



RESEARCH ARTICLE

Multivariate Analysis of Electron Detachment Dissociation and Infrared Multiphoton Dissociation Mass Spectra of Heparan Sulfate Tetrasaccharides Differing Only in Hexuronic acid Stereochemistry

Han Bin Oh,¹ Franklin E. Leach III,² Sailaja Arungundram,^{2,3} Kanar Al-Mafraji,^{2,3} Andre Venot,³ Geert-Jan Boons,^{2,3} I. Jonathan Amster²

¹Department of Chemistry, Sogang University, Seoul, 121-742, Korea

²Department of Chemistry, University of Georgia, Athens, GA, 30602, USA

³Complex Carbohydrate Research Center, University of Georgia, Athens, GA, USA

Abstract

The structural characterization of glycosaminoglycan (GAG) carbohydrates by mass spectrometry has been a long-standing analytical challenge due to the inherent heterogeneity of these biomolecules, specifically polydispersity, variability in sulfation, and hexuronic acid stereochemistry. Recent advances in tandem mass spectrometry methods employing threshold and electron-based ion activation have resulted in the ability to determine the location of the labile sulfate modification as well as assign the stereochemistry of hexuronic acid residues. To facilitate the analysis of complex electron detachment dissociation (EDD) spectra, principal component analysis (PCA) is employed to differentiate the hexuronic acid stereochemistry of four synthetic GAG epimers whose EDD spectra are nearly identical upon visual inspection. For comparison, PCA is also applied to infrared multiphoton dissociation spectra (IRMPD) of the examined epimers. To assess the applicability of multivariate methods in GAG mixture analysis, PCA is utilized to identify the relative content of two epimers in a binary mixture.

Key words: Multivariate analysis, Principal Component Analysis (PCA), Electron Detachment Dissociation (EDD), Glycosaminoglycans (GAGs), Epimers, Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS)

Introduction

Proteoglycans (PGs) play an important role in regulating a variety of physiologic and pathologic processes, cellular

communication, and cell signaling [1–4]. PGs consist of a protein core and carbohydrates known as glycosaminoglycans (GAGs). GAGs are linear biopolymers consisting of alternating hexuronic acid and hexosamine sugar residues. Variations are found in their polysaccharide type, composition, functional group substitution, stereochemistry, and distribution of polysaccharide chain lengths, resulting in great structural diversity. For example, heparin and heparan sulfate (HS), the most structurally complex GAGs, consist of a repeating disaccharide of hexuronic acid (HexA) and glucosamine (GlcN) and vary in

Electronic supplementary material The online version of this article (doi:10.1007/s13361-010-0047-y) contains supplementary material, which is available to authorized users.

Correspondence to: Han Bin Oh; e-mail: hanbinoh@sogang.ac.kr, I. Jonathan Amster; e-mail: jamster@uga.edu

Received: 3 September 2010
Revised: 17 November 2010
Accepted: 17 November 2010
Published online: 28 January 2011

the C5 stereochemistry with acidic residues assigned as either glucuronic acid (GlcA) or iduronic acid (IdoA). The acidic residues may also be 2-O-sulfated. Furthermore, the glucosamine may be N-sulfated or N-acetylated, 6-O-sulfated and occasionally 3-O-sulfated. These variable modifications and stereochemistry make the characterization of GAG oligosaccharides a very challenging analytical task.

Many analytical methods have been used for GAG sequencing, which include one-dimensional (1D) and two-dimensional (2D) NMR, radioactive/fluorescent labeling, polyacrylamide gel electrophoresis, or LC/CE separation [5–8]. These methods are based on the analysis of the GAG oligosaccharides that are obtained from enzymatic digestion of longer GAG chains. Recently, mass spectrometry and, in particular, tandem mass spectrometry has emerged as a more sensitive and molecular specific tool for the analysis of GAG oligosaccharides [8–21]. Collision induced dissociation (CID) of GAG oligosaccharides yielded glycosidic bond and cross-ring cleavages, which provided information about the oligosaccharide sequence and position of sulfation [18, 19]. It has also been shown that the CID product ion patterns of GAG oligosaccharides are dependent upon the charge states of the precursor ion [14, 19] and that the abundances of product ions from glycosidic bond cleavage are indicative of the position of sulfation on the N-acetylgalactosamine (GalNAc) residue [10, 14, 15, 17, 19], for example, 4S/6S sulfation in chondroitin sulfate (CS) oligosaccharides [13]. Interestingly, even hexuronic acid epimerization in CS oligosaccharides could be determined by product ion abundances generated from tandem mass spectrometry [19, 20]. Previously, Zaia and coworkers have shown that the hexuronic acid stereochemistry in CS oligomers could be determined based on the relative abundance of specific X and Y ions produced by CID [19]. Additionally, the relative abundance of key fragment ions was used to determine the fractional abundance of IdoA versus GlcA in tetrasaccharides and hexasaccharides from an enzymatic digestion of chondroitin/dermatan sulfate GAGs [20, 22]. Tandem mass spectrometry ion abundances have also been used to quantify GAG oligosaccharide mixtures [15, 16, 20]. In these previous studies, univariate analysis, which involves a single diagnostic product ion representing each spectrum, was employed.

In recent years, electron detachment dissociation (EDD) has been demonstrated to be a very powerful tool for the characterization of negatively charged GAG oligosaccharides [23–29] ranging from tetrasaccharides to decasaccharides [23, 24, 27]. Irradiation of multiply charged precursor anions with moderate energy (~19 eV) electrons results in extensive glycosidic and cross-ring bond cleavages, which facilitates the determination of the sites of acetylation/sulfation. Compared with threshold dissociation methods such as CID or infrared multiphoton dissociation (IRMPD), the loss of SO₃ from the sites of sulfation was reduced in EDD. For the analysis and interpretation of complicated EDD spectra, it is necessary to employ more reliable, objective statistical tool that makes use of a multiple number of product ions such as multivariate analysis.

Furthermore, it was demonstrated that EDD distinguishes GlcA from IdoA in heparan sulfate (HS) tetrasaccharides based on the occurrence of diagnostic product ions [24]. It was proposed that the diagnostic products were due to a radical-initiated mechanism, with the initial radical site generated by electron detachment from the carboxylate group on acidic residues and subsequent migration through hydrogen atom transfer. This hydrogen atom transfer is dependent upon the proximity of the initial carboxyl radical to hydrogen atoms on adjacent carbons, for example, C2, C3, or C4, thus inducing fragmentation that is sensitive to stereochemistry [24].

