Published in final edited form as:

Curr Opin Struct Biol. 2018 June ; 50: 144–154. doi:10.1016/j.sbi.2018.03.011.

Deciphering functional glycosaminoglycan motifs in development

Robert A. Townley[#] and Hannes E. Bülow^{*,†}

[#]Department of Biological Sciences, Columbia University, New York, New York, 10027

^{*}Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, 10461

[†]Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, New York, 10461

Abstract

Glycosaminoglycans (GAGs) such as heparan sulfate, chondroitin/dermatan sulfate and keratan sulfate are linear glycans, which when attached to protein backbones form proteoglycans. GAGs are essential components of the extracellular space in metazoans. Extensive modifications of the glycans such as sulfation, deacetylation and epimerization create structural GAG motifs. These motifs regulate protein-protein interactions and are thereby responsible for many of the essential functions of GAGs. This review focusses on recent genetic approaches to characterize GAG motifs and their function in defined signaling pathways during development. We discuss a coding approach for GAGs that would enable computational analyses of GAG sequences such as alignments and the computation of position weight matrices to describe GAG motifs.

Graphical Abstract

Correspondence should be addressed to: Hannes E. Bülow Telephone 718 430 3621 Fax 718 430 8778, hannes.buelow@einstein.yu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A new coding approach

Position weight matrix to visualize protein binding sites in GAGs

example EV3D1 (scFv)

Keywords

heparan sulfate; chondroitin sulfate; dermatan sulfate; keratan sulfate; glycosaminoglycan; GAG; alignment; position weight matrix

Introduction

Glycosaminoglycans (GAGs) are linear glycans of repeating disaccharides and are ubiquitous components of cellular membranes and the extracellular space in metazoans. They play important biological roles in many aspects of development and physiology of multicellular organisms [1–8]. GAGs are classified based on the type of disaccharide repeats into heparan sulfates (HS, hexuronic acid/*N*-acetyl-glucosamine), chondroitin/dermatan sulfates (CS/DS, hexuronic acid/*N*-acetyl-galactosamine) and keratan sulfate (KS, galactose/*N*-acetyl-hexosamine)(Figure 1a–c). Hyaluronic acid (HA) comprises glucuronic-acid/*N*-acetyl-glucosamine disaccharides (although with different linkages), yet is not attached to a protein backbone. All GAGs (except HA) show diverse and molecularly complex modification patterns. These modification patterns include epimerizations, sulfations and deacetylations of select positions along the glycan chains, which create motifs that can function as protein binding sites (Figure 1d). Modification patterns are introduced in the Golgi, and not based on a template. Consequently, GAG structures display substantial heterogeneity within tissues and, an enormous molecular diversity across different tissues and over developmental time.

Determining direct structure-function relationships of GAG motifs *in vivo* remains challenging, because no tools exist that allow either the manipulation or removal of defined GAG motifs *in vivo* or, the routine characterization of GAG motifs *in vitro*. Progress has been made in the biochemical characterization of some GAG motifs, although these approaches remain difficult and time consuming. Similarly, the function of some more or less defined GAG sequences *in vitro* has been described using cell culture experiments.

Here, we will review recent progress to establish characteristics of GAG sequences and their functions in defined signaling pathways during development. We will focus on HS, where most is known about the structural characteristics of motifs that mediate function *in vivo*. In a second part, we will briefly review established concepts of functional CS/DS motifs, although some of this data comes from *in vitro* experiments. We will not review KS, for which very little is known about functional domains during development, or HA, which is unmodified and thus does not contain motifs. In a final section we will discuss a novel coding approach that can be used to (1) depict and characterize functional GAG motifs, (2) describe the range of sequences bound by a protein and (3) capture the variation that is inherent in all GAG preparations from natural sources.

HS motifs in development

HS motifs can range significantly in length from a few disaccharides to dodecamers. Their binding to their cognate proteins can be broadly categorized as occurring via at least three different mechanisms [9,10]. The first mechanism is dominated by electrostatic interactions and is mediated primarily by negatively charged groups (carboxyl and sulfate groups) on the glycans and basic amino-acid residues in the protein ligands. Second, HS sequences can bind to proteins based on specific sequences (discussed in more detail below). Finally, HS can bind proteins through two motifs, separated by an unmodified domain of variable length. We focus our review on HS motifs that likely function as specific sequences, or more likely, a set of closely related sequences.

Biochemical studies have successfully identified a few well defined HS motifs, such as the pentasaccharide HS sequence that binds antithrombin III to regulate coagulation or, the HS sequence that binds the fibroblast growth factor 1 (reviewed in [9,10]). While these examples of protein binding HS motifs are structurally very well defined, many proteins may be more tolerant to structural variation in HS motifs for binding. Despite the advances in our structural understanding in a few isolated instances, less is known about the GAG motifs that function *in vivo* during development. The best insight into structures of GAG motifs in development has come from genetic studies that have suggested that distinct sequence characteristics of GAG motifs serve specific and instructive functions in different developmental contexts and signaling systems [11,12•,13-13••,15•,16•,17•,18-21,22••,23,24••,25-26], detailed below.

HS and Slit/Robo signaling

In vitro experiments suggested that the Slit/Robo ligand-receptor pair requires heparan sulfates for function [27]. These findings were confirmed *in vivo* by showing that genetically removing HS polymerizing enzymes result in phenotypes similar to Slit/Robo-mutants. In addition, HS and Slit/Robo genes display genetic interactions [28,29]. Genetic studies in the nematode *Caenorhabditis elegans* established that Slit/Robo signaling in certain cellular contexts required an HS motif comprising iduronic acid (at least one of which must be 2-*O*-sulfated) and 6-*O*-sulfated glucosamine [14](Figure 2). These studies also suggested that a different HS motif may coordinate Slit/Robo signaling in a different cellular context (Figure 2). At least some of structural characteristics of HS motifs important for Slit/Robo signaling

may be conserved, because genetic studies in mice show that 2-*O*-sulfation functions with Slit1 to pattern the optic chiasm [18]. In addition, Slit/Robo signaling has been suggested to depend on the level *N*-sulfation, 2-*O*-sulfation, and 6-*O*-sulfation for vascular development (Figure 2) [30].

