binding. Unexpectedly, many interactions appear to depend more on the overall organization of HS domains than on their fine structure.

Introduction

Almost every type of animal cell has the capacity to synthesize heparan sulfate (HS). The HS polysaccharide is composed of alternating hexuronic acid and \( \text{N} \)-glucosamine units and is substituted with sulfate groups in various positions. The sulfated saccharide domains provide numerous docking sites for protein ligands and are abundantly expressed at cell surfaces and in the extracellular matrix as part of proteoglycans (PGs). Two main types of cell surface–bound PG core proteins have been identified: the glycosylphosphatidylinositol–linked glypicans and the transmembrane syndecans; both types are thought to be expressed in high copy numbers (up to \( 10^6 \) per cell; Bernfeld et al., 1992). The diverse ligands include growth factors/morphogens and their receptors, enzymes, enzyme inhibitors, cell adhesion molecules, chemokines, various extracellular matrix proteins, and microbial proteins (Bernfeld et al., 1999). Interactions with HS contribute to or modify the various protein functions, which are of particular interest in relation to growth factor/morphogen translocation and signaling (Fig. 1). Thus, HSPGs are essential for normal embryonic development but are also implicated in homeostasis as well as in pathological processes of growing and adult individuals (Hacker et al., 2005).

The HS backbone typically contains 50–400 monosaccharide units. The hexuronic acid moieties are of two kinds: \( \text{d} \)-glucuronic acid (GlcA) and \( \text{l} \)-iduronic acid (IdoA). The extreme structural diversity typical of HS species is a result of the variable distribution of these residues as well as of sulfate substituents along the chain. Because this diversity appears to be strictly regulated, it is currently believed to enable selective interaction with proteins in a topologically and temporally controlled manner (Esko and Lindahl, 2001). The aim of this present review is to reassess this concept in view of recent findings.

Regulated polymer modification in HS biosynthesis

Nascent HS chains evolve by the stepwise addition of alternating GlcA and \( \text{N} \)-acetylg glucosamine (GlcNAc) residues to acceptor linkage region oligosaccharides, which are substituted on PG core proteins. These chains are subsequently modified through several consecutive steps, including \( \text{N} \)-deacetylation and \( \text{N} \)-sulfation of GlcNAc residues, C5 epimerization of GlcA to IdoA, and \( \text{O} \)-sulfation in various positions (Esko and Lindahl, 2001). The process, which is outlined in Fig. 2, is generally incomplete, such that the final products display a domain-type arrangement of more or less modified saccharide sequences. Because of the substrate specificities of the enzymes involved (at least in part), IdoA and \( \text{O} \)-sulfate residues are accumulated in domains of contiguous \( \text{N} \)-sulfated (NS) disaccharide units (NS domains), are less abundant in regions of mixed \( \text{N} \)-substitution, and are essentially lacking in contiguous \( \text{N} \)-acyetylated sequences. Heparin, which is produced by connective tissue–type mast cells, is the result of extensive biosynthetic polymer modification and may be conceived as an extended, highly \( \text{N} \)- and \( \text{O} \)-sulfated IdoA-rich NS domain.

Analysis of HS from different mammalian tissues revealed the tissue-specific composition of samples, pointing to strict regulation of biosynthetic polymer modification (Maccarana et al., 1996; Ledin et al., 2004). Anion exchange chromatograms of products obtained by chemical or enzymatic degradation of mouse HS thus differed for samples from different tissues but were virtually superimposable for corresponding samples from different individuals. Moreover, immunohistochemical analysis showed selective, reproducible distribution of distinct HS epitopes within a given tissue (van Kuppevelt et al., 1998). These findings could reflect differences in saccharide domain composition, sequence, and/or overall organization, which are potentially required to selectively bind different protein ligands.

Abbreviations used in this paper: EXT, GlcA/GlcNAc transferase; GlcA, glucuronic acid; GlcNAc, N-acetylg glucosamine; HS, heparan sulfate; IdoA, iduronic acid; NS, N-sulfated; PG, proteoglycan; SAS, N-sulfated/acyetylated/sulfated.