In the present study, EDD experiments were performed on four synthetic di-sulfated GAG tetrasaccharides with two epimeric sites; *GlcA*-GlcNAc6S-*GlcA*-GlcNAc6S (GG), *GlcA*-GlcNAc6S-*IdoA*-GlcNAc6S (GI), *IdoA*-GlcNAc6S-*IdoA*-GlcNAc6S (II), and *IdoA*-GlcNAc6S-*GlcA*-GlcNAc6S (IG) (shown as Scheme 1 in the Supplementary Material). Unlike tetrasaccharides prepared by enzymatic digestion of longer GAG oligosaccharides, the non-reducing end (NRE) has a saturated sugar ring, and thus the tetrasaccharides contain a stereocenter on each hexuronic acid residue. The EDD spectra of the doubly deprotonated precursor ion for each epimer are nearly identical, varying only in product ion intensity. To facilitate the differentiation of the four tetrasaccharide EDD spectra, we apply a multivariate analysis (MVA) method [30–32]. MVA has been widely applied in many scientific disciplines to extract valuable information from complicated data sets. Factor analysis has been used to help elucidate basic chemical interactions in gas chromatography and to interpret fragmentation patterns in electron impact mass spectrometry [33–36]. Furthermore, it has been shown that factor analysis can separate the data from mass spectra of mixtures into the mass spectra of the pure components and give their respective concentrations [36]. Recently, MVA has been successfully applied to discern the characteristic features of large proteomics data sets [37–39] and carbohydrate structure analysis [40, 41]. For example, comparative multivariate statistical analysis of 2D LC-MS data for proteomic samples dictated the peptide masses that were differentially expressed, and subsequently targeted during LC-MS/MS as peptides of interest [37, 38]. In the current study, we expand the use of MVA to the analysis of GAG tandem mass spectra; specifically, we use principal component analysis (PCA) to differentiate the acid stereochemistry in four synthetic HS epimers through the analysis of EDD and IRMPD spectra that contain peaks assigned to both glycosidic and cross-ring cleavages.

Experimental

Tetrasaccharide Synthesis

Four HS tetrasaccharides, *GlcA*-GlcNAc6S-*GlcA*-GlcNAc6S (GG), *GlcA*-GlcNAc6S-*IdoA*-GlcNAc6S (GI), *IdoA*-GlcNAc6S-*IdoA*-GlcNAc6S (II), and *IdoA*-

GlcNAc6S-GlcA-GlcNAc6S (IG), were synthesized as described previously and purified by silica gel column chromatography [42]. Prepared structures were confirmed by ^1H NMR and accurate mass measurement by FTICR-MS.

Mass Spectrometry

Experiments were performed on a 9.4 T Bruker Apex Ultra QhFTMS (Billerica, MA, USA) equipped with an Apollo II dual source, a 25 W CO_2 laser (Synrad model J48-2; Mukilteo, WA, USA) for IRMPD, and an indirectly heated hollow cathode for generating electrons for EDD. Solutions of each tetrasaccharide were introduced at a concentration of 0.1 mg/mL in 50:50:0.1 methanol:H₂O:formic acid (Sigma, St. Louis, MO, USA) and ionized by nanospray using a pulled fused silica tip (model FS360-75-15-D-20; New Objective, Woburn, MA, USA). Formic acid was added to the ESI spray solvent to reduce Na/H heterogeneity [28]. The sample solutions were infused at a rate of 10 $\mu\text{L}/\text{h}$ and were examined in negative ion mode.

In EDD experiments, precursor ions were mass-selected in the external quadrupole and accumulated for 1 to 2 s in an rf-only hexapole before injection into the FTICR analyzer cell. The precursor ion selection was further refined by a coherent harmonic excitation frequency (CHEF) event [43]. The isolated precursor ions were then irradiated with electrons for 1 s. During EDD, the cathode was biased at -19.0 V, the extraction lens was set to -18.5 V, and the cathode heater to 1.4 A. In IRMPD experiments the same experimental setup as in EDD experiments was used, except the electron irradiation period was replaced with a laser pulse. The isolated precursor ions were irradiated for 0.02 to 0.25 s with the laser beam attenuated to 60% of full laser power. In both experiments, 24 acquisitions were signal averaged per mass spectrum. For each mass spectrum, 512 K points were acquired, padded with one zero fill, and apodized using a sinebell window.

Data Analysis

Product ion assignments were made based on accurate mass measurement and aided by Glycoworkbench [44]. Monoisotopic peaks corresponding to assigned glycosidic and cross-ring cleavages and their respective abundances were compiled for each mass spectrum in Bruker Data Analysis ver. 3.4. External calibration of mass spectra produced mass accuracy of 5 ppm. Internal calibration was also performed using confidently assigned glycosidic bond cleavage products as internal calibrants, providing mass accuracy of <1 ppm. Due to the larger number of low intensity products formed by EDD, only peaks with $S/N > 10$ are reported. All products are reported using the Domon and Costello nomenclature [45]. The annotation of glycosidic fragmentation products that occur with the loss of additional hydrogen (compared with products found in conventional ion activation experiments) are modified by appending single or

double quotes to the B, C, Y, or Z designation, to indicate the loss of 1 or 2 hydrogen atoms, respectively, consistent with the EDD literature [23–28].

Principal component analysis (PCA) was performed using PLS Toolbox (Eigenvector Research, Inc., Wenatchee, WA). The abundances of 59 assigned fragment ions were normalized with respect to total ion abundance in each EDD spectrum. An input data matrix was constructed with each row containing the mass spectrum of a single tetrasaccharide epimer (samples) and each column, the normalized abundance of an assigned fragment ion (variables). For each tetrasaccharide, five EDD spectra were obtained in the same day. Prior to PCA, each data set was mean-centered and cross-validated. In brief, PCA analysis explains the variance in the examined datasets by generation of a system of linear equations, i.e., principal components. The plot of principal component scores shows the relationship between the samples, e.g., similar samples will be grouped. Examination of the PCA loading plots reveals the variables that contribute to the sample distinction, e.g., minimally contributing variables will be located near the origin whereas variables with high contribution will have relatively large values.

Results and Discussion

EDD of HS Tetrasaccharide Epimers

The doubly deprotonated anion, $[\text{M} - 2\text{H}]^{2-}$, of each tetrasaccharide shown in Scheme 1 (Supplementary Material) was subjected to a 1.0 s, 19 eV electron irradiation to obtain their EDD mass spectra. As previously stated, these compounds differ only in their hexuronic acid stereochemistry; therefore, the four potential doubly deprotonated precursor ions are isobaric and cannot be distinguished at the MS level.

EDD of each pure tetrasaccharide produced extensive fragmentation shown in Figure 1. Each fragment in Figure 1 arose from the cleavage of a glycosidic bond (B, C, Y, or Z ions), cross-ring (A or X ions), and/or an accompanying neutral loss. A total of 59 fragment ions were assigned for each tetrasaccharide. Cleavage of all glycosidic bonds was observed, e.g., Y_1^- , Y_3^{2-} , C_2^- , Z_2^- , B_2^- , C_3^- , Z_3^- , and Y_3^- . Based on these glycosidic bond cleavages, the extent of sulfation and N-acetylation can be determined at the residue level. Glycosidic bond fragmentation is, however, not unique to EDD, but also prevalent in CID or IRMPD. Many of the fragment ions found in Figure 1 were assigned to cross-ring cleavages, which is consistent with EDD results for other GAGs [23–29]. These cleavages enable the localization of modifications to a particular site on a saccharide ring, in this case the 6-O-sulfation of each N-acetyl glucosamine residue. In some cases, fragment ions were accompanied with the neutral loss of SO_3 and CO_2 : for example, $(\text{B}_3\text{-SO}_3)^-$ and $(\text{Y}_3\text{-CO}_2)^-$.

