LIMNOLOGY and OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 20, 2022, 683–700

doi: 10.1002/lom3.10513

Quantification of dissolved metabolites in environmental samples through cation-exchange solid-phase extraction paired with liquid chromatography–mass spectrometry

Joshua S. Sacks,[1] **Katherine R. Heal**,[1,2] **Angela K. Boysen**,[1,3] **Laura T. Carlson**,[1] **Anitra E. Ingalls**[1]*

[1]School of Oceanography, University of Washington, Seattle, Washington
[2]Integral Consulting Inc, Seattle, Washington
[3]Department of Geophysical Sciences, University of Chicago, Chicago, Illinois

Abstract

Small, biologically produced, organic molecules called metabolites play key roles in microbial systems where they directly mediate exchanges of nutrients, energy, and information. However, the study of dissolved polar metabolites in seawater and other environmental matrices has been hampered by analytical challenges including high inorganic ion concentrations, low analyte concentrations, and high chemical diversity. Here we show that a cation-exchange solid-phase extraction (CX-SPE) sample preparation approach separates positively charged and zwitterionic metabolites from seawater and freshwater samples, allowing their analysis by liquid chromatography–mass spectrometry. We successfully extracted 69 known compounds from an in-house compound collection and evaluated the performance of the method by establishing extraction efficiencies (EEs) and limits of detection (pM to low nM range) for these compounds. CX-SPE extracted a range of compounds including amino acids and compatible solutes, resulted in very low matrix effects, and performed robustly across large variations in salinity and dissolved organic matter concentration. We compared CX-SPE to an established SPE procedure (PPL-SPE) and demonstrate that these two methods extract fundamentally different fractions of the dissolved metabolite pool with CX-SPE extracting compounds that are on average smaller and more polar. We use CX-SPE to analyze four environmental samples from distinct aquatic biomes, producing some of the first CX-SPE dissolved metabolomes. Quantified compounds ranged in concentration from 0.0093 to 49 nM and were composed primarily of amino acids (0.15–16 nM) and compatible solutes such as trimethylamine N-oxide (0.89–49 nM) and glycine betaine (2.8–5.2 nM).

Metabolomics is commonly used to measure intracellular pools of small molecules where it is a powerful tool in characterizing cellular phenotype and biochemical pathways (Bundy et al. 2008). In aquatic environments, dissolved metabolites (operationally defined as passing through a 0.2-μm filter) mediate microbial nutrient and energy exchanges and can serve as critical controls on microbial community composition and activity (Amin et al. 2015; Heal et al. 2019; Shibl et al. 2020; Moran et al. 2022). However, our ability to measure extracellular metabolomes in environmental samples is hampered by low metabolite concentrations, high inorganic ion concentrations, and the complex matrix of dissolved organic matter (DOM) (Moran et al. 2022). This problem is particularly acute in seawater since salt ($\sim$ 0.6 M) and background DOM ($\sim$ 40–80 μM) concentrations exceeds those of individual metabolites (low pM to low nM) by roughly nine and four orders of magnitude, respectively (Emerson and Hedges 2008; Moran et al. 2022). In liquid chromatography mass spectrometry (LC–MS) analyses, salts and background DOM can cause severe ion suppression, preventing the detection and quantification of compounds of interest (Johnson et al. 2017; Boysen et al. 2018). Furthermore, metabolites are extremely chemically diverse, spanning a range of formal charges (positive to negative), polarities (nonpolar to polar), functional classifications, and masses (10s to 1000s of

*Correspondence: aingalls@uw.edu

Additional Supporting Information may be found in the online version of this article.

Author Contribution Statement: A.E.I., J.S.S., and K.R.H. designed the study. J.S.S. and L.T.C. performed laboratory analyses. J.S.S. analyzed the data, with substantial guidance and input from K.R.H. and A.K.B. A.E.I. and J.S.S wrote the manuscript. A.E.I., J.S.S., K.R.H., L.T.C., and A.K.B. provided edits on the manuscript.

Daltons). This diversity makes measuring all metabolites with a singular analytical approach nearly impossible as the techniques used to separate and concentrate analytes from the matrix rely on differences in charge, polarity, size, and functional groups. Therefore, specific extraction techniques are needed to target the polar, low-molecular-weight fraction of metabolites and remove them from the salts and background DOM present in aquatic samples to enable their analysis via LC–MS.

In recognition of this need, a variety of methods targeting low molecular weight, polar compounds dissolved in seawater are now available (Johnson et al. 2017; Sogin et al. 2019; Pontrelli and Sauer 2021; Widner et al. 2021; Xu et al. 2021), but critical gaps remain in the analytical windows of these techniques (Moran et al. 2022). The commercially available modified styrene divinyl benzene polymer solid-phase extraction Bond Elute PPL column (PPL-SPE) (Agilent Technologies) is widely used in low-molecular-weight DOM analysis, retains 43–62% of DOM from marine samples, and has been used widely for dissolved metabolite studies (Dittmar et al. 2008; Fiore et al. 2015; Johnson et al. 2016, 2017; Petras et al. 2017; Weber et al. 2020). The PPL-extracted DOM, while containing some metabolites, is comprised primarily of non-polar molecules that are representative of the large, recalcitrant ocean DOM pool that can cause substantial matrix effects, limiting the ionization of polar metabolites of interest during electrospray ionization MS (Dittmar et al. 2008; Johnson et al. 2017). Benzyl chloride (BC) derivatization of primary amine, secondary amine, and alcohol containing functional groups before extraction via PPL-SPE captures a range of metabolites not extracted using just PPL-SPE (Widner et al. 2021). However, polar compounds without those functional groups, such as glycine betaine, are not derivatized with BC, cannot be extracted with PPL (Johnson et al. 2017), and often co-elute with other compounds with a reversed-phase (RP) LC–MS run (Boysen et al. 2021; Heal et al. 2021). In addition to the metabolites of interest, the BC-derivatization/PPL approach also has the potential to extract a large number of unidentified compounds present in DOM including hydrophobic DOM that binds to the PPL column without derivatization as well as unknown molecules that bind to the PPL column post derivatization. These molecules may introduce matrix effects, particularly in high DOM samples such as those from coastal environments or laboratory cultures. Derivatization paired with gas chromatography (Sogin et al. 2019) and with liquid/liquid separations (Xu et al. 2021) as well as salt tolerant RP columns (Pontrelli and Sauer 2021) have also been used, but without sufficient sensitivities to measure compounds at the low nanomolar to picomolar concentrations expected in seawater.

A critical gap in understanding dissolved organic compounds is in the ability to measure polar zwitterionic compounds containing positively charged quaternary amine or sulfonium groups including dimethylsulfoniopropionate (DMSP), gonyol, dimethylsulfonioacetate (DMS-Ac), glycine betaine, homarine, trigonelline, and trimethylamine N-oxide (TMAO). These compounds have been identified as quantitatively important components of marine particulate metabolite pools (Johnson et al. 2020; Boysen et al. 2021; Heal et al. 2021), are often accumulated as osmolytes to help microbes respond to osmotic stress (Welsh 2000; Gebser and Pohnert 2013; Dawson et al. 2020*a*,*b*), mediate microbial interactions such as phytoplankton-bacteria symbioses (Johnson et al. 2016), and play important roles in the cycling of carbon and other elements (Welsh 2000; Yoch 2002; Boysen et al. 2022). Many polar charged and zwitterionic compounds are largely unextractable with existing dissolved metabolomics extraction approaches and invisible to some LC–MS analytical methods, requiring the development of new metabolomics methods to study the distribution and roles of these compounds in marine systems alongside other key dissolved metabolite groups such as amino acids.

Cation-exchange SPE (CX-SPE) has the potential to broaden our view of the polar components in DOM by targeting compounds with a permanent positive charge or that can be protonated under acidic conditions. CX-SPE has been successfully used for the extraction of small organoarsenic compounds in seawater (Glabonjat et al. 2018), therapeutic and illicit drugs from aqueous sewage samples (Fontanals et al. 2014), and drugs and metabolites from urine samples (Logan et al. 1990). In CX-SPE, protonated analytes and cationic salts in an acidified sample bind to negatively charged binding sites on the stationary-phase resin through electrostatic attractions. The analytes are then eluted through the addition of a basic solution which deprotonates the analytes or alters their interactions with the stationary-phase while the salts remain attached to the resin, separating the analytes from the salts. The eluted analytes can then be concentrated, separated, and analyzed using LC–MS. The stationary phase is regenerated by flushing the columns with a concentrated strong acid solution to displace the salts on the column.

Here, we present a CX-SPE approach for extraction of dissolved metabolites from seawater and pair it with established LC–MS methods optimized for the separation and quantification of small, polar molecules (Boysen et al. 2018). We identify compounds that are extracted using this approach, characterize their EEs, and examine the robustness of this new method across a range of salinities and DOM concentrations. We also compare this new approach to the common PPL-SPE approach for extracting DOM from seawater. Finally, we demonstrate the utility of the method by quantifying dissolved metabolite pools in a variety of environmental samples.

Methods

Materials

Most metabolite standards were purchased from Sigma Aldrich, Santa Cruz Biotechnology, Toronto Research

Chemicals, Cambridge Chemicals, Spectrum Chemical, and Cambridge Isotope Laboratories. The sulfonate compound 2,3-dihydroxypropane-1-sulfonate (DHPS) was provided by A. Bourdon and S. Champagna (University of Tennessee). DMS-Ac and gonyol were provided by Dr. G. Pohnert (Friedrich Schiller University). *N*-acetyltaurine was provided by A. Cook and K. Denger (University of Konstanz). 2-(3,5-Dichlorophenylcarbamoyl)-1,2-dimethylcyclopropane-1-carboxylic acid (cinnamoyl-HSL) was provided by C. Harwood (University of Washington). TMAP (β-alanine betaine) was synthesized in house using the procedure detailed in Gebser and Pohnert (2013). Full supplier information is provided in Supporting Information Table S1.

Sample collection

Natural water samples for method assessment and environmental analysis were collected from four different locations, Station ALOHA in the North Pacific Subtropical Gyre (ALOHA, 22°45′N, 158°W), the North Pacific Transition Zone (NP, 41°24′N, 158° W), Puget Sound (PS, 47°41′N, 122°25′W), and Lake Washington (LW, 47°38′N, 122°18′W). ALOHA and NP samples were collected at a depth of 15 m on the research cruise KOK1606 aboard the R.V. Ka'imikai-O-Kanoloa between 20 April 2016 and 02 May 2016 from Niskin bottles. The PS sample was collected on 07 May 2017 at a depth of 8 m from aboard the R.V. Rachel Carson using Niskin bottles. The ALOHA, PS, and NP samples were collected by filtering seawater through 142-mm, 0.2-μm polytetrafluoroethylene (PTFE) filters (Omnipore Membrane Filters, Merck Millipore Ltd) using a peristaltic pump into 2-L polycarbonate bottles and stored at −20°C until analysis. The LW sample was collected at a depth of 1 m from the University of Washington Waterfront Activities Center Dock on 20 June 2020. It was filtered through a 47-mm 0.2-μm PTFE filter (Omnipore Membrane Filters, Merck Millipore Ltd) using a glass vacuum filtration setup before being stored in 45-mL polypropylene falcon tubes at −20°C until analysis. Samples were stored at −20°C for a duration of several weeks to 4 yr. We cannot rule out the possibility that storage influenced the measured metabolite concentrations in natural samples. The main objectives of our method development and assessment were to determine the EE and detection limits of metabolites within samples with a variety of salt and DOM matrices. We expect the sample matrices (inorganic ions and background DOM) to be unaffected by storage time. GO-SHIP protocols allow for multi-year storage of samples prior to analysis of bulk DOM, suggesting that while there could be some structural alteration, there is no detectable net loss of the DOM matrix over multiple years of storage (Halewood et al. 2022).To remove trace organic contaminants, all plastic used in collection and sample storage were soaked in a 10% HCl acid bath for 24 h and rinsed three times with MilliQ H_2O and all glass was combusted to 450°C for 5 h before use.

