# Constraints on isomers of dissolved organic matter in aquatic environments: Insights from ion mobility mass spectrometry

Kaijun Lu[1], Xiaolin Li[2], Hongmei Chen[2] and Zhanfei Liu[1*]

[1]Marine Science Institute, The University of Texas at Austin, Port Aransas, Texas, USA

[2]State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China

*Corresponding author

**Abstract**

Elucidating the chemical structure of dissolved organic matter (DOM) is key to understanding this large yet enigmatic carbon pool. Over the last two decades much progress has been made in assigning exceptionally accurate molecular formulas owing to the application of ultrahigh resolution mass spectrometry, but little is known about the number of isomers in each molecular formula, a question essentially related to the total number of organic molecules in DOM. Such information is critical for a further understanding of the formation and long-term stability of refractory DOM in the ocean. In this work, we used ion mobility quadrupole time of flight liquid chromatography tandem mass spectrometry (IM Q-TOF LC/MS/MS) to analyze DOM samples collected in different aquatic environments including south Texas rivers, Gulf of Mexico, and South China Sea. Our data showed that generally less than 23% of all detected DOM formulas, which shared a small fraction of common molecules (ca. 12%) detected by direct-infusion Fourier transform ion cyclotron resonance MS, contained structurally distinct isomers (represented by "isomer clusters"). In addition, isomer diversity, in terms of how similar the structures are, decreased with degradation in both natural and incubation samples. Specifically, the number of structurally distinct isomers was lower at river mouth and open ocean than coastal waters, where organic matter tends to be fresher due to high primary production; and with depth in water column, isomer diversity of DOM also decreased. Results from a set of incubation experiments also showed that the percentages of riverine DOM molecules that have multiple isomer clusters decreased with time, suggesting that biodegradation decreases the diversity of molecules from an isomer perspective. Overall, these results suggested that isomers, at least in a certain fraction of DOM, are highly constrained, and that degradation decreases its isomeric diversity.

## 1. Introduction

As one of the most important reduced carbon pools on the Earth, dissolved organic matter (DOM) in the ocean contains $660 \times 10^{15}$ g of carbon, similar in magnitude to atmospheric carbon and over 1000 times more than all living organisms in the oceans combined (Hansell et al., 2009). The active carbon of this pool is constantly exchanging with atmospheric carbon through various processes such as photosynthesis and respiration, thus even minor disturbances of the DOM pool could impact global climate substantially (Hedges, 1992). Given its potential climatic significance, there is a need to characterize DOM chemically, yet less than 30% of DOM in surface ocean and only about 5% in deep water, based on the carbon content, has been identified at molecular level (Benner, 2002). Also unclear is why a major fraction of this reduced, low molecular weight (≤ 1000 Da), organic carbon is stable for millennial time scales in an environment enriched with electron acceptors and microbes (Dittmar and Kattner, 2003; Dittmar and Paeng, 2009; Dittmar, 2015).

Two major hypotheses have been proposed to explain the paradox of DOM persistence in the ocean. The intrinsic stability hypothesis links the reactivity of DOM to its molecular structure, i.e., the persisting molecules have intrinsic recalcitrant structures that prevent degradation. These structures may be biosynthesized by organisms, or originated from secondary abiotic/biotic modifications (Shen and Benner, 2019). Alternatively, the dilution hypothesis, or molecular diversity hypothesis, states that the seemingly recalcitrant nature of DOM is a result of extremely diluted concentrations of each individual DOM molecule (e.g., Arrieta et al., 2015), because decomposition rate of an organic substrate is hypothesized to be a function of concentration (e.g., Barber, 1968; Kattner et al., 2011; Dittmar, 2015). Although results to support either hypothesis can be found, the evidence is far from being conclusive and both hypotheses have their own caveats (Dittmar, 2015). For the intrinsic stability hypothesis, though several recalcitrant molecular structures have been proposed, it still remains unclear the manner in which these known structures are connected to the total refractory DOM pool, and how much the inherent recalcitrance contribute to the preservation of DOM. For example, the exact mechanism of the persistence of carboxyl-rich alicyclic molecules (CRAM; Hertkorn et al., 2006), as well as material derived from linear terpenoids (MDLT; Lam et al., 2007; Arakawa et al., 2017), remains elusive, as it might come from its alicyclic rings and branching structure (Hertkorn et al., 2006), or from the

formation of gel due to metal-ligand binding (Chin et al., 1998). A rather similar scenario applies for the dilution hypothesis. With limited information on the total number of DOM compounds, as well as a lack of knowledge in the diversity of DOM molecules (Mentges et al., 2017), the concentration of each individual molecule, which is arguably the most crucial parameter in the hypothesis, remains unclear. Therefore, revealing molecular diversity of DOM is an important step in explaining the long-term stability of this large pool of reduced carbon.

Molecular formula data generated by ultrahigh-resolution mass spectrometry such as Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) and Orbitrap MS have been widely applied to calculate molecular diversities of DOM (Hawkes et al., 2016), including the molecular formula richness (i.e., the total number of different molecular formulas; e.g., Schmidt et al., 2009; Roth et al., 2015), abundance-based formula diversity (e.g., Kellerman et al., 2014; Seidel et al., 2015), or formula-based functional diversity (Mentges et al., 2017). Previous studies showed that molecular formula richness (e.g., number of formulas assigned) tends to increase with biotic/abiotic degradation (e.g., Lechtenfeld et al., 2015; Osterholz et al., 2015; Mentges et al., 2017), but molecular diversity (e.g., number of compound types) tends to decrease with degradation, reflected by an increase in homogeneity in molecular formula for more degraded DOM (Mentges et al., 2017; O'Connor et al., 2020).

Even though ultrahigh-resolution mass spectrometers can resolve up to 20,000 molecular formulas from the intricate mixture of DOM sample (Riedel and Dittmar, 2014), the isomeric information of DOM molecules remains unclear. A few studies have investigated DOM isomers, mainly through mathematical calculation (Hertkorn et al., 2007; Hertkorn et al., 2008) and/or modelling (Zark et al., 2017; Hawkes et al., 2018; Zark and Dittmar, 2018) based on MS or tandem MS results. The estimated number of isomers ranged from over a hundred per formula (Zark et al., 2017; Hawkes et al., 2018) to several million, assuming that the arrangement of atoms (i.e., C, H, N, O) is random as long as the structure is chemically reasonable (Hertkorn et al., 2007; Dittmar and Stubbins, 2014). However, the disagreement between the unimodal molecular weight distribution of DOM (maximum intensity at an *m/z* of ~400) and the exponentially increasing number of isomers with molecular weight, as pointed out by a previous review (Dittmar and Stubbins, 2014), implies

99 that the arrangement of atoms is not totally random, and that isomers in natural DOM are
100 under certain constraints. Therefore, there is a critical need to obtain an actual measurement
101 of isomers in natural DOM.
102 Ion mobility mass spectrometry (IM-MS) can directly quantify structural and
103 isomeric aspects of DOM molecules. In drift-time IM-MS (DTIM-MS), ions are separated in
104 the IM chamber through low-energy interactions with inert buffer gas ($N_2$ or He) during the
105 transit (May et al., 2014; May and McLean, 2015), as ions with larger collision cross section
106 (CCS) areas are delayed with longer drift times than smaller ones due to more frequent
107 collisions with the buffer gas molecules (McDaniel, 1964; McDaniel and Mason, 1973; Fenn
108 et al., 2009). Therefore, molecular-level structural information in the form of CCS values can
109 be obtained to facilitate the characterization of targeted molecules and to differentiate
110 isomers (May et al., 2014). IM-MS has been applied in metabolomics studies to resolve
111 structure and isomers of targeted biomolecules (e.g., McLean et al., 2005; Baker et al., 2007;
112 Tao et al., 2007; Fenn et al., 2009; Stow et al., 2015, 2017; Hines et al., 2016). A pioneering
113 study in the field of natural DOM has demonstrated the power of IM-MS to unravel the
114 existence of multiple charged compounds in standard Suwannee River Fulvic Acids (SRFA;
115 Gaspar et al., 2009). With recent commercialization of trapped IM-MS (TIM-MS; Leyva et
116 al., 2019) and DTIM-MS (Lanucara et al., 2014; May and McLean, 2015), and the
117 application of IM calibration standards (e.g., Agilent Tuning Mix), current IM-MS offers a
118 standardized measurement with high resolving power and precision (Hines et al., 2016),
119 making cross-lab comparisons possible (Stow et al., 2017a). Following this advancement,
120 IM-MS was first successfully applied to elucidate the isomer diversity of water
121 accommodated fractions of crude oil (Benigni et al., 2017), and quickly recognized in the
122 DOM field (Tose et al., 2018; Lu et al., 2018; Lu and Liu, 2019; Leyva et al., 2019; Gao et
123 al., 2019; Leyva et al., 2020). For instance, previous work demonstrated that in riverine and
124 wetland systems only a small fraction of the detected DOM formulas (ca. 15%) possesses
125 isomers, and the number of isomers for each formula is generally less than 10 (Lu et al.,
126 2018; Lu and Liu, 2019; Morrison et al., 2020), even though the detected formulas may
127 represent only a small fraction of the total DOM pool. Each peak detected in IM-MS can be
128 further resolved into several sub-peaks mathematically by applying a Software Assisted
129 Molecular Elucidation package, which utilizes noise removal, mean gap filling, “asymmetric

least squares smoothing" baseline correction, peak detection by continuous wavelet transform, as well as Gaussian fitting with nonlinear least-squares functions (Benigni et al., 2017). Based on this calculation, the number of isomers for each detected riverine DOM formula (e.g., SRFA Standard or Paraguay River DOM) was estimated to range from 6 to 10 (Tose et al., 2018; Leyva et al., 2019). However, current studies are isolated, thus there is a need for a systematic evaluation of the isomers in DOM, especially how isomers in DOM change in response to environmental parameters (e.g., degradation status of DOM) and in marine DOM including that in the deep ocean.

The goals of this study are to: (1) elucidate isomeric information of natural DOM in broader aquatic environments, and (2) relate the isomeric information to environmental degradation. We report the results of isomers in DOM using an online combination of DTIM-MS, high-performance liquid chromatogram (HPLC), and Quadrupole Time-Of-Flight mass spectrometry (Q-TOF MS), and have compared the results with ultrahigh resolution mass spectrometer FT-ICR MS. Natural DOM samples were collected from several different environments, including rivers, coastal ocean, surface and deep waters of open ocean, as well as from bioassays. Specifically, we aim to determine how isomers in DOM change in different environments and with environmental degradation.

## 2. Materials and Methods

### 2.1. Sample Collection

Field samples were collected mainly from northern Gulf of Mexico (GOM) in July 2017 (Figure 1). Specifically, river samples were collected from the ship channel of Atchafalaya River at three sites (ARSC1, ARSC2, and ARSC3) following a salinity gradient from ARSC2 (0.3 ‰) to ARSC1 (3.0 ‰), and to ARSC3 (15.8 ‰). Surface water was collected at stations F1 (N 29.15°, W 91.62°), F2 (N 29.00°, W 91.62°), and F3(N 28.88°, W 91.62°) along the F-transect (Rabalais et al., 2001), respectively. Open ocean samples were collected from three depths (0.5 m, 500 m, and 1240 m) at station D11 (N 27.54°, W 90.83°; D11 0.5, D11 500, and D11 1240). In addition to the GOM field samples, a depth profile of samples at station DC6 of South China Sea (N 15.24°, E 114.85°; 5 m, 100 m, 500 m, 1000 m, 2000 m, 3000 m, and 4000 m) was also collected.

Coastal and open ocean water samples from GOM were collected using Niskin bottles mounted on a conductivity-temperature-depth (CTD) rosette (Seabird 911) and transferred to

acid-washed 10 L polyethylene carboys. River water was directly pumped with an underway system (e.g., Zhai et al., 2005) on *R/V Pelican*. Upon collection, 10 L samples were immediately filtered on board through in-series filter cartridges (Whatman™ Polycap HD 75 Capsule Filter, 5.0 µm, 2712M; Whatman™ Polycap AS 75 Capsule filter, 0.2 µm, 2706T) to remove suspended particles. An aliquot of the filtered water (30 mL) was preserved in 30-mL polycarbonate bottles at -20 °C until further DOC concentration analysis. Following the established protocol (Dittmar et al., 2008), DOM in the filtered water was acidified with HCl to a final pH of 2, and was extracted by solid-phase extraction (SPE) using PPL cartridges (Agilent Bond Elut PPL cartridge, 500 mg, P/N 12105006; five cartridges were used for one 10 L-sample). DOM extracts were eluted with methanol (four cartridge volumes for each cartridge), re-concentrated to a final volume of 10 mL, and stored at -20 °C after returning to lab until further analysis. The extracts were diluted 10-fold with methanol before LC/MS analysis. The carbon-based recovery of DOM extraction, calculated as DOC of extract/DOC of filtered water × 100%, ranged from 45% to 48%. DOC concentrations were measured with a Shimadzu total organic carbon analyzer (TOC-L), with organic carbon stock solution (RICCA) as standards, and the calibration curves were independently checked with low carbon water, surface water, and deep water references from Hansell Lab, University of Miami. All solvents and chemicals were LC/MS grade (Fisher or Sigma Aldrich).

A bioassay experiment was conducted in July 2016 (details can be found in parallel experiments of Wu et al., 2019). Briefly, 1 L river water was collected from lower-river region of San Antonio River (SR; N 28.48°, W 96.86°) and Lavaca River (LR; N 28.83°, W 96.58°) using a Van Dorn sampler (Figure S1), and the sample was transferred to high-density polyethylene bottles (4 L) that were first rinsed several times with the sample water. Samples were stored on ice and transported to lab the same day, where they were filtered through 0.2 µm Capsule filter (Whatman™ Polycap AS 75) to remove particles and microbes, and inoculated with 50 mL of natural coastal bacterial assemblages (30 ‰; pre-filtered through 0.8 µm pore size filter to remove algae and large eukaryote predators) from Port Aransas ship channel. The inoculation ratio (1/20) in this work was similar to other incubation experiments (Rochelle-Newall et al., 2004; Petrone et al., 2009), and was sufficient to support the development of bacteria community (Wu et al., 2019). The incubation lasted for 3 weeks at close to *in situ* temperature (31 °C) in the dark. Pre- and

post-incubation samples (500 mL) were collected and filtered through 0.2 μm pore size filters (Whatman™ 47 mm 0.2 μm membrane filter, 7402-004) prior to SPE. The carbon-based SPE recovery ranged from 44% to 52%.

## 2.2. Molecular level characterization of DOM

DOM was analyzed using an Ion Mobility Quadrupole Time of Flight Liquid Chromatography Mass Spectrometer (IM Q-TOF LC/MS, Agilent 6560) with an orthogonal electrospray ionization (ESI) source. Electrospray ionization negative mode (ESI-) has been commonly applied in the molecular characterization of complex DOM mixtures (e.g., Kim et al., 2003; Kujawinski, 2002; Wagner et al., 2019), which contain abundant molecules with carboxyl groups that can be ionized efficiently in negative mode. However, the presence of high carboxyl-content aromatic compounds, especially in terrestrial DOM (Kramer et al., 2012), can suppress ionization of aliphatic and carbohydrate-like DOM (Ohno et al., 2016). Therefore, both ESI- and ESI+ modes were applied for Q-TOF in this work, aiming at providing a more comprehensive picture of natural DOM.

The instrument and data acquisition methods followed our previous studies (Lu et al., 2018; Lu and Liu, 2019), with HPLC settings kept constant for MS, MS/MS, and IM-MS acquisitions. Briefly, mobile phase A was $H_2O$ with 0.1% (v/v) formic acid, and B was acetonitrile for ESI+ mode. One-μL of sample or 10-μL of standards (DOM standards and peptide standard) was eluted through a StableBond $C_{18}$ column (Poroshell 120 SB-C18; 2.7 μm, 2.1 × 100 mm; Agilent P/N 685775-902) at a flow rate of 0.5 $mL \cdot min^{-1}$. During the 21-min run, mobile phase B was increased from 3% to 90% in the first 15 min, held at 90% from 15 min to 20 min, and then dropped to 3% at 21 min. A post-run time of 4 min allowed the column to reach equilibrium. For ESI- mode, mobile phase A was $H_2O$ with 10 $mmol \cdot L^{-1}$ ammonium acetate, and B was acetonitrile. One-μL of samples or 10-μL of DOM standards was eluted through a HILIC column (2.7 μm, 15 cm × 4.6 mm SUPELCO, P/N 53981-U) at a flow rate of 0.5 $mL \cdot min^{-1}$. During the 10-min run, mobile phase B was held at 98% during the first 1 min, and dropped to 95% from 1 min to 10 min. A post-run period of 15 min allowed the column to reach equilibrium before the injection of next sample. For each sample (field samples and incubation samples) a total amount of ca. 0.1 – 0.2 μg C was injected during the analysis. For the DOM standards ca. 5 μg C was injected, and for the peptide isomer mixture ca. 0.02 μg C was injected for either peptide. An injection volume of 10 μL

was used for DOM standards in order to achieve a better ionization, as DOM standards had a much lower ionization efficiency in our test runs when compared with SPE-extraction samples, even though a much higher amount of materials was injected (Figure S2).

