analytical chemistry

ACS AuthorChoice

Article

pubs.acs.org/ac

Terms of Use

Conformational Ordering of Biomolecules in the Gas Phase: Nitrogen Collision Cross Sections Measured on a Prototype High Resolution Drift Tube Ion Mobility-Mass Spectrometer

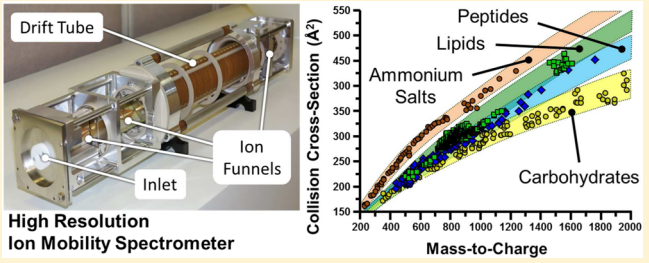
Jody C. May,[†] Cody R. Goodwin,[†] Nichole M. Lareau,[†] Katrina L. Leaptrot,[†] Caleb B. Morris,[†] Ruwan T. Kurulugama,[‡] Alex Mordehai,[‡] Christian Klein,[‡] William Barry,[‡] Ed Darland,[‡] Gregor Overney,[‡] Kenneth Imatani,[‡] George C. Stafford,[‡] John C. Fjeldsted,*[,‡] and John A. McLean*[,†]

[†]Department of Chemistry, Vanderbilt Institute of Chemical Biology, Vanderbilt Institute for Integrative Biosystems Research and Education, Vanderbilt University, Nashville, Tennessee 37235, United States

[‡]Agilent Technologies, Santa Clara, California 95051, United States

Supporting Information

ABSTRACT: Ion mobility-mass spectrometry measurements which describe the gas-phase scaling of molecular size and mass are of both fundamental and pragmatic utility. Fundamentally, such measurements expand our understanding of intrinsic intramolecular folding forces in the absence of solvent. Practically, reproducible transport properties, such as gas-phase collision cross-section (CCS), are analytically useful metrics for identification and characterization purposes. Here, we report 594 CCS values obtained in nitrogen drift gas on an electrostatic drift tube ion mobility-mass spectrometry (IM-MS) instrument. The instrument platform is a newly developed prototype incorporating a uniform-field drift tube bracketed by electrodynamic ion funnels and coupled to a high resolution quadrupole time-of-flight mass spectrometer. The CCS values reported here are of high experimental precision (±0.5% or better) and represent four chemically distinct classes of molecules (quaternary ammonium salts, lipids, peptides, and carbohydrates), which enables structural comparisons to be made between molecules of different chemical compositions for the rapid "omni-omic" characterization of complex biological samples. Comparisons made between helium and nitrogen-derived CCS measurements demonstrate that nitrogen CCS values are systematically larger than helium values; however, general separation trends between chemical classes are retained regardless of the drift gas. These results underscore that, for the highest CCS accuracy, care must be exercised when utilizing helium-derived CCS values to calibrate measurements obtained in nitrogen, as is the common practice in the field.



With the rising demand for high-throughput analyses of increasingly complex samples, ion mobility-mass spectrometry (IM-MS) has found broad application in the analysis of biological systems, as this rapid 2D separation (ms and μs, respectively) provides comprehensive molecular information regarding analyte size, mass, and relative abundance. In ion mobility, separation is achieved by low-energy interactions of charged analytes with an inert buffer gas (conventionally helium or nitrogen), where analyte size-to-charge ratio is measured as a function of the time required to traverse the mobility region.[1] As a means of comparison with other laboratory measurements, drift time values are either normalized to standard temperature and pressure as a reduced mobility (K_0) or converted to a collision cross-section (CCS) value, the latter of which is a size parameter related to the averaged momentum transfer impact area of the molecule.[2] Structural information in the form of CCS values assists in the characterization of analytes by biomolecular class, as these classes are known to separate in IM-MS space and adopt conformational correlations due to prevailing class-specific structural folding in the gas phase.[3,4] These class-specific mobility-mass correlations can be used as a predictor for molecule class, demonstrating the potential value of IM-MS structural separations for life sciences research which seek systems biology level information. Expanding upon this concept, CCS-based molecular prediction has previously been explored for peptides, utilizing intrinsic size parameter calculations[5,6] and machine learning algorithms[7] for sequence prediction, but no detailed study of other biochemical classes has yet been undertaken.

The separation and characterization of biological samples by IM-MS has been achieved using both commercial and laboratory built instrumentation. Virtually all contemporary commercial IM-MS instruments utilize nitrogen as the buffer gas for IM separations, motivated by practical considerations of cost, availability, and technical considerations for pumping requirements and electrical discharge. The most common

Received: November 7, 2013
Accepted: January 22, 2014
Published: January 22, 2014

