Analyst

Broadband Ion Mobility Deconvolution for Rapid Analysis of Complex Mixtures

Journal:	*Analyst*
Manuscript ID	AN-ART-01-2018-000193.R1
Article Type:	Paper
Date Submitted by the Author:	23-Mar-2018
Complete List of Authors:	Pettit, Michael; Baylor University, Chemistry and Biochemistry Brantley, Matthew; Baylor University, Chemistry and Biochemistry Donnarumma, Fabrizio; Louisiana State University, Chemistry Murray, Kermit; Louisiana State University, Department of Chemistry Solouki, Touradj; Baylor University, Department of Chemistry & Biochemistry

Broadband Ion Mobility Deconvolution for Rapid Analysis of Complex Mixtures

Michael E. Pettit[1], Matthew R. Brantley[1], Fabrizio Donnarumma[2], Kermit K. Murray[2], and Touradj Solouki[1*]

[1]Department of Chemistry and Biochemistry, Baylor University, Waco, TX 76798, USA

[2]Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803, USA

ABSTRACT (Word Count = 205): High resolving power ion mobility (IM) allows for accurate characterization of complex mixtures in high-throughput IM mass spectrometry (IM-MS) experiments. We previously demonstrated that pure component IM-MS data can be extracted from IM unresolved post-IM/collision-induced dissociation (CID) MS data using an Automated Ion Mobility Deconvolution (AIMD) software [*J. Am. Soc. Mass Spectrom.,* 2014, **25**, 1810-1819]. In our previous reports, we utilized a quadrupole ion filter for *m/z*-isolation of IM unresolved monoisotopic species prior to post-IM/CID MS. Here, we utilize a broadband IM-MS deconvolution strategy to remove the *m/z*-isolation requirement for successful deconvolution of IM unresolved peaks. Broadband data collection has throughput and multiplexing advantages; hence, elimination of the ion isolation step reduces experimental run times and thus expands the applicability of AIMD to high-throughput bottom-up proteomics. We demonstrate broadband IM-MS deconvolution of two separate and unrelated pairs of IM unresolved isomers (*viz.*, a pair of isomeric hexapeptides and a pair of isomeric trisaccharides) in a simulated complex mixture. Moreover, we show that broadband IM-MS deconvolution improves high-throughput bottom-up characterization of a proteolytic digest of rat brain tissue. To our knowledge, this manuscript is the first to report successful deconvolution of pure component IM and MS data from an IM-assisted data-independent analysis (DIA) or HDMS^E dataset.

Keywords: broadband deconvolution, co-elution, ion mobility, mass spectrometry (MS), resolving power,

*Please address reprint requests to: Professor Touradj Solouki, Department of Chemistry and Biochemistry, Baylor University Sciences Building, One Bear Place #97348, Waco, TX 76798 Telephone number: 254-710-2678, E-mail: Touradj_Solouki@baylor.edu

Introduction

Mass spectrometry (MS) is utilized in a wide variety of omics fields (*e.g.*, lipidomics,[1] petroleomics,[2] proteomics,[3] metabolomics,[4] and glycomics[5]) to study complex sample mixtures. Frequently, to distinguish between isobaric species present in complex samples and improve data quality, high-resolution MS instruments such as Fourier transform ion cyclotron resonance (FT-ICR) MS and Orbitrap MS are used.[6] Although specialized MS analysis techniques, for example, ion-molecule reactions,[7] can be used to distinguish isomers, differentiating them is difficult to address by MS alone. Hence, high-throughput MS analysis of complex mixtures containing isomeric and isobaric species often requires some form of analyte-separation prior to MS analysis.

High performance liquid chromatography (HPLC)[8] and gas chromatography (GC)[9] can be readily interfaced with mass spectrometers for multidimensional analyses; however, these chromatographic techniques are occasionally insufficient to resolve isobaric or isomeric species and thus tandem use of additional separation systems is an attractive alternative. For example, ion mobility (IM) has seen recent growth for complex sample separations[10-31] and has been combined with other chromatographic techniques for multidimensional separations such as LC-IM-MS[25,32] and GC-IM-MS.[14] In IM experiments, ions are separated based on their collisional cross sections (CCSs) as they move through an (typically) inert gas environment.[33,34]

Despite the recent improvements in MS instrumentation[6,35] and separation technologies,[8,9,36] co-elution continues to impede adequate characterization of complex mixtures in areas such as petroleomics[37] and biological sample analyses.[38-40] Previously, methods such as energy-resolved IM-MS[22] and species-specific fragment identification[24] have been used to address IM convolution issues. In 2012, we utilized a chemometric data analysis technique that allows extraction of pure component IM profiles and their associated collision-induced dissociation (CID) mass spectra from convoluted post-IM/CID MS data.[41] In 2014, we automated the IM-MS deconvolution approach by developing the Automated Ion Mobility Deconvolution (AIMD) software[11] and subsequently used AIMD to investigate several IM unresolved systems.[27,42-45] Although the acronym "AIMD" is also utilized to mean "ab initio molecular dynamics",[46] here it strictly refers to "Automated Ion Mobility Deconvolution". To date, pure component IM-MS deconvolution techniques have not been applied to MS-based proteomic workflows.

MS-based proteomic workflows are typically classified as either top-down[47-49] analysis of intact proteins or bottom-up[50-52] characterization of proteolytic digests. Top-down approaches can yield in-depth primary structure information, including site-specific mutations and post-translational modifications[50] that are often lost during chemical and enzymatic proteolysis.[53] However, because peptides from proteolytic digests are typically easier to solubilize for LC separation and more readily analyzed with MS than intact large proteins, bottom-up approaches are more commonly utilized in high-throughput MS-based proteomics.[51]

Bottom-up proteomic workflows typically utilize HPLC for peptide separation prior to introduction into a mass spectrometer. Gas-phase ion fragmentations, using CID[54-56] or other ion activation methods,[57] such as electron capture dissociation (ECD)[58] or electron transfer dissociation (ETD),[59] are employed to generate product ions that are subsequently correlated to their corresponding precursor ions. The resulting list of precursor-product ion assignments can be searched manually or automatically queried[60-63] through protein databases to identify amino acid sequences for elucidation of protein primary structures.

Peptide fragmentations and subsequent precursor-product ion assignments are typically achieved using an either data-dependent acquisition (DDA) or data-independent acquisition (DIA) strategy. DDA methods[64] utilize sequential ion isolation for fragmenting specific ions of interest. In contrast, DIA methods[65-68] do not require sequential isolations of single ionic species and can be implemented by alternating the CID activation energies between low and high energy regimes for all ions. The former produces minimal ion fragmentation for generation of precursor ion data whereas the latter yields increased ion fragmentation for generation of product ion data. Precursor-product ion assignments are more difficult to generate in DIA compared to DDA and require time alignment[69-71] by accurate mass retention time (AMRT) correlations[72] or by ion mobility drift-time alignments[73] as is done in HDMS^E data processing. However, DIA offers multiplexing (similar to Fellgett's advantage[74]) and throughput (similar to Jacquinot's advantage[75]) benefits for high-throughput proteomics. Alternating the energy regime in a DIA approach for simultaneous fragmentation of all ions is markedly faster than scanning through individual precursor ions as is done in DDA. In addition, elimination of the ion isolation step in DIA yields higher sensitivity by avoiding unnecessary ion losses and improving efficient use of instrumental duty cycle.[66]

The benefits in throughput and sensitivity for DIA[66,76] are often outweighed by the difficulty of establishing reliable precursor-product ion assignments using AMRT or drift-time alignment. Specifically, IM co-elution limits accurate assignment of drift times in complex mixtures.[11,22,24,27,31,41,43,77] Although custom methods and instrument modifications have resulted in increased IM resolving powers[13,16,19,23,26,28,35,78-82] (defined as peak arrival time (AT)/ΔAT$_{50\%}$ or CCS/ΔCCS$_{50\%}$),[83-85] these improvements have not been widely applied to proteomic workflows. For example, several groups have proposed experimental strategies to probe for the presence of unresolved species in IM-MS.[15,22,24,30] Chemometric data processing has also been applied for untargeted precursor-product ion assignment; however, this approach does not allow IM-MS deconvolution.[20] These strategies do not allow for extraction of pure component IM profiles and/or MS data from co-eluting species,[15,20,22,24,30] and thus have not been used to enhance precursor-product ion assignment in current proteomic workflows.

We previously demonstrated that deconvolution of IM unresolved post-IM/CID MS data using AIMD allowed for rapid generation of both pure component IM profiles and CID mass spectra of

co-eluting components.[11] These deconvoluted IM profiles and associated CID mass spectra can potentially be used to establish precursor-product ion assignments from unresolved IM-MS data in IM-enhanced DIA (*e.g.*, HDMS^E) bottom-up proteomics experiments. However, to date, all reported AIMD results have used an *m/z*-isolation step prior to post-IM/CID analysis; this single *m/z*-isolation step precludes the use of AIMD in MS-based proteomics. Our IM-MS deconvolution workflows are not compatible with IM-enhanced DDA (*e.g.*, HD-DDA) approaches that utilize pre-IM/CID.[86] On the other hand, HDMS^E approaches employ broadband post-IM/CID.[73] Consequently, broadband IM-MS deconvolution is distinct from previous IM deconvolution approaches as it provides an exclusive avenue for combining AIMD with MS-based bottom-up proteomics.