Previously, EDD of HS tetrasaccharide epimers has produced diagnostic product ions that can be used to

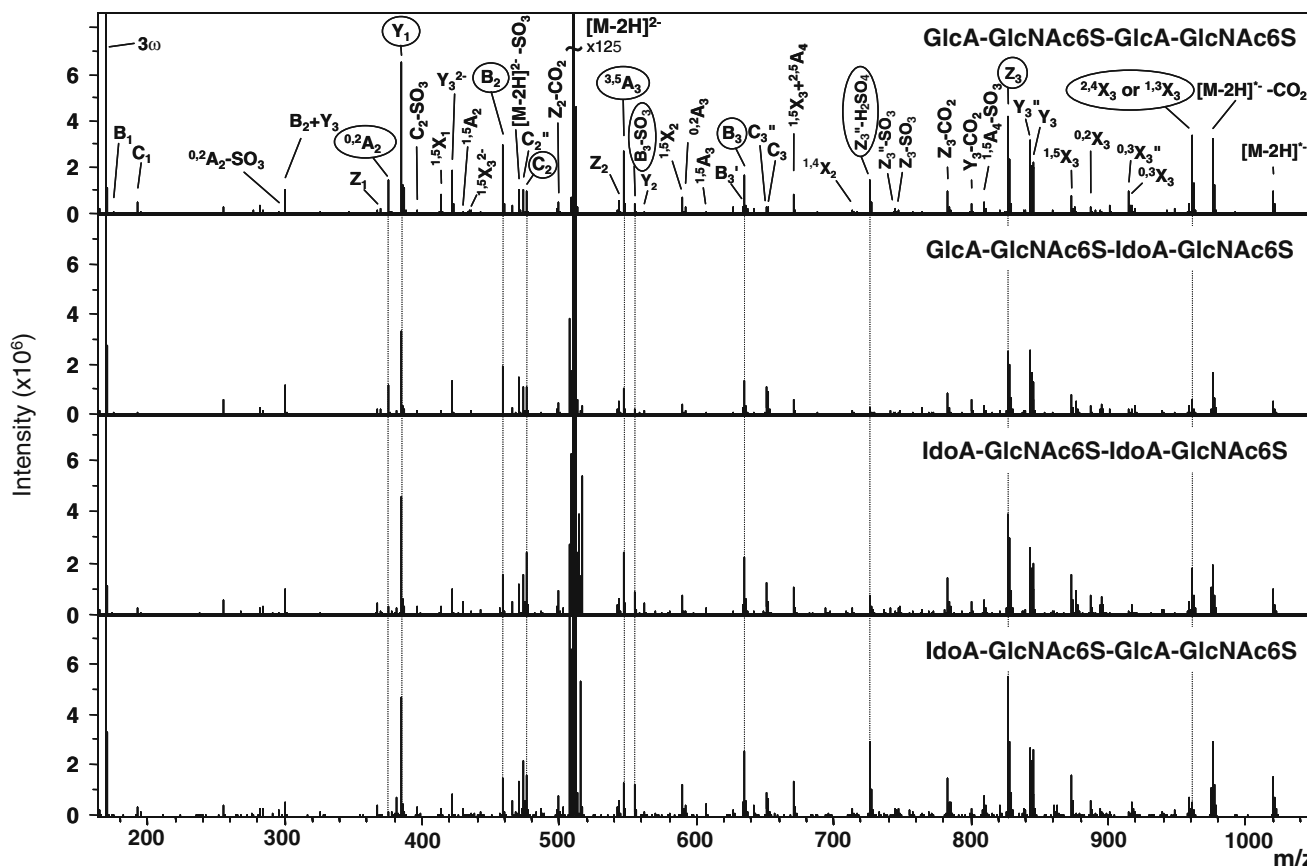


Figure 1. EDD mass spectra of the four different HS tetrasaccharide epimers: (a) GlcA-GlcNAc6S-GlcA-GlcNAc6S (GG), (b) GlcA-GlcNAc6S-IdoA-GlcNAc6S (GI), (c) IdoA-GlcNAc6S-IdoA-GlcNAc6S (II), and (d) IdoA-GlcNAc6S-GlcA-GlcNAc6S (IG). Fragments of high loading values in PCA (see Figure 2 and 3) are circled and denoted with dotted lines. (Differences in the appearance of precursor ions between epimers are due to adjacent fragment ions of varying intensity.)

distinguish IdoA from GlcA [24]. From a mechanistic aspect, the proximity of the initially generated carboxy radical to a neighboring hydroxyl group on the sugar ring directs different hydrogen transfer rates and, therefore, distinct cross-ring cleavage patterns. With this in mind, the EDD mass spectra in Figure 1 are compared. A striking distinction cannot be found between the four spectra in terms of the overall fragment ion distributions. (This result is based on the examination of EDD spectra generated from the activation of a doubly deprotonated ion generated from a tetrasaccharide with two sulfate groups. The $[M - 2H]^{2-}$ precursor ion is the most easily obtained species and presents two ionized sulfates based on pKa. To obtain EDD spectra containing products consistent with prior results, it is necessary to ionize and detach from a carboxyl group. This mechanistic aspect is the subject of a manuscript in preparation.) However, a close inspection of relative abundances of fragments shows a number of diagnostic product ion peaks that may enable the differentiation of one epimer from another. Notable fragment peaks are the followings: $(B_2 + Y_3)^-$, $^{0,2}A_2^-$, Y_3^{2-} , B_2^- , C_2^- , $^{3,5}A_3^-$, $(B_3 - SO_3)^-$ at, $C_3^{''}$, $(Z_3^{''} - H_2SO_4)^-$, Z_3^- , Y_3^- , and $^{1,3}X_3$ or $^{2,4}X_3^-$. (Masses are provided as Supplementary Data.) More

specifically, $^{0,2}A_2^-$ fragment appears in high abundance for the tetrasaccharides containing GlcA at the NRE. In the case of the C_2^- and $^{3,5}A_3^-$ fragments, their abundances stand out for II. As another example, the abundance of the $C_3^{''}$ peak is relatively higher when IdoA is located at the third saccharide residue from the NRE, which is consistent with prior EDD results when IdoA is located at this position [24]. Although visual inspection of these spectra can provide some insight into the assignment of hexuronic acid stereochemistry, it is necessary to consider as many diagnostic peaks as possible in order to characterize an unknown EDD mass spectrum in a reliable manner.