Correspondence to Ulf Lindahl: Ulf.Lindahl@imbim.uu.se

Published July 31, 2006

© The Rockefeller University Press $8.00

The Journal of Cell Biology, Vol. 174, No. 3, July 31, 2006 323–327
http://www.jcb.org/cgi/doi/10.1083/jcb.200604035
What do we know regarding specificity in HS–protein interactions? Early evidence for specificity was the antithrombin-binding sequence, which is responsible for the clinically exploited blood anticoagulant activity of heparin (Lindahl et al., 1980). Each one of the sulfate substituents of this GlcNAc6S-GlcA-GlcNS3S-IdoA-GlcNS pentasaccharide structure is essential for high affinity interaction with antithrombin and, thus, for anticoagulant activity. Notably, this sequence contains the rare glucosamine 3-O-sulfate group in addition to the more common 2-O-sulfate of IdoA and N- and 6-O-sulfate groups of GlcN residues. Another unusual structure based on a 3-O-sulfated N-unsubstituted GlcNS unit mediates the apparently specific binding of herpes simplex gD protein to the cell surface HS during viral infection (Shukla et al., 1999). What about the multitude of other heparin-binding proteins that depend on the interaction with endogenous HS ligands for biological activity? We previously proposed that the common sulfate substituents could also be arranged in a sequence-specific manner to provide selective protein binding (Salmivirta et al., 1996). Such epitopes would be masked in the highly sulfated heparin molecule, thus explaining the apparently nonselective protein binding to this polysaccharide. However, although recent binding studies with selected proteins have shown that a particular kind of sulfate group (e.g., 6-O-sulfates) might contribute more to interaction than others (Ashikari-Hada et al., 2004), there is yet no clear evidence of distinct sequence specificity based on the distribution of common sulfate residues (Powell et al., 2004). Various signaling proteins may be considered in this context.

Different members of the FGF family share binding sites on the HS chain, and their affinities for HS-related oligosaccharides generally correlate with the overall degree of saccharide sulfation (Jemth et al., 2002; Kreuger et al., 2005; however, see Ashikari-Hada et al., 2004 regarding preferential binding of variously O-sulfated oligosaccharides). Our recent experiments suggest that a relatively nonspecific charge interaction may also prevail in the formation of FGF–HS–FGF receptor complexes (Fig. 2). Using a variety of oligo- and polysaccharide probes, we found that the complex formation of FGF1 or FGF2 with their various receptors was increasingly promoted by saccharide sequences of increasing overall sulfate content in an apparently nonspecific fashion. Heparin oligosaccharides were generally the most efficient complex promoters, whereas less sulfated HS species were less efficient (Jastrebova et al., 2006). These findings suggest that the dependence of FGF signaling on HS fine structure may be less critical than previously anticipated. However, we note studies claiming that NS oligosaccharide fractions derived from authentic HS contained receptor-activating as well as nonactivating species (Pye et al., 1998; Guimond and Turnbull, 1999). Of course, many ligand–receptor combinations remain to be examined, and we cannot exclude elements of selectivity here that are currently unrecognized.

The role of VEGF as a key regulator of vascular development is well documented (Carmeliet et al., 1996). The interaction of the long splice variant VEGF-A165 with HS is essential for proper signaling, as recently demonstrated in variously designed cell culture systems (Ashikari-Hada et al., 2005; Jakobsson et al., 2006). Studies aimed at defining the structural features of HS that are required for VEGF-A165 binding implicated all common sulfate groups (N, 2-O, and 6-O), although with different emphasis on their relative importance (Ashikari-Hada et al., 2005; Robinson et al., 2006). Moreover, the pleiotropic hepatocyte growth factor binds a variety of glycosaminoglycan structures without any clear preference (Catlow et al., 2003).

**Knockout clues from embryology**

Mice that were genetically deficient in enzymes involved in HS biosynthesis provided novel insight into the question of specificity in HS–protein interactions. Phenotype analysis revealed developmental events that require the involvement of HS, and structural analysis of the corresponding polysaccharides could, in some cases, pinpoint molecular features that are of critical importance to such events. Equally important, however, is to identify HS-dependent events that remain unperturbed in spite of deranged HS structure.

