Heparan sulfate-protein interactions – A concept for drug design?

Ulf Lindahl
Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

Summary
The glycosaminoglycan, heparan sulfate (HS) is composed of alternating units of hexuronic acid and glucosamine, that are variously sulfate-substituted at different positions. Proteoglycans carrying HS chains are ubiquitously expressed at cell surfaces and in the extracellular matrix. The structures of these chains are highly variable, yet under strict biosynthetic control. Due to their high negative charge, HS chains interact with a multitude of proteins, including growth factors/morphogens and their receptors, chemokines, and extracellular-matrix proteins. These interactions regulate key events in embryonic development and in homeostasis. HS-protein interactions vary with regard to specificity, and often seem to depend primarily on charge density rather than on strict carbohydrate sequence. The organization of sulfated domains along the HS chain appears to be of importance. HS-protein interactions are involved in a variety of pathophysiological processes, including inflammation, angiogenesis, and amyloid deposition. Drugs targeting such interactions may be useful in treatment of disease conditions as diverse as cancer, inflammatory bowel disease, and Alzheimer’s disease. Potential drugs may mimic HS oligosaccharides, but could also be peptides blocking the protein-binding domains of HS chains. Drug generation requires a firm understanding of the pathophysiological role of a given HS-protein interaction, and of the aspect of specificity. Even inhibition of HS biosynthesis may be considered.

Keywords
Sulfation, oligosaccharide, glycomimetic

Introduction
Heparin has been used as a drug for more than 60 years, and remains the drug of choice in prophylaxis against postoperative thromboembolic complications. Current annual world production, largely based on pig intestinal mucosa as starting material, exceeds 20 tons. The mechanism of action of heparin, poorly understood at the time of clinical introduction, was unraveled through decades of work in several laboratories, including our own, and is now explained in reasonable detail (1). However, it was also gradually realized that heparin is merely one member of the vast heparan sulfate (HS) family of glycosaminoglycans. While heparin is confined to connective-tissue type mast cells, HSS are produced by most cells throughout the animal kingdom down to comparatively simple organisms. HS polysaccharides interact with a multitude of proteins, thereby affecting their biological functions, and thus profoundly influence important processes in development, homeostasis and disease (2, 3). Several otherwise unrelated disease conditions have been shown to involve HS-protein interactions that are increasingly being elucidated at the molecular level. Given the heparin-antithrombin precedence case, it would seem logical to probe also HS-protein interactions for clues toward generation of HS-based drugs. I will attempt to discuss some requisites for such projects, without any ambition to comprehensively cover all aspects of this complex subject area.

Heparan sulfate – basic features and functions
HS polysaccharides are composed of alternating units of hexuronic acid [D-glucuronic acid (GlcA) or L-iduronic acid (IdoA)] and D-glucosamine in linear sequence. The glucosamine residues are N-sulfated (GlcNS), N-acetylated (GlcNAc) or, rarely, N-unsubstituted and may in addition carry O-sulfate groups at C6 or C3. The hexuronic acid moieties are either unsubstituted or O-sulfated at C2. The biosynthesis of these polymers is initiated by formation of a [GlcA-GlcNAc]n precursor, which then undergoes sequential N-deacetylation and N-sul-
fation, GlcA C5-epimerization to yield IdoA residues, 2-O-, 6-O-, and 3-O-sulfation (Fig. 1). The enzymes involved in polymer formation and modification have all been cloned and expressed in recombinant form (see reviews [4, 5]). Due to constraints imposed by substrate specificity (and as yet unknown factors) the various modifications do not occur randomly along the polysaccharide chain, but show typical domain distribution. Consecutive N-sulfated disaccharide units (NS-domains), usually ≤5-6, rich in IdoA and O-sulfate groups, provide clusters of negative charge that are separated by regions that remain N-acetylated and thus essentially lack IdoA and sulfate residues (NA-domains). Yet other regions are composed of alternating N-acetylated and N-sulfated disaccharide units with IdoA and 6-O-sulfate but no 2-O-sulfate groups (NA/NS-domains) (6, 7). Murine HS preparations showed tissue-specific differences in composition that were conserved between individuals (8). Moreover, immunohistochemical analysis revealed highly selective display of different HS epitopes within and between tissues (9, 10). Together, these observations point to highly regulated HS biosynthesis. To further complicate the issue, HS chains may be "edited" after completed biosynthesis by endo-6-O-sulfatasess that have been functionally implicated in various signaling systems (see e.g. [11]). They may also be subjected to endolytic cleavage by heparanase, an endo-β-D-glucuronidase capable of generating HS fragments for release either extracellularly or in the course of lysosomal HS degradation (12).

