Environmental Science & Technology

pubs.acs.org/est

Article

Two-Dimensional Liquid Chromatography Tandem Mass Spectrometry Untangles the Deep Metabolome of Marine Dissolved Organic Matter

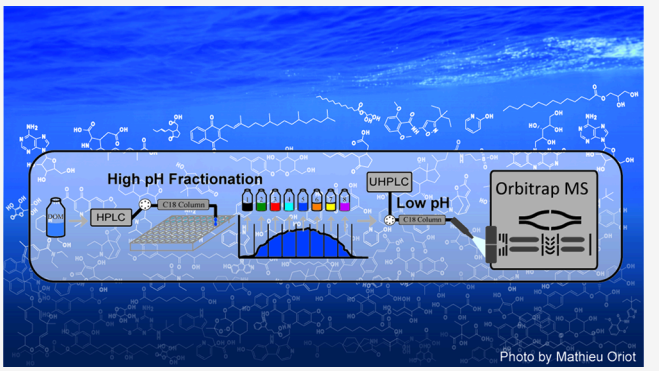
Stilianos Papadopoulos Lambidis, Tilman Schramm, Karoline Steuer-Lodd, Shane Farrell, Paolo Stincone, Robin Schmid, Irina Koester, Ralph Torres, Thorsten Dittmar, Lihini Aluwihare, Carsten Simon, and Daniel Petras*

Cite This: *Environ. Sci. Technol.* 2024, 58, 19289−19304

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Dissolved organic matter (DOM) is an ultracomplex mixture that plays a central role in global biogeochemical cycles. Despite its importance, DOM remains poorly understood at the molecular level. Over the last decades, significant efforts have been made to decipher the chemical composition of DOM by high-resolution mass spectrometry (HR-MS) and liquid chromatography (LC) coupled with tandem mass spectrometry (MS/MS). Yet, the complexity and high degree of nonresolved isomers still hamper the full structural analysis of DOM. To address this challenge, we developed an offline two-dimensional (2D) LC approach using two reversed-phase dimensions with orthogonal pH levels, followed by MS/MS data acquisition and molecular networking. 2D-LC-MS/MS reduced the complexity of DOM, enhancing the quality of MS/MS spectra and increasing spectral annotation rates. Applying our approach to analyze coastal-surface DOM from Southern California (USA) and open-ocean DOM from the central North Pacific (Hawaii), we annotated in total more than 600 structures via MS/MS spectrum matching, which was up to 90% more than that in iterative 1D LC-MS/MS analysis with the same total run time. Our data offer unprecedented insights into the molecular composition of marine DOM and highlight the potential of 2D-LC-MS/MS approaches to decipher the chemical composition of ultracomplex samples.



KEYWORDS: *dissolved organic matter, 2D chromatography, 2D-LC-MS/MS, tandem mass spectrometry, environmental metabolomics, molecular networking*

Downloaded via NATL UNIV OF IRELAND on August 20, 2025 at 11:20:35 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Dissolved organic matter (DOM) is one of the most complex natural chemical mixtures. Despite its ubiquity and major role in biogeochemical cycling, DOM remains poorly understood at the structural level. Especially in marine systems, DOM is of fundamental importance, representing the largest and most actively cycling pool of reduced carbon.[1−3] For example, DOM affects nutrient retention,[4,5] trace metal complexation,[6] long-term carbon sequestration,[3] and aerosol formation.[7−9] Missing molecular information, however, hampers our understanding of ecosystem functioning and potential feedback loops under climate change scenarios.[1,3,10]

Many studies investigate the molecular composition of DOM by direct injection (DI) ultrahigh-resolution mass spectrometry (HR-MS) via Fourier transform ion cyclotron resonance (FT-ICR) or Orbitrap mass spectrometers. While these techniques allow unprecedented resolution at the molecular formula level, complementary coupled techniques such as tandem mass spectrometry (MS/MS),[11−14] liquid chromatography−mass spectrometry (LC-MS),[15−18] and ion mobility mass spectrometry (IMS)[19,20] all suggest a high degree of nonresolved isomeric complexity. Although prior attempts to resolve this complexity via LC-MS have shown promising results, they typically resolve the more abundant components of DOM and thus are more suitable for freshly produced metabolites and xenobiotics and less suitable for the diverse refractory components of DOM that are present at lower concentrations. In addition to inherent problems of subsampling due to a limited number of possible MS/MS scans

Received: July 13, 2024
Revised: October 6, 2024
Accepted: October 8, 2024
Published: October 16, 2024



ACS Publications

© 2024 American Chemical Society

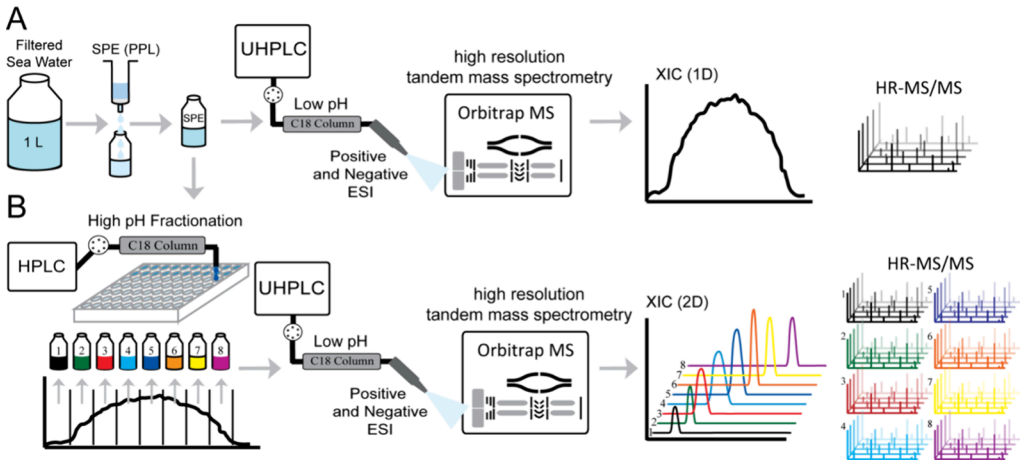


Figure 1. Schematic workflow of 1D and 2D-LC-MS/MS analysis. (A) 1D LC-MS/MS setup. Filtered seawater is extracted using a PPL cartridge and separated in a UHPLC with a C18 column and low-pH mobile phase before positive and negative electrospray ionization (ESI) and subsequent high-resolution tandem mass spectrometry. Exemplary extracted ion chromatogram (XIC) indicated insufficient separation of isobaric compounds ("DOM-Mountain"). (B) 2D-LC-MS/MS setup. After solid-phase extraction, samples were separated using an HPLC with a high-pH mobile phase and a C18 column (1D) and fractionated into multiple fractions. Fractions were further separated in a UHPLC with a low-pH mobile phase and a C18 column before subsequent positive and negative ESI spectra were recorded in a high-resolution tandem mass spectrometer. Exemplary XICs showing the separation of isobaric compounds.

per LC-MS/MS run, a central bottleneck is the annotation of the resulting spectra.[21] Here, the use of molecular networking (MN) offers a promising avenue to resolve this complexity.[22] MN allows for the internal comparison of MS/MS spectra based on their spectral similarity as well as the matching of spectra against existing data sets and growing community-based reference libraries.[23−25] This way, large LC-MS/MS data sets can be processed, and similar molecular features can be grouped to propagate chemical class-level annotations.[21,22,26,27] By making use of the concept of annotation propagation, comprehensively annotated reference samples could lay the foundation for a detailed exploration of DOM data sets, as reference MS/MS spectra for most compounds present in DOM are still limited.[26,28]

DOM analyzed by LC-MS/MS setups suffers from insufficient chromatographic resolution, leading to so-called DOM "mountains", "humps", or "unresolved complex mixture" of mass features[16,29,30] (Figure 1A). Furthermore, the structural-level annotation of DOM often remains less than 10% of detected features.[15,21,31] This reflects not only the insufficient coverage of MS/MS duty cycles and chemodiversity in public spectral libraries but also the insufficient separation of molecular features, leading to the acquisition of low-quality "chimeric" tandem mass spectra.[13,15,32] To improve the chromatographic resolution of DOM, multidimensional separations via gas chromatography[33,34] and liquid chromatography[30,35] have been attempted, but they have not yet led to high-throughput routines.

In theory, the problem of insufficient chromatographic resolution can be overcome by the use of two-dimensional (2D) LC. 2D LC approaches can be divided into two major methodologies, heart-cutting and comprehensive 2D LC. Briefly, heart-cutting LC includes the collection of a single fraction of 1D eluent and its subsequent injection into a 2D column for further separation, whereas comprehensive 2D LC collects discrete fractions at regular intervals and these fractions are transferred to a 2D column.[36] Furthermore, 2D LC can be divided into online and offline methods.[37] In online 2D LC approaches, the second-dimension separation is directly coupled to the first LC step, before reaching the detector. Therefore, it requires two LC systems connected by a valve with sample loops and a fast second dimension to prevent large volumes and diffusion effects in the 1D sample loop. Major limitations of online 2D LC include the short duration of the second dimension and the need of a second HPLC instrument with a suitable modulator (e.g., multiport valves). Offline 2D LC, on the other hand, can be performed using a standard HPLC instrument, as the effluent of the first LC step is collected and reinjected at any given time point into the same instrument after changing the LC condition. This increases the number of usable orthogonal second-dimension LC methods, making it possible to reach the highest peak capacities.[38]

Among the available 2D setups, separation by two reversed-phase LC steps at varying pH has shown to be most effective due to the high peak capacity, orthogonality, and high compatibility of mobile phases with evaporation and electrospray ionization.[39−41] Gilar and co-workers evaluated the orthogonality of different 2D setups working with peptides, which share structural features (organic acids and amines) with marine DOM, and found varying pH (high/low) to yield the highest orthogonality.[42] Such 2D setups have been widely used in the proteomics field to improve the identification depth of peptides.[43] For example, Wang and co-workers reported 2.2- and 2.7-fold increases in protein and proteoform identifications, respectively, compared to a traditional 1D LC-MS approach.[40] Similarly, McIlvin and Saito reported an online 2D LC-MS setup that led to a 1.5−2-fold higher protein identification rate in marine meta-proteome samples, arguing that the diverse assemblages of organisms and ecological interactions led to a wide array of low-abundance protein features.[44] Combined high-pH/low-pH separations for the MS/MS analysis of marine DOM represent a promising application in environmental metabolomics, as this field faces similar problems as in (meta)proteomics samples, such as high sample complexity and large numbers of rare features. The orthogonality of pH over the two chromatographic dimensions might be especially suitable for DOM, as many compounds contain carboxylic acids or amines, which would drastically

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

change their protonation and thereby their retention to the stationary phase with different pH.

Making use of this concept, we adapted a 2D separation with a first (1D) offline high-pH fractionation, followed by a second (2D) low-pH separation in both negative and positive ESI ionization, followed by data-dependent acquisition (DDA) of MS/MS spectra (Figure 1B). Applying and comparing these workflows to our standard 1D LC-MS/MS approaches from a coastal, open ocean surface, and a deep-sea DOM sample, we show that the 2D approach effectively reduces sample complexity while enhancing tandem MS data quality and coverage, leading to a higher coverage and annotation rate of DOM.

MATERIALS AND METHODS

Sample Collection and Processing of Coastal DOM. As a DOM reference sample, we collected 200 L of surface seawater with a 20 L bucket from the Ellen Browning Scripps Memorial Pier (32°52′001.500N 117°15′026.900W) in La Jolla, Southern California, USA, on the 26th of February 2021 between 11:00 and 19:00 PDT. Seawater was filtered through 0.7 μm GF/F filters. 25 L aliquots of the sample were acidified with 26 mL of hydrochloric acid (37% p.a., trace metal grade, J.T.Baker, pH 2). 25 L of seawater was then extracted via PPL cartridges with a bed mass of 5 g (Bond Elut, Agilent) according to previous studies.[45,46] Before use, the PPL cartridges were activated with 30 mL of MeOH and then rinsed with 30 mL of H_2O (pH 2, LC-MS grade) and 30 mL of MeOH (LC-MS grade) followed by an equilibration with 30 mL of H_2O (pH 2, LC-MS grade). For sample loading, acidified seawater was pulled through eight SPE PPL cartridges in parallel. We used a vacuum SPE station (Agilent 20 port SPE station) to maintain a flow rate of approximately 20 mL/min/cartridge for a total loading time of 20 h. A process blank was collected using 4 L of acidified H_2O (pH2, LC-MS grade) onto a 5 g cartridge using the same SPE protocol as that above.