In this manuscript, we report results from a modified AIMD approach that avoids *m/z*-isolation and yields broadband ion fragmentation. Moreover, we demonstrate the suitability of this new and rapid broadband approach for improved HDMS^E data processing capabilities. Prior to this report, precursor ions that were unresolved in the LC and IM dimensions could not be correlated to their respective product ions using currently available HDMS^E data processing algorithms. Here, we report the use of broadband IM-MS deconvolution for generating precursor-product ion correlations from UPLC-IM-MS unresolved HDMS^E data.

In direct infusion IM-MS, HDMS^E can be emulated with separate low and high collision-energy experiments. As examples of broadband ion fragmentation and IM peak deconvolution, we show extracted pure component IM profiles and mass spectra for two IM unresolved binary isomer mixtures: a hexapeptide mixture containing MGRYGF and FRMYGG peptide isomers and a trisaccharide mixture containing D-(+)-raffinose and D-(+)-isomaltotriose trisaccharide isomers. To confirm the validity of AIMD analyses in broadband mode, we compare deconvolution results to validation sets comprising IM-MS data of corresponding single component solutions. Additionally, we evaluate the efficacy of broadband IM-MS deconvolution in the presence of high-intensity background ions by deconvoluting IM-MS data of a simulated complex mixture that contains both abovementioned binary isomer mixtures and polypropylene glycol species. Furthermore, we apply broadband IM-MS deconvolution to HDMS^E analysis of a proteolytic digest of rat brain tissue.

Experimental

Preparation of Simulated Complex Mixtures

Hexapeptide isomers, MGRYGF and FRMYGG, were synthesized by Peptide 2.0, Inc. (Chantilly, VA, USA). D-(+)-raffinose, D-(+)-isomaltotriose, and lithium chloride (LiCl) were purchased from Sigma-Aldrich (St. Louis, MO, USA). A mixture of polypropylene glycol (PPG) 425 and PPG 1000 was purchased from Agilent Technologies (Santa Clara, CA, USA). Optima grade methanol and acetic acid were purchased from Fisher-Scientific (Waltham, MA, USA). Water, with an overall ionic concentration of < 0.1 ppb and a resistivity of ~18.2 MΩ·cm at 25

°C, was purified in-house using a Direct-Q 3 UV water purification system (EMD Millipore Corporation, Billerica, MA, USA). All commercial samples and chemical solvents were used as received and without further purification.

All analytes used in sample solutions were initially prepared as stock solutions in a methanol:water solvent (1:1, v/v) with 0.1% acetic acid. Analyte concentrations for single component samples (*i.e.*, pure isomer solutions used as validation sets) were optimized such that comparable ion counts were measured for all monitored precursor ions. Pure MGRYGF and pure FRMYGG hexapeptide isomers were individually prepared as ~1.7 μM solutions; pure D-(+)-raffinose and pure D-(+)-isomaltotriose trisaccharide isomers were each prepared as ~3.3 μM solutions and spiked with 15:1 excess LiCl to ensure the formation of Li-adducts. The binary hexapeptide mixture contained ~1.7 μM each of MGRYGF and FRMYGG isomers (~3.4 μM total peptide concentration), and the binary trisaccharide mixture contained ~3.3 μM each of D-(+)-raffinose and D-(+)-isomaltotriose isomers (~6.6 μM total trisaccharide concentration) with an excess (~15:1 per isomer) of LiCl. The simulated complex sample solution contained ~1.7 μM MGRYGF, ~1.7 μM FRMYGG, ~3.3 μM D-(+)-raffinose, ~3.3 μM D-(+)-isomaltotriose, ~0.2 mM LiCl and ~0.3 μM PPG. It was experimentally determined (data not shown) that the formation of Li-adducts in the complex mixture required increasing the LiCl to trisaccharide isomer ratio from ~15:1 (as used for the pure trisaccharide solutions) to ~60:1 per isomer, presumably due to the formation of PPG and peptide Li-adducts.

Instrumentation

Direct Infusion IM-MS Data Acquisition

All experiments were performed in positive-ion mode electrospray ionization[87] (ESI+) using a Synapt G2-S HDMS system (Waters, Milford, MA, USA) set to resolution mode (*i.e.*, V mode). Argon (Ar) was used in the trap and transfer cells as buffer and collision gases, respectively. Helium (He) gas was used for collisional cooling prior to IM, and nitrogen (N_2) was used as the IM drift gas. "Generic" ESI conditions and default instrument parameters were used for all experiments (*viz*., Table S1; reported pressures are direct measurement readouts without correcting for geometry[88] or sensitivity[89] factors) except the transfer collision-energies (*i.e.*, electric potential differences applied between the exit of the IM cell and entrance of the transfer cell) used during the post-IM CID experiments. Sample introduction was performed *via* direct infusion using a Standard Infusion 11 Plus Syringe Pump (Harvard Apparatus, Holliston, MA, USA) at a sample flow rate of 0.5 μL/min. The pre-IM quadrupole was operated in transmission mode (rf-only, *i.e.*, no *m/z*-isolation in the quadrupole assembly prior to IM). Broadband data were collected for all detected ions in the *m/z* range of 50-1500. After IM separation, ions were fragmented *via* CID in the transfer cell (*i.e.*, post-IM/CID). The Synapt G2-S is configured such that HDMS^E data collection only operates in conjunction with LC separation and is not compatible with direct infusion sample introduction. Thus, HDMS^E was emulated for direct

infusion IM-MS experiments by acquiring separate datasets using low (*i.e.*, minimal post-IM/CID) and high transfer collision-energy (*i.e.*, increased post-IM/CID) experiments. Previously optimized[27] transfer collision-energies of 26 V and 45 V were used to induce post-IM/CID for the binary hexapeptide mixture and binary trisaccharide mixture, respectively. Data for high-energy post-IM/CID of the simulated complex sample solution were acquired at a 35 V transfer collision-energy. The total ion accumulation time for all experiments was 13.8 ms (corresponding to a maximum IM drift time of 13.8 ms). Data acquisition times for direct infusion IM-MS experiments were 1 minute (Figures 1-7 and Table 1) or 2 seconds (Table 1). All direct infusion IM-MS experiments reported here were performed in triplicate to confirm reproducibility of broadband deconvolution using AIMD.

UPLC-HDMS^E of Rat Brain Tissue Samples

UPLC-HDMS^E data acquisition was used to collect, in parallel, broadband UPLC-IM-MS and UPLC-IM-MS/MS data of a rat brain tissue digest. Experimental details for biological sample preparation,[90] UPLC-HDMS^E data acquisition, data processing, and data searching against a UniprotKB/Swiss-Prot database[91] are provided in the Supplementary Information. Briefly, a rat brain tissue digest was measured using both (i) 60-min. and (ii) 15-min. UPLC gradients of 5 to ~45% mobile phase B (acetonitrile with 0.1% formic acid).

Data Analysis

Raw IM-MS data were analyzed using MassLynx (Ver. 4.1; Waters). Unresolved high collision-energy IM-MS data (used for deconvolution) were the average of a 1 min. acquisition time (or 2 s for comparisons presented in Table 1) which provided signal-to-noise ratios > 3 for all reported ions in the AIMD deconvoluted spectra; 30 s acquisition times were adequate for generating validation datasets for all pure component samples. Data preprocessing and AIMD analysis for IM-MS data deconvolutions were performed as previously described[11] with the exception of data-notching for isolation of individual IM AT regions. The data-notching windows were centered on the mobiligraphic apexes (from low collision-energy experiments) and set to 200% of the IM peak widths at base to accommodate for the expected IM shifts[20,44] between low and high collision-energy post-IM/CID experiments; peak widths were approximated from low collision-energy experiments as four times the standard deviation (4σ) of the Gaussian-like IM drift distribution of interest. Analysis of a notched dataset improved the efficiency of AIMD by reducing data sizes of both the *m/z* and mobility axes, thus reducing data matrix build times.[11] For each IM deconvolution using AIMD, the number of components (*i.e.*, presumed convoluted peaks) and offset values[11] were manually optimized to minimize negative values in the deconvoluted data[41] and to maximize the degree of spectral matching[27] between the deconvoluted and pure post-IM/CID mass spectra. Although the numbers of convolved components were manually optimized for the deconvolutions reported herein, the Malinowski's factor indicator function[92] (previously integrated with AIMD[11]) can be used for assigning the number of possible chemical "components" of post-IM/CID MS data matrices; the maximum

number of possible components is limited to the number of IM bins input to AIMD. As previously described, the spectral matching *R*-values can range from 0 for no correlation between the pure and deconvoluted mass spectra to 1 for perfect match between the pure and deconvoluted mass spectra.[27]

Deconvolution of the isomers' IM profiles and their corresponding post-IM/CID mass spectra was classified as successful if: (1) deconvoluted data had no negative values more than 15% of the corresponding base peak intensity, (2) the deconvoluted IM AT peak maximum was within 1 IM bin (69 μs) of the corresponding pure sample's IM AT peak maximum when individually measured under identical experimental settings, and (3) *R*-values for deconvoluted isomers were 0.70 or higher. For the simulated complex mixture, experimental *R*-values and IM ATs (IM peak maxima) for each deconvoluted isomer are reported in Table 1. Please note that AIMD successfully regenerates the IM profile and post-IM/CID spectrum of single component IM peaks that are solved to a single component. As expected, AIMD yields non-real IM profiles and mass spectra when single component IM peaks are deconvoluted to two or more components (see Figure S2).

