IDENTIFICATION OF GLYCOSAMINOGLYCANS WITH MASS SPECTROMETRY-BASED ANALYTICAL TECHNIQUES

by

JANDI KIM

(Under the Direction of I. Jonathan Amster)

ABSTRACT

Glycosaminoglycans (GAGs) are linear anionic polysaccharides that play important roles in modulating biological functions in most mammalian organisms. The characterization of GAGs has been the subject of numerous analytical techniques over the years, allowing researchers to gain a better understanding of the functions of GAG molecules. Nonetheless, highly sulfated GAGs pose an analytical challenge because of a high level of heterogeneity inherent from their non-template biosynthesis, and also because of the chemical lability of the sulfation modifications that are the target of these analyses. In this study, MS-based techniques, tandem mass spectrometry (MS/MS) and capillary zone electrophoresis-mass spectrometry (CZE-MS), have allowed for GAG analysis in a fast, accurate, and reliable manner. The stereochemistry of the C5 carbon of GAGs was distinguished by different types of MS/MS methods and multivariate statistical analysis. Workflows have been evaluated to examine GAG biomarkers with multivariate statistical analysis. In addition, applications of our methods have been

explored, such as the high-throughput analysis of clinical GAG samples (e.g., complex and/or low levels of GAGs), synthetic GAG oligosaccharides, and CRISPR/Cas9-mediated GAG biosynthetic products.

INDEX WORDS: Glycosaminoglycans, Fourier transform ion cyclotron resonance mass spectrometry, Capillary zone electrophoresis-mass spectrometry, Tandem mass spectrometry, Multivariate statistical analysis

IDENTIFICATION OF GLYCOSAMINOGLYCANS WITH MASS SPECTROMETRY-BASED ANALYTICAL TECHNIQUES

by

JANDI KIM

B.S., Changwon National University, Republic of Korea, 2012

M.S., Changwon National University, Republic of Korea, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2023

IDENTIFICATION OF GLYCOSAMINOGLYCANS

WITH MASS SPECTROMETRY-BASED ANALYTICAL TECHNIQUES

by

JANDI KIM

Major Professor:	I. Jonathan Amster
Committee:	Ron Orlando
	Kelly Hines

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2023

DEDICATION

I dedicate this dissertation to God and to my family who has been praying for me in God.

ACKNOWLEDGEMENTS

I am deeply appreciative of the people, experience, and opportunities that have shaped this dissertation and my graduate school journey as a whole.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REIVEW

1.1 Glycosaminoglycans

Glycosaminoglycans (GAGs) are linear anionic polysaccharides that typically occur as protein conjugates. They regulate a variety of physio/pathological processes in mammalian systems by through specific interactions with target proteins.[1-3] GAGs are often found on the surface of cells where they can mediate a variety of biological processes via their specific interactions with proteins on the cell surface, on the surface of other cells, or in the extracellular matrix.[1, 4-6] They can also act as receptors for viruses and toxins, thus demonstrating their pathological role.[7-9]

GAGs are composed of a repeating disaccharide. These are pyranose rings that are covalently connected by glycosidic linkages between anomeric carbons to form GAGs as a long linear chain structure.[5] The repeating disaccharide unit and the glycosidic linkage between units defines a GAG family, of which there are four main types, shown in Figure 1: Heparin/Heparan sulfate (Hp/HS), chondroitin sulfate/dermatan sulfate (CS/DS), keratan sulfate (KS), and hyaluronan (HA).[3] With the exception of KS, the repeating disaccharide is an amino sugar and an acidic sugar.

Hp/HS has been the most widely studied GAG, because of its diverse biological activity. The Hp/HS disaccharide unit is glucuronic acid (GlcA)β1-4/iduronic acid (IdoA)α1-4GlcNAcα1-4. Polymerization of this disaccharide results in GAG chains of variable length. The 3-*O* or/and 6-*O* positions on the hexosamine

residue or the 2-*O* position of the hexuronic acid can be enzymatically modified by sulfation. In addition, the amino sugar can be enzymatically deacetylated, and the amine can be sulfated. In general, the difference between Hp and HS is qualtitative, and refers to the extent of sulfation per disaccharide unit (Hp > HS).[5, 10] Hp/HS are generally biosynthesized as protein conjugates, connected via serine residues via a common tetrasaccharide, -GlcAβ1–3Galβ1–3Galβ1–4Xylβ1-*O*-(Ser).

CS possesses GlcAβ1-3GalNAcβ1-4 as a disaccharide repeating unit, and is conjugated to serine residues in proteins via the same tetrasaccharide linker as Hp/HS. Enzymatic epimerization of the acidic sugar produces DS, with a repeat unit of IdoAα1-3GalNAcβ1-4. CS comprises more than 80 percent of the GAGs in articular cartilage, so it has received attention as a potential treatment for osteoarthritis.[11] CS can be subclassified by the major constituent of the repeating unit/sulfation: CSA is sulfated at 4-*O* position on the hexosamine residue, CSC is sulfated at 6-*O* position on the hexosamine residue, CSD refers to as sulfation at 6-*O* position on the hexosamine residue as well as 2-*O* position on the hexuronic acid, CSE refers to as sulfation at both 4-*O* and 6-*O* position on the hexosamine residue. CSB is specifically known as dermatan sulfate which has the same chemical composition as CSA while the GlcA is epimerized to IdoA, playing an important role in wound healing.[5, 6, 10]

KS is different from all other GAGs in that it does not contain an acidic residue in its repeating disaccharide. Instead, it is composed of a Galβ1-4GlcNAcβ1-3 disaccharide unit and can be sulfated at 6-*O* position on the Gal hexose or GlcNAc hexosamine residues. KS is bound to a protein core via different linkers than CS/DS or Hp/HS, and is divided into three subcategories based on its core structures. The linkers are branched N-linked or O-linked glycans. KS keeps the corneal water at a

constant level as part of its function and small leucine-rich PGs are mostly bound to KS in corneas, regulating corneal development and function.[12, 13]

Although a type of GAG, HA is neither sulfated nor covalently attached to a core protein, and is composed of the repeating unit, GlcAβ1-4GlcNAcβ1-3. A typical molecule of HA has a molecular weight of 100 to 4000 kDa, but can be further degraded by enzymes into even smaller polysaccharides in the extracellular environment.[14] Since HA can absorb up to 1000-fold its weight in water, it maintains and cushions tissue cells, as well as other mechanical and biological properties.[15, 16]

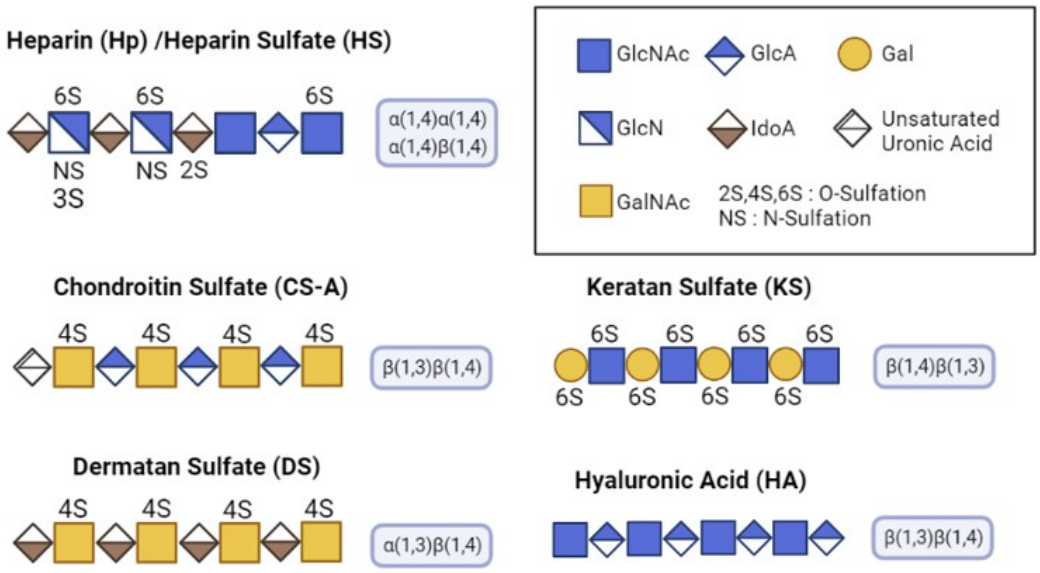

Figure 1. Symbolic structures of four subclasses of GAGs.

Unlike the biosynthesis of DNA and protein that utilize a template to maintain structural fidelity, GAGs are synthesized by a non-template process.[13] GAG biosynthesis takes place in the endoplasmic reticulum and Golgi apparatus.[17] The repeating disaccharides of CS/DS and Hp/HS GAGs are extended from a tetrasaccharide linkage region that occur on a serine residue of the core protein. The

stepwise activity of numerous enzymes causes polymerization, *O*-sulfation, deacetylation/*N*-sulfation, and epimerization spatiotemporally.[12, 18] A summary of the two major types of GAGs, HS and CS, based on their modification and epimerization is provided in Table 1.

Alteration	HS	CS
Hexosamine	3-*O*-Sulfation, 6-*O*-Sulfation, N-Deacetylation/Sulfation	4-*O*-Sulfation, 6-*O*-Sulfation
Hexuronic Acid	2-*O*-Sulfation Epimerization (GlcA→IdoA)	2-*O*-Sulfation, 3-*O*-Sulfation Epimerization (GlcA→IdoA)

Table 1. Types of modification and epimerization in HS and CS.

In addition to being highly complex and heterogeneous, GAGs are also highly information-dense biopolymers. For instance, the most distinguished unique pentasaccharide sequence (GlcNAc/NS**6S**-GlcUA-GlcNS**3S**6S-IdoUA2S-GlcNS6S, critical sulfation for GAG-protein interactions in bold) has shown in HS induces anticoagulant activity due to the interaction between HS and antithrombin.[10] Over the past few decades, many studies have revealed that biosynthesis of GAGs and their unique patterns are dynamically regulated by cellular enzyme and environmental factors.[18-20] In addition, the structural alterations in GAGs and GAG-protein binding have been the focus of numerous studies aimed at finding potential biomarkers for

human diseases, with the potential applications of diagnosis and treatment.[6, 21] To date, considerable effort has been made to analyze GAG compositions and sequences using a variety of analytical approaches.

1.2 Mass Spectrometry for GAG Analysis

GAGs are present in small quantities in biological specimen in general, and cannot also be amplified like DNA or overexpressed like proteins. Therefore, a principal challenge for GAG analysis is their low abundance and heterogenous compositions and structures.[13] The focus of analytical research for GAGs is to develop sensitive and specific analytical techniques. Among them, mass spectrometry (MS) stands out because of its high sensitivity, accuracy, and compatibility to orthogonal analytical methods. [13, 22, 23] MS has a significant advantage in sensitivity for analyzing GAGs compared to nuclear magnetic resonance spectroscopy (NMR). The latter is also a useful tool for analyzing GAGs that have been isolated in significant quantities and that have been highly purified. NMR spectra of GAGs contain many overlapping signals which need to be deconvoluted, and it is much more challenging to get complete structural characterization on each individual GAG in a mixture.[24] In contrast, MS techniques are sensitive enough to detect pmol quantities of GAGs and can analyze complex mixtures with the help of off-line separations or on-line separations including various chromatographic methods and capillary electrophoresis, among others.[13, 24] Additionally, a partial/entire digestion process and reliable extraction process are essential to characterizing GAGs accurately and sensitively.[23]

The most basic application of MS is to determine the molecular weight of a GAG or all GAGs present in a mixture. Using the molecular weight of an intact GAG chain provides information about its composition, but not its structure.[13] There are informative components in a GAG, including its degree of polymerization (dp), number of SO_3 modifications, number of NAc present, and any terminal modifications such as the non-reducing end (NRE) unsaturation or the reducing end (RE) derivatization. This can be represented in the [D, U, A, N, S] notation, where D represents the number of Δ-hexuronic acids (0 or 1), U represents the number of hexuronic acids, A represents the number of amino sugars, N represents the number of N-Acetyl groups (must be equal to or less than A), and S represents of sulfo groups. The composition analysis of a GAG mixture can provide insights into chain lengths and sulfate occupancy levels. However, the heterogeneity associated with sodium/hydrogen exchange is one challenge in assigning composition based on accurate mass measurements. Molecules containing multiple acidic sites, such as GAGs or nucleic acids, are susceptible to replacement of ionizable protons with alkali cations. Thus, a wide range of molecular compositions can be produced differing by sequential Na/H exchange.[25] A broad distribution of alkali exchange states for the molecular ion reduces the intensity of the peaks and complicates compositional assignment. Besides, chromatographic analysis for GAG mixtures may be more challenging with high alkali exchange degrees. This problem can be remediated by desalting the sample and adding diluted formic acid or diethylamine to the electrospray solvent to reduce the heterogeneity of the ions.[13, 26]

In order to gain a deeper understanding of how GAGs function, not only the composition but also the locations of modifications are required. This information can be provided by tandem mass spectrometry, also known as MS2 or MS/MS.[13]

Moreover, the MS-based analysis has the limitation that high abundance ions may conceal low abundance ions, since the number of ions that can be accumulated in the analyzer is limited. A mixture sample requires additional separation techniques in order to prevent co-elution and differentiate isomers. [13] Therefore, MS combined with a separation method provides a general framework for characterizing GAGs and determining the next step in their analysis.

1.3 Tandem Mass Spectrometry for GAGs

Tandem mass spectrometry (MS/MS or MS2) is a powerful tool for the structural elucidation of GAGs. Fragment ions from a mass selected precursor ion can provide information about the structure of that GAG. Fragment types are divided into two main categories, glycosidic bond cleavage and cross-ring bond cleavage. Early work in the field by Domon and Costello [27] established a standard nomenclature for these fragments that intuitively resemble protein MS/MS nomenclature. Glycosidic fragments (B, C, Y, Z) arise from breakages to either side of the oxygen that forms the glycosidic bond. B and C fragments contain the non-reducing end (NRE) of the GAG, while Y and Z fragments contain the reducing end (RE) of the GAG. Glycosidic cleavage is also accompanied by a hydrogen migration that results in complementary glycosidic pairs (B/C, or Y/Z) differing in composition by H_2O and in mass by 18 Da. This also forms a double bond on the residue that loses the hydrogen. For Z fragments on hexuronic acids, the double bond mimics the structure and mass of Δ-hexuronic acids. So, if the sample had Δ-hexuronic acids at the NRE, the Z and C ions can be hard to distinguish due to their identical composition.

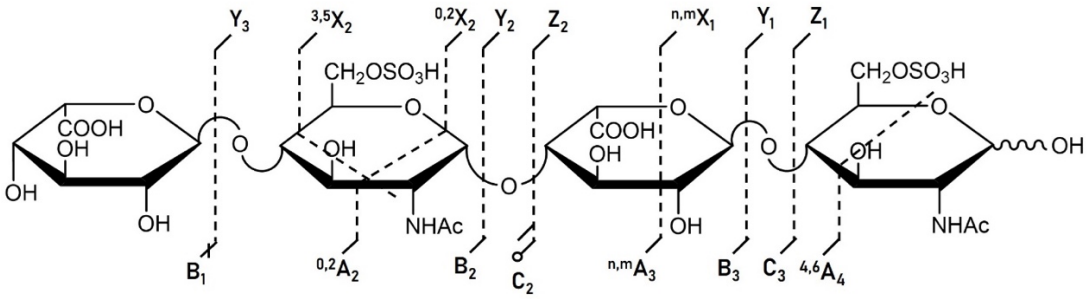

Figure 2. GAG cleavage types and Domon nomenclature. A circle at the end of each cleavage means the sulfate group decomposition. (○: single sulfo group loss, ●: multiple sulfo group losses)

MS/MS experiments have the potential for the complete structural characterization of a GAG in terms of the locations of sulfation, (both residue and exact linkage site), locations of NAc/NS modifications, and the C5 stereochemistry of hexuronic acid residues, all of which are important to the molecule's biological activity.

Using glycosidic fragments alone, the number and type of modifications to each sugar can be determined. SO_3 modifications produce a predictable mass shift of 79.956 Da and NAc groups have a mass shift of 42.047 Da. So, if a series of B and Y ions are produced between every residue, the number and type of modification can be determined per residue, but the sites of SO_3 occupancy cannot. There is some evidence to support that SO_3 linkage (4S/6S) can influence the relative abundance of B, Y and Z fragments, but the scope of this approach is limited.[28, 29]

Cross-ring fragments are more informative than glycosidic fragments in locating sites of modification but require more energy to produce. Cross-ring fragments are generated when two bonds within an individual monosaccharide ring break and are denoted $^{n,m}A$ and $^{n,m}X$ for NRE and RE fragments, respectively. The n

and m correspond to the bond within a ring that was broken, and can have values from 0 to 5. If sufficient cross-ring fragments are generated, the site of modification can be localized. As an example, a 1,3 linked GlcNAc with 6-*O* sulfation would require at least two cross-ring fragments on the GlcNAc, one on either side of C5 (^{2,5}X/A, ^{1,4}X/A) to confidently assign the modification.

In addition to *N*- and *O*-sulfation, Hp/HS and CS/DS can be modified by epimerization of C5 in an uronic acid residue. This subtle modification produces no change in mass, and thus is a particular challenge to characterization by mass spectrometry. However, a combination of glycosidic and cross-ring fragments has been used to infer C5 stereochemistry for hexuronic acid residues. IdoA and GlcA are structural isomers, and as such their cross-ring fragments are all isomers. Yet, the stereochemistry at the C5 position influences the shape of the ring and the conformation of the GAG chain and can promote differences in fragmentation pathways. This can influence the intensities of diagnostic fragments (both glycosidic and cross-ring) that can be used to distinguish between the isomers. [28, 30-33]

The number and type of each fragment generated by MS/MS is dependent on a number of factors, but the single highest influence comes from the fragmentation method that is chosen. As discussed in more detail below, collision-based activation and electron-based activation are the primary means of fragmenting ions. Another MS/MS method that has attracted attention is photodissociation, which will be briefly discussed in this chapter. Herein, most common techniques and their variations as they pertain to the analysis of GAGs are described in this chapter, providing a comparison of each technique in terms of the fragmentation information that can be obtained, the availability of each technique, as well as their utility in coupling to chromatography or capillary electrophoresis techniques.

MS/MS method	Examples	Property
Collision-based activation	Collision-induced dissociation (CID) Higher-energy collision dissociation (HCD)	Glycosidic linkage cleavages are prevalent SO_3 loss Neutral loss
Electron-based activation	Electron detachment dissociation (EDD) Negative electron transfer dissociation (NETD) Electron-induced dissociation (EID) Charge transfer dissociation (CTD)	Glycosidic and cross-ring linkage cleavages are prevalent SO_3 retained High sequence coverage
Photodissociation	Ultraviolet photodissociation (UVPD) Infrared multiphoton dissociation (IRMPD)	Glycosidic and cross-ring linkage cleavages SO_3 retained High sequence coverage

Table 2. The most common activation methods used to characterize GAGs.

Collisional-induced dissociation (CID) is the most general fragmentation technique across all MS platforms and was naturally one of the first to be applied to GAGs.[34, 35] CID is performed by accelerating analyte ions through a region of the instrument that contains a high pressure of an inert gas, usually Ar, N_2 or He. Collisions between the ions and atoms of neutral gas increase the internal energy of the ion. The excess energy drives unimolecular dissociation processes that lead to the generation of fragment ions and neutral losses.[36] CID, as implemented in most commercial mass spectrometers, is a threshold activation technique, meaning that it will induce fragmentation through the lowest energy pathway available. In other words, the most labile bonds on the molecule will break first. With regards to GAGs, CID will produce an abundance of glycosidic cleavages, mainly B and Y ions that contain the non-reducing end, and Y and Z ions that contain the non-reducing end. Sulfate half-esters are also labile. In their protonated form, they can undergo a proton rearrangement and easily leave as a neutral (SO_3) under CID.[33] To minimize these losses, careful precursor selection must be performed prior to fragmentation.[37] The optimal GAG precursor for a CID experiment has one charge per sulfate group, thus increasing the odds that all sulfates are deprotonated and less susceptible to fragmentation. For relatively short GAGs (DP < 10) with a low degree of sulfation (one SO_3/disaccharide) a fully ionized precursor can be selected. This makes CS a more approachable target for MS/MS characterizations than Hp/HS.[33]

Higher-energy collision dissociation (HCD) is a similar activation method to CID that is available on Thermo-fisher linear ion trap (LIT) instruments. Despite its name, HCD is actually a low energy form of CID (ion translation energies of 100 eV or less.) As with CID, HCD also produces primarily glycosidic cleavages and sulfate loss with a small degree of cross ring fragmentation. [13, 38] With regard to the

determination of linkage position in GAGs, early HCD work showed that the NRE residues for cationized HS precursors generated more cross-ring fragments than was usually observed using CID. This allowed for SO_3 linkage sites to be determined on NRE and NRE adjacent residues.[33] Later studies even showed that leaving a single ionizable site on the GAG protonated can promote cross ring fragmentation that can be used to further distinguish GlcA and IdoA residues from each other in CS and DS.[32] This strategy have been also further developed and applied for the analysis of various GAGs using even different types of MS/MS later. [39, 40]

Electronic activation (ExD) is an umbrella term meant to encompass the many varieties of electron-based dissociation techniques that have been developed.[13] While ExD techniques all utilize electrons, they differ by how the electron is introduced (either by an electron-emitting cathode or by ion-ion reactions), the energy of the incident electrons, and the irradiation or reaction time. These parameters subsequently control the dominant fragmentation mechanism. Unlike CID, which is a threshold activation method, ExD dissociation occurs through radical-driven reaction mechanisms or by direct electronic excitation.[41, 42] Because of this, ExD methods tend to leave more sulfates retained and produce a much higher proportion of cross-ring fragments when compared to CID and HCD. ExD provides a highly valuable set of techniques to determine sulfo-modification locations on highly sulfated GAGs and has significantly developed the field of the structural characterization for other biomolecules.

Electron capture dissociation (ECD) was the first ExD method to be developed. In an ECD experiment, an electron gun is used to irradiate multiply charged cations with electrons. Some cations capture an electron to form a charge reduced radical, which will then undergo a radical-driven reaction to generate

fragment ions.[43] For proteins and peptides, ECD was able to dramatically increase the number of backbone cleavages that could be obtained.[44]

Unfortunately, ECD is not useful for GAG analysis. Due to their acidity, GAGs are more stable as negative ions, and more efficiently ionized using negative electrospray ionization. This makes electron capture particularly difficult, although Hakansson and co-workers have been able to perform ECD on negative ions, dubbed negative ion ECD or niECD.[45, 46] Electron detachment dissociation (EDD) was later developed to analyze acidic molecule such as sulfated acidic polypeptides, oligosaccharides, and nucleic acids in negative ion mode and is currently only commercially available in Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS).[41, 47, 48] It can be thought of as the negative ion complement to ECD. Electrons of a moderate kinetic energy (15-20 eV) from the heated hollow cathode are used to irradiate multiply charged precursor anions stored in the analyzer cell. When an electron passes sufficiently close to an anion, it can detach an electron and form a radical anion. The carboxylate group on hexuronic acid residues is thought to be the primary site of electron detachment, as the energy required for electron detachment from a carboxylate is significantly lower than that required from a sulfate anion. The radical anion then fragments through radical rearrangement pathways.[41, 49] As a result, radical species of fragments are produced, and labile functional groups are ideally retained. Radical species products, also known as odd-electron products, can be considered as a yardstick whether EDD has occurred or not. In addition, some products are formed by direct electronic excitation of the precursor by its collision with an energetic electron. Products can be formed via even-electron dissociation of the electronically-excited precursor. As previously mentioned, EDD was exclusively available on FT-ICR MS instruments although recent studies

have shown the implementation of ECD cell on both an Orbitrap and a QTOF system, although this setup has not been investigated for negative precursor ions yet. [50]

The Amster group has examined EDD for the dissociation of GAGs for about fifteen years. They have reported key experimental parameter values for EDD with a variety of GAG compounds.[28, 40, 41, 49-51] While irradiating GAG precursors with electrons, pulse length, extraction lens voltage, and cathode bias were be optimized to improve a fragmentation yield. As the pulse length and lens voltage determine how many electrons are extracted from the cathode, the bias is associated with the kinetic energy of the electrons that are accelerated; the typically value is one second, the extraction voltage is 17 to -20 volts, and the bias is -19 volts. Additionally, the typical cathode heater current is 1.5-1.6 A, as the maximum to control the number of electrons produced by the cathode.[50] The difference between EDD and ECD can be found clearly by knowing the details of the EDD mechanism and roles of each parameter. Prior work monitored the electron flux through the ICR cell while changing other important parameters, leading to the understanding that a heater current above the threshold for electron production does not affect the product ion yield. A defined electron current is obtained by adjustment of the extraction lens voltage.[51] EDD was first used to determine the structural features of HS tetrasaccharide standards (fig. 8) by Wolff et al., then extended to longer, more highly sulfated GAGs.[28, 31, 40, 41, 52]

After Syka et al. had pioneered electron transfer dissociation (ETD) based on the discovery of ion-ion reaction by McLuckey and coworkers[53, 54], negative electron transfer (NETD) was developed for the characterization of peptide anions in negative ion mode by Coon et al[55]. NETD achieves electron transfer to the precursor anion via low-energy ion-ion reactions, while minimizing sulfate decomposition. Fluoranthene

or xenon can serve as radical cation reagents, while polyanions of the analytes serve as electron donors.[50] During the reaction, the NETD reagent radical cation reacts with a polyanion to snatch an electron and produce an odd electron charge-reduced species similar to EDD. Then, the charge-reduced ion fragments through internal rearrangements to make odd-electron and even-electron product ions with a similar abundance of cross ring fragments to EDD. [42]. NETD, like ECD requires highly charged precursor ions in order to obtain charge-reduced product ions and extensive fragmentation. Also, the quality of fragmentation that can be obtained with NETD can be affected by precursor cation/proton exchange. Like CID, replacing protons by alkali cations (Na^+) stabilizes sulfo modifications in the precursor and reduces the amount of SO_3 that is lost during fragmentation.[52]

NETD has several merits for the analysis of GAGs. While EDD has a very small cross-section for the incident electron to interact with the negatively charged precursor ion, NETD relies on a reaction of oppositely-charged species, with an attractive potential, and thus has a much larger reaction cross-section. This allows faster ion activation, shorter experimental sequences, and more efficient fragmentation. ETD also produces an abundance of cross-ring fragments.[50] NETD was initially applied to the structural characterization of GAGs using a quadrupole ion trap platform, however, mass accuracy and resolving power limited the study of GAGs to those with short chain lengths.[42] Since that initial study, NETD has been implemented on other types of mass spectrometers, including Orbitrap and FTICR. This serves as a clear advantage of NETD over EDD, which is restricted to less accessible FTICR instruments.[42, 52, 56, 57] Recent studies have extended the use of NETD to the sequencing of GAGs with varying levels of sulfation.[50, 52, 58-60] The other benefit is that the ion-ion reaction requires only milliseconds (compared to seconds

for EDD), so NETD can be coupled to on-line separation technology such as high performance liquid chromatography (HPLC) or capillary zone electrophoresis (CZE).[57, 58, 60]

As an example, Stickney et al. showed the coupling of CZE with NETD to facilitate separation of complex GAG mixtures including the pharmaceutical product, Enoxaparin.[57] CZE can separate GAG molecules with overlapping m/z from a complex mixture based on their charge and shape, allowing the assignment of the degree of polymerization, sulfate modification, and isomers.[26] Huang et al. demonstrated that high-resolution tandem mass spectra for Hp/HS oligosaccharides was achieved by EDD and NETD, showing the possibility of on-line LC-MS/MS analysis of GAGs. Using the highly informative EDD and NETD fragment ions, they developed a *de novo* sequencing algorithm for analyzing oligosaccharide structures using both a molecular weight and MS/MS data, demonstrating a great potential for GAG characterization with high confidence in the assigned structures.[61]

The Zaia group has employed hydrophilic interaction liquid chromatography (HILIC) with NETD to profile 3-*O* and 6-*O* sulfation on HS. LC-MS has been simplified to enable GAG characterization, omitting permethylation. The use of an online ion suppressor after the HILIC column reduced metal cation adduction for molecular ions while increasing their charge states, improving their fragmentation, and simplifying their assignment.[56, 60] More recently, NETD has been applied to gated-trapped ion mobility tandem mass spectrometry (gated-TIMS MS/MS) to sequence highly sulfated Hp/HS isomers, allowing for the first-time relative quantification of Hp/HS.[58] It has been shown that NETD is an efficient method for characterizing isomers, including heparan sulfate and chondroitin sulfate species, by

using the abundance measurement for cross-ring linkage cleavages, but a more refined statistical method needs to be developed.[42, 52, 59]

MS-based analysis for GAGs–Distinguishing GAG epmiers

The stereochemistry at the C5 position of hexuronic acid residues distinguishes IdoA from GlcA, and identifies the class of GAG, such as CS versus DS. MS/MS has distinguished stereoisomers because stereochemistry at the C5 position influences the shape of the ring as well as the conformation of the chain, thus bonds strain differentially, which leads to different fragmentation mechanisms. Moreover, the carboxyl group can serve as a radical site in EDD or NETD experiments, and its position above or below the ring can lead to differences in radical-site driven hydrogen rearrangement reactions that can distinguish these epimeric sugars. In MS/MS, precursor isomers demonstrate different patterns of glycosidic and cross-ring fragmentation, demonstrating that MS/MS can distinguish these two isomers.[28, 31-33]

Zaia et al. have compared CID daughter ion abundances to delineate not only hexuronic acid epimerization but also sulfate modification isomers for CS oligosaccharides (e.g. sulfo modification at the 4-*O* position versus 6-*O* position on hexosamine residues).[33] In addition, multivariate statistical approaches, including principal component analysis (PCA), have been used to differentiate stereoisomeric GAGs. Oh et al. utilized EDD and IRMPD spectra and chemometrics to distinguish the C5 stereochemistry of hexuronic acid residues in four different HS epimers, obtaining a set of stereo-specific ions.[40] As another EDD example, Leach et al. were able to discriminate the CS from the DS tetrasaccharides based on the formation of indicative fragment ions. It is noted that they examined the stereochemistry of C5 by

looking at the relative intensities of charge-conserving products caused by electron excitation rather than radical formation induced by detachment.[31]

The identification of biomarker GAGs

A number of physiological and pathological processes are modulated by the fine-structural details of GAGs, which control their interaction with their ligands.[62] In this regard, GAG and GAG-derived oligosaccharides have attracted the interest of researchers in the biomedical and pharmaceutical field. For example, mucopolysaccharidosis is an accumulation disorder of glycosaminoglycans, leading to severe neurological diseases. A more sensitive and accurate method of quantifying GAGs is necessary for diagnosis and treatment of this disease. Tanaka et al. have previously employed LC-MS/MS to quantify HS and DS in the central nervous system and brain of mice, demonstrating that HS levels in the central nervous system are correlated with the prognosis and efficacy of treatment in MPS patients.[63] More recently, Forni et al. developed a simple and reproducible workflow for the simultaneous quantification of DS and HS with urinary samples by LC-MS/MS.[64] Discovering cancer-associated GAGs has been pursued because the specific structures of GAGs or the biological process of GAGs are highly involved in expression of tumor cells through their growth and metastasis.[65] These trends have led to the development of detection methods and data analysis for detecting and profiling cancer-related GAG biomarkers. Detection of GAGs without invasive procedures has also become more available as a method of maximizing the benefits of this method over conventional biopsies.[66, 67]

GAGs are known to be associated with virus invasion mechanisms. Since the COVID-19 (SARS-CoV-2) outbreak in 2019, the role of GAGs and PGs in the virus

infectivity has been extensively studied, from the point of view of binding to the cellular receptor, and their sulfate pattern that may affect the interaction.[7, 8, 68] Although the exact mechanisms of invasion and the structural specificities of COVID-19 are still contested, we need to explore strategies to treat COVID-19 and other viruses that may arise in the future. At this center, GAG analysis is an essential and cannot be overlooked.

Lastly, further pilot studies and systematic clinical trials with variation among individuals are needed in order to more deeply understand the role of GAGs as valuable biomarkers for many diseases.[69, 70] GAGs have been extensively investigated, but their heterogeneous structures and binding affinity/specificity with protein ligands still require the use of state-of-the-art analytical methods, which will be discussed in detail in the following chapter.

1.4 Capillary Zone Electrophoresis-Mass Spectrometry (CZE-MS) for Complex Mixtures of GAGs

On-line separations combined with MS have been widely used to characterize GAGs from biologically active samples after chemical synthesis and enzymatic depolymerization. Glycosaminoglycans naturally have heterogeneous oligosaccharide forms and even anomeric compositions. Thus, on-line separation techniques are essential for the proper analysis of complex GAG mixtures.[13, 58, 71] Several of these separation techniques are known as chromatography, including reversed phase-ion paring (RP-IP), hydrophilic interaction liquid chromatography (HILIC), strong anion exchange (SAX), and porous graphitized carbon liquid chromatography (PGC-LC) analysis, each of which presents its own challenges. It may be difficult to couple RP-

IP directly with MS because of the clustering of ions with anionic analytes resulting from the use of alkylammonium salts in mobile phase modifiers of RP-IP.[72] In addition, most RPLCs have been used to analyze disaccharide GAGs or small oligosaccharides because of their poor resolution when analyzing polysaccharides.[73] HILIC resolves analytes based on their degree of polymerization, which results in an incomplete separation of certain isomers.[74] Separating GAG oligosaccharides into fractions was achieved using SEC and SAX coupled with volatile ammonium bicarbonate.[75] PGC has had reproducibility and PGC column stability issues despite its good chemical stability.[76]

Capillary zone electrophoresis (CZE)

Capillary zone electrophoresis (CZE) is a separation technique for biomolecules based on their charge, size, and shape (Equation 1.1). CZE is specifically a sub-CE technique and alternative to chromatographic approaches. Charged analytes migrate differentially within a capillary with an applied electric field on the order of 500 V/cm. In general, this is due to two forces: electrophoretic flow (EF, Equation 1.2 and 1.3) and electroosmotic flow (EOF, Equation 1.4).

$$\mu_{obs} = \mu_{ep} + \mu_{eo} \qquad \text{Equation 1.1}$$

$$\mu_{ep} = v_{ep}/E \qquad \text{Equation 1.2}$$

$$\mu_{ep} = q / 6\pi\eta r \qquad \text{Equation 1.3}$$

μ_{ep} : electrophoretic mobility

v_{ep} : electrophoretic velocity

E : electrode potential

q : the charge of an analyte

η : the viscosity of the background electrolyte

r : the radius of an analyte

$$\mu_{eo} = \varepsilon\zeta / 4\pi\eta \qquad \text{Equation 1.4}$$

μ_{eo} : electroosmotic flow

ζ : zeta potential

η : the viscosity of the background electrolyte

EOF is affected by the chemical nature of the inner wall of a capillary. In a bare fused silica capillary (BFS), an electrical double layer forms at the inner wall, with solvated mobile cations moving away from the positive potential, or toward a negative potential, creating a flow of solvent with a flat velocity profile. As a result of the electroosmotic flow, the mixture of the background electrolyte (BGE), as well as the analytes migrate towards the cathode. Normal (also known as forward or direct) polarity is applied to bare fused silica capillaries in which EOF is induced to flow in the direction of an emitter tip in the case of CZE-MS. Early studies of CZE analysis

of GAGs focused primarily on the normal mode polarity, which applies a positive potential to the capillary column while limiting longer migration times.[77, 78]

There are several advantages associated with CZE that make it an attractive separation method for analyzing GAGs.[79, 80] First, short migration times can be achieved when using a high electric field (~± 10~30 kV). High throughput analysis can be performed with high efficiency and reliability. Second, superior separation performance can be achieved due to the theoretical plate number and high-resolution that CZE instruments represent. In comparison to other separation technologies coupling with MS, CZE-MS and CZE-MS/MS separated sulfated GAG isomers, allowing for a sequence analysis including the localization of the sulfate group. Third, injections within the nanoliter range are possible for CZE, which is ideal for biological samples with limited volume. Similarly, compared with LC, organic solvent consumption for buffers, BGE, and sheath liquid is required in much smaller amounts.

CZE has been widely employed as an adequate MS-coupled approach for characterization of biomolecules in proteomics, peptidomics, glycomics, and lipidomics.[79, 81, 82] Electrospray is a common ionization technique for CZE-MS.[83] An electric field applied to the liquid phase generates charged droplets out of the spray tip, creating ions in the gas phase. Then, individual droplet rapidly shrinks to smaller size through evaporation as the charge density within a droplet increases. The surface tension of a droplet that has a high charge density is then subject to charge repulsion on its surface, called Rayleigh limit, resulting in the droplet exploding and forming smaller droplets of nanometer radius.[84]

A typical CZE-MS interface insists of two types: including sheath flow or sheathless ESI.[85] The difference between these two types of interfaces is the use of sheath liquid

and electrical contact status. The sheath liquid flow type contains a reservoir in the coaxial direction to facilitate the electrosprays for CZE as adequate potential is applied by an external power supply (~2 kV). After that, the sheath liquid is combined at the end of an emitter spray tip by the mixture of BGE and analytes migrated from the capillary (Figure 3). It is also necessary to maintain a stable electrical potential for electrophoretic voltage delivery across a capillary, as well as for electrospray with the aid of sheath flow toward the inlet of a mass spectrometer.[86]

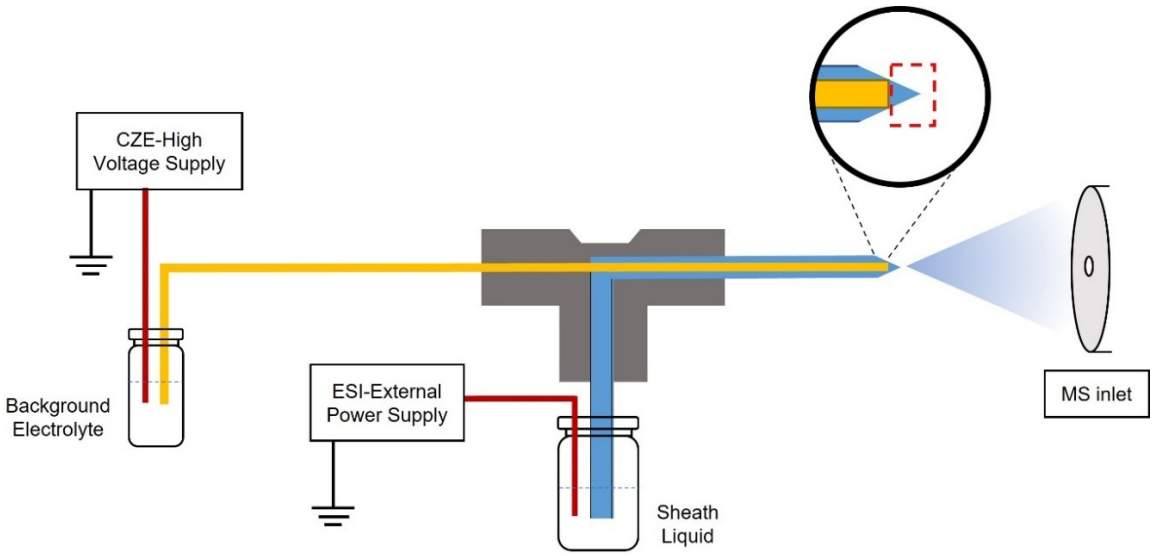

Figure 3. Illustration of CZE-MS sheath flow interface. High voltage (-30 kV) was set to CZE-High voltage supply source and the external power supply was set to -2.0 kV to establish a stable flow of the sheath liquid through an emitter tip. Separation capillary contact with emitter sheath act as electrical circuit, mixing background electrolyte with sheath liquid in small volumes (e.g., 200 pL) at the end of the capillary (red dots).

CZE-MS has investigated various biomolecules with high resolution, accuracy, and sensitivity over the past four decades along with the development of CZE and MS technology.[79, 87] In the meantime, the features of CZE-MS have also been enhanced in terms of sensitivity and resolution by improving the CE-MS interface and developing sample pre-concentration methods (e.g. sample stacking and solid-phase extraction).[88] For example, solid-phase extraction CZE-MS (SPE-CZE-MS) method has proven to be highly effective as a method for online preconcentration and for routine protein analysis.[89] Also, the introduction of high-resolution mass spectrometers has improved its resolution as a second dimension, enabling precise measurements beyond migration time assumptions from other detection methods (e.g. UV/fluorescence detection).[78, 90]

Recently, Sanderson et al. developed reverse polarity CZE-MS for profiling sulfated GAG mixtures. Specially coated with N-(6-aminohexyl) aminomethyltriethoxysilane (AHS), a cation material, capillaries have been effectively separated HS oligosaccharides from tetrasacchrides to dodecasacchrides.[91] Stolz et al. achieved high sensitivity and productivity in detecting 368 glycoforms for the human alpha-1-acid glycoprotein (AGP) by a nanoLC-CZE-MS interface as a two-dimensional separational approach, whereas CZE-MS could identify only 186 glycoforms, even for the 20 times more concentrated sample (50 ug/mL vs. 1 mg/mL).[92]

CZE-MS/MS

CZE-MS/MS has been used to generate structurally informative data for the identification of biomolecules using MS/MS technique (e.g. collision-based activation or electron-based activations).[93, 94] Recently, McCool et al. used CZE-activated ion

electron transfer dissociation (CZE-AI-ETD) to analyze Escherichia coli (E.coli) cells, and identified 3028 proteoforms and 387 proteins, resulting in the largest proteomics dataset available.[95] Considered all advantages of CZE technique, it is well-tailored to the highly anionic nature of glycosaminoglycans (GAGs) and can separate them based on their type, length, charge density, degree of modifications, motif arrangement. As oligosaccharides contain deprotonated residues, they are especially suitable for CZE-MS and CZE-MS/MS.[26] Moreover, coupling CZE-MS is more efficient in identifying heterogeneous structural information pertaining to GAGs since UV/fluorescence analyzers require oligosaccharide standards to assign their exact structures based on CE migration times.[78]

CZE-MS has been used starting in the late 1990s, and has proved to be a very attractive tool for the characterization of heterogenous structure of GAGs. [26, 57, 78, 96-100] In the early literature on GAG analysis, the normal polarity mode of CZE-MS was employed in which positive potential is applied to a capillary. While much research has performed with CZE in normal polarity using direct UV detection, several studies have also used CZE in normal polarity connected to MS.[101, 102] For instance, Duteil et al. investigated CZE-MS to analyze heparinases-depolymerized heparins under both normal and reverse CZE polarity.[103]

Reverse polarity CZE has upgraded the application of CZE-MS to identify glycoforms, including GAG analysis[26, 57, 98, 99, 104]. CZE-MS/MS in reverse polarity has also emerged as a useful tool in GAG analysis. In addition, NETD is more compelling as an MS/MS method because it produces not only informative fragment ions, but also a relatively quick reaction with fluoranthene, subject to online separation, that sequences GAG sulfation patterns. Stickney et al. applied CZE-NETD-MS/MS to the highly sulfated pharmaceutical, Enoxaparin. The CZE-NETD-

MS/MS data has effectively assigned 37 heparin compositions and 9 structural identifications for sulfate groups located on heparin.[57] Han et al. employed CZE-NETD-MS/MS to urinary GAG samples from different ages and genders with a focus on profiling HS oligosaccharides that contained sulfate group modifications. This study established a normal range for GAG concentration and structure despite the fact that the need to further examine controlled individuals (e.g., hydration, diet, health status) for clinical applications of GAG analysis was implied. [99] Yu et al. developed a multifaceted methodology that included a filter binding assay, LC-MS, and CZE-MS, which led to the discovery of the relationship between the antithrombin affinity domain and the oligosaccharide chains of low molecular weight heparins through CZE-MS and CZE-NETD-MS/MS. [100]

Despite significant research development in CZE-MS and CZE-MS/MS for complex GAG samples like highly sulfated GAGs and proteoglycans, fragmentation and chemometrics still require further improvement. It is vital for researchers to improve a versatile CZE-MS interface with other developed technologies, such as orthogonal separation, preconcentration, and other MS/MS dissociation methods, for the structural elucidation of biomolecules. Consequently, CZE-MS and CZE-MS/MS are highly promising tools for the structural elucidation of GAGs.

1.5 Summary

Glycosaminoglycans (GAGs) are linear polysaccharides expressed in mammalian tissues. They are classified into heparin, heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate, and hyaluronan based on the repeating disaccharide and linkages between them. GAGs have been a subject of significant

interest in both academia and industry. Their structures, sequences, and specific sulfation patterns serve as keys to interact with their ligands and perform essential functions during biological processes.

Numerous analytical methods have been used to characterize enzymatically depolymerized GAGs as well as synthesized GAGs. In particular, mass spectrometry and tandem mass spectrometry have been widely used in GAG analysis. Electrospray ionization (ESI) is compatible with GAG separation techniques due to the hydrophilic nature and multiple negative charges. GAGs can be identified at the MS1 level based on their intact mass, but not their structure. The composition information includes the dp for the GAG, the number of sulfo modifications, the number of NAc present, as well as any terminal modifications. It is possible to elucidate the structural properties of GAGs based on glycosidic bond cleavages and cross-ring bond cleavages by tandem mass spectrometry. The type of activation technique is chosen based on the sample properties, such as complexity of sulfation, chain lengths, purity, etc., to induce fragmentation for GAG analysis. For example, electron detachment dissociation (EDD) and negative electron transfer dissociation (NETD) as MS/MS approaches have been significantly achieved to delineate the structures of GAGs, and to sequence the exact sulfate modification on GAGs with informative-rich fragmentation in the past two decades. Both have attracted attention; however, NETD has been chosen because it has a shorter activation time (both reagent accumulation and reaction time) compared to EDD, which may lead to a successful combination with separation methods.

In many applications, on-line separation coupled with mass spectrometry has proven to be beneficial for GAG mixtures. The capillary zone electrophoresis-mass spectrometry (CZE-MS) is a separation technique for biomolecules based on their

charge, size, and shape. GAG molecules are migrated differently, even isomers that have the same mass-to-charge ratio, through a narrow capillary by a total of electrophoretic flow and electroosmotic flow. Capillary inner wall functionality also directly influences electroosmotic flow direction, which determines migration speed in general. Since neutral/negatively functioned capillaries have high separation efficiency and robustness, they are ideal for MS analysis of GAGs. CZE is well comparable to ESI. CZE-MS, therefore, offers several advantages to separate complex GAG oligomers and analyze each GAG molecule in terms of sensitivity, accuracy, separation efficiency, rapid migration, and low sample consumption. However, CZE-MS/MS methods still have some unmet challenges. The fragmentation of GAG molecules requires a high-efficiency MS/MS to occur within a short migration window, so MS/MS techniques need further exploration. The fragmentation of GAG molecules requires a high-efficiency MS/MS to occur within a short migration window, so the development of MS/MS techniques is expected. Regarding CZE-NETD-MS/MS, highly charged precursor ions allow for better fragmentation. In addition, future developments in hybrid techniques and data analysis tools will allow us to better understand GAG activity.

REFERENCES

(1) Jackson, R. L.; Busch, S. J.; Cardin, A. D. Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. *Physiological reviews* **1991**, *71* (2), 481-539.

(2) Weinbaum, S.; Tarbell, J. M.; Damiano, E. R. The structure and function of the endothelial glycocalyx layer. *Annual Review of Biomedical Engineering* **2007**, *9*, 121-167.

(3) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H. Essentials of Glycobiology **2015**.

(4) Huang, Y.-F.; Aoki, K.; Akase, S.; Ishihara, M.; Liu, Y.-S.; Yang, G.; Kizuka, Y.; Mizumoto, S.; Tiemeyer, M.; Gao, X.-D.; et al. Global mapping of glycosylation pathways in human-derived cells. *Developmental Cell* **2021**, *56* (8), 1195-1209.e1197.

(5) Gandhi, N. S.; Mancera, R. L. The Structure of Glycosaminoglycans and their Interactions with Proteins. *Chemical Biology & Drug Design* **2008**, *72* (6), 455-482.

(6) Zhang, B.; Chi, L. Chondroitin Sulfate/Dermatan Sulfate-Protein Interactions and Their Biological Functions in Human Diseases: Implications and Analytical Tools. *Frontiers in Cell and Developmental Biology* **2021**, *9*, Review.