HS and KAL1 (Kallmann Syndrome protein 1) signaling

KAL1 is a secreted neural cell adhesion molecule important for neuronal migration and development, which is mutated in some patients with Kallmann Syndrome [31]. In *C. elegans*, KAL-1-dependent neurite branching relies on HS 6-*O*-sulfation, HS C5-epimerization (i.e. iduronic acid), HS 3-*O*-sulfation and, to a lesser extent, on HS 2-*O*-sulfation [13•,14,15•,32,33•]. Thus, KAL-1 neuronal function is coordinated by an HS motif comprising iduronic acid and glucosamine with sulfation on the C6 carbon atom (Figure 2). In addition, the motif must contain a HS 3-*O*-sulfate on a glucosamine, introduced by a HS 3-*O*-sulfotransferase predicted to be of the gD-type. Based on sequence HS 3-*O*-sulfotransferases can be classified into two types, which exhibit different substrate specificities. One class (AT-type) introduces a 3-*O*-sulfate in the antithrombin III binding HS sequence, whereas the other class (gD-type) introduces 3-*O*-sulfate in a motif that binds to the gD protein of herpes simplex [34]. The substrate specificity of the gD-type HS 3-*O*-sulfotransferase requires a 2-*O*-sulfated iduronic acid on the non-reducing side of the glucosamine that is to be 3-*O*-sulfated [34]. Since KAL-1 function is dependent on C5-epimerization but less so on 2-*O*-sulfation [15•], these studies imply that the KAL-1-relevant neuronal HS motif consists of at least two iduronic acids, one of which must be 2-*O*-sulfated and on the non-reducing side of a 3-*O*-sulfated glucosamine, and a 6-*O*-sulfated glucosamine(s). Interestingly, some HS motif characteristics important for KAL1 function seem conserved in humans. Mutations in HS6ST1, a human homolog of the *C. elegans* HS 6-*O*-sulfotransferase *hst-6*, have been found in patients with Kallmann Syndrome suggesting that loss of KAL1 or HS6ST1 results in similar pathologies [35]. Lastly, KAL-1 may function in concert with other HS motifs in a context dependent manner, because in a different cellular context KAL-1 function is dependent on C5-epimerization but less on 6-*O*- or 2-*O*-sulfation [14].

HS and FGF (fibroblast growth factor) signaling

Pioneering studies showed (1) that HS is required for FGF signaling and (2) that HS with different modification patterns can result in distinct FGF-signaling outcomes *in vitro* and *in vivo* [36–38]. Yet, only recently genetic studies have begun to investigate structural characteristics of HS in this process *in vivo*. Loss of function mutations in the HS sulfotransferases Hs2st and Hs6st1 in mice resulted in elevated FGF signaling during formation of the corpus callosum, consistent with the notion that certain HS modification patterns inhibit signaling through FGF8 [17•]. In contrast, in the developing telencephalon Hs2st promoted signaling through FGF2 [16•]. Moreover, studies in the developing lens and lacrimal gland established that *N*-sulfation, HS 6-*O*-sulfation and HS 2-*O*-sulfation promote signaling through FGF3/FGFR1 and FGF10/FGFR2b, respectively [39–41]. Experiments with explants also support a role for HS 6-*O*-sulfation and HS 2-*O*-sulfation in FGF signaling [42•]. In zebrafish, HS 3-*O*-sulfation introduced by a HS 3-*O*-sulfotransferase of the AT-type (see above) is important to induce FGFR target gene expression, implying a role

for HS 3-*O*-sulfation in FGF signaling [22••]. While these studies do not reveal detailed information about the HS motif that mediates function in each case, they nonetheless establish that either similar HS motifs have distinct functions on different FGF/FGFR ligand receptor pairs or, more likely, different HS motifs function in each developmental context, possibly by promoting different ligand receptor pairs.

HS and Wg/Wnt (wingless/Wnt) signaling

The dependence and interaction between Wnt/Wg signaling and heparan sulfate has long been known (reviewed in [43]), yet structural information is only recently emerging. Work in quail showed that QSulf1, a secreted endosulfatase that selectively removes HS 6-*O*-sulfates, appears to promote Wnt signaling [44,45]. In flies, DSulf1 inhibits Wg/Wnt signaling, possibly through modulation of the morphogen gradient [46,47•], whereas loss of the HS 6-*O*-sulfotransferase, which introduces HS 6-*O*-sulfates can promote Wg/Wnt signaling [48]. Collectively, these studies suggest that HS 6-*O*-sulfates may inhibit Wg/Wnt signaling in some contexts, possibly through tighter binding of the Wg/Wnt protein by 6-*O*-sulfated HS or through other mechanisms that result in changes of Wnt gradients. Alternatively, the compensatory upregulation of HS 2-*O*-sulfation seen in Hs6st mutants [49,50] could promote Wg/Wnt signaling. In support of the latter scenario, HS 2-*O*-sulfates have been shown to promote Wnt signaling during zebrafish [23] and possibly mouse development [51]. Nonetheless, HS 6-*O*-sulfation may also be promoting Wnt signaling in other contexts, and the differences may be either species specific or context-dependent.

HS and Hedgehog/Sonic hedgehog and other signaling pathways

Work in flies has long implicated HS proteoglycans with Hh/SHH hedgehog signaling (reviewed in [52]), but relatively little information has been gleaned in regard to HS motifs. Nonetheless, there were some indications that HS may require specific sequence characteristics to regulate Hh/SHH hedgehog signaling in flies [53]. Recently, it was shown that similar to Wg/Wnt signaling, the DSulf1 HS 6-*O*-endosulfatase regulates hedgehog (Hh) signaling [54,55•,56] further supporting this notion. Experiments in mouse primary cell culture suggest that 2-*O*-sulfated HS is important for hedgehog signaling during cerebellar development [57]. However, overall the structural information regarding GAG motifs important for Hh signaling remains limited. The same is true for other pathways such as the bone morphogenetic protein (BMP) pathways. While early experiments clearly established the importance of HS proteoglycans for BMP signaling in vertebrates and invertebrates (e.g. [58,59]), the importance of specific sequences is less clear. On the one hand some reports hint at negative charge [55] or HS domain structure [60] as important determinants of HS to regulate BMP signaling, while other studies imply the importance of specific modifications [24••]. Further studies will be required to resolve these questions.

CS/DS motifs in development

CS/DS disaccharides are categorized based on the modification status of the hexuronic acid/*N*-acetylgalactosamine disaccharides. For example, a disaccharide in which both the 4-*O*-position and 6-*O*-position of glucosamine are sulfated is termed CS-E (Figure 1b). CS/DS oligosaccharides are usually named after the predominant disaccharide, i.e. a polysaccharide termed CS-E may also contain additional differently modified CS disaccharides.

Transcriptional analyses of enzymes that modify CS/DS and, expression analyses of CS/DS epitopes using immunoreagents, suggest that CS/DS displays a molecular diversity in multicellular organisms comparable to HS (reviewed in [61–63]).

The best understood aspects of CS/DS are likely their prominent inhibitory roles during neural regeneration. These aspects of CS/DS function have been reviewed elsewhere [3,64,65]. However, like HS, CS/DS and the associated molecular diversity serves crucial functions during development of multicellular organisms. For example, genetic removal of the *C4st1*/CS 4-*O*-sulfotransferase showed that CS 4-*O*-sulfation is important for modulating the function of morphogens such as TGF β and BMP during bone development in mice [66]. On the other hand, knockdown experiments of the UST/CS uronyl-2-*O*-sulfotransferase and GalNAc4S-6ST/CS GalNAc4S-6-*O*-sulfotransferase in mice revealed that both CS uronyl-2-*O*-sulfation and galactosamine 4,6-disulfation are important for the migration of neuronal precursors in the brain [67], whereas only the latter restricts the formation of multiple axons in dissociated hippocampal cells [68]. However, structure function relationships of CS/DS have been addressed largely through *in vitro* work rather than genetic experiments. In the following we will briefly summarize some of these findings.

