ORIGINAL ARTICLE

An integrated approach using orthogonal analytical techniques to characterize heparan sulfate structure

Daniela Beccati[1] · Miroslaw Lech[1] · Jennifer Ozug[1] · Nur Sibel Gunay[1] · Jing Wang[1] ·
Elaine Y. Sun[1] · Joël R Pradines[1] · Victor Farutin[1] · Zachary Shriver[1] ·
Ganesh V. Kaundinya[1] · Ishan Capila[1]

Abstract Heparan sulfate (HS), a glycosaminoglycan present on the surface of cells, has been postulated to have important roles in driving both normal and pathological physiologies. The chemical structure and sulfation pattern (domain structure) of HS is believed to determine its biological function, to vary across tissue types, and to be modified in the context of disease. Characterization of HS requires isolation and purification of cell surface HS as a complex mixture. This process may introduce additional chemical modification of the native residues. In this study, we describe an approach towards thorough characterization of bovine kidney heparan sulfate (BKHS) that utilizes a variety of orthogonal analytical techniques (*e.g.* NMR, IP-RPHPLC, LC-MS). These techniques are applied to characterize this mixture at various levels including composition, fragment level, and overall chain properties. The combination of these techniques in many instances provides orthogonal views into the fine structure of HS, and in other instances provides overlapping / confirmatory information from different perspectives. Specifically, this approach enables quantitative determination of natural and modified saccharide residues in the HS chains, and identifies unusual structures. Analysis of partially digested HS chains allows for a better understanding of the domain structures within this mixture, and yields specific insights into the non-reducing end and reducing end structures of the chains. This approach outlines a useful framework that can be applied to elucidate HS structure and thereby provides means to advance understanding of its biological role and potential involvement in disease progression. In addition, the techniques described here can be applied to characterization of heparin from different sources.

Introduction

Heparan sulfate (HS) is a member of the glycosaminoglycan family, ubiquitously found at the surface of mammalian cells, where it plays a role in mediating the cell's interaction with its extracellular environment. HS is generally present as a proteoglycan (HSPG), *i.e.* covalently attached to a core protein, and it is involved in different biological processes, including cell-cell communication, binding of lipoprotein lipases and lipoproteins to cell surfaces, regulation of growth factors and cytokines, and binding to antithrombin to support blood flow across the vascular endothelium [1–5]. HS also has a regulatory role in pathological conditions, such as inflammation, tumor onset, progression, and metastasis [6, 7].

Heparan sulfate is a polydisperse mixture of linear polysaccharides consisting of repeating units of 1 → 4-linked glucosamine and uronic acids that can be variably sulfated during biosynthesis. Heparan sulfate biosynthesis is a complex process that starts with the action of D-glucuronyl and N-acetyl-D-glucosaminyl transferase to generate a copolymer of alternating H_{NAc} and G, which subsequently undergoes modifications by at least four families of sulfotransferases and one epimerase. Enzyme specificity during HS biosynthesis

Daniela Beccati and Miroslaw Lech contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10719-016-9734-7) contains supplementary material, which is available to authorized users.

✉ Ishan Capila
icapila@momentapharma.com

1 Momenta Pharmaceuticals Inc., 675 West Kendall Street, Cambridge, MA 02142, USA

dictates the positioning of sulfate groups; also, the expression levels of different isoforms of HS biosynthetic enzymes contribute to the synthesis of specific saccharide sequences characteristic of each cell type [8].

Despite the significant potential for diversity in HS, the concerted action of the biosynthetic machinery seemingly generates putative block structures [9, 10]. In this model, HS is thought to be composed of NA domains, which are dominated by repeating units of alternating G and H_{NAc}, and NS domains, constituted of I_{2S} and H_{NS} residues. Mixed or transition domains that alternate sulfated and non-sulfated monosaccharides separate the NA domains from the NS domains [11, 12].

Due to their higher charge density, the NS and transition domains constitute the preferred protein-binding regions of HS [13–15]. The major role of the NA domains is to confer structural flexibility to the HS backbone and facilitate interaction of more sulfated regions of HS with proteins [16]. NA domains are believed to lack specific binding properties, although heparanases seem to use differences in sulfate content between NS and NA domains to orient themselves and reach the sites of cleavage [17].

In addition, specific HS structures and sequences have been identified as critical for interaction with antithrombin, selectin, FGF-1 and FGF-2 [18, 19].

Many HS-protein interactions are believed to depend on the overall organization and spacing of NS domains, which regulate oligomerization or formation of ternary complexes [20, 21]. So far, little is known about the distribution of NS domains along or at the non-reducing end of the HS chains, while several studies indicate that NA sequences are located close to the protein-linkage region [22, 23].

A detailed understanding of the structure-function relationship for HS would require elucidation of fine structure, overall composition, spacing, and distribution of NS, NA, and transition domains along the HS chain. In this paper, multiple analytical tools are applied to provide a comprehensive understanding of bovine kidney HS (BKHS) structure. First, the overall disaccharide composition of BKHS is examined and quantified using orthogonal analytical methods (2D NMR on intact chains, and enzymatic digestion followed by IP-RPHPLC and LC-MS). Identification of minor residues such as epoxide, galacturonic acid, non-sulfated glucosamine, and glucuronic acid at the non-reducing end of the HS chain is described in detail.

Next, BKHS is separated into chains of different length to identify a potential relationship between degrees of polymerization and distribution of NA, NS, and transition domains. Finally, BKHS is digested with *heparinases* (*Hep*) I or III, and the digested mixture is analyzed by gel permeation chromatography-mass spectrometry (GPC-MS) to allow characterization of NA, NS, and transition domains both in terms of chain length and finer structure. The results of GPC-MS,

LC-MS and NMR techniques are used to perform structural analysis of non-reducing ends and reducing ends of HS chains, and to identify unusual structures, such as phosphorylated residues within the linkage region.

This analysis builds on previous studies that have demonstrated sequencing of "simple" glycosaminoglycans, such as bikunin [24], and provides a systematic analysis of the more structurally complex HS. Given the complexity of HS, employment of multiple, orthogonal analytical approaches to obtain distinct levels of information on the polysaccharide chain provides important structural insights in a more comprehensive manner than is possible from a single analytical approach. Therefore, this approach outlines a useful framework that can be applied to elucidate structural properties of a heparan sulfate mixture.

Materials and methods

Enzymatic digestion of BKHS samples for IP-RPHPLC and LC-MS analysis

BKHS (1 mg, Seikagaku) was digested at 30 °C for 16 h using a cocktail of *Bacteroides Heparinases* I, III, and IV manufactured in house. The digested sample was then incubated for 6 h at 30 °C simultaneously with 2-*O*-sulfatase and $\Delta^{4,5}$-glycuronidase. On completion of digestion, the sample was frozen and lyophilized. *Flavobacterium Heparinases* (I, II and III) may also be used instead of the cocktail of *Bacteroides Heparinases* I, III, and IV [25].