After sample loading, the cartridges were desalted with 60 mL of H_2O (pH 2, LC-MS grade) and dried under N_2 gas. After drying, the cartridges were eluted with 20 mL MeOH per cartridge resulting in a total of 200 mL eluted sample to which an internal standard mix was added that contained 5 μg each of domoic acid, kanic acid, isoxaben, irgarol, imazapyr, heroin, methamphetamine, and cocaine, which have been observed in similar studies in that region.[47] The pooled sample was then aliquoted to 100 individual 2 mL HPLC vials. All vials were dried in a vacuum centrifuge overnight at room temperature, resulting in 1.8 mg of total organic matter per vial.

Two representative aliquots were redissolved in H_2O (pH 2) for TOC measurements, resulting in 30.0 μM (for 2 L) and 0.72 mg C/vial. The overall extraction efficiency was 43.3% considering the DOC concentration of the initial seawater sample of 69.3 μM carbon.

For 1D and 2D-LC-MS/MS samples were redissolved in 200 μL of 80/20 MeOH/H_2O.

Sample Collection and Processing of Open Ocean DOM. We used in-house reference DOM samples from North Equatorial Pacific Intermediate Water (NEqPIW) and the surface oligotrophic North Pacific Ocean. NEqPIW is one of the oldest oceanic water masses. Samples were isolated via solid-phase extraction after filtration and at pH 2 using the same PPL resin as that for the other samples considered in this study. Very large volumes (up to 30,000 L) of NEqPIW and surface water were obtained at the pumping station at the Natural Energy Laboratory of Hawaii Authority (NELHA) from 674 and 21 m water depths, respectively. Details on the large-scale DOM isolation are described in Green et al.[48] The original DOC concentrations were 44 μM (NEqPIW) and 71 μM (surface waters), from which 61 ± 3% was recovered via PPL-SPE. These in-house DOM reference samples have been extensively characterized previously, e.g., via tandem FT-ICR-MS[14] or two-dimensional ^{1}H nuclear magnetic resonance spectroscopy. 1 mg solid-phase extracted DOM was redissolved in 200 μL of 80/20 MeOH/H_2O for 1D and 2D-LC-MS/MS analyses.

Prefractionation by High-pH Reversed-Phase Liquid Chromatography. Separation of samples in the first dimension was performed using a 1260 Infinity II semi-preparative HPLC (Agilent) system with mobile phase A: H_2O + 0.05% NH_4OH and mobile phase B: ACN + 0.05% NH_4OH, with the mobile phases having a pH around 10. For the chromatographic separation, a reversed-phase C18 porous core column (Kinetex EVO C18, 150 mm × 4.6 mm, 2,6 μm particle size and 100 A pore size, Phenomenex) was used. The flow rate was 1 mL/min, and the elution was performed with a linear gradient, from 0 to 8 min from 95 to 50% A, and then from 8 to 10 min 50−1% A, holding 1% A for 3 min to wash out the column followed by a 3 min re-equilibration phase at 95% A. A total of 250 μL was injected in 10 iterative HPLC runs (each 25 μL) and fractionated using the integrated fraction collector. Fractions were collected in a time-dependent manner into a 96-well deep well plate (Brand), changing the fraction after every minute, yielding 10 fractions. For an easier workflow, we also collected the first two fractions of the washout, 11 and 12, so that the same fractions are in the same column of the plate. Corresponding fractions between runs were pooled into vials before evaporating the solvent in a SpeedVac (Concentrator plus, Eppendorf). The samples were redissolved in 100 μL of 80:20 MeOH:H_2O.

Low-pH Reversed-Phase Liquid Chromatography Tandem Mass Spectrometry. DOM samples and the fractions (first dimension) were separated using the same LC gradient as in the prefractionation, but different mobile phases, with mobile phase A: H_2O + 0.1% formic acid and mobile phase B: ACN + 0.1% formic acid (pH ∼ 2), were used. The eluting molecules were detected by the coupled Q Exactive HF Orbitrap mass spectrometer. DOM constituents were ionized in positive or negative electrospray ionization (ESI) mode. ESI settings were as described previously.[15,18] In short, gas flows (in AU) were set to 52 AU (sheath gas), 14 AU (auxiliary gas), and 0 AU (sweep gas); spray voltage was set to 3.5 kV; inlet capillary was heated to 320 °C; and S-Lens level was at 50%. MS1 spectra were acquired with a scan range of 150−1500 m/z at a nominal resolution of 140,000 at m/z 200 and automatic gain control (AGC) set to one million charges. Maximum injection time was set at 100 ms. MS/MS spectra were acquired at a lower AGC target of 300,000 charges with a minimum 10% C-trap filling. Precursors were selected within a 1 m/z window width. Collision-induced dissociation (HCD) was conducted stepwise at 25, 35, and 45% and at a charge state (z) of 1. MS/MS acquisition was triggered at the peak apex within 2−15 s after the first precursor detection. Precursors were then dynamically excluded for 5 s, as well as precursors with unassigned charge states or isotope peaks. Retention time drifts were controlled with a quality control mix of six standards evenly distributed in the chromatogram (sulfamethazine, sulfamethizole, sulfachloropyridazine, sulfadi-

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

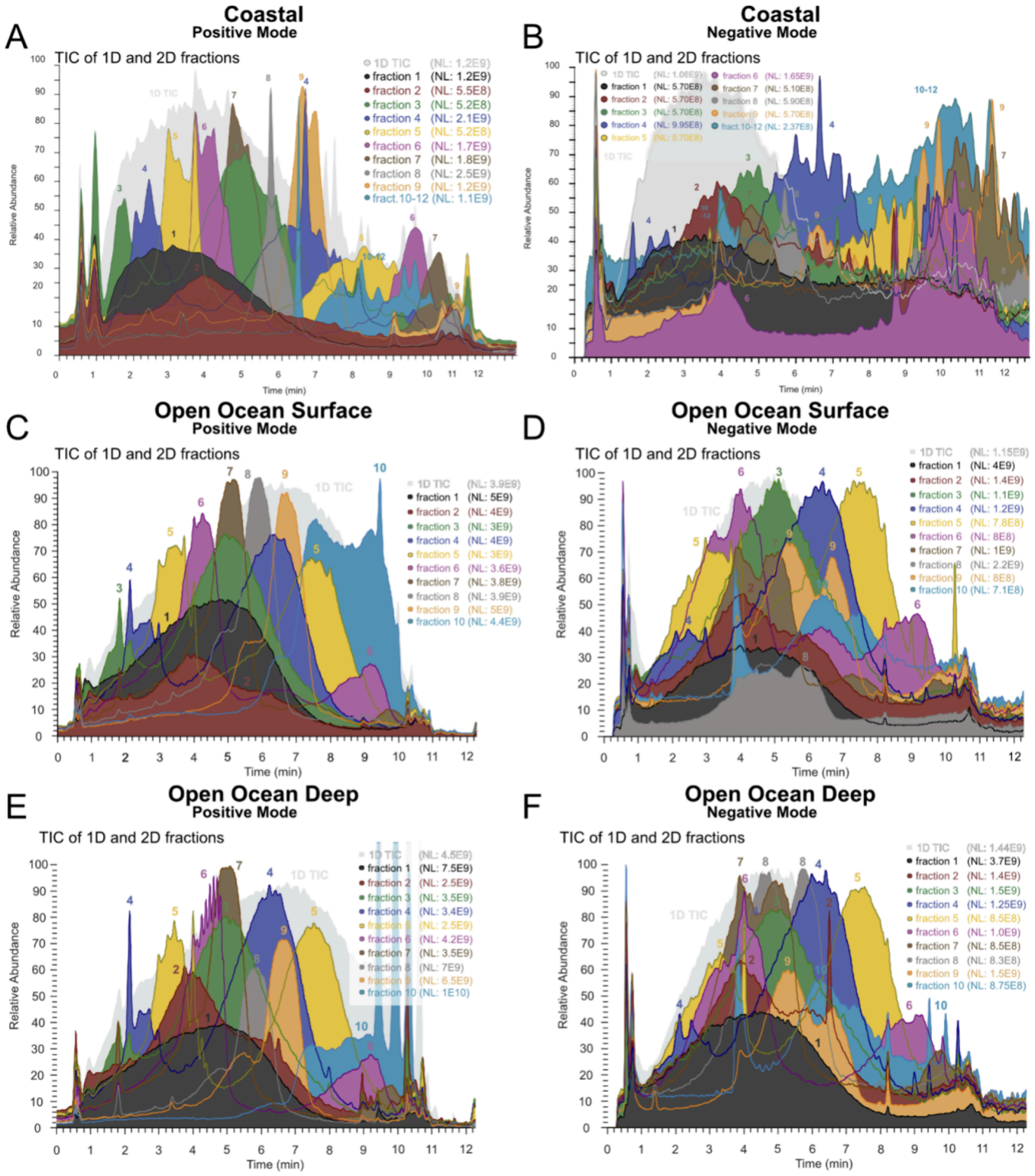


Figure 2. Comparison of 1D and 2D total ion chromatograms showing improved separation of marine DOM.(A) Overlay of the total ion chromatograms (TICs) of the single fractions of the coastal sample (A, B) and open ocean surface (C, D) and open ocean deep water sample (E, F) in both positive and negative ionization modes.

methoxine, amitriptyline, and Coumarin 314) at the beginning and the end of the sequence.

Identification and Quantification Using Authentic Standards. Authentic standards for triallyl cyanurate (Thermo Scientific), isoxaben (Sigma-Aldrich), *N*-lauroylsarcosine (Thermo Scientific), simetryn (Sigma-Aldrich), amitriptyline (Sigma-Aldrich), and hydroxybupropion (Sigma-Aldrich) were used to prepare stock solutions in methanol (LC-MS grade, Honeywell) with a concentration of 1 mg/mL. Stock solutions were mixed and diluted into glass vials to yield a final stock solution of all standards at a concentration of 100 ng/mL. Standard solutions with 10, 1, 0.1, and 0 ng/mL were