In post-IM/CID, the precursor ion is fragmented after traversing the IM cell, thus the high collision-energy IM profile of a precursor ion cannot be directly extracted. However, drift time alignment between precursor-product ion pairs can be performed in post-IM/CID experiments. Therefore, the selected IM (SIM) profile of a fragmented precursor ion species can be indirectly extracted from post-IM/CID experiments by summing the IM data of corresponding post-IM/CID product ions. For example, accurate SIM profiles for each pure isomer were generated by summing the IM data of the 10 most abundant product ions (rather than using a single ion) from the corresponding pure post-IM/CID MS data.

Please note that AIMD does not utilize curve fitting algorithms to generate peak shapes of deconvoluted IM profiles.[11] As previously described,[42] the 69 μs sampling interval used for collecting IM data is insufficient for accurately assigning IM ATs (*i.e.*, peak centroids). Therefore, cubic spline interpolation was used on AIMD output (deconvoluted IM profiles) and raw data (pure SIM profiles) for more accurate comparisons of deconvoluted and pure IM ATs.

The 60-minute and 15-minute UPLC-HDMS^E datasets were processed and searched using ProteinLynx Global Server (PLGS Ver. 2.5.2; Waters). Because pure stock solutions of rat brain tissue digests do not exist for use as validation sets, the 60-minute UPLC-HDMS^E run was used to generate a well resolved three-dimensional dataset. IM-MS unresolved ions were found by cross-referencing PLGS results from the 60-minute UPLC-HDMS^E validation dataset with the 15-minute UPLC-HDMS^E dataset. We hypothesized that peptides detected by PLGS from the 60-minute dataset, but absent from the 15-minute dataset, could correspond to precursor ions that were IM-MS unresolved. Our experimental findings confirmed our hypothesis, and we were able to identify and deconvolute IM profiles and post-IM/CID mass spectra for two IM-MS

unresolved peptides present in the 15-minute UPLC-HDMS^E run. Deconvolution of IM-MS unresolved species in the 15-minute dataset was validated using results from PLGS processing of the 60-minute dataset. Please note that database searching of IM-MS deconvoluted data, using our current proteomics data processing workflow, would not be impacted or hindered by the presence of negative values in AIMD output. The Apex3D algorithm employed by PLGS for precursor and product ion detections (during HDMS^E data processing) imposes minimum intensity thresholds of >1 count. Therefore, only positive values are considered by Apex3D when determining the LC, IM, and MS peak properties of ions. In cases where negative values could adversely impact proteomics data processing and databank searching, non-negativity constraints could be utilized.[93]

Results & discussion

To demonstrate the advantages of broadband IM-MS deconvolution, we performed three sets of experiments. Firstly, we performed broadband IM-MS deconvolutions at optimized collision-energies[27] for two separate binary mixtures of IM unresolved compounds: (i) MGRYGF and FRMYGG hexapeptide isomers and (ii) D-(+)-raffinose and D-(+)-isomaltotriose trisaccharide isomers. Secondly, we performed broadband IM-MS deconvolution for a simulated complex mixture containing the abovementioned peptide and sugar isomer mixtures in the presence of high-intensity background PPG species. In two separate experiments, IM-MS deconvolutions of this simulated complex mixture (containing hexapeptide and trisaccharide isomers as well as the PPG background ions) were performed using data acquisition times of (a) 1 minute and (b) 2 seconds. Isomerically pure post-IM/CID MS data (*i.e.*, single component data) of each species were used to validate all direct infusion broadband IM-MS deconvolution results presented herein. Thirdly, we utilized IM-MS broadband deconvolution to demonstrate its advantage for analyzing high-throughput bottom-up HDMS^E proteomics data from a rat brain tissue sample.

Broadband IM-MS Deconvolution of Binary Isomer Mixtures

As an initial demonstration of broadband IM-MS deconvolution, low and high collision-energy IM-MS data were collected for a binary hexapeptide isomer mixture. The low collision-energy datasets contained IM-MS information for the intact precursor ions. The high collision-energy datasets contained product ion information and were used for IM-MS deconvolution. Broadband IM-MS deconvolution was validated by comparing each isomer's deconvoluted IM-MS data to its isomerically pure IM-MS data according to the three criteria outlined in the Data Analysis section.

At a low collision-energy of 4 V, both pure $[MGRYGF + 2H]^{2+}$ and pure $[FRMYGG + 2H]^{2+}$ had an IM AT of 2.56 ms (Figure S1a). Based on alkali metal adduct formation under identical experimental settings, it was apparent that FRMYGG had more intense sodium and potassium adduct peaks than MGRYGF. Additionally, ion fragmentation was observed for the hexapeptide isomers despite using only a 4 V transfer collision-energy. Namely, the $[y_5]^+$ fragment of

MGRYGF and the $[b_2]^+$, $[b_3]^+$, $[b_4]^+$, and $[b_5]^+$ fragments of FRMYGG were unintentionally generated prior to the IM cell *via* in-source CID and/or other pre-IM/CID and metastable decay events.[94] Thus, in addition to having a mobiligraphic apex at 2.56 ms, the mobiligrams of pure $[MGRYGF + 2H]^{2+}$ (Figure S1a) and pure $[FRMYGG + 2H]^{2+}$ (Figure S1c) showed several other IM peaks (at later ATs) corresponding to singly-charged fragment ions from pre-IM dissociation events. For both hexapeptide isomers, detailed discussions of unintentional adduct formation and pre-IM dissociation fragments, including their IM ATs and fragment identities, are provided in the Supplementary Information.

Figure 1 shows the convoluted broadband IM-MS data of the binary hexapeptide mixture (*i.e.*, containing isomers MGRYGF and FRMYGG) prior to broadband deconvolution. Representative low and high collision-energy mobiligrams of the binary hexapeptide mixture are shown in Figures 1a and 1c. The highlighted regions in Figures 1a and 1c denote the data-notching windows from IM AT of 2.08 ms to 3.18 ms and correspond to the regions of interest containing IM unresolved hexapeptide isomers. The extracted mass spectrum (XMS) corresponding to the highlighted IM region in Figure 1a (or Figure 1c) is shown in Figure 1b (or Figure 1d).

The mobiligraphic apex at 2.56 ms in Figure 1a corresponds to the precursor ion species at *m/z* 365 in Figure 1b. Additional IM peak features shown in Figure 1a, for example the IM peaks at 3.05 ms, 3.88 ms, 4.50 ms, 4.85 ms, 5.47 ms, 6.23 ms, and 6.86 ms, are the result of in-source fragmentation and other pre-IM/CID and metastable decay from either of the two isomeric hexapeptides (further discussions on fragment ion identities are available in Supplementary Information).

In Figure 1b, the base peak at *m/z* 365 corresponds to a combination of the doubly-charged hexapeptide precursor ions $[MGRYGF + 2H]^{2+}$ and $[FRMYGG + 2H]^{2+}$. Because the success of AIMD depends on the presence of detectable differences in post-IM/CID mass spectra as a function of IM bin number,[27] low collision-energy IM-MS data (*e.g.*, as for the binary hexapeptide mixture) is not always suitable for IM peak deconvolution. However, as discussed below, high collision-energy IM-MS experiments can provide additional fragment ions for IM peak deconvolution.

Because ion fragmentation from post-IM/CID occurs after the ions pass through the IM drift tube, the high collision-energy mobiligram of the binary hexapeptide mixture at 26 V in Figure 1c has a similar shape to the low collision-energy mobiligram at 4 V in Figure 1a. However, as expected, there is a small IM shift of ~0.14 ms in Figure 1d to earlier ATs due to higher ion acceleration used for post-IM/CID.[20,44]

The unresolved IM peaks and mass spectrum in Figures 1c and 1d were deconvoluted using AIMD, and the resulting IM profiles and mass spectra are shown in **Figures 2a** and **2b-c,** respectively. Deconvoluted mobiligrams in Figure 2a correspond to $[MGRYGF + 2H]^{2+}$ (orange) and $[FRMYGG + 2H]^{2+}$ (purple) and have IM ATs of 2.40 ms and 2.49 ms, respectively. For

comparison, isomerically pure IM profiles were individually measured using the same 26 V collision-energy and their SIM profiles are shown in Figure 2d (with IM ATs of 2.43 ms for $[MGRYGF + 2H]^{2+}$ and 2.45 ms for $[FRMYGG + 2H]^{2+}$). The low intensity peak at ~2.91 ms in Figure 2d is due to pre-IM dissociation products of $[FRMYGG + 2H]^{2+}$ (*viz.*, $[b_2 - NH_3]^+$ and $[b_2]^+$ fragments labelled in Figure 2f) and should not be misinterpreted as an additional isomer (see Figure S1). Centroids of deconvoluted mobiligrams in Figure 2a match their isomerically pure counterparts (pure mobiligrams' centroids) in Figure 2d within 1 IM bin.

