ELSEVIER

Chemical approaches to deciphering the glycosaminoglycan code

Cristal I Gama and Linda C Hsieh-Wilson

Glycosaminoglycans are sulfated biopolymers with rich chemical diversity and complex functions *in vivo*, contributing to processes ranging from cell growth and neuronal development to viral invasion and neurodegenerative disease. Recent studies suggest that glycosaminoglycans may encode information in the form of a 'sulfation code,' whereby discrete modifications to the polysaccharide backbone may direct the location or activities of proteins.

Addresses
Howard Hughes Medical Institute, Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

Corresponding author: Hsieh-Wilson, Linda C (lhw@caltech.edu)

Current Opinion in Chemical Biology 2005, **9**:609–619

This review comes from a themed issue on
Biopolymers
Edited by Scott A Strobel and Tom W Muir

Available online 18th October 2005

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2005.10.003

Introduction

Glycosaminoglycans are a family of sulfated polysaccharides involved in diverse biological processes such as neuronal development, tumor growth and metastasis, viral invasion and spinal cord injury [1–3]. For example, glycosaminoglycans modulate key signaling pathways essential for proper cell growth and angiogenesis [1,2]. They are also important for axon pathfinding in the developing brain and have been linked to the pathology of Alzheimer's disease [4••,5,6]. The remarkable ability of glycosaminoglycans to regulate various processes is only beginning to be understood at a molecular level. Increasing evidence suggests that glycosaminoglycans encode information in the form of a 'sulfation code.' Namely, discrete sulfation motifs along the carbohydrate backbone carry instructions to direct proteins and regulate complex processes such as neuronal wiring. Deciphering this code and the mechanisms by which it coordinates biological events is crucial for understanding diverse aspects of biology and could reveal new therapeutic opportunities.

In this review, we describe evidence for the sulfation code in the context of heparan and chondroitin sulfate glycosaminoglycans. We also discuss the potential for synthetic, chemoenzymatic and technological approaches to advance a molecular-level understanding of this important class of biopolymers.

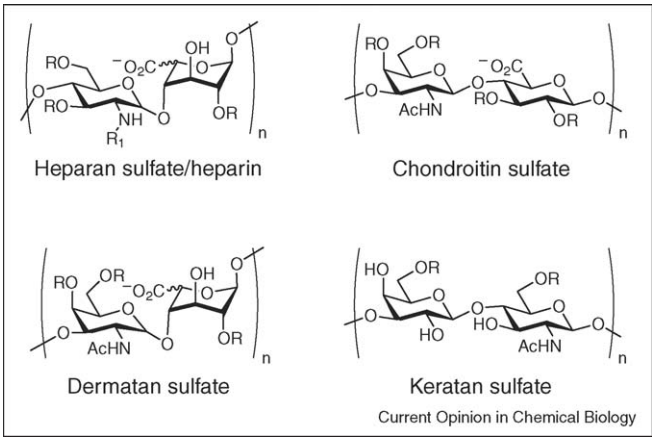
The chemical diversity of glycosaminoglycans

Glycosaminoglycans are composed of repeating disaccharide subunits that are assembled into linear polysaccharide chains (Figure 1). These polysaccharides are often covalently attached to proteoglycan proteins at the cell surface or in the extracellular matrix. There are several major classes of glycosaminoglycans, including heparan sulfate (HS) and heparin, chondroitin sulfate (CS), dermatan sulfate and keratan sulfate, which differ in their core disaccharide subunit. HS and heparin contain D-glucosamine (GlcN) and either D-glucuronic acid (GlcA) or L-iduronic acid (IdoA) subunits joined by α(1,4) and β(1,4) linkages. CS has *N*-acetylgalactosamine (GalNAc) and GlcA subunits and alternating β(1,3) and β(1,4) linkages.

Diverse sulfation patterns are generated *in vivo* through extensive modification of the carbohydrate backbone [3]. For example, sulfation of HS and heparin can occur at the C2 hydroxyl of IdoA and the C3 and C6 hydroxyls of GlcN. The C2 amine can also be either sulfated, acetylated or remain unmodified. Although HS and heparin are structurally related, HS has greater overall chemical complexity, exhibiting more varied sulfation patterns, lower IdoA content and longer polysaccharide chains. Moreover, HS is ubiquitously expressed *in vivo* and has a broader range of physiological targets than heparin, which is localized primarily to specialized granule cells. Diverse sulfation motifs are also found on CS, with sulfation occurring at each of the free hydroxyls [7]. As a result, a simple tetrasaccharide of CS has the potential to encode 256 sulfation sequences, whereas an HS tetrasaccharide, which has greater complexity due to the presence of IdoA and *N*-sulfation, can display over 2000 sulfation motifs. Although it remains to be seen whether all of these possible sulfation patterns occur *in vivo*, a large number of distinct sulfated structures have been identified to date [3,7].

On a macromolecular level, HS and CS polysaccharides exhibit various chain lengths (~10 to 100 disaccharide units) and clustered regions of high or low sulfation [3]. Structural studies have shown that glycosaminoglycans adopt helical structures whose pitch can vary with the associated counterion [8,9]. Moreover, the conformational flexibility of the pyranose ring of IdoA, which exists in

Figure 1



Representative classes of glycosaminoglycans, with potential sites of sulfation indicated. R = SO_3^- or H; R_1 = SO_3^-, H or Ac; n = ~10–100.

equilibrium between different chair and skew-boat conformations when sulfated at the C2 position, has been postulated to enhance the specificity of HS for its protein targets [3]. Thus, the combination of sequence, charge distribution, sugar conformation and three-dimensional structure endows glycosaminoglycans with rich structural diversity.

The potential 'sulfation code': implications for growth factor signaling and neuronal wiring