Embryos lacking GlcA/GlcNac transferase 1 (EXT1), which generates the initial (GlcA-GlcNac)₉ polysaccharide chain (Fig. 2), failed to undergo proper gastrulation in accord with the recognized need for HS in early patterning events (Lin et al., 2000). In contrast, the targeted disruption of genes encoding selected enzymes involved in the later stages of HS biosynthesis resulted in strikingly varied phenotypes. These mice displayed variously perturbed HS structures and a variety of developmental abnormalities but also displayed features assumed...
to be HS dependent that were surprisingly normal. For example, brain-selective, conditional EXT1 knockout, which was aimed at identifying HS-dependent processes in cerebral development, resulted in multiple brain defects, as predicted from the established involvement of various HS-dependent growth factors (in particular FGF8; Inatani et al., 2003). Yet, mice lacking the C5 epimerase catalyzing the conversion of GlcA to IdoA residues (Li et al., 2003) or the 2-O-sulfotransferase required for IdoA 2-O sulfation (Bullock et al., 1998; Merry et al., 2001) showed no obvious brain phenotype irrespective of the severely deranged HS structures. Moreover, assessment of the cardiovascular system pointed to VEGF signaling that is compatible with adequate vasculogenesis and angiogenesis. The elevated N- and 6-O-sulfation, which is characteristic of these mutated HS species, apparently sufficed to satisfy the requirement for HS in important VEGF and FGF signaling events (Fig. 2). Also, other organ systems, such as the gastrointestinal tract, which is known to require HS for growth factor/morphogen signaling in development, appeared normal in the C5 epimerase and 2-O-sulfotransferase knockout mice.

Certain other events failed, however, as indicated by the kidney agenesis, the skeletal malformations, and other problems leading to early postnatal death of the animals. Interestingly, the kidney agenesis was not observed in N-deacetylase/N-sulfotransferase-1-null mice, which were deficient in the early N-deacetylation/N-sulfuration step (Ringvall et al., 2000); the overall poorly modified but IdoA-containing HS that resulted apparently functioned at a critical stage of kidney induction. Moreover, the lung phenotype caused by C5 epimerase deficiency (HS lacking IdoA and IdoA2S residues) was not seen in mice lacking 2-O-sulfotransferase (HS containing IdoA but no IdoA2S units; Fig. 2). Redundant O-sulfation was also observed in Drosophila melanogaster 2-O- or 6-O-sulfotransferase-null mutants that were able to complete development with apparently normal FGF signaling and morphology (Kaminura, K., and H. Nakato, personal communication). Lack of both enzymes, on the other hand, led to impaired FGF signaling and multiple patterning deficiencies. We conclude that several functionally important HS–protein interactions depend primarily on charge distribution, whereas others may require the presence of specific saccharide components. In this perspective, what is the functional purpose of the strict regulation of polymer modification in HS biosynthesis?

Regulated domain organization?

The rationale for stringent HS biosynthesis may relate primarily to the domain organization of HS chains. Various designed interaction studies implicate saccharide sequences of up to 12-mer size (in some cases even longer) for efficient interaction with proteins (Schlessinger et al., 2000; Gallagher, 2001). Contiguous NS domains of >8-mer size are generally rare in HS. Instead, composite binding sites involving short NS domains separated by N-acetylated disaccharide units (N-sulfated/acyetylated/sulfated [SAS] domains) can mediate interactions with monomeric (e.g., endostatin; Kreuger et al., 2002) as well as oligomeric (e.g., interferon-γ [Lortat-Jacob et al., 1995], interleukin-8 [Spillmann et al., 1998], and platelet factor 4 [Stringer and Gallagher, 1997]) protein ligands. Also, VEGF-A_{165} occurs as a dimer that interacts with SAS domains in HS chains (Fig. 2; Robinson et al., 2006). Given the size of saccharide domains that are implicated in complex formation with growth factors and their receptors, we predict that SAS-type structures may be involved or even required in various signaling complexes. In fact, the differential complex formation of endogenous HS in mouse embryos with defined/given FGF–FGF receptor pairs (Allen and Rapraeger, 2003) may well reflect selective domain spacing rather than precisely tailored saccharide sequences.