   dx.doi.org/10.1021/ac4038448 | *Anal. Chem.* 2014, 86, 2107−2116

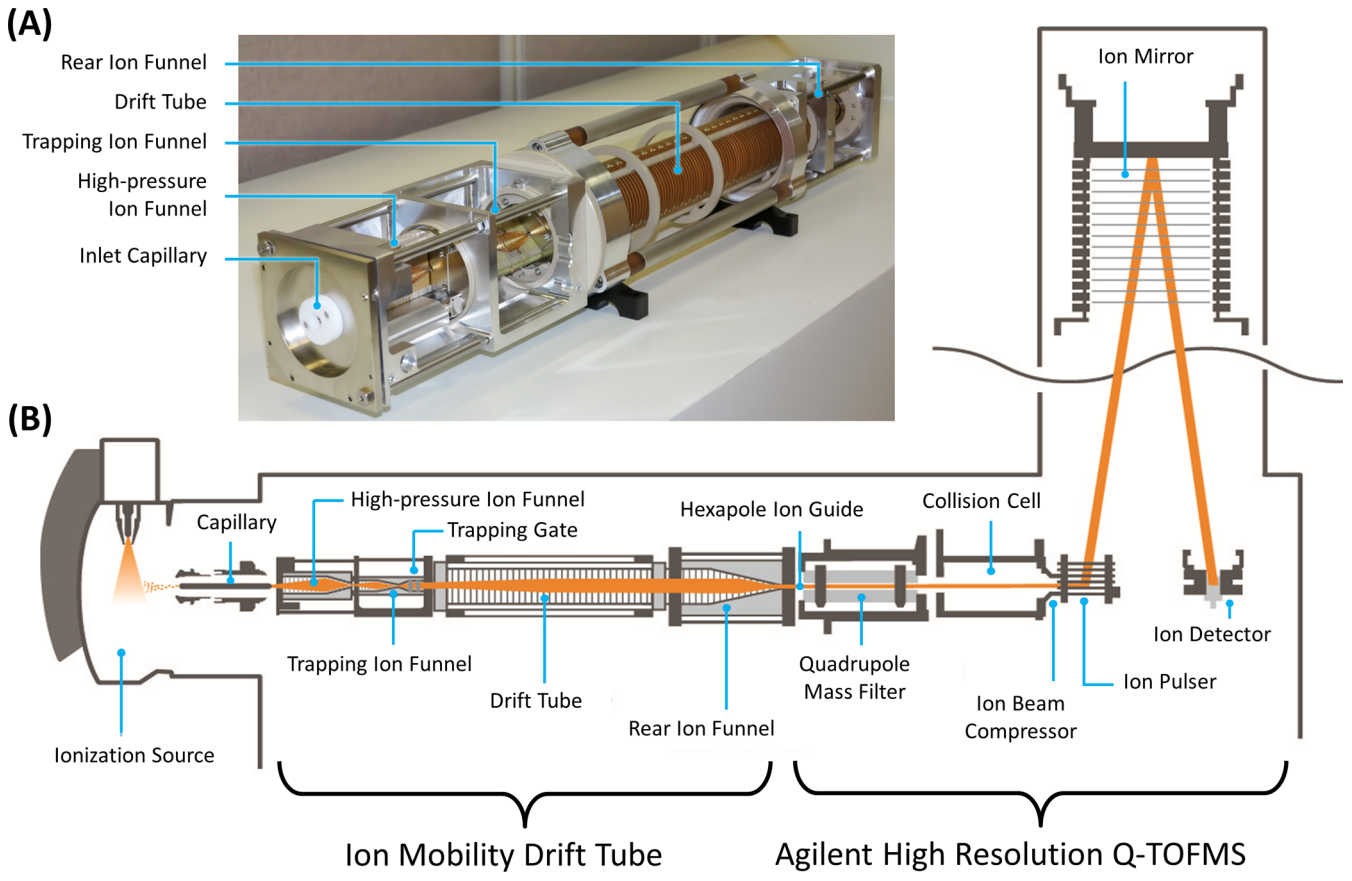


Figure 1. Details of the prototype IM-MS instrumentation used in this study. (A) A picture of the ion optical elements of the ion mobility component. (B) A representative schematic of the instrumentation used with significant components annotated.

commercial IM-MS platform utilizes an electrodynamic field (*i.e.*, a traveling wave potential) for mobility separation,[8] and drift time measurements must be calibrated against electrostatic drift tube data in order to convert these measurements to CCS values.[9,10] Conversely, many independently constructed instruments incorporate uniform electrostatic field mobility regions utilizing helium as the buffer gas. Uniform field measurements serve as the benchmark for electrodynamic CCS value determination, as the CCS obtained from a uniform field drift tube can be determined empirically through kinetic theory.[11,12]

One common practice among researchers utilizing IM-MS is calibration of nitrogen-based traveling wave ion mobility measurements against helium-based CCS values reported in the literature.[13,14] The use of helium-based CCS values to calibrate nitrogen-based drift time measurements results in calibrated "helium-equivalent" CCS values, which can be useful for comparing with literature values and correlating measurements to theory.[15,16] There is, however, concern that this practice introduces added experimental error, as nitrogen vs helium mobility measurements differ substantially in magnitude, and the success of calibration strategies relies heavily on careful selection of calibrants that accurately describe the sample conditions, charge state, mass range, and chemical class of the system of interest.[9,14,17] Differences in CCS values in helium versus nitrogen arise due to several factors including intrinsic size differences between the buffer gases, mass effects which factor into the momentum transfer cross-section (the experimental CCS), and the over 8-fold difference in gas polarizability between helium and nitrogen (0.21×10^{-24} and 1.74×10^{-24} cm^3, respectively).[12,18]

Recently, a prototype IM-MS instrument utilizing nitrogen drift gas was developed (Agilent Technologies, Santa Clara, CA). This instrument incorporates a uniform electrostatic field ion mobility separator bracketed by electrodynamic focusing devices (ion funnels), which allows for high sensitivity and direct measurements of CCS values in nitrogen.[7,19] Presented in this report is an extensive and diverse database of empirically derived nitrogen CCS measurements (594 values), which comprises four molecular classes and expands upon several previous databases for the structural characterization of biological molecules.[5,7,9,20−23] This affords the opportunity to explore the fundamental considerations of buffer gas composition and the subsequent effects on ion mobility parameters (reduced mobility and CCS) across different molecular classes.

EXPERIMENTAL METHODS

Preparation of Standards. *Lipids.* All solvents and buffers were purchased as HPLC grade from Sigma-Aldrich (St. Louis, MO, USA). Dry lipid extracts were purchased from Avanti Lipids (Birmingham, AL, USA) and constituted in chloroform prior to analysis. Lipid extracts include sphingomyelins (SM, porcine brain), glycosphingolipids (GlcCer, porcine brain), phosphatidylcholines (PC, chicken egg), phosphatidylserines (PS, porcine brain), and phosphatidylethanolamines (PE, chicken egg). For analysis, lipid standards were diluted in 90% chloroform/10% methanol (v/v) with 10 mM sodium

dx.doi.org/10.1021/ac4038448 | *Anal. Chem.* 2014, 86, 2107−2116

acetate to a final concentration of 10 μg/mL. Putative identification of lipids was performed using the exact mass measurement through the Lipid Metabolites and Pathways Strategy (LIPID MAPS) Structural Database (LMSD).[24] A full list of identified lipids can be found in the Supporting Information.

Carbohydrates. Carbohydrate dextrins (linear and cyclic) and sugar alcohol standards were purchased from Sigma-Aldrich. Lacto-*N*-difucohexaose I and II and lacto-*N*-fucopentaose I and II were purchased from Dextra Laboratories (Reading, UK). All carbohydrate standards were prepared as received and reconstituted in water with 10 mM ammonium acetate to final concentrations of 10 μg/mL. For cationization, 10 mM NaCl, 10 mM LiCl, 10 mM CsCl, 10 mM KCl, and 10 mM RbCl solutions were prepared in water to a final concentration of *ca.* 10 μM. A full list of identified carbohydrates can be found in the Supporting Information.

Peptides. Predigested peptide standards (MassPREP) were purchased from Waters (Milford, MA, USA). Peptide standards (SDGRG and GRGDS) were purchased from Sigma-Aldrich. All peptide standards were received as a lyophilized powder and reconstituted in 10 mM ammonium acetate in water to a final concentration of 10 μg/mL. The MassPREP digestion standard mix contained approximately equimolar concentrations of four tryptically digested proteins: Alcohol Dehydrogenase (ADH, yeast), Serum Albumin (BSA, bovine), Phosphorylase B (PHOSPH, Rabbit) and Enolase (ENOLASE, yeast). Peptide identifications were assigned on the basis of exact mass of all possible tryptic peptides (no missed cleavages) produced by the Expert Protein Analysis System (ExPASy) PeptideMass proteomics tool[25] (Swiss Institute of Bioinformatics, Lausanne, Switzerland) using the SWISS-PROT database entry number for each intact protein (P00330, P02769, P00924, and P00489, respectively). A full list of identified peptides can be found in the Supporting Information.

Quaternary Ammonium Salts. Tetraalkylammonium (TAA) salts with alkyl chain lengths between 3 and 18 carbons (TAA3 to TAA18) were purchased from the following sources: TAA4, TAA6, TAA7, TAA10, TAA12, and TAA16 from Sigma-Aldrich; TAA3, TAA5, and TAA8 from Acros Organics; and TAA18 from Alfa Aesar. All TAA salts were supplied with a stated purity of greater than 98% and were prepared as received. TAA3 to TAA8 were prepared in 50% methanol/50% water, while TAA10, TAA12, TAA16, and TAA18 were prepared in 50% methanol/50% isopropanol. Final concentrations were *ca.* 1 μg/mL. A full list of primary TAA salt standards and concomitant ions identified in the samples can be found in the Supporting Information.