Sample preparation

Cation-exchange solid-phase extraction

The CX-SPE approach to measuring trace organic molecules requires that the water used in all steps of the process (column equilibration, extraction, elution, regeneration, and blank preparation) is free of analytes. In addition to the purification provided by a Milli-Q system (Sigma-Aldrich), further purification was required to ensure the complete removal of all analytes of interest (glycine betaine in particular) (Supporting Information Fig. S1). To prepare analyte-free water, Milli-Q water was rinsed through CX-SPE columns before being used, producing what we term "Cation-Exchange Clean Water" or "CXC H_2O." The washing columns were set up and equilibrated as follows. First, 35 g of strong cation-exchange resin (Dowex 50WX8; H+ form, 100–200 mesh, Sigma-Aldrich) was placed in a glass chromatography column with a fritted disk and a PTFE stopcock (1 in ID × 12 in, Chemglass Life Sciences). The resin was then equilibrated for use by washing with 50 mL of H_2O, 100 mL of 1 M NH_3, 50 mL of H_2O, 100 mL of 3 M HNO_3, and 50 mL of H_2O using gravity flow. CXC H_2O was then generated by flowing Milli-Q H_2O through the washing columns and storing in combusted 2-liter glass bottles with polypropylene caps and kept in the dark until use. CXC H_2O was then used to prepare CXC 1 M NH_3 and 3 M HNO_3 solutions. The CXC H_2O and solutions were then used to prepare a second set of CXC CX-SPE columns for sample analysis in the same manner as the washing columns. The CX-SPE columns were prepared in-house rather than using pre-packed columns to enable thorough cleaning for trace organic compounds, the reuse of columns, and the flexibility to control column design to meet our needs.

Sample processing (outlined in Fig. 1) was performed by acidifying 40 mL of sample to pH 2 using CXC 3 M HNO_3. This sample volume was chosen for convenience during sample collection, storage, and transport, efficient thawing during analysis, and to minimize the chemical waste generated during the extraction procedure. The sample was added to the CXC columns along with a 20-μL spike of our "Extraction" isotopically labeled internal standards dissolved in HPLC grade water and allowed to stand for 5 min (internal standard information detailed in Supporting Information Table S2). Our internal standard concentrations were chosen to be within or slightly above the range of expected natural values for dissolved metabolites (Table 1). The sample was then drained from the column and the column was rinsed with 50 mL CXC H_2O by trickling along the side with a combusted glass Pasteur pipette so as not to disturb the resin. NH_3 (1 M) was added to the column (once again trickled down the side using a glass pipette to avoid disturbing the resin) and eluted in 10 mL fractions into combusted 20-mL glass vials. The pH of each fraction was measured by dabbing a small amount of sample onto a pH strip (Panpeha pH indicator strips, Sigma-Aldrich) using a glass Pasteur pipette. pH strips were chosen over pH probes

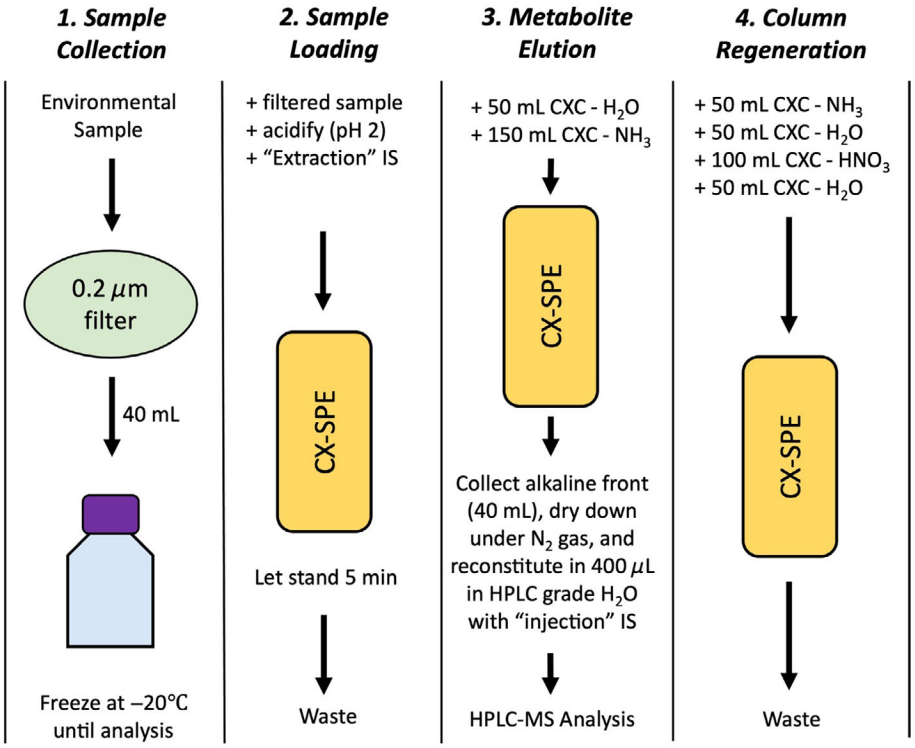

Fig. 1. Visual outline of the CX-SPE workflow for dissolved metabolomics. IS, internal standards.

to minimize cross-contamination among samples. The pH of the fractions was measured until the alkaline front was reached (pH change from 2–4 to 9–11) and the alkaline front fraction, the fraction before, and the two fractions after were collected and combined (40 mL total collected of approximately 150 mL total NH_3 used). These fractions were selected for as they contained, on average, > 90% of the total recoverable signal for compounds in our standards library from the CX-SPE extraction process (Supporting Information Data S1; Table S3; Fig. S2). The combined fractions were dried down in the dark in a TurboVap Evaporator (Biotage) under N_2 gas (flow rate of 1.8 L min^{-1}) in a water bath at 32°C. The dried fractions were stored at −20°C until reconstitution. The column was regenerated through sequential washing with 50 mL 1 M CXC NH_3, 50 mL CXC H_2O, 100 mL 3 M CXC HNO_3, and 50 mL of CXC H_2O. The Teflon-coated stopcock pieces were removed and solvent rinsed sequentially with 10% formic acid, methanol, and dichloromethane to remove any residual trace organics. When in storage, columns were filled with 0.01 M CXC HNO_3.

PPL extraction

PPL-SPE extraction was performed following established literature protocols (Dittmar et al. 2008; Johnson et al. 2017). Agilent Bond Elut PPL cartridges (1 g bed mass, 6 mL volume) were rinsed with 1 cartridge volume of methanol and 1 cartridge volume of 0.01 M HCl. Forty milliliters of sample were acidified to pH 2 with 3 M HCl, spiked with 20 μL "Extraction" internal standards, and pushed through the column using a peristaltic pump (flow rate: 10 mL min^{-1}) followed by 2 cartridge volumes 0.01 M HCl. Finally, the sample was eluted with 1 cartridge volume of methanol, dried down under N_2 gas, and stored at −20°C until reconstitution.

Reconstitution

All samples were reconstituted in 380 μL of HPLC grade H_2O and 20 μL of our "Injection" isotopically labeled internal standards mix dissolved in HPLC grade H_2O for a total volume of 400 μL H_2O (Supporting Information Table S2). As a standard precaution to prevent clogging of the LC or MS systems, each sample was then syringe filtered through a 13-mm syringe filter with 0.22-μm PTFE membranes to remove any potential particulates, and stored at −80°C until analysis. There were no visible particulates in the vials after reconstitution of dry residues. The removal of salts during the CX-SPE extraction was verified by measuring the salinity of reconstituted samples of ALOHA seawater which had a starting salinity of 36 PPT and a post-reconstitution salinity of 0 PPT (Portable Refractometer, Agriculture Solutions). A similar CX-SPE approach using a smaller resin bed volume for an equal sample volume also found that < 0.5% of total dissolved solids remained following SPE (Glabonjat et al. 2018).

Table 1. Concentrations of dissolved metabolites in environmental samples measured in this study with CX-SPE and in other studies. Note that ornithine concentrations are higher than expected, potentially reflecting contamination introduced during sample processing.

Compound or compound class	Concentration measured with CX-SPE (nM)	Locations measured with CX-SPE	Literature concentrations (nM)	Literature locations	References
Individual DFAA	0.15–17	ALOHA, PS, NP, LW	< 0.01–62.5	South Atlantic Transect	Sabadel et al. (2017)*
			0.06–3.0	Bermuda Atlantic Time Series	Widner et al. (2021)†
			0–16.9	Jardines de la Reina, Cuba Reef System	Weber et al. (2020)‡
			0–20	Equatorial Pacific	Lee and Bada (1975)*
			1.0–15	New York Bight	Fuhrman and Ferguson (1986)*
TMAO	0.89–49	ALOHA, PS, NP, LW	< 1.65–76.9	Antarctic Coastal Waters	Gibb and Hatton (2004)§
DMSP	0.013–6.7	ALOHA, PS, NP, LW	1.1–2.8	Sargasso Sea	Kiene and Slezak (2006)§
Creatine	0.20–0.90	ALOHA, PS, NP, LW	6.3–57	North Pacific Subtropical Gyre	Wawrik et al. (2017)‖
Arsenobetaine	0.015	ALOHA	0.0067–0.13	North Atlantic	Glabonjat et al. (2018)¶
Xanthine	0.073–0.092	ALOHA, PS	0–1.06	Jardines de la Reina, Cuba Reef System	Weber et al. (2020)‡
Ectoine	0.046–0.054	ALOHA, NP	0.10–0.37	Bermuda Atlantic Time Series	Widner et al. (2021)†
Citrulline	2.0–2.6	ALOHA, LW	0.28–0.31	Bermuda Atlantic Time Series	Widner et al. (2021)†
Ornithine	8.2–8.3	ALOHA, LW	0.42–0.54	Bermuda Atlantic Time Series	Widner et al. (2021)†
Sarcosine	0.54–1.2	ALOHA, NP, PS	0.09–0.22	Bermuda Atlantic Time Series	Widner et al. (2021)†

*DFAA specific approach.
†BC derivatization/PPL-SPE.
‡PPL-SPE.
§Chemical or enzymatic degradation to volatile analyte.
‖Select SCX-SPE.
¶CX-SPE.

