

Review

Glycosaminoglycan-Protein Interactions by Nuclear Magnetic Resonance (NMR) Spectroscopy

Vitor H. Pomin ^{1,*}  and Xu Wang ^{2,*}

¹ Department of BioMolecular Sciences, Division of Pharmacognosy and Research Institute of Pharmaceutical Sciences, School of Pharmacy, University of Mississippi, Oxford, MS 38677-1848, USA

² School of Molecular Sciences, Arizona State University, Tempe, AZ 85287-0112, USA

* Correspondence: vpomin@olemiss.edu (V.H.P.); xuwang@asu.edu (X.W.); Tel.: +1-662-915-3114 (V.H.P.); +1-480-727-8256 (X.W.)

Received: 8 August 2018; Accepted: 5 September 2018; Published: 11 September 2018



Abstract: Nuclear magnetic resonance (NMR) spectroscopy is one of the most utilized and informative analytical techniques for investigating glycosaminoglycan (GAG)-protein complexes. NMR methods that are commonly applied to GAG-protein systems include chemical shift perturbation, saturation transfer difference, and transferred nuclear Overhauser effect. Although these NMR methods have revealed valuable insight into the protein-GAG complexes, elucidating high-resolution structural and dynamic information of these often transient interactions remains challenging. In addition, preparation of structurally homogeneous and isotopically enriched GAG ligands for structural investigations continues to be laborious. As a result, understanding of the structure-activity relationship of GAGs is still primitive. To overcome these deficiencies, several innovative NMR techniques have been developed lately. Here, we review some of the commonly used techniques along with more novel methods such as waterLOGSY and experiments to examine structure and dynamic of lysine and arginine side chains to identify GAG-binding sites. We will also present the latest technology that is used to produce isotopically enriched as well as paramagnetically tagged GAG ligands. Recent results that were obtained from solid-state NMR of amyloid's interaction with GAG are also presented together with a brief discussion on computer assisted modeling of GAG-protein complexes using sparse experimental data.

Keywords: glycosaminoglycan; GAG binding site; GAG-protein interactions; isotopic labeling; NMR; paramagnetic labeling

1. Introduction

Glycosaminoglycans (GAGs) are linear, anionic, high molecular weight (MW) polysaccharides that are composed of repeating disaccharide building blocks of alternating hexosamine and uronic acid/galactose units (Table 1, Figure 1). The monosaccharide composition and sulfation patterns of the disaccharide units of GAGs are used to classify GAGs into different categories (Table 1, Figure 1). GAGs have potent and varied biological activity. The formation of intermolecular complexes between GAGs and functional proteins is a requirement in most physiological and certain pathological events (Figure 2) [1–6]. These GAG-protein interactions and their consequential functions are prevalent in all tissues of the human body.

Among all glycoconjugates in the glycocalyx, GAGs are likely to be the most important player of the extracellular matrix in terms of protein interaction and regulation. This is because of two essential characteristics of these sulfated heteropolysaccharides: (i) they are ubiquitous to all eukaryotic cells; and, (ii) their highly complex structure gives rise to unique motifs that may have specific interactions with proteins. Hence, knowledge about the structural biology of GAG-protein interactions

is extremely relevant. Three major goals exist in the current science of GAG-protein interactions: (i) identification of novel GAG sequences recognized specifically by proteins; (ii) identification of the potential GAG-binding site(s) in GAG-binding proteins; and, (iii) understand how GAG-induced changes in protein oligomerization and conformation regulate physiological activity.

Table 1. Commonest monosaccharides and sulfation patterns of the different (sub)classes of glycosaminoglycans (GAGs).

GAG Type or Sub-Type	Hexosamine *	Uronic Acid or Neutral Sugar *	Sulfation Pattern *
Heparin	4-linked α -GlcNAc	4-linked α -IdoA	GlcNS6S, IdoA2S
Heparan sulfate	4-linked α -GlcNAc	4-linked β -GlcA	non-sulfated
Chondroitin sulfate A	3-linked β -GalNAc	4-linked β -GlcA	GalNAc4S
Chondroitin sulfate B (dermatan sulfate)	3-linked β -GalNAc	4-linked α -IdoA	GalNAc4S, IdoA2S
Chondroitin sulfate C	3-linked β -GalNAc	4-linked β -GlcA	GalNAc6S
Chondroitin sulfate D	3-linked β -GalNAc	4-linked β -GlcA	GalNAc6S, GlcA2S
Chondroitin sulfate E	3-linked β -GalNAc	4-linked β -GlcA	GalNAc4S6S
Keratan sulfate	4-linked β -GlcNAc	3-linked α -Gal	GlcNAc6S, Gal6S
Hyaluronan	3-linked β -GlcNAc	3-linked β -GlcA	non-sulfated

* The abbreviation of the sugars are GlcNAc, *N*-acetylglucosamine; IdoA, iduronic acid; GlcNS6S, *N*,6-di-sulfated glucosamine; IdoA2S, 2-sulfated iduronic acid; GalNAc, *N*-acetylgalactosamine, GalNAc4S, 4-sulfated *N*-acetylgalactosamine; GalNAc6S, 6-sulfated *N*-acetylgalactosamine; GlcA2S, 2-sulfated glucuronic acid; GalNAc4S6S, 4,6-di-sulfated galactosamine; Gal, galactose; GlcNAc6S, 6-sulfated *N*-acetylglucosamine.

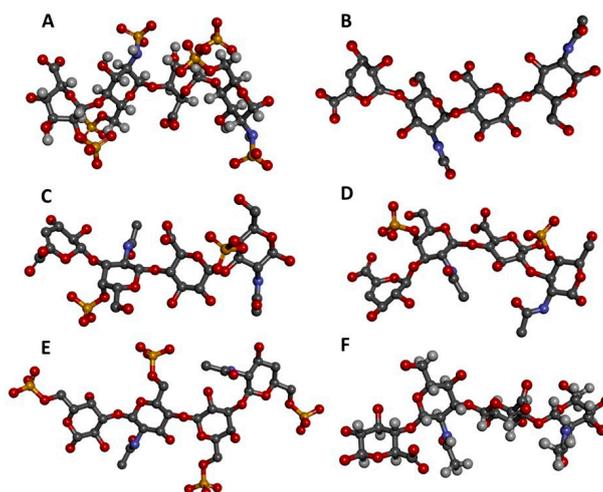


Figure 1. Three-dimensional representations of GAG tetrasaccharides taken from the protein data bank (PDB). (A) Heparin [IdoA2S-(α 1-4)-GlcNS6S-(α 1-4)-IdoA2S-(α 1-4)-GlcNS6S] from 1HPN [7]; (B) Heparan sulfate [GlcA-(β 1-4)-GlcNAc-(β 1-4)-GlcA-(β 1-4)-GlcNAc] from 3E7 [8]; (C) Chondroitin 4-sulfate (CS-A) [GlcA-(β 1-3)-GalNAc4S-(β 1-4)-GlcA-(β 1-3)-GalNAc4S] from 1OFM [9]; (D) Dermatan sulfate (CS-B) [IdoA-(α 1-3)-GalNAc4S-(β 1-4)-IdoA-(α 1-3)-GalNAc4S] from 1OFL [9]; (E) Keratan sulfate [Gal6S-(β 1-4)-GlcNAc6S-(β 1-3)-Gal6S-(β 1-4)-GlcNAc6S] from 1KES [10]; and (F) Hyaluronan [GlcA-(β 1-3)-GlcNAc-(β 1-4)-GlcA-(β 1-3)-GlcNAc] from 2BVK [11]. The atoms in the ball-stick representations are carbon (grey), nitrogen (blue), hydrogen (light grey); oxygen (red) and sulfur (yellow). Representations A and F are nuclear magnetic resonance (NMR) solution structures, and non-exchangeable protons are shown; the other representations are crystal structures so are shown without protons. Reprint with permission from MDPI *Pharmaceuticals* (Basel), 2018, 11(1) [12].

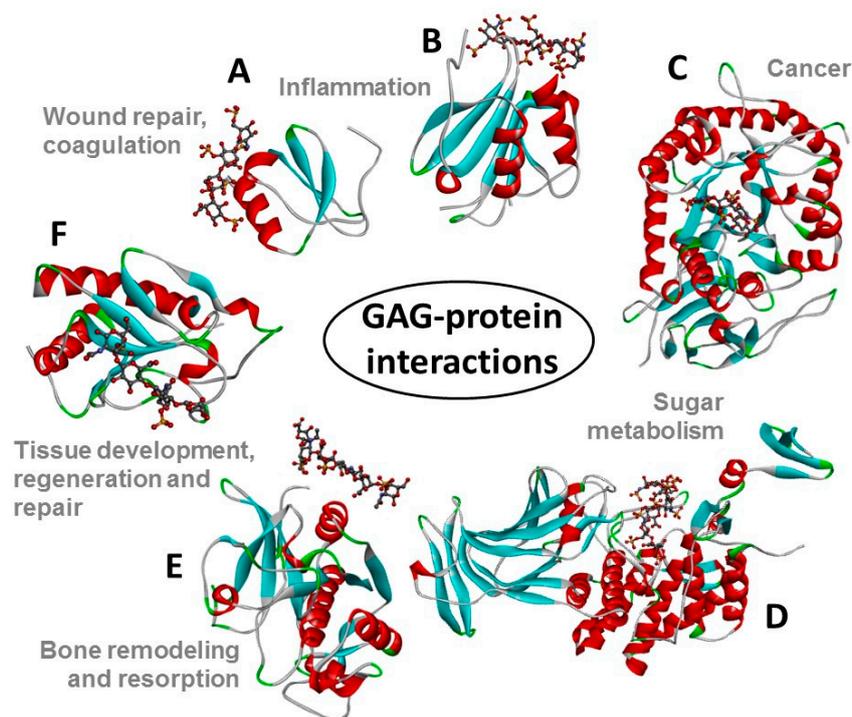


Figure 2. Structural representations from the crystal structures of some illustrative GAG-protein complexes in PDB: (A) macrophage inflammatory protein 1 α (CCL3) monomer + heparin tetrasaccharide (from 5D65) [13], (B) Platelet factor 4 (CXCL4) dimer + fondaparinux (antithrombin-high affinity heparin pentasaccharide) (from 4R9W) [14], (C) human heparanase complex + heparin tetrasaccharide (from 5E9C) [15], (D) D-glucuronyl C5-epimerase + heparin hexasaccharide (from 4PXQ) [16], (E) Cathepsin K monomer + dermatan sulfate hexasaccharide (from 4N79) [17], and (F) Sonic Hedgehog monomer + CS-A tetrasaccharide (from 4C4M) [18]. The atoms of the GAG ligands represented in the ball-stick view are carbon (grey), nitrogen (blue), hydrogen (light grey); oxygen (red) and sulfur (yellow). In the proteins, the α -helices, β -sheets, loops and random coils are represented respectively in red, blue, green, and grey. The pathophysiological systems in which these complexes play a role are indicated by grey fonts in the panel. Reprint with permission from MDPI *Pharmaceuticals* (Basel), 2018, 11(1) [12].

One analytical technique that is capable of studying the various aspects of GAG-protein interactions including the three above-mentioned goals is nuclear magnetic resonance (NMR), whose versatility has made it a popular choice among researchers interested in GAG-binding proteins. NMR-oriented studies have been designed to elucidate multiple aspects of GAG-protein interactions, not only through the individual protein and ligand point-of-views but also through the perspective of the resultant complexes in solution, including their conformational fluctuations from the ns-to-ms timescale.

In this review, we want to offer an overview of the potential of this analytical technique for advancing studies of GAG-protein interactions. Although discussion on the most common NMR-based methods that are utilized to investigate GAG-protein interactions, such as chemical shift perturbation, saturation transferred difference, and transferred nuclear Overhauser effect, are presented, attention is also given to newer methods that have been developed lately. These up-to-date methods include water-ligand observed via gradient spectroscopy, isotopic enrichment and paramagnetic labeling of GAG ligands of well-defined chemical structures, characterization of lysine and arginines of GAG-binding sites, dynamics of the GAG-protein complexes, and solid-state NMR for resultant complexes with low solubility. Finally, we discuss how these NMR-based findings of GAG-protein interactions can be used to create molecular models of the complexes to better understand the structure-activity relationship of GAG-protein interactions.

2. Common NMR Experiments

2.1. Chemical Shift Perturbation

Chemical shift perturbation (CSP) or chemical shift mapping is perhaps the mostly used NMR method in GAG-protein interaction studies [19]. It is generally carried out on ^{15}N -labeled proteins and allows for GAG-binding sites in proteins to be identified. The protein atomic chemical shifts of interest are usually obtained through the ^{15}N -edited heteronuclear single quantum coherence (HSQC) spectra of ^{15}N -labeled proteins and medium-sized GAG oligosaccharides whose composition can be heterogeneous. To determine the GAG-binding site, multiple HSQC spectra of the protein with increasing concentrations of the ligand are recorded. These spectra allow for the ligand induced chemical shift changes in protein atoms to be quantified. In particular, the ^{15}N -HSQC experiment measures the chemical shift changes of backbone amide proton and nitrogen ($\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{N}}$) in a residue specific fashion. The magnitudes of the changes are often diagnostic of ligand binding sites as larger changes are interpreted as a consequence of closeness of the atoms to the ligand. But, exceptions can exist. For example, when aromatic rings or charges approach the amide groups, this can sometimes exhibit a larger chemical shift change than expected just from the distances. Nonetheless, the magnitude of CSP for amide ^1H and ^{15}N resonances is usually quantified while using the empirical nonlinear Equation (1), where $\Delta\delta_{\text{HN}}$ and $\Delta\delta_{\text{N}}$ are the observed CSP for ^1H and ^{15}N , respectively.

$$\Delta\delta = [(\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{N}} \times 0.12)^2]^{1/2} \quad (1)$$

The weighting factor 0.12 reflects the difference in chemical shift dispersion of ^1H and ^{15}N in folded proteins [20]. These data can be easily used to determine the dissociation constant (K_{d}) of the intermolecular complex with a binding stoichiometry of one-to-one using Formula (2), where x is the total ligand concentration, $\Delta\delta$ is the observed chemical shift change, $[\text{Protein}]$ is the total protein concentration, and $\Delta\delta_{\text{max}}$ is the maximum GAG-induced chemical shift change for the atom.

$$\Delta\delta = (\Delta\delta_{\text{max}}/2[\text{Protein}]) \cdot \{(K_{\text{d}} + x + [\text{Protein}]) - [(K_{\text{d}} + x + [\text{Protein}])^2 - (4[\text{Protein}]x)]^{1/2}\} \quad (2)$$

As discussed in Section 7, data derived from CSP can be used as input in computational docking for generating the molecular model of the resultant complex.