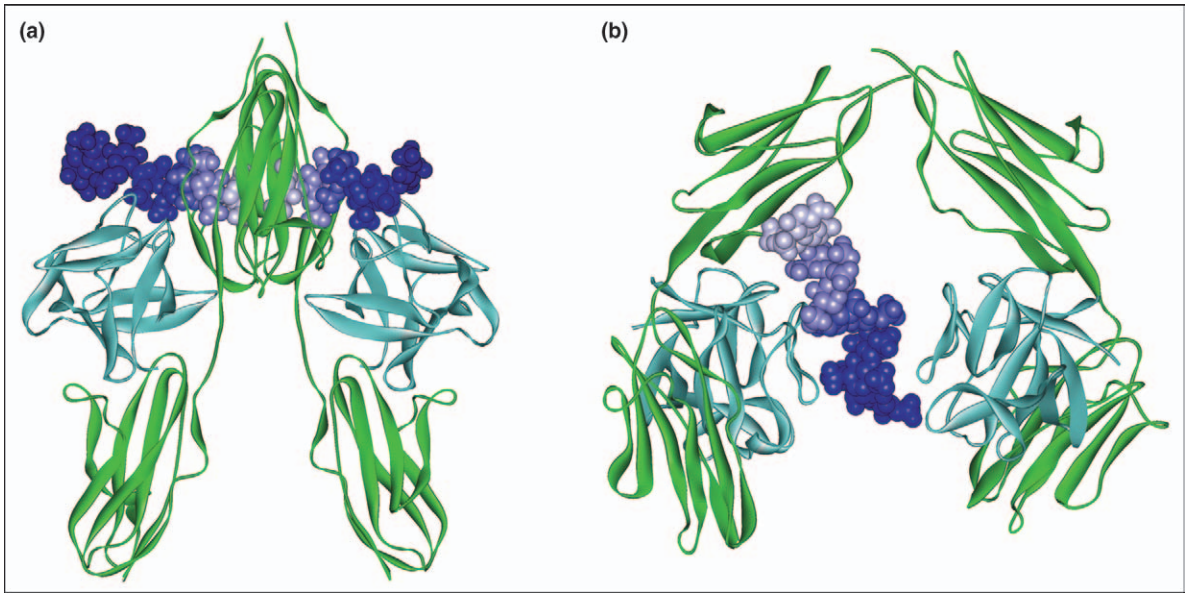
Evidence suggests that the fine structure of glycosaminoglycans is crucial for their functions *in vivo*. Genetic studies have established the importance of various sulfotransferase enzymes in the glycosaminoglycan biosynthetic pathway. For example, deletion of an HS 2-*O*-sulfotransferase gene in mice led to complete failure of kidney development, and mutation of the *N*-deacetylase–*N*-sulfotransferase gene in *Drosophila* inhibited growth factor signaling and disrupted embryonic development [5,10]. Consistent with their essential roles, the sulfation patterns of glycosaminoglycans are tightly regulated *in vivo*. Distinct sulfated forms are associated with particular tissues: for example, differentially sulfated CS motifs are localized to specific brain regions and found along axonal growth tracts [11,12]. The sulfation patterns of HS and CS are also altered during embryonic brain development, as are specific sulfotransferase activities [13,14]. Furthermore, distinct HS motifs have been linked to the development of several diseases, including the pathology of Alzheimer's disease and cancer metastasis [2,6].

The molecular mechanisms by which glycosaminoglycans contribute to these biological events are only beginning to be understood. However, many studies suggest that glycosaminoglycans coordinate complex processes by regulating the activities of growth factors and other target proteins. One of the most studied examples is the binding of HS to the fibroblast growth factors (FGFs). The FGFs comprise a large family of growth factors (23 members to date) and have been shown to have crucial roles in morphogenesis, development, angiogenesis and wound healing [15]. They induce the dimerization of FGF receptor tyrosine kinases (FGFRs), thereby activating intracellular signaling pathways. Two distinct models have been proposed to explain the essential contribution of HS to FGF–FGFR signaling. A crystal structure of the HS–FGF2–FGFR1 complex obtained by Schlessinger *et al.* [8] suggests that two ternary complexes of HS–FGF–FGFR come together with the non-reducing ends of each sugar chain facing one another upon activation (Figure 2a). By contrast, the HS–FGF1–FGFR2 structure of Pellegrini *et al.* [16] suggests that a single HS chain initiates the assembly of two FGF–FGFR complexes (Figure 2b). Although these models propose distinct roles for HS in coordinating receptor dimerization, both agree on the formation of an activated FGF–FGFR–HS complex.

Importantly, the specific sulfation pattern of HS appears to be crucial for FGF binding and assembly of the complex. Structural analyses have shown that many of the interactions between HS and the FGFs involve salt bridges and hydrogen-bonding contacts between the sulfate and carboxylate groups of the oligosaccharide with polar residues of the protein [17]. Optimal van der Waals contacts and the flexibility of HS chains further enhance the interaction [18]. Interestingly, none of the residues in the heparin-binding region, including the polar side chains, are completely conserved throughout the FGF family [17]. This raises the intriguing possibility that variations in HS sequence or sulfation pattern might specify the binding of particular FGFs, enabling the selective activation of signaling pathways. Consistent with this view, biochemical studies have shown that FGF2 requires 2-*O*-sulfation but not 6-*O*-sulfation for HS binding, whereas FGF10 has the reverse preference, and FGF1 requires both 2-*O*-sulfation and 6-*O*-sulfation [19]. Distinct sulfation preferences are also exhibited by the FGFRs; for example, 6-*O*-sulfation was required for FGFR2 IIIb but not FGFR1 activation [20]. Unfortunately, the heterogeneity of biochemical HS preparations has made it difficult to examine specific sulfation sequences. As discussed below, it is anticipated that homogeneous oligosaccharide libraries of defined sequence will provide additional insights into the importance of the sulfation code. Resolution of the code and precise activation mechanism will be crucial for understanding growth factor signaling, and may reveal common mechanistic themes utilized by both HS and CS glycosaminoglycans.

Recent studies have also uncovered striking roles for glycosaminoglycans and their sulfation patterns during

Figure 2



Crystallographic models of the FGF–FGFR–HS signaling complex. **(a)** The crystal structure of an FGF2–FGFR1–heparin complex has a 2:2:2 stoichiometry, with the non-reducing ends of each sugar facing one other [8]. **(b)** The crystal structure of an FGF1–FGFR2–heparin complex has a 2:2:1 stoichiometry, with a single sugar chain initiating complex formation [16]. Coloring: FGF (light blue), FGFR (green); heparin is shown in space-fill representation and colored from the reducing (purple) to the non-reducing (light purple) end. Reproduced from [51] with permission. Copyright 2001, Elsevier, Ltd.

neuronal development. Axons are guided to their target locations by diffusible and cell surface-bound cues that either attract or repel the growing tip of the axon. One such cue is the chemorepellent protein Slit. Slit repels axon growth upon binding to specific cell surface receptors. Studies have shown that HS is required for the interaction of Slit with its receptors. For example, removal of HS by treatment of cells with heparinase abolished Slit binding to Robo receptors [21]. Although the precise sulfation motifs have not been identified, *O*-sulfation of HS was found to be crucial for Slit binding to glypican-1 [22]. In related studies, Bülow and Hobert [4••] used genetic approaches to probe the role of HS sulfation in axon guidance. Abolishing the activity of three HS-modifying enzymes, C-5 epimerase, 2-*O*-sulfotransferase and 6-*O*-sulfotransferase, in *Caenorhabditis elegans* revealed that particular neuron types require specific HS motifs for normal growth. Some axons required all three modifying enzymes, whereas others required either C-5 epimerase or 2-*O*-sulfotransferase activity, and still other neuron types did not require any of the enzymes. These studies support the idea that distinct modifications to HS structure are essential for neuronal development and may encode instructions that guide neurons to their proper targets *in vivo*.

In all, the above studies highlight the importance of glycosaminoglycan structure in regulating crucial biological processes. The molecular diversity of glycosaminoglycans may provide a powerful means to influence complex signaling pathways *in vivo*. The spatial and temporal regulation of HS and CS modifications may facilitate or inhibit ligand–receptor interactions in a highly localized manner. With the considerable diversity that exists in glycosaminoglycan chains, the sulfation code would represent an elegant means of molecular-level control. It will be exciting to discover the extent to which Nature utilizes this potential.

Unlocking the code using chemistry

Deciphering the sulfation code will require the development of new strategies for manipulating and evaluating specific glycosaminoglycan structures. At present, there are no methods for the rapid identification of biologically active sulfation motifs. Genetic and biochemical approaches have established crucial roles for glycosaminoglycans in particular biological contexts. However, deletion of a sulfotransferase gene leads to global changes throughout the carbohydrate chain, making it difficult to pinpoint the impact of a specific structural motif. Glycosaminoglycans have also been isolated from natural sources but their structural complexity and heterogeneity are a significant limitation. The presence of multiple sulfation motifs in biochemical preparations complicates efforts to attribute a biological function to a specific sulfation motif. Moreover, studies with purified natural glycosaminoglycans are biased toward abundant, readily isolable sequences. As such, it can be difficult to study

physiologically important sulfation patterns that are present in low cellular abundance.