Figure 1: Biosynthesis of HS and models of molecular phenotypes resulting from deficient biosynthetic enzymes. A) HS chains grow by alternate action of GlcNAc- and GlcA-transferases, while attached to core protein serine residues through a GlcA-Gal-Gal-Xyl-linkage region. The linear polymer is modified by partial N-deacetylation/N-sulfation (catalyzed by NDST enzymes) to yield N-sulfated disaccharide units. Consecutive sequences of such units (NS domains, blue boxes) are preferred targets for further modifications: a C5 epimerase converts GlcA to IdoA, followed by variable O-sulfation at C2 (yellow circles) of IdoA (and some GlcA) and at C6 (red circles) and (rarely) C3 of GlcN residues. Completed chains may be further modified by endo-6-O-sulfatases (11). 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 PDGF-BB, (32) and FGF-HS-FGF-receptor complexes). B) These models depict the molecular phenotypes of NDST1-/− HS that contains all constituents of wild-type HS but is overall poorly modified, and of C5-epimerase−/− HS that lacks IdoA and IdoA 2-O-sulfation but is extensively N- and 6-O-sulfated. Analysis of the corresponding animal phenotypes (described in the text) suggests that also severely perturbed HS structures may engage in functional interactions with selected protein ligands.
Importantly, HS is synthesized and expressed in proteoglycan form, with polysaccharide chains covalently bound to various distinct core proteins (13). HS proteoglycans (HSPGs) occur both at cell surfaces (syndecans, glypicans) and in the extracellular matrix (perlecan, agrin, collagen XVIII). Contrary to the structural differences observed between HS species derived from different tissues, the HS constituents of the various proteoglycans produced by a given cell appear similar. Heparin substitutes the intracellular proteoglycan, serglycin, and can be conceived as a HS chain essentially consisting of unusually extended NS-domains.

The negatively charged regions, in particular NS-domains, provide interaction sites for a variety of proteins (Fig. 1), including growth factors/morphogens and their receptors (Fig. 2A), chemokines, enzymes/enzyme inhibitors, and various extracellular-matrix proteins (4, 5). HS thus has essential functions in development, serving as co-receptor in protein-mediated cell signaling, and as stabilizer of morphogen gradients along epithelial surfaces. HS regulates diverse processes essential to homeostasis, such as transport of large and small molecules across plasma membranes into cells, or across basement membranes, cell migration in inflammation, food intake. In fact, HS has been ascribed essential roles in most physiological systems (3). Notably, several pathogens have been shown to use cell-surface HS proteoglycans as primary "docking sites" during host invasion.