As an example of the CSP application, the binding region of fondaparinux (Fx) on human beta-defensin 6 (hBD6) mapped by this method is here described [21]. Fx is a synthetic and highly sulfated heparin pentasaccharide that is commonly exploited as heparan sulfate mimetic in NMR-based GAG-protein interaction studies. HS is believed to be the major physiological target for most of the GAG-binding proteins, including the defensin hBD6. Isolation of well-defined oligosaccharides of heparan sulfate is difficult and since Fx is commercially available, this ligand becomes a preferable choice for research. Figure 3A shows a partial but detailed region of the ^{15}N -HSQC NMR spectra of hBD6 recorded as a function of the Fx concentration: from a 0 (black) to 4:1 Fx:hBD6 ratio (cyan). NMR titration experiments revealed large changes in hBD6 chemical shifts in a fast exchange equilibrium between free and Fx-bound hBD6, as demonstrated by a gradual shift in peak positions seen by the different ^{15}N -HSQC spectra recorded with increasing concentrations of heparin ligand. The quantitative analysis using the ^1HN and ^{15}N chemical shift changes for a subset of hBD6 resonances upon titration with Fx using the above-discussed formulas yielded a K_{d} value of 4.1 (± 2.9) μM (Figure 3B). Figure 3C depicts the CSP map of Fx on hBD6 upon complete saturation (Fx:hBD6 molar ratio of 4:1) plotted as a function of the residue number of hBD6 for the Fx titration.

From the work depicted in Figure 3, the hBD6 residues most sensitive to Fx binding were determined. They are located in the α -helix (amino acids F1, F2, D3, E4, K5, C6, N7) and in part of the $\beta 2$ and $\beta 3$ strands (amino acids C27, Q28, K29, S30, L31, K32) of the defensin. Three lysines (K5, K29, and K32) out of seven have backbone ^{15}N and ^1H resonances significantly affected by the

binding of Fx (Figure 3). These three perturbed lysine residues were reported to be part of the same heparin binding motif in hBD6 [21]. These basic amino acids perturbed in the CSP method indicate that electrostatic interactions are a major contributor to Fx binding to hBD6. Figure 3D shows the Fx binding site mapped onto the hBD6 structure (PDB ID: 2LWL). This dataset was revisited herein with permission from [21].

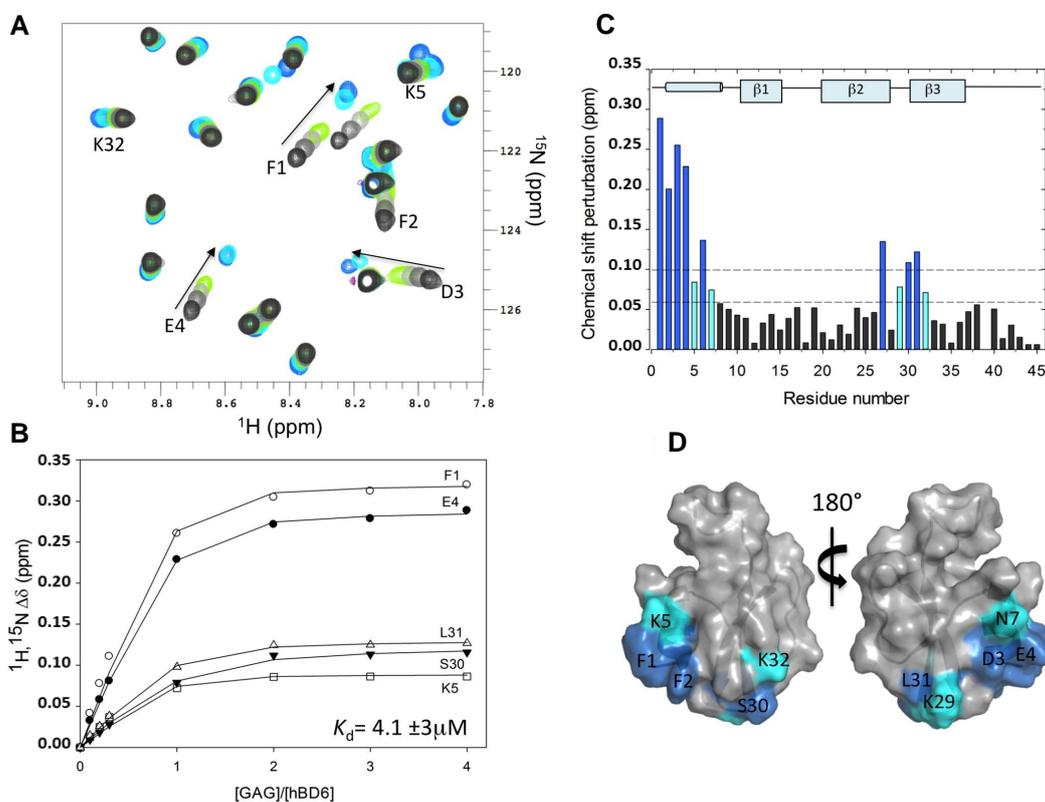


Figure 3. Interaction of human beta-defensin 6 (hBD6) with fondaparinux (Fx) by chemical shift perturbation (CSP) analysis. **(A)** A section of the ^{15}N -edited heteronuclear single quantum coherence (^{15}N -HSQC) spectra of the free hBD6 (black) and in the presence of heparin pentasaccharide Fx. The arrow indicates the direction of binding-induced chemical shift changes; **(B)** Representative binding profile for calculating the binding constant of hBD6-Fx interactions; **(C)** CSP map of hBD6-Fx interactions. The change seen for individual residues were judged in cutoffs (dotted lines) of 0.06 (light cyan) and 0.10 (dark cyan) ppm. The sequence-specific secondary structural elements are shown on the top of the CSP map; **(D)** Surface representation of hBD6 residues that are significantly perturbed on heparin pentasaccharide binding. The perturbed residues are shown in blue and cyan. Republished with permission of ASBMB *J. Biol. Chem.* from Unique properties of human β -defensin 6 (hBD6) and glycosaminoglycan complex: sandwich-like dimerization and competition with the chemokine receptor 2 (CCR2) binding site. de Paula V.S., Pomin V.H., Valente A.P.; 289, 33, 2014; permission conveyed through Copyright Clearance Center, Inc. [21].

It is worth pointing out that CSP is not only reflective of direct protein-ligand interactions since CSP can be also influenced by other factors, such as changes in protein conformation due to binding and/or on the chemical environment of the mapped region of the protein. In addition to that, as commented above, when aromatic rings and/or charged groups approach amides, magnitude of the CSP may not be solely related to physical contact and this would increase complexity during data interpretation, leading to artifactual results. An alternative to enhance precision in the GAG-protein interaction study is to perform additional paramagnetic relaxation enhancement (PRE) experiments to complement the CSP methodology. The PRE method is described below at Section 3.3. Experiments

designed to look at dynamics and orientations of side chains of lysines and arginines should also be helpful in determining the GAG-binding residues.

2.2. Saturation Transfer Difference

Saturation transfer difference (STD) is a NMR experiment designed to map ligand protons that are involved in the interaction with protein [22–25]. In this method a group of protein protons, whose resonance frequencies do not overlap with any resonance frequency of the ligand, is selectively saturated. The extensive dipole-dipole interaction network in a protein allows the saturation energy to be spread to other protein protons as well as bound ligands, resulting in decreases in the signal intensities of the ligand protons in contact with the protein. Figure 4 illustrates the principle behind this NMR method. Because ligand proton signals are far easier to detect and STD is usually carried out under excess ligand conditions, STD is most frequently used to investigate the GAG-protein complexes through the GAG (ligand) perspective.

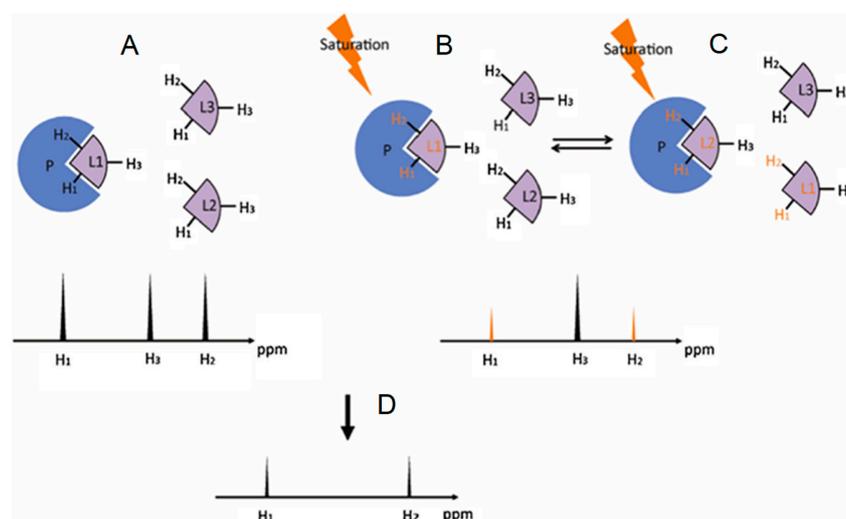


Figure 4. Schematic representation of the saturation transfer difference (STD) experiment. (A) The atoms of the ligand free in solution, in this case H_1 , H_2 and H_3 , generate NMR signals whose intensities are proportional to their abundance; (B) Upon interaction, the saturated energy-excitement (orange arrow) on the protein will be transferred to the bound atoms (orange fonts); (C) The on-off behavior of the protein-ligand complex in solution leads to free but magnetized ligands. The intensities of the NMR signals will reflect the difference between the non-saturated (black peak) and the saturated atoms (orange peaks); The resultant STD spectrum (D) will be generated by subtracting the saturated spectrum (B,C) from the reference spectrum (A). Modified with permission from Springer *Anal. Bioanal. Chem.* License in conditions provided by Springer Nature and Copyright Clearance Center, Inc. [25].

STD has been used to investigate the GAG specificity of numerous GAG-binding proteins, including basic fibroblast growth factor (bFGF or FGF2). Crystal structure of FGF2 with heparin showed IdoA2S-GlcNS disaccharide units are essential for strong FGF2-GAG interactions and this was later confirmed by Yu et al. through STD (Figure 5) [25]. In Figure 5A, the 1D 1H reference spectrum of 30 μM of FGF2 and 370 μM of a synthetic heparin octasaccharide composed of only IdoA2S-GlcNS disaccharide units (Figure 5B) is shown. The STD spectrum of the same sample is displayed in Figure 5C. Intense and narrow peaks can be noted in this STD spectrum, indicating significant interactions between FGF2 and the ligand. The assignments of the ligand residues that are involved in the interaction revealed FGF2 has the strongest interaction with the internal region (dashed box) of the ligand. In particular, atoms on the 2-O- and N-sulfation side of the internal tetrasaccharide made of the units labeled as C, D, E, and F produced the strongest STD signals.

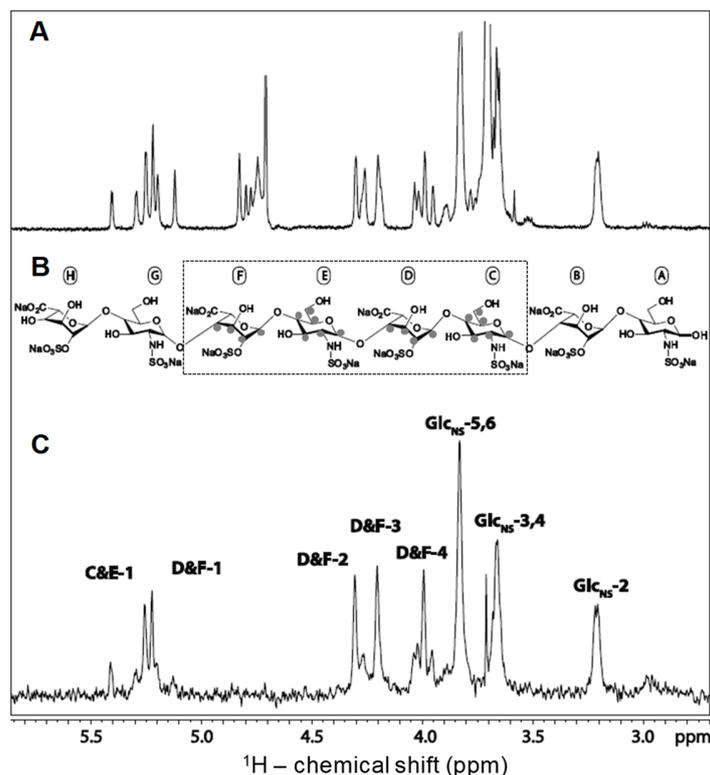


Figure 5. NMR assignments of the STD spectrum obtained from the complex of a heparin octasaccharide and fibroblast growth factor 2 (FGF2). (A) Reference spectrum recorded in the presence of FGF2, but without saturating the protein with magnetization; (B) Structure of the heparin octasaccharide utilized in the analysis. This ligand is composed of alternating IdoA2S and GlcNS units. See Table 1 for abbreviations. The internal tetrasaccharide highlighted by the dashed box indicates the binding motif recognized by the STD experiment; and, (C) The grey dots on the structural units indicate the signals assigned in the STD spectrum. Adapted with permission from Springer *Anal. Bioanal. Chem.* License in conditions provided by Springer Nature and Copyright Clearance Center, Inc. [25].