Chromatography and MS

In this study, two complementary types of LC were used to enable the separation and detection of a greater range of compound polarities. RP chromatography uses a non-polar stationary phase and polar mobile phase, resulting in more polar compounds eluting early on while less polar compounds elute later in the run, making it ideal for separating more non-polar compounds for analysis. Hydrophilic interaction liquid chromatography (HILIC) is a type of normal-phase approach that uses a polar stationary phase and a less polar mobile phase. In HILIC, less polar compounds elute early in the run while more polar compounds are better retained on the column, enabling their separation. Each technique has benefits and drawbacks. HILIC chromatography tends to be more variable, with changes in retention time (RT) for compounds over the course of a run. However, RP chromatography is more vulnerable to ion suppression, particularly for more polar compounds that elute with inorganic salts at the beginning of the run and does not achieve separation of important polar metabolites that are structural isomers (same m/z) such as homarine and trigonelline, two compounds that have been highlighted as components of particulate polar metabolite pools (Boysen et al. 2018; Heal et al. 2021). The chromatography and MS methods detailed here are adapted from Boysen et al. 2018 with small changes in the instrument scan range and the HILIC column solvent timings.

HILIC analysis

HILIC chromatography was performed using a SeQuant ZIC-pHILIC column (5 μm particle size, 2.1 mm $\times$ 150 mm,

from Millipore) with 10 mM ammonium carbonate in 85 : 15 water to acetonitrile (Solvent A) and 10 mM ammonium carbonate in 85 : 15 acetonitrile to water (Solvent B) at a flow rate of 0.15 mL min^{-1}. The column was held at 100% B for 2 min, ramped to 64% A over 18 min, ramped up to 100% A over 1 min, held at 100% A for 7 min, and equilibrated at 100% B for 22 min (total time is 50 min). The column was maintained at 30°C throughout the analysis. The injection volume was 2 μL.

RP analysis

RP chromatography was performed using a Waters Acquity UPLC HSS Cyano column (1.8 μm particle size, 2.1 mm × 100 mm) equipped with an Acquity UPLC HSS Cyano guard column (1.8 μm particle size, 2.1 mm × 5 mm) with 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B) at a flow rate of 0.4 mL min^{-1}. The column was held at 5% B for 2 min, ramped to 100% B over 18 min, held at 100% B for 2 min, and equilibrated at 5% B for 5 min (total run time is 25 min). The column was maintained at 35°C throughout the analysis. The injection volume was 15 μL.

Mass spectrometry

Targeted and untargeted MS was performed using a Thermo Orbitrap Q-Exactive HF Mass Spectrometer (QE) using the settings detailed in Boysen et al. 2018. For HILIC analysis, a capillary temperature of 320°C, an Heated-Electrospray Ionization probe (H-ESI) spray voltage of 3.5 kV, an auxiliary gas heater temperature of 90°C, an S-lens RF level of 65, a sheath gas flow rate of 16 L h^{-1}, an auxiliary gas flow rate of 3 L h^{-1}, and a sweep gas flow rate of 1 L h^{-1} were used. Polarity switching was employed with a scan range of 60–900 m/z and a resolution of 60,000. For RP analysis, a capillary temperature of 320°C, an H-ESI spray voltage of 3.8 kV, an auxiliary gas heater temperature of 90°C, an S-lens RF level of 65, a sheath gas flow rate of 40 L h^{-1}, an auxiliary gas flow rate of 10 L h^{-1}, and a sweep gas flow rate of 1 L h^{-1} were used. A full scan method was used in positive mode with a scan range of 60–900 m/z and a resolution of 120,000. As in Boysen et al. 2018, polarity switching was used with the HILIC analysis as many compounds in our targeted compound lists ionized well in either positive or negative ionization modes, allowing us to successfully measure these compounds. However, for RP analysis, most compounds only ionized well in positive ionization mode so a single ionization mode was chosen to allow for higher resolution data collection.

Data processing and quality control

Targeted

MS data from the QE were converted to .mzML files using MSConvert (Chambers et al. 2012). All peaks of targeted compounds and internal standards were manually integrated using Skyline, with peak identification performed by comparison to authentic standards in matrix for both RT and peak shape (Adams et al. 2020). Quality control was then performed to remove low-quality peaks from the data by requiring that peaks met requirements for minimum area (40,000 for HILIC, 5000 for RP) and exact mass (< 6 ppm difference). Best-matched internal standard normalization (B-MIS) was applied to account for changes in instrument response throughout the run (Boysen et al. 2018). All data analysis and visualization was performed in the R statistical environment (version 4.2.0) using R Studio (version 2022.02.0) with the the "readr," "dyplr," "stringr," "ggplot2," and "RaMS" packages.

Untargeted

To expand the analytical window and further compare CX-SPE and PPL-SPE extraction approaches beyond our standards library, an untargeted metabolomics approach was employed. Untargeted metabolomics decreases the analytical bias inherent in targeted metabolomics approaches by allowing for the comparison of a greater number of mass features (MFs) within each sample. ALOHA and PS samples were prepared using both CX-SPE and PPL-SPE and analyzed using both HILIC and RP chromatography. MFs were then detected and integrated using MSDial (*see* Supporting Information Table S4 for parameters) separately for HILIC-positive, HILIC-negative, and RP-positive modes, normalized using B-MIS normalization and filtered using the quality control filters detailed below to curate a final list of high-quality MFs (Tsugawa et al. 2015). All samples were aligned to pooled samples made up of both the ALOHA and PS samples analyzed with both the CX-SPE and PPL-SPE enabling the detection of the same MFs across samples and methods. MFs were required to have a signal-to-noise ratio of 5, a minimum normalized peak area of 5000, a relative standard deviation (RSD) within the pooled samples of < 0.3, and a peak area at least 3 times greater than the average value for that MF in the corresponding methodological blank. HILIC MFs with RTs less than 2 min and greater than 20 min and RP MFs with RTs less than 0.5 min and greater than 17.8 min were discarded as these regions demonstrate high levels of ion suppression (Boysen et al. 2018). These parameters were applied to each analytical triplicate independently and MFs not passing all quality control parameters in each individual analytical triplicate were discarded.

Assessment

Identification of extracted compounds, extraction efficiencies, and linearity

We identify compounds extracted by our CX-SPE method by measuring EEs, RSD of the EEs (RSD_{EE}), and the linearity of method and instrument response to standard spikes at different environmentally relevant concentrations (R^2) of our in-house library of authentic standards. Compounds included in the compound library are molecules involved in fundamental biological processes (e.g., amino acids, intermediates in central carbon metabolism, nucleobases), signaling, cross feeding, or bioactivity (e.g., B vitamins, homoserine lactones) (Sañudo-Wilhelmy et al. 2014; Hmelo 2017). Compounds involved in

microbial adaptation to salt or other stressors (e.g., compatibles solutes such as glycine betaine or DMSP) that enable organisms to handle osmotic stress and have been documented to have high concentrations in marine phytoplankton and bacteria (Welsh 2000) are also a part of our compound library. While not exhaustive, the compound library employed in this study includes many of the most abundant compounds documented in marine particulate community metabolomes (Johnson et al. 2020; Boysen et al. 2021; Heal et al. 2021) as are compounds that are suggested by previous work to be actively cycled and potentially abundant in marine dissolved metabolomes (Gibb and Hatton 2004; Kiene and Slezak 2006; Poretsky et al. 2010; Fiore et al. 2015; Sabadel et al. 2017; Mayali and Weber 2018; Vorobev et al. 2018; Weber et al. 2020; Ferrer-González et al. 2021; Widner et al. 2021; Boysen et al. 2022). Our in-house library consists of 179 standards separated and analyzed using HILIC chromatography (HILIC Standards) and 70 standards separated and analyzed using RP chromatography (RP Standards) (Supporting Information Table S1). Using HILIC chromatography, leucine and isoleucine do not achieve chromatographic separation and therefore are integrated together and referred to collectively as (iso)leucine.

The EE of a compound is the percentage of a compound recovered at the end of sample preparation compared to the initial amount present in the sample. The median spike concentration for HILIC Standards was 25 nM, and the median spike concentration for RP Standards was 2.5 nM (spike values for each compound are presented in Supporting Information Table S1). For method evaluation, standard spike concentrations were required to be at least as concentrated as the most abundant environmental concentrations of common metabolites (~ 1–10 nM, Table 1) and high enough such that a high-quality signal on the instrument was achieved after reconstitution, as determined previously for this combination of chromatography and instrument (Boysen et al. 2018).

To determine EEs, three sets of samples were run on the columns in triplicate: a "spike before" sample where the standard spike was added to the natural seawater or lake water before processing the sample, a "spike after" sample where the standard spike was added during reconstitution, and a "no spike" sample to which no standard spike was added. Standards were dissolved in 1 mL HPLC H_2O with a concentration of each compound of approximately 1000 nM for most HILIC standards and 100 nM for most RP standards. These standard spikes result in the final concentrations reported in Supporting Information Table S1 when combined with 40 mL of sample. We are not aware of any potential interferences in our standards mixes due to cross reactions. EEs were then determined using the following equation:

$$\text{Extraction Efficiency (EE)} = \frac{\text{Area}_{\text{spike before}} - \text{Area}_{\text{no spike}}}{\text{Area}_{\text{spike after}} - \text{Area}_{\text{no spike}}} \times 100\%. \quad (1)$$

Subtracting the "no spike" sample from both the "spike before" and "spike after" samples removes background noise as well as the signal contributed by the natural abundance of the compound from the EE calculation. EEs were determined in triplicate for all four environmental sample matrices (ALOHA, NP, PS, LW) for the HILIC standards and in ALOHA and PS sample matrices for the RP standards. An EE value was calculated separately for each triplicate in each sample and the RSD_{EE} for each compound was calculated from the standard deviation of all EE values ($n = 12$ for HILIC standards, $n = 6$ for RP standards), divided by the mean EE value, and multiplied by 100 to attain a percentage value.

The linearity of the EE of the compounds in ALOHA seawater was evaluated by spiking in our standards mix at 100%, 50%, or 10% of our original spike concentration (e.g., 25, 12.5, and 2.5 nM) in triplicate to build a calibration curve for each compound. The goal of this test was to determine if there was any variability in EE in response to changes in concentrations. The dynamic ranges of these compounds on this instrument were established in Boysen et al. (2018). The R^2 value was calculated using a linear model to assess the linearity of the relationship between spike concentration and peak area. A strongly linear relationship between these two values indicates that EE is not changing in response to changes in analyte concentration. This range of concentrations was chosen as the lowest values that still allowed for a spike signal to be detected despite the natural abundance of many compounds. We did not test higher concentrations as we did not expect many environmental metabolite concentrations to be above 25 nM.

Compounds were considered successfully extracted by CX-SPE when they met three extraction quality control thresholds indicating that the method could be used to generate reproducible metabolomics data in environmental samples. Compounds were required to have EEs greater than 1% and below 150%, RSD_{EE} values of less than 0.5, and R^2 values greater than 0.7. These values were selected to remove compounds that were not extracted through CX-SPE (EEs < 1%), demonstrated very high levels of variability in calculated EEs (RSD_{EE} values > 0.5 or $R^2 < 0.7$), or demonstrated increases in concentration potentially related to the breakdown of other standards into these molecules (e.g., acetylcarnitine into carnitine) during the CX-SPE procedure (EE > 150%). However, we also attempted to avoid over penalizing compounds that may have suffered from higher or lower levels of ion-suppression in the "spike after" or "spike before" samples due to the unnaturally high total analyte concentrations in these samples introduced by the standard spike. The final selection of compounds meeting the extraction quality control thresholds were designated as "reproducibly extracted" compounds and were used for further analysis to compare method performance across different sample matrices. A total of 69 targeted compounds (Table 2) were deemed reproducibly extracted using CX-SPE after passing quality control thresholds for EE,

Table 2. *Analysis fraction*, EE, RSD_{EE}, R^2 values, and LODs for compounds meeting quality control thresholds using CX-SPE. Isoleucine and leucine are not fully separated with our chromatography and are treated together as (iso)leucine.