(7) Yu, M.; Zhang, T.; Zhang, W.; Sun, Q.; Li, H.; Li, J. Elucidating the Interactions Between Heparin/Heparan Sulfate and SARS-CoV-2-Related Proteins—An Important

Strategy for Developing Novel Therapeutics for the COVID-19 Pandemic. *Frontiers in Molecular Biosciences* **2021**, *7*, Mini Review.

(8) Kim, S. Y.; Jin, W.; Sood, A.; Montgomery, D. W.; Grant, O. C.; Fuster, M. M.; Fu, L.; Dordick, J. S.; Woods, R. J.; Zhang, F.; et al. Glycosaminoglycan binding motif at S1/S2 proteolytic cleavage site on spike glycoprotein may facilitate novel coronavirus (SARS-CoV-2) host cell entry. *bioRxiv* **2020**, 2020.2004.2014.041459.

(9) Casu, B.; Lindahl, U. Structure and biological interactions of heparin and heparan sulfate. *Advances in Carbohydrate Chemistry and Biochemistry* **2001**, *57*, 159-206.

(10) Zhang, L.; Lawrence, R.; Schwartz, J. J.; Bai, X.; Wei, G.; Esko, J. D.; Rosenberg, R. D. The Effect of Precursor Structures on the Action of Glucosaminyl 3-O-Sulfotransferase-1 and the Biosynthesis of Anticoagulant Heparan Sulfate. *Journal of Biological Chemistry* **2001**, *276* (31), 28806-28813.

(11) Plaas, A. H. K.; West, L. A.; Wong-Palms, S.; Nelson, F. R. T. Glycosaminoglycan Sulfation in Human Osteoarthritis: DISEASE-RELATED ALTERATIONS AT THE NON-REDUCING TERMINI OF CHONDROITIN AND DERMATAN SULFATE. *Journal of Biological Chemistry* **1998**, *273* (20), 12642-12649.

(12) Sasarman, F.; Maftei, C.; Campeau, P. M.; Brunel-Guitton, C.; Mitchell, G. A.; Allard, P. Biosynthesis of glycosaminoglycans: associated disorders and biochemical tests. *Journal of Inherited Metabolic Disease* **2016**, *39* (2), 173-188.

(13) Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J. Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Molecular & cellular proteomics : MCP* **2021**, *20*, 100025.

(14) Goncalves, J. P.; Ghebosu, R. E.; Tan, X. N. S.; Iannotta, D.; Koifman, N. a.; Wolfram, J. Hyaluronic acid: An overlooked extracellular vesicle contaminant. *Journal of Extracellular Vesicles* **2023**, *12* (9), 12362.

(15) Turley, E. A.; Noble, P. W.; Bourguignon, L. Y. W. Signaling Properties of Hyaluronan Receptors. *Journal of Biological Chemistry* **2002**, *277* (7), 4589-4592.

(16) Snetkov, P.; Zakharova, K.; Morozkina, S.; Olekhnovich, R.; Uspenskaya, M. Hyaluronic Acid: The Influence of Molecular Weight on Structural, Physical, Physico-Chemical, and Degradable Properties of Biopolymer. *Polymers* **2020**, *12* (8), 1800.

(17) Prydz, K.; Dalen, K. T. Synthesis and sorting of proteoglycans. *Journal of Cell Science* **2000**, *113* (2), 193-205.

(18) Huang, Y.-F.; Mizumoto, S.; Fujita, M. Novel Insight Into Glycosaminoglycan Biosynthesis Based on Gene Expression Profiles. *Frontiers in Cell and Developmental Biology* **2021**, *9*, Brief Research Report.

(19) Sugahara, K.; Kitagawa, H. Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Current Opinion in Structural Biology* **2000**, *10* (5), 518-527.

(20) Zhang, L. Glycosaminoglycan (GAG) Biosynthesis and GAG-Binding Proteins. In *Progress in Molecular Biology and Translational Science*, Zhang, L. Ed.; Vol. 93; Academic Press, 2010; pp 1-17.

(21) Wang, Q.; Chi, L. The Alterations and Roles of Glycosaminoglycans in Human Diseases. *Polymers (Basel)* **2022**, *14* (22).

(22) O Staples, G.; Zaia, J. Analysis of glycosaminoglycans using mass spectrometry.

Current proteomics **2011**, *8* (4), 325-336.

(23) Zaia, J. Glycosaminoglycan glycomics using mass spectrometry. *Molecular and Cellular Proteomics* **2013**, *12* (4), 885-892.

(24) Perez, S.; Makshakova, O.; Angulo, J.; Bedini, E.; Bisio, A.; de Paz, J. L.; Fadda, E.; Guerrini, M.; Hricovini, M.; Hricovini, M. Glycosaminoglycans: what remains to be deciphered? *JACS Au* **2023**, *3* (3), 628-656.

(25) Chi, L.; Wolff, J. J.; Laremore, T. N.; Restaino, O. F.; Xie, J.; Schiraldi, C.; Toida, T.; Amster, I. J.; Linhardt, R. J. Structural analysis of bikunin glycosaminoglycan. *Journal of the American Chemical Society* **2008**, *130* (8), 2617-2625.

(26) Sanderson, P.; Stickney, M.; Leach, F. E., 3rd; Xia, Q.; Yu, Y.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Heparin/heparan sulfate analysis by covalently modified reverse polarity capillary zone electrophoresis-mass spectrometry. *Journal of chromatography. A* **2018**, *1545*, 75-83.

(27) Domon, B.; Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal* **1988**, *5* (4), 397-409.

(28) Agyekum, I.; Patel, A. B.; Zong, C.; Boons, G.-J.; Amster, I. J. Assignment of hexuronic acid stereochemistry in synthetic heparan sulfate tetrasaccharides with 2-*O*-sulfo uronic acids using electron detachment dissociation. *International journal of mass spectrometry* **2015**, *390*, 163-169.

(29) Zaia, J.; Costello, C. E. Tandem mass spectrometry of sulfated heparin-like glycosaminoglycan oligosaccharides. *Analytical Chemistry* **2003**, *75* (10), 2445-2455.

(30) Agyekum, I.; Zong, C.; Boons, G.-J.; Amster, I. J. Single Stage Tandem Mass Spectrometry Assignment of the C-5 Uronic Acid Stereochemistry in Heparan Sulfate Tetrasaccharides using Electron Detachment Dissociation. *Journal of the American Society for Mass Spectrometry* **2017**, *28* (9), 1741-1750.

(31) Leach III, F. E.; Ly, M.; Laremore, T. N.; Wolff, J. J.; Perlow, J.; Linhardt, R. J.; Amster, I. J. Hexuronic acid stereochemistry determination in chondroitin sulfate glycosaminoglycan oligosaccharides by electron detachment dissociation. *Journal of the American Society for Mass Spectrometry* **2012**, *23* (9), 1488-1497.

(32) Kailemia, M. J.; Patel, A. B.; Johnson, D. T.; Li, L.; Linhardt, R. J.; Amster, I. J. Differentiating chondroitin sulfate glycosaminoglycans using collision-induced dissociation; uronic acid cross-ring diagnostic fragments in a single stage of tandem mass spectrometry. *European Journal of Mass Spectrometry* **2015**, *21* (3), 275-285.

(33) Zaia, J.; Li, X.-Q.; Chan, S.-Y.; Costello, C. E. Tandem mass spectrometric strategies for determination of sulfation positions and uronic acid epimerization in chondroitin sulfate oligosaccharides. *Journal of the American Society for Mass Spectrometry* **2003**, *14* (11), 1270-1281.

(34) Zamfir, A.; Seidler, D. G.; Kresse, H.; Peter-Katalinić, J. Structural characterization of chondroitin/dermatan sulfate oligosaccharides from bovine aorta by capillary electrophoresis and electrospray ionization quadrupole time-of-flight tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **2002**, *16* (21), 2015-2024.

(35) Zaia, J.; Costello, C. E. Compositional Analysis of Glycosaminoglycans by Electrospray Mass Spectrometry. *Analytical Chemistry* **2001**, *73* (2), 233-239.

(36) Zaia, J. Principles of mass spectrometry of glycosaminoglycans. *Journal of Biomacromolecular Mass Spectrometry* **2005**, *1* (1), 3-36.

(37) Kailemia, M. J.; Li, L.; Ly, M.; Linhardt, R. J.; Amster, I. J. Complete mass spectral characterization of a synthetic ultralow-molecular-weight heparin using collision-induced dissociation. *Analytical Chemistry* **2012**, *84* (13), 5475-5478.

(38) Klein, D. R.; Leach, F. E., 3rd; Amster, I. J.; Brodbelt, J. S. Structural Characterization of Glycosaminoglycan Carbohydrates Using Ultraviolet Photodissociation. *Analytical Chemistry* **2019**, *91* (9), 6019-6026.

(39) Agyekum, I.; Pepi, L.; Yu, Y.; Li, J.; Yan, L.; Linhardt, R. J.; Chen, S.; Amster, I. J. Structural elucidation of fucosylated chondroitin sulfates from sea cucumber using FTICR-MS/MS. *European Journal of Mass Spectrometry* **2018**, *24* (1), 157-167.

(40) Oh, H. B.; Leach III, F. E.; Arungundram, S.; Al-Mafraji, K.; Venot, A.; Boons, G.-J.; Amster, I. J. Multivariate analysis of electron detachment dissociation and infrared multiphoton dissociation mass spectra of heparan sulfate tetrasaccharides differing only in hexuronic acid stereochemistry. *Journal of the American Society for Mass Spectrometry* **2011**, *22* (3), 582-590.

(41) Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. Electron detachment dissociation of glycosaminoglycan tetrasaccharides. *Journal of the American Society for Mass Spectrometry* **2007**, *18* (2), 234-244.

(42) Wolff, J. J.; Leach III, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. Negative electron transfer dissociation of glycosaminoglycans. *Analytical Chemistry* **2010**, *82* (9), 3460-3466.

(43) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society* **1998**, *120* (13), 3265-3266.

(44) Kelleher, N. L. Peer reviewed: top-down proteomics. *Analytical Chemistry* **2004**, *76* (11), 196-A.

(45) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Hakansson, K. Negative-ion electron capture dissociation: radical-driven fragmentation of charge-increased gaseous peptide anions. *Journal of the American Society for Mass Spectrometry* **2011**, *133* (42), 16790-16793.

(46) Borotto, N. B.; Ileka, K. M.; Tom, C.; Martin, B. R.; Hakansson, K. Free Radical Initiated Peptide Sequencing for Direct Site Localization of Sulfation and Phosphorylation with Negative Ion Mode Mass Spectrometry. *Analytical Chemistry* **2018**, *90* (16), 9682-9686.

(47) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. Electron detachment dissociation of peptide di-anions: an electron-hole recombination phenomenon. *Chemical Physics Letter* **2001**, *342* (3-4), 299-302.

(48) Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. Characterization of oligodeoxynucleotides by electron detachment dissociation fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* **2005**, *77* (6), 1876-1882.

(49) Wolff, J. J.; Laremore, T. N.; Aslam, H.; Linhardt, R. J.; Amster, I. J. Electron-induced dissociation of glycosaminoglycan tetrasaccharides. *Journal of the American Society for Mass Spectrometry* **2008**, *19* (10), 1449-1458.

(50) Pepi, L. E.; Amster, I. J. Electron Detachment Dissociation (EDD) and Negative Electron Transfer Dissociation (NETD). *Advanced Fragmentation Methods in Biomolecular Mass Spectrometry* **2020**, *9*, 134.

(51) Leach Iii, F. E.; Wolff, J. J.; Laremore, T. N.; Linhardt, R. J.; Amster, I. J. Evaluation of the experimental parameters which control electron detachment dissociation, and their effect on the fragmentation efficiency of glycosaminoglycan carbohydrates. *International Journal of Mass Spectrometry* **2008**, *276* (2-3), 110-115.

(52) Leach Iii, F. E.; Wolff, J. J.; Xiao, Z.; Ly, M.; Laremore, T. N.; Arungundram, S.; Al-Mafraji, K.; Venot, A.; Boons, G.-J.; Linhardt, R. J. Negative electron transfer dissociation Fourier transform mass spectrometry of glycosaminoglycan carbohydrates. *European Journal of Mass Spectrometry* **2011**, *17* (2), 167-176.

(53) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. Gas-phase electron transfer reactions from multiply-charged anions to rare gas cations. *Journal of the American Chemical Society* **1995**, *117* (46), 11555-11562.

(54) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences* **2004**, *101* (26), 9528-9533.

(55) Coon, J. J.; Ueberheide, B.; Syka, J. E. P.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **2005**, *102* (27), 9463-9468.

(56) Wu, J.; Wei, J.; Hogan, J. D.; Chopra, P.; Joshi, A.; Lu, W.; Klein, J.; Boons, G. J.; Lin, C.; Zaia, J. Negative Electron Transfer Dissociation Sequencing of 3-*O*-Sulfation-

Containing Heparan Sulfate Oligosaccharides. *Journal of the American Society for Mass Spectrometry* **2018**, *29* (6), 1262-1272.

(57) Stickney, M.; Sanderson, P.; Leach, F. E.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Online capillary zone electrophoresis negative electron transfer dissociation tandem mass spectrometry of glycosaminoglycan mixtures. *International Journal of Mass Spectrometry* **2019**, *445*, 116209.

(58) Wei, J.; Wu, J.; Tang, Y.; Ridgeway, M. E.; Park, M. A.; Costello, C. E.; Zaia, J.; Lin, C. Characterization and quantification of highly sulfated glycosaminoglycan isomers by gated-trapped ion mobility spectrometry negative electron transfer dissociation MS/MS. *Analytical Chemistry* **2019**, *91* (4), 2994-3001.

(59) Leach, F. E.; Riley, N. M.; Westphall, M. S.; Coon, J. J.; Amster, I. J. Negative Electron Transfer Dissociation Sequencing of Increasingly Sulfated Glycosaminoglycan Oligosaccharides on an Orbitrap Mass Spectrometer. *Journal of the American Society for Mass Spectrometry* **2017**, *28* (9), 1844-1854.

(60) Wu, J.; Wei, J.; Chopra, P.; Boons, G.-J.; Lin, C.; Zaia, J. Sequencing Heparan Sulfate Using HILIC LC-NETD-MS/MS. *Analytical Chemistry* **2019**, *91* (18), 11738-11746.

(61) Huang, R.; Liu, J.; Sharp, J. S. An Approach for Separation and Complete Structural Sequencing of Heparin/Heparan Sulfate-like Oligosaccharides. *Analytical Chemistry* **2013**, *85* (12), 5787-5795.

(62) Afratis, N.; Gialeli, C.; Nikitovic, D.; Tsegenidis, T.; Karousou, E.; Theocharis, A. D.; Pavão, M. S.; Tzanakakis, G. N.; Karamanos, N. K. Glycosaminoglycans: key players in cancer cell biology and treatment. *The FEBS Journal* **2012**, *279* (7), 1177-

1197.

(63) Tanaka, N.; Kida, S.; Kinoshita, M.; Morimoto, H.; Shibasaki, T.; Tachibana, K.; Yamamoto, R. Evaluation of cerebrospinal fluid heparan sulfate as a biomarker of neuropathology in a murine model of mucopolysaccharidosis type II using high-sensitivity LC/MS/MS. *Molecular Genetics and Metabolism* **2018**, *125* (1-2), 53-58.

(64) Forni, G.; Malvagia, S.; Funghini, S.; Scolamiero, E.; Mura, M.; Della Bona, M.; Villanelli, F.; Damiano, R.; la Marca, G. LC-MS/MS method for simultaneous quantification of heparan sulfate and dermatan sulfate in urine by butanolysis derivatization. *Clinica Chimica Acta* **2019**, *488*, 98-103.

(65) Li, G.; Li, L.; Joo, E. J.; Son, J. W.; Kim, Y. J.; Kang, J. K.; Lee, K. B.; Zhang, F.; Linhardt, R. J. Glycosaminoglycans and glycolipids as potential biomarkers in lung cancer. *Glycoconjugate Journal* **2017**, *34* (5), 661-669.

(66) Wei, J.; Hu, M.; Huang, K.; Lin, S.; Du, H. Roles of proteoglycans and glycosaminoglycans in cancer development and progression. *International Journal of Molecular Sciences* **2020**, *21* (17), 5983.

(67) Bratulic, S.; Limeta, A.; Dabestani, S.; Birgisson, H.; Enblad, G.; Stålberg, K.; Hesselager, G.; Häggman, M.; Höglund, M.; Simonson, O. E. Noninvasive detection of any-stage cancer using free glycosaminoglycans. *Proceedings of the National Academy of Sciences* **2022**, *119* (50), e2115328119.

(68) Cerezo-Magaña, M.; Bång-Rudenstam, A.; Belting, M. Proteoglycans: a common portal for SARS-CoV-2 and extracellular vesicle uptake. *American Journal of Physiology-Cell Physiology* **2023**, *324* (1), C76-C84.

(69) Kubaski, F.; Mason, R. W.; Nakatomi, A.; Shintaku, H.; Xie, L.; van Vlies, N. N.; Church, H.; Giugliani, R.; Kobayashi, H.; Yamaguchi, S.; et al. Newborn screening for mucopolysaccharidoses: a pilot study of measurement of glycosaminoglycans by tandem mass spectrometry. *Journal of Inherited Metabolic Disease* **2017**, *40* (1), 151-158.

(70) Pál, D.; Tóth, G.; Sugár, S.; Fügedi, K. D.; Szabó, D.; Kovalszky, I.; Papp, D.; Schlosser, G.; Tóth, C.; Tornóczky, T.; et al. Compositional Analysis of Glycosaminoglycans in Different Lung Cancer Type-A Pilot Study. *International Journal of Molecular Sciences* **2023**, *24* (8), 7050.

(71) Zaia, J. On-line separations combined with MS for analysis of glycosaminoglycans. *Mass Spectrometry Reviews* **2009**, *28* (2), 254-272.

(72) Kuberan, B.; Lech, M.; Zhang, L.; Wu, Z. L.; Beeler, D. L.; Rosenberg, R. D. Analysis of Heparan Sulfate Oligosaccharides with Ion Pair-Reverse Phase Capillary High Performance Liquid Chromatography-Microelectrospray Ionization Time-of-Flight Mass Spectrometry. *Journal of the American Chemical Society* **2002**, *124* (29), 8707-8718.

(73) Lin, T.-S.; Hsieh, C.-H.; Kuo, C.; Juang, Y.-P.; Hsieh, Y. S. Y.; Chiang, H.; Hung, S.-C.; Jiang, C.-C.; Liang, P.-H. Sulfation pattern of chondroitin sulfate in human osteoarthritis cartilages reveals a lower level of chondroitin-4-sulfate. *Carbohydrate Polymers* **2020**, *229*, 115496.

(74) Gill, V. L.; Aich, U.; Rao, S.; Pohl, C.; Zaia, J. Disaccharide analysis of glycosaminoglycans using hydrophilic interaction chromatography and mass spectrometry. *Analytical Chemistry* **2013**, *85* (2), 1138-1145.

(75) 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. *Analytical Chemistry* **2016**, *88* (23), 11542-11550.

(76) Gray, C. J.; Sánchez-Ruíz, A.; Šardzíková, I.; Ahmed, Y. A.; Miller, R. L.; Reyes Martinez, J. E.; Pallister, E.; Huang, K.; Both, P.; Hartmann, M.; et al. Label-Free Discovery Array Platform for the Characterization of Glycan Binding Proteins and Glycoproteins. *Analytical Chemistry* **2017**, *89* (8), 4444-4451.

(77) Prabhakar, V.; Capila, I.; Sasisekharan, R. The structural elucidation of glycosaminoglycans. *Glycomics: Methods and Protocols* **2009**, 147-156.

(78) Volpi, N.; Maccari, F.; Linhardt, R. J. Capillary electrophoresis of complex natural polysaccharides. *Electrophoresis* **2008**, *29* (15), 3095-3106.

(79) Kaur, H.; Beckman, J.; Zhang, Y.; Li, Z. J.; Szigeti, M.; Guttman, A. Capillary electrophoresis and the biopharmaceutical industry: Therapeutic protein analysis and characterization. *TrAC Trends in Analytical Chemistry* **2021**, *144*, 116407.

(80) Khan, S. A.; Mason, R. W.; Kobayashi, H.; Yamaguchi, S.; Tomatsu, S. Advances in glycosaminoglycan detection. *Molecular Genetics and Metabolism* **2020**, *130* (2), 101-109.

(81) Oefner, P. J.; Chiesa, C. Capillary electrophoresis of carbohydrates. *Glycobiology* **1994**, *4* (4), 397-412.

(82) Lu, G.; Crihfield, C. L.; Gattu, S.; Veltri, L. M.; Holland, L. A. Capillary Electrophoresis Separations of Glycans. *Chemical Reviews* **2018**, *118* (17), 7867-7885.

(83) Wu, H.; Tang, K. Highly sensitive and robust capillary electrophoresis-electrospray ionization-mass spectrometry: Interfaces, preconcentration techniques and applications. *Reviews in Analytical Chemistry* **2020**, *39* (1), 45-55.

(84) Lapizco-Encinas, B. H.; Zhang, Y. V.; Gqamana, P. P.; Lavicka, J.; Foret, F. Capillary electrophoresis as a sample separation step to mass spectrometry analysis: A primer. *TrAC Trends in Analytical Chemistry* **2023**, *164*, 117093.

(85) Maxwell, E. J.; Chen, D. D. Y. Twenty years of interface development for capillary electrophoresis–electrospray ionization–mass spectrometry. *Analytica chimica acta* **2008**, *627* (1), 25-33.

(86) Furter, J. S.; Hauser, P. C. Compact automated capillary electrophoresis instrument for coupling with mass spectrometry by using sheathless electrospray ionization. *Journal of Chromatography A* **2021**, *1656*, 462533.

(87) Olivares, J. A.; Nguyen, N. T.; Yonker, C. R.; Smith, R. D. On-line mass spectrometric detection for capillary zone electrophoresis. *Analytical Chemistry* **1987**, *59* (8), 1230-1232.

(88) Stroink, T.; Paarlberg, E.; Waterval, J. C. M.; Bult, A.; Underberg, W. J. M. On-line sample preconcentration in capillary electrophoresis, focused on the determination of proteins and peptides. *Electrophoresis* **2001**, *22* (12), 2374-2383.

(89) Figeys, D.; Zhang, Y.; Aebersold, R. Optimization of solid phase microextraction - capillary zone electrophoresis - mass spectrometry for high sensitivity protein identification. *Electrophoresis* **1998**, *19* (13), 2338-2347.

(90) Shen, X.; Liang, Z.; Xu, T.; Yang, Z.; Wang, Q.; Chen, D.; Pham, L.; Du, W.; Sun,

L. Investigating native capillary zone electrophoresis-mass spectrometry on a high-end quadrupole-time-of-flight mass spectrometer for the characterization of monoclonal antibodies. *International Journal of Mass Spectrometry* **2021**, *462*, 116541.

(91) Sanderson, P.; Stickney, M.; Leach Iii, F. E.; Xia, Q.; Yu, Y.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Heparin/heparan sulfate analysis by covalently modified reverse polarity capillary zone electrophoresis-mass spectrometry. *Journal of Chromatography A* **2018**, *1545*, 75-83.

(92) Stolz, A.; Neusüß, C. Characterisation of a new online nanoLC-CZE-MS platform and application for the glycosylation profiling of alpha-1-acid glycoprotein. *Analytical and Bioanalytical Chemistry* **2022**, *414* (5), 1745-1757.

(93) Jin, X.; Kim, J.; Parus, S.; Lubman, D. M.; Zand, R. On-line capillary electrophoresis/microelectrospray ionization-tandem mass spectrometry using an ion trap storage/time-of-flight mass spectrometer with SWIFT technology. *Analytical Chemistry* **1999**, *71* (16), 3591-3597.

(94) Tong, W.; Link, A.; Eng, J. K.; Yates, J. R. Identification of Proteins in Complexes by Solid-Phase Microextraction/Multistep Elution/Capillary Electrophoresis/Tandem Mass Spectrometry. *Analytical Chemistry* **1999**, *71* (13), 2270-2278.

(95) McCool, E. N.; Lodge, J. M.; Basharat, A. R.; Liu, X.; Coon, J. J.; Sun, L. Capillary Zone Electrophoresis-Tandem Mass Spectrometry with Activated Ion Electron Transfer Dissociation for Large-scale Top-down Proteomics. *Journal of The American Society for Mass Spectrometry* **2019**, *30* (12), 2470-2479.

(96) Duteil, S.; Gareil, P.; Girault, S.; Mallet, A.; Feve, C.; Siret, L. Identification of heparin oligosaccharides by direct coupling of capillary electrophoresis/Ionspray-Mass

spectrometry. *Rapid Communications in Mass Spectrometry* **1999**, *13* (19), 1889-1898.

(97) Zamfir, A. D. Applications of capillary electrophoresis electrospray ionization mass spectrometry in glycosaminoglycan analysis. *Electrophoresis* **2016**, *37* (7-8), 973-986.

(98) Lin, L.; Liu, X.; Zhang, F.; Chi, L.; Amster, I. J.; Leach, F. E.; Xia, Q.; Linhardt, R. J. Analysis of heparin oligosaccharides by capillary electrophoresis–negative-ion electrospray ionization mass spectrometry. *Analytical and Bioanalytical Chemistry* **2017**, *409*, 411-420.

(99) Han, X.; Sanderson, P.; Nesheiwat, S.; Lin, L.; Yu, Y.; Zhang, F.; Amster, I. J.; Linhardt, R. J. Structural analysis of urinary glycosaminoglycans from healthy human subjects. *Glycobiology* **2019**, *30* (3), 143-151.

(100) 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.

(101) Honda, S.; Ueno, T.; Kakehi, K. High-performance capillary electrophoresis of unsaturated oligosaccharides derived from glycosaminoglycans by digestion with chondroitinase ABC as 1-phenyl-3-methyl-5-pyrazolone derivatives. *Journal of Chromatography A* **1992**, *608* (1-2), 289-295.

(102) Volpi, N. Separation of capsular polysaccharide K4 and defructosylated K4 by high-performance capillary electrophoresis. *Electrophoresis* **2004**, *25* (4-5), 692-696.

(103) Duteil, S.; Gareil, P.; Girault, S.; Mallet, A.; Feve, C.; Siret, L. Identification of heparin oligosaccharides by direct coupling of capillary electrophoresis/Ionspray-Mass

spectrometry. *Rapid Communications in Mass Spectrometry* **1999**, *13* (19), 1889-1898.

(104) Toida, T.; Linhardt, R. J. Detection of glycosaminoglycans as a copper (II) complex in capillary electrophoresis. *Electrophoresis* **1996**, *17* (2), 341-346.

CHAPTER 2

MULTIVARIATE ANALYSIS OF TANDEM MASS SPECTROMETRY DATA DISTINGUISHES DIASTEREOMERIC GLYCOSAMINOGLYCANS

2.1 Introduction

Glycosaminoglycans (GAGs) are linear anionic polysaccharides that engage in many roles in physiological processes. [1] Their propensity to interact strongly and specifically with proteins allows them to play a biological regulatory role in physiological and pathological processes such as cell growth and development, cancer, inflammation, neurodegenerative diseases, and more. [2-5] They are divided into four main groups that differ in their repeating unit composition and linkage: heparin/heparan sulfate (Hp/HS), chondroitin/dermatan sulfate (CS/DS), keratan sulfate (KS), and hyaluronan (HA), also known as hyaluronic acid. This paper focuses on CS species; CS/DS was the first reported GAG and is the most abundant GAG in the human body. CS has several subtypes, including the most common forms, CS-A, CS-B, and CS-C, as well as a number of less common CS types such as CS-D and CS-E. CS-A and CS-C consist of repeating units of N-acetyl-D-galactosamine (GalNAc) and D-glucuronic acid (GlcA), with sulfation at 4-*O* and 6-*O* positions, respectively, on the amino sugar. In contrast, DS (also known as CS-B) contains GalNAc and D-iduronic acid (IdoA) alternating units, with 4-*O* sulfation on the amino sugar. Therefore, CS differs from DS by the stereochemistry of the hexuronic acid component. There are many less-common patterns of modification, including those

with 2-*O* sulfation on the uronic acid, and those with two sulfation modifications per disaccharide.

CS biosynthesis takes place in the endoplasmic reticulum (ER) and Golgi apparatus. [1] A generic tetrasaccharide linkage region is first initiated in the ER, with xylose tethered to a serine residue of the protein portion of a proteoglycan. Further chain elongation, sulfation, and epimerization occurs in the Golgi via a diverse set of glycosyltransferases and sulfotransferases, adding to the complexity of the structure of CS/DS. [6] Subtle alterations in GAG structure have implicated in human diseases or health status since GAG assembly are unable to function with their normal biological activity. [7] Interestingly, certain GAG structures are found to be involved in cancer invasion, progression, angiogenesis, and metastasis. [8] For instance, it has been shown that overexpression of CS-E (4-*O*, 6-*O* disulfo GlcNAc) is associated with certain types of cancer. [9] Small differences due to heterogeneity inherent in GAG chains make structure assignment more challenging, and can be aided by separation approaches suitable for resolving complex GAG mixtures. [10-12]

A variety of analytical methods have been developed to ensure structural information regarding GAGs such as derivatizing fluorescent functional groups, protein binding methods, liquid chromatography-mass spectrometry, capillary electrophoresis-mass spectrometry and ion mobility MS. [13-18] Mass spectrometry and tandem mass spectrometry have been found to be highly effect techniques for sequencing the structure of GAGs of both intact and depolymerized glycan chains in order to confirm the domain distribution as well as detailed composition and structure within GAG domains. [17-20]

Tandem mass spectrometry has facilitated the characterization of GAGs analyzed in negative ionization mode, consisting of two ion activation approaches:

vibrational excitation and electronic excitation. [19] Vibrational excitation methods include collisional induced dissociation (CID) and infrared multiphoton dissociation (IRMPD), while the electronic excitation includes electron detachment dissociation (EDD), negative electron transfer dissociation (NETD), and ultraviolet photodissociation (UVPD). These ion activation techniques reveal site-specific sulfate modification as well as other features of GAG structure, such as deacetylation and uronic acid epimerization. [21-25]

The sulfation modifications present in GAGs are labile once ionized, and it can be challenging to dissociate the glycan by glycosidic or cross-ring cleavage without also inducing a neutral loss of SO_3, particularly when using vibrational activation. Also, vibrational activation tends to favor glycosidic cleavage. In contrast, electronic excitation can provide abundant cross-ring cleavages that can decode subtle structural differences within oligosaccharides. [29, 21] Structural information including modifications at the residue level as well as within a residue are facilitated by the production of both glycosidic bond cleavage and cross-ring cleavage. Selection of activation method is an important influence in the abundance of cross-ring versus glycosidic bond cleavage. [19] Hence, tandem mass spectrometry spectra with complementary ion activation methods provide more complete information than a single activation method for determining the structure of GAGs.

The application of various electronic activation approaches has advanced the capability of tandem mass spectrometry for determining the structure of GAGs. Because CS and DS have such similar structures, varying only by stereochemistry of their uronic acids, or by 4-*O* versus 6-*O* sulfation, their tandem mass spectra share many common features. Finding statistically significant differences in their tandem mass spectra requires an unbiased approach, such as multivariate statistical methods.

For example, principal component analysis (PCA), has been used fine differences in the tandem mass spectra of small HS epimers.[26-28] CID and EDD have been shown to be capable of distinguishing CS and DS.[26, 27, 29] Diagnostic fragment ions to assign isomers have been established depending on the type of tandem mass spectrometry methods.[26, 27, 29] It was noted by Zaia et al. that CS and DS show distinct differences in their CID mass spectra, including cross-ring fragment ions, demonstrating great potential for providing details about the structure of GAG oligomers.[21] Later, Kailemia et al. introduced PCA to visualize the differences between the spectra of CS and DS, finding salient differences between the two.[29] PCA was also used by Agyekum *et al.* to assess differences in the tandem mass spectra of synthetic HS tetrasacharide diastereomers using EDD.[28] They developed a formula that used fragment ion abundances to assign the presence of IdoA versus GlcA, laying a critical foundation for later research on PCA applied to GAG stereochemistry. Later, PCA using EDD revealed that tetrasaccharides with GlcA at the non-reducing end had higher levels of $^{1,3}X_3$, $^{2,4}X_3$, Y_1, B_2, and $^{0,2}A_2$ ions. In contrast, tetrasaccharides containing IdoA at the non-reducing end had higher abundances of Z_3-H_2SO_4, B_3-SO_3, B_3, and Z_3.[26] Thus, indicative fragment ions can be identified, determining the hexuronic acid stereochemistry of GAG epimers whose tandem mass spectra are nearly indistinguishable by manual inspection.[27]

Here, we explore the application of PCA using a multimodal approach to tandem mass spectrometry analysis of CS/DS hexasaccharides, by combining data from more than one method of ion activation. We extend the capability of statistical approaches to distinguish GAG diastereomers to CID and NETD for the first time, validating the efficacy to use multivariate statistical analysis to characterize GAG epimerization. In addition, analysis of the variability in the dataset explained by each

component (principal components) provides insight into how PCA recognizes the C5-stereochemistry of GAGs and how fragment ions are correlated based on tandem mass spectra obtained from each activation method to profile GAG structures.

2.2 Experimental Method

Preparation of Chondroitin Sulfate Oligosaccharides

Chondroitin sulfate A oligosaccharide was prepared by partial enzymatic depolymerization of bovine trachea chondroitin sulfate A (Celsus Laboratories, Cincinnati, OH). Chondroitin sulfate B (dermatan sulfate) was collected from porcine intestinal mucosa dermatan sulfate (Celsus Laboratories, Cincinnati, OH). and digested by partial enzymatic depolymerization. A 20 mg/mL solution of each, in 50 mM Tris-HCl/60 mM sodium acetate buffer at pH 8 was incubated at 37°C with chondroitin ABC lyase from *Proteus vulgaris*, EC 4.2.2.4. (Seikagaku, Japan).

After the absorbance at 232 nm indicated the digestion was 50% completed, the digestion mixture was heated at 100°C for 3 min. An MWCO 5000 membrane was used for removing high-molecular-weight oligosaccharides and enzymes. In the following step, the oligosaccharide mixture was rotary evaporated and fractionated by low-pressure gas phase separation (GPS) using Bio-Gel P10 (Bio-Rad, Richmond, CA). Fractions of our oligosaccharide samples were desalted by GPC on a Bio-Gel P2 column and lyophilized. Additional purification was carried out using strong anion exchange high-pressure liquid chromatography (SAX-HPLC) on a semi-preparative SAX S5 Spherisorb column (Waters Corp, Milford, MA). The SAX-HPLC fractions containing > 90% of selected oligosaccharides were collected, desalted by GPC, and lyophilized. The solid sample was reconstituted in water and purified a second time

by SAX-HPLC. Only the top 30% of the chromatographic peak was obtained, desalted, and freeze-dried. Concentration of the oligosaccharide solutions was determined by measuring the absorbance at 232 nm (e = 3800 $M^{-1}cm^{-1}$). The resulting fractions containing individual oligosaccharides were characterized by PAGE, ESI-MS, and high-field nuclear magnetic resonance (NMR) spectroscopy.

Mass Spectrometry

Experiments were performed on a 9.4 T Bruker SolariX FTMS instrument (Bruker Daltonics, Bremen, Germany) equipped with an ESI source. The prepared sample were injected by the volume of 250 μL syringe at a rate of 120 μL/h and ionized by electrospray. The CS or DS solutions, both hexasaccharide and octasaccharide, were infused at a concentration of 0.1 mg/mL in the HPLC grade of 50:50 methanol:H_2O (Fisher Scientific, Hampton, NH). All samples were investigated in negative ion mode.

For all tandem mass spectrometry measurements, precursor ions were isolated by a quadrupole mass filter with a 1 Da isolation window that selected the complete isotopic envelope. For CID, precursor ions were dissociated using 7-10 V of collision energy in a RF-hexapole collision cell, and then transferred through a RF-quadrupole ion guide to a dynamically harmonized FTICR cell. For EDD experiments, precursor ions were accumulated for 1-2 seconds before being transferred to the analyzer cell of the FTICR. Here, ions underwent EDD by exposure to a 1.5 s pulse of energetic electrons produced by by a hollow cathode emitter. The hollow-cathode used a heater current of 1.5 A. The bias value of the cathode was set to -19.0 eV, which determines the energy of the electrons as they interact with trapped ions in the analyzer cell. The extraction lens was adjusted to -18.5 V to control the flux of electrons that enter the

ICR cell. Background spectra were collected by setting the cathode vias to 0 V to ensure that no electrons reached the cell. For NETD experiments, precursor ions were mass-selected by the external quadrupole mass filter and accumulated for 1 to 2 s in the hexapole collision cell before their reaction with the reagent ions. Reagent tune mode was used to confirm the generation of the fluoranthene radical cation which was used the electron acceptor reagent ion. The reagent radical cation was created in a chemical ionization (CI) source mounted on the side of an ion funnel, and then directed to the hexapole collision cell to react with pre-isolated analyte ions for 500 to 1000 ms. NETD products and residual precursor ion were then introduced into the analyzer cell through the RF-only quadrupole ion guide. For all tandem mass spectra, 24 transients were summed together for each mass spectrum.

Data Analysis and Principal Component Analysis

For each transient, 1M points were acquired, padded with one zero fill, and apodized using a sine bell window prior to a magnitude-mode Fourier transform. Internal calibration was performed using accurate reference values of glycosidic bond cleavage from the samples, resulting in mass accuracy of < 1 ppm. Fragment ions were assigned using Glycoworkbench annotation function. [30] All product ions have been annotated using Wolff and Amster's annotation scheme derived from the Domon and Costello nomenclature. [31] For this process, the molecular structure is drawn with dashed lines to indicate fragmentation. Product ions at reducing end or non-reducing end are indicated by a slanted line at the top or bottom of the dashed line, respectively. In addition to lines showing glycosidic or cross-ring cleavage, an empty circle indicates that fragmentation occurs with the loss of one sulfate group from a

sulfate half ester, whereas a black circle represents the loss of two or more sulfate groups. [27]

For each tandem mass spectrum, principal component analysis (PCA) was performed using PLS Toolbox (Eigenvector Research, Inc., Wenatchee, WA, USA). The abundances of all assigned fragment ions were normalized with respect to a total ion abundance in each spectrum. To generate an input data matrix the observation of each oligosaccharide diastereomer was represented in each row and the normalized abundance of an assigned fragment ion was reported into each column. A total of three to four tandem mass spectra were acquired for each oligosaccharide during an intra-day measurement. Three observations of each diastereomer were used for PCA for each of the three tandem mass spectrometry methods. All data were mean-centered and cross-validated. For combined analysis of all three tandem mass spectrometry methods, the data were mean-centered, log-scaled, and cross-validated.

2.3 Results

Distinguishing CS from DS hexassacharides using each activation method

In previous work, CID, EDD, and NETD using FTICR mass spectrometry have been shown to provide a set of fragment ions that can be used to characterize many of the important structural features of GAG samples. A fully ionized precursor ion, without any sodium/proton exchange was selected for this study to ensure complete deprotonation of sulfate groups on GAGs. Generally, the number of repeat units is equal to the most abundant charge state for CS/DS GAGs composed of a repeat of mono-sulfated disaccharides.

In fact, there has been a great need for determining the stereochemistry of uronic acid in GAG epimers that exhibit nearly identical tandem mass spectra in manual analysis as a complementary analysis to investigate structural information of GAGs. For example, Figure 4 shows how difficult to recognize two chondroitin sulfate isomeric oligosaccharides with a direct inspection on mass spectrum.

Here, a multivariate statistical approach has been applied to the tandem mass spectrometry data for hexasaccharide CSA and DS standards produced by enzymatic depolymerization. As shown in Figure 5, PCA of the data for two isomeric hexasaccharides demonstrates clear differences by each activation method. The first principal components of CID, EDD and NETD explained over 92% of total variance and NETD mode accounted for the largest percent variance explained by the first PC. Figure 6 depicts bioplot analyses to demonstrate explanation extents and loading information for the hexasacchrides. In the loading plot, each fragment ion is represented by its coefficient, allowing identification of the key variables that can be used to distinguish CS from DS to assign the stereochemistry of CS/DS as shown in Figure 7.

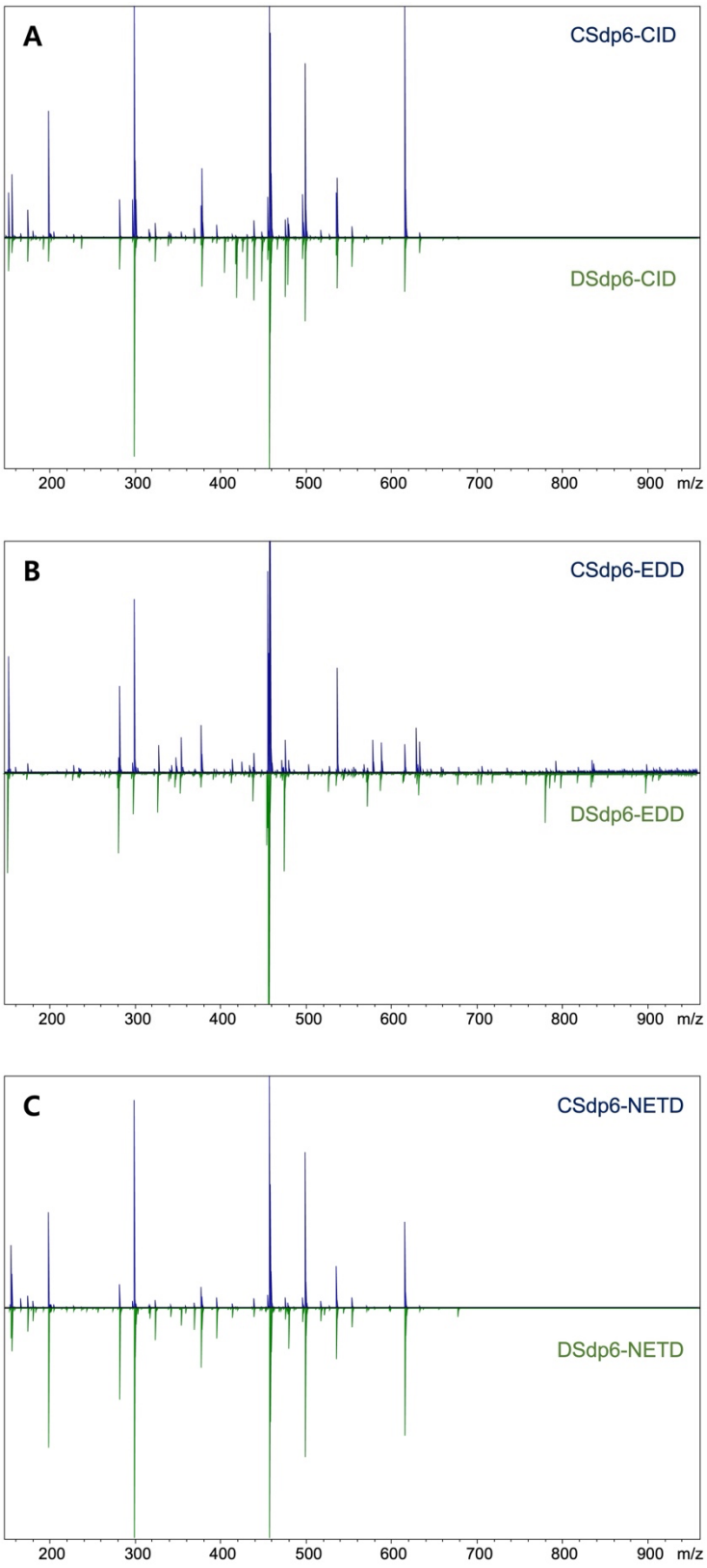

Figure 4. Comparison of the MS/MS for the $[M-3H]^{3-}$ precursor ion of the hexasaccharide epimer pair, CS and DS: (A) CID, (B) EDD, and (C) NETD.

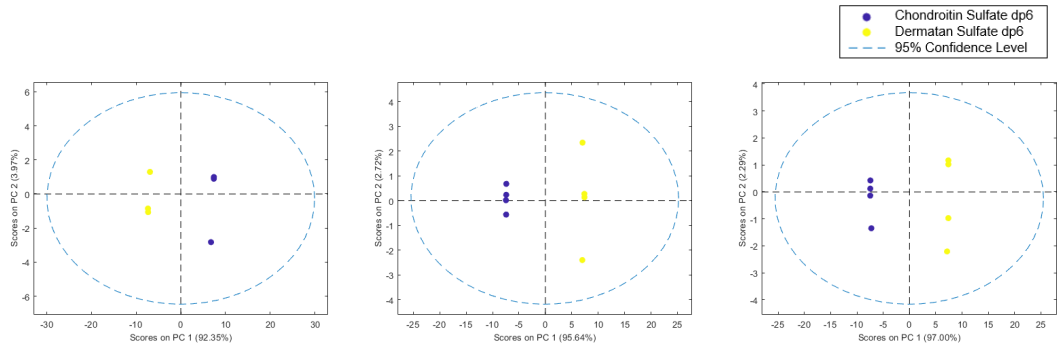

Figure 5. PCA results for CID (left), EDD (middle), and NETD (right) tandem mass spectra obtained from the precursor ion of triply-charged hexasaccharides, $[M-3H]^{3-}$, of CS and DS hexasaccharides. PC2 versus PC1 plots of scores distinguish two diastereomeric hexasaccharides, CSA and DS, with more than 95% of the total variance explained by two principal components: CID (92.35% and 3.97%), EDD (95.64% and 2.72%), and NETD (97.00% and 2.29%).

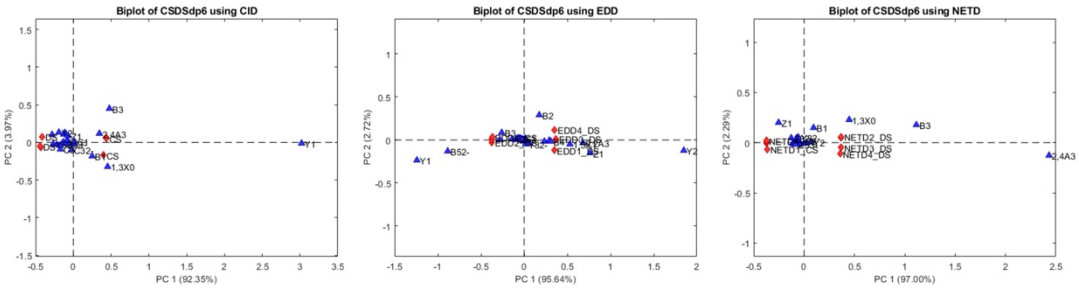

Figure 6. Biplot displays both the explanation extend of PCs and the loading information. Each plot clearly indicates that the two diasteromeric hexassacharides (red) are well separated into two classified groups accounting for a variety of fragment ions as shown in blue: CID (left), EDD (middle), and NETD (right).