Chemical approaches provide a powerful solution to these challenges. Virtually any desired glycosaminoglycan structure can be generated using synthetic chemistry, with exquisite control over stereochemistry, length and pattern of sulfation. Access to homogeneous, well-characterized structures facilitates the identification of biologically active sequences and enables systematic investigations into structure–activity relationships. The ability to obtain defined glycosaminoglycan structures and related analogs should also accelerate investigations into the therapeutic potential of glycosaminoglycans, in areas such as cancer biology, neurobiology and virology.

Early work demonstrated the potential of chemical synthesis to reveal important insights into glycosaminoglycans. Choay *et al.* [23] synthesized a key sulfated pentasaccharide that interacts with antithrombin III (ATIII) to inhibit blood coagulation (Figure 3a). This compound helped to elucidate the mechanism of heparin anticoagulant activity of heparin and led to the development of the pentasaccharide drug Arixtra^{TM}, an effective alternative to low molecular weight heparin for the treatment of deep vein thrombosis.

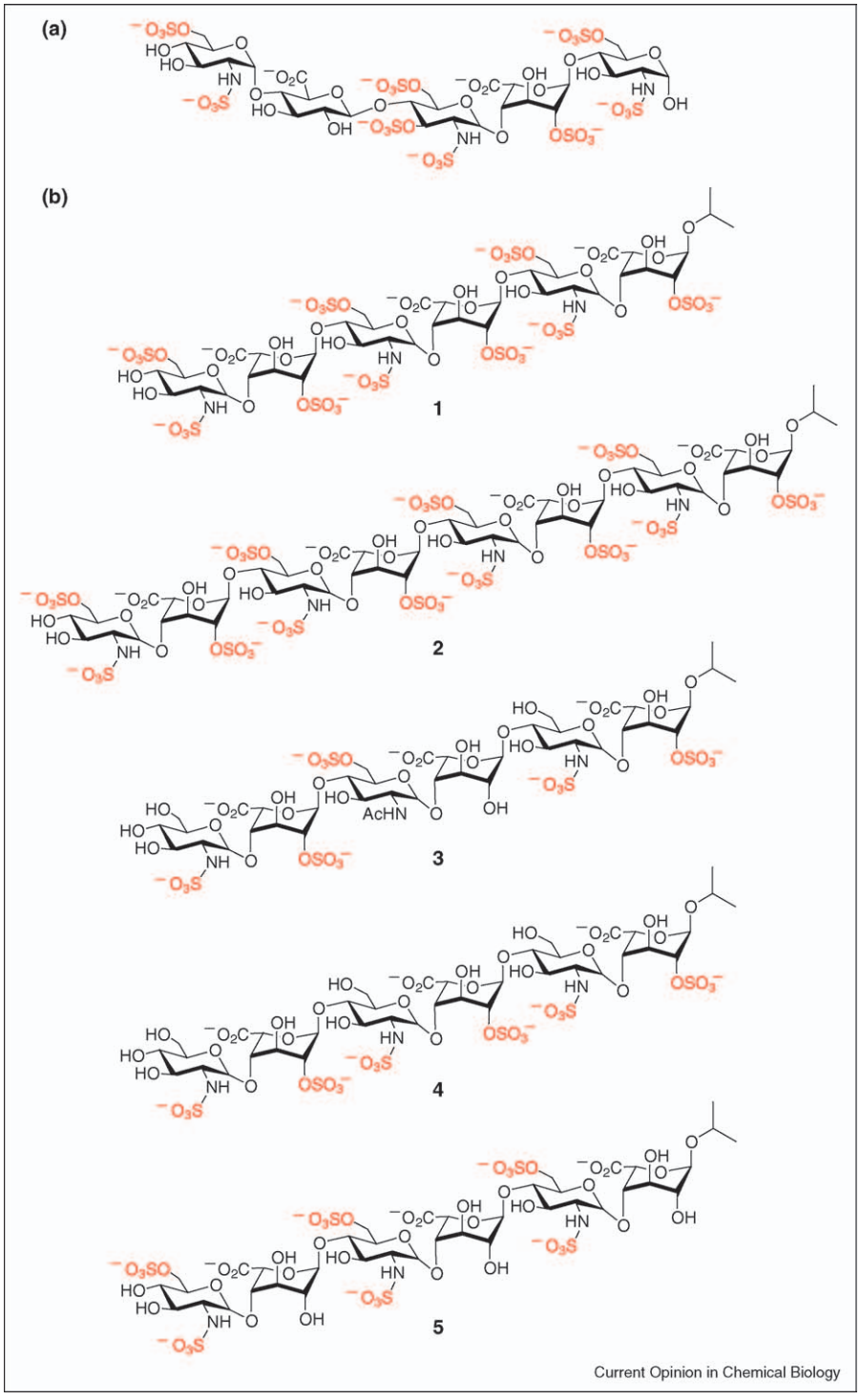
More recently, Angulo *et al.* [24••] synthesized several oligosaccharides of varying length and sulfation pattern to study the influence of HS structure on the activity of FGF1 (Figure 3b). Hexasaccharide **1** and octasaccharide **2** contain the major sulfation motif found in heparin and are predicted by molecular modeling and NMR studies to distribute negatively charged groups on both sides of the heparin helix. By contrast, hexasaccharide **3**, which has a lower overall negative charge than **1** and **2**, orientates the sulfate groups along one side of the helix. Remarkably, hexasaccharide **3** was more effective than **1** and equally effective as **2** and heparin at promoting FGF1-mediated cell proliferation. Moreover, hexasaccharides **4** and **5**, which possess anionic charge similar to that of **3** but distribute the sulfate groups along both sides of the helix, have poor mitogenic activity. These studies show that subtle changes in the sulfation pattern can modulate the biological activity of HS, and they lend further support to the notion that a precise arrangement of sulfate groups may be required for FGF1 activity.

Historically, studies of glycosaminoglycans have focused mainly on the HS and heparin class. However, recent evidence suggests that CS glycosaminoglycans also have essential roles in biological processes such as neuronal growth and regeneration. Tully *et al.* [25••] have reported the first biological studies using well-defined CS molecules. A tetrasaccharide (**6**) bearing a specific sulfation motif, CS-E, was synthesized and shown to stimulate the outgrowth of embryonic hippocampal neurons (Figure 4). For these studies, neurons were grown on CS-coated glass surfaces to mimic CS proteoglycans on the cell surface or in the extracellular matrix. Whereas the CS-E tetrasaccharide promoted neurite outgrowth by 40%, the corresponding disaccharide (**7**) and the unsulfated CS tetrasaccharide (**8**) had no stimulatory activity. These findings establish that sulfation is required for the growth-promoting properties of CS and define a tetrasaccharide as a minimum functional domain. Recently, our laboratory also has shown that altering the precise sulfation pattern of the CS tetrasaccharide has a significant impact on its biological activity (Gama *et al.*, unpublished data). Together, these studies provide the first direct demonstration that CS activity can be controlled at a molecular level through its sulfation patterns. Synthetic CS oligosaccharides are currently being exploited to elucidate the molecular mechanisms underlying CS-E activity, including the role of multivalency and the proteins and signaling pathways activated by CS-E.