Compound	Fraction	EE (%)	RSD of EE (%)	R^2	LOD (nM)
(3-Carboxypropyl)trimethylammonium	HILIC positive	46.8	27.6	0.964	0.012
(Iso)leucine	HILIC positive	59.7	17.6	0.949	0.23
3′,5′-Cyclic adenosine monophosphate	HILIC negative	5.92	39.7	0.886	0.0071
4-Aminobutyric acid	HILIC positive	89.2	16.9	0.934	0.67
5-Hydroxyectoine	HILIC positive	47.2	16.0	0.994	0.0074
5-Methylcytosine	HILIC positive	21.0	15.1	0.962	0.042
5-Oxoproline	HILIC negative	18.1	38.5	0.912	17
Abscisic acid	RP	4.77	38.1	0.925	0.62
Adenine	HILIC positive	49.1	18.4	0.954	0.23
Adenosine	HILIC positive	10.6	20.9	0.953	0.12
Adenosine monophosphate	HILIC negative	23.0	39.3	0.961	0.027
Allopurinol	HILIC negative	46.0	38.6	0.859	1.2
Arsenobetaine	HILIC positive	88.9	14.1	0.992	0.0013
β-Alanine	HILIC positive	82.1	15.4	0.977	1.3
β-Alaninebetaine	HILIC positive	66.2	15.1	0.991	0.014
β-Glutamic acid	HILIC positive	52.9	12.7	0.843	1.0
Betonicine	HILIC positive	89.2	14.1	0.996	0.0059
Butyrylcarnitine	HILIC positive	5.56	39.2	0.930	0.0090
Citrulline	HILIC positive	40.2	27.7	0.857	1.5
Creatine	HILIC positive	50.4	19.5	0.980	0.047
Cytidine	HILIC positive	19.9	21.3	0.970	0.038
Cytosine	HILIC positive	30.8	15.2	0.969	0.050
Desthiobiotin	RP	66.1	38.3	0.784	0.0042
Dimethylglycine	HILIC positive	65.1	14.9	0.993	0.70
DMS-Ac	HILIC positive	94.8	12.6	0.996	0.00013
DMSP	HILIC positive	55.5	15.5	0.973	0.00010
Ectoine	HILIC positive	47.6	12.5	0.990	0.025
Glucosamine	HILIC positive	49.4	15.9	0.988	0.21
Glutamylphenylalanine	HILIC positive	26.0	23.4	0.923	0.012
Glycerophosphocholine	HILIC positive	2.24	47.8	0.931	0.25
Glycine betaine	HILIC positive	100	20.6	0.969	0.38
Gonyol	HILIC positive	71.0	26.2	0.974	0.17
Guanine	HILIC positive	47.1	26.0	0.955	0.35
Guanosine	HILIC positive	13.2	30.0	0.952	0.17
Homarine	HILIC positive	50.0	15.2	0.995	0.15
Hordenine	HILIC positive	7.29	41.9	0.732	0.060
Hydroxyisoleucine	HILIC positive	63.2	39.1	0.916	0.25
Hydroxyproline	HILIC positive	54.5	16.1	0.965	0.041
Hypoxanthine	HILIC positive	59.6	18.7	0.978	0.39
Inosine	HILIC negative	36.7	28.2	0.950	0.023
L-alanine	HILIC positive	106	30.3	0.974	1.8
L-asparagine	HILIC positive	82.7	14.4	0.916	0.49
L-aspartic acid	HILIC positive	67.2	18.5	0.963	3.4
L-glutamic acid	HILIC positive	53.7	22.4	0.850	2.0
L-glutamine	HILIC positive	131	16.8	0.988	0.024
L-histidine	HILIC positive	33.0	36.8	0.969	0.66
L-Homoserine	HILIC positive	68.8	18.1	0.986	0.18
L-Hydroxylysine	HILIC positive	38.2	23.4	0.732	0.0026

(Continues)

Table 2. Continued

Compound	Fraction	EE (%)	RSD of EE (%)	R^2	LOD (nM)
L-lysine	HILIC positive	56.2	39.1	0.888	1.5
L-methionine S-oxide	HILIC positive	72.6	18.6	0.973	0.16
L-ornithine	HILIC positive	47.2	30.7	0.907	6.8
L-phenylalanine	RP	128	21.0	0.914	0.20
L-proline	HILIC positive	104	16.4	0.984	0.51
L-serine	HILIC positive	75.3	38.4	0.919	4.0
L-threonine	HILIC positive	62.2	17.5	0.942	0.21
L-tyrosine	HILIC positive	28.6	26.7	0.938	0.88
Melamine	HILIC positive	60.2	18.2	0.975	0.36
Muramic acid	HILIC positive	14.6	25.5	0.968	0.046
N6-acetyl-L-lysine	HILIC positive	37.8	26.4	0.984	0.015
N6-Methyladenine	HILIC positive	15.8	29.2	0.917	0.020
Nicotinic acid	RP	102	12.2	0.894	1.1
O-Methylmalonyl-L-carnitine	HILIC positive	47.4	20.7	0.989	0.0064
Ophthalmic acid	HILIC positive	23.0	27.8	0.959	0.020
Proline betaine	HILIC positive	58.2	17.3	0.985	0.022
Sarcosine	HILIC positive	41.0	14.4	0.968	0.32
Trigonelline	HILIC positive	53.3	15.7	0.987	0.057
TMAO	HILIC positive	54.7	33.0	0.946	0.13
Urocanic acid	HILIC negative	65.6	17.6	0.939	0.73
Xanthine	RP	69.2	22.7	0.965	0.055

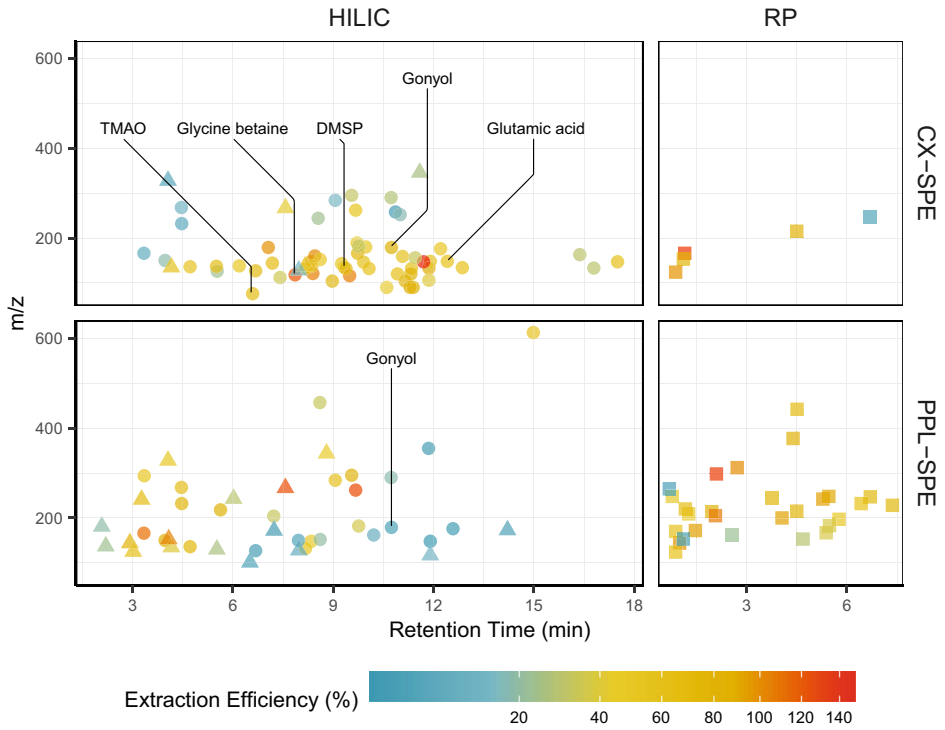

Fig. 2. Extraction efficiency of targeted compounds extracted by CX-SPE and PPL-SPE shown by the compounds' m/z and RT for both types of chromatography. Different analytical fractions are denoted by shape (HILIC positive, circle; HILIC negative, triangle; RP, square). Several compounds representative of various functional classes of metabolites extracted by CX-SPE (TMAO, amine oxides; homarine and glycine betaine, betaines; DMSP and gonyol, sulfoniums; glutamic acid, amino acids) are annotated. Note that the color bar has a square root transformation.

RSD_{EE}, and R^2 values. Seventy of the compounds were HILIC standards (6 negative mode, 58 positive mode) while 5 were RP standards. For compounds passing quality control thresholds, EEs range from 2.24% for glycerophosphocholine to 131% for L-glutamine (Table 2, Fig. 2). RSD_{EE} values ranged from 12.2% for nicotinic acid to 47.8% for glycerophosphocholine. R^2 values ranged from 0.731 for hordenine to 0.996 for betonicine. Example chromatograms of several compounds in ALOHA seawater at natural abundance and with a range of standard spike concentrations are shown in Fig. 3.

Method performance across environmental samples

To assess the robustness of the method to differences in environmental sample matrix, the EEs and response factors (RFs) of the reproducibly extracted compounds in the four different sample types were compared. RFs are the amount of signal detected on the instrument relative to concentration of that compound present in the sample and can vary over orders of magnitude in electrospray ionization-mass spectrometry (ESI-MS). Four samples (ALOHA, NP, PS, LW) representing a range of DOM concentrations (oligotrophic [low DOM]–eutrophic [high DOM]) and salinities (freshwater to open-ocean seawater) were chosen. The EEs and RFs of the HILIC standards were compared for all four environmental samples in triplicate while the EEs and RFs of the RP standards were compared for just ALOHA and Puget Sound samples in triplicate (Supporting Information Table S5). RF values were calculated using eq. 2.

$$RF = \frac{\text{Area}_{\text{spike after}} - \text{Area}_{\text{no spike}}}{\text{Spike Concentration}}. \quad (2)$$

The EE and RF values for each sample matrix were then compared by one-way ANOVA and the *p*-values were then adjusted to control the false discovery rate (FDR) through the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995) (Supporting Information Table S5).