Heparan sulfate-protein interactions – Aspects of specificity

The first protein-binding site in a "HS" chain to be characterized in detail was the antithrombin-binding pentasaccharide sequence in heparin (Fig. 3). This structure consists of three GlcN residues, two of which need to be N-sulfated, one GlcA and one IdoA unit. Further, two O-sulfate groups are required for productive antithrombin binding, a 6-O-sulfate group and a 3-O-sulfate residue. The latter component is a rare constituent that was initially believed to be unique to the antithrombin-binding sequence (14). More recent research has shown that 3-O-sulfate groups may be selectively expressed in HS, and has identified a family of 3-O-sulfotransferases all capable of catalyzing specific incorporation of this particular residue (15). Additional "rare" HS components have been identified, such as 2-O-sulfated GlcA (16–18) and N-unsulfated GlcN (10, 19, 20). In fact, the latter residue in combination with a 3-O-sulfate group was implicated in apparently specific binding of herpes simplex virus gD protein to cell surface HS during viral infection (21). These findings, along with the demonstrated strict regulation of HS biosynthesis, suggested that HS-protein interaction in general is mediated by specifically tailored saccharide domains with restricted binding specificity. The majority of HS-binding proteins were initially detected through their ability to interact with the highly sulfated heparin chain. Indeed, we speculated that "specific" HS-sequences required for binding distinct proteins would be expressed also in heparin, although masked by redundant sulfate groups, thus explaining the apparently nonselective protein binding to this polysaccharide (22). Notably, however, whereas binding studies with selected proteins could highlight a particular kind of sulfated group (e.g. 6-O-sulfates) as being more important to interaction than others (23), there is yet no clear evidence of distinct sequence specificity based on the distribution of "common" sulfate residues (24).

Recently, we designed experiments to assess in more direct manner the aspect of specificity in HS-protein interactions. Application of libraries of HS-related oligosaccharides, generated by chemo-enzymatic methods to probe interactions with various fibroblast growth factors (FGFs) suggested that different members of the FGF family share binding sites on HS chains (25, 26). Interaction affinities generally correlated with the overall degree of saccharide sulfation. Moreover, relatively non-specific charge interaction appeared to prevail also in the formation of FGF–HS–FGF-receptor complexes, i.e. the signaling unit at the cell surface (27). Complex formation of FGF1 or FGF2 with

![Figure 2: Potential strategies in drug development based on HS-protein interactions.](https://example.com figure2.png)

Most of the examples shown relate to HS (red)-dependent binding of a protein ligand (dark blue, e.g. a growth factor) to its cell-surface receptor (light blue); however, similar principles would apply to a variety of interaction systems. A) Binding of protein ligand to receptor; assisted by endogenous HS. B) Activation of receptor by "HS drug" (oligosaccharide, glycomimetic) that forms a ternary complex with protein ligand and receptor, and displaces endogenous HS. Direct binding of such drug to target protein may promote or inhibit bioactivity. C) Inhibition of receptor signaling by drug displacing protein ligand from receptor. D) Inhibition of receptor signaling by peptide drug blocking protein binding site on HS. This model may also represent the mode of action of peptide drugs preventing HS-dependent aggregation of amyloid proteins. E) Drugs (not indicated) interfering with HS biosynthesis. For further information see the text.

![Figure 3: Structure of the antithrombin-binding region in heparin.](https://example.com figure3.png)

The pentasaccharide sequence is composed of three GlcN (blue) units, one GlcA (green) and one IdoA (magenta) residue. R’’, -H or -SO$_3$; R’’, -SO$_2$ or -COCH$_3$. The four marked (yellow rectangle) sulfate groups are all essential for high-affinity interaction with antithrombin; the 3-O-sulfate group on the internal GlcN unit 3 is a marker for the antithrombin-binding sequence, and is rare or absent elsewhere in the heparin molecule (ref. [11]).
their various receptors thus was increasingly promoted by sac- 
ccharide sequences of increasing overall sulfate content. Heparin 
oligosaccharides were generally the most efficient complex pro-

tomers, whereas less sulfated HS analogs were less efficient. We 
concluded that the dependence of FGF signaling on HS fine 
structure is less critical than previously anticipated. Would this 
conclusion apply also to other protein-HS interactions – and to 
conditions in vivo?