The ability of synthetic oligosaccharides to recapitulate the activities of the natural polysaccharides has opened the possibility of generating simplified analogs that share the high charge density of glycosaminoglycans and their ability to modulate proteins. In the case of the heparin–ATIII interaction, minor perturbations to the carbohydrate structure are tolerated. For example, replacement of the *N*-sulfonate groups with *O*-sulfonate groups, methylation of the free hydroxyls, or introduction of a carbon-based interglycosidic bond between two residues in the pentasaccharide preserved or enhanced the affinity of the pentasaccharide for ATIII [26•,27]. More drastic changes to the carbohydrate scaffold have also been explored, including the development of peptides that mimic the nonsulfated glycosaminoglycan hyaluronan; dextran polymers modified with carboxymethyl and sulfate groups; polyacrylamide-based polymers; dendritic polyglycerol sulfate structures; and nonsugar-based, sulfated small molecules [28–31]. These variants show promise at mimicking the activities of glycosaminoglycans, although the selectivity of the molecules toward particular biological targets remains to be demonstrated.

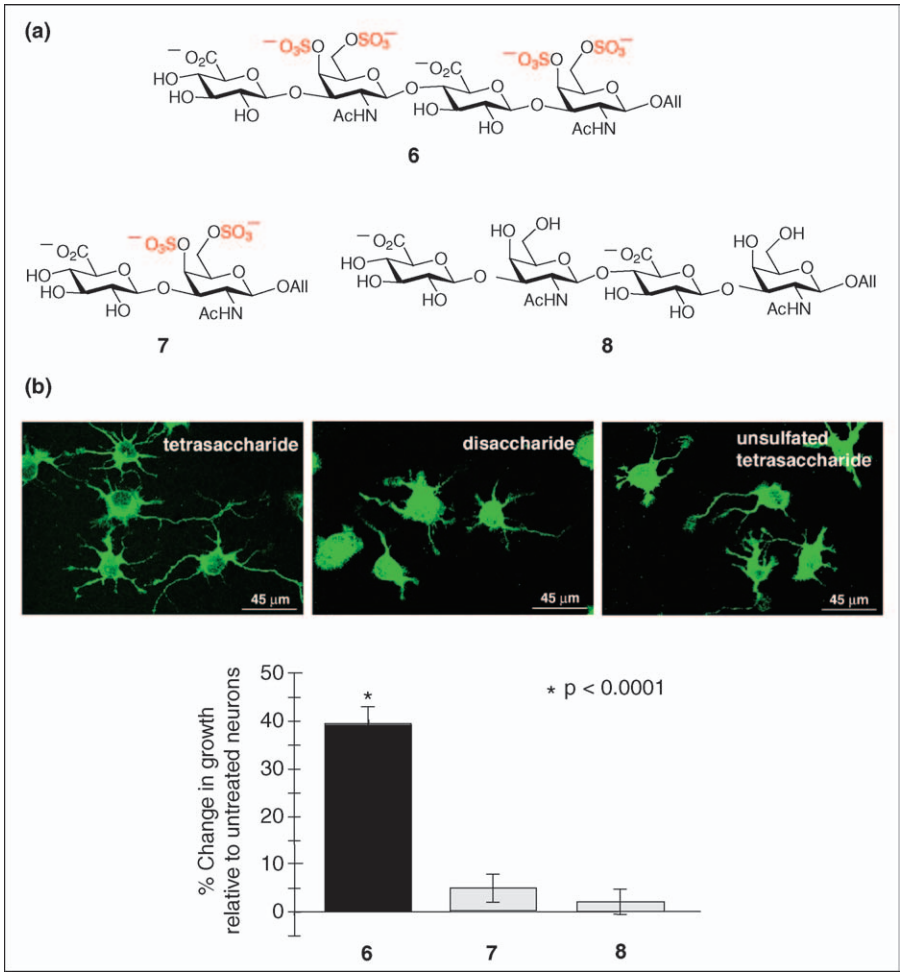
Elegant work has been reported by Petitou and van Boeckel [26•] on heparin-like analogs for the inhibition of thrombin. In addition to modulating ATIII, heparin exerts its anticoagulant activity by targeting thrombin. The mechanism involves bridging thrombin and ATIII in a ternary complex and requires a minimum of 15 saccharide units. Studies suggest that heparin attracts thrombin electrostatically and then guides thrombin to a position where it complexes with ATIII. Effective inhibitors were generated by joining the ATIII-binding pentasaccharide to thrombin-binding oligosaccharides using neutral linkers, including flexible polyethylene glycol or rigid polyglucose spacers (Figure 5). In contrast to the highly specific binding requirements of ATIII, persulfated di-

Figure 3



Synthetic glycosaminoglycans of defined structure. **(a)** The pentasaccharide synthesized by Choay *et al.*, which selectively recognizes ATIII [23]. **(b)** Hexasaccharide **3** and octasaccharide **2** are more effective at promoting FGF1-mediated cell proliferation than hexasaccharides **1**, **4** and **5**. Hexasaccharide **3** was proposed by molecular modeling studies to display the sulfate groups along one side of the HS helix and thus to stimulate FGF1 activity [24••].

Figure 4



A synthetic chondroitin sulfate tetrasaccharide bearing a specific sulfation motif, CS-E, stimulates the outgrowth of hippocampal neurons [25••].
(a) Structures of the CS-E tetrasaccharide (**6**), CS-E disaccharide (**7**) and unsulfated tetrasaccharide (**8**) compared in the study.
(b) Immunofluorescence images of neurons after treatment with the indicated compound. The number of neurites emanating from the cell body was enhanced, and the growth of the major extension was stimulated by 39.3 ± 3.6% relative to the poly-DL-ornithine control. In contrast, sulfated disaccharide **7** and unsulfated tetrasaccharide **8** had no significant effect on neuronal outgrowth. Copyright 2004, The American Chemical Society.