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

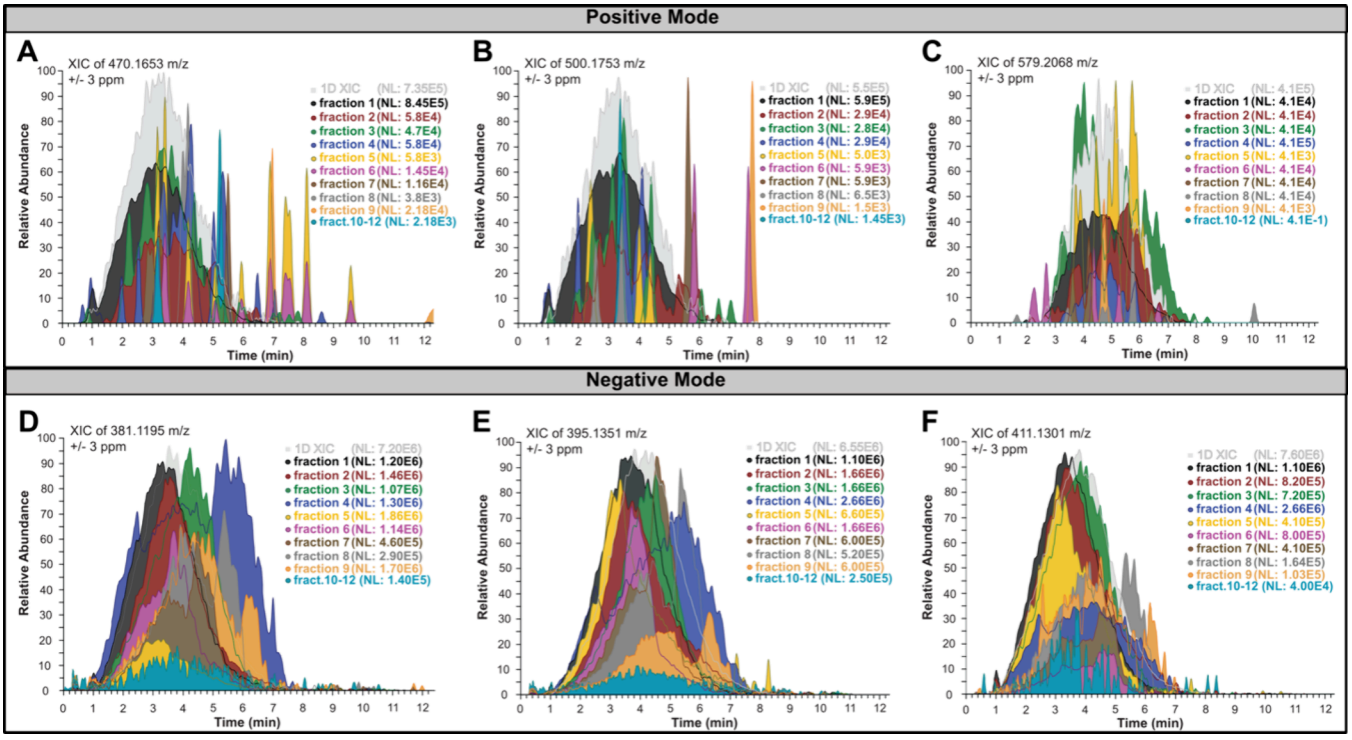


Figure 3. Comparison of 1D and 2D extracted ion chromatograms showing improved separation of frequently observed molecular features in coastal DOM. Overlay of the extracted ion chromatograms (XIC) for (A) m/z 470.1653, (B) m/z 500.1753, and (C) m/z 579.2068 of the single fractions of the coastal sample of the 2D analysis over the XIC of the 1D analysis in positive mode (gray, in background). XICs from negative mode are shown for (D) m/z 381.1195 (E) m/z 395.1351 (F) m/z 411.1301 of the single fractions of the coastal sample of the 2D analysis over the XIC of the 1D analysis (light gray, in background). XICs of the same masses for the open ocean samples are shown in Figure S1.

prepared using a dilution series, transferring 100 μL stock solution into glass vials containing 900 μL of methanol. Stock solutions were measured in triplicate using the same LC-MS/MS method used for the 1D measurements and 2D fractions.

To unambiguously validate these compounds in the samples, extracted ion chromatograms (XICs) of the standards and the samples were overlaid before comparing the retention time and MS2 spectra. Feature intensities of the standards were extracted and used to construct standard curves to estimate the approximate concentration of the compounds in the sample. Taking into account the dilution factor from the extraction and the extraction efficiency of PPL cartridges, we could estimate the approximate concentration of selected compounds in the seawater sample.

LC-MS/MS Feature Detection and Molecular Networking. MS1 features were extracted with MZmine 3[49] at a signal threshold of 10,000 and MS2 features with a signal threshold of 1000. MS1 features were then transformed to XICs using five consecutive scans at a minimum signal threshold of 100,000 and 5 ppm mass tolerance. Thresholds for deconvolution were set to 300,000 minimum absolute height, a 1.5 min peak top-to-edge ratio, and a maximal duration of 5 min. Isotope peaks were grouped at tolerances of 3 ppm for mass and 0.1 min for retention time. Alignment of XICs between samples was conducted using 5 ppm tolerance, and only XICs grouped with >1 isotope peaks were considered for further analysis. The aligned XIC list was checked for duplicates at 3 ppm mass and 0.1 min retention time tolerances.

From MZmine, the precursor purity information was exported and used to evaluate the presence of the chimeric spectra. Classical MN was used to evaluate the spectral quality and annotate by matching MS/MS spectra with the GNPS libraries.[23] The MZmine 3 output quant table and .mgf files were used to create feature-based molecular networking (FBMN) in GNPS.[24] Detailed settings and urls of processed jobs are provided in Tables S2−S5.

***In Silico* Spectrum Annotation.** SIRIUS (5.8.5)[50] was used for *in silico* annotation of tandem mass spectrometry data. Using the SIRIUS module, molecular formulas were computed by matching experimental with predicted isotopic patterns from the fragmentation tree analysis of MS/MS. Parameters for SIRIUS are as follows: instrument—Orbitrap, MS/MS ppm—5, isotope scorer—ignore, candidates stored—10, min candidates per ion—1, databases used—Bio Database, possible ionizations—pos/neg, tree timeout—0, compound timeout—0, use heuristics above m/z—300, use heuristics only above m/z—650, and only molecular formulas present in the Bio Databases were considered. *In silico* structure annotations were performed with CSI:FingerID[51] and structures from the Bio Database.

Visualization of Molecular Networks. The visualization and analysis of the networks were carried out using Cytoscape (Version 3.9.1).[52] For that, the obtained table from the feature-based network analysis was thinned out, eliminating redundant SMILES codes and structural formulas. Nodes represent detected compounds and their drawn chemical structure. Chemical similarity was calculated using Tanimoto similarity (cutoff 0.8), creating a connection between the nodes using the ChemViz add-on.

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

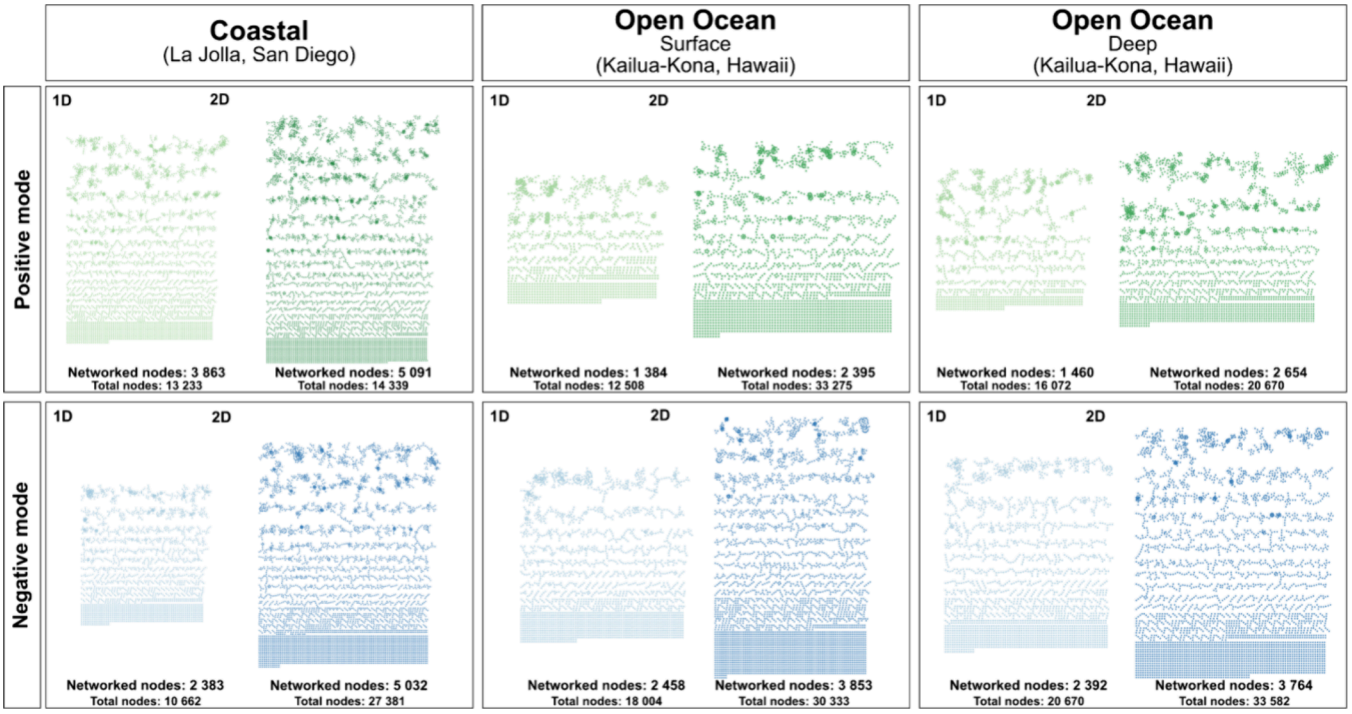


Figure 4. 1D and 2D LC-MS/MS-based molecular networks, showing larger numbers of networked and overall nodes for the 2D setup. Classic molecular networks of both positive mode (green, upper row) and the negative mode (blue, lower row) with the 1D and 2D methods (lighter and darker colors, respectively). Data contained 20 replicate injections for 1D and duplicated injections for the 10 fractions for 2D (20 total injections). Only networked nodes shown, i.e., nodes connected to at least one other node.

RESULTS AND DISCUSSION

Improved Chromatographic Resolution Extends the Depth of Annotated Molecules in Marine DOM. To improve the chromatographic resolution and analytical depth for the molecular characterization of DOM, we developed an offline 2D UHPLC-MS/MS method. Our assumption herein is that improved chromatographic resolution should lead to cleaner and more unique MS2 spectra, especially in the data-dependent acquisition (DDA) mode. In DDA, precursors are selected for isolation and fragmentation by their intensities from the previous MS1 scans. With higher chromatographic resolution, different compounds are selected for isolation and fragmentation, and spectra should be cleaner, as fewer isomers and compounds with similar masses should be present in the quadrupole isolation window.

Our 2D-LC workflow relies on the fractionation of the DOM sample with a C18 column with a high-pH mobile phase with 0.05% NH_4OH as a modifier for the first dimension. Technical duplicates of corresponding fractions are then injected into a second dimension with a C18 column and low-pH mobile phase with 0.1% formic acid, before MS/MS data acquisition (20 injections per ESI+ and ESI− mode).

To compare the 2D results to our classic 1D LC-MS/MS workflow, we injected the same samples from coastal California (Scripps Pier, San Diego, 2021) and an open ocean environment (NELHA Pipeline, Hawaii, 2009) in each 20 consecutive LC-MS/MS runs for ESI+ and ESI− mode, yielding the same MS/MS analysis time and theoretical maximum number of spectra. To assess possible improvements in the chromatographic resolution, we first compared the total ion chromatograms (TICs) of a classical 1D LC-MS/MS run and our 2D approach and subsequently compared XICs of representative high abundant masses (470.1653; 500.1753; 579.2068 m/z in pos mode and 381.1195; 395.1351; 411.1301 m/z in neg mode), as shown in Figures 2 and 3 and Figure S1.

The TIC obtained from the 1D runs (Figure 2, light-gray TICs) shows the aforementioned "DOM-Mountain", showing poor resolution of the peaks, whereas the TICs obtained from the technical replicates from the 2D runs show narrower and more defined peaks that resemble the fraction order from the first dimension. Importantly, the retention time of the main peaks in all fractions is clearly distinct, indicating a constant and reproducible interaction of the DOM component with the stationary phase and likely less colloidal interactions between DOM itself, as proposed in previous studies.[30,53−55]

The effect of this sample decomplexation becomes more pronounced when looking at specific XICs. In Figure 3A−C, the XICs of three highly abundant masses in positive mode, m/z 470.1653 (± 3 ppm (Figure 3 A)), m/z 500.1753 (± 3 ppm (Figure 3 B)), and m/z 579.2068 (± 3 ppm (Figure 3 C)), are shown. The XIC obtained from the 1D run is shown in gray in the background, and the XICs obtained from the 2D run are overlaid. While the 1D XICs show poor chromatographic resolution and a "Mountain-like" shape again, the XICs obtained from the 2D run show at least partially resolved peaks, with detection of molecules with the same m/z in multiple different fractions. While this effect is more pronounced in the positive mode, clearer peaks and a better resolution are also visible for frequently observed features in the negative mode, shown in Figure 3D−F for m/z 381.1195 (± 3 ppm (Figure 3D)), m/z 395.1351 (± 3 ppm (Figure 3E)), and m/z 411.1301 (± 3 ppm (Figure 3F). Looking at the same XICs in the more complex open ocean samples (Figure SI1), similar chromatographic profiles to those of the individual fractions can be seen. Together, these results suggest improved resolution of isomeric and isobaric