Mass spectrum (MS), tandem MS (MS/MS) and ion mobility-mass spectrum (IM-MS) data were acquired with software MassHunter LC/MS Data Acquisition (Version B.07.00 and B.09.00) in both ESI- and ESI+ modes. The resolving power of drift time ion mobility is calculated as $t/\Delta t$, in which $t$ is the drift time and $\Delta t$ is the peak width measured in milliseconds at half peak height. The estimated resolving power of ion mobility alone exceeds 60 at drift time of 30 ms. Combined with ion mobility, resolving power of the whole instrument (including LC, IM, and MS) can be over 2,000,000 (Lu et al., 2018). Each sample was analyzed in duplication in this work.

The acquisition for MS, MS/MS and IM-MS followed the published protocol (Lu et al., 2018; Lu and Liu, 2019). For ESI+ MS data acquisition, the orthogonal electrospray ionization source (Dual Agilent Jet Stream ESI) was operated with $N_2$ sheath gas temperature of 350 °C at a flow rate of 12 $L \cdot min^{-1}$. $N_2$ drying gas applied at the source entrance was at the temperature of 225 °C, and maintained at a flow rate of 13 $L \cdot min^{-1}$ with a nebulizer pressure of 45 psig. The source operated in positive mode was set to a VCap voltage of 3500 V, and a nozzle voltage of 0 V. Q-TOF was in positive ion polarity in MS mode, with an MS mass range of 70-1200 *m/z*, and an acquisition rate of 1 spectrum/s. Reference masses of 121.050873 and 922.009798 (Agilent Tuning Mix) were used for mass calibration. For ESI- MS data acquisition, the ion source was also the Dual AJS ESI. Sheath gas parameters were the same as in ESI+ mode. Drying gas temperature was also 225 °C, but the flow rate was adjusted to 5 $L \cdot min^{-1}$ and nebulizer pressure was maintained at 20 psig. The source operated in negative mode was maintained at 3500 V VCap voltage, but the nozzle voltage was raised to 2000 V. Q-TOF settings of ESI- MS resembled those of ESI+ mode, except that Q-TOF was operated under negative ion polarity, with reference masses of 112.985587 and 1033.988109 (Agilent Tuning Mix). The acquisition settings for MS/MS, including ion source settings and reference ions, generally followed those of MS, except for the instrument operated under MS/MS mode, with an MS/MS acquisition time of 500 ms/spectrum. A preferred ID list containing the information of compounds *m/z* and retention time was applied to guide each sample run in MS/MS. The threshold to trigger MS/MS was at least 2000

counts for the precursor ion under ESI+ mode, and 6000 counts under ESI- mode. The collision energy was 30 V for ESI+, and 40 V for ESI-.

For ESI+ IM-MS data acquisition, Q-TOF was operated under IM Q-TOF Acquisition mode. General parameters of IM followed those described in ESI+ MS. Funnel parameters were maintained at default voltages. For ESI- IM-MS, general parameters of IM also followed those described for ESI- MS. Most IM funnel parameters were default voltages except for a few optimizations: high pressure funnel RF of -150 V, trap funnel RF of -125 V, trap entrance grid delta of -4 V, trap exit grid 1 delta of -6 V, trap exit grid 2 delta of -12 V, drift tube entrance voltage of -1500 V, and rear funnel RF of -100 V.

Data analysis of IM-MS also followed the published protocol (Lu et al., 2018; Lu and Liu, 2019). Specifically, IM-MS data was analyzed with MassHunter IM-MS Browser (Version B.07.01 and B.09.00). The *m/z* of features detected were calibrated with IM-MS Data File Reprocessing Utility (Version B.08.00). Drift times, representing the total transit time of the ions (May et al., 2014), were acquired using the function "Find Features (IMFE)" with the following settings: isotope model common organic (no halogens); charge state ≤ 1. Compound drift times were calibrated with reference ions (Tune Mix, G2421A, Agilent Technologies, Santa Clara, CA) to determine the gas-phase momentum transfer collision cross section (CCS; Mason and Schamp, 1958; May et al., 2014), which can be used as a proxy of the three-dimensional (3D) structure of the molecule. IM-MS provides insights into structural isomers of DOM molecules and thus a direct measurement of structural diversity (Benigni et al., 2017; Lu et al., 2018; Lu and Liu, 2019; Leyva et al., 2019). Features with same *m/z* but different LC retention time (RT) and/or different CCS are characterized as different isomers clusters (Lu et al., 2018; Lu and Liu, 2019). This DTIM-MS can differentiate isomers with difference in CCS as low as 2% at the resolving power of 60 (Nichols et al., 2018). A major loss in ions, ~90%, was observed during IM-MS data acquisition when compared to the acquisition of MS data. This loss was attributed to the interactions of molecules with nitrogen gas during the transit in drift tube (Lu et al., 2018; Lu and Liu, 2019).

To achieve a better understanding of the separation processes of LC and drift-time IM-MS (LC-DTIM-MS), physical properties of the modeled DOM molecules were calculated with software Marvin Sketch (Version 19.20.0). Specifically, the octanol-water

partition coefficient (log P), which is often used as a measurement of molecular hydrophobicity, was calculated based on Viswanadhan et al. (1989). The maximal and minimal projection areas of a molecule, as a proxy representing CCS, were calculated based on the van der Waals radius of the modeled DOM structure. In addition, the van der Waals surface area of the modeled molecule was calculated based on Ferrara (2002).

### 2.3. Validation of the molecular level characterization method

Concerns have been raised that analytical instruments differing in performances, such as resolving powers and source conditions, and operating conditions, would result in inconsistent results regarding the molecular level composition of DOM (Hawkes et al., 2020). Therefore, there is a need to investigate the difference in DOM assessments between the IM Q-TOF LC/MS and other instruments on the basis of a same set of standards or samples.

To compare with the published results from other high resolution mass spectrometers (e.g., Patriarca et al., 2018; Kim et al., 2019; Hawkes et al., 2020), DOM standards, Suwannee River Humic Acid (SRHA, 3S101H) and Suwannee River Natural Organic Matter (SRNOM, 2R101N) purchased from International Humic Substances Society (IHSS), were analyzed by IM Q-TOF LC/MS. Briefly, approximately 10 mg of DOM standards were dissolved in LC/MS grade $H_2O$-methanol solvent (50/50, v/v), resulting in a final concentration of ca. 500 ppm C, given the known elemental composition of the standards (Green et al., 2014). Note that the concentrations of DOM standards were much higher than typical direct-infusion setups (e.g., ca. 20 ppm in Hawkes et al., 2020), but were on a par with those used in LC/MS studies (e.g., ca. 200 ppm in Kim et al., 2019; and 500 ppm in Patriarca et al., 2018).

In addition to DOM standards, the DC6 sample set (collected from 7 depths at station DC6 in South China Sea) was run by both IM Q-TOF LC/MS and FT-ICR MS for a direct comparison. Specifically, the FT-ICR data was acquired from a 12 Tesla Apex Qe FT-ICR MS at Old Dominion University, and the instrument settings generally followed the published protocol (Liu et al., 2011; Waggoner et al., 2015), with samples continuously infused into the FT-ICR MS at a flow rate of 120 $\mu L \cdot h^{-1}$ by a syringe pump. Ions accumulated in a hexapole for 0.5 s before being transferred to the ICR cell. Three hundred transients, collected with a 4 MWord time domain, were added, giving about a 30 min total

316 run time. The summed free induction decay (FID) signal was zero-filled once and Sine-Bell
317 apodized prior to fast Fourier transformation and magnitude calculation using the Bruker
318 Daltonics Data Analysis software. Only negative ion mode data was used in this work. The
319 formula assignment also followed the rules described previously (Liu et al., 2011; Waggoner
320 et al., 2015). The MS results from FT-ICR MS were compared with the MS, MS/MS, and
321 IM-MS results from IM Q-TOF LC/MS to find the common features.
322 Finally, a pair of peptides, Serine-Aspartic Acid-Glycine-Arginine-Glycine (SDGRG)
323 and Glycine-Arginine-Glycine-Aspartic Acid-Serine (GRGDS) purchased from Sigma
324 Aldrich, was analyzed as isomer standards to validate the resolving ability for isomers of the
325 IM Q-TOF LC/MS. One-mg of the peptide standard was reconstituted in 10 $mmol \cdot L^{-1}$
326 ammonium acetate solution (in LC/MS grade $H_2O$) to a concentration of 1000 $\mu g \cdot L^{-1}$. The
327 working solution of each peptide was further diluted and mixed to create an isomer standard
328 with a final concentration of 50 $\mu g \cdot L^{-1}$ for each peptide. The isomer standard was analyzed
329 under ESI+ mode, consistent with published protocol (May et al., 2014).
330 **3. Results**
331 3.1. Evaluation of IM Q-TOF LC/MS.
332 The ability of IM Q-TOF LC/MS to process DOM samples was first assessed through
333 a direct comparison with the classic FT-ICR MS on the same set of samples, as well as
334 through analyses of DOM and isomer standards. The direct comparison with FT-ICR MS
335 was made under the ESI- mode on the same set of samples from South China Sea. On
336 average, over 3100 unique formulas were assigned with FT-ICR MS, as compared to ca.
337 1500 with IM Q-TOF LC/MS (Table S1). The number of shared formulas ranged from 318 to
338 436 across different samples, representing ca. 12% and ca. 25% of FT-ICR MS and IM Q-
339 TOF LC/MS formulas, respectively. These shared molecules were mainly located in a region
340 known as “island of stability” (Lechtenfeld et al., 2014; Koch et al., 2014) in the van
341 Krevelen diagrams (Figure 2). These shared molecules, though only representing a small
342 fraction of total detected features with respective techniques, verify the capability of
343 molecular level measurement by IM Q-TOF LC/MS, and further suggest that the analytical
344 window of IM Q-TOF LC/MS was different from commonly used ultrahigh resolution MS,
345 such as FT-ICR MS.

As for DOM standards, the LC chromatogram showed evident difference between SRHA and SRNOM (Figure S3). On average, 539 and 583 unique formulas were assigned for SRHA and SRNOM under ESI+ mode, respectively, and 1219 and 1154 unique formulas were assigned for SRHA and SRNOM under ESI- mode, respectively (Figure S4). By comparing with results from direct-infusion ultrahigh resolution mass spectrometry including FT-ICR MS and Orbitrap MS (e.g., Hertkorn et al., 2013; Hawkes et al., 2016, 2020; Spencer et al., 2016), a similar pattern was revealed: even though the assigned formulas covered some common regions on the van Krevelen diagrams, the overall pattern from IM Q-TOF LC/MS was still different from direct-infusion MS. Combined with the aforementioned direct comparison results, the most striking differences between IM Q-TOF LC/MS and direct-infusion MS from the perspective of the van Krevelen diagram, included: (1) a less dense cluster of molecules in the center of the diagram (H/C ratio of ca. 1.0 – 1.3 and an O/C ratio of ca. 0.4 – 0.6; the "island of stability" region) from IM Q-TOF LC/MS, and (2) molecules occupying the upper and lower left region of the diagram (H/C ratio > 1.5 and/or O/C ratio < 0.5) were present from IM Q-TOF LC/MS results, but mostly absent from direct-infusion MS results.

These differences in the assignment of DOM molecules can be partially attributed to the addition of an online LC system. LC provides a preliminary chromatographic separation prior to MS analysis, which greatly alleviates the ionization suppression issue during direct-infusion analyses (Kim et al., 2019), and thus could reveal compounds that may be otherwise occluded in conventional studies. The chromatographic dilution may also explain the lower density of the island of stability. Interestingly, though the observed pattern of IM Q-TOF LC/MS resembled the result of an LC-TOF MS study (Rathgeb et al., 2017), it was still somewhat different from the most recent LC-ultrahigh resolution MS work (Patriarca et al., 2018; Kim et al., 2019), suggesting that other factors such as the difference in ionization techniques (e.g., the orthogonal electrospray used in Q-TOF) and the ways of measuring *m/z* (i.e., FT-ICR vs. Orbitrap vs. TOF) may have also played important roles. Nevertheless, compounds in the apolar region of the van Krevelen space (i.e., low O/C region) were also observed in the work of both Patriarca et al. (2018) and Kim et al. (2019), consistent with our results. Overall, the comparison with commonly applied direct-infusion MS further confirms that IM Q-TOF LC/MS provides a reliable, though quite different, window to look into the

intricate pool of DOM molecules. However, given the different analytical windows, cautions must be taken before extrapolating the presented result to the whole DOM pool.

The peptide mixture, consisting of two peptide isomers, was made to validate the resolving capability of IM Q-TOF LC/MS. Two features with an *m/z* of 491.2195 were detected under ESI+ mode, consistent with published protocol (May et al., 2014). One feature had a LC retention time (LC RT) of 0.533 min and a drift time (DT) of 25.50 ms, while the other feature had a LC RT of 0.537 min and a DT of 27.36 ms (Figure 3). The separation, by a DT difference of 7%, of two peptide isomers SDGRG (25.50 ms) and GRGDS (27.36 ms) is consistent with May et al. (2014), demonstrating the resolving power of IM-MS for isomer analysis.

### 3.2. Environmental parameters and MS results

Along the GOM salinity transect, dissolved organic carbon (DOC) concentrations ranged from 134.2 μmol C $L^{-1}$ at the offshore site (F3) to 364.2 μmol C $L^{-1}$ in the Atchafalaya river ship channel (ARSC1; Table S2). Dissolved oxygen (DO) concentrations remained relatively constant (207.5 – 221.3 μmol $O_2$ $L^{-1}$) for samples collected along the F-transect (Figure 1). In contrast, DO concentrations varied more with depth, dropping from 196.9 μmol $O_2$ $L^{-1}$ at the surface to 111.9 μmol $O_2$ $L^{-1}$ at 500 m, and back to 198.8 μmol $O_2$ $L^{-1}$ in bottom water (1240 m; Table S2). Though no DO data was collected for the Atchafalaya river ship channel due to the setup of the underway sampling system, it is reasonable to assume that the surface water samples were well oxygenated due to constant mixing and air sea exchange. For bioassay incubations, DOC concentrations decreased from 463.9 μmol C $L^{-1}$ to 448.7 μmol C $L^{-1}$ in Lavaca River incubation, and from 208.5 μmol C $L^{-1}$ to 197.2 μmol C $L^{-1}$ in San Antonio River incubation, as reported in previous work (Table S3; Wu et al., 2019). Vertical distribution of DOC at the station DC6 in the South China Sea showed a typical distribution pattern of open ocean with relative higher concentration in the surface (71.5 μmol C $L^{-1}$) and lower level at 4000 m (40.0 μmol C $L^{-1}$), where the water is originated from Philippine Sea in Northern Pacific (Li and Qu, 2006).

Number of assigned formulas ranged from 1238 to 3337 under ESI+ mode, and from 199 to 748 under ESI- mode (Table 1) for the GOM samples. For the incubation samples, the number of assigned formulas ranged from 477 to 577 under ESI+ mode, and from 236 to 462 under ESI- mode (Table 2). Generally, coastal and open ocean samples have higher H/C

ratios than river samples under both ESI modes (ESI+: 1.47 – 1.53 vs. 1.41 – 1.46, $p = 0.01$ single factor ANOVA test; ESI-: 1.11 – 1.19 vs. 1.07 – 1.12, $p = 0.001$ single factor ANOVA test; positive correlations between H/C and salinity were also verified with linear mixed model, with a $p$ value of 0.03 under ESI+ and 0.1 under ESI-; Table 1). In addition, H/C ratios were lower for samples from depth than those from surface (e.g., ca. 1.50 vs. 1.53 under ESI+, $p = 0.05$, t-test; Table 1). In the bioassay incubation, H/C ratios decreased after incubation (ESI+: 1.58 to 1.56 for LR, 1.59 to 1.55 for SR; Table 2), along with the decrease of DOC concentrations (Table S4). These results are consistent with previous findings that more refractory DOM molecules generally exhibit lower H/C ratios (e.g., D'Andrilli et al., 2015; Seifert et al., 2016).

Similar to DOM standards, the van Krevelen diagrams of the field and incubation samples from this work using Q-TOF LC/MS (Figures S5, S6, and S7) differed from those obtained by direct-infusion ultrahigh resolution MS. However, they agreed well with the DOM standards run by the same instrument (Figure S4), suggesting the consistency of the method used in this work. This further indicates that the molecules detected with Q-TOF LC/MS employed in the current work may represent a different pool of natural DOM than what are detected by other direct-infusion mass spectrometry techniques. Since the same technique was applied to all samples, the trend of elemental stoichiometry changes in DOM observed among different samples in this work are expected to be robust.