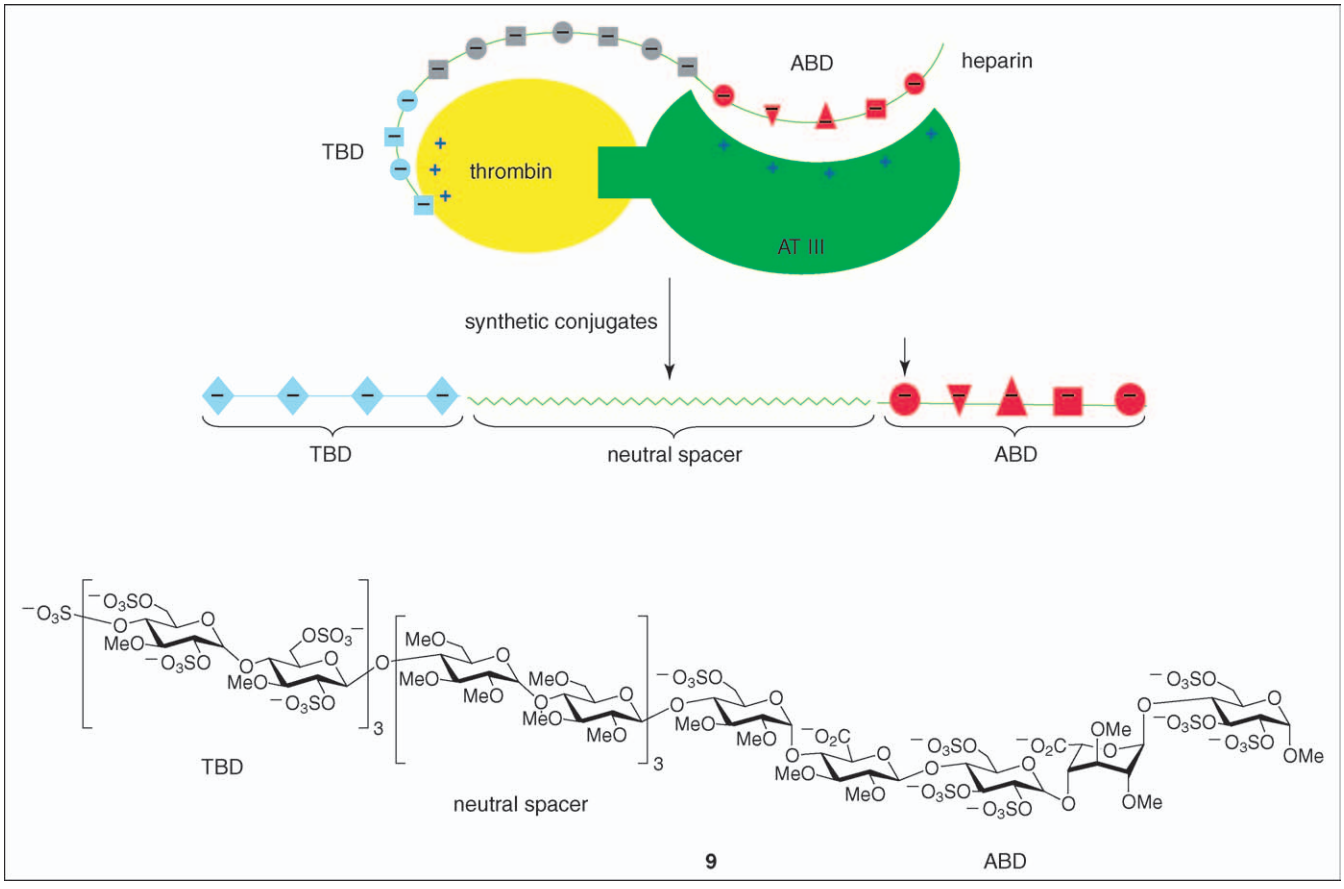
and trisaccharides, oligodeoxythymidine nucleotides or phosphate monoesters were sufficient to attract thrombin by mimicking the charge density of heparin [26•,32]. Replacement of the heparin domains with these simpler mimics prevented nonspecific interactions with platelet factor 4 and thus may eliminate thrombocytopenia, a major undesirable side effect of heparin therapy. Importantly, these studies show that portions of the heparin backbone can be altered in certain cases, without compromising biological activity.

Harnessing the biosynthetic machinery

The ability to create a wide range of chemical architectures has provided key insights into the structural determinants and mechanisms of glycosaminoglycans. Despite the continued development of new synthetic methodologies, however, the synthesis of HS and CS oligosaccharides remains challenging. The chemical complexity of glycosaminoglycan structures necessitates multiple chemical steps and the development of sophisticated protecting group strategies and stereo- and regiocontrolled glycosylation reactions. In response to these challenges, several groups have co-opted enzymes for the biosynthesis of glycosaminoglycans.

Kuberan *et al.* [33••] assembled an ATIII-binding pentasaccharide in only six steps, using enzymes from the HS biosynthetic pathway (Figure 6a). These studies built upon extensive previous work on the cloning, expression and characterization of various enzymes in the pathway [34]. A bacterial polysaccharide resembling the unmodified HS chain was incubated with *N*-deacetylase–*N*-sulfotransferase 2 (NDST2) and 3′-phosphoadenosine 5′-phosphosulfate (PAPS), followed by partial cleavage with