CS/DS and FGF (fibroblast growth factor) signaling

Like HS, CS/DS appears to serve important functions in regulating FGF signaling. Enzymatic removal of CS resulted in reduced responsiveness to FGF2 but not EGF in an *ex vivo* model of neural differentiation [69], suggesting that FGF signaling depends on CS. In support of this conclusion, CS/DS, possibly in conjunction with HS, participates in a FGF2-evoked mitogenic response in cell culture models [70]. Moreover, different FGF proteins have been shown to bind distinct CS structures *in vitro* with different affinities, thereby also raising the possibility that distinct CS/DS motifs may serve specific functions [71,72].

CS/DS and axon guidance factors

Semaphorins, an important class of axon guidance molecules that function through neuropilin/plexin receptors, have also been functionally implicated with CS. Enzymatic removal of CS, reduced Sema3a binding on perineuronal nets [73] and Sema3a strongly binds to CS-E [74]. Other semaphorins (such as Sema3e and Sema6b) have been shown to bind CS-A *in vitro*, suggesting that semaphorins in general, are functioning or being modulated by CS with specific structural characteristics.

A recurring feature of CS/DS is its antagonistic relationship with HS. The phenomenon was first described by Kantor and colleagues, who showed that during axonal patterning semaphorin-mediated attraction by Sema5a through HS could be converted into axonal repulsion by CS/DS [75]. This observation is not limited to semaphorin signaling, but also observed for the receptor tyrosine phosphatase sigma (RPTP σ) RPTPs is one of the three 317 LAR-type phosphatases (named after the founding member leukocyte common antigen-related protein (LAR)), of which two (LAR and RPTP σ are known to bind CS [62] Axonal extension of sensory neurons was promoted by HS, but inhibited by CS [76]. Mechanistically this appears to be the result of conformational changes in the RPTP σ receptor, which upon binding to HS adopts a conformation conducive to oligomerization,

whereas binding to CS inhibits oligomerization [76]. Interestingly, HS and CS functioned through the same binding site on the RPTP σ receptor. It is not clear whether or which specific GAG sequence characteristics are required in HS or CS for these different outcomes.

CS/DS and neuronal growth factors

Further evidence for a role of CS/DS in general and CS-E in particular during neuronal development comes from other *in vitro* experiments. For example, a tetrasaccharide containing the two CS-E motifs strongly promoted neurite outgrowth of hippocampal and cortical neuron, among others [77]. This effect was not mimicked by CS saccharides with the same number of sulfates (i.e. negative charges), providing strong evidence for sequence specific effects of the CS oligosaccharide [77]. These effects were blocked by antibodies against neuronal growth factors such as brain-derived neurotrophic factor (BDNF), midkine or their receptors arguing for a function of specific CS oligosaccharides in mediating BDNF-dependent signaling [77,78].

CS/DS and Wg/Wnt (wingless/Wnt) signaling and others

Wg/Wnt (wingless/Wnt) signaling is also regulated by CS/DS. For example, CS-E has been shown to strongly bind protein ligands, such as Wnt-3a. However, beyond binding there is also evidence that CS/DS sulfation patterns are functionally important. The early differentiation of embryonic stem cells into cardiac progenitors is promoted, whereas the later differentiation of cardiac progenitors into cardiomyocytes is inhibited by Wg/Wnt signaling. CS has been shown to inhibit Wg/Wnt signaling and consequently, early enzymatic removal of CS resulted in more cardiac progenitors whereas later removal increased the number of cardiomyocytes [79•]. Conversely, exogenous addition of CS-E to embryonic stem cell cultures had the opposite effect on differentiation. Early addition of CS-E inhibited the differentiation of embryonic stem cells into cardiac progenitors whereas later addition promoted the differentiation of cardiac progenitors into cardiomyocytes [79•], suggesting that the inhibitory functions of CS/DS on Wg/Wnt signaling are mediated by CS-E [79•].

CS/DS and RAGE (Receptor of glycation end products)

CS-E has also plays roles in other contexts. For instance, CS-E regulates metastasis in tumor models by binding to the immunoglobulin domain containing receptor for glycation end products (RAGE) [80]. While RAGE is clearly a receptor for CS-E in these studies, other CS disaccharides exhibited weaker or no binding such as CS-D or CS-A/CS-C, respectively, suggesting that both CS 4-*O*-sulfation and CS 6-*O*-sulfation are important for binding. Additional structural studies suggested that RAGE recognizes a CS/DS oligosaccharide of possibly nine or ten saccharide units in length, including CS-E. Removing CS/DS or competing through exogenous addition of CS-E inhibited RAGE-dependent metastasis in tumor models (reviewed in [62]).

A position weight matrix for glycosaminoglycan sequences

Traditional approaches to describe the sequence of GAG chains of HS and CS/DS have used the characteristic disaccharide as the basic unit, likely because chemical or enzymatic depolymerization of GAGs results in disaccharides that can be uniquely identified and quantified. Disaccharide compositional analyses are conceptually analogous to Erwin Chargaff's studies of DNA [81], which gave insight into basic features of DNA, but failed to provide sequence information. Enzymatic degradation as well as mass spectrometric approaches have been successful to determine the sequence of purified oligosaccharides [82,83].

From a biological perspective, there is no need to treat GAGs as polymers of disaccharides, because proteins that bind GAGs do not necessarily interact with a multiple of disaccharides. For example, one of the best characterized HS sequences, the antithrombin binding sequence, is a pentasaccharide (2.5 disaccharides). Moreover, each (or many) of the modifications likely coordinate(s) interactions with proteins through specific molecular interactions [9,10].

Computational approaches to analyze different GAG sequences are challenging. A HS octasaccharide would allow for almost 17 million theoretical combinations (assuming 32 different disaccharides). Even, if in nature only about twenty different disaccharides have been identified, HS chains can be as long as 150 disaccharides, resulting in an astronomical number of combinations. Here, we propose an alternative approach to encode GAG sequences continuously by building on the coding system by Lawrence *et al.* [84] (Figure 3a). Using HS as an example, the first digit designates the non-reducing sugar (e.g. G or I for glucuronic acid or iduronic acid) and thereby determines the conformation at the C5 of the hexuronic acid, followed by 0 or 2 to designate the sulfation status on the hexuronic acid. The hexosamine with its modification status on the nitrogen is defined by the third digit (H, A, S for NH₂, for NHAc, or NHSO₃, respectively). The fourth and fifth digits determine sulfation status at position 3 and 6 of glucosamine, respectively (Figure 3a). For HS this approach would result in a code with eleven digits to describe every disaccharide through a 'quintuplet' code (Figure 3b). Corresponding approaches to encode CS/DS or KS would also result in a 'quintuplet' code (drawing from nine digits) or a 'quadruplet' code (drawing from six digits), respectively (Figure 3c). As proposed by [84], GAG sequences can be differentiated from each other by the presence of different digits, just as RNA sequences can be discerned from DNA sequences by the presence of U (uracil). The advantages of this approach are obvious. First, the coding does not combine modifications in different positions in a single digit and is continuous, rather than discontinuous as in the case of multimers of disaccharides. Second, the coding allows computational approaches to align glycan sequences. Such computational approaches may become particularly pertinent in the future. As technologies evolve for the generation of HS and CS/DS oligosaccharides [78,85–87], more and more oligosaccharides with defined sequences are becoming available and are used on microarray platforms [78,88–90]. Microarrays will produce large amounts of massive parallel binding data for a given protein to different oligosaccharides. Thus, the third advantage of the new approach is that position weight matrices can be calculated much in analogy to binding sites of transcription factors in DNA (Figure 3c). Such position weight

matrices will allow to both visually and computationally capture the sequence diversity that characterizes the inherent heterogeneity of biological GAG samples as well as the range of sequences bound by a given protein.

