



**HAL**  
open science

## Contribution of separation sciences to the deciphering of the binding sites of glycosaminoglycans to proteins

Frédéric Jeanroy, Claire Demesmay, Vincent Dugas

### ► To cite this version:

Frédéric Jeanroy, Claire Demesmay, Vincent Dugas. Contribution of separation sciences to the deciphering of the binding sites of glycosaminoglycans to proteins. *Separation and Purification Reviews*, In press, 10.1080/15422119.2023.2293027 . hal-04579046

**HAL Id: hal-04579046**

**<https://hal.science/hal-04579046v1>**

Submitted on 17 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Contribution of separation sciences to the deciphering of the binding sites of glycosaminoglycans to proteins.

Frédéric Jeanroy, Claire Demesmay, Vincent Dugas\*

*Universite Claude Bernard Lyon 1, CNRS, ISA UMR 5280 , 5 rue de la Doua, F-69100 VILLEURBANNE, France*

\* Corresponding author. Tel.: +33437423552

E-mail address: [vincent.dugas@univ-lyon1.fr](mailto:vincent.dugas@univ-lyon1.fr)

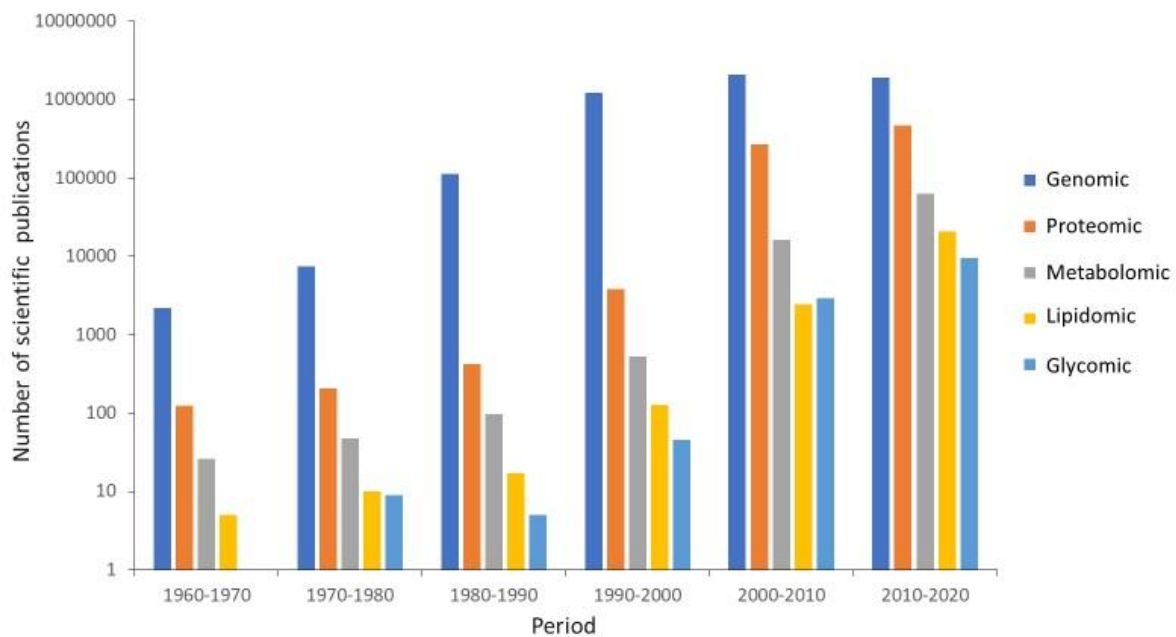
## Abstract

Following the rapid developments in genomics and proteomics in recent decades, the analysis of sugars is currently the focus of much effort, particularly in the field of analytical chemistry. Indeed, their synthetic route, diversity and complexity make these molecules, and in particular glycosaminoglycans, a real analytical puzzle. These polymers are involved in a wide range of biological processes through their particular interaction with proteins. This article aims to show the contribution of separation sciences to the identification and characterisation of interaction sites on GAGs. After a brief review of the structure and diversity of GAGs, the question was raised as to whether or not GAGs have a specific interaction motif with proteins. A review of the literature shows that this question is still controversial and that the reality is more nuanced. The following section looks at the different strategies for identifying such a site if it exists, and in particular the contribution of separative methods to these strategies.

**Keywords :** Affinity chromatography, Glycosaminoglycans, Binding Sites, Interactions, Sugars

## Introduction

For several decades, the omics sciences have aimed to characterise and quantify pools of biological molecules as a whole, and are at the heart of our understanding of how biological systems work. These multidisciplinary sciences, which combine biochemistry, biology, chemistry and physics, gained interest in the 1990s and especially in the 2000s with the improvement of technologies and the sequencing of the genome. A distinction can be made between genomics, proteomics, metabolomics, lipidomics and glycomics, which deal respectively with the genome, proteome, metabolome, lipidome and glycome, the respective sets of genes, proteins, metabolites, lipids and carbohydrates. Figure 1 illustrates the evolution of scientific research in the field of omics.



*Figure 1: Evolution of scientific research in the last decades illustrated by the number of scientific publications in the field of genomics, proteomics, metabolomics, lipidomics and glycomics in 1960-1970, 1970-1980, 1980-1990, 1990-2000, 2000-2010 and 2010-2020 periods of time. Source: google scholar in November 2023. Key words: “genomic”, “proteomic”, “metabolomics”, “lipidomic” and “glycomic”*

Advances in these sciences are essential if we are to gain a better understanding of human biology and improve performance in areas such as medical diagnosis and the search for new drugs. For example, proteomics has allowed great advances in deciphering post-translational modifications and many proteomics studies are looking for biomarkers of disease<sup>[1]</sup>. Metabolomics are crucial for studying, for example, the genetic and environmental contributions to metabolites variation that occur within living organism as well as for early cancer diagnosis with the identification biomarkers. Lipidomics is also of great interest in the search for biomarkers for the early diagnosis of cancer, as lipids play an essential role in cellular functions<sup>[2]</sup>. Glycomics studies are also of particular importance, as is the role of carbohydrates in many biological processes, including cell migration, intercellular adhesion, signalling or the adhesion of toxins and pathogens<sup>[3-5]</sup>. However, the glycome is still a relatively

unexplored area compared to the proteome, probably because of its great complexity, linked to the huge diversity of the molecules that make it up.

The aim of this review is to discuss the challenges associated with the analysis of a particular class of molecules, the glycosaminoglycans (GAGs) <sup>[6, 7]</sup>. Their study began towards the end of the 19th century with the discovery of chondroitin sulphate in cartilage. Other families of GAGs were subsequently identified, and in the 1950s the foundations of their structure were laid. Since the 1980s, with the development of new characterisation techniques, these compounds have been increasingly studied for their multiple roles in regulating biological processes.

We will begin with a brief review of the structure of GAGs and their biological interest to illustrate the complexity of this family of molecules and to highlight their importance. The second part will address the question of whether or not there is a short specific interaction sequence. Finally, the third part will focus on the global strategies implemented to identify such a sequence, focusing on the depolymerisation, affinity sequence isolation and analysis stages.

## **Structure and biological interest of glycosaminoglycans**

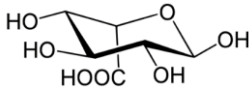



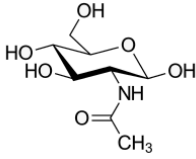



### *Structure of GAGs*

Glycosaminoglycans are a family of complex polysaccharides with a highly charged and sulphated linear (except for hyaluronic acid which is unsulfated) and unbranched structure. They are composed of 10 to 200 disaccharide units and their molecular mass can reach 100,000 Da <sup>[8]</sup>. Each disaccharide unit consists of :

- a uronic acid, which can be either L-iduronic acid (IdoA) or D-glucuronic acid (GlcA), these two monosaccharides being C5 epimers <sup>[9]</sup>, with the exception of keratin sulphate, in which the uronic acid is replaced by glucose (Glc).
- a hexosamine, which may be either N-acetyl-D-glucosamine (GlcNac) or N-acetyl-D-galactosamine (GalNac), these two monosaccharides being C4 epimers <sup>[9]</sup>.

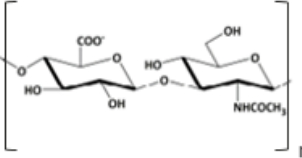
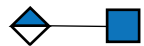
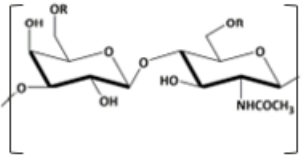

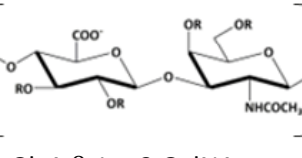

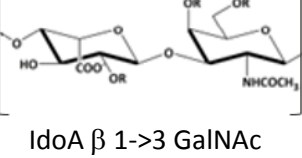

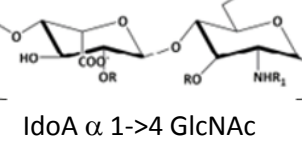

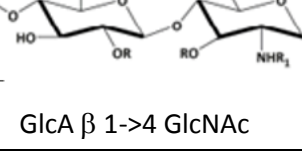

The structures of uronic acids and hexosamines are given in Table 1.

Table 1: Structure of the different uronic acids and hexosamines based building blocks of glycosaminoglycans.

Acides uroniques			
	IdoA		GlcA
Acide iduronique		Acide Glucuronique	
Hexosamine			
	GlcNAc		GalNAc
N-acetyl-D-Glucosamine		N-acetyl-D-Galactosamine	

There are six families of glycosaminoglycans based on their uronic acid and hexosamine composition: hyaluronic acid, keratan sulphate, chondroitin sulphate, dermatan sulphate, heparin and heparan sulphate. The structure of the different GAG families is given Table 2.

Table 2: Structure of the different glycosaminoglycans families

Glycosaminoglycane	Acide uronique	Hexosamine	Symboles
Acide hyaluronique			
	GlcA $\beta$ 1->3 GlcNAc		
Kératane sulfate			
	Gal $\beta$ 1->4 GalNAc		R = H ou SO <sub>3</sub> <sup>-</sup>
Chondroïtine sulfate			
	GlcA $\beta$ 1-> 3 GalNAc		R = H ou SO <sub>3</sub> <sup>-</sup>
Dermatane sulfate			
	IdoA $\beta$ 1->3 GalNAc		R = H ou SO <sub>3</sub> <sup>-</sup>
Héparine			
	IdoA $\alpha$ 1->4 GlcNAc		R = H ou SO <sub>3</sub> <sup>-</sup> R = H, SO <sub>3</sub> <sup>-</sup> R <sub>1</sub> = H, SO <sub>3</sub> <sup>-</sup> , COCH <sub>3</sub>
Héparane sulfate			
	GlcA $\beta$ 1->4 GlcNAc		R = H ou SO <sub>3</sub> <sup>-</sup> R <sub>1</sub> = H ou SO <sub>3</sub> <sup>-</sup> ou COCH <sub>3</sub>

*Biosynthesis responsible for significant heterogeneity*

GAGs are synthesised on a protein core from a tetrasaccharide sequence which is the same for chondroitin sulphate, dermatan sulphate, heparin and heparan sulphate [10]. From this anchor sequence, the elongation and polymerisation of the tetrasaccharide sequence depends on the enzymes involved (EXT1 and EXT2 from exostosin family for heparin and heparan sulfate). For example, the 4GlcA( $\beta$ 1-4)GlcNAc $\alpha$ 1 unit is added for heparin and heparan sulfate [11, 12]. After or during

polymerisation, several modifications induced by different enzymes occur, leading to epimerisation and sulphation. By way of illustration, four modifications are likely to occur in the heparin/heparan sulphate chains:

- N-deacetylation followed by N-sulphation of the N-acetylated glucosamine GlcNAc is induced by two enzymes, N-deacetylases/N-sulphotransferases (NDST) <sup>[10-12]</sup>. In rare cases, GlcNAc may remain N-deacetylated <sup>[13]</sup>.
- C5-epimerase mediated epimerisation of glucuronic acid GlcA to iduronic acid IdoA <sup>[10, 12]</sup>.
- 2-O-sulphation (2-OS) induced by 2-O-sulphotransferase (2-OST) at the C2 position of iduronic acids. Although this sulphation preferentially occurs on iduronic acids, it can also occur on glucuronic acids <sup>[10, 12]</sup>.
- 6-O-sulphation (6-OS), in which a sulphate group is added to the C6 position of N-sulphated or N-acetylated glucosamine using 6-O-sulphotransferase (6-OST) <sup>[10, 12]</sup>.
- Finally, and very rarely, 3-O-sulphation (3-OS) at the C3 position of a glucosamine, induced by 3-O-sulphotransferase (3-OST) <sup>[10, 12]</sup>.

Due to the various possible modifications, heparin can have 48 disaccharide units and the heterogeneity increases exponentially with the length of the chain. This increases the difficulty of analysing these molecules tenfold.

### *Biological interests of GAGs*

GAGs can interact with a large number of proteins. To date, several hundred proteins have been identified whose biological action involves a GAG. Depending on the proteins involved, GAGs play different roles:

- Change in protein conformation and involvement in protein-protein interactions,

One of the mechanisms induced by the interaction of a GAG with a protein is a change in the conformation of the target protein. This is particularly true of the interaction between heparin and antithrombin III, which is arguably the most extensively studied interaction to date. Antithrombin III is an inhibitor of blood clotting factors. When antithrombin III interacts with heparin at a pentasaccharide sequence, a reactive loop in its protein structure is exposed to Factor Xa, allowing it to be recognised. The formation of an antithrombin III - Factor Xa complex results in the irreversible inhibition of Factor Xa <sup>[14-16]</sup>.

- Role as co-receptor for growth factors

Fibroblast Growth Factors (FGFs) are a family of growth factors with about twenty members, of which FGF2 is undoubtedly the most studied. These proteins play a role in cell proliferation, differentiation and migration by binding to specific cell surface receptors known as FGFRs (Fibroblast Growth Factor Receptors). However, FGFRs are only active when they are dimerised. Heparan sulphate then acts as a co-receptor, bringing the FGF-HS-FGFR complexes together and stabilising them <sup>[17-20]</sup>.

- Role in cell signalling

GAGs are also involved in cell signalling by modulating the bioavailability of messengers such as growth factors, cytokines, chemokines, etc. Certain proteins can be stored and protected by their interactions with GAG chains. They are released

when they are close to their receptors <sup>[21]</sup>. In the case of chemokines, GAGs allow them to be concentrated, creating concentration gradients that guide circulating cells <sup>[22, 23]</sup>.

#### - GAG-pathogen interactions

In addition to their regulatory functions, which are essential for the body, GAGs can interact with pathogens and enable their activity within cells. This is the case with viruses, bacteria and parasites that use GAGs as anchorage sites on the cell surface <sup>[24-27]</sup>. Viruses, including HIV, herpes and dengue, use GAGs to attach to the surface of cells and fuse with them <sup>[27-29]</sup>.