Figure 5

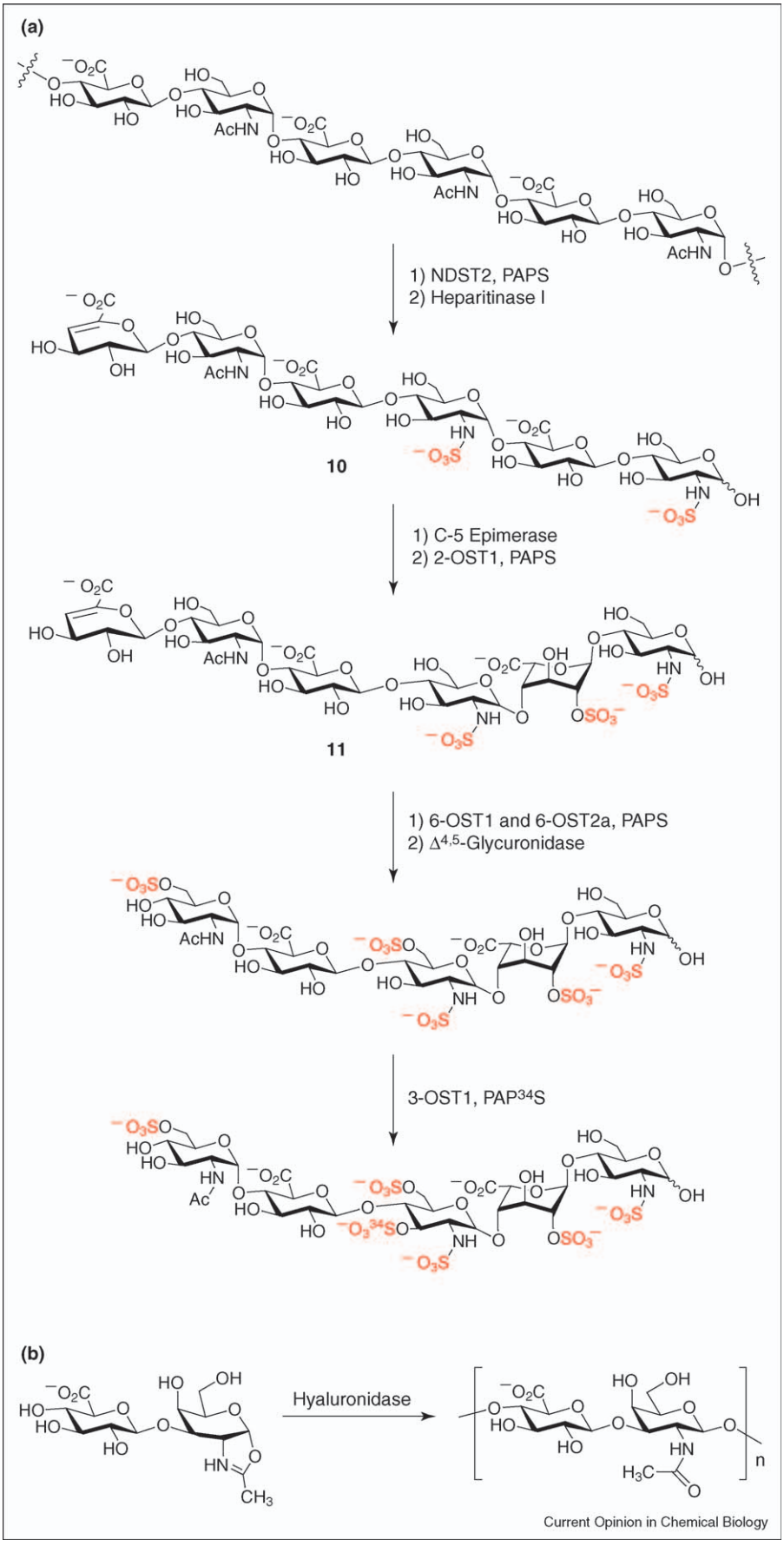


Heparin-like analogues such as **9** were designed to inhibit thrombin by bridging ATIII and thrombin in a ternary complex. Compound **9** joins a negatively charged ATIII-binding pentasaccharide (ABD, ATIII-binding domain) with a thrombin-binding oligosaccharide (TBD, thrombin-binding domain) through a neutral polyglucose spacer. The compound displayed potent thrombin inhibition while avoiding interaction with platelet factor 4. Platelet factor 4 in complex with heparin can be immunogenic and induce thrombocytopenia. Adapted from [26•] with permission. Copyright 2004, Wiley.

heparitinase to produce hexasaccharide **10**. The hexasaccharide was then sequentially treated with C-5 epimerase, which acts only on GlcA residues flanked by *N*-sulfoglucosamine units, and *2-O*-sulfotransferase 1 (2-OST1), which preferentially sulfates the IdoA residues at the reducing end of *N*-sulfoglucosamine, to give hexasaccharide **11**. Treatment with 6-OST1 and 6-OST2a was followed by $\Delta^{4,5}$-glycuronidase to remove selectively the terminal unsaturated uronic acid residue. The crucial 3-*O*-sulfation step was performed in the final reaction using 3-OST1 to produce the ATIII-binding pentasaccharide in microgram quantities. In the future, it will be exciting to examine whether these approaches are scalable and can be generalized to other sulfation motifs.

Biosynthetic approaches to sulfated CS oligosaccharides have yet to be reported, although many enzymes in the biosynthetic pathway have been cloned and expressed [34]. Recently, Kitagawa *et al.* [35] demonstrated an *in vitro* polymerization reaction to generate unmodified CS using a recombinant chondroitin synthase enzyme in conjunction with a chondroitin polymerizing factor protein. CS chains of ~100 units in length were produced on an analytical scale using a tetrasaccharide acceptor, UDP-GlcA and UDP-GalNAc sugars. An alternative approach capitalizes on the ability of hyaluronidases to function in reverse. Kobayashi *et al.* [36•] devised a clever strategy that utilizes disaccharide oxazoline derivatives and commercially available hyaluronidase enzymes to generate hyaluronan and unmodified CS in milligram quantities (Figure 6b). The oxazoline is proposed to serve as a transition state analog and thus is readily recognized and activated by the enzyme. When an *N*-acetylchondrosine oxazoline derivative is provided as a substrate and hyaluronidase is used at suboptimal pH (neutral versus acidic), the enzymatic reaction runs in reverse to generate CS chains of average molecular weight 2500 in a 35% yield. Although exploiting biosynthetic enzymes to access defined glycosaminoglycan structures is still a relatively young field compared with carbohydrate synthesis, the

Figure 6



Biosynthesis of **(a)** an ATIII-binding pentasaccharide [33••] and **(b)** unmodified chondroitin sulfate of mean molecular weight 2500 [36•].

approach holds promise for the rapid generation of diverse structures.

Technology development: analysis and sequencing of glycosaminoglycans

With growing interest in understanding the fine structure of glycosaminoglycans, there is an increasing demand for new technologies to analyze glycosaminoglycans and study their interactions with proteins. Several excellent methods to characterize glycosaminoglycan–protein interactions have been developed, including gel mobility shift assays [37], filter-binding assays [38,39], isothermal titration calorimetry (ITC) [40] and surface-plasmon resonance (SPR) [40]. SPR and ITC are most commonly used for quantitative analyses. Kinetic on–off rates can be measured using SPR, whereas ITC affords kinetic dissociation constants, thermodynamic parameters and binding stoichiometry. Gel mobility shift assays have been particularly valuable for studying the formation of higher-order protein complexes. For example, the ternary complex of HS, FGF1 and FGFR1 was readily distinguishable from HS–FGF1 and HS–FGFR1 binary complexes based on their relative mobility in native gels [37]. Interestingly, ternary complex formation was induced by oligosaccharides as small as tetrasaccharides, and the molar ratio of HS in the ternary complex was suggested to be 1:1:1, consistent with the model proposed by Schlessinger *et al.* (Figure 2a). Despite numerous methods, high-throughput approaches for the rapid screening of glycosaminoglycan–protein interactions are generally lacking. In this regard, it is anticipated that recent developments in small molecule and protein microarrays may provide effective solutions [41,42].

Advances in mass spectrometry analysis have also enabled significant developments in the study of glycosaminoglycans. The standard method for characterizing glycosaminoglycans involves chemical or enzymatic digestion of the polysaccharide chains and HPLC analysis to determine the percent composition of various sulfated disaccharides [13]. Unfortunately, the linear sequence of the polysaccharide is lost in this process. By contrast, matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and electrospray ionization mass spectrometry can provide important structural information on heparin, HS and CS oligosaccharides, including elemental composition, molecular weight and sequence information [43–47]. The complex fragmentations that arise upon tandem mass spectrometry can be used in certain cases to determine the position of sulfate groups and distinguish between IdoA and GlcA epimers [44,45].

To date, purified HS oligosaccharide fragments have been successfully sequenced. One promising methodology utilizes exosulfatases and exoglycosidases to remove specific sulfates and monosaccharides from the non-reducing end of the oligosaccharide [48,49]. The resulting product is then analyzed by polyacrylamide gel electrophoresis or HPLC. Using the various exoenzymes, the glycosaminoglycan sequence can be iteratively read from the non-reducing end. This approach works for most HS sulfation sequences and has been applied to purified decasaccharides. Another approach, called property-encoded nomenclature-MALDI, involves a combination of MALDI-MS, compositional analysis and enzymatic or chemical degradation to deduce the sequence of the HS oligosaccharide [46,47]. First, MALDI-MS is used to determine the length and total number of sulfates and acetates present in the oligosaccharide. Compositional analysis then establishes the number and type of disaccharide building blocks. Based on this information, a computer generates a master list of all possible sequences. The masses of fragments obtained by enzymatic digestion or chemical degradation are subsequently used to eliminate sequences from the master list until a unique solution emerges. The property-encoded nomenclature-MALDI strategy has been applied successfully to several purified oligosaccharides and can be integrated with other analytical methods such as NMR and capillary electrophoresis [50].