Mice genetically deficient in enzymes involved in HS bio-
synthesis provided novel insight into this question. Embryos 
either lacking HS chains (28), or being essentially unable to 
modify the initial \([\text{GlcA-GlcNAC}]_n\) polysaccharide formed (Fig. 
1) (29), failed to undergo proper gastrulation in accord with 
the recognized need for HS in early patterning events. By contrast, 
elimination of enzymes involved in the later stages of HS biosyn-
thesis resulted in strikingly varied phenotypes. Mice deficient in 
the C5-epimerase catalyzing the conversion of GlcA to IdoA 
residues thus generated severely perturbed HS, with no IdoA and 

essentially no 2-O-sulfate groups but elevated N- and 6-O-sul-
fation (Fig. 1) (30). We were surprised, in the first place, to find 
this defect to be compatible with a completed pregnancy. The 
cubs displayed a variety of developmental abnormalities, but 
also features assumed to be HS dependent that appeared normal. 
Developmental failures included skeletal malformations, kidney 
agenesis and other problems leading to early postnatal death of 
the animals. On the other hand, the gross anatomical features of 
the brain seemed normal (30), in spite of findings by others (31) 
implicating HS with signaling mechanisms (in particular 
FGF8-dependent) essential for brain development. Similarly, 
the HS-dependent action of platelet-derived growth factor-BB 

in vascular development appeared essentially unperturbed in 
C5-epimerase deficient mice (32). By contrast, both brain and 

vasculature development were affected by loss of N-deacetylase/ 
N-sulfotransferase-1, a key enzyme in the initial modification of 
the HS precursor polymer (Fig. 1). The compromised HS dis-
played short and sparsely distributed sulfated domains, that 
nevertheless contained all components typical of the wild-type 
polysaccharide (33). Interestingly, the N-deacetylase/N-sulfo-

transferase-1-deficient mice were capable of kidney develop-
ment, suggesting that the IdoA residues in HS are essential at 
a critical stage of kidney induction. These observations, along 
with studies targeting other enzymes (34, 35) collectively implic-
ate HS in an array of processes critical to normal embryonic de-
velopment, but also suggest that many, maybe most, of the requi-
site HS-protein interactions depend primarily on charge distribu-
tion, maybe on the presence of specific saccharide components, 
but not on precisely defined sequence of variously substituted 
sugar residues (36). This notion raises intriguing questions re-
garding the functional purpose of regulated polymer modifica-
tion in HS biosynthesis.

We propose that regulation of HS biosynthesis relates pri-
marily to the domain organization of the polysaccharide chains. 
Studies of interactions between HS and selected proteins suggest 
that the protein-binding saccharide domains can be of variable 
size, ranging from 4–5-mer to >12-mer sequences (36). The 

binding sites may be composed of one, two or more NS-domains 
interspersed by NA-domains, designed for interaction with 
single HS-binding regions on a protein ligand, with two or 
multiple HS-binding regions on a single protein, or with oligo-
meric proteins (e.g. in a FGF signaling complex). Extended, con-
tiguous NS-domains are generally rare in HS (but not in hepa-
rin), and may be substituted by composite binding sites either be-
cause of the relative abundance of such sites, or because they 
provide more favorable interactions with a given protein ligand 
or ligand complex. HS biosynthesis is primarily regulated with 
regard to size, spacing and overall degree of sulfation of the vari-
ous domains, but not with regard to precise sequence.

Heparan sulfate and disease

Mammalian embryos completely devoid of HS die early in de-
velopment (28). Localized lack of HS is seen in multiple hered-
itary exostoses, characterized by formation of benign bone tu-
mors (37). Further, patients with various protein-losing enter-
opathies show diminished amounts of HSPG (syndecan-1) in 
their intestinal epithelial cells (38). Proteinuria in diabetic nep-
hropathy is associated with loss of HS from the glomerular base-
ment membrane (39). Intriguingly, no human disease has yet 
been primarily ascribed to effects of perturbed HS structure.