### **The specific interaction motif of GAGs: myth or reality?**

Although GAGs are compounds with great therapeutic potential, the use of intact GAGs to regulate specific processes can lead to problems with side effects. For example, the use of heparin as a blood clotting agent can be associated with side effects. The use of low-molecular-weight heparins (enoxaparin type) or even a specific sequence of interaction with antithrombin III (fondaparinux) has made it possible to limit these. Identifying the sequence in which a GAG interacts with proteins could therefore be the key to new treatments that are more effective and less prone to side effects.

The question of whether there is always a specific site for all GAGs interacting with proteins is still open to debate. The reality is more nuanced: while there is a specific site for certain interactions, in other cases there is more intermediate specificity or 'hidden specificity', or even non-specific interaction sites.

It is therefore interesting to ask whether the GAG/protein interaction depends on the existence of an interaction site whose charge density and monosaccharides and their substituents are determined. Mosier et al. and Sarkar et al. define this notion of specificity either biologically, with a unique mode of interaction, or chemically, with the existence of a unique interaction sequence from a wide range of oligosaccharide sequences <sup>[30, 31]</sup>. The notion of specificity here will be based essentially on the chemical concept.

#### *Existence of a specific site for certain interactions*

Very early on, it became clear that certain GAGs were capable of interacting with proteins in a very strong and selective way. The best example is the discovery, after almost a decade of research, of a specific interaction sequence between heparin and antithrombin III <sup>[32]</sup>. Antithrombin III is a protein of the serpin family involved in blood clotting mechanisms <sup>[33]</sup>. The antithrombin interaction site on the heparin chain is a pentasaccharide [GlcNAc6S-GlcA-GlcNS3S6S-IdoA2S-GlcNS6S] with a very high affinity for antithrombin III.

Each monosaccharide and each substituent play a role in the interaction <sup>[34]</sup>, in particular 4 sulphate groups, the absence of which increases the dissociation constant by a factor of 2 to 3 <sup>[35]</sup>, or the presence of a GlcA instead of an IdoA at the non-reducing end <sup>[36]</sup>. This structural information is therefore encoded. The presence



of this epitope, specific to the position of the substituents and the rate of sulphation, is also known as the "sulphation code" or "sulphation pattern" [37].

In addition, the rarity of certain substituents present in the interaction sequence is a parameter likely to generate a high degree of sequence specificity. In the case of the heparin-antithrombin III interaction, the presence of the sulphate group on a 3-OS glucosamine, a very rare sulphation in the polysaccharide chain (about 1/3 of the oligosaccharide chains in a heparin preparation have this substitution), is not only crucial for the interaction, as its absence drastically reduces the strength of the interaction (increase in Kd by a factor of  $10^3$  [38]), but also ensures a high degree of specificity [39].

Other groups are also considered to be very rare in a heparin chain, such as an unsubstituted GlcN monosaccharide, and their very low occurrence in GAGs helps to make the interaction sequence unique when they are present [39, 40]. Other examples of interactions have shown a correlation between the specificity of certain oligosaccharide sequences towards certain proteins and the presence of very rare substituents, such as the interactions of heparan sulphate with herpes simplex virus 1 entry, neuropilin 1 or stabilin 2 [41–44].

However, in addition to this glycode, and as shown by the example of the heparin-antithrombin complex, other criteria are likely to be involved in the specificity of the interaction site, as non-ionic interactions and the conformations of the GAG and the protein appear to be essential for the specificity of the interaction sequence [40]. From the point of view of GAGs, Raman et al. in 2003 showed that ionic interactions alone cannot explain the specificity of the interactions [45, 46] by studying the structure of several complexes obtained by X-ray diffraction between heparan sulphate and different proteins such as growth factors (FGF: Fibroblast Growth Factor) or antithrombin III.

The spatial folding of the GAG within the protein is a structural signature that allows very strong interaction via an optimal fit between ionic interactions and Van der Waals interactions [46]. This folding is enabled by certain sulphate groups, such as those of the trisaccharide [HNS,3S,6S-12S-HNS,6S] of the heparin-antithrombin interaction sequence, via hydrogen bonds between the 3-OS group and the adjacent NS group [47]. This type of folding has been demonstrated for other GAG protein co-crystals such as heparin and FGF1, FGF2 or FMDV [46].

The example of the interaction between heparin and antithrombin also shows that the conformation of the protein can play a role [30]. This has been studied using computer simulations on crystals of the heparin-antithrombin complex: in addition to the presence of basic amino acids on the surface of antithrombin, the presence of a cavity in the structure of antithrombin III at the heparin interaction site allows occupation by a ligand [30, 48]. This cavity, linked to basic amino acids, is 5 to 7 Å deep and 15 to 20 Å long and is large enough to accommodate the trisaccharide GlcNS3S6S-IdoA2S-GlcNS6S, and more specifically residues GlcNS3S6S and GlcNS6S. However, this example of the spatial conformation of the protein playing a role in the specificity of the interaction between the pentasaccharide and antithrombin appears to be unique [30] and no other interaction of this type has yet been demonstrated.

The example of the interaction between heparin and antithrombin therefore seems to show that the existence of a highly specific sequence of a protein is the combination

of several parameters, whether they are linked to the GAG or to the protein of interest. Another example of an interaction whose level of specificity appears to be close to that of heparin/antithrombin, although less studied, is the interaction between heparan sulphate and FGF2 (Fibroblast Growth Factor 2). The heparan sulphate interaction site appears to be an IdoA2S-GlcNS disaccharide, the presence or absence of which produces differences in affinity between low, medium and high affinity sequences [49, 50]. Other sulphate groups, such as the 6-OS of glucosamines, appear to be essential for the affinity between heparan sulphate and FGF2 [49]. The importance of this disaccharide was demonstrated by studying the crystals of heparin/FGF2 complexes. This study showed that this disaccharide is involved in the formation of 8 of the 10 hydrogen bonds between the GAG sequence and the protein [18]. In addition, similar to the heparin/antithrombin interaction, the study of GAG-protein co-crystals revealed a folding of the GAG sequence indicating a high degree of specificity for the trisaccharide involved [HNS,6S-I2S- HNS,6S] [46]. Thus, there is strong evidence that this interaction is highly specific for the GAG sequence.

### *Existence of hidden sites of specificity*

However, studies of the interaction between heparan sulphate and FGF2 have also shown that different GAG sequences are likely to have similar affinities and activate FGF2 as long as only a few groups are present on the chain [51]. Although this interaction cannot be considered strictly specific from a chemical point of view, it appears to be an interaction with 'hidden specificity'. This type of interaction shows that there is indeed a sulphation code (or sulphation motif) for the GAG sequence with a few sulphate groups whose presence and position are critical. However, the other substituents can be quite diverse. There can therefore be several GAG sequences with similar affinities, provided the critical sulphation pattern is present. This idea of intermediate specificity first appeared in 2005 [52] and in 2006 with the study by Gama et al. on the interaction between chondroitin sulphate and growth factors [37]. Thus, the CS of dp4 containing the interaction sequence appeared to be much more effective in activating growth factors than CSs containing part of the interaction sequence or dp4 but with a different sulphation pattern (same charge density but different positions of the sulphate groups, Figure 2). This result therefore demonstrates the existence of a sulphation pattern with an intermediate level of specificity.

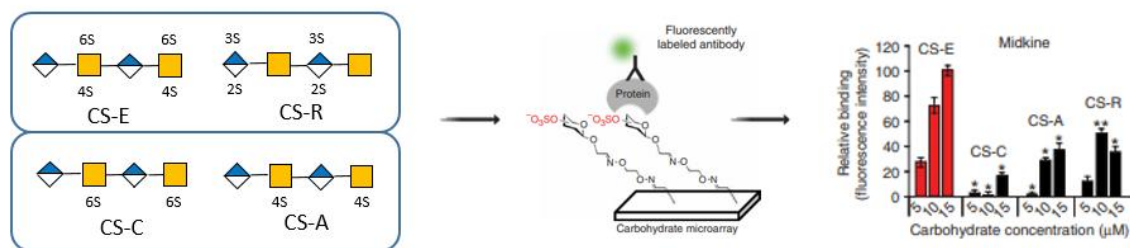


Figure 2: Diagram, adapted from [30], of the strategy demonstrating the presence of a sulphation motif which promotes the interaction of CS with neuronal growth factors. Four CS sequences are immobilised on the surface of a microarray and incubated with midkine. The CS-E sequence shows a higher affinity than the CS-R sequence

*with the same charge density. CSs-C and CS-A with only a part of the sulphation motif of cs-e show lower affinity.*

A large number of proteins also appear to be able to interact with a wide variety of GAG sequences, provided they have this very specific sulphation code: FGF10, FGF7, FGF4, which require either a 2-OS group or a 6-OS group, or both <sup>[53]</sup>. This type of intermediate specificity interaction seems to correlate with the absence of rare sulphate groups, but it is still difficult to precisely define the minimum epitope required.

#### *Existence of non-specific interaction sites with proteins*

In contrast to the highly specific interaction between heparin and antithrombin, there are also non-specific interactions between GAGs and proteins. The best known example is the interaction between heparin and thrombin. In 1991, Olson et al. showed that the interaction between thrombin and heparin is non-specific and that only the number of charges plays a role in the interaction <sup>[54]</sup>. They were able to characterise the non-specific aspect of the interaction by three features: 1) the strong influence of ionic strength on the affinity between thrombin and heparin, 2) the increase in affinity with the size of the oligosaccharide chains and 3) the very weak influence of non-ionic interactions. These three characteristics can be contrasted with the specific affinity between antithrombin and heparin or even with intermediate affinities such as that of heparin with FGF2: a moderate dependence on ionic strength, the essential role of non-ionic interactions (hydrogen bonds, Van der Waals) and the independence of affinity from chain size <sup>[36, 54]</sup>.

In the case of the thrombin-heparin interaction, a sufficient density of anionic charges allows the interaction: only 5 to 6 negative charges are required, regardless of the position of the sulphate groups. Furthermore, the structural study of the interaction between thrombin and heparin confirms the difference between antithrombin and thrombin: unlike the very specific antithrombin-heparin interaction, it is not possible to observe the presence of a cavity similar to that of antithrombin. This difference with antithrombin seems to confirm the non-specific nature of the interaction <sup>[30]</sup>.

Other examples of non-specific interactions have since been identified. In particular, hepatocyte growth factor (HGF) is able to interact with several sequences of dermatan sulphate or heparan sulphate <sup>[55]</sup>. Furthermore, the interaction appears to depend only on charge density, as Catlow et al. showed that 2-OS, 6-OS or GlcNS groups are not essential for affinity <sup>[56]</sup>. However, with a dissociation constant in the order of 1-3 nM, it is interesting to note that even a non-specific interaction can have a high affinity <sup>[36]</sup>.

#### *Conclusion*

In conclusion, it appears that the special properties of GAGs (flexible and polyanionic linear chain) are associated with three types of interaction sites with proteins: (i) a highly specific site, such as the interaction between antithrombin and heparin, in which each substituent and each monosaccharide participates in the interaction and in which a very rare substituent, such as the 3-OS group, is involved; (ii) a site of hidden specificity, in which a specific epitope, is required, consisting of common

substituents such as 2-OS or 6-OS groups, but where this may be surrounded by a wide variety of groups, and (iii) a non-specific site, where multiple oligosaccharide sequences are likely to interact with the protein, provided the charge density is sufficient<sup>[39]</sup>.

The example of antithrombin and heparin is undoubtedly the most studied GAG-protein interaction. Its study shows that the existence of such a highly specific site requires several conditions in addition to the presence of a highly specific epitope: the presence of a very rare 3-OS group and the conformation of the GAG sequence or of the protein, which seem to play a role in this interaction. Is this specific interaction unique? As far as we know, there are no other specific GAG sequences like this one. However, numerous examples suggest that there is very often at least some intermediate specificity. The identification of minimal epitopes would provide a better understanding of the interaction mechanisms involved. However, given the great diversity of GAGs, this remains a real challenge, although analytical strategies have been developed over the last two decades.

### **Research and identification of oligosaccharide sequences interacting with proteins**

The identification of a specific interaction motif implicitly requires the existence of a specific oligosaccharide sequence on the GAGs. If this sequence exists, the challenge is to identify the interaction zone on the GAG and to obtain its fine structure (sugar sequence, epimerisation, acetylation and sulphation).

GAGs are polyanionic compounds that carry a large number of negative charges due to the presence of acidic functions (uronic acids) and sulphates throughout their chain. Proteins may contain positively charged amino acids such as arginine, lysine and histidine. It therefore appears that the interactions between oligosaccharides and proteins are essentially electrostatic, although other types of interaction (Van der Waals, hydrogen bonding) may also be involved<sup>[57]</sup>.

While protein interaction sites are easily identified by NMR or XRD techniques, this is much more complex for GAGs due to their very heterogeneous nature and their different sulphation domains (highly sulphated and non-sulphated domains). The charge density is likely to be very different depending on the sequence.

In addition, the identification and characterisation of such a sequence is a difficult analytical challenge due to the extreme diversity and heterogeneity of oligosaccharides and the presence of a very large number of isomers. Their characterisation generally requires specific techniques and analytical strategies that combine several aspects.

The first specific sequence of interaction between a GAG and a protein was identified in the early 1980s through the combined work of several teams over almost a decade<sup>[32]</sup>. These discoveries were made possible by the use of purification techniques such as affinity chromatography and characterisation techniques such as NMR.

While NMR and X-ray crystallography techniques have been and continue to be used to characterise GAG sequences, the advent of mass spectrometry has revolutionised the structural analysis of GAGs thanks to advantages such as sensitivity, small

sample size and very high resolution. MSMS techniques such as NETD and IRMPD can now even distinguish between epimeric monosaccharides <sup>[58, 59]</sup>.

However, the complexity of oligosaccharide mixtures still requires the development of strategies to identify the specific sequences that interact with a protein. This can be achieved by various approaches, both theoretical (computer modelling) and experimental, such as the use of microarrays or separation methods. Our study will focus in particular on experimental approaches using separation methods in strategies to isolate and characterise oligosaccharide sequences with high affinity for a target protein prior to structural analysis.

## Global strategies for identifying a specific sequence ("sulphation code") using combined analytical techniques

The different strategies used for the selection, isolation and identification of oligosaccharide sequences with affinity for a target protein are summarised in Figure 3.