Principal Component Analysis (PCA) of EDD Mass Spectra

In the present study, a statistical method is employed as a reliable, objective tool for distinguishing the EDD spectra obtained from four HS tetrasaccharide epimers. As briefly described above, it is difficult to gain discriminant information using a univariate analysis. Multivariate analysis (MVA) is therefore used for the differentiation of the obtained mass spectra. Specifically, principal component analysis (PCA), one of the most common multivariate analysis tools, is used in the

present study. This method is known to be very useful for the extraction and visualization of the most important features from complicated data sets [30, 32]. PCA was carried out for 20 EDD mass spectra obtained from quintuplicate experiments for each of the four tetrasaccharides. The relative abundances of 59 fragment ions shown in Figure 1 were normalized with respect to the total abundance and used to construct an input data matrix for PCA.

Selection of three principal components (PCs) explains 99% of the total variance in the data set (PC1=51%, PC2=33%, and PC3=15%). The obtained PCA scores and loading plots are shown in Figure 2 and 3. Grouping of the replicates for each HS tetrasaccharide is visualized in the 2D score plot of PC2 vs. PC1, shown in Figure 2a, with each epimer located in a unique quadrant. To further visualize these three PCs, a 2D score plot for PC3 versus PC1 and a three dimensional (3D) plot for all three PCs are also shown in Figure 3a, b, respectively. The

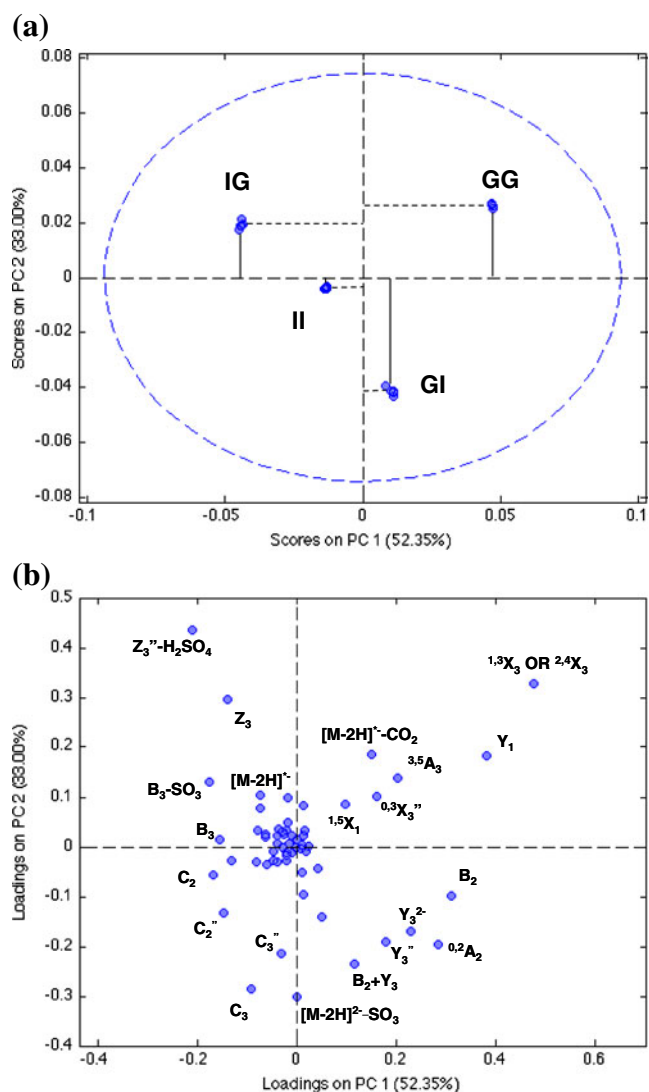


Figure 2. PC2 versus PC1 plots of (a) scores and of (b) loadings obtained from PCA of fragment ions that arose from EDD of the four HS epimers: GG, GI, IG, and II

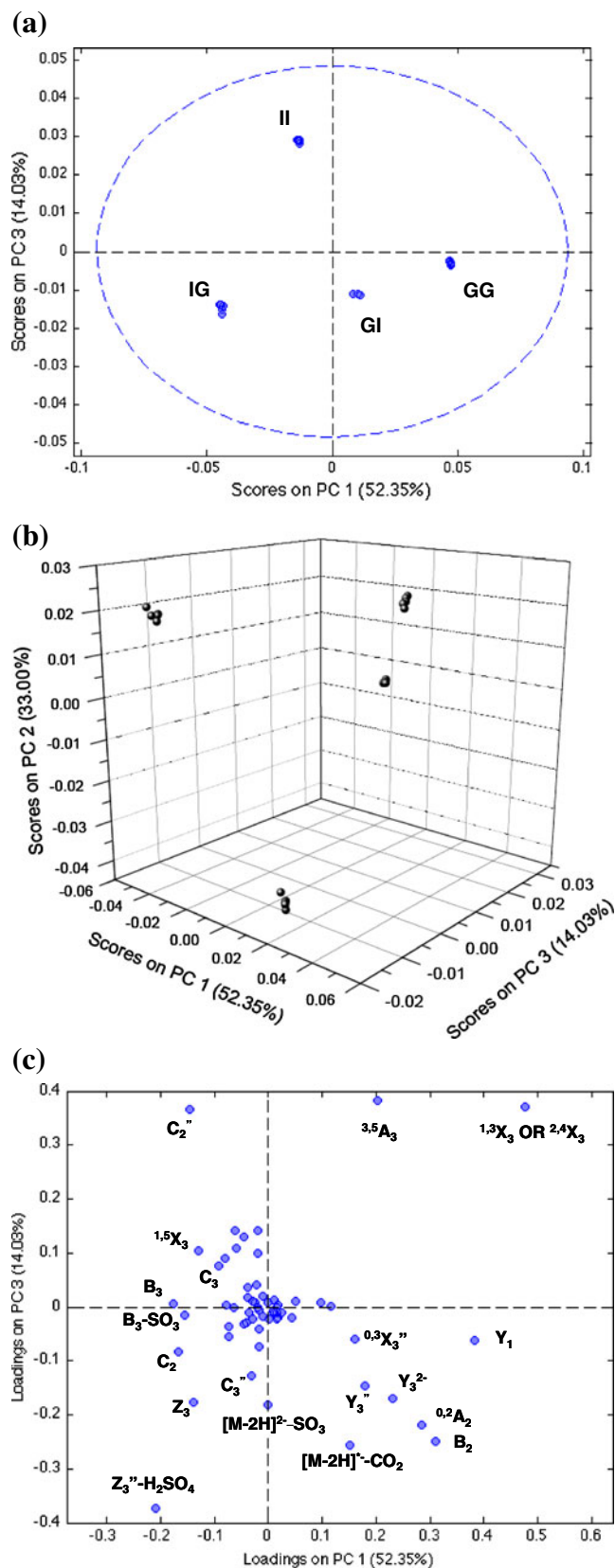


Figure 3. (a) 2D score plot of PC3 versus PC1, (b) 3D score plot of PC1:PC2:PC3, and (c) loading plot of PC3 versus PC1. EDD fragment ions obtained for the four HS epimers (GG, GI, IG, and II) are analyzed

score plots show the effective clustering of replicates originating from the same epimer, with a clear separation between the replicates of different epimers. For clear distinction of the isomers, the separation between clusters of replicates should be larger than that of the replicates themselves. Indeed, as shown in Figure 2a, 3a, the separations between the clusters are much larger than those between replicates within a cluster. Furthermore, variations in the PCA scores of each replicate are generally negligible, indicating that the EDD experiments were highly reproducible.