Inflammatory responses associated with tissue injury are 
seen in a variety of inherently different diseases, such as rheuma-
toid arthritis, inflammatory bowel disease and microbial infec-
tions. HSPGs have important roles in these processes, as ad-
hesion ligand in selectin-mediated leukocyte extravasation, and 
carrier/presenter of chemokines and growth factors (40, 41). 
Moreover, HSPGs contribute to cancer development, by promo-
ting growth-factor-dependent signalling that increases tumor 
growth and associated angiogenesis, but potentially also through 
additional mechanisms that will not be discussed in detail here 
(42). The complexity of these interactions are illustrated by the 
finding of strong correlation between metastatic potential and 
expression of heparanase that acts by limited endolysosomal cleavage 
of HS chains (43). This correlation may reflect functions of he-
paranase to promote angiogenesis, to degrade extracellular ma-
trices or basement membranes required for mobilization of 
tumor cells, or to release HS oligosaccharides carrying growth 
factors that stimulate proliferation of tumor (or stromal) cells.

Amyloid diseases are characterized by deposition in tissues of 
fibrillar aggregates of polypeptides that share certain structural 
and biophysical properties, but are otherwise unrelated. This 
heterogeneous group includes AA-amyloidosis, Alz-
heimer’s disease, type-2 diabetes, Parkinson’s disease, prion dis-

cases and yet other conditions. Remarkably, virtually all cor-
responding amyloidogenic peptides bind HS \textit{in vitro}, and are 
codeposited with HS in tissue lesions \textit{in vivo} (44). HS (or hepa-
rin) appears capable of promoting amyloid fibrillogenesis \textit{in vitro}, in support of the notion that interaction of amyloidogenic 
peptides with HS is important to disease progression. Con-
versely, transgenic overexpression of heparanase rendered mice 
resistant to experimental AA amyloidosis, presumably due to se-
estration of amyloid peptide by released HS oligosaccharides 
(45). 

Finally, many pathogenic microorganisms express on their 
surface proteins capable of binding to HS, and these interactions 
appear important for infectivity, at least \textit{in vitro} (3).
Drugs interfering with heparan sulfate pathophysiology?
The only heparin/HS-related bioactivity so far employed in routine clinical application is the anticoagulant/antithrombotic activity. However, scattered observations point to effects of heparin therapy unrelated to anticoagulant activity. Notably, the plasma concentrations achieved after subcutaneous injection of (unfractionated) heparin at conventional dose levels generally exceed those required to release proteins such as selectins from endogenous carbohydrate ligands (46). Such disruption may contribute to unexpected beneficial effects of heparin on various pathological conditions, e.g. transient regression of metastatic tumors or inflammatory conditions. Also HS-protein interactions may be similarly disrupted. HS-based pathophysiology thus offers possibilities for drug intervention that are still poorly exploited. The following section is an attempt to survey potential strategies. Drugs may potentially be manufactured by chemical synthesis, isolation from naturally occurring polysaccharides (with or without chemical modification), but also using recombinant enzymes (GlcA C5-epimerase; sulfotransferases) with appropriate saccharide substrates (47). Problems related to pharmacokinetics, access of drugs to various compartments etc. will not be considered here.

Activation/inactivation of target proteins
Interaction of a saccharide or glycomimetic with a target protein may directly modulate an inherent bioactivity (Fig. 2B), as shown with heparin and antithrombin (Fig. 3). The saccharide drug binds to antithrombin and potentiates its ability to inhibit the serine proteases involved in blood coagulation. This interaction provides one of the few examples of heparin/HS-based bioactivity that strictly depends on specific carbohydrate sequence (1). Notably, such sequences can be reproduced through chemical synthesis (48, 49). Similar potentiation of other activities of clinical interest would seem within reach, for instance in growth-factor-dependent wound healing. Such drugs should presumably substitute for endogenous HS and form stable, long-lasting ternary complexes with growth factors (e.g. FGFs) and their tyrosine-kinase type cell-surface receptors (Fig. 2B).

Heparin is known to directly inhibit several enzymes, including heparanase, the endo-β-D-glucuronidase that selectively cleaves HS chains and correlates with the metastatic and angiogenic potential of tumor cells (43). Several heparanase inhibitors, based on chemical modification of naturally occurring saccharides have been described recently (50–52).