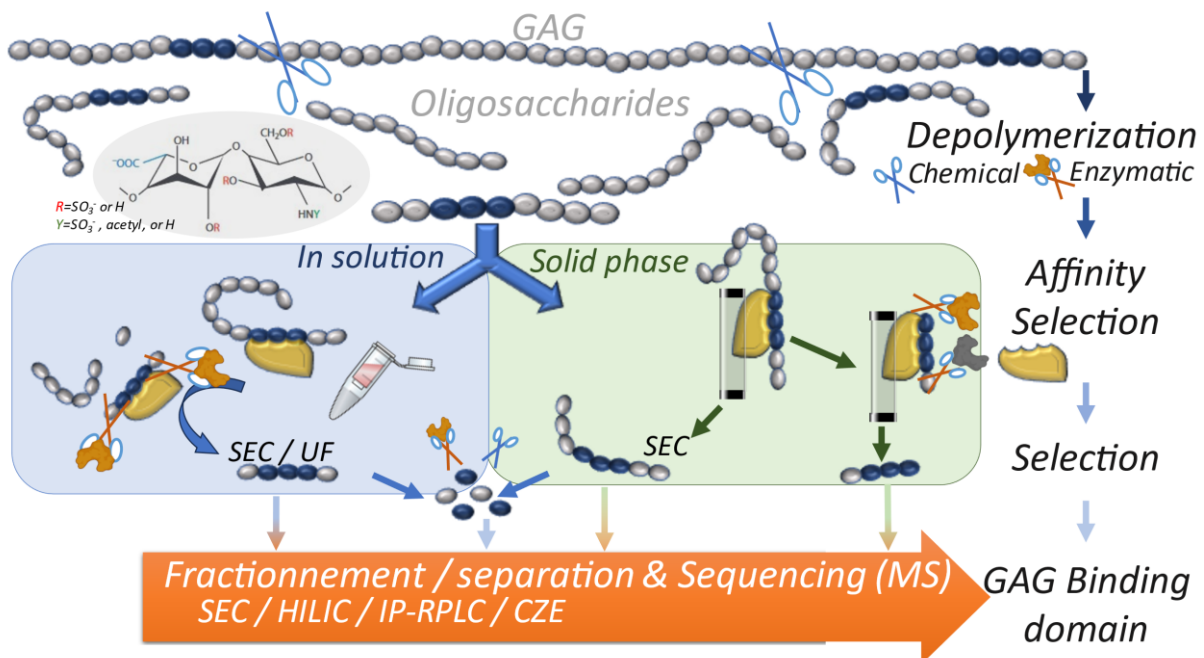


Figure 3: Strategies used for the selection, isolation and identification of oligosaccharide sequences with affinity for a target protein. The first step is a chemical or enzymatic depolymerisation. Then affinity selection is realised by bringing GAG sequences into contact with target protein. This step can take place in solution before isolation by SEC / Ultrafiltration or on a solid support (MALDI plate, SPR chip, affinity column, magnetic beads) sometimes before SEC fractionation. Finally, GAG sequences can be depolymerized before chromatographic fractionation and sequencing.

The first step is always a GAG depolymerisation step to obtain smaller oligosaccharide sequences, some of which contain the interaction motif(s). The selection of the sequences containing the interaction motif(s) is based on an affinity mechanism by bringing the oligosaccharide sequences into contact with the target protein. This step can be performed in solution or on a solid support (MALDI plate, SPR sensor, affinity column, magnetic beads) after immobilisation of the target protein. In the case of selection in solution, affinity sequences can be isolated by ultrafiltration or size exclusion chromatography. In the case of solid phase selection, oligosaccharides captured on the support can be (i) fractionated by SEC then separated before sequencing, (ii) fractionated by SEC then depolymerised and separated before sequencing, (iii) depolymerised on solid support then sequenced.

An intermediate strategy was used by Sheng et al. in 2020. They performed depolymerisation of GAGs in gel prior to gel electrophoresis separation<sup>[60]</sup>. However, hydrogels have to be dehydrated before digestion and gel like agarose (AGE) does

not allow it due to its hydrophilic properties and structure. In addition, depolymerisation of high molecular weight GAGs (MW > 120 kDa) is difficult in Polyacrylamide Gel (PAGE). This type of strategy is under development and it is a promising strategy for direct digestion in gel electrophoresis medium.

### *Depolymerisation Methods*

The identification and analysis of affinity sequences requires a reduction in the size of the initial GAGs, which is achieved by depolymerisation. This depolymerisation step can be performed at two levels: before and/or after the affinity sequence isolation step. Applied to the starting GAG, for example heparin (dp40), controlled depolymerisation results in low molecular weight heparin mixtures (LMWH: Low Molecular Weight Heparin,  $12 < dp < 24$ ). Applied to the selected oligomers to be sequenced, it leads to exhaustive depolymerisation (dp2). These steps can be performed by various enzymatic or chemical means (Figure 4).

Controlled chemical depolymerization (generally performed prior to the binding sequence isolation step) can be achieved by, (i) alkaline conditions ( $\beta$  elimination resulting in the formation of an unsaturation at the non-reducing end and in the formation of a 1,6-anhydro group at the reducing end), (ii) oxidation with hydrogen peroxide (formation of an HO<sup>•</sup> radical capable of cleaving glycosidic bonds), and (iii) nitrous deamination in the presence of nitrous oxide, (ii) oxidation with hydrogen peroxide (formation of an HO<sup>•</sup> radical capable of cleaving glycosidic bonds), and (iii) nitrous deamination in the presence of nitrous acid (contraction of the ring from 6 to 5 carbon atoms, formation of an aldehyde function, and cleavage of the glycosidic bond). When applied to heparin, these chemical depolymerizations result in the LMWHs known as enoxaparin, ardeparin, dalteparin <sup>[61, 62]</sup>.

Only enzymatic depolymerization can be performed in the presence of the target protein so as not to degrade it. In enzymatic depolymerization, enzymes called heparinases catalyze the cleavage of the glycosidic bond of heparin by  $\beta$  elimination. This type of cleavage results in the formation of a double bond on the sugar at the non-reducing end (uronic acid) of the cleavage. Three types of enzymes are distinguished according to their specificity:

- Heparinase I cleaves the glycosidic bonds of heparin between a glucosamine GlcNS3S and an iduronic acid IdoA2S, 2)
- Heparinase II cleaves the glycosidic bond between a glucosamine and a uronic acid with no particular specificity related to sulfation or epimerization.
- Heparinase III cleaves the bond between an unsulfated glucuronic acid and a glucosamine. The low-molecular-weight heparin formed is called tinzaparin.

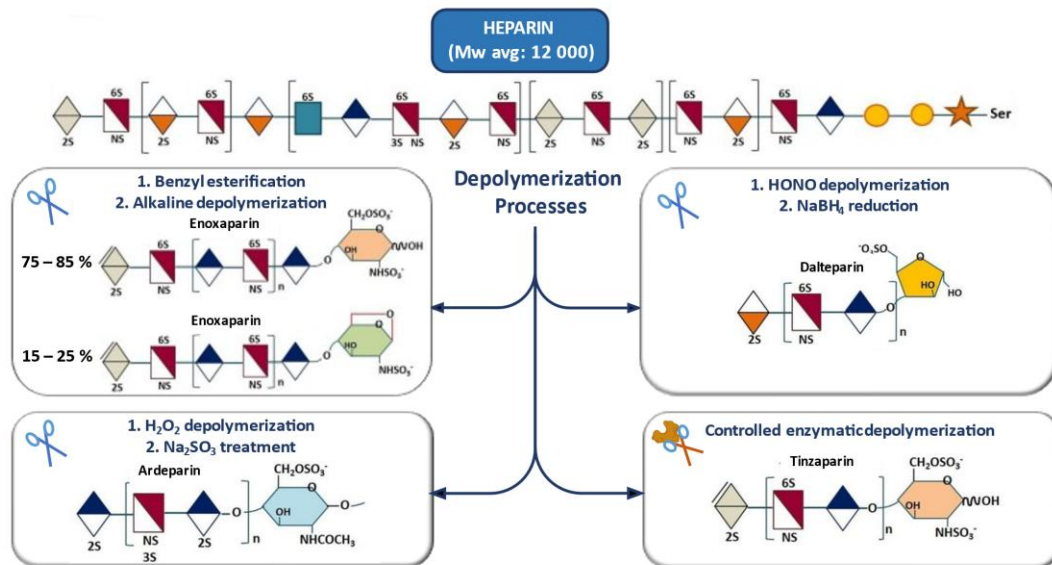


Figure 4: Depolymerisation of GAGs by various enzymatic or chemical routes. Heparin can be depolymerized by 1) benzyl esterification and alkaline depolymerization which results in Enoxaparin. A 4,5-anhydro uronic acid is formed at the non-reducing end and 15% to 25% of the mixture is composed of a 1,6-anhydro derivative at the reducing end, 2)  $H_2O_2$  depolymerization and  $Na_2SO_3$  treatment which results in Ardeparin, 3) HONO depolymerization and  $NaBH_4$  reduction which results in Dalteparin. This way of depolymerization produces a contraction of the ring from 6 to 5 carbons, the formation of an aldehyde function and a cleavage of the glycosidic bond at the reducing-end and 4) enzymatic depolymerization which results in Tinzaparin. A 4,5-anhydro uronic acid is formed at the non-reducing end. Scheme adapted from <sup>[55]</sup>.

### Separation or Fractionation Methods for Affinity-Selected Interaction Motifs Prior to Structural Analysis

The oligosaccharide mixtures obtained after isolation of interaction sequences consist of a large number of epimeric, isomeric or isobaric structures. Therefore, it is often necessary to perform a separation/fractionation step for selected affine oligosaccharide mixtures prior to mass-spectrometric sequencing. This separation step is usually performed by liquid chromatography or capillary electrophoresis.

#### - Size Exclusion Chromatography

Size Exclusion Chromatography (SEC Size Exclusion Chromatography) is often used to fractionate oligosaccharides according to their size or degree of polymerization. Figure 5 shows the separation of enoxaparin oligosaccharides (obtained by alkaline depolymerization of heparin) before and after affinity enrichment <sup>[63]</sup>.

However, the main drawback of SEC is its lack of resolution <sup>[64]</sup>. Therefore, it is rarely used alone for preliminary purification of oligosaccharide mixtures prior to structural analysis. In addition, highly saline mobile phases are often used to limit secondary electrostatic interactions, which limits coupling to mass spectrometry, although



desalting can be performed after fractionation for mass-spectrometric analysis. In many cases, separation by size exclusion is only the first stage of fractionation before a second stage of liquid chromatography <sup>[6]</sup>.

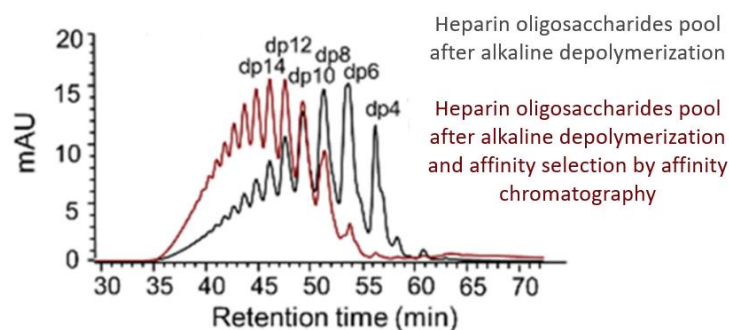


Figure 5: Chromatogram adapted from <sup>[56]</sup>. Size exclusion separation of enoxparin before and after affinity enrichment adapted from <sup>[55]</sup>. Separation was performed on two serially connected Waters 1.7  $\mu$ m SEC 125 Å columns (4.6 mm  $\times$  150 mm and 4.6 mm  $\times$  300 mm) at a flow rate of 0.075 mL/min. The mobile phase consisted of 50 mM ammonium acetate and 20% (v/v) methanol. Heparin oligosaccharides are obtained by benzyl extraction and alkaline depolymerisation of heparin. High affinity oligosaccharides were obtained by affinity chromatography.

#### - Strong Anion eXchange (SAX) Chromatography

Since oligosaccharides are negatively charged compounds, SAX is a technique naturally suited to the separation of sequences with different charges. The main limitation of this technique is the difficulty in separating oligosaccharides with the same number of charges (same number of sulfate groups and uronic acids), especially for large compounds with a large number of charges <sup>[65]</sup>. While it is possible to separate small oligosaccharides (dp 2) that differ only in the position of the sulfate groups with good resolution <sup>[66]</sup>, the difficulty increases with the size of the oligosaccharides and the complexity of the mixtures. For disaccharides, retention increases with the number of sulfate groups. For compounds with the same sulfation, the sulfate group carried by the azide (NS) induces less retention than that carried by the alcohol carried by carbon 6 (6-OS) followed by the 2-OS group (sulfate carried by the alcohol carried by carbon 2). For more complex mixtures, Miller et al. showed that even with a long gradient, peaks from hexasaccharide separation contained several co-eluted compounds <sup>[67]</sup>.

Figure 6 shows the separation by anion-exchange chromatography of (A) 8 disaccharides <sup>[66]</sup> and (B) a mixture of hexasaccharides isolated by SEC <sup>[67]</sup>.

Again, the use of mobile phases rich in inorganic salts presents a difficulty when coupled directly to a mass spectrometer, although the use of ion suppressors or mobile phases containing volatile salts such as ammonium carbonate <sup>[67]</sup> can overcome this difficulty.

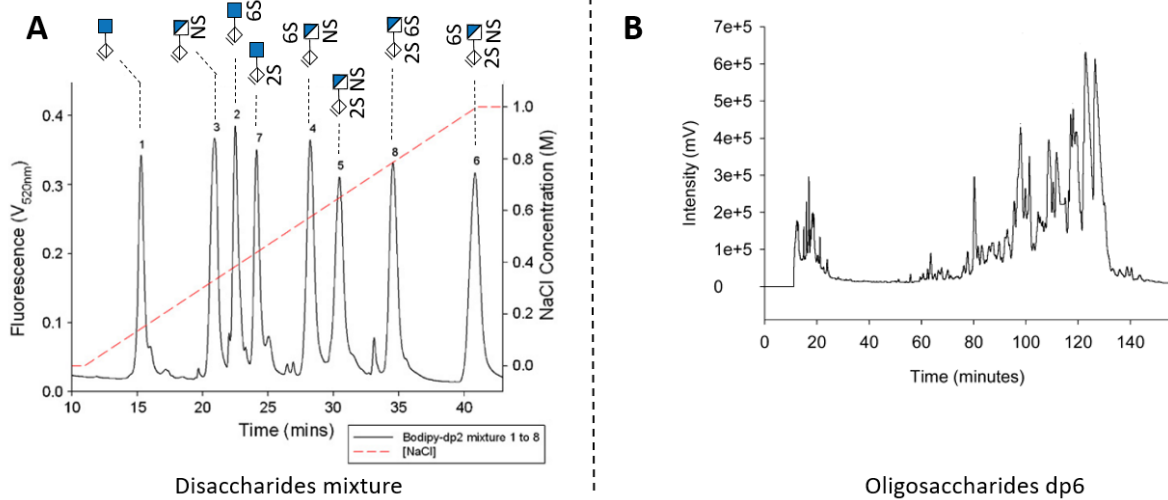


Figure 6: Chromatogram adapted from <sup>[60]</sup> of disaccharides mixture separated on a quaternary amine stationary phase 250x9 mm and eluted with a NaCl gradient from 0 to 1 M at 2 ml/min. Detection is performed by fluorescence, exciting at 488 nm and detecting at 520 nm (A) and chromatogram adapted from <sup>[61]</sup> of hexasaccharides mixtures isolated by SEC and separated on a quaternary amine stationary phase 250x4.6 mm and eluted with a NaCl gradient from 0 to 1.4 M in 90 minutes. Detection is performed by UV at 232 nm.