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

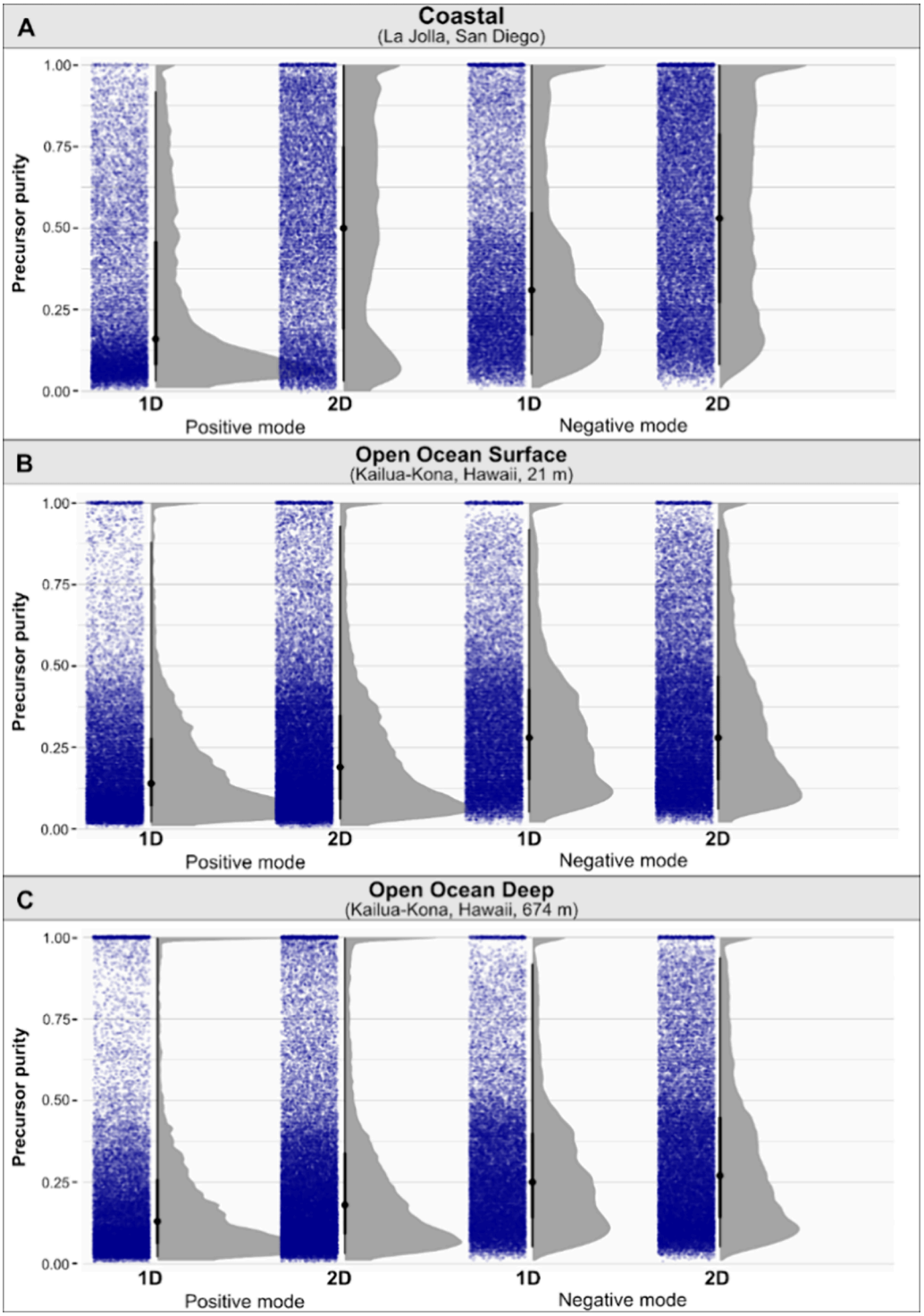


Figure 5. Distribution of MS/MS precursor purity. MS/MS precursor purity was evaluated using the precursor purity checker module in MZmine 3.9.0, which ranks the purity between 1 and 0, with 1 representing a "clean" precursor selection for MS/MS. Precursor purity is shown as a raincloud plot, comparing between the 1D and 2D analysis of positive (left) and negative modes (right), with (A) showing the coastal sample, (B) the surface, and (C) the deep sea sample. The single spectra are depicted as dots, the distribution is depicted as a half-violin plot, the black dot indicates the median, the thick black line the interquartile range (IQR), and the thin black line the 95% confidence interval.

compounds by the orthogonal 2D LC, which is crucial to increase MS/MS purity and, ultimately, MS/MS-based identification.

Evaluation of Improved MN and Spectral Quality. We employed two forms of molecular networking: classical MN and feature-based molecular networking (FBMN). Briefly, classical MN is a widely used bioinformatics tool for the visualization and annotation of nontargeted MS data. In addition to spectral matching against reference spectra, it aligns all experimental spectra of a data set against one another and connects related molecules by their spectral similarity. This is achieved by clustering MS2 spectra and using the obtained

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

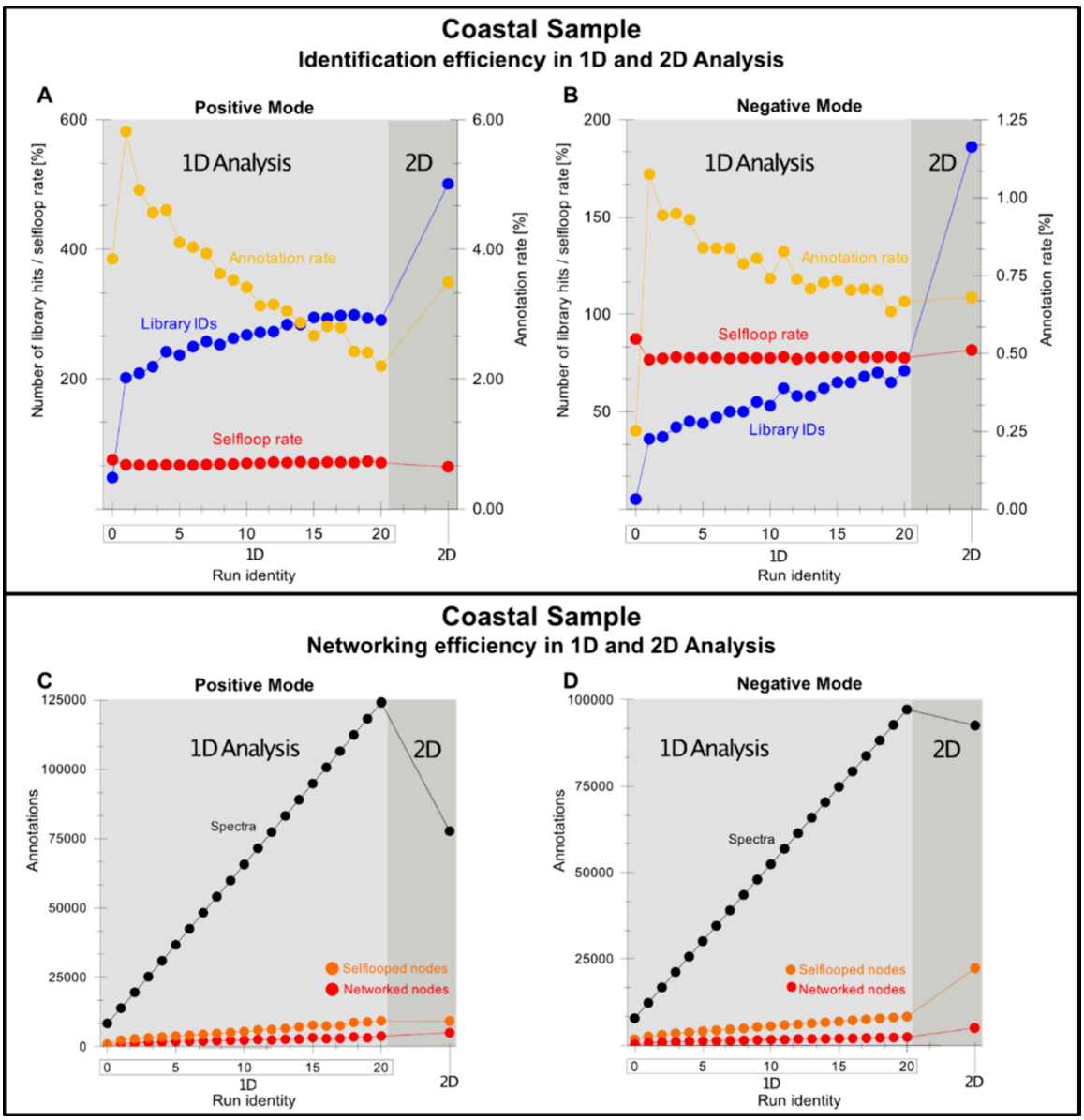


Figure 6. Increase in Identification and Networking efficiency between 1D and 2D analysis. (A, B) Comparison of annotation efficiency from classic molecular networking results of the coastal sample's 1D and 2D analysis in positive mode (A) and negative mode (B). (C, D) Comparison of networking efficiency in 1D and 2D analysis in positive mode (C) and negative mode (D). Self-looped nodes are nodes not connected to any other node; self-loop rate is the percentage of overall nodes being self-looped.

consensus MS2 spectra for MN.[56] We employed classical molecular networking using the GNPS infrastructure and created classical molecular networks from either the 20 2D runs or the 20 1D runs (Figure 4).

As a first result, the classic molecular networks obtained from the 2D analysis show more nodes overall but also more complex networks, having more connected nodes than the 1D networks, as shown in Figure 4. The difference is far more pronounced in the negative mode, where an additional strong increase in self-looped nodes (nodes not connected to other nodes) is visible. This result could be indicative of the improvement of carboxylic acid-rich refractory components that are better ionized and typically more dominant in negative ESI mode. Larger and more complex molecular networks indicate a better overall spectral quality of the MS2 spectra, as chimeric spectra would be more random and thus less likely to be similar to other spectra (e.g., single nodes).

While coastal DOM typically includes more labile components and higher abundant features, e.g., from terrestrial runoff or recent algal production, we anticipate that the 2D approach also improves spectral quality and annotation from refractory, deep-sea DOM. To test this, we included surface and deep-sea samples from the oligotrophic North Pacific. The deep-sea sample is from North Equatorial Pacific Intermediate waters, one of the oldest water masses in the global ocean, and contains largely refractory DOM. The surface sample is from the oligotrophic North Pacific Gyre and does therefore contain less freshly produced biomolecules compared to the coastal sample.

To assess the spectral quality, we used the precursor purity checker module available in MZmine 3, which provides the precursor purity of each MS/MS feature, giving a score between 0 and 1, with 1 being a "clean" spectrum, which means that no other precursor was present within the isolation

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

window in the MS1 survey scan closest to the MS/MS scan. The spectral purity scores for all spectra between the 1D and 2D experiments are visualized in Figure 5. The median spectral purity in the coastal sample increased over threefold in positive mode, from 0.16 for the 1D analysis to 0.50 for the 2D analysis, whereas the improvement in the negative mode was not as pronounced, increasing only from 0.31 to 0.53 (Figure 5 A). Looking at the surface and deep sea samples from the central North Pacific Ocean (Figure 5B,C), the improvement is not as pronounced but still clearly visible, especially concerning the emergence of a substantial amount of features having a score above 0.50 in the 2D analyses compared to the 1D analyses. The improvement is more pronounced in the positive mode—the median of the surface sample increased from 0.14 to 0.19 and for the deep-sea sample from 0.13 to 0.18, while the median in the negative mode stayed the same for the surface sample and slightly increased from 0.25 to 0.27 in the deep-sea sample, although the average precursor purity increased for both samples in the negative mode (surface 0.325 to 0.342, deep-sea 0.303 to 0.332). A possible explanation for the stronger improvement of the precursor purity in the coastal sample is that, due to influences from the mainland and the presence of biomolecules from recent production, the chemical composition is more diverse and contains many labile compounds, whereas the open-ocean samples will have more similar compounds. Labile compounds degrade and stable compounds accumulate, leading to a higher proportion of structurally closely related compounds. Overall, in the distribution, a clear shift is visible in both modes: In the 1D analysis, the distribution shows a high abundance of spectra having low purity, which shifts toward higher purity in the 2D analysis. This shows a clear improvement of the spectral purity by the improved chromatographic resolution obtained through the 2D LC approach. Interestingly, we observed a general higher spectral purity of MS/MS in negative mode, which are mainly attributed to a set of features that had high precursor purity (∼1.0) in both the 1D and 2D analysis. This could also be observed in the open-ocean samples. Yet, the general trend of improved precursor purity in the 2D run also holds true for the remaining features (<0.99) in negative mode and the open ocean samples.