2.3. Transferred Nuclear Overhauser Effect

The bound-state conformation of a ligand, such as a GAG oligosaccharide, can be deduced through a method called exchange-transferred nuclear Overhauser effect (trNOE) [26]. Through the use of this NMR method, protein-induced changes in the intramolecular NOE signals of the ligand can be identified. This provides valuable information about the conformational changes of the GAG oligosaccharide upon interaction with the protein. Similar to STD, trNOE takes advantage of the changes in the rotational motion of the ligand upon binding to the larger macromolecule [27]. The rotational correlation time of small molecules such as medium- and short-sized GAG oligosaccharides lies on the ps timescale and the resultant NOE signals are small and positive. The correlation time of larger molecules, such as most of the GAG-binding protein and the resultant GAG-protein complexes, lies on the nanosecond time range and the associated NOEs are large and negative (Figure 6). Hence, the correlation time of a small ligand, GAG oligosaccharide in our case, changes considerably upon interaction with the protein partner. The consequential change of NOE cross-peak sign and magnitude can be used to elucidate the conformation of protein-bound ligands, even in the presence of excess unbound ligands.

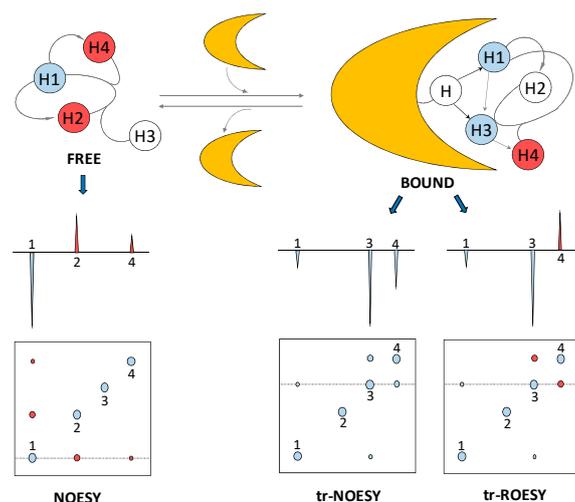


Figure 6. Schematic representation of the trNOESY and trROESY methods for an exchanging ligand-protein (yellow) complex. The NOE cross-peaks of a free ligand are positive (red) and display opposite sign to the negative (blue) diagonal (auto) peaks, but upon interaction with the protein in a trNOE experiment, these peaks also become negative due to the increase of effective rotational correlation time. The analysis of the negative cross-peaks provides information on the bound conformation of the ligand. The possible spin-diffusion-related trNOE cross-peaks can be identified by trROE where the alternation of the sign is diagnostic of the real and artifact peaks. Adapted with permission of RSC *MedChemComm* and Elsevier *Curr Opin Struct Biol* from NMR and molecular recognition. The application of ligand-based NMR methods to monitor molecular interactions. Unione L., Galante S., Diaz D., Cañada J., Jiménez-Barbero J. 5, 2014, and New structural insights into carbohydrate-protein interactions from NMR spectroscopy. Kogelberg H., Solís D., Jiménez-Barbero J. 13, 5, 2013; permission conveyed through Copyright Clearance Center, Inc. [27,28].

Two of the biggest challenges for a trNOE experiment are the very narrow accessible kinetic window and the possible existence of spin-diffusion-associated effects. Regarding the first obstacle, for a trNOE experiment to be successful, the on-off rates must be fast relative to the spin relaxation rates. Otherwise, the NOE information acquired by the ligand at the protein interaction will be lost or decreased by relaxation properties before the ligand gets dissociated from the protein. This phenomenon is also dependent on affinity and thus the K_d of the complex. To optimize the conditions for the trNOE experiment, parameters such as temperature, ionic strength and protein:GAG oligosaccharide molar ratio must be carefully controlled. The second obstacle can be circumvented while using rotating-frame Overhauser effect (ROE). ROE cross-peaks have different signs depending on whether they are produced by direct or three-spin indirect effects [29] (Figure 6), allowing cross-peaks of direct interactions to be identified. Like the STD experiment, trNOE investigates the GAG-protein complexes through the GAG (ligand) perspective.

An example of the use of trNOE and trROE in a protein-GAG oligosaccharide investigation is the work of Künze et al., who studied the binding of a heparin tetrasaccharide to interleukin-10 (IL-10) [30]. The structure of the heparin fragment that was used in this work is Δ UA2S-GlcNS6S-IdoA2S-GlcNS6S (see Table 1 for abbreviations). Δ UA2S stands for 4,5-unsaturated 2-sulfated uronic acid. Both NOESY and ROESY spectra were collected at 150 ms mixing time on heparin tetrasaccharide alone and in the presence of IL-10. Experiments were recorded with 20-fold excess of the ligand. NOE/ROE signals of the GAG ligand at the free-state were close to zero and did not interfere with the signals observed in the bound state. In the absence of protein, positive NOEs were observed between protons H₂ and H₃, H₄, and H₅ of GlcNS6S at the non-reducing end disaccharide and between H₄ and H₅ of GlcNS6S of the reducing end. In addition, some small negative NOEs appeared between H₃ and H₄ of the IdoA2S, between H₃ and H₄ of GlcNS6S of the non-reducing end disaccharide and between H₆ protons of both GlcNS6S. Multiple strong negative NOEs were observed in the heparin tetrasaccharide in the

presence of 5 mol % of IL-10, aside from the single positive NOE between H₂ and H₃ of the GlcNS6S at the non-reducing end disaccharide. The set of trNOE observed in the experiment indicates proper molecular interaction between IL-10 and heparin since the average correlation time of the ligand has obviously increased significantly to produce negative NOE signals. trROE was also used to identify artifacts produced by spin-diffusion in the same study.

Another successful application of the trNOE technique is the recent publication of Gao et al. about the intermolecular interaction of Fx with the two N-terminal immunoglobulin domains (Ig1-2) of leukocyte common antigen-related protein (LAR) [31]. The temperatures and magnetic field was optimized so that the free ligand in solution produced no NOE cross-peaks (~30 °C at 800 MHz for Fx), therefore allowing for cross-peak intensities in trNOE to be interpreted while using the simple $1/r^6$ relationship. A set of NOE and trNOE-derived interproton distances (Table 2) were generated through these experiments and used for assessing conformational change of the ligand from the free- to the bound-state. Besides assessing information regarding the ¹C₄ chair and ²S₀ skew-boat configurations of the IdoA2S ring of Fx, the trNOE-derived data was also useful as conformational restraints in the docking process [31].

Table 2. Nuclear Overhauser effect (NOE)- and trNOE-derived interproton distances of Fx free and bound to LAR-Ig1-2. Adapted with permission from Gao et al. Biochemistry 2018, 57, 2189–2199. Copyright. 2018 American Chemical Society [31].

¹ H- ¹ H Connection ^a	Free ^b	Bound
AH1-BH4	2.62 ± 0.02 ^c	2.86 ± 0.21
CH1-DH4	2.41 ± 0.02	2.65 ± 0.20
CH3-DH1	2.43 ± 0.11	3.05 ± 0.25
DH1-DH2	3.16 ± 0.01	-
DH1-DH3	3.27 ± 0.01	2.72 ± 0.19
DH1-EH4	2.71 ± 0.02	3.14 ± 0.25
DH1-EH6	2.85 ± 0.02	2.91 ± 0.23
DH1-EH6'	3.22 ± 0.03	2.49 ± 0.19
OMe-EH1	2.96 ± 0.02	3.22 ± 0.20

^a The structure and nomenclature used were GlcNS6S-GlcA-GlcNS3S6S-IdoA2S-GlcNS6S-OMe and A-B-C-D-E-OMe.

^b The distance of 2.50 Å between H₂ and H₄ of the GlcNS6S ring A was used reference to calculate the other NOE- and trNOE-derived interatomic distances. ^c Errors were estimated based on the signal-to-noise of the individual peaks.

2.4. WaterLOGSY

Water-ligand observed via gradient spectroscopy (waterLOGSY) is an experiment whose mechanism is akin to STD and trNOE, but it is focused on identifying interactions between water and biomacromolecules [32]. The experiment was originally devised to confirm protein-ligand interactions mediated by water. It has since become a valuable tool in protein-glycan studies because interactions between proteins and glycans are often mediated by water, especially when the interactions are electrostatic in nature. Figure 7 illustrates the mechanism of waterLOGSY experiment. The experiment relies on the fact that trapping of water molecules in intermolecular niches between glycan and the target protein slows the motions of water molecules, leading to cross-relaxation with adjacent glycan and protein protons and the production of detectable cross-peaks in NOESY. In practice, the experiment is implemented as a one-dimensional selective NOE experiment where only the water magnetization is inverted. Although use of waterLOGSY in studies of protein-glycan interactions is common, an example involving GAG has been published only recently. Interestingly, in that study, Mandaliti et al. did not use the experiment to probe the ligand, but to identify protein residues that are involved in hyaluronan binding [33], providing a good demonstration of the versatility of the experiment. Although information from waterLOGSY is limited to hydration, it can often be used to create more accurate models and provide a deeper understanding of the energetics of binding. A quantitative utilization of the technique, referred to as LOGSY titration, was published recently. In this study the authors examined the waterLOGSY signal of the ligand at different protein concentrations to accurately

identify ligand protons involved in water mediated interactions in the protein-ligand complex [34]. Such an experiment can be easily applied to the investigation of protein-GAG systems to pinpoint locations of water at the protein-GAG interface.

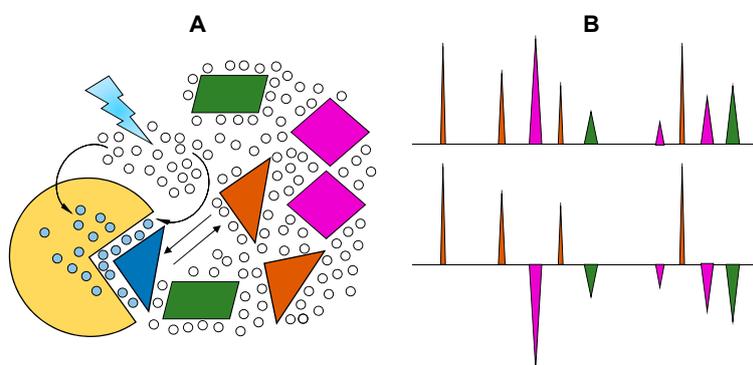


Figure 7. Schematic description of the waterLOGSY experiment. (A) The magnetization of the bulk water is selectively transferred, via the protein-ligand complex, to binding ligands (triangles). (B) Hypothetical representations of NMR spectra from a mixture (top) and from the resultant one-dimensional (1D) waterLOGSY spectrum (bottom), where binding and non-binding ligands can be straightforwardly identified. The positive-phase signals (brown peaks) are produced from binding ligands interacting with protein (blue and brown triangles) and cross-relax with protein-bound water, whose magnetization is selectively inverted. The peaks in negative-phase (green and purple peaks) come from the non-binding ligands which do not interact with inverted water molecules bound to the protein. Adapted with permission of RSC *MedChemComm* from NMR and molecular recognition. The application of ligand-based NMR methods to monitor molecular interactions.Unione L., Galante S., Diaz D., Cañada J., Jiménez-Barbero J. 5, 2014; permission conveyed through Copyright Clearance Center, Inc. [27].

3. Production of GAG Ligands for NMR

As mentioned above, native GAGs have two structural characteristics that make them not ideal for structural studies: (i) very complex structures with variable lengths (polydispersity); and, (ii) high MW chains that lead to multivalent binding and protein oligomerization and/or precipitation. As a result, only small oligosaccharides with defined structure can be used in high resolution structural studies. Currently, two methods are being employed to generate these defined GAG fragments: (i) purification of oligosaccharides derived from the enzymatic depolymerization of GAG from natural sources; and, (ii) synthesis of GAG-derived oligosaccharides through multiple chemical and enzymatic steps. Recently, we discussed the protocols that were used for chemoenzymatic synthesis and the GAG-like oligosaccharide libraries [35]. In the next section, we briefly discuss the three-step process used to prepare GAG oligosaccharides derived from natural sources.

Another obstacle for the NMR study of GAGs is that, because prokaryotic organisms that are capable of growing in minimal media do not produce sulfated GAGs, there is yet no simple method to produce isotopically enriched GAG polysaccharides. Despite these difficulties, some progress has been made to produce isotopically labeled GAG ligands for NMR study. The following sections are dedicated to explaining some of these methods as well as techniques for paramagnetically labeled GAG oligosaccharides for NMR studies.

3.1. Unlabeled Oligosaccharides

Structurally homogeneous GAG fragments can be obtained by a general protocol involving three basic steps (Figure 8). First, the native GAG polymers (Figure 8A) are partially digested by GAG-degrading enzymes, such as hydrolases or lyases (Figure 8B). In this initial step, digestions must be very controlled in order to produce large amounts of medium-sized oligosaccharides, such as

hexasaccharides and octasaccharides. Otherwise, the over-digested materials will produce mostly short-sized fragments, such as disaccharides and tetrasaccharides. Then, the mixture of digested materials is separated based on size while using size-exclusion chromatography (SEC) (Figure 8C). Lastly, the fragments of uniform sizes are re-fractionated based on structure and charge through strong-anion exchange (SAX) chromatography, which is usually performed in automated systems, such as high-pressure liquid chromatography (HPLC) (Figure 8D).