The pure hexapeptide isomer solutions contained ~1.7 μM of MGRYGF or FRMYGG, and the binary hexapeptide mixture contained ~1.7 μM of each peptide. The IM peak ratio of $[FMRYGG + 2H]^{2+}$ to $[MGRYGF + 2H]^{2+}$ in Figure 2a was 0.75 ± 0.46 whereas the corresponding ratio in Figure 2d was 0.55 ± 0.23 (value reported as 95% confidence interval for n = 3). This small peak ratio variation is statistically insignificant at the 95% confidence level (for $n_1 = n_2 = 3$, using Student's *t*-test, Case 2).

The deconvoluted post-IM/CID mass spectra of the isomeric peptides (Figures 2b and 2c) were quantitatively compared with the individual post-IM/CID mass spectra (Figures 2e and 2f) using mass spectral matching by calculating *R*-values, where $R = 1$ represents a perfect match and $R = 0$ indicates no spectral correlation.[27] The deconvoluted mass spectra matched the corresponding individual component mass spectra with *R*-values of 0.91 and 0.78 for MGRYGF and FRMYGG, respectively. *R*-values are calculated using deviations in relative intensities of peaks that are present in deconvoluted and/or pure mass spectra; misassigned peaks can lower the calculated *R*-values.[27] For instance, erroneous assignment of higher relative intensity to *m/z* 166 by AIMD,[11] denoted by a black asterisk in deconvoluted mass spectrum of $[FRMYGG +2H]^{2+}$ in Figure 2c, contributes to the calculated *R*-value of 0.78.

Similar to the binary mixture of peptide isomers, broadband IM-MS deconvolution was used to extract IM profiles and mass spectra from a binary trisaccharide isomer mixture (*i.e.*, containing $[D\text{-}(+)\text{-raffinose} + Li]^+$ and $[D\text{-}(+)\text{-isomaltotriose} + Li]^+$). When separately measured at a 4 V collision-energy, the trisaccharide ions had distinct IM ATs of 4.43 ms and 4.57 ms, respectively (data not shown). However, when mixed, these ions yielded a single, convoluted IM peak. For example, **Figure 3** shows the convoluted IM profiles (3a and 3c) and mass spectra (3b and 3d) of the trisaccharide mixture. The highlighted regions in Figures 3a and 3c denote the data-notching windows from IM AT of 3.74 ms to 5.12 ms; this region of interest contained the IM unresolved peak for the trisaccharide isomers.

The convoluted broadband mobiligram of the binary trisaccharide isomer mixture in Figure 3a has a mobiligraphic apex at an IM AT of 4.50 ms corresponding to the trisaccharide isomer precursor ion species at *m/z* 511 in Figure 3b. The XMS in Figure 3b was constructed by summing mass spectra of ions with mobility times between 3.74 ms to 5.12 ms in Figure 3a.

The apex of the mobiligram shown in Figure 3c is at 4.29 ms and corresponds to the $[M_T + Li]^+$ precursor ions. As expected, a ~0.21 ms shift to a shorter AT was observed for $[M_T + Li]^+$ after increasing the transfer collision-energy from 4 V (Figure 3a) to 45 V (Figure 3c).[20,44] The high collision-energy XMS for the binary trisaccharide isomer mixture in Figure 3d was constructed by summing mass spectra of ions with mobility times between 3.74 ms to 5.12 ms in Figure 3c; this average mass spectrum contains post-IM/CID fragment ions generated from [D-(+)-raffinose + Li]$^+$ and/or [D-(+)-isomaltotriose + Li]$^+$.

The unresolved IM peak at AT = 4.29 ms (Figure 3c) and mass spectrum (Figure 3d) were deconvoluted using AIMD, and the deconvoluted IM profiles and mass spectra are shown in **Figures 4a** and **4b-c**, respectively. The IM AT region 3.74 ms to 5.12 ms (*viz*., highlighted regions in Figures 3a and 3c) is expanded in Figures 4a (after performing AIMD) and 4d (for individually ran pure samples). The deconvoluted IM profiles in Figure 4a correspond to [D-(+)-raffinose + Li]$^+$ (blue) and [D-(+)-isomaltotriose + Li]$^+$ (red) and have IM ATs of 4.26 ms and 4.40 ms, respectively. To validate the deconvolution results, isomerically pure IM profiles were individually measured using the same 45 V collision-energy and are displayed in Figure 4d as SIM profiles. IM deconvolution results for the binary sugar mixture (Figure 4a) are in agreement with results from the analysis of their isomerically pure counterparts (Figure 4d), and deconvoluted IM peak centroids for both isomers match their pure counterparts' values to within 1 IM bin.

Each pure trisaccharide sample was a ~3.3 μM solution of either isomer and concentrations of each sugar in the binary trisaccharide mixture were ~3.3 μM. The IM peak ratio of [D-(+)-isomaltotriose + Li]$^+$ to [D-(+)-raffinose + Li]$^+$ was 0.90 ± 0.40 (value reported as 95% confidence interval for n = 3) for the deconvoluted data in Figure 4a. When independently measured as pure isomer solutions, the IM peak ratio was 0.78 ± 0.47 (Figure 4d). At the 95% confidence level, this observed decrease in sugar isomer peak ratio was statistically insignificant (Student's *t*-test, Case 2 for $n_1 = n_2 = 3$). Deconvoluted post-IM/CID mass spectra in Figures 4b and 4c matched their isomerically pure counterparts with *R*-values of ~1.00 and 0.90, respectively. Please note that broadband IM-MS deconvolution using AIMD is successful even when unresolved components are not prepared at equimolar concentrations (see Figure S3).

Broadband IM-MS Deconvolution of a Complex Mixture

Broadband AIMD deconvolution was performed on a simulated complex mixture to evaluate the performance of our broadband approach in the presence of high-intensity background ions. The complex mixture contained the two hexapeptide isomers, the two trisaccharide isomers, PPG 425, and PPG 1000. Although only LiCl was added to the simulated complex mixture (to ensure the formation of Li-trisaccharide adducts), various other alkali metal complexes and charge-states of the peptides, sugars, and PPG species were also observed.

The convoluted IM peaks and mass spectra of the simulated complex mixture are shown in **Figure 5**. The highlighted data-notching windows (Figure 5a, central panel) spanning from IM AT of 2.08 ms to 3.18 ms and 3.74 ms to 5.12 ms were utilized for AIMD analysis of $[M_P + 2H]^{2+}$ and $[M_T + Li]^{+}$ isomers, respectively.

The convoluted high collision-energy mobiligram of the complex mixture is shown in Figure 5a. The IM peaks at 2.42 ms and 4.36 ms correspond to isomeric peptide and trisaccharide precursor ions, respectively, and include additional interfering background species. The convoluted XMS shown in Figures 5b and 5c were generated by summing mass spectra of all peptide, sugar and PPG ions with mobility times between 2.08 ms to 3.18 ms (for Figure 5b) and 3.74 ms to 5.12 ms (for Figure 5c). Sixty-five of the 100 most abundant MS peaks corresponding to mobility times 2.08 ms to 3.18 ms were not unique to either one of the two isomeric hexapeptides (these assignments were confirmed by both raw IM and MS data and AIMD deconvolution). Likewise, for the IM AT region of 3.74 ms to 5.12 ms, seventy-six of the 100 MS peaks were not unique to either one of the two trisaccharide isomers. In other words, despite the presence of a large number of unrelated MS peaks within each selected window (*i.e.*, Figures 5b and 5c), AIMD analyses of multiple IM regions corresponding to two sets of co-eluting isomers were successful.

The unresolved peptide data (Figures 5a and 5b) were deconvoluted from the simulated complex mixture without prior knowledge of the total number of convolved components, and the results are shown in **Figure 6**. The 2.08 ms to 3.18 ms region in Figure 5a is expanded in Figures 6a (deconvoluted IM for peptide isomers) and 6d (IM for pure peptide isomers) to show the correlation between deconvoluted and pure component data. The deconvoluted IM profiles in Figure 6a, corresponding to $[MGRYGF + 2H]^{2+}$ (orange solid trace) and $[FRMYGG + 2H]^{2+}$ (purple solid trace), have IM ATs of 2.37 and 2.49 ms, respectively. IM profiles of nine other interfering background species are also shown in Figure 6a as dotted line IM profiles. Representative mass spectra for these background species are provided and further discussed in the Supplementary Information (Figure S4). AIMD does not utilize any curve fitting for generating IM profiles;[11] therefore, the observed Gaussian-like shapes of background IM profiles match the expected physical reality for peak broadening in ion mobility separation[42] and support the validity of broadband deconvolution. Isomerically pure peptide isomers were individually measured using a 35 V collision-energy to validate their deconvolution from the complex mixture, and IM AT values of 2.43 ms and 2.46 ms were observed for pure $[MGRYGF + 2H]^{2+}$ and pure $[FRMYGG + 2H]^{2+}$, respectively. The deconvoluted IM ATs matched their corresponding single components within a single IM bin (drift time differences of less than 69 μs).

Deconvolution of data collected at a 26 V collision-energy yielded an IM AT of 2.40 ms for $[MGRYGF + 2H]^{2+}$ (orange trace in Figure 2a). As expected,[20,44] deconvolution of data collected at a higher collision-energy of 35 V yielded a shorter IM AT of 2.37 ms for $[MGRYGF + 2H]^{2+}$ (orange solid trace in Figure 6a) which corresponded to a difference of less than one IM bin.