### 3.3. Changes in CCS of DOM across different sites.

In DTIM-MS, molecules of similar *m/z* but with higher charges and more compacted 3D structures will traverse faster through the drift tube, and vice versa. As the majority of extractable DOM molecules are singly charged (Leenheer et al., 2003; Rostad and Leenheer, 2004), the geometric configuration of the molecules, which determines collision section area (Collision Cross Section, CCS) values, is the sole factor affecting the drift time. CCS values of the DOM molecules generally increased with increasing *m/z* under both ESI modes (Figure 4). Overall, ca. 100 and 200 features per sample were detected in DTIM-MS under ESI- and ESI+ modes, respectively.

The molecular mass and *m/z* are reasonable approximations of molecular volume (Potts and Guy, 1992), and CCS can be considered as a descriptor of the surface area of a molecule. Thus, the relationship between CCS values and *m/z* is expected to be depicted in a

nonlinear power fit (May et al., 2014), as in the case of Van der Waals molecular volume and surface area (Moldoveanu and David, 2017). Under ESI+ mode, the relationship between CCS and *m/z* follows a power function of CCS = 21.794$(m/z)^{0.354}$ ($R^2$ = 0.774) for the composite of all samples (Figure 4A). When the samples were grouped based on sampling sites, the power function fit still holds (Figure S8). No obvious difference among sites was detected, with the exponents ranging from 0.332 to 0.381 and $R^2$ ranging from 0.736 to 0.805 (Figure S8; Table S5). Under ESI- mode, a similar power function relationship between CCS and *m/z* was found, but this relationship was better illustrated by two distinct groups of molecules, with one group possessing much lower CCS (Figure 4B). The upper group has a power-fit of CCS = 15.982$(m/z)^{0.383}$ ($R^2$ = 0.830), while the lower group has a function of CCS = 37.255$(m/z)^{0.156}$ ($R^2$ = 0.760). Similar to ESI+, no clear difference in the distribution was observed among different sites under ESI- mode (Figure S9; Table S6). It remains unclear why DOM molecules under ESI- mode were in two groups, but the MS analysis showed that the lower group of compounds in the CCS-*m/z* diagram had lower H/C ratios but elevated O/C ratios, suggesting that the unexpected lower CCS values under ESI- mode may be correlated with the degradation status of DOM (Lu et al., 2018; Lu and Liu, 2019). Compared with biomolecule standards of lipids, peptides and carbohydrates, which have an exponent of 0.47 – 0.60 (May et al., 2014), natural DOM molecules had a much lower exponent (ca. 0.35 under ESI+), suggesting their more compacted geometric configuration, consistent with our previous findings (Lu et al., 2018; Lu and Liu, 2019).

### 3.4. Changes in isomers of DOM across different sites.

Isomers, by definition, have the same *m/z* (same molecular formula) but different structures, which can be reflected in DTIM-MS. Isomers mainly include constitutional isomers (structural isomers) and stereoisomers (spatial isomers), and the stereoisomers can be further categorized into diastereomers and enantiomers. If the differences in structure are large enough (e.g., different functional groups), isomers can be separated by their different DTs, and/or LC RTs, and can be further visualized as different peaks in the CCS-*m/z* diagram along with the drift time chromatogram (Figures 4 and 5). However, stereoisomers, in particular enantiomers, could possess exactly the same DT and CCS values, thus cannot be differentiated by LC IM-MS. Also, isomers with subtle structural differences may not be able to be separated in IM due to the limited resolution. With these in consideration, here we

define each peak from DTMS as an isomer cluster, with each cluster containing structurally similar isomers. As an illustration, two different isomer clusters were detected for the *m/z* of 311.1650 ($C_{20}H_{24}O_3$; Figure 5). Isomer cluster 1 has a lower CCS value of ca. 159.6 Å$^2$, while the other cluster has a CCS value of ca. 168.6 Å$^2$.

Overall, the majority of the detected *m/z*s in all GOM samples had only one isomer cluster (Figures 6, S10, S11, and S12). For those formulas with more than one isomer cluster, the highest number of isomer clusters per formula was 8 (only one formula) under ESI+ and 9 (only one formula) under ESI-, respectively (Figure 6), which agrees with previous findings in river water systems (Lu et al., 2018; Lu and Liu, 2019), and wetland systems (Morrison et al., 2020), and is within the isomer range reported by TIM-MS (6 – 10; Tose et al., 2018; Leyva et al., 2019). Even though a simple correlation between the number of isomer clusters and the *m/z* cannot be drawn, it is still evident that the *m/z*s with the highest number of isomer clusters (i.e., different structure isomers) were in the molecular weight range of 150 – 400, and the number of isomer clusters did not increase with *m/z*, consistent with the molecular size distribution of DOM.

"Isomer cluster percentage", slightly modified from previous studies (Lu et al., 2018; Lu and Liu, 2019), was applied to quantify the isomeric information on DOM. It is defined as the percentage of detected formulas that have more than one isomer cluster:

$$Isomer\ Cluster\ \% = \frac{Number\ of\ different\ formulas\ possessing\ more\ than\ one\ isomer\ cluster}{Total\ number\ of\ different\ formulas} \times 100\% \quad (1)$$

The isomer cluster percentage depicts the number of isomers with different structures, and likely only represents a lower estimate of the number of isomers in DOM. The isomer cluster percentages of GOM samples ranged from ca. 7% to ca. 13% under ESI+ mode (Table 3), and from ca. 11% to ca. 22% under ESI- mode (Table S7). With less than a quarter of the formulas detected in DTIM-MS possessing more than one cluster of isomers, the majority of DOM formulas detected only had a single peak in DTIM-MS. Similarly, in the South China Sea samples, isomer cluster percentages were in the range of ca. 5% to ca. 10% under ESI+ mode, suggesting that only a small fraction of DOM possessed structurally distinct isomers, while the majority had structurally similar isomers, if any. Moreover, the isomer cluster percentage of field samples agreed with that of SRHA (16.6% under ESI+; 5.7% under ESI-) and SRNOM (8.6% under ESI+; 10.6% under ESI-). This isomer cluster percentage of DOM is consistent with a recent study using a similar instrument, in which ca. 4 – 7% of the IM-

detected features had isomers under ESI+ mode, while less than 1% of the observed unique mass had isomers under ESI- mode (Morrison et al., 2020).

Under ESI+ mode, samples with the lowest isomer cluster percentage were from the river mouth (ARSC2), while those with the highest were from surface water (F1 and D11 0.5 m). The averaged isomer cluster percentages from river mouth samples (9.0% under ESI+, 12.2% under ESI-) were lower than those of coastal ocean (9.9% under ESI+, 19.8% under ESI-) and open ocean samples (11.2% under ESI+, and 17.3% under ESI-) under both ESI modes (Tables 3 and S7). The trend in isomer cluster percentage with environmental gradient was much clearer under ESI+. For instance, a decrease in isomer cluster percentage was detected from near shore to offshore (13.1% at F1 to 8.0% at F3; Table 3). Isomer cluster percentage also decreased with water depth, from 12.9% at surface to 9.0% in deep water. This trend was also observed in the South China Sea samples, in which the isomer cluster percentages decreased from 7.1% at 5 m to 5.7% at 1000 m, and to 6.6% at 4000 m.

Since the analytical window of IM Q-TOF LC/MS is different from the direct-infusion ultrahigh-resolution FT-ICR MS, there may be concerns that the features detected with IM are not representative of DOM. To address this, IM-MS data of DC6 samples were compared with their corresponding FT-ICR MS data, and shared features (i.e., shared mass detected in both FT-ICR MS and IM-MS) were further extracted. The isomer cluster percentages of these shared features ranged from 1.9% to 18.4% (Table S8). Even though no clear trend with depth was observed (probably due to the high variation in the number of common features), the isomer cluster percentages of shared features was within the same order compared with the results obtained with IM Q-TOF LC/MS alone, suggesting that the estimated overall isomer percentage is reliable.

Formulas with multiple isomer clusters were further analyzed based on the HPLC RT of each cluster. Physical and chemical properties, such as hydrophobic/hydrophilic interactions and hydrogen-bonding of molecules, can be inferred from the separation of HPLC depending on the type of column and mobile phases used (Alpert, 1990; Yoshida, 2004; Guo et al., 2007; Hao et al., 2008). Close LC RTs do not necessarily reflect similar chemical structure, but different LC RTs are surely indicative of different structures. The percentage of isomer clusters with similar RTs (i.e., RT-similar isomers with RT difference < 0.1 min) increased as water depth increased under both ESI modes (ESI+: 12.5% to 39.0%,

Table 3; ESI-: 11.7% to 17.9%, Table S7). On the other hand, the percentage of isomer clusters with different RT (i.e., RT-different isomers with RT difference > 0.5 min) decreased as water depth increased, from 70.0% in surface to 47.2% in deep water under ESI+ (Table 3), and from 83.8% to 60.7% under ESI- (Table S7). Taken together, these data showed that as DOM becomes less labile or more degraded (i.e., water depth increases), isomers in DOM become less diverse.

### 3.5. Changes in CCS and isomers of DOM in bioassays

Although DOM in the bioassay experiment was collected from south Texas rivers, its CCS-*m/z* relationship is similar to those of field samples from GOM, as CCS and *m/z* follow a power-fit equation of CCS = $22.257(m/z)^{0.343}$ ($R^2 = 0.678$) under ESI+ mode (Figure 7A), and an equation of CCS = $7.333(m/z)^{0.457}$ ($R^2 = 0.594$) under ESI- mode (Figure 7B). Some features disappeared during the incubation while others were produced as shown in the CCS-*m/z* diagrams (Figure S13), but the CCS-*m/z* distribution was not affected by the biodegradation. No significant difference between the distribution pattern of before- and after-incubation samples was detected under either ESI mode (Figure S13).

The distribution of number of isomer clusters across different *m/z*s followed the same pattern as described in field samples (Figure 8), with most detected *m/z*s possessing only one isomer cluster (Figure S14), while the *m/z* with the highest number of isomer clusters in the mid *m/z* range (ca. 200 – 400). The isomer cluster percentages of the riverine samples at initial time point ($T_i$) were less than 10%. Changes in the isomeric parameters throughout the 24-day incubation were more evident under ESI+ mode. Isomer cluster percentage decreased from 7.9% (LR) and 7.1% (SR) at $T_i$ to 4.9% (LR) and 4.3% (SR) at final time point ($T_f$) under ESI+ mode ($p = 0.009$ for LR; $p = 0.05$ for SR). The bioassay incubation results are consistent with natural DOM results – as DOM becomes more degraded, DOM become homogenous in its isomeric aspect, indicated by a decrease in isomer cluster percentage.

## 4. Discussion

### 4.1. A critical evaluation of the isomer data

The data presented here provide a unique angle to evaluate structural characteristics of natural DOM. However, given the limitation of this technique, such as the resolution and the loss of ions during the IM process, it is important to have a critical evaluation of whether the features detected in this work are representative before drawing any conclusion,

particularly when considering that the analytical window for this work differs from those of FT-ICR MS. However, because roughly 25% of the features detected with IM Q-TOF LC/MS (Table S1) were also in the "island of stability" region in the van Krevelen diagram from FT-ICR MS (Lechtenfeld et al., 2014; Koch et al., 2014), together with the consistency between different samples (Figures S5, S6, and S7) and DOM standards (Figure S4), suggests that the DOM molecules identified by MS alone in this work represent at least a shared fraction (ca. 12%) of the previously analyzed DOM pool.

Since a large fraction of ions (ca. 90%) was lost during DTIM-MS, there may be concern that the features detected in DTIM-MS were biased compared with MS. However, the chromatograms of MS and DTIM-MS suggested that features detected in DTIM-MS were not biased to certain groups of compounds compared with the MS without DTIM (Figure S15). Several peaks detected under the MS mode (e.g., peak at ca. 0.3 min, peaks from 6 to 7 min, etc.), were also captured under the DTIM-MS mode (Figure S15). In addition, the *m/z* distribution of the DTIM-MS results generally follows a typical normal distribution, with a range of 100 – 1000 and the highest signal intensity observed at ca. 400 – 600 (Figures 2&5; Figures S8&S9), consistent with MS-alone results, and with other LC-MS work, in which the molecular weight range of DOM has a unimodal distribution with the maximum intensity at around 400 Da (Dittmar and Stubbins, 2014; Kim et al., 2019). This similar distribution suggests that IM did not heavily skew the analytical window compared with MS alone. Furthermore, the fact that isomer cluster percentages measured in field samples and incubation samples agreed well with the DOM standards SRNOM and SRHA, indicated the consistency of our results and the observed trends. Finally, a comparison between FT-ICR MS and DTIM-MS based on the DC6 samples from South China Sea showed that, even if the isomer cluster percentages were calculated only within those molecular features that were detected by both MS techniques, they did not differ significantly from the isomer cluster percentages calculated with all molecular features detected by DTIM-MS alone (Table S7 and S8), suggesting that the results obtained by DTIM-MS are reliable. Taken together, DOM molecules detected through DTIM-MS are consistent with the results of MS, and are also representative of the SPE-extracted DOM.

The resolving power of available IM instruments generally ranges from 50 to ca. 100 (Dodds et al., 2017a; Tose et al., 2018; Leyva et al., 2019). With a typical resolving power of

60 (i.e., separation of isomers with a 0.5 ms difference in drift time when the target compounds have a drift time of 30 ms), the current DTIM-MS cannot resolve all isomers in a complex mixture (May and McLean, 2015), but isomers with large structural difference can be separated. This has been verified with various standards by previous pioneering work in IM. DTIM-MS can differentiate most constitutional isomers even with subtle structural difference (Dwivedi et al., 2007; Zhu et al., 2009), with a resolution of 2% difference in CCS (Nichols et al., 2018). For example, tertrasaccharide alditols (Glcβ1-4 Glcβ1-4 Glcβ1-4 Glc-ol) can be well separated from maltotetraitol (Glcα1-4 Glcα1-4 Glcα1-4 Glc-ol) by a difference of 1.20 ms in drift time in a similar instrument (Zhu et al., 2009) due to the structural difference in α- and β-glycosidic bonds. A similar system (LC-ESI-IM-TOF-MS) resolved over 70% of the constitutional isomers from a 4000-peptide mixture (Srebalus Barnes et al., 2002). A more recent study reported that with a resolving power of 60, IM-MS alone (without LC) resolved over 30% of the isomers of leucine ($C_6H_{13}NO_2$), with most constitutional isomers (i.e., ethyl ester vs. tertleucine vs. norleucine) resolved, but some diastereomers (i.e., L-allo-isoleucine vs. L-isoleucine) and most enantiomers (L-leucine vs. D-leucine and L-isoleucine vs. D-isoleucine) unresolvable (Dodds et al., 2017b). Our results of the peptide isomer standards also confirmed the resolving power of this instrument to be capable of resolving structurally similar isomers (Figure 3), in which two peptide isomers (SDGRG and GRGDS) were separated by a DT difference of 7%. Indeed, more technical aspects are still needed to enhance the resolving power of DTIM-MS, but the results here represent real measurement and offer at least a lower estimation of isomer diversity of DOM (i.e., identification of structurally different isomers in DOM), and thus a unique angle to view the tip of the iceberg of DOM.

### 4.2. The number of structurally distinct isomers in DOM is constrained

As mentioned previously, the resolving power of current IM-MS is incapable of separating all possible isomers of DOM. For instance, peaks detected in IM-MS generally follow a normal distribution (e.g., Figure 5; Edelson et al., 1967; Revercomb and Mason, 1975; Spangler and Collins, 1975; Dodds et al., 2017a), with a width of ca. 1 ms (e.g., Figure 5B&C), consistent with the typical peak widths of single standards (e.g., May et al., 2014; Stow et al., 2015, 2017a; and Figure 3). With each peak consisting of about 10 sampling points, the sampling frequency is in the level of 0.1 ms. Under the circumstances where the

structural differences among isomers were so small, reflected by minute differences in both the IM drift time and LC retention time, each detected peak could be a composite of multiple smaller ones (e.g., Figure S16). Therefore, each peak detected in LC-DTIM-MS should be viewed as an "isomer cluster" or "isomer clade", representing a cluster of isomers with similar structure. Meanwhile, multiple peaks sharing the same molecular mass should be viewed as different clusters of isomers with distinct structures, and within each cluster there could be multiple inseparable isomers (Hertkorn et al., 2007; Hertkorn et al., 2008). This is consistent with several recent studies, in which one *m/z* could have multiple separated IM-MS peaks, but a further separation of the seemingly one-apex peaks from TIM-MS generally relied on a Software Assisted Molecular Elucidation (SAME) package (Benigni et al., 2017; Tose et al., 2018; Leyva et al., 2019; Gao et al., 2019).