Since the resultant fragments are products of enzymatic digestions, sulfation patterns that are observed in the generated products are controlled by enzyme specificity and they may not reflect distribution of structural motifs in native GAG. This is a downside of the protocol because a particular structure of interest with important biological relevance may not be obtainable. Due to this limitation, GAG-protein investigations performed in the laboratory may not reflect the full spectrum of GAG motifs the protein is exposed to in vivo. Nonetheless, this method still remains the most practical way for generating structurally homogeneous ligands for further GAG-protein studies [36]. Another obstacle is the availability of GAG-digestive enzymes in reasonable amounts for studies. Although many GAG hydrolases and lyases are commercially available, their preparations are not consistent since they are isolates from natural sources. The inhomogeneous preparations of these enzymes impair reproducibility in structure and size distributions of the products. Although recombinant GAG-digestive enzymes exist and are more reliable in terms of homogeneity, they are not widely commercially available.

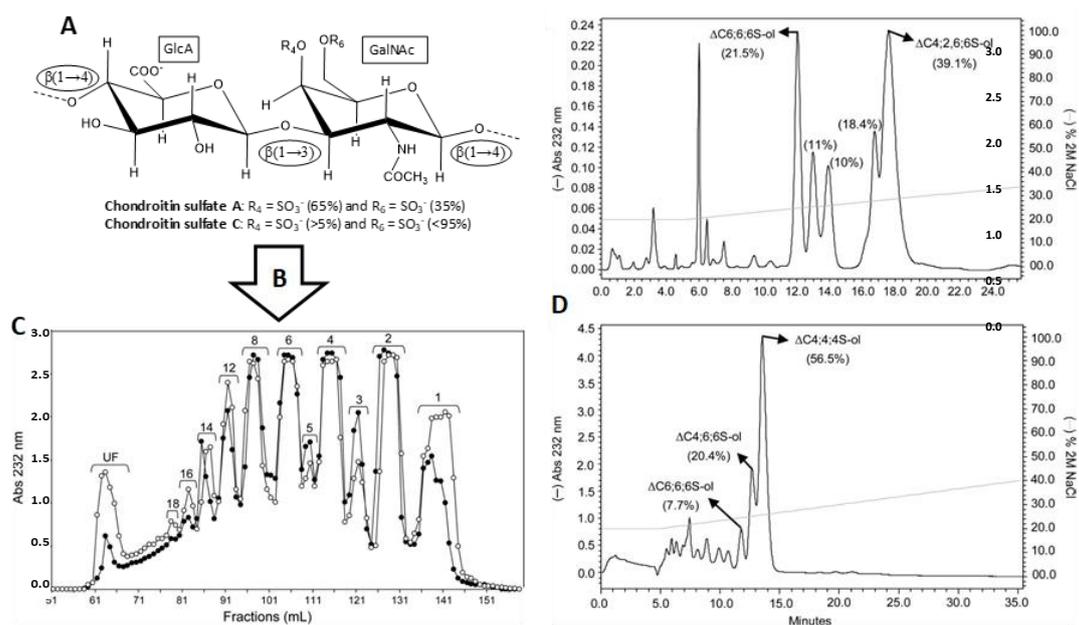


Figure 8. General scheme for production of GAG oligosaccharides of defined size and structures, exemplified by the various hexasaccharides of chondroitin sulfates A and C. (A) Structural representation of the chondroitin sulfate A and C. GlcA and GalNAc stand for glucuronic acid and N-acetylgalactosamine, respectively. Glycosidic linkages are depicted inside the ellipses. (B) After a 2-day digestion of chondroitin sulfate A derived from bovine trachea (btCS-A) or chondroitin sulfate C derived from shark cartilage (scCS-C) with a commercial preparation of chondroitinase C from *Flavobacterium heparinum*, the resultant mixture of fragments were applied in a SEC Bio-Gel P-10 column (C) for oligosaccharide fractionation (open circles for scCS-C and filled circles for btCS-A). Then, the pool of hexasaccharides was applied to a HPLC-SAX column for fractionation of isomers. (D) Both mixtures of hexasaccharides were obtained from peak 6 of their respective Bio-Gel P-10 chromatograms.

The characterized structures of the fractionated isomers are the following: Top chromatogram: $\Delta C6;6;6S$ -ol ($\Delta U A-(\beta 1-3)$ -GalNAc6S-($\beta 1-4$)-GlcA-($\beta 1-3$)-GalNAc6S-($\beta 1-4$)-GlcA($\beta 1-3$)-GalNAc6S-ol) and $\Delta C4;2,6;6S$ -ol ($\Delta U A-(\beta 1-3)$ -GalNAc4S-($\beta 1-4$)-GlcA2S-($\beta 1-3$)-GalNAc6S-($\beta 1-4$)-GlcA-($\beta 1-3$)-GalNAc6S-ol) for scCS-C digestion. Bottom chromatogram: $\Delta C6;6;6S$ -ol, $\Delta C4;6;6S$ -ol ($\Delta U A-(\beta 1-3)$ -GalNAc4S-($\beta 1-4$)-GlcA-($\beta 1-3$)-GalNAc6S-($\beta 1-4$)-GlcA($\beta 1-3$)-GalNAc6S-ol) and $\Delta C4;4;4S$ -ol ($\Delta U A-(\beta 1-3)$ -GalNAc4S-($\beta 1-4$)-GlcA-($\beta 1-3$)-GalNAc4S-($\beta 1-4$)-GlcA($\beta 1-3$)-GalNAc4S-ol) for btCS-A digestion. The nomenclatures $\Delta U A$, "S", and -ol used for the structures stand, respectively, for $\Delta 4,5$ unsaturated uronic acid, sulfation group, which the number before "S" represent the ring position, and reduced sugars (open rings at the reducing-ends). The percentage of material yielded in each peak is given in parentheses. The NaCl gradient is shown with the continuous light grey line. Modified with permission from Oxford University Press. License in conditions provided by Oxford University Press and Copyright Clearance Center, Inc. [37].

3.2. Isotopically Labeled Oligosaccharides

Although there are great barriers in total chemical synthesis of GAGs, fortunately GAG precursors, such as heparosan [-4]-GlcNAc-($\alpha 1-4$)-GlcA-($\beta 1$ -)]_n, can be found in capsules surrounding *Escherichia coli* K5 and *Pasteurella multocida*. This has been exploited in biotechnology for chemoenzymatic synthesis of heparin- and heparin sulfate-like polymers [38]. Since the growth of *E. coli* K5 in minimal media supplemented with ^{13}C -glucose and ^{15}N - $(NH_4)_2SO_4$ is easily achievable, $^{13}C,^{15}N$ -doubly labeled heparosan-derived heparan sulfate and heparin can be produced in small but sufficient quantities [39], and can be used as ligands in protein interaction studies [40]. No prokaryotic system has been reported so far to obtain chondroitin sulfate precursors, but efforts are in progress to obtain metabolically engineered *E. coli* for production of such precursors [41].

Another alternative is to generate isotopically labeled GAGs while using mammalian cells culture. An example uses commercially available side chain ^{15}N -labeled glutamine to produce ^{15}N -labeled amino sugars and consequently GAGs in Chinese Hamster Ovary cells and human lung endothelial cells. (Figure 9) [42]. These ^{15}N -labeled GAGs were structurally characterized via ^{15}N -NMR, whose signals have proven to be sensitive and diagnostic to a series of GAG structural features, such as lengths (disaccharides, oligosaccharides, and even high-order polymers), anomericity, glycosylation sites, hexosamine types, sulfation types (*O*-linked or *N*-linked) and positions, and neighboring uronate types (IdoA or GlcA) [43]. Although the GAG quantity produced in the above example was never sufficient for biochemical studies with GAG-binding proteins, attempts are underway to produce GAG in large scale cell cultures [44]. The ^{15}N -labeling on these ligands can be exploited as a structural probe in GAG-protein interaction studies. In particular, ^{15}N chemical shifts in these ligands were very sensitive to the surrounding environment, therefore protein-induced chemical shift perturbations on ^{15}N resonances of these GAG ligands would offer a strategy to monitor interactions with GAG-binding proteins.

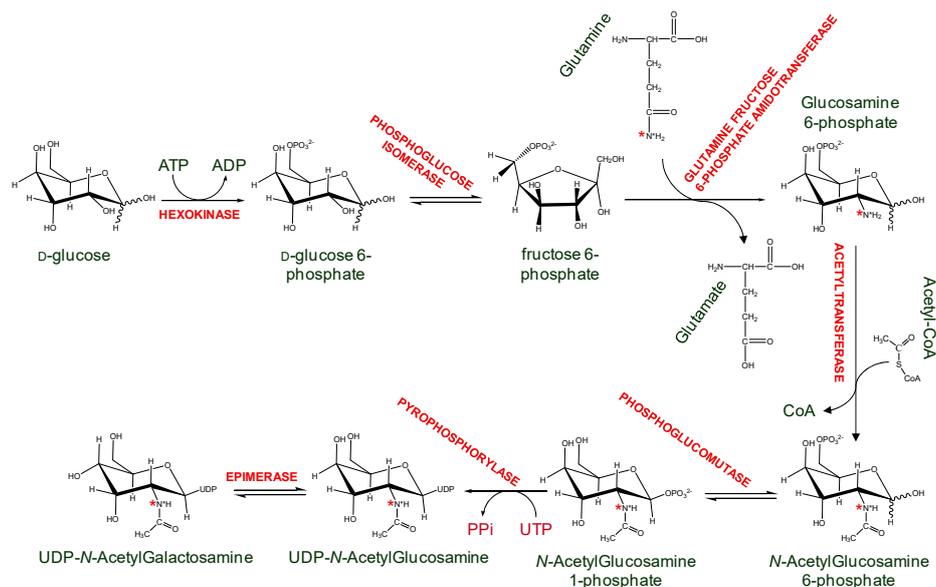


Figure 9. Schematic representation of the cytosolic biosynthesis of ^{15}N -labeled UDP-*N*-acetylgalactosamine and UDP-*N*-acetylglucosamine, using side chain ^{15}N -labeled glutamine as ^{15}N donor. These nucleotide-*N*-acetyl-hexosamines will be further used inside the Golgi apparatus to synthesize GAG backbones in proteoglycans. The names of the enzymes are fully written in red while the substrates and products are mostly written in black. The ^{15}N -labeling are indicated with the red asterisks throughout the reaction steps. Adapted with permission from Pomin et al. Analytical Chemistry 2010, 82, 4078–4088. Copyright 2018 American Chemical Society [42].

3.3. Paramagnetically Labeled Oligosaccharides

Paramagnetically labeled GAG ligands have proven to be useful in structural studies of GAGs' interactions with several different proteins [31,45–49]. The paramagnetic effect relies on the dipole-dipole interactions between unpaired electrons and surrounding atoms. Such interactions produce relaxation effects on nearby nuclei. The extent of the paramagnetism-induced increase in transverse relaxation of proton can be calculated using the following Equation (3):

$$R_{2,\text{pre}} = \frac{1}{15} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\gamma_H^2 g_s^2 \mu_B^2 S(S+1)}{r^6} \left(4\tau_c + \frac{3\tau_c}{1 + \omega_H^2 \tau_c^2} \right) \quad (3)$$

where $R_{2,\text{pre}}$ is the increase in transverse relaxation rates due to the paramagnetic center, r is the distance between the paramagnetic center and the proton under analysis, γ_H is the gyromagnetic ratio of proton, g_s is the electron g factor for the unpaired electron, ω_H is the resonance frequency of the proton under analysis, and τ_c is the overall rotational correlation time of the whole system. This effect is called PRE and it can be accurately quantified using specific experiments [50,51]. However, paramagnetism can change more than the relaxation rate of a NMR signal. As seen in Figure 10, all paramagnetic labels can produce NMR signal line-broadening ($\Delta\nu_2$ is bigger than $\Delta\nu_1$) on nearby atoms, but paramagnetic center with asymmetric g -tensors and favorable relaxation properties can also produce chemical shift displacements ($\Delta\delta$) in these atoms. These chemical shift changes are referred to as pseudo-contact shifts.

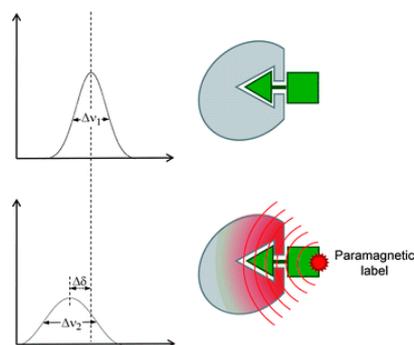


Figure 10. Paramagnetically tagged ligands affect the relaxation properties of their binding partners causing two major effects: pseudo-contact shifts ($\Delta\delta$) and line-broadening ($\Delta\nu_2 > \Delta\nu_1$). These effects can be useful for mapping the binding site of the ligand. Adapted with permission of RSC *Nat. Prod. Rep.* from Application of NMR methods to the study of the interaction of natural products with biomolecular receptors. Calle L.P., Cañada F.J., Jiménez-Barbero J. 28, 6, 2011; permission conveyed through Copyright Clearance Center, Inc. [52].