Competition with endogenous ligands
HS-protein interactions have been implicated with a multitude of pathological conditions, as briefly outlined above. Complexes may be disrupted by addition of competitive saccharide ligands that substitute for the HS moiety (Fig. 2C). Alternatively, peptide competitors may be applied, that block protein-binding domains of HS chains (Fig. 2D). In fact, both strategies offer prospects for drug development.

Identification of structural features conducive to interaction of HS oligosaccharides with a given protein ligand might provide cues to design of drugs with selective action (Fig. 2C). As noted above, however, many functionally important HS-protein interactions appear to be relatively non-specific, in the sense that they depend primarily on overall charge density rather than on precise sequence of variously substituted sugar residues (36). In such case, a drug should encompass, or reflect an oligosaccharide structure of appropriate size with a high content of N- and O-sulfate groups. Interaction studies with selected protein targets (chemokine, growth factor etc.) may highlight certain sub- stituent types (e.g. 6-O-sulfate groups) as being more important for binding than others. Moreover, certain (usually oligomeric, but also monomeric) proteins may preferentially interact with multiple domains (“SAS-domains”) along a HS chain (32, 53–55). Such information may provide selectivity in drug design, for exploitation in either oligosaccharide-type products or glycomimetics.

A recently developed anti-malaria drug, designed to block interaction between the plasmoidium-induced protein PfEMP1 and HS on erythrocyte or vascular endothelial surfaces, contains heparin oligosaccharide lacking the antithrombin-binding site (56). The active oligomers are of sufficient size (12-mers) to efficiently bind the malaria protein. Glycomimetics may be applied in principally similar fashion. Synthetic, low-molecular-weight (130–1,000) anionic sulfonate compounds, administered orally, thus substantially reduced murine splenic AA amyloid progression (57). Similar strategy is currently applied in clinical trial against Alzheimer’s disease, aimed at perturbing the interaction between the amyloidogenic Aβ peptide and endogenous HS in the brain. We may anticipate further drug applications of HS oligosaccharides or mimetics. Appreciation of the precise role of HS-protein interactions in specific pathology will be essential to such development.

HS-binding polypeptides have been implicated as potential antiangiogenic drugs in cancer therapy. Active compounds include endostatin (54), latent antithrombin (58), histidine-rich glycoprotein and fragments thereof (59). Most or all of these polypeptides bind HS and may interfere with binding of endogenous protein ligands (Fig. 2D). Whereas several HS-binding growth factors (FGF-2, VEGF, PDGF-B) are known to participate in angiogenesis, the precise effects of antiangiogenic peptides on the various relevant signaling systems remain to be defined. Also amyloid disorders are potentially amenable to therapy involving HS-binding peptide drugs, as indicated by the inhibitory effect of small peptides containing the Aβ HHQK sequence on Aβ-induced neurotoxicity (60).

Inhibition of heparan sulfate biosynthesis
General inhibition of HS biosynthesis for therapeutic purposes may seem an adventurous enterprise, given the diverse roles of HS in homeostasis, but is nevertheless worthy of consideration (Fig. 2E). A variety of xylosides have been applied as competitive inhibitors of the xylosyltransferase that initiates polysaccharide substitution on proteoglycan core proteins. Some of these inhibitors that specifically target HS generation have been implicated in cancer therapy, together with inhibitors of polyamine biosynthesis. This strategy builds on the notion that polyamines, essential for cell growth, may be taken up by cells through participation of cell-surface HSPGs, which thus may provide a salvage pathway upon inhibition of endogenous poly-
amine generation (61). Formation of the HS precursor polysaccharide can also be inhibited by glucoamine analogs, that offer prospects for treatment of various amyloid diseases (62). Finally, modulation of HS sulfation may be considered as yet another way to target HS-protein interactions of pathophysiological importance. Small-molecular compounds should be tested as (possibly selective) inhibitors of the various N- and O-sulfotransferases involved in HS biosynthesis. To my knowledge neither the benefits nor the potential hazards of such strategy have yet been assessed.

References