Ion-pairing reverse phase HPLC analyses

Digested BKHS samples were analyzed by IP-RPHPLC using 30 mM tetra-*n*-butyl ammonium chloride (TBA) as the ion-pairing reagent in 15 % acetonitrile (ACN). The digested samples were separated using an analytical C18 Discovery column (4.6 × 250 mm, Supelco) maintained at 25 °C, with a flow rate of 0.7 mL/min over 130 min of total run time, with a gradient ranging from 0 to 1 M NaCl. The elution profile was monitored by UV absorption at 232 nm.

LC-MS analysis

Chromatographic separation was performed on a 0.3 mm × 250 mm C18 polymeric silica column (thermostated at 17 °C) at a flow rate of 4 μL/min using a Ultimate 3000 capillary HPLC system (Dionex). The mobile phase consisted of (A) 8 mM dibutylammonium acetate in 100 % water, (B) 8 mM dibutylammonium acetate in 70 % methanol/30 % water and (C) 90 % acetonitrile/10 % water. Samples were eluted with the following step gradient: 0 % B for 4 min, 9 % B for 20 min, 20 % B for 21 min, 32 % B for 19 min, 48 % B for

17 min, 63 % B for 14 min, 100 % C for 14 min, and 100 % A for 20 min. To improve ionization, post-column addition of aqueous methanol was carried out using a secondary binary solvent system. Mass spectra were acquired in negative mode with a quadrupole time-of-flight mass spectrometer (Waters-Micromass, UK). Data acquisition was controlled by Masslynx 4.1, according to the following settings: source temperature 100 °C, desolvation temperature 200 °C, capillary voltage 2.4 V, cone voltage 15 V and collision energy 1.7 V.

Size fractionation

10 mg of BKHS were fractionated on a Superdex 75 column (100 mL column volume) and eluted with 5 mM sodium phosphate dibasic and 150 mM sodium chloride, pH 7.0, at a linear flow rate of 15 cm/h. Fractions were buffer-exchanged with 10 % ethanol using Sephadex G-10 and lyophilized to a dry powder.

NMR analysis

The mono- and two-dimensional spectra of BKHS were recorded at 298 K, with a 600 MHz Bruker Avance spectrometer equipped with a 5-mm triple-resonance inverse cryoprobe. Before spectra acquisition, HS was dissolved at 2 mg/150 μL of D_2O (99.9 %) and sonicated for 30 s to remove air bubbles.

^{1}H–NMR spectrum was acquired with presaturation of the residual water signal, with a recycle delay of 5 s, for 64 scans. COSY spectrum was recorded with presaturation of the water signal, for 8 scans of 320 increments. TOCSY spectrum was acquired in phase sensitive mode with 80 ms of DIPSI-2 mixing, for 24 scans of 320 increments. For COSY and TOCSY spectra, the matrix size was zero filled to 2 K × 2 K prior to Fourier transformation. HSQC spectrum was recorded with sensitivity enhancement and carbon decoupling during acquisition, for 64 scans of 256 increments. The polarization transfer delay was set with a $^1J_{C-H}$ coupling value of 155 Hz. For HSQC spectra, the matrix was zero filled to 2 K × 1 K prior to Fourier transformation.

GPC-MS analysis

BKHS (0.6 mg) was digested at 30 °C for 16 h with *Bacteroides Heparinases* I (5 IU/mg) or III (2 IU/mg). On completion of digestion, the sample was frozen, lyophilized, and separated on a chromatographic system consisting of the Ultimate 3000 with Dual-Ternary Capillary HPLC pumps (Dionex) and two 4.6 mm × 300 mm, 4 μm gel permeation TOSOH TSKgel SuperSW2000 columns placed in series and equilibrated at 25 °C. 100 mM ammonium acetate was used as eluent at the constant flow-rate of 100 μl/min delivered by ternary solvent system. The wavelength of the UV detector

(Dionex) was set at 232 nm, followed by the Dionex 2 mm Anion Self-Regenerating Suppressor 300 (ASRS-300 (2-mm)) using water as regenerant. Electrolysis of water in the ASRS-300 was driven by the Dionex SRS controller using the 100 mA current. The flow was post-column controlled by QuickSplit adjustable flow splitters (Analytical Scientific Instrument) to a constant splitting ratio of 1:10.

MS analysis was carried out on a quadrapole time-of-flight (Q-TOF) mass spectrometer (Waters) equipped with an electrospray ionization source (ESI). Mass spectra were acquired in negative "W″ ion mode; N_2 was used as a desolvation gas as well as a nebulizer. Conditions for ESI-MS were as follows: source temperature 100 °C, desolvation temperature 300 °C, capillary voltage 2.4 kV, desolvation gas flow 250 L/h, cone voltage 7 V, cone gas flow 100 L/lr, and collision energy set to 1.2 V. MS spectra were generated by scanning the range of m/z 70–1400, scan time 1.0 s, interscan time 0.1 s. and pusher cycle time 99 μs.

Digestion with alkaline phosphatase

HS (200 μg), previously digested with *Bacteroides Heparinase* III, was diluted in 1 mL of 50 mM Tris, pH 8.8, containing 1 mM $MgCl_2$. Six units of alkaline phosphatase (from *E coli*, Sigma-Aldrich) were added and the solution was incubated at 37 °C overnight.

Results

Composition analysis by IP-RPHPLC, NMR, and LC-MS
In this study, ion-pairing reverse phase high performance liquid chromatography (IP-RPHPLC) was used to determine disaccharide composition of BKHS after exhaustive digestion with a cocktail of *heparinases*, as reported previously. Peaks were identified by co-injection with commercial standards, or through peak isolation and characterization by mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR). The disaccharide composition of BKHS is reported in Table 1. Consistent with characterization of HS from other sources [26], IP-RPHPLC data indicate that six disaccharides constitute more than 90 % of BKHS (based on area percentage): ΔUH_{NAc}, ΔUH_{NS}, ΔUH_{NAc6S}, ΔUH_{NS6S}, $\Delta U_{2S}H_{NS}$, and $\Delta U_{2S}H_{NS6S}$ where ΔU indicates the formation of a $\Delta^{4,5}$ glucuronidate as a result of enzymatic cleavage. A tetrasaccharide corresponding to the anti-thrombin binding site, *i.e.* $\Delta UH_{NAc6S}GH_{NS3S6S}$ [27], was also identified. As previously demonstrated, this structure is resistant to further enzymatic cleavage due to the presence of a reducing glucosamine carrying a sulfate group at the 3-*O* position [28].