Since the magnitude of PRE is primarily dependent on the distance between the paramagnetic center and nearby atoms, data from this technique are simpler to interpret than CSP, which is also sensitive to allosteric effects. Therefore, PRE is a reliable validation of CSP data. Another key advantage of PRE is that the effective range of the paramagnetic effect can be up to 20 Å. This is ideal for detecting long-range or transient contacts between GAGs and proteins. In several works [45–47], GAG ligands have been functionalized with the paramagnetic tag TEMPO (2,2,6,6-tetramethylpiperidine-1-oxyl) through the reductive amination of the reducing end of the oligosaccharide (Figure 11). This method has the distinct advantage of being specific to the reducing end of the GAG ligand, allowing for the location of the reducing end of the ligand to be identified.

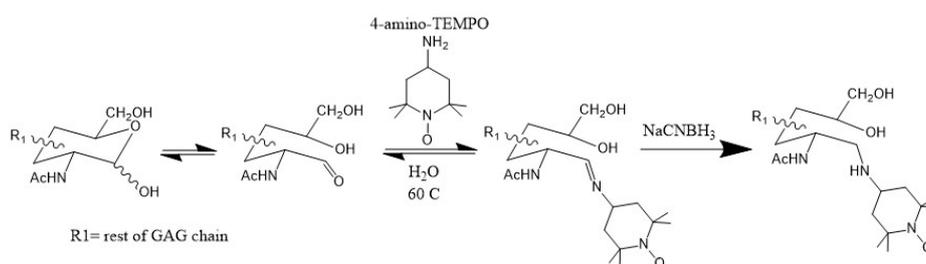


Figure 11. Reductive amination of the reducing end of glycans using 4-amino-TEMPO.

A particular interesting application of the technique was provided by Deshauer et al [45]. In this study, TEMPO-tagged chondroitin sulfate hexasaccharides were used to investigate GAG interactions of the chemokine CCL5/RANTES. Figure 12 revisits the data. The ^{15}N HSQC spectra of RANTES in the presence of either oxidized or reduced TEMPO-labeled chondroitin sulfate hexasaccharide CS6;4;4 (GlcA-(β 1-3)-GalNAc6S-(β 1-4)-GlcA-(β 1-3)-GalNAc4S-(β 1-4)-GlcA(β 1-3)-GalNAc4S-TEMPO) are shown in Figure 10A. When comparing the ^{15}N -HSQC spectra, several signals have disappeared upon the oxidation of TEMPO. Those belonging to residues A16 and R21 are highlighted in the upper spectrum. This indicates these residues are in close contact with the paramagnetically tagged ligand. Quantitative measurements of the ^1H PRE of the amide protons were carried out (graph on Figure 12B) and used to identify the binding site on RANTES (Figure 12B). This approach was also similarly employed in investigations of GAG binding by the bacterial adhesion decorin-binding proteins A and B [47,53].

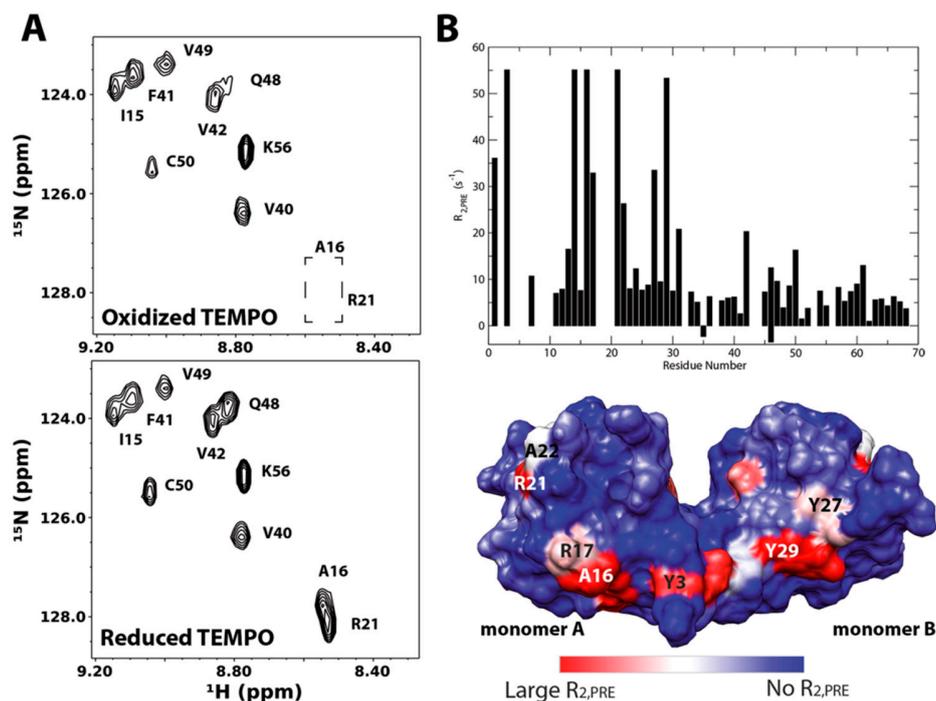


Figure 12. (A) ^{15}N -HSQC spectra of RANTES in the presence of oxidized (top panel) and reduced (bottom panel) forms of TEMPO-labeled chondroitin sulfate hexasaccharide CS6;4;4. (B) Quantitative measurements of paramagnetic TEMPO-induced transverse relaxation ($R_{2,\text{pre}}$). Graph and surface plot of amide proton $R_{2,\text{pre}}$ produced by TEMPO-labeled CS6;4;4. Residues Y14, A16 and R21 disappeared upon addition of the ligand, but re-appeared upon its reduction. They are given the same magnitude of $R_{2,\text{pre}}$ as residue Y3, which showed the largest PRE effect. Reproduced with permission of RSC New developments in NMR. NMR in Glycoscience and Glycotechnology. Chapter 11. Xu Wang, page 264, 2017 [36].

4. Characterizing Lysine and Arginine Residues in GAG-Binding Sites

4.1. Direct Observation of Lysine Amine Signals

Because of their ability to participate in ionic interactions with GAGs, lysines and arginines are the most important residues in GAG-binding sites. However, studying these side chains is not routine. Specifically, the positively charged functional groups at the termini of their side chains contain labile protons that are not usually NMR detectable. Since interactions are mediated by these functional groups, their invisibility in NMR meant that much of GAG-protein interactions remains undetectable. This problem is especially acute with lysines since protons on their terminal amine are highly exchangeable, whereas some information on arginine can usually be gleaned from $\text{H}\epsilon\text{-N}\epsilon$ of arginine while using only ^{15}N -HSQC.

The group of Junji Iwahara has contributed greatly to NMR investigations of lysine side chain amine groups. In particular, Iwahara and co-workers have shown that the heteronuclear in-phase single quantum coherence experiment (HISQC) is more sensitive at detecting side chain amine signals because of its ability to reduce the relaxation effect of hydrogen exchange on ^{15}N atoms [54]. The Iwahara group also developed much of the theory behind the relaxation of ^{15}N atoms in $-\text{NH}_3^+$ groups and showed that DNA binding to a protein changes dynamics of some of its lysine and arginine residues [55,56]. Their methodology is easily adaptable to GAG-binding proteins. The first instance of these is reported by Sepuru et al. in the study of the interactions of chemokines CXCL1 and CXCL5 with heparin. They noted that, in the absence of GAG octasaccharides, very few lysine side chain amine signals from the two chemokines can be seen even at pH 5.7 and 10 °C. However, the addition of several molar

equivalents of GAG octasaccharides greatly increased the number and intensity of detectable amine signals [57].

We have applied the same methodology to the cytokine pleiotrophin. Figure 13 shows lysine side chain amino HSQC and HISQC of pleiotrophin. The HISQC showed significant enhancements in intensity when compared to the HSQC. However, in both cases, signals can only be observed at pH 6.0 and low temperature (~10 °C). Unlike CXCL1/5, there is little dispersion in the lysine side chain signals of pleiotrophin, which reflect the fact that many of the lysines in pleiotrophin are located in the unstructured C-terminal tail. Similar to CXCL1/5, the presence of GAG dramatically enhanced the intensities of the lysine signals of pleiotrophin, indicating that interactions between GAG and protein likely reduced solvent exchange rates of amine protons. Despite the change in signal intensity, there is no significant change in the chemical shifts of the signals, which is indicative of highly dynamic interactions.

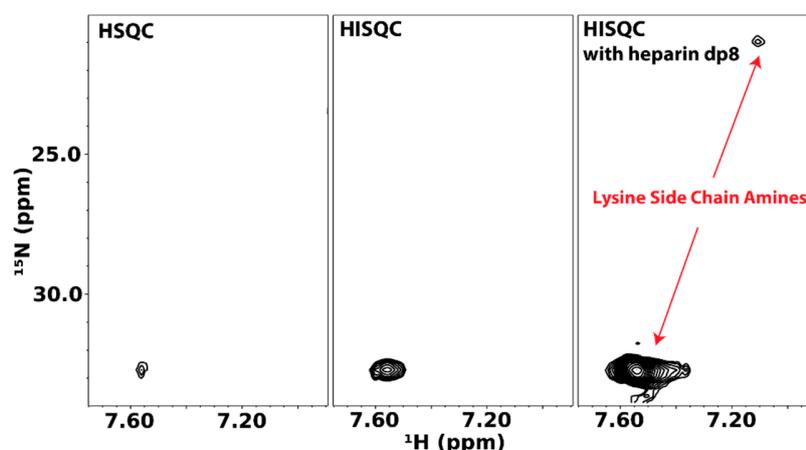


Figure 13. Lysine side chain amine ^{15}N -HSQC and HISQC of pleiotrophin in the presence and absence of heparin octasaccharides. HISQC of the protein produced significantly stronger signals than HSQC. Addition of heparin octasaccharides also induced appearance of additional signals.

Because GAG-induced chemical shift changes in lysine side chain signals can be small in many instances, it can be difficult to identify lysines that are involved in GAG binding by direct examination of the HISQC spectrum. Uhrin and co-workers showed that a more sensitive way to identify GAG-binding lysines is through direct cross-linking with GAGs [58]. To do this, they prepared GAG ligands whose carboxyl groups are activated with sulfo-*N*-hydroxysuccinimide. Activation of carboxyl groups on these ligands led to amidation reactions between the activated carboxyls on GAGs and nearby lysine side chains on proteins. The conversion of amine to amide eliminated signals of these lysines from the HISQC spectrum, allowing for them to be identified easily. It should be noted that the identity of these modified lysines can also be obtained through mass spectrometry.

4.2. Specific Labeling of Lysine and Arginine Side Chains

Detecting intermolecular NOE cross-peaks between protein and GAG is usually quite difficult. Many factors contribute to this, including a lack of non-labile protons at the interface and the often dynamic nature of interactions between protein and GAGs. Compounding the difficulty is that the traditional $^{13}\text{C}/^{15}\text{N}$ -filtered/edited NOESY experiment used to obtain intermolecular contacts is low in sensitivity as a result of multiple isotopic filters and the need for ^{13}C -labeling. To increase the sensitivity of the NOESY experiments, we have been exploring the use of deuterated protein containing specifically protonated lysines and arginines as a way to improve the sensitivity of the NOESY experiment to intermolecular contacts. By avoiding ^{13}C -labeling, we increase the transverse relaxation time of protons and improve resolution in the both direct and indirect dimensions. However, a lack of ^{13}C -labeling means spectral crowding cannot be resolved with ^{13}C -edited experiments.

Fortunately, specifically protonating the amino acids of interest while deuterating other amino acids mostly alleviate the problem. This method of specific labeling has the additional benefit of reducing spin diffusion and longitudinal relaxation, which can further enhance the sensitivity of NOESY to crucial long range intermolecular NOEs [59]. It also reduces transverse relaxation and eliminates some ^1H homonuclear coupling, leading to additional increases in the signal-to-noise. However, it should be noted that the reduction in longitudinal relaxation requires increases in recycle delay time in order to re-establish the equilibrium magnetization.

To examine the efficiency and specificity of labeling, we have prepared ^2H pleiotrophin containing either protonated lysines or arginines. Figure 14 shows the NOESY spectra of these samples. All lysines in the structured region of the protein can be identified. Most lysines have strong $\text{H}\alpha$ signals, indicating that the deamination of the amino acid is not extensive. Attempts at specific labeling of arginine also produced relatively clean spectra, but $\text{H}\alpha$ correlations of arginines are generally weak, a sign that the amino acid underwent significant deamination. Despite the weak $\text{H}\alpha$ signals, correlations between other protons in the side chain are visible and three out of five arginines can be identified. Addition of heparin tetrasaccharide to the specifically lysine labeled sample showed that heparin can induce drastic changes in the $\text{H}\alpha$ chemical shifts of K92 and K93, two lysines identified as crucial for GAG binding in mutagenesis studies. The higher resolution of homonuclear NOESY also allowed identification of K93 as having strong interactions with the uronic acids in the tetrasaccharide, but not glucosamine residues. In fact, both the non-reducing end unsaturated uronate and the iduronate close to the reducing end have similar contacts to K93. This indicates the tetrasaccharide is likely sliding along the GAG binding site and it does not have a single binding position. One remaining obstacle to the present methodology is that single lysine labeling is not possible. However, this can potentially be solved while using the amber codon technology and pyrrolysyl-tRNA synthetase, which allow for the incorporation of a single lysine terminated with the removable protecting group BOC.