These IM AT shifts may not be assigned accurately if IM data are under-sampled. To better identify peak centroids and extract more accurate IM arrival times from the AIMD generated outputs, we utilized cubic spline interpolation for IM peak fitting.[42] Please note that small IM AT shifts (*i.e.*, < 0.01 ms) might go unnoticed when rounding fitted data (*e.g.*, deconvoluted IM ATs for $[FRMYGG + 2H]^{2+}$ at 26 V *vs* 35 V in Figures 2a and 6a, respectively).

The deconvoluted post-IM/CID mass spectrum corresponding to $[MGRYGF + 2H]^{2+}$ is shown in Figure 6b; comparison between this deconvoluted mass spectrum (Figure 6b) and its corresponding pure mass spectrum (Figure 6e) yields a mass spectral matching factor of 0.85. Figure 6c shows the deconvoluted post-IM/CID mass spectrum of $[FRMYGG + 2H]^{2+}$; the mass spectral matching factor between the deconvoluted (Figure 6c) and pure (Figure 6f) mass spectra for $[FRMYGG + 2H]^{2+}$ was 0.75. Spectral matching factors of 0.85 and 0.75 indicate successful deconvolution of both peptides at a 35 V collision-energy.

The unresolved IM peaks and mass spectrum of $[M_T + Li]^+$ (Figures 5a and 5c, respectively) were deconvoluted using AIMD without prior knowledge of the total number of convolved components, and results are shown in **Figure 7**. The IM regions of 3.74 ms to 5.12 ms for deconvoluted and pure isomers of the two trisaccharides are expanded in Figures 7a and 7d. Figure 7a shows the deconvoluted IM profiles for raffinose (blue solid trace) and isomaltotriose (red solid trace) with IM ATs of 4.31 ms and 4.46 ms, respectively. Although negative values can be present in AIMD deconvoluted mobiligrams,[95] we consider IM deconvolution successful if negative values are less than 15% of the base peak intensity (*e.g.*, ~4.20 ms to ~4.25 ms for $[isomaltotriose + Li]^+$ in Figure 7a). IM profiles of ten other interfering background species are also shown in Figure 7a as dotted line IM profiles. Representative mass spectra for these background species are provided and further discussed in the Supplementary Information (Figure S5). Again, the Gaussian-like shapes of background IM profiles support the validity of broadband deconvolution. Moreover, the truncated IM profiles of background species 1, 2, 9, and 10, as generated by AIMD, match the physical reality of these background species not being completely within the notched window and further support the validity of broadband deconvolution.

The deconvoluted ATs (Figure 7a) are in agreement with those of pure components (Figure 7d). The mass spectral matching factors of raffinose (Figure 7b) and isomaltotriose (Figure 7c) were 0.83 and 0.88, respectively. According to the criteria for broadband deconvolution outlined in the Data Analysis section, all four components were successfully extracted *via* broadband deconvolution of the complex mixture's IM-MS data.

A 2 second data acquisition time, rather than 1 minute, was also used with a 35 V collision-energy for AIMD analysis of the simulated complex mixture. **Table 1** provides deconvoluted IM ATs and *R*-values for each isomer using a 1 minute and 2 second acquisition. All values reported in Table 1 are from deconvoluted IM profiles and mass spectra that comply with the three criteria for successful broadband IM-MS deconvolution. Values in Table 1 were calculated as 95%

confidence intervals with n = 3. The deconvoluted IM ATs for each data acquisition time were compared using Student's *t*-test (Case 2), and the difference between the two measurements at the 95% confidence level was insignificant. As expected, spectral matching factors decreased at the shorter acquisition time (with exception of raffinose) but remained within the designated limit (of $1 \geq R \geq 0.70$) for successful deconvolution.

Comparison of Broadband and *m/z*-isolated Approaches

A broadband or DIA deconvolution approach (a) eliminates ion isolation losses and (b) reduces the analysis time *via* multiplexed detection, both of which should enhance sensitivity. The magnitude of such enhancements for both (a) and (b) will depend on the width of the ion isolation window: narrow windows can result in both larger ion losses and require more scans to cover the *m/z* range of interest. In other words, sensitivity enhancement of a broadband deconvolution approach will be greater for complex mixtures that require high resolution ion isolation. As explained below, combined ion transmission and multiplexing improvements in broadband deconvolution can (theoretically) yield three orders of magnitude sensitivity enhancement for a binary mixture.

To demonstrate the sensitivity advantage of DIA deconvolution with broadband acquisition compared to DDA deconvolution with *m/z*-isolated acquisitions, the binary hexapeptide mixture was analyzed at concentrations of ~340 nM, ~34 nM, and ~3.4 nM. Deconvolution using AIMD was ten times more sensitive using broadband IM-MS data acquisition (see Figure S6). Using AIMD, broadband acquisition allowed IM-MS deconvolution of both hexapeptide isomers from the ~34 nM solution; however, *m/z*-isolated acquisition of $[M_P + 2H]^{2+}$ at *m/z* 365 (using the same ~34 nM solution) did not allow IM-MS deconvolution. It should be noted that this tenfold sensitivity enhancement for DIA deconvolution is due to elimination of ion losses and does not include benefits of multiplexing. Based on the experimental MS data, compared to *m/z*-isolation, the use of broadband acquisition improved ion transmission of $[M_P + 2H]^{2+}$ at *m/z* 365 by more than 140%; this sensitivity enhancement is analogous to Jacquinot's throughput advantage.[75] Use of *m/z*-isolation in DDA-like experiments limited our data acquisition to a 2 Da window which would require 725 *m/z*-isolation events to cover the entire experimental *m/z* range from 50 to 1500 amu (or 725 minutes, based on a 1-minute MS data acquisition time per experiment). However, we used a single 1-minute broadband experiment to measure the entire *m/z* and drift time ranges; this time saving aspect of DIA deconvolution is analogous to Fellgett's multiplex advantage[74] and scales as narrower *m/z*-isolation windows are employed. In this particular example, the combined sensitivity enhancements of DIA deconvolution would exceed three orders of magnitude (*i.e.*, $10 \times 725 = 7250$) when compared to a comparable DDA experiment of the entire *m/z* range. Because convoluted IM drift time regions are selected *after* data collection, broadband data collection is faster and allows AIMD interrogation of multiple IM regions from the available drift time range using a single IM-MS experiment.

Broadband Deconvolution of UPLC-HDMS^E Proteomics Data

Figure 8 shows IM profiles and associated CID mass spectra of IM-MS unresolved peptides from broadband IM-MS deconvolution of UPLC-HDMS^E proteomics data from a rat brain tissue digest. The broadband IM profile in Figure 8a represents the IM data that corresponded to the UPLC retention time region of 10.5 minutes to 11.5 minutes (Figure S7b) from a relatively short 15-minute UPLC-HDMS^E dataset. The highlighted region in Figure 8a spans the IM AT region of 2.69 ms to 3.80 ms and corresponds to the IM AT data that was submitted to AIMD for deconvolution of the peptides; the averaged XMS corresponding to this IM drift time region is displayed in Figure 8b. The corresponding deconvoluted IM profiles (Figure 8c) and deconvoluted post-IM/CID mass spectra (Figures 8d and 8e) indicate IM-MS deconvolution of the two peptides LIETYFSK and VLSIGDGIAR. The AIMD results in Figures 8c-e were validated by PLGS results from a 60-minute UPLC-HDMS^E dataset where these peptides were UPLC resolved.

From the 60-minute UPLC-HDMS^E dataset, the isobaric peptides $[LIETYFSK]^{2+}$ at *m/z* 500.7725 and $[VLSIGDGIAR]^{2+}$ at *m/z* 500.7941 (Figure S7) were detected and used by PLGS to identify two separate proteins (UniProtKB accession numbers of P60203 and P15999, respectively). Despite being UPLC-IM-MS unresolvable in the 15-minute dataset (Figure 8a-b), $[LIETYFSK]^{2+}$ and $[VLSIGDGIAR]^{2+}$ were identified in the 60-minute dataset by PLGS because they had distinct UPLC retention times of 22.38 minutes and 21.98 minutes, respectively (Figure S7a). In the 15-minute UPLC-HDMS^E dataset, UPLC co-elution prevented PLGS identification of $[VLSIGDGIAR]^{2+}$ (Figure S7b); in other words, a unique retention time, drift time, or *m/z* value was no longer available for $[VLSIGDGIAR]^{2+}$as required for successful identification by PLGS. Because $[VLSIGDGIAR]^{2+}$ was not identified from the 15-minute dataset, the sequence coverage of its associated protein (*i.e.*, P15999) fell from 18.6% in the 60-minute dataset to 15.4% in the 15-minute dataset. However, AIMD was used to extract IM (Figure 8c, blue trace) and post-IM/CID MS data of $[VLSIGDGIAR]^{2+}$ (Figure 8e) from the 15-minute UPLC-HDMS^E dataset. The deconvoluted IM ATs of 3.17 ms for $[LIETYFSK]^{2+}$ and 3.24 ms for $[VLSIGDGIAR]^{2+}$ (red and blue traces in Figure 8c, respectively) matched the PLGS results from the 60-minute dataset (3.167 ms and 3.241 ms, respectively).