To complement and extend results obtained by IP-RPHPLC, intact BKHS was analyzed by mono- and di-dimensional NMR. ^{1}H–NMR and ^{1}H-^{13}C heteronuclear

Table 1 Disaccharide composition of BKHS determined by IP-RPHPLC

	Relative Area (%)
ΔUH_{NAc}	52.3
$\Delta U_{gal}H_{NS}$	0.5
$\Delta UH_{NS}(\alpha)$	14.7
$\Delta UH_{NS}(\beta)$	1.8
ΔUH_{NAc6S}	10.4
$\Delta U_{2S}H_{NAc}$ (α)	0.3
$\Delta U_{2S}H_{NAc}$ (β)	0.5
ΔUH_{NS3S}	0.1
ΔUH_{NS6S} (α)	5.4
$\Delta U_{gal}H_{NS6S}$ (α)	0.1
ΔUH_{NS6S} (β)	0.6
$\Delta U_{2S}H_{NS}$ (α)	6.9
$\Delta U_{2S}H_{NS}$ (β)	0.9
$\Delta U_{2S}H_{NAc6S}(\alpha)$	0.1
$\Delta U_{2S}H_{NAc6S}(\beta)$	0.1
$\Delta UHN_{Ac6S}GH_{NS,3S}$	0.5
$\Delta U_{2S}H_{NS6S}$ (α)	4.4
$\Delta UH_{NAc6S}GH_{NS3S6S}$ (α)	0.1
$\Delta U_{2S}H_{NS6S}$ (β)	0.4

single quantum coherence (HSQC) spectra of BKHS are reported in Fig. 1a, b, respectively. Assignment of HSQC cross peaks was based on literature data [29, 30] and analysis of COSY, TOCSY, and ROESY spectra (data not shown). Taken together, the NMR data identify H_{NS}-(I_{2S}/I) and $H_{NS/NAc}$-(G) as the major glucosamine components of BKHS, and H_{NS3S}, H_{NH2}, and ($H_{NAc\alpha red}$) as minor glucosamine components (Table 2). The most prevalent uronic acids include I_{2S}-(H_{NS}), I-($H_{NS/NAc}$), and G-($H_{NS/NAc}$). Although most of these residues are identified by the chemical shifts of H1/C1, the position of the H2/C2 peaks allows discrimination between I_{2S}-(H_{NS6OH}) and I_{2S}-(H_{NS6S}) (4.349/77.32 and 4.341/78.72 ppm, respectively) [31]. The presence of small amount of epoxide is indicated by a weak signal at 5.41/97.1 ppm. Epoxide is an unnatural residue, generated by the basic reagents used during isolation of HS from the protein core. Treatment with bases is known to cause 2-*O*-desulfation of uronic acids with concomitant formation of epoxide [32].

Besides mono/disaccharide composition, NMR analysis provides useful information regarding the non-reducing and reducing end of BKHS chains. Glucuronic acid located at the non-reducing end is differentiated from internal glucuronic acid by the distinct chemical shift of its H3/C3 and H4/C4 signals. While peaks corresponding to internal glucuronic acid resonate at 3.73/79.2 and 3.83/79.5 ppm in a spectral region crowded with signals, non-reducing glucuronic acid exhibits clear resonances at 3.55/77.9 (H4/C4) and 3.53/74.7 ppm (H3/C3) [33].

Signals of monosaccharides belonging to the linkage region are observed at 4.67/106.8 ppm (GlcA-β1–3- and Gal-β1–3-) and 4.55/104.1 ppm (Gal-β1–4-). The presence of a xylose residue is indicated by H5,5'/C5 cross-peaks at 3.42/65.8 and 4.10/65.8 ppm. Signals at 5.22/94.8 ppm and 4.62/99.3 ppm are assigned to the H1/C1 of the α and β form of xylose, respectively, which is present as a reducing sugar. This observation suggests that the conditions commercially used to isolate BKHS release the saccharide chain from the Ser residue of the proteoglycan core. Additional signals at 5.23/93.4 ppm and 6.37/98.9 ppm are attributed to reducing *N*-acetyl glucosamine (α and β, respectively), providing evidence of the presence of BKHS chains lacking the linkage region sequence. Quantitative composition analysis is performed by integrating the relevant peaks in the HSQC spectrum of BKHS (Table 2). The amount of linkage region and *N*-acetyl glucosamine at the reducing position is used to estimate an average chain length of 16 disaccharides. The relative amount of glucuronic acid at the non-reducing end of the chains is comparable to the amount of reducing end residues (*i.e.*, linkage region plus H_{NAc} redox), indicating that in commercial BKHS the majority of the chains start with a G residue.

The monosaccharide composition determined either by IP-RPHPLC and 2D NMR is compared in Supplementary Table 1. Although IP-RPHPLC analysis has the advantage of requiring less material compared to NMR, the latter provides additional structural information, including epimerization at the C5 position of the uronic acid within different disaccharide moieties. In addition, unlike IP-RPHPLC, 2D NMR detects residues lacking UV absorbance, *e.g.*, glucuronic acid present at the non-reducing end of chains. Finally, NMR is better suited to characterize the linkage region and to determine the presence of residual serine, oxidized serine, or the lack of it. Identification by HPLC is in this case hampered by the lack of suitable standards and by the fact that the current HPLC conditions are optimized to separate sulfated disaccharides and, consequently, are not efficient in isolating neutral residues like the linkage region. This is also the reason why, although NMR shows presence of small amounts of unsubstituted glucosamine (H_{NH2}), such a residue could not be detected by IP-RPHPLC (the standard ΔUH_{NH2} elutes in the void volume).

Analysis of IP-RPHPLC and NMR data indicates general agreement between the two techniques regarding sulfation level and overall composition. However, IP-RPHPLC analysis shows presence of galacturonic acid (ΔU_{gal}), while NMR detects 2,3-epoxide [34]. The root cause for this apparent discrepancy could be attributed to the use of heparin lyase enzymes prior to HPLC analysis, which may result initially in hydrolysis of the epoxide to galacturonic acid, then in conversion of galacturonic acid to ΔU_{gal} through H5 abstraction and glycosidic bond cleavage. Preliminary experiments performed
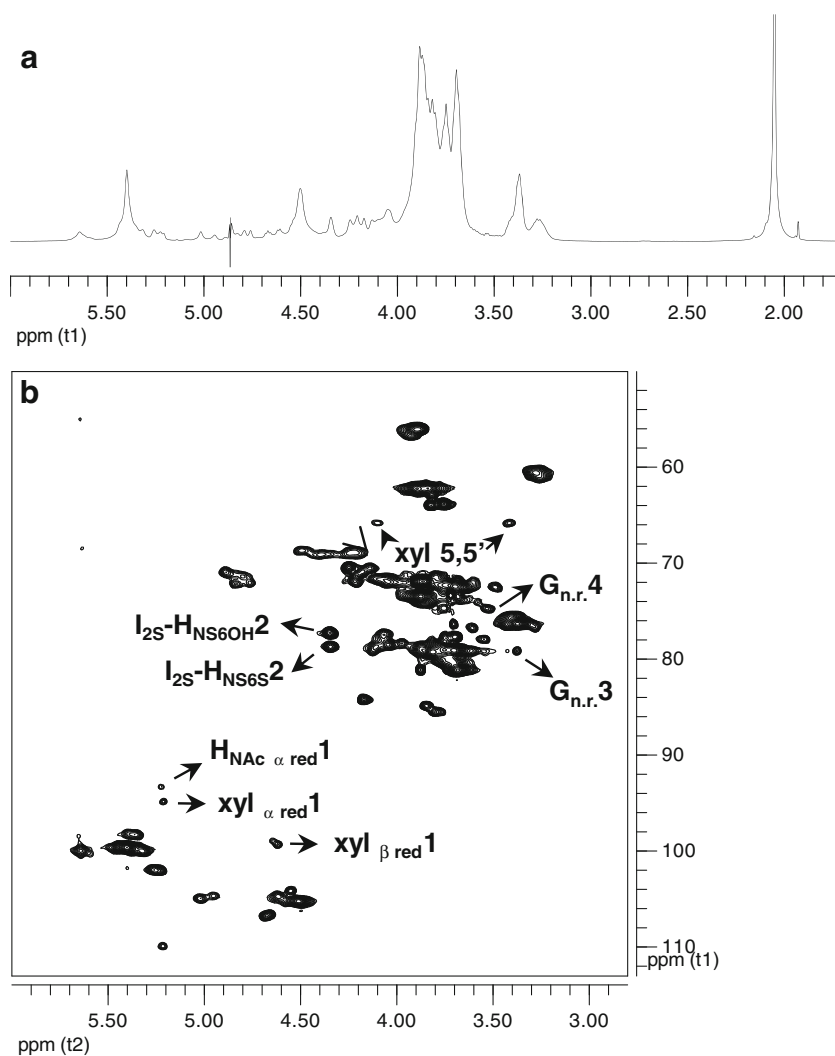
with various enzyme cocktails seem to confirm this hypothesis (data not shown).