Figure 2. Biosynthesis of HS and molecular phenotypes resulting from deficient HS biosynthetic enzymes. See Esko and Lindahl (2001) and Hacker et al. (2005). HS chains are synthesized while attached to core protein serine residues through a GlcA-Gal-Gal-Xyl linkage region. The linear HS chain is thereafter polymerized through the action of GlcNAc- and GlcA-transferases belonging to the EXT family and further modified by partial N-deacetylation/N-sulfation (N-deacetylase/N-sulfotransferase) to yield NS disaccharide units. Consecutive stretches of such units (NS domains) are hotspots for further modifications: a C5 epimerase converts GlcA to IdoA followed by variable O-sulfation at C-3 and C-6 (red circles) of GlcN and at C-2 (yellow circles) of IdoA (and some GlcA) units. Completed chains may be further edited by endo-6-O-sulfatases (Ai et al., 2003). Protein ligands interact with single NS domains (e.g., FGFs) or with NS domains separated by N-acetylated disaccharide residues (SAS domains; illustrated here for VEGF-A_{165} [Robinson et al., 2006]; and FGF–HS–FGF receptor complexes). The bottom model depicts a molecular phenotype of C5 epimerase−/− HS that lacks IdoA and IdoA 2-O-sulfation but is more extensively N- and 6-O-sulfated than the corresponding wild-type product. The 2-O-sulfotransferase−/− HS is similar to the C5 epimerase−/− polysaccharide except for the presence of IdoA (Merry et al., 2001). HS from C5 epimerase−/− or 2-O-sulfotransferase−/− cells may still interact more or less efficiently with many protein ligands (see Knockout clues from embryology). Blue boxes, NS domains containing GlcA and/or IdoA; yellow boxes, NS domains containing GlcA but no IdoA.
We propose that polymer modification in HS biosynthesis is primarily regulated with regard to domain distribution and degree of sulfation (i.e., the distribution of N-substituents and the levels of 2-O- and 6-O-sulfation). Such regulation would presumably suffice to explain the observed consistent differences in composition between HS species from different cellular or tissue sources (Maccarana et al., 1996; Ledin et al., 2004). The resulting clusters of negative charge will determine interactions with proteins that may be relatively nonselective with sharing/overlap of saccharide target sequences between different protein ligands. Given the excessive number of possible saccharide epitopes, specific sequences based on common constituents that provide somewhat stronger binding of a particular protein ligand than other sequences may well be preferentially formed in the course of such regulated polymer modification. More selective interactions would require either sequences containing rare components or precise spacing of two (or more) sulfated domains (SAS arrangement). Notably, there is no method readily available to reveal the detailed distribution of various domains along a native HS chain, although some progress has been reported based on selective lyase degradation of the polymer (Murphy et al., 2004). In addition, we still do not understand the regulatory mechanisms in HS biosynthesis that determine domain generation or localization.

We wish to acknowledge funding by the Swedish Research Council, the Swedish Cancer Foundation, the Swedish Foundation for Strategic Research, and the Wenner-Gren Foundations.

Submitted: 10 April 2006
Accepted: 22 June 2006

References

Ai, X., A.T. Do, O. Lozynska, M. Kusche-Gullberg, U. Lindahl, and C.P. Emerson Jr. 2003. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J. Cell Biol. 162:341–351.

Allen, B.L., and A.C. Rapraeger. 2003. Spatial and temporal expression of Ai, X., A.T. Do, O. Lozynska, M. Kusche-Gullberg, U. Lindahl, and C.P. Emerson Jr. 2003. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J. Cell Biol. 162:341–351.

Allen, B.L., and A.C. Rapraeger. 2003. Spatial and temporal expression of Ai, X., A.T. Do, O. Lozynska, M. Kusche-Gullberg, U. Lindahl, and C.P. Emerson Jr. 2003. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J. Cell Biol. 162:341–351.

Allen, B.L., and A.C. Rapraeger. 2003. Spatial and temporal expression of Ai, X., A.T. Do, O. Lozynska, M. Kusche-Gullberg, U. Lindahl, and C.P. Emerson Jr. 2003. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J. Cell Biol. 162:341–351.

Allen, B.L., and A.C. Rapraeger. 2003. Spatial and temporal expression of Ai, X., A.T. Do, O. Lozynska, M. Kusche-Gullberg, U. Lindahl, and C.P. Emerson Jr. 2003. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J. Cell Biol. 162:341–351.
ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell.* 6:743–750.

Shukla, D., J. Liu, P. Blaiklock, N.W. Shworak, X. Bai, J.D. Esko, G.H. Cohen, R.J. Eisenberg, R.D. Rosenberg, and P.G. Spear. 1999. A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell.* 99:13–22.

Spillmann, D., D. Witt, and U. Lindahl. 1998. Defining the interleukin-8-binding domain of heparan sulfate. *J. Biol. Chem.* 273:15487–15493.

Stringer, S.E., and J.T. Gallagher. 1997. Specific binding of the chemokine platelet factor 4 to heparan sulfate. *J. Biol. Chem.* 272:20508–20514.

van Kuppevelt, T.H., M.A. Dennissen, W.J. van Venrooij, R.M. Hoet, and J.H. Veerkamp. 1998. Generation and application of type-specific anti-heparan sulfate antibodies using phage display technology. Further evidence for heparan sulfate heterogeneity in the kidney. *J. Biol. Chem.* 273:12960–12966.

Vlodavsky, I., and Y. Friedmann. 2001. Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J. Clin. Invest.* 108:341–347.