Instrumentation. A schematic of the instrumentation used to obtain the cross-section measurements is shown in Figure 1. The instrument used in this work is a commercial prototype IM-MS which incorporates a drift tube coupled to a quadrupole time-of-flight mass spectrometer (IM-Q-TOFMS, Agilent Technologies, Santa Clara, CA). For this work, an orthogonal electrospray ionization (ESI) source (Agilent Jet Stream) was utilized which incorporates a heated sheath gas nebulizer to aerodynamically focus and desolvate ions prior to introduction into the vacuum system. Ions from the ESI are introduced to a single-bore glass capillary tube which is resistively coated across its length, allowing the nebulizer to be maintained at ground potential, while the exit end of the capillary can be biased to around 2100 V.[26] Ions exiting the capillary are introduced into a tandem ion funnel interface consisting of a high-pressure transmission ion funnel in the first stage,[27] followed by a second stage trapping ion funnel which incorporates a dual-grid ion gate.[28] The second stage ion funnel trap operates as an ion focusing and accumulation region whereby temporally narrow (typically 100 to 150 μs) ion pulses are gated into the IM spectrometer. Mobility separation occurs in a 78 cm uniform field drift tube composed of a series (*ca.* 150) of 50 mm internal diameter gold-plated ring electrodes. The buffer gas is high purity nitrogen. Ions traverse the drift tube under the influence of a weak electric field (10 to 20 $V \cdot cm^{-1}$) and consequently drift under low-field conditions. The combination of extended drift length, precision electronics, and high drift voltages enables high resolution ion mobility separations in excess of 60 resolving power ($t/\Delta t$, observed for a +1 ion, m/z 294). Resolving power values can vary depending on the analyte charge state and also the inherent structural rigidity of the ion and do not depend on the class of molecules being investigated. Ions exiting the drift region are refocused axially using an ion funnel and traverse a differential pressure interface region by means of a resistively coated hexapole ion guide. Following the hexapole, ions are introduced into a modified Q-TOFMS (Agilent 6550), which incorporates a quadrupole mass filter and collision cell to enable mass-selective ion fragmentation experiments. The TOFMS is capable of greater than 40 000 mass resolving power and can acquire MS spectra at a rate of up to 8.3 kHz (120 μs transients at m/z 1700). Additional instrumentation details are provided in Figure 1.

Experimental Parameters. All 2D IM-MS spectra were acquired via direct infusion using positive mode electrospray ionization (Agilent Jet Stream Source) with a flow rate of *ca.* 10 μL/min. The Jet Stream source was operated with a nitrogen sheath gas temperature between 400 and 600 K (solvent dependent) at a flow rate of 12 L/min. Nitrogen drying gas applied at the source entrance was heated to *ca.* 570 K at a flow rate of 10 L/min. The source was operated in positive mode with the following voltages: ground potential emitter, −4.5 kV capillary entrance, and −1.8 kV nozzle. The three ion funnels were operated as follows: high-pressure funnel RF 100 Vpp (peak-to-peak) at 1.5 MHz, 150 V DC; trapping funnel RF 100 Vpp at 1.2 MHz, 180 V DC; rear funnel RF 100 Vpp at 1.2 MHz, 200 V DC. The IM drift gas pressure (nitrogen) was maintained at *ca.* 4 Torr and *ca.* 300 K, while the drift potential varied from 750 to 1450 V, which represents an E/N ratio of 7 to 15 Td. In this E/N range, the mobility operates under low field conditions as all analytes investigated exhibited a linear change in drift times with respect to the electric field. Data was acquired with a modified version of the MassHunter software (Agilent Technologies). The mass measurement was calibrated externally using a series of homogeneously substituted fluorinated triazatriphosphorines (Agilent tuning mixture, *ca.* 100 to 3000 m/z), which are characterized as being amphoteric and nonreactive. Additionally, a mixture of tetraalkylammonium salts (TAA3 to TAA18) was added to all samples as an internal mass and mobility calibration standard for positive mode analysis.

Collision Cross-Section Calculations. Uncorrected drift times are extracted as centroid values using a beta version of the IM-MS Browser (Agilent Technologies). This uncorrected drift time represents the total transit time of the ions, including the mobility drift time and the flight time through the interfacing IM-MS ion optics and MS. Because the nonmobility flight time component (the transit time of ions outside the drift region) is independent of the drift voltage, this value can be determined

dx.doi.org/10.1021/ac4038448 | *Anal. Chem.* 2014, 86, 2107−2116

Table 1. Summary of Statistics Related to the CCS Database

	collision cross-section statistics					fits to empirical data		
	number of CCS values	mass range [Da]	CCS range [Å^2]	average CCS precision[a]	average N for each value	fit equation coefficients ($y = Ax^B$)	coefficient of determination[b]	amount of data included within ±5% of fit[c]
peptides	92	430−1760	200−450	0.2% (±0.1%)	7 (±2)	A = 6.8440 B = 0.5547	R^2 = 0.975	91%
carbohydrates	125	190−2150	140−410	0.3% (±0.1%)	12 (±3)	A = 11.553 B = 0.4656	R^2 = 0.983	89%
lipids	314	500−1600	220−460	0.2% (±0.1%)	10 (±2)	A = 5.2469 B = 0.6000	R^2 = 0.949	96%
tetraalkyl-ammonium salts	63	130−1030	140−400	0.4% (±0.1%)	18 (±8)	A = 8.2631 B = 0.5561	R^2 = 0.991	98%

[a]The precision reported here represents the reproducibility across replicate measurements. The total precision due to propagation of uncertainty in experimental parameters is estimated to be less than 2%. [b]The observed R^2 value for the nonlinear power fit. [c]The data inclusion band chosen is based on the smallest sized band which incorporates the most amount of data (refer to Figure 2B, inset).

Table 2. Measured CCS Values for the TAA Salts Compared with Literature Values

name		exact mass [Da]	CCS (this work[a]) [Å^2]	CCS (literature[b]) [Å^2]	abs. percent difference[c] [%]
tetramethylammonium	TAA1	74.14		107.40	
tetraethylammonium	TAA2	130.25		122.20	
tetrapropylammonium	TAA3	186.36	144.1 ± 0.7 (23)	143.80	0.22
tetrabutylammonium	TAA4	242.46	166.6 ± 0.9 (16)	166.00	0.36
tetrapentylammonium	TAA5	298.57	190.1 ± 1.0 (28)	190.10	0.02
tetrahexylammonium	TAA6	354.68	213.5 ± 1.0 (31)	214.00	0.23
tetraheptylammonium	TAA7	410.78	236.4 ± 0.4 (31)	236.80	0.17
tetraoctylammonium	TAA8	466.54	256.6 ± 0.7 (31)	258.30	0.64
tetradecylammonium	TAA10	579.11	293.5 ± 0.7 (24)		
tetradodecylammonium	TAA12	691.32	319.0 ± 0.9 (24)		
tetrahexadecylammonium	TAA16	915.04	361.5 ± 0.9 (24)		
tetraoctadecylammonium	TAA18	1027.16	379.0 ± 1.7 (21)		

[a]The number of measurements is reported in parentheses. The error due to experimental uncertainty is reported next to each value and is less than 0.5% for all measurements. The total error based on propagating the limits of precision in experimental parameters is estimated to be less than 2%. [b]Literature values from ref 16. [c]The absolute percent difference is the difference in CCS compared to the average of both values.