#### - Ion Pairing Chromatography

In reversed-phase chromatography, the addition of cationic ion-pairing agents such as triethylamine, tributylamine, and tetrabutylammonium hydroxide to the mobile phase allows the retention of polar and anionic oligosaccharides <sup>[68]</sup>. This chromatographic mode offers good potential for the separation of small and large oligosaccharides. For example, in 2013, Galeotti et al. proposed the separation of a mixture of disaccharides and a tetrasaccharide using tetrabutylammonium as an ion-pairing agent (

Figure 7A) <sup>[69]</sup>. The retention of the mixture increases with the number of sulfate groups. In addition, as observed in SAX, for disaccharides with the same sulfation rate, the N-sulfated group is less retained than the 6-O-sulfate group, followed by the 2-O-sulfate group.

Much more complex mixtures have also been analyzed in this chromatographic mode. Doneanu et al. reported the separation of oligosaccharides from dp 6 to dp 18, while Patel et al. presented the separation of oligosaccharides present in complex mixtures of depolymerized heparin containing enoxaparin <sup>[68, 70]</sup>. In both cases (

Figure 7B), the authors succeeded in separating oligosaccharides according to their size, with larger oligosaccharides showing higher retentions due to a higher total number of charges.

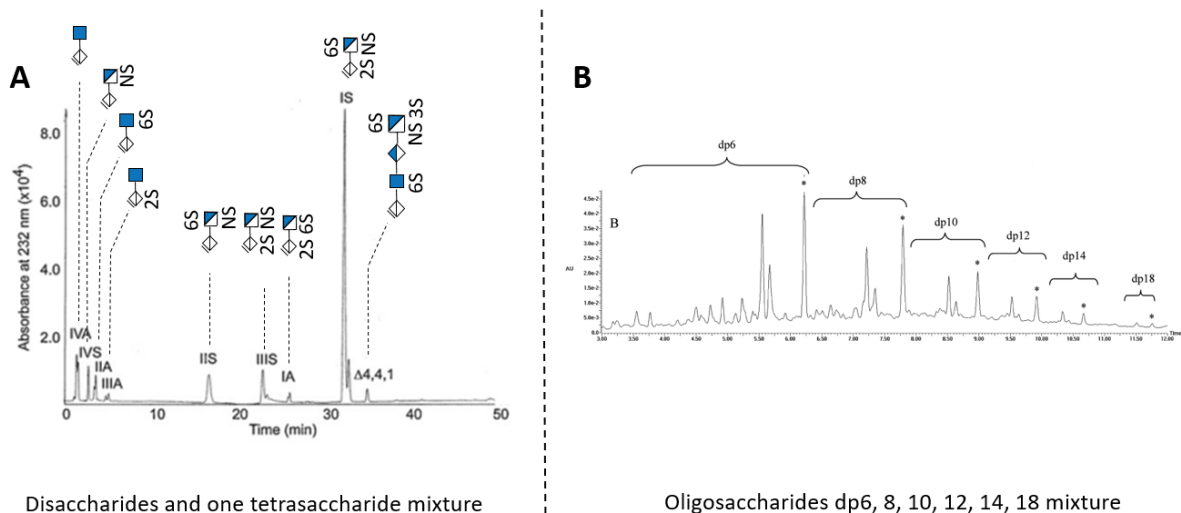
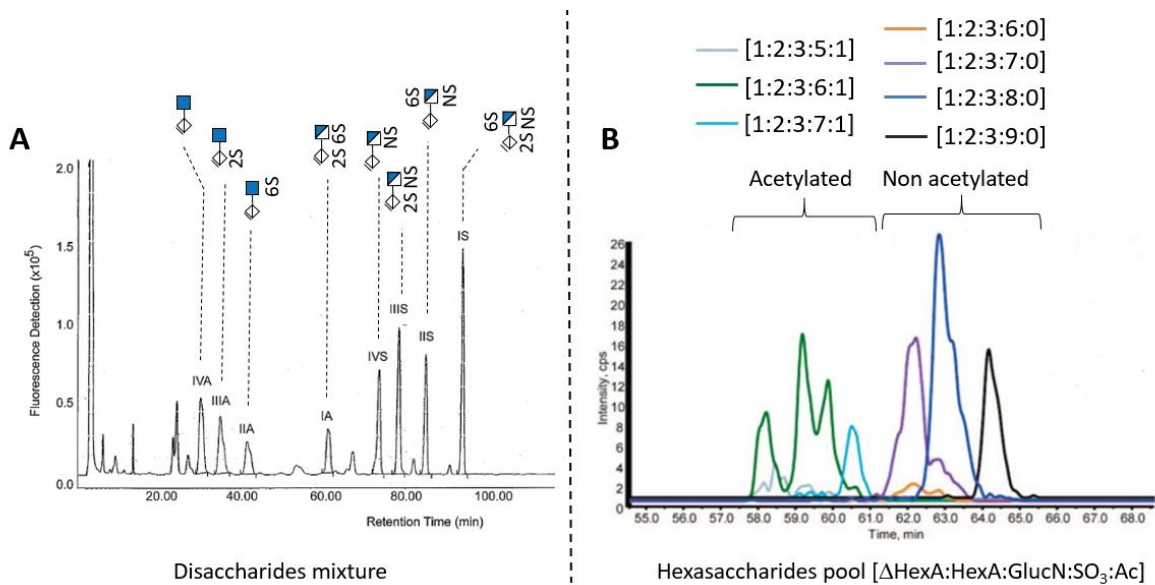


Figure 7: Chromatogram adapted from [63]. Separation of 7 disaccharides and 1 tetrasaccharide mixture by IPRP on a C18 stationary phase 100x4.6 mm eluted with a gradient between phase A (TBA 10 mM in H<sub>2</sub>O/acetonitrile 91.5/8.5) and phase B composed (TBA 10 mM and NaCl 300 mM in H<sub>2</sub>O/acetonitrile 91.5/8.5) from 0% to 98% of phase B at a flow rate of 1.0 mL/min. UV detection is performed at 232 nm (A) and chromatogram adapted from [64]. Separation of oligosaccharides dp6, 8, 10, 12, 14, 18 mixture on a C18 stationary phase 150x2.1 mm and eluted with a gradient between phase A (PTA 15mM, HFIP 50mM pH 8.8 in water) and phase B (PTA) 15mM, HFIP 50mM in water / acetonitrile 25/75) at a flow rate of 0.4 mL/min. UV detection is performed at 232 nm.

In the case of MS detection, it is necessary to select a pairing agent that can be easily removed by evaporation to avoid ionization suppression.

- Hydrophilic Interaction Liquid Chromatography (HILIC)

This chromatographic mode allows the retention of highly polar compounds, such as oligosaccharides<sup>[71]</sup>, with polar stationary phases, initially grafted with amine groups and now replaced by amide<sup>[72, 73]</sup>, diol<sup>[74]</sup> or zwitterionic<sup>[75]</sup> phases. These stationary phases allow the separation of isomeric disaccharides<sup>[75, 76]</sup> with a different selectivity (Figure 8A). In fact, even if the compounds are eluted globally according to their sulfation rate, the N-sulfate group alone induces a higher retention than a disaccharide di-sulfated in 2-OS and 6-OS. However, the effect of salt in the mobile phase on the retention should be evaluated<sup>[77]</sup>. This technique can also be used to separate oligosaccharides from more complex mixtures<sup>[74, 78]</sup> (Figure 8B).



**Figure 8:** Chromatogram adapted from <sup>[69]</sup>. Separation of 8 disaccharides on amide stationary phase 250x4.6 mm, eluted with a gradient between mobile phase A (ammonium acetate 50 mM pH 9.0) and mobile phase B (acetonitrile) from 15% to 90% of mobile phase B at a flow rate of 1.0 ml/min. Detection is performed by fluorescence exciting at 425 nm and emitting at 520 nm (A) and chromatogram adapted from <sup>[70]</sup>. Separation of a hexasaccharides pool (fractionated by SEC) on an amide stationary phase 150x0.25 mm, eluted with a gradient between mobile phase A (formic acid 50 mM in water pH 4.4) and mobile phase B (acetonitrile/mobile phase A 95/5) from 5% to 60% of mobile phase A in 60 minutes. Detection is performed by ESI-MS (B).

This chromatographic mode has the advantage of being perfectly suited for detection by mass spectrometry. Unlike the SAX and IPRP modes, the mobile phases present little or no risk of ionization suppression, except when buffers are used to control pH. In this case, the use of highly volatile inorganic salts (ammonium acetate) is recommended. In addition, the use of phases with high solvent content ensures good sensitivity <sup>[6, 64]</sup>.

#### - Capillary Electrophoresis

Gel electrophoresis allows the separation of GAGs based on charge, conformation and size <sup>[79]</sup> with higher resolution than LC methods. However, it is rarely used for GAGs due to the difficulty of coupling to mass spectrometry, and therefore structural characterisation. Nevertheless, in 2020 Sheng et al. established in-gel separation prior to in-gel digestion and MSMS analysis of a complex mixture of GAGs <sup>[60]</sup>. They used polyacrylamide gel electrophoresis (PAGE), where the migration distance is inversely proportional to the logarithm of their molecular weight without significant influence of the sulphation rate, and agarose gel electrophoresis (AGE), where the migration also depends on the molecular weight and in less extent the number of negative charges.

Due to the high negative charge density of oligosaccharides, capillary electrophoresis is particularly suitable for separating them according to their charge/size ratio under

the influence of an electric field. In normal polarity mode, oligosaccharides are separated in silica capillaries at high pH in order to increase the electroosmotic flow allowing negatively charged oligosaccharides to migrate in long analysis time [6, 80]. However, the polarity of the electric field is usually reversed and electroosmotic flow is reduced or eliminated [81–84]. Capillary electrophoresis has the advantage of being a high resolution technique with very short analysis times and low consumption of samples [64, 85]. Therefore, it has been used for oligosaccharide analysis for many years. As early as the 1990s, the separation of disaccharides derived from the depolymerization of heparin was reported several times, with good resolution even for the separation of isomers differing only in the position of the sulfate groups [82, 86]. Under conditions of reversed polarity and reduced electroosmotic flow, highly sulfated disaccharides migrate faster than slightly charged disaccharides (Figure 9) [83]. The order of elution is therefore reversed compared to SAX, IPRP and HILIC liquid chromatography techniques. On the other hand, the analysis of low molecular weight oligosaccharide mixtures is much more difficult due to the complexity of the mixtures. Due to the extreme heterogeneity of the compounds in terms of size or charge, the separation of oligosaccharides shows a poorly resolved ensemble.

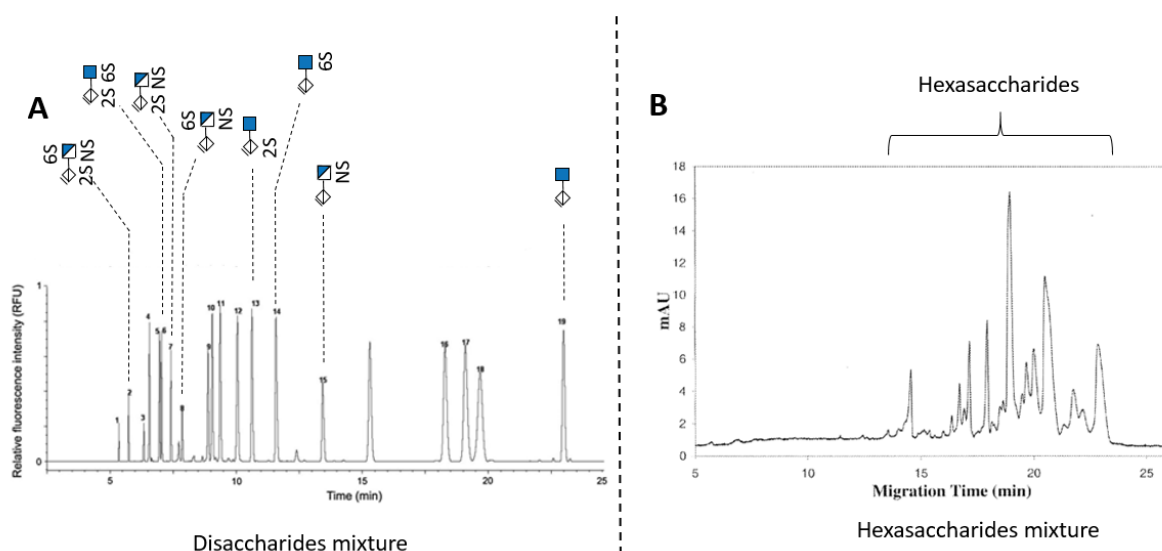


Figure 9: Electropherogram adapted of [75]. Separation in reverse polarity of 19 disaccharides issued from depolymerization of heparin, chondroitin sulfate and hyaluronan (heparin disaccharides are pointed) in silica capillary  $L=60$  cm,  $l=45$  cm,  $d_i=50$   $\mu$ m, with the use of sodium phosphate 50mM ph 3.3 electrolyte and a voltage of -25kv (A) and electropherogram adapted from [72]. Separation of hexasaccharides mixture in a silica capillary  $L=57$  cm,  $l=50$  cm,  $d_i=75$   $\mu$ m, with the use of ammonium hydrogencarbonate 30 mM ph 8.5 and triethylamine 10 mM electrolyte and a voltage of 22 kV (B)

Conclusion :With the exception of SEC, all these techniques allow the separation of disaccharides with sufficient resolution for compound identification. Similarly, all techniques are limited by the complexity of the mixtures once the oligosaccharides are large ( $dp > 6$ ). Their pros and cons are summarized in Table 3.

Table 3: Comparison of the pros and cons of the separation methods used for oligosaccharides GAGs separation.

	<b>Pros</b>	<b>Cons</b>
Size Exclusion Chromatography (SEC)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Simple size separation</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Low separation resolution</li> </ul>
Strong Anion Exchange Chromatography (SAX)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Separation of disaccharides according to the number of charges and the position of the sulfates</li> <li><input type="checkbox"/> Size separation of oligosaccharides in complex mixtures</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Low resolution of large oligosaccharides separation</li> <li><input type="checkbox"/> Difficult to couple to mass spectrometry without desalting</li> <li><input type="checkbox"/> Time of analysis</li> </ul>
Ion-Pair Chromatography (IPC)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Separation of disaccharides according to the number of charges and the position of the sulfates</li> <li><input type="checkbox"/> Coupling with mass spectrometry possible using volatile ion pairing agent</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Low resolution of large oligosaccharides separation</li> </ul>
Hydrophilic Interaction Liquid Chromatography (HILIC)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Separation of disaccharides according to the number of charges and the position of the sulfates</li> <li><input type="checkbox"/> Selectivity slightly different than with SAX</li> <li><input type="checkbox"/> Easily coupled to mass spectrometry</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Low resolution of large oligosaccharides separation</li> </ul>
Capillary Zone Electrophoresis (CE)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Low sample consumption</li> <li><input type="checkbox"/> High disaccharides separation resolutions</li> <li><input type="checkbox"/> Coupling to mass spectrometry possible</li> <li><input type="checkbox"/> Time of analysis</li> <li><input type="checkbox"/> Adapted to low sample volume</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Low resolution of large oligosaccharides separation</li> <li><input type="checkbox"/> Low sensitivity</li> </ul>

## Affinity Selection and Isolation of Interaction Motifs by Physicochemical Techniques Prior to Structural Analysis

The critical step in identifying the interaction motif(s) is the affinity selection and isolation of these motifs. Whatever the selection and isolation strategy, it always involves bringing oligosaccharide sequences of different sizes and structures (obtained after GAG depolymerization [62]) into contact with the target protein. This affinity-based selection can be performed in solution or on solid supports.