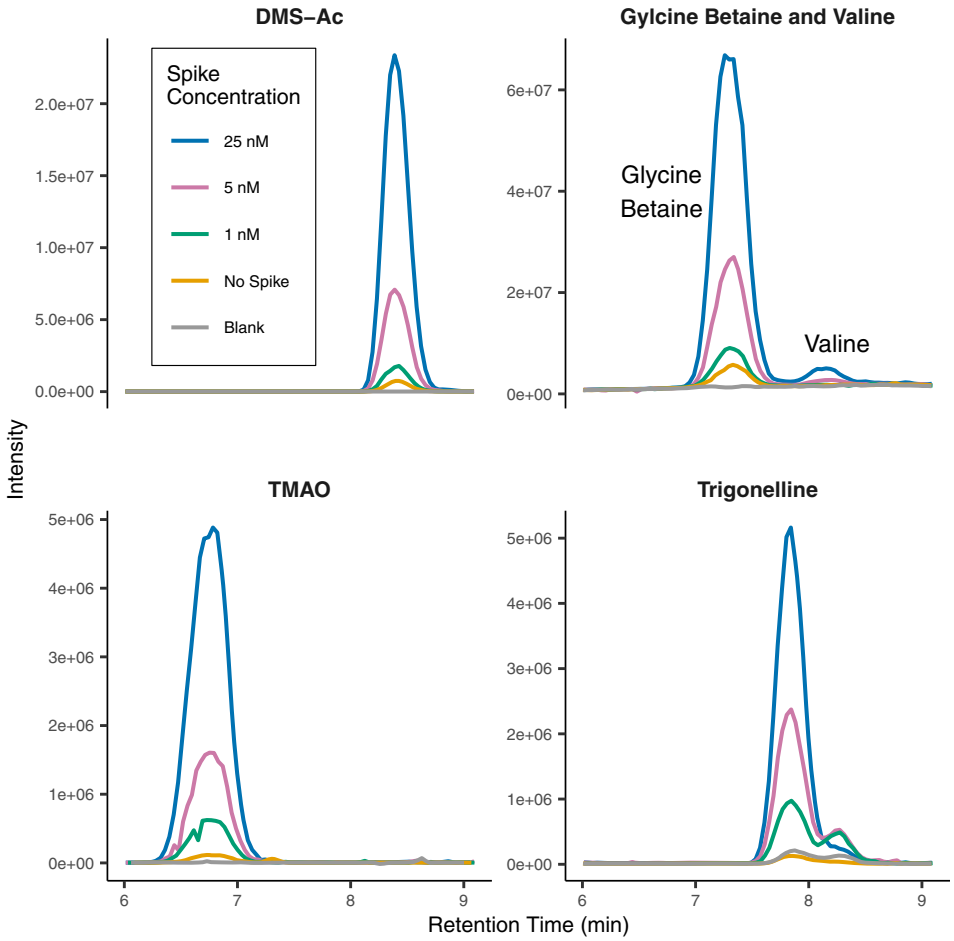

Fig. 3. Extracted ion chromatograms (EIC) of selected compounds at different standard spike concentrations in ALOHA seawater processed using CX-SPE. "Blank" is a CXC H_2O blank. "No spike" is an unamended sample of ALOHA seawater; "1 nM," "5 nM," and "25 nM" are samples where the depicted standards were spiked into a sample of ALOHA seawater at 1, 5, and 25 nM concentrations before processing with CX-SPE. These chromatograms were not used in any of the method assessment presented in this study.

Differences in EE and RF were considered significant if the FDR-adjusted *p*-value was below 0.05. No compounds displayed significant differences in EE, suggesting that CX-SPE EEs are robust to changes in salinity and DOM concentration. Only two compounds, adenosine monophosphate and L-glutamic acid, had significant differences in RF values across environmental matrices, suggesting that CX-SPE helps minimize differences in instrument response caused by sample variation for the majority of compounds.

Matrix effects

Matrix effects in LC–MS refer to the impact of the sample matrix (solvent, salts, and coeluting organic compounds present in the solution) on the ionization efficiency of target analytes. Ionization efficiencies are affected by the presence of matrix components that influence the extent of ionization of target analytes in the electrospray source. One goal of SPE is to reduce matrix effects by separating analytes from the sample matrix. To assess the ability of CX-SPE to remove matrix effects, a RF ratio for each compound was calculated. RF ratios equal to 1 are indicative of no matrix effects, those below 1 indicate ion suppression caused by the matrix, and those above 1 indicate ionization enhancement. Our HILIC and RP standard mixes were spiked at equal concentrations into a pooled sample matrix of our four environmental samples (matrix) and into HPLC grade water (water). Background concentrations of the analytes in the matrix were assessed by spiking CXC H_2O water into the matrix at an equal dilution factor to that of the authentic standards. RFs were calculated for each compound in the pooled matrix and in water and the ratio of these two RFs was determined.

$$\text{RF ratio} = \frac{\text{RF}_{\text{matrix}}}{\text{RF}_{\text{water}}}. \quad (3)$$

The RF ratio of 68 of the 69 successfully extracted compounds was between 0.89 and 1.16 with one RP standard, nicotinic acid, having a much higher RF ratio (1.56) (Supporting Information Fig. S3). The RF ratios determined here are much closer to 1 than for particulate metabolite analyses using an identical LC–MS analysis where many compounds exhibited substantial ion suppression (RF ratio < 0.5), indicating that CX-SPE excels at removing matrix effects, particularly for compounds analyzed with HILIC chromatography (Boysen et al. 2018).

Limits of detection

We define the limit of detection (LOD) as the concentration at which the chance of a false positive is 5% (Currie 1968). We determined the LODs for each compound by calculating the 95% confidence interval of signal in the CXC H_2O methodological blanks using the mean and standard deviation of the integrated peak area of the m/z and RT window of each compound. LODs can be difficult to assess for ESI-MS measurements because of the difficulty in attaining a matrix matched blank while also being analyte free. However, CX-SPE demonstrated low matrix-effects (Supporting Information Fig. S3), suggesting that CXC H_2O serves as an appropriate sample for determining the "true" blank value of the compound (Glabonjat et al. 2018). Furthermore, we observed no indication of significant changes in EE or RF under different salinities or DOM concentrations for most metabolites, suggesting that these factors do not impact CX-SPE analysis. CXC H_2O blanks were run throughout the analysis of the environmental samples for a total of 15 blanks. LODs were calculated for each compound using the following equation where B is the average value of the blank, S_B is the standard deviation of the blank, 1.761 is Student's T-value for a 95% confidence interval, and n is the number of blanks. The RF, RF ratio, and EE are used to convert the LOD from a signal value to an environmental concentration. For two compounds, cytidine and inosine, no signal was observed in the blanks, initially resulting in a calculated LOD of 0 nM which is not a true representation of method performance. To control for this, the signal blank value (B) was set at half the quality control threshold for a real peak (20,000) to enable a LOD calculation for these compounds. This adjustment resulted in LOD values for these three compounds that fell well within the range of the LOD values calculated for the other compounds.

$$\text{LOD} = \left(B + \frac{1.761(s_B)}{\sqrt{n}}\right) \times \frac{1}{\text{Response Factor}} \times \frac{1}{\text{RF ratio}} \times \frac{1}{\text{EE}/100}. \quad (4)$$

LOD ranged from the low-pM range for DMSP and DMS-Ac (0.000103 and 0.000131 nM, respectively) to the low nM for L-serine and L-ornithine (4.04 and 6.84 nM, respectively) (Table 2). The large variation in these values is due to the large range in instrument RFs and background ion concentrations in the blanks.

Comparison of CX-SPE and PPL-SPE

Targeted

To compare the performance and analytical windows of CX-SPE and the commonly used PPL-SPE approach, we compared the compounds extracted from our compound-library using each approach in ALOHA and PS seawater. PPL-SPE compounds were considered successfully extracted if they met the same quality control thresholds of EE (1% < EE < 150%) and RSD_{EE} (RSD < 50%) as in the CX-SPE analysis. PPL-SPE successfully extracted 71 compounds, 25 in the HILIC-positive fraction, 18 in HILIC negative, and 28 in RP (Supporting Information Table S5). EEs ranged from 1.8% to 136.2% with RSD_{EE} values between 4.4% and 49.1% (Fig. 2, Supporting Information Table S5).

Between CX-SPE and PPL-SPE, a total of 116 compounds were extracted but only 24 compounds (21%) were extracted by both methods, indicating that the two methods are suited to extracting unique types of compounds (Supporting

Information Table S1). We compared CX-SPE and PPL-SPE compounds using the cheminformatics classification system Chem-Ont (Djoumbou Feunang et al. 2016) (Supporting Information Fig. S4) and elemental composition using H/C and O/C ratios (Supporting Information Fig. S5) to attempt to distinguish between these two groups but did not identify clear differences between the methods, potentially as a result of the limited coverage and potential bias of our finite compound library.

Untargeted

The curated lists of untargeted MFs for CX-SPE and PPL-SPE were compared by RT and mass to charge ratio (m/z) to characterize and distinguish the fractions of the untargeted, dissolved metabolome extracted by each SPE approach. It is worth noting that while each unique m/z and RT identifier likely represents a unique compound, some isomers may not be fully separated using our chromatography, some compounds may be duplicated in both HILIC and RP chromatography or in both positive and negative ion modes, and some compounds may result in multiple mass spectral peaks due to having multiple ionization states or adducts. However, we expect that the highly conservative quality control thresholds we applied removed the vast majority of duplicated peaks for a single compound. RT serves as a proxy for polarity where higher HILIC RTs and lower RP RTs indicate greater polarity. The m/z of a MF is primarily a measure of molecular weight, particularly with singly charged ions. Comparisons were made using the non-parametric Kruskal–Wallis approach for samples that do not meet assumptions of normality or homogeneity of variances. In addition, the sample/method pairs (e.g., PS/CX-SPE) were compared by the presence and absence of MFs determined using each type of chromatography to identify the degree of overlap between the two SPE approaches and to determine if the different samples contained unique metabolites.

The untargeted analysis of both CX-SPE and PPL-SPE resulted in a total of 244, 91, and 1458 MFs for the HILIC-positive, HILIC-negative, and RP fractions, respectively, for a total of 1793 total MFs that passed all quality control thresholds (Supporting Information Fig. S6; Table S7). This set of untargeted MFs includes compounds characterized using the targeted approach. HILIC-positive and HILIC-negative features were combined (hereafter HILIC features) for further comparisons (resulting in a total of 335 HILIC features) while acknowledging that there may be some compounds that are represented in both HILIC-positive and HILIC-negative fractions. We also tentatively identified 31 MFs (1 HILIC-negative feature, 25 HILIC-positive features, and 5 RP features) as compounds within our standards library through comparisons to m/z and RT (Supporting Information Table S7).

PPL-SPE resulted in a greater total number of MFs compared to CX-SPE for both forms of chromatography and in both samples. For the RP features, 1036 (71.1%) were uniquely present in PPL-SPE samples, 143 (9.8%) were uniquely present in CX-SPE samples, and 279 (19.1%) were present in both PPL-SPE and CX-SPE samples. For the HILIC features, 219 (65.3%) were uniquely present in PPL-SPE samples, 87 (26%) were uniquely present in CX-SPE samples, and 29 (8.7%) were present in samples from both methods. The two samples had similar numbers of total MFs (PPL-SPE and CX-SPE combined) with the ALOHA and PS samples having 1324 and 1419 MFs, respectively (Supporting Information Fig. S6). ALOHA and PS HILIC MFs were not significantly different in terms of average RT or average m/z for both CX-SPE (Kruskal–Wallis, $p = 0.34$; $p = 0.85$, respectively) and PPL-SPE (Kruskal–Wallis, $p = 0.46$; $p = 0.19$, respectively) and so ALOHA and PS HILIC MFs were grouped together for further comparisons of the CX-SPE and PPL-SPE methods. ALOHA and PS RP MFs were significantly different in terms of average RT and average m/z for both CX-SPE (Kruskal–Wallis, $p = 0.031$; $p = 0.000081$, respectively) and PPL-SPE (Kruskal–Wallis, $p = 0.0026$; $p = 0.00014$, respectively) and so the two samples were compared separately.

