Detection and Structural Elucidation of Copper Binding Tri- and Tetrapyrrole Ligands Produced by the Marine Diatom *Phaeodactylum Tricornutum*

Lydia Babcock-Adams[a,b], Jingxuan Li[b], Amy M. McKenna[a,c], Christopher L. Hendrickson[a], and Daniel J. Repeta[b*]

[a] Ion Cyclotron Resonance Program, National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310, United States
[b] Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, United States
[c] Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, 80521, United States

*Corresponding Author, drepeta@whoi.edu

Running Title: Copper Binding Ligands from *P. tricornutum*

Address reprint requests to Daniel Repeta, 360 Woods Hole Rd. MS #51 Watson Laboratory, Woods Hole, MA, 02540, (508) 289-2635, drepeta@whoi.edu

Abstract

In seawater most dissolved copper (Cu) is complexed by organic ligands, many of which are thought to be produced by phytoplankton. Although very little is known about the composition and structure of these ligands, they play an important role in determining the reactivity and bioavailability of Cu. In this study, *Phaeodactylum tricornutum*, a marine diatom known to produce Cu ligands (CuLs), was grown in laboratory pure culture and the CuLs recovered from the growth media. Using liquid chromatography coupled to ultrahigh resolution tandem mass spectrometry, eleven Cu ligand complexes were identified and assigned molecular formulae. Molecular formulae were confirmed by comparing the expected and observed relative abundances of ^{15}N, ^{13}C, ^{65}Cu, and ^{18}O isotopologues. The CuLs had molecular weights from 520 to 719 Da and molecular formulae of $C_{26\text{-}35}H_{23\text{-}36}O_{5\text{-}9}N_{3\text{-}4}Cu$ with an average assignment error of 56 ppb. High-resolution tandem mass spectrometry of the Cu-bound and metal-free ligands revealed these to be a suite of tri- and tetrapyrroles stabilized through complexation of Cu by N. The ligands share similar parent structures but differ in the number, type, and arrangement of functional groups that decorate the pyrroles. The similarity of CuL structures with known catabolites of chlorophyll suggests these ligands may be widely produced by marine photoautotrophs.

Introduction

Copper (Cu) is a redox active metal, which makes it suitable for various essential cellular processes such electron transport in photosynthesis and respiration,[1,2] the breakdown of reactive oxygen species,[3,4] and uptake of iron.[5] However, the same redox properties that make Cu beneficial can be toxic when concentrations exceed what can be properly handled by the cell, resulting in decreased cell growth and ultimately leading to cell death.[6–9] The most bioavailable and toxic form of Cu is the cupric ion (Cu^{2+}).[5,6,10] The toxicity threshold for especially sensitive microbes, such as cyanobacteria,[6,11] can be as low as 10^{-13} M, which is within the range of environmental Cu^{2+} concentrations in seawater (10^{-15} to 10^{-12} M).[12,13] One strategy microbes employ to mitigate the toxicity of Cu^{2+} is to produce organic ligands that complex Cu to make it less (or not at all) bioavailable.[14,15] For example, phytochelatins (oligomers of glutathione) and metallothioneins (a family of low molecular weight proteins) form strong complexes with Cu^{2+} and have been shown to be produced in response to Cu toxicity.[16–18]

In seawater, dissolved Cu is almost entirely complexed by organic ligands (CuL).[19] Remobilization of the ligand-bound Cu is governed by the structure of the ligand which plays a role in determining the strength of complexation, influencing its photo-reactivity and ultimately the bioavailability of Cu.[20] The detection and identification of CuLs in seawater has been plagued by analytical challenges common to marine samples – high concentrations of salts, low concentrations of individual CuLs, and an ultracomplex background organic matrix. A few studies have characterized CuLs in seawater using Fourier Transform Ion Cyclotron Resonance mass spectrometry (FT-ICR MS).[21,22] These studies, which distinguish CuLs from other organic compounds by the mass difference and relative abundance of ^{63}Cu and ^{65}Cu ions, found more than 500 putative Cu-containing molecular ions. However, molecular formulae could only be assigned to 66 of these ions. This study used several conservative knock-out criteria to determine the molecular formula in the event of unambiguous assignment, leading to the low percent of putative Cu-containing ions with a formula assignment. A contributing factor is the allowed assignment error of ± 0.5 ppm – with higher resolution instruments this error range can