Isomers can be structurally distinct (represented by multiple peaks with same *m/z* in IM-MS) or similar (represented by a single peak at given *m/z* in IM-MS). In this work, the isomer cluster percentages, which represented the percentages of structurally different isomers, of measured DOM molecules ranged from 4.31% to 13.11% under ESI+ mode, and from 6.65% to 22.08% under ESI- mode from a variety of environments (Tables 3, 4&S7). Following the aforementioned logic, this rather low percentage suggests that the number of structurally distinct isomers in the detected formulas was constrained.

It has to be acknowledged that "constrained number of structurally distinct isomer" does not necessarily mean "constrained number of total isomers". Even though achieving an exact number of isomers per cluster was well beyond the current analytical power, a rough estimate can be inferred based on computational chemistry. We used a measured formula from LC-DTIM-MS, $C_{18}H_{19}NO_5$, to investigate the hydrophobicity and geometric properties of its possible isomers. A N-containing CRAM-like structure was tentatively assigned based on current understanding of DOM molecules (Hertkorn et al., 2006; Cao et al., 2017), and it could have numerous isomers in theory based on random arrangements of functional groups (i.e., the positions of functional groups). However, this formula was only separated into two different isomer clusters in LC-DTIM-MS, with one having an LC RT of 3.66 min and CCS of 152.7 $Å^2$, and the other having a RT of 3.22 min and CCS of 153.3 $Å^2$. In this case, we would expect each detected cluster to represent a large number of isomers. The log P values and projection areas, though obtained through calculation, could provide insights into

whether these molecules could be resolved in LC-DTIM-MS. One would expect almost identical log P values and projection areas for isomers aggregated in one LC-DTIM-MS peak. However, calculations suggested that even subtle changes in structure (e.g., position of double bond) introduced distinguishable differences in molecule properties (Table 5, structures 1 – 5; log P from 0.20 to 0.50), while larger changes such as the rearrangement of carboxyl group greatly decreased the hydrophobicity and the projection area of the isomer (Table 5, structure 6). Another typical CRAM compound was investigated. With a formula of $C_{18}H_{20}O_9$, this compound was assigned with a core structure of tetradecahydrophenanthrene, and functional groups of carboxyl groups, carbonyl group and a double bond (Hertkorn et al., 2006). Likewise, this compound could have thousands of isomers as a result of different positions of functional groups. Yet again, calculations suggested the hydrophobicity and geometric properties of the isomers were not identical, as shown by some examples (Table S9). Structures 1 to 14 were only differentiated by the position of the double bond (Table S9, line 1 – 14), and even such subtle changes resulted in a quite wide range in log P from -0.07 (hydrophilic) to 0.44 (hydrophobic). For isomers with same log P, their projection areas were generally quite different (ca. 1 – 3% difference, on the same level of the CCS; e.g., structures 7 and 8). Similarly, changes in the positions of the carboxyl group (e.g., molecules 1, 15, 18, 19, 20, 21), were also well reflected in both hydrophobicity and geometric properties. Based on these calculations, the differences in hydrophobicity and projection areas of these isomers make it impossible for one peak from LC-DTIM-MS to contain thousands of isomers in the present work, though the possibility that a small number of isomers exist within one peak cannot be ruled out. In other words, the number of isomers in one isomer cluster is also restricted, by the hydrophobicity and geometric properties of the isomers, at least for the DOM detected in this study.

Previous work has assumed that each formula in DOM could represent thousands of different isomers as long as these isomers are chemically possible (e.g., Hertkorn et al., 2007, 2008). This would naturally result in large amount of structurally distinct isomers at the same *m/z*, which would be reflected in DTIM-MS as either a jagged shape, multi-apex wide joint peak (like the IM-MS peaks of petroleum in Benigni et al., 2017), or a large peak with a large width. However, neither of these patterns was observed in the DOM molecules detected in this work. Instead, most detected *m/z* had a peak width of ca. 1 ms (Figure 3), which was on

686 the same magnitude of a single standard compound (Edelson et al., 1967; Spangler and
687 Collins, 1975; Revercomb and Mason, 1975; Dodds et al., 2017a), and only a small fraction
688 had well-separated IM-MS peaks. Furthermore, the MS/MS product ion spectrum offers
689 insights into the quantity of isomers in a given formula. For a detected *m/z* with multiple
690 structurally distinct isomers, fragmentation would generate a wide range of and complex
691 product ions. On the other hand, if a detected *m/z* is constrained to limited numbers of
692 structurally similar isomers, it would only have product ion fragments with certain *m/z*s, and
693 would be more recognizable and cleaner. As an example of this, a feature with an *m/z* of
694 423.1651 ($C_{21}H_{28}O_9$) had two distinct clusters of isomers, with one eluted at ca. 3.50 min and
695 a CCS of 162.1 $Å^2$, and the other eluted at ca. 3.87 min and a CCS of 166.5 $Å^2$. Due to the
696 prior separation on LC, the product ion spectrum at RT 3.5 min represented only one isomer
697 cluster (within the cluster there could be multiple structurally similar isomers), and most of
698 its product ion peaks were assigned with structures (Figure S17). The fact that this clean
699 pattern has been commonly seen in the features which were detected by both MS/MS and
700 FT-ICR MS (data not shown), further suggested the similarity of structures of the isomers
701 within one isomer cluster, or just one compound.
702 Together, these results suggest that isomers in this specific fraction of DOM within
703 the analytical window of IM Q-TOF LCMS, were under certain constraints in addition to
704 chemical laws (e.g., Hertkorn et al., 2007). Isomers within the same isomer cluster, if any,
705 constrained by LC retention time and geometric configuration, are expected to share very
706 similar core chemical structures, and therefore can be further considered as homologous
707 compounds with similar function (Mentges et al., 2017). Note that because the molecules
708 detected in this work represented a certain fraction of the natural DOM pool, it is possible
709 that those undetected DOM molecules could possess a high number of isomers, although
710 there is no particular reason for such argument.
711 The selectivity of isomers, in fact, is not uncommon in biochemistry. Biological
712 homochirality, for example, results in L-forms of amino acids and D-forms of glucose, as the
713 dominant enantiomers in living organisms (Ribó et al., 2017; Hochberg et al., 2019), and this
714 selectivity in chirality is strictly maintained through the metabolic pathways (known as
715 Spontaneous Mirror Symmetry Breaking; e.g., Plasson et al., 2007; Ribó et al., 2017;
716 Hochberg et al., 2019). As microbial reworking is an important process in the degradation of

DOM (e.g., Zark and Dittmar, 2018), selectivity in isomers can be expected to some extent. Note that this constraint or selectivity in isomers does not necessarily contradict previous work, which modelled a large number of isomers in DOM (Zark et al., 2017; Hawkes et al., 2018; Zark and Dittmar, 2018), as this constraint in isomer clusters, which is more on a function level, is focused on structural diversity not the actual number of isomers.

### 4.3. Trend of DOM isomers with environmental degradation

Our results also offer insights into how isomer diversity of DOM changes among different environments and through biodegradation. Isomer cluster percentages of DOM was lowest in samples collected from the river mouth, and decreased from surface to deep water (Table 3), suggesting that isomers in DOM became more structurally similar with degradation. Samples collected from surface ocean (F-sites, and D11 0.5), where DOM is expected to be most labile due to primary production, have more diverse isomers (higher percentages RT-different isomer clusters, and lower percentages of RT-similar isomer clusters; Table 3). Concurrently, samples collected from deep water (e.g., D11 1240) and the river mouth (particularly ARSC1 and ARSC2), where DOM may have been heavily processed, generally have more homogenous isomers (i.e., structurally more similar) compared with coastal and surface ocean sites (e.g., F1, D11 0.5), as indicated by a lower percentage in RT-different isomer cluster (e.g., less than 50% for ARSC1, ARSC2 and D11 1240, but over 70% for D11 0.5 and F1 under ESI+, Table1; generally less than 60% for D11 1240, ARSC1, ARSC2, but over 65% for F1 and D11 0.5, Table S7), and a higher percentage in RT-similar isomer clusters (e.g., over 20% for ARSC1, ARSC2 and D11 1240, but less than 15% for D11 0.5 and F1 under ESI+, Table1; generally over 15% for D11 1240, ARSC1, ARSC2, but only ca. 12% for F1 and D11 0.5, Table S7). These changes in isomer cluster percentage and isomer properties suggest that as DOM becomes more refractory, it becomes more homogenous in its isomeric diversity.

The results from incubation experiment further support this idea. As DOM was degraded through microbial metabolism, isomer cluster percentage of DOM decreased from ca. 7.5% to less than 5% under ESI+ (Table 4). This decrease in isomer cluster diversity with incubation, is consistent with previous findings that DOM becomes less diverse or more homogenous with degradation (e.g., Mentges et al., 2017; O'Connor et al., 2020). However,

it should be noted that since SPE preferentially extracts hydrophobic compounds, the "labile" fraction in SPE-DOM may not necessarily represent the labile fraction of the whole DOM.

As mentioned previously, the decreasing isomer cluster diversity with increasing DOM recalcitrance may be attributed to the highly regulated nature of biochemical processes, which are important in the formation of natural DOM when processed by microbes. Unlike abiotic reactions, which can form a large number of structurally diverse isomers as products as long as the reactions follow the laws of thermodynamics, the diversity of isomers in the products of biochemical reactions is generally constrained. For instance, abiotic reactions such as the formation of petroleum hydrocarbon and the abiotic racemization of amino acids tend to form a large number of isomers (e.g., Vandenbroucke and Largeau, 2007; Ribó et al., 2017). On the other hand, biotic reactions such as the biodegradation of peptide, during which extracellular peptidase can hydrolyze constitutional isomers of peptide (same amino acids composition but in different sequences) into same small fragments or single amino acids (Liu et al., 2013; Liu and Liu, 2015), results in a loss in isomer diversity. Another example would be the glycolysis of different forms of $C_6$ carbohydrate (e.g., glucose, fructose, and mannose), in which isomers of $C_6$ carbohydrate are all converted into pyruvate. As microbial reworking of DOM has been thought to play a major role in the formation of deep ocean DOM (e.g., Jiao et al., 2010), and it is still unclear to which extent abiotic reactions contribute to the formation of DOM, therefore a decreasing isomer diversity of natural DOM with biodegradation is not unexpected.

## 5. Conclusions

We applied multi-dimension molecular level analysis of natural DOM via IM Q-TOF LCMS, with the focus on the isomeric aspect of this intricate mixture. By a critical comparison with FT-ICR MS, we demonstrated that the formulas detected with IM Q-TOF LCMS represented an important fraction, though different from the widely accepted FT-ICR MS results, of the DOM pool, representing roughly 12% of the commonly recognized DOM pool. From "the tip of iceberg", our results indicate that only a small fraction of the detected DOM formulas (less than 23%) possesses more than one isomer clusters, or structurally distinct isomers. Multidimensional pieces of evidence, from tandem MS, IM, and computational chemistry, have shown that the majority of DOM only had structurally similar isomer clusters, if any. This constraint in the functional diversity of isomers in DOM could

be attributed to the biological nature of DOM. With DOM samples collected from different environments and through bioassay experiments, we further demonstrate that isomer clusters in natural DOM become more homogenous, or less diverse with environmental degradation. However, more data, including but not limited to DOM collected from different sources as well as investigations on possible processes (e.g., both biotic and abiotic, such as photochemical reactions) that could affect isomeric diversities of DOM, are required to further elucidate the isomeric information on DOM. New analytical tools and improvements on existing approaches are also in need to further probe into the isomeric characteristics of DOM. For instance, the combination of pre-reduction and comprehensive gas chromatography × gas chromatography (GC × GC) MS offered another unique angle to investigate the C-backbone and isomeric diversity of DOM molecules (Arakawa et al., 2017). To conclude, this work with IM-MS only represents an initial, but pivotal, step of the endeavor to a better understanding of the intricate natural DOM.

**Acknowledgement**

We thank Dr. Ryan Hladyniuk from UTMSI Core Facility Lab for helping assemble and maintain the equipment, and Drs. Carol Haney Ball, Dawn Stickle, Caroline S. Chu, and Daniel Cuthbertson from Agilent Technologies for invaluable technical support. Special thanks to Drs. Robert Dickey and Ed Buskey for making the instrumentation available at UTMSI. Thanks to Dr. Wei-Jun Cai, and to the crew members of *R/V Pelican* for the GOM cruise. We also would like to thank Dr. Kai Wu, and John O'Connor for their help in sampling trips and the incubation experiments. This study was supported by NOAA (NA15NOS4780185), National Natural Science Foundation of China (NSFC 41676059 & 41890801) an Institutional Grant (NA14OAR4170102) to the Texas Sea Grant College Program from the National Sea Grant Office, NOAA, and the NSF Chemical Oceanography program (OCE 1763167).

The data that support the findings of this study are available on request from the corresponding author, ZL.

807 Reference


808 Alpert A. J. (1990) Hydrophilic-interaction chromatography for the separation of peptides,
809 nucleic acids and other polar compounds. *J. Chromatogr. A* **499**, 177–196.

810 Arakawa N., Aluwihare L. I., Simpson A. J., Soong R., Stephens B. M. and Lane-Coplen D.
811 (2017) Carotenoids are the likely precursor of a significant fraction of marine dissolved
812 organic matter. *Sci. Adv.* **3**, e1602976.

813 Arrieta J. M., Mayol E., Hansman R. L., Herndl G. J., Dittmar T. and Duarte C. M. (2015)
814 Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* **348**, 331–
815 333.

816 Baker E. S., Clowers B. H., Li F., Tang K., Tolmachev A. V., Prior D. C., Belov M. E. and
817 Smith R. D. (2007) Ion mobility spectrometry–mass spectrometry performance using
818 electrodynamic ion funnels and elevated drift gas pressures. *J. Am. Soc. Mass Spectrom.*
819 **18**, 1176–1187.

820 Barber R. T. (1968) Dissolved organic carbon from deep waters resists microbial oxidation.
821 *Nature* **220**, 274–275.

822 Benigni P., Sandoval K., Thompson C. J., Ridgeway M. E., Park M. A., Gardinali P. and
823 Fernandez-Lima F. (2017) Analysis of photoirradiated water accommodated fractions of
824 crude oils using tandem TIMS and FT-ICR MS. *Environ. Sci. Technol.* **51**, 5978–5988.

825 Benner R. (2002) Chemical composition and reactivity. In *Biogeochemistry of Marine Dissolved*
826 *Organic Matter (First Edition)* (eds. D. A. Hansell and C. A. Carlson). Academic Press,
827 Boston. pp. 59–90.

828 Cao X., Mulholland M. R., Helms J. R., Bernhardt P. W., Duan P., Mao J. and Schmidt-Rohr K.
829 (2017) A major step in opening the black box of high-molecular-weight dissolved organic
830 nitrogen by isotopic labeling of *Synechococcus* and multibond two-dimensional NMR.
831 *Anal. Chem.* **89**, 11990–11998.

832 Chin W.-C., Orellana M. V. and Verdugo P. (1998) Spontaneous assembly of marine dissolved
833 organic matter into polymer gels. *Nature* **391**, 568–572.

834 D'Andrilli J., Cooper W. T., Foreman C. M. and Marshall A. G. (2015) An ultrahigh-resolution
835 mass spectrometry index to estimate natural organic matter lability. *Rapid Commun.*
836 *Mass Spectrom.* **29**, 2385–2401.

837 Dittmar T. (2015) Reasons behind the long-term stability of dissolved organic matter. In
838 *Biogeochemistry of Marine Dissolved Organic Matter (Second Edition)* (ed. C. A.
839 Carlson). Academic Press, Boston. pp. 369–388. Available at:
840 http://www.sciencedirect.com/science/article/pii/B9780124059405000078 [Accessed
841 May 3, 2016].

Dittmar T. and Kattner G. (2003) Recalcitrant dissolved organic matter in the ocean: major contribution of small amphiphilics. *Mar. Chem.* **82**, 115–123.

Dittmar T., Koch B., Hertkorn N. and Kattner G. (2008) A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol Ocean. Methods* **6**, 230–235.

Dittmar T. and Paeng J. (2009) A heat-induced molecular signature in marine dissolved organic matter. *Nat. Geosci.* **2**, 175–179.

Dittmar T. and Stubbins A. (2014) Dissolved organic matter in aquatic systems. *Treatise Geochem. Second Ed. Elsevier*, 125–156.

Dodds J. N., May J. C. and McLean J. A. (2017a) Correlating resolving power, resolution, and collision cross section: unifying cross-platform assessment of separation efficiency in ion mobility spectrometry. *Anal. Chem.* **89**, 12176–12184.

Dodds J. N., May J. C. and McLean J. A. (2017b) Investigation of the complete suite of the leucine and isoleucine isomers: toward prediction of ion mobility separation capabilities. *Anal. Chem.* **89**, 952–959.

Dwivedi P., Bendiak B., Clowers B. H. and Hill Jr. H. H. (2007) Rapid resolution of carbohydrate isomers by electrospray ionization ambient pressure ion mobility spectrometry-time-of-flight mass spectrometry (ESI-APIMS-TOFMS). *J. Am. Soc. Mass Spectrom.* **18**, 1163–1175.