To investigate if the improved precursor purity relates to improved chromatographic separation and peak shape, we plotted the precursor purity against the full width at half-maximum (FWHM) of the chromatographic peak for the coastal sample, which showed a higher density of cleaner spectra (precursor purity >0.8) in the 2D runs (Figure S9). In both the 1D and 2D runs, a general relation between high-purity spectra and sharper peaks (FWHM <0.1 min) could be observed. Yet, we noted a shift in this pattern in the 2D runs, where an increasing amount of high-purity MS/MS spectra come from chromatographic features with wider peaks (FWHM >0.1 min).

To further assess the influence of spectral quality on annotation and chemical coverage in the coastal DOM sample, we plotted spectral identification and networking efficiency of the 20 cumulative 1D LC-MS/MS and the 2D LC-MS/MS runs (Figure 6 and Figure S2). To also visualize the effect of the iterative 1D analysis, we plotted the total IDs in an increasing order (1, 2, 3, 4 runs, etc.) and the 2D analysis from the classic molecular networking results (Figure 6A,B). As expected, we see that with an increasing amount of 1D runs, an increase in the number of total Library IDs from 48 to 291 in positive mode and from 5 to 71 in negative mode can be observed. Strikingly, for the 2D approach (a total of 20 runs in the second dimension with the same total run time as 20 1D runs), a total of 501 Library IDs were obtained in positive mode and 186 in negative mode, increasing the Library IDs by 72% or even 162%, respectively. Taken together, the overall increase in Library IDs is 90% (from 362 to 687) in the combined positive and negative mode data. The annotation rate of all features in negative mode is comparable between the 20 1D runs and the 2D approach, whereas the annotation rate in positive mode is at 2.2% with 20 1D runs and increases to 3.5% for the 2D approach.

Interestingly, the 2D runs did not yield a higher number of total spectra, and for positive mode, the overall spectra decreased by 37.4% (from 124010 to 77712) and by 4.8% (from 97236 to 92591) in negative mode. Conversely, the total number of nodes (clustered spectra) increased slightly from 13233 to 14339 in positive mode and significantly from 10662 to 27381 in negative mode. We assume that the larger number of total spectra in the 1D runs could be explained by the general higher signal throughout the LC gradient, so the number of MS/MS spectra acquired in every DDA duty cycle will be maxed out. For the 2D runs on the other hand, as the sample is decomplexed, in some areas of the LC gradient, the topN (5 iterative MS2 events after MS1 survey scan) of the DDA duty cycle will not be maxed out, and thus, less total MS/MS spectra will be acquired. While a decrease in the number of total spectra could indicate less total coverage, we argue that it could also indicate an improvement regarding MS/MS subsampling and less redundant spectra. Investigating the average amount of clustered spectra per network node (i.e., redundant MS/MS spectra), we see a clear decrease from an average of 9.4 spectra per node in the 1D run vs an average of 5.4 spectra per node in the 2D runs in positive mode as well as in negative mode with 9.1 vs 3.4 spectra per node.

To identify coeluting adduct features with defined mass differences (e.g., sodium, potassium, ACN, chloride, or formate adducts), we used ion-identity networking.[57] In the positive mode, ion-identities could be determined for 75.5% of the observed features, with 26.8% protonated species, 20.3% sodium adducts, 8.9% ammonium adducts, 7.5% potassium adducts, 6% formate adducts, and 6% acetonitrile adducts. In the negative mode, 53.6% of features could be assigned an ion-identity, including 18.1% deprotonated species, 15.7% formate adducts, 10.6% sodium adducts, and 9.2% chloride adducts.

Assessing the number of networks that include at least one identified compound showed that, although the consecutive runs led to an increase (from 57 to 105 in positive mode and from 13 to 23 in negative mode), the 2D analysis was able to identify at least one compound in 128 networks for positive mode and in 42 networks for negative mode (Figure S2). This shows a broader distribution of identified compounds throughout the molecular networks, which, as connected nodes share high structural similarity and have a known mass difference, facilitates structure elucidation and compound identification of unknowns.

Comparing the unique library IDs (based on unique library entries with smiles codes) between the 1D and 2D LC-MS/MS-based molecular networks of all samples (Figure S3), an average of ∼13.5% were only annotated in the 1D analysis, whereas ∼27.1% were identified in both analyses and ∼9.4% only in the 2D analysis.

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

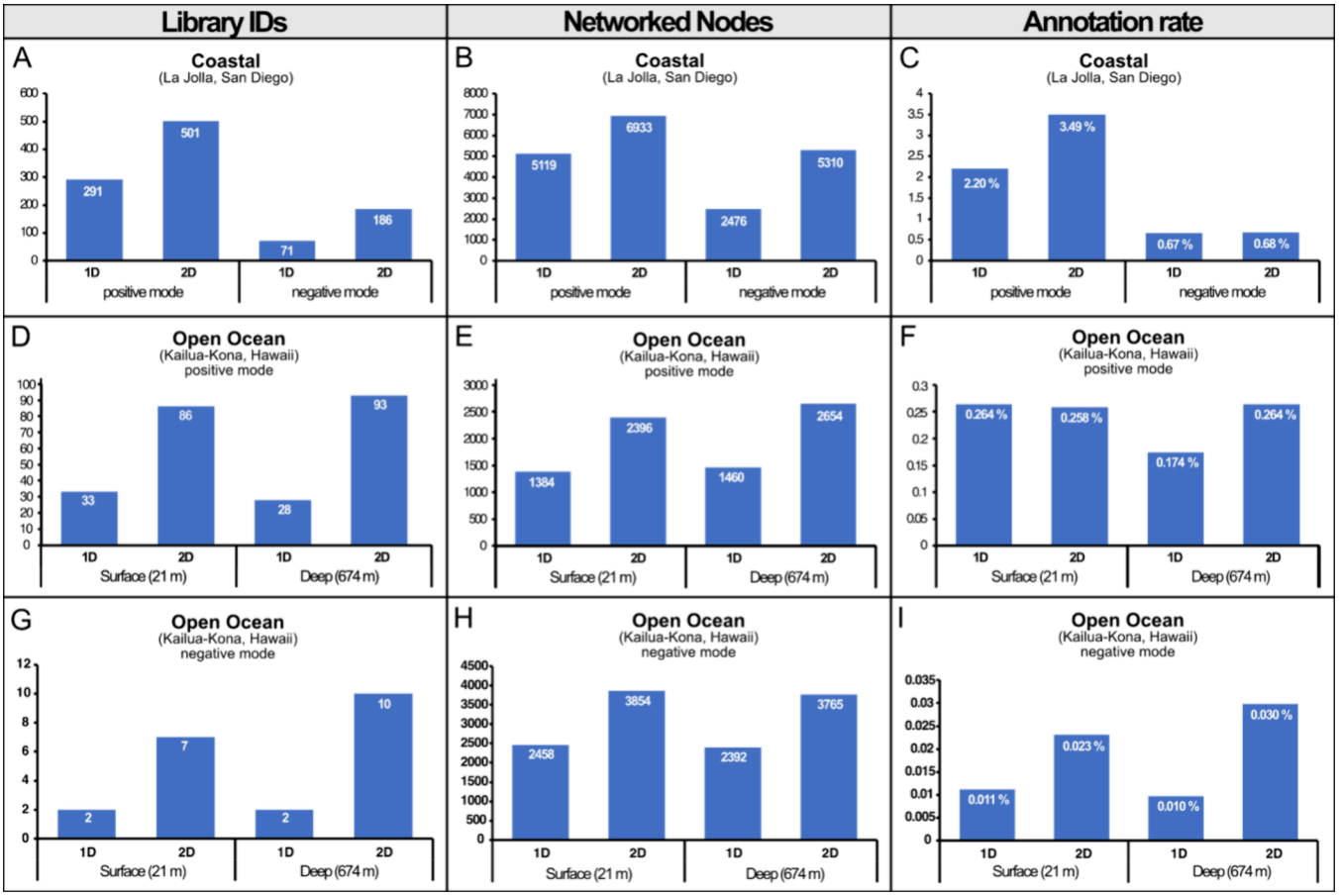


Figure 7. Increase in identification and networking efficiency in classical molecular networking. Comparison of obtained library annotations (A, D, G), networked notes (B, E, H), and annotation rate (C, F, I) between 1D and 2D analysis. (A−C) Positive and negative mode analysis of the coastal sample, (D−F) positive mode of the surface and deep-sea samples, and (G, H) negative mode of the surface and deep-sea samples.

Investigating the influence of the 2D separation on the open-ocean and deep-sea samples from the open ocean site, we observed similar improvements in the classic MN results with regard to total number of library IDs, networked nodes, and annotation rates (Figure 7). The absolute numbers are lower, which can be expected, due to incomplete libraries and less reference spectra, but looking at the unique library identifications in the open ocean samples (Figure 7D), an approximately fourfold increase in both samples can be observed. Although the annotation rate is still far lower than in the coastal sample (Figure 7F).

To further assess the chemical space of the compounds detected, we used FBMN and SIRIUS. Briefly, FBMN differs from classic MN in that it uses processed spectral information and therefore includes MS1 information like retention time or ion mobility, thereby facilitating the distinction between isomers with similar mass and MS2 spectra and spectral annotation in general, as well as incorporating relative quantitative information, which enables more robust downstream metabolomics statistical analysis. In all samples, 2D analysis yielded more nodes and more library IDs. In the coastal sample, 10,843 features were observed, of which 151 were annotated in the positive mode 1D analysis, whereas the 2D analysis yielded 12,110 features, of which 440 were annotated, improving the annotation rate from 1.4 to 3.6%. The improvement observed in the open-ocean samples was more pronounced. The positive mode 1D analysis showed 16,296 features for the surface sample and 16,290 features for the deep-sea sample, of which 43 (surface) and 27 (deep) could be annotated. The 2D analysis increased the number of features dramatically, to 27,070 for the surface and 26,500 for the deep-sea sample. The obtained library IDs increased over threefold (to 158) for the surface sample and over sixfold (to 178) for the deep-sea sample. In the negative mode, the trends are comparable. In the 2D analysis, the coastal sample contained 15,628 features with 93 annotations (11,497 features and 51 LibIDs in 1D), the surface sample showed 20,163 features with 13 annotations (14,714 features and 5 LibIDs in 1D), and the deep-sea sample contained 21,161 features with 17 annotations (15,813 features and 1 LibID in 1D).