be decreased, resulting in a higher percentage of unambiguous formula assignments. The Cu-containing ions ranged in *m/z* from 237 to 689 and all contained 1-5 N atoms. In addition to N, 12, 17, and 34 of the molecular ions also contained O (without S), S (without O), or both O and S, respectively. To date, only one study has provided deeper insight to the structural characteristics of a CuL isolated from seawater. Using a combination of FT-ICR MS and inductively coupled plasma mass spectrometry (ICP MS) coupled to liquid chromatography (LC), an abundant CuL with the molecular formula $[C_{20}H_{21}N_4O_8S_2Cu]^+$ was isolated from seawater collected in the Eastern Tropical Pacific Ocean.[23] Based on high resolution fragmentation data, the authors concluded that the ligand had heavily conjugated, cyclic, azole-like functional groups, that the sulfur was present as thiols, and that the Cu was complexed by N. The O:C (0.4) and N:C (0.2) ratios of this CuL are consistent with a peptide-like compound,[24] with a high degree of unsaturation.

Diatoms, a globally important group of eukaryotic phytoplankton, have many metabolic uses for Cu, including in iron acquisition[5,25,26] and electron transport.[27] Diatoms are known to produce ligands that complex copper,[28] and recently it has been shown that the diatoms *Thalassiosira oceanica* and *Phaeodactylum tricornutum* can acquire Cu complexed to ethylenediaminetetraacetic acid (EDTA) through a reductive pathway,[29,30] which suggests that at least some portion of organically bound Cu in seawater could be bioavailable to diatoms. Although diatoms have a metabolic requirement for Cu, they are still susceptible to cellular damage and death due to high concentrations of Cu, and some diatoms have been shown to respond to high Cu through the intracellular production of phytochelatins.[31,32] Here we report the structural characterization of CuLs produced by the widely studied marine diatom *P. tricornutum* using the custom-built hybrid linear ion trap 21 tesla FT-ICR MS at the National High Magnetic Field Laboratory.[33] This instrument achieves the high mass resolving power, acquisition speed, dynamic range, and mass accuracy needed for analyzing complex organic matricies,[34,35] and is uniquely suited to provide structural characterization of marine CuLs.

Methods

Materials and Reagents.

Polycarbonate culture bottles, PTFE tubing, and silicone tubing were soaked in 1% detergent (Citranox) for one week, rinsed with deionized water, soaked in 10% HCl (Trace Metal Grade, Fisher Scientific) for another week, then rinsed with ultrapure water (qH_2O, 18 MΩ cm^{-1}) before use. Solid phase extraction columns (1 g, 6 mL Bond Elut ENV, Agilent) were activated by passing 6 mL of methanol (MeOH) through the resin, rinsed with 6 mL of pH 2 qH_2O followed by 6 mL of qH_2O. A 1 µM aqueous cyanocobalamin (Sigma Aldrich) solution was used as an internal standard. For analyses coupling liquid chromatography (LC) with inductively coupled plasma MS (ICP MS), LC-MS grade methanol, ammonium formate (Optima, Fisher Scientific), and qH_2O were used. The methanol was redistilled using a polytetrafluoroethlene (PTFE) still to reduce metal contaminants. For LC-FT-ICR MS analyses, LC-MS grade water, methanol, and ammonium formate (Honeywell) were used.

Phaeodactylum tricornutum CCMP 632 Culture.

An axenic culture of *P. tricornutum* CCMP 632 (National Center for Marine Algae and Microbiota, Bigelow Laboratory, Boothbay, Maine) was grown at 18 °C in polycarbonate bottles under continuous illumination (90 µmol m^{-2} s^{-1}). The culture medium was prepared using 0.2 µm

filtered coastal seawater that was autoclaved to sterilize then amended with macro- and micro-nutrients as detailed in the supplementary information (**SI Table 1**). All nutrients were sterile filtered (0.2 μm PES) prior to addition to sterilized seawater. Cultures were maintained by sterile technique and the absence of bacteria was confirmed via DAPI stained samples. Relative chlorophyll fluorescence was monitored (Turner TD-700) as a proxy for growth (**SI Fig. 1**). Cultures were harvested during the exponential growth phase.