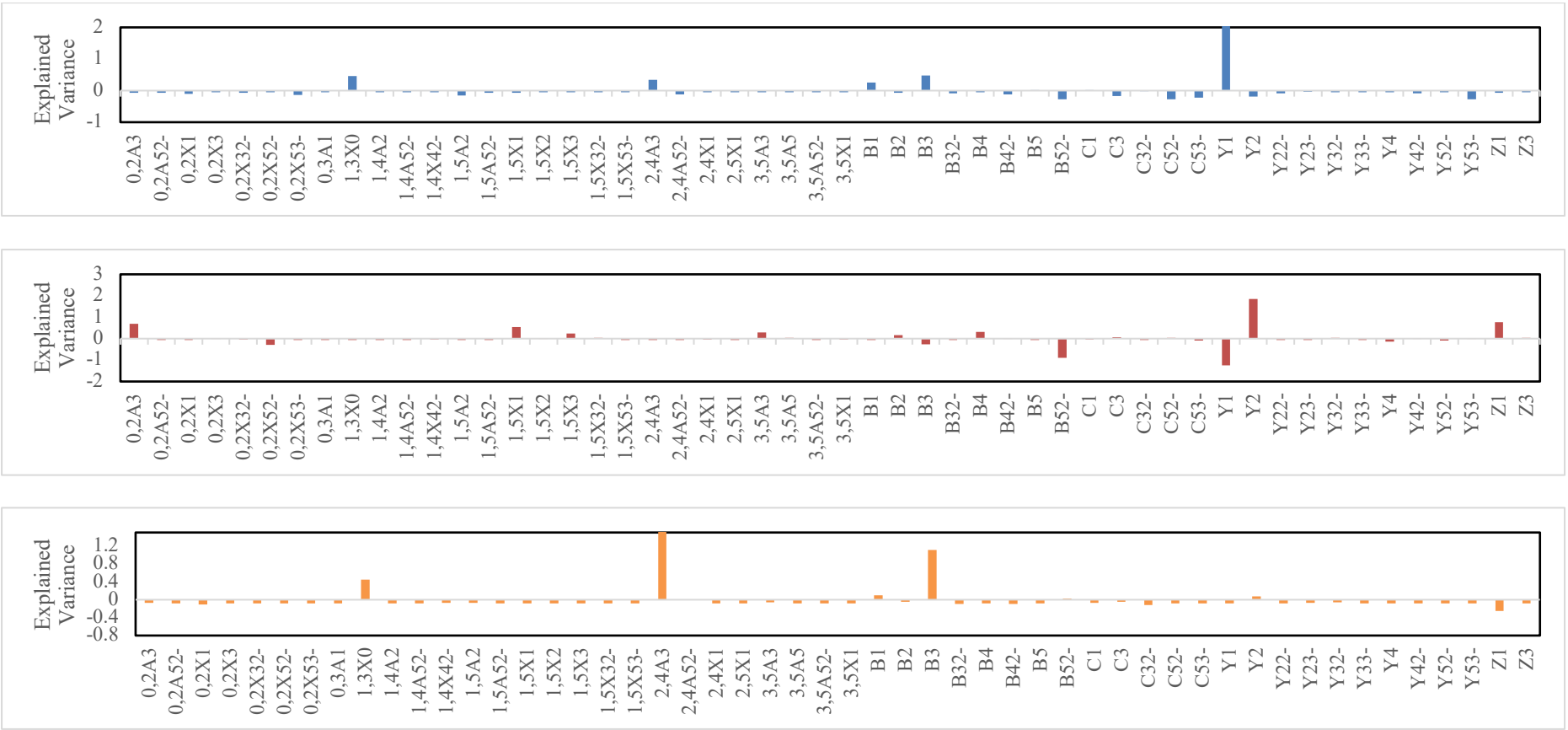

Figure 7. Comparison of each MS/MS analysis loadings (CID, EDD, and NETD from the top) from Figure 4. All loadings (fragment ions) are present on the x-axis and the magnitude for each variable loaded on the component is represented on the y-axis. Scale-exceeded explained variances were Y_1 in CID was 4.5 and $^{2,4}A_3$ was 2.4 in NETD.

A Simultaneous PCA of CS/DS hexassacharides using CID, EDD, and NETD

Product ions generated from the three types of tandem mass spectrometry methods were then statistically analyzed altogether by PCA. An overall analysis was conducted in which each observation was represented in a row and each product ion labeled with a corresponding tandem mass spectrometry was represented in a column. Taking into account intrinsic scales of intensities of product ions that vary from method to method, log transformation was necessary to allow for a broader distribution of variables. As a result, fragment ions caused by CID were prominent in PC1 and PC2, while NETD had significant horizontal variability but not vertical variation (Figure 8). The score plot shows the reasonable clustering of triplicates originating from the same diastereomer (either CS or DS) and CS replicates have clearly separated from DS replicates as seen in Figure 8. This empirical evidence confirmed that the separation between two clusters of diastereomers is larger than that of the diastereomers themselves with PC1.[27]

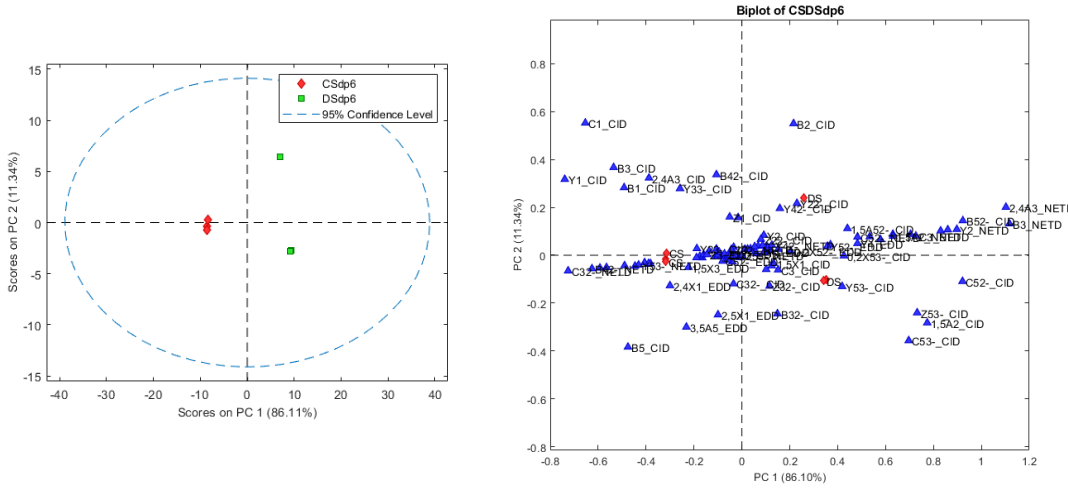

Figure 8. PCA results from the combination of all three tandem mass spectra of hexasaccharides of CS and DS. (left) 2D score plot of PC1 versus PC1. chondroitin sulfate was differentiated clearly from dermatan sulfate by PC1. (right) Biplot shows the score and the loading information in combination. These data clearly indicate that the two diasteromeric hexassacharides (red) are well separated into two classified groups accounting for a variety of fragment ions as shown in blue.

A biplot is an effective way to represent the sample and variables as a two-dimensional PCA projection, indicating certain relationships between observations and variables. [32] The relationship between PC1 and PC2 and their contribution on the loadings in two PCs were visualized in Figure 8. The first two principal components explained the majority of the data set variance greater than 97% by the summation of PC1 and PC2. In addition, three tandem mass spectrometry methods ensured the most information-rich product ions at once, providing a full picture of the fragmentation distribution. Points in the biplot that lie far from the origin indicate peaks that contribute the most to the differences in the spectra. The biplot shows not only the

fragments that distinguish CSA from DS, but also the type of ion activation that provides the most significant diagnostic ions.

Variables in the loading plot can contribute subsequent multivariate analysis to determine distinct fragment ion features to assign GAG diastereomers. [27] Figure 9 illustrates how much variance is explained by product ions through each tandem method. Comparison of these loading plots provide an interesting assessment of the way each activation method distinguishes CSA from DS. Interestingly, these data suggest that when combined, CID and NETD contribute equally to differences that distinguish CSA from DS, while the EDD data contribute to a less significant way. For CID, the majority of diagnostic fragments are from glycosidic bond cleavages. EDD exhibits far more diagnostically relevant cross-ring cleavages than CID, which is consistent with earlier results. [24, 26, 29, 33] For NETD, a large number of diagnostically relevant cross-ring cleavages with high variance are observed, which is in agreement with previous NETD results. [25, 27]

Some unexpected results from these data are the diagnostic relevance of fragments that are remote from the GlcA or IdoA residues, such as Y_5^{2-} (EDD) or Y_5^{3-} (NETD), which occur in the glycosidic bond between the non-reducing end residue, a D-uronic acid, and the adjacent amino sugar. Also, $^{0,2}X_5^{3-}$ (CID) and $^{0,2}X_5^{2-}$ (EDD) appear to be diagnostic of differences between CSA and DS, even though these fragments occur in the reducing end D-uronic acid which identical for both hexasaccharides. Less surprising are the larger number of diagnostic cross-ring fragments in the EDD and NETD mass spectra but absent from the CID data, such as $^{0,2}X_3^{1-}$ and $^{0,2}X_3^{2-}$ ions (EDD) or the $^{2,4}A_3^{1-}$ and $^{2,4}A_5^{2-}$ (NETD). Additional diagnostic X and A cross-ring cleavage ions in the NETD spectra include $^{1,4}X_4^{2-}$, $^{1,5}A_2^{1-}$, and $^{3,5}A_3^{1-}$. Interestingly, Figure 9 shows that the explained variations on the y-axis for

three different MS/MS methods were more distributed across fragment ions with fewer extreme diagnostic ion values than Figure 7 (loading of each MS/MS).

Distinguishing CS from DS octassacharides using MS/MS

Figure 10 illustrates that two chondroitin sulfate isomeric octasaccharides were indistinguishable with manual interpretation. PCA analysis of longer glycan chains, namely octasaccharides, has been carried out in a manner similar to that used in the analysis of hexasaccharides. In general, precursors arising via Na^+/H^+ exchange tend to form abundant fragment ions that are resistant to sulfo decomposition. [34] However, $[M-4H]^{4-}$ was utilized here as a precursor to discriminate epimers using tandem mass spectrometry since this was the most abundant charge state. PCA was used to cluster isomers based on each tandem mass spectrum data set. Sample triplicates and quadruplicates for both CS and DS were well separated as shown in Figure 11. The degrees of variance explained by PC1 were over 98% in EDD and NETD, while PC1 accounted for only 76.49% in CID. CID showed 16.60% of the second principal component, and the sum of the two principal components was 93.09%, which was less than the summed PCs from other tandem mass spectrometry methods (99.71% in EDD, 99.98% in NETD). Figure 12 depicts biplot analyses to demonstrate explanation extents and loading information for the octasacchrides.

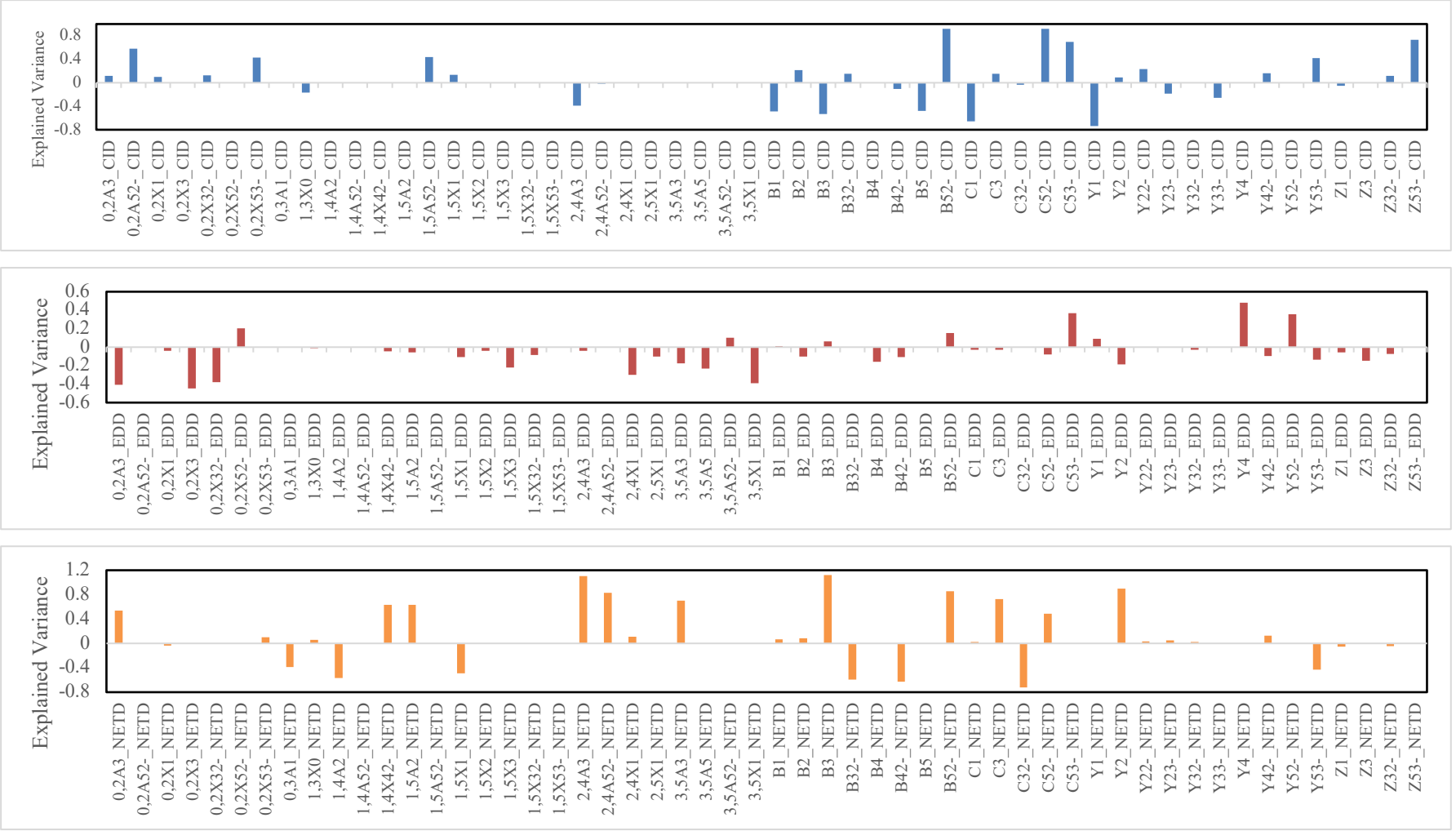

Figure 9. Comparison of each MS/MS analysis loadings (CID, EDD, and NETD from the top) from Figure 6. All loadings (fragment ions) are present on the x-axis and the magnitude for each variable loaded on the component is represented on the y-axis.

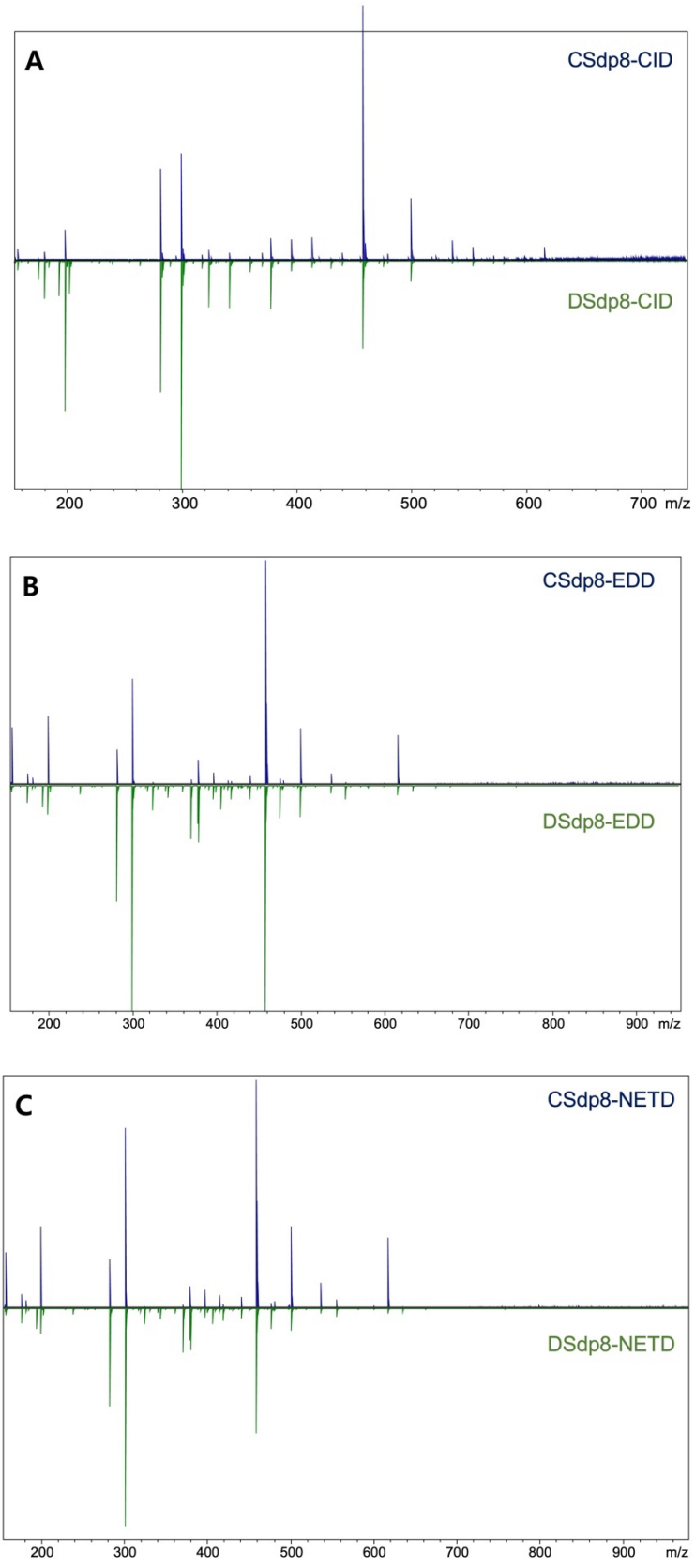

Figure 10. Comparison of the MS/MS for the $[M-4H]^{4-}$ precursor ion of the octasaccharide epimer pair, CS and DS: (A) CID, (B) EDD, and (C) NETD.

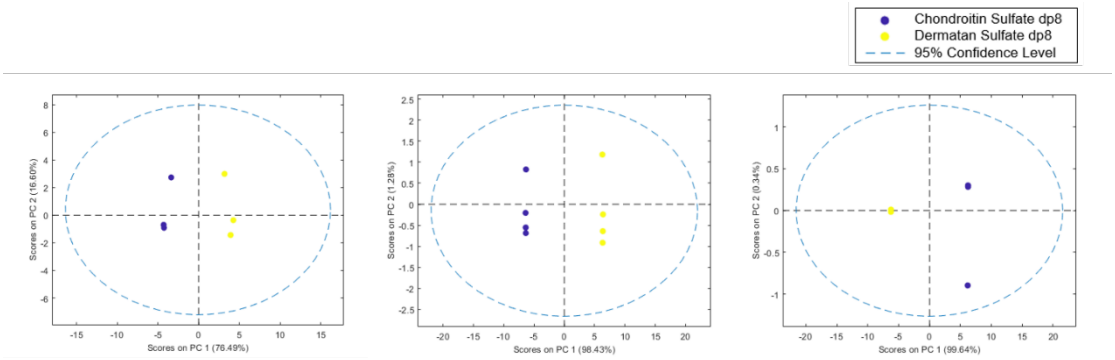

Figure 11. PCA results for CID (left), EDD (middle), and NETD (right) tandem mass spectra obtained from the precursor ion of $[M-4H]^{4-}$. PC2 versus PC1 plots of EDD and NETD scores distinguish two diastereomeric octasaccharides, CSA and DS, with more than 99% of the total variance explained by two principal components: EDD (98.43% and 1.28%), and NETD (99.64% and 0.34%). For CID, PC1 and PC2 explain 76.49% and 16.60% of the variance, respectively.

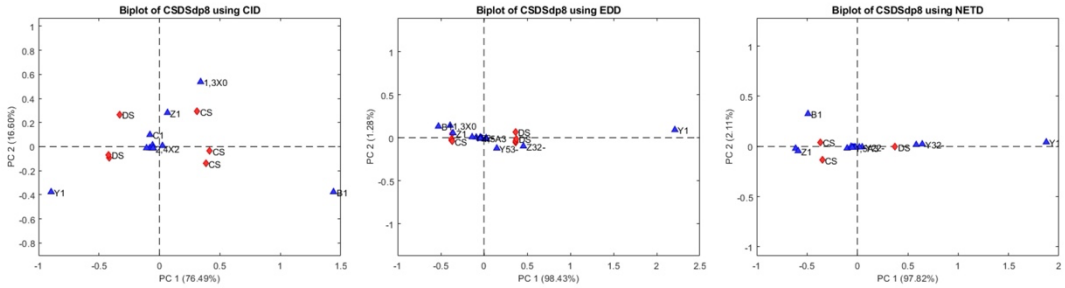

Figure 12. Biplot displays both the explanation extend of PCs and the loading information. Each plot clearly indicates that the two diasteromeric octaassacharides (red) are well separated into two classified groups accounting for a variety of fragment ions as shown in blue: CID (left), EDD (middle), and NETD (right).

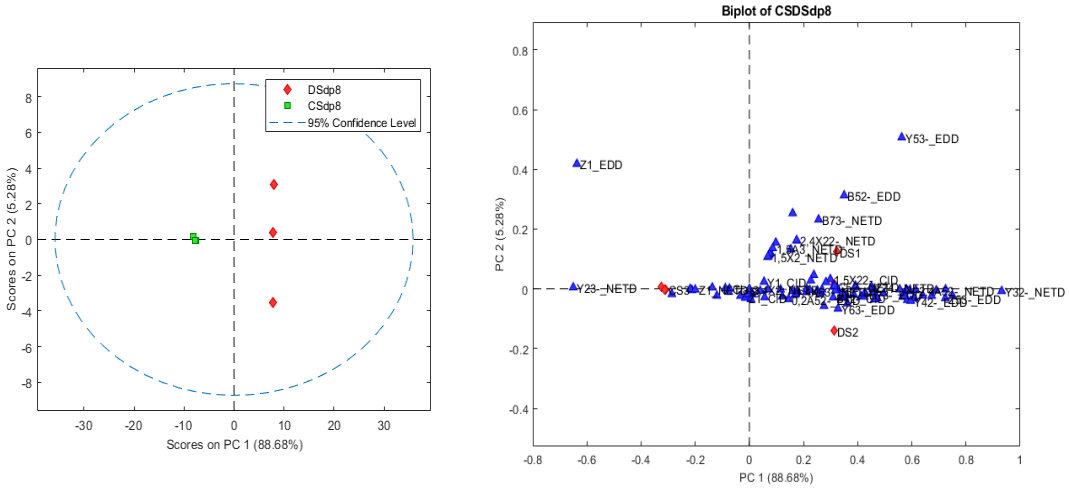

Figure13. PCA results from the combination of all three tandem mass spectra of octasaccharides of CSA and DS. (left) 2D score plot of PC1 versus PC1. Chondroitin sulfate was differentiated clearly from dermatan sulfate by PC1. (right) Biplot provides the information of the score and the loading. This demonstrates the two octasaccharides are well separated into two classified groups (red) accounted for a variety of fragment ions as shown in blue.

Simultaneous PCA of CS/DS octassacharides using CID, EDD, and NETD

As shown in Figure 13, PCA is capable of distinguishing CS from DS by allocating all fragmented ions derived from three tandem mass spectrometry methods simultaneously into one data set. Figure 13 (left) illustrates that the first two principal components, PC1 and PC2, explained 86% of the total variance. As a result of the difference between the two epimer groups along the PC1 axis, PC1 was able to discriminate stereoisomers, while PC2 yielded a span of CS and DS triplicates as ordinate. Figure 13 (right) shows the resulting biplot. Blue triangles are variables for the samples, while red diamonds point out the loadings. Each loading shows how strongly each fragment ion influences a PC (weights) to distinguish the epimers. [35] In

most cases, fragment ions were observed near the x-axis (PC1), with little dispersion along the y-axis (PC2). A number of fragments that can distinguish the CSA from DS can be found in this plot at points remote from the origin. DS samples showed more positive projections on the first PC, whereas CSA samples were plotted with more negative projections.

Shown in Figure 14, the B and Y product ions are found to be diagnostically relevant in CID, EDD, and NETD mass spectra. Similar to the PCA result of hexasaccharides, C and Z ions contributed more significantly to the differences between CSA and DS by EDD and NETD than by CID. Diagnostic cross-ring cleavages are far more prevalent by EDD and NETD than CID, as expected. Nevertheless, CID does display some diagnostic ^{0,2}A and ^{2,4}A fragment ions that contribute to distinguishing CSA from DS.

Our PCA results for CID, EDD, and NETD spectra revealed that more highly positive fragment projections were present than negative ones for CID, nevertheless the negative loadings still contributed to CS separation. Our findings are consistent with previous EDD and NETD results based on peak abundance analysis and serve to justify the chemometrics to epimer characterization.

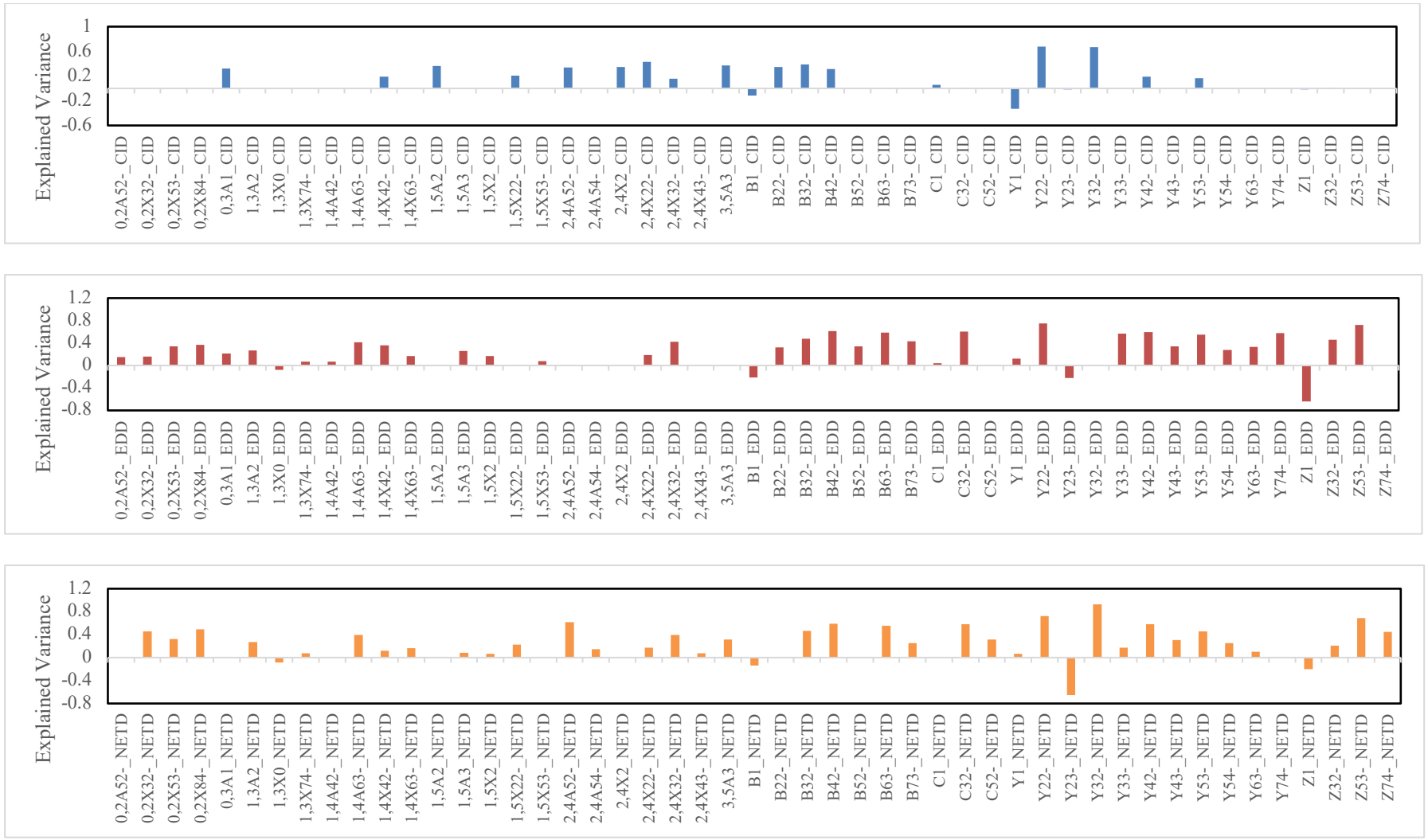

Figure 14. Comparison of each MS/MS analysis loadings (CID, EDD, and NETD from the top) from Figure13. All loadings as fragment ions are shown on the x-axis and the magnitude for each variable loaded on the component are present on the y-axis.

2.4 Conclusions

GAGs have complex structures, which necessitates sophisticated analytical approaches for their characterization. Considerable research effort has been directed at examining a variety of ion activation methods to improve the quality of the resulting tandem mass spectrometry data. Until now, no attempts have been made to combine various types of dissociation statistically in order to determine the stereochemistry of GAGs, specifically for CS and DS. In terms of structural similarity, CSA and DS GAGs differ only in the stereochemistry of the uronic acid in the repeating disaccharides. Multivariate statistical analysis has been applied hereto distinguish CSA from DS using two different lengths of GAG oligomers, specially dp6 and dp8. As the length of GAGs became longer, manual interpretation becomes more challenging, particularly for EDD and NETD which provide rich fragmentation spectra. Multivariate statistical analysis makes the assignment of important structural features more tractable.

PCA analysis of individual and combined data sets show that all three types of ion activation (CID, EDD, NETD) provide diagnostically significant fragment ions that can distinguish CSA from DS, both for dp6 and dp8 oligomers. In addition, EDD and NETD have the added benefit of producing informative fragment ions to assign the locations of sulfate modification. CID activation does not sufficiently distinguish GlcA from IdoA in HS oligomers, it appears to provide diagnostic ions for this assignment for CS/DS. The other two ion activation methods, EDD and NETD could be complementary to CID in differentiating CS from DS. Furthermore, a PCA analysis of combined data from three ion activation methods simultaneously provided the possibility of not only how the results of each tandem mass spectrometry analysis

are associated with, but also that of global fragment ions being present. More cross-ring cleavage ions were abundant in this case; however, glycosidic bond cleavage ions were able to still influence to determine an isomer of either CS or DS.

The assessment of hybrid/longer GAG chains is still under development as our dataset was limited up to octasaccharides of chondroitin sulfate species that were uniform in their uronic acid composition (all IdoA or all GlcA). Future efforts should be made with oligomers that contain a mix of uronic acids, to see if these can be assigned on a residue by residue basis.

Combining tandem mass spectrometry with multivariate statistical analysis identified informative fragment ions that contribute significantly to identifying epimeric uronic acids based on tandem mass spectral data, thereby enhancing our understanding of the relationship between variables and tandem mass spectrometry. Based on this method, subtle changes in stereochemistry could be detected in biological GAG samples with the potential to identify improper GAG structures present as a result of disease states. These results provide preliminary evidence that fragment ions reflect on the characteristics of each tandem mass spectrometry method and are beneficial to diagnose subtle stereochemical differences present in the glycan chains of proteoglycans.

REFERENCES

(1) Esko, J. D.; Kimata, K.; Lindahl, U. Proteoglycans and sulfated glycosaminoglycans. In *Essentials of Glycobiology*, 3rd ed.; Cold Spring Harbor Laboratory Press, 2017.

(2) Linhardt, R. J.; Toida, T. Role of glycosaminoglycans in cellular communication. *Accounts of Chemical Research* **2004**, *37* (7), 431-438.

(3) Kamhi, E.; Joo, E. J.; Dordick, J. S.; Linhardt, R. J. Glycosaminoglycans in infectious disease. *Biological Reviews* **2013**, *88* (4), 928-943.

(4) Hitchcock, A. M.; Yates, K. E.; Costello, C. E.; Zaia, J. Comparative glycomics of connective tissue glycosaminoglycans. *Journal of Proteomics* **2008**, *8* (7), 1384-1397.

(5) Sasisekharan, R.; Shriver, Z.; Venkataraman, G.; Narayanasami, U. Roles of heparan-sulphate glycosaminoglycans in cancer. *Nature Reviews Cancer* **2002**, *2* (7), 521-528.

(6) Mikami, T.; Kitagawa, H. Biosynthesis and function of chondroitin sulfate. *Biochimica et Biophysica Acta* **2013**, *1830* (10), 4719-4733.

(7) Basu, A.; Patel, N. G.; Nicholson, E. D.; Weiss, R. J. Spatiotemporal diversity and regulation of glycosaminoglycans in cell homeostasis and human disease. *American Journal of Physiology-Cell Physiology* **2022**, *322* (5), C849-C864.

(8) Li, F.; Gerdy, B.; Murugan, S.; Yamada, S.; Hashiguchi, T.; Mizumoto, S.; Oguri, K.; Okayama, M.; van Kuppevelt, T. H.; Sugahara, K. Involvement of highly sulfated chondroitin sulfate in the metastasis of the Lewis lung carcinoma cells. J*ournal of Biological Chemistry* **2008**, *283* (49), 34294-34304.

(9) Afratis, N.; Gialeli, C.; Nikitovic, D.; Tsegenidis, T.; Karousou, E.; Theocharis, A. D.; Pavão, M. S.; Tzanakakis, G. N.; Karamanos, N. K. Glycosaminoglycans: key players in cancer cell biology and treatment. *The FEBS journal* **2012**, *279* (7), 1177-1197.

(10) 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.

(11) Bülow, H. E.; Hobert, O. The molecular diversity of glycosaminoglycans shapes animal development. *Annual Review of Cell and Developmental Biology* **2006**, *22*, 375-407.

(12) Zaia, J. Glycosaminoglycan glycomics using mass spectrometry. *Molecular and Cellular Proteomics* **2013**, *12* (4), 885-892.

(13) Wang, W.; Han, N.; Li, R.; Han, W.; Zhang, X.; Li, F. Supercharged fluorescent protein as a versatile probe for the detection of glycosaminoglycans in vitro and in vivo. *Analytical Chemistry* **2015**, *87* (18), 9302-9307.

(14) Calabro, A.; Midura, R.; Wang, A.; West, L.; Plaas, A.; Hascall, V. Fluorophore-assisted carbohydrate electrophoresis (FACE) of glycosaminoglycans. *Osteoarthritis and Cartilage* **2001**, *9*, S16-S22.

(15) Nyren-Erickson, E. K.; Haldar, M. K.; Gu, Y.; Qian, S. Y.; Friesner, D. L.; Mallik, S. Fluorescent liposomes for differential interactions with glycosaminoglycans. *Analytical Chemistry* **2011**, *83* (15), 5989-5995.

(16) Auray-Blais, C.; Bhérer, P.; Gagnon, R.; Young, S. P.; Zhang, H. H.; An, Y.; Clarke,

J. T.; Millington, D. S. J. M. g.; metabolism. Efficient analysis of urinary glycosaminoglycans by LC-MS/MS in mucopolysaccharidoses type I, II and VI. *Molecular genetics and metabolism* **2011**, *102* (1), 49-56.

(17) Huang, R.; Pomin, V. H.; Sharp, J. S. LC-MSn Analysis of Isomeric Chondroitin Sulfate Oligosaccharides Using a Chemical Derivatization Strategy. *Journal of the American Society for Mass Spectrometry* **2011**, *22* (9).

(18) Miller, R. L.; Guimond, S. E.; Schwörer, R.; Zubkova, O. V.; Tyler, P. C.; Xu, Y.; Liu, J.; Chopra, P.; Boons, G.-J.; Grabarics, M.; et al. Shotgun ion mobility mass spectrometry sequencing of heparan sulfate saccharides. *Nature Communications* **2020**, *11* (1), 1481.

(19) Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J. Developments in mass spectrometry for glycosaminoglycan analysis: A review. *Molecular & Cellular Proteomics* **2021**, *20*.

(20) Chi, L. L.; Wolff, J. J.; Laremore, T. N.; Restaino, O. F.; Xie, J.; Schiraldi, C.; Toida, T.; Amster, I. J.; Linhardt, R. J. Structural analysis of bikunin glycosaminoglycan. *Journal of the American Chemical Society* **2008**, *130* (8), 2617-2625.

(21) Zaia, J.; Costello, C. E. Tandem mass spectrometry of sulfated heparin-like glycosaminoglycan oligosaccharides. *Analytical Chemistry* **2003**, *75* (10), 2445-2455.

(22) Saad, O. M.; Leary, Delineating mechanisms of dissociation for isomeric heparin disaccharides using isotope labeling and ion trap tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2004**, *15* (9), 1274-1286.

(23) Saad, O. M.; Leary, J. A. Heparin Sequencing Using Enzymatic Digestion and ESI-MSn with HOST: A Heparin/HS Oligosaccharide Sequencing Tool. *Analytical*

Chemistry **2005**, *77* (18), 5902-5911.

(24) Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. Electron detachment dissociation of glycosaminoglycan tetrasaccharides. *Journal of the American Society for Mass Spectrometry* **2007**, *18* (2), 234-244.

(25) Wolff, J. J.; Leach III, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. Negative electron transfer dissociation of glycosaminoglycans. *Analytical Chemistry* **2010**, *82* (9), 3460-3466.

(26) Oh, H. B.; Leach III, F. E.; Arungundram, S.; Al-Mafraji, K.; Venot, A.; Boons, G.-J.; Amster, I. J. Multivariate analysis of electron detachment dissociation and infrared multiphoton dissociation mass spectra of heparan sulfate tetrasaccharides differing only in hexuronic acid stereochemistry. *Journal of the American Society for Mass Spectrometry* **2011**, *22* (3), 582-590.

(27) Leach III, F. E.; Ly, M.; Laremore, T. N.; Wolff, J. J.; Perlow, J.; Linhardt, R. J.; Amster, I. J. Hexuronic acid stereochemistry determination in chondroitin sulfate glycosaminoglycan oligosaccharides by electron detachment dissociation. *Journal of the American Society for Mass Spectrometry* **2012**, *23* (9), 1488-1497.

(28) Agyekum, I.; Patel, A. B.; Zong, C.; Boons, G.-J.; Amster, I. J. Assignment of hexuronic acid stereochemistry in synthetic heparan sulfate tetrasaccharides with 2-*O*-sulfo uronic acids using electron detachment dissociation. *International journal of mass spectrometry* **2015**, *390*, 163-169.

(29) Kailemia, M. J.; Patel, A. B.; Johnson, D. T.; Li, L.; Linhardt, R. J.; Amster, I. J. Differentiating chondroitin sulfate glycosaminoglycans using collision-induced dissociation; uronic acid cross-ring diagnostic fragments in a single stage of tandem mass spectrometry. *European Journal of Mass Spectrometry* **2015**, *21* (3), 275-285.

(30) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *Journal of proteome research* **2008**, *7* (4), 1650-1659.

(31) Domon, B.; Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal* **1988**, *5* (4), 397-409.

(32) Sleighter, R. L.; Liu, Z.; Xue, J.; Hatcher, P. G. Multivariate statistical approaches for the characterization of dissolved organic matter analyzed by ultrahigh resolution mass spectrometry. *Environmental Science & Technology* **2010**, *44* (19), 7576-7582.

(33) Pepi, L. E.; Sasiene, Z. J.; Mendis, P. M.; Jackson, G. P.; Amster, I. J. Structural characterization of sulfated glycosaminoglycans using charge-transfer dissociation. *Journal of the American Society for Mass Spectrometry* **2020**, *31* (10), 2143-2153.

(34) Wolff, J. J.; Laremore, T. N.; Busch, A. M.; Linhardt, R. J.; Amster, I. J. Influence of charge state and sodium cationization on the electron detachment dissociation and infrared multiphoton dissociation of glycosaminoglycan oligosaccharides. *Journal of the American Society for Mass Spectrometry* **2008**, *19* (6), 790-798.

(35) Ringnér, M. What is principal component analysis? *Nature Biotechnology* **2008**, *26* (3), 303-304.

CHPATER 3

OPTIMIZATION AND VALIDATION OF A CAPILLARY ZONE ELECTROPHORESIS MASS SPECTROMETRY METHOD FOR GLYCOSAMINOGLYCANS ANALYSIS

3.1 Introduction

Glycosaminoglycans (GAGs) play important roles in modulating biological functions in most mammalian cells. Their structures are heterogenous and have a variety of modifications to their repeating disaccharide structures which influence cellular function. [1, 2] GAGs play a role with several pathophysiological conditions, including inflammation, angiogenesis, osteoarthritis, cardiovascular disorders, viral infection, and cancer. [3-6] The characterization of GAGs has been the subject of numerous analytical developments over the years, allowing researchers to gain a better understanding of the functions of GAG molecules. [2, 7-9] Among them, capillary zone electrophoresis-mass spectrometry (CZE-MS) allows for GAG analysis in a fast, accurate, and reliable manner. [10-12]

In this chapter, we investigate several experimental factors that influence the effectiveness of CZE-MS, including the length of a capillary column, the inner capillary coating type, the etching of the capillary outlet, the composition of the background electrolyte (BGE), the distance between an MS inlet and a spray tip, and the reproducibility for both inter and intraday measurements. The goal of this work is to optimize CZE-MS for the analysis of GAG samples using standards to gain insight into the performance, versatility, and future potential of this technique.

3.2 Experimental method

Sample Preparation

CHO cells were grown in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% (v/v) FBS and 1% (v/v) penicillin/streptomycin at 37 °C under an atmosphere of 5% CO2/95% air. Each cell was subcultured every three to four days and was revived from liquid nitrogen after ten passages. The cells were seeded at 0.5 x 106 cells per ml in a 10 cm plate and harvested when confluent. In brief, cells were washed in PBS, lifted with trypsin (Gibco), and digested with Pronase (0.5 mg/ml, Sigma) overnight at 37°C. To desalt the product, a DEAE-Sephacel (Cytiva) column was equilibrated in 50 mM sodium acetate buffer, pH 6.0, containing 200 mM NaCl, and then passed through a PD-10 column (Cytiva). DNase was then applied to the desalted product, which was then passed once again through the DEAE-Sephacel and PD-10 columns. Lyophilized GAGs were incubated with two mU each of the heparin lyases I, II, and III for 16 hours at 37 °C in a buffer containing 40 mM ammonium acetate and 3.3 mM calcium acetate, pH 7. The desalting process of HS samples for the use of CZE-MS was conducted by the 0.5mL 3 K Amicon Ultra spin filters. HPLC-grade water was used to dissolve the dry HS samples. The dissolved samples were centrifuged three times at 14,000 g for 25 min while being filtered through cellulose.[13] DSdp10 standard samples were gifted from Dr. Linhardt and Dr. Zhang. Enoxaparin was purchased from USP (Rockville, MD) and were identically prepared by the method in Chapter 2, but different concentration (0.01 mg/mL in water). All GAG samples were transferred to a 250 μL CE auto sampler vial capped with a pre-cut septa cap.

Preparation of Functionalized Capillary

Functionalized capillaries were prepared from polyimide-coated bare fused silica capillaries with a 363 µm outer diameter (OD) and a 50 µm inner diameter (ID) (Polymicro Technologies). Linear polyacrylamide coating (LPA) and dimethylsiloxane (DMS) functionalization and the HF etching used procedures that were reported previously. [14] A 130 cm long functionalized capillary was cut into two columns after etching the middle section with HF, Figure 15A, to maintain the uniformity of internal coating. The etched end of the capillary yielded a tapered end with an outer diameter at the outlet of <100 µm. This tapered end is placed inside of a glass capillary nanospray tip to provide a sheath flow interface for CZE-MS. In order to obtain even outer diameters, two-step etching was employed rather than static etching (Figure 15).

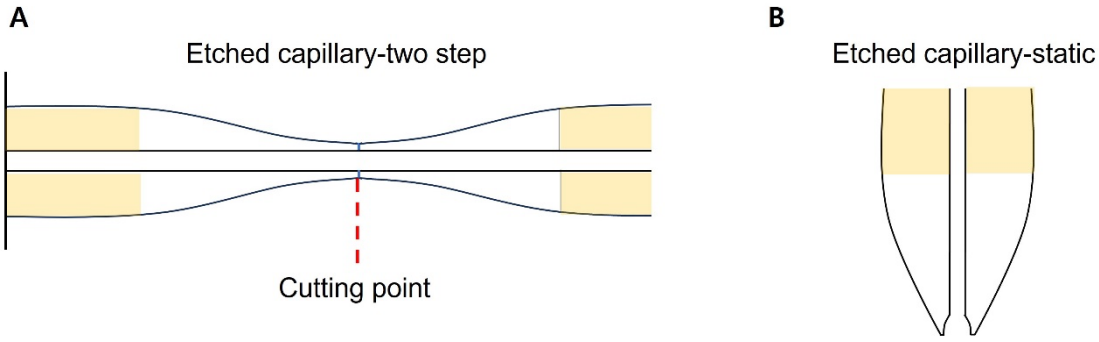

Figure 15. A schematic of HF etching. (A) As a result of two-step HF etching, the glass window in the middle of a capillary is affected. HF solution cannot be absorbed into the capillaries, resulting in unaffected inner diameters and improved migration. (B) Static etching is the process of removing the outer coating of a capillary. A capillary end immersed in HF solution causes corrosion of the inner capillary wall, which results in a poor resolution of migration peaks.

CZE-FT-ICR MS Analysis

Capillary zone electrophoresis (CZE) was performed using on a CMP ECE-001 CZE system (CMP scientific, Brooklyn, NY). The CZE was interfaced to the mass spectrometer with an electrokinetically pumped sheath flow CE-MS interface (EMASS-II interface, CMP Scientific). Mass spectra were collected in negative mode on Bruker 9.4 T SolariX FTICR mass spectrometer (Bruker Daltonics, Bremen, Germany). Ion accumulation time was set to 0.3 s with 1 M data points and the time of flight (TOF) was set to 0.8 ms. The flow rate of the drying gas was set to 2.0 L/min at 180 C. The inlet capillary voltage of the FT-ICR was set to 0 V.

CZE separations were performed on fused silica capillaries (60 cm x 360 μm OD x 50 μm ID) functionalized with DMS and LPA neutral coated capillary. 25 mM Ammonium formate in 70% (v/v) methanol/water was used as a sheath liquid (SL) and a background electrolyte (BGE). The etched end of both functionalized capillaries was positioned 0.5 mm from the tip of a borosilicate glass emitter orifice (0.75 mm ID, 5.0 cm length and 20 μm opening diameter of tip). The distance between the emitter opening and the inlet of MS was ca. 2.5 mm. All experiments were conducted at -2.0 kV external high voltage power supply along with the EMASS-II ion source.

Total CZE-MS running time was set to 60 min for GAG identification of GAG samples. Injection of samples was carried out using a pressure of 400 mbar for 10 s, resulting in circa 115 nL volume and 10.6% of the total capillary volume. The capillary was completely rinsed with fresh BGE after each run to remove residual carryover for the next run. Injection of samples was carried out using a pressure of 400 mbar for 10 s, resulting in circa 115 nL injected volume which occupies approximately 10% of the total capillary volume.

Experiments were carried out on a 9.4 T Bruker SolariX FTMS apparatus equipped with an ESI source (Bruker Daltonics, Bremen, Germany). For auto MS/MS acquisitions using collisionally-induced dissociation (CID), both preferential and exclusion lists were implemented. An MS1 scan was acquired first followed by three MS/MS scans using the external ion accumulation time of 0.5 s. The collision voltage was fixed between 13 and 15 V for each mass spectrum. Mass spectra were analyzed using Compass Data Analysis v4.1 software (Bruker Scientific, Bremen, Germany), MZmine3[44], in-house software developed in MATLAB. MetaboAnalyst 5.04 was used to generate PCA plots. [45]

Data processing

Raw files were converted to mzML format using DataAnalysis. The converted data files were uploaded to Mzmine3 for initial data processing and mass feature identification. Mzmine3 is a software suite designed to process batches of metabolomics datasets collected using LC-MS. Therefore, many of the data processing modules mention "chromatogram" which is the main graph generated from LC-MS data, but these can also be applied to CZE-MS datasets, however, the traces that show CZE separations are referred to as "electropherograms." These two terms may be used interchangeably in this context.

To obtain an accurate mass-to-charge ratio, mass tolerance used for data extraction was defined as 10 parts per million mass resolution. The noise threshold was set to 2.0×10^5 and only peaks above this threshold were selected for further processing. [46] Next, the ADAP (Automated Data Analysis Pipeline) chromatogram builder, one of the LC-MS feature detection algorithms supported by Mzmine 3, was used a minimum absolute peak height set to 6.0×10^5 and a minimum number of scan

span set to 5 to be recognized as an electropherogram. Electropherogram deconvolution was then performed with the local minimum resolver to remove noisy features and separate features with the same m/z but different migration times. The chromatographic threshold of 10% and a minimum peak top/edge ratio of 1.5. A minimum absolute intensity and a maximum peak duration range were established with 6.0×10^5 and 10 min, respectively. Isotopic peaks grouper algorithm was chosen to deisotope features with a retention time tolerance of 0.1 min. The allowed maximum charge was four, and charge detection was performed while keeping only the lowest isotope, which is the monoisotopic peak. Columns m/z, intensity, migration time, and charge were exported as features for each dataset.