To detect structures present in trace amounts, and complete compositional analysis of BKHS, a LC-MS method based on ion-pairing reversed phase high performance liquid chromatography (IP-RPHPLC) coupled to electrospray ionization time-of-flight mass spectrometry (ESI-TOF-MS) was applied to BKHS digested with heparin lyases. This analysis confirmed the major constituents previously detected by IP-RPHPLC (see Supplementary Table 2). Furthermore, the higher degree of sensitivity of ESI-MS enabled identification of additional species below the detection limit of IP-RPHPLC, *e.g.*, unsaturated tetrasaccharides with different levels of sulfation and acetylation. These tetrasaccharides contain a 3-*O*-sulfated glucosamine or an epoxide moiety, which explains their resistance to complete enzymatic digestion. Although epoxide residues may be converted into galacturonic acid by *heparinases*, trace amounts can be detected by ESI-MS. Importantly, through the detection of disaccharides that do not contain a $\Delta^{4,5}$ glucuronate, LC-MS analysis provides

information regarding the non-reducing end of chains. Two types of non-reducing end structures are identified: saturated trisaccharides starting with a glucosamine residue and saturated di/tetrasaccharides starting with an uronic acid moiety. These saccharides present different levels of acetylation and sulfation, ranging from zero to three sulfates per disaccharide and from two to five sulfates per trisaccharide. A fragment present at the reducing end of the chain is also identified, *i.e.* ΔU-Gal-Gal-Xyl-OH. A molecular mass of 337.10 Da assigned to the building block ΔUH$_{NH2}$ provides evidence of the presence of unsubstituted glucosamine, as previously shown by 2D NMR data.

Size fractionation To determine whether disaccharide composition changes as a function of chain length, BKHS was separated by SEC into seven fractions with different elution times. These fractions were then digested with a heparinase cocktail and analyzed by IP-RPHPLC (Table 3). Disaccharide composition data indicate that shorter chains contain a higher amount of residues belonging to the NA region (*i.e.* enriched

Table 2 Relative percentage of variously substituted glucosamine and uronic acid in BKHS determined by HSQC analysis

Monosaccharides[1]	Composition (mole %)
Glucosamine[2]	
H_{NS}-(I_{2S})	11.5
H_{NS}-(I)	9.6
H_{NS}-(G)	10.6
H_{NAc}-(G)	64.8
H_{NS3S}	0.6
H_{NH2}	1.5
H_{NAc} αredox	1.4
H_{NX6OH}[3]	75.0
H_{NX6S}[3]	25.0
Linkage region[2]	
LR	4.7
Uronates[4]	
I_{2S}-(H_{NS6OH})	10.1
I_{2S}-(H_{NS6S})	7.7
I-(H_{NX6S})[3]	7.4
I-(H_{NX6OH})[3]	4.3
G-(H_{NS})[5]	12.8
G-(H_{NAc})[5]	57.0
G-(H_{NS3S})[5]	n.d.
Epoxide	0.7
GalA	n.d.

[1] Residues reported in brackets indicate the monosaccharide downstream of the quantified glucosamine or uronic acid residue

[2] Mole % of the total glucosamine content

[3] X = S or Ac

[4] Mole % of the total uronic acid content

[5] The amount of G at the non-reducing end is calculated as 6.4 % of the total uronic acid content

in ΔUH_{NAc}) as well as residues attributed to the NS region (*i.e.* $\Delta U_{2S}H_{NS6S}$ and $\Delta U_{2S}H_{NS}$). Shorter chains contain fewer residues of the transition regions, as indicated by lower levels of ΔUH_{NS}, ΔUH_{NS6S}, and ΔUH_{NAc6S}. Comparison of *N*-sulfation, 6-*O*-sulfation, and 2-*O*-sulfation levels among chains of different length indicates that shorter chains contain less 6-*O* and *N*-sulfates than longer chains, but are more sulfated at the 2-*O* position (Table 4).

GPC and GPC-MS analysis of HS digested with HepI or HepIII To examine chain composition in finer detail, BKHS was initially digested with either *Hep* I, an enzyme that cleaves highly sulfated regions, or *Hep* III, which cleaves *N*-acetylated and undersulfated domains [36, 37]. After overnight digestion, reaction products were separated by GPC and the UV elution profile at 232 nm was integrated to determine the fragments composition and length distribution (Supplementary Figs. 1 and 2, Supplementary Tables 3 and 4). Digestion of BKHS with *Hep* I degrades NS domains into disaccharides while substantially preserving the NA and the transition domains of the BKHS chains. GPC of *Hep* I-digested BKHS indicates that fragments containing 20 or more monosaccharides represent a significant portion of this sample. Conversely, digestion of BKHS with *Hep* III provides mainly disaccharides; tetra- to dodeca-saccharides constitute in this case a small portion of the total mixture.

Fragments generated by *Hep* I and *Hep* III digestion were consequently analyzed by GPC-MS (Supplementary Tables 5 and 6, respectively). The relative abundance of each fragment was derived from peak intensities determined by applying the MaxEnt3 (Maximum Entropy) algorithm and after application of correction factors to selected residues (see Supplementary Information). Correction factors are introduced to compensate for variations in ionization efficiency of saccharides containing different molar quantities of sulfate groups.