Conclusions and future directions

A molecular-level understanding of glycosaminoglycans has begun to emerge, which suggests the importance of fine structure in controlling the biological properties of glycosaminoglycans. Future developments in this field will require well-defined structures to facilitate the identification of biologically active sequences and investigations into structure–activity relationships. Synthetic organic and biosynthetic approaches will be valuable in this regard, providing access to molecules with defined stereochemistry, length and patterns of sulfation. In the future, it should be possible to create libraries of oligosaccharide structures to enable exploration of the sequence and functional space of glycosaminoglycans. The identification of discrete, biologically active sequences should also enable the assembly of increasingly complex molecular architectures, in which crucial features of the natural polysaccharide, such as multivalency and distance between active motifs, can be controlled and evaluated. Defined structures should also allow for the generation of selective antibodies for monitoring the spatiotemporal expression of specific sulfation motifs in biological systems. These efforts, combined with structural and computational studies of glycosaminoglycan–protein interactions, should lead to a more comprehensive understanding of the mechanisms by which glycosaminoglycans contribute to diverse biological processes. Finally, new technologies for analyzing glycosaminoglycans and their molecular interactions will continue to have a significant impact on the field. With the enormous chemical potential of glycosaminoglycans, it will be exciting to discover the extent to which the structure of glycosaminoglycans is used to encode information, similar to DNA and the other biopolymers.

Acknowledgements

We thank the Hsieh-Wilson laboratory for helpful discussions. This research was supported by the Human Frontiers Science Program (RGY0072), National Institutes of Health (RO1 NS045061) and Alfred P. Sloan Foundation.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Laabs T, Carulli D, Geller HM, Fawcett JW: **Chondroitin sulfate proteoglycans in neural development and regeneration**. *Curr Opin Neurobiol* 2005, **15**:116-120. [Erratum: *Curr Opin Neurobiol* 2005, **15**:252.]

2. Sasisekharan R, Shriver Z, Venkataraman G, Narayanasami U: **Roles of heparan-sulphate glycosaminoglycans in cancer**. *Nat Rev Cancer* 2002, **2**:521-528.

3. Capila I, Linhardt RJ: **Heparin-protein interactions**. *Angew Chem Int Ed Engl* 2002, **41**:391-412.

4. •• Bülow HE, Hobert O: **Differential sulfations and epimerization define heparan sulfate specificity in nervous system development**. *Neuron* 2004, **41**:723-736.
The authors generated *C. elegans* mutants that lacked three HS-modifying enzymes (glucuronyl C-5 epimerase, 6-*O*-sulfotransferase, 2-*O*-sulfotransferase) to study the effects of HS modification on nervous system development. Analysis of the mutants revealed defects in neuronal development, in which distinct neuron types required specific HS modifications for normal growth. Moreover, the authors were able to link the effects of single or multiple mutations to specific signaling pathways.

5. Merry CL, Bullock SL, Swan DC, Backen AC, Lyon M, Beddington RS, Wilson VA, Gallagher JT: **The molecular phenotype of heparan sulfate in the Hs2st−/− mutant mouse**. *J Biol Chem* 2001, **276**:35429-35434.

6. Li JP, Escobar Galvis ML, Gong F, Zhang X, Zcharia E, Metzger S, Vlodavsky I, Kisilevsky R, Lindahl U: ***In vivo* fragmentation of heparan sulfate by heparanase overexpression renders mice resistant to amyloid protein A amyloidosis**. *Proc Natl Acad Sci USA* 2005, **102**:6473-6477.

7. Kinoshita-Toyoda A, Yamada S, Haslam SM, Khoo KH, Sugiura M, Morris HR, Dell A, Sugahara K: **Structural determination of five novel tetrasaccharides containing 3-*O*-sulfated D-glucuronic acid and two rare oligosaccharides containing a β-D-glucose branch isolated from squid cartilage chondroitin sulfate E**. *Biochemistry* 2004, **43**:11063-11074.

8. Schlessinger J, Plotnikov AN, Ibrahimi OA, Eliseenkova AV, Yeh BK, Yayon A, Linhardt RJ, Mohammadi M: **Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization**. *Mol Cell* 2000, **6**:743-750.

9. Millane RP, Mitra AK, Arnott S: **Chondroitin 4-sulfate: comparison of the structures of the potassium and sodium salts**. *J Mol Biol* 1983, **169**:903-920.

10. 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.

11. Fernaud-Espinosa I, Nieto-Sampedro M, Bovolenta P: **Developmental distribution of glycosaminoglycans in embryonic rat brain: relationship to axonal tract formation**. *J Neurobiol* 1996, **30**:410-424.

12. Properzi F, Carulli D, Asher RA, Muir E, Camargo LM, van Kuppevelt TH, ten Dam GB, Furukawa Y, Mikami T, Sugahara K *et al.*: **Chondroitin 6-sulphate synthesis is up-regulated in injured CNS, induced by injury-related cytokines and enhanced in axon-growth inhibitory glia**. *Eur J Neurosci* 2005, **21**:378-390.

13. Brickman YG, Ford MD, Gallagher JT, Nurcombe V, Bartlett PF, Turnbull JE: **Structural modification of fibroblast growth factor-binding heparan sulfate at a determinative stage of neural development**. *J Biol Chem* 1998, **273**:4350-4359.

14. Kitagawa H, Tsutsumi K, Tone Y, Sugahara K: **Developmental regulation of the sulfation profile of chondroitin sulfate chains in the chicken embryo brain**. *J Biol Chem* 1997, **272**:31377-31381.

15. Ornitz DM, Itoh N: **Fibroblast growth factors**. *Genome Biol* 2001, **2**:1-12.

16. Pellegrini L, Burke DF, von Delft F, Mulloy B, Blundell TL: **Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin**. *Nature* 2000, **407**:1029-1034.

17. Faham S, Linhardt RJ, Rees DC: **Diversity does make a difference: fibroblast growth factor-heparin interactions**. *Curr Opin Struct Biol* 1998, **8**:578-586.

18. Raman R, Venkataraman G, Ernst S, Sasisekharan V, Sasisekharan R: **Structural specificity of heparin binding in the fibroblast growth factor family of proteins**. *Proc Natl Acad Sci USA* 2003, **100**:2357-2362.

19. Ashikari-Hada S, Habuchi H, Kariya Y, Itoh N, Reddi AH, Kimata K: **Characterization of growth-factor binding structures in heparin/heparan sulfate using an octasaccharide library**. *J Biol Chem* 2004, **279**:12346-12354.

20. Ostrovsky O, Berman B, Gallagher J, Mulloy B, Fernig DG, Delehedde M, Ron D: **Differential effects of heparin saccharides on the formation of specific fibroblast growth factor (FGF) and FGF receptor complexes**. *J Biol Chem* 2002, **277**:2444-2453.

21. Hu H: **Cell-surface heparan sulfate is involved in the repulsive guidance activities of Slit2 protein**. *Nat Neurosci* 2001, **4**:695-701.

22. Ronca F, Andersen JS, Paech V, Margolis RU: **Characterization of Slit protein interactions with glypican-1**. *J Biol Chem* 2001, **276**:29141-29147.

23. Choay J, Petitou M, Lormeau JC, Sinaÿ P, Casu B, Gatti G: **Structure-activity relationship in heparin: a synthetic pentasaccharide with high affinity for anti-thrombin-III and eliciting high anti-factor Xa activity**. *Biochem Biophys Res Commun* 1983, **116**:492-499.

24. •• Angulo J, Ojeda R, de Paz JL, Lucas R, Nieto PM, Lozano RM, Redondo-Horcajo M, Giménez-Gallego G, Martín-Lomas M: **The activation of fibroblast growth factors (FGFs) by glycosaminoglycans: influence of the sulfation pattern on the biological activity of FGF-1**. *Chembiochem* 2004, **5**:55-61.
The authors use synthetic heparin-like oligosaccharides to investigate the effects of HS sulfation on FGF1 signaling and dimerization. This study emphasizes the role of the sulfate groups in the context of the helical structure of HS and suggests that a specific arrangement of sulfate groups is important for FGF–FGFR activation.