3.3 Results

Coating

In order to achieve the MS and MS/MS characterization of complex GAGs with minimal runs, an appropriate separation window and a sufficient detection time for identifying GAGs corresponding to each electropherogram peak must be established. [15] The separation window is a function of both the electroosmotic flow (EOF) and the electrophoretic flow (EF). For capillaries with a DMS neutral coating, the EOF is small or non-existent, and the migration time is determined principally by the EF. This stands in contrast to a bare-fused silica capillary, for which the EOF and the EF oppose each other leading to long migration times. Another capillary preparation to produce a neutral inner surface with little or no EOF uses linear polyacrylamide (LPA). Early work with LPA coated capillaries were performed in normal mode (with a positive voltage at the inlet). This enabled the separation of

CZE to be slowed down, and the separation window to be increased. [10, 11] For example, Sun and coworkers demonstrated the robust CZE-MS platform with an LPA-coated capillary that increases the separation capacity in the field of proteomics. [15, 16] However, for reverse mode CZE (negative voltage at the inlet), reducing the EOF actually speeds up the migration time of anionic analytes. Therefore, it is necessary to take EOF into account, which results in a compromise between the resolution, the analysis time (migration time) required, and the robustness of CZE-MS system. [17] Figure 16 shows migration times and separation capacities for two types of coatings for Enoxaparin, one of the most common and complex low molecular weight heparin pharmaceuticals. Compared to the bottom set of electropherograms of the BFS capillary, the LPA functionalized capillary had a shorter migration time and narrower peak widths.

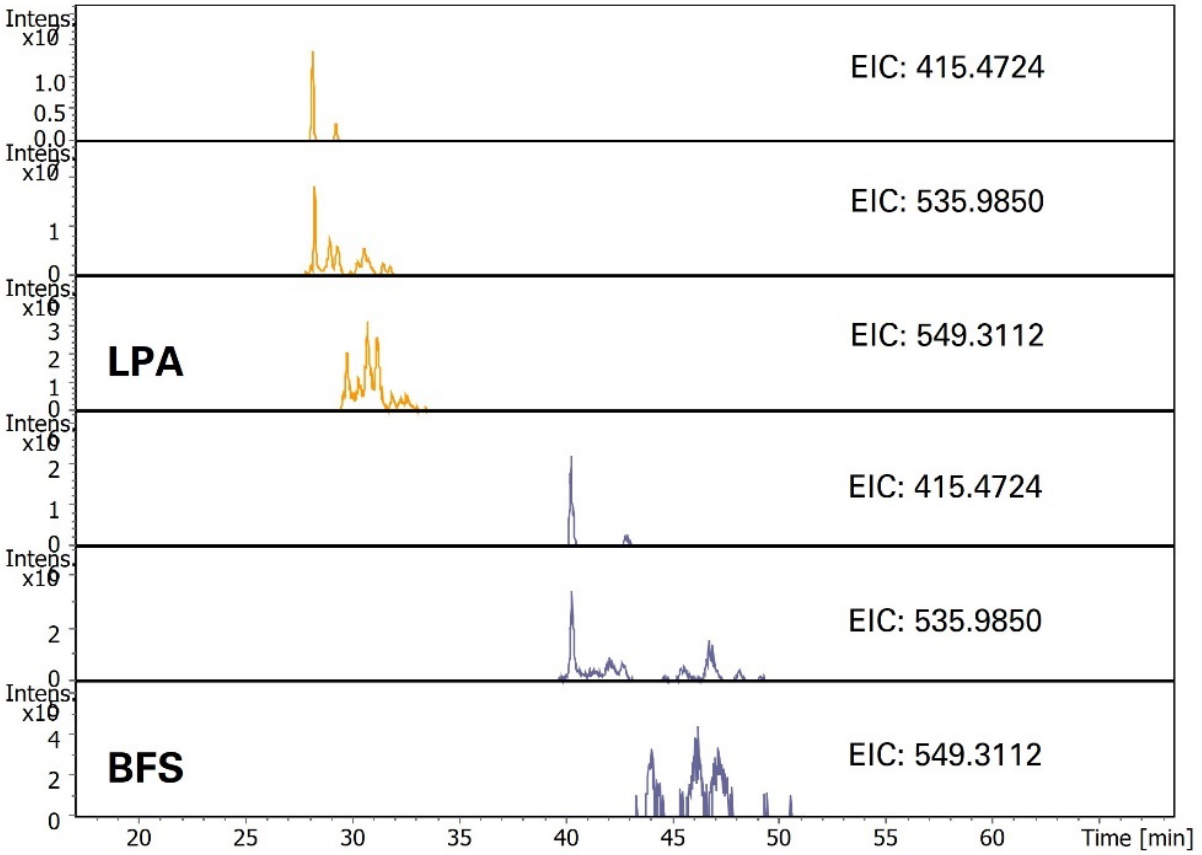

Figure 16. Extracted ion electropherograms of Enoxaparin major components. A neutral linear polyacrylamide (LPA) coating and anion bare fused silica (BFS) exhibited different migration times and peak widths.

Length

Migration times are expected to be directly proportional to the square of the capillary length, for a fixed applied voltage and constant inner diameter. LPA-functioned capillaries of different lengths, with inner diameters of 50 μm, were prepared and tested in reverse polarity of CZE operated in negative mode to analyze an Enoxaparin sample. Enoxaparin is a pharmaceutical product from depolymerization of heparin, and consists of a complex mixture of oligosaccharides ranging from dp 3 to dp 20. Figure 17 illustrates an electropherogram showing migration of the analyte, with a base peak electropherogram and corresponding extracted ion electropherogram for lengths ranging from 50 to 62 cm. Accordingly, a length of 60 cm was chosen as the optimal compromise of adequate separation and a reasonable analysis time.

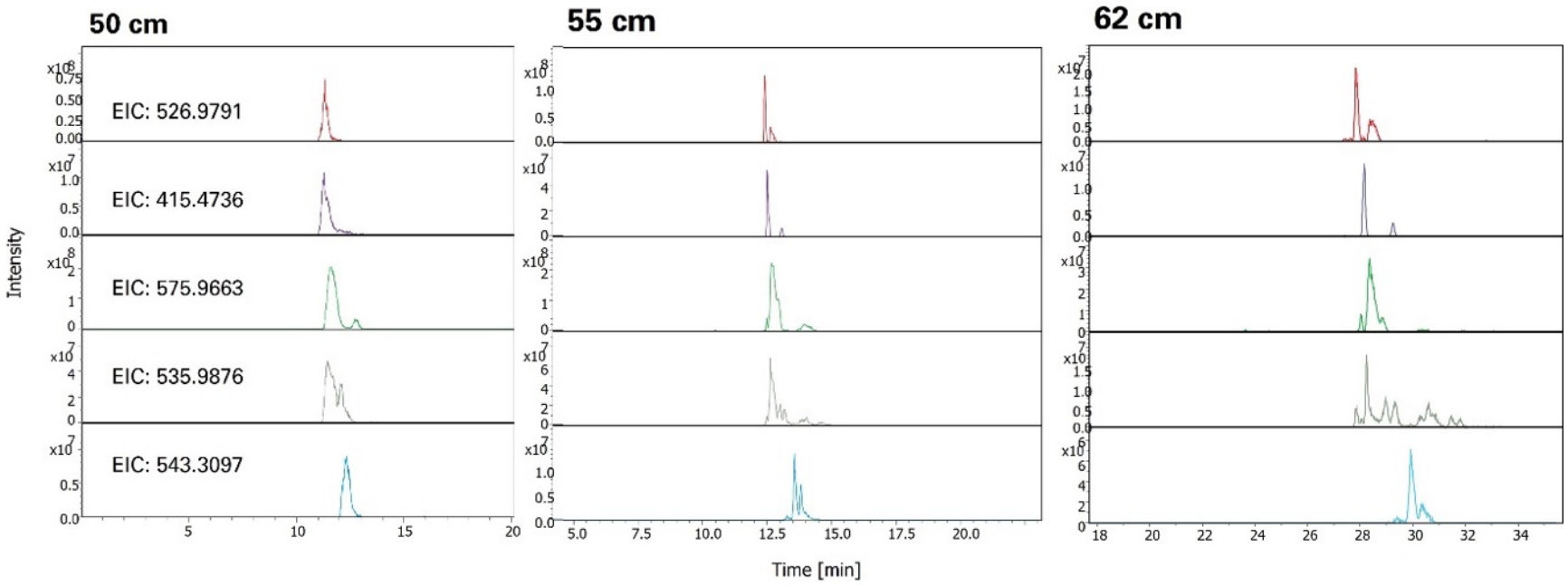

Figure 17. The extracted ion electropherograms (EIE) of main components of Enoxaparin using LPA-coated capillaries in different lengths. The colors in the figure indicate the same species of extracted ions.

Etching

The outer diameter of BFS capillaries was chemically etched by hydrofluoric acid solution (HF, 48% - 51% in water) at one end for use at the electro-kinetically pumped sheath flow interface. With this option, the distance between the end of the etched capillary and the spray tip orifice in the ESI spray tip greatly influences CZE-MS sensitivity, resulting in low flow rates on the nanoliter per minute scale and higher sensitivity with a shorter distance. Moreover, a sophisticated etched capillary reduces the mixing volume between the sheath liquid flow and the sample flow, thus preventing sample dilution in CE-MS interfaces. [18]

To etch the capillary, 20.0 mm of polyimide coating was removed from the middle of a 130 cm length of capillary, using the flame of an ordinary lighter to expose the bare fused silica. It is important to leave the capillary coating on either side of the bare fused silica. A small hole matching the size of the capillary was drilled into the bottom of an Eppendorf tube. The capillary was threaded through the hole until the clear window portion of the capillary approached the bottom of the capillary, yielding a watertight junction. HF solution was added to the Eppendorf tube to etch the outside of the capillary uniformly for 90 min. The capillary column was then washed with water to remove any residual HF. Cleaving stones were utilized to cut the etched part of the capillary at the middle of the narrow portion. Figure 18 shows a cleaved chemically etched capillary as observed by a scanning electron microscope.

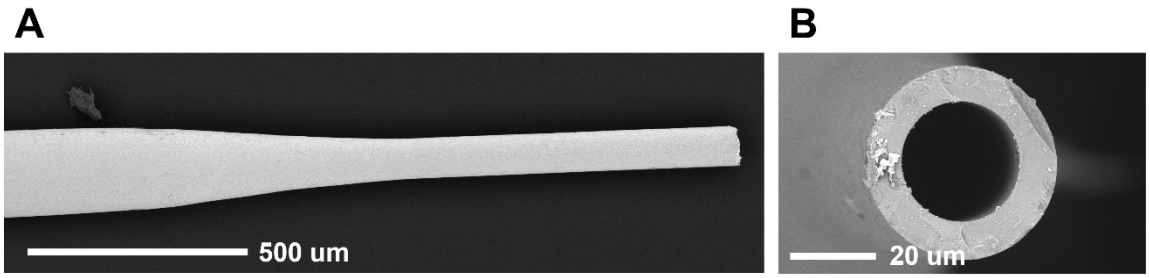

Figure 18. (A) Scanning electron microscopy (SEM) micrograph showing etched outer diameter of a BFS capillary for use in the sheath flow CE interface. (B) The open end of the capillary having a 50 μm aperture by our chemical etching process, narrowing the outer diameter of the capillary toward the end.

In practice, the etched capillary reaches the orifice of a nanospray emitter tip and significantly decreases sample dilution, yielding high sensitivity and electrophoretic resolution compared to a non-etched capillary. [12, 19] The distance between the end of the etched side of the capillary in the emitter and the orifice of the emitter tip was approximately 150 μm. In contrast, the distance between the non-etched separation capillary and the emitter orifice was approximately 0.5 mm as shown in Figure 19. When the volume between the outer wall of the separation capillary and the inner wall of the emitter tip increases and the outer diameter is reduced with a metal emitter because reduced the electric resistance in the space filled with the sheath liquid results in robust decoupling of the spray and separation potentials. [20]

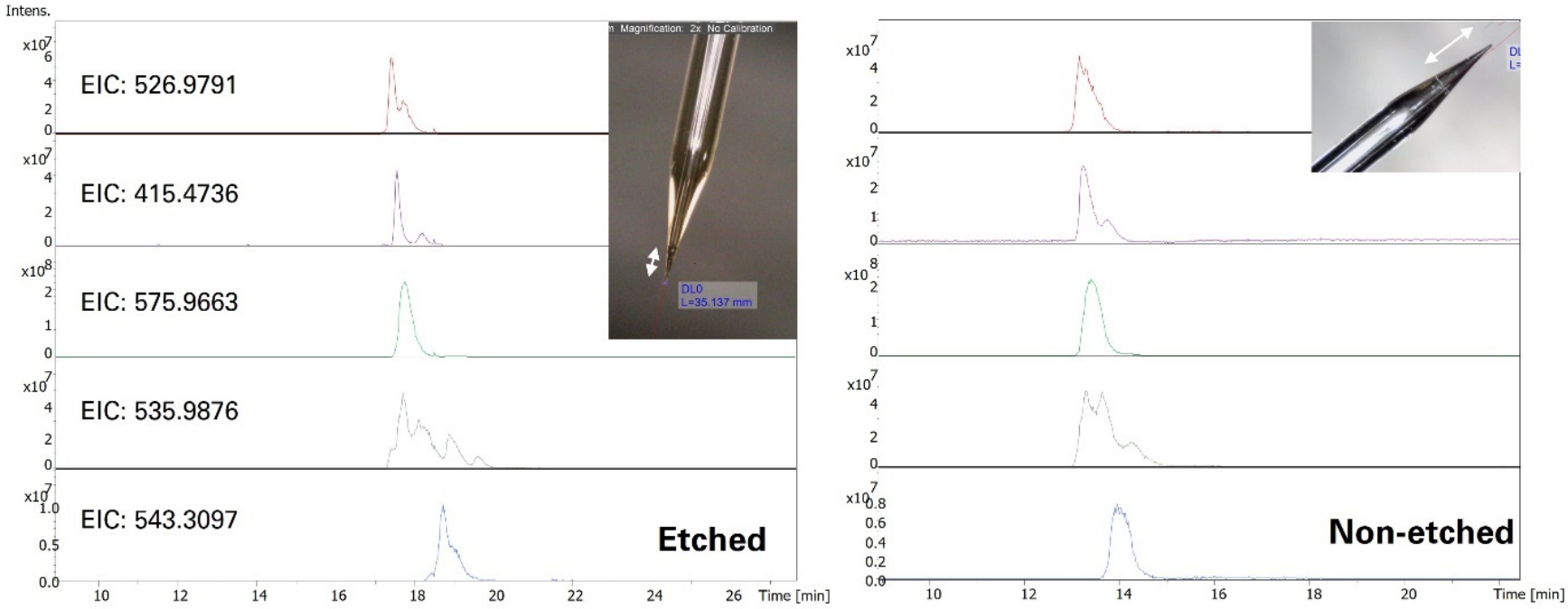

Figure 19. Sample separation optimization with either an etched or non-etched capillary. The major extracted ion electropherograms of Enoxaparin standard samples; (Left) using an etched separation capillary and (Right) a non-etched separation capillary. Photograph of each interface were displayed by insets.

Distance

The effect of distance between the inlet of a mass spectrometer and the orifice of the spray emitter tip was investigated to optimize interface performance. Figure 20 illustrates the influence of different distances between the LPA-coated capillary and the MS inlet on the abundance of the $[M-5H]^{5-}$ ion in a dermatan sulfate decasaccharide standard sample in BGE. The distance from inlet was from 1.0 to 5.0 mm while keeping other parameter the same including the spray voltage of 2 kV. The spray voltage of 2 kV value was found to be the best value in our CZE-MS system for avoiding spray destabilization, noise increases, and reduced sensitivity. The ion intensity was monitored for 24 scans and plotted against the distance with two accumulation times 0.1 sec and 0.2 sec. As illustrated in Figure 20 and 21, positioning the spray between 1.0 and 3.0 mm in front of the MS inlet resulted in high sensitivity with 2.0 kV of spray voltage. Furthermore, electrical discharge at the emitter tip is more likely to occur at distances less than 1.0 mm due to the decrease in resistance across the gap and the increase in electrical current. The optimum value was determined to be at 2.0 mm when the accumulation time was 0.2 s. The final parameters selected were a separation voltage of -30 kV, a power supply voltage of 2.0 kV, and a distance between the MS inlet and the spray tip of 2.0 mm.

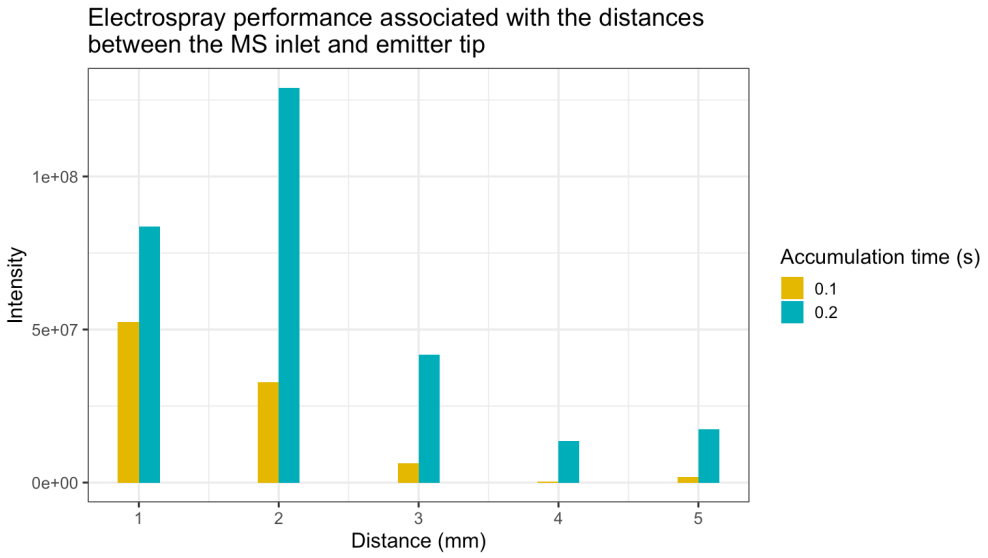

Figure 20. Electrospray performance assessment using an LPA-coated capillary for DSdp10. Ion abundance of $[M-5H]^{5-}$ trace ion of is plotted versus distance between the MS inlet and the ESI spray tip in a CZE-MS interface. The optimum value was determined at 2.0 mm when the accumulation time was 0.2 for the suitable intensity of the ion. The ion intensity was monitored for 24 scans and plotted against the distance with two accumulation time in different color bars, gold and turquoise, respectively. The power supply voltage was fixed to 2.0 kV.

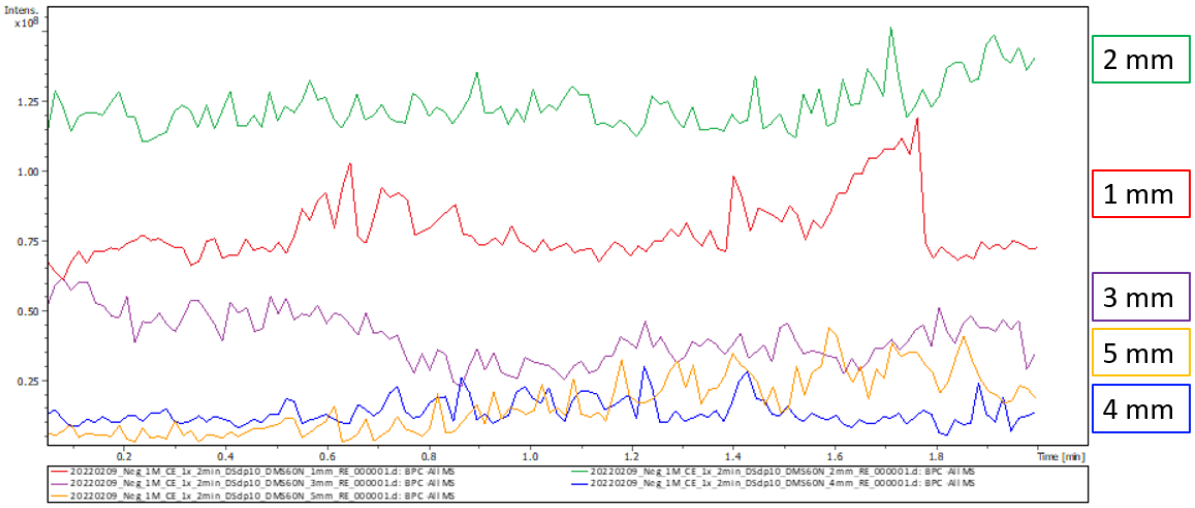

Figure 21. The base peak electropherograms of DSdp10 in the dissolved state in BGE with varying distances from 1.0 to 5.0 mm between the MS inlet and the spray tip. The accumulation time was set to 0.2 s during the data acquisition period.

Injection Volume

CZE injection adopted the widely used method of field-amplified sample stacking. This is the simplest and most common method of sample stacking that is requires only that the analyte solvent is less conductive than the running electrolyte. [15] As shown in Figure 22 for Enoxaparin, the heparin oligosaccharides are well separated by injections ranging from 58-176 nL. These injections correspond to 5.4-16.3% of the total capillary volume based on Poiseuille's law, but were not separated well by injection of 470 nL, which corresponds to 43.5% of the total capillary volume. For this experiment, a 55 cm LPA-coated capillary was used, and the injection time and pressure were set at 15 seconds.

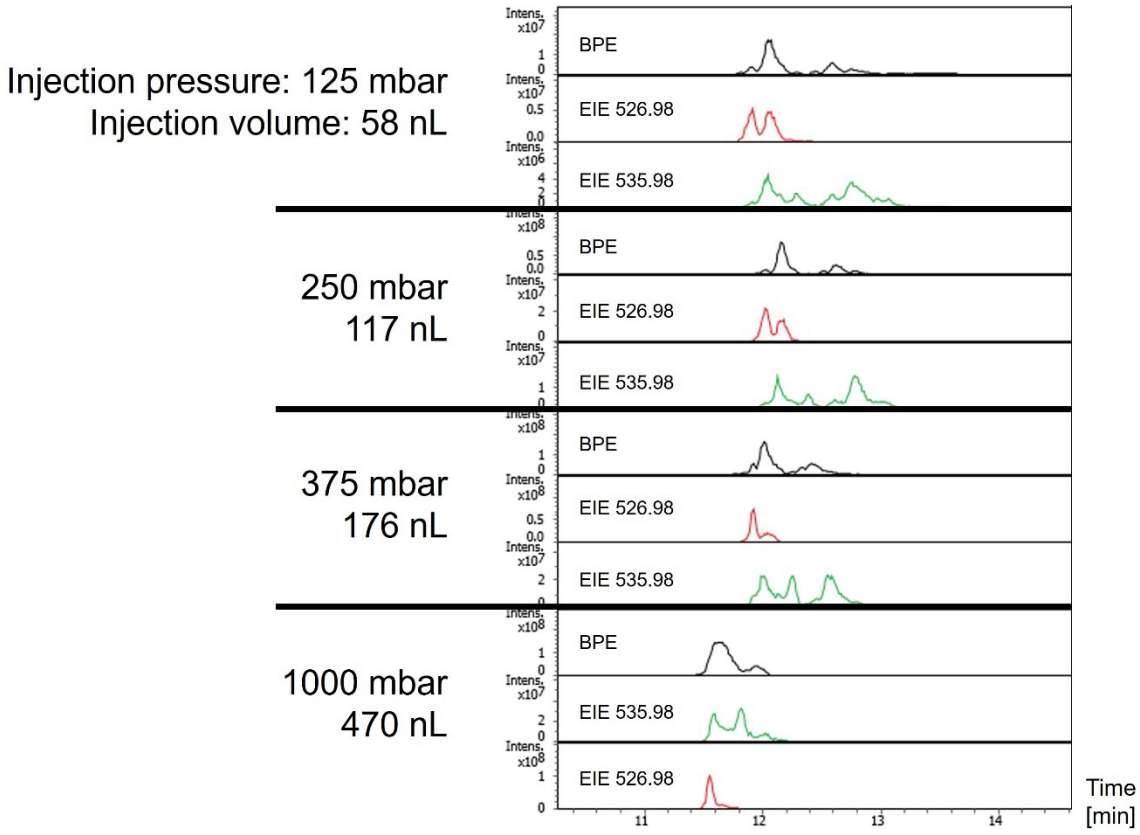

Figure 22. With each volume injection, base peak electropherograms are shown in black, and extracted electropherograms of major heparin tetrasaccharides are shown in red and green. Using Poiseuille's law, each injection volume was calculated for a capillary of 55 cm length and 50 µm inner diameter coated with dichlorodimethylsilane (DMS). Each pressure was applied for 15 s.

The detected ion abundances were improved as the injection volume increased until the base peak electropherogram reached a plateau at 1.0×10^8 intensity units. The extracted ions, m/z 526.98 and 535.98, were not separated when 176 nL volume or larger was injected, compared to electropherograms using a lower injection volume. 110 nL sample volume, approximately 10% of the total capillary volume, provided the best compromise of signal intensity of electrophoretic resolution.

Reproducibility

One advantage of the CZE-MS platform is to analyze sulfated GAG oligosaccharide/mixtures in a fast and reproducible manner using a coated capillary. [10, 21] In order to assess the reproducibility of the optimized CZE-MS platform for GAG analysis, 25 mM Ammonium formate in 70% (v/v) methanol was used as the BGE and 120 nL of sample was injected as the injection volume over a three-week period of the experiment as shown in Figure 23. Table 3 shows the method utilized for CZE analysis.

Order	**Injection material**	**Pressure (mbar)**	**Duration (s)**	**Value (kV)**
1	Sample	300	15	0
2	BGE; CZE Separation	0	3600	-20
3	BGE; Cleaning/Conditioning	300	600	0

Table 3. The CZE method set in a CZE instrument.

The reproducibility of a newly LPA-coated capillary with a 60 cm length was validated by measuring the migration time of Enoxaparin with multiple injections over a three-week period. The capillary was only utilized to run the same Enoxaparin sample for the reproducibility test. Table 4 shows relative standard deviations (RSDs) of migration time and major extracted ion intensities over a period of three weeks for three measurement points. Notably, there was reproducible separation of the LPA-coated capillary with low RSD values, but not the intensity of the heparin component. Perhaps, the glass emitter conditions were varied in use or unavoidable sample decomposition during the experiment might have affected RSDs of the intensity because migration times

were not very different according to inner wall quality, which may result in analytes being consistent in migration. In a similar manner to Sun and colleagues' finding, where LPA-coated capillaries performed well for at least a week without significant degradation, Enoxaparin samples were separated and analyzed for five days without significant decomposition or loss of intensity (data not shown) despite more than 10 runs during a week. [16]

Ion (m/z)	Migration time (min)			SD	Mean	RSD (%)
	03-26-2020	04-10-2020	04-16-2020			
526.98	27.86	28.57	27.52	0.44	27.98	1.56
415.47	27.82	28.63	27.86	0.37	28.10	1.33
575.96	28.07	29.21	28.37	0.48	28.55	1.69

Ion (m/z)	Intensity			SD	Mean	RSD (%)
	03-26-2020	04-10-2020	04-16-2020			
526.98	20426462	5498889	2116382	7954962	9347244	85.10
415.47	13556357	5161804	1189002	5155398	6635721	77.69
575.96	32009288	10253574	7544096	10950383	16602319	65.96

Table 4. The reproducibility data from three different CZE-MS runs on for 3 weeks period. (Top) Migration time. (Bottom) Ion abundance.

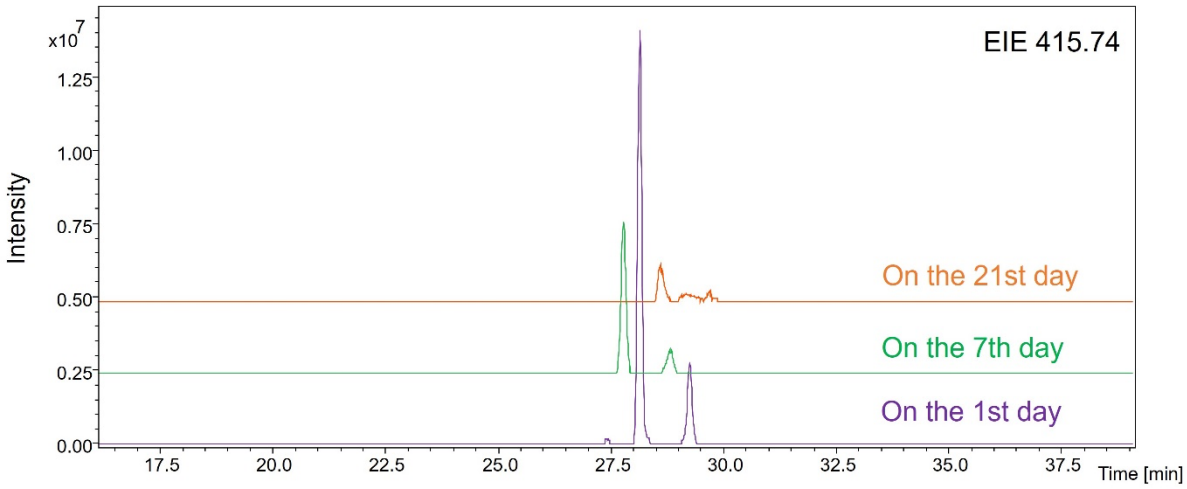

Figure 23. A stacked electropherograms of m/z 415.74 on different day acquisitions. Sample concentration of 0.01 mg/mL; injected volume of 120.0 nL; all other conditions of the CZE remained the same.

HS Profiling of CHO Cells using CZE-MS

To optimize conditions and parameters for reverse polarity CZE-MS, a HS complex mixture of GAG oligosaccharides isolated from CHO cells was used to test the methodology. The CZE-MS platform was also compared with the nano-ESI-MS method in order for a complementary GAG profiling, especially the purified/digested HS component isolated from CHO cells. Then, the HS components were assigned, and the matched features were counted for each method. Figure 24 shows the complete workflow developed for this analysis. GAGs were isolated from CHO cells with an enzymatic digestion. The purified fractions were then utilized for two experiments. First, the isolated HS mixture was subjected to a nano-ESI-MS analysis. Second, a CZE-MS analysis was performed on the heparan sulfate mixture in order to determine the number

of HS-forms that might have been overlooked during nano-ESI-MS analysis. After all outputs were analyzed using GAGwizard, a house-built MATLAB script, to characterize HS compositions, two methodologies were compiled and compared in terms of feature matches.

Figure 25 shows the mass spectrum of HS mixtures isolated from CHO cells using a nano-ESI-MS as the form of direct infusion. In this platform, components in the sample are co-eluted, resulting in ion suppression and difficulties in peak identification. [22] Direct infusion offers short analysis times but cannot provide structural elucidation of positional isomers of the sulfate group.

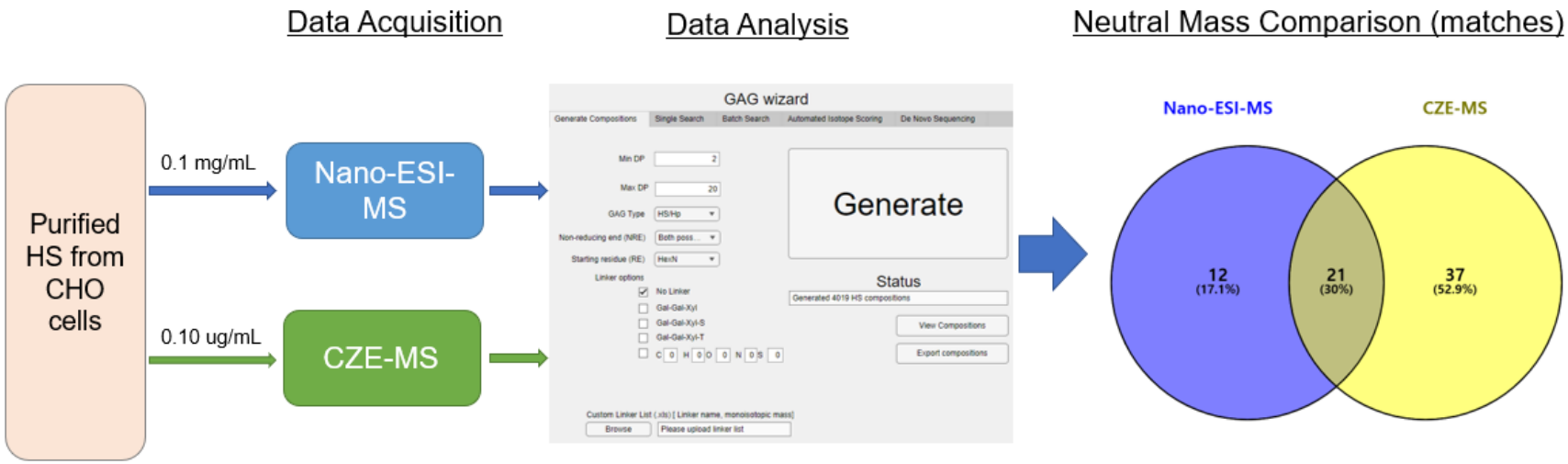

Figure 24. Overview of experimental workflow, including the sensitive CZE-MS and nano-ESI-MS platforms to characterize HS components from CHO cells.

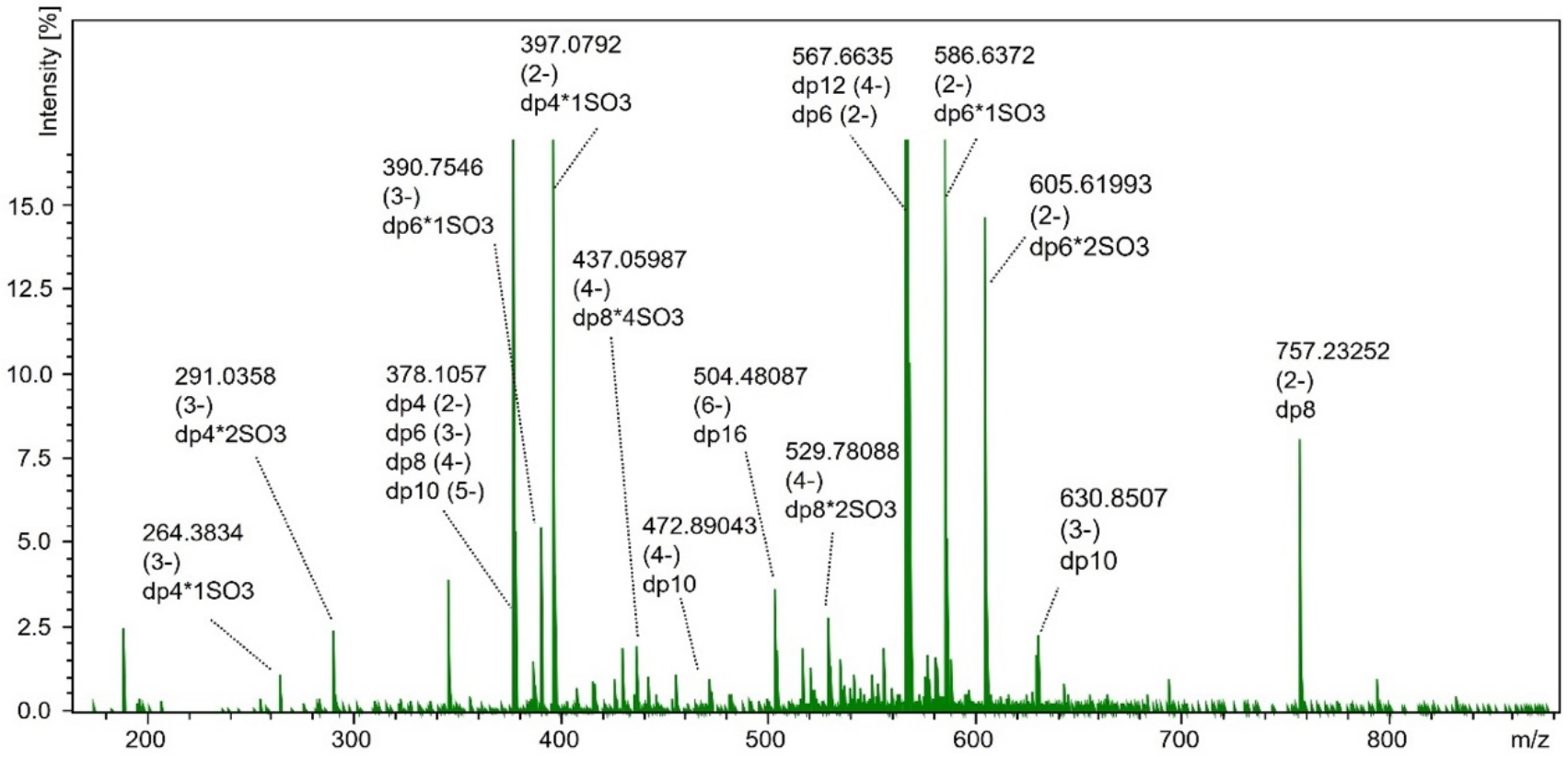

Figure 25. HS identification isolated from CHO cells using direct infusion MS. Identifiable ions were represented. Each peak displays the information regarding neutral mass and charges (the spectrum is the deconvoluted spectrum and the charge listed is the charge state before deconvolution), along with the degree of polymerization and sulfate groups (if applicable).

With CZE-MS, we demonstrated that peaks at m/z 757.21 occur at 32 minutes and 36 minutes (top and bottom mass spectrum), but these compounds are not identical HS compounds, but rather tetramers and octamers according to GAGwizard interpretation on the peaks as shown in Figure 27. In comparison to the nano-ESI-MS experiment (Figure 25 and Figure 27), CZE-MS experiments produced fewer major peaks over time, simplifying interpretation. Based on these observations, it is concluded from these observations that direct infusion MS and CZE-MS offer complementary analysis for the complex HS samples.

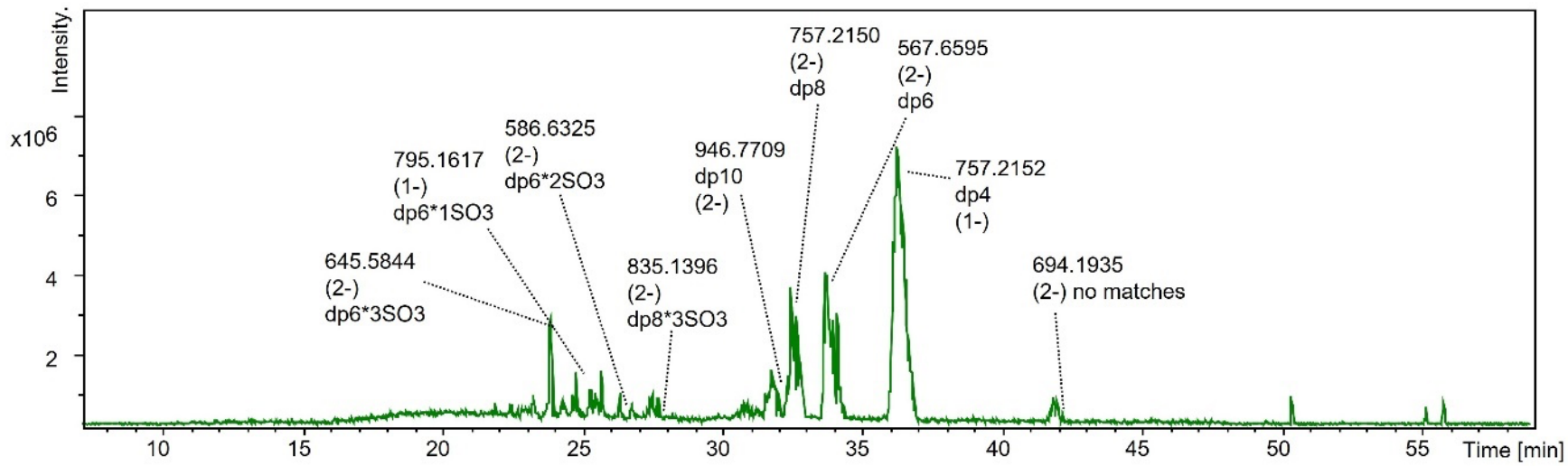

Figure 26. Base peak electropherograms of the purified HS mixture from CHO cells using CZE-MS. A summary of the neutral mass, the charge state, the chain length, and the degree of sulfation for major peaks can be found.

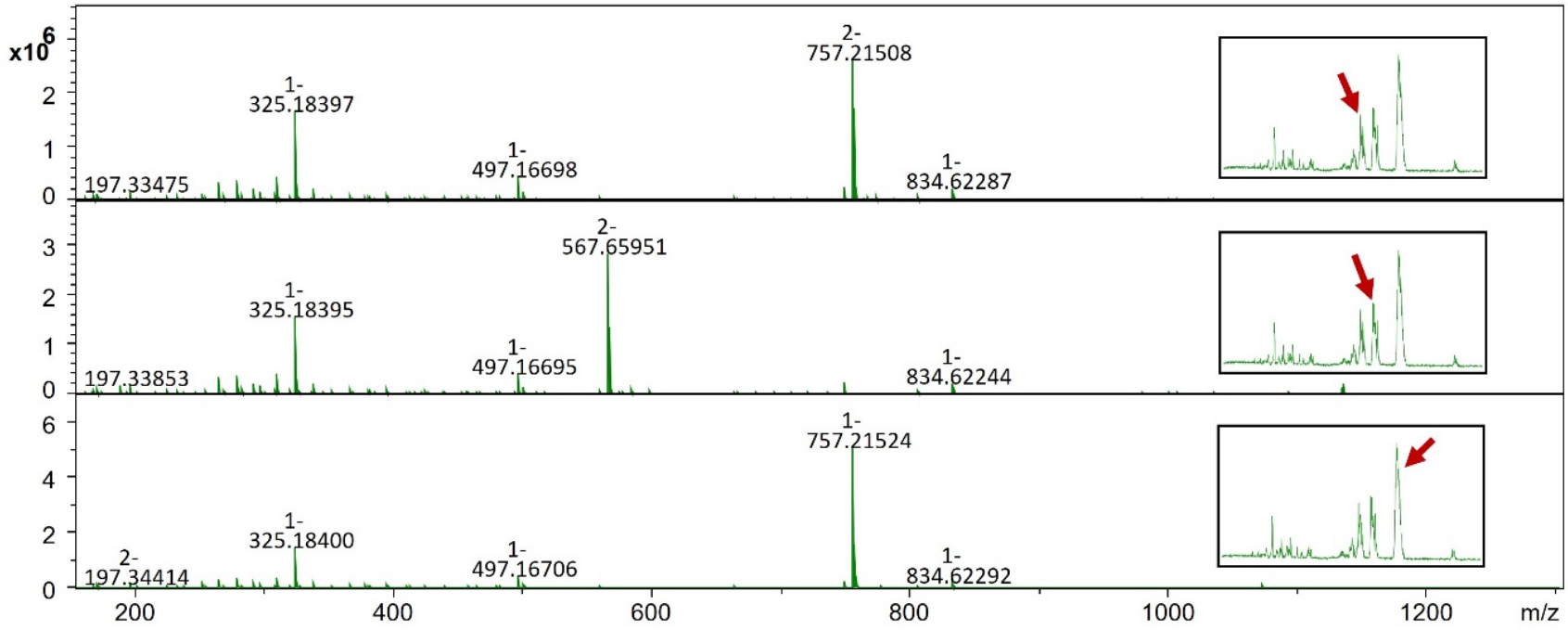

Figure 27. Mass spectrum of the electropherogram in Figure 26 demonstrating the major HS components clearly. An inset of red arrows indicates the corresponding peaks.

MS/MS Sequencing

Tandem MS is an important tool in elucidating the sequence of GAGs with regard to sulfation and the epimerization of hexuronic acid because of the structural information from fragmentations. It produces fragments depending on activation types, vibrational-activation method, and electron-activation method. Fragment ions generated by glycosidic bond cleavage and cross-ring cleavage provide information on not only compositions (the number of modifications per monosaccharide) as from MS1 but locations the location of modification.

Glycosidic cleavages provide a series of B and C ions containing the nonreducing end of a GAG chain or Y and Z ions for those containing the reducing end of a chain. Hydrogen migrates between two fragment ions, showing the oxygen at the cleaved glycosidic bond and difference of 18 Da (H_2O). Another type of fragmentation is cross-ring cleavage, which occurs when two bonds in a monosaccharide residue are broken as the name means. According to Domon and Costello nomenclature for glycosaminoglycan fragmentation, all fragment ions are denoted as $A_n/B_n/C_n$ or $X_{m-1}/Y_m/Z_m$ for nonreducing end and reducing end fragment ions, respectively (Figure 28). [23] In general, production of cross-ring fragmentation requires more energy than glycosidic cleavage because more bonds are broken. In accordance with the Symbol Nomenclature for Glycans (SNFG), colored-coded geometric shapes and modification of text have been utilized for ease of molecule presentation [24] It also has been updated to facilitate mass spectrometric depictions more, for example, in which only the mass is known, generic monosaccharide class was designated by an all-empty symbol. These are useful tools to

interpret and deliver MS/MS data. Therefore, in terms of interpreting and delivering MS/MS data, these are excellent tools.

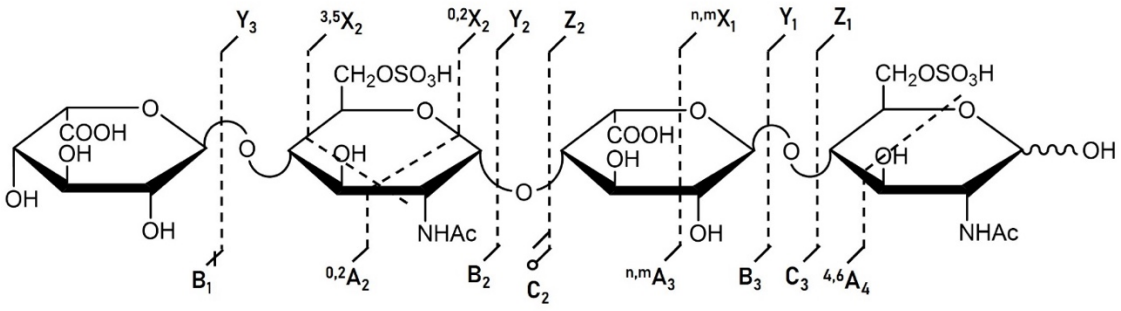

Figure 28. GAG cleavage types and Domon nomenclature. A dash at the end of each cleavage means the sulfate group decomposition. (○: single sulfate group loss, ●: multiple sulfate group losses)

Collision induced dissociation (CID) spectra revealed that CZE-MS can identify several HS components that contain a sulfate group and no sulfate group as shown in Figure 28. According to the extracted ion electropherograms, neither of these components showed sulfate group isomers, so direct infusion method was conducted to assign the structure of HS. In addition, data dependent acquisition of MS/MS spectra was also performed using AutoMS/MS mode, while applying the fixed collision voltage between 13 V to 15 V (Since each voltage value had several isomers with different dissociation degrees, CID had to be repeated for each voltage value). As shown in Figure 29A, HS mixtures isolated from CHO cells were migrated according to their size and charge over a separation window ranging from 18 min to 50 min. A certain extracted ion electropherogram of m/z 795 was followed the base peak electropherogram (Figure 29A, bottom). This shows the six distinct migration peaks within 32 minutes, however, only

three of these peaks are associated with the relevant HS tetrasacchride, whose charge state is singly charged status. In other words, the minor peaks were not monoisotopic or of different charge states. An interesting finding is that the tetrasaccharide whose N-sulfated glucosamine is located on the reducing end migrates later than the tetrasaccharide whose N-sulfated glucosamine is located closer to the non-reducing end (Figure 29A). It is noteworthy that the mixture had two components that differed in a disaccharide unit order but were chemically identical.