Table 3 IP-RPHPLC analysis of BKHS and its fractions (Rel. Area %). Fractions are reported in order of elution: Pool 1 + 2 corresponds to the fraction with the highest MW, Pool 10 to the fraction with the lowest MW. The MW progressively decreases form Pool 1 + 2 to Pool 10

	Chain Region	BKHS	Pool 1 + 2	Pool 3	Pool 4	Pool 5	Pool 7	Pool 8	Pool 10
ΔUH_{NAc}	NA domain	55.2	52.0	53.5	53.5	54.3	56.4	58.1	58.1
ΔUH_{NS}	Transition	17.7	20.1	19.5	19.0	18.3	17.3	16.4	15.9
ΔUH_{NAc6S}[1]	Transition	10.3	12.0	12.0	11.0	10.6	9.2	8.9	7.9
ΔUH_{NS6S}	Transition	6.1	6.9	6.0	6.3	6.3	6.1	5.6	6.0
$\Delta U_{2S}H_{NS}$	NS domain	6.1	4.6	4.8	5.6	5.9	6.2	6.4	6.9
$\Delta U_{2S}H_{NS6S}$	NS domain	4.6	4.4	4.2	4.5	4.5	4.8	4.6	5.2

[1] Glucosamine 6-*O*-sulfotransferase sulfates H_{NAc} residues provided that there is a H_{NS} residue on either the upstream or the downstream disaccharides [35]. According to the enzyme specificity, we can assume that all the ΔUH_{NAc6S} residues are in the transition region or just at the border of it

[2] The relative standard deviation (RSD) for the peaks measured by this method is ≤1 %, which basically corresponds to a change in peak area of ±0.6 for the most abundant peak and about ±0.1 or less for the lower abundance peaks. Therefore, the changes observed across fractions are larger than what would be expected by method variability

Table 4 Relative percentage of *N*-sulfation vs *N*-acetylation, 6-*O*-sulfation, and 2-*O*-sulfation in BKHS and its fractions (IP-RPHPLC data). Fractions are reported in order of elution: the MW progressively decreases form Pool 1 + 2 to Pool 10

	Intact	Pool 1+ 2	Pool 3	Pool 4	Pool 5	Pool 7	Pool 8	Pool 10
N-Sulfation	34	37	35	36	36	34	33	33
N-Acetylation	66	64	66	64	65	66	67	66
6-*O*-Sulfation	21	24	22	22	22	20	19	19
2-*O*-Sulfation	11	9	9	10	10	11	11	12

Experimentally, using isolated saccharide standards, it was established that differences in ionization due to sulfation are more significant for shorter fragments (*i.e.*, disaccharides). Given this result, it was considered more robust to base disaccharide quantification on UV absorbance; therefore, the disaccharide ratio determined by MS was adjusted to reflect abundances determined by UV. The same correction factors were applied to both saturated and unsaturated disaccharides.

Fragments listed in Supplementary Tables 5 and 6 can be traced back to their position within the BKHS chains: saturated residues derive from the non-reducing end of the chain; residues with a $\Delta^{4,5}$ moiety are originally internal to the chain; residues of formula ΔU-Gal-Gal-Xyl-OH belong to the linkage region at the reducing end of the chain. Monosaccharides and residues of formula H-(U-H)$_n$ are believed to derive from the non-reducing end of the chain. As will be described below, digestion of HS followed by GPC-MS analysis is a useful tool to provide information regarding the NS, NA, and transition domains internal to the chain, as well as the non-reducing and reducing ends of the chain.

Characterization of transition and NA domains Digestion with *Hep* I largely preserves NA domains and transition domains, while it hydrolyzes NS domains to component disaccharides. The GPC profile of BKHS digested with *Hep* I indicates that fragments with a degree of polymerization (Dp) from 8 to 20 are approximately equally represented, and chains longer than Dp20 constitute about 20 % of the entire mixture. Because NA domains are defined as repeating units of G-HNAc residues, fragments generated by *Hep* I can be undoubtedly classified as belonging to NA if they contain a number of acetyl groups equal to (Dp-N)/2 and at least two sulfated residues, which supposedly reside at the border of the NA region and constitute the site of action of *Hep* I. Residues that are assigned to NA domains are marked with an asterisk in Supplementary Table 5. The presence of fragments with a Dp ≥ 20 indicates that BKHS contains NA domains longer than 18 saccharides. Fragments that possess a number of acetyl groups ≤ (Dp-N)/2 are assigned to the transition regions. Transition regions are characterized by significantly higher 2-*O* sulfation and/or 6-*O* sulfation than NA domains, as demonstrated by the presence of structures of formula Dp20,Ac7,S8,Δ and Dp18,Ac7,S6,Δ. Interestingly, the

transition domains appear to be of comparable length to the NA domains.

Characterization of NS domains By preserving sulfated regions, digestion of BKHS with *Hep* III allows characterization of NS domains (Supplementary Table 6). The maximum length of fragments obtained after digestion with *Hep* III is Dp12, providing evidence that within BKHS the NS domains are significantly shorter than the NA or the transition domains. The number of sulfate groups per fragment of the same length indicates that 6-*O*-sulfation is highly variable: some fragments are completely 6-*O*-sulfated (*e.g.*, Dp8,10S,0Ac,Δ) whereas others are 6-*O*-desulfated (*e.g.*, Dp8,7S,0Ac,Δ). Interestingly, shorter fragments (Dp ≤8) are more likely to present acetyl groups at the border of the NS domains.

Characterization of non-reducing end Characterization of BKHS non-reducing ends is performed by GPC-MS analysis of saturated residues obtained after digestion of BKHS with *Hep* I and *Hep* III (Supplementary Tables 5 and 6, respectively). It was previously reported that non-reducing ends of HS chains are mainly sulfated [38]. The GPC-MS and HSQC data reported here indicate that they also contain *N*-acetylglucosamine in significant amounts, and that the most abundant residue observed at the non-reducing end of BKHS chains is glucuronic acid (G). Despite the presence of *N*-acetylglucosamine, non-reducing ends do not contain NA domains, but NS sequences ranging from 8-mer to 12-mer, or transition domains up to 8-mer. Although complete structural elucidation goes beyond the MS capabilities of this experimental set-up, assignment of the predominant non-reducing end structures is proposed by combining information derived from GPC-MS, NMR, the known substrate specificities of *Hep* I and *Hep* III [36, 37], and observations from previous studies on the biosynthetic machinery of HS. Given this set of information, assignment of the non-reducing end structures is proposed in Tables 5 and 6.