Characterization of culture extracts by LC-ICPMS.

Filtered medium (0.22 μm; Sterivex, Millipore) was pumped through an activated 1 g Bond Elut ENV SPE column at a flow rate of 10 mL min^{-1}. A medium blank was prepared with the culture medium without inoculation, incubated alongside the culture, and processed in the same way. SPE columns were rinsed with 6 mL of qH_2O to remove salts, and CuLs eluted with 12 mL of MeOH. The MeOH eluents were concentrated to approximately 1 mL using a SpeedVac concentrator (Thermo Scientific) at 35 °C for 3 hours. MeOH extracts from five, 1 L cultures were combined and stored in 100% MeOH at -20 °C until analysis.

CuLs were separated on a Zorbax SB-C_{18} column (0.5 *x* 150 mm, 5 μm particles), at a flow rate of 40 μL min^{-1} and a 30-minute gradient from 95/5% to 5/95% solvent A/B followed by a 5-minute isocratic hold at 5/95% solvent A/B (solvent A = 5 mM ammonium formate in qH_2O, solvent B = 5 mM ammonium formate in MeOH). The column flow was directed into an iCAP Q MS (Thermo Scientific) fitted with a perfluoroalkoxy micronebulizer (PFA LC-2040, Elemental Scientific) and a cyclonic spray chamber cooled to 4 °C. Oxygen was introduced at 25 mL min^{-1} to minimize formation of reduced carbon deposits on the skimmer and sampler cones. Data was collected in kinetic energy discrimination mode with helium as the collision gas. The ^{63}Cu and ^{65}Cu ions were monitored with integration times of 0.1 and 0.05 s respectively. Analysis of the culture and medium blank extracts showed that all CuLs were produced by *P. tricornutum* as there were no detectable CuLs in the medium blank (**Fig. 1**). Additionally, there were no coeluting features in the chromatograms of other bioactive metals such as iron and nickel.

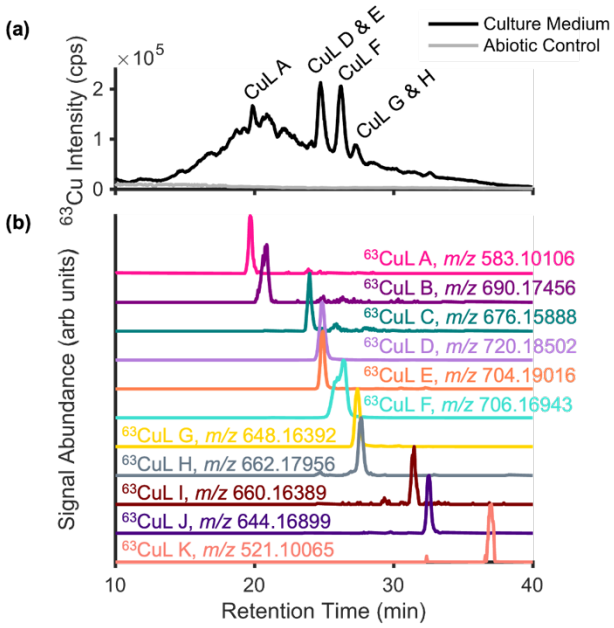


Figure 1. (a) LC-ICP MS ^{63}Cu chromatogram for the *P. tricornutum* culture medium (black) and abiotic control (gray). The abiotic control was incubated alongside the culture with no inoculum to ensure that CuLs were produced by the diatom. There were no detectable ^{63}Cu peaks in the abiotic control, confirming that the CuLs were produced by the diatom. **(b)** LC-ESI MS chromatogram of the ^{63}CuL extracted ion chromatograms. The retention times of the four most abundant peaks in the LC-ICP MS ^{63}Cu chromatogram aligned with Cu-containing molecular ions identified by the LC-ESI MS analyses.