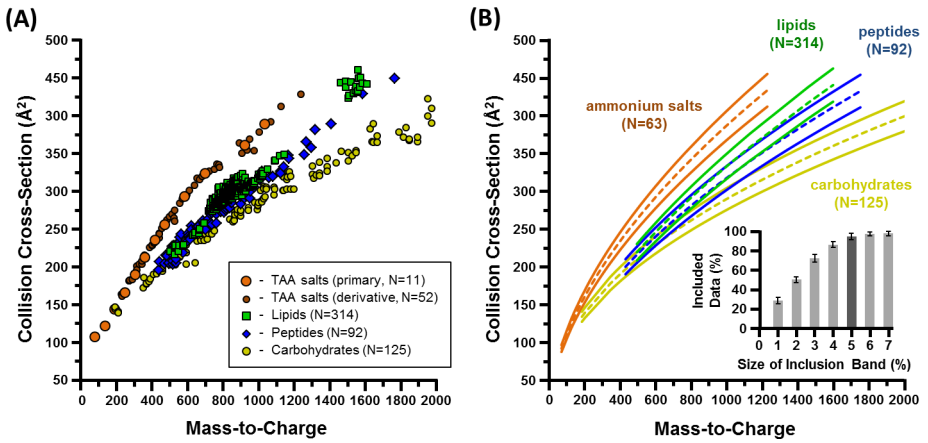


Figure 2. (A) A scatter plot of the CCS values measured in this study, separated by chemical class. (B) Best fit lines of the data, separated into class and fit to a power-law function. Also shown are data inclusion bands representing ±5% deviation from the best fit line. The inset bar graph represents the amount of data included within different sized inclusion bands. Fit equations and their corresponding coefficients of determination (R^2) can be found in Table 1.

from a plot of the measured drift time versus the inverse drift voltage,[29] where a linear fit to the data will indicate the nonmobility time component (y-intercept) in the limit of infinite electric field ($1/V$ of zero). Time measurements are obtained from a minimum of six different drift voltages, ranging from 750 to 1450 V. The determined nonmobility time is subtracted from the uncorrected drift times in order to obtain the corrected ion mobility drift time. Corrected drift times are used to determine the gas-phase momentum transfer collision cross-section (CCS) using the Mason-Schamp relationship,[30] incorporating the scaling terms for standard temperature and pressure. On the basis of a propagation-of-error analysis

dx.doi.org/10.1021/ac4038448 | Anal. Chem. 2014, 86, 2107−2116

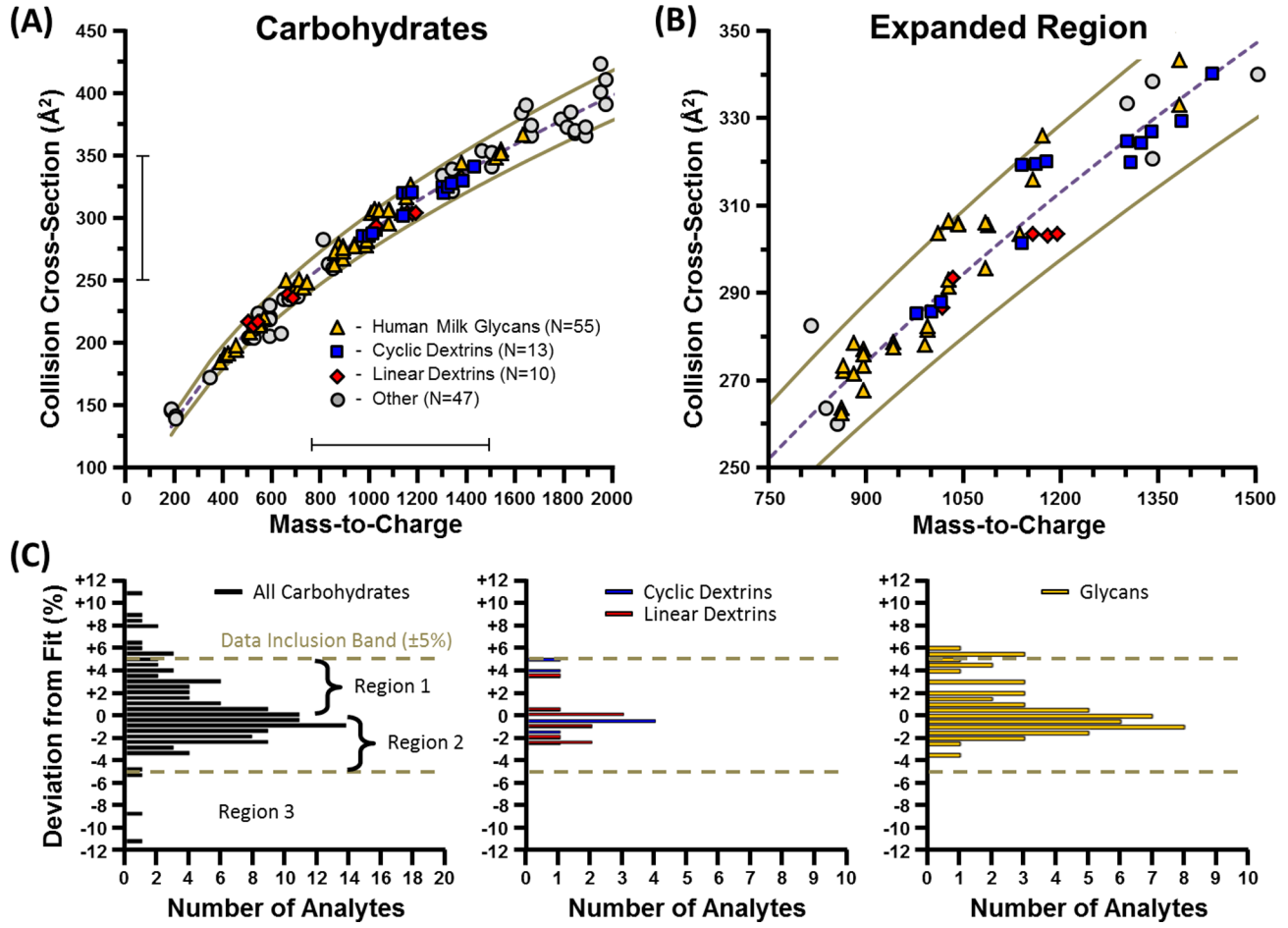


Figure 3. A subclass analysis of carbohydrates, with subclasses composed of human milk derived glycans, cyclic, and linear dextrins. (A) A scatter plot of the relative location of carbohydrate subclasses in 2D IM-MS conformational space. (B) An expanded region of the scatter plot where all three subclasses of carbohydrates are observed. (C) A histogram analysis of carbohydrate subclass deviation in 2D IM-MS space relative to the best fit line. In general, the carbohydrate subclasses do not differentiate into distinct regions of conformational space.

incorporating the limits of precision for individual experimental parameters, we estimate the accuracy of all CCS values to be better than 2% (see Supporting Information).

RESULTS AND DISCUSSION

Database Description and General Cross-Section Trends in Nitrogen. A total of 594 nitrogen collision cross-section values were measured empirically in this study, representing three biomolecular classes (lipids, carbohydrates, and peptides) and TAA salts. This includes 92 peptides, 125 carbohydrates, 314 lipids, and 63 TAA salts and TAA salt derivatives. All CCS values were measured in positive ion mode, and all represent singly charged analytes, of which 63 are molecular ions, 111 are protonated species, 273 are sodiated, 124 are potassiated, and the remaining representing other cations (lithium, rubidium, and cesium). The range of CCS values measured spans from ca. 150 to 450 Å^2, covering a mass range of ca. 150 to 2200 Da. Summary statistics regarding the CCS database are provided in Table 1. The average RSD of all database values was 0.3% (±0.1%), with each CCS value representing an average of 11 (±4) measurements. A complete list of all analytes and respective CCS measurements is provided as Supporting Information.

TAA salts ranging from tetrapropylammonium (TAA3) to tetraoctadecylammonium (TAA18) were analyzed, and a subset of these measured CCS values were compared with literature values in order to estimate the CCS measurement accuracy.[16] Results of this comparison are summarized in Table 2. Where CCS literature values existed for nitrogen, the absolute differences were found to be less than 1% and, in most cases, less than 0.5% deviation was observed. All TAA salts investigated exhibited excellent CCS measurement reproducibility (less than 0.5% RSD).