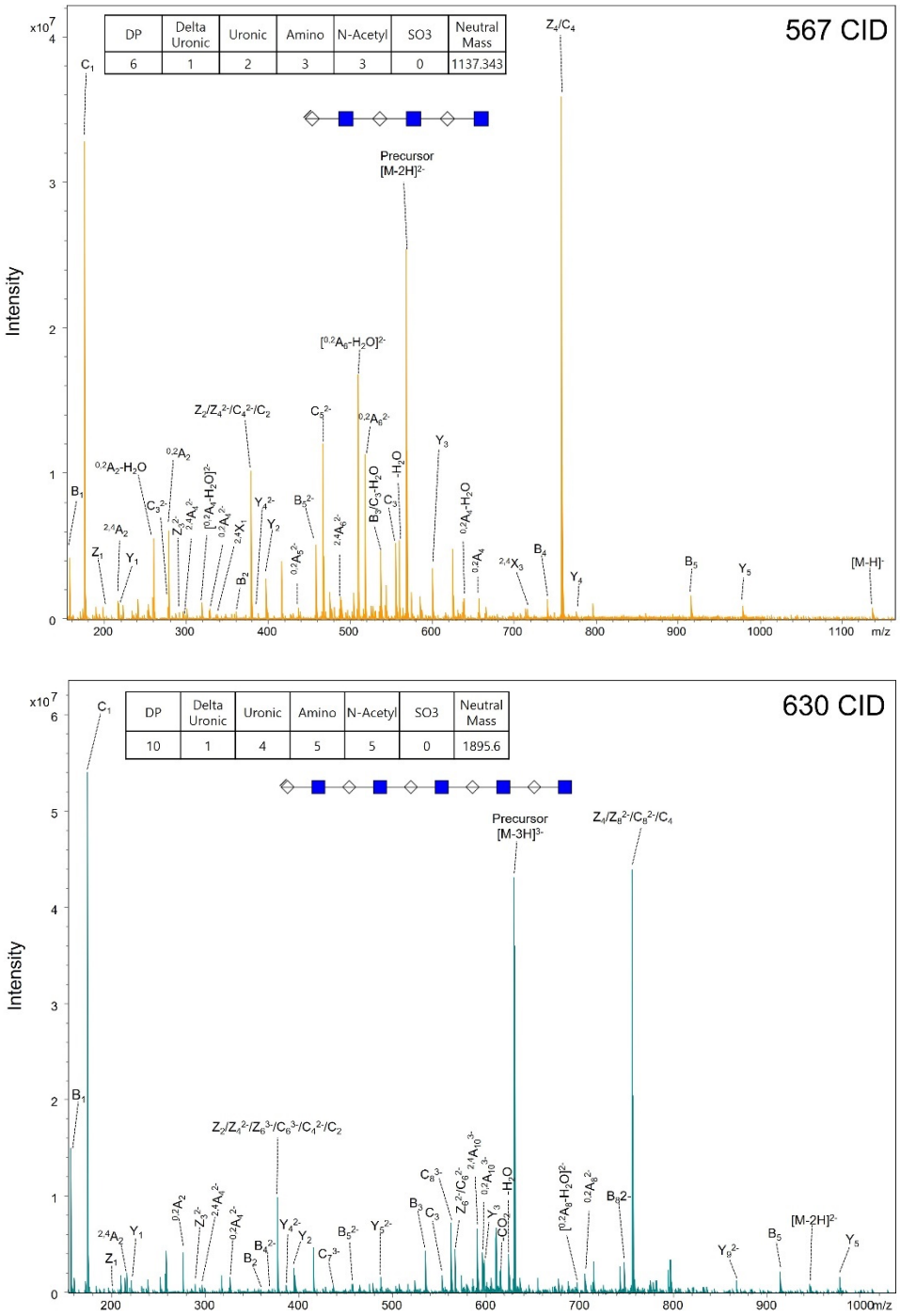

Figure 29. Tandem mass spectrometry (MS/MS) facilitated characterization of various HS components in the CHO sample mixtures. The inset tables provide the details of the HS compositions assigned by GAGWzard. (Top) the mass spectrum of a hexasaccharide of HS as determined by collision induced dissociation (CID). (Bottom) the mass

spectrum of a decasaccharide of HS as obtained by collision induced dissociation. Both GAGs were represented by the Symbol Nomenclature for Glycans (SNFG): blue ■, N-acetylglucosamine; white ◇, either glucuronic acids or iduronic acids (unsure C-5 epimerization). DP; the degree of polymerization, Delta uronic; unsaturated hexuronic acid residues, Uronic: hexuronic acid residues, Amino; amino sugar residues, N-Acetyl; N-Acetylated amino sugar residues, SO3; sulfate group modified residues, Neutral mass; a neutral mass assigned by GAGwizard search.

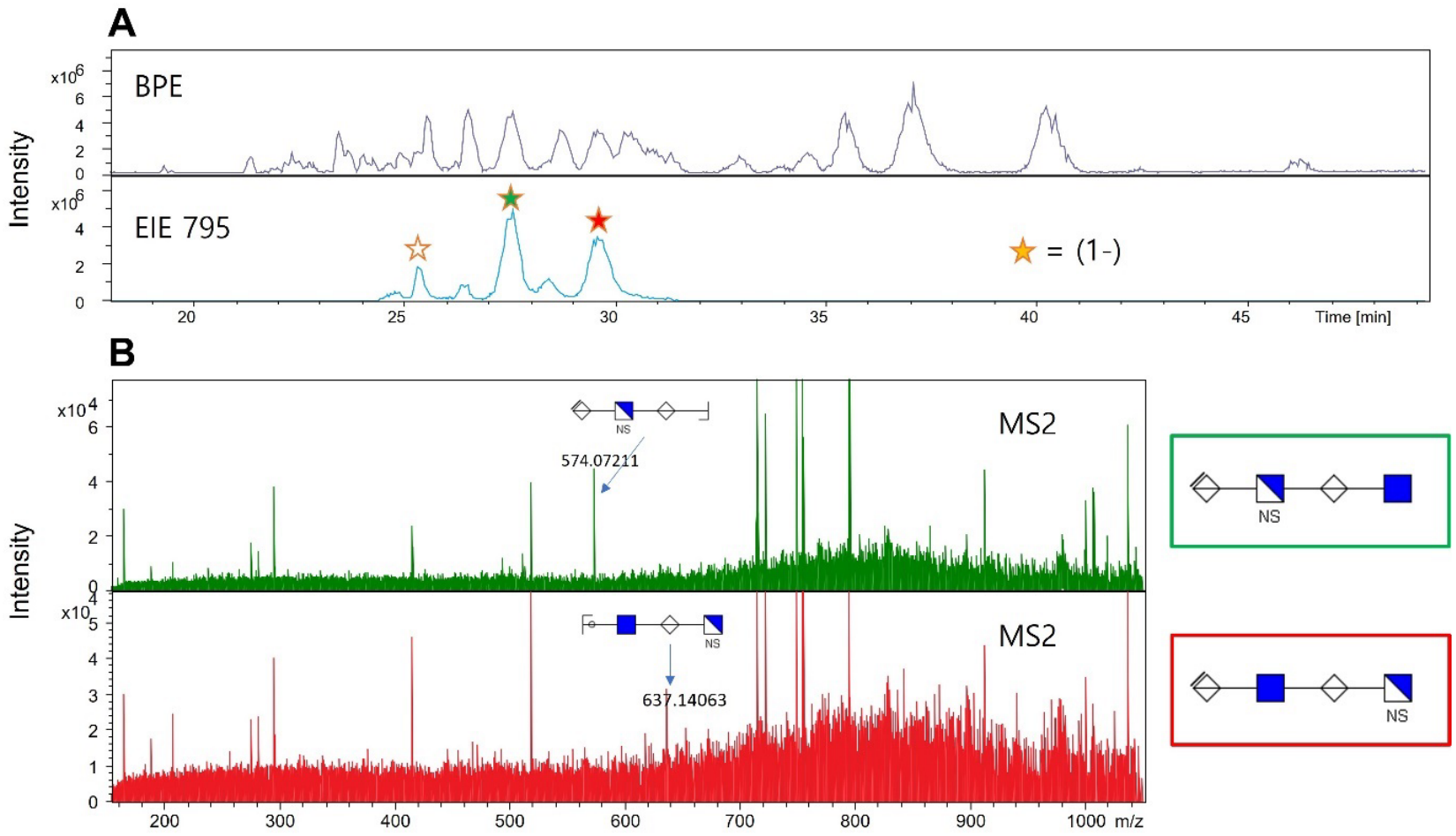

Figure 30. (A) A base peak electropherogram and an extracted ion electropherogram of m/z 795. (B) Mass Spectra of two tetrasacchrides containing a hexuronic acid, a delta hexuronic acid, an N-acetyl glucosamine, and an N-sulfated glucosamine.

Comparisons of CZE-MS and nano-ESI-MS for the Identification of HS

On the basis of the number of HS features conducted from a GAGWizard assignment, Figure 30A illustrates that CZE-MS had a 75% improvement (31 vs. 12) in HS characterization in comparison with nano-ESI-MS, and overlapping regions indicate the same HS features were identified by both methods. To be specific, when the CZE-MS and nano-ESI-MS results were combined, a total of 64 HS compositions were identified at different migration times. These results highlighted the robust performances of the two methods which are complementary. As well as that, CZE-MS provides the ability to analyze complex HS samples from cells and tissues to which it is difficult to detect specific sulfated GAGs because of their low abundance in biological samples. Figure 30B shows that an unused CZE glass emitter tip enabled a better coverage than a used tip in HS characterization and the trends in migration times and molecular weights of the matched features by GAGwizard assignment. A new tip was able to identify a higher number of HS oligosaccharides than an old tip because of the better electrospray conditions. A tip that was deemed old/used in this context was one that had been used for at least two consecutive days and at least ten separate runs. The old tip was characterized by scanning electron microscopy (SEM), which revealed obvious damage and irregularities at the end of an emitter tip. The installation of an EMASS-II coupling ion source may lead to damage to the emitter tip over time, particularly the end of the tip, although no significant physical impacts are caused by a researcher's carelessness.

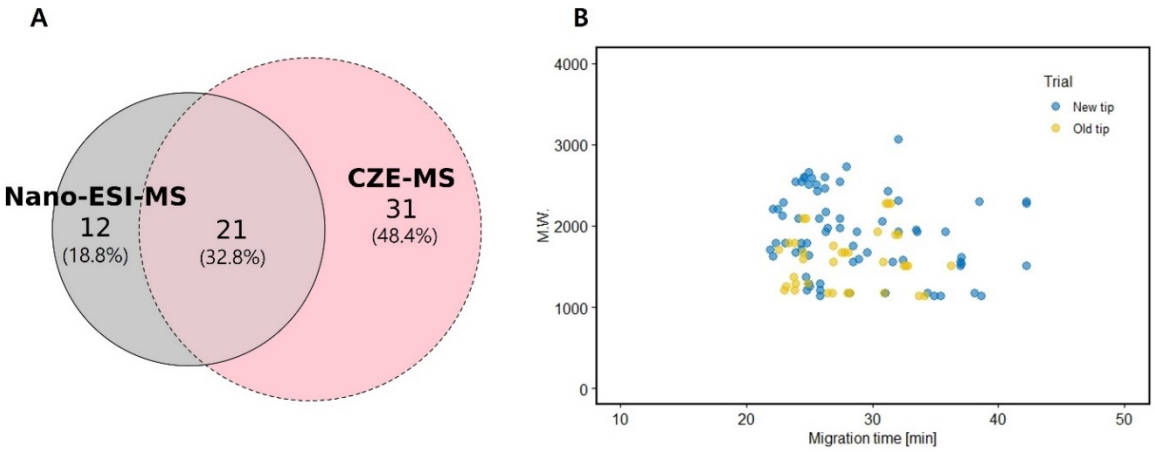

Figure 31. Assessment of HS matches for CHO cell samples. (A) Identification of HS features using CZE-MS and nano-ESI-MS. Comparison between CE and nano ESI of HS identifications from ~0.1 mg/mL of standard samples. Each method was utilized different concentration of samples: 0.1 mg/mL (nano-ESI-MS) and 0.01 mg/mL (CZE-MS) of digested and purified HS from CHO cells standard were injected. Venn diagram demonstrated the number of assigned HS oligosaccharides, and the calculated percentages represented the proportion of all HS oligosaccharides represented by each region. (B) Comparison of the HS oligosaccharide matches from the same sample but different qualities of CZE glass emitter tips.

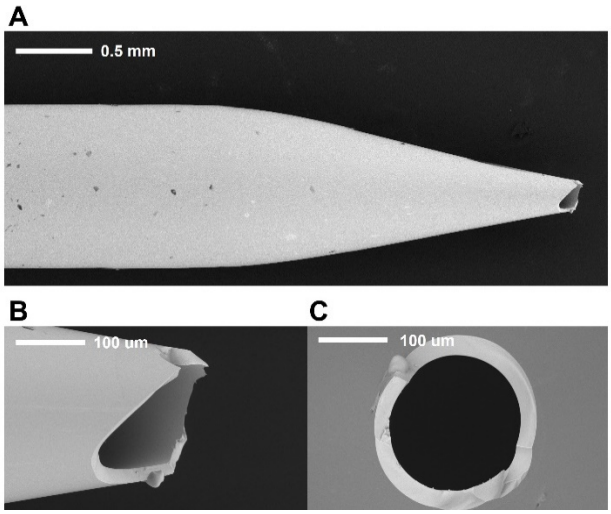

Figure 32. A scanning electron microscope (SEM) image of a damaged CZE electrospray glass emitter tip with an inner diameter of the tip hole of 120 μm. Scale bars are as indicated in the respective SEM images.

3.4 Conclusions

We reported an optimized CZE-MS method using DSdp10, a low-molecular weight heparin (LMWH) pharmaceutical product, and HS oligosaccharide mixtures derived from CHO cells. Several factors were evaluated to improve the resolution and sensitivity of our method for the best performance, including the type of coating on the inner wall of the capillary, the length of the capillary, the status of the etching on the capillary tip, the position of the emitter tip in relation to the MS inlet, and the sample injection volume. As a result, CZE-MS has been demonstrated to be superior to direct injection of GAG in identifying HS oligosaccharides in a GAG mixture, suggesting that it can be used to characterize complex engineered cell-isolated HS mixtures, especially to resolve biosynthesis pathways of GAGs.

In addition, The CZE-MS performance is also compared with nano-ESI-MS to show their incorporation so as to broaden its scope of application. Our results demonstrate that CZE-MS provides significantly more identifications for mixtures of GAG oligomers compared to nanoESI-MS. However, there is a need for further research in order to consider both additive effects in CZE-MS conditions as well as develop a CZE-MS/MS methodology that is effective. In this work, CZE-MS/MS and our annotation workflow were used to separate and assign positional isomers at baseline level as a pilot study. MS/MS analyses has proven effective in unraveling the distinct regulatory mechanisms that are associated with specific GAG motifs. Therefore, a combination of CZE-MS with electron-based activation methods has promising future potential for GAG analysis. Therefore, future applications will include intact GAG analysis and bottom-up workflows to distinguish different sulfation patterns in GAGs using the hyphenation between electron-based activation MS/MS and CZE.

REFERENCES

(1) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H. Essentials of Glycobiology **2015**.

(2) Perez, S.; Makshakova, O.; Angulo, J.; Bedini, E.; Bisio, A.; de Paz, J. L.; Fadda, E.; Guerrini, M.; Hricovini, M.; Hricovini, M. Glycosaminoglycans: what remains to be deciphered? *JACS Au* **2023**, *3* (3), 628-656.

(3) Wang, Q.; Chi, L. The Alterations and Roles of Glycosaminoglycans in Human Diseases. *Polymers (Basel)* **2022**, *14* (22).

(4) Plaas, A. H. K.; West, L. A.; Wong-Palms, S.; Nelson, F. R. T. Glycosaminoglycan Sulfation in Human Osteoarthritis: DISEASE-RELATED ALTERATIONS AT THE NON-REDUCING TERMINI OF CHONDROITIN AND DERMATAN SULFATE. *Journal of Biological Chemistry* **1998**, *273* (20), 12642-12649.

(5) Sasarman, F.; Maftei, C.; Campeau, P. M.; Brunel-Guitton, C.; Mitchell, G. A.; Allard, P. Biosynthesis of glycosaminoglycans: associated disorders and biochemical tests. *Journal of Inherited Metabolic Disease* **2016**, *39* (2), 173-188.

(6) Zhang, B.; Chi, L. Chondroitin Sulfate/Dermatan Sulfate-Protein Interactions and Their Biological Functions in Human Diseases: Implications and Analytical Tools. *Frontiers in Cell and Developmental Biology* **2021**, *9*, Review.

(7) Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J. Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Molecular & cellular proteomics : MCP* **2021**, *20*, 100025.

(8) O Staples, G.; Zaia, J. Analysis of glycosaminoglycans using mass spectrometry. *Current proteomics* **2011**, *8* (4), 325-336.

(9) Zaia, J. Principles of mass spectrometry of glycosaminoglycans. *Journal of Biomacromolecular Mass Spectrometry* **2005**, *1* (1), 3-36.

(10) Stickney, M.; Sanderson, P.; Leach, F. E.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Online capillary zone electrophoresis negative electron transfer dissociation tandem mass spectrometry of glycosaminoglycan mixtures. *International Journal of Mass Spectrometry* **2019**, *445*, 116209.

(11) Sanderson, P.; Stickney, M.; Leach, F. E., 3rd; Xia, Q.; Yu, Y.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Heparin/heparan sulfate analysis by covalently modified reverse polarity capillary zone electrophoresis-mass spectrometry. *Journal of chromatography. A* **2018**, *1545*, 75-83.

(12) Wu, H.; Tang, K. Highly sensitive and robust capillary electrophoresis-electrospray ionization-mass spectrometry: Interfaces, preconcentration techniques and applications. *Reviews in Analytical Chemistry* **2020**, *39* (1), 45-55.

(13) Basu, A.; Weiss, R. J. Glycosaminoglycan Analysis: Purification, Structural Profiling, and GAG-Protein Interactions. *Methods Mol Biol* **2023**, *2597*, 159-176.

(14) McCool, E. N.; Lubeckyj, R.; Shen, X.; Kou, Q.; Liu, X.; Sun, L. Large-scale Top-down Proteomics Using Capillary Zone Electrophoresis Tandem Mass Spectrometry. *J Vis Exp* **2018**, (140).

(15) Shen, X.; Yang, Z.; McCool, E. N.; Lubeckyj, R. A.; Chen, D.; Sun, L. Capillary zone electrophoresis-mass spectrometry for top-down proteomics. *TrAC Trends in Analytical Chemistry* **2019**, *120*, 115644.

(16) Lubeckyj, R. A.; McCool, E. N.; Shen, X.; Kou, Q.; Liu, X.; Sun, L. Single-Shot Top-Down Proteomics with Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry for Identification of Nearly 600 Escherichia coli Proteoforms. *Analytical Chemistry* **2017**, *89* (22), 12059-12067.

(17) Huhn, C.; Ramautar, R.; Wuhrer, M.; Somsen, G. W. Relevance and use of capillary coatings in capillary electrophoresis–mass spectrometry. *Analytical and Bioanalytical Chemistry* **2010**, *396* (1), 297-314.

(18) Maxwell, E. J.; Zhong, X.; Zhang, H.; van Zeijl, N.; Chen, D. D. Y. Decoupling CE and ESI for a more robust interface with MS. *Electrophoresis* **2010**, *31* (7), 1130-1137.

(19) Sun, L.; Zhu, G.; Yan, X.; Zhang, Z.; Wojcik, R.; Champion, M. M.; Dovichi, N. J. Capillary zone electrophoresis for bottom-up analysis of complex proteomes. *Proteomics* **2016**, *16* (2), 188-196.

(20) Konášová, R.; Koval, D.; Hošek, J.; Kašička, V. Investigating the position of the separation capillary and emitter tube tips in a nanoflow sheath-liquid CE-ESI-MS interface to decouple the ESI potential. *Talanta* **2021**, *228*, 122212.

(21) 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.

(22) Kirwan, J. A.; Weber, R. J. M.; Broadhurst, D. I.; Viant, M. R. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Scientific Data* **2014**, *1* (1), 140012.

(23) Domon, B.; Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal* **1988**, *5* (4), 397-409.

(24) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25* (12), 1323-1324.

CHAPTER 4

GLYCOINFORMATIC PROFILING OF LABEL-FREE INTACT HEPARAN SULFATE OLIGOSACCHARIDES[1]

4.1 Introduction

Glycans are a major class of molecules that are rich in information about the health status of an individual and play vital roles in protein interactions, stability, solubility, biological half-life, and targeting, among others. Clinical interest in the glycome and glycomics has exponentially increased in recent years, owing to the rapid development of novel methods, tools, technologies and glycoinformatics. However, glycomics and glycoinformatics for certain types of glycans remain sparse. Glycosaminoglycans such as heparan sulfates (HS), are particularly underrepresented in glycomics, hampered by inherent logistical challenges posed by their large size, structural similarities, high negative change, and labile sulfate groups. HS are a group of long, linear negatively charged polysaccharides decorated with sulfate groups and are found on the cell surface, in the extracellular matrix and in biofluids such as plasma and saliva. HS chains are synthesized in the Golgi of mammalian cells on serine residues of select proteins (termed proteoglycans, PGs). Each HS chain is built by a myriad of enzymes that collectively produce the characteristic heterogenic structures of HS (Figure 33). The positioning of

[1] Marissa L. Maciej-Hulme, Jandi Kim, Elijah Roberts, Yiqing Zhang, Anouk van der Velden, Dirk den Brancker, Cansu Yanslir, Mark de Graaf, Ton Rabelink, Bernard van den Berg, Anne-Els Logt, Ellen van Omen, Rutger Maas, I. Jon Amster, Johan van der Vlag. *To be submitted to PNAS.*

sulfate groups along the HS chain give rise to distinct 3-dimensional structural and chemical properties (Figure 33). Local biochemistry and architecture within the chain encode the interaction of HS with hundreds of different HS-binding proteins (HSBPs), enabling orchestration of a wide variety of biological activities including: 1) co-receptor roles in growth factor signaling, as 2) scaffolds for protein dimerization /multi-oligomerization, and 3) in morphogen gradient dynamics and ligand sequestration mechanisms. HS heterogeneity is further expanded after biosynthesis at several points throughout its lifecycle, including potential modification by extra-Golgi sulfotransferases[1] sulfatases and heparanase-1 (HPSE)[2]. Activated HPSE cleaves portions of HS chains; shortening the parent chain(s) left on the proteoglycan and releasing HS fragments into the extracellular milieu. Shedding of whole HSPGs from cells and the extracellular matrix can also occur during processes such as tissue remodeling, infection and/or inflammation[3].

Mass spectrometry (MS)-based glycomics offers an overview of structural glycan changes in a sample, providing potential insight into the considerable heterogeneity of HS by enabling multiple HS species from a single sample to be analyzed simultaneously (profiling). To date, almost all clinical MS-based HS analyses conducted so far have focused on HS disaccharides, which are derived from HS chains by enzymatic depolymerization using bacterial lyases. This step conveniently reduces sample heterogeneity to 12 distinct disaccharide species but simultaneously destroys vital information encoded within oligosaccharide architecture about HS biological function exerted through HS-binding partners. Despite this, compositional analysis of plasma heparan sulphate disaccharides by MS were recently demonstrated as a potential diagnostic for early stage cancer[4], potentiating the use of HS glycoprofiling for personalized medicine

from a simple blood draw even though different cancer types could not be distinguished from disaccharide data alone.

We and others have previously established a prominent role for HS in the development of several kidney diseases[5-10]. In some of these increased circulating plasma HS precedes kidney damage, which suggests that plasma HS may be a potential biomarker for kidney injury and/or disease progression. Since HS oligosaccharides contain more biological information than disaccharides, we sought to develop a complete workflow pipeline from plasma sample preparation, through MS data acquisition, and onto glycoinfomatics that retained the natural heterogeneity of HS, as a more information-rich glycoprofiling approach for probing the health status of a patient. To demonstrate the concept, we purified and analyzed plasma HS from healthy controls and groups of kidney disease patients, providing the first data set of plasma-derived biologically intact oligosaccharides[6]. These data underpin the identification of plasma HS oligosaccharides associated with specific diseases and/or patients with potential disturbances in HS oligosaccharide structures.

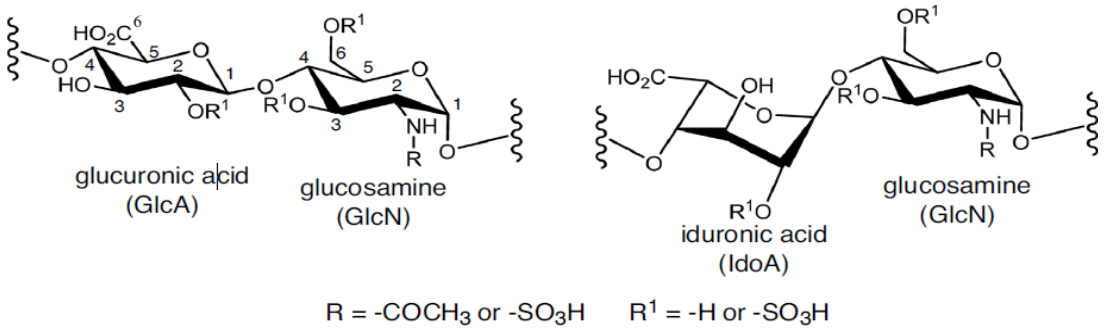

Figure 33. Heparan sulfate biosynthesis and structure

4.2 Experimental method

Preparation of human plasma and HS oligosaccharide purification

Written informed consent was obtained from healthy volunteers and patients prior to blood collection. EDTA-collected blood was centrifuged at 4ºC for 15 mins at 2000 *xg*. The plasma (supernatant) was heat inactivated at 60ºC for 30 mins, then rapidly cooled on ice. 1 mL plasma was supplemented with cABCase buffer (25 mM Tris, 2 mM $Mg(Ac)_2$ pH 8), then treated with 2 mU of chondroitinase ABC (Sigma), 7.5 U/mL DNAse I (Qiagen) and 10 U/mL RNAse I (Thermo Scientific) for 4 hours at 37ºC, followed by addition of 125 µg/mL proteinase K (Merck chemicals B.V., Amsterdam, The Netherlands) and continued digestion at 37ºC overnight. To inactivate proteinase K, samples were heated at 95ºC for 10 mins, then cooled on ice. Samples were 5x diluted with room temperature (RT) sterile PBS immediately prior to HS isolation to adjust the temperature and pH. Anion exchange chromatography was performed using individual 1 mL bed volume DEAE-Sepharose CL-6B beads (Sigma) prepared as previously described[11] and equilibrated in PBS at RT. Bound HS was washed with 0.25 M NaCl/PBS, pH 7.4 and then eluted with 4x 1 mL of 2 M NaCl/PBS, pH 7.4. Isolated HS was desalted in 1 mL aliquots using PD10 desalting columns (GE Healthcare, Sephadex G25) with CHROMASOLV™ LC-MS H_2O (Fisher Scientific), dried by centrifugal evaporation and stored at -20ºC.

HS quantification

Isolated HS was quantified in black walled 96-well microplates for fluorescence-based assays, (Invitrogen) using the Heparin Red assay (Red Probes, Germany) according to the manufacturer's instructions. Porcine HSBK (Sigma) was used to prepare the standard

curve. Fluorescence was measured after linear shaking for 20 secs ($_{ex}$ 570 nm, $_{em}$ 605 nm) using an Infinite M200 Pro fluorescence plate reader and software.

CE-MS analysis and HS oligosaccharide structure predictions

Desalting HS samples and DMS-functionalized capillary

The desalting process of HS samples was conducted by the 0.5 mL 3 K Amicon Ultra spin filters. Dry HS sample was dissolved in 400 μL HPLC grade water and further de-salted with 0.5mL 3 K Amicon Ultra spin filters. The cellulose membrane filter was pre-moisturized for 1 min before application of the rehydrated samples and centrifugation at 14,000 x *g* for 25 mins three times. The desalted aqueous GAG sample was transferred to a 250 μL CE vial with a pre-slit septum cap on.

CZE-Orbitrap MS analysis

Mass spectra were collected in negative mode on an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). The automatic gain control (AGC) target value was set to $1e^6$. All mass spectra were collected under the profile mode. To minimize sulfate loss, the source region stacked-ring ion guide (S-lens) was set to 20%. Capillary zone electrophoresis instrument (Agilent G1600 HP 3D CE) was coupled to the Orbitrap mass spectrometer with an EMASS-II CE-MS Ion Source (CMP Scientific, Brooklyn, NY). The ion transfer tube was maintained at ground potential and at 200°C. A voltage of -1.70 to -1.85 kV was applied to the sheath liquid (SL) to achieve electrospray. The scans were acquired at a resolution of 120,000 at m/z 200, one micro scan with a maximum injection 200 ms.

Prior to CZE-MS experiments, a semi-automatic optimization of source parameters was performed using sucrose octasulfate to improve sensitivity of sulfated GAGs and reduce sulfate decomposition during ion transfer prior to MS analysis. The inlet of ion transfer tube was aligned ca. 2.5 mm away from the outlet of glass emitter. CZE separations were performed on fused silica capillaries (60 cm x 360 μm OD x 50 μm ID) functionalized with dichlorodimethylsilane (DMS). The outlet end of the capillaries were etched with hydrofluoric acid (HF) to reduce the outer diameter to <100 μm and improve performance of the sheath flow interface. DMS functionalization and HF etching procedures are reported previously[12, 13]. As the electrokinetic nanospray ion source, the flow rate of the interface was ca. 10 L/h with 25 mM ammonium acetate in 70% (v/v) methanol/water used as a SL[14]. The glass emitter was fabricated with Sutter P-1000 pipette puller (Sutter instrument, Novota, California) in the UGA Biomembranes Engineering Facility.

The aqueous GAG sample was injected for 6 s at 950 mbar followed by a background electrolyte (BGE) injection for 10 s at 50 mbar. The injected volume ended as 156 nL (sample plug was 14.5% for a 60 cm length column). The etched capillary was positioned 0.3-0.5 mm from the end of the emitter orifice to create a mixing volume of approximately 2 nL. The small mixing volume significantly reduces the dilution of analytes, compared to the conventional CZE-MS sheath interface.

The HS samples were run in triplicate and the sequence order was randomized to minimize system bias[15]. Every tenth sample injection was followed by a blank injection and standard Dermatan Sulfate to monitor the system for carryover and mass accuracy

shifts. Total CZE-MS running time was set to 60 min. After each run, 5 min fresh BGE flush was followed to remove residual contaminants in the capillary.

Data processing of Orbitrap mass spectra

Mass spectra were visualized in Thermo Xcalibur 2.2 software. MS convert (developed by Jerry Holman[16]) in ProteoWizard 3.02 package was utilized to convert Thermo raw files to open-format mzML files. None of filter is applied in MS convert GUI. Binary encoding precision is 32 bits. Zlib compression is used for file size deduction and TPP compatibility is allowed.

MZmine 3.1 package was employed to extract MS features (m/z, charge, retention time, and intensity)[17]. Part of key parameters used in MZmine 3.1 were optimized with Paramounter R package developed by Guo[9]. All the in-house data processing with MATLAB and MZmine 3.1 run on the custom desktop with 64GB RAM and twelve-threaded CPU. After comparing the running time and RAM consumption of three different deconvolution algorithm for one GTP017-01.mzML file (XCMS 20 min, ADAP 5 min, Local minimum 1 min), the local minimum algorithm was selected for the deconvolution.

The setting of MZmine 3.1 parameters is described below. Mass tolerance for chromatogram builder was assigned to ±5 parts per million mass resolution to seek an accurate mass m/z. The wavelet transform was selected as the mass detector with scale level 5 and wavelet window size 30%. The signal/noise threshold was set to the estimation of local minima by in-house software[18]. Next, the ADAP chromatogram builder parameters were set as follows: minimum absolute peak height 2.0×10^2 and

minimum number of scans 5. Electropherogram deconvolution parameters were established as follow: Chromatographic threshold 85%, minimum peak top/edge ratio 1.7, minimum absolute intensity 4.0×10^2, and maximum peak duration range 10 min. Isotopic peaks grouper algorithm was chosen to deisotope features with retention time tolerance 0.41 min, maximum charge five, and the lowest isotope selected, which is the monoisotopic peak.

Common data processing; MZmine2 and GAGWizard

MZmine2GAGWizard is an in-house MATLAB script that calculates the neutral mass for each feature based on its m/z and charge, and also reorders the columns [neutral mass, m/z, charge, intensity, RT].

$$\text{Neutral mass} = (m/z * \text{charge}) + (1.0078 * \text{charge})$$

Healthy control filter

In house software was used to identify features that were associated with healthy control patients. To meet this criterion, a neutral mass feature must appear in at least two technical replicates of a healthy patient, and only has to appear in at least one patient sample. These features were then removed from the remaining datasets for non-healthy patients.

Technical Replicate Condenser

For each set of technical replicates (in triplicate), their feature lists were condensed into a single list. Features that were found in at least two of the three technical

replicates were allowed to survive on the condensed list, and a 5 ppm error tolerance was used for neutral mass variation. The neutral masses and intensities were saved on the final list.

GAG Wizard

Likely HS type GAG compositions were found for the remaining neutral mass features using GAG Wizard. GAG Wizard generates a combinatorial database of theoretical GAG compositions and their neutral masses which can be used to search against the data. For this study, a database of HS type GAGs was generated to resemble fully intact HS that are found on HSPGs. The DPs generated were between dp2 and dp50. "Gal-Gal-Xyl-Ser" was tenable linker region considered because of the complete proteinase k digestion in this protocol. This generated a list of 35,444 theoretical compositions ranging from 940 Da to 16.9 kDa.

The mass lists for each patient were uploaded to GAG wizard, and the software searched the database for theoretical masses that matched the data within 5 ppm mass error. If multiple compositions fell within the 5 ppm error tolerance, each of them was returned on the results file. The output of GAG wizard contains information about the input data (m/z, z, I, RT, neutral mass, RT) as well as information about the theoretical composition that it matched, which includes DP, #HexA, #HexN, #SO_3, #NAc, the number and type of adduct, as well as the elemental formula. Lastly the ppm error is calculated for each matching composition.

Principal Component Analysis (PCA)

Neutral masses that were found to match known GAG compositions were extracted from the GAG wizard results files and collated into a list. A PCA matrix was constructed with in house software in GAG data analysis pipelines, where each row is a neutral mass feature, and each column represents a patient. The intensity of every neutral mass feature was filled in the middle of the matrix. If a neutral mass feature was not found for a particular patient, that cell was filled with a zero, and if a neutral mass matched multiple HS compositions, that neutral mass was only added to the matrix once. The matrix was uploaded to MetaboAnalyst which is a web-based platform for comprehensive data analysis[19]. Pre-processing steps included mean centering and normalization by the highest peak intensity for each neutral mass. Considering that there are a significant number of variables in Orbitrap dataset, inbuilt interquartile range (IQR) filter of MetaboAnalyst was automatically applied to keep only top maximum features, which is strongly recommended for untargeted metabolomics dataset.

HS biomarker discovery

Neutral mass lists with GAG composition match in the same disease's cohort were collated into one group neutral mass list. Six diseases cohorts neutral mass list were input into Venneuler R package to make the HS hits database set and were calculated from the intersections of the neutral mass list. The UpSet diagram represents unique and shared HS hits within six kidney disease cohorts, sorted by frequency-proportional manner.

4.3 Results

First, we collected plasma from healthy individuals with informed consent and developed a sample preparation workflow compatible with either EDTA- or citrate- plasma (Figure 34). Sample preparation was adapted from previously established methods [2-5]. Successful purification of plasma HS by the adapted method was confirmed by HS disaccharide analysis from 3 pooled healthy samples (Table 5). Total HS isolated from 11 healthy individuals varied from 2.25 µg/mL to 7.58 µg/mL, with an average of 5.37 µg/mL. The number of detected species by CE-MS did not correlate with the quantity of total HS (data not shown). Using our in-house code, GAGWizard, information about each neutral mass was assigned including: m/z, z, I, RT, neutral mass) as well as the theoretical composition (degrees of polymerization (dp), #HexA, #HexN, #SO3, #NAc, the number and type of adduct, elemental formula). HS dp size ranged between 2-34 (although predicted parameters were limited at 50) and the number of sulphate groups varied from 0-28, demonstrating versatile detection of a structurally diverse range of HS species within each sample. All matched HS data of neutral masses from MS spectra were combined onto a master list, which was subsequently used in the data processing pipeline to filter out healthy HS matches from kidney disease patient data (Figure 35).

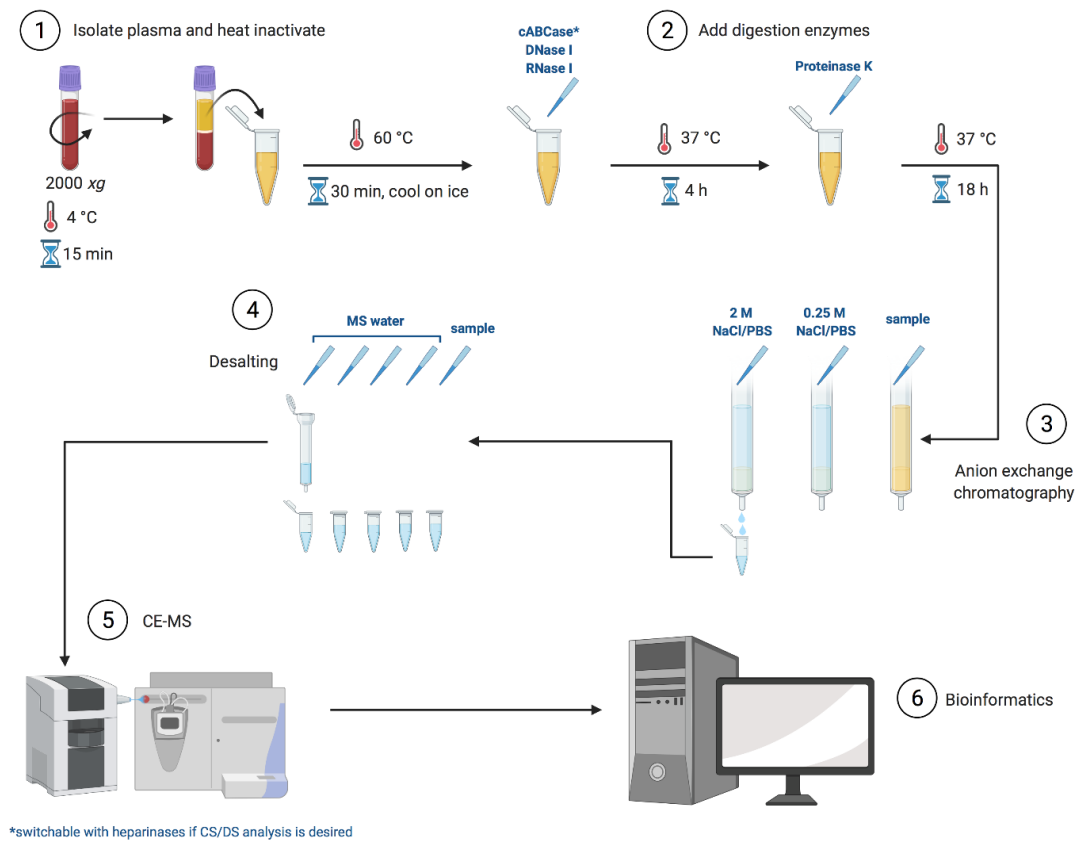

Figure 34. A workflow ranging from sample preparation to data processing by bioinformatic tools.

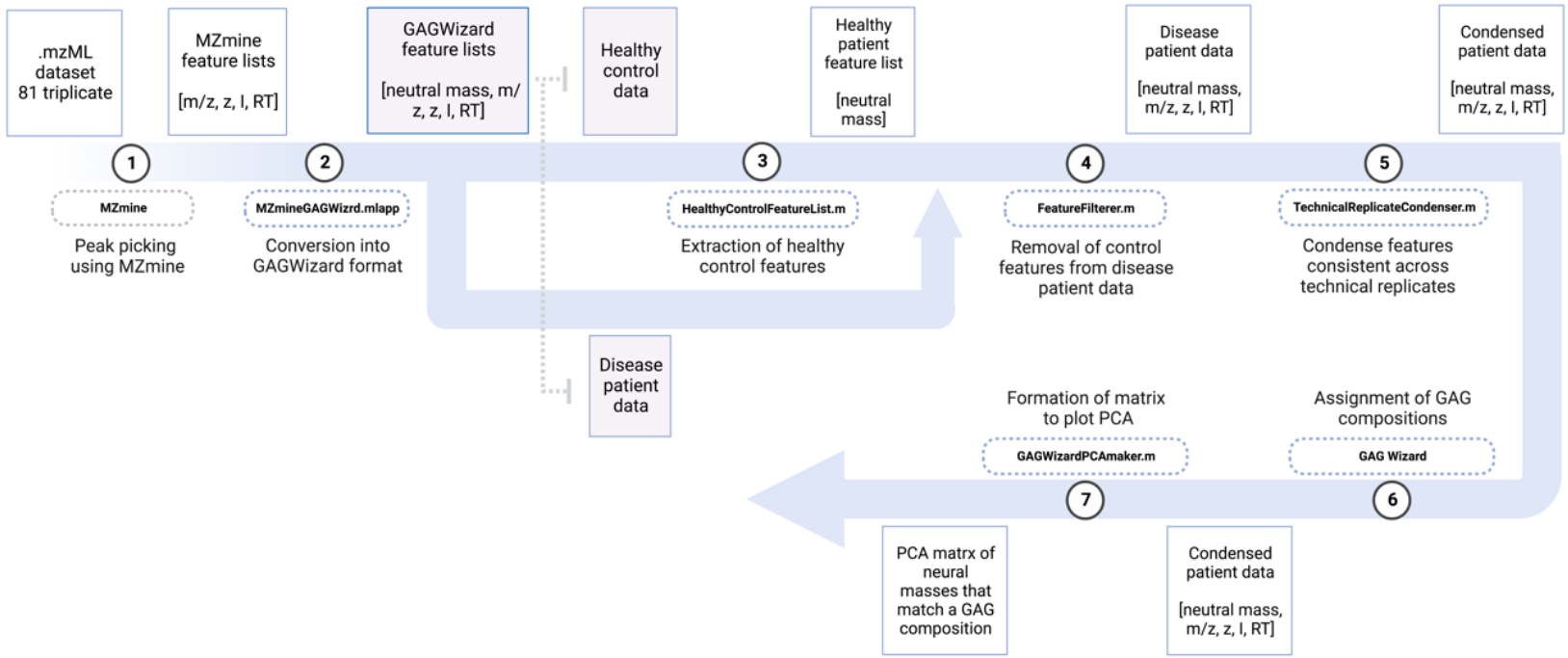

Figure 35. Data processing and bioinformatics workflow in detail. Square boxes show collected data and their format styles. Dotted boxes with numbers inside of the blue arrow mean MATLAB scripts and the objectives. 1~2: MZmine and a MATLAB script provide raw data processing and GAGWizard feature lists. 3~4: Screening features that are only useful to the disease group using the healthy control filter script. 5~6: Completion of representative features, GAG composition assignment, and PCA application.

Sample number	Disease	HS (ug/mL)	P value (*vs*. healthy)	HPSE activity (ImU/mL)	P value (*vs.* healthy)
11	healthy	5.11 (1.75)	n.s	0.13 (0.008)	n.s
15/2	FSGS/MCD	9.77 (3.43)	0.0002	0.22 (0.080)	0.0074
11	SLE	3.61 (1.08)	n.s	0.15 (0.034)	n.s
20	MG	6.9 (2.55)	n.s	0.13 (0.003)	<0.0001
5	T2D	Not tested	-	Not tested	-

n.s, not significant

Table 5. Healthy and patient metadata characteristics.

HS Disaccharides	O-sulfation	N-substituents	Weight (µg)	Percentage (%)
D0A0	-	acetate	0.12	57
D0S0	-	sulfate	0.04	18
D0A6	6	acetate	0.02	9
D2A0	2	acetate	0.02	9
D0S6	6	sulfate	ND	0
D2S0	2	sulfate	ND	0
D2A6	2,6	acetate	ND	0
D2S6	2,6	sulfate	0.02	8
Total HS			0.22	100

*Lawrence et al. (2008) Nature Methods 5:291

Table 6. Heparin Disaccharide SAX-HPLC Analysis of Pooled Healthy Control Group.

Previously we successfully demonstrated HS oligosaccharide analysis on small scale sample sets from human urine[12, 20] and mouse cell culture preparations[6]. However, application to larger datasets posed significant data processing problems using a manual approach. Therefore, we designed a data processing and glycoinformatics workflow for curated HS glycoprofiling suitable for large scale datasets (Figure 35). We then applied the same plasma purification and analysis method to 56 kidney disease patients. Once the matched HS lists for each patient were obtained, data were processed and curated to reflect the presence of any potential unique HS species within each patient compared with healthy samples. Grouped data for each disease showed shared and distinct features for each disease type; the most striking being for membranous glomerulopathy (MG), which contained the most total number of HS species detected (88) of which 59 were unique to MG (Figure 36A). Each kidney disease group had at least 6 or more unique HS species, which may represent potential biomarkers for specific kidney disease diagnostics. Shared HS species between kidney disease groups may be indicative of general mechanisms such as acute inflammation that result in tissue injury when not resolved. Multiple HS species were found in more than one disease group, with 2 HS species common to all kidney diseases. Principle component analysis (PCA) analysis revealed two groups where 4 MG patient HS profiles were clustered away from all other patients (Figure 36B).

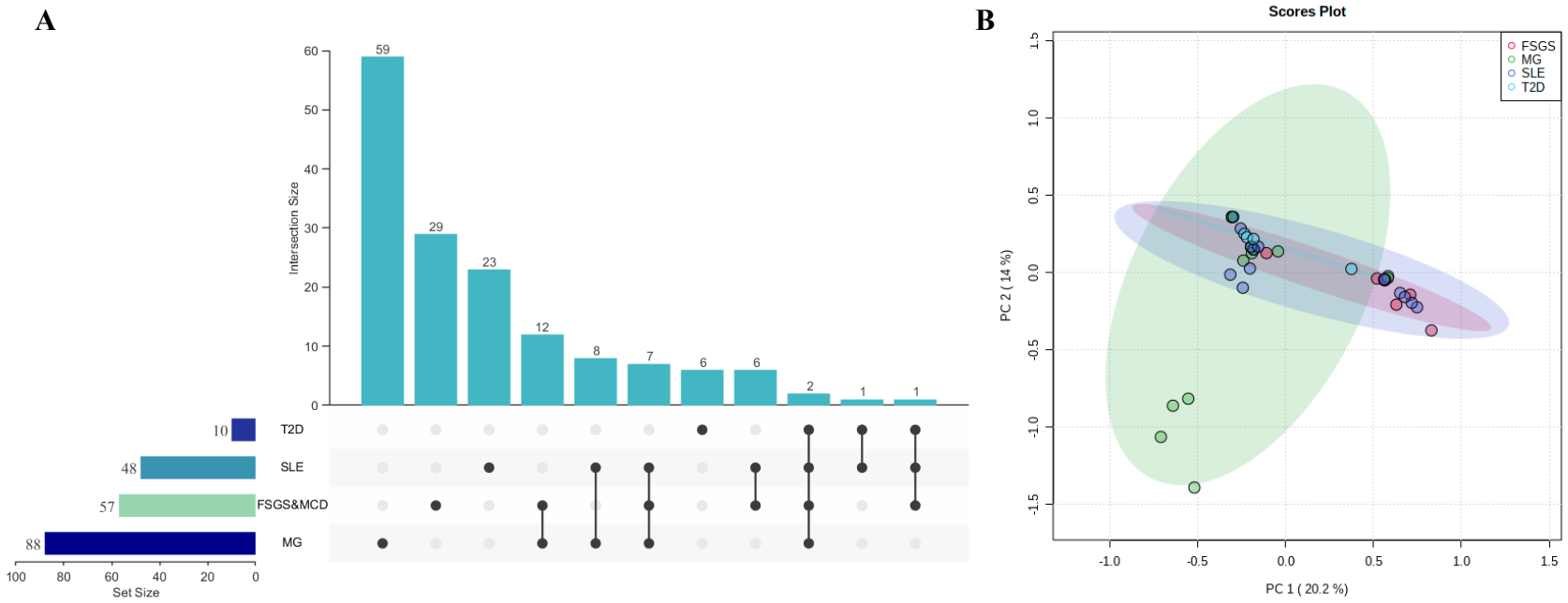

Figure 36. Statistical analysis to identify HS features and classify groups for the disease cohorts. (A) UpSet plot of shared and unique HS hits for kidney diseases groups. (B) PCA of HS profiles. PCA stratifies data classifications based on disease type. Each dot represents a patient HS profile, with different colors representing different disease groups. Four MG samples were separated from other patients in which most of the variables were gathered in the PCA plot.

One common mechanism that may modify the HS profiles of kidney disease patients is HPSE activity. Active HPSE degrades HS chains and releases HS oligosaccharide fragments from proteoglycans. Damage to the HS glycocalyx plays a prominent role in the development of kidney disease[21]. Therefore, we also tested the plasma samples for HPSE activity to see if significant HPSE activity might produce a common HS oligosaccharide profile. Of the 6 disease groups, only FSGS/MCD patients showed a significant increase in plasma HPSE activity accompanied by a significant increase in circulating total plasma HS. In contrast, MG patients had statistically significantly lower plasma HPSE activity than healthy controls and total HS levels within range of healthy controls. To investigate which HS species may be common to patients with increased HPSE activity, we stratified the patients irrespective of disease type into HPSE- and HPSE+ groups, where HPSE+ represented patients with enzyme activity values above those found in healthy controls (Figure 37).