These chromatographic conditions were additionally coupled to an Orbitrap Fusion MS (Thermo Scientific) equipped with a heated electrospray ionization (HESI) source operated in positive mode at 3500 V. The sheath and aux gas were set to 25 and 5 arbitrary units respectively. The ion transfer tube was set to 275 °C and the vaporizer temperature was set to 75 °C. MS^1 spectra were collected with a resolving power of 500,000 at *m/z* 200. Cyanocobalamin was spiked into the sample at a concentration of 50 nM and used to align the LC-ICP and LC-Orbitrap mass spectra. Data files were converted to an mzXML format using MSconvert (Proteowizard) and processed using an algorithm to identify Cu-containing compounds (https://github.com/rboiteau/LC-ESIMS-isotope-pattern-algorithm).[36] Briefly, the algorithm searches for molecular ions occurring at the same retention time with a mass difference of 1.9982

Da, the mass difference between ^{63}Cu and ^{65}Cu. An error of ± 1.5 mDa was allowed. Manual validation of isotopologue relative abundance (RA) ratios of 2.24 confirms the presence of Cu.

Characterization of Cu Ligands by LC-FT-ICR MS.

A Dionex Ultimate 3000 (Thermo Scientific) LC was coupled to the front end of a custom-built hybrid linear ion trap FT-ICR MS equipped with a 21 T superconducting magnet.[33] CuLs were separated using a 30 min gradient from 95/5% to 5/95% solvent A/B (solvent A = 5 mM ammonium formate in water, solvent B = 5 mM ammonium formate in MeOH) on a Zorbax SB-C_{18} column (40 °C; 0.5 *x* 150 mm, 5 µm particles, Agilent) at 40 µL min^{-1}. The eluent was coupled to a HESI source operated in positive mode (4.5 kV). The inlet capillary and source heater temperatures were set to 350 °C and 75 °C, and sheath and auxiliary gas flow rates were set to 5 and 3 (arbitrary units). MS^1 spectra were collected from 150 to 1500 *m/z*. All spectra were collected with a resolution of 1,200,000 at 400 *m/z*, an automatic gain control (AGC) target of $1x10^6$ charges, and a maximum ion injection time of 1500 ms. The achieved mass spectral resolving power at *m/z* 200 was ~1,700,000-3,000,000 across the LC gradient.

CuL candidates identified by LC-Orbitrap analyses were extracted from the LC-FT-ICR MS data to obtain ultrahigh resolution *m/z* to allow for the unambiguous assignment of molecular formulae. Metal-free ligands were found by searching for the presence of a molecular ion at a *m/z* occurring at 61.92177 Da or 60.91395 Da lower than the *m/z* of CuL complex, which correspond to the Cu(I) or Cu(II) oxidation states. All metal-free ligands identified were present at *m/z* 60.91395 lower than the CuL, indicating a +2 oxidation state of the Cu. Elemental compositions were assigned using the Predator Molecular Formula Calculator (v.1.3.3), with elemental constraints of $C_{\infty}H_{\infty}O_{0\text{-}15}N_{0\text{-}10}Cu_1$ ± 2.5 ppm for CuLs and $C_{\infty}H_{\infty}O_{0\text{-}15}N_{0\text{-}10}$ ± 2.5 ppm for metal-free ligand assignments. No ^{34}S isotope peak was observed for any Cu or metal-free ligands, therefore sulfur was excluded from elemental constraints.

Cu and metal-free ligands were targeted for tandem mass spectrometry (MS^2) experiments to elucidate structural characteristics. LC and HESI source settings were the same as in MS^1 analyses, except the voltage was set to 3.75 kV and MS^1 spectra were collected from 455 to 1015 *m/z* with a resolution of 600,000 at *m/z* 400. MS^2 spectra were collected in the ICR cell using an isolation width of 0.6 Da and a maximum ion injection time of 500 ms at a resolution of 600,000 at *m/z* 400. Ions were fragmented via collision induced dissociation (CID) with a normalized collisional energy of 40, activation Q of 0.25, and an activation time of 10 ms. Select fragment ions were targeted for MS^3 experiments in the linear ion trap using the same settings as for the MS^2 experiments.