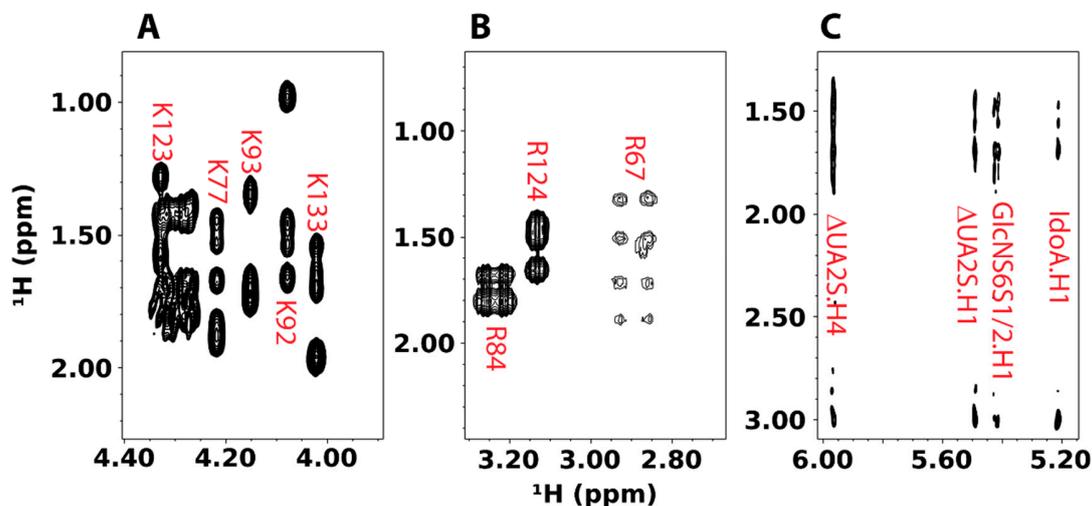


Figure 14. 2D homonuclear NOESY spectra of deuterated pleiotrophin containing either ^1H -labeled lysine or arginine. (A) $\text{H}\alpha$ region of the NOESY spectrum of ^2H -pleiotrophin with specific ^1H lysine labeling. Signals from lysines in the structured regions of pleiotrophin are labeled in red. (B) $\text{H}\delta$ region of the NOESY spectrum of ^2H -pleiotrophin with specific ^1H arginine labeling. Signals from assigned arginines are labeled in red. (C) NOESY cross-peaks arising from intermolecular contacts involving GAG anomeric protons of 0.2 mM ^2H -pleiotrophin with ^1H lysine labeling in the presence of 2 mM heparin tetrasaccharide, $\Delta\text{UA}2\text{S-GlcNS}6\text{S-IdoA}2\text{S-GlcNS}6\text{S}$.

5. Probing GAG-Induced Changes in Protein Dynamics and Oligomerization

Interactions of GAG with target proteins often result in oligomerization and aggregation of the target. GAG-induced oligomerization is an important aspect of GAG's activity and is crucial to the

activation of many signaling proteins, including chemokines, growth factors, and membrane receptors. Knowledge of the protein oligomer formed in the presence of GAG is often vital to understanding the mechanism of signaling events. NMR is very sensitive to changes in protein oligomerization, which increases the transverse relaxation time and leads to peak broadening. In fact, by measuring spin longitudinal (R_1) and transverse (R_2) relaxation rates of protein complexes, the total correlation time of the complex can be calculated, and because MW is known to be correlated with the total correlation time, the size of the GAG-induced oligomer can be estimated. The detailed instruction on the procedure can be found on the website of the Northeast Structural Genomics Consortium (http://www.nmr2.buffalo.edu/nesc/wiki/NMR_determined_Rotational_correlation_time). In some instances, such as the case of the chemokine CCL5/RANTES, the monomeric and oligomeric species are in slow exchange on the NMR timescale under the right conditions, and therefore can be separately identified in the NMR spectrum (Figure 15). In this particular case, the disappearance of the monomeric species in the presence of chondroitin sulfate hexasaccharide offered conclusive evidence that GAG-induced oligomerization has occurred.

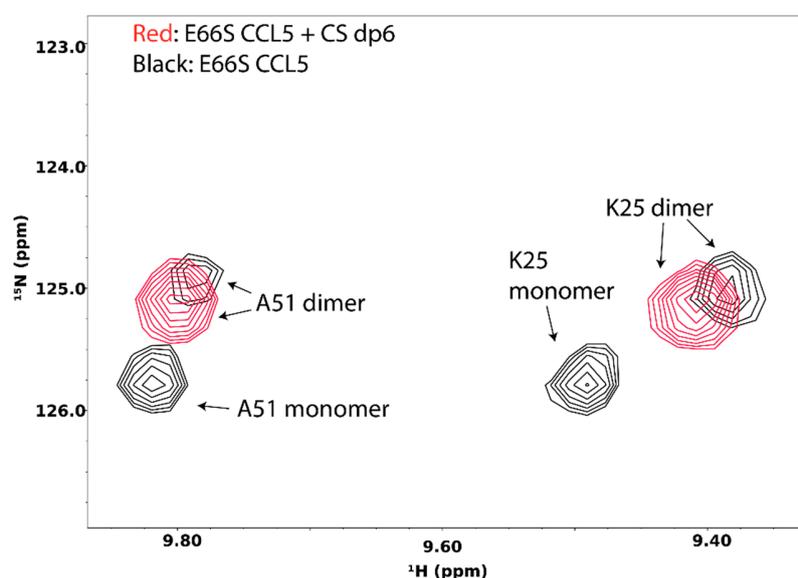


Figure 15. ^{15}N -HSQC spectra of CCL5 in the presence (red) and absence (black) of chondroitin sulfate hexasaccharide. At 40 °C, CCL5 exists in both dimer and monomer forms. However, addition of chondroitin sulfate hexasaccharide induces the formation of CCL5 dimer, leaving only dimeric CCL5 signals in the ^{15}N -HSQC spectrum.

Besides the use of global dynamics in the characterization of GAG-induced oligomerization, quantitative characterization of GAG-induced local changes in protein dynamics can also be done. NMR is an excellent tool for studying protein dynamics at a wide range of timescales. For fast motion on the ps-to-ns timescale in ^{15}N atoms, the model-free formalism can be used to estimate local motion of individual atoms using their heteronuclear NOE, R_2 , and R_1 relaxation rates [60]. For slow motion on the μs -to- ms timescale, relaxation dispersion experiments can be used to estimate the frequency of the motion as well as the change in chemical shift of the atom as a result of the motion. Such analyses have been carried out on several protein-GAG systems. One example comes from the study of vascular endothelial growth factor-165 (VEGF₁₆₅) by Kim et al. [61]. They looked at the effects of GAGs on slow timescale motions of ^{15}N atoms in the protein backbone in heparin binding sites of VEGF₁₆₅ while using relaxation dispersion experiments. They noticed that many residues in the GAG-binding sites exhibit dynamic motion in the ms timescale, but these motions were decreased dramatically by the presence of heparin octasaccharide, indicating heparin has stabilized the GAG-binding site.

On the other hand, Feng et al. looked at GAG-induced slow timescale motion of decorin-binding protein B (DBPB), a GAG-binding surface adhesin that is found on the Lyme disease bacterium

Borrelia burgdorferi. They found that GAG binding does not change the ps-to-ns timescale motion of DBPB. However, dermatan sulfate oligosaccharides were able to induce slow timescale motion of the residues in the GAG-binding site [46]. Interestingly, this phenomenon is only seen with dermatan sulfate and not other GAGs, such as heparin, indicating that the kinetics of interaction between DBPB and GAG is GAG-type specific. Methods also exist to look at the relaxation of the ^{15}N atom in lysine side chain amine groups. They have been productively applied to the study of DNA-binding proteins [55,56]. Although there is currently no report of application of these methods to GAG-binding proteins, the utilization of these methods on GAG-binding proteins should shed important insights into protein-GAG interactions.

6. Solid-State NMR Studies of GAG-Protein Interactions

One important biological activity of GAG is their ability to induce formation of amyloids. In fact, GAG is known to promote amyloid formation of a number of proteins, including A β 40/42, α -synuclein, tau, amylin, and calcitonin [62–66]. Because the aggregates of these proteins are beyond the size regime of solution NMR and are not capable of forming crystals, solid state NMR has become a favored tool for investigating interactions of GAG with amyloids. One particular notable series of studies was carried out by the groups of David Middleton and Sheena Radford [67–70]. They discovered that low-MW heparin only associates strongly with A β 40 amyloid, and not A β 42 or shorter peptides. The structure of the amyloid was also found to be important as amyloid with three-fold cross-sectional symmetry has significantly higher affinity for heparin than those with two-fold symmetry. Using solid-state NMR and a synthetic ^{13}C -enriched heparin octasaccharide, they determined that the octasaccharide remains rigid and is in contact with the amyloid throughout. They also examined amino acids involved in binding heparin by looking at signal dephasing caused by dipolar interactions between ^{15}N atoms in lysine, arginine and histidines, and ^{13}C atoms in the heparin octasaccharide. Surprisingly, the contact between heparin and the amyloid was mediated mostly by histidine residues. Since most of the histidines are located at the junction between peptides in the same cross-sectional layer, they were able to construct a model of these GAG-A β 40 amyloid complexes. The results also explained the lower heparin affinity of A β 40 amyloids with two-fold symmetry, since the arrangement of amino acids in two-fold symmetric structures did not allow for correct clustering of the basic amino acids. It should be noted that the availability of ^{13}C -labeled heparin ligands was crucial to their studies. Without this ligand, the important frequency selective rotational echo double resonance experiment used to measure dipole-dipole interactions between ^{15}N and ^{13}C atoms would not be possible.

7. Computational Modeling of Protein-GAG Complexes Using Sparse Constraints

Obtaining a large quantity of experimental high-resolution structural information on GAG-protein systems is usually not feasible. As a result, constructing of models of protein-GAG complexes cannot be accomplished while using conventional NMR structure determination protocols. This is because the default simulation conditions in these procedures use simplified force-fields in either implicit solvent or vacuum models. Although these parameters greatly expedite the calculation, they require more experimental information to arrive at the correct structure, which is unattainable for most protein-GAG systems. The available force field parameters for creating GAG ligands are also lacking in these software. For instance, molecular topology specification in XPLOR-NIH [71] and CNS [72] cannot easily be used to define the puckering of the sugar ring. However, IdoA residues in most GAG ligands adopt the $^1\text{C}_4$ rather than the traditional $^4\text{C}_1$ puckering. The inability to consistently specify ring puckers means that the generation of the correct ligand structure using the software is usually cumbersome. The imprecision in the carbohydrate parameters also means longer calculation times are required to find the minimal energy conformation of the ligand in some cases. This makes obtaining ligand structures that are consistent with experimental data difficult. On the other hand, although fully fledged molecular dynamics simulation software usually have more accurate force field and realistic

simulation conditions, the relatively large size of most protein-GAG systems means that unrestricted molecular dynamics simulation is not likely to produce accurate structures in a realistic time frame.

To overcome these limitations, we have been using a hybrid approach to create realistic models of protein-GAG complexes. In particular, we first obtain the coordinate of the GAG ligands while using AMBER and its GLYCAM force field, which has just been updated to include monosaccharides specific to GAGs [73]. This step also allows us to put the ligand in the correct bound-conformation prior to structure calculation. We then use the structure in XPLOR-NIH or CNS to dock the glycan onto protein using NMR constraints. The lowest energy structures are then further refined in AMBER. Other groups have adopted similar protocols in which an approximate model is constructed while using traditional NMR software, such as XPLOR-NIH, CNS, and HADDOCK [74], and further refined under more realistic environments in molecular dynamics simulations if necessary. Figure 16 displays a general scheme for the workflow in computational modeling of GAG-protein complexes while using NMR data. GAG-binding proteins studied using such an approach include CXCL1 [75], CXCL5 [76], β -defensin [21], Robo-1 [77], and LAR [31]. HADDOCK has proven to be especially popular because of its ability to incorporate a large array of ambiguous data from NMR, including information from chemical shift perturbation, ambiguous NOEs, residual dipolar couplings and pseudo-contact shifts. It should be noted that there are several docking software that are designed specifically for protein-GAG systems [78–81]. However, it is not clear whether they can easily incorporate experimental constraints from NMR.

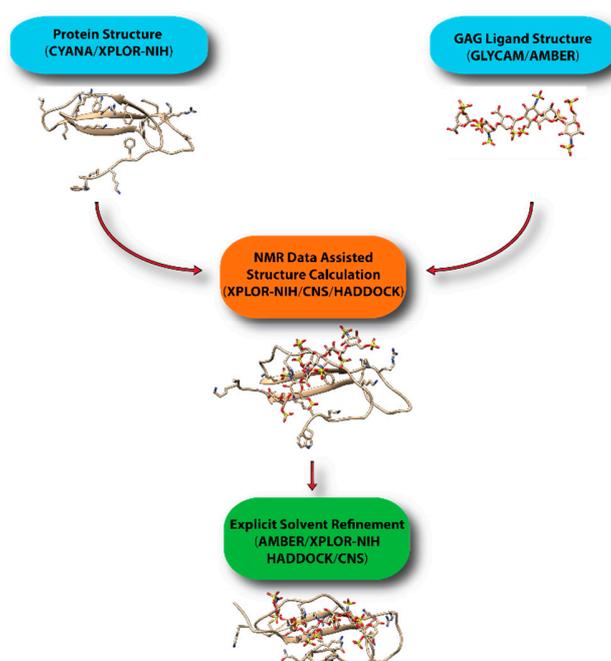


Figure 16. Common workflow in determining structures of protein-GAG complexes using sparse information. To ensure proper ring puckering and stereo chemistry, the initial ligand structure should be created with AMBER using the updated GLYCAM force field. After docking the ligand using experimental data in either XPLOR-NIH or HADDOCK/CNS, an explicit solvent refinement is highly desirable.