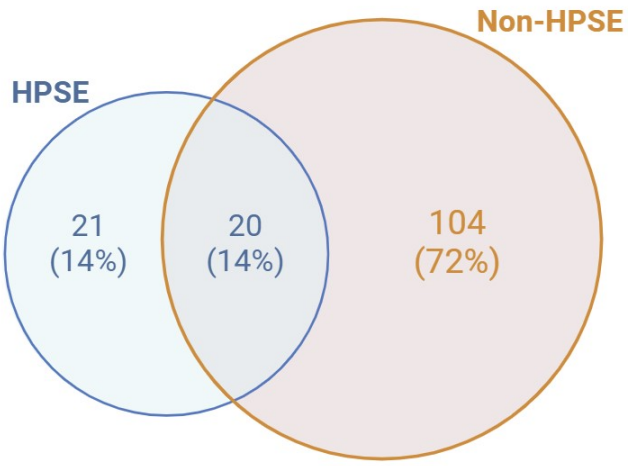

Figure 37. HS Features comparison between HPSE- and HPSE+ groups.

4.4 Conclusions

Human plasma is a biofluid rich in information about patient health status and a routinely biobanked sample, offering diverse sample collections from a plethora of diseases and easy integration into diagnostic processes in healthcare systems. Various diseases have been reported to show increases in circulating plasma HS, including COVID-19[4], sepsis [22], cancer[23], and kidney diseases[24, 25]. Current heparan sulphate purification methods for MS glycomics rely on simplifying the complexity of heparan sulphates to determine disaccharide composition, which reduces heterogeneity but also severely limits structural and sequence information that underpin biological function. Here we took a different approach and developed an MS-compatible preparation method and systematic glycoinformatics of biologically relevant heparan sulphate oligosaccharides. This method represents a core method for application to any plasma sample (excluding heparinized plasma, which are not compatible for any type of heparan sulphate analysis due to its inherently similar biochemical structure with heparan sulphate) derived from patients with diverse conditions or diseases. Furthermore, the method can easily be adapted to profile chondroitin/dermatan sulphate through omission of the cABCase buffer and exchange of chondroitinase ABC for heparinase enzymes (I-III)[11].

Kidney disease affects more than 10% of the global population (>800 million individuals) and represents one of the leading causes of mortality worldwide, with associated deaths continuing to increase. MG is an autoimmune kidney disease. Our data also identifies a subset of MG patients. Recent data in the field has hinted towards a role for heparan sulphate in disease development and/or progression. Antibodies in MG patients to heparan sulfate[26] have been found in MG patients as well as the abnormal presence of the heparan sulfate copolymerase complex enzymes,

EXT1 and EXT2, in the glomerular basement membrane of a subset of MG patients with autoimmune disease features, suggesting disturbances in heparan sulfate expression may underpin facets of MG. In conclusion, we describe the first method for the profiling of intact HS oligosaccharides from plasma of healthy and kidney disease patients. Using these samples, we demonstrate how HS glycoprofiling generates an individual glyco-'fingerprint' for patients that may serve, along with other parameters, as a future clinical diagnostic tool. Our analyses also highlight heparan sulphate as a potential novel biomarker in MG patients that warrants further exploration.

REFERENCES

1. Delos, M., et al., Heparan sulfate 3-O-sulfotransferase 2 (HS3ST2) displays an unexpected subcellular localization in the plasma membrane. *Biochim Biophys Acta Gen Subj* **2018,** *1862* (7), 1644-1655.

2. Maciej-Hulme, M. L., et al., The role of heparan sulfate in host macrophage infection by Leishmania species. *Biochem Soc Trans* **2018,** *46* (4), 789-796.

3. Nam, E. J.; Park, P. W., Shedding of cell membrane-bound proteoglycans. *Methods Mol Biol* **2012,** *836*, 291-305.

4. Buijsers, B., et al., Increased Plasma Heparanase Activity in COVID-19 Patients. *Front Immunol* **2020,** *11*, 575047.

5. Loeven, M. A., et al., Selective Binding of Heparin/Heparan Sulfate Oligosaccharides to Factor H and Factor H-Related Proteins: Therapeutic Potential for C3 Glomerulopathies. *Front Immunol* **2021,** *12*, 676662.

6. Maciej-Hulme, M. L., et al., Glomerular endothelial glycocalyx-derived heparan sulfate inhibits glomerular leukocyte influx and attenuates experimental glomerulonephritis. *Front Mol Biosci* **2023,** *10*, 1177560.

7. Garsen, M., et al., Heparanase Is Essential for the Development of Acute Experimental Glomerulonephritis. *Am J Pathol* **2016,** *186* (4), 805-15.

8. Rops, A. L., et al., Modulation of heparan sulfate in the glomerular endothelial glycocalyx decreases leukocyte influx during experimental glomerulonephritis. *Kidney Int* **2014,** *86* (5), 932-42.

9. Laboux, T., et al., Hemolysis is associated with altered heparan sulfate of the endothelial glycocalyx and with local complement activation in thrombotic

microangiopathies. *Kidney Int* **2023,** *104* (2), 353-366.

10. Kim, H. J., et al., Anti-heparan sulfate antibody and functional loss of glomerular heparan sulfate proteoglycans in lupus nephritis. *Lupus* **2017,** *26* (8), 815-824.

11. Maciej-Hulme, M. L., et al., High sensitivity (zeptomole) detection of BODIPY-labelled heparan sulfate (HS) disaccharides by ion-paired RP-HPLC and LIF detection enables analysis of HS from mosquito midguts. *Anal Methods* **2023,** *15* (11), 1461-1469.

12. Sanderson, P., et al., Heparin/heparan sulfate analysis by covalently modified reverse polarity capillary zone electrophoresis-mass spectrometry. *J Chromatogr A* **2018,** *1545*, 75-83.

13. McCool, E. N., et al., Large-scale top-down proteomics using capillary zone electrophoresis tandem mass spectrometry. **2018,** (140), e58644.

14. Stickney, M., et al., Investigation of electrospray for a capillary electrophoresis–mass spectrometry interface in reverse polarity and negative ion mode. **2019,** *25* (1), 157-163.

15. Broadhurst, D., et al., Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. **2018,** *14* (6), 1-17.

16. Holman, J. D., et al., Employing ProteoWizard to convert raw mass spectrometry data. **2014,** *46* (1), 13.24. 1-13.24. 9.

17. Pluskal, T., et al., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. **2010,** *11* (1), 1-11.

18. Zhurov, K. O., et al., Distinguishing analyte from noise components in mass spectra of complex samples: where to cut the noise? **2014,** *86* (7), 3308-3316.

19. Pang, Z., et al., MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. **2021,** *49* (W1), W388-W396.

20. Han, X., et al., Structural analysis of urinary glycosaminoglycans from healthy human subjects. *Glycobiology* **2020,** *30* (3), 143-151.

21. Rabelink, T. J., et al., Heparanase: roles in cell survival, extracellular matrix remodelling and the development of kidney disease. *Nat Rev Nephrol* **2017,** *13* (4), 201-212.

22. Hippensteel, J. A., et al., Circulating heparan sulfate fragments mediate septic cognitive dysfunction. *J Clin Invest* **2019,** *129* (4), 1779-1784.

23. Bratulic, S., et al., Noninvasive detection of any-stage cancer using free glycosaminoglycans. *Proc Natl Acad Sci U S A* **2022,** *119* (50), e2115328119.

24. Barbas, A. S., et al., Heparan sulfate is a plasma biomarker of acute cellular allograft rejection. *PLoS One* **2018,** *13* (8), e0200877.

25. Koch, J., et al., Plasma syndecan-1 in hemodialysis patients associates with survival and lower markers of volume status. *Am J Physiol Renal Physiol* **2019,** *316* (1), F121-F127.

26. Boilard, E., et al., Interaction of low molecular weight group IIA phospholipase A2 with apoptotic human T cells: role of heparan sulfate proteoglycans. *Faseb j* **2003,** *17* (9), 1068-80.

CHAPTER 5

APPLICATIONS–HIGHLY SENSITIVE GLYCOSAMINOGLYCAN ANALYSIS FOR VARIOUS TYPES OF SAMPLES BY CAPILLARY ZONE ELECTROPHORESIS-MASS SPECTROMETRY

5.1 Introduction

Glycosaminoglycans, Heparan Sulfate, and Chondroitin Sulfate

Glycosaminoglycans (GAGs) are linear anionic polysaccharides that are present in the extracellular matrix and on the cell surface of mammalian cells. GAGs play many critical roles in regulating biological function and implicated in disease states as primary mediators of cell-environment interactions. [1, 2] These complex oligosaccharides are made up of disaccharides that contain amino sugar and hexuronic acid as building blocks. GAGs are classified into four major groups based on their building blocks and biosynthesis. Heparin/heparan sulfate (Hp/HS), Chondroitin/dermatan sulfate (CS/DS), keratan sulfate (KS), and hyaluronic acid (HA). [1] HS accounts for 50-90% of the total pool of GAGs and is among the most abundant GAG type. Hp is most frequently found in mast cells and is known to be an anticoagulant. Both Hp and HS, play important roles in cell growth/development, inflammation, angiogenesis, cardiovascular disorders, viral infection, and cancer. [3] CS plays a number of important biological functions in the cell, such as cell-to-cell adhesion, cell signaling, and the development of the central nervous system due to its inherent sulfate group motif. Besides, more than 80% of the GAGs in articular cartilage are composed of CS, so it has gained attention as a pharmacological product

for treating osteoarthritis and cataracts. [4, 5] For example, in old cartilage samples, higher levels of sulfation were observed at O6 than O4 positions of GalNAc suggesting that sulfation distribution is critical to its function. [6] Further, a structure-specific nonreducing end terminating with GalNAc 6-sulfate can be utilized as an effective diagnostic biomarker for mucopolysaccharidosis type IVA, a genetic disorder caused by lysosomal hydrolase degradation of GAGs (6-*O*-exosulfatase) which is a more sensitive and specific MS approach than colorimetric approaches. [7]

Biosynthesis of Heparan Sulfate and Chondroitin Sulfate

GAG biosynthesis is a non-template-driven process that occurs by a number of enzymatic steps, such as chain elongation, deacetylation, sulfation, and epimerization, resulting in highly heterogeneous oligosaccharides. Assembly of either HS or CS chains on a protein, to form a proteoglycan, begins with the addition a tetrasaccharide linker in the endoplasmic reticulum and continues with glycan elongation and modification in the Golgi apparatus. The linker tetrasacchride, GlcAβ1–3Galβ1–3Galβ1–4Xylβ1–O-Ser, is common to HS and CS. The next transferred hexosamine to the non-reducing end side of the linker is critical in determining which family of GAG will be produced; GalNAc by GalNAc transferase I for CS or GlcNAc by GlcNAc transferase for HS. Following the GAG tetrasaccharide linker synthesis, modifications are made in a spatially and temporally defined manner. As the elongating chain passes through lamella of the Golgi, enzymes that are located in specific regions can modify the growing chain in sequential fashion, resulting in mature and functional GAG molecules that carry out specific biological function. In addition, extracellular sulfatase and/or glycosidase enzymes can remodel the GAGs by removing sulfate groups or sugar residues. The

final pattern of modifications creates specific binding partners for various proteins on the cell surface or in the intracellular space, and play a significant role in binding and modulating biological processes. [8, 9]

Hp possesses the most dense sulfation motif of all GAGs. HS is produced by the same biosynthetic pathways as Hp, but bears fewer modifications overall. HS presents regions of high sulfation interspersed with less or non-sulfated regions, giving rise to characteristic sequenced sulfation patterns and biological functions. [10, 11] There has been considerable progress in identifying and understanding the modification process involved in GAG biosynthesis. For instance, N-deacetylase/N-sulfotransferases design the HS chain based on the N-sulfation process. [8] As in case of cell surface sulfatases, endosulfatases, such as SULF1 and SULF2, regulate HS-protein interaction, signaling pathways, and growth factor activation by removing 6-O-sulfate groups from HS chains. As a result, SULF-mediated HS sulfate group decomposition plays an important role in cell signaling resulting in tissue-specific HS fine structures. [12] CS sulfatases participate in specific sulfation motif in the chain for biological functions. For example, endosulfatases, catalyze the hydrolysis of the sulfate ester on GalNAc/unsaturated hexuronic acid residues. [13, 14]

Indeed, the polymerization and modification of GAGs are controlled by enzymes encoded by more than 40 genes in mammalian cells. [15] Several researchers have discovered not only the activity of biosynthesis enzymes, through gene targeting experiments to reveal substrate specificity, but also the diversity of GAG chains and how spatiotemporal differences are caused by genetic regulatory factors from the precise decoding of GAG informative fragments. Researchers have used clustered regularly interspersed short palindromic repeat gene editing technology employing the RNA-guided nuclease Cas9, referred to as CRISPR/Cas9, to selectively inactivate

specific enzymes in the Hp/HS biosynthetic pathway, to gain a better understanding of how the pattern of modifications in these GAGs contribute to their biological activity. [16, 17] Since biological GAG samples possess heterogeneous properties and are complex in nature, structural identification can be challenging. [18-20] The middle-down method to characterize polysaccharide GAGs involves partial digestion into oligosaccharides of suitable length, from disaccharides to decasaccharides, in contrast to top-down analysis of full-length sulfated GAG chains without depolymerization. [21] Over the past few decades, a number of studies have been conducted to examine the structure of GAGs, particularly by mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR). However, NMR has been hampered by low sample abundance and low sensitivity. As opposed to NMR, MS techniques are sensitive enough to detect GAGs in pmol amounts and can identify complex mixtures without deconvolving overlapping signals as in NMR. [2] Therefore, MS can detect and analyzing GAG assemblies in biological samples with high accuracy and sensitivity.

Analytical Separation Methods for GAGs and CZE-MS

The pre-separation of complex GAG mixtures can be performed either on-line or off-line. [22] Several methods have been utilized for separating heterogenous GAGs, such as size exclusion chromatography (SEC), strong anion exchange (SAX) chromatography[23], liquid chromatography (LC) ─including hydrophilic interaction liquid chromatography (HILIC) [24] and reversed phase ion paring chromatography (RP-IP) [25]─, porous graphitized carbon chromatography (PGC) [26], and capillary zone electrophoresis (CZE). [27, 28] Considering that different levels of sulfation and even isomeric GAGs can coexist in mixtures, multidimensional separation techniques can be used to separate them into distinct components. SAX utilizes reagent cations at

high concentrations that can suppress ions prior to the MS analysis if not removed. [29] In this context, the CZE-MS method has the advantage of not requiring high concentrations of cation salt, which results in ion suppression that prevents direct coupling with MS regardless of the high isomer resolution. [27, 28, 30] Besides, HILIC resolves analytes according to their degree of polymerization (dp), it causes an incomplete separation of some isomers. [31, 32] Alkylammonium salts derived from mobile phase modifiers of RP-IP often make difficult to couple RP-IP directly with MS because of clustering ions with anionic analytes, degrading its performance. [33] Despite its good chemical stability, PGC has recovery issues. [34]

CZE has been utilized to examine anionic biomolecules for many years, including metabolites and nucleic acids. [35, 36] Although CZE analysis of GAGs offers a number of advantages as mentioned above, there is still a need for further development of this methodology. In a normal (called also as forward or direct) polarity mode, a positive separation voltage is applied, creating an electroosmotic force (EOF) flowing towards the outlet of a bare fused silica (BFS) capillary. Early studies of CZE analysis of GAGs focused primarily on the normal mode polarity, to allow the EOF to reduce migration times. [37, 38] However, GAGs are highly ionic, and are more easily analyzed using reverse polarity, in which a negative potential is used to create an electrophoretic force (EF) that pushes the analyte toward the outlet of the capillary. However, the migration is hampered in BFS capillaries by the EOF, which is attracted toward the negative potential of the capillary inlet, and therefore opposes the migration of the anionic analytes. Neutral and cationic coated separation capillaries have been used to improve separation speed and enhance CZE resolution by eliminating EOF (neutral coating) or reversing the EOF (cationic coating) so that it flows toward the outlet. [27, 28] In the latter case, the EF and EOF both move analytes in

the direction of the outlet, shortening the migration time of anionic GAGs and enabling the detection of narrow peaks. [39, 40] The migration time of analytes is affected by the electrophoretic force as well as EOF, caused by charge separation between the separation medium and capillary inner walls. In the reverse polarity with a bare fused silica capillary, the movement of EOF opposes that of the EF because the double layer present on the inner surface of the capillary, leading to slow-migrating analytes or low resolution (because of inadequate spray conditions and inefficient analyte ionization). [39, 41] In order to improve the repeatability of migration times of analytes and robustness of capillaries, chemically modified capillaries have been utilized, such as neutral (dichlorodimethylsilane; DMS or linear polyacrylamide; LPA) and cationic (aminomethyltriethoxysilane; AHS) coating materials through covalent bonds since the movement of EOF is added to that of the EF in the same direction towards the MS inlet, leading to comparatively short migration times compared to a non-coated capillary separation. [28, 42, 43] This allows the EOF movement to be inactivated/severely reduced with a neutral coated capillary or to be added along with the EF of anionic analytes in the direction of the MS interface at the capillary outlet, thus facilitating fast analysis of anionic analytes with a cation coated capillary as shown in Figure 15. CZE is also compatible with electrospray ionization (ESI) which is commonly used in GAG analysis. Besides, GAGs are hydrophilic and anionic biomolecules. These have led to the successful use of CZE to separate and characterize GAG oligosaccharides.

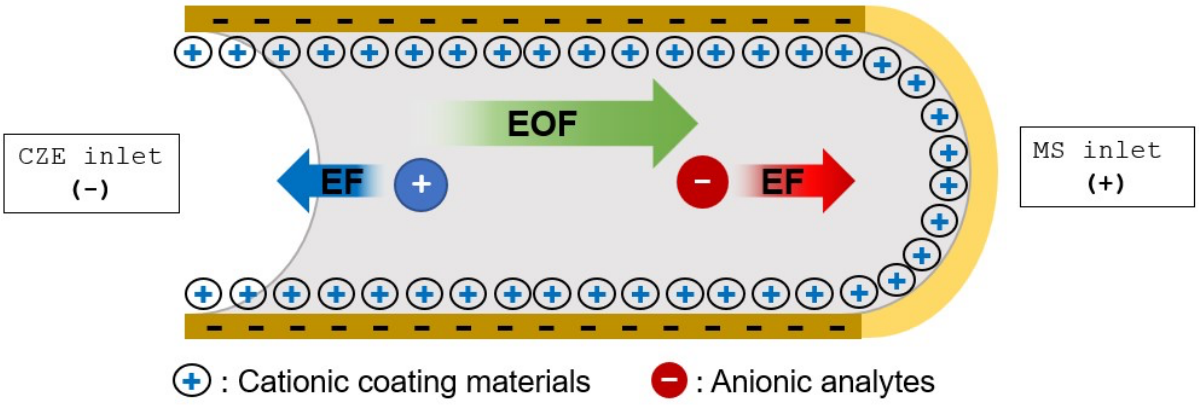

Figure 38. Schematic of the interior of a cation coated capillary in the system of capillary zone electrophoresis-mass spectrometry (CZE-MS).

In this study, we utilized a reverse-polarity CZE-MS system to analyze HS oligosaccharides derived from A375 human melanoma cells (both wild-type and knockout). In addition, GAGwizard, an in-house MATLAB code, assigned migrated GAGs to demonstrate that HS the sulfate group modification and HS biosynthesis were significantly different among the two samples.

5.2 Experimental Method

Preparation of GAG Samples for CZE-MS Analysis-Heparan Sulfate

A375 cells (ATCC CRL-1619) were grown in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% (v/v) FBS and 1% (v/v) penicillin/streptomycin at 37 °C under an atmosphere of 5% CO2/95% air. Each sample was subcultured every three to four days and was revived from liquid nitrogen after ten passages. In order to store transfected cell lines in liquid nitrogen, all transfected lines were cloned. Lentiviral particles were collected and added to A375 cells followed by selection with 2 μg/ml blasticidin to generate Cas9-expressing cells. A375 EZH2 mutant cell lines were generated by ligation of sgRNAs targeting human

EZH2 into the LentiGuide-Puro vector (a gift from Feng Zhang and purchased from Addgene, #52963). Cell harvest and chromatographic extraction were the same as in Chapter 3. All samples were desalted based on the previous method and were transferred to a 250 µL CE auto sampler vial. [105]

Chondroitin Sulfate

For the preparation of chemoenzymatically synthesized CS,10 mg of native bikunin peptide (AVLPQEEEGSGGGQLVTEVTKKEDS, 3.86 µmol) was treated with 3,4-Diethoxy-3-cyclobutene-1,2-dione (9 equiv.) in MeOH/carbonate buffer (v/v 1:1, 50 µL, pH 8), mixture was incubated for 6h until no starting material left, resulting mixture was dried in vacuo then redissolve in 1 mL of carbonate buffer (pH 8). To a carbonate buffer containing peptide conjugate, 0.85 mL of drained EAH sepharose (2 equiv.) was added and it was shaken in RT for 1 day until no starting material left. To construct the tetrasaccharide linkage region to initiate a glycopeptide, peptide conjugated EAH sepharose was sequentially rotated at RT with XT-1 (100 µg), UDP-xylose (2 equiv.) in 4 mL of XT-1 buffer; B4GALT7 (250 µg), UDP-galactose (2 equiv.) in 4 mL of B4GALT7 buffer; FAM20B (200 µg, ATP (3 equiv.) in 4 mL of FAM20B buffer; B3GALT6 (200 µg), UDP-galactose (1.1 equiv.) in 4 mL of B3GALT6 buffer; rSAP (100 U) in 4 mL of rCutSmart™ Buffer; B3GAT3 (500 µg), UDP-glucuronic acid (2 equiv.) in B3GAT3 buffer. To further construct chondroitin sulfate containing glycopeptide, sepharose from previous steps was sequentially treated with KfoC (100 µg), UDP-N-acetylgalactosamine (2 equiv.) and KfoC (100 µg), UDP-glucuronic acid for 2,3 ,4 rounds, resulting in 8, 10, 12 mer chondroitin sulfate chains covalently attached to the bikunin peptide. These glycopeptides were further sulfated with CS4OST (400 µg) and 3’ -

Phosphoadenosine-5'-phosphosulfate (2 equiv.). Upon completion, agarose was treated with 1% hydrazine, resulting liquid were dried in vacuo then purify with C18 equipped HPLC.

2-Aminobenzamide Derivatization on CS samples

Solvents used in the experiment were 0.047 grams 2-aminobenzamide (2-AB), 300 mL DMSO, 700 mL acetic acid, and 0.062 grams sodium cyanoborohydride and were dissolved in a 1.5 mL Eppendorf tube. Disaccharide standards and biological CS samples from canine cartilage were mixed into 10 mL of solvent solution. The disaccharide sample was left to react at 65 °C for 2.5 hours to react. A LudgerClean S Cartridge was used to extracted and wash the mixture to get rid of excess reagents and ensure that only the GAGs were present at the end. To prepare the CZE-MS solution, 50 μL of water was added to the condensed sample. All samples were received as a gift from the University of Pennsylvania except disaccharide standards (from Iduron, Alderley Edge, UK).

5.3 Results

5.3.1. Probing the Biosynthetic Pathway of Heparan Sulfate using Capillary Zone Electrophoresis Mass Spectrometry

Approach to Identifying HS Isolated from Cancer Cells Based on CZE-MS

In addition to enzymatic factors in GAG biosynthesis, nonenzymatic factors have also been linked to elaborate regulatory pathway and expression of the enzymes and core proteins, leading to heterogeneous GAG structures with certain sulfate group modification motifs. Many biological functions are primarily modulated by the degree

of sulfation and the position of the sulfate groups based on their selective binding to proteins. [56-58] The application of genomics and bioinformatics has enabled the study of GAG regulatory pathways, which modulate the epigenetic gene expression that regulates the function of GAGs in relation to disease progression and development. CRISPR genome-wide screening has emerged as an important tool for identifying significant regulatory factors for HS expression in human cells. [11, 17] Here we have focused the development of our CZE-MS platform for identifying HS isolated from human cancer cells. As shown in Figure 39, a wild type of cell lines and a mutant cell lines samples were prepared then washed in PBS, digested with Pronase, and desalted. Regarding enzymatic depolymerization for HS, DNase I was applied to the desalted products, and the samples were purified. Lyophilized GAGs were partially depolymerized with a mixture of heparin lyases I-III (for HS samples) or Chondroitinase ABC (for CS samples). Lastly, electropherograms obtained for the two cell-isolated HS samples through CZE-MS are illustrated at the last stage in Figure 39. Data analysis using GAGwizard provides HS assignment regarding the type/length of chains, the number of modification and more.

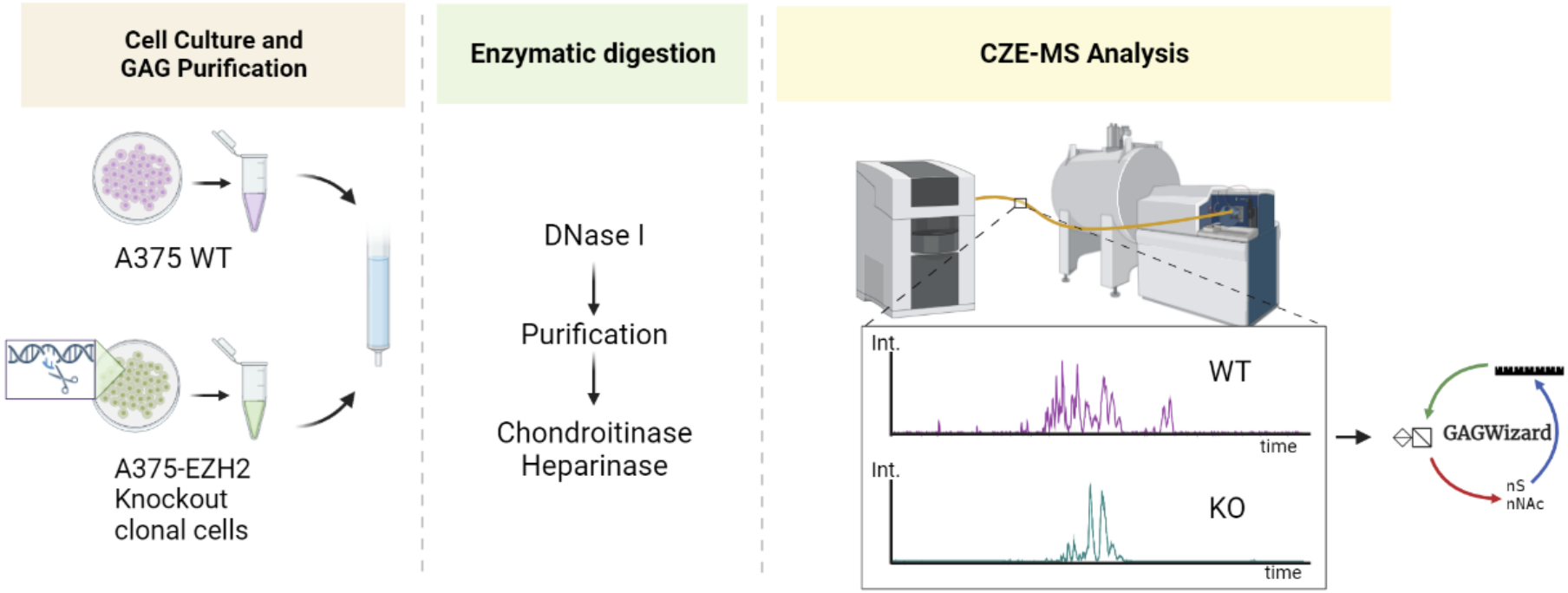

Figure 39. Schematic representation of our CZE-MS workflow to elucidate the distinct regulatory factors involved in GAG assembly isolated from cells with genomic technology.

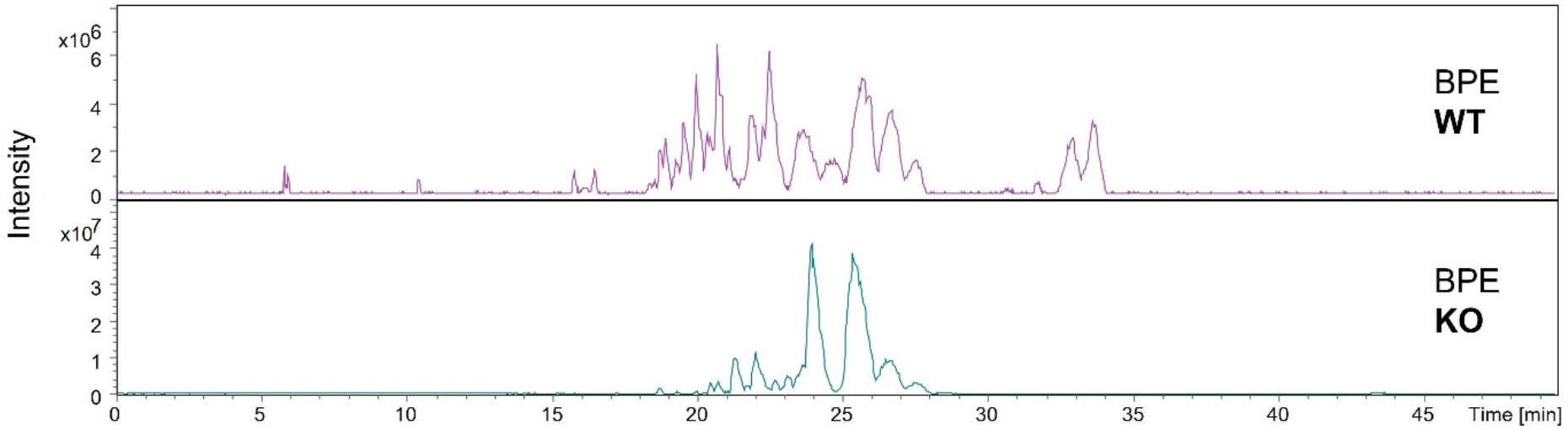

Figure 40. Base peak electropherograms (BPEs) obtained for HS samples isolated from the wild-type and knockout-type human cancer cells. Detected ions were ranging from m/z 400 to 2000. A 9.4T Bruker SolariX XR mass spectrometer coupled with CMP ECE-001 capillary zone electrophoresis was employed for both.

Characterization of HS Isolated from Purified Wild type Cells

CZE-MS was utilized to examine GAG samples from both WT and KO in negative mode of the mass analyzer. The quality of separation and the abundances of detected ions were dependent on the adjustment of an in-house pulled a glass emitter tip. To ensure the reproducibility of the experiment, the glass emitter tip had an opening of 30-40 μm. Optimal parameters for the FT-ICR mass spectrometer and CZE system were employed for the best performance based on results discussed in the previous subchapters.

Figure 41 shows a base peak electropherogram measured for the isolated GAG from the WT sample using a DMS-functioned capillary, and most analytes were baseline separated in less than 35 min with a sufficient resolution. As a result, our CZE-MS system using 60 cm long DMS-functioned capillary produced a 20 min separation window for the analysis of WT HS sample and identified 48 HS features in a single run as in Table 7. Most of the ions detected during separation were doubly charged species; however, singly charged ions were detected at the beginning and end of migration time. Besides, few positional isomers of triply charged ions were observed (m/z 748.75).

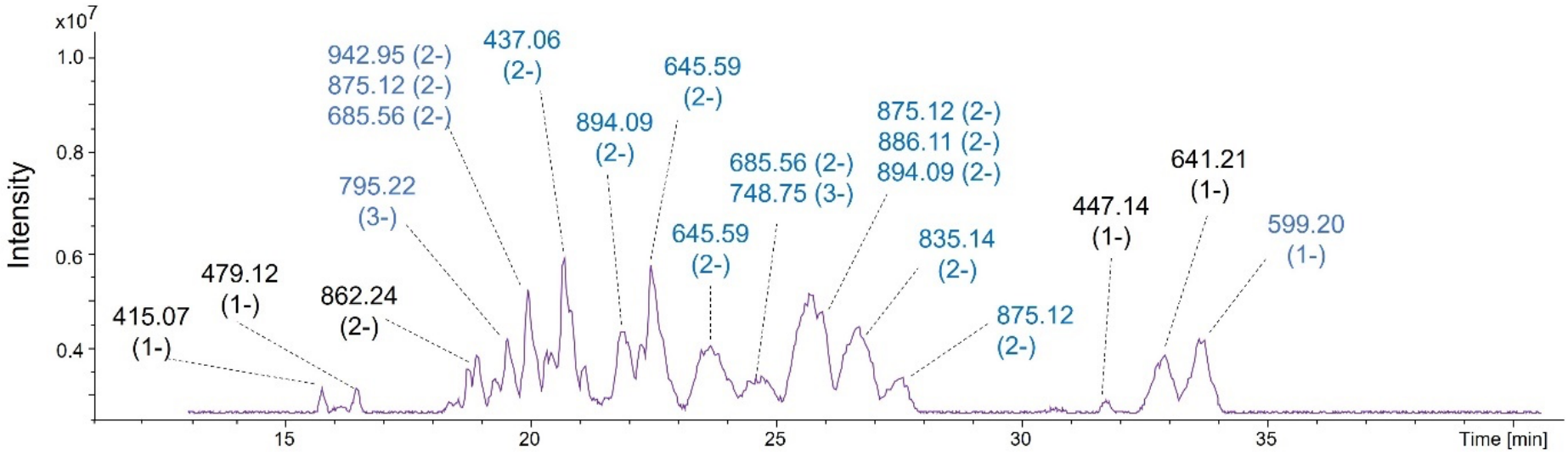

Figure 41. A base peak electropherogram of a wide-type HS sample after CZE-MS analysis with a 60 cm long DMS-coated capillary. Assigned ions to GAG compositions were represented in blue. Each peak includes the information of neutral mass and charges.

m/z	z	Intensity	RT	Neutral Mass	DP	ΔU	U	AS	NAc	SO3	Mass Error	# Adducts	Adduct Type	Linker	Formula
437.0564	2	6459552	20.71	876.13	4	1	1	2	1	2	2.73	0	none	No linker	C26H40O27N2S2
437.0566	2	3931299	19.96	876.13	4	1	1	2	1	2	3.11	0	none	No linker	C26H40O27N2S2
599.1955	1	3259956	33.61	600.20	3	0	1	2	2	0	3.24	0	none	No linker	C22H36O17N2S0
645.5851	2	2958554	23.69	1293.19	6	1	2	3	1	3	1.56	0	none	No linker	C38H59O40N3S3
645.5854	2	6183991	22.48	1293.19	6	1	2	3	1	3	2.09	0	none	No linker	C38H59O40N3S3
645.5864	2	1973031	19.88	1293.19	6	1	2	3	1	3	3.59	0	none	No linker	C38H59O40N3S3
666.5919	2	1071845	19.88	1335.20	6	1	2	3	2	3	3.78	0	none	No linker	C40H61O41N3S3
666.5920	2	1683277	19.29	1335.20	6	1	2	3	2	3	3.93	0	none	No linker	C40H61O41N3S3
685.5637	2	1752068	24.72	1373.14	6	1	2	3	1	4	1.74	0	none	No linker	C38H59O43N3S4
685.5642	2	3033048	22.28	1373.14	6	1	2	3	1	4	2.52	0	none	No linker	C38H59O43N3S4
685.5645	2	1714663	20.59	1373.14	6	1	2	3	1	4	2.92	0	none	No linker	C38H59O43N3S4
685.5649	2	1822620	19.92	1373.15	6	1	2	3	1	4	3.55	0	none	No linker	C38H59O43N3S4
722.0972	3	683003	22.71	2169.31	10	1	4	5	2	5	2.38	0	none	No linker	C64H99O67N5S5
734.7464	3	839655	20.90	2207.26	10	1	4	5	1	6	3.03	0	none	No linker	C62H97O69N5S6
748.7490	3	599953	23.20	2249.27	10	1	4	5	2	6	1.77	0	none	No linker	C64H99O70N5S6
748.7496	3	538815	22.48	2249.27	10	1	4	5	2	6	2.49	0	none	No linker	C64H99O70N5S6
748.7501	3	535244	20.90	2249.27	10	1	4	5	2	6	3.19	0	none	No linker	C64H99O70N5S6
761.3986	3	893986	21.88	2287.22	10	1	4	5	1	7	2.85	0	none	No linker	C62H97O72N5S7
761.3988	3	882049	21.53	2287.22	10	1	4	5	1	7	3.04	0	none	No linker	C62H97O72N5S7
761.3990	3	1592562	20.31	2287.22	10	1	4	5	1	7	3.31	0	none	No linker	C62H97O72N5S7
795.1633	1	1847504	26.47	796.17	4	1	1	2	1	1	2.38	0	none	No linker	C26H40O24N2S1
835.1421	2	3749236	26.70	1672.30	8	1	3	4	2	3	2.71	0	none	No linker	C52H80O51N4S3
854.1158	2	612820	26.78	1710.25	8	1	3	4	1	4	3.38	0	none	No linker	C50H78O53N4S4
854.1161	2	2159135	21.14	1710.25	8	1	3	4	1	4	3.71	0	none	No linker	C50H78O53N4S4
861.1164	3	569553	22.75	2586.37	12	1	5	6	2	6	1.99	0	none	No linker	C76H118O80N6S6

875.1202	2	5061762	25.68	1752.26	8	1	3	4	2	4	2.26	0	none	No linker	C52H80O54N4S4
875.1204	2	1700932	27.59	1752.26	8	1	3	4	2	4	2.55	0	none	No linker	C52H80O54N4S4
875.1209	1	3287724	20.86	876.13	4	1	1	2	1	2	3.14	0	none	No linker	C26H40O27N2S2
875.1214	1	5221122	19.96	876.13	4	1	1	2	1	2	3.66	0	none	No linker	C26H40O27N2S2
887.7696	3	1237155	22.40	2666.33	12	1	5	6	2	7	2.96	0	none	No linker	C76H118O83N6S7
894.0939	2	3534185	21.92	1790.20	8	1	3	4	1	5	2.92	0	none	No linker	C50H78O56N4S5
894.0942	2	2763331	20.35	1790.20	8	1	3	4	1	5	3.25	0	none	No linker	C50H78O56N4S5
894.0942	2	1663912	20.79	1790.20	8	1	3	4	1	5	3.25	0	none	No linker	C50H78O56N4S5
894.0944	2	893130	25.72	1790.20	8	1	3	4	1	5	3.38	0	none	No linker	C50H78O56N4S5
894.0944	2	922003	21.30	1790.20	8	1	3	4	1	5	3.38	0	none	No linker	C50H78O56N4S5
894.0949	2	1419704	20.11	1790.21	8	1	3	4	1	5	3.99	0	none	No linker	C50H78O56N4S5
920.4313	3	449919	22.40	2764.32	12	0	6	6	2	8	9.31	0	none	No linker	C76H120O87N6S8
934.0719	2	620646	21.65	1870.16	8	1	3	4	1	6	2.28	0	none	No linker	C50H78O59N4S6
1083.6493	2	565084	25.09	2169.31	10	1	4	5	2	5	2.07	0	none	No linker	C64H99O67N5S5
886.1116	2	633682	25.80	1774.24	8	1	3	4	2	4	2.70	1	Na	No linker	C52H79O54N4S4Na1
920.4313	3	449919	22.40	2764.32	12	1	5	6	5	5	6.06	6	Na	No linker	C82H118O80N6S5Na6
781.0782	3	752559	20.19	2346.26	10	1	4	5	2	7	3.51	1	NH3	No linker	C64H102O73N6S7
942.5863	2	911514	21.61	1887.19	8	1	3	4	1	6	3.51	1	NH3	No linker	C50H81O59N5S6
942.5869	2	2200669	20.07	1887.19	8	1	3	4	1	6	4.14	1	NH3	No linker	C50H81O59N5S6
795.2167	3	2931733	19.56	2388.67	12	1	5	6	6	1	2.06	2	NH3	No linker	C84H132O69N8S1
888.7704	3	598964	22.08	2669.33	11	0	5	6	0	9	2.67	5	NH3	No linker	C66H123O82N11S9
1042.5077	3	467803	21.36	3130.55	14	1	6	7	3	7	6.88	5	NH3	No linker	C90H154O94N12S7
1074.8404	3	520909	25.96	3227.54	14	1	6	7	3	8	2.09	6	NH3	No linker	C90H157O97N13S8

Table 7. Assigned GAG compositions for the WT HS sample using CZE-MS. MZmine 3 was utilized for electropherogram deconvolution and peak picking.

All 48 distinct m/z values that were assigned are shown in Table 7. Our results indicate that GAG isomers with the same m/z value and composition are observed, suggesting that the sulfate group modifications are located at different residue locations. Specifically, the composition of oligosaccharides with dp 3 to dp 12 were assigned, with the number of sulfate group modification ranging from 0 to 8. This is in good agreement with the heparinase digestion that was optimized to yield GAG chains ranging from hexasaccharides to decasaccharides.[59] The neutral masses of the assigned GAG chains ranged from 600 to 3227 Da with less than 2.0 ppm mass error. Previously, Sanderson et al. reported that the length of GAG oligosaccharides influences their migration time, with shorter GAG chains having shorter migration times.[28] Overall longer GAGs migrated later in this experiment. Multiple migration peaks of the same composition were observed, arising from the isotopic complexity of sulfate modification. Additionally, it was common to observe single-charged ions with low intensity near the beginning and end of separations that correspond to short chains, such as an unusual trisaccharides.[60] CZE-MS involves the formation of sodium and ammonium adducts due to its BGE formulation, sodium and ammonia adduct were detected in nine of the 48 assigned compositions.[28] A lower level of two types of adduct was observed: (1) Only two sodium adducts were identified as the samples were desalted in the sample preparation step. (2) The use of an ammonium formate BGE can explain seven ammonia adducts. Although ammonium adduction leads to the reduction of the charge state, alkali ion heterogeneity (Na^+/H^+) is capable of controlling ion activation to improve the efficiency of MS/MS with on-line separation methods.[2, 61] Future research should try to exploit the appearance of adducts for CZE with tandem MS.

Characterization of HS Isolated from Purified Knock-out Cells

In Figure 42, the nano-ESI mass spectrum of the HS sample isolated from EZH2 KO was obtained without any prior separation, revealing the mixture's inherent complexity. Figure 43 shows the result of CZE-MS analysis using a DMS-coated capillary. The base peak electropherogram of the HS sample isolated from the KO sample showing a number of features.

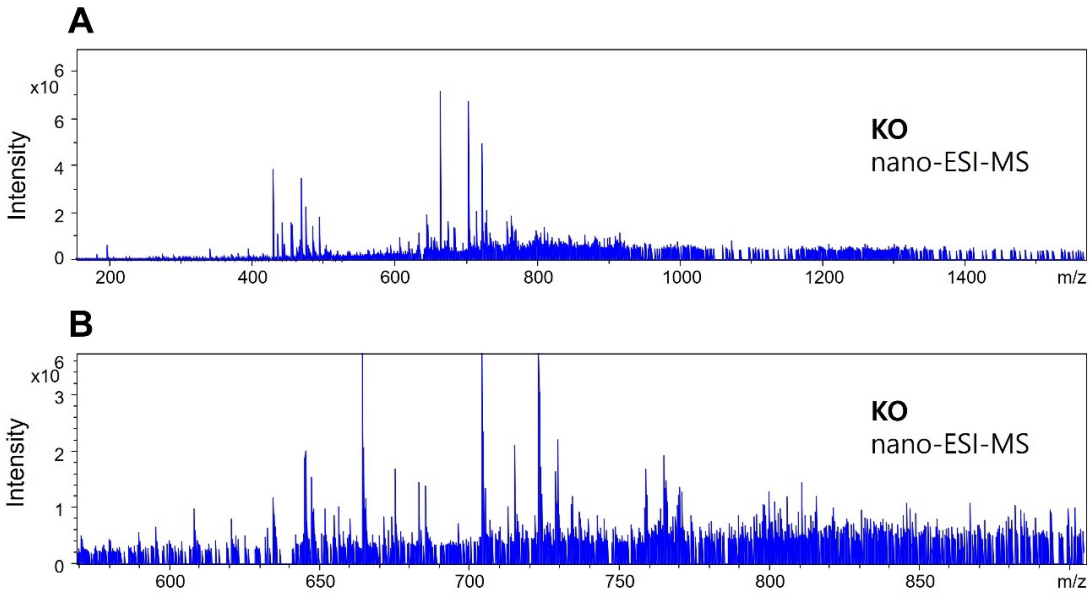

Figure 42. (A) an nano-ESI mass spectrum for a HS mixture isolated from the EZH2 knockout sample. (B) The bottom spectrum shows expanded spectra of A ranging in mass-to-charge ratio (m/z) from 550 to 900.

Separation was observed between approximately 18 min to 30 min with several major peaks exhibiting good separation efficiency. In accordance with the previous WT sample, more doubly charged ions were present than singly charged ions. Triply charged ions, on the other hand, were more prevalent in the KO sample suggesting a higher level of sulfation. Two major peaks are discernible at 24 min and

25.5 min, respectively. The masses of the components were used as input to GAGwizard, to assist with assigning GAG compositions to the long list of candidates.