Results and Discussion

Characterization of Cu Ligands by LC-ICP-ESI MS

Copper ligands (CuLs) isolated from spent *P. tricornutum* medium include an unresolved complex mixture of CuLs that appear as a gradual rise and fall in the baseline of the LC-ICP MS ^{63}Cu chromatogram topped with a number of defined peaks (**Fig. 1a**). To identify molecular ions corresponding to CuL complexes, samples were analyzed by LC- Orbitrap using the same chromatographic conditions. CuLs were identified based on the mass difference and RA of ^{63}Cu and ^{65}Cu peaks and alignment with the LC-ICP MS chromatogram. Eleven putative CuLs (**A-K**) were identified, with *m/z* ratios between 521 and 720 that elute between 19 and 38 minutes (**Fig.**

1b). In order to confirm the presence of Cu and assign a molecular formula to these putative CuLs, higher resolution *m/z* were required.

Isotopic Fine Structure Confirms Elemental Composition of Formula Assignments

The high mass accuracy achievable by FT-ICR MS at 21T allows for the unambiguous assignment of molecular formulae to molecular ions. The presence of Cu was confirmed in all eleven CuLs, with an experimental mass difference between ^{63}Cu and ^{65}Cu of 1.99821 ± 0.00001 Da ($\Delta m_{theoretical}$, 1.99819 Da) and a ^{65}Cu to ^{63}Cu ratio of 2.3 ± 0.4 (**SI Fig. 2**). Molecular formulae were assigned to all CuLs with absolute errors between 6 and 100 ppb (**Table 1**).

Table 1. CuLs retention times (detected by LC-Orbitrap) and corresponding ultrahigh resolution m/z (LC-FT-ICR MS at 21T) with assigned molecular formula.

Ligand Identifier	Retention Time (min)	Molecular Formula	Theoretical *m/z*	Measured *m/z*	Error (ppb)
A	19.68	$[C_{27}H_{25}O_8N_3Cu + H]^+$	583.101013	583.10107	100
B	20.87	$[C_{34}H_{34}O_8N_4Cu + H]^+$	690.174561	690.17458	30
C	23.95	$[C_{33}H_{32}O_8N_4Cu + H]^+$	676.158875	676.15888	8
D	24.86	$[C_{35}H_{36}O_9N_4Cu + H]^+$	720.185059	720.18502	-50
E	24.90	$[C_{35}H_{36}O_8N_4Cu + H]^+$	704.190186	704.19016	-40
F	26.40	$[C_{34}H_{34}O_9N_4Cu + H]^+$	706.169434	706.16942	-20
G	27.39	$[C_{32}H_{32}O_7N_4Cu + H]^+$	648.163940	648.16389	-80
H	27.65	$[C_{33}H_{34}O_7N_4Cu + H]^+$	662.179565	662.17947	100
I	31.47	$[C_{33}H_{32}O_7N_4Cu + H]^+$	660.163940	660.16389	-80
J	32.53	$[C_{33}H_{32}O_6N_4Cu + H]^+$	644.169067	644.16899	-100
K	37.02	$[C_{26}H_{23}O_5N_3Cu + H]^+$	521.100647	521.10065	6

All molecular formula assignments were confirmed by their isotope fine structure – the presence of heavy isotope peaks (e.g. ^{15}N, ^{13}C, ^{18}O, ^{65}Cu) at a RA predicted from the number of each element in the formula obtained using LC-FT-ICR MS at 21T. The theoretical isotope fine structure for each formula assignment was simulated using IsoPro 3.1 at the achieved resolving power for that *m/z* and compared to the measured isotope fine structure (**Fig. 2**).