8. Conclusions

GAG-dependent regulation of proteins is a pivotal biochemical process in many physiological and pathological events. The understanding of the resultant GAG-protein complexes at the atomic level is necessary to better comprehend the underlying mechanisms that are involved in these events and makes the design of new drugs targeting these proteins possible. NMR spectroscopy is a versatile and robust analytical technique that is often employed in the studies of GAG-protein interactions.

Through isotopic labeling, not only is NMR capable of examining the physicochemical properties of the complexes as a whole, but it also excels at examining changes affected by its binding partner in each individual component. Multiple NMR methods are available or under development for assessing the characteristics of the GAG-protein complexes. Here, we reviewed the theoretical and practical aspects of the most common methods that are found in the literature. The experiments discussed herein were CSP, STD, trNOE , PRE, and waterLOGSY. In addition, more recent methods utilized to assess contribution of arginine and lysine side chains in interactions with GAGs and computational approaches commonly employed to model GAG-protein interactions while using NMR-derived information are also presented. We also discussed the importance of measuring spin relaxation rates for assessing dynamic information of the GAG-protein complexes in solution as well as the preparation of GAG oligosaccharides of defined chemical structures for isotopic and paramagnetic labeling so they can be used as functional probes in GAG-protein interaction studies. The number of investigations in this field has increased considerably during the last few years. These methods have been proven to be essential not only to the structural biologists, but also offer benefits to those professionals working in drug discovery. In this regard, we believe that NMR spectroscopy will continue to play a key role in future investigations of GAG-protein interactions.

Author Contributions: V.H.P. and X.W. searched, critically analyzed and discussed the literature, made the figures and wrote the manuscript.

Funding: V.H.P. is supported by startup funds from the University of Mississippi. X.W. is funded by grants from NIGMS (GM118518) and NCI (U01CA221235).

Acknowledgments: We acknowledge Brian Volkman, guest-editor of the Special Issue “NMR as a Tool to Investigate Biomolecular Structure, Recognition and Dynamics” in MDPI Molecules, and Doris Xiao, managing editor of MDPI *Molecules*, for their kind invitation to collaborate a paper of our authorship to the Special Issue. V.H.P. is grateful to Fatima M.C. Pomin for making Figures 6 and 7. X.W. wants to thank Eathen O. Ryan for preparing the pleiotrophin samples used in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pomin, V.H.; Mulloy, B. Current structural biology of the heparin interactome. *Curr. Opin. Struct. Biol.* **2015**, *34*, 17–25. [[CrossRef](#)] [[PubMed](#)]
2. Gandhi, N.S.; Mancera, R.L. The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **2008**, *72*, 455–482. [[CrossRef](#)] [[PubMed](#)]
3. Proudfoot, A.E.I.; Johnson, Z.; Bonvin, P.; Handel, T.M. Glycosaminoglycan Interactions with Chemokines Add Complexity to a Complex System. *Pharmaceuticals (Basel)* **2017**, *10*, 70. [[CrossRef](#)] [[PubMed](#)]
4. Imberty, A.; Lortat-Jacob, H.; Pérez, S. Structural view of glycosaminoglycan-protein interactions. *Carbohydr. Res.* **2007**, *342*, 430–439. [[CrossRef](#)] [[PubMed](#)]
5. Canales, A.; Angulo, J.; Ojeda, R.; Bruix, M.; Fayos, R.; Lozano, R.; Giménez-Gallego, G.; Martín-Lomas, M.; Nieto, P.M.; Jiménez-Barbero, J. Conformational flexibility of a synthetic glycosylaminoglycan bound to a fibroblast growth factor. FGF-1 recognizes both the (1)C(4) and (2)S(O) conformations of a bioactive heparin-like hexasaccharide. *J. Am. Chem. Soc.* **2005**, *127*, 5778–5779. [[CrossRef](#)] [[PubMed](#)]
6. Pomin, V.H. Heparin-Binding Proteins (Chemokines and Defensins) and their complexes with glycosaminoglycans from the solution NMR perspective. *Curr. Protein Pept. Sci.* **2014**, *15*, 738–744. [[CrossRef](#)] [[PubMed](#)]
7. Mulloy, B.; Forster, M.J.; Jones, C.; Davies, D.B. N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.* **1993**, *293*, 849–858. [[CrossRef](#)] [[PubMed](#)]
8. Shaya, D.; Zhao, W.; Garron, M.L.; Xiao, Z.; Cui, Q.; Zhang, Z.; Sulea, T.; Linhardt, R.J.; Cygler, M. Catalytic mechanism of heparinase II investigated by site-directed mutagenesis and the crystal structure with its substrate. *J. Biol. Chem.* **2010**, *285*, 20051–20061. [[CrossRef](#)] [[PubMed](#)]
9. Michel, G.; Pojasek, K.; Li, Y.; Sulea, T.; Linhardt, R.; Raman, R.; Prabhakar, V.; Sasisekharan, R.; Cygler, M. The Structure of Chondroitin B Lyase Complexed with Glycosaminoglycan Oligosaccharides Unravels a Calcium-Dependent Catalytic Machinery. *J. Biol. Chem.* **2004**, *279*, 32882–32896. [[CrossRef](#)] [[PubMed](#)]

10. Arnott, S.; Gus, J.M.; Hukins, D.W.; Dea, I.C.; Rees, D.A. Conformation of keratan sulphate. *J. Mol. Biol.* **1974**, *88*, 175–184. [[CrossRef](#)]
11. Almond, A.; Deangelis, P.L.; Blundell, C.D. Hyaluronan: The Local Solution Conformation Determined by NMR and Computer Modeling is Close to a Contracted Left-Handed 4-Fold Helix. *J. Mol. Biol.* **2006**, *358*, 1256–1269. [[CrossRef](#)] [[PubMed](#)]
12. Pomin, V.H.; Mulloy, B. Glycosaminoglycans and Proteoglycans. *Pharmaceuticals (Basel)* **2018**, *11*, 27. [[CrossRef](#)] [[PubMed](#)]
13. Liang, W.G.; Triandafillou, C.G.; Huang, T.Y.; Zulueta, M.M.; Banerjee, S.; Dinner, A.R.; Hung, S.C.; Tang, W.J. Structural basis for oligomerization and glycosaminoglycan binding of CCL5 and CCL3. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 5000–5005. [[CrossRef](#)] [[PubMed](#)]
14. Cai, Z.; Yarovoi, S.V.; Zhu, Z.; Rauova, L.; Hayes, V.; Lebedeva, T.; Liu, Q.; Poncz, M.; Arepally, G.; Cines, D.B.; et al. Atomic description of the immune complex involved in heparin-induced thrombocytopenia. *Nat. Commun.* **2015**, *6*, 8277. [[CrossRef](#)] [[PubMed](#)]
15. Wu, L.; Viola, C.M.; Brzozowski, A.M.; Davies, G.J. Structural characterization of human heparanase reveals insights into substrate recognition. *Nat. Struct. Mol. Biol.* **2015**, *22*, 1016–1022. [[CrossRef](#)] [[PubMed](#)]
16. Qin, Y.; Ke, J.; Gu, X.; Fang, J.; Wang, W.; Cong, Q.; Li, J.; Tan, J.; Brunzelle, J.S.; Zhang, C.; et al. Structural and Functional Study of d-Glucuronyl C5-epimerase. *J. Biol. Chem.* **2015**, *290*, 4620–4630. [[CrossRef](#)] [[PubMed](#)]
17. Aguda, A.H.; Panwar, P.; Du, X.; Nguyen, N.T.; Brayer, G.D.; Bromme, D. Structural basis of collagen fiber degradation by cathepsin K. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 17474–17479. [[CrossRef](#)] [[PubMed](#)]
18. Whalen, D.M.; Malinauskas, T.; Gilbert, R.J.C.; Siebold, C. Structural Insights into Proteoglycan-Shaped Hedgehog Signaling. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16420–16425. [[CrossRef](#)] [[PubMed](#)]
19. Pomin, V.H. Biological findings from the recent NMR-based studies of glycosaminoglycan-protein interactions. *Glycobiology* **2014**, *24*, 991–1003. [[CrossRef](#)] [[PubMed](#)]
20. Farmer, B.T., 2nd; Constantine, K.L.; Goldfarb, V.; Friedrichs, M.S.; Wittekind, M.; Yanchunas, J., Jr.; Robertson, J.G.; Mueller, L. Localizing the NADP⁺ binding site on the MurB enzyme by NMR. *Nat. Struct. Mol. Biol.* **1996**, *3*, 995–997. [[CrossRef](#)] [[PubMed](#)]
21. De Paula, V.S.; Pomin, V.H.; Valente, A.P. Unique Properties of Human β -Defensin 6 (hBD6) and Glycosaminoglycan Complex: Sandwich-like dimerization and competition with the chemokine receptor 2 (CCR2) binding site. *J. Biol. Chem.* **2014**, *289*, 22969–22979. [[CrossRef](#)] [[PubMed](#)]
22. Bhunia, A.; Bhattacharjya, S.; Chatterjee, S. Applications of saturation transfer difference NMR in biological systems. *Drug Discov. Today* **2012**, *17*, 505–513. [[CrossRef](#)] [[PubMed](#)]
23. Cala, O.; Guillière, F.; Krimm, I. NMR-based analysis of protein-ligand interactions. *Anal. Bioanal. Chem.* **2014**, *406*, 943–956. [[CrossRef](#)] [[PubMed](#)]
24. Haselhorst, T.; Lamerz, A.-C.; von Itzstein, M. Saturation transfer difference NMR spectroscopy as a technique to investigate protein-carbohydrate interactions in solution. *Methods Mol. Biol.* **2009**, *534*, 375–386. [[PubMed](#)]
25. Yu, F.; Roy, S.; Arevalo, E.; Schaeck, J.; Wang, J.; Holte, K.; Gunay, N.S.; Capila, I.; Kaundinya, G.V. Characterization of heparin-protein interaction by saturation transfer difference (STD) NMR. *Anal. Bioanal. Chem.* **2014**, *406*, 3079–3089. [[CrossRef](#)] [[PubMed](#)]
26. Post, C.B. Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Curr. Opin. Struct. Biol.* **2003**, *13*, 581–588. [[CrossRef](#)] [[PubMed](#)]
27. Unione, L.; Galante, S.; Diaz, D.; Cañada, J.; Jiménez-Barbero, J. NMR and molecular recognition. The application of ligand-based NMR methods to monitor molecular interactions. *Med. Chem. Commun.* **2014**, *5*, 1280–1289. [[CrossRef](#)]
28. Kogelberg, H.; Solís, D.; Jiménez-Barbero, J. New structural insights into carbohydrate-protein interactions from NMR spectroscopy. *Curr. Opin. Struct. Biol.* **2003**, *13*, 646–653. [[CrossRef](#)] [[PubMed](#)]
29. Bax, A.; Sklenar, V.; Summers, M.F. Direct identification of relayed nuclear Overhauser effects. *J. Magn. Reson.* **1986**, *70*, 327–331. [[CrossRef](#)]
30. Künze, G.; Gehrcke, J.-P.; Pisabarro, M.T.; Huster, D. NMR characterization of the binding properties and conformation of glycosaminoglycans interacting with interleukin-10. *Glycobiology* **2014**, *24*, 1036–1049. [[CrossRef](#)] [[PubMed](#)]
31. Gao, Q.; Yang, J.-Y.; Moremen, K.W.; Flanagan, J.G.; Prestegard, J.H. Structural characterization of a heparin sulfate pentamer interacting with LAR-Ig1-2. *Biochemistry* **2018**, *57*, 2189–2199. [[CrossRef](#)] [[PubMed](#)]

32. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *J. Biomol. NMR* **2001**, *21*, 349–359. [[CrossRef](#)] [[PubMed](#)]
33. Mandaliti, W.; Nepravishta, R.; Pica, F.; Vallebona, P.S.; Garaci, E.; Paci, M. Thymosin alpha1 Interacts with Hyaluronic Acid Electrostatically by Its Terminal Sequence LKEKK. *Molecules* **2017**, *22*, 1843. [[CrossRef](#)] [[PubMed](#)]
34. Geist, L.; Mayer, M.; Cockcroft, X.L.; Wolkerstorfer, B.; Kessler, D.; Engelhardt, H.; McConnell, D.B.; Konrat, R. Direct NMR Probing of Hydration Shells of Protein Ligand Interfaces and Its Application to Drug Design. *J. Med. Chem.* **2017**, *60*, 8708–8715. [[CrossRef](#)] [[PubMed](#)]
35. Pomin, V.H.; Wang, X. Synthetic Oligosaccharide Libraries and Microarray Technology: A Powerful Combination for the Success of Current Glycosaminoglycan Interactomics. *Chem. Med. Chem.* **2018**, *13*, 648–661. [[CrossRef](#)] [[PubMed](#)]
36. Wang, X. NMR studies of protein-glycosaminoglycan interactions. In *NMR in Glycoscience and Glycotechnology*, 1st ed.; Kato, K., Peters, T., Eds.; Royal Society of Chemistry: Croydon, UK, 2017; Chapter 11.
37. Pomin, V.H.; Park, Y.; Huang, R.; Heiss, C.; Sharp, J.S.; Azadi, P.; Prestegard, J.H. Exploiting enzyme specificities in digestions of chondroitin sulfates A and C: Production of well-defined hexasaccharides. *Glycobiology* **2012**, *22*, 826–838. [[CrossRef](#)] [[PubMed](#)]
38. Xiong, J.; Bhaskar, U.; Li, G.; Fu, L.; Li, L.; Zhang, F.; Dordick, J.S.; Linhardt, R.J. Immobilized enzymes to convert N-sulfo, N-acetyl heparosan to a critical intermediate in the production of bioengineered heparin. *J. Biotechnol.* **2013**, *167*, 241–247. [[CrossRef](#)] [[PubMed](#)]
39. Zhang, Z.; McCallum, S.A.; Xie, J.; Nieto, L.; Corzana, F.; Jiménez-Barbero, J.; Chen, M.; Liu, J.; Linhardt, R.J. Solution structures of chemoenzymatically synthesized heparin and its precursors. *J. Am. Chem. Soc.* **2008**, *130*, 12998–13007. [[CrossRef](#)] [[PubMed](#)]
40. Laguri, C.; Sapay, N.; Simorre, J.P.; Brutscher, B.; Imberty, A.; Gans, P.; Lortat-Jacob, H. ¹³C-labeled heparan sulfate analogue as a tool to study protein/heparan sulfate interactions by NMR spectroscopy: Application to the CXCL12α chemokine. *J. Am. Chem. Soc.* **2011**, *133*, 9642–9645. [[CrossRef](#)] [[PubMed](#)]
41. He, W.; Fu, L.; Li, G.; Andrew Jones, J.; Linhardt, R.J.; Koffas, M. Production of chondroitin in metabolically engineered *E. coli*. *Metab. Eng.* **2015**, *27*, 92–100. [[CrossRef](#)] [[PubMed](#)]
42. Pomin, V.H.; Sharp, J.S.; Li, X.; Wang, L.; Prestegard, J.H. Characterization of glycosaminoglycans by 15N NMR spectroscopy and in vivo isotopic labeling. *Anal. Chem.* **2010**, *82*, 4078–4088. [[CrossRef](#)] [[PubMed](#)]
43. Pomin, V.H. Advances in glycosaminoglycanomics by 15N-NMR spectroscopy. *Anal. Bioanal. Chem.* **2013**, *405*, 3035–3048. [[CrossRef](#)] [[PubMed](#)]
44. Baik, J.Y.; Gasimli, L.; Yang, B.; Datta, P.; Zhang, F.; Glass, C.A.; Esko, J.D.; Linhardt, R.J.; Sharfstein, S.T. Metabolic engineering of Chinese hamster ovary cells: Towards a bioengineered heparin. *Metab. Eng.* **2012**, *14*, 81–90. [[CrossRef](#)] [[PubMed](#)]
45. Deshauer, C.; Morgan, A.M.; Ryan, E.O.; Handel, T.M.; Prestegard, J.H.; Wang, X. Interactions of the Chemokine CCL5/RANTES with Medium-Sized Chondroitin Sulfate Ligands. *Structure* **2015**, *23*, 1066–1077. [[CrossRef](#)] [[PubMed](#)]
46. Feng, W.; Wang, X. Structure of decorin binding protein B from *Borrelia burgdorferi* and its interactions with glycosaminoglycans. *Biochim. Biophys. Acta* **2015**, *1854*, 1823–1832. [[CrossRef](#)] [[PubMed](#)]
47. Morgan, A.; Sepuru, K.M.; Feng, W.; Rajarathnam, K.; Wang, X. Flexible Linker Modulates Glycosaminoglycan Affinity of Decorin Binding Protein, A. *Biochemistry* **2015**, *54*, 5113–5119. [[CrossRef](#)] [[PubMed](#)]
48. Morgan, A.M.; Wang, X. Structural mechanisms underlying sequence-dependent variations in GAG affinities of decorin binding protein A, a *Borrelia burgdorferi* adhesin. *Biochem. J.* **2015**, *467*, 439–451. [[CrossRef](#)] [[PubMed](#)]
49. Park, Y.; Jowitt, T.A.; Day, A.J.; Prestegard, J.H. Nuclear Magnetic Resonance Insight into the Multiple Glycosaminoglycan Binding Modes of the Link Module from Human TSG-6. *Biochemistry* **2016**, *55*, 262–276. [[CrossRef](#)] [[PubMed](#)]
50. Koehler, J.; Meiler, J. Expanding the utility of NMR restraints with paramagnetic compounds: Background and practical aspects. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 360–389. [[CrossRef](#)] [[PubMed](#)]
51. Iwahara, J.; Tang, C.; Marius Clore, G. Practical aspects of (1) H transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn. Reson.* **2007**, *184*, 185–195. [[CrossRef](#)] [[PubMed](#)]

52. Calle, L.P.; Cañada, F.J.; Jiménez-Barbero, J. Application of NMR methods to the study of the interaction of natural products with biomolecular receptors. *Nat. Prod. Rep.* **2011**, *28*, 1118–1125. [[CrossRef](#)] [[PubMed](#)]
53. Morgan, A.; Wang, X. The novel heparin-binding motif in decorin-binding protein A from strain B31 of *Borrelia burgdorferi* explains the higher binding affinity. *Biochemistry* **2013**, *52*, 8237–8245. [[CrossRef](#)] [[PubMed](#)]
54. Iwahara, J.; Jung, Y.S.; Clore, G.M. Heteronuclear NMR spectroscopy for lysine NH (3) groups in proteins: Unique effect of water exchange on (15) N transverse relaxation. *J. Am. Chem. Soc.* **2007**, *129*, 2971–2980. [[CrossRef](#)] [[PubMed](#)]
55. Esadze, A.; Chen, C.; Zandarashvili, L.; Roy, S.; Pettitt, B.M.; Iwahara, J. Changes in conformational dynamics of basic side chains upon protein-DNA association. *Nucleic Acids Res.* **2016**, *44*, 6961–6970. [[CrossRef](#)] [[PubMed](#)]
56. Nguyen, D.; Hoffpauir, Z.A.; Iwahara, J. Internal Motions of Basic Side Chains of the Antennapedia Homeodomain in the Free and DNA-Bound States. *Biochemistry* **2017**, *56*, 5866–5869. [[CrossRef](#)] [[PubMed](#)]
57. Sepuru, K.M.; Iwahara, J.; Rajarathnam, K. Direct detection of lysine side chain NH₃⁺ in protein–heparin complexes using NMR spectroscopy. *Analyst* **2018**, *143*, 635–638. [[CrossRef](#)] [[PubMed](#)]
58. Blaum, B.S.; Deakin, J.A.; Johansson, C.M.; Herbert, A.P.; Barlow, P.N.; Lyon, M.; Uhrin, D. Lysine and arginine side chains in glycosaminoglycan-protein complexes investigated by NMR, cross-linking, and mass spectrometry: A case study of the factor H-heparin interaction. *J. Am. Chem. Soc.* **2010**, *132*, 6374–6381. [[CrossRef](#)] [[PubMed](#)]
59. Koharudin, L.M.; Bonvin, A.M.; Kaptein, R.; Boelens, R. Use of very long-distance NOEs in a fully deuterated protein: An approach for rapid protein fold determination. *J. Magn. Reson.* **2003**, *163*, 228–235. [[CrossRef](#)]
60. Kay, L.E.; Torchia, D.A.; Bax, A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* **1989**, *28*, 8972–8979. [[CrossRef](#)] [[PubMed](#)]
61. Jeong, K.-W.; Jeong, M.-C.; Jin, B.; Kim, Y. Relationship between Structural Flexibility and Function in the C-Terminal Region of the Heparin-Binding Domain of VEGF165. *Biochemistry* **2013**, *52*, 8823–8832. [[CrossRef](#)] [[PubMed](#)]
62. Abedini, A.; Tracz, S.M.; Cho, J.H.; Raleigh, D.P. Characterization of the heparin binding site in the N-terminus of human pro-islet amyloid polypeptide: Implications for amyloid formation. *Biochemistry* **2006**, *45*, 9228–9237. [[CrossRef](#)] [[PubMed](#)]
63. Cohlberg, J.A.; Li, J.; Uversky, V.N.; Fink, A.L. Heparin and other glycosaminoglycans stimulate the formation of amyloid fibrils from alpha-synuclein in vitro. *Biochemistry* **2002**, *41*, 1502–1511. [[CrossRef](#)] [[PubMed](#)]
64. Hasegawa, M.; Crowther, R.A.; Jakes, R.; Goedert, M. Alzheimer-like changes in microtubule-associated protein Tau induced by sulfated glycosaminoglycans. Inhibition of microtubule binding, stimulation of phosphorylation, and filament assembly depend on the degree of sulfation. *J. Biol. Chem.* **1997**, *272*, 33118–33124. [[CrossRef](#)] [[PubMed](#)]
65. Jha, S.; Patil, S.M.; Gibson, J.; Nelson, C.E.; Alder, N.N.; Alexandrescu, A.T. Mechanism of amylin fibrillization enhancement by heparin. *J. Biol. Chem.* **2011**, *286*, 22894–22904. [[CrossRef](#)] [[PubMed](#)]
66. Malmos, K.G.; Bjerring, M.; Jessen, C.M.; Nielsen, E.H.; Poulsen, E.T.; Christiansen, G.; Vosegaard, T.; Skrydstrup, T.; Enghild, J.J.; Pedersen, J.S.; et al. How Glycosaminoglycans Promote Fibrillation of Salmon Calcitonin. *J. Biol. Chem.* **2016**, *291*, 16849–16862. [[CrossRef](#)] [[PubMed](#)]
67. Madine, J.; Clayton, J.C.; Yates, E.A.; Middleton, D.A. Exploiting a (13) C-labelled heparin analogue for in situ solid-state NMR investigations of peptide-glycan interactions within amyloid fibrils. *Org. Biomol. Chem.* **2009**, *7*, 2414–2420. [[CrossRef](#)] [[PubMed](#)]
68. Madine, J.; Pandya, M.J.; Hicks, M.R.; Rodger, A.; Yates, E.A.; Radford, S.E.; Middleton, D.A. Site-specific identification of an abeta fibril-heparin interaction site by using solid-state NMR spectroscopy. *Angew. Chem. Int. Ed.* **2012**, *51*, 13140–13143. [[CrossRef](#)] [[PubMed](#)]
69. Stewart, K.L.; Hughes, E.; Yates, E.A.; Akien, G.R.; Huang, T.Y.; Lima, M.A.; Rudd, T.R.; Guerrini, M.; Hung, S.C.; Radford, S.E.; et al. Atomic Details of the Interactions of Glycosaminoglycans with Amyloid-beta Fibrils. *J. Am. Chem. Soc.* **2016**, *138*, 8328–8331. [[CrossRef](#)] [[PubMed](#)]
70. Stewart, K.L.; Hughes, E.; Yates, E.A.; Middleton, D.A.; Radford, S.E. Molecular Origins of the Compatibility between Glycosaminoglycans and Abeta40 Amyloid Fibrils. *J. Mol. Biol.* **2017**, *429*, 2449–2462. [[CrossRef](#)] [[PubMed](#)]

71. Schwieters, C.D.; Kuszewski, J.J.; Tjandra, N.; Clore, G.M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Res.* **2003**, *160*, 65–73. [[CrossRef](#)]
72. Brunger, A.T.; Adams, P.D.; Clore, G.M.; DeLano, W.L.; Gros, P.; Grosse-Kunstleve, R.W.; Jiang, J.S.; Kuszewski, J.; Nilges, M.; Pannu, N.S.; et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **1998**, *54*, 905–921. [[PubMed](#)]
73. Singh, A.; Tessier, M.B.; Pederson, K.; Wang, X.; Venot, A.P.; Boons, G.J.; Prestegard, J.H.; Woods, R.J. Extension and validation of the GLYCAM force field parameters for modeling glycosaminoglycans. *Can. J. Chem.* **2016**, *94*, 927–935. [[CrossRef](#)] [[PubMed](#)]
74. Dominguez, C.; Boelens, R.; Bonvin, A.M. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737. [[CrossRef](#)] [[PubMed](#)]
75. Poluri, K.M.; Joseph, P.R.; Sawant, K.V.; Rajarathnam, K. Molecular basis of glycosaminoglycan heparin binding to the chemokine CXCL1 dimer. *J. Biol. Chem.* **2013**, *288*, 25143–25153. [[CrossRef](#)] [[PubMed](#)]
76. Sepuru, K.M.; Nagarajan, B.; Desai, U.R.; Rajarathnam, K. Molecular Basis of Chemokine CXCL5-Glycosaminoglycan Interactions. *J. Biol. Chem.* **2016**, *291*, 20539–20550. [[CrossRef](#)] [[PubMed](#)]
77. Gao, Q.; Chen, C.-Y.; Zong, C.; Wang, S.; Ramiah, A.; Prabhakar, P.; Morris, L.C.; Boons, G.-J.; Moremen, K.W.; Prestegard, J.H. Structural Aspects of Heparan Sulfate Binding to Robo1–Ig1–2. *ACS Chem. Biol.* **2016**, *11*, 3106–3113. [[CrossRef](#)] [[PubMed](#)]
78. Gehrcke, J.P.; Pisabarro, M.T. Identification and characterization of a glycosaminoglycan binding site on interleukin-10 via molecular simulation methods. *J. Mol. Gr. Mod.* **2015**, *62*, 97–104. [[CrossRef](#)] [[PubMed](#)]
79. Griffith, A.R.; Rogers, C.J.; Miller, G.M.; Abrol, R.; Hsieh-Wilson, L.C.; Goddard, W.A. Predicting glycosaminoglycan surface protein interactions and implications for studying axonal growth. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13697–13702. [[CrossRef](#)] [[PubMed](#)]
80. Sankaranarayanan, N.V.; Desai, U.R. Toward a robust computational screening strategy for identifying glycosaminoglycan sequences that display high specificity for target proteins. *Glycobiology* **2014**, *24*, 1323–1333. [[CrossRef](#)] [[PubMed](#)]
81. Forster, M.; Mulloy, B. Computational approaches to the identification of heparin-binding sites on the surfaces of proteins. *Biochem. Soc. Trans.* **2006**, *34*, 431–444. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).