Edelson D., Morrison J. A., McKnight L. G. and Sipler D. P. (1967) Interpretation of ion-mobility experiments in reacting systems. *Phys. Rev.* **164**, 71–75.

Fenn L. S., Kliman M., Mahsut A., Zhao S. R. and McLean J. A. (2009) Characterizing ion mobility-mass spectrometry conformation space for the analysis of complex biological samples. *Anal. Bioanal. Chem.* **394**, 235–244.

Ferrara P., Apostolakis J. and Caflisch A. (2002) Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins Struct. Funct. Bioinforma.* **46**, 24–33.

Gao Y., Wang W., He C., Fang Z., Zhang Y. and Shi Q. (2019) Fractionation and molecular characterization of natural organic matter (NOM) by solid-phase extraction followed by FT-ICR MS and ion mobility MS. *Anal. Bioanal. Chem.* Available at: https://doi.org/10.1007/s00216-019-01943-7 [Accessed June 27, 2019].

Gaspar A., Kunenkov E. V., Lock R., Desor M., Perminova I. and Schmitt-Kopplin P. (2009) Combined utilization of ion mobility and ultra-high-resolution mass spectrometry to identify multiply charged constituents in natural organic matter. *Rapid Commun. Mass Spectrom.* **23**, 683–688.

876 Green N. W., McInnis D., Hertkorn N., Maurice P. A. and Perdue E. M. (2014) Suwannee river
877 natural organic matter: isolation of the 2R101N reference sample by reverse osmosis.
878 *Environ. Eng. Sci.* **32**, 38–44.

879 Guo Y., Srinivasan S. and Gaiki S. (2007) Investigating the effect of chromatographic conditions
880 on retention of organic acids in hydrophilic interaction chromatography using a design of
881 experiment. *Chromatographia* **66**, 223–229.

882 Hansell D., Carlson C., Repeta D. and Schlitzer R. (2009) Dissolved organic matter in the ocean:
883 a controversy stimulates new insights. *Oceanography* **22**, 202–211.

884 Hao Z., Xiao B. and Weng N. (2008) Impact of column temperature and mobile phase
885 components on selectivity of hydrophilic interaction chromatography (HILIC). *J. Sep.*
886 *Sci.* **31**, 1449–1464.

887 Hawkes J. A., D'Andrilli J., Agar J. N., Barrow M. P., Berg S. M., Catalán N., Chen H., Chu R.
888 K., Cole R. B., Dittmar T., Gavard R., Gleixner G., Hatcher P. G., He C., Hess N. J.,
889 Hutchins R. H. S., Ijaz A., Jones H. E., Kew W., Khaksari M., Lozano D. C. P., Lv J.,
890 Mazzoleni L., Noriega-Ortega B. E., Osterholz H., Radoman N., Remucal C. K., Schmitt
891 N. D., Schum S., Shi Q., Simon C., Singer G., Sleighter R. L., Stubbins A., Thomas M.
892 J., Tolic N., Zhang S., Zito P. and Podgorski D. C. (2020) An international laboratory
893 comparison of dissolved organic matter composition by high resolution mass
894 spectrometry: are we getting the same answer? *Limnol. Oceanogr. Methods* **n/a**.
895 Available at: https://aslopubs.onlinelibrary.wiley.com/doi/abs/10.1002/lom3.10364
896 [Accessed June 10, 2020].

897 Hawkes J. A., Dittmar T., Patriarca C., Tranvik L. J. and Bergquist J. (2016) Evaluation of the
898 Orbitrap mass spectrometer for the molecular fingerprinting analysis of natural dissolved
899 organic matter (DOM). *Anal. Chem.* **88**, 7698–7704.

900 Hawkes J. A., Patriarca C., Sjöberg P. J. R., Tranvik L. J. and Bergquist J. (2018) Extreme
901 isomeric complexity of dissolved organic matter found across aquatic environments.
902 *Limnol. Oceanogr. Lett.*, n/a-n/a.

903 Hedges J. I. (1992) Global biogeochemical cycles: progress and problems. *Mar Chem* **39**, 67–93.

904 Hertkorn N., Benner R., Frommberger M., Schmitt-Kopplin P., Witt M., Kaiser K., Kettrup A.
905 and Hedges J. I. (2006) Characterization of a major refractory component of marine
906 dissolved organic matter. *Geochim. Cosmochim. Acta* **70**, 2990–3010.

907 Hertkorn N., Frommberger M., Witt M., Koch B. P., Schmitt-Kopplin Ph. and Perdue E. M.
908 (2008) Natural organic matter and the event horizon of mass spectrometry. *Anal. Chem.*
909 **80**, 8908–8919.

910 Hertkorn N., Harir M., Koch B. P., Michalke B. and Schmitt-Kopplin P. (2013) High-field NMR
911 spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular
912 level characterization of marine dissolved organic matter. *Biogeosciences* **10**, 1583–1624.

Hertkorn N., Ruecker C., Meringer M., Gugisch R., Frommberger M., Perdue E. M., Witt M. and Schmitt-Kopplin P. (2007) High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal. Bioanal. Chem.* **389**, 1311–1327.

Hines K. M., May J. C., McLean J. A. and Xu L. (2016) Evaluation of collision cross section calibrants for structural analysis of lipids by traveling wave ion mobility-mass spectrometry. *Anal. Chem.* **88**, 7329–7336.

Hochberg D., Blanco C. and Stich M. (2019) Spontaneous mirror symmetry breaking from recycling in enantioselective polymerization. In *Biological Systems: Nonlinear Dynamics Approach* (eds. J. Carballido-Landeira and B. Escribano). SEMA SIMAI Springer Series. Springer International Publishing, Cham. pp. 39–57. Available at: https://doi.org/10.1007/978-3-030-16585-7_3 [Accessed January 13, 2020].

Jiao N., Herndl G. J., Hansell D. A., Benner R., Kattner G., Wilhelm S. W., Kirchman D. L., Weinbauer M. G., Luo T., Chen F. and Azam F. (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* **8**, 593–599.

Kattner G., Simon M. and Koch B. (2011) Molecular characterization of dissolved organic matter and constraints for prokaryotic utilization. *Microb. Carbon Pump Ocean N Jiao F Azam Sanders Eds Wash. DC Sci.* Available at: http://epic.awi.de/24456/ [Accessed October 18, 2014].

Kellerman A. M., Dittmar T., Kothawala D. N. and Tranvik L. J. (2014) Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nat. Commun.* **5**, 3804.

Kim D., Kim Sungjune, Son S., Jung M.-J. and Kim Sunghwan (2019) Application of online liquid chromatography 7 T FT-ICR mass spectrometer equipped with quadrupolar detection for analysis of natural organic matter. *Anal. Chem.* **91**, 7690–7697.

Kim S., Kramer R. W. and Hatcher P. G. (2003) Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Anal. Chem.* **75**, 5336–5344.

Koch B., Kattner G., Witt M. and Passow U. (2014) Molecular insights into the microbial formation of marine dissolved organic matter: recalcitrant or labile? *Biogeosciences* **11**, 4173–4190.

Kramer M. G., Sanderman J., Chadwick O. A., Chorover J. and Vitousek P. M. (2012) Long-term carbon storage through retention of dissolved aromatic acids by reactive particles in soil. *Glob. Change Biol.* **18**, 2594–2605.

Kujawinski E. B. (2002) Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex environmental mixtures. *Environ. Forensics* **3**, 207–216.

951 Lam B., Baer A., Alaee M., Lefebvre B., Moser A., Williams A. and Simpson A. J. (2007) Major
952 structural components in freshwater dissolved organic matter. *Environ. Sci. Technol.* **41**,
953 8240–8247.

954 Lanucara F., Holman S. W., Gray C. J. and Eyers C. E. (2014) The power of ion mobility-mass
955 spectrometry for structural characterization and the study of conformational dynamics.
956 *Nat. Chem.* **6**, nchem.1889.

957 Lechtenfeld O. J., Hertkorn N., Shen Y., Witt M. and Benner R. (2015) Marine sequestration of
958 carbon in bacterial metabolites. *Nat. Commun.* **6**, 1–8.

959 Lechtenfeld O. J., Kattner G., Flerus R., McCallister S. L., Schmitt-Kopplin P. and Koch B. P.
960 (2014) Molecular transformation and degradation of refractory dissolved organic matter
961 in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* **126**, 321–337.

962 Leenheer J. A., Ferrer I., Furlong E. T. and Rostad C. E. (2003) Charge characteristics and
963 fragmentation of polycarboxylic acids by electrospray ionization—multistage tandem
964 mass spectrometry. In *Liquid Chromatography/Mass Spectrometry, MS/MS and Time of*
965 *Flight MS* ACS Symposium Series. American Chemical Society. pp. 312–324. Available
966 at: https://doi.org/10.1021/bk-2003-0850.ch018 [Accessed December 1, 2019].

967 Leyva D., Jaffe R. and Fernandez-Lima F. (2020) Structural characterization of dissolved
968 organic matter at the chemical formula level using TIMS-FT-ICR MS/MS. *Anal. Chem.*
969 **92**, 11960–11966.

970 Leyva D., Valadares L., Porter J., Wolff J., Jaffè R. and Fernandez-Lima F. (2019)
971 Understanding the structural complexity of dissolved organic matter: isomeric diversity.
972 *Faraday Discuss.* Available at:
973 https://pubs.rsc.org/en/content/articlelanding/2019/fd/c8fd00221e [Accessed April 10,
974 2019].

975 Li L. and Qu T. (2006) Thermohaline circulation in the deep South China Sea basin inferred
976 from oxygen distributions. *J. Geophys. Res. Oceans* **111**. Available at:
977 https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2005JC003164@10.1002/(ISSN)21
978 69-9291.CHINASEAS1 [Accessed January 16, 2020].

979 Liu S. and Liu Z. (2015) Comparing extracellular enzymatic hydrolysis between plain peptides
980 and their corresponding analogs in the northern Gulf of Mexico Mississippi River plume.
981 *Mar. Chem.* **177, Part 2**, 398–407.

982 Liu Z., Liu S., Liu J. and Gardner W. S. (2013) Differences in peptide decomposition rates and
983 pathways between hypoxic and oxic coastal environments. *Mar. Chem.* **157**, 67–77.

984 Liu Z., Sleighter R. L., Zhong J. and Hatcher P. G. (2011) The chemical changes of DOM from
985 black waters to coastal marine waters by HPLC combined with ultrahigh resolution mass
986 spectrometry. *Estuar. Coast. Shelf Sci.* **92**, 205–216.

Lu K., Gardner W. S. and Liu Z. (2018) Molecular structure characterization of riverine and coastal dissolved organic matter with ion mobility quadrupole time-of-flight LCMS (IM Q-TOF LCMS). *Environ. Sci. Technol.* **52**, 7182–7191.

Lu K. and Liu Z. (2019) Molecular level analysis reveals changes in chemical composition of dissolved organic matter from south Texas rivers after high flow events. *Front. Mar. Sci.* **6**. Available at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00673/full?&utm_source=Email_to_authors_&utm_medium=Email&utm_content=T1_11.5e1_author&utm_campaign=Email_publication&field=&journalName=Frontiers_in_Marine_Science&id=488705 [Accessed November 14, 2019].

Mason E. A. and Schamp H. W. (1958) Mobility of gaseous lons in weak electric fields. *Ann. Phys.* **4**, 233–270.

May J. C., Goodwin C. R., Lareau N. M., Leaptrot K. L., Morris C. B., Kurulugama R. T., Mordehai A., Klein C., Barry W., Darland E., Overney G., Imatani K., Stafford G. C., Fjeldsted J. C. and McLean J. A. (2014) Conformational ordering of biomolecules in the gas phase: nitrogen collision cross sections measured on a prototype high resolution drift tube ion mobility-mass spectrometer. *Anal. Chem.* **86**, 2107–2116.

May J. C. and McLean J. A. (2015) Ion mobility-mass spectrometry: time-dispersive instrumentation. *Anal. Chem.* **87**, 1422–1436.

McDaniel E. W. (1964) Collision phenomena in ionized gases. *N. Y. Wiley 1964*. Available at: http://adsabs.harvard.edu/abs/1964cpig.book.....M [Accessed December 11, 2016].

McDaniel E. W. and Mason E. A. (1973) *The mobility and diffusion of ions in gases*., Available at: https://ntrs.nasa.gov/search.jsp?R=19740034402 [Accessed December 12, 2016].

McLean J. A., Ruotolo B. T., Gillig K. J. and Russell D. H. (2005) Ion mobility–mass spectrometry: a new paradigm for proteomics. *Int. J. Mass Spectrom.* **240**, 301–315.

Mentges A., Feenders C., Seibt M., Blasius B. and Dittmar T. (2017) Functional molecular diversity of marine dissolved organic matter is reduced during degradation. *Front. Mar. Sci.* **4**. Available at: https://www.frontiersin.org/articles/10.3389/fmars.2017.00194/full [Accessed June 29, 2018].

Moldoveanu S. C. and David V. (2017) Properties of analytes and matrices determining HPLC selection. In *Selection of the HPLC Method in Chemical Analysis* (eds. S. C. Moldoveanu and V. David). Elsevier, Boston. pp. 189–230. Available at: http://www.sciencedirect.com/science/article/pii/B9780128036846000056 [Accessed January 4, 2020].

Morrison E. S., Shields M. R., Bianchi T. S., Liu Y., Newman S., Tolic N. and Chu R. K. (2020) Multiple biomarkers highlight the importance of water column processes in treatment wetland organic matter cycling. *Water Res.* **168**, 115153.

Nichols C. M., Dodds J. N., Rose B. S., Picache J. A., Morris C. B., Codreanu S. G., May J. C., Sherrod S. D. and McLean J. A. (2018) Untargeted molecular discovery in primary metabolism: collision cross section as a molecular descriptor in ion mobility-mass spectrometry. *Anal. Chem.* **90**, 14484–14492.

O'Connor J. A., Lu K., Guo L., Rosenheim B. E. and Liu Z. (2020) Composition and lability of riverine dissolved organic matter: insights from thermal slicing ramped pyrolysis GC–MS, amino acid, and stable isotope analyses. *Org. Geochem.* **149**, 104100.

Ohno T., Sleighter R. L. and Hatcher P. G. (2016) Comparative study of organic matter chemical characterization using negative and positive mode electrospray ionization ultrahigh-resolution mass spectrometry. *Anal. Bioanal. Chem.* **408**, 2497–2504.

Osterholz H., Niggemann J., Giebel H.-A., Simon M. and Dittmar T. (2015) Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nat. Commun.* **6**, 1–8.

Patriarca C., Bergquist J., Sjöberg P. J. R., Tranvik L. and Hawkes J. A. (2018) Online HPLC-ESI-HRMS method for the analysis and comparison of different dissolved organic matter samples. *Environ. Sci. Technol.* **52**, 2091–2099.

Petrone K. C., Richards J. S. and Grierson P. F. (2009) Bioavailability and composition of dissolved organic carbon and nitrogen in a near coastal catchment of south-western Australia. *Biogeochemistry* **92**, 27–40.

Plasson R., Kondepudi D. K., Bersini H., Commeyras A. and Asakura K. (2007) Emergence of homochirality in far-from-equilibrium systems: mechanisms and role in prebiotic chemistry. *Chirality* **19**, 589–600.

Potts R. O. and Guy R. H. (1992) Predicting skin permeability. *Pharm. Res.* **9**, 663–669.

Rabalais N. N., Turner R. E. and Wiseman W. J. (2001) Hypoxia in the Gulf of Mexico. *J. Environ. Qual.* **30**, 320.

Rathgeb A., Causon T., Krachler R. and Hann S. (2017) From the peat bog to the estuarine mixing zone: common features and variances in riverine dissolved organic matter determined by non-targeted analysis. *Mar. Chem.* **194**, 158–167.

Revercomb H. E. and Mason E. A. (1975) Theory of plasma chromatography/gaseous electrophoresis - a review. *Anal. Chem.* **47**, 970–983.

Ribó J. M., Hochberg D., Crusats J., El-Hachemi Z. and Moyano A. (2017) Spontaneous mirror symmetry breaking and origin of biological homochirality. *J. R. Soc. Interface* **14**, 20170699.

Riedel T. and Dittmar T. (2014) A method detection limit for the analysis of natural organic matter via Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **86**, 8376–8382.

1059 Rochelle-Newall E. J., Pizay M.-D., Middelburg J. J., Boschker H. T. S. and Gattuso J.-P. (2004)
1060 Degradation of riverine dissolved organic matter by seawater bacteria. *Aquat. Microb.*
1061 *Ecol.* **37**, 9–22.

1062 Rostad C. E. and Leenheer J. A. (2004) Factors that affect molecular weight distribution of
1063 Suwannee river fulvic acid as determined by electrospray ionization/mass spectrometry.
1064 *Anal. Chim. Acta* **523**, 269–278.