PLGS processing of the 15-minute UPLC-HDMS^E dataset yielded 10 total product ion assignments for $[LIETYFSK]^{2+}$; 9 of these 10 product ions were assigned using AIMD deconvolution. Moreover, AIMD allowed 13 additional product ion assignments for $[LIETYFSK]^{2+}$ that were absent from PLGS results. The 22 product ions of $[LIETYFSK]^{2+}$ identified using AIMD had an average mass accuracy error of 9.30 ppm. PLGS did not identify $[VLSIGDGIAR]^{2+}$ in the 15-minute dataset. However, from the 60-minute dataset, PLGS identified 9 product ions for $[VLSIGDGIAR]^{2+}$. In total, broadband deconvolution of 15-minute dataset using AIMD allowed identification of 25 product ions for $[VLSIGDGIAR]^{2+}$ (including all 9 product ions identified by PLGS processing of the 60-minute dataset) with an average mass measurement error of 8.45 ppm.

Agreement between PLGS and AIMD, in terms of equivalent IM peak arrival times and common precursor-product assignments, indicates that AIMD can be used to (i) improve protein identification from PLGS processing of HDMS^E proteomics data and (ii) decrease UPLC separation time requirements for high-throughput bottom-up experiments. Based on the reduced number of peptides detected using a shortened UPLC separation, it is likely that there are many other IM-MS unresolved peptides in the 15-minute dataset that can be deconvoluted.

Conclusions

Previously, we demonstrated that *m/z*-isolation and subsequent post-IM/CID could be used in combination with AIMD to extract pure IM and MS data from IM unresolved species.[11,27,43,45] However, the use of *m/z*-isolation limited the suitability of AIMD for use with high-throughput MS-based proteomics. One drawback of using *m/z*-isolation with AIMD is that only one IM AT region is interrogated per deconvolution experiment. Using broadband IM-MS deconvolution, AIMD analysis of a single convoluted IM-MS dataset allowed extraction of pure IM peaks and post-IM/CID mass spectra for two unrelated pairs of isomers. Additionally, we demonstrated that broadband IM-MS deconvolution using AIMD was successful despite the presence of unrelated high-intensity background ions such as PPG species. By deconvoluting doubly-charged hexapeptides *and* singly-charged trisaccharides using a single dataset, we further demonstrated that AIMD was successful even when ion fragmentation *via* CID involved fragmenting different chemical bond types with varied bond dissociation energies (*i.e.*, amide bonds *versus* glycosidic bonds[96-98]).

In addition to interrogating multiple IM ATs from a single dataset, broadband data collection also improved AIMD throughput by reducing the data acquisition time for deconvolution. A quadrupole device's ion transmission efficiency decreases when operating as a mass filter (as opposed to operating as an rf-only ion guide). Therefore, the use of *m/z*-isolation in combination with AIMD may require long data acquisition times[27] (*i.e.*, 5 to 45 mins) to achieve successful deconvolution. On the other hand, due to its fundamental throughput and multiplex advantages, broadband data collection effectively reduced the amount of time required to achieve appropriate S/N ratios for all ions in the deconvoluted mass spectra. For instance, here, 2 second broadband data acquisition was sufficient for AIMD deconvolution of unresolved post-IM/CID data.

We also demonstrated that broadband IM-MS deconvolution can improve results from an IM-assisted DIA proteomics workflow (*i.e.*, HDME^E). AIMD is attractive for integration with IM-assisted DIA workflows, such as HDMS^E, because it allows for: (i) deconvolution of various IM AT regions (and molecular classes/charge-states) from a broadband post-IM/CID MS dataset, (ii) deconvolution of unresolved IM-MS data collected at UPLC compatible timescales, and (iii) extraction of precursor-product ion assignments from (previously) mischaracterized UPLC-IM-MS regions for improved characterization of real proteomics samples. However, two questions must be addressed before AIMD can be combined with IM-assisted DIA methodologies. The first question is how to identify IM AT regions that contain co-eluting species for subsequent

targeted deconvolution. We recently reported on an IM peak fitting methodology that facilitates detection of IM co-elution by identifying *m/z*-values whose IM peak widths are inappropriately wide.[42] Once IM AT regions that are suspected of containing multiple components have been tagged, the number of IM unresolved species can be ascertained using chemical rank determination techniques.[11,27,31,41] The second question is related to characterization of the AIMD method's dynamic range, as it pertains to concentration differences between co-eluting species and will be addressed in future reports. In summary, the results presented in this manuscript demonstrate the advantages of broadband IM-MS deconvolution for potential high-throughput applications such as IM-assisted DIA methodologies.

Conflict of interest

There are no conflicts to declare.

Acknowledgements:

The authors would like to acknowledge the financial support provided by the National Science Foundation (NSF) (NSF-IDBR Award # 1455668 and CHE-1709526). The authors thank Dr. D. G. Baker (School of Veterinary Medicine, Louisiana State University) for providing rat brain tissue samples.

Figure 1. Unresolved broadband mobiligram and mass spectra of the binary hexapeptide isomer mixture (containing MGRYGF and FRMYGG) using a data acquisition time of 1 minute. **(a)** Total ion mobiligram from AT of 0 ms to 13.8 ms measured using a 4 V collision-energy and **(b)** a representative mass spectrum generated using MS data corresponding to IM AT region 2.08 ms to 3.18 ms in Fig. 1a. **(c)** Total ion mobiligram from AT of 0 ms to 13.8 ms measured using a 26 V CE and **(d)** a representative mass spectrum generated using MS data corresponding to IM AT region 2.08 ms to 3.18 ms in Fig. 1c (please refer to text (and previous report[44]) for details on observed IM shifts between low- and high-collision-energy experiments). Yellow rectangles in **(a)** and **(c)** denote data-notching windows from IM AT 2.08 ms to 3.18 ms. The asterisk with red dashed line in **(d)** shows the precursor ion $[M_P + 2H]^{2+}$ region at *m/z* 365 and indicates a complete signal reduction for $[M_P + 2H]^{2+}$due to post-IM/CID.

Figure 2. Broadband deconvolution of $[M_P + 2H]^{2+}$ using a 26 V collision-energy and a data acquisition time of 1 minute. **(a)** Deconvoluted IM profiles for the two peptide isomers and corresponding post-IM/CID mass spectra for **(b)** $[MGRYGF + 2H]^{2+}$ (orange) and **(c)** $[FRMYGG + 2H]^{2+}$ (purple). **(d)** Pure SIM profiles and pure CID mass spectra for **(e)** $[MGRYGF + 2H]^{2+}$ and **(f)** $[FRMYGG + 2H]^{2+}$ were individually measured using a 26 V collision-energy. Black asterisk above the MS peak at *m/z* 166 in **(c)** denotes an example of partially misassigned MS peak intensity by AIMD.

Figure 3. Unresolved broadband mobiligram and mass spectra of binary trisaccharide isomer mixture (containing [D-(+)-raffinose + Li]$^+$ and [D-(+)-isomaltotriose + Li]$^+$) using a data acquisition time of 1 minute. **(a)** Total ion mobiligram from AT 0 ms to 13.8 ms measured using a 4 V collision-energy and **(b)** a representative mass spectrum generated using MS data corresponding to IM AT region 3.74 ms to 5.12 ms in Fig. 3a. **(c)** Total ion mobiligram from AT of 0 ms to 13.8 ms measured using a 45 V collision-energy and **(d)** a representative high collision-energy CID mass spectrum generated using MS data corresponding to IM AT region 3.74 ms to 5.12 ms in Fig. 3c. Yellow rectangles **(a)** and **(c)** denote data-notching windows from IM AT of 3.74 ms to 5.12 ms.

Figure 4. Broadband deconvolution of [M_T + Li]$^+$ using a 45 V collision-energy and a data acquisition time of 1 minute. **(a)** Deconvoluted IM profiles for the two trisaccharide isomers and their corresponding post-IM/CID mass spectra for **(b)** [D-(+)-raffinose + Li]$^+$ (blue) and **(c)** [D-(+)-isomaltotriose + Li]$^+$ (red). **(d)** Pure SIM profiles and pure CID mass spectra for **(e)** [D-(+)-raffinose + Li]$^+$ and **(f)** [D-(+)-isomaltotriose + Li]$^+$ were individually measured using a 45 V collision-energy.

Figure 5. Unresolved broadband mobiligram and mass spectra of the simulated complex mixture collected using a 35 V collision-energy and a data acquisition time of 1 minute. **(a)** Total ion mobiligram from AT of 0 to 13.8 ms. Yellow rectangles in **(a)** denote data-notching windows from IM ATs of 2.08 ms to 3.18 ms (left) and 3.74 ms to 5.12 ms (right). Representative post-IM/CID mass spectra at 35 V for **(b)** ions with IM ATs ranging from 2.08 ms to 3.18 ms and **(c)** ions with IM ATs ranging from 3.74 ms to 5.12 ms).