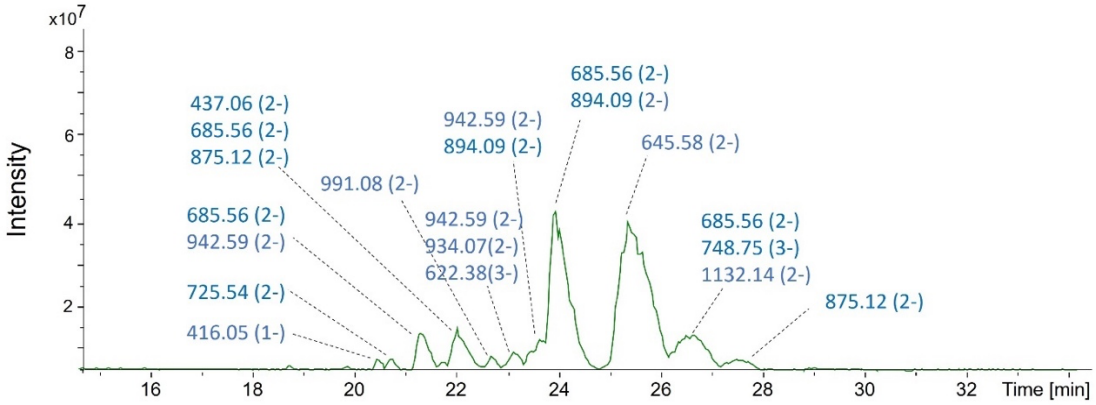

Figure 43. A base peak electropherogram of a KO HS sample after CZE-MS analysis. Assigned ions to GAG compositions were in blue. Each peak includes the information of neutral mass and charges. A star mark denotes that the theoretical number of plates was 25,600 for the highest peak.

m/z	z	Intensity	RT	Neutral Mass	DP	ΔU	U	AS	NAc	SO3	Mass Error	# Adducts	Adduct Type	Linker	Formula
416.0506	1	3323308	20.44	417.10	2	1	0	1	0	1	2.26	0	none	No linker	C12H19O13N1S1
437.0560	2	11300000	22.02	876.10	4	1	1	2	1	2	1.89	0	none	No linker	C26H40O27N2S2
456.7068	3	1356602.5	23.91	1373.10	6	1	2	3	1	4	2.46	0	none	No linker	C38H59O43N3S4
595.7265	3	1057676.8	24.07	1790.20	8	1	3	4	1	5	2.65	0	none	No linker	C50H78O56N4S5
622.3784	3	1187924.8	21.27	1870.20	8	1	3	4	1	6	1.89	0	none	No linker	C50H78O59N4S6
622.3785	3	850235.3	22.53	1870.20	8	1	3	4	1	6	2.01	0	none	No linker	C50H78O59N4S6
622.3786	3	2681599.2	23.67	1870.20	8	1	3	4	1	6	2.21	0	none	No linker	C50H78O59N4S6
645.5851	2	38500000	25.37	1293.20	6	1	2	3	1	3	1.57	0	none	No linker	C38H59O40N3S3
645.5853	2	14700000	24.11	1293.20	6	1	2	3	1	3	1.89	0	none	No linker	C38H59O40N3S3
649.0308	3	2589969.8	22.69	1950.10	8	1	3	4	1	7	2.00	0	none	No linker	C50H78O62N4S7
664.5583	2	1357306.2	25.60	1331.10	6	1	2	3	0	4	1.63	0	none	No linker	C36H57O42N3S4
685.5636	2	9930246	21.31	1373.10	6	1	2	3	1	4	1.68	0	none	No linker	C38H59O43N3S4
685.5637	2	9934126	22.02	1373.10	6	1	2	3	1	4	1.75	0	none	No linker	C38H59O43N3S4
685.5638	2	41000000	23.95	1373.10	6	1	2	3	1	4	1.88	0	none	No linker	C38H59O43N3S4
685.5638	2	2525267	23.00	1373.10	6	1	2	3	1	4	1.99	0	none	No linker	C38H59O43N3S4
685.5639	2	4174789.8	23.52	1373.10	6	1	2	3	1	4	2.09	0	none	No linker	C38H59O43N3S4
704.5372	2	1267308.9	23.99	1411.10	6	1	2	3	0	5	2.30	0	none	No linker	C36H57O45N3S5
725.5420	2	3421556.5	20.72	1453.10	6	1	2	3	1	5	1.54	0	none	No linker	C38H59O46N3S5
725.5421	2	2114064.5	25.09	1453.10	6	1	2	3	1	5	1.72	0	none	No linker	C38H59O46N3S5
725.5423	2	3078493.8	22.06	1453.10	6	1	2	3	1	5	2.00	0	none	No linker	C38H59O46N3S5
748.7494	3	1059741.1	26.35	2249.30	10	1	4	5	2	6	1.77	0	none	No linker	C64H99O70N5S6
761.3978	3	699981.56	21.66	2287.20	10	1	4	5	1	7	1.72	0	none	No linker	C62H97O72N5S7
761.3979	3	2224467	22.14	2287.20	10	1	4	5	1	7	1.88	0	none	No linker	C62H97O72N5S7
761.3979	3	966722	23.12	2287.20	10	1	4	5	1	7	1.91	0	none	No linker	C62H97O72N5S7

761.3981	3	2006063.1	23.55	2287.20	10	1	4	5	1	7	2.20	0	none	No linker	C62H97O72N5S7
775.4014	3	479602.12	25.13	2329.20	10	1	4	5	2	7	1.85	0	none	No linker	C64H99O73N5S7
788.0503	3	1531770.6	21.35	2367.20	10	1	4	5	1	8	2.04	0	none	No linker	C62H97O75N5S8
788.0503	3	686068.1	22.81	2367.20	10	1	4	5	1	8	2.05	0	none	No linker	C62H97O75N5S8
788.0505	3	981539.3	23.20	2367.20	10	1	4	5	1	8	2.26	0	none	No linker	C62H97O75N5S8
875.1198	1	4545366	22.02	876.10	4	1	1	2	1	2	1.83	0	none	No linker	C26H40O27N2S2
875.1198	1	4587531	22.17	876.10	4	1	1	2	1	2	1.87	0	none	No linker	C26H40O27N2S2
875.1209	2	3239393	27.49	1752.30	8	1	3	4	2	4	3.12	0	none	No linker	C52H80O54N4S4
887.7695	3	898484.3	23.75	2666.30	12	1	5	6	2	7	2.82	0	none	No linker	C76H118O83N6S7
894.0932	2	2575108	21.74	1790.20	8	1	3	4	1	5	2.06	0	none	No linker	C50H78O56N4S5
894.0932	2	5370115	23.48	1790.20	8	1	3	4	1	5	2.10	0	none	No linker	C50H78O56N4S5
894.0932	2	1459534.9	21.31	1790.20	8	1	3	4	1	5	2.14	0	none	No linker	C50H78O56N4S5
894.0933	2	6949969.5	24.11	1790.20	8	1	3	4	1	5	2.21	0	none	No linker	C50H78O56N4S5
894.0934	2	1590133.2	22.25	1790.20	8	1	3	4	1	5	2.30	0	none	No linker	C50H78O56N4S5
934.0711	2	2785940.2	23.12	1870.20	8	1	3	4	1	6	1.49	0	none	No linker	C50H78O59N4S6
934.0713	2	1098529.4	21.27	1870.20	8	1	3	4	1	6	1.68	0	none	No linker	C50H78O59N4S6
934.0715	2	1541967.1	22.49	1870.20	8	1	3	4	1	6	1.84	0	none	No linker	C50H78O59N4S6
934.0718	2	1943070.4	23.59	1870.20	8	1	3	4	1	6	2.25	0	none	No linker	C50H78O59N4S6
1145.3365	2	549437.5	43.39	2292.70	12	0	6	6	6	0	4.04	0	none	No linker	C84H128O67N6S0
656.5765	2	1193867.6	25.68	1315.20	6	1	2	3	1	3	2.19	1	Na	No linker	C38H58O40N3S3Na1
696.5555	2	920501.56	23.91	1395.10	6	1	2	3	1	4	2.87	1	Na	No linker	C38H58O43N3S4Na1
793.7258	3	1160397.1	21.31	2384.20	10	1	4	5	1	8	2.00	1	NH3	No linker	C62H100O75N6S8
793.7258	3	1484137.5	22.85	2384.20	10	1	4	5	1	8	2.01	1	NH3	No linker	C62H100O75N6S8
793.7259	3	2431803.5	23.28	2384.20	10	1	4	5	1	8	2.11	1	NH3	No linker	C62H100O75N6S8
820.3779	3	891032.4	20.99	2464.20	10	1	4	5	1	9	1.67	1	NH3	No linker	C62H100O78N6S9

820.3782	3	903769.6	22.92	2464.20	10	1	4	5	1	9	2.07	1	NH3	No linker	C62H100O78N6S9
902.6070	2	1344402.9	24.07	1807.20	8	1	3	4	1	5	2.66	1	NH3	No linker	C50H81O56N5S5
920.0963	3	837700.2	23.44	2763.30	12	1	5	6	2	8	1.70	1	NH3	No linker	C76H121O86N7S8
923.6111	2	1217847	25.09	1849.20	8	1	3	4	2	5	1.37	1	NH3	No linker	C52H83O57N5S5
932.7454	3	754962.1	23.36	2801.30	12	1	5	6	1	9	2.06	1	NH3	No linker	C74H119O88N7S9
942.5850	2	5044258.5	21.27	1887.20	8	1	3	4	1	6	2.12	1	NH3	No linker	C50H81O59N5S6
942.5850	2	5053913.5	23.12	1887.20	8	1	3	4	1	6	2.15	1	NH3	No linker	C50H81O59N5S6
942.5853	2	8262750.5	23.63	1887.20	8	1	3	4	1	6	2.46	1	NH3	No linker	C50H81O59N5S6
942.5854	2	1210266.5	22.53	1887.20	8	1	3	4	1	6	2.58	1	NH3	No linker	C50H81O59N5S6
982.5634	2	1218432.8	22.69	1967.10	8	1	3	4	1	7	2.02	1	NH3	No linker	C50H81O62N5S7
1132.1414	2	538765.9	26.19	2266.30	10	1	4	5	2	6	2.36	1	NH3	No linker	C64H102O70N6S6
965.0736	3	766241.25	23.08	2898.20	12	1	5	6	1	10	2.48	2	NH3	No linker	C74H122O91N8S10
991.0768	2	4035387	22.69	1984.20	8	1	3	4	1	7	2.16	2	NH3	No linker	C50H84O62N6S7

Table 8. Assigned GAG compositions for the KO HS sample using CZE-MS. MZmine 3 was utilized for electropherogram deconvolution and peak picking.

The 62 m/z values that were provided by GAGwizard are shown in Table 4. This included several neutral mass values with the same composition, which could indicate the presence of isomers because of multiple migration times for the same m/z value (Figure 44). The observed HS chains ranged from dp 2 to 10, and the number of sulfate group modifications ranged from 0 to 10. The neutral masses for the assigned GAG compositions ranged from 876 to 2898 Da with a mass error of 2.5 ppm. Similar to the WT sample, the majority of assigned GAG chains ranged from hexasaccharides to decassacharides, indicating that the sample preparation step was properly carried out. In addition, several sodium adducts, and 17 ammonia adducts were detected, despite the desalting step and the use of ammonium formate as a BGE.

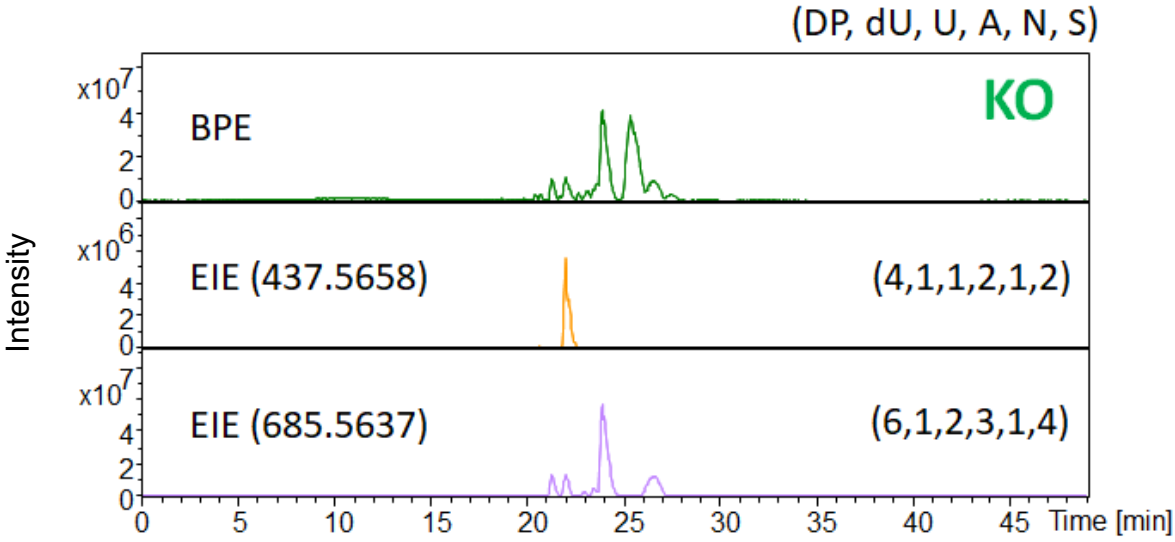

Figure 44. In the HS analysis in KO cells, multiple isomers were present in many assigned features. (Top) The base peak electropherogram of the sample. (Middle) The extracted ion electropherogram of m/z 437.57, which appears to show a single component. (Bottom) The extracted ion electropherogram of m/z 685.56, showing 4-6 different components. [DP, dU, U, A, N, S] means the code regarding the degree of polymerization, the number of unsaturated hexuronic acids, hexuronic acids, hexosamines, N-Acetylation, and sulfate group modification, respectively.

MS signals of anion analytes in negative mode were generally not as stable as those in positive mode due to the direction of electroosmosis, which was directed away from the opening of the emitter with an untreated capillary/emitter. [39] Sarver et al. modified the interior wall of an emitter to control electroosmotic flow forward to the emitter orifice under negative ion operation, demonstrating stable spray ionization. [39] Sanderson et al. coated inner wall of a BFS capillary with cationic molecules to produce electroosmotic flow in the direction of a MS analyzer, producing stable flow with high sensitivity. [28] We therefore suspected that the total ion electropherogram was stable and reproducible under the condition of an appropriate equilibrium and a uniform orifice of glass emitter for each cycle. Therefore, this study may present a new opportunity to improve a methodology for analyzing complex GAG-type mixtures that are well-suited to negative-ion electrospray in MS.

HS Assignment via Whisker Box Plots

Figure 45 exhibits a set of box plots for a quick visual representation of the original table (Table 7 and 8) using in-house developed MATLAB code. The box plots represent a central dot with a dark blue box range showing the distribution of results for the CZE-MS method, including the lower quartile, media, and upper quartile value. White empty circles are outliers defined as values which exceed the upper and lower quartile by at least 1.5 times between the two quartiles. Notably, WT and KO samples did not differ significantly in neutral mass or chain length, as these were a function of the depolymerization step, which was identical for both samples. The sulfation levels and N-acetylation levels do differ significantly, showing that the EZH2 KO has an altered biosynthetic pathway for GAG production. The EZH2 KO is expected to have a reduction in sulfatase production. The higher degree of sulfate

group modification for the KO versus the WT sample is consistent with an increase in 6-O sulfation on the amino sugar residues in the KO sample that lacked sulfatases. Other analytical methods revealed that 6-O sulfate modification increased, as did fibroblast growth factor binding (not shown here). Furthermore, the extent of N-acetyl group modification in the WT sample not only was greater, but also has a higher median value than that of in the KO sample. These suggest a reduction in N-deacetylase in the genetically altered cell line.

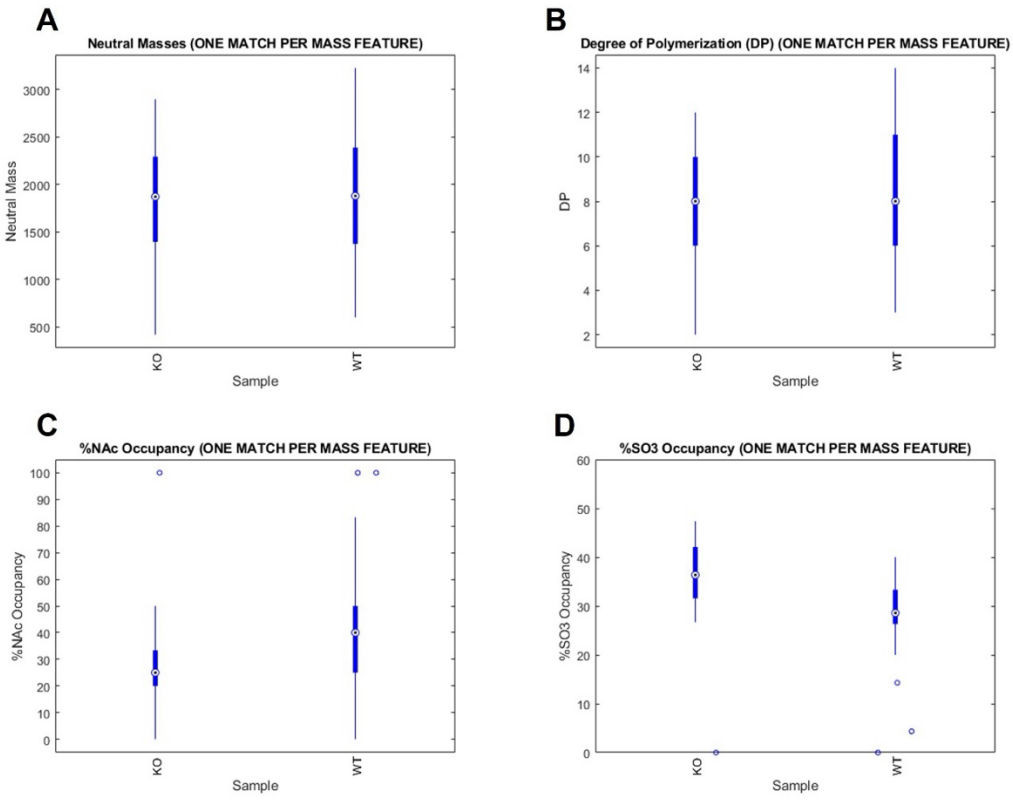

Figure 45. Box plots depicting comparisons of two samples, KO and WT, in terms of four characteristics of HS (A) Neutral masses (B) Degree of polymerization or chain lengths, (C) Degree of N-acetylation on 2-O carbons of amino sugar residues, (D) Degree of sulfate modification in its structure. To avoid duplicates, only one HS match was counted per mass feature.

HS Identifications from WT and KO Samples and Venn Diagram Analysis

We observed that the electropherogram and corresponding mass spectra included GAG component peaks, including adduct and solvent signals due to the nature of ESI. Based on the parameters that used during peak picking process, isotope peaks were removed except for the monoisotopic peak. To compare the GAG assignments for both samples, Venn diagram analysis for neural masses (matched features) with 0.1 Da as the m/z tolerance was used. Based on the combined results of the two samples, 46 HS features could be identified, with only 10 overlapping GAGs as seen in the figure 46. Further research can be conducted using tandem mass spectrometry with electron-based activation to gain a better understanding of the GAG profile, including those structures. The identification of precursor ions that are unique to either the WT or KO samples can provide novel insights into the HS biosynthesis pathway.

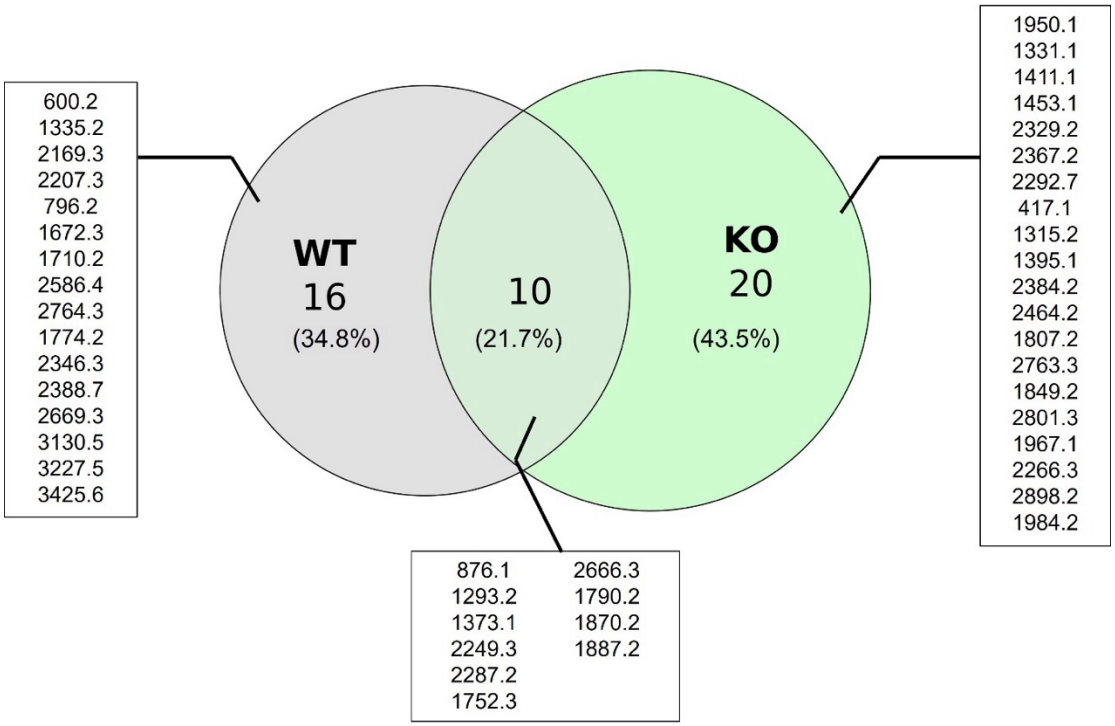

Figure 46. The numbers of HS neutral mass from CZE-MS experiments were represented by Venn diagram. The overlap of assigned to HS via GAGwizard between duplicate CZE-MS analyses of WT and KO samples shows 10 unique neutral masses.

We also found that HS oligosaccharides identified in the WT sample contained not only higher molecular weight HS components as compared to those in the KO sample but also sodium and ammonium adducts in response to relatively longer HS species. As a result of the unusual number of adducts on those HS molecules (> 3k Da), the results are likely to be false positives as shown in Table 7.

However, we do not consider retention times (migration times) here and expect that migration time correction which align peaks using internal markers during separation may improve the analytical utility of separations. [62, 63] By doing this, identifying isomers of the same mass/charge ratio multiple times will help assign sulfate modifications to them to answer question about the biosynthesis of HS.

4.3.2. Analysis of Chondroitin Sulfate Oligosaccharides in Cartilage with Mucopolysaccharidosis VII by Capillary Zone Electrophoresis-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

CZE-MS Analysis

All samples were first subjected to CZE-MS in order to separate and analyze the mixtures. As illustrated in Figure 47, base peak electropherograms were obtained within a five min separation window for the CS analysis. Figure 48A shows the number of unique compositions per sample, 21 at the maximum and 5 at the minimum CS compositions in a single CZE-MS/MS. This was achieved with only 115 nL injection volume (calculated based on the Poiseuille's law) for each sample. As shown in Figure 48B, observed neutral masses can confidently assign the degree of polymerization in each sample. A total of three samples (M3175, M3182, 3195) ranked as the widest distribution of chain lengths from disaccharides to decasaccharides (M3182 was dodecasaccharides), but there was no correlation between the control and the disease. Also, Figure 48C shows neutral mass ranges that overlap with the distribution pattern in Figure 48B, with the exception of a few data points, possibly faults positive. As shown in Figure 48D, the most percent SO_3 group modifications were determined from 0% to 45%, suggesting varying sulfate group modification on CS in the disease compared to the control. As compared to the control, M3182 in the disease showed a high level of agreement with each other composition, but still had several outliers. In other words, the bottom and top edges were very close for four control CS samples. This resulted in a narrow interquartile range (between 25th and 75th percentile) and a tendency for data points outside of the blue box to be outliers.

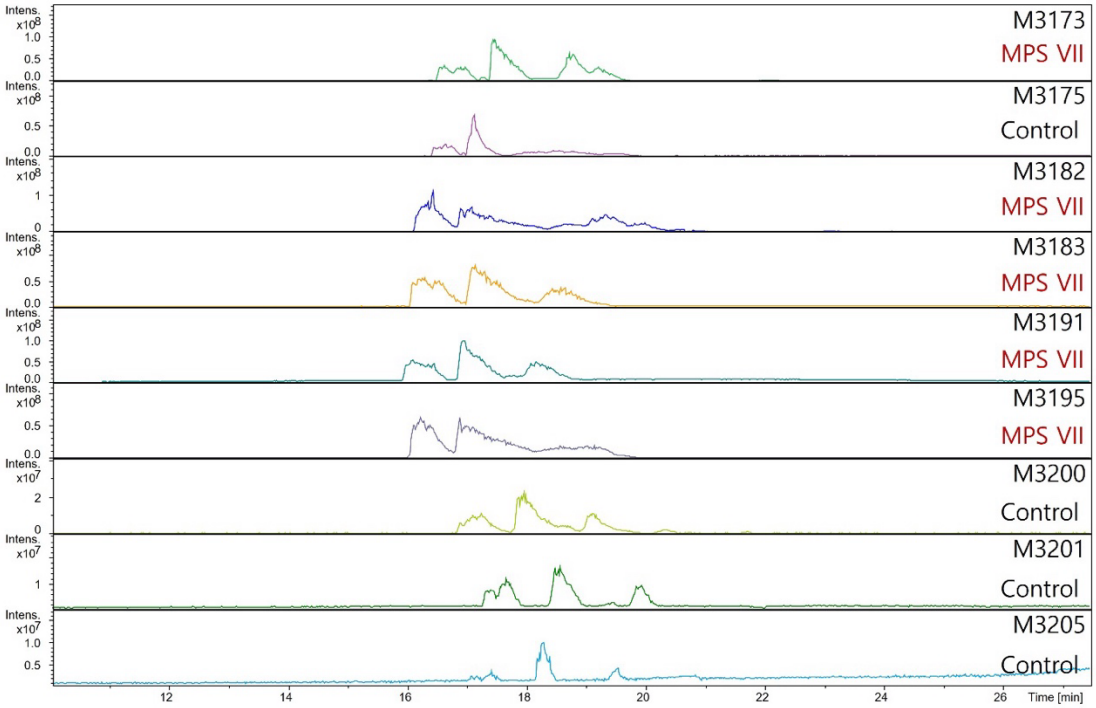

Figure 47. Base peak electropherograms for a total of nine CS samples. Using Poiseuille's law, each injection volume was calculated for a capillary of 60 cm length and 50 µm inner diameter coated with dichlorodimethylsilane (DMS). Each pressure was applied at 400 mbar for 15 s, resulting in 115 nL.

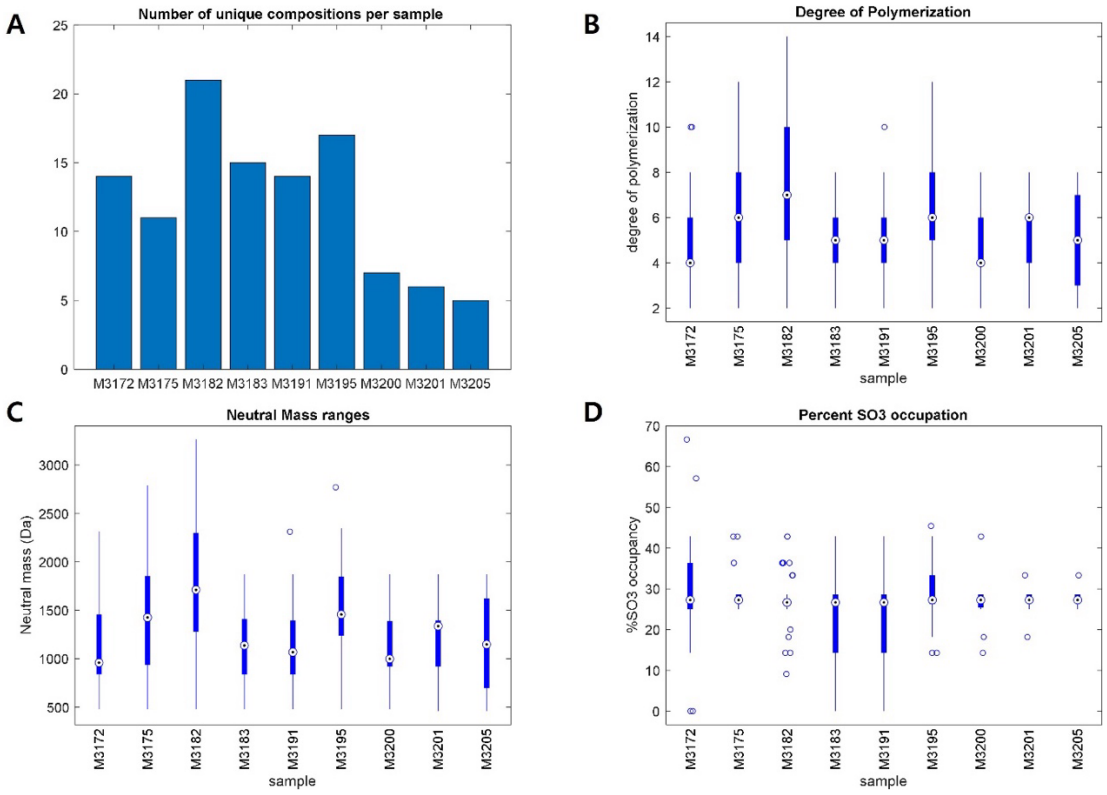

Figure 48. Box plots illustrating comparisons of the control and the disease. (A) The number of unique compositions per sample. Redundant were ejected to be count. (B) The degree of polymerization ranging from disaccharides to14mers. Empty dots refer to as outliers that located outside the whiskers of the box plot (e.g., in the case of 1.5 interquartile ranges below the first quartile or 1.5 interquartile ranges above the third quartile. (C) The neutral mass ranges of matches CS features for each sample. (D) The percent of sulfate modification on each CS feature. Considering that 2-*O* at iduronic acid residues and 2-*O*, 4-*O*, and/or 6-*O* at N-acetyl-galactosamine residues.

Additionally, a spiking test was performed in order to determine the structure and location of the sulfate group modification within CS compositions. As a result, CS dp4 with 4-O sulfate group modification was added to one of the original samples of CS, which provided structural information about the mid-peak in electropherograms. As illustrated in Figure 49, the abundance of migration peak in the

middle was larger in the spiked sample than that in the unspiked sample. Because of the limited number of standard sample stocks available, CS dp4(4S) was only confirmed and it is likely that the remaining two major migrated peaks contain 2-*O*, 4-*O*, or/and 6-*O*, except for the 4-*O*/4-*O* combination.

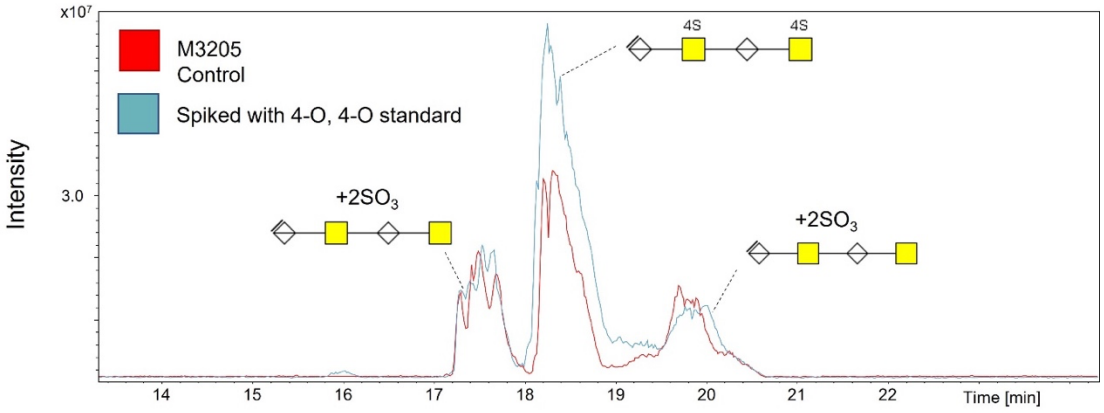

Figure 49. Based on the overlapped base peak electropherograms, the spiking standard test provides structural information regarding the middle peak as follows: ΔGlcA-β1,3-GalNAc(4S)-β1,4-GlcA-β1,3-GalNAc(4S)β1,4. As for the composition of the rest peaks, the backbone structures are the same, but its modification by the sulfate group is unclear.

Analysis of 2-AB Derivatized CS Standards

CS oligosaccharides across all samples were observed to range from two to fourteen chains in length, but the most abundant form was CS tetrasaccharides containing only one sulfate group.

However, the number of charges corresponds to the number of disaccharides by losing one proton per repeat unit for the fully ionized but not Na exchanged CS, showing the same mass-to-charge ratio independent of length. (e.g., doubly charged

CS tetrasaccharide versus triply charged CS hexasaccharide) If all repeating units are in GlcA-GalNAc-4S, the charge will scale in proportion to chain length, and the same mass-to-charge ratio will be observed at m/z 458.0610). To solve the m/z 458 overlap issue, 2-aminobenzamide (2-AB), a reducing terminal derivatizing reagent, was used to generate unique molecular weights and masses for each different length of CS. [64] As shown in Figure 50, a reactive carbonyl group of a disaccharide GAG and amine forms Schiff base that is subsequently reduced to secondary amine by reductant such as sodium cyanoborohydride. [64]

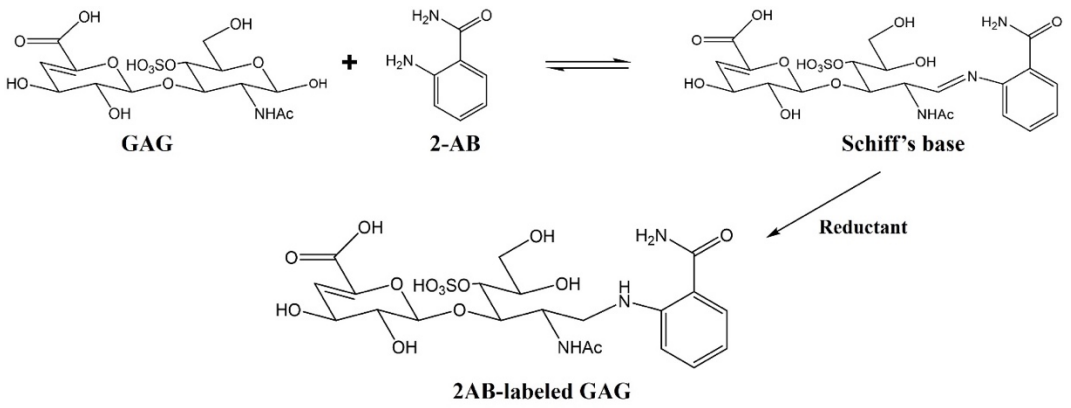

Figure 50. A labeling approach to attach 2-AB reagent to the reducing end of GAG.

Before analyzing our canine CS samples with 2-AB derivatization, CS/DS dp4 were functionalized with 2-AB and analyzed by CZE-MS/MS. As shown in Figure 51A, we observed two separately migrating peaks in the CS dp4 sample. In order to prove that the standard sample contains DS dp4 in a small proportion, we used a multivariate statistical approach (shown in later this chapter). The relative abundances of CS dp4 and DS dp4 were shown in Figure 52, leading to the relative area of 100 % and 20.16 %, respectively. Another benefit of the derivatization is that C_2 and Z_2 ions that have identical masses for unlabeled tetrasaccharides will have

different masses when labeled, allowing accurate fragment assignment (Figure 51B, 51C, and 52B).

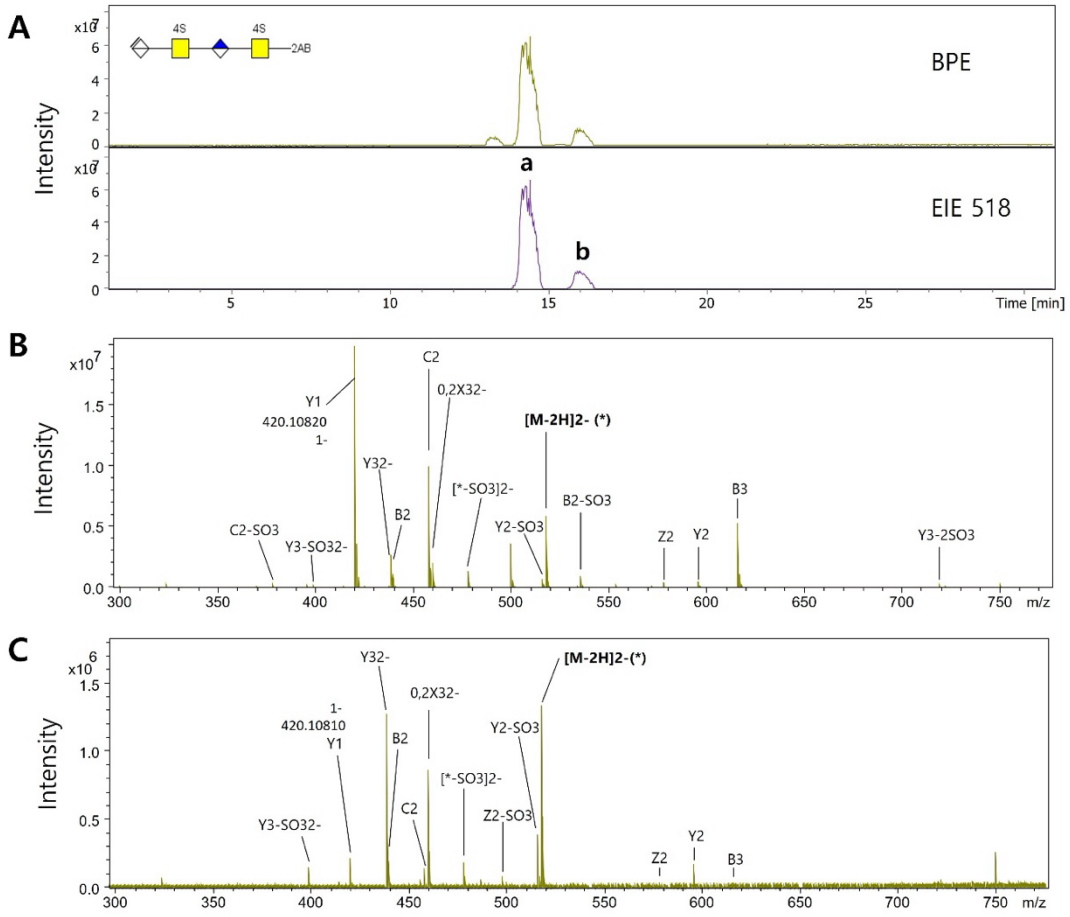

Figure 51. (A) Top: the base peak electropherogram for the CS dp4 standard as shown in the inset, including two sulfate group. Bottom: the extract ion electropherogram for m/z 518 (B) An annotated MS/MS mass spectrum for the first (a) peak in Figure 51A. Since it was derivatized with 2-AB, C and Z ions were identified. (C) A mass spectrum for the second (b) peak. Asterisks insist a precursor molecule.

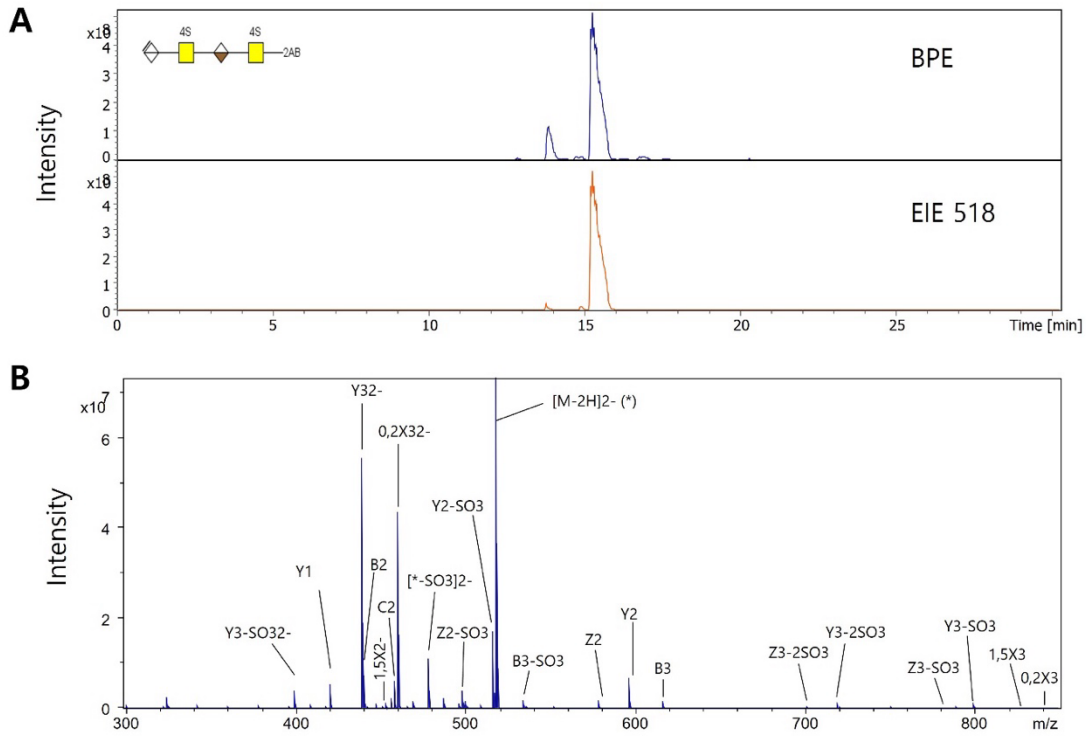

Figure 52. (A) Top: the base peak electropherogram for the DS dp4 standard as shown in the inset, including two sulfate group. Bottom: the extract ion electropherogram for m/z 518 (B) An annotated MS/MS mass spectrum for the peak in the extracted ion electropherogram in Figure 52A. Since it was derivatized with 2-AB, C and Z ions were identified. Asterisks insist a precursor molecule.

PCA was performed in order to further analyze the side product of the CS dp4 standard based on multivariate statistical analysis. Figure 53 illustrates that each major component was represented by either CS or DS, and the small side component within the CS dp4 sample was referred to as unknown to be classified. All sample were injected to a CZE-MS platform as triplicates. As a result, a clear separation was observed between the CS and DS samples in Figure 53. In addition, the first PC corresponds to 98.9% of the differences between the spectra, which means the variation in the CS isomers comparison can be explained by only one principal

component (PC1). The second PC addresses only 0.7% of the remaining variance. Additionally, based on PC1, the MS/MS spectra of a small side product from the unknown sample were clearly identified as DS molecules as shown in Figure 53. Dendrogram analysis also insisted that MS/MS analysis enabled distinguishing CS from DS. From the heatmap and PCA loading plot, diagnostic fragment ions were able to be distinguish two isomers as well as the component of a side product present in the CS sample (Figure 54).

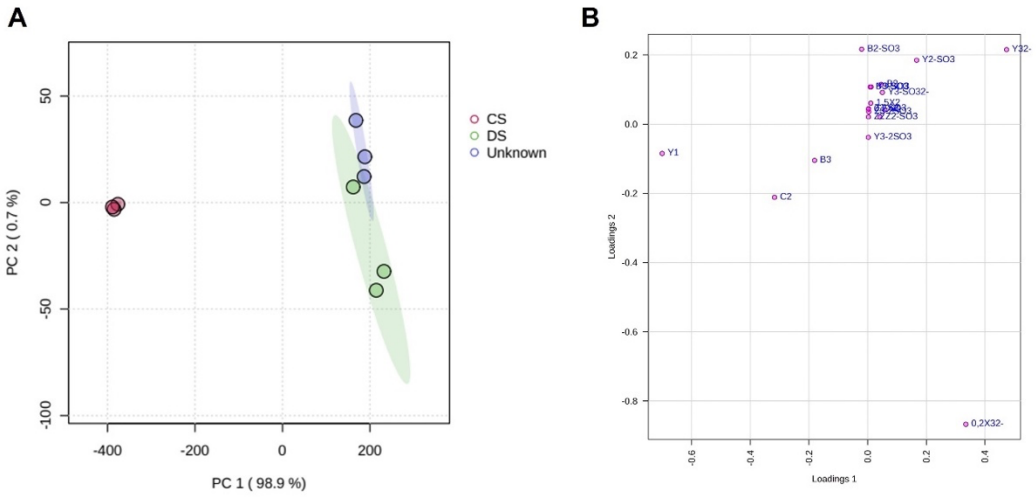

Figure 53. (A) PCA for CS and DS samples by PC1 and PC2. CS was clearly distinguished from DS to an explanation extent of 98.9% by PC1. (B) An illustration of the loading plot indicates the features of fragment ions that determine a class of isomers.

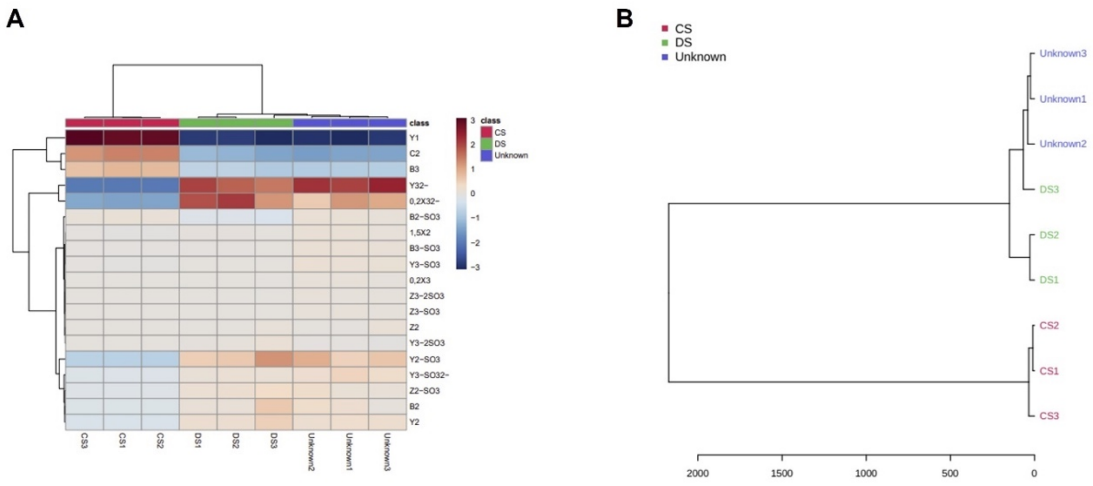

Figure 54. (A) Heatmap visualizes the correlation of fragment ion matrix as well as classification for CS isomers. It was combined with dendrogram (top) to show clustering of samples by the type of isomers, including the unknown side product peak (B) Dendrograms summarize the information concerning each isomer value and sort them based on the degree of similarity between them.

Analysis of 2-AB Derivatized Biological CS Samples

2-AB was used was used in order to generate unique molecular weights and masses for each different lengths of biologically-derived CS samples. [64] The CS samples were dissolved in water and analyzed by CZE-MS to obtain electropherograms and corresponding mass spectra.

A mucopolysaccharidosis dog CS sample was analyzed using CZE-MS/MS in figure 55. The migration window was ranging from 14.6 min to 16.8 min and CID was utilized to fragment the precursor ion, 518.0954, which is a 2-AB labeled doubly charged CS with two sulfation. Figure 56 shows annotated fragment ions to characterize the derivatized peaks, resulting in several glycosidic bond cleavages as well as a few cross-ring cleavages. Although the MS/MS result was not able to

sufficiently provide the detailed structural information on each peak, there is still room for utilizing CID fragmentation for multivariate statistical analysis to distinguish the disease from the control. The CZE-MS/MS system is currently being applied to other samples with triplicates that remain in our group for this study. In this regard, CZE-MS/MS is capable of characterizing GAGs for biological samples for diagnostic purposes.

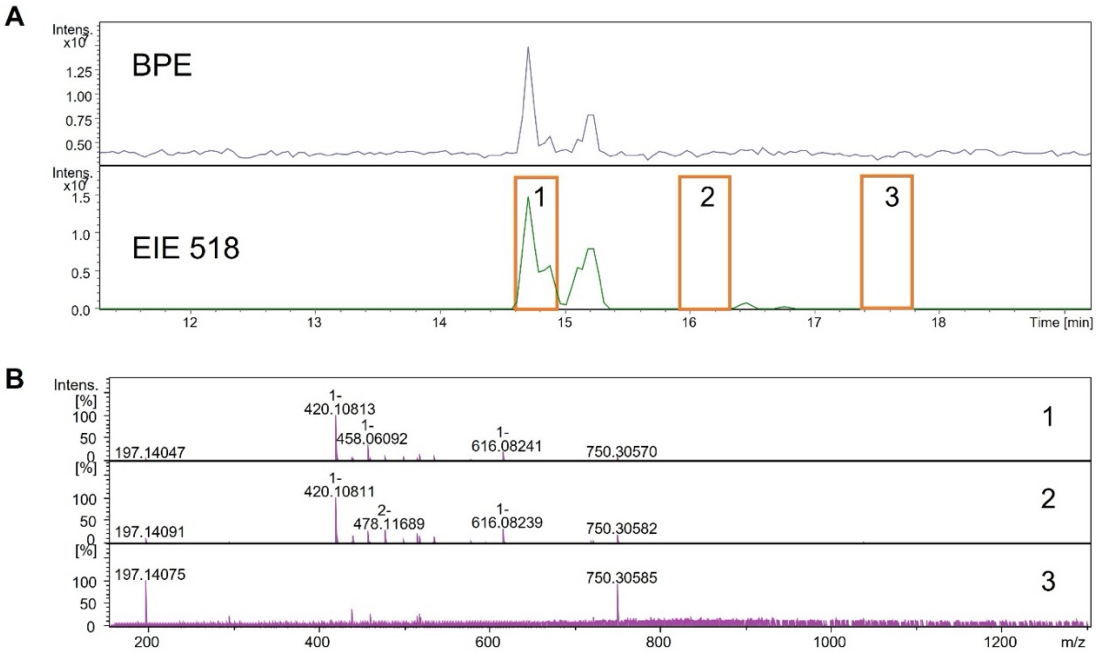

Figure 55. (A) Top: the base peak electropherogram for CS sample (M3183) in the disease Bottom; the extract ion electropherogram for m/z 518, which is 2-AB labelled. (B) Annotated MS/MS mass spectra for the corresponding peak in the EIE in Figure 55A.

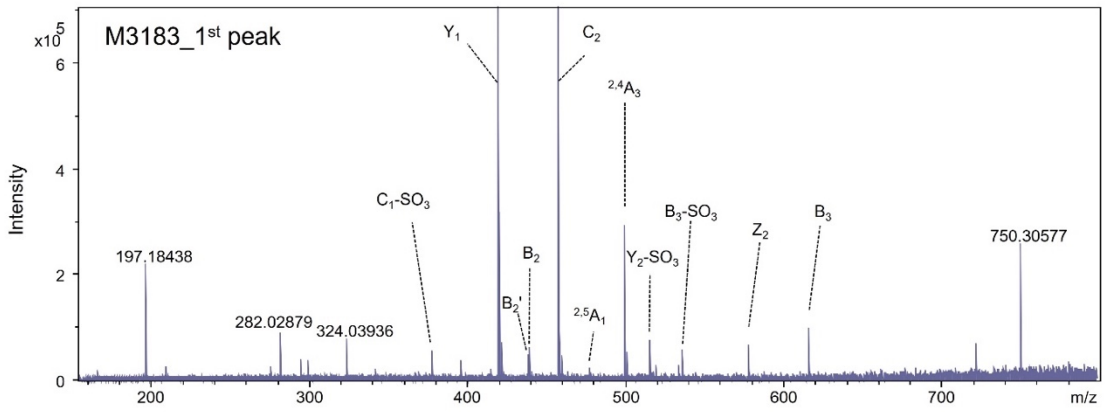

Figure 56. A mass spectrum for the first peak in the EIE in Figure 55A. The sample was derivatized with 2-AB so that Y ions can be recognized from Z ions.