Characterization of reducing end Confirming NMR data, LC-MS and GPC-MS identify ΔU/G-Gal-Gal-Xyl-OH as the most abundant residue attributed to the linkage region (Supplementary Table 6). A residue of mass 712.14 Da, corresponding to the mass of ΔU-Gal-Gal-Xyl-OH plus 80 mass

Table 5 GPC-MS Analysis of non-reducing end residues of BKHS digested by *Hep* I

Composition	Relative intensity	Proposed Sequence
Dp2, Ac0, S1, saturated	1.14	G-H$_{NS}$
Dp2, Ac0, S2, saturated	2.48	G-H$_{NS6S}$
Dp4, Ac1, S2, saturated	0.94	G-H$_{NAc}$-G^1-H$_{NS6S}$
Dp4, Ac0, S3, saturated	0.42	G-H$_{NS6X}$–G/I-H$_{NS6X}$
Dp4, Ac1, S3, saturated	0.38	G-H$_{NAc6S}$–G^1-H$_{NS6S}$
Dp4, Ac0, S4, saturated	0.17	G-H$_{NS6S}$–G/I-H$_{NS6S}$
Dp6, Ac1, S3, saturated	0.13	G-H$_{NAc6X}$^{2,3}-G^1-H$_{NS6X}$^{2,3}-G/I-H$_{NS6X}$
Dp6, Ac1, S4, saturated	0.09	G-H$_{NAc6X}$^{2,3}-G^1-H$_{NS6S}$^{2,3}-G/I-H$_{NS6S}$
Dp6, Ac2, S2, saturated	0.10	G-H$_{NAc}$-G^1-H$_{NAc6X}$^{2}-G-H$_{NS6X}$
Dp8, Ac2, S4, saturated	0.02	Not determined

X = S or OH

[1] The specificity of the enzymes that synthesize HS suggests that immediately downstream from HNAc there must be a G [51]. Exception have been found, but these structures should constitute a minority

[2] Glucosamine 6-*O*-sulfotransferase sulfates H$_{NAc}$ residues provided that there is a H$_{NS}$ residue on either the upstream or the downstream disaccharides

[3] the positions of H$_{NAc}$ and H$_{NS,6X}$ can be interchanged

units is also observed in BKHS digested with *Hep* III. To determine whether the 80 mass units could be attributed to a phosphate (PO$_3^-$) or to a sulfate (SO$_3^-$) group, the sample was digested with alkaline phosphatase and analyzed by GPC-MS. Disappearance of the peak at *m/z* 711.14 [M-H]$^-$ and concomitant increase of the peak at *m/z* 631.21 [M-H]$^-$ provided evidence of the presence of a phosphorylated residue. Fragmentation analysis showed that the phosphate group is carried by xylose (data not shown). Interestingly, the presence of phosphorylated linkage region has been previously reported only in bovine lung HS [39]. Phosphorylated linkage region is present only in one of the two commercial lots of BKHS analyzed. This may indicate that the amount of phosphate is highly variable, or that the isolation procedure may not consistently preserve these structures.

Table 6 GPC-MS Analysis of non-reducing end residues of HS digested by *Hep* III

Composition	Relative intensity	Proposed Sequence
Dp2, 0S, 1Ac, saturated	0.83	G-H$_{NAc}$
Dp2, 2S, 0Ac, saturated	0.25	G-H$_{NS6S}$
Dp8, 6S, 2Ac, saturated	0.11	Not determined
Dp8, 7S, 0Ac, saturated	0.14	G-(H$_{NS}$-I$_{2S}$)$_3$-H$_{NS}$
Dp8, 8S, 0Ac, saturated	0.21	G-(H$_{NS6X}$–I$_{2S}$)$_3$-H$_{NS6X}$
Dp8, 9S, 0Ac, saturated	0.12	G-(H$_{NS6X}$–I$_{2S}$)$_3$-H$_{NS6X}$
Dp8, 10S, 0Ac, saturated	0.07	G-(H$_{NS6S}$–I$_{2S}$)$_3$-H$_{NS}$
Dp10, 9S, 0Ac, saturated	0.07	G-(H$_{NS}$-I$_{2S}$)$_4$-H$_{NS}$
Dp10, 10S, 0Ac, saturated	0.09	G-(H$_{NS6X}$–I$_{2S}$)$_4$-H$_{NS6X}$
Dp10, 11S, 0Ac, saturated	0.04	G-(H$_{NS6X}$–I$_{2S}$)$_4$-H$_{NS6X}$
Dp12, 11S, 0Ac, saturated	0.02	G-(H$_{NS}$-I$_{2S}$)$_5$-H$_{NS}$
Dp12, 12S, 0Ac, saturated	0.02	G-(H$_{NS6X}$–I$_{2S}$)$_5$-H$_{NS6X}$

X = S or OH

No residues containing the linkage region were observed by GPC-MS analysis of BKHS digested with *Hep* I. Based on this analysis, the disaccharide sequences close to the linkage region are believed to be mainly non-sulfated, and therefore not easily degraded by *Hep* I. The linkage region may be present in long fragments not easily detectable by GPC-MS analysis [40].

Discussion

Elucidation of GAG chain sequences is a challenging task due to the inherent variability imparted by a non-template driven biosynthetic process coupled with the difficulty of isolating individual HS chains for structural characterization. The theory that GAGs may possess a unique sequence is a subject of considerable debate. Although it is known that biosynthetic enzymes are responsible for organizing HS chains into regions of high sulfation separated by domains of low sulfation [41], it is still unclear whether they possess enough specificity to arrange such domains according to a specific pattern. Recent studies conducted on bikunin GAG indicate that this simple proteoglycan possesses a defined sequence [24]: because different members of the glycosaminoglycan family share a common biosynthetic pathway, results obtained on bikunin GAGs seem to suggest that other GAGs could have well-defined sequences.

To provide additional characterization of HS and to develop a protocol for analysis of complex polysaccharide mixtures, orthogonal analytical methods were optimized and applied to BKHS structural elucidation. The initial efforts focused on implementation of orthogonal analytical methods to determine the fine structure of BKHS, *i.e.*, its mono

−/disaccharide composition. Three different techniques were evaluated: 2D NMR HSQC analysis of intact BKHS, and IP-RPHPLC and LC-MS analysis of BKHS digested with a cocktail of heparinases. By detecting major components as well as structures present in small amount such as H_{NH2}, H_{NS3S}, $\Delta U_{2S}H_{NAc6X}$, or $\Delta UH_{NAc6S}GH_{NS3S6X}$, HSQC and IP-RPHLC provide a snapshot of the structural microheterogeneity of BKHS. These techniques also identify unnatural moieties generated by the chemical treatment used to release HS from the proteoglycan core, *i.e.* 2,3-epoxide and GalA, allowing control over the extent of modification caused by the isolation process [32]. HSQC and IP-RPHLC methods can easily be applied to determine sulfation levels at specific positions (*i.e.*, at the 2-*O* position of uronic acids, or at the *N*- and 6-*O* position of glucosamine residues), making them suitable techniques to compare HS from different sources or pathologic states [42, 43]. In addition, detailed monosaccharide analysis could provide insight into the post-biosynthetic fate of BKHS: the presence of significant amount of glucuronic acid at the non-reducing end of the chains rules out extensive digestion by *β*-D-glucuronidase or heparanase, enzymes that would generate chains starting with a glucosamine residue.

LC-MS analysis of HS chains digested with a cocktail of heparinases is a valuable technique to detect structures that are present in trace amount, below the limit of detection for NMR and IP-RPHPLC. When applied to the study of non-reducing end structures, LC-MS identifies free glucosamine (H_{NH2}) and saturated trisaccharides starting with a glucosamine residue, in addition to saturated di/tetrasaccharides starting with a uronic acid moiety. The presence of saturated residues with different grade of sulfation indicates that non-reducing ends possess a highly variable composition. A recently developed analytical approach based on LC-MS analysis of non-reducing ends of enzymatically digested GAGs isolated from cells, blood and urine is intended to be used as a diagnostic tool to detect lysosomal enzyme deficiency [44].