### - Selection and isolation in solution

The protein in solution is added to the oligosaccharide mixture (molar ratio  $\geq 1$ ), which allows the interaction to occur. Bound and free sequences can be separated by size or charge/size ratio.

*Separation by size.* After an affinity selection step in solution, free or protein-bound oligosaccharide sequences vary widely in size and can be easily separated by ultrafiltration using membranes with cutoffs between 5 and 100 kDa, membranes that have been used for protein purification for several decades [87].

The strategy developed by Yu et al [88] uses this approach to isolate oligosaccharide sequences with affinity for antithrombin III (58 kDa). Membranes with cut-offs of 50 kDa are used to separate unbound from bound oligosaccharide sequences. In a second step, protein-bound sequences are released from the complex by increasing the ionic strength (addition of NaCl), and isolated in a further ultrafiltration step. The advantage of this strategy is that affinity sequences can be fractionated (by controlling the ionic strength) into two pools, the first containing low-affinity sequences and the second containing high-affinity sequences (**Erreur ! Source du renvoi introuvable.**).

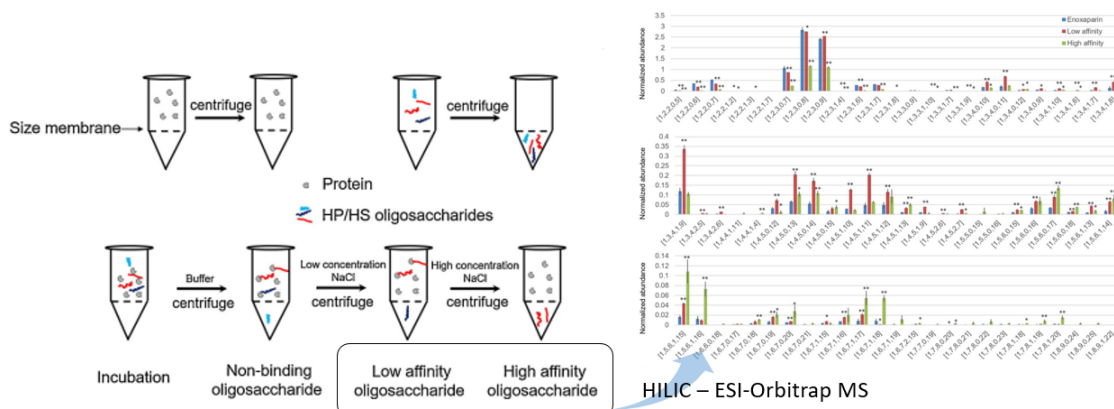


Figure 10: Ultrafiltration strategy adapted from [80]. A 50 kDa membrane is used to separate GAG sequences (enoxaparin) not retained on the protein membrane from GAG-protein complexes. GAG sequences complexed with weak interaction are released with a low concentration of NaCl. High affinity sequences are released by a high concentration of NaCl. Harvested fractions are characterized by HILIC-ESI-

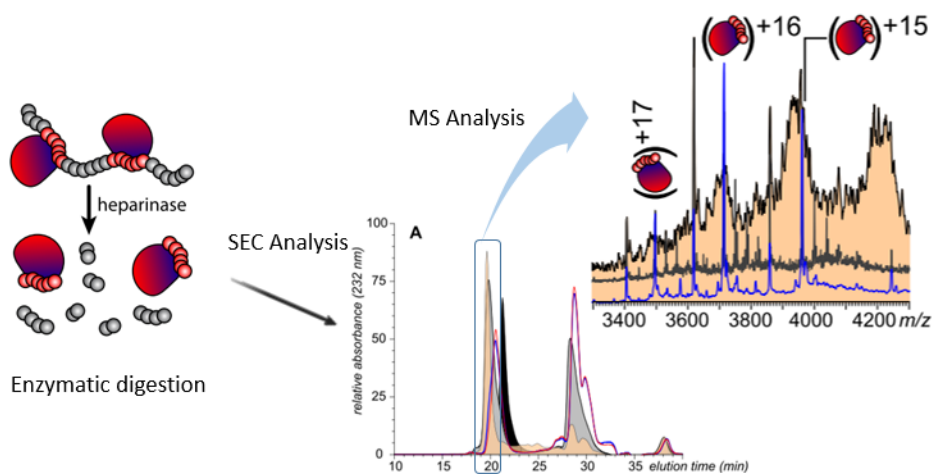
*Orbitrap MS. Identified sequences are compared between enoxaparin, the low affinity fraction and the high affinity fraction.*

Proof of concept was first demonstrated with fondaparinux, a pentasaccharide with high affinity for antithrombin III, which was detected by LC-MS in the high-affinity fraction. Secondly, enoxaparin, was used to identify low- and high-affinity sequences. The oligosaccharides isolated from the low- and high-affinity fractions were then analyzed by HILIC-FT-MS.

The results showed that (i) both fractions were depleted in oligosaccharides smaller than a pentasaccharide, (ii) the low-affinity fraction, unlike the high-affinity fraction, was enriched in sequences between the hexasaccharide and decasaccharide sizes, (iii) the high-affinity fraction was enriched in large sequences (dp12 to dp16). Finally, after digestion to tetrasaccharides, both fractions were analyzed by CZE-MSMS, which revealed the predominant presence of the structure  $\Delta\text{UA}2\text{S-GlcNS-GlcUA-GlcNS3S6S}$  and its isomeric form  $\Delta\text{UA}2\text{S-GlcNS6S-GlcUA-GlcNS3S}$ .

These results highlighted the potential of this strategy to at least enrich fractions with high affinity sequences. However, the authors were not able to identify a sequence close to the very high affinity pentasaccharide, but were only able to enrich for large oligosaccharides (dp12 to dp16).

After selection by in-solution affinity, steric exclusion chromatography was used by Niu et al. in 2020 [89] to separate the free protein, in this case antithrombin III, from a protein bound to an oligosaccharide chain. The authors formed the antithrombin-heparin complex from dp20 oligosaccharide sequences in solution. Since the heparin sites complexed with antithrombin III are inaccessible to enzymes, the oligosaccharide sequence involved in the complex was then depolymerized by enzymatic digestion. Sufficient time for enzymatic digestion allows the formation of antithrombin III-oligosaccharide complexes with the shortest affinity sequences. The change in complex size with depolymerization time can be observed by SEC, with retention related to molecule size. Finally, the complexes are characterized by high-resolution FTICR mass spectrometry (Figure 10).



*Figure 10: : Diagram of the strategy adapted from [89]. GAGs from GAG-protein are depolymerized by enzymatic digestion to produce complex with the lowest GAG size.*



The mixture is separated by SEC on a TSK gel 2000SWXL column with ammonium acetate pH 6.9 as mobile phase at a flow rate of 0.3 mL/min. Detection is performed by ESI-FTICR-MS

The authors were able to identify a majority complex formed between antithrombin and a hexasaccharide containing 6 sulfate groups and 1 acetyl group. Two other oligosaccharides bound to antithrombin were detected in the minority and with low intensity: an octasaccharide and a decasaccharide containing 9 and 12 sulfate groups and 1 acetyl group, respectively. The potential of this approach was demonstrated with this hexasaccharide, whose structure seems to be very close to that of the high-affinity pentasaccharide.

More recently, the use of native ESI-MS for heparin-antithrombin complexes in solution studies has emerged. Using ion mobility coupling, Zhao et al. identified the stoichiometry and length of the heparin chains involved in the complexes. However, the mass spectra are still difficult to interpret and coupling with a chromatographic step could be considered <sup>[90]</sup>.

*Separation by charge/size ratio.* After an affinity selection step in solution, free or protein-bound oligosaccharide sequences have very different charge/size ratios and can be separated by zone capillary electrophoresis. In 2002, Militopoulou et al. used a special strategy to identify affine sequences by comparing the electrophoretic profiles obtained for a mixture of oligosaccharides (prepared from enzymatically depolymerized heparan sulfate) in the presence and absence of protein <sup>[91]</sup>.

The authors first developed a method for separating oligosaccharide sequences by capillary electrophoresis. By eliminating electroosmotic flow in combination with polarity reversal, they were able to separate free oligosaccharide sequences (migrating toward the UV detector) from protein or oligosaccharide-protein complexes migrating in the opposite direction. Five oligosaccharide clusters, differentiated by charge density, were separated and detected under UV light at 232 nm (Figure 11).

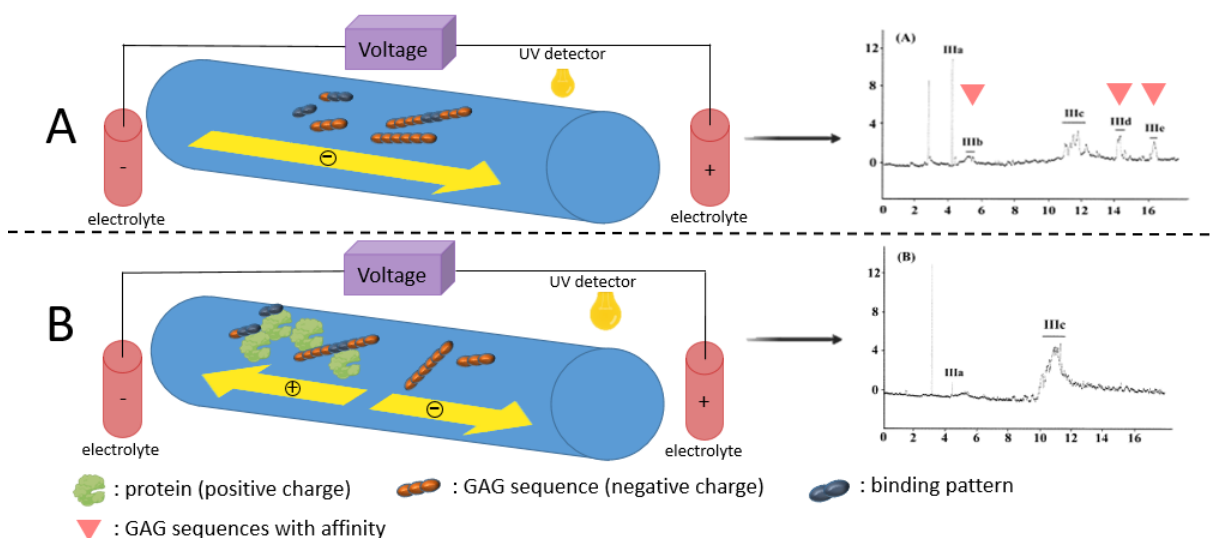


Figure 11: Diagram adapted from <sup>[82]</sup> of the protein affinity sequence detection strategy based on the separation of free gag sequences from gag-protein complexes by CE in a silica capillary (L=64.5 cm, l=56 cm, di=50  $\mu$ m) with a phosphate 50 mM pH 3.5 electrolyte and an applied voltage of -30kV. Detection is performed at 232 nm.

*Gag sequences migrate towards the detector and are separated according to their charge/size ratio into 5 groups iia, iib, iic, iid and iie (A). In the presence of a target protein, complexes do not migrate towards the detector, unlike non-interacting gag sequences. only iia and iic groups are detected (B).*

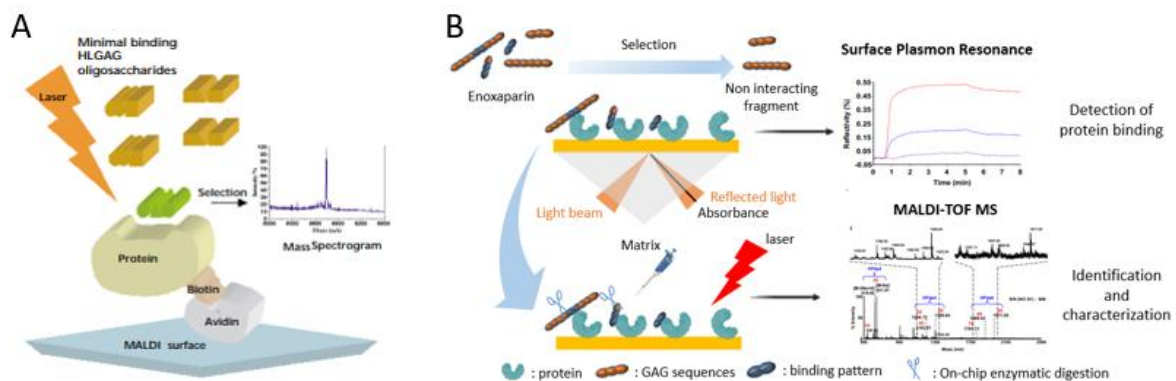
- Solid Support Selection and Isolation

Another way to isolate oligosaccharide sequences with affinity for a target protein is to immobilize the protein before the protein-oligosaccharide complex is formed. The protein can be immobilized on a flat support (MALDI plate, SPR sensor) or on a three-dimensional support (chromatography column).

In 2001, Keiser et al. developed a method called Surface Noncovalent Association Mass Spectrometry (SNA-MS) [92], which consists of immobilizing a protein on a MALDI plate where the protein-oligosaccharide complexes are formed directly prior to analysis by mass spectrometry (Figure 12 A). This is achieved by first adsorbing avidin onto the hydrophobic surface of a MALDI plate.

The target protein, in this case antithrombin III or FGF 2, is then biotinylated and immobilized on the plate thanks to the very strong interaction between avidin and biotin ( $K_d = 10^{-14}$  M). A mixture of oligosaccharides is then placed into contact with the protein, allowing affinity sequences to form protein-oligosaccharide complexes. Sequences with low or no affinity are washed out with water and then a saline solution. Since the complex is formed directly on a MALDI plate, the addition of a matrix allows MS analysis. An enzymatic digestion step can be added prior to analysis to retain only protein-bound sequences.

A similar immobilization strategy was later implemented by Przybylski et al. in 2019 [93], using an SPR sensor instead of a MALDI plate. This type of coupling had already been used to study protein-protein interactions [94, 95]. For this purpose, the biological sensors used in SPR are grafted with a short hydrophobic chain of polyethylene oxide, which allows the non-covalent immobilization of a target protein, in this case the chemokines IFN- $\gamma$  and SDF-1 $\alpha$  (Figure 12 B). Oligosaccharide mixtures derived from heparan sulfate or heparin are then released onto the sensors carrying the target proteins. Coupling with SPRi (SPR Imaging) at this point offers the advantage of following the formation of protein-oligosaccharide complexes to saturation and measuring their dissociation constants. The oligosaccharide chains captured by the sensor are then enzymatically digested and analyzed by MALDI-TOF MS.