4.3.3. Analysis of Chemo-enzymatically Synthesized Chondroitin Sulfates by Capillary Zone Electrophoresis-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

CZE-MS/MS analysis

At the MS1 level, the intact mass of the GAG gives information about its DP, number of SO_3, number of NAc, and any terminal modifications like non-reducing-end Δ-unsaturation or reducing-end derivatization. Based on the results of the initial analysis for three CS samples using CZE-MS, it was possible to ascertain whether the synthesized CS products had been purified and to identify appropriate precursors. As shown in Figure 57, each oligosaccharide CS separated through a DMS functioned capillary and major components of three CS migrated at around 30 min. As shown in Figure 58, the three synthesized samples contained a few side-products in addition to each of the major CS molecules. However, the relative intensities of these side-products varied depending on the synthesized sample. With regard to the

characterization of CS using CZE-MS, observed mass to charge ratios were analyzed by an in-house MATLAB script, called GAGWizard.

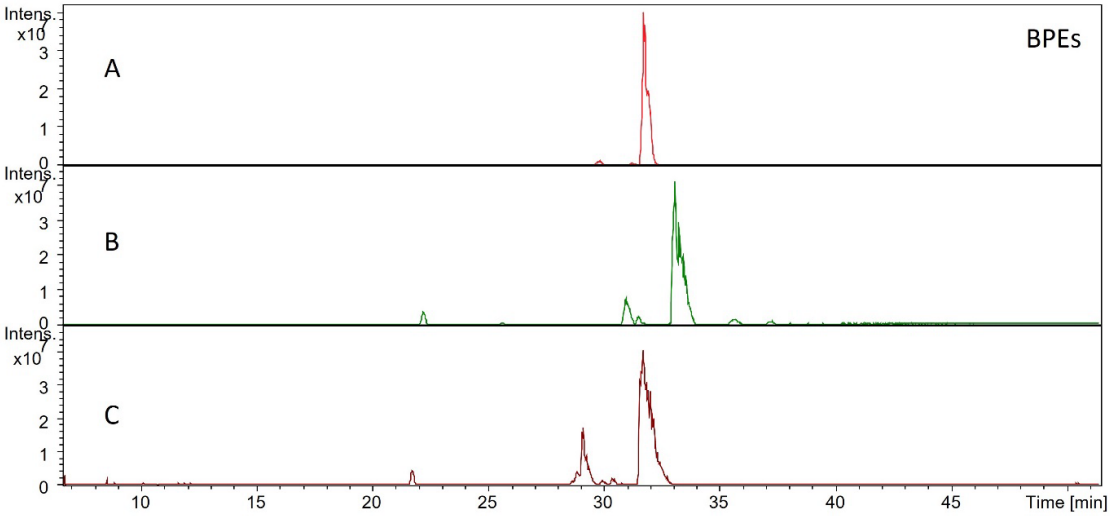

Figure 57. Base peak electropherograms of three chemoenzymatically synthesized chondroitin sulfates. (A) CS tetrasaccharide with Glu-Gal-Gal-Xyl-S linker; (B) CS hexassaccharide with Glu-Gal-Gal-Xyl-Serine linker; (C) CS octassaccharide with Glu-Gal-Gal-Xyl-Serine linker.

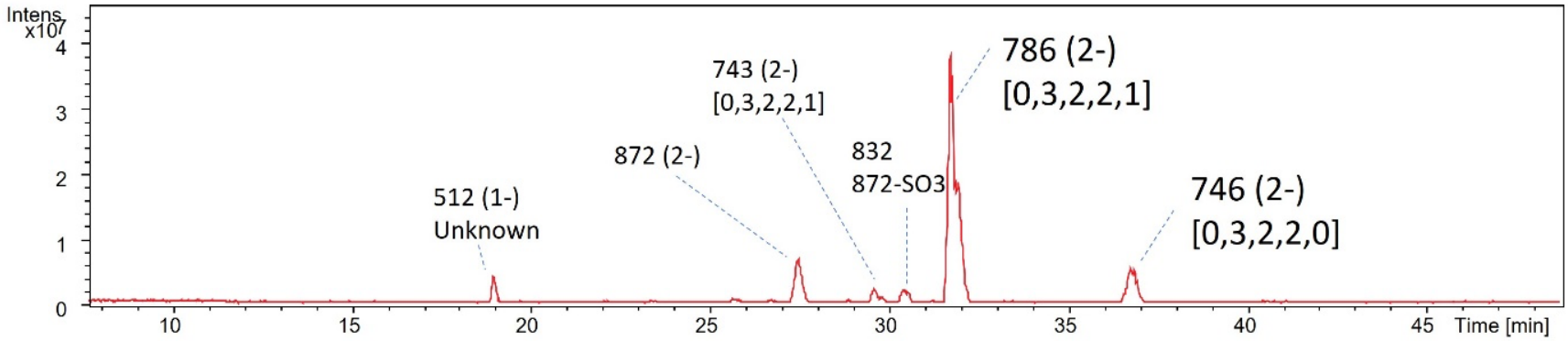

Figure 58. Base peak electropherogram for the tetrasaccharide CS sample. [DP, dU, U, A, N, S] means the code regarding the degree of polymerization, the number of unsaturated hexuronic acids, hexuronic acids, hexosamines, N-Acetylation, and sulfate group modification, respectively.

m/z	z	Intensity	RT	Neutral Mass	DP	Delta Uronic	Uronic	Amino	N-Acetyl	SO3	Theor Mass	Mass Error (ppm)	Linker	Formula
743.181	2	2437150	29.6305	1488.378	5	0	3	2	2	1	1488.37	4.876071	Gal-Gal-Xyl	C51H80O46N2S1
746.7171	2	5649345	36.7558	1495.45	5	0	3	2	2	0	1495.446	2.840893	Gal-Gal-Xyl-S	C54H85O45N3S0
786.6901	2	38200000	31.7428	1575.396	5	0	3	2	2	1	1575.402	-4.16732	Gal-Gal-Xyl-S	C54H85O48N3S1

Table 9. Assigned GAG compositions for the tetrasaccharide CS sample using CZE-MS. MZmine 3 was utilized for electropherogram deconvolution and peak picking. m/z 786.6901 was designated to the most abundant peak at 31.74 min.

A house-built program also showed that small chains migrate before and after the major oligosaccharides, resulting in less sulfated CS or CS with partial linkers (Gal-Gal-Xyl). The precursor ion of m/z 786.6901(2-) was isolated to conduct CZE-MS/MS (CID) as shown in Figure 59. The annotated MS/MS spectra was showed that the precursor ion matched with our desired structure with one sulfation on the GalNAc residue that lies closes to the linker.

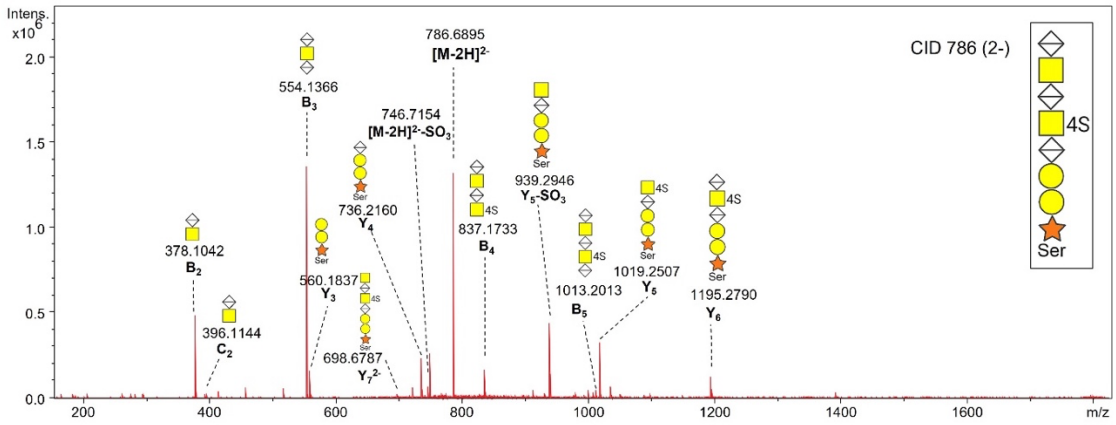

Figure 59. Tandem mass spectrometry (MS/MS) facilitated characterization of the major component of CS in the tetrasaccharide sample. The mass spectrum was obtained by collision induced dissociation. GAGs, including fragment ions, were represented by the Symbol Nomenclature for Glycans (SNFG): yellow ■, N-acetylgalactosamine; white ◇, either glucuronic acids or iduronic acids (unsure C-5 epimerization, assuming glucuronic acids based on synthesis). Each mass-to-charge ratio and fragment type were presented with dotted lines to insist corresponding peaks.

Hexasaccharide CS Analysis using CZE-MS and CZE-MS/MS

The base peak electropherogram of the hexasaccharide shows that the m/z 1016.22 (2-) ion was the major component from the injected CS sample, and assigned as a hexassacharide CS. As illustrated in Figure 60, the sample also contained a few side-products besides the major CS molecule, such as m/z 1101 (2-) and 976 (2-). Other minor components were discovered as sulfate group loss species or a shorter linker without serine attached as shown in Table 6 (e.g., m/z 976.2511 for sulfate group loss and m/z 972.7119 for a small linker region (Gal-Gal-Xyl). Figure 61 shows how CZE-MS/MS (CID) was used to characterize the CS structural information. Due to the redundant MS/MS data from the doubly and triply charged precursor ions, the CZE-MS/MS assignments are more reliable.

Octasaccharide CS Analysis using CZE-MS and CZE-MS/MS

A similar method was used to analyze biosynthetically prepared CS octasaccharide sample. CZE-MS and CZE-MS/MS analysis were carried out and the base peak electropherogram is shown in Figure 62. In fact, the side-products were greater in abundance than the other two CS samples, and one of them contained an additional sulfate group modified CS compound (m/z 856, (3-)) than the major component. Interestingly, there were multiple charge states for some molecules, e.g., 830 (3-) and 1245 (2-). Future analysis using electron-based activation applications could further characterize the structures of synthesized compounds.

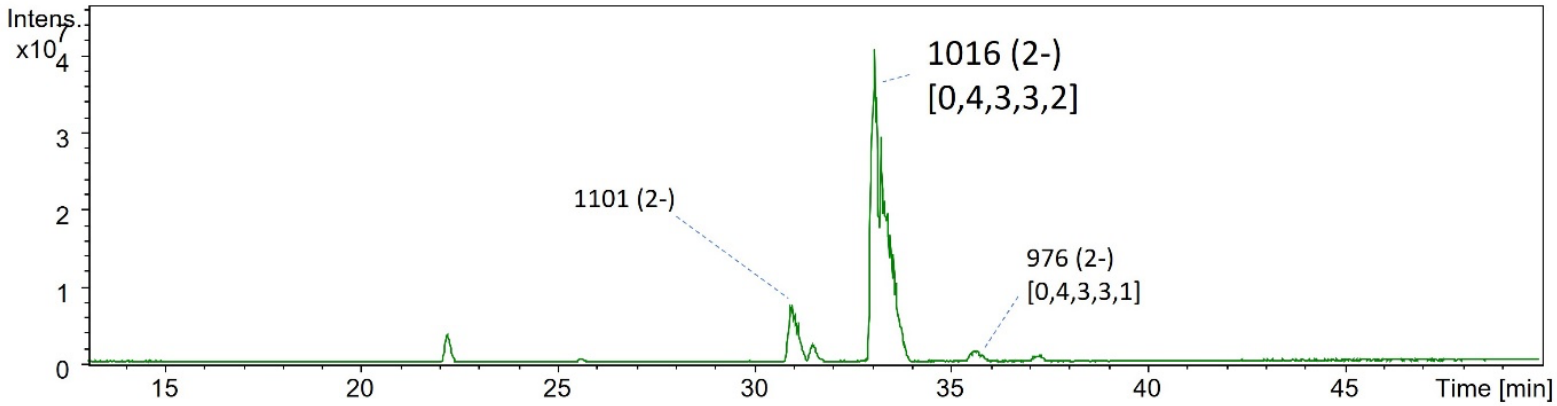

Figure 60. Base peak electropherogram for the hexamer CS sample. [DP, dU, U, A, N, S] means the code regarding the degree of polymerization, the number of unsaturated hexuronic acids, hexuronic acids, hexosamines, N-Acetylation, and sulfate group modification, respectively.

m/z	z	Intensity	RT	Neutral Mass	DP	Delta Uronic	Uronic	Amino	N-Acetyl	SO3	Theor Mass	Mass Error (ppm)	Adduct Type	Linker	Formula
677.1492	3	3601451	33.0571	2034.471	7	0	4	3	3	2	2034.471	0.197447	none	Gal-Gal-Xyl-S	C68H106O62N4S2
972.7119	2	2617263	31.5292	1947.439	7	0	4	3	3	2	1947.439	0.380911	none	Gal-Gal-Xyl	C65H101O60N3S2
976.2511	2	1901747	35.665	1954.518	7	0	4	3	3	1	1954.514	2.070899	none	Gal-Gal-Xyl-S	C68H106O59N4S1
1016.22	2	40800000	33.0795	2034.457	7	0	4	3	3	2	2034.471	-6.92042	none	Gal-Gal-Xyl-S	C68H106O62N4S2

Table 10. Assigned GAG compositions for the hexamer CS sample using CZE-MS. MZmine 3 was utilized for electropherogram deconvolution and peak picking. m/z 1016.22 was designated to the most abundant peak at 33.0795 min.

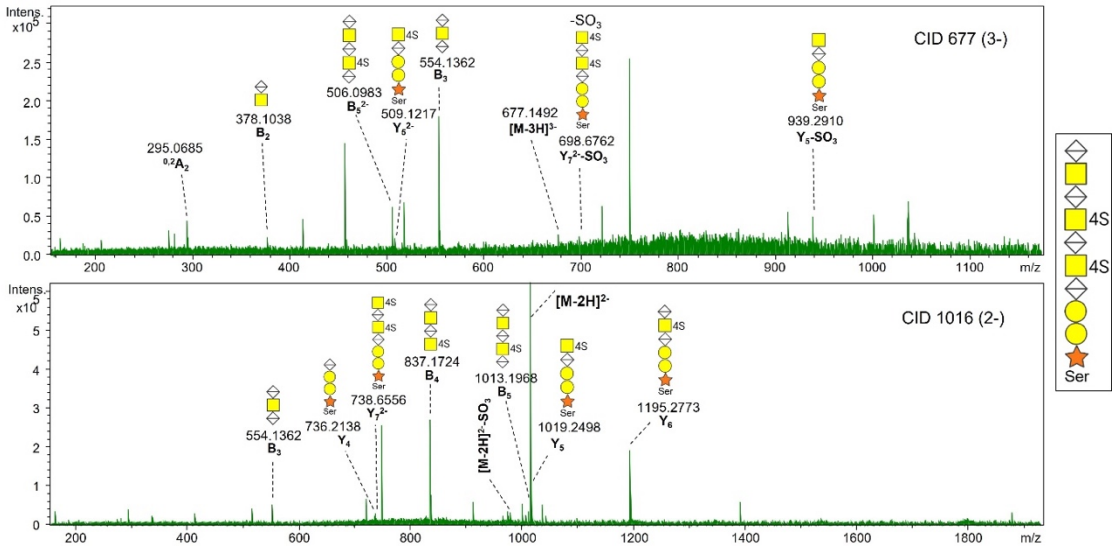

Figure 61. Tandem mass spectrometry (MS/MS) facilitated characterization of the major component of CS in the hexasaccharide sample. There were two different charge states of ions isolated, proving its structure in a highly convincing way. Mass spectra were obtained by collision induced dissociation. GAGs, including fragment ions that were represented by the Symbol Nomenclature for Glycans (SNFG): yellow ■, N-acetylgalactosamine; white ◇, either glucuronic acids or iduronic acids (unsure C-5 epimerization, assuming glucuronic acids based on synthesis). Each mass-to-charge ratio and fragment type were presented with dotted lines to insist corresponding peaks.

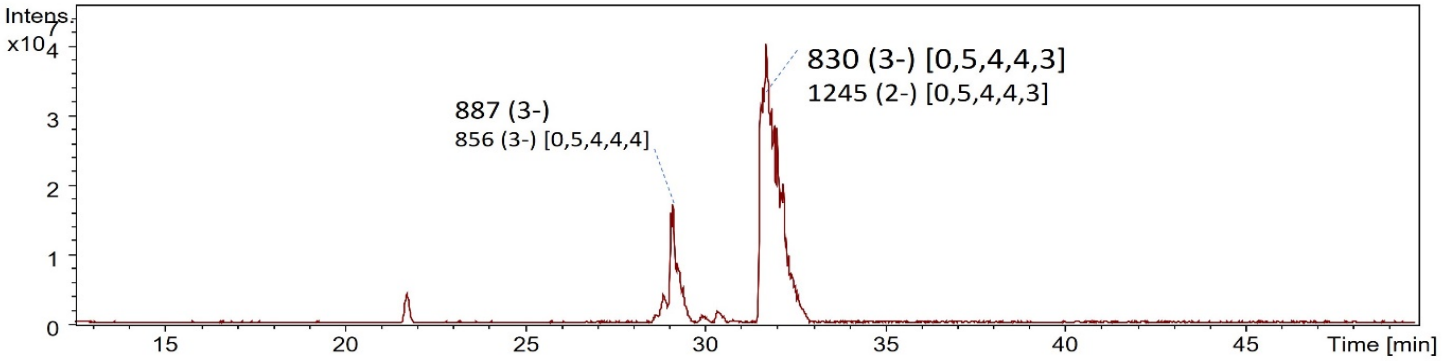

Figure 62. Base peak electropherogram for the octasaccharide CS sample. [DP, dU, U, A, N, S] means the code regarding the degree of polymerization, the number of unsaturated hexuronic acids, hexuronic acids, hexosamines, N-Acetylation, and sulfate group modification, respectively.

m/z	z	Intensity	RT	Neutral Mass	DP	Delta Uronic	Uronic	Amino	N-Acetyl	SO3	Theor Mass	Mass Error (ppm)	Adduct Type	Linker	Formula
801.1605	3	2052953	30.3813	2406.505	9	0	5	4	4	3	2406.507	-0.80752	none	Gal-Gal-Xyl	C79H122O74N4S3
830.1651	3	40300000	31.7065	2493.519	9	0	5	4	4	3	2493.539	-8.13258	none	Gal-Gal-Xyl-S	C82H127O76N5S3
856.8205	3	4097538	28.8744	2573.485	9	0	5	4	4	4	2573.496	-4.19904	none	Gal-Gal-Xyl-S	C82H127O79N5S4
1245.758	2	5161571	31.7291	2493.532	9	0	5	4	4	3	2493.539	-2.69833	none	Gal-Gal-Xyl-S	C82H127O76N5S3
837.4966	3	2658451	31.7065	2515.513	9	0	5	4	4	3	2515.521	2.989679	1Na	Gal-Gal-Xyl-S	C82H126O76N5S3Na1

Table 11. Assigned GAG compositions for the octasaccharide CS sample using CZE-MS. MZmine 3 was utilized for electropherogram deconvolution and peak picking. m/z 830.1651 (3-) was designated to the most abundant peak at 31.71 min.

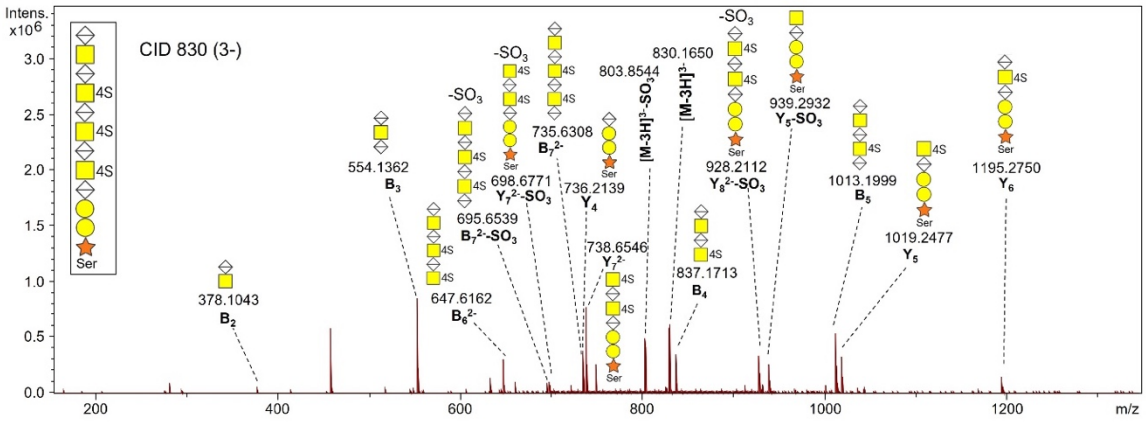

Figure 63. Tandem mass spectrometry (MS/MS) facilitated characterization of the major component of CS in the octasaccharide sample. The mass spectrum were obtained by collision induced dissociation. GAGs. Fragment ions were represented by the Symbol Nomenclature for Glycans (SNFG): yellow ■, N-acetylgalactosamine; white ◇, either glucuronic acids or iduronic acids (unsure C-5 epimerization, assuming glucuronic acids based on synthesis). Each mass-to-charge ratio and fragment type were presented with dotted lines to insist corresponding peaks.

Other unlabeled side products were also analyzed using CZE-MS/MS (CID), shown in Figures 64 and 65, The script did not find the expected structures for them (m/z 872 of CS tetramer, 1101 of CS hexamer, and 887 of CS octamer). Further fragment ions from MS/MS experiments concluded that partial CS oligosaccharides without sulfate groups were produced in small amounts as side-products. As compared with other separation-MS combinations, CZE-MS exhibits low sample consumption coupled with high sensitivity, making it useful for examining minor products.

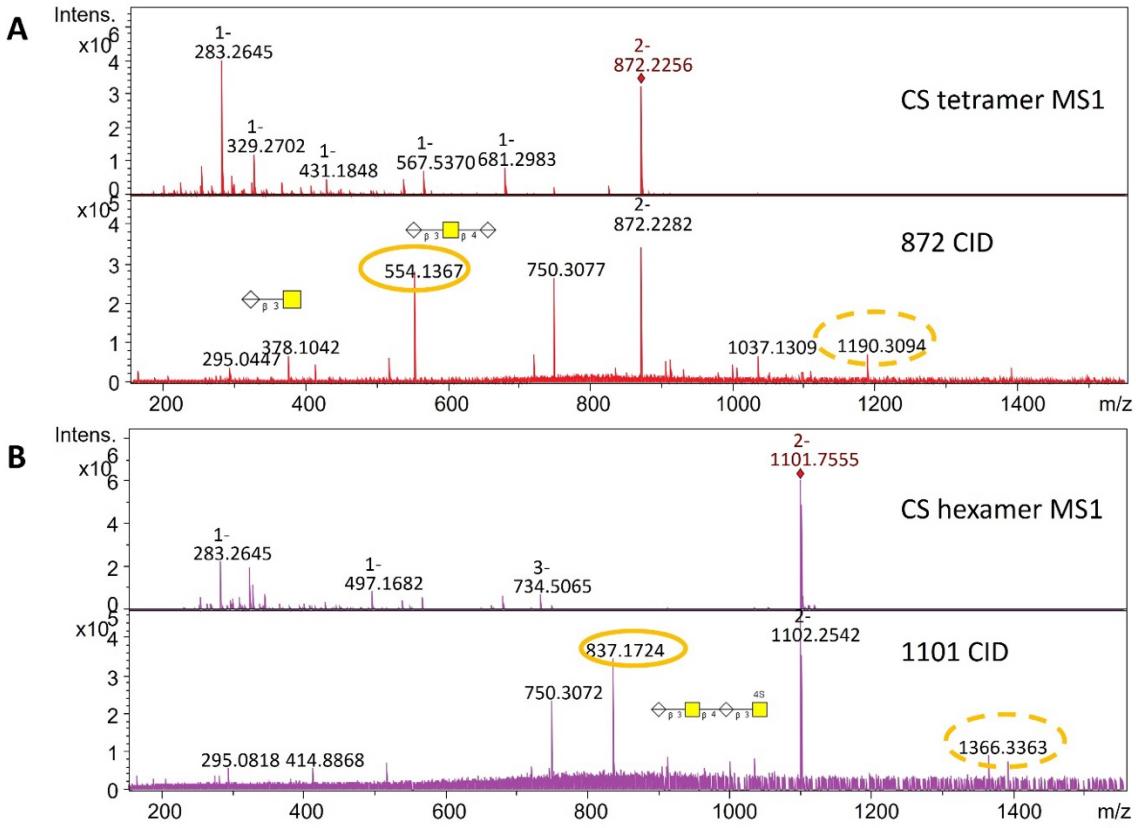

Figure 64. (A) top; MS1 using CZE-MS for the CS tetrasaccharide sample. 872.2256 (2-) was isolated and fragmented by CID. Bottom; The MS/MS mass spectrum for the precursor ion. A complementary ion of m/z 554.1367 in a line circle to m/z 1190.3094 in a dotted circle proved its molecular weight. (B) top; MS1 using CZE-MS for the CS hexasaccharide sample. 837.1724 (2-) was isolated and fragmented by CID. Bottom; MS/MS mass spectrum for the precursor ion. A complementary ion of m/z 837.1724 in a line circle to m/z 1366.3363 in a dotted circle proved its molecular weight.

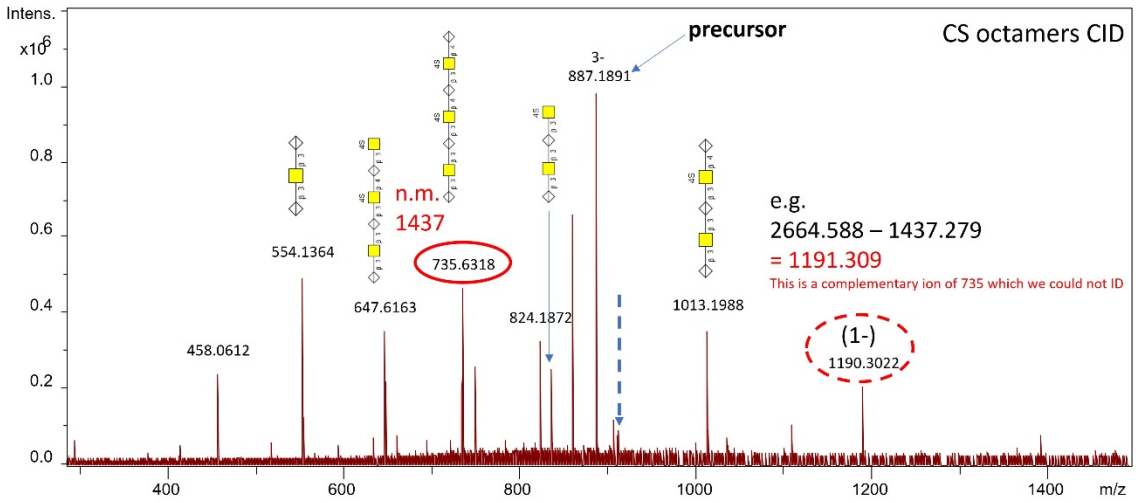

Figure 65. The CS octasaccharide MS/MS mass spectrum for the precursor ion of m/z 887.1891. A complementary ion of m/z 735.6318 in a line circle to m/z 1190.3022 in a dotted circle proved its molecular weight. As a result of other fragment ion peaks, the unannotated precursor ion might have partial CS oligosaccharides but different end compositions.

5.4 Conclusions

In this chapter, we examined how GAG samples can be analyzed using CZE-MS and provided three applications as examples to evaluate the performance and understand a future potential of CZE-MS. HS and CS GAGs derived from biological medium or chemo-enzymatical synthesis were characterized by CZE-MS and processed by in-house software as previously validated in Chapter 3.

With regards to HS analysis, CZE-MS has successfully identified HS oligosaccharide mixtures derived from A375 human melanoma cells that have been edited using the CRISPR/Cas9 genome editing technology to knockout genes involved in

regulating HS assembly. The CZE-MS method that was previously demonstrated in Chapter 3 identified HS oligosaccharides in a GAG mixture, suggesting that it can be used to characterize complex engineered cell-isolated HS mixtures,

especially to resolve biosynthetic pathways of GAGs. Based on a comparison of WT and KO samples using CZE-MS, a significant difference was found between cell types in terms of sulfation. Thus, our study provides insights into the glycobiology of heterogenous GAGs through a series of investigations with sophisticatedly engineered cells producing HS oligosaccharides, that can be used to address questions concerning their biological pathways and potentially leading to improved therapeutics for GAG-based diseases.

In CS analysis, CZE-MS and CZE-MS/MS were used for enzymatically treated as well as chemo-enzymatically synthesized samples, with automated data processing using in-house software for GAG annotation. CZE-MS and CZE-MS/MS were able to determine the purity and complexity of CS samples despite the small sample qunatity. In proof-of-concept experiments with canine cartridge CS, derivatization was used to overcome a significant issue in distinguishing C ions from Z ions as well as an issue with overlapping signals from oligomers of different lengths for CS extracted from canine cartilage samples. Future effort will utilize multivariate statistical analysis on CZE-MS/MS mass spectra to distinguish the control from the disease as described in the Chapter 3. CZE-MS/MS was also utilized to analyze chemoenzymatically synthesized CS in order to confirm the reaction product, and to assign the structures of side products. In response to a increasing demand for synthetic GAGs, chemoenzymatic synthesis will be an important route for GAG production, providing flexibility and specificity through

chemical derivatization and enzyme-catalyzed reactions. CZE-MS/MS can aid this development by providing important data on yield and side products, leading to improvements in synthesis with better defined structures from a more controlled reaction. As demonstrated by these three applications, MS-based analytical methods coupled with CZE techniques have become more useful and offer powerful tools for characterization of GAGs.

REFERENCES

(1) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H. Essentials of Glycobiology **2015**.

(2) Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J. Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Molecular & cellular proteomics : MCP* **2021**, *20*, 100025.

(3) Matyjaszczyk-Gwarda, K.; Kij, A.; Olkowicz, M.; Fels, B.; Kusche-Vihrog, K.; Walczak, M.; Chlopicki, S. Simultaneous quantification of selected glycosaminoglycans by butanolysis-based derivatization and LC-SRM/MS analysis for assessing glycocalyx disruption in vitro and in vivo. *Talanta* **2022**, *238*, 123008.

(4) Mikami, T.; Kitagawa, H. Biosynthesis and function of chondroitin sulfate. *Biochimica et Biophysica Acta* **2013**, *1830* (10), 4719-4733.

(5) Lin, T.-S.; Hsieh, C.-H.; Kuo, C.; Juang, Y.-P.; Hsieh, Y. S. Y.; Chiang, H.; Hung, S.-C.; Jiang, C.-C.; Liang, P.-H. Sulfation pattern of chondroitin sulfate in human osteoarthritis cartilages reveals a lower level of chondroitin-4-sulfate. *Carbohydrate Polymers* **2020**, *229*, 115496.

(6) Bayliss, M. T.; Osborne, D.; Woodhouse, S.; Davidson, C. Sulfation of chondroitin sulfate in human articular cartilage: the effect of age, topographical position, and zone of

cartilage on tissue composition. *Journal of Biological Chemistry* **1999**, *274* (22), 15892-15900.

(7) Lawrence, R.; Prill, H.; Vachali, P. P.; Adintori, E. G.; de Hart, G.; Wang, R. Y.; Burton, B. K.; Pasquali, M.; Crawford, B. E. Characterization of disease-specific chondroitin sulfate nonreducing end accumulation in mucopolysaccharidosis IVA. *Glycobiology* **2020**, *30* (7), 433-445.

(8) Kusche-Gullberg, M.; Kjellén, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Current Opinion in Structural Biology* **2003**, *13* (5), 605-611.

(9) van Kuppevelt, T. H.; Oosterhof, A.; Versteeg, E. M. M.; Podhumljak, E.; van de Westerlo, E. M. A.; Daamen, W. F. Sequencing of glycosaminoglycans with potential to interrogate sequence-specific interactions. *Scientific Reports* **2017**, *7* (1), 14785.

(10) Billings, P. C.; Pacifici, M. Interactions of signaling proteins, growth factors and other proteins with heparan sulfate: mechanisms and mysteries. *Connective Tissue Research* **2015**, *56* (4), 272-280.

(11) Basu, A.; Patel, N. G.; Nicholson, E. D.; Weiss, R. J. Spatiotemporal diversity and regulation of glycosaminoglycans in cell homeostasis and human disease. *American Journal of Physiology-Cell Physiology* **2022**, *322* (5), C849-C864.

(12) Ai, X.; Kitazawa, T.; Do, A.-T.; Kusche-Gullberg, M.; Labosky, P. A.; Emerson, C. P., Jr. SULF1 and SULF2 regulate heparan sulfate-mediated GDNF signaling for esophageal innervation. *Development* **2007**, *134* (18), 3327-3338.

(13) Chin, S. J.; Saville, J. T.; McDermott, B. K.; Zankl, A.; Fletcher, J. M.; Fuller, M. Chondroitin sulfate disaccharide is a specific and sensitive biomarker for mucopolysaccharidosis type IVA. *JIMD Reports* **2020**, *55* (1), 68-74.

(14) Saville, J. T.; McDermott, B. K.; Fletcher, J. M.; Fuller, M. Disease and subtype specific signatures enable precise diagnosis of the mucopolysaccharidoses. *Genetics in Medicine* **2019**, *21* (3), 753-757.

(15) Zhang, L. Glycosaminoglycan (GAG) Biosynthesis and GAG-Binding Proteins. In *Progress in Molecular Biology and Translational Science*, Zhang, L. Ed.; Vol. 93; Academic Press, 2010; pp 1-17.

(16) Petitjean, O.; Girardi, E.; Ngondo, R. P.; Lupashin, V.; Pfeffer, S. Genome-wide CRISPR-Cas9 screen reveals the importance of the heparan sulfate pathway and the conserved oligomeric Golgi complex for synthetic double-stranded RNA uptake and Sindbis virus infection. *Msphere* **2020**, *5* (6), e00914-00920.

(17) Weiss, R. J.; Spahn, P. N.; Chiang, A. W. T.; Liu, Q.; Li, J.; Hamill, K. M.; Rother, S.; Clausen, T. M.; Hoeksema, M. A.; Timm, B. M. Genome-wide screens uncover KDM2B as a modifier of protein binding to heparan sulfate. *Nature chemical biology* **2021**, *17* (6), 684-692.

(18) Zaia, J. Principles of mass spectrometry of glycosaminoglycans. *Journal of Biomacromolecular Mass Spectrometry* **2005**, *1* (1), 3-36.

(19) Wolff, J. J.; Leach III, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. Negative electron transfer dissociation of glycosaminoglycans.

Analytical Chemistry **2010**, *82* (9), 3460-3466.

(20) Leach III, F. E.; Xiao, Z.; Laremore, T. N.; Linhardt, R. J.; Amster, I. J. Electron detachment dissociation and infrared multiphoton dissociation of heparin tetrasaccharides. *International Journal of Mass Spectrometry* **2011**, *308* (2-3), 253-259.

(21) Zaia, J. Mass spectrometry of oligosaccharides. *Mass Spectrometry Reveiws* **2004**, *23* (3), 161-227.

(22) Jaravel, L.; Schindler, B.; Randon, J.; Compagnon, I.; Demesmay, C.; Dugas, V. Off-line coupling of capillary isotachophoresis separation to IRMPD spectroscopy for glycosaminoglycans analysis: Application to the chondroitin sulfate disaccharides model solutes. *Journal of Chromatography A* **2020**, *1617*, 460782.

(23) Wang, Z.; Zhang, T.; Xie, S.; Liu, X.; Li, H.; Linhardt, R. J.; Chi, L. Sequencing the oligosaccharide pool in the low molecular weight heparin dalteparin with offline HPLC and ESI-MS/MS. *Carbohydrate Polymymers* **2018**, *183*, 81-90.

(24) Hitchcock, A. M.; Yates, K. E.; Costello, C. E.; Zaia, J. Comparative glycomics of connective tissue glycosaminoglycans. *Proteomics* **2008**, *8* (7), 1384-1397.

(25) Korir, A. K.; Limtiaco, J. F. K.; Gutierrez, S. M.; Larive, C. K. Ultraperformance ion-pair liquid chromatography coupled to electrospray time-of-flight mass spectrometry for compositional profiling and quantification of heparin and heparan sulfate. *Analytical Chemistry* **2008**, *80* (4), 1297-1306.

(26) Stavenhagen, K.; Kolarich, D.; Wuhrer, M. Clinical glycomics employing graphitized

carbon liquid chromatography–mass spectrometry. *Chromatographia* **2015**, *78*, 307-320.

(27) Stickney, M.; Sanderson, P.; Leach, F. E.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Online capillary zone electrophoresis negative electron transfer dissociation tandem mass spectrometry of glycosaminoglycan mixtures. *International Journal of Mass Spectrometry* **2019**, *445*, 116209.

(28) Sanderson, P.; Stickney, M.; Leach, F. E., 3rd; Xia, Q.; Yu, Y.; Zhang, F.; Linhardt, R. J.; Amster, I. J. Heparin/heparan sulfate analysis by covalently modified reverse polarity capillary zone electrophoresis-mass spectrometry. *Journal of chromatography. A* **2018**, *1545*, 75-83.

(29) Zaia, J.; Khatri, K.; Klein, J.; Shao, C.; Sheng, Y.; Viner, R. Complete molecular weight profiling of low-molecular weight heparins using size exclusion chromatography-ion suppressor-high-resolution mass spectrometry. *Analytical Chemistry* **2016**, *88* (21), 10654-10660.

(30) Ruiz-Calero, V.; Moyano, E.; Puignou, L.; Galceran, M. T. Pressure-assisted capillary electrophoresis–electrospray ion trap mass spectrometry for the analysis of heparin depolymerised disaccharides. *Journal of Chromatography A* **2001**, *914* (1), 277-291.

(31) Kirwan, J. A.; Weber, R. J. M.; Broadhurst, D. I.; Viant, M. R. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Scientific Data* **2014**, *1* (1), 140012.

(32) Gill, V. L.; Aich, U.; Rao, S.; Pohl, C.; Zaia, J. Disaccharide analysis of glycosaminoglycans using hydrophilic interaction chromatography and mass spectrometry.

Analytical Chemistry **2013**, *85* (2), 1138-1145.

(33) Kuberan, B.; Lech, M.; Zhang, L.; Wu, Z. L.; Beeler, D. L.; Rosenberg, R. D. Analysis of Heparan Sulfate Oligosaccharides with Ion Pair-Reverse Phase Capillary High Performance Liquid Chromatography-Microelectrospray Ionization Time-of-Flight Mass Spectrometry. *Journal of the American Chemical Society* **2002**, *124* (29), 8707-8718.

(34) Staples, G. O.; Zaia, J. Analysis of Glycosaminoglycans Using Mass Spectrometry. *Current proteomics* **2011**, *8* (4), 325-336.

(35) Hernández-Borges, J.; Neusüß, C.; Cifuentes, A.; Pelzing, M. On-line capillary electrophoresis-mass spectrometry for the analysis of biomolecules. *Electrophoresis* **2004**, *25* (14), 2257-2281.

(36) Helena, H.; Ivona, V.; Roman, Ř.; František, F. Current applications of capillary electrophoresis-mass spectrometry for the analysis of biologically important analytes in urine (2017 to mid-2021): A review. *Journal of Separation Science* **2022**, *45* (1), 305-324.

(37) Volpi, N.; Maccari, F.; Linhardt, R. J. Capillary electrophoresis of complex natural polysaccharides. *Electrophoresis* **2008**, *29* (15), 3095-3106.

(38) Sun, X.; Lin, L.; Liu, X.; Zhang, F.; Chi, L.; Xia, Q.; Linhardt, R. J. Capillary Electrophoresis–Mass Spectrometry for the Analysis of Heparin Oligosaccharides and Low Molecular Weight Heparin. *Analytical Chemistry* **2016**, *88* (3), 1937-1943.

(39) Sarver, S. A.; Schiavone, N. M.; Arceo, J.; Peuchen, E. H.; Zhang, Z.; Sun, L.; Dovichi, N. J. Capillary electrophoresis coupled to negative mode electrospray ionization-mass

spectrometry using an electrokinetically-pumped nanospray interface with primary amines grafted to the interior of a glass emitter. *Talanta* **2017**, *165*, 522-525.

(40) Lubeckyj, R. A.; McCool, E. N.; Shen, X.; Kou, Q.; Liu, X.; Sun, L. Single-Shot Top-Down Proteomics with Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry for Identification of Nearly 600 Escherichia coli Proteoforms. *Analytical Chemistry* **2017**, *89* (22), 12059-12067.

(41) Huhn, C.; Ramautar, R.; Wuhrer, M.; Somsen, G. W. Relevance and use of capillary coatings in capillary electrophoresis–mass spectrometry. *Analytical and Bioanalytical Chemistry* **2010**, *396* (1), 297-314.

(42) Stolz, A.; Jooß, K.; Höcker, O.; Römer, J.; Schlecht, J.; Neusüß, C. Recent advances in capillary electrophoresis-mass spectrometry: Instrumentation, methodology and applications. *Electrophoresis* **2019**, *40* (1), 79-112.

(43) Peuchen, E. H.; Zhu, G.; Sun, L.; Dovichi, N. J. Evaluation of a commercial electro-kinetically pumped sheath-flow nanospray interface coupled to an automated capillary zone electrophoresis system. *Analytical and Bioanalytical Chemistry* **2017**, *409* (7), 1789-1795.

(44) Basu, A.; Weiss, R. J. Glycosaminoglycan Analysis: Purification, Structural Profiling, and GAG-Protein Interactions. *Methods in Molecular Biology* **2023**, *2597*, 159-176.

(45) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nature biotechnology* **2023**, 1-3.

(46) Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D. A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P.-É.; Li, S.; Xia, J. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic acids research* **2021**, *49* (W1), W388-W396.

(47) Zhurov, K. O.; Kozhinov, A. N.; Fornelli, L.; Tsybin, Y. O. Distinguishing Analyte from Noise Components in Mass Spectra of Complex Samples: Where to Cut the Noise? *Analytical Chemistry* **2014**, *86* (7), 3308-3316.

(48) Shen, X.; Yang, Z.; McCool, E. N.; Lubeckyj, R. A.; Chen, D.; Sun, L. Capillary zone electrophoresis-mass spectrometry for top-down proteomics. *TrAC Trends in Analytical Chemistry* **2019**, *120*, 115644.

(49) Maxwell, E. J.; Zhong, X.; Zhang, H.; van Zeijl, N.; Chen, D. D. Y. Decoupling CE and ESI for a more robust interface with MS. *Electrophoresis* **2010**, *31* (7), 1130-1137.

(50) Wu, H.; Tang, K. Highly sensitive and robust capillary electrophoresis-electrospray ionization-mass spectrometry: Interfaces, preconcentration techniques and applications. *Reviews in Analytical Chemistry* **2020**, *39* (1), 45-55.

(51) Sun, L.; Zhu, G.; Yan, X.; Zhang, Z.; Wojcik, R.; Champion, M. M.; Dovichi, N. J. Capillary zone electrophoresis for bottom-up analysis of complex proteomes. *Proteomics* **2016**, *16* (2), 188-196.

(52) Konášová, R.; Koval, D.; Hošek, J.; Kašička, V. Investigating the position of the separation capillary and emitter tube tips in a nanoflow sheath-liquid CE-ESI-MS interface to decouple the ESI potential. *Talanta* **2021**, *228*, 122212.

(53) 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.

(54) Domon, B.; Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal* **1988**, *5* (4), 397-409.

(55) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25* (12), 1323-1324.

(56) van Boeckel, C. A. A.; Petitou, M. The unique antithrombin III binding domain of heparin: a lead to new synthetic antithrombotics. *Angewandte Chemie International Edition in English* **1993**, *32* (12), 1671-1690.

(57) Gama, C. I.; Tully, S. E.; Sotogaku, N.; Clark, P. M.; Rawat, M.; Vaidehi, N.; Goddard Iii, 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.

(58) Soares da Costa, D.; Reis, R. L.; Pashkuleva, I. Sulfation of glycosaminoglycans and its implications in human health and disorders. *Annual review of biomedical engineering* **2017**, *19*, 1-26.

(59) Pye, D. A.; Vives, R. R.; Turnbull, J. E.; Hyde, P.; Gallagher, J. T. Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity. *Journal of Biological Chemistry* **1998**, *273* (36), 22936-22942.

(60) Okada, Y.; Yamada, S.; Toyoshima, M.; Dong, J.; Nakajima, M.; Sugahara, K. Structural Recognition by Recombinant Human Heparanase That Plays Critical Roles in Tumor Metastasis: HIERARCHICAL SULFATE GROUPS WITH DIFFERENTIAL EFFECTS AND THE ESSENTIAL TARGET DISULFATED TRISACCHARIDE SEQUENCE. *Journal of Biological Chemistry* **2002**, *277* (45), 42488-42495.

(61) Staples, G. O.; Bowman, M. J.; Costello, C. E.; Hitchcock, A. M.; Lau, J. M.; Leymarie, N.; Miller, C.; Naimy, H.; Shi, X.; Zaia, J. A chip-based amide-HILIC LC/MS platform for glycosaminoglycan glycomics profiling. *Proteomics* **2009**, *9* (3), 686-695.

(62) Zhang, H.; Chen, H. Correction of migration time of Paeoniae Radix in capillary electrophoresis by powerful one-marker technology. *Analytical Methods* **2011**, *3* (3), 745-750, 10.1039/C0AY00273A.

(63) Huang, Z.-A.; Tan, J.; Li, Y.; Miao, S.; Scotland, K. B.; Chew, B. H.; Lange, D.; Chen, D. D. Y. Migration time correction for dual pressure capillary electrophoresis in semi-targeted metabolomics study. *Electrophoresis* **2022**, *43* (15), 1626-1637.

(64) Šimek, M.; Lemr, K.; Hermannová, M.; Havlíček, V. Analysis of hyaluronan and its derivatives using chromatographic and mass spectrometric techniques. *Carbohydrate Polymers* **2020**, *250*, 117014.

CHAPTER 6

CONCLUSIONS

GAGs are anionic linear polysaccharides that play a significant role in biological functions in organisms. They covalently attach to the core proteins to form proteoglycans except HA, interacting with ligands/co-receptors to convey many biological functions. Their heterogeneity (e.g., length, types, and sulfation patterns) implicates the flexibility in biosynthesis and the specificity of cellular interactions, showing various tissue distribution with tissue-specific isomers.

GAG structures have been sensitively and accurately characterized by mass spectrometry. MS has been also proven useful when combined with chromatographic separation methods. However, there have been several challenges: (1) discrimination of C5 hexuronic acid stereochemistry (2) unintended sulfate group decomposition during MS/MS (3) a high level of heterogeneity inherent to a non-template biosynthesis in biological mixtures. To overcome these challenges, we have developed MS-based analytical techniques using MS/MS and online CZE coupled to MS.

In Chapter 2, distinguishing CS from DS was executed by ion activation methods in MS/MS and multivariate statistical analysis. Various activation methods for MS/MS were able to separate CS diastereomers using PCA. This approach can also be applied to biological samples for the purpose of distinguishing isomers differing only by sites of sulfation. In Chapter 3, we optimized a CZE-MS platform using standard GAG samples

for the application. In Chapter 4, we developed a glycoprofiling approach for intact HS oligosaccharides from human plasma samples, which retains the structural complexity of each individual HS chain and generates HS features for each patient. The workflow developed in this study can lead to improvements in clinical diagnostics for HS-mediated diseases. In addition, CZE-MS was used to illustrate a variety of applications in Chapter 5. Several GAG samples (derived from cell lines, from animal tissue, and synthesized chemoenzymatically) were discussed in subsections, and each sample preparation was examined closely by CZE-MS. Providing sufficient concentrations of CZE in clinical samples can improve the efficiency and reliability of CZE-MS/MS analysis. As a result, we reason that these MS-based methods are versatile and reproducible for a variety of samples. Therefore, MS-based analytical technologies can be further applied for the analysis of complex GAG structures to address not only their structural diversity and provide important data for unravelling challenging biological questions.