We exported the output of the FBMN to SIRIUS, an *in silico* annotation tool, predicting the molecular formula by calculating possible fragmentation trees for MS2 spectra and used the obtained molecular formulas. For every 2D analysis, more molecular formulas could be obtained compared to the respective 1D analysis (all FBMN and SIRIUS stats are provided in the Supporting Information). Furthermore, we used CANOPUS, a tool integrated in SIRIUS, to predict the chemical classes of the observed features. The CANOPUS output was used to create sunburst plots resembling the chemical composition of the samples (Figure 7F for the coastal sample and Figures S10 and S11 for all samples). Investigating the score distribution of the SIRIUS result, we observed significant improvement of SIRIUS scores in positive mode, yet SIRIUS scores in negative mode and CSI:FingerID confidence scores between 1D and 2D runs (Figure S5) did

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

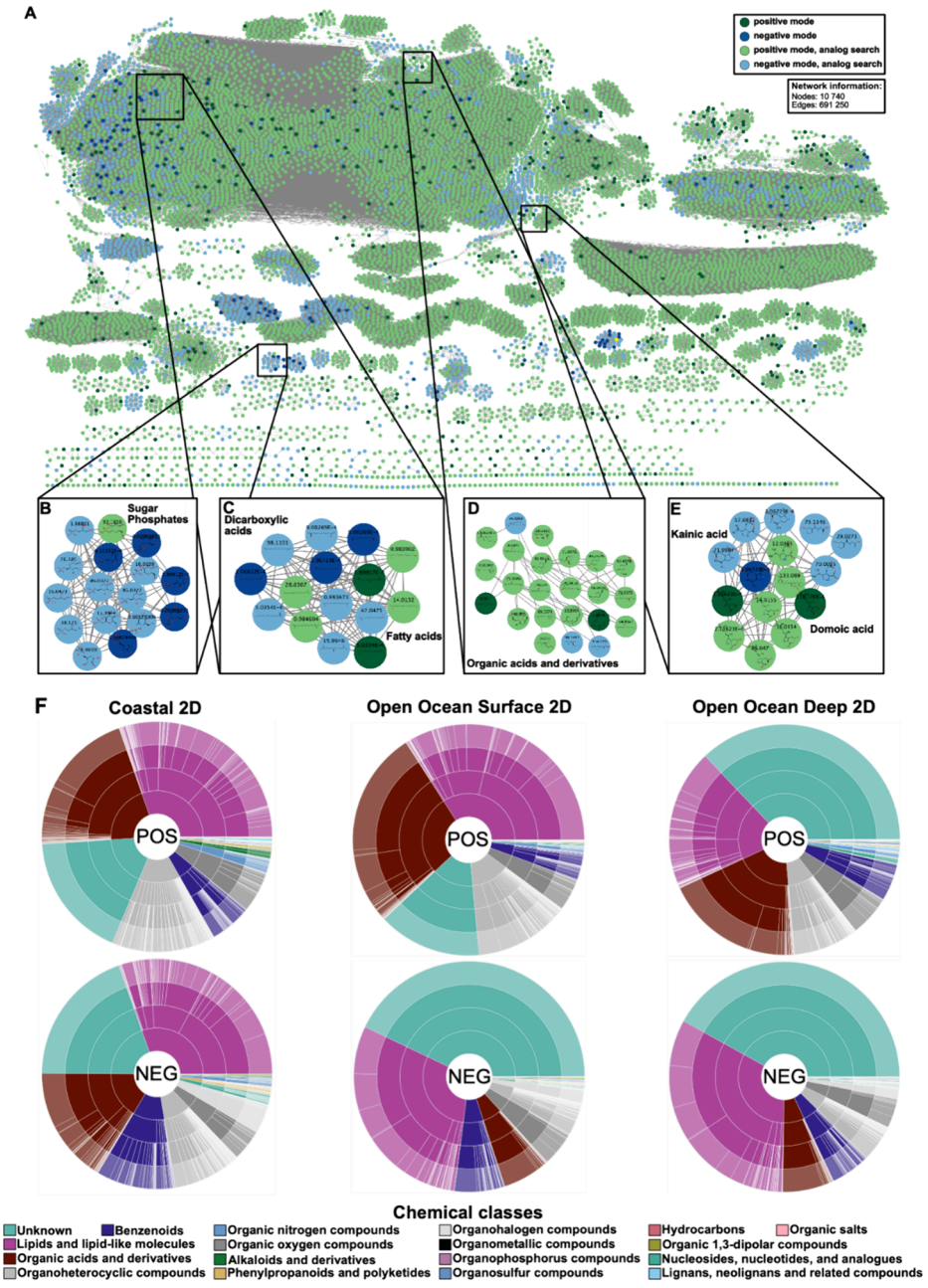


Figure 8. Structure−similarity network of annotated features (level 2 and level 3 IDs) from marine DOM. (A) Chemical similarity network between the annotated features by using the Tanimoto similarity metric. Each node represents an annotated compound. The colors of the nodes indicate the mode of analysis (green, positive mode; blue, negative mode). After removal of redundant nodes, the network was expanded including the analogue search provided by GNPS2 advanced library search, which is indicated by the respective lighter nodes. The connections between the nodes were calculated based on the chemical similarity between the compound's structural fingerprints using a threshold of 0.75. The network contains 500/127 unique library matches and 8669/1558 analogue library matches in the positive/negative ion mode. (B−E) Zoomed subnetworks; for (C) and (D) only, a subset of the subnetwork is shown. (B) Zoom-in on a subnetwork containing sugar phosphates. (C) Zoom-in on a subset of nodes containing dicarboxylic acids and fatty acids. (D) Subset of nodes containing organic acids and derivatives. (E) Zoom-in on a subnetwork containing domoic and kainic acid derivatives. (F) Sunburst plots showing the chemical classes of all features observed in the 2D analysis. The plot area represents the number of features.

not show significant improvement. This indicates that MS/MS spectral purity might have less of an effect on molecular formula calculation with SIRIUS, which relies heavily on MS1 mass accuracy and isotope patterns.

Untangling the Metabolome of Marine Dissolved Organic Matter. To obtain a status quo of annotated chemical space of DOM, we combined the structural level 2 identification (spectral matches to MS/MS libraries[58]) obtained from the FBMN using 2D analysis in positive and negative modes and removed redundant annotations. Comparing coastal DOM to open-ocean refractory DOM, we observed a strong difference in the number of annotated compounds (501 IDs in the coastal positive mode 2D data vs 80 and 85 in the open ocean samples), which we attribute to the higher refractory nature and lower concentrations, which also yielded lower network connectivity. Yet, the improvement between 1D

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

and 2D analysis was also significant in the open-ocean data and provides a so far unprecedented level of structural annotation of DOM.

Focusing on all samples, after the removal of redundant nodes, often originating from library entries having isomeric or canonical smiles codes, 500 and 127 uniquely annotated compounds in positive mode and negative mode were level 2 IDs, respectively.

An overview of all level 2 annotated compounds in the 2D analyses is provided in the tables in the Supporting Information. Expanding the library search to putative analogues (spectra with high similarity but different precursor masses), we obtained 8669 analogue hits (level 3 ID, meaning spectral similarity to known compounds of a chemical class[58]) for the compounds detected in positive mode, and 1558 in negative mode. In order to display the annotated chemical space, we generated a structure-based network using the Tanimoto similarity metric, shown in Figure 8. The network, containing 10,740 nodes and 691 250 edges, yielded one big subnetwork, containing approximately 5000 nodes, four larger networks containing between 900 and 500 nodes, some subnetworks with 100 to 200 nodes, and many (>40) subnetworks between 10 and 100 nodes, as well as individual unconnected nodes (215). As expected, the nodes in these clusters share a high structural similarity and/or are analogue matches with similar frequently occurring delta mass patterns.

The largest cluster contained 557 compounds, mainly highly oxidized molecules, such as di- and tricarboxylic acids and fatty acids. The second largest cluster, containing almost 864 nodes, also harbors highly oxidized compounds, most of which contain carboxylic acids as well as aromatic and alicyclic rings, which are structural properties representatives for carboxyl-rich alicyclic molecules (CRAM), a common class of refractory DOM components (Figure 8D).

Overall, the annotated compounds include 3519 (32.8%) nitrogen-containing compounds. Hundreds of metabolites were detected within those classes including disaccharides, fatty acids, peptides, and bile acids. In addition to the spiked standards, we could annotate a multitude of lipids and amino acids as well as terpenes such as loliolide and the signaling compound dihydrojasmonic acid. In addition to the vast amount of natural products, dozens of xenobiotics including herbicides such as simetryn, and 2-hydroxyatrazine, the hydrolysis product of atrazine, drugs such as amitriptyline, drug metabolites such as benzoylecgonine and hydroxybupropion, and the industrially used triallyl cyanurate were annotated as level 2 IDs (mirror matches inspected for all annotated compounds through the GNPS links provided in Tables S3 and S4). To unambiguously identify those compounds, we used authentic standards for triallyl cyanurate, isoxaben, *N*-lauroylsarcosine, simetryn, amitriptyline, and hydroxybupropion and analyzed them with the same LC-MS/MS method used for the 1D runs and the second dimension in the 2D runs. We prepared standard curves for the compounds and used them to estimate the concentrations of those compounds. Taking into account the extraction efficiency of PPL cartridges using seawater, assessed in an earlier study,[45] the approximate concentration in seawater for those compounds is isoxaben >48.7 ng/L; *N*-lauroylsarcosine 4.9−48.7 ng/L; triallyl cyanurate, simetryn, and hydroxybupropion 0.49−4.9 ng/L; and amitriptyline 0.049−0.49 ng/L. Notably, most of the mentioned xenobiotics were only identified in the 2D analysis and would have remained unannotated using our conventional 1D LC-MS/MS approach.

Going beyond library annotations, the improved spectral purity also increases the confidence of *in silico* spectrum annotation. To visualize the observable chemotype as a whole, we generated sunburst plots of the CANOPUS chemical class annotations using the ClassyFire chemical ontology (Figure 8F and Figure S10). While around 20% remained unannotated in both positive and negative mode data, our data provide detailed insights into the chemical makeup of coastal DOM. The largest observed components in the coastal sample were lipids and lipid-like compounds, followed by organic acids and derivatives, benzenoids, organoheterocycles, and others (Figure 8F). Class-level insights are particularly useful when no library matches are observed for a feature and provide an overall compositional information on the chemotype. Such information is particularly useful to study global elemental flux in temporal or spatial context.

Together, the compounds annotated in this study (as level 2 IDs) as well as large amounts of derivatives and class-level annotations (level 3 IDs) provide an unprecedented view into the chemical composition, vast diversity, and human influences on the coastal DOM pool, which can be used as a baseline for future nontargeted metabolomics studies of marine DOM.

Practical Considerations to Implement 2D-LC-MS/MS in Routine Analysis Workflows. A key advantage of the suggested protocol is its ease of integration with high-throughput 1D setups to both conserve fast data acquisition and improved annotation rates[28] and does not require changes in terms of hardware and gradient time to improve peak capacities.[42] This outweighs the comparatively small additional effort of an added fractionation step of a representative or pooled sample(s). Through the different pH values in the mobile phases, different functional groups are influenced differently and change their interaction with the stationary phase. This leads to an orthogonal separation, increasing the chromatographic resolution, especially for closely related compounds. While offline 2D-LC-MS/MS approaches are inherently more time- and work-intensive compared to a routine LC-MS/MS approach, they clearly increase the chemical coverage and analytical depth. A possible strategy to implement 2D-LC-MS/MS analysis into large-scale DOM studies at a feasible total runtime could be a hybrid 1D−2D approach. In such a workflow, the 2D analysis would be performed on representative or pooled sample sets to obtain high-quality MS/MS spectra and increased depth, and the quantitative comparison of these features between all samples would be performed on the MS1 level of the 1D runs. An important point here is that the chromatographic condition of the 1D run and the second dimension of the 2D runs must be identical to align features by retention time and MS1.

2D LC-MS/MS Improves Structural Insights into Dissolved Organic Matter. To elucidate molecular signatures that drive ecosystem processes, such as global carbon cycling, including human impacts, it is essential to be able to detect and identify very low-abundance molecules to capture a more complete picture. With both iterative injections via DDA-based LC-MS/MS as well as 2D-LC-MS/MS, we could increase the spectral coverage and annotation of marine DOM, through which we obtained unprecedented structural insights into the chemical space of exemplary coastal and open ocean samples. We demonstrate that 2D-LC-MS/MS analysis, using reversed-phase C18 UHPLC and orthogonal mobile

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

phase modifier and pH, effectively improves MS/MS spectral purity and spectrum library matching results from DOM in comparison to standard nontargeted LC-MS/MS. Specifically, the obtained Library IDs increased by 90% and the median MS/MS spectral purity increased by 188 and 53% in positive and negative modes, respectively. The improved data quality and depth of the 2D analysis are clearly evidenced by the higher MS/MS purity, increased number of library annotations, and larger network size in comparison to the 1D analysis.

While obtaining detailed molecular inventories is only the beginning of mechanistic interrogations, we believe this information is paramount to understanding the role of DOM in ecosystem function.

There is little doubt that with the constant increase of public DOM LC-MS/MS data sets and reference samples, and the inclusion of new *in silico* annotation tools, molecular insight into DOM will further improve. We envision that the here presented 2D LC-MS/MS method will be broadly applicable in the field and will contribute to a deeper annotation of DOM. Together with FAIR sharing of raw and processed data, we hope that these analytical advances will contribute to a better understanding of biogeochemical dynamics and human influences on DOM in the ocean as well as other ecosystems.