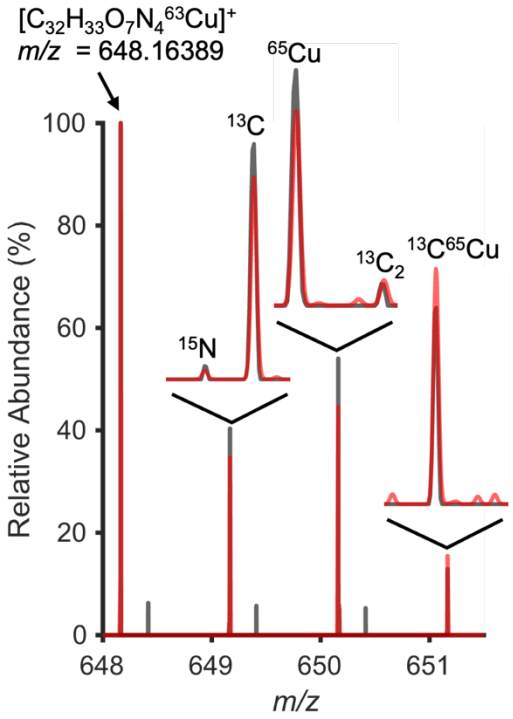


Figure 2. Mass spectrum for Cu ligand **G** (*m/z* 648) shows the theoretical (red) overlaid atop the observed (gray) isotope fine structure acquired by FT-ICR MS at 21 T with detected heavy isotope peaks labeled.

To assess how well the measured values matched the theoretical values, the difference between the measured and theoretical isotope split (mass difference between the heavy isotope and the monoisotopic peak) and RA for each isotope peak was calculated. This analysis showed that the isotope split was accurate to ± 0.1 mDa for all heavy isotope peaks and the RA was accurate to ± 5% for heavy isotope peaks with S/N > 30 (**SI Fig. 2**). Heavy isotope peaks with S/N < 30 showed larger deviations from the theoretical RA, from -21.8 to 9.6%. The CuLs in this study all had 26-35 carbon atoms, 5-9 oxygen atoms, and 3-4 nitrogen atoms. The corresponding metal-free ligands had H:C ratios from 0.96 to 1.09, O:C ratios from 0.18 to 0.30, and N:C ratios from 0.11 to 0.13 (**Table 2**). The ligands produced by *P. tricornutum* fall within a narrow compositional space, based on their elemental ratios. In fact, many of the ligands are related by systematic differences, such as a CH_2, O, and H_2.

Table 2. Metal-free ligand ultrahigh resolution m/z detected by LC-FT-ICR MS at 21 T with assigned molecular formula.

Ligand Identifier	Molecular Formula	Theoretical *m/z*	Measured *m/z*	Error (ppb)
A	$[C_{27}H_{27}O_8N_3 + H]^+$	552.187073	522.18709	3
B	$[C_{34}H_{36}O_8N_4 + H]^+$	629.260559	629.26058	3
C	$[C_{33}H_{34}O_8N_4 + H]^+$	615.244941	615.24494	0.3
D	$[C_{35}H_{38}O_9N_4 + H]^+$	659.271179	659.27108	-20
E	$[C_{35}H_{38}O_8N_4 + H]^+$	643.276241	643.27622	3
F	$[C_{34}H_{36}O_9N_4 + H]^+$	645.255493	645.25549	-0.5
G	$[C_{32}H_{34}O_7N_4 + H]^+$	587.250000	587.24996	-7
H	$[C_{33}H_{36}O_7N_4 + H]^+$	601.265676	601.26564	6
I	$[C_{33}H_{34}O_7N_4 + H]^+$	599.250026	599.25000	-4
J	$[C_{33}H_{34}O_6N_4 + H]^+$	583.255111	583.25506	9
K	$[C_{26}H_{25}O_5N_3 + H]^+$	460.186697	460.18667	5