1065 Roth V.-N., Dittmar T., Gaupp R. and Gleixner G. (2015) The molecular composition of
1066 dissolved organic matter in forest soils as a function of pH and temperature. *PLOS ONE*
1067 **10**, e0119188.

1068 Schmidt F., Elvert M., Koch B. P., Witt M. and Hinrichs K.-U. (2009) Molecular
1069 characterization of dissolved organic matter in pore water of continental shelf sediments.
1070 *Geochim. Cosmochim. Acta* **73**, 3337–3358.

1071 Seidel M., Yager P. L., Ward N. D., Carpenter E. J., Gomes H. R., Krusche A. V., Richey J. E.,
1072 Dittmar T. and Medeiros P. M. (2015) Molecular-level changes of dissolved organic
1073 matter along the Amazon River-to-ocean continuum. *Mar. Chem.* **177**, 218–231.

1074 Seifert A.-G., Roth V.-N., Dittmar T., Gleixner G., Breuer L., Houska T. and Marxsen J. (2016)
1075 Comparing molecular composition of dissolved organic matter in soil and stream water:
1076 Influence of land use and chemical characteristics. *Sci. Total Environ.* **571**, 142–152.

1077 Shen Y. and Benner R. (2019) Molecular properties are a primary control on the microbial
1078 utilization of dissolved organic matter in the ocean. *Limnol. Oceanogr.* **n/a**. Available at:
1079 https://aslopubs.onlinelibrary.wiley.com/doi/abs/10.1002/lno.11369 [Accessed January 2,
1080 2020].

1081 Spangler G. E. and Collins C. I. (1975) Peak shape analysis and plate theory for plasma
1082 chromatography. *Anal. Chem.* **47**, 403–407.

1083 Spencer R. G. M., Mann P. J., Dittmar T., Eglinton T. I., McIntyre C., Holmes R. M., Zimov N.
1084 and Stubbins A. (2016) Detecting the signature of permafrost thaw in Arctic rivers. *J.*
1085 *Geophys. Res. Oceans*, 2830–2835.

1086 Srebalus Barnes C. A., Hilderbrand A. E., Valentine S. J. and Clemmer D. E. (2002) Resolving
1087 isomeric peptide mixtures: a combined HPLC/ion mobility-TOFMS analysis of a 4000-
1088 component combinatorial library. *Anal. Chem.* **74**, 26–36.

1089 Stow S. M., Causon T. J., Zheng X., Kurulugama R. T., Mairinger T., May J. C., Rennie E. E.,
1090 Baker E. S., Smith R. D., McLean J. A., Hann S. and Fjeldsted J. C. (2017a) An
1091 interlaboratory evaluation of drift tube ion mobility–mass spectrometry collision cross
1092 section measurements. *Anal. Chem.* **89**, 9048–9055.

1093 Stow S. M., Crescentini T. M., Forsythe J. G., May J. C., McLean J. A. and Hercules D. M.
1094 (2017b) Structural characterization of methylenedianiline regioisomers by ion mobility-

mass spectrometry, tandem mass spectrometry, and computational strategies. 3. MALDI spectra of 2-ring isomers. *Anal. Chem.* **89**, 9900–9910.

Stow S. M., Onifer T. M., Forsythe J. G., Nefzger H., Kwiecien N. W., May J. C., McLean J. A. and Hercules D. M. (2015) Structural characterization of methylenedianiline regioisomers by ion mobility-mass spectrometry, tandem mass spectrometry, and computational strategies. 2. Electrospray spectra of 3-ring and 4-ring isomers. *Anal. Chem.* **87**, 6288–6296.

Tao L., McLean J. R., McLean J. A. and Russell D. H. (2007) A collision cross-section database of singly-charged peptide ions. *J. Am. Soc. Mass Spectrom.* **18**, 1232–1238.

Tose L. V., Benigni P., Leyva D., Sundberg A., Ramírez C. E., Ridgeway M. E., Park M. A., Romão W., Jaffé R. and Fernandez-Lima F. (2018) Coupling trapped ion mobility spectrometry to mass spectrometry: trapped ion mobility spectrometry–time-of-flight mass spectrometry versus trapped ion mobility spectrometry–Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* **32**, 1287–1295.

Vandenbroucke M. and Largeau C. (2007) Kerogen origin, evolution and structure. *Org. Geochem.* **38**, 719–833.

Viswanadhan V. N., Ghose A. K., Revankar G. R. and Robins R. K. (1989) Atomic physicochemical parameters for three-dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. *J. Chem. Inf. Comput. Sci.* **29**, 163–172.

Waggoner D. C., Chen H., Willoughby A. S. and Hatcher P. G. (2015) Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. *Org. Geochem.* **82**, 69–76.

Wagner S., Fair J. H., Matt S., Hosen J. D., Raymond P., Saiers J., Shanley J. B., Dittmar T. and Stubbins A. (2019) Molecular hysteresis: hydrologically driven changes in riverine dissolved organic matter chemistry during a storm event. *J. Geophys. Res. Biogeosciences* **0**. Available at: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JG004817 [Accessed April 25, 2019].

Wu K., Lu K., Dai M. and Liu Z. (2019) The bioavailability of riverine dissolved organic matter in coastal marine waters of southern Texas. *Estuar. Coast. Shelf Sci.* **231**, 106477.

Yoshida T. (2004) Peptide separation by hydrophilic-interaction chromatography: a review. *J. Biochem. Biophys. Methods* **60**, 265–280.

Zark M., Christoffers J. and Dittmar T. (2017) Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: evidence from tandem FT-ICR-MS. *Mar. Chem.* **191**, 9–15.

1132 Zark M. and Dittmar T. (2018) Universal molecular structures in natural dissolved organic
1133 matter. *Nat. Commun.* **9**, 3178.

1134 Zhai W., Dai M., Cai W.-J., Wang Y. and Wang Z. (2005) High partial pressure of $CO_2$ and its
1135 maintaining mechanism in a subtropical estuary: The Pearl River estuary, China. *Mar.*
1136 *Chem.* **93**, 21–32.

1137 Zhu M., Bendiak B., Clowers B. and Hill H. H. (2009) Ion mobility-mass spectrometry analysis
1138 of isomeric carbohydrate precursor ions. *Anal. Bioanal. Chem.* **394**, 1853–1867.

Table 1. Assigned formula number and elemental stoichiometry of GOM samples from Q-TOF LC/MS

| Samples | ESI+ | | | | ESI- | | | |
|---|---|---|---|---|---|---|---|---|
| | Formula Number* | H/C* | O/C* | N/C* | Formula Number* | H/C* | O/C* | N/C* |
| ARSC2 | 1659 ± 26 | 1.41 ± 0.03 | 0.28 ± 0.01 | 0.07 ± 0.00 | 652 ± 48 | 1.07 ± 0.04 | 0.33 ± 0.01 | 0.08 ± 0.00 |
| ARSC1 | 3337 ± 4 | 1.46 ± 0.01 | 0.26 ± 0.00 | 0.06 ± 0.00 | 748 ± 9 | 1.11 ± 0.00 | 0.34 ± 0.00 | 0.08 ± 0.00 |
| ARSC3 | 1328 ± 48 | 1.41 ± 0.02 | 0.28 ± 0.00 | 0.07 ± 0.00 | 550 ± 6 | 1.12 ± 0.00 | 0.33 ± 0.01 | 0.08 ± 0.00 |
| F1 | 1620 ± 43 | 1.47 ± 0.00 | 0.26 ± 0.00 | 0.07 ± 0.00 | 727 ± 170 | 1.19 ± 0.01 | 0.33 ± 0.01 | 0.08 ± 0.00 |
| F2 | 1262 ± 36 | 1.51 ± 0.01 | 0.24 ± 0.01 | 0.08 ± 0.00 | 199 ± 23 | 1.13 ± 0.02 | 0.34 ± 0.01 | 0.08 ± 0.00 |
| F3 | 1296 ± 0 | 1.49 ± 0.00 | 0.24 ± 0.00 | 0.08 ± 0.00 | 381 ± 79 | 1.11 ± 0.03 | 0.37 ± 0.01 | 0.08 ± 0.00 |
| D11 0.5m | 2875 ± 164 | 1.53 ± 0.01 | 0.23 ± 0.00 | 0.07 ± 0.00 | 574 ± 29 | 1.15 ± 0.04 | 0.32 ± 0.02 | 0.08 ± 0.00 |
| D11 500m | 2600 ± 49 | 1.47 ± 0.07 | 0.25 ± 0.03 | 0.08 ± 0.00 | 572 ± 121 | 1.14 ± 0.03 | 0.34 ± 0.01 | 0.08 ± 0.00 |
| D11 1240m | 1238 ± 34 | 1.50 ± 0.01 | 0.25 ± 0.00 | 0.07 ± 0.00 | 364 ± 6 | 1.17 ± 0.03 | 0.35 ± 0.01 | 0.08 ± 0.00 |

*: Number of formulas and elemental ratios were calculated by averaging two duplicate MS runs for each sample, with the precision between the two runs listed.

Table 2. Elemental stoichiometry of bioassay samples from Q-TOF LC/MS

| Samples | ESI+ | | | | ESI- | | | |
|---|---|---|---|---|---|---|---|---|
| | Formula Number* | H/C* | O/C* | N/C* | Formula Number* | H/C* | O/C* | N/C* |
| Lavaca River (LR) $T_i$ | 540 ± 16 | 1.58 ± 0.01 | 0.19 ± 0.00 | 0.05 ± 0.00 | 462 ± 21 | 1.22 ± 0.06 | 0.28 ± 0.02 | 0.07 ± 0.00 |
| Lavaca River (LR) $T_f$ | 477 ± 2 | 1.56± 0.02 | 0.19 ± 0.01 | 0.05 ± 0.00 | 262 ± 9 | 1.33 ± 0.00 | 0.25 ± 0.00 | 0.06 ± 0.00 |
| San Antonio River (SR) $T_i$ | 577 ± 34 | 1.59 ± 0.01 | 0.19 ± 0.00 | 0.06 ± 0.00 | 325 ± 51 | 1.33 ± 0.03 | 0.27 ± 0.01 | 0.06 ± 0.00 |
| San Antonio River (SR)$T_f$ | 530 ± 10 | 1.55 ± 0.01 | 0.18 ± 0.01 | 0.05 ± 0.00 | 236 ± 15 | 1.43 ± 0.00 | 0.24 ± 0.01 | 0.05 ± 0.00 |

*: Number of formulas and elemental ratios were calculated by averaging two duplicate MS runs for each sample, with the precision between two runs listed.

1149 Table 3. Isomer cluster percentage across different sites under ESI+ mode.

| Samples | Isomer Percentage (%) * | Percentage (%) of isomer clusters with similar LC-RT (RT < 0.1 min or no difference) | Percentage (%) of isomer clusters with distinct LC-RT (> 0.5 min) |
|---|---|---|---|
| ARSC1 | 9.85 ± 0.40 | 38.83 ± 2.83 | 48.83 ± 7.17 |
| ARSC2 | 7.26 ± 0.07 | 27.98 ± 13.69 | 44.64 ± 19.64 |
| ARSC3 | 9.92 ± 0.83 | 9.09 ± 9.09 | 67.13 ± 5.59 |
| F1 | 13.11 ± 1.30 | 12.92 ± 0.42 | 80.63 ± 0.63 |
| F2 | 8.50 ± 1.98 | 43.16 ± 12.39 | 34.19 ± 11.97 |
| F3 | 8.03 ± 0.10 | 25.00 ± 15.00 | 55.00 ± 25.00 |
| D11 0.5 | 12.92 ± 0.50 | 12.50 ± 2.50 | 70.00 ± 10.00 |
| D11 500 | 11.65 ± 0.28 | 12.69 ± 2.69 | 56.92 ± 3.08 |
| D11 1240 | 9.02 ± 0.64 | 38.97 ± 13.97 | 47.21 ± 17.79 |

1150 *: The standard error here stood for the standard error between duplications.

1151

1152 Table 4. Isomer cluster percentage before and after incubation.

| Sites | Time | Isomer Percentage (%) * | |
|---|---|---|---|
| | | ESI+ | ESI- |
| LR | $T_i$ | 7.88 ± 0.66 | 6.65 ± 0.96 |
| | $T_f$ | 4.95 ± 0.38 | 7.26 ± 2.26 |
| SR | $T_i$ | 7.11 ± 0.19 | 9.16 ± 3.07 |
| | $T_f$ | 4.31 ± 0.58 | 8.78 ± 3.22 |

1153 *: The standard error here stood for the standard error between duplications.

1154

1155

Table 5. Properties of isomers of modelled molecules with the molecular formula $C_{18}H_{19}NO_5$

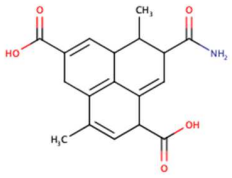
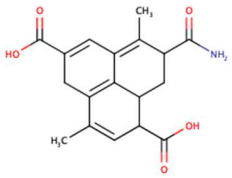
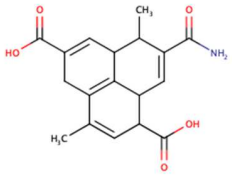
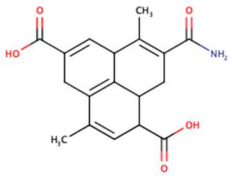
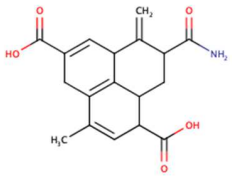
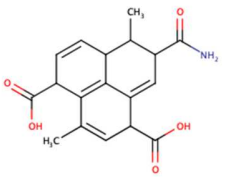
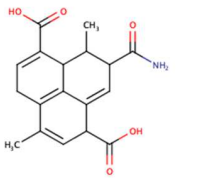
| Number | Structure | Log P | Molecular Surface Area (Å$^2$) | Min projection Area (Å$^2$) | Max projection Area (Å$^2$) | Average projection Area (Å$^2$) |
|---|---|---|---|---|---|---|
| 1 |  | 0.25 | 407.66 | 57.74 | 88.98 | 73.36 |
| 2 |  | 0.20 | 410.53 | 53.98 | 91.41 | 72.70 |
| 3 |  | 0.50 | 404.30 | 51.89 | 95.46 | 73.68 |
| 4 |  | 0.46 | 406.83 | 55.53 | 88.69 | 72.11 |
| 5 |  | 0.30 | 405.17 | 57.78 | 85.77 | 71.78 |
| 6 |  | 0.03 | 403.52 | 53.84 | 81.02 | 67.43 |
| 7 |  | 0.25 | 406.70 | 59.32 | 84.67 | 72.00 |

Figure 1. Sampling sites in Gulf of Mexico (GOM).

Figure 2. Comparison of van Krevelen diagrams of DC6 samples (collected from N 15.24°, E 114.85°, South China Sea) across different depths between FT-ICR MS and IM Q-TOF LC/MS: unique formulas assigned with FT-ICR MS (blue), unique formulas assigned with IM Q-TOF LC/MS (green), and common formulas shared by two techniques (red).

Figure 3. (A) The drift time vs. *m/z* diagram and (B) the drift time chromatogram showing the separation of two peptide isomers, SDGRG and GRGDS, under IM-MS mode.