IP-RPHPLC analysis of BKHS fractionated according to chain length suggests that disaccharide composition, as well as domains organization, change as a function of chain polymerization. Shorter chains seem to contain higher relative amount of disaccharides belonging to the NA and NS regions, and less residues belonging to the transition regions. Previous reports indicated that 2-*O* and 6-*O*-sulfotransferase enzymes are independently regulated during HS biosynthesis [45]. Currently, a comparison of 6-*O*-sulfation, 2-*O*-sulfation, and *N*-sulfation levels among chains of different length suggests that *N*-deacetylases, *N*-sulfotranferases, and 6-*O*-sulfotransferases are more active on longer chains, while 2-*O*-sulfotranferases modify shorter chains more extensively than longer chains. It was previously acknowledged that chains of different lengths may have different composition, and that chain length may be indicative of different biological roles of HS: neuroepithelial cells secrete shorter and simpler chains when they interact with FGF-2, and longer and more complex chains when they interact with FGF-1 [46].

To determine length, distribution, and composition of different domains, BKHS is digested with either *Hep* I or *Hep* III, and the obtained mixtures are analyzed by GPC and GPC-MS. Digestion with *Hep* I, an enzyme that preserves non-sulfated sequences, provides fragments with a degree of depolymerization ranging from 2 to more than 20 saccharides, which can be assigned by MS to either NA or transition domains. Taken together, MS data, degree of polymerization, and acetylation/sulfation levels of fragments obtained by *Hep* I digestion suggest substantial structural diversity within BKHS. These results seem inconsistent with the hypothesis that commercially available BKHS possesses well defined sequences, although this may hold true for chains isolated from a single glycosylation site in specific cells. NS domains obtained after digestion of BKHS with *Hep* III present a maximum length of 12 saccharides, and show variable degrees of sulfation. Overall, the results indicate that in BKHS, NS domains are shorter, while NA and transition domains are longer and of comparable length. Analysis of saturated fragments obtained by *Hep* I or *Hep* III digestion shows that non-reducing ends of BKHS contain NS sequences ranging from 8-mer to 12-mer, or transition domains up to 8-mer long. Although quantitative GPC-MS analysis indicates that non-reducing ends contain *N*-acetylglucosamine in significant amounts, no extended NA domains are detected. The longest NS domain identified at the non-reducing end of BKHS is a dodecasaccharide [47], which is also the maximum length of the NS domains observed within the BKHS chain.

Differently from non-reducing ends, which are constituted by NS or transition domains, reducing ends of BKHS are believed to contain mainly NA domains [48]. In support of this conclusion GPC-MS analysis of BKHS digested with *Hep* III shows fragments belonging to the linkage region, indicating that such residues are most likely immediately preceded by non-sulfated disaccharides. GPC-MS also reveals traces of linkage region fragment where xylose carries a phosphate group. To our knowledge, although phosphorylated linkage region was previously observed in bovine lung HS [39], this is the first report of the presence of phosphate groups in BKHS.

Implementation of orthogonal analytical techniques, such as outlined here, provides a coherent strategy to address HS structure, even in the context of a mixture of chains. Furthermore, we have demonstrated in a separate study (recently published in Scientific Reports [49]) how mathematical modeling can be applied to orthogonal analytical measurements, like those described here, to estimate mixture properties of a complex mixture like heparan sulfate. Combined with structural tools, including array technologies, analysis of HS topology, as well as analysis of HS-protein interactions through crystallography [50] and NMR [51], the ability to

elucidate more general structure-activity relationships for this class of molecules should be increased substantially by the application of such an approach.

Compliance with ethical standards

Conflict of interest All authors are or at some point were employees of Momenta Pharmaceuticals Inc. with stock compensation.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Walker A., Gallagher J.T.: Structural domains of heparan sulphate for specific recognition of the C-terminal heparin-binding domain of human plasma fibronectin (HEPII). Biochem. J. **317**, 871–877 (1996)

2. Sasisekharan R., Venkataraman G.: Heparin and heparan sulfate: biosynthesis, structure and function. Curr. Opin. Chem. Biol. **4**, 626–631 (2000)

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

4. Witt D.P., Lander A.D.: Differential binding of chemokines to glycosaminoglycan subpopulations. Curr. Biol. **4**, 394–400 (1994)

5. Casu B., Lindahl U.: Structure and biological interactions of heparin and heparan sulfate. Adv. Carbohydr. Chem. Biochem. **57**, 159–206 (2001)

6. Liu D., Shriver Z., Qi Y., Venkataraman G., Sasisekharan R.: Dynamic regulation of tumor growth and metastasis by heparan sulfate glycosaminoglycans. Semin. Thromb. Hemost. **28**, 67–78 (2002)

7. Norgard-Sumnicht K.E., Varki N.M., Varki A.: Calcium-dependent heparin-like ligands for L-selectin in nonlymphoid endothelial cells. Science. **261**, 480–483 (1993)

8. Liu J. et al.: Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities. J. Biol. Chem. **274**, 5185–5192 (1999)

9. Gallagher J.T., Turnbull J.E., Lyon M.: Patterns of sulphation in heparan sulphate: polymorphism based on a common structural theme. Int. J. Biochem. **24**, 553–560 (1992)

10. Rudd T.R., Yates E.A.: A highly efficient tree structure for the biosynthesis of heparan sulfate accounts for the commonly observed disaccharides and suggests a mechanism for domain synthesis. Mol. BioSyst. **8**, 1499–1506 (2012)

11. Capila I., Linhardt R.J.: Heparin-protein interactions. Angew. Chem. Int. Ed. Eng. **41**, 391–412 (2002)

12. Murphy K.J. et al.: A new model for the domain structure of heparan sulfate based on the novel specificity of K5 lyase. J. Biol. Chem. **279**, 27239–27245 (2004)

13. Fromm J.R., Hileman R.E., Caldwell E.E., Weiler J.M., Linhardt R.J.: Differences in the interaction of heparin with arginine and lysine and the importance of these basic amino acids in the binding of heparin to acidic fibroblast growth factor. Arch. Biochem. Biophys. **323**, 279–287 (1995)

14. Guglieri S. et al.: Minimum FGF2 binding structural requirements of heparin and heparan sulfate oligosaccharides as determined by NMR spectroscopy. Biochemistry. **47**, 13862–13869 (2008)

15. Fath M.A. et al.: Interaction of secretory leukocyte protease inhibitor with heparin inhibits proteases involved in asthma. J. Biol. Chem. **273**, 13563–13569 (1998)

16. Khan S. et al.: The solution structure of heparan sulfate differs from that of heparin: implications for function. J. Biol. Chem. **286**, 24842–24854 (2011)

17. Bame K.J., Robson K.: Heparanases produce distinct populations of heparan sulfate glycosaminoglycans in Chinese hamster ovary cells. J. Biol. Chem. **272**, 2245–2251 (1997)