P. tricornutum CuLs as tri- and tetrapyrroles

The elemental composition of CuLs ligands ($C_{26\text{-}35}$, $H_{23\text{-}36}$, $O_{5\text{-}8}$, $N_{3\text{-}4}$) suggests similarities to breakdown products of the well-known and ubiquitous biomolecule, chlorophyll (C_{35}, H_{34}, O_5, N_4 for chlorophyllide a). Chlorophyll has a macrocyclic tetrapyrrole ring system of four substituted pyrroles (labeled A-D) connected by methine (meso-carbon) bridges (**Fig. 3**). There is also a ring (E) that is not a pyrrole positioned exocyclic to the macrocycle. Enzymatic catabolism of chlorophylls a and b begins with demetallation (loss of Mg^{2+}) and ester hydrolysis (loss of phytol), followed by oxygenolytic opening of the macrocycle between rings A and D to form a red chlorophyll catabolite (RCC; **Fig. 3**).[37] RCC has an H:C ratio of 1.09, O:C ratio of 0.2, and N:C ratio of 0.11, values that all fall within the ranges observed for the CuLs produced by *P. tricornutum*. Subsequent catabolism of RCC produces di-, tri-, and tetrapyrroles decorated with a suite of different functional groups. These catabolites have characteristic fragmentation pathways that include: (1) the loss of CH_4O from a methoxycarbonyl functional group on ring E, (2) loss of H_2O and CO, (3) fragmentation at the meso-carbon positions with loss of pyrroles and (4) decarboxylation reactions (loss of CO_2, CH_2O_2) **(Fig. 3).**[38–40]

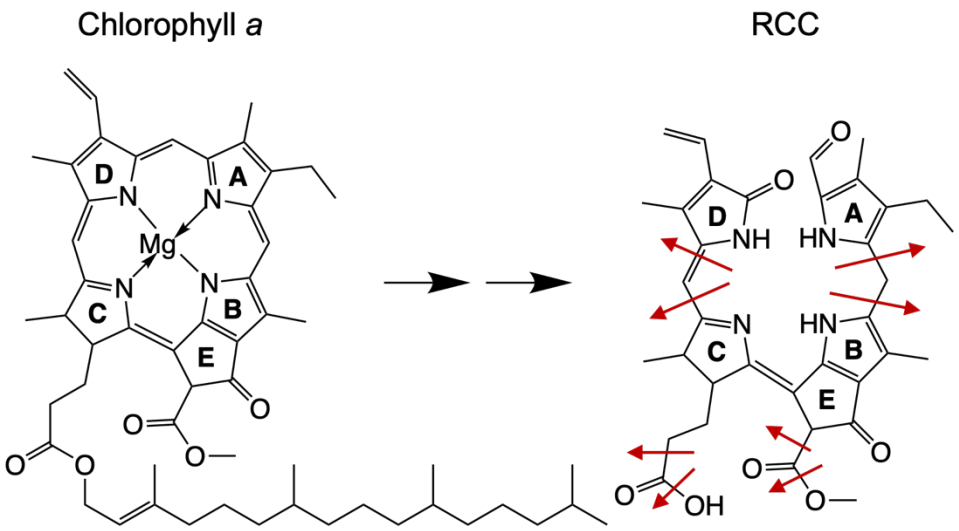


Figure 3. Chlorophyll *a* is degraded to red chlorophyll catabolite (RCC) through a series of enzymatically mediated steps.[37] Characteristic fragmentation sites for the protonated ion, $[M + H]^+$, shown for RCC.[38-40]

Due to low analyte abundance and ion suppression from a complex organic matrix (10,369 features across the gradient) that co-elute with CuLs, it was not possible to acquire high quality MS^2 spectra for CuLs **A, B,** and **J**. However, the MS^2 spectra of the 8 remaining CuLs contain major fragment ions that correspond to losses of CH_4O and/or CO_2, and minor fragment ions arising from the loss of substituted pyrrole(s). For example, the MS^2 spectrum of CuL **G**, $[C_{32}H_{32}O_7N_4Cu + H]^+$, with *m/z* of 648.16389 (**Fig. 4**) was dominated by two fragment ions at *m/z* 633.14055, $[C_{31}H_{30}O_7N_4Cu]^+$, and 576.17925, $[C_{30}H_{33}O_4N_4Cu]^+$, which represent small neutral losses $[M - CH_3]^{\bullet+}$ and $[M - (CO_2 + CO)]^+$ (**Table 3**). There are also three fragment ions of lower relative abundance (< 5%) at *m/z* 495.12138, 483.12138, and 468.09790 with molecular formulae $[C_{25}H_{26}O_4N_3{}^{63}Cu]^+$, $[C_{24}H_{26}O_4N_3{}^{63}Cu]^+$, $[C_{23}H_{23}O_4N_3Cu]^+$, representing the loss of an N-containing moiety, $C_7H_7O_3N$, $C_8H_7O_3N$, and $C_9H_{10}O_3N$, respectively. The low abundance of ions indicating loss of pyrrole(s) relative to the high abundance of ions arising from the loss of methyl, carboxyl, and carbonyl suggests the stabilization of the ligand through coordination with Cu.