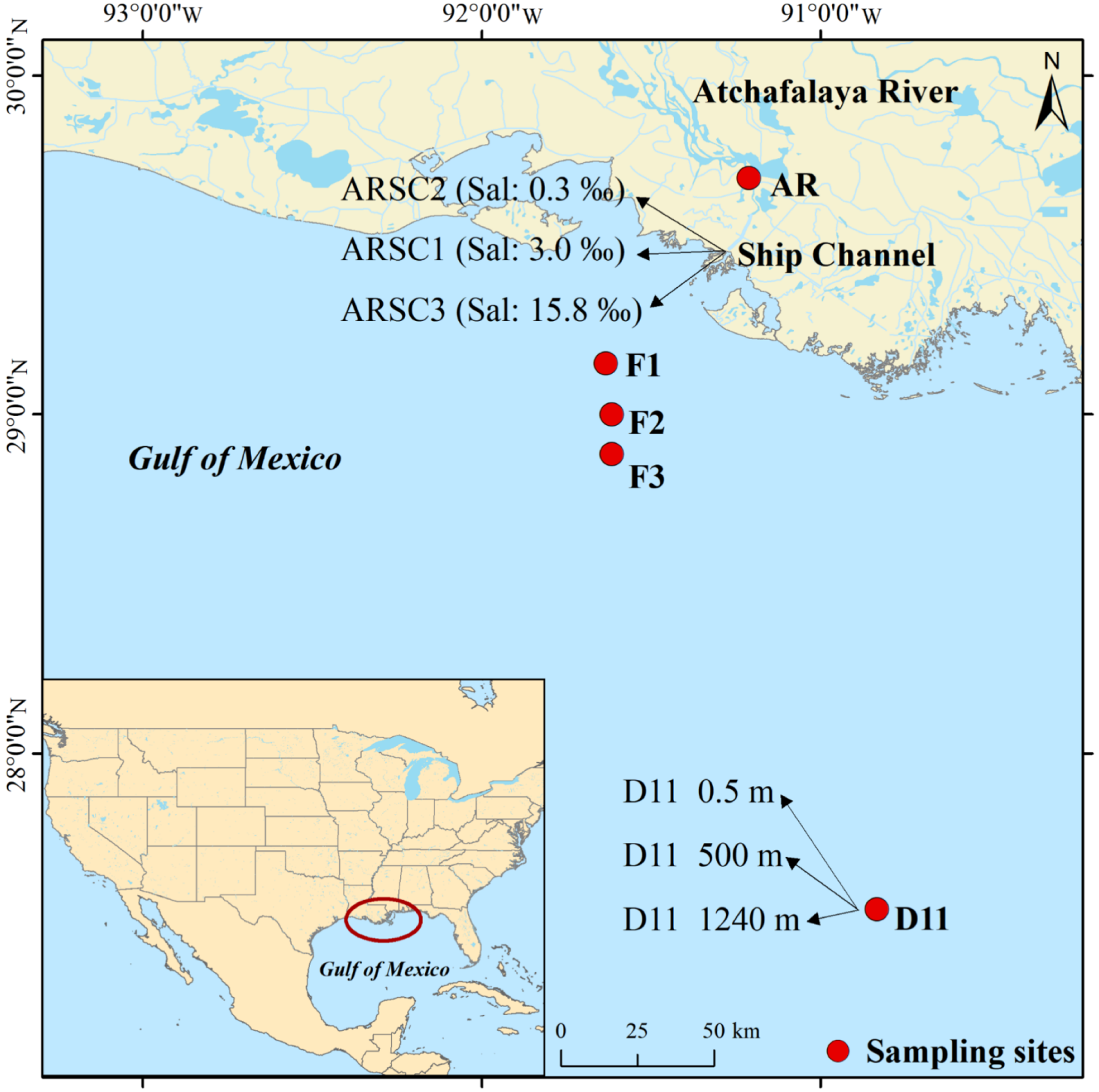
Figure 4. (A) Collision Cross Section (CCS) vs. *m/z* for GOM samples under ESI+ mode; (B) CCS vs. *m/z* for GOM samples under ESI- mode. The locations of the GOM samples are provided in Figure 1. The shaded area indicated 95% interval of prediction.

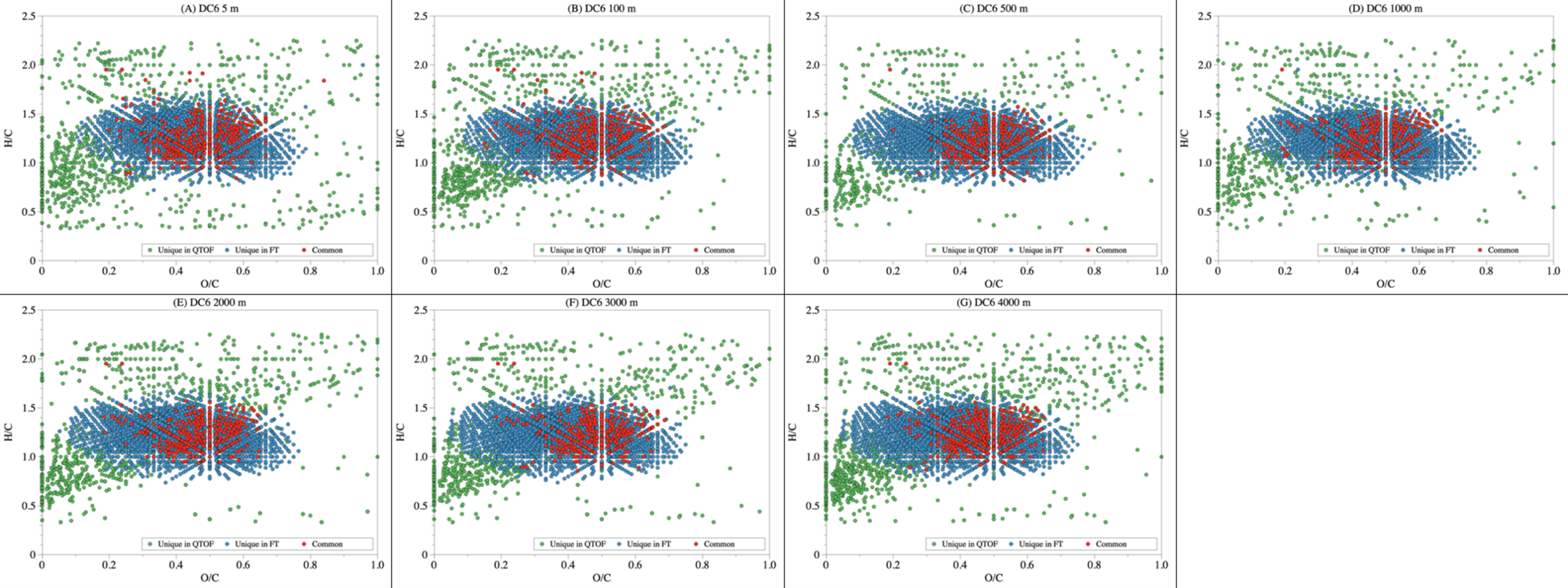
Figure 5. Example of isomer separation in DTIMS under ESI- mode. (A) CCS-*m/z* diagram of sample D11 1240m. The isomer pair is shown in blue; (B) Two isomers of *m/z* 311.1650 ($C_{20}H_{24}O_3$), shown in orange and green, respectively. The red outline around the peaks in (B) show the cumulative profile of all isomers.

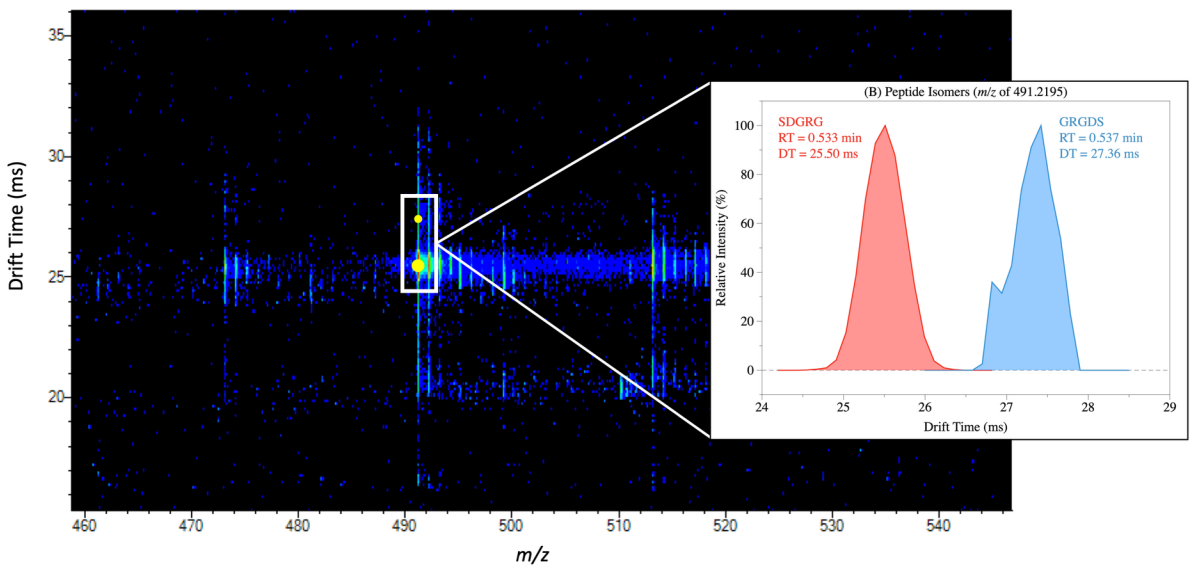
Figure 6. The max number of isomer clusters detected for the given *m/z* across different GOM samples (from ARSC1 to D11 1240) under: (A) ESI+; and (B) ESI-.

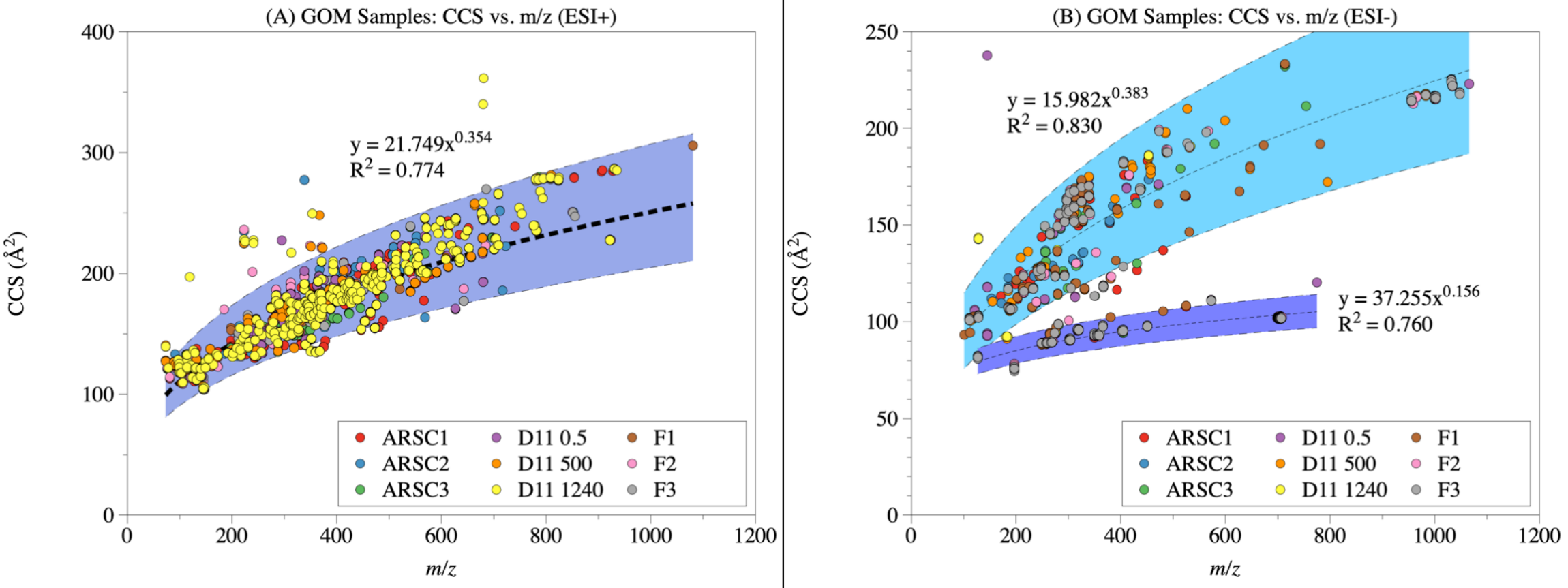
Figure 7. (A) CCS vs. *m/z* for incubation samples under ESI+ mode; (B) CCS vs. *m/z* for incubation samples under ESI- mode. The locations of the incubation samples are provided in Figure S1. The shaded area indicated 95% interval of prediction.

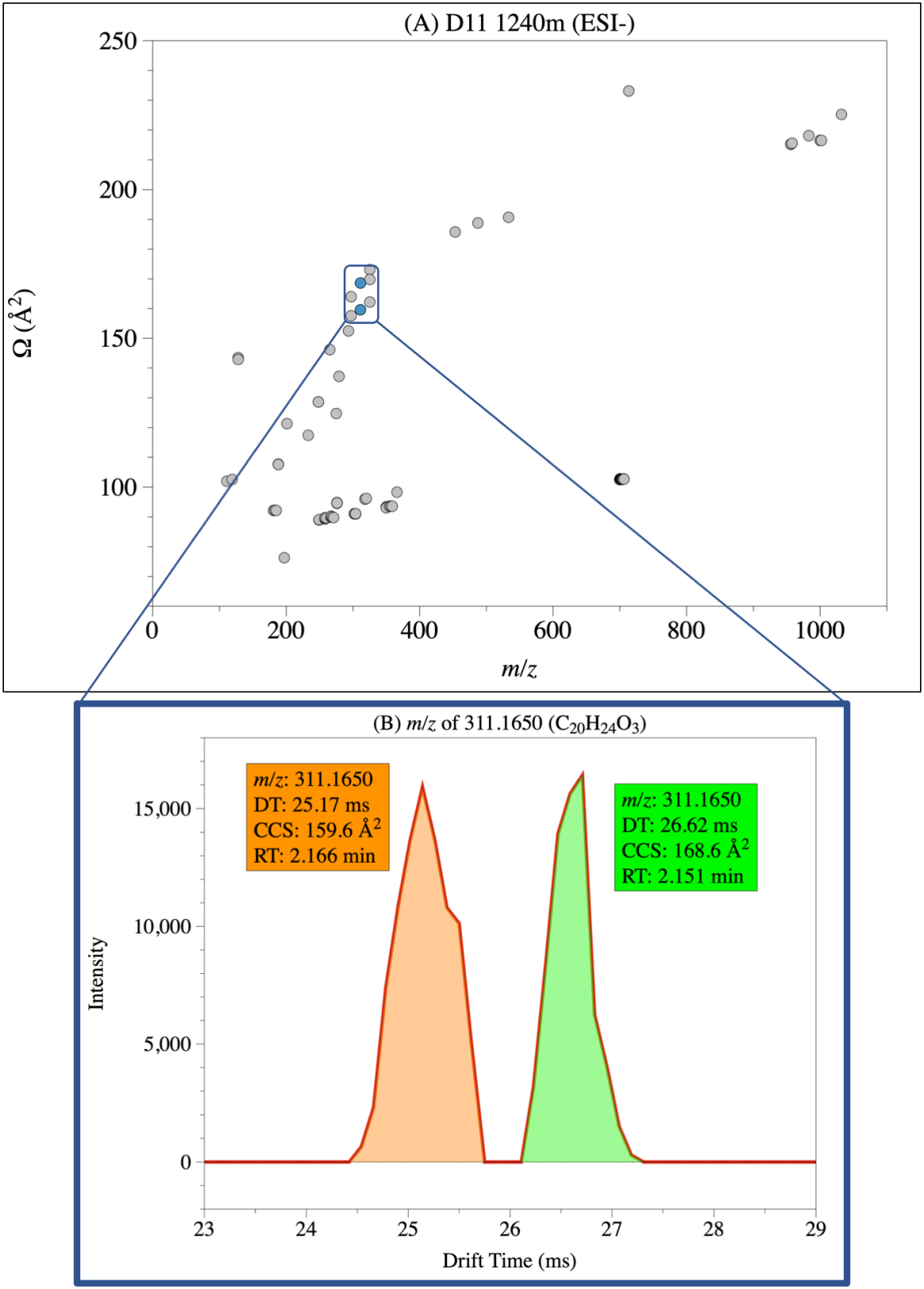
Figure 8. The max number of isomer clusters detected for the given *m/z* across different incubation samples (including SR and LR) at: (A) $T_i$ under ESI+; (B) $T_f$ under ESI+; (C) $T_i$ under ESI-; (B) $T_f$ under ESI-.