As shown in Figure 2a, when the four tetrasaccharides are projected onto the PC1 axis (solid lines), they are all well separated from each other, as the high value of variance explained by PC1 (51%) indicates. Specifically, the GG replicates have the highest positive scores along the PC1 axis, while IG replicates have the most negative scores. These two epimers are most readily distinguished by PC1 alone. Closer examination of the PCA score plot reveals differentiation of the stereochemistry of the non-reducing end hexuronic acid by PC1. Tetrasaccharides containing IdoA at the NRE have negative PCA scores whereas GlcA results in positive scores for PC1. Further differentiation can be made in PC2 for the stereochemistry of the third saccharide residue. The epimers containing GlcA at the third position have positive scores in PC2 while the tetrasaccharides with IdoA have negative scores. In this plot, however, II and GI have the lowest degree of separation in PC1, but the separation between II and GI is more distinct in the plot of PC3 versus PC1 visualized in Figure 3a. The separation between the clustered groups can be more clearly demonstrated in a 3D plot displayed in Figure 3b.

In general, principal components are constructed by a weighted combination of fragment ion intensities in a manner that best explains the variance in the data, revealing the internal structure of the data. The relative contribution of each fragment in constructing principal components can be determined from the loading plots. For example, Figure 2b shows that the following peaks have high positive loading values on PC1: $^{1,3}X_3$ or $^{2,4}X_3^-$, Y_1^- , B_2^- , and $^{0,2}A_2^-$. High abundances of these peaks can be collectively used as characteristics of epimers containing GlcA at the NRE. In contrast, high negative loading values on PC1 were observed for $(Z_3''\text{-H}_2\text{SO}_4)^-$, $(B_3\text{-SO}_3)^-$, B_3^- , and Z_3^- , are diagnostic for epimers containing IdoA at the NRE.

As illustrated in Figure 3a, II and GI isomers are well separated on PC3, on which $^{1,3}X_3$ or $^{2,4}X_3^-$, $^{3,5}A_3^-$, and C_2^- have high positive loadings, while B_2^- , $(Z_3''\text{-H}_2\text{SO}_4)^-$, and $^{0,2}A_2^-$ have high negative loadings (Figure 3c). To make the above-mentioned important loadings easily noticeable, these loadings are denoted with dotted lines and assignments are circled in Figure 1.

It may be possible that relative abundances of fragments in EDD product mass spectra can vary day to day to such a degree that it may significantly affect PCA results. To shed light on this possibility, EDD was carried out for the GI

tetrasaccharide, chosen as a representative example, on several different days with almost identical experimental parameters. To examine the effect of EDD parameter variation on PCA results, the extraction lens was varied on d 3. Newly acquired EDD results for GI isomers were projected onto the PCA plots previously shown in Figure 2a, 3a. As shown in the projected plots of Figure 4a, GI PCA scores from spectra acquired on the same day are generally well clustered in the vicinity of the original GI scores, but are generally distinct from those acquired on different days. However, they are still well separated from the other GG, IG, and II points. We assume that the deviations are dependent on instrument and experimental factors, and that the spread in principal component values will be similar for all epimers. These results indicate that there were some variations in the relative abundances of fragments in GI EDD mass spectra when acquired on different days. The basis of the statistical variance in the GI PCA scores is not

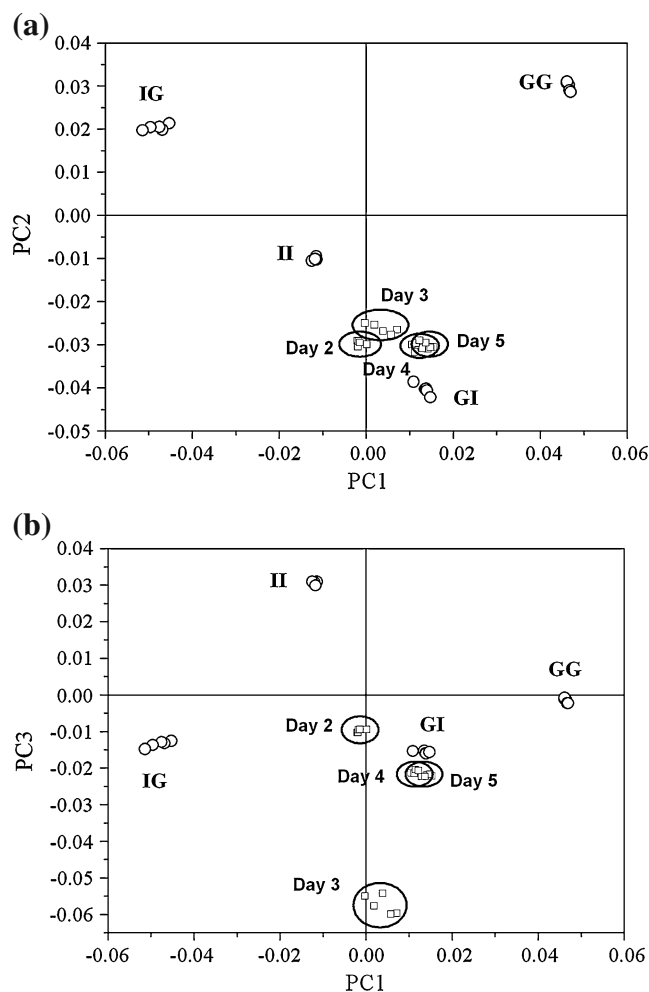


Figure 4. Projection of GI EDD results obtained on several different days onto the plots of (a) PC2 versus PC1 from Figure 2(a) and (b) PC3 versus PC1 Figure 3(a). The data points representing those obtained in the same day are separately circled

discernable by visual inspection. Even though the electron current was altered from the optimal value of 15 μA [29] on d 3, the projection of the PC2 versus PC1 results is in good agreement with other day-to-day variation. Only when PC3 is examined does PCA reveal a significant difference in the data set when compared to the daily variation, shown in Figure 4b.

Binary Mixture Analysis

The determination of mixture composition is one of the primary aims during the analysis of real world samples. Here it is examined whether PCA can be employed to determine the composition of a binary mixture based on EDD mass spectra. As a demonstrative example, four binary mixtures of IG and GG, 7:3, 5:5, 3:7, 1:9 (IG:GG), were separately subjected to EDD mass spectrometry and the resulting mass spectra analyzed using PCA.