A scatter plot of CCS versus m/z for all database values is presented in Figure 2A, separated into chemical classes. We refer to this type of 2D IM-MS projection as conformational space analysis,[4,31] as the differential scaling of mass (m/z) and size (CCS) between molecular classes is indicative of differences in gas-phase packing efficiency.[20]

Description of the Fits to the Empirical Data. Several different equation functional forms were evaluated in order to determine which expression best described molecular class correlations between CCS and m/z values, and it was found that the data sets were adequately described by a power-law relationship ($y = Ax^B$), based upon the coefficient of determination (R^2). Conceptually, power-law equations are descriptors for several phenomena related to mass-size scaling, including allometric scaling laws in biology,[32] stellar velocity dispersion relative to black hole mass (M-sigma relation),[33] and the well-known square-cube law, first described by Galileo,[34]

dx.doi.org/10.1021/ac4038448 | Anal. Chem. 2014, 86, 2107−2116

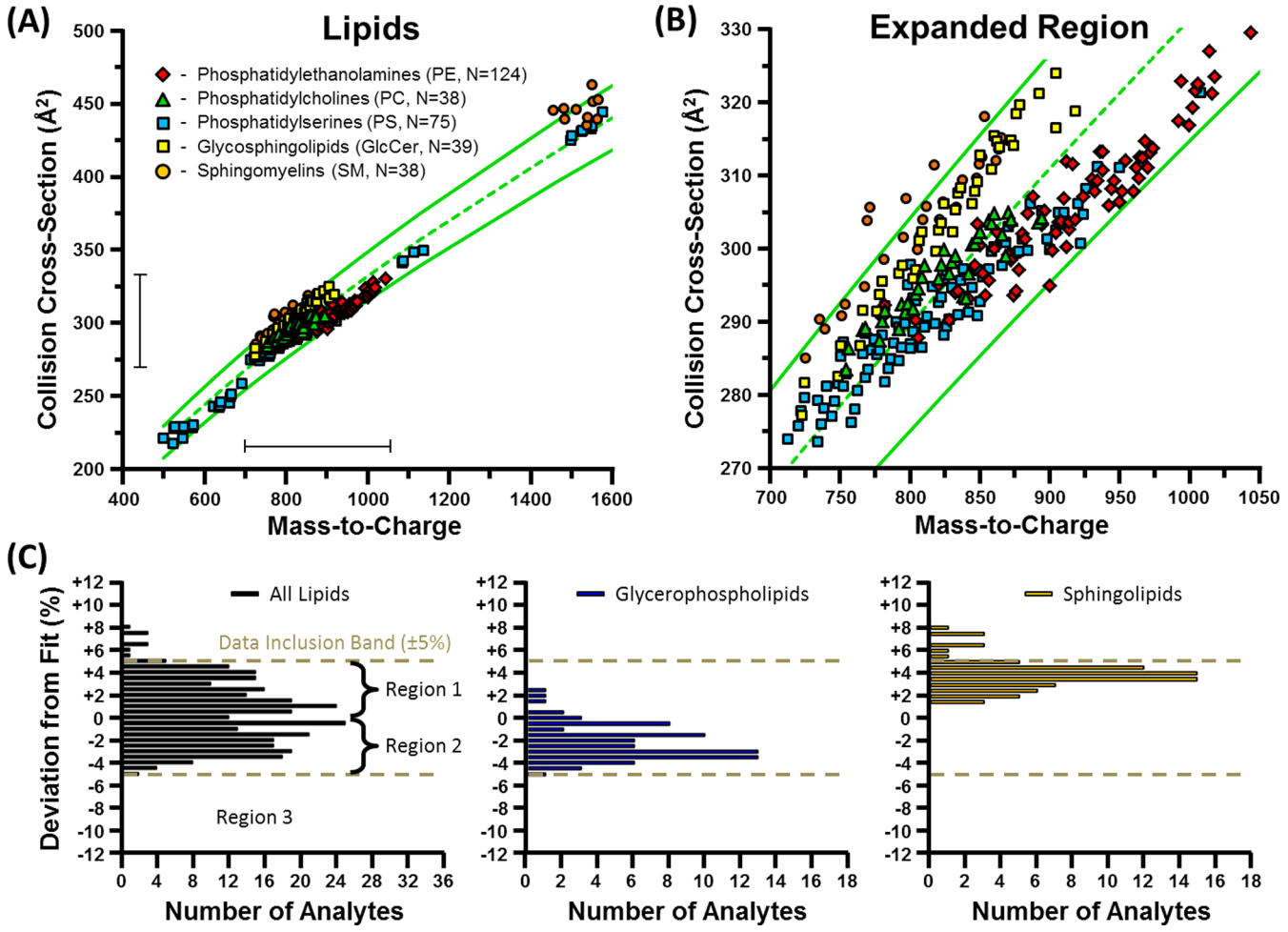


Figure 4. A subclass analysis of lipids composed of PE, PC, PS, GlcCer, and SM lipids. These lipids are further categorized into two general structural groups: glycerophospholipids (PE, PC, PS) and sphingolipids (GlcCer, SM). (A) A scatter plot of the conformational ordering of each subclass of lipid. (B) An expanded region of the scatter plot detailing a preferentially ordering of the different lipid subclasses in conformational space. (C) A histogram analysis and locations of general lipid structural groups relative to the best fit line. Unlike carbohydrates, individual lipid subclasses partition into distinct regions of 2D IM-MS space, allowing for finer structural information to be extracted from the conformational space analysis.

which universally relates any shape's increase in volume relative to its surface area. Additionally, power-law relationships are scale-invariant such that different power-law functions can be related by a simple scaling factor, which has implications for describing universal relationships independent of the specific details of the measurement.

The resulting power-law fits to the empirical data are presented in Figure 2B. Coefficients and associated R^2 values are summarized in Table 1. The data inclusion bands projected in Figure 2B represent ±5% deviation from the line of best fit. Other inclusion band sizes are summarized in Figure 2B, inset, averaged across the four data sets. For all data sets, a ±5% inclusion band incorporated an average of 94% (±4%) of data. Decreasing the band to ±4% results in an average of 86% (±3%) of data being included (a decrease of *ca.* 8% data inclusion), whereas increasing the band to ±6% only incorporated an additional 3% (±2%) of data on average. Thus, the ±5% data inclusion band represents an optimal balance between specificity and data incorporation. Interestingly, the ±5% band describes all data sets similarly, regardless of chemical class.

Several observations can be made from the data contained in Figure 2. The TAA salts were found to exhibit the highest CCS values relative to m/z and were located in a region of 2D IM-MS space which was disparate from the biomolecules. Previously, TAA salts were recommended as an ion mobility calibrant due to their low propensity for forming clusters, which otherwise complicates the interpretation of mobility data.[35] Here, it is found that, in addition to the lack of clustering, the TAA salts are useful mobility-mass calibrants as the complete series (1 to 18 carbons) span a wide range of CCS values (107 to 400 Å^2) and m/z values (75 to 1027 Da) and occupy a region of 2D IM-MS space where biomolecules are not predicted to occur. Carbohydrates were observed to have the lowest CCS values relative to their mass, while peptides and lipids occupy similar regions of conformational space. In general, all of the biochemical classes surveyed were readily separated above a mass of *ca.* 1200 Da, indicating that differences in relative gas-phase packing scale with molecular size and mass.