Concluding remarks

The structural information we have to date for functional GAG motifs during development is rather limited. While the genetic approach has provided valuable information in many cases, it often falls short of providing detailed structural and sequence information about functional GAG motifs in specific contexts. This limitation is to some extent conceptual. Genetic removal of a single HS modification from an animal will affect not only the motif functioning in the process of interest, but all motifs containing this modification. Moreover, loss of a single HS modifying enzymes can results in effects on other modifications [49,50,91]. Thus, all genetic results are potentially confounded by pleiotropic effects. Nonetheless, the genetic approach has allowed drawing several general conclusions for HS, and likely for CS/DS function *in vivo*. First, distinct combinations of modifications, i.e. different GAG motifs, function in different cellular contexts. Second, a given protein or signaling pathway may employ different GAG motifs in different cellular contexts. Third, specific GAG motifs may display conservation of structural characteristics and anatomical localization over evolutionary time [92•].

The genetic approach has been primarily used to understand HS motifs in development, because fewer knock out models for enzymes that modify CS/DS or KS exist. With the exciting discovery of CS in model organisms [93•4,94•], and knock out models for KS modifying enzymes [95•,96•], there should be convenient new entry points for genetic analyses in analogy to the approaches that have been successful for HS.

In the future, there is a need to devise techniques to specifically interfere with defined motifs in specific developmental contexts, rather than to compromise individual modifications. Such approaches could include the transgenic expression or injection of immunoreagents (many of which exist) that recognize and neutralize specific GAG motifs. This could allow to directly test the function of a given motif in a particular context. Second, the use of more advanced biochemical and computational approaches will have to be harnessed in microarray studies to define the range of GAG sequences that function in conjunction with a particular signaling pathway. We expect that the combination of advanced biochemical and genetic approaches will be synergistic to better understand structure function relationships of GAG motifs.

Acknowledgments

We thank K. Saied-Santiago for comments on the manuscript and, together with D. Zheng, for discussions. Work in the Bülow laboratory is funded by the NIH (RC1 GM090825 and R01GM01313). H.E.B. is an Irma T. Hirschl/Monique Weill-Caullier research fellow.

We apologize for omission of literature due to space constraints and the focus of this review on primarily *in vivo* evidence for structural characteristics of GAG motifs in developmental processes.

References

1. Bishop JR, Schuksz M, Esko JD. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature*. 2007; 446:1030–1037. [PubMed: 17460664]
2. Saied-Santiago K, Bülow HE. Diverse roles for glycosaminoglycans in neural patterning. *Dev Dyn*. 2017; doi: 10.1002/dvdy.24555
3. Silver DJ, Silver J. Contributions of chondroitin sulfate proteoglycans to neurodevelopment, injury, and cancer. *Curr Opin Neurobiol*. 2014; 27:171–178. [PubMed: 24762654]
4. Masu M. Proteoglycans and axon guidance: a new relationship between old partners. *J Neurochem*. 2016; 139(Suppl 2):58–75. [PubMed: 26709493]
5. Poulain FE, Yost HJ. Heparan sulfate proteoglycans: a sugar code for vertebrate development? *Development*. 2015; 142:3456–3467. [PubMed: 26487777]
6. Kwok JC, Warren P, Fawcett JW. Chondroitin sulfate: a key molecule in the brain matrix. *Int J Biochem Cell Biol*. 2012; 44:582–586. [PubMed: 22265655]
7. Smith PD, Coulson-Thomas VJ, Foscarin S, Kwok JC, Fawcett JW. “GAG-ing with the neuron”: The role of glycosaminoglycan patterning in the central nervous system. *Exp Neurol*. 2015; 274:100–114. [PubMed: 26277685]
8. Coulson-Thomas VJ. The role of heparan sulphate in development: the ectodermal story. *Int J Exp Pathol*. 2016; 97:213–229. [PubMed: 27385054]
9. Xu D, Esko JD. Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem*. 2014; 83:129–157. [PubMed: 24606135]
10. Meneghetti MC, Hughes AJ, Rudd TR, Nader HB, Powell AK, Yates EA, Lima MA. Heparan sulfate and heparin interactions with proteins. *J R Soc Interface*. 2015; 12:0589. [PubMed: 26289657]
11. Kinnunen T, Huang Z, Townsend J, Gatdula MM, Brown JR, Esko JD, Turnbull JE. Heparan 2-O-sulfotransferase, hst-2, is essential for normal cell migration in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 2005; 102:1507–1512. [PubMed: 15671174]
12. Kinnunen TK. Combinatorial roles of heparan sulfate proteoglycans and heparan sulfates in *Caenorhabditis elegans* neural development. *PLoS One*. 2014; 9:e102919. This paper together with [15] demonstrates the redundancy of HSPG core proteins in developmental processes. [PubMed: 25054285]
13. Teclé E, Díaz-Balzac CA, Bülow HE. Distinct 3-O-sulfated heparan sulfate modification patterns are required for kal-1-dependent neurite branching in a context-dependent manner in *Caenorhabditis elegans*. *G3 (Bethesda)*. 2013; 3:541–552. This paper together with [22,24] provides the first in vivo evidence for HS 3-O-sulfation in connection with specific signaling pathways. [PubMed: 23451335]
14. Bülow HE, Hobert O. Differential sulfations and epimerization define heparan sulfate specificity in nervous system development. *Neuron*. 2004; 41:723–736. [PubMed: 15003172]
15. Díaz-Balzac CA, Lázaro-Peña MI, Teclé E, Gomez N, Bülow HE. Complex Cooperative Functions of Heparan Sulfate Proteoglycans Shape Nervous System Development in *Caenorhabditis elegans*. *G3 (Bethesda)*. 2014; 4:1859–1870. This paper together with [33] provides a detailed genetic analysis of KAL-1 function and shows, together with [12], the redundancy of HSPG function. [PubMed: 25098771]
16. Chan WK, Howe K, Clegg JM, Guimond SE, Price DJ, Turnbull JE, Pratt T. 2-O Heparan Sulfate Sulfation by Hs2st Is Required for Erk/Mapk Signalling Activation at the Mid-Gestational Mouse Telencephalic Midline. *PLoS One*. 2015; 10:e0130147. Together with [17,42] this paper provides convincing evidence that HS 6-O-sulfation and HS 2-O-sulfation are important for FGF signaling. [PubMed: 26075383]
17. Clegg JM, Conway CD, Howe KM, Price DJ, Mason JO, Turnbull JE, Basson MA, Pratt T. Heparan sulfotransferases Hs6st1 and Hs2st keep Erk in check for mouse corpus callosum development. *J Neurosci*. 2014; 34:2389–2401. Together with [16,42] this paper provides convincing evidence that HS 6-O-sulfation and HS 2-O-sulfation are important for FGF signaling. [PubMed: 24501377]