■ ASSOCIATED CONTENT

Data Availability Statement

Raw data was converted to .mzML format with MSConvert using centroid mode. All raw and converted data can be found on the Mass spectrometry Interactive Virtual Environment (http://massive.ucsd.edu/) with the accession number MSV000092520 (coastal sample) and MSV000092729 (open- and deep-ocean sample). Molecular network analysis was performed through the GNPS and GNPS2 data analysis environment (gnps.ucsd.edu and gnps2.org) with the parameters found in the SI (Tables S2, S5, S6, and S7). URLs to the jobs submitted to GNPS with the obtained results can be found in the SI (Tables S3 and S4).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c07173.

Comparison of number of networks with library annotations between 1D and 2D (XLSX)

Comparison of number of networks with library annotations between 1D and 2D (XLSX)

Comparison of number of networks with library annotations between 1D and 2D (XLSX)

Comparison of extracted ion chromatograms of open ocean samples; comparison of number of networks with library annotations between 1D and 2D; Venn diagrams of unique SMILES codes from library IDs; Van Krevelen plot of the coastal sample; raincloud plots of SIRIUS results for the coastal sample; identification and quantification of level 2 annotated compounds with authentic standards, including XICs, mirror matches, and quantification curves; distribution of MS/MS precursor purity and FWHM for the coastal sample; sunburst plots for all samples and every analysis, with a detailed view on sunburst plots shown in Figure 8; detailed listing of parameters used for classical and feature-based molecular networking, as well as MZmine processing; and tables with the URLs to all molecular networking jobs and network statistics of FBMNs (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Daniel Petras − *Interfaculty Institute of Microbiology and Medicine, University of Tuebingen, Tuebingen 72076, Germany; Department of Biochemistry, University of California Riverside, Riverside, California 92521, United States;* orcid.org/0000-0002-6561-3022; Email: dpetras@ucr.edu

Authors

Stilianos Papadopoulos Lambidis − *Interfaculty Institute of Microbiology and Medicine, University of Tuebingen, Tuebingen 72076, Germany*

Tilman Schramm − *Interfaculty Institute of Microbiology and Medicine, University of Tuebingen, Tuebingen 72076, Germany; Department of Biochemistry, University of California Riverside, Riverside, California 92521, United States*

Karoline Steuer-Lodd − *Interfaculty Institute of Microbiology and Medicine, University of Tuebingen, Tuebingen 72076, Germany; Department of Biochemistry, University of California Riverside, Riverside, California 92521, United States*

Shane Farrell − *Interfaculty Institute of Microbiology and Medicine, University of Tuebingen, Tuebingen 72076, Germany; Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine 94544, United States*

Paolo Stincone − *Interfaculty Institute of Microbiology and Medicine, University of Tuebingen, Tuebingen 72076, Germany;* orcid.org/0000-0002-2214-6655

Robin Schmid − *Institute of Organic Chemistry and Biochemistry of the CAS, Prague 166 10, Czech Republic*

Irina Koester − *Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92037, United States*

Ralph Torres − *Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92037, United States*

Thorsten Dittmar − *Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Oldenburg 26382, Germany; Helmholtz Institute for Functional Marine Biodiversity (HIFMB), Carl von Ossietzky University Oldenburg, Oldenburg 26129, Germany;* orcid.org/0000-0002-3462-0107

Lihini Aluwihare − *Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92037, United States*

Carsten Simon − *Department of Environmental Analytical Chemistry, Helmholtz-Centre for Environmental Research, Leipzig 04318, Germany;* orcid.org/0000-0003-3882-3013

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.4c07173

Author Contributions

D.P. conceptualized and designed the study. C.S., I.K., and D.P. performed initial proof-of-concept experiments. R.T., T.D., L.A., and D.P. provided materials, collected the sample, and performed the extraction. S.P.L., P.S., and D.P. performed

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

the MS data acquisition. S.P.L., T.S., K.S.-L., S.F., and D.P. performed data analysis and generated figures. R.S. contributed software. S.P.L., C.S., and D.P. wrote the manuscript. All authors read, edited, and approved the final manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

D.P. was supported by Deutsche Forschungsgemeinschaft through the CMFI Cluster of Excellence (EXC 2124) and the Collaborative Research Center CellMap (TRR 261). P.S. was supported by the European Union Horizon 2020 Program through a Marie Skłodowska-Curie fellowship (101108450-MeStaLeM).

■ REFERENCES

(1) Moran, M. A.; Kujawinski, E. B.; Stubbins, A.; Fatland, R.; Aluwihare, L. I.; Buchan, A.; Crump, B. C.; Dorrestein, P. C.; Dyhrman, S. T.; Hess, N. J.; Howe, B.; Longnecker, K.; Medeiros, P. M.; Niggemann, J.; Obernosterer, I.; Repeta, D. J.; Waldbauer, J. R. Deciphering Ocean Carbon in a Changing World. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113* (12), 3143−3151.

(2) Vorobev, A.; Sharma, S.; Yu, M.; Lee, J.; Washington, B. J.; Whitman, W. B.; Ballantyne, F.; Medeiros, P. M.; Moran, M. A. Identifying Labile DOM Components in a Coastal Ocean through Depleted Bacterial Transcripts and Chemical Signals. *Environmental Microbiology* **2018**, *20* (8), 3012−3030.

(3) Zhang, C.; Dang, H.; Azam, F.; Benner, R.; Legendre, L.; Passow, U.; Polimene, L.; Robinson, C.; Suttle, C. A.; Jiao, N. Evolving Paradigms in Biological Carbon Cycling in the Ocean. *National Science Review* **2018**, *5* (4), 481−499.

(4) Kujawinski, E. B.; Longnecker, K.; Barott, K. L.; Weber, R. J. M.; Soule, M. C. K. Microbial Community Structure Affects Marine Dissolved Organic Matter Composition. *Front. Mar. Sci.* **2016**, *3*, 45.

(5) Wang, W.-L.; Moore, J. K.; Martiny, A. C.; Primeau, F. W. Convergent Estimates of Marine Nitrogen Fixation. *Nature* **2019**, *566* (7743), 205−211.

(6) Decho, A. W.; Gutierrez, T. Microbial Extracellular Polymeric Substances (EPSs) in Ocean Systems. *Front. Microbiol.* **2017**, *8*, 922.

(7) Beaupré, S. R.; Kieber, D. J.; Keene, W. C.; Long, M. S.; Maben, J. R.; Lu, X.; Zhu, Y.; Frossard, A. A.; Kinsey, J. D.; Duplessis, P.; Chang, R. Y.-W.; Bisgrove, J. Oceanic Efflux of Ancient Marine Dissolved Organic Carbon in Primary Marine Aerosol. *Sci. Adv.* **2019**, *5* (10), No. eaax6535.

(8) Malfatti, F.; Lee, C.; Tinta, T.; Pendergraft, M. A.; Celussi, M.; Zhou, Y.; Sultana, C. M.; Rotter, A.; Axson, J. L.; Collins, D. B.; Santander, M. V.; Anides Morales, A. L.; Aluwihare, L. I.; Riemer, N.; Grassian, V. H.; Azam, F.; Prather, K. A. Detection of Active Microbial Enzymes in Nascent Sea Spray Aerosol: Implications for Atmospheric Chemistry and Climate. *Environ. Sci. Technol. Lett.* **2019**, *6* (3), 171−177.

(9) Ossola, R.; Tolu, J.; Clerc, B.; Erickson, P. R.; Winkel, L. H. E.; McNeill, K. Photochemical Production of Sulfate and Methanesulfonic Acid from Dissolved Organic Sulfur. *Environ. Sci. Technol.* **2019**, *53* (22), 13191−13200.

(10) Kujawinski, E. B. The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annu. Rev. Mar. Sci.* **2011**, *3* (1), 567−599.

(11) Zark, M.; Dittmar, T. Universal Molecular Structures in Natural Dissolved Organic Matter. *Nat. Commun.* **2018**, *9* (1), 3178.

(12) Leyva, D.; Tariq, M. U.; Jaffé, R.; Saeed, F.; Lima, F. F. Unsupervised Structural Classification of Dissolved Organic Matter Based on Fragmentation Pathways. *Environ. Sci. Technol.* **2022**, *56* (2), 1458−1468.

(13) Simon, C.; Dührkop, K.; Petras, D.; Roth, V.-N.; Böcker, S.; Dorrestein, P. C.; Gleixner, G. Mass Difference Matching Unfolds Hidden Molecular Structures of Dissolved Organic Matter. *Environ. Sci. Technol.* **2022**, *56* (15), 11027−11040.

(14) Zark, M.; Christoffers, J.; Dittmar, T. Molecular Properties of Deep-Sea Dissolved Organic Matter Are Predictable by the Central Limit Theorem: Evidence from Tandem FT-ICR-MS. *Marine Chemistry* **2017**, *191*, 9−15.

(15) Petras, D.; Koester, I.; Da Silva, R.; Stephens, B. M.; Haas, A. F.; Nelson, C. E.; Kelly, L. W.; Aluwihare, L. I.; Dorrestein, P. C. High-Resolution Liquid Chromatography Tandem Mass Spectrometry Enables Large Scale Molecular Characterization of Dissolved Organic Matter. *Front. Mar. Sci.* **2017**, *4*, 405.

(16) Patriarca, C.; Bergquist, J.; Sjöberg, P. J. R.; Tranvik, L.; Hawkes, J. A. Online HPLC-ESI-HRMS Method for the Analysis and Comparison of Different Dissolved Organic Matter Samples. *Environ. Sci. Technol.* **2018**, *52* (4), 2091−2099.

(17) Han, L.; Kaesler, J.; Peng, C.; Reemtsma, T.; Lechtenfeld, O. J. Online Counter Gradient LC-FT-ICR-MS Enables Detection of Highly Polar Natural Organic Matter Fractions. *Anal. Chem.* **2021**, *93* (3), 1740−1748.

(18) Petras, D.; Minich, J. J.; Cancelada, L. B.; Torres, R. R.; Kunselman, E.; Wang, M.; White, M. E.; Allen, E. E.; Prather, K. A.; Aluwihare, L. I.; Dorrestein, P. C. Non-Targeted Tandem Mass Spectrometry Enables the Visualization of Organic Matter Chemotype Shifts in Coastal Seawater. *Chemosphere* **2021**, *271*, No. 129450.

(19) Lu, K.; Gardner, W. S.; Liu, Z. 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.* **2018**, *52* (13), 7182−7191.

(20) Leyva, D.; Tose, L. V.; Porter, J.; Wolff, J.; Jaffé, R.; Fernandez-Lima, F. Understanding the Structural Complexity of Dissolved Organic Matter: Isomeric Diversity. *Faraday Discuss.* **2019**, *218*, 431−440.

(21) Koester, I.; Quinlan, Z. A.; Nothias, L.; White, M. E.; Rabines, A.; Petras, D.; Brunson, J. K.; Dührkop, K.; Ludwig, M.; Böcker, S.; Azam, F.; Allen, A. E.; Dorrestein, P. C.; Aluwihare, L. I. Illuminating the Dark Metabolome of *Pseudo-nitzschia* − Microbiome Associations. *Environmental Microbiology* **2022**, *24* (11), 5408−5424.

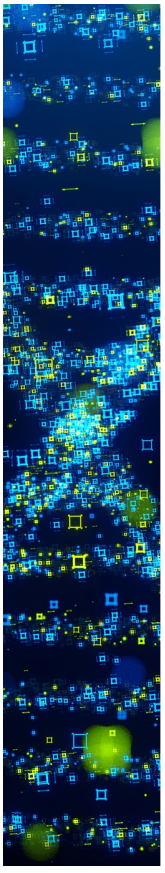
(22) Petras, D.; Caraballo-Rodríguez, A. M.; Jarmusch, A. K.; Molina-Santiago, C.; Gauglitz, J. M.; Gentry, E. C.; Belda-Ferre, P.; Romero, D.; Tsunoda, S. M.; Dorrestein, P. C.; Wang, M. Chemical Proportionality within Molecular Networks. *Anal. Chem.* **2021**, *93* (38), 12833−12839.