*Figure 12: Affinity sequence identification strategy after immobilization of the target protein on a solid surface by a: immobilization on a maldi plate and maldi-tof ms detection of protein-gag complexes. Scheme adapted from <sup>[83]</sup> (A) and immobilization of the protein on a SPR sensor followed by maldi-tof ms analysis directly on the sensor. Diagram adapted from <sup>[84]</sup> (B).*

The strategy of Keiser et al. was first validated by immobilizing antithrombin III and using a mixture containing the pentasaccharide with high affinity for antithrombin III and two hexasaccharides. After washing with 0.2M NaCl, the pentasaccharide was the only sequence detected. The strategy was then applied to FGF-2, in three different ways: (i) with a mixture of two hexasaccharides, one of which is known to have a high affinity for FGF-2, (ii) with a complex mixture of hexasaccharides obtained by heparinase I depolymerization of porcine heparin, and (iii) with heparan sulfate extracted from vascular smooth muscle and depolymerized by heparinase I or heparinase III. In each case, a hexasaccharide sequence was identified: (i) the hexasaccharide known for its affinity with FGF 2, (ii) two hexasaccharides with sequences  $\Delta U2SHNS,6SI2SHNS,6SI2SHNS,6S$  and  $\Delta U2SHNS,6SI2SHNS,6SGHNS,6S$  and (iii) one hexasaccharide whose structure was not identified. Thus, this study highlighted the potential of this strategy by identifying hexasaccharides with high affinity for FGF 2.

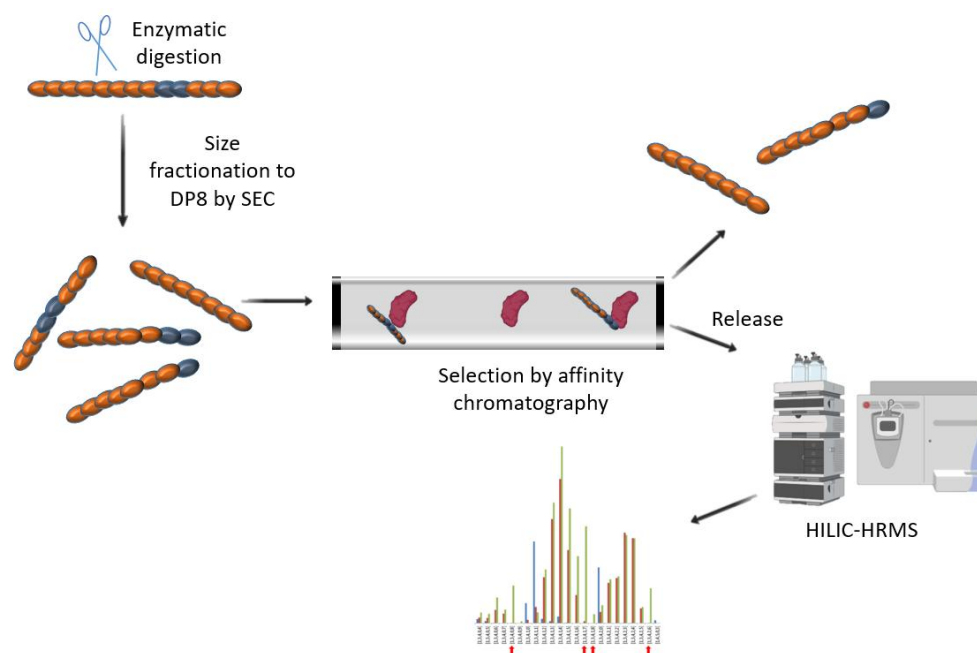
The strategy of Przybylski et al. was applied to four proteins: IFN- $\gamma$ , SDF-1 $\alpha$ , MCP-1 and aprotinin, which was used as a control because of its low affinity for heparin. Two types of oligosaccharide mixtures were used with these sensors, prior to analysis by MALDI-TOF MS: a complex mixture of heparin-derived decasaccharides and a solution of fondaparinux, the pentasaccharide with high affinity for antithrombin III. The results showed that protein-affinity oligosaccharide sequences are detectable. However, the complexity of the oligosaccharide mixtures proved to be a limiting factor in the identification of affinity sequences.

#### - Immobilization on Affinity Columns

Instead of using a flat support such as a MALDI plate or an SPR sensor, protein immobilisation can be carried out on a three-dimensional support such as a chromatographic support to assess the affinity of ligands for the grafted protein. Affinity chromatography is often used to separate low affinity depolymerized heparin fractions from high affinity fractions. As early as 2008, Guerini et al. used this technique to isolate octasaccharide fractions with affinity for antithrombin prior to fractionation by anion exchange chromatography and NMR structural analysis <sup>[96]</sup>. The authors were able to identify an octasaccharide containing the pentasaccharide sequence, although the measured dissociation constant did not agree with previous data. On the other hand, they were able to show that the monosaccharide residues linked to this pentasaccharide are likely to improve the affinity, in particular the presence of a glucuronic acid rather than an iduronic acid.

This technique was used by Zong et al. in 2016 <sup>[97]</sup> in an integrated strategy that combined affinity chromatography to isolate affinity sequences and HILIC-HRMS for their structural characterization. This strategy was applied to the Robo1 (human roundabout receptor 1) protein and heparan sulfate. The authors first depolymerized heparan sulfate and then purified it by steric exclusion chromatography to obtain a mixture of octasaccharides. The octasaccharide affinity sequences were then isolated and enriched by capture on an affinity column, and then collected by rinsing

with saline. The affinity sequences were then separated by HILIC before detection by MSMS mass spectrometry (Figure 12). Four octasaccharides were identified with the compositions  $[\Delta\text{HexA}, \text{HexA}, \text{GlcN}, \text{Ac}, \text{SO}_3]$ : [1,3,4,0,8], [1,3,4,1,7], [1,3,4,1,8], and [1,3,4,2,6], with the [1,3,4,1,7] composition being in the majority.



*Figure 13: Diagram adapted from the method used by [88]. After enzymatic digestion and fractionation by SEC, dp8 oligosaccharides are selected by affinity chromatography grafted with the target protein (Robo1-GFP) and then released in an ammonium acetate solution. The released compounds are fractionated by hilic on an amide stationary phase 110x0.08 mm. Elution is performed with a gradient between mobile phase A (ammonium formate 55mM in water / acetonitrile 80/20) and mobile phase B (acetonitrile / mobile phase A 95/5) from 65% to 70% of mobile phase B at a flow rate of 0.135 mL/min. Detection is performed by HRMS on an orbitrap in negative ion mode.*

A similar strategy was used by Shi et al. in 2022 [63]. The authors used an original approach based on the retention factors of compounds in anion exchange chromatography, which they compared with predicted ones. This strategy allowed them to identify the pentasaccharide with the highest affinity for antithrombin. In addition, they applied this strategy to the study of the interaction between heparin and IFN- $\gamma$  and were able to identify a pentasaccharide [GlcNS6S-IdoA2S-GlcNS6S-IdoA2S-GlcNS6S] as the sequence with the lowest affinity.

### Summary of isolation strategies

The various isolation strategies developed over the last two decades have shown great potential for the identification of interaction sequences. Although none of the strategies used to study the interaction between heparin and antithrombin has been able to identify a single high affinity sequence with certainty, they have been able to approximate the known pentasaccharide structure by identifying high levels of sulfation or even specific sulfate groups. Strategies combining immobilization on solid

surfaces or affinity columns, depolymerization and structural analysis seem likely to overcome the challenges still posed by the extreme complexity of GAG mixtures.

Although the interaction between heparin and antithrombin most of the time serves as a model to demonstrate the feasibility of strategies, other interactions are of interest and have already been the subject of studies, without the interaction site having been clearly identified to date. Interaction studies with proteins such as IFN- $\gamma$  [63, 93], FGF2 [92], and Robo1 [97] have led to advances in the identification of an interaction sequence. Potential ligands have already been proposed, corresponding to highly sulfated chains of 4 to 8 monosaccharides. While the high degree of sulfation of heparin or heparan sulfate allows for numerous interactions, making them more amenable to studies, chondroitin sulfate also presents a particular interest for its interactions with growth factors [30] and viruses such as SARS-CoV-2 [98], for example. However, apart from microarray-based strategies, there are no proposed studies with this GAG based on affinity isolation before structural analysis.

Confirming the feasibility of a strategy using the antithrombin-heparin model could expand the application domain to other GAGs besides heparin and heparan sulfate. In this regard, strategies for isolation after immobilization of a protein appear to be the most promising today. In particular, affinity chromatography has recently shown interesting potential for the isolation and, most importantly, enrichment of high-affinity sequences. Given the number of interactions of GAGs with proteins and their therapeutic potential, this type of strategy, directly coupled with mass spectrometry characterization techniques that have made significant advancements over the past two decades [58], could become increasingly significant in the future. This would enable the identification of new oligosaccharide ligands and, consequently, a better understanding of the mechanisms in which they are involved, including the regulation by cells of their production, which remains a mystery in the biological field [40].

## References

- [1] Altelaar, A. F. M.; Munoz, J.; Heck, A. J. R. Next-Generation Proteomics: Towards an Integrative View of Proteome Dynamics. *Nature Reviews Genetics*, **2013**, *14* (1), 35–48. <https://doi.org/10.1038/nrg3356>.
- [2] Perrotti, F.; Rosa, C.; Cicalini, I.; Sacchetta, P.; Del Boccio, P.; Genovesi, D.; Pieragostino, D. Advances in Lipidomics for Cancer Biomarkers Discovery. *International Journal of Molecular Sciences*, **2016**, *17* (12). <https://doi.org/10.3390/ijms17121992>.
- [3] Griffin, M. E.; Hsieh-Wilson, L. C. Glycan Engineering for Cell and Developmental Biology. *Cell Chemical Biology*, **2016**, *23* (1), 108–121. <https://doi.org/10.1016/j.chembiol.2015.12.007>.
- [4] Kulkarni, A. A.; Weiss, A. A.; Iyer, S. S. Glycan-Based High-Affinity Ligands for Toxins and Pathogen Receptors. *Medicinal Research Reviews*, **2010**, *30* (2), 327–393. <https://doi.org/10.1002/med.20196>.
- [5] Varki, A. Glycan-Based Interactions Involving Vertebrate Sialic-Acid-Recognizing Proteins. *Nature*, **2007**, *446* (7139), 1023–1029. <https://doi.org/10.1038/nature05816>.
- [6] Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J. Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Molecular & Cellular Proteomics*, **2021**, *20*, 100025. <https://doi.org/10.1074/mcp.R120.002267>.
- [7] Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Evers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; et al. Advancing Solutions to the Carbohydrate Sequencing Challenge. *J. Am. Chem. Soc.*, **2019**, *141* (37), 14463–14479. <https://doi.org/10.1021/jacs.9b06406>.

- [8] Kamhi, E.; Joo, E. J.; Dordick, J. S.; Linhardt, R. J. Glycosaminoglycans in Infectious Disease. *Biological Reviews*, **2013**, *88* (4), 928–943. <https://doi.org/10.1111/brv.12034>.
- [9] Morla, S. Glycosaminoglycans and Glycosaminoglycan Mimetics in Cancer and Inflammation. *International Journal of Molecular Sciences*, **2019**, *20* (8). <https://doi.org/10.3390/ijms20081963>.
- [10] Kusche-Gullberg, M.; Kjellén, L. Sulfotransferases in Glycosaminoglycan Biosynthesis. *Current Opinion in Structural Biology*, **2003**, *13* (5), 605–611. <https://doi.org/10.1016/j.sbi.2003.08.002>.
- [11] Grobe, K.; Ledin, J.; Ringvall, M.; Holmborn, K.; Forsberg, E.; Esko, J. D.; Kjellén, L. Heparan Sulfate and Development: Differential Roles of the N-Acetylglucosamine N-Deacetylase/N-Sulfotransferase Isozymes. *Biochimica et Biophysica Acta (BBA) - General Subjects*, **2002**, *1573* (3), 209–215. [https://doi.org/10.1016/S0304-4165\(02\)00386-0](https://doi.org/10.1016/S0304-4165(02)00386-0).
- [12] Rabenstein, D. L. Heparin and Heparan Sulfate: Structure and Function. *Nat. Prod. Rep.*, **2002**, *19* (3), 312–331. <https://doi.org/10.1039/B100916H>.
- [13] TOIDA, T.; YOSHIDA, H.; TOYODA, H.; KOSHIISHI, I.; IMANARI, T.; HILEMAN, R. E.; FROMM, J. R.; LINHARDT, R. J. Structural Differences and the Presence of Unsubstituted Amino Groups in Heparan Sulphates from Different Tissues and Species. *Biochemical Journal*, **1997**, *322* (2), 499–506. <https://doi.org/10.1042/bj3220499>.
- [14] Björk, I.; Olson, S. T. Antithrombin. In *Chemistry and Biology of Serpins*; Church, F. C., Cunningham, D. D., Ginsburg, D., Hoffman, M., Stone, S. R., Tollefsen, D. M., Eds.; Springer US: Boston, MA, 1997; pp 17–33. [https://doi.org/10.1007/978-1-4615-5391-5\\_3](https://doi.org/10.1007/978-1-4615-5391-5_3).
- [15] Olson, S. T.; Richard, B.; Izaguirre, G.; Schedin-Weiss, S.; Gettins, P. G. W. Molecular Mechanisms of Antithrombin–Heparin Regulation of Blood Clotting Proteinases. A Paradigm for Understanding Proteinase Regulation by Serpin Family Protein Proteinase Inhibitors. *Biochimie*, **2010**, *92* (11), 1587–1596. <https://doi.org/10.1016/j.biochi.2010.05.011>.
- [16] Rezaie, A. R.; Giri, H. Anticoagulant and Signaling Functions of Antithrombin. *Journal of Thrombosis and Haemostasis*, **2020**, *18* (12), 3142–3153. <https://doi.org/10.1111/jth.15052>.
- [17] Pellegrini, L.; Burke, D. F.; von Delft, F.; Mulloy, B.; Blundell, T. L. Crystal Structure of Fibroblast Growth Factor Receptor Ectodomain Bound to Ligand and Heparin. *Nature*, **2000**, *407* (6807), 1029–1034. <https://doi.org/10.1038/35039551>.
- [18] Schlessinger, J.; Plotnikov, A. N.; Ibrahimi, O. A.; Eliseenkova, A. V.; Yeh, B. K.; Yayon, A.; Linhardt, R. J.; Mohammadi, M. Crystal Structure of a Ternary FGF-FGFR-Heparin Complex Reveals a Dual Role for Heparin in FGFR Binding and Dimerization. *Molecular Cell*, **2000**, *6* (3), 743–750. [https://doi.org/10.1016/S1097-2765\(00\)00073-3](https://doi.org/10.1016/S1097-2765(00)00073-3).
- [19] Mohammadi, M.; Olsen, S. K.; Ibrahimi, O. A. Structural Basis for Fibroblast Growth Factor Receptor Activation. *Cytokine & Growth Factor Reviews*, **2005**, *16* (2), 107–137. <https://doi.org/10.1016/j.cytogfr.2005.01.008>.
- [20] Ornitz, D. M.; Yayon, A.; Flanagan, J. G.; Svahn, C. M.; Levi, E.; Leder, P. Heparin Is Required for Cell-Free Binding of Basic Fibroblast Growth Factor to a Soluble Receptor and for Mitogenesis in Whole Cells. *Molecular and Cellular Biology*, **1992**, *12* (1), 240–247. <https://doi.org/10.1128/mcb.12.1.240-247.1992>.
- [21] Vlodavsky, I.; Goldshmidt, O.; Zcharia, E.; Atzmon, R.; Rangini-Guatta, Z.; Elkin, M.; Peretz, T.; Friedmann, Y. Mammalian Heparanase: Involvement in Cancer Metastasis, Angiogenesis and Normal Development. *Seminars in Cancer Biology*, **2002**, *12* (2), 121–129. <https://doi.org/10.1006/scbi.2001.0420>.
- [22] Bernfield, M.; Götte, M.; Park, P. W.; Reizes, O.; Fitzgerald, M. L.; Lincecum, J.; Zako, M. Functions of Cell Surface Heparan Sulfate Proteoglycans. *Annu. Rev. Biochem.*, **1999**, *68* (1), 729–777. <https://doi.org/10.1146/annurev.biochem.68.1.729>.
- [23] Proudfoot, A. E. I.; Handel, T. M.; Johnson, Z.; Lau, E. K.; LiWang, P.; Clark-Lewis, I.; Borlat, F.; Wells, T. N. C.; Kosco-Vilbois, M. H. Glycosaminoglycan Binding and Oligomerization Are Essential for the in Vivo Activity of Certain Chemokines. *Proc Natl Acad Sci U S A*, **2003**, *100* (4), 1885–1890. <https://doi.org/10.1073/pnas.0334864100>.