18. Conway CD, Howe KM, Nettleton NK, Price DJ, Mason JO, Pratt T. Heparan sulfate sugar modifications mediate the functions of slits and other factors needed for mouse forebrain commissure development. *J Neurosci.* 2011; 31:1955–1970. [PubMed: 21307234]
19. Conway CD, Price DJ, Pratt T, Mason JO. Analysis of axon guidance defects at the optic chiasm in heparan sulphate sulphotransferase compound mutant mice. *J Anat.* 2011; 219:734–742. [PubMed: 21951307]
20. McLaughlin D, Karlsson F, Tian N, Pratt T, Bullock SL, Wilson VA, Price DJ, Mason JO. Specific modification of heparan sulphate is required for normal cerebral cortical development. *Mech Dev.* 2003; 120:1481–1488. [PubMed: 14654220]
21. Pratt T, Conway CD, Tian NM, Price DJ, Mason JO. Heparan sulphation patterns generated by specific heparan sulfotransferase enzymes direct distinct aspects of retinal axon guidance at the optic chiasm. *J Neurosci.* 2006; 26:6911–6923. [PubMed: 16807321]
- 22•. Neugebauer JM, Cadwallader AB, Amack JD, Bisgrove BW, Yost HJ. Differential roles for 3-OSTs in the regulation of cilia length and motility. *Development.* 2013; 140:3892–3902. This paper together with [13,24] provides the first in vivo evidence for HS 3-O-sulfation in connection with specific signaling pathways. [PubMed: 23946439]
23. Cadwalader EL, Condic ML, Yost HJ. 2-O-sulfotransferase regulates Wnt signaling, cell adhesion and cell cycle during zebrafish epiboly. *Development.* 2012; 139:1296–1305. [PubMed: 22357927]
- 24•. Samson SC, Ferrer T, Jou CJ, Sachse FB, Shankaran SS, Shaw RM, Chi NC, Tristani-Firouzi M, Yost HJ. 3-OST-7 regulates BMP-dependent cardiac contraction. *PLoS Biol.* 2013; 11:e1001727. This paper together with [13,22] provides the first in vivo evidence for HS 3-O-sulfation in connection with specific signaling pathways. [PubMed: 24311987]
25. Saied-Santiago K, Townley RA, Attonito JD, da Cunha DS, Diaz-Balzac CA, Tecle E, Bülow HE. Coordination of Heparan Sulfate Proteoglycans with Wnt Signaling To Control Cellular Migrations and Positioning in *Caenorhabditis elegans*. *Genetics.* 2017; doi: 10.1534/genetics.116.198739
26. Bülow HE, Tjoe N, Townley RA, Didiano D, van Kuppevelt TH, Hobert O. Extracellular Sugar Modifications Provide Instructive and Cell-Specific Information for Axon-Guidance Choices. *Curr Biol.* 2008; 18:1978–1985. [PubMed: 19062279]
27. Hu H. Cell-surface heparan sulfate is involved in the repulsive guidance activities of Slit2 protein. *Nat Neurosci.* 2001; 4:695–701. [PubMed: 11426225]
28. Inatani M, Irie F, Plump AS, Tessier-Lavigne M, Yamaguchi Y. Mammalian brain morphogenesis and midline axon guidance require heparan sulfate. *Science.* 2003; 302:1044–1046. [PubMed: 14605369]
29. Lee JS, von der Hardt S, Rusch MA, Stringer SE, Stickney HL, Talbot WS, Geisler R, Nusslein-Volhard C, Selleck SB, Chien CB, et al. Axon sorting in the optic tract requires HSPG synthesis by ext2 (dackel) and extl3 (boxer). *Neuron.* 2004; 44:947–960. [PubMed: 15603738]
30. Zhang B, Xiao W, Qiu H, Zhang F, Moniz HA, Jaworski A, Condac E, Gutierrez-Sanchez G, Heiss C, Clugston RD, et al. Heparan sulfate deficiency disrupts developmental angiogenesis and causes congenital diaphragmatic hernia. *J Clin Invest.* 2014; 124:209–221. [PubMed: 24355925]
31. Boehm U, Bouloux PM, Dattani MT, de Roux N, Dode C, Dunkel L, Dwyer AA, Giacobini P, Hardelin JP, Juul A, et al. Expert consensus document: European Consensus Statement on congenital hypogonadotropic hypogonadism--pathogenesis, diagnosis and treatment. *Nat Rev Endocrinol.* 2015; 11:547–564. [PubMed: 26194704]
32. Bülow HE, Berry KL, Topper LH, Peles E, Hobert O. Heparan sulfate proteoglycan-dependent induction of axon branching and axon misrouting by the Kallmann syndrome gene *kal-1*. *Proc Natl Acad Sci U S A.* 2002; 99:6346–6351. [PubMed: 11983919]
- 33•. Schwiertman AA, Steves AN, Yee V, Donelson CJ, Bentley MR, Santorella EM, Mehlenbacher TV, Pital A, Howard AM, Wilson MR, et al. The *Caenorhabditis elegans* Ephrin EFN-4 Functions Non-cell Autonomously with Heparan Sulfate Proteoglycans to Promote Axon Outgrowth and Branching. *Genetics.* 2016; 202:639–660. This paper together with [15] provides a detailed genetic analysis of *KAL-1* function. [PubMed: 26645816]
34. Thacker BE, Xu D, Lawrence R, Esko JD. Heparan sulfate 3-O-sulfation: a rare modification in search of a function. *Matrix Biol.* 2014; 35:60–72. [PubMed: 24361527]