Figure 5 shows the projection of the PCA scores obtained by EDD mass spectrometry of the four IG:GG mixtures onto the PCA score plot generated from the EDD of the pure compounds, previously shown in Figure 2a. In this plot, small, unfilled squares located between the data points associated with the pure IG and GG represent the PC1 and PC2 scores of the four binary mixtures. To determine whether the PC1 scores of the unfilled squares well reflect the mixture ratio of IG/GG, these squares are further projected onto the PC1 axis. As shown in Figure 5, the projection of 7:3, 5:5, 3:7, and 1:9 mixtures indicate 7.3:2.7, 4.8:5.2, 3.4:6.6, and 0.7:9.3 mixture ratios, respectively, which are in a rough agreement with the prepared ratios. This result is largely due to the explanation of greater than 99% of the total variance by three principal components and specifically 51% of the variance by PC1 alone. Data sets that are not as reproducible would encompass more principal

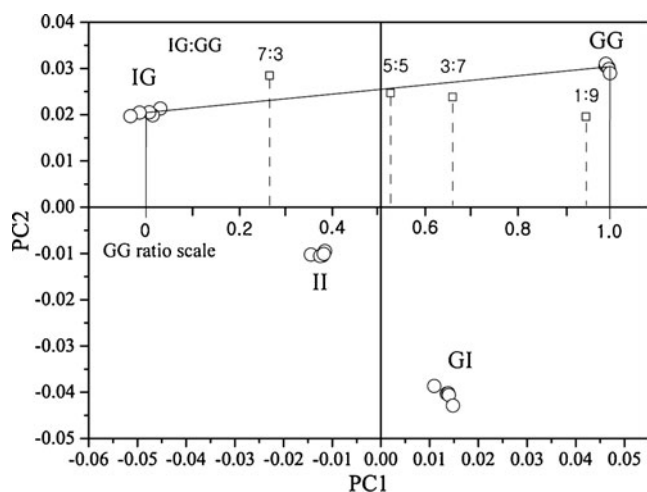


Figure 5. Projection of EDD results for IG/GG binary mixtures of 7:3, 5:5, 3:7, 1:9 (IG:GG) ratios onto the plot of Figure 2(a). Small unfilled squares represent the PCA results for the four binary mixtures. Along the PC1 axis, GG ratio scale is displayed, onto which PCA results for the four binary mixtures are projected

components and would require alternative MVA methods such as partial least squares (PLS) to examine mixtures. Considering non-negligible day-to-day variations of EDD mass spectrometry results as described above, it seems difficult to obtain better estimates of mixture ratios using this approach. The same approach was also applied to PC2, but was not successful due to less explained variance in PC2 for IG and GG. Nevertheless, this approach of projecting the PCA scores of mixtures onto the PCA axis of the already-established score plot of pure components appears to be a useful way to provide a rough estimate of mixing ratio, at least for binary mixtures.

PCA Analysis of IRMPD Spectra

We have shown above that PCA analysis is a very useful tool for the analysis of complex EDD mass spectra of GAGs, but the question may arise whether PCA analysis is also beneficial for the analysis of tandem mass spectra of GAGs due to threshold activation methods such as CID or IRMPD, which generally produce fewer product ions. Included in the Supplemental Material are the IRMPD mass spectra for the four synthetic HS epimeric isomers. Compared with the EDD mass spectra in Figure 1, IRMPD mass spectra exhibit fewer fragments. The dominant fragment ions are Y_1^- , Y_3^{-2} , B_2^- , $[(M - 2H) - SO_3]^{2-}$, and B_3^- . When the four IRMPD mass spectra are compared, it is clearly visible that the relative abundances of B_3^- vary significantly from one spectrum to another; the relative abundance was the highest for II isomer, while the lowest for GG isomer. In contrast, B_2^- has a low abundance for the II isomer, but high abundance for GG isomer.

The PCA results for the IRMPD spectra are shown in Figure 6. Eighteen fragments were assigned in each spectrum, acquired in quintuplicate for each epimer, resulting in 20 total spectra for PCA analysis. PCA indicates that three principal components explain 99% of the total variance: PC1, 68%; PC2, 22%; PC3, 9%. As Figure 6a, b clearly indicate, the four epimers are well separated into four clustered groups. In the score plot of PC2 versus PC1 the isomers of GG and II are most separated, while in the plot of PC3 versus PC1 the isomers of GG and IG are relatively more distinguishable. The loading plots of Figure 6c and d show the relative contributions in constructing principal components. As seen in Figure 6c, B_3^- has the highest positive value in the PC1 coordinate, which suggests that the high abundance of this peak is a signature of the II isomer. In contrast, the B_2^- fragment shows the highest negative value along the PC1 coordinate, a signature of the GG isomer. This result is consistent with the visual inspection of the IRMPD mass spectra discussed previously. On the other hand, the IG and GI isomers are well separated in the PC3 coordinate shown in Figure 6b. In distinguishing these two isomers, C_1^- , $(B_3-SO_3)^-$, and Y_3^{-2} are shown to contribute significantly. This result can be easily confirmed by examination of the IRMPD mass spectra (included in Supplemental Material).