- [24] RUIZ-BUSTOS, E.; OCHOA, J. L.; WADSTRÖM, T.; ASCENCIO, F. Isolation and Characterisation of Putative Adhesins from *Helicobacter Pylori* with Affinity for Heparan Sulphate Proteoglycan. *Journal of Medical Microbiology*, 2001, 50, 215–222. <https://doi.org/10.1099/0022-1317-50-3-215>.
- [25] Chen, Y.; Hayashida, A.; Bennett, A. E.; Hollingshead, S. K.; Park, P. W. Streptococcus Pneumoniae Sheds Syndecan-1 Ectodomains through ZmpC, a Metalloproteinase Virulence Factor\*. *Journal of Biological Chemistry*, 2007, 282 (1), 159–167. <https://doi.org/10.1074/jbc.M608542200>.
- [26] Carruthers Vern B.; Håkansson Sebastian; Giddings Olivia K.; Sibley L. David. Toxoplasma Gondii Uses Sulfated Proteoglycans for Substrate and Host Cell Attachment. *Infection and Immunity*, 2000, 68 (7), 4005–4011. <https://doi.org/10.1128/iai.68.7.4005-4011.2000>.
- [27] Delehedde, M.; Lortat-Jacob, H.; Gallagher, T. J.; Bonnaffe, D.; Adam, E.; Querolle, O.; Lequien, S.; Degove, S.; Lassalle, P.; Bechard, D. Proteoglycan Involvement in Inflammatory Diseases. New Developments in GAG-Based Therapies. *Medicinal Chemistry Reviews - Online (Discontinued)*, 2005, 2 (4), 345–357. <https://doi.org/10.2174/1567203054637560>.
- [28] Ye Chen; Martin Goette; Jian Liu and Pyong Woo Park. Microbial Subversion of Heparan Sulfate Proteoglycans. *Mol. Cells*, 2008, 26 (5), 415–426.
- [29] Mondor Isabelle; Ugolini Sophie; Sattentau Quentin J. Human Immunodeficiency Virus Type 1 Attachment to HeLa CD4 Cells Is CD4 Independent and Gp120 Dependent and Requires Cell Surface Heparans. *Journal of Virology*, 1998, 72 (5), 3623–3634. <https://doi.org/10.1128/jvi.72.5.3623-3634.1998>.
- [30] Mosier, P. D.; Krishnasamy, C.; Kellogg, G. E.; Desai, U. R. On the Specificity of Heparin/Heparan Sulfate Binding to Proteins. Anion-Binding Sites on Antithrombin and Thrombin Are Fundamentally Different. *PLOS ONE*, 2012, 7 (11), e48632. <https://doi.org/10.1371/journal.pone.0048632>.
- [31] Sarkar, A.; Desai, U. R. A Simple Method for Discovering Druggable, Specific Glycosaminoglycan-Protein Systems. Elucidation of Key Principles from Heparin/Heparan Sulfate-Binding Proteins. *PLOS ONE*, 2015, 10 (10), e0141127. <https://doi.org/10.1371/journal.pone.0141127>.
- [32] Petitou, M.; Casu, B.; Lindahl, U. 1976–1983, a Critical Period in the History of Heparin: The Discovery of the Antithrombin Binding Site. *Biochimie*, 2003, 85 (1), 83–89. [https://doi.org/10.1016/S0300-9084\(03\)00078-6](https://doi.org/10.1016/S0300-9084(03)00078-6).
- [33] Quinsey, N. S.; Greedy, A. L.; Bottomley, S. P.; Whisstock, J. C.; Pike, R. N. Antithrombin: In Control of Coagulation. *The International Journal of Biochemistry & Cell Biology*, 2004, 36 (3), 386–389. [https://doi.org/10.1016/S1357-2725\(03\)00244-9](https://doi.org/10.1016/S1357-2725(03)00244-9).
- [34] Desai, U. R.; Petitou, M.; Björk, I.; Olson, S. T. Mechanism of Heparin Activation of Antithrombin: ROLE OF INDIVIDUAL RESIDUES OF THE PENTASACCHARIDE ACTIVATING SEQUENCE IN THE RECOGNITION OF NATIVE AND ACTIVATED STATES OF ANTITHROMBIN\*. *Journal of Biological Chemistry*, 1998, 273 (13), 7478–7487. <https://doi.org/10.1074/jbc.273.13.7478>.
- [35] PETITOU, M.; LORMEAU, J.-C.; CHOAY, J. Interaction of Heparin and Antithrombin III. *European Journal of Biochemistry*, 1988, 176 (3), 637–640. <https://doi.org/10.1111/j.1432-1033.1988.tb14324.x>.
- [36] Xu, D.; Esko, J. D. Demystifying Heparan Sulfate–Protein Interactions. *Annu. Rev. Biochem.*, 2014, 83 (1), 129–157. <https://doi.org/10.1146/annurev-biochem-060713-035314>.
- [37] Gama, C. I.; Tully, S. E.; Sotogaku, N.; Clark, P. M.; Rawat, M.; Vaidehi, N.; Goddard, W. A.; Nishi, A.; Hsieh-Wilson, L. C. Sulfation Patterns of Glycosaminoglycans Encode Molecular Recognition and Activity. *Nature Chemical Biology*, 2006, 2 (9), 467–473. <https://doi.org/10.1038/nchembio810>.
- [38] Richard, B.; Swanson, R.; Olson, S. T. The Signature 3-O-Sulfo Group of the Anticoagulant Heparin Sequence Is Critical for Heparin Binding to Antithrombin but Is Not Required for Allosteric Activation\*. *Journal of Biological Chemistry*, 2009, 284 (40), 27054–27064. <https://doi.org/10.1074/jbc.M109.029892>.

- [39] Lindahl, U.; Li, J. Chapter 3 Interactions Between Heparan Sulfate and Proteins—Design and Functional Implications. In *International Review of Cell and Molecular Biology*; Academic Press, 2009; Vol. 276, pp 105–159. [https://doi.org/10.1016/S1937-6448\(09\)76003-4](https://doi.org/10.1016/S1937-6448(09)76003-4).
- [40] Kjellén, L.; Lindahl, U. Specificity of Glycosaminoglycan–Protein Interactions. *Current Opinion in Structural Biology*, **2018**, *50*, 101–108. <https://doi.org/10.1016/j.sbi.2017.12.011>.
- [41] Danielsson, A.; Kogut, M. M.; Maszota-Zieleniak, M.; Chopra, P.; Boons, G.-J.; Samsonov, S. A. Molecular Dynamics-Based Descriptors of 3-O-Sulfated Heparan Sulfate as Contributors of Protein Binding Specificity. *Computational Biology and Chemistry*, **2022**, *99*, 107716. <https://doi.org/10.1016/j.compbiolchem.2022.107716>.
- [42] Shukla, D.; Liu, J.; Blaiklock, P.; Shworak, N. W.; Bai, X.; Esko, J. D.; Cohen, G. H.; Eisenberg, R. J.; Rosenberg, R. D.; Spear, P. G. A Novel Role for 3-O-Sulfated Heparan Sulfate in Herpes Simplex Virus 1 Entry. *Cell*, **1999**, *99* (1), 13–22. [https://doi.org/10.1016/S0092-8674\(00\)80058-6](https://doi.org/10.1016/S0092-8674(00)80058-6).
- [43] Thacker, B. E.; Seamen, E.; Lawrence, R.; Parker, M. W.; Xu, Y.; Liu, J.; Vander Kooi, C. W.; Esko, J. D. Expanding the 3-O-Sulfate Proteome—Enhanced Binding of Neuropilin-1 to 3-O-Sulfated Heparan Sulfate Modulates Its Activity. *ACS Chem. Biol.*, **2016**, *11* (4), 971–980. <https://doi.org/10.1021/acscchembio.5b00897>.
- [44] Pempe, E. H.; Xu, Y.; Gopalakrishnan, S.; Liu, J.; Harris, E. N. Probing Structural Selectivity of Synthetic Heparin Binding to Stabilin Protein Receptors\*. *Journal of Biological Chemistry*, **2012**, *287* (25), 20774–20783. <https://doi.org/10.1074/jbc.M111.320069>.
- [45] Raman, R.; Venkataraman, G.; Ernst, S.; Sasisekharan, V.; Sasisekharan, R. Structural Specificity of Heparin Binding in the Fibroblast Growth Factor Family of Proteins. *Proceedings of the National Academy of Sciences*, **2003**, *100* (5), 2357–2362. <https://doi.org/10.1073/pnas.0437842100>.
- [46] Raman, R.; Sasisekharan, V.; Sasisekharan, R. Structural Insights into Biological Roles of Protein-Glycosaminoglycan Interactions. *Chemistry & Biology*, **2005**, *12* (3), 267–277. <https://doi.org/10.1016/j.chembiol.2004.11.020>.
- [47] Langeslay, D. J.; Young, R. P.; Beni, S.; Beecher, C. N.; Mueller, L. J.; Larive, C. K. Sulfamate Proton Solvent Exchange in Heparin Oligosaccharides: Evidence for a Persistent Hydrogen Bond in the Antithrombin-Binding Pentasaccharide Arixtra. *Glycobiology*, **2012**, *22* (9), 1173–1182. <https://doi.org/10.1093/glycob/cws085>.
- [48] Singh, P.; Singh, K.; Jairajpuri, M. A. Energetics of Hydrogen Bond Switch, Residue Burial and Cavity Analysis Reveals Molecular Basis of Improved Heparin Binding to Antithrombin. *Journal of Biomolecular Structure and Dynamics*, **2011**, *29* (2), 339–350. <https://doi.org/10.1080/07391102.2011.10507389>.
- [49] Turnbull, J. E.; Fernig, D. G.; Ke, Y.; Wilkinson, M. C.; Gallagher, J. T. Identification of the Basic Fibroblast Growth Factor Binding Sequence in Fibroblast Heparan Sulfate. *Journal of Biological Chemistry*, **1992**, *267* (15), 10337–10341. [https://doi.org/10.1016/S0021-9258\(19\)50023-0](https://doi.org/10.1016/S0021-9258(19)50023-0).
- [50] Kreuger, J.; Salmivirta, M.; Sturiale, L.; Giménez-Gallego, G.; Lindahl, U. Sequence Analysis of Heparan Sulfate Epitopes with Graded Affinities for Fibroblast Growth Factors 1 and 2\*. *Journal of Biological Chemistry*, **2001**, *276* (33), 30744–30752. <https://doi.org/10.1074/jbc.M102628200>.
- [51] Jastrebova, N.; Vanwildemeersch, M.; Lindahl, U.; Spillmann, D. Heparan Sulfate Domain Organization and Sulfation Modulate FGF-Induced Cell Signaling\*. *Journal of Biological Chemistry*, **2010**, *285* (35), 26842–26851. <https://doi.org/10.1074/jbc.M109.093542>.
- [52] Mulloy, B. The Specificity of Interactions between Proteins and Sulfated Polysaccharides. *Anais da Academia Brasileira de Ciências*, **2005**, *77*.
- [53] Ashikari-Hada, S.; Habuchi, H.; Kariya, Y.; Itoh, N.; Reddi, A. H.; Kimata, K. Characterization of Growth Factor-Binding Structures in Heparin/Heparan Sulfate Using an Octasaccharide Library\*. *Journal of Biological Chemistry*, **2004**, *279* (13), 12346–12354. <https://doi.org/10.1074/jbc.M313523200>.
- [54] Olson, S. T.; Halvorson, H. R.; Björk, I. Quantitative Characterization of the Thrombin-Heparin Interaction. Discrimination between Specific and Nonspecific Binding Models. *Journal of*