35. Tornberg J, Sykietis GP, Keefe K, Plummer L, Hoang X, Hall JE, Quinton R, Seminara SB, Hughes V, Van Vliet G, et al. Heparan sulfate 6-O-sulfotransferase 1, a gene involved in extracellular sugar modifications, is mutated in patients with idiopathic hypogonadotrophic hypogonadism. *Proc Natl Acad Sci U S A*. 2011; 108:11524–11529. [PubMed: 21700882]
36. Guimond SE, Turnbull JE. Fibroblast growth factor receptor signalling is dictated by specific heparan sulphate saccharides. *Curr Biol*. 1999; 9:1343–1346. [PubMed: 10574766]
37. Allen BL, Rapraeger AC. Spatial and temporal expression of heparan sulfate in mouse development regulates FGF and FGF receptor assembly. *J Cell Biol*. 2003; 163:637–648. [PubMed: 14610064]
38. Lin X, Buff EM, Perrimon N, Michelson AM. Heparan sulfate proteoglycans are essential for FGF receptor signaling during *Drosophila* embryonic development. *Development*. 1999; 126:3715–3723. [PubMed: 10433902]
39. Qu X, Hertzler K, Pan Y, Grobe K, Robinson ML, Zhang X. Genetic epistasis between heparan sulfate and FGF-Ras signaling controls lens development. *Dev Biol*. 2011; 355:12–20. [PubMed: 21536023]
40. Qu X, Carbe C, Tao C, Powers A, Lawrence R, van Kuppevelt TH, Cardoso WV, Grobe K, Esko JD, Zhang X. Lacrimal gland development and Fgf10-Fgfr2b signaling are controlled by 2-O- and 6-O-sulfated heparan sulfate. *J Biol Chem*. 2011; 286:14435–14444. [PubMed: 21357686]
41. Cai Z, Grobe K, Zhang X. Role of heparan sulfate proteoglycans in optic disc and stalk morphogenesis. *Dev Dyn*. 2014; 243:1310–1316. [PubMed: 24753163]
42. Tillo M, Charoy C, Schwarz Q, Maden CH, Davidson K, Fantin A, Ruhrberg C. 2- and 6-O-sulfated proteoglycans have distinct and complementary roles in cranial axon guidance and motor neuron migration. *Development*. 2016; 143:1907–1913. Together with [16,17] this paper provides convincing evidence that HS 6-O-sulfation and HS 2-O-sulfation are important for FGF signaling. [PubMed: 27048738]
43. Niehrs C. The complex world of WNT receptor signalling. *Nat Rev Mol Cell Biol*. 2012; 13:767–779. [PubMed: 23151663]
44. Dhoot GK, Gustafsson MK, Ai X, Sun W, Standiford DM, Emerson CP Jr. Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. *Science*. 2001; 293:1663–1666. [PubMed: 11533491]
45. Ai X, Do AT, Lozynska O, Kusche-Gullberg M, Lindahl U, Emerson CP Jr. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. *J Cell Biol*. 2003; 162:341–351. [PubMed: 12860968]
46. Kleinschmit A, Koyama T, Dejima K, Hayashi Y, Kamimura K, Nakato H. *Drosophila* heparan sulfate 6-O endosulfatase regulates Wingless morphogen gradient formation. *Dev Biol*. 2010; 345:204–214. [PubMed: 20637191]
47. Kleinschmit A, Takemura M, Dejima K, Choi PY, Nakato H. *Drosophila* heparan sulfate 6-O-endosulfatase Sulf1 facilitates wingless (Wg) protein degradation. *J Biol Chem*. 2013; 288:5081–5089. This paper together with [55] provides a detailed genetic analysis of DSulf and signaling. [PubMed: 23300081]
48. Dani N, Nahm M, Lee S, Broadie K. A targeted glycan-related gene screen reveals heparan sulfate proteoglycan sulfation regulates WNT and BMP trans-synaptic signaling. *PLoS Genet*. 2012; 8:e1003031. [PubMed: 23144627]
49. Kamimura K, Koyama T, Habuchi H, Ueda R, Masu M, Kimata K, Nakato H. Specific and flexible roles of heparan sulfate modifications in *Drosophila* FGF signaling. *J Cell Biol*. 2006; 174:773–778. [PubMed: 16966419]
50. Townley RA, Bülow HE. Genetic analysis of the heparan modification network in *Caenorhabditis elegans*. *J Biol Chem*. 2011; 286:16824–16831. [PubMed: 21454666]
51. Shah MM, Sakurai H, Sweeney DE, Gallegos TF, Bush KT, Esko JD, Nigam SK. Hs2st mediated kidney mesenchyme induction regulates early ureteric bud branching. *Dev Biol*. 2010; 339:354–365. [PubMed: 20059993]
52. Nakato H, Li JP. Functions of Heparan Sulfate Proteoglycans in Development: Insights From *Drosophila* Models. *Int Rev Cell Mol Biol*. 2016; 325:275–293. [PubMed: 27241223]

53. The I, Bellaiche Y, Perrimon N. Hedgehog movement is regulated through tout velu-dependent synthesis of a heparan sulfate proteoglycan. *Mol Cell*. 1999; 4:633–639. [PubMed: 10549295]
54. Wojcinski A, Nakato H, Soula C, Glise B. DSulfatase-1 fine-tunes Hedgehog patterning activity through a novel regulatory feedback loop. *Dev Biol*. 2011; 358:168–180. [PubMed: 21806980]
55. Dejima K, Kleinschmit A, Takemura M, Choi PY, Kinoshita-Toyoda A, Toyoda H, Nakato H. The role of *Drosophila* heparan sulfate 6-O-endosulfatase in sulfation compensation. *J Biol Chem*. 2013; 288:6574–6582. This paper together with [47] provides a detailed genetic analysis of DSulf and signaling. [PubMed: 23339195]
56. Takemura M, Nakato H. *Drosophila* Sulf1 is required for the termination of intestinal stem cell division during regeneration. *J Cell Sci*. 2017; 130:332–343. [PubMed: 27888216]
57. Witt RM, Hecht ML, Pazyra-Murphy MF, Cohen SM, Noti C, van Kuppevelt TH, Fuller M, Chan JA, Hopwood JJ, Seeberger PH, et al. Heparan sulfate proteoglycans containing a glypican 5 core and 2-O-sulfo-iduronic acid function as Sonic Hedgehog co-receptors to promote proliferation. *J Biol Chem*. 2013; 288:26275–26288. [PubMed: 23867465]
58. Jackson SM, Nakato H, Sugiura M, Jannuzi A, Oakes R, Kaluza V, Golden C, Selleck SB. dally, a *Drosophila* glypican, controls cellular responses to the TGF-beta-related morphogen, Dpp. *Development*. 1997; 124:4113–4120. [PubMed: 9374407]
59. Paine-Saunders S, Viviano BL, Zupicich J, Skarnes WC, Saunders S. glypican-3 controls cellular responses to Bmp4 in limb patterning and skeletal development. *Dev Biol*. 2000; 225:179–187. [PubMed: 10964473]
60. Viviano BL, Paine-Saunders S, Gasiunas N, Gallagher J, Saunders S. Domain-specific modification of heparan sulfate by Qsulf1 modulates the binding of the bone morphogenetic protein antagonist Noggin. *J Biol Chem*. 2004; 279:5604–5611. [PubMed: 14645250]
61. Purushothaman A, Sugahara K, Faissner A. Chondroitin sulfate “wobble motifs” modulate maintenance and differentiation of neural stem cells and their progeny. *J Biol Chem*. 2012; 287:2935–2942. [PubMed: 22094467]
62. Mizumoto S, Yamada S, Sugahara K. Molecular interactions between chondroitin-dermatan sulfate and growth factors/receptors/matrix proteins. *Curr Opin Struct Biol*. 2015; 34:35–42. [PubMed: 26164146]
63. Miller GM, Hsieh-Wilson LC. Sugar-dependent modulation of neuronal development, regeneration, and plasticity by chondroitin sulfate proteoglycans. *Exp Neurol*. 2015; 274:115–125. [PubMed: 26315937]
64. Rauvala H, Paveliev M, Kuja-Panula J, Kuleshkaya N. Inhibition and enhancement of neural regeneration by chondroitin sulfate proteoglycans. *Neural Regen Res*. 2017; 12:687–691. [PubMed: 28616017]
65. Fawcett JW. The extracellular matrix in plasticity and regeneration after CNS injury and neurodegenerative disease. *Prog Brain Res*. 2015; 218:213–226. [PubMed: 25890139]
66. Klüppel M, Wight TN, Chan C, Hinek A, Wrana JL. Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis. *Development*. 2005; 132:3989–4003. [PubMed: 16079159]
67. Ishii M, Maeda N. Oversulfated chondroitin sulfate plays critical roles in the neuronal migration in the cerebral cortex. *J Biol Chem*. 2008; 283:32610–32620. [PubMed: 18819920]
68. Nishimura K, Ishii M, Kuraoka M, Kamimura K, Maeda N. Opposing functions of chondroitin sulfate and heparan sulfate during early neuronal polarization. *Neuroscience*. 2010; 169:1535–1547. [PubMed: 20600662]
69. Sirko S, von Holst A, Weber A, Wizenmann A, Theodoridis U, Götz M, Faissner A. Chondroitin sulfates are required for fibroblast growth factor-2-dependent proliferation and maintenance in neural stem cells and for epidermal growth factor-dependent migration of their progeny. *Stem Cells*. 2010; 28:775–787. [PubMed: 20087964]
70. Nikitovic D, Assouti M, Sifaki M, Katonis P, Krasagakis K, Karamanos NK, Tzanakakis GN. Chondroitin sulfate and heparan sulfate-containing proteoglycans are both partners and targets of basic fibroblast growth factor-mediated proliferation in human metastatic melanoma cell lines. *Int J Biochem Cell Biol*. 2008; 40:72–83. [PubMed: 17706452]