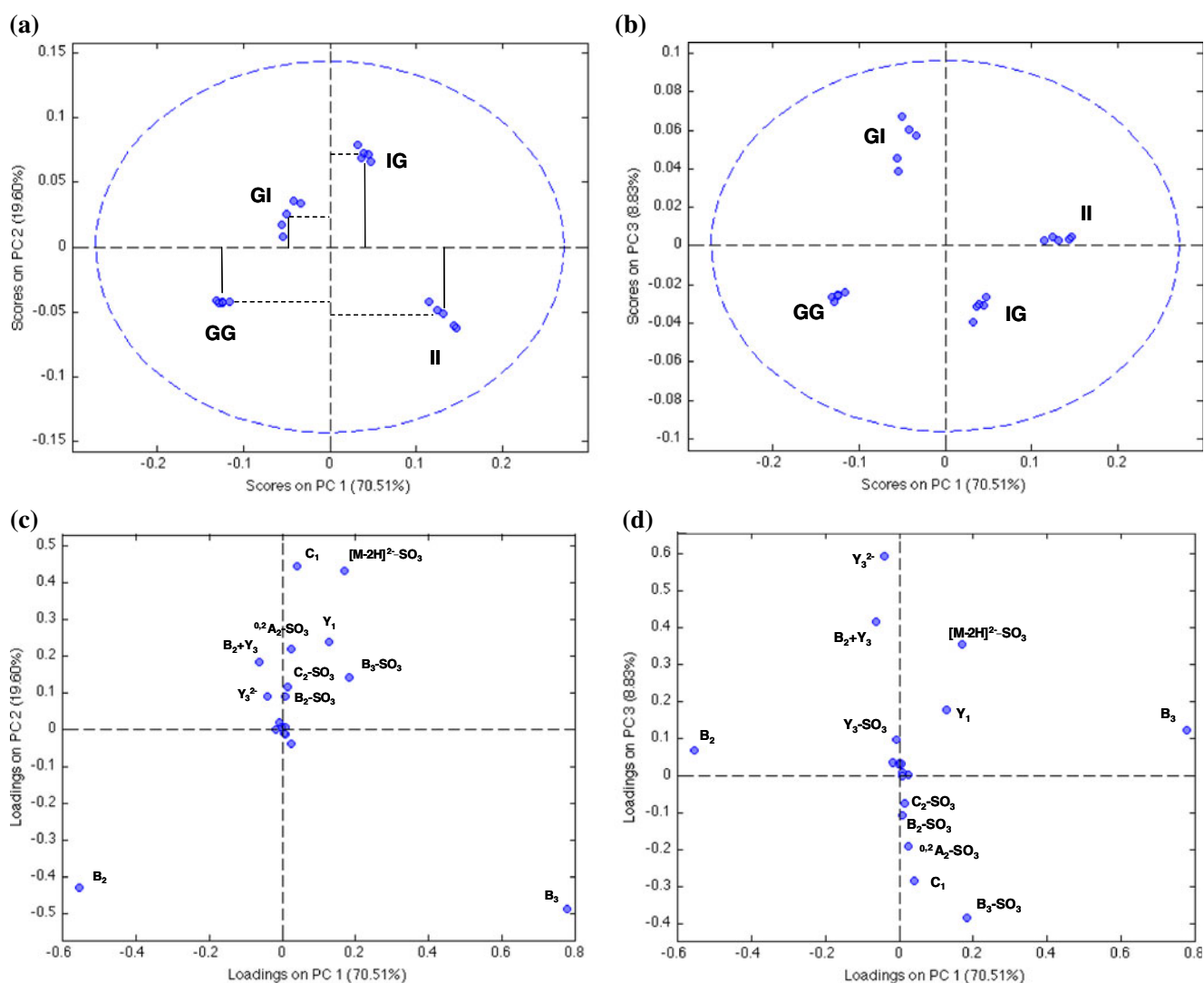


Figure 6. Score plots of (a) PC2 versus PC1 and of (b) PC3 versus PC1. Loading plots of (c) PC2 versus PC1 and of (d) PC3 versus PC1. PCA was made for fragment ions arising from IRMPD of four HS epimers (GG, GI, IG, and II)

Conclusions

A MVA approach to the analysis of EDD mass spectra of GAG tetrasaccharide anions is described. In general, EDD spectra of GAG oligomer anions generally exhibit a large number of fragment ions that arose from cleavages of glycosidic and/or cross-ring bonds. Indeed, in the present study for synthetic HS tetrasaccharide anions that contain two epimeric sites, i.e., *GlcA*-*GlcNAc6S*-*GlcA*-*GlcNAc6S* (GG), *GlcA*-*GlcNAc6S*-*IdoA*-*GlcNAc6S* (GI), *IdoA*-*GlcNAc6S*-*GlcA*-*GlcNAc6S* (IG), and *IdoA*-*GlcNAc6S*-*IdoA*-*GlcNAc6S* (II), EDD produced highly complex product ion spectra. Due to the high spectral complexity, it was neither simple nor straightforward to distinguish one spectrum from another by visual inspection. To aid the differentiation of the obtained EDD spectra, PCA was performed to extract the information necessary to characterize each spectrum. In the PCA score plots, EDD spectra of

the four tetrasaccharides were well grouped on the basis of their epimeric states. Furthermore, the loading plots indicated the characteristic fragments that contributed significantly to differentiating each of EDD spectra. It was also demonstrated that the compositions of binary tetrasaccharide mixtures could be broadly analyzed by projecting their EDD spectra results onto the score plot of 'pure' tetrasaccharides. PCA was also shown to be very useful for analyzing the IRMPD spectra of the four synthetic HS tetrasaccharide anions. This study clearly demonstrates that PCA is a promising tool for characterizing GAG oligosaccharides, particularly, isobaric or epimeric GAG mixtures, in combination with EDD and IRMPD tandem mass spectrometry.

Acknowledgments

H.B.O. thanks the National Research Foundation of Korea (NRF-2009-0075245) for the financial support. F.E.L. III and I.

J.A. gratefully acknowledge financial support from the National Institutes of Health grant no. 2R01-GM038060-20. F.E.L. III, I. J.A., S.A., K.A.-M, A.V., and G.-J.B. gratefully acknowledge financial support from the Center for Research Resource of the National Institutes of Health grant no P41RR005351.