CX-SPE MFs were more polar than PPL-SPE MFs. CX-SPE MFs had a significantly higher average RT in HILIC chromatography (Kruskal–Wallis, $p = < 2.2 \times 10^{-16}$) and a significantly lower average RT in RP chromatography in both the ALOHA and PS samples (Kruskal–Wallis, $p = 0.00033$; $p = 4.24 \times 10^{-15}$, respectively) (Fig. 4). Both types of chromatography suggest that CX-SPE extracts a more polar set of compounds than PPL-SPE. CX-SPE MFs also had significantly lower average m/z values than PPL-SPE MFs, suggesting that CX-SPE extracts a set of compounds with a lower average molecular weight than PPL-SPE (Fig. 4). The mean m/z values of CX-SPE MFs were significantly lower for both HILIC chromatography (Kruskal–Wallis, $p = <2.2 \times 10^{-16}$) and in RP chromatography for both ALOHA and PS samples (Kruskal–Wallis, $p = 0.036$; $p = 0.00074$, respectively). These results show that CX-SPE and PPL-SPE extract different subsets of the low-molecular-weight DOM pool.

Field application of CX-SPE for targeted dissolved metabolomics

To assess the utility of CX-SPE for characterizing and quantifying dissolved metabolomes, we characterized four distinct environmental samples (ALOHA, PS, NP, LW) using HILIC chromatography and two samples using RP chromatography (ALOHA, PS). Calculations were performed using eq. 4 where A_{norm} is normalized peak area.

$$\text{Concentration} = A_{\text{norm}} \times \frac{1}{\text{Response Factor}} \times \frac{1}{\text{RF ratio}} \times \frac{\text{Vol}_{\text{reconstituted}}}{\text{Vol}_{\text{sample}}} \times \frac{1}{\text{EE}/100}. \quad (4)$$

Using our CX-SPE approach we quantified 21, 34, 36, and 41 dissolved metabolites in the NP, LW, PS, and ALOHA samples, respectively, that were above LODs in our four environmental

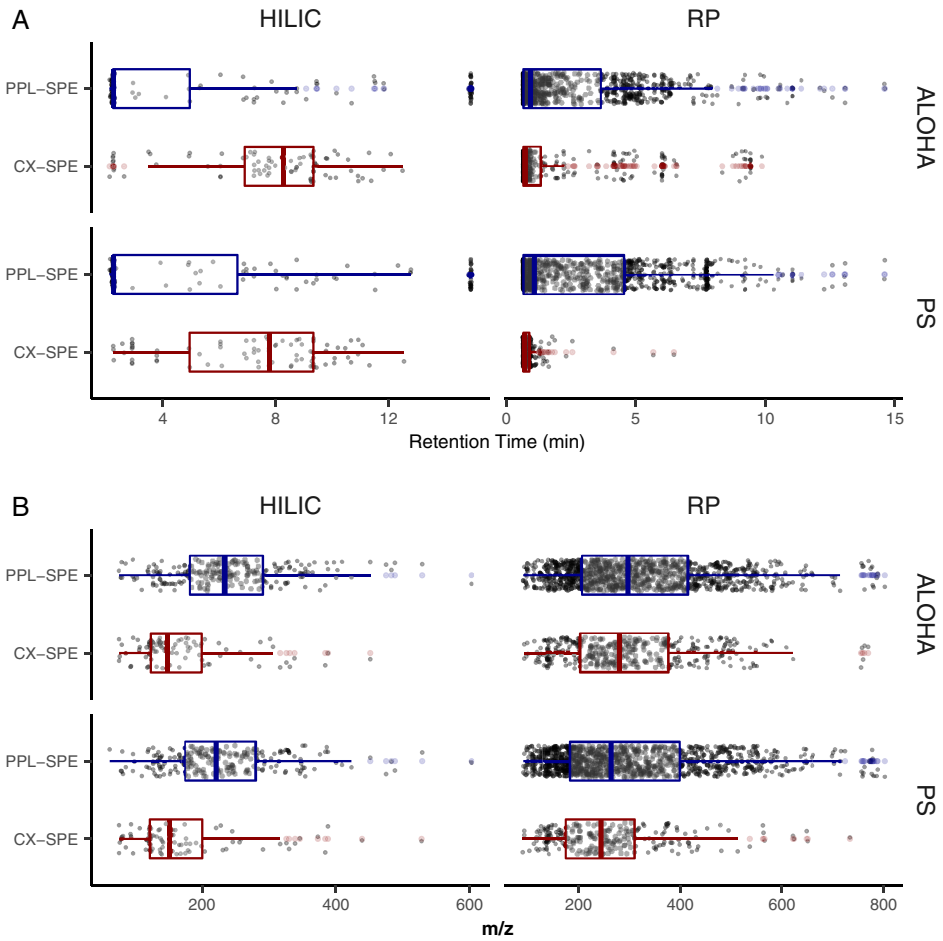

Fig. 4. Scatter and box plots showing the distribution of RTs (**A**) and m/z values (**B**) of untargeted, high-quality MFs generated by CX-SPE and PPL-SPE using HILIC and RP chromatography from ALOHA and PS samples.

samples, ranging in concentration from 0.0093 nM desthiobiotin in the ALOHA sample to 48.5 nM for and TMAO in the ALOHA sample (Fig. 5, Supporting Information Fig. S7; Table S8). Amino acids such as glutamic acid and proline as well as other osmolytes such as the nitrogen-containing compounds glycine betaine and TMAO and the sulfoniums DMSP, DMS-Ac, and gonyol were among the most abundant compounds measured (Fig. 5, Supporting Information Table S8). The freshwater LW sample was noticeably different than the marine samples as many of the major osmolytes that contributed significantly to the marine samples were below detection limits and LW contained higher relative percentages of many amino acids (Fig. 5). We also characterized two environmental samples using PPL-SPE using the same quantification approach as was used with CX-SPE. Although we did not attempt to calculate a true LOD, we determined a conservative "applied LOD threshold" using eq. 4 and just three methodological blanks of CXC H_2O analyzed with PPL-SPE (Supporting Information Figs. S8, S9; Table S9). The PPL-SPE approach resulted in the quantification of 20 and 21 targeted compounds above the applied LOD threshold in ALOHA and PS seawater, respectively. Concentrations ranged from 0.00017 nM methylthioadenosine in the ALOHA sample to 9.31 nM tryptophan, also in the ALOHA sample. It should be noted that the targeted compound library used in this study is focused on small, polar metabolites and likely lead to more targeted compounds being successfully measured with CX-SPE compared with PPL-SPE.

Discussion

In the ocean, the dissolved polar metabolite pool is a small but rapidly cycled pool of carbon and other essential elements through which roughly half of net primary production flows (Moran et al. 2022).The magnitude of this flux makes measuring and understanding the cycling of these highly labile molecules essential for understanding marine biogeochemical cycles and marine microbiology. The CX-SPE method presented here provides an effective way to extract small, polar, organic molecules from seawater that either have a positive

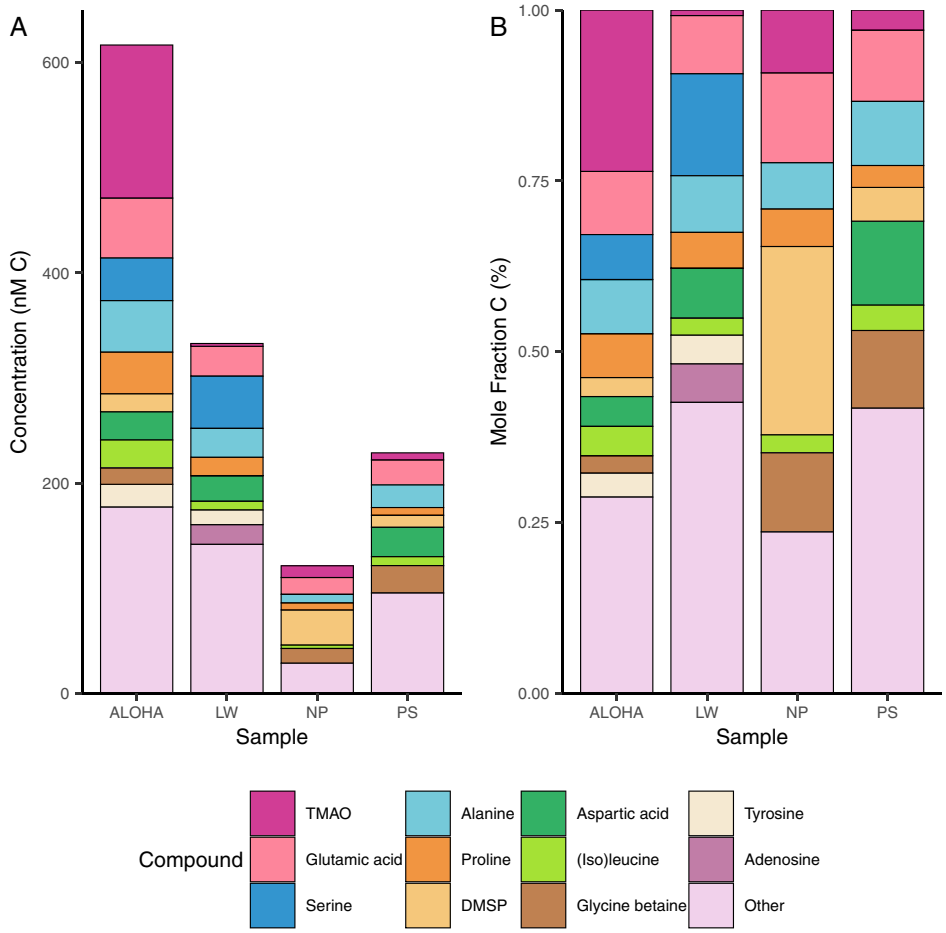

Fig. 5. Average concentration in nM C (**A**) and average relative concentration in mole fraction C (**B**) of compounds quantified by CX-SPE in nM in four discrete environmental samples. The 11 compounds that contribute the most carbon are shown individually while all other compounds are grouped together.

charge or can be protonated by acidifying the sample. This sample preparation approach paired with a combination of HILIC and RP LC and high-resolution MS enables the measurement of dissolved metabolites at environmentally relevant concentrations. Key aspects that led to the success of the method presented here include the trace-organics clean protocols (the use of CXC H_2O for all reagents and blanks), the near complete removal of salts and background organic matter from the analytes of interest resulting in very low matrix effects, and the use of HILIC chromatography which enables the separation of very polar metabolites such as betaines (e.g., glycine betaine and homarine), amino acids (e.g., glutamic acid and proline), and sulfonium ions (e.g., DMSP, DMS-Ac, and gonyol) that are highly abundant components of marine particulate metabolite pools (Johnson et al. 2020; Boysen et al. 2021; Heal et al. 2021).