18. Westling C., Lindahl U.: Location of N-unsubstituted glucosamine residues in heparan sulfate. J. Biol. Chem. **277**, 49247–49255 (2002)

19. Jastrebova N. et al.: Heparan sulfate-related oligosaccharides in ternary complex formation with fibroblast growth factors 1 and 2 and their receptors. J. Biol. Chem. **281**, 26884–26892 (2006)

20. Gallagher J.T.: Heparan sulfate: growth control with a restricted sequence menu. J. Clin. Invest. **108**, 357–361 (2001)

21. Kreuger J., Spillmann D., Li J.P., Lindahl U.: Interactions between heparan sulfate and proteins: the concept of specificity. J. Cell Biol. **174**, 323–327 (2006)

22. Turnbull J.E., Gallagher J.T.: Sequence analysis of heparan sulphate indicates defined location of N-sulphated glucosamine and iduronate 2-sulphate residues proximal to the protein-linkage region. Biochem. J. **277**, 297–303 (1991)

23. Parthasarathy N., Spiro R.G.: Isolation and characterization of the heparan sulfate proteoglycan of the bovine glomerular basement membrane. J. Biol. Chem. **259**, 12749–12755 (1984)

24. Ly M. et al.: The proteoglycan bikunin has a defined sequence. Nat. Chem. Biol. **7**, 827–833 (2011)

25. Karamanos N.K., Vanky P., Tzanakakis G.N., Tsegenidis T., Hjerpe A.: Ion-pair high-performance liquid chromatography for determining disaccharide composition in heparin and heparan sulphate. J. Chromatogr. A. **765**, 169–179 (1997)

26. Toida T. et al.: Structural differences and the presence of unsubstituted amino groups in heparan sulphates from different tissues and species. Biochem. J. **322**, 499–506 (1997)

27. Shriver Z. et al.: Sequencing of 3-O sulfate containing heparin decasaccharides with a partial antithrombin III binding site. Proc. Natl. Acad. Sci. U. S. A. **97**, 10359–10364 (2000)

28. Shriver Z. et al.: Cleavage of the antithrombin III binding site in heparin by heparinases and its implication in the generation of low molecular weight heparin. Proc. Natl. Acad. Sci. U. S. A. **97**, 10365–10370 (2000)

29. Guerrini M., Naggi A., Guglieri S., Santarsiero R., Torri G.: Complex glycosaminoglycans: profiling substitution patterns by two-dimensional nuclear magnetic resonance spectroscopy. Anal. Biochem. **337**, 35–47 (2005)

30. Guerrini M. et al.: A novel computational approach to integrate NMR spectroscopy and capillary electrophoresis for structure assignment of heparin and heparan sulfate oligosaccharides. Glycobiology. **12**, 713–719 (2002)

31. Yates E.A. et al.: 1H and 13C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. Carbohydr. Res. **294**, 15–27 (1996)

32. Rej R.N., Perlin A.S.: Base-catalyzed conversion of the a-l-iduronic acid 2-sulfate unit of heparin into a unit of a-l-galacturonic acid, and related reactions. Carb. Res. **200**, 437–447 (1990)

33. Guerrini M., Guglieri S., Naggi A., Sasisekharan R., Torri G.: Low molecular weight heparins: structural differentiation by bidimensional nuclear magnetic resonance spectroscopy. Semin. Thromb. Hemost. **33**, 478–487 (2007)

34. Hricovíni M., Guerrini M., Torri G., Piani S., Ungarelli F.: Conformational analysis of heparin epoxide in aqueous solution. An NMR relaxation study. Carbohydr. Res. **277**, 11–23 (1995)

35. Conrad H.E.: Heparin-binding proteins ch. 2. Academic Press, San Diego (1998)

36. Ernst S., Rhomberg A.J., Biemann K., Sasisekharan R.: Direct evidence for a predominantly exolytic processive mechanism for depolymerization of heparin-like glycosaminoglycans by heparinase I. Proc. Natl. Acad. Sci. U. S. A. **95**, 4182–4187 (1998)

37. Pojasek K., Shriver Z., Hu Y., Sasisekharan R.: Histidine 295 and histidine 510 are crucial for the enzymatic degradation of heparan sulfate by heparinase III. Biochemistry. **39**, 4012–4019 (2000)

38. Wu Z.L., Lech M.: Characterizing the non-reducing end structure of heparan sulfate. J. Biol. Chem. **280**, 33749–33755 (2005)

39. Fransson L.A., Silverberg I., Carlstedt I.: Structure of the heparan sulfate-protein linkage region. Demonstration of the sequence galactosyl-galactosyl-xylose-2-phosphate. J. Biol. Chem. **260**, 14722–14726 (1985)

40. Lyon M., Steward W.P., Hampson I.N., Gallagher J.T.: Identification of an extended N-acetylated sequence adjacent to the protein-linkage region of fibroblast heparan sulphate. Biochem. J. **242**, 493–498 (1987)

41. Gallagher J.T.: Heparan sulphate and protein recognition. Binding specificities and activation mechanisms. Adv. Exp. Med. Biol. **376**, 125–134 (1995)

42. Parra A. et al.: Heparin-like heparan sulfate from rabbit cartilage. Glycobiology. **22**, 248–257 (2012)

43. Feyzi E., Saldeen T., Larsson E., Lindahl U., Salmivirta M.: Age-dependent modulation of heparan sulfate structure and function. J. Biol. Chem. **273**, 13395–13398 (1998)

44. Lawrence R. et al.: Disease-specific non-reducing end carbohydrate biomarkers for mucopolysaccharidoses. Nat. Chem. Biol. **8**, 197–204 (2012)

45. Maccarana M., Sakura Y., Tawada A., Yoshida K., Lindahl U.: Domain structure of heparan sulfates from bovine organs. J. Biol. Chem. **271**, 17804–17810 (1996)

46. Brickman Y.G. et al.: Structural modification of fibroblast growth factor-binding heparan sulfate at a determinative stage of neural development. J. Biol. Chem. **273**, 4350–4359 (1998)

47. Staples G.O., Shi X., Zaia J.: Extended N-sulfated domains reside at the nonreducing end of heparan sulfate chains. J. Biol. Chem. **285**, 18336–18343 (2010)

48. Lyon M., Deakin J.A., Gallagher J.T.: Liver heparan sulfate structure. A novel molecular design. J. Biol. Chem. **269**, 11208–11215 (1994)

49. Pradines J.R., Beccati D., Lech M., Ozug O., Farutin V., Huang H., Gunay N.S., Capila I.: Combining measurements to estimate properties and characterization extent of complex biochemical mixtures; applications to heparan sulfate. Sci. Report. **6**, 24829 (2016). doi:10.1038/srep24829

50. DiGabriele A.D. et al.: Structure of a heparin-linked biologically active dimer of fibroblast growth factor. Nature. **393**, 812–817 (1998)

51. Levy-Adam F. et al.: Identification and characterization of heparin/heparan sulfate binding domains of the endoglycosidase heparanase. J. Biol. Chem. **280**, 20457–20466 (2005)