Extraction of Sub-Trend Information from the Data. From a cursory analysis of the CCS database described in this report, it is evident that the general chemical class information is retained through the specific mobility-mass correlation trends in the 2D IM-MS projection. While class separations are unambiguous at the higher m/z values (beyond *ca.* 1200), class-

dx.doi.org/10.1021/ac4038448 | *Anal. Chem.* 2014, 86, 2107−2116

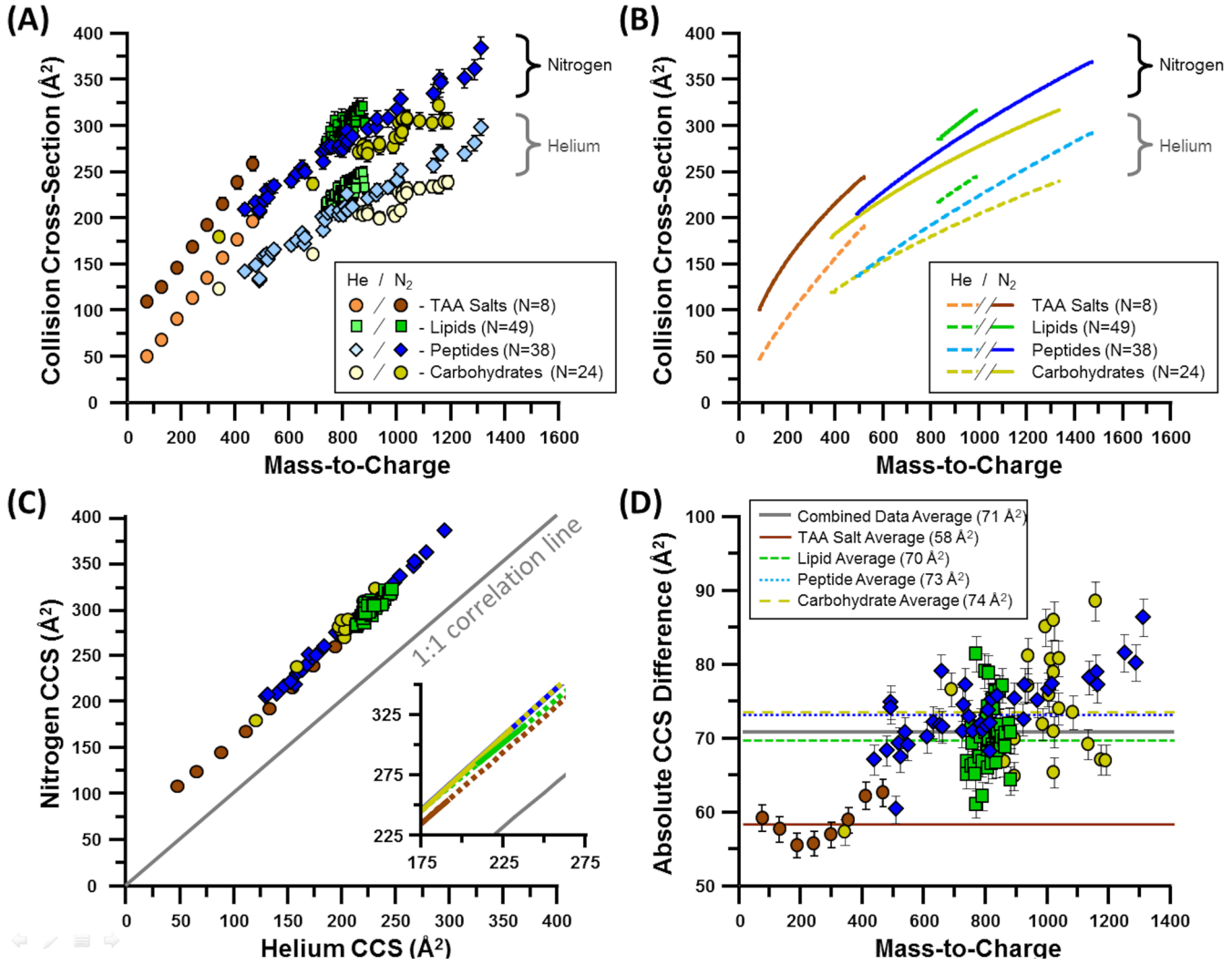


Figure 5. Comparisons between helium and nitrogen-derived CCS values. (A) A scatter plot of class-specific subsets of CCS data measured in both helium and nitrogen. (B) Power fits to the data projected in panel A. (C) Correlation plot of helium vs nitrogen CCS values. (D) Absolute differences in CCS between helium and nitrogen measurements, plotted as a function of mass-to-charge. In general, nitrogen CCS values are significantly larger than helium, with subtle differences being observed between different chemical classes.

specific trend information is still largely retained within the regions of overlap. For example, within the intermediate region where the majority of signals occur (m/z 700 to 1000), the class-specific mobility-mass correlations partition into distinct bands which can be subjected to a probability distribution analysis for molecular class information (see, for example, Figure S2, Supporting Information). The molecular information derived from such trends holds promise for conducting comprehensive omics experiments whereby unknown analytes originating from a complex sample (*e.g.,* blood, tissue, whole cell lysate) can be prioritized on the basis of their likely chemical class. This biomolecular filtering would allow for the sorting of unknown analytes into distinct identification workflows, as lipid, peptide, metabolite, and glycan identification methods often warrant searching of specific databases. In order to determine the detail of class-specific information obtained from the conformational space analysis, select coarse biomolecular classes were further categorized into finer specific subclasses. Figure 3 contains a detailed analysis of carbohydrates, which were further delineated into glycans (human milk oligosaccharides), cyclic dextrins (cyclodextrins), and linear dextrins (maltose polysaccharides). Figure 3A,B illustrates the relative location of each carbohydrate subclass in conformational space, while Figure 3C describes the data as a histogram relative to the best fit line. In general, there is no strong correlation between the carbohydrate subclasses, with all signals distributed in relatively the same locations with respect to the power-law fit. This suggests that the carbohydrates surveyed do not adopt strong structural differences which can be easily differentiated in the 2D analysis. On the other hand, the subclasses chosen here represent broad descriptors for carbohydrate structure and as such are not structurally descriptive subclassifications. For example, glycans can represent both linear and branched oligosaccharides and thus occupy a broad region of the total carbohydrate conformational trend. Interestingly, the cyclization of sugars (cyclodextrins) does not seem to enhance gas-phase packing efficiency as compared with their linear analogues. A more comprehensive carbohydrate data set may engender subclass differentiation, or differences may bear out for more limited situations such as positional and structural isomers or various metal-coordinated species.[36] It should also be noted that the data projected in Figure 2 includes various alkali cationized species. While previous work has indicated that carbohydrate gas-phase ion

dx.doi.org/10.1021/ac4038448 | *Anal. Chem.* 2014, 86, 2107−2116

structure is strongly influenced by the cation,[37] it is difficult to draw any definitive cation-specific effects in this work due to the structurally diverse nature of the analytes (the cation-specific carbohydrate analysis is provided in Figure S1, Supporting Information).

Application of a similar subclass analysis to the lipid data set is illustrated in Figure 4. In this case, the lipid data set is substantially larger than the carbohydrate data set (N = 314 vs N = 125, respectively), and measurements were obtained from five distinct lipid structural classes. These lipid subclasses can be broadly categorized into two structural classes as sphingolipids (SM, GlcCer) and glycerophospholipids (PE, PC, PS). It is qualitatively evident in Figure 4A,B that each class of lipid exists in a distinct region of conformational space. The histogram distribution analysis in Figure 4C (right panel) indicates that sphingolipids fall predominantly above the best fit line (97% in region 1), whereas glycerophospholipids (Figure 4C, middle panel) are more broadly dispersed around the mobility-mass correlation (33% in region 1, 65% in region 2), and adopt denser gas phase conformations than sphingolipids. These results suggest that, with proper structural subclass descriptors, conformational space analysis is capable of differentiating finer structural detail beyond general biomolecular class.