- Biological Chemistry*, **1991**, 266 (10), 6342–6352. [https://doi.org/10.1016/S0021-9258\(18\)38124-9](https://doi.org/10.1016/S0021-9258(18)38124-9).
- [55] Lyon, M.; Deakin, J. A.; Gallagher, J. T. The Mode of Action of Heparan and Dermatan Sulfates in the Regulation of Hepatocyte Growth Factor/Scatter Factor\*. *Journal of Biological Chemistry*, **2002**, 277 (2), 1040–1046. <https://doi.org/10.1074/jbc.M107506200>.
- [56] Catlow, K. R.; Deakin, J. A.; Wei, Z.; Delehedde, M.; Fernig, D. G.; Gherardi, E.; Gallagher, J. T.; Pavão, M. S. G.; Lyon, M. Interactions of Hepatocyte Growth Factor/Scatter Factor with Various Glycosaminoglycans Reveal an Important Interplay between the Presence of Iduronate and Sulfate Density\*. *Journal of Biological Chemistry*, **2008**, 283 (9), 5235–5248. <https://doi.org/10.1074/jbc.M706589200>.
- [57] Gallagher, J. Fell–Muir Lecture: Heparan Sulphate and the Art of Cell Regulation: A Polymer Chain Conducts the Protein Orchestra. *International Journal of Experimental Pathology*, **2015**, 96 (4), 203–231. <https://doi.org/10.1111/iep.12135>.
- [58] Wolff, J. J.; Leach, F. E. I.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. Negative Electron Transfer Dissociation of Glycosaminoglycans. *Anal. Chem.*, **2010**, 82 (9), 3460–3466. <https://doi.org/10.1021/ac100554a>.
- [59] Schindler, B.; Renois-Predelus, G.; Bagdadi, N.; Melizi, S.; Barnes, L.; Chambert, S.; Allouche, A.-R.; Compagnon, I. MS/IR, a New MS-Based Hyphenated Method for Analysis of Hexuronic Acid Epimers in Glycosaminoglycans. *Glycoconjugate Journal*, **2017**, 34 (3), 421–425. <https://doi.org/10.1007/s10719-016-9741-8>.
- [60] Sheng, A.; Chen, Q.; Yu, M.; Xiao, R.; Zhang, T.; Wang, Z.; Linhardt, R. J.; Sun, X.; Jin, L.; Chi, L. Coupling Liquid Chromatography and Tandem Mass Spectrometry to Electrophoresis for In-Depth Analysis of Glycosaminoglycan Drugs: Heparin and the Multicomponent Sulodexide. *Anal. Chem.*, **2021**, 93 (3), 1433–1442. <https://doi.org/10.1021/acs.analchem.0c03330>.
- [61] Linhardt, R.; Gunay, N. Production and Chemical Processing of Low Molecular Weight Heparins. *Semin Thromb Hemost*, **1999**, 25 Suppl 3, 5–16.
- [62] Iqbal, Z.; Sadaf, S. Commercial Low Molecular Weight Heparins — Patent Ecosystem and Technology Paradigm for Quality Characterization. *Journal of Pharmaceutical Innovation*, **2022**. <https://doi.org/10.1007/s12247-022-09665-7>.
- [63] Shi, D.; Sheng, A.; Bu, C.; An, Z.; Cui, X.; Sun, X.; Li, H.; Zhang, F.; Linhardt, R. J.; Zhang, T.; et al. A Cluster Sequencing Strategy To Determine the Consensus Affinity Domains in Heparin for Its Binding to Specific Proteins. *Anal. Chem.*, **2022**, 94 (40), 13987–13994. <https://doi.org/10.1021/acs.analchem.2c03267>.
- [64] Wang, Z.; Chi, L. Recent Advances in Mass Spectrometry Analysis of Low Molecular Weight Heparins. *Chinese Chemical Letters*, **2018**, 29 (1), 11–18. <https://doi.org/10.1016/j.ccllet.2017.08.050>.
- [65] Zappe, A.; Miller, R. L.; Struwe, W. B.; Pagel, K. State-of-the-Art Glycosaminoglycan Characterization. *Mass Spectrometry Reviews*, **2022**, 41 (6), 1040–1071. <https://doi.org/10.1002/mas.21737>.
- [66] Skidmore, M. A.; Guimond, S. E.; Dumax-Vorzet, A. F.; Atrih, A.; Yates, E. A.; Turnbull, J. E. High Sensitivity Separation and Detection of Heparan Sulfate Disaccharides. *Journal of Chromatography A*, **2006**, 1135 (1), 52–56. <https://doi.org/10.1016/j.chroma.2006.09.064>.
- [67] Miller, R. L.; Guimond, S. E.; Shivkumar, M.; Blocksidge, J.; Austin, J. A.; Leary, J. A.; Turnbull, J. E. Heparin Isomeric Oligosaccharide Separation Using Volatile Salt Strong Anion Exchange Chromatography. *Anal. Chem.*, **2016**, 88 (23), 11542–11550. <https://doi.org/10.1021/acs.analchem.6b02801>.
- [68] Patel, R. P.; Narkowicz, C.; Jacobson, G. A. Effective Reversed-Phase Ion Pair High-Performance Liquid Chromatography Method for the Separation and Characterization of Intact Low-Molecular-Weight Heparins. *Analytical Biochemistry*, **2009**, 387 (1), 113–121. <https://doi.org/10.1016/j.ab.2009.01.007>.
- [69] Galeotti, F.; Volpi, N. Novel Reverse-Phase Ion Pair-High Performance Liquid Chromatography Separation of Heparin, Heparan Sulfate and Low Molecular Weight-Heparins Disaccharides and

- Oligosaccharides. *Journal of Chromatography A*, **2013**, *1284*, 141–147. <https://doi.org/10.1016/j.chroma.2013.02.013>.
- [70] Doneanu, C. E.; Chen, W.; Gebler, J. C. Analysis of Oligosaccharides Derived from Heparin by Ion-Pair Reversed-Phase Chromatography/Mass Spectrometry. *Anal. Chem.*, **2009**, *81* (9), 3485–3499. <https://doi.org/10.1021/ac802770r>.
- [71] Alpert, A. J. Electrostatic Repulsion Hydrophilic Interaction Chromatography for Isocratic Separation of Charged Solutes and Selective Isolation of Phosphopeptides. *Anal. Chem.*, **2008**, *80* (1), 62–76. <https://doi.org/10.1021/ac070997p>.
- [72] Antia, I. U.; Mathew, K.; Yagnik, D. R.; Hills, F. A.; Shah, A. J. Analysis of Procainamide-Derivatized Heparan Sulphate Disaccharides in Biological Samples Using Hydrophilic Interaction Liquid Chromatography Mass Spectrometry. *Analytical and Bioanalytical Chemistry*, **2018**, *410* (1), 131–143. <https://doi.org/10.1007/s00216-017-0703-1>.
- [73] Gill, V. L.; Aich, U.; Rao, S.; Pohl, C.; Zaia, J. Disaccharide Analysis of Glycosaminoglycans Using Hydrophilic Interaction Chromatography and Mass Spectrometry. *Anal. Chem.*, **2013**, *85* (2), 1138–1145. <https://doi.org/10.1021/ac3030448>.
- [74] Li, L.; Zhang, F.; Zaia, J.; Linhardt, R. J. Top-Down Approach for the Direct Characterization of Low Molecular Weight Heparins Using LC-FT-MS. *Anal. Chem.*, **2012**, *84* (20), 8822–8829. <https://doi.org/10.1021/ac302232c>.
- [75] Poyer, S.; Seffouh, I.; Lopin-Bon, C.; Jacquinet, J.-C.; Neira, J. L.; Salpin, J.-Y.; Daniel, R. Discrimination of Sulfated Isomers of Chondroitin Sulfate Disaccharides by HILIC-MS. *Analytical and Bioanalytical Chemistry*, **2021**, *413* (28), 7107–7117. <https://doi.org/10.1007/s00216-021-03679-9>.
- [76] Galeotti, F.; Volpi, N. Oligosaccharide Mapping of Heparinase I-Treated Heparins by Hydrophilic Interaction Liquid Chromatography Separation and Online Fluorescence Detection and Electrospray Ionization-Mass Spectrometry Characterization. *Journal of Chromatography A*, **2016**, *1445*, 68–79. <https://doi.org/10.1016/j.chroma.2016.03.078>.
- [77] Alpert, A. J. Effect of Salts on Retention in Hydrophilic Interaction Chromatography. *Journal of Chromatography A*, **2018**, *1538*, 45–53. <https://doi.org/10.1016/j.chroma.2018.01.038>.
- [78] Naimy, H.; Leymarie, N.; Bowman, M. J.; Zaia, J. Characterization of Heparin Oligosaccharides Binding Specifically to Antithrombin III Using Mass Spectrometry. *Biochemistry*, **2008**, *47* (10), 3155–3161. <https://doi.org/10.1021/bi702043e>.
- [79] Volpi, N.; Maccari, F.; Suwan, J.; Linhardt, R. J. Electrophoresis for the Analysis of Heparin Purity and Quality. *ELECTROPHORESIS*, **2012**, *33* (11), 1531–1537. <https://doi.org/10.1002/elps.201100479>.
- [80] Gunay, N. S.; Linhardt, R. J. Capillary Electrophoretic Separation of Heparin Oligosaccharides under Conditions Amenable to Mass Spectrometric Detection. *Journal of Chromatography A*, **2003**, *1014* (1), 225–233. [https://doi.org/10.1016/S0021-9673\(03\)01288-3](https://doi.org/10.1016/S0021-9673(03)01288-3).
- [81] Patel, R. P.; Narkowicz, C.; Hutchinson, J. P.; Hilder, E. F.; Jacobson, G. A. A Simple Capillary Electrophoresis Method for the Rapid Separation and Determination of Intact Low Molecular Weight and Unfractionated Heparins. *Journal of Pharmaceutical and Biomedical Analysis*, **2008**, *46* (1), 30–35. <https://doi.org/10.1016/j.jpba.2007.10.009>.
- [82] Ruiz-Calero, V.; Puignou, L.; Galceran, M. T. Use of Reversed Polarity and a Pressure Gradient in the Analysis of Disaccharide Composition of Heparin by Capillary Electrophoresis. *Journal of Chromatography A*, **1998**, *828* (1), 497–508. [https://doi.org/10.1016/S0021-9673\(98\)00662-1](https://doi.org/10.1016/S0021-9673(98)00662-1).
- [83] Ucakturk, E.; Cai, C.; Li, L.; Li, G.; Zhang, F.; Linhardt, R. J. Capillary Electrophoresis for Total Glycosaminoglycan Analysis. *Analytical and Bioanalytical Chemistry*, **2014**, *406* (19), 4617–4626. <https://doi.org/10.1007/s00216-014-7859-8>.
- [84] Volpi, N.; Maccari, F. Electrophoretic Approaches to the Analysis of Complex Polysaccharides. *Journal of Chromatography B*, **2006**, *834* (1), 1–13. <https://doi.org/10.1016/j.jchromb.2006.02.049>.

- [85] Zamfir, A.; Peter-Katalinić, J. Capillary Electrophoresis-Mass Spectrometry for Glycoscreening in Biomedical Research. *ELECTROPHORESIS*, **2004**, *25* (13), 1949–1963. <https://doi.org/10.1002/elps.200405825>.
- [86] Damm, J. B. L.; Overklift, G. T.; Vermeulen, B. W. M.; Fluitsma, C. F.; Dedem, G. W. K. van. Separation of Natural and Synthetic Heparin Fragments by High-Performance Capillary Electrophoresis. *Journal of Chromatography A*, **1992**, *608* (1), 297–309. [https://doi.org/10.1016/0021-9673\(92\)87136-V](https://doi.org/10.1016/0021-9673(92)87136-V).
- [87] Cui, Z. Protein Separation Using Ultrafiltration — an Example of Multi-Scale Complex Systems. *China Particuology*, **2005**, *3* (6), 343–348. [https://doi.org/10.1016/S1672-2515\(07\)60213-9](https://doi.org/10.1016/S1672-2515(07)60213-9).
- [88] Yu, Y.; Zhang, F.; Renois-Predelus, G.; Amster, I. J.; Linhardt, R. J. Filter-Entrapment Enrichment Pull-down Assay for Glycosaminoglycan Structural Characterization and Protein Interaction. *Carbohydrate Polymers*, **2020**, *245*, 116623. <https://doi.org/10.1016/j.carbpol.2020.116623>.
- [89] Niu, C.; Zhao, Y.; Bobst, C. E.; Savinov, S. N.; Kaltashov, I. A. Identification of Protein Recognition Elements within Heparin Chains Using Enzymatic Foot-Printing in Solution and Online SEC/MS. *Anal. Chem.*, **2020**, *92* (11), 7565–7573. <https://doi.org/10.1021/acs.analchem.0c00115>.
- [90] Zhao, Y.; Abzalimov, R. R.; Kaltashov, I. A. Interactions of Intact Unfractionated Heparin with Its Client Proteins Can Be Probed Directly Using Native Electrospray Ionization Mass Spectrometry. *Anal. Chem.*, **2016**, *88* (3), 1711–1718. <https://doi.org/10.1021/acs.analchem.5b03792>.
- [91] Militopoulou, M.; Lamari, F.; Karamanos, N. K. Capillary Electrophoresis: A Tool for Studying Interactions of Glycans/Proteoglycans with Growth Factors. *Journal of Pharmaceutical and Biomedical Analysis*, **2003**, *32* (4), 823–828. [https://doi.org/10.1016/S0731-7085\(03\)00185-7](https://doi.org/10.1016/S0731-7085(03)00185-7).
- [92] Keiser, N.; Venkataraman, G.; Shriver, Z.; Sasisekharan, R. Direct Isolation and Sequencing of Specific Protein-Binding Glycosaminoglycans. *Nature Medicine*, **2001**, *7* (1), 123–128. <https://doi.org/10.1038/83263>.
- [93] Przybylski, C.; Gonnet, F.; Saesen, E.; Lortat-Jacob, H.; Daniel, R. Surface Plasmon Resonance Imaging Coupled to On-Chip Mass Spectrometry: A New Tool to Probe Protein-GAG Interactions. *Analytical and Bioanalytical Chemistry*, **2020**, *412* (2), 507–519. <https://doi.org/10.1007/s00216-019-02267-2>.
- [94] Bellon, S.; Buchmann, W.; Gonnet, F.; Jarroux, N.; Anger-Leroy, M.; Guillonneau, F.; Daniel, R. Hyphenation of Surface Plasmon Resonance Imaging to Matrix-Assisted Laser Desorption Ionization Mass Spectrometry by On-Chip Mass Spectrometry and Tandem Mass Spectrometry Analysis. *Anal. Chem.*, **2009**, *81* (18), 7695–7702. <https://doi.org/10.1021/ac901140m>.
- [95] Remy-Martin, F.; El Osta, M.; Lucchi, G.; Zeggari, R.; Lebouis, T.; Bellon, S.; Ducoroy, P.; Boireau, W. Surface Plasmon Resonance Imaging in Arrays Coupled with Mass Spectrometry (SUPRA-MS): Proof of Concept of on-Chip Characterization of a Potential Breast Cancer Marker in Human Plasma. *Analytical and Bioanalytical Chemistry*, **2012**, *404* (2), 423–432. <https://doi.org/10.1007/s00216-012-6130-4>.
- [96] Guerrini, M.; Guglieri, S.; Casu, B.; Torri, G.; Mourier, P.; Boudier, C.; Viskov, C. Antithrombin-Binding Octasaccharides and Role of Extensions of the Active Pentasaccharide Sequence in the Specificity and Strength of Interaction: EVIDENCE FOR VERY HIGH AFFINITY INDUCED BY AN UNUSUAL GLUCURONIC ACID RESIDUE\*. *Journal of Biological Chemistry*, **2008**, *283* (39), 26662–26675. <https://doi.org/10.1074/jbc.M801102200>.
- [97] Zong, C.; Huang, R.; Condac, E.; Chiu, Y.; Xiao, W.; Li, X.; Lu, W.; Ishihara, M.; Wang, S.; Ramiah, A.; et al. Integrated Approach to Identify Heparan Sulfate Ligand Requirements of Robo1. *J. Am. Chem. Soc.*, **2016**, *138* (39), 13059–13067. <https://doi.org/10.1021/jacs.6b08161>.
- [98] Watanabe, T.; Takeda, K.; Hiemori, K.; Minamisawa, T.; Tateno, H. A Glycosaminoglycan Microarray Identifies the Binding of SARS-CoV-2 Spike Protein to Chondroitin Sulfate E. *FEBS Letters*, **2021**, *595* (18), 2341–2349. <https://doi.org/10.1002/1873-3468.14173>.