We show in both the targeted and untargeted comparisons that CX-SPE and PPL-SPE extract fundamentally different portions of dissolved metabolite pools, suggesting that each approach provides a complimentary window into marine DOM. In the targeted study, among the 116 extracted compounds, only 24 compounds were successfully extracted by both methods (Supporting Information Table S1). PPL-SPE resulted in a higher total number of MFs in both HILIC and RP LC–MS analyses which fit our expectations given that PPL columns are favored for DOM analysis, where they retain the greatest total percentage of DOM on SPE resins that have been tested (Dittmar et al. 2008). PPL-SPE also resulted in a set of MFs that had a higher average molecular weight and lower polarity than CX-SPE (Fig. 4), suggesting that these features are more representative of the much higher concentration, less-polar, refractory DOM pool (Moran et al. 2016). Compared to PPL-SPE, CX-SPE extracted a smaller number of MFs that had a lower molecular weight and were more polar, making these features likely more representative of the dissolved polar metabolite pool rather than the background DOM pool (Fig. 4). The targeted features uniquely extracted by CX-SPE include highly abundant known groups of marine

particulate metabolites (free amino acids, betaines, sulfoniums) that are not successfully extracted using PPL-SPE (Johnson et al. 2020; Boysen et al. 2021; Heal et al. 2021). CX-SPE and PPL-SPE are complimentary analyses, which, when combined, provide an analytical window into metabolites with a range of polarities.

CX-SPE compared favorably in method performance to other dissolved metabolomics approaches, with each approach enabling access to a unique set of compounds. CX-SPE LODs ranged from 0.00010 to 6.8 nM, which is similar to the derivatization/PPL method (0.010–150 nM) with both methods enabling measurements of compounds in the pM to low nM range expected of dissolved metabolites in seawater (Widner et al. 2021). CX-SPE LODs were lower than the derivatization/GC–MS method (60–1000 nM) (Sogin et al. 2019). CX-SPE did not achieve comparable LODs for amino acids (0.025–6.8 nM) compared with focused dissolved free amino acid (DFAA) methods (0.009–0.163 nM) (Sabadel et al. 2017), but CX-SPE did still enable the quantification of many DFAAs in environmental samples. The CX-SPE LODs for DFAAs and other compounds can potentially be improved in the future through the use of larger sample volumes and micro or nanoflow HPLC systems.

Different dissolved metabolomics approaches provide different analytical windows, enabling measurements of unique sets of compounds. For example, of the 69 compounds successfully extracted and measured by CX-SPE, 15 (22%) do not contain primary amine, secondary amine, or alcohol functional groups and therefore would not be derivatized by the BC reagent used in the BC derivatization/PPL method (Widner et al. 2021). Compounds that would not be derivatized using the BC approach include many of the most abundant metabolites measured such as glycine betaine and TMAO (Fig. 5). CX-SPE was developed in part to focus on positively charged and polar, zwitterionic metabolites and successfully extracted most positively charged compounds from our standards library (betaines, sulfonium ions). However, some positively charged compounds were not extracted reproducibly with high confidence because other compounds decomposed into them during analysis (carnitine) or poor performance on the instrument. While a good number of zwitterionic compounds were successfully extracted (amino acids), some were not, either through EEs below threshold values or poor reproducibility (such as the amino acid glycine). CX-SPE also failed to extract a large number of other compounds within our in-house standards library including sugars such as trehalose and sucrose, some B vitamins, compounds with negatively charged functional groups such as DHPS. These molecules often do not develop strong positive charges in acidic solutions, preventing their extraction via cation exchange; have highly variable RFs, preventing confident quantification; or have very low ionization efficiencies, preventing their detection via LC–MS at environmentally relevant concentrations. Derivatization type approaches are likely the best option for studies focusing on these compounds (Sogin et al. 2019; Widner et al. 2021).

The concentrations of many of the compounds measured in environmental samples by CX-SPE analysis are within or close to the environmental concentration ranges reported by other studies of dissolved metabolites such as amino acids (Lee and Bada 1975; Fuhrman and Ferguson 1986; Sabadel et al. 2017), TMAO (Gibb and Hatton 2004), creatine (Wawrik et al. 2017), DMSP (Kiene and Slezak 2006), and arsenobetaine (Glabonjat et al. 2018) (Table 1). Furthermore, the concentrations of some compounds measured for the first time using CX-SPE such as glycine betaine were of roughly the same order of magnitude (low nM) as those predicted by uptake kinetics studies (Kiene and Slezak 2006; Boysen et al. 2022; Mausz et al. 2022). The overall composition of the compounds in the environmental samples fit our expectations based on observed particulate metabolite pools, with the most abundant particulate metabolites also contributing significantly to the quantifiable dissolved metabolite pool. We cannot exclude the possibility that the dissolved concentrations of compounds measured are impacted by cell breakage during filtration, resulting in metabolites moving from the particulate into the dissolved phase, impacting our dissolved-phase measurements and contributing to these similarities. We were surprised by the high concentration of many metabolites at station ALOHA relative to other stations. We predicted the concentrations of dissolved metabolites would be the lowest in this oligotrophic environment. The RFs and EEs of CX-SPE are robust to changes in environmental parameters, suggesting that these values are accurate. However, it is important to note that each sample presents a single snapshot of a dissolved metabolome that is likely constantly and rapidly remodeled by a range of dynamic uptake and release processes such as viral lysis and sloppy feeding by grazers (Connell et al. 2020; Bandara et al. 2021; Mruwat et al. 2021). More samples from a variety of locations and timepoints will be required to accurately constrain the geographic, temporal, and environmental variability of dissolved metabolomes.

The results from the analysis of these four environmental samples underscores the importance of CX-SPE as a tool for enhancing our understanding of microbial communities and the cycling of DOM in marine and aquatic environments. For example, glycine betaine and TMAO are key molecules in marine ecosystems that serve as osmolytes for marine microbes and animals (Welsh 2000; Yancey et al. 2014), fixed C and N sources to marine microbes (Welsh 2000; Lidbury et al. 2014; Boysen et al. 2022), or precursors of methane and atmospheric aerosols (Lidbury et al. 2015; Jones et al. 2019). Until now, studies of these molecules have been restricted to the particulate phase (Heal et al. 2021), molecular biology techniques (Ngugi et al. 2020), indirect approximations through uptake kinetics studies (Boysen et al. 2022; Mausz et al. 2022), or a single, very targeted set of measurements (Gibb and Hatton 2004). CX-SPE enables some of the first measurements of these molecules, paving the way for future studies in understanding the distribution and cycling of these

compounds. Finally, CX-SPE can be paired with untargeted metabolomics approaches to discover previously unidentified dissolved metabolites that may play key roles in shaping microbial systems.

Conclusion

The CX-SPE method presented and evaluated in this study is a valuable addition to the suite of sample preparation approaches aimed at the challenging problem of measuring polar metabolites that are dissolved in complex matrices. Paired with LC–MS, this extraction technique can be utilized in a wide range of marine and freshwater environments to improve our understanding of microbial community interactions and labile DOM cycling. We compare CX-SPE to the popular PPL-SPE and demonstrate that the two approaches capture fundamentally different components of DOM, with CX-SPE excelling at extracting small, positively charged and zwitterionic molecules that often contain nitrogen or sulfur. Finally, we present the first CX-SPE dissolved environmental metabolomes, including some of the first measurements of several compounds with documented importance in aquatic microbial systems such as homarine, glycine betaine, and TMAO, providing baseline measurements to facilitate future research.

Data availability statement

Metabolomics data are available through Metabolomics Workbench (https://www.metabolomicsworkbench.org) under project PR0014. All data processing and analysis code is publicly available at https://github.com/jssacks/CX_SPE_Method.

References

Adams, K. J., and others. 2020. Skyline for small molecules: A unifying software package for quantitative metabolomics. J. Proteome Res. **19**: 1447–1458. doi:10.1021/acs.jproteome.9b00640

Amin, S. A., and others. 2015. Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. Nature **522**: 98–101. doi:10.1038/nature14488

Bandara, K., Ø. Varpe, L. Wijewardene, V. Tverberg, and K. Eiane. 2021. Two hundred years of zooplankton vertical migration research. Biol. Rev. **96**: 1547–1589. doi:10.1111/brv.12715

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Ser. B Methodol. **57**: 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Boysen, A. K., and others. 2021. Particulate metabolites and transcripts reflect diel oscillations of microbial activity in the surface ocean. mSystems **6**: e00896-20. doi:10.1128/mSystems.00896-20

Boysen, A. K., and others. 2022. Glycine betaine uptake and metabolism in marine microbial communities. Environ. Microbiol. **24**: 2380–2403. doi:10.1111/1462-2920.16020

Boysen, A. K., K. R. Heal, L. T. Carlson, and A. E. Ingalls. 2018. Best-matched internal standard normalization in liquid chromatography–mass spectrometry metabolomics applied to environmental samples. Anal. Chem. **90**: 1363–1369. doi:10.1021/acs.analchem.7b04400

Bundy, J. G., M. P. Davey, and M. R. Viant. 2008. Environmental metabolomics: A critical review and future perspectives. Metabolomics **5**: 3–21. doi:10.1007/s11306-008-0152-0

Chambers, M. C., and others. 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nat. Biotechnol. **30**: 918–920. doi:10.1038/nbt.2377

Connell, P., F. Ribalet, E. Armbrust, A. White, and D. Caron. 2020. Diel oscillations in the feeding activity of heterotrophic and mixotrophic nanoplankton in the North Pacific Subtropical Gyre. Aquat. Microb. Ecol. **85**: 167–181. doi:10.3354/ame01950

Currie, L. A. 1968. Limits for qualitative detection and quantitative determination. Application to radiochemistry. Anal. Chem. **40**: 586–593. doi:10.1021/ac60259a007

Dawson, H. M., K. R. Heal, A. K. Boysen, L. T. Carlson, A. E. Ingalls, and J. N. Young. 2020*a*. Potential of temperature- and salinity-driven shifts in diatom compatible solute concentrations to impact biogeochemical cycling within sea ice. Element. Sci. Anthropocene **8**: 25. doi:10.1525/elementa.421

Dawson, H. M., K. R. Heal, A. Torstensson, L. T. Carlson, A. E. Ingalls, and J. N. Young. 2020*b*. Large diversity in nitrogen- and sulfur-containing compatible solute profiles in polar and temperate diatoms. Integr. Comp. Biol. **60**: 1401–1413. doi:10.1093/icb/icaa133

Dittmar, T., B. Koch, N. Hertkorn, and G. Kattner. 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol. Oceanogr. Methods **6**: 230–235. doi:10.4319/lom.2008.6.230

Djoumbou Feunang, Y., and others. 2016. ClassyFire: Automated chemical classification with a comprehensive, computable taxonomy. J. Chem. **8**: 61. doi:10.1186/s13321-016-0174-y

Emerson, S., and J. Hedges. 2008. Chemical oceanography and the marine carbon cycle. Cambridge Univ. Press.