Comparisons between Helium and Nitrogen CCS Values. The diverse compilation of CCS values described in this report allows for direct comparisons against helium-derived CCS values reported in the literature. Of the over 3000 singly charged helium CCS values surveyed from the literature, overlapping measurements exist for 119 nitrogen CCS values in the current database (8 TAA salts, 49 lipids, 38 peptides, and 24 carbohydrates; refer to Supporting Information). Differences between helium and nitrogen-derived CCS measurements have been previously noted for atomic species,[38] small molecules, and peptides,[39] and more recently, proteins and large protein complexes.[9,23] Here, we add the differences observed for TAA salts, lipids, and carbohydrates, in addition to corroborating previous peptide observations.

A scatter plot of the overlapping helium and nitrogen CCS values is provided in Figure 5A. Vertical error bars representing ±2% are also included, although this error is sufficiently small such that most of the error bars are obscured within the scale of individual data points. Figure 5B contains the power fits to the data, which are useful in visualizing differences between data sets. In general, gross separation trends between chemical classes are retained within the helium and nitrogen-based data sets, with qualitatively similar conformational space ordering being exhibited regardless of the drift gas (*i.e.*, carbohydrate density > peptide density > lipid density > TAA salt density). Figure 5C contains the same overlap data as projected on a plot of nitrogen versus helium CCS values. In Figure 5C, all of the class-specific data reside within the same region of the projection, indicating that overall differences between helium and nitrogen CCS are systematic within this range and thus can be accounted for to allow conversion of one data set to another, with some loss in precision associated with error propagation. This possibility of generating effective helium-based CCS values from nitrogen measurements was previously noted by Bush et al. for peptides and proteins.[9,22] Recently, Pagel and Harvey noted good correlation (less than 1.5% error) between helium and nitrogen CCS measurements for singly charged carbohydrates, though significant error was introduced when multiply charged values were incorporated into the calibration.[17] Here, we confirm a strong correlation between singly charged helium and nitrogen CCS values for lipids, peptides, carbohydrates, and TAA salts. It should be cautioned, however, that the relationship between helium and nitrogen-based CCS values are both charge state and mass dependent,[40] and it is expected that any correlation between the two measurements would deviate at the extremes of low and high mass. In fact, Bush et al. previously noted that cross-calibration error from nitrogen to helium CCS is higher at lower masses (up to 15% error) where the magnitude of the CCS value is small, while at higher masses, the error can be reduced to as low as 2.2% for predicting helium CCS from nitrogen measurements.[9] It was also noted in this study and elsewhere that calibration across different chemical classes (*e.g.*, using literature peptide values to calibrate lipids[14]) introduces additional and significant error (*ca.* 7%), further underscoring the importance of compiling a chemically diverse set of empirical drift tube CCS values. Figure 5C, inset, contains the linear best fits to the data, with the axes rescaled to a region where data exists for all four chemical classes. Linear fits are extrapolated (dotted lines) for visualization purposes. Here, the small but notable differences between chemical classes can be observed as offset correlation lines, which corroborate the absolute CCS differences between helium and nitrogen noted previously for each chemical class. Specifically, peptides, carbohydrates, and lipids fall along a similar helium−nitrogen CCS correlation trend, while the TAA salts exhibit a slightly lower correlation. Interestingly, all class correlations exhibit similar slopes (*ca.* 1), suggesting that the factors which give rise to the cross-sectional differences between helium and nitrogen (buffer gas size, mass, and polarizability) affect different chemical classes in a similar manner across a broad range of both size and mass.

Absolute CCS differences between the helium and nitrogen data sets are plotted as a function of mass in Figure 5D, with error bars representing ±2% CCS uncertainty. Average absolute CCS differences are projected as a horizontal line through each class distribution, with the following values: TAA salts, 58 (±3) Å^2; lipids, 70 (±4) Å^2; carbohydrates, 74 (±8) Å^2; and peptides, 73 (±5) Å^2. Cross-sectional differences are lowest for the TAA salts, while lipids, carbohydrates, and peptides differ by approximately the same amount. Overall, there is a small but notable increase in the helium−nitrogen CCS difference with increasing mass for all classes except lipids where a limited mass range is surveyed. This suggests that the nitrogen and helium CCS are not increasing at the same rate relative to the mass of the analyte, with the greater CCS increase occurring in nitrogen. Wyttenbach et al. recently noted that ion systems up to *ca.* 760 Da (sodiated PEG_{17}) still exhibit strong contributions from the ion-neutral interaction potential in their measured CCS.[41] From their atomic superposition argument, it would be expected that, with nitrogen buffer gas, the combined effect of each atomic potential for large polyatomic systems would give rise to a steeper increase in CCS than with helium buffer gas, since the atom−nitrogen interaction potential is stronger than the atom−helium interaction potential. In other words, the stronger interaction potential of nitrogen would be expected to scale with the number of atoms in the ionic system being measured, at least to a first approximation. Ion systems with different heteroatom compositions (*e.g.*, lipids vs peptides) would also be expected to exhibit different scaling of mass to CCS between helium and nitrogen; this effect cannot be definitively observed in the relatively narrow mass range surveyed in this work, though cursory effects of gas polarization seem to be present in the

enhanced high-mass separation of lipids and peptides in nitrogen vs helium. Such class-specific CCS differences may bear out as more overlapping measurements are obtained in future studies.

CONCLUSIONS

The large database of nitrogen-derived CCS values presented here offers a glimpse at the intrinsic intermolecular packing forces of four chemically different molecular classes across a relatively wide range of both size (*ca.* 150 to 450 Å^2) and mass (*ca.* 150 to 2200 Da). Four molecular classes were investigated in this study, with relative gas-phase densities observed as follows, from least to most efficient packing: TAA salts, lipids, peptides, and carbohydrates. The biopolymers (carbohydrates and peptides) demonstrated the highest efficiency for gas-phase packing, and among these, carbohydrates tend to adopt the most compact gas-phase CCS values. This observation is somewhat intuitive in that carbohydrates have considerable degrees of freedom and can adopt both linear and branched primary structures. In contrast, lipids exhibit the largest CCS values among the biomolecules investigated, and this observation appears to be intrinsic to the inability of lipids for forming compact, self-solvated structures in the gas phase. Noteworthy among these findings is that despite the significant differences between helium and nitrogen in terms of mass, degrees-of-freedom (atomic vs diatomic), and polarization, the biomolecular class trends observed here for the nitrogen-based ion mobility are qualitatively the same as those previously observed in helium.[3,20] We do observe evidence that these qualitative trends between the two drift gases are not retained at low mass, and a more detailed investigation of helium and nitrogen-based ion mobility studies for low mass analytes (less than 200 Da) will be the subject of future studies.

We emphasize that these studies are only possible by the remarkable advances made over the past decade in the development of biological IM-MS instrumentation. The IM-MS described in this report can achieve high resolving powers with high sensitivity, making it possible to observe and characterize low abundance isomeric species in highly complex samples with unprecedented scale and throughput. While we have purposely chosen to report only the highest abundant species, we note that the observation of multiple ion mobility peak features (*i.e.*, structural and positional isomers) is routine with this instrumentation. As the analytical capabilities of distinguishing low-abundance isomeric species become widely accessible, we begin to move toward a new paradigm whereby it no longer becomes the question of if a particular isomer exists but rather how much of it is present and in what context.