71. Shipp EL, Hsieh-Wilson LC. Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem Biol.* 2007; 14:195–208. [PubMed: 17317573]
72. Deepa SS, Umehara Y, Higashiyama S, Itoh N, Sugahara K. Specific molecular interactions of oversulfated chondroitin sulfate E with various heparin-binding growth factors. Implications as a physiological binding partner in the brain and other tissues. *J Biol Chem.* 2002; 277:43707–43716. [PubMed: 12221095]
73. Vo T, Carulli D, Ehlert EM, Kwok JC, Dick G, Mecollari V, Moloney EB, Neufeld G, de Winter F, Fawcett JW, et al. The chemorepulsive axon guidance protein semaphorin3A is a constituent of perineuronal nets in the adult rodent brain. *Mol Cell Neurosci.* 2013; 56:186–200. [PubMed: 23665579]
74. Dick G, Tan CL, Alves JN, Ehlert EM, Miller GM, Hsieh-Wilson LC, Sugahara K, Oosterhof A, van Kuppevelt TH, Verhaagen J, et al. Semaphorin 3A binds to the perineuronal nets via chondroitin sulfate type E motifs in rodent brains. *J Biol Chem.* 2013; 288:27384–27395. [PubMed: 23940048]
75. Kantor DB, Chivatakarn O, Peer KL, Oster SF, Inatani M, Hansen MJ, Flanagan JG, Yamaguchi Y, Sretavan DW, Giger RJ, et al. Semaphorin 5A is a bifunctional axon guidance cue regulated by heparan and chondroitin sulfate proteoglycans. *Neuron.* 2004; 44:961–975. [PubMed: 15603739]
76. Coles CH, Shen Y, Tenney AP, Siebold C, Sutton GC, Lu W, Gallagher JT, Jones EY, Flanagan JG, Aricescu AR. Proteoglycan-specific molecular switch for RPTPsigma clustering and neuronal extension. *Science.* 2011; 332:484–488. [PubMed: 21454754]
77. Gama CI, Tully SE, Sotogaku N, Clark PM, Rawat M, Vaidehi N, Goddard WA 3rd, Nishi A, Hsieh-Wilson LC. Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat Chem Biol.* 2006; 2:467–473. [PubMed: 16878128]
78. Rogers CJ, Clark PM, Tully SE, Abrol R, Garcia KC, Goddard WA 3rd, Hsieh-Wilson LC. Elucidating glycosaminoglycan-protein-protein interactions using carbohydrate microarray and computational approaches. *Proc Natl Acad Sci U S A.* 2011; 108:9747–9752. [PubMed: 21628576]
79. Prinz RD, Willis CM, van Kuppevelt TH, Klüppel M. Biphasic role of chondroitin sulfate in cardiac differentiation of embryonic stem cells through inhibition of Wnt/beta-catenin signaling. *PLoS One.* 2014; 9:e92381. A detailed study suggesting that CS-E inhibits Wg/Wnt signaling. [PubMed: 24667694]
80. Mizumoto S, Takahashi J, Sugahara K. Receptor for advanced glycation end products (RAGE) functions as receptor for specific sulfated glycosaminoglycans, and anti-RAGE antibody or sulfated glycosaminoglycans delivered in vivo inhibit pulmonary metastasis of tumor cells. *J Biol Chem.* 2012; 287:18985–18994. [PubMed: 22493510]
81. Chargaff E, Zamenhof S, Green C. Composition of human desoxypentose nucleic acid. *Nature.* 1950; 165:756–757.
82. Venkataraman G, Shriver Z, Raman R, Sasisekharan R. Sequencing complex polysaccharides. *Science.* 1999; 286:537–542. [PubMed: 10521350]
83. Turnbull JE. Strategies for sequencing bioactive glycosaminoglycan saccharides. In: Lander ANH, Selleck SB, Turnbull JE, Coath C, editors *Cell Surface Proteoglycans in Signalling and Development*. HSFP; 1999. 58–67.
84. Lawrence R, Lu H, Rosenberg RD, Esko JD, Zhang L. Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans. *Nat Methods.* 2008; 5:291–292. [PubMed: 18376390]
85. Xu Y, Masuko S, Takieddin M, Xu H, Liu R, Jing J, Mousa SA, Linhardt RJ, Liu J. Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science.* 2011; 334:498–501. [PubMed: 22034431]
86. Dulaney SB, Xu Y, Wang P, Tiruchinapally G, Wang Z, Kathawa J, El-Dakdouki MH, Yang B, Liu J, Huang X. Divergent Synthesis of Heparan Sulfate Oligosaccharides. *The Journal of Organic Chemistry.* 2015; 80:12265–12279. [PubMed: 26574650]