References

- Linhardt, R.J., Toida, T.: Role of Glycosaminoglycans in Cellular Communication. *Acc Chem Res* **37**, 431–438 (2004)
- Fannon, M., Forsten, K.E., Nugent, M.A.: Potentiation and Inhibition of bFGF Binding by Heparin: A Model for Regulation of Cellular Response. *Biochemistry* **39**, 1434–1445 (2000)
- Wu, Z.L., Zhang, L., Yabe, T., Kuberan, B., Beeler, D.L., Love, A., Rosenberg, R.D.: The Involvement of Heparan Sulfate (HS) in FGF1/HS/FGFR1 Signaling Complex. *J Biol Chem* **278**, 17121–17129 (2003)
- Handel, T.M., Johnson, Z., Crown, S.E., Lau, E.K., Sweeney, M., Proudfoot, A.E.: Regulation of Protein Function by Glycosaminoglycans; as Exemplified by Chemokines. *Annu Rev Biochem* **74**, 385–410 (2005)
- Horne, A., Gettins, P.: ¹H-N.M.R. Spectral Assignments for Two Series of Heparin-Derived Oligosaccharides. *Carbohydr Res* **225**, 43–57 (1992)
- Rudd, P.M., Dwek, R.A.: Rapid, Sensitive Sequencing of Oligosaccharides from Glycoproteins. *Curr Opin Biotechnol* **8**, 488–497 (1997)
- Jackson, P.: The Analysis of Fluorophore-Labeled Carbohydrates by Polyacrylamide Gel Electrophoresis. *Mol Biotechnol* **5**, 101–123 (1996)
- Saad, O.M., Leary, J.A.: Heparin Sequencing Using Enzymatic Digestion and ESI-MSⁿ with HOST: A Heparin/HS Oligosaccharide Sequencing Tool. *Anal Chem* **77**, 5902–5911 (2005)
- Reinhold, V.N., Carr, S.A., Green, B.N., Petitou, M., Choay, J., Sinay, P.: Structural Characterization of Sulfated Glycosaminoglycans by Fast Atom Bombardment Mass Spectrometry: Application to Heparin Fragments Prepared by Chemical Synthesis. *Carbohydr Res* **161**, 305–313 (1987)
- Lamb, D.J., Wang, H.M., Mallis, L.M., Linhardt, R.J.: Negative Ion Fast-Atom Bombardment Tandem Mass Spectrometry to Determine Sulfate and Linkage Position in Glycosaminoglycan-Derived Disaccharides. *J Am Soc Mass Spectrom* **3**, 797–803 (1992)
- Ii, T., Kubota, M., Hirano, T., Ohashi, M., Yoshida, K., Suzuki, S.: FAB CID-MS/MS Characterization of Tetrasaccharide Tri- and Tetrasulfate derived from the Antigenic Determinant Recognized by the Anti-Chondroitin Sulfate Monoclonal Antibody MO-225. *Glycoconj J* **12**, 282–289 (1995)
- Zaia, J., Costello, C.E.: Compositional Analysis of Glycosaminoglycans by Electrospray Mass Spectrometry. *Anal Chem* **73**, 233–239 (2001)
- Zaia, J., McClellan, J.E., Costello, C.E.: Tandem Mass Spectrometric Determination of the 4S/6S Sulfation Sequence in Chondroitin Sulfate Oligosaccharides. *Anal Chem* **73**, 6030–6039 (2001)
- McClellan, J.E., Costello, C.E., O'Conno, P.B., Zaia, J.: Influence of Charge State on Product Ion Mass Spectra and the Determination of 4S/6S Sulfation Sequence of Chondroitin Sulfate Oligosaccharides. *Anal Chem* **74**, 3760–3771 (2002)
- Desaire, H., Leary, J.A.: Detection and Quantification of the Sulfated Disaccharides in Chondroitin Sulfate by Electrospray Tandem Mass Spectrometry. *J Am Soc Mass Spectrom* **11**, 916–920 (2000)
- Saad, O.M., Leary, J.A.: Compositional Analysis and Quantification of Heparin and Heparan Sulfate by Electrospray Ionization Ion Trap Mass Spectrometry. *Anal Chem* **75**, 2985–2995 (2003)
- Zaia, J., Costello, C.E.: Tandem Mass Spectrometry of Sulfated Heparin-Like Glycosaminoglycan Oligosaccharides. *Anal Chem* **75**, 2445–2455 (2003)
- Saad, O.M., Leary, J.A.: Delineating Mechanisms of Dissociation for Isomeric Heparin Disaccharides Using Isotope Labeling and Ion Trap Tandem Mass Spectrometry. *J Am Soc Mass Spectrom* **15**, 1274–1286 (2004)
- 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. *J Am Soc Mass Spectrom* **14**, 1270–1281 (2003)
- Miller, M.J.C., Costello, C.E., Malmstrom, A., Zaia, J.: A Tandem Mass Spectrometric Approach to Determination of Chondroitin/Dermatan Sulfate Oligosaccharide Glycoforms. *Glycobiology* **16**, 502–513 (2006)
- Sweeney, M.D., Yu, Y., Leary, J.A.: Effects of Sulfate Position on Heparin Octasaccharide Binding to CCL2 Examined by Tandem Mass Spectrometry. *J Am Soc Mass Spectrom* **17**, 1114–1119 (2006)
- Hitchcock, A.M., Costello, C.E., Zaia, J.: Glycoform Quantification of Chondroitin/Dermatan Sulfate Using a Liquid Chromatography-Tandem Mass Spectrometry Platform. *Biochemistry* **45**, 2350–2361 (2006)
- Wolff, J.J., Amster, I.J., Chi, L., Linhardt, R.J.: Electron Detachment Dissociation of Glycosaminoglycan Tetrasaccharides. *J Am Soc Mass Spectrom* **18**, 234–244 (2007)
- Wolff, J.J., Chi, L., Linhardt, R.J., Amster, I.J.: Distinguishing Glucuronic from Iduronic Acid in Glycosaminoglycan Tetrasaccharides by Using Electron Detachment Dissociation. *Anal Chem* **79**, 2015–2022 (2007)
- Wolff, J.J., Laremore, T.N., Aslam, H., Linhardt, R.J., Amster, I.J.: Electron-Induced Dissociation of Glycosaminoglycan Tetrasaccharides. *J Am Soc Mass Spectrom* **19**, 1449–1458 (2008)
- 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. *J Am Chem Soc* **130**, 2617–2625 (2008)
- Wolff, J.J., Laremore, T.N., Busch, A.M., Linhardt, R.J., Amster, I.J.: Electron Detachment Dissociation of Dermatan Sulfate Oligosaccharides. *J Am Soc Mass Spectrom* **19**, 294–304 (2008)
- 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. *J Am Soc Mass Spectrom* **19**, 790–798 (2008)
- 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. *Int J Mass Spectrom* **276**, 110–115 (2008)
- Feinstein, A.R.: Multivariable Analysis. Yale University Press, New Haven (1996)
- Kowalski, B.R.: Chemometrics. *Anal Chem* **52**, 112R–122R (1980)
- Brereton, R.G.: Introduction to Multivariate Calibration in Analytical Chemistry. *Analyst* **125**, 2125–2154 (2000)
- Rozett, R.W., Petersen, E.M.: Classification of Compounds by the Factor Analysis of Their Mass Spectra. *Anal Chem* **48**, 817–825 (1976)
- Justice, J.B., Isenhour, T.L.: Factor Analysis of Mass Spectra. *Anal Chem* **47**, 2286–2288 (1975)
- Ritter, G.L., Lowry, S.R., Isenhour, T.L., Wilkins, C.L.: Factor Analysis of the Mass Spectra of Mixtures. *Anal Chem* **48**, 591–595 (1976)
- Knorr, F.J., Futrell, J.H.: Separation of Mass Spectra of Mixtures by Factor Analysis. *Anal Chem* **51**, 1236–1241 (1979)
- Gaspari, M., Verhoeckx, K.C.M., Verheij, E.R., van der Greef, J.: Integration of Two-Dimensional LC-MS with Multivariate Statistics for Comparative Analysis of Proteomic Samples. *Anal Chem* **78**, 2286–2296 (2006)
- America, A.H.P., Cordewener, J.H.G., van Geffen, M.H.A., Lommen, A., Vissers, J.P.C., Bino, R.J., Hall, R.D.: Alignment and Statistical Difference Analysis of Complex Peptide Data Sets Generated by Multidimensional LC-MS. *Proteomics* **6**, 641–653 (2006)
- Gottfried, J., Sjogren, M., Holmberg, B., Rosengren, L., Davidsson, P., Blennow, K.: Proteomics for Drug Target Discovery. *Chemom Intell Lab Syst J* **73**, 47–53 (2004)
- Berman, E.S.F., Kulp, K.S., Knize, M.G., Wu, L., Nelson, E.J., Nelson, D.O., Wu, K.J.: Distinguishing Monosaccharide Stereo- and Structural Isomers with TOF-SIMS and Multivariate Statistical Analysis. *Anal Chem* **78**, 6497–6503 (2006)
- Fangmark, I., Jansson, A., Nilsson, B.: Determination of Linkage Position and Anomeric Configuration in Glucose-Containing Disaccharide Alditols by Multivariate Analysis of Data from Mass Spectrometry. *Anal Chem* **71**, 1105–1110 (1999)
- Arungundram, S., Al-Mafraji, K., Asong, J., Leach, F.E., Amster, I.J., Venot, A., Turnbull, J.E., Boons, G.-J.: Modular Synthesis of Heparan Sulfate Oligosaccharides for Structure-Activity Relationship Studies. *J Am Chem Soc* **131**, 17394–17405 (2009)
- Heck, A.J.R., de Koning, L.J., Pinkse, F.A., Nibbering, N.M.M.: Mass-Specific Selection of Ions in Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry. Unintentional Off-Resonance Cyclotron Excitation of Selected Ions. *Rapid Commun Mass Spectrom* **5**, 406–414 (1991)
- 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. *J Prot Res* **7**, 1650–1659 (2008)
- Domon, B., Costello, C.E.: A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconj J* **5**, 397–409 (1988)