Figure 6. Broadband deconvolution of [M_P + 2H]$^{2+}$ from the simulated complex mixture using a 35 V collision-energy and a data acquisition time of 1 minute. **(a)** Deconvoluted IM profiles for the two peptide isomers and corresponding post-IM/CID mass spectra for **(b)** [MGRYGF + 2H]$^{2+}$ (orange solid trace) and **(c)** [FRMYGG + 2H]$^{2+}$ (purple solid trace). **(d)** Pure SIM profiles and pure CID mass spectra for **(e)** [MGRYGF + 2H]$^{2+}$ and **(f)** [FRMYGG + 2H]$^{2+}$ were individually measured using a 35 V collision-energy. In **(a)**, dotted line IM profiles correspond to background species 1 to 9 that were also present in the deconvoluted IM AT region 2.08 ms to 3.18 ms.

Figure 7. Broadband deconvolution of [M_T + Li]$^+$ from the simulated complex mixture using a 35 V collision-energy and a data acquisition time of 1 minute. **(a)** Deconvoluted IM profiles for the two trisaccharide isomers and corresponding post-IM/CID mass spectra for **(b)** [D-(+)-raffinose + Li]$^+$ (blue solid trace) and **(c)** [D-(+)-isomaltotriose + Li]$^+$ (red solid trace). **(d)** Pure SIM profiles and pure CID mass spectra for **(e)** [D-(+)-raffinose + Li]$^+$ and **(f)** [D-(+)-isomaltotriose + Li]$^+$ were individually measured using a 35 V collision-energy. In **(a)**, dotted line IM profiles correspond to background species 1 to 10 that were also present in the deconvoluted IM AT region 3.74 ms to 5.12 ms.

Table 1. Ion mobility arrival times (IM ATs) and spectral matching factors (*R*-values) from broadband IM-MS deconvolution of the simulated complex mixture. Data were collected using a 35 V transfer collision-energy and are compared at 1 minute and 2 second data acquisition times. Values were calculated as 95% confidence intervals (n = 3). Ion mobility arrival times were compared using Student's *t*-test at the 95% confidence level ($n_1 = n_2 = 3$) and no statistical differences were observed.

Figure 8. Broadband deconvolution of peptides from UPLC-HDMS^E proteomics analysis of a rat brain tissue digest. The broadband mobiligram of the proteolytic digest in **(a)** was generated by extracting IM data that corresponded to the convolved UPLC retention time region 10.5 min. to 11.5 min. in the 15-minute UPLC-HDMS^E dataset. The yellow rectangle in **(a)** spans the IM AT region 2.69 ms to 3.80 ms; this IM AT region was deconvoluted using AIMD. The convoluted XMS in **(b)** was generated by integrating the highlighted IM AT region of interest shown in **(a)**. **(c)** Deconvoluted IM profiles of the isobaric peptides $[LIETYFSK]^{2+}$ and $[VLSIGDGIAR]^{2+}$ have IM ATs (3.17 ms and 3.24 ms, respectively) that match the results from PLGS analysis (3.167 ms and 3.241 ms, respectively). Deconvoluted post-IM/CID MS data for $[LIETYFSK]^{2+}$ and $[VLSIGDGIAR]^{2+}$ are shown in **(d)** and **(e)**, respectively.

References

1. M. Wang and X. Han, *Mass Spectrometry in Metabolomics: Methods and Protocols*, 2014, **1198** 203-220.
2. A. G. Marshall and R. P. Rodgers, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 18090-18095.
3. A. Bensimon, A. J. Heck and R. Aebersold, *Annu. Rev. Biochem.*, 2012, **81**, 379-405.
4. R. J. Mishur and S. L. Rea, *Mass Spectrom. Rev.*, 2012, **31**, 70-95.
5. K. Aoki, M. Perlman, J.-M. Lim, R. Cantu, L. Wells and M. Tiemeyer, *J. Biol. Chem.*, 2007, **282**, 9127-9142.
6. F. Xian, C. L. Hendrickson and A. G. Marshall, *Anal. Chem.*, 2012, **84**, 708-719.
7. M. Miladi, B. Zekavat, S. M. Munisamy and T. Solouki, *Int. J. Mass Spectrom.*, 2012, 316, 164-173.
8. L. R. Snyder, J. J. Kirkland and J. W. Dolan, *Introduction to modern liquid chromatography*, John Wiley & Sons, Hoboken, NJ, 2011.
9. R. L. Grob and E. F. Barry, *Modern practice of gas chromatography*, John Wiley & Sons, Hoboken, NJ, 2004.
10. G. R. Asbury and H. H. Hill, *J. Microcolumn Sep.*, 2000, **12**, 172-178.
11. M. Brantley, B. Zekavat, B. Harper, R. Mason and T. Solouki, *J. Am. Soc. Mass Spectrom.*, 2014, **25**, 1810-1819.
12. B. Harper, B. Zekavat, M. Brantley, M. E. Pettit and T. Solouki, Proceedings of the 62nd ASMS Conference on Mass Spectrometry and Allied Topics, Baltimore, MD, 2014.
13. B. Clowers, P. Dwivedi, W. Steiner, H. Hill and B. Bendiak, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 660-669.

14. C. L. Crawford, S. Graf, M. Gonin, K. Fuhrer, X. Zhang and H. H. Hill Jr, *Int. J. Ion Mobility Spectrom.*, 2011, **14**, 23-30.
15. A. J. Creese and H. J. Cooper, *Anal. Chem.*, 2012, **84**, 2597-2601.
16. P. Dwivedi, C. Wu, L. M. Matz, B. H. Clowers, W. F. Siems and H. H. Hill, *Anal. Chem.*, 2006, **78**, 8200-8206.
17. R. G. Ewing, D. A. Atkinson, G. A. Eiceman and G. J. Ewing, *Talanta*, 2001, **54**, 515-529.
18. M. Fasciotti, P. M. Lalli, C. c. F. Klitzke, Y. E. Corilo, M. A. Pudenzi, R. C. Pereira, W. Bastos, R. J. Daroda and M. N. Eberlin, *Energy Fuels*, 2013, **27**, 7277-7286.
19. M. Fasciotti, G. B. Sanvido, V. G. Santos, P. M. Lalli, M. McCullagh, G. F. de Sá, R. J. Daroda, M. G. Peter and M. N. Eberlin, *J. Mass Spectrom.*, 2012, **47**, 1643-1647.
20. S. Garmón-Lobato, B. Abad-García, M. B. Sánchez-Ilárduya, M. Romera-Fernández, L. A. Berrueta, B. Gallo and F. Vicente, *Anal. Chim. Acta*, 2013, **771**, 56-64.
21. E. Hanozin, D. Morsa and E. De Pauw, *Proteomics*, 2015, **15**, 2823-2834.
22. W. Hoffmann, J. Hofmann and K. Pagel, *J. Am. Soc. Mass Spectrom.,* 2014, **25**, 471-479.
23. Y. M. Ibrahim, A. A. Shvartsburg, R. D. Smith and M. E. Belov, *Anal. Chem.*, 2011, **83**, 5617-5623.
24. S. Lee, Z. Li, S. J. Valentine, S. M. Zucker, N. Webber, J. P. Reilly and D. E. Clemmer, *Int. J. Mass Spectrom.*, 2012, **309**, 154-160.
25. X. Liu, M. Plasencia, S. Ragg, S. J. Valentine and D. E. Clemmer, *Briefings Funct. Genomics Proteomics*, 2004, **3**, 177-186.
26. Y. Liu and D. E. Clemmer, *Anal. Chem.*, 1997, **69**, 2504-2509.
27. M. E. Pettit, B. Harper, M. R. Brantley and T. Solouki, *Analyst*, 2015, **140**, 6886-6896.
28. A. A. Shvartsburg and R. D. Smith, *Anal. Chem.*, 2011, **83**, 9159-9166.
29. R. T. Vinopal, J. R. Jadamec, P. deFur, A. L. Demars, S. Jakubielski, C. Green, C. P. Anderson, J. E. Dugas and R. F. DeBono, *Anal. Chim. Acta*, 2002, **457**, 83-95.
30. Y. Xuan, A. J. Creese, J. A. Horner and H. J. Cooper, *Rapid Commun. in Mass Spectrom.*, 2009, **23**, 1963-1969.
31. B. Zekavat, M. Miladi, C. Becker, S. Munisamy and T. Solouki, *J. Am. Soc. Mass Spectrom.*, 2013, **24**, 1355-1365.
32. A. E. Hilderbrand, S. Myung, C. A. S. Barnes and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2003, **14**, 1424-1436.
33. S. N. Lin, G. Griffin, E. Horning and W. Wentworth, *J. Chem. Phys.*, 1974, 60, 4994-4999.
34. J. C. May and J. A. McLean, *Anal. Chem.*, 2015, **87**, 1422-1436.
35. S. Zucker, S. Lee, N. Webber, S. Valentine, J. Reilly and D. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2011, **22**, 1477-1485.
36. L. R. Snyder, *Anal. Chem.*, 2000, **72**, 412 A-420 A.
37. F. A. Fernandez-Lima, C. Becker, A. M. McKenna, R. P. Rodgers, A. G. Marshall and D. H. Russell, *Anal. Chem.*, 2009, **81**, 9941-9947.
38. M. D. Plasencia, D. Isailovic, S. I. Merenbloom, Y. Mechref and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2008, **19**, 1706-1715.
39. R. D. Smith, G. A. Anderson, M. S. Lipton, C. Masselon, L. Paša-Tolic, Y. Shen and H. R. Udseth, *Omics: a journal of integrative biology*, 2002, **6**, 61-90.
40. S. J. Valentine, M. D. Plasencia, X. Liu, M. Krishnan, S. Naylor, H. R. Udseth, R. D. Smith and D. E. Clemmer, *J. Proteome Res.*, 2006, **5**, 2977-2984.