87. Arungundram S, Al-Mafraji K, Asong J, Leach FE 3rd, Amster IJ, Venot A, Turnbull JE, Boons GJ. Modular synthesis of heparan sulfate oligosaccharides for structure-activity relationship studies. *J Am Chem Soc.* 2009; 131:17394–17405. [PubMed: 19904943]
88. de Paz JL, Noti C, Seeberger PH. Microarrays of synthetic heparin oligosaccharides. *J Am Chem Soc.* 2006; 128:2766–2767. [PubMed: 16506732]
89. de Paz JL, Spillmann D, Seeberger PH. Microarrays of heparin oligosaccharides obtained by nitrous acid depolymerization of isolated heparin. *Chem Commun (Camb).* 2006; :3116–3118.doi: 10.1039/b605318a [PubMed: 16855704]
90. Zong C, Venot A, Li X, Lu W, Xiao W, Wilkes JL, Salanga CL, Handel TM, Wang L, Wolfert MA, et al. Heparan Sulfate Microarray Reveals That Heparan Sulfate-Protein Binding Exhibits Different Ligand Requirements. *J Am Chem Soc.* 2017; 139:9534–9543. A systematic microarray study to probe HS-protein interactions. [PubMed: 28651046]
91. Ringvall M, Kjellen L. Mice deficient in heparan sulfate N-deacetylase/N-sulfotransferase 1. *Prog Mol Biol Transl Sci.* 2010; 93:35–58. [PubMed: 20807640]
92. Attreed M, Saied-Santiago K, Bülow HE. Conservation of anatomically restricted glycosaminoglycan structures in divergent nematode species. *Glycobiology.* 2016; 26:862–870. The first report showing that specific HS epitopes may be anatomically and structurally conserved during evolution. [PubMed: 26976619]
93. Izumikawa T, Dejima K, Watamoto Y, Nomura KH, Kanaki N, Rikitake M, Tou M, Murata D, Yanagita E, Kano A, et al. Chondroitin 4-O-Sulfotransferase Is Indispensable for Sulfation of Chondroitin and Plays an Important Role in Maintaining Normal Life Span and Oxidative Stress Responses in Nematodes. *J Biol Chem.* 2016; 291:23294–23304. Together with [94] the first report that CS exists in nematodes. This should open the door for systematic genetic analyses. [PubMed: 27645998]
94. Dierker T, Shao C, Haitina T, Zaia J, Hinas A, Kjellen L. Nematodes join the family of chondroitin sulfate-synthesizing organisms: Identification of an active chondroitin sulfotransferase in *Caenorhabditis elegans*. *Sci Rep.* 2016; 6:34662. Together with [93] the first report that CS exists in nematodes. This should open the door for systematic genetic analyses. [PubMed: 27703236]
95. Hoshino H, Foyez T, Ohtake-Niimi S, Takeda-Uchimura Y, Michikawa M, Kadomatsu K, Uchimura K. KSGal6ST is essential for the 6-sulfation of galactose within keratan sulfate in early postnatal brain. *J Histochem Cytochem.* 2014; 62:145–156. Together with [96] this paper reports a mouse knock out model for a key KS modification enzyme. [PubMed: 24152993]
96. Hashimoto H, Ishino Y, Jiang W, Yoshimura T, Takeda-Uchimura Y, Uchimura K, Kadomatsu K, Ikenaka K. Keratan Sulfate Regulates the Switch from Motor Neuron to Oligodendrocyte Generation During Development of the Mouse Spinal Cord. *Neurochem Res.* 2016; 41:450–462. Together with [95] this paper reports a mouse knock out model for a key KS modification enzyme. [PubMed: 26869039]
97. Kreuger J, Salmivirta M, Sturiale L, Gimenez-Gallego G, Lindahl U. Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2. *J Biol Chem.* 2001; 276:30744–30752. [PubMed: 11406624]
98. Dennissen MA, Jenniskens GJ, Pieffers M, Versteeg EM, Petitou M, Veerkamp JH, van Kuppevelt TH. Large, tissue-regulated domain diversity of heparan sulfates demonstrated by phage display antibodies. *J Biol Chem.* 2002; 277:10982–10986. [PubMed: 11790764]

Highlights

- Specific GAG sequences can function in defined cellular contexts
- A given protein can interact with different HS motifs in different contexts
- Protein-protein interactions can be modulated by defined GAG sequences
- A new coding approach to compute position weight matrices for GAG sequences

Figure 1. General structures of GAGs

- a. Schematic of a heparan sulfate disaccharide with all possible modifications and the respective enzymatic activities required to introduce them are indicated: NDST (*N*-deacetylase-sulfotransferase), GLCE (*C*5-glucuronyl-epimerase), HS2ST (HS-2-*O*-sulfotransferase), HS3ST (HS-3-*O*-sulfotransferase) and HS6ST (HS-6-*O*-sulfotransferase).
- b. Schematic of a chondroitin/dermatan sulfate disaccharide with all possible modifications indicated. Dermatan sulfate is characterized by the presence of iduronic acid. The respective enzymatic activities required to introduce modifications are indicated: CS/DS5epi (*C*5-glucuronyl-epimerase), UST (uronyl-2-*O*-sulfotransferase), C4ST/D4ST (chondroitin/dermatan-4-*O*-sulfotransferase), C6ST (chondroitin-6-*O*-sulfotransferase), GalNAc4S-6ST (GalNAc4S-6-*O*-sulfotransferase). A table lists the structural characteristics of specific CS disaccharides using the coding system of Lawrence *et al.* [85] as well as our extended Lawrence system (cf. Figure 3).
- c. Schematic of a keratan sulfate disaccharide with all possible modifications and the respective enzymatic activities required to introduce them are indicated: KSGal6ST (KS galactose-6-*O*-sulfotransferase), GlcNAc6ST (KS *N*-acetyl-glucosamine-6-*O*-sulfotransferase).
- d. An HS chain attached to a protein is shown. Red circles indicate sulfates in various positions. Note the characteristic tetrasaccharide linker by which the chain is attached to a

serine in a protein and, the structural segregation into domains with a high degree (NS), low degree (NS/NA) or no (NA) modification. A putative FGF1 binding motif [97] is shown.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Summary of functional HS epitopes during development

Summary of structural characteristics and findings for HS and CS motifs functioning with different signaling pathways in the species indicated: Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Mm, *Mus musculus*; Dr, *Danio rerio*; Hs, *Homo sapiens*.

Figure 3. A coding system to create position weight matrices of GAG epitopes

a. A glycosaminoglycan coding system to describe glycosaminoglycan sequences (modified after [84]), which we term “extended Lawrence”. This coding system differs from the original proposed by [84] in that the modifications on the hexosamine are indicated by two rather than one digit. As noted by [84], the stereochemistry of linkages can be indicated if desired.

b. HS disaccharide with five positions that can be modified are shown for both the hexuronic acid and the glucosamine using the “extended Lawrence code”. The resulting 11 digits can describe a HS disaccharide in a quintuplet code. IdoUA: iduronic acid, GlcUA: glucuronic acid.

c. CS/DS disaccharide with five positions that can be modified are shown for both the hexuronic acid and the galactosamine using the “extended Lawrence code”. The resulting 9 digits can describe a CS/DS disaccharide in a quintuplet code. IdoUA: iduronic acid, GlcUA: glucuronic acid. A conceptually analogous code could be used for other linear glycans including KS.

d. Sample position weight matrix of an HS epitope recognized by a single chain variable fragment antibody (scFv, EV3D1), calculated from competitive ELISA data with 12 defined HS oligos [98]. The methodology to calculate such position weight matrices will be

described elsewhere. Note the presence of invariant and variable positions. The accuracy of these position weight matrices is a function of the number and structural diversity of oligosaccharides used in obtaining binding data.