Ferrer-González, F. X., B. Widner, N. R. Holderman, J. Glushka, A. S. Edison, E. B. Kujawinski, and M. A. Moran. 2021. Resource partitioning of phytoplankton metabolites that support bacterial heterotrophy. ISME J. **15**: 762–773. doi:10.1038/s41396-020-00811-y

Fiore, C. L., K. Longnecker, M. C. Kido Soule, and E. B. Kujawinski. 2015. Release of ecologically relevant metabolites by the cyanobacterium *Synechococcus elongatus* CCMP 1631. Environ. Microbiol. **17**: 3949–3963. doi:10.1111/1462-2920.12899

Fontanals, N., N. Miralles, N. Abdullah, A. Davies, N. Gilart, and P. A. G. Cormack. 2014. Evaluation of strong cation-exchange polymers for the determination of drugs by solid-phase extraction–liquid chromatography–tandem mass spectrometry. J. Chromatogr. A **1343**: 55–62. doi:10.1016/j.chroma.2014.03.068

Fuhrman, J., and R. Ferguson. 1986. Nanomolar concentrations and rapid turnover of dissolved free amino acids in seawater: Agreement between chemical and microbiological measurements. Mar. Ecol. Prog. Ser. **33**: 237–242. doi:10.3354/MEPS033237

Gebser, B., and G. Pohnert. 2013. Synchronized regulation of different zwitterionic metabolites in the osmoadaption of phytoplankton. Mar. Drugs **11**: 2168–2182. doi:10.3390/md11062168

Gibb, S. W., and A. D. Hatton. 2004. The occurrence and distribution of trimethylamine-N-oxide in Antarctic coastal waters. Mar. Chem. **91**: 65–75. doi:10.1016/j.marchem.2004.04.005

Glabonjat, R. A., G. Raber, B. A. S. Van Mooy, and K. A. Francesconi. 2018. Arsenobetaine in seawater: Depth profiles from selected sites in the North Atlantic. Environ. Sci. Technol. **52**: 522–530. doi:10.1021/acs.est.7b03939

Halewood, E., K. Opalk, L. Custals, M. Carey, D. Hansell, and C. A. Carlson. 2022. GO-SHIP repeat hydrography: Determination of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) in seawater using high temperature combustion analysis. Repository of Ocean Best Practices **2022**: 32. doi:10.25607/OBP-1745

Heal, K. R., N. A. Kellogg, L. T. Carlson, R. M. Lionheart, and A. E. Ingalls. 2019. Metabolic consequences of cobalamin scarcity in the diatom *Thalassiosira pseudonana* as revealed through metabolomics. Protist **170**: 328–348. doi:10.1016/j.protis.2019.05.004

Heal, K. R., and others. 2021. Marine community metabolomes carry fingerprints of phytoplankton community composition. mSystems **6**: e01334-20. doi:10.1128/mSystems.01334-20

Hmelo, L. R. 2017. Quorum sensing in marine microbial environments. Ann. Rev. Mar. Sci. **9**: 257–281. doi:10.1146/annurev-marine-010816-060656

Johnson, W. M., M. C. Kido Soule, and E. B. Kujawinski. 2016. Evidence for quorum sensing and differential metabolite production by a marine bacterium in response to DMSP. ISME J. **10**: 2304–2316. doi:10.1038/ismej.2016.6

Johnson, W. M., M. C. K. Soule, and E. B. Kujawinski. 2017. Extraction efficiency and quantification of dissolved metabolites in targeted marine metabolomics. Limnol. Oceanogr. Methods **15**: 417–428. doi:10.1002/lom3.10181

Johnson, W. M., K. Longnecker, M. C. K. Soule, W. A. Arnold, M. P. Bhatia, S. J. Hallam, B. A. S. V. Mooy, and E. B. Kujawinski. 2020. Metabolite composition of sinking particles differs from surface suspended particles across a latitudinal transect in the South Atlantic. Limnol. Oceanogr. **65**: 111–127. doi:10.1002/lno.11255

Jones, H. J., E. Kröber, J. Stephenson, M. A. Mausz, E. Jameson, A. Millard, K. J. Purdy, and Y. Chen. 2019. A new family of uncultivated bacteria involved in methanogenesis from the ubiquitous osmolyte glycine betaine in coastal saltmarsh sediments. Microbiome **7**: 120. doi:10.1186/s40168-019-0732-4

Kiene, R. P., and D. Slezak. 2006. Low dissolved DMSP concentrations in seawater revealed by small-volume gravity filtration and dialysis sampling. Limnol. Oceanogr. Methods **4**: 80–95. doi:10.4319/lom.2006.4.80

Lee, C., and J. L. Bada. 1975. Amino acids in equatorial Pacific Ocean water. Earth Planet. Sci. Lett. **26**: 61–68. doi:10.1016/0012-821X(75)90177-6

Lidbury, I., J. C. Murrell, and Y. Chen. 2014. Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. PNAS **111**: 2710–2715. doi:10.1073/pnas.1317834111

Lidbury, I. D., J. C. Murrell, and Y. Chen. 2015. Trimethylamine and trimethylamine N-oxide are supplementary energy sources for a marine heterotrophic bacterium: Implications for marine carbon and nitrogen cycling. ISME J. **9**: 760–769. doi:10.1038/ismej.2014.149

Logan, B. K., D. T. Stafford, I. R. Tebbett, and C. M. Moore. 1990. Rapid screening for 100 basic drugs and metabolites in urine using cation exchange solid-phase extraction and high-performance liquid chromatography with diode array detection. J. Anal. Toxicol. **14**: 154–159. doi:10.1093/jat/14.3.154

Mausz, M. A., and others. 2022. Microbial uptake dynamics of choline and glycine betaine in coastal seawater. Limnol. Oceanogr. **67**: 1052–1064. doi:10.1002/lno.12056

Mayali, X., and P. K. Weber. 2018. Quantitative isotope incorporation reveals substrate partitioning in a coastal microbial community. FEMS Microbiol. Ecol. **94**: fiy047. doi:10.1093/femsec/fiy047

Moran, M. A., and others. 2016. Deciphering ocean carbon in a changing world. PNAS **113**: 3143–3151. doi:10.1073/pnas.1514645113

Moran, M. A., and others. 2022. Microbial metabolites in the marine carbon cycle. Nat. Microbiol. **7**: 508–523. doi:10.1038/s41564-022-01090-3

Mruwat, N., and others. 2021. A single-cell polony method reveals low levels of infected *Prochlorococcus* in oligotrophic waters despite high cyanophage abundances. ISME J. **15**: 41–54. doi:10.1038/s41396-020-00752-6

Ngugi, D. K., M. Ziegler, C. M. Duarte, and C. R. Voolstra. 2020. Genomic blueprint of glycine betaine metabolism in coral metaorganisms and their contribution to reef nitrogen budgets. iScience **23**: 101120. doi:10.1016/j.isci.2020.101120

Petras, D., and others. 2017. High-resolution liquid chromatography tandem mass spectrometry enables large scale

molecular characterization of dissolved organic matter. Front. Mar. Sci. **4**: 4.

Pontrelli, S., and U. Sauer. 2021. Salt-tolerant metabolomics for exometabolomic measurements of marine bacterial isolates. Anal. Chem. **93**: 7164–7171. doi:10.1021/acs.analchem.0c04795

Poretsky, R. S., S. Sun, X. Mou, and M. A. Moran. 2010. Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. Environ. Microbiol. **12**: 616–627. doi:10.1111/j.1462-2920.2009.02102.x

Sabadel, A. J. M., T. J. Browning, D. Kruimer, R. L. Airs, E. M. S. Woodward, R. Van Hale, and R. D. Frew. 2017. Determination of picomolar dissolved free amino acids along a South Atlantic transect using reversed-phase high-performance liquid chromatography. Mar. Chem. **196**: 173–180. doi:10.1016/j.marchem.2017.09.008

Sañudo-Wilhelmy, S. A., L. Gómez-Consarnau, C. Suffridge, and E. A. Webb. 2014. The role of B vitamins in marine biogeochemistry. Ann. Rev. Mar. Sci. **6**: 339–367. doi:10.1146/annurev-marine-120710-100912

Shibl, A. A., and others. 2020. Diatom modulation of select bacteria through use of two unique secondary metabolites. PNAS **117**: 27445–27455. doi:10.1073/pnas.2012088117

Sogin, E. M., E. Puskás, N. Dubilier, and M. Liebeke. 2019. Marine metabolomics: A method for nontargeted measurement of metabolites in seawater by gas chromatography–mass spectrometry. mSystems **4**: e00638–e00619. doi:10.1128/mSystems.00638-19

Tsugawa, H., and others. 2015. MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat. Methods **12**: 523–526. doi:10.1038/nmeth.3393

Vorobev, A., and others. 2018. Identifying labile DOM components in a coastal ocean through depleted bacterial transcripts and chemical signals. Environ. Microbiol. **20**: 3012–3030. doi:10.1111/1462-2920.14344

Wawrik, B., D. Bronk, S. Baer, L. Chi, M. Sun, J. Cooper, and Z. Yang. 2017. Bacterial utilization of creatine in seawater. Aquat. Microb. Ecol. **80**: 153–165. doi:10.3354/ame01850

Weber, L., M. Armenteros, M. Kido Soule, K. Longnecker, E. B. Kujawinski, and A. Apprill. 2020. Extracellular reef metabolites across the protected Jardines de la Reina, Cuba Reef System. Front. Mar. Sci. **7**: 582161. doi:10.3389/fmars.2020.582161

Welsh, D. T. 2000. Ecological significance of compatible solute accumulation by micro-organisms: From single cells to global climate. FEMS Microbiol. Rev. **24**: 263–290. doi:10.1111/j.1574-6976.2000.tb00542.x

Widner, B., M. C. Kido Soule, F. X. Ferrer-González, M. A. Moran, and E. B. Kujawinski. 2021. Quantification of amine- and alcohol-containing metabolites in saline samples using pre-extraction benzoyl chloride derivatization and ultrahigh performance liquid chromatography tandem mass spectrometry (UHPLC MS/MS). Anal. Chem. **93**: 4809–4817. doi:10.1021/acs.analchem.0c03769

Xu, C., and others. 2021. MetFish: A metabolomics pipeline for studying microbial communities in chemically extreme environments. mSystems **6**: e01058–e01020. doi:10.1128/mSystems.01058-20

Yancey, P. H., M. E. Gerringer, J. C. Drazen, A. A. Rowden, and A. Jamieson. 2014. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. PNAS **111**: 4461–4465. doi:10.1073/pnas.1322003111

Yoch, D. C. 2002. Dimethylsulfoniopropionate: Its sources, role in the marine food web, and biological degradation to dimethylsulfide. Appl. Environ. Microbiol. **68**: 5804–5815. doi:10.1128/AEM.68.12.5804-5815.2002

Acknowledgments

The authors thank Alec Meyers, Everetta Rasyid, and Leland Wood for assistance with laboratory analyses and preliminary data processing; William Kumler, Natalie Kellogg, and Regina Lionheart for feedback and assistance with data analysis and interpretation; and the science party and crew of KOK1606 and the R.V. Rachel Carson for support during sample collection. The authors thank G. Pohnert, C. Harwood, A. Bourdon, S. Champagna, A. Cook, and K. Denger for generously providing metabolite standards. This work was supported by grants from the Simons Foundation (LS award ID: 385428, A.E.I.; SCOPE award ID 329108, A.E.I.; award ID 598819, K.R.H.; award ID 731716, A.K.B.) and the National Science Foundation (OCE-2125886 to A.E.I.).

Submitted 13 May 2022
Revised 02 August 2022
Accepted 17 August 2022

Associate editor: Krista Longnecker