41. B. Zekavat and T. Solouki, *J. Am. Soc. Mass Spectrom.*, 2012, **23**, 1873-1884.
42. M. R. Brantley, M. E. Pettit, B. Harper, B. Brown and T. Solouki, *Anal. Chim. Acta*, 2016, **941**, 49-60.
43. B. Harper, M. Miladi and T. Solouki, *J. Am. Soc. Mass Spectrom.*, 2014, **25**, 1716-1729.
44. B. Harper, E. K. Neumann, S. M. Stow, J. C. May, J. A. McLean and T. Solouki, *Anal. Chim. Acta*, 2016, **939**, 64-72.
45. M. Miladi, A. D. Olaitan, B. Zekavat and T. Solouki, *J. Am. Soc. Mass Spectrom.*, 2015, **26**, 1938-1949.
46. M. E. Tuckerman, *J. Phys. Condens. Matter*, 2002, **14**, R1297-R1355.
47. A. M. Brunner, P. Lössl, F. Liu, R. Huguet, C. Mullen, M. Yamashita, V. Zabrouskov, A. Makarov, A. M. Altelaar and A. J. Heck, *Anal. Chem.*, 2015, **87**, 4152-4158.
48. H. Ye, R. Mandal, A. Catherman, P. M. Thomas, N. L. Kelleher, C. Ikonomidou and L. Li, *PloS one*, 2014, **9**, e92831.
49. Y. Zhao, N. M. Riley, L. Sun, A. S. Hebert, X. Yan, M. S. Westphall, M. J. Rush, G. Zhu, M. M. Champion and F. M. Medie, *Anal. Chem.*, 2015, **87**, 5422-5429.
50. B. Bogdanov and R. D. Smith, *Mass Spectrom. Rev.*, 2005, **24**, 168-200.
51. L. C. Gillet, A. Leitner and R. Aebersold, *Annu. Rev. Anal. Chem.*, 2016, **9**, 449-472.
52. B. T. Chait, *Science*, 2006, **314**, 65-66.
53. A. I. Nesvizhskii and R. Aebersold, *Mol. Cell. Proteomics*, 2005, **4**, 1419-1440.
54. K. R. Jennings, *Int. J. Mass Spectrom. Ion Phys.*, 1968, **1**, 227-235.
55. F. McLafferty, P. Bente, R. Kornfeld, S.-C. Tsai and I. Howe, *J. Am. Chem. Soc.*, 1973, **95**, 2120-2129.
56. J. M. Wells and S. A. McLuckey, *Methods Enzymol.*, 2005, **402**, 148-185.
57. L. Sleno and D. A. Volmer, *J. Mass Spectrom.*, 2004, **39**, 1091-1112.
58. R. A. Zubarev, N. L. Kelleher and F. W. McLafferty, *J. Am. Chem. Soc.*, 1998, **120**, 3265-3266.
59. J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz and D. F. Hunt, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 9528-9533.
60. J. S. Cottrell and U. London, *Electrophoresis*, 1999, **20**, 3551-3567.
61. J. K. Eng, A. L. McCormack and J. R. Yates, *J. Am. Soc. Mass Spectrom.*, 1994, **5**, 976-989.
62. J. K. Eng, B. C. Searle, K. R. Clauser and D. L. Tabb, *Mol. Cell. Proteomics*, 2011, **10**, R111. 009522.
63. W. W. B. Goh and L. Wong, *Drug Discovery Today*, 2014, **19**, 266-274.
64. D. C. Stahl, K. M. Swiderek, M. T. Davis and T. D. Lee, *J. Am. Soc. Mass Spectrom.,* 1996, **7**, 532-540.
65. T. Geiger, J. Cox and M. Mann, *Mol. Cell. Proteomics*, 2010, **9**, 2252-2261.
66. S. J. Geromanos, J. P. Vissers, J. C. Silva, C. A. Dorschel, G. Z. Li, M. V. Gorenstein, R. H. Bateman and J. I. Langridge, *Proteomics*, 2009, **9**, 1683-1695.
67. S. Purvine, E. C. Yi and D. R. Goodlett, *Proteomics*, 2003, **3**, 847-850.
68. M. Wrona, T. Mauriala, K. P. Bateman, R. J. Mortishire-Smith and D. O'Connor, *Rapid Commun. Mass Spectrom.*, 2005, **19**, 2597-2602.
69. N. J. Bond, P. V. Shliaha, K. S. Lilley and L. Gatto, *J. Proteome Res.*, 2013, **12**, 2340-2353.
70. L. C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner and R. Aebersold, *Mol. Cell. Proteomics*, 2012, **11**, O111. 016717.

71. G. Z. Li, J. P. Vissers, J. C. Silva, D. Golick, M. V. Gorenstein and S. J. Geromanos, *Proteomics*, 2009, **9**, 1696-1719.
72. J. C. Silva, R. Denny, C. A. Dorschel, M. Gorenstein, I. J. Kass, G.-Z. Li, T. McKenna, M. J. Nold, K. Richardson and P. Young, *Anal. Chem.*, 2005, **77**, 2187-2200.
73. S. J. Geromanos, C. Hughes, S. Ciavarini, J. P. Vissers and J. I. Langridge, *Anal. Bioanal. Chem.*, 2012, **404**, 1127-1139.
74. P. Fellgett, *JOSA*, 1949, **39**, 970-976.
75. P. Jacquinot, *JOSA*, 1954, **44**, 761-765.
76. R. Bateman, R. Carruthers, J. Hoyes, C. Jones, J. Langridge, A. Millar and J. Vissers, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 792-803.
77. B. Zekavat, M. Miladi, A. Al-Fdeilat, A. Somogyi and T. Solouki, *J. Am. Soc. Mass Spectrom.*, 2014, **25**, 226-236.
78. E. Badman, C. Hoaglund-Hyzer and D. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 719-723.
79. K. Giles, J. P. Williams and I. Campuzano, *Rapid Commun. Mass Spectrom.*, 2011, **25**, 1559-1566.
80. D. R. Hernandez, J. D. DeBord, M. E. Ridgeway, D. A. Kaplan, M. A. Park and F. Fernandez-Lima, *Analyst*, 2014, **139**, 1913-1921.
81. X. Liang, X. Wang, W. Wang, Q. Zhou, C. Chen, L. Peng, M. Wen, T. Qu, Z. Wang, K. Zhao, J. Li and H. Li, *Talanta*, 2014, **121**, 215-219.
82. R. Purves, A. R. Ozog, S. Ambrose, S. Prasad, M. Belford and J.-J. Dunyach, *J. Am. Soc. Mass Spectrom.*, 2014, **25**, 1274-1284.
83. S. Rokushika, H. Hatano, M. A. Baim and H. H. Hill, *Anal. Chem.*, 1985, **57**, 1902-1907.
84. A. A. Shvartsburg and R. D. Smith, *Anal. Chem.*, 2008, **80**, 9689-9699.
85. W. F. Siems, C. Wu, E. E. Tarver, H. H. Hill, Jr., P. R. Larsen and D. G. McMinn, *Anal. Chem.*, 1994, **66**, 4195-4201.
86. D. Helm, J. P. Vissers, C. J. Hughes, H. Hahne, B. Ruprecht, F. Pachl, A. Grzyb, K. Richardson, J. Wildgoose and S. K. Maier, *Mol. Cell. Proteomics*, 2014, **13**, 3709-3715.
87. S. Wong, C. Meng and J. Fenn, *J. Phys. Chem.*, 1988, **92**, 546-550.
88. J. E. Bartmess and R. M. Georgiadis, *Vacuum*, 1983, **33**, 149-153.
89. T. Solouki and J. Szulejko, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 2026-2039.
90. F. Donnarumma and K. K. Murray, *J. Mass Spectrom.,* 2016, **51**, 261-268.
91. U. Consortium, *Nucleic Acids Research*, 2014, DOI: 10.1093/nar/gku989.
92. E. R. Malinowski, *Anal. Chem.*, 1977, **49**, 612-617.
93. R. Bro and S. De Jong, *J. Chemom.*, 1997, **11**, 393-401.
94. B. Spengler, D. Kirsch, R. Kaufmann and R. Cotter, *Rapid Commun. Mass Spectrom.*, 1991, **5**, 198-202.
95. B. Zekavat, B. Harper, M. R. Brantley, M. E. Pettit and T. Solouki, Proceedings of the 62nd ASMS Conference on Mass Spectrometry and Allied Topics, Baltimore, MD, 2014.
96. K. Fukui, A. Kameyama, Y. Mukai, K. Takahashi, N. Ikeda, Y. Akiyama and H. Narimatsu, *Carbohydr. Res.*, 2006, **341**, 624-633.
97. L. Pauling, *J. Am. Chem. Soc.*, 1931, **53**, 1367-1400.
98. M. J. Polce, D. Ren and C. Wesdemiotis, *J.Mass Spectrom.*, 2000, **35**, 1391-1398.