(23) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community Curation of Mass Spectrometry Data with Global

Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828−837.
(24) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin H, C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; Van Der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Feature-Based Molecular Networking in the GNPS Analysis Environment. *Nat. Methods* **2020**, *17* (9), 905−908.
(25) Wang, M.; Jarmusch, A. K.; Vargas, F.; Aksenov, A. A.; Gauglitz, J. M.; Weldon, K.; Petras, D.; Da Silva, R.; Quinn, R.; Melnik, A. V.; Van Der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Nothias, L. F.; Aceves, C. M.; Panitchpakdi, M.; Brown, E.; Di Ottavio, F.; Sikora, N.; Elijah, E. O.; Labarta-Bajo, L.; Gentry, E. C.; Shalapour, S.; Kyle, K. E.; Puckett, S. P.; Watrous, J. D.; Carpenter, C. S.; Bouslimani, A.; Ernst, M.; Swafford, A. D.; Zúñiga, E. I.; Balunas, M. J.; Klassen, J. L.; Loomba, R.; Knight, R.; Bandeira, N.; Dorrestein, P. C. Mass Spectrometry Searches Using MASST. *Nat. Biotechnol.* **2020**, *38* (1), 23−26.
(26) Quinlan, Z. A.; Koester, I.; Aron, A. T.; Petras, D.; Aluwihare, L. I.; Dorrestein, P. C.; Nelson, C. E.; Wegley Kelly, L. ConCISE: Consensus Annotation Propagation of Ion Features in Untargeted Tandem Mass Spectrometry Combining Molecular Networking and In Silico Metabolite Structure Prediction. *Metabolites* **2022**, *12* (12), 1275.
(27) Stephens, B.; Stincone, P.; Petras, D.; English, C.; Opalk, K.; Giovannoni, S.; Carlson, C. Oxidation State of Bioavailable Dissolved Organic Matter Influences Bacterioplankton Respiration and Growth Efficiency. 2023. *Research Square*.
(28) Da Silva, R. R.; Wang, M.; Nothias, L.-F.; Van Der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. Propagating Annotations of Molecular Networks Using in Silico Fragmentation. *PLoS Comput. Biol.* **2018**, *14* (4), No. e1006089.
(29) Sandron, S.; Rojas, A.; Wilson, R.; Davies, N. W.; Haddad, P. R.; Shellie, R. A.; Nesterenko, P. N.; Kelleher, B. P.; Paull, B. Chromatographic Methods for the Isolation, Separation and Characterisation of Dissolved Organic Matter. *Environ. Sci.: Processes Impacts* **2015**, *17* (9), 1531−1567.
(30) Brown, T. A.; Jackson, B. A.; Bythell, B. J.; Stenson, A. C. Benefits of Multidimensional Fractionation for the Study and Characterization of Natural Organic Matter. *Journal of Chromatography A* **2016**, *1470*, 84−96.
(31) Zhang, F.; Harir, M.; Moritz, F.; Zhang, J.; Witting, M.; Wu, Y.; Schmitt-Kopplin, P.; Fekete, A.; Gaspar, A.; Hertkorn, N. Molecular and Structural Characterization of Dissolved Organic Matter during and Post Cyanobacterial Bloom in Taihu by Combination of NMR Spectroscopy and FTICR Mass Spectrometry. *Water Res.* **2014**, *57*, 280−294.
(32) Stincone, P.; Shah, A. K. P.; Schmid, R.; Graves, L.; Lambidis, S. P.; Torres, R.; Xia, S.-N.; Minda, V.; Aron, A.; Wang, M.; Hughes, C. C.; Petras, D. Evaluation of Data-Dependent MS/MS Acquisition Parameters for Non-Targeted Metabolomics and Molecular Networking of Environmental Samples: Focus on the Q Exactive Platform. *ChemRxiv* 2023.
(33) Arakawa, N.; Aluwihare, L. I.; Simpson, A. J.; Soong, R.; Stephens, B. M.; Lane-Coplen, D. Carotenoids Are the Likely Precursor of a Significant Fraction of Marine Dissolved Organic Matter. *Sci. Adv.* **2017**, *3* (9), No. e1602976.
(34) Spranger, T.; Pinxteren, D. V.; Reemtsma, T.; Lechtenfeld, O. J.; Herrmann, H. 2D Liquid Chromatographic Fractionation with Ultra-High Resolution MS Analysis Resolves a Vast Molecular Diversity of Tropospheric Particle Organics. *Environ. Sci. Technol.* **2019**, *53* (19), 11353−11363.
(35) Woods, G. C.; Simpson, M. J.; Simpson, A. J. Oxidized Sterols as a Significant Component of Dissolved Organic Matter: Evidence from 2D HPLC in Combination with 2D and 3D NMR Spectroscopy. *Water Res.* **2012**, *46* (10), 3398−3408.
(36) Stoll, D. R.; Carr, P. W. Two-Dimensional Liquid Chromatography: A State of the Art Tutorial. *Anal. Chem.* **2017**, *89* (1), 519−531.
(37) Mondello, L.; Dugo, P.; Donato, P.; Herrero, M.; Montero, L.; Schmitz, O. J. Comprehensive Two-Dimensional Liquid Chromatography. *Nat. Rev. Methods Primers* **2023**, *3* (1), 86.
(38) Horváth, K.; Fairchild, J.; Guiochon, G. Optimization Strategies for Off-Line Two-Dimensional Liquid Chromatography. *Journal of Chromatography A* **2009**, *1216* (12), 2511−2518.
(39) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. Two-dimensional Separation of Peptides Using RP-RP-HPLC System with Different pH in First and Second Separation Dimensions. *J. of Separation Science* **2005**, *28* (14), 1694−1703.
(40) Wang, Z.; Ma, H.; Smith, K.; Wu, S. Two-Dimensional Separation Using High-pH and Low-pH Reversed Phase Liquid Chromatography for Top-down Proteomics. *Int. J. Mass Spectrom.* **2018**, *427*, 43−51.
(41) Baghdady, Y. Z.; Schug, K. A. Online Comprehensive High pH Reversed Phase × Low pH Reversed Phase Approach for Two-Dimensional Separations of Intact Proteins in Top-Down Proteomics. *Anal. Chem.* **2019**, *91* (17), 11085−11091.
(42) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. Orthogonality of Separation in Two-Dimensional Liquid Chromatography. *Anal. Chem.* **2005**, *77* (19), 6426−6434.
(43) Vinh, J. Proteomics and Proteoforms: Bottom-up or Top-down, How to Use High-Resolution Mass Spectrometry to Reach the Grail. *Fundam. Appl. Fourier Transform Mass Spectrom.* **2019**, 529−567.
(44) McIlvin, M. R.; Saito, M. A. Online Nanoflow Two-Dimension Comprehensive Active Modulation Reversed Phase−Reversed Phase Liquid Chromatography High-Resolution Mass Spectrometry for Metaproteomics of Environmental and Microbiome Samples. *J. Proteome Res.* **2021**, *20* (9), 4589−4597.
(45) Cancelada, L.; Torres, R. R.; Garrafa Luna, J.; Dorrestein, P. C.; Aluwihare, L. I.; Prather, K. A.; Petras, D. Assessment of Styrene-divinylbenzene Polymer (PPL) Solid-phase Extraction and Non-targeted Tandem Mass Spectrometry for the Analysis of Xenobiotics in Seawater. *Limnology & Ocean Methods* **2022**, *20* (2), 89−101.
(46) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A Simple and Efficient Method for the Solid-phase Extraction of Dissolved Organic Matter (SPE-DOM) from Seawater. *Limnology & Ocean Methods* **2008**, *6* (6), 230−235.
(47) Pendergraft, M. A.; Belda-Ferre, P.; Petras, D.; Morris, C. K.; Mitts, B. A.; Aron, A. T.; Bryant, M.; Schwartz, T.; Ackermann, G.; Humphrey, G.; Kaandorp, E.; Dorrestein, P. C.; Knight, R.; Prather, K. A. Bacterial and Chemical Evidence of Coastal Water Pollution from the Tijuana River in Sea Spray Aerosol. *Environ. Sci. Technol.* **2023**, *57* (10), 4071−4081.
(48) Green, N. W.; Perdue, E. M.; Aiken, G. R.; Butler, K. D.; Chen, H.; Dittmar, T.; Niggemann, J.; Stubbins, A. An Intercomparison of Three Methods for the Large-Scale Isolation of Oceanic Dissolved Organic Matter. *Marine Chemistry* **2014**, *161*, 14−19.
(49) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Dührkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdžija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L.-F.; Wang, M.; Orešič, M.; Weng, J.-K.; Böcker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T. Integrative Analysis of Multimodal Mass Spectrometry Data in MZmine 3. *Nat. Biotechnol.* **2023**, *41* (4), 447−449.

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304

(50) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299−302.
(51) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching Molecular Structure Databases with Tandem Mass Spectra Using CSI:FingerID. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (41), 12580−12585.
(52) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* (11), 2498−2504.
(53) Woods, G. C.; Simpson, M. J.; Koerner, P. J.; Napoli, A.; Simpson, A. J. HILIC-NMR: Toward the Identification of Individual Molecular Components in Dissolved Organic Matter. *Environ. Sci. Technol.* **2011**, *45* (9), 3880−3886.
(54) Baalousha, M.; Motelica-Heino, M.; Coustumer, P. L. Conformation and Size of Humic Substances: Effects of Major Cation Concentration and Type, pH, Salinity, and Residence Time. *Colloids Surf., A* **2006**, *272* (1−2), 48−55.
(55) Hawkes, J. A.; Sjöberg, P. J. R.; Bergquist, J.; Tranvik, L. J. Complexity of Dissolved Organic Matter in the Molecular Size Dimension: Insights from Coupled Size Exclusion Chromatography Electrospray Ionisation Mass Spectrometry. *Faraday Discuss.* **2019**, *218*, 52−71.
(56) Aron, A. T.; Gentry, E. C.; McPhail, K. L.; Nothias, L. F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F.; van der Hooft, J. J. J.; Ernst, M.; Kang, K. B.; Aceves, C. M.; Caraballo-Rodríguez, A. M.; Koester, I.; Weldon, K. C.; Bertrand, S.; Roullier, C.; Sun, K.; Tehan, R. M.; Boya P., C. A.; Christian, M. H.; Gutiérrez, M.; Ulloa, A. M.; Tejeda Mora, J. A.; Mojica-Flores, R.; Lakey-Beitia, J.; Vásquez-Chaves, V.; Zhang, Y.; Calderón, A. I.; Tayler, N.; Keyzers, R. A.; Tugizimana, F.; Ndlovu, N.; Aksenov, A. A.; Jarmusch, A. K.; Schmid, R.; Truman, A. W.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Reproducible Molecular Networking of Untargeted Mass Spectrometry Data Using GNPS. *Nat. Protoc.* **2020**, *15* (6), 1954−1991.
(57) Schmid, R.; Petras, D.; Nothias, L.-F.; Wang, M.; Aron, A. T.; Jagels, A.; Tsugawa, H.; Rainer, J.; Garcia-Aloy, M.; Dührkop, K.; Korf, A.; Pluskal, T.; Kameník, Z.; Jarmusch, A. K.; Caraballo-Rodríguez, A. M.; Weldon, K. C.; Nothias-Esposito, M.; Aksenov, A. A.; Bauermeister, A.; Albarracin Orio, A.; Grundmann, C. O.; Vargas, F.; Koester, I.; Gauglitz, J. M.; Gentry, E. C.; Hövelmann, Y.; Kalinina, S. A.; Pendergraft, M. A.; Panitchpakdi, M.; Tehan, R.; Le Gouellec, A.; Aleti, G.; Mannochio Russo, H.; Arndt, B.; Hübner, F.; Hayen, H.; Zhi, H.; Raffatellu, M.; Prather, K. A.; Aluwihare, L. I.; Böcker, S.; McPhail, K. L.; Humpf, H.-U.; Karst, U.; Dorrestein, P. C. Ion Identity Molecular Networking for Mass Spectrometry-Based Metabolomics in the GNPS Environment. *Nat. Commun.* **2021**, *12* (1), 3832.
(58) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed Minimum Reporting Standards for Chemical Analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3* (3), 211−221.


CAS BIOFINDER DISCOVERY PLATFORM™
STOP DIGGING THROUGH DATA —START MAKING DISCOVERIES
CAS BioFinder helps you find the right biological insights in seconds
Start your search
CAS
A division of the American Chemical Society

https://doi.org/10.1021/acs.est.4c07173
Environ. Sci. Technol. 2024, 58, 19289−19304