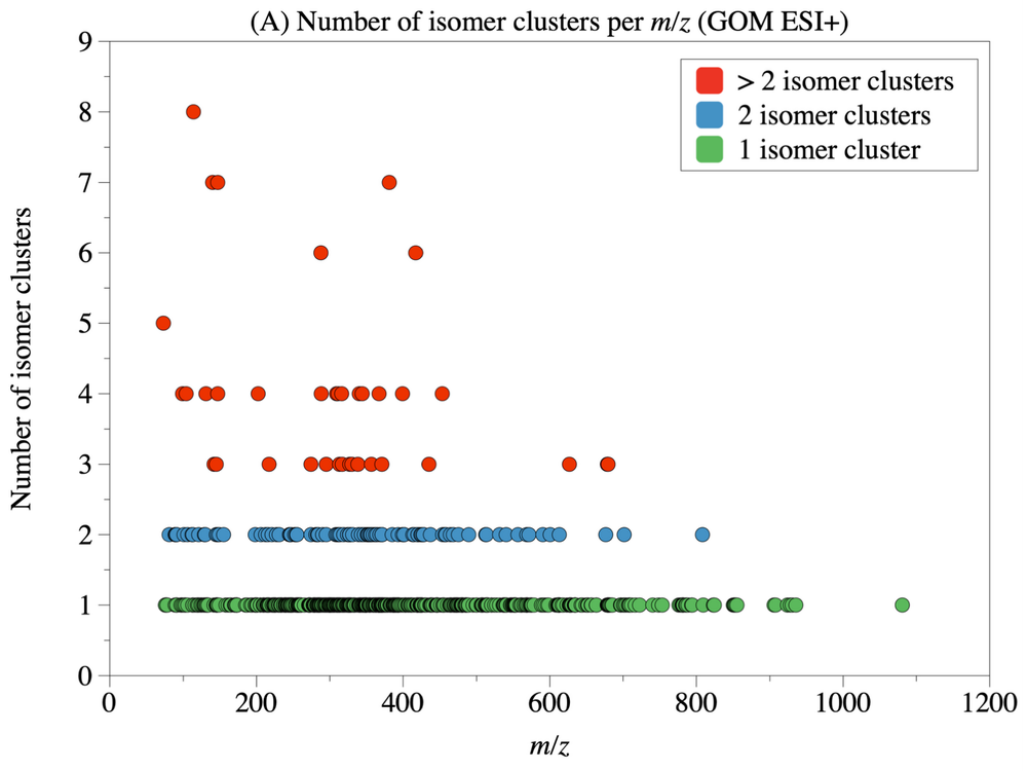

93°0'0"W
92°0'0"W
91°0'0"W
30°0'0"N
29°0'0"N
28°0'0"N
N
Atchafalaya River
AR
ARSC2 (Sal: 0.3 ‰)
ARSC1 (Sal: 3.0 ‰)
ARSC3 (Sal: 15.8 ‰)
Ship Channel
F1
F2
F3
Gulf of Mexico
D11 0.5 m
D11 500 m
D11 1240 m
D11
0
25
50 km
Sampling sites
Gulf of Mexico

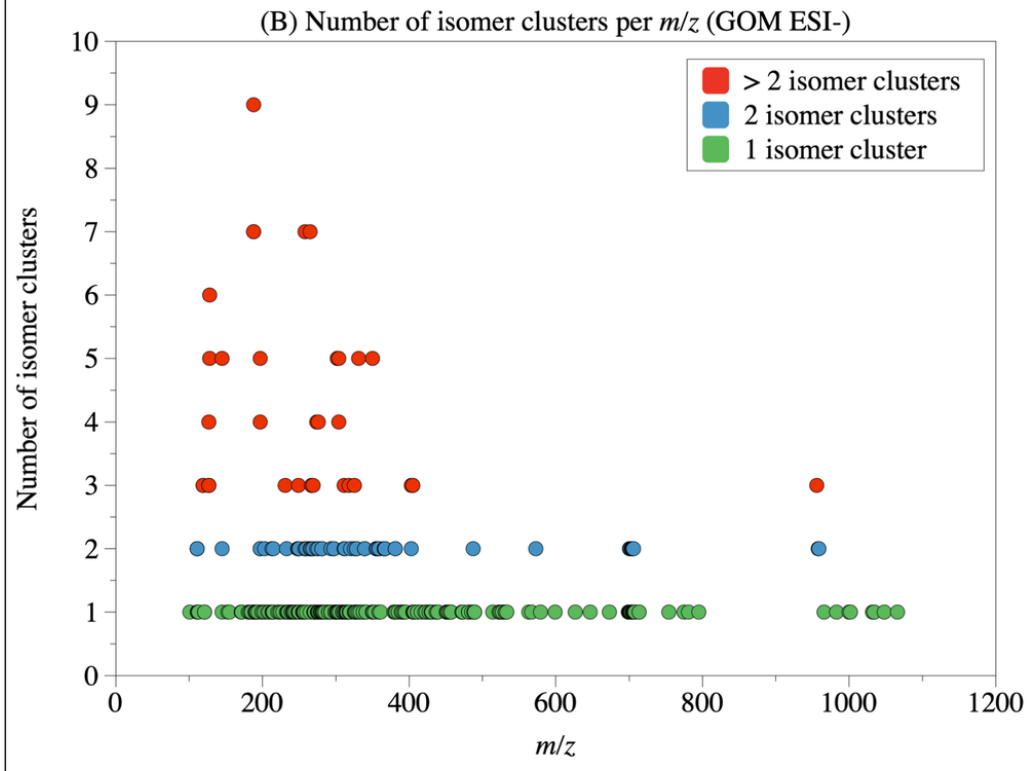

(A) DC6 5 m
(B) DC6 100 m
(C) DC6 500 m
(D) DC6 1000 m
(E) DC6 2000 m
(F) DC6 3000 m
(G) DC6 4000 m
H/C
O/C
Unique in QTOF
Unique in FT
Common

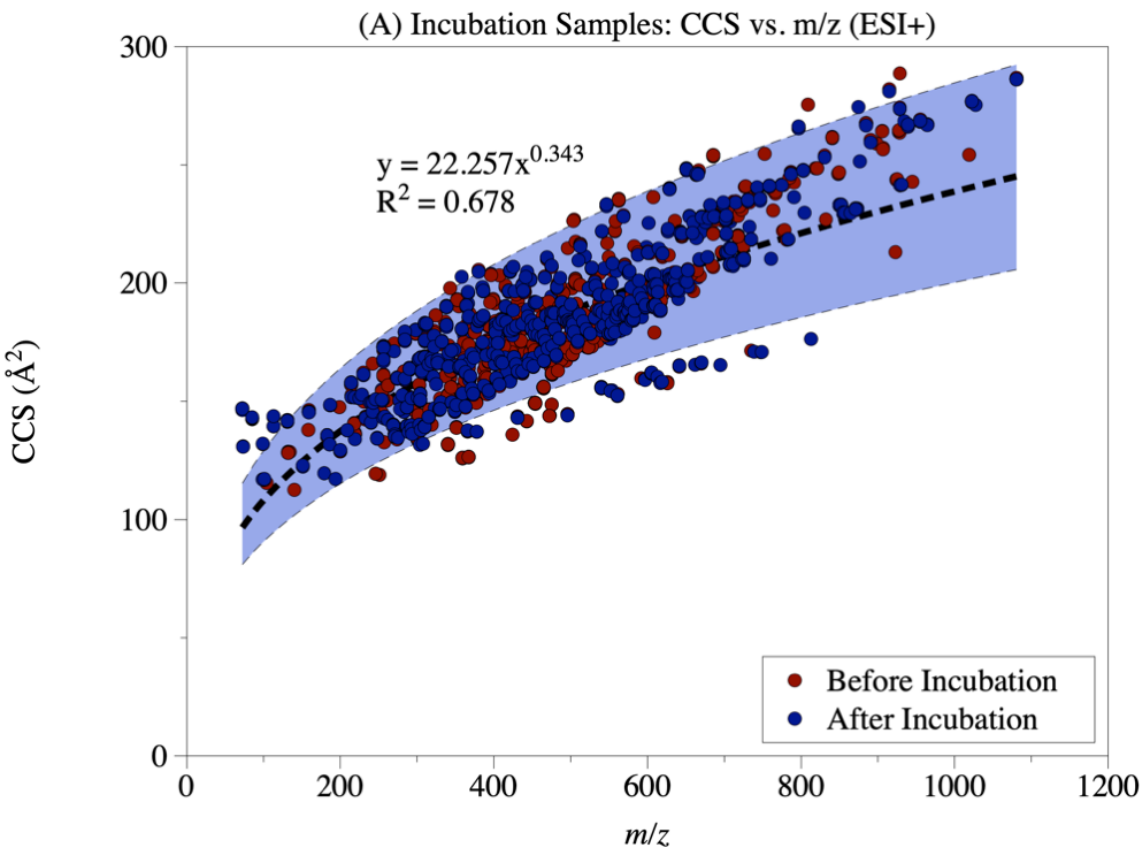

(A) Drift Time (ms) vs. m/z
Drift Time (ms)
35
30
25
20
460
470
480
490
500
510
520
530
540
m/z
(B) Peptide Isomers (m/z of 491.2195)
SDGRG
RT = 0.533 min
DT = 25.50 ms
GRGDS
RT = 0.537 min
DT = 27.36 ms
Relative Intensity (%)
100
80
60
40
20
0
24
25
26
27
28
29
Drift Time (ms)

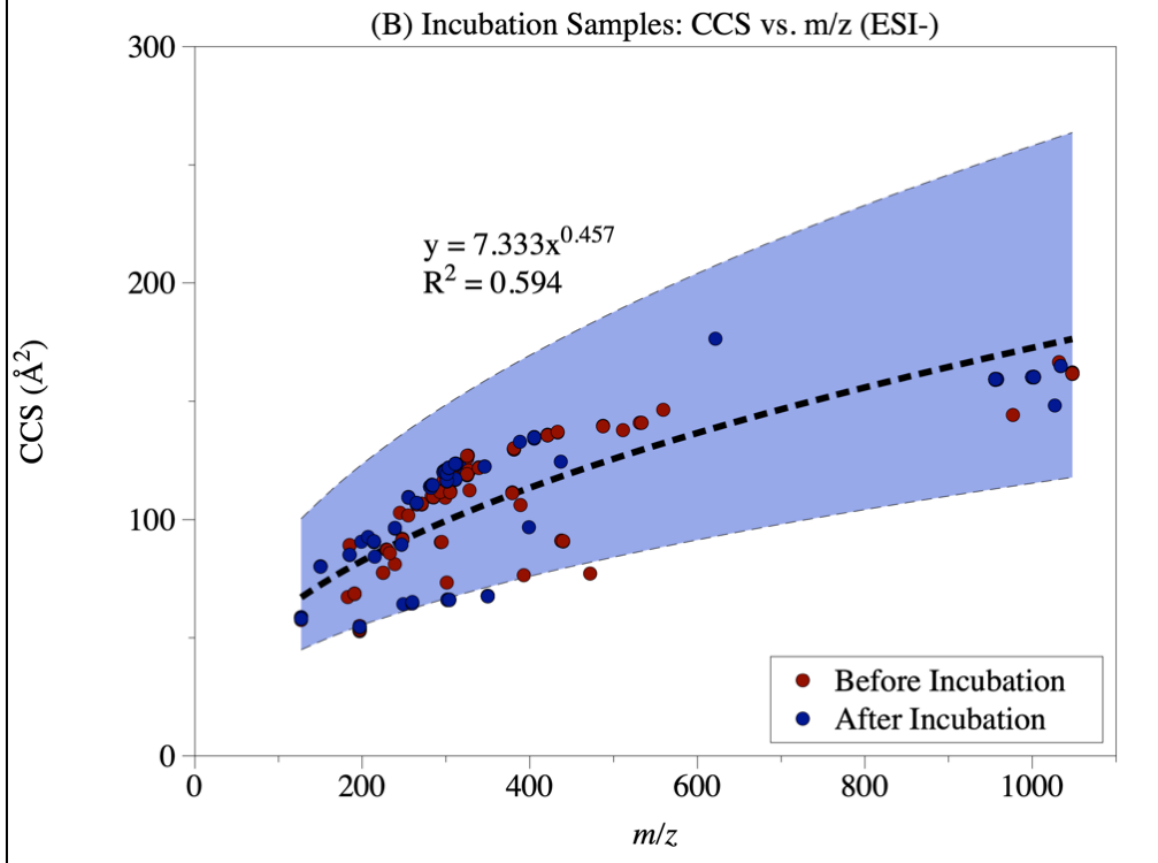

(A) GOM Samples: CCS vs. m/z (ESI+)
y = 21.749x^0.354
R^2 = 0.774
CCS (Å^2)
m/z
ARSC1
ARSC2
ARSC3
D11 0.5
D11 500
D11 1240
F1
F2
F3
(B) GOM Samples: CCS vs. m/z (ESI-)
y = 15.982x^0.383
R^2 = 0.830
y = 37.255x^0.156
R^2 = 0.760
CCS (Å^2)
m/z

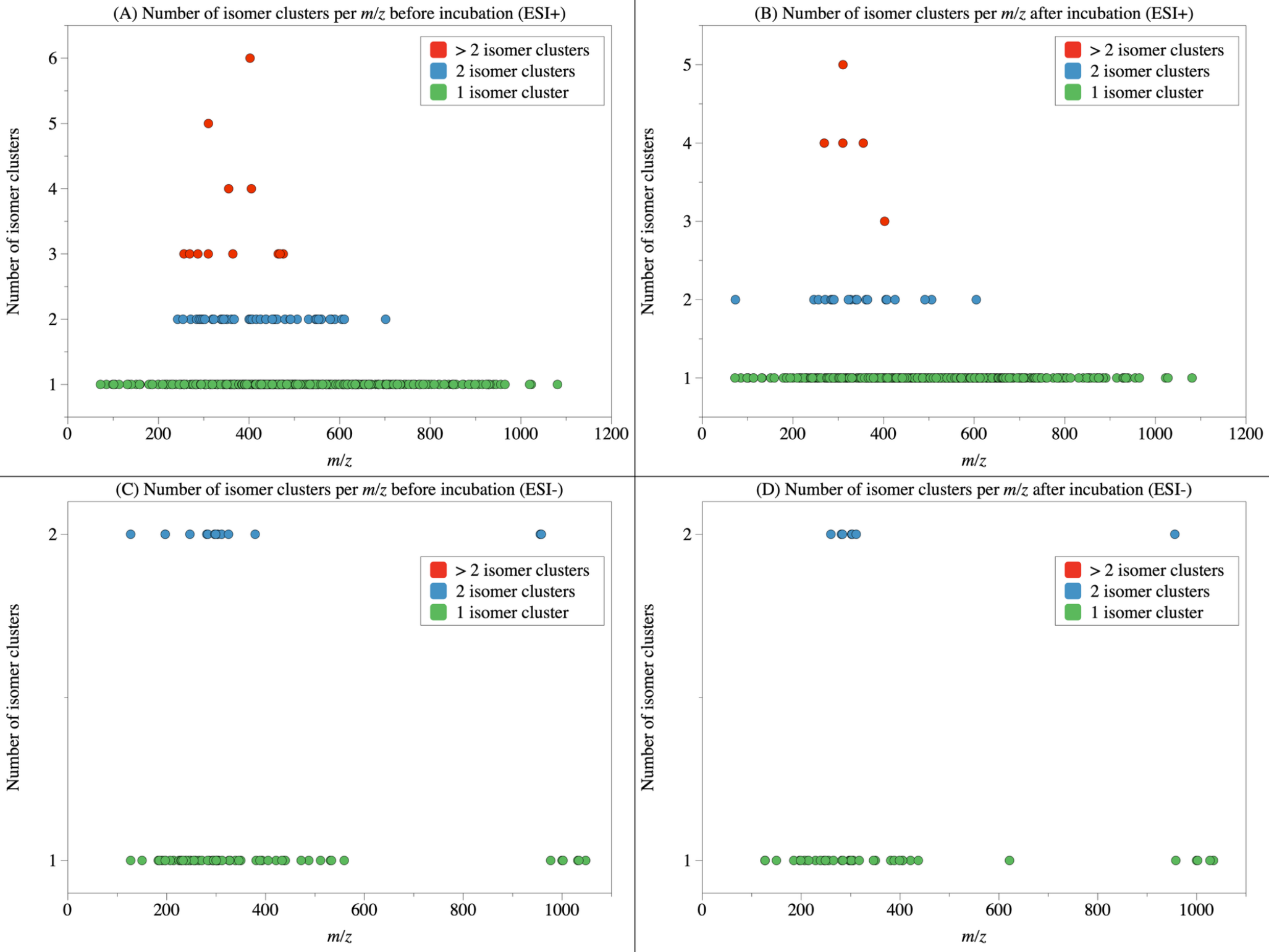

(A) D11 1240m (ESI-)
250
200
150
100
Ω (Å²)
0
200
400
600
800
1000
m/z
(B) m/z of 311.1650 (C20H24O3)
m/z: 311.1650
DT: 25.17 ms
CCS: 159.6 Å²
RT: 2.166 min
m/z: 311.1650
DT: 26.62 ms
CCS: 168.6 Å²
RT: 2.151 min
15,000
10,000
5,000
0
Intensity
23
24
25
26
27
28
29
Drift Time (ms)


(A) Number of isomer clusters per m/z (GOM ESI+)
> 2 isomer clusters
2 isomer clusters
1 isomer cluster
Number of isomer clusters
m/z
0
1
2
3
4
5
6
7
8
9
0
200
400
600
800
1000
1200


(B) Number of isomer clusters per m/z (GOM ESI-)
> 2 isomer clusters
2 isomer clusters
1 isomer cluster
Number of isomer clusters
m/z
0
1
2
3
4
5
6
7
8
9
10
0
200
400
600
800
1000
1200


(A) Incubation Samples: CCS vs. m/z (ESI+)
y = 22.257x^0.343
R^2 = 0.678
Before Incubation
After Incubation
CCS (Å^2)
m/z
0
100
200
300
0
200
400
600
800
1000
1200


(B) Incubation Samples: CCS vs. m/z (ESI-)
y = 7.333x^0.457
R^2 = 0.594
Before Incubation
After Incubation
CCS (Å^2)
m/z
0
100
200
300
0
200
400
600
800
1000


(A) Number of isomer clusters per m/z before incubation (ESI+)
(B) Number of isomer clusters per m/z after incubation (ESI+)
(C) Number of isomer clusters per m/z before incubation (ESI-)
(D) Number of isomer clusters per m/z after incubation (ESI-)
> 2 isomer clusters
2 isomer clusters
1 isomer cluster
Number of isomer clusters
m/z