25. •• Tully SE, Mabon R, Gama CI, Tsai SM, Liu XW, Hsieh-Wilson LC: **A chondroitin sulfate small molecule that stimulates neuronal growth**. *J Am Chem Soc* 2004, **126**:7736-7737.
This study reports the first biological investigations using homogeneous, defined CS molecules. The authors demonstrate that a specific sulfation motif, CS-E, stimulates the outgrowth of cultured hippocampal neurons and define a tetrasaccharide as a minimum function domain. The findings raise the possibility of using synthetic small molecules to understand the CS sulfation code and its roles in neuronal development and injury.

26. • Petitou M, van Boeckel CA: **A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next?** *Angew Chem Int Ed Engl* 2004, **43**:3118-3133.
This review discusses the development of synthetic HS pentasaccharide analogs as antithrombotic drugs and the design of analogs with longer-lasting activity, including novel conjugates and long oligosaccharides with anti-factor Xa and antithrombin activities.

27. Petitou M, Hérault JP, Lormeau JC, Helmboldt A, Mallet JM, Sinaÿ P, Herbert JM: **Introducing a C-interglycosidic bond in a biologically active pentasaccharide hardly affects its biological properties**. *Bioorg Med Chem* 1998, **6**:1509-1516.

28. Ziebell MR, Zhao ZG, Luo B, Luo Y, Turley EA, Prestwich GD: **Peptides that mimic glycosaminoglycans: high-affinity ligands for a hyaluronan binding domain**. *Chem Biol* 2001, **8**:1081-1094.

29. Monien BH, Desai UR: **Antithrombin activation by nonsulfated, non-polysaccharide organic polymer**. *J Med Chem* 2005, **48**:1269-1273.

30. Guan R, Sun XL, Hou S, Wu P, Chaikof EL: **A glycopolymer chaperone for fibroblast growth factor-2**. *Bioconjug Chem* 2004, **15**:145-151.

31. Gunnarsson GT, Desai UR: **Exploring new non-sugar sulfated molecules as activators of antithrombin**. *Bioorg Med Chem Lett* 2003, **13**:679-683.

32. Buijsman RC, Basten JEM, Dreef-Tromp CM, van der Marel GA, van Boeckel CAA, van Boom JH: **Synthesis of heparin-like antithrombotics having perphosphorylated thrombin binding domains**. *Bioorg Med Chem* 1999, **7**:1881-1890.

33. •• Kuberan B, Lech MZ, Beeler DL, Wu ZL, Rosenberg RD: **Enzymatic synthesis of antithrombin III-binding heparan sulfate pentasaccharide**. *Nat Biotechnol* 2003, **21**:1343-1346.
The authors present the first enzymatic synthesis of an ATIII-binding HS pentasaccharide. A single, defined structure was obtained in six steps on a microgram-scale, raising the possibility of harnessing the biosynthetic machinery to access various oligosaccharide structures of defined size and sequence.

34. Kusche-Gullberg M, Kjellén L: **Sulfotransferases in glycosaminoglycan biosynthesis**. *Curr Opin Struct Biol* 2003, **13**:605-611.

35. Kitagawa H, Izumikawa T, Uyama T, Sugahara K: **Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization**. *J Biol Chem* 2003, **278**:23666-23671.

36. • Kobayashi S, Itoh R, Morii H, Fujikawa SI, Kimura S, Ohmae M: **Synthesis of glycosaminoglycans via enzymatic polymerisation**. *J Polym Sci [A]* 2003, **41**:3541-3548.
The authors show that oxazoline derivatives with β-D-glucuronyl-(1→3)-*N*-acetyl-D-glucosamine or β-D-glucuronyl-(1→3)-*N*-acetyl-D-galactosamine structures can be polymerized by hyaluronidases to generate hyaluronic acid (MW 13500, 52% yield) or unmodified chondroitin sulfate (MW 2500, 35% yield), respectively. The enzyme-catalyzed ring-opening occurred with complete regio- and stereoselectivity.

37. Wu ZL, Zhang L, Yabe T, Kuberan B, Beeler DL, Love A, Rosenberg RD: **The involvement of heparan sulfate (HS) in FGF1/HS/FGFR1 signaling complex**. *J Biol Chem* 2003, **278**:17121-17129.

38. 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.

39. Loo BM, Kreuger J, Jalkanen M, Lindahl U, Salmivirta M: **Binding of heparin/heparan sulfate to fibroblast growth factor receptor 4**. *J Biol Chem* 2001, **276**:16868-16876.

40. Dong J, Peters-Libeu CA, Weisgraber KH, Segelke BW, Rupp B, Capila I, Hernáiz MJ, LeBrun LA, Linhardt RJ: **Interaction of the N-terminal domain of apolipoprotein E4 with heparin**. *Biochemistry* 2001, **40**:2826-2834.

41. Feizi T, Fazio F, Chai W, Wong CH: **Carbohydrate microarrays – a new set of technologies at the frontiers of glycomics**. *Curr Opin Struct Biol* 2003, **13**:637-645.

42. LaBaer J, Ramachandran N: **Protein microarrays as tools for functional proteomics**. *Curr Opin Chem Biol* 2005, **9**:14-19.

43. Thanawiroon C, Rice KG, Toida T, Linhardt RJ: **Liquid chromatography/mass spectrometry sequencing approach for highly sulfated heparin-derived oligosaccharides**. *J Biol Chem* 2004, **279**:2608-2615.

44. Zaia J, Li XQ, Chan SY, Costello CE: **Tandem mass spectrometric strategies for determination of sulfation positions and uronic acid epimerization in chondroitin sulfate oligosaccharides**. *J Am Soc Mass Spectrom* 2003, **14**:1270-1281.

45. Saad OM, Leary JA: **Delineating mechanisms of dissociation for isomeric heparin disaccharides using isotope labeling and ion trap tandem mass spectrometry**. *J Am Soc Mass Spectrom* 2004, **15**:1274-1286.

46. Venkataraman G, Shriver Z, Raman R, Sasisekharan R: **Sequencing complex polysaccharides**. *Science* 1999, **286**:537-542.

47. Keiser N, Venkataraman G, Shriver Z, Sasisekharan R: **Direct isolation and sequencing of specific protein-binding glycosaminoglycans**. *Nat Med* 2001, **7**:123-128.

48. Turnbull JE, Hopwood JJ, Gallagher JT: **A strategy for rapid sequencing of heparan sulfate and heparin saccharides**. *Proc Natl Acad Sci USA* 1999, **96**:2698-2703.

49. Merry CLR, Lyon M, Deakin JA, Hopwood JJ, Gallagher JT: **Highly sensitive sequencing of the sulfated domains of heparan sulfate**. *J Biol Chem* 1999, **274**:18455-18462.

50. Guerrini M, Raman R, Venkataraman G, Torri G, Sasisekharan R, Casu B: **A novel computational approach to integrate NMR spectroscopy and capillary electrophoresis for structure assignment of heparin and heparan sulfate oligosaccharides**. *Glycobiology* 2002, **12**:713-719.

51. Pellegrini L: **Role of heparan sulfate in fibroblast growth factor signalling: a structural view**. *Curr Opin Struct Biol* 2001, **11**:629-634.