ASSOCIATED CONTENT

Supporting Information

Empirically measured transport properties for the analytes evaluated in this work (Tables S1−S4). A summary of the overlapping helium and nitrogen CCS measurements compared in this study (Table S5). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: john.a.mclean@vanderbilt.edu.
*E-mail: john_fjeldsted@agilent.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support for this research to Vanderbilt University authors was in part from the NIH National Center for Advancing Translational Sciences (UH2TR000491); the Defense Threat Reduction Agency (HDTRA1-09-1-00-13 and DTRA100271 A-5196); the Defense Advanced Research Projects Agency (W911NF-12-2-0036); the Vanderbilt Institute of Chemical Biology; and the Vanderbilt Institute for Integrative Biosystems Research and Education. N.M.L. acknowledges a Vanderbilt Chemical Biology Interface training grant (T32GM065086). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies and organizations.

REFERENCES

(1) McDaniel, E. W. *Collision phenomena in ionized gases*; Wiley: New York, 1964; Vol. *1*.
(2) McDaniel, E. W.; Mason, E. A. *Mobility and diffusion of ions in gases*; John Wiley and Sons, Inc., New York, 1973.
(3) Fenn, L.; McLean, J. *Anal. Bioanal. Chem.* **2008**, *391*, 905−909.
(4) McLean, J. A. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1775−1781.
(5) Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1188−1211. Valentine, S. J.; Counterman, A. E.; Hoaglund-Hyzer, C. S.; Clemmer, D. E. *J. Phys. Chem. B* **1999**, *103*, 1203−1207.
(6) Shvartsburg, A. A.; Siu, K. W. M.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 885−888.
(7) Shah, A. R.; Agarwal, K.; Baker, E. S.; Singhal, M.; Mayampurath, A. M.; Ibrahim, Y. M.; Kangas, L. J.; Monroe, M. E.; Zhao, R.; Belov, M. E.; Anderson, G. A.; Smith, R. D. *Bioinformatics* **2010**, *26*, 1601−1607.
(8) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401−2414. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. *Int. J. Mass Spectrom.* **2007**, *261*, 1−12.
(9) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. *Anal. Chem.* **2010**, *82*, 9557−9565.
(10) Williams, J. P.; Grabenauer, M.; Holland, R. J.; Carpenter, C. J.; Wormald, M. R.; Giles, K.; Harvey, D. J.; Bateman, R. H.; Scrivens, J. H.; Bowers, M. T. *Int. J. Mass Spectrom.* **2010**, *298*, 119−127.
(11) Revercomb, H. E.; Mason, E. A. *Anal. Chem.* **1975**, *47*, 970−983.
(12) Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; John Wiley & Sons: New York, 1988; p 560.
(13) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**, *310*, 1658−1661. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protoc.* **2008**, *3*, 1139−1152. Knapman, T. W.; Berryman, J. T.; Campuzano, I.; Harris, S. A.; Ashcroft, A. E. *Int. J. Mass Spectrom.* **2010**, *298*, 17−23.
(14) Ridenour, W. B.; Kliman, M.; McLean, J. A.; Caprioli, R. M. *Anal. Chem.* **2010**, *82*, 1881−1889.
(15) Shvartsburg, A. A.; Mashkevich, S. V.; Siu, K. W. M. *J. Phys. Chem. A* **2000**, *104*, 9448−9453. Shvartsburg, A. A.; Hudgins, R. R.; Dugourd, P.; Jarrold, M. F. *Chem. Soc. Rev.* **2001**, *30*, 26−35. Larriba, C.; Hogan, C. J. *J. Phys. Chem. A* **2013**, *117*, 3887−3901.
(16) Campuzano, I.; Bush, M. F.; Robinson, C. V.; Beaumont, C.; Richardson, K.; Kim, H.; Kim, H. I. *Anal. Chem.* **2011**, *84*, 1026−1033.
(17) Pagel, K.; Harvey, D. J. *Anal. Chem.* **2013**, *85*, 5138−5145.
(18) Haynes, W. M.; Lide, D. R.; Bruno, T. J. *CRC Handbook of Chemistry and Physics 2012−2013*, 93 ed.; CRC Press: Boca Raton, FL, 2012.

(19) Baker, E. S.; Clowers, B. H.; Li, F.; Tang, K.; Tolmachev, A. V.; Prior, D. C.; Belov, M. E.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1176−1187.
(20) Fenn, L.; Kliman, M.; Mahsut, A.; Zhao, S.; McLean, J. *Anal. Bioanal. Chem.* **2009**, *394*, 235−244.
(21) Tao, L.; McLean, J. R.; McLean, J. A.; Russell, D. H. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1232−1238.
(22) Bush, M. F.; Campuzano, I. D. G.; Robinson, C. V. *Anal. Chem.* **2012**, *84*, 7124−7130.
(23) Salbo, R.; Bush, M. F.; Naver, H.; Campuzano, I.; Robinson, C. V.; Pettersson, I.; Jørgensen, T. J. D.; Haselmann, K. F. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 1181−1193.
(24) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Subramaniam, S. *Nucleic Acids Res.* **2007**, *35*, D527−D532.
(25) Wilkins, M. R.; Lindskog, I.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Hochstrasser, D. F.; Appel, R. D. *Electrophoresis* **1997**, *18*, 403−408.
(26) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, *246*, 64−71.
(27) Ibrahim, Y.; Tang, K.; Tolmachev, A. V.; Shvartsburg, A. A.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1299−1305.
(28) Ibrahim, Y.; Belov, M. E.; Tolmachev, A. V.; Prior, D. C.; Smith, R. D. *Anal. Chem.* **2007**, *79*, 7845−7852.
(29) Kemper, P. R.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 197−207. von Helden, G.; Hsu, M.-T.; Kemper, P. R.; Bowers, M. T. *J. Chem. Phys.* **1991**, *95*, 3835−3837.
(30) Mason, E. A.; Schamp, H. W. *Ann. Phys.* **1958**, *4*, 233−270.
(31) McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *Int. J. Mass Spectrom.* **2005**, *240*, 301−315.
(32) West, G. B.; Brown, J. H.; Enquist, B. J. *Science* **1997**, *276*, 122−126.
(33) Shen, Y.; Greene, J. E.; Strauss, M. A.; Richards, G. T.; Schneider, D. P. *Astrophys. J.* **2008**, *680*, 169.
(34) Galilei, G. *Dialogues Concerning Two New Sciences, translated by Henry Crew and Alfonso De Salvio*; Macmillan: New York, New York, 1914.
(35) Viidanoja, J.; Sysoev, A.; Adamov, A.; Kotiaho, T. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 3051−3055.
(36) Fenn, L. S.; McLean, J. A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2196−2205.
(37) Huang, Y.; Dodds, E. D. *Anal. Chem.* **2013**, *85*, 9728−9735.
(38) Ellis, H. W.; Pai, R. Y.; McDaniel, E. W.; Mason, E. A.; Viehland, L. A. *At. Data Nucl. Data Tables* **1976**, *17*, 177−210. Ellis, H. W.; McDaniel, E. W.; Albritton, D. L.; Viehland, L. A.; Lin, S. L.; Mason, E. A. *At. Data Nucl. Data Tables* **1978**, *22*, 179−217. Ellis, H. W.; Thackston, M. G.; McDaniel, E. W.; Mason, E. A. *At. Data Nucl. Data Tables* **1984**, *31*, 113−151. Viehland, L. A.; Mason, E. A. *At. Data Nucl. Data Tables* **1995**, *60*, 37−95.
(39) Matz, L. M.; Hill, H. H., Jr.; Beegle, L. W.; Kanik, I. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 300−307.
(40) Berant, Z.; Karpas, Z. *J. Am. Chem. Soc.* **1989**, *111*, 3819−3824.
(41) Wyttenbach, T.; Bleiholder, C.; Bowers, M. T. *Anal. Chem.* **2013**, *85*, 2191−2199. von Helden, G.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass Spectrom. Ion Processes* **1995**, *146−147*, 349−364.