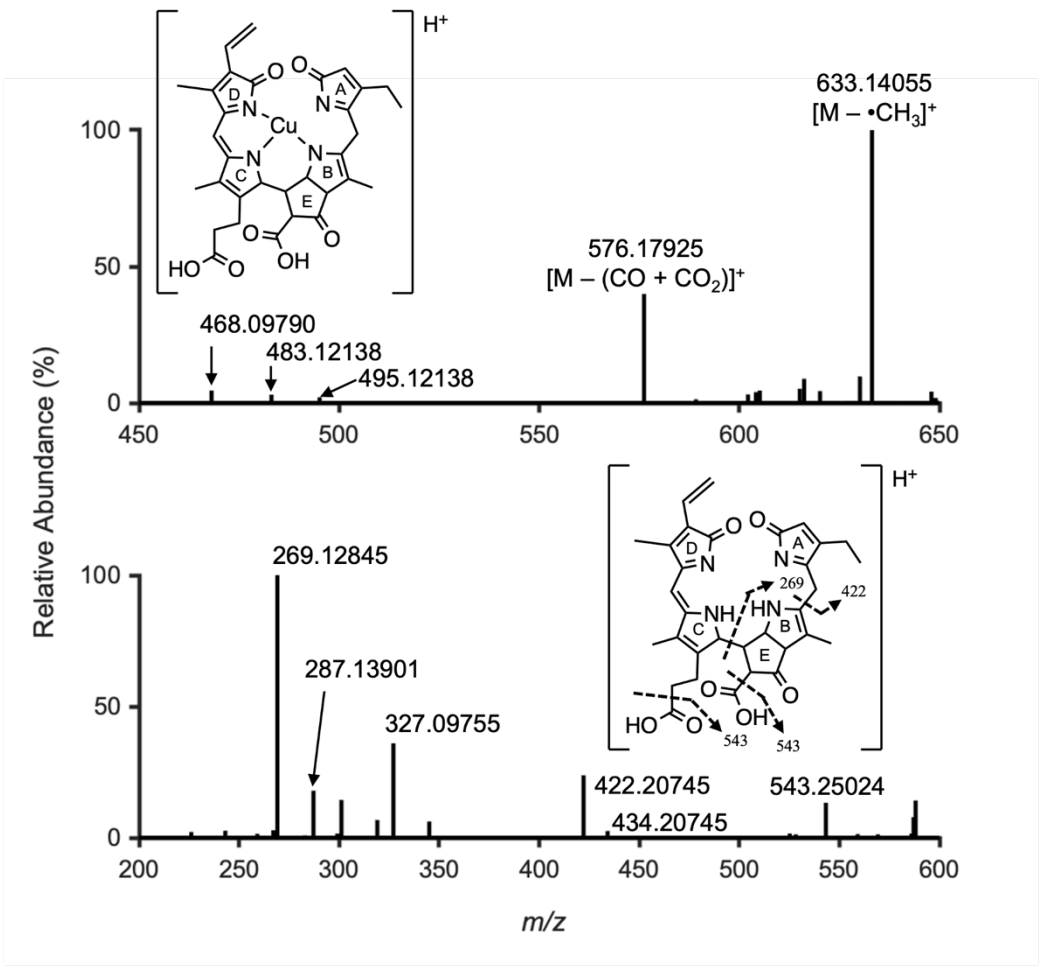


Figure 4. MS^2 spectrum of *m/z* 648.16385 and draft structure for CuL **G**. Fragments with *m/z* 495, 483, and 468 arise due to the loss of pyrrole ring(s). The same neutral losses are observed in the MS^2 spectrum of the metal-free ligand. MS^2 spectrum of *m/z* 587.24992 and draft structure (bottom). Dashed lines indicate location of fragmentations. Arrow direction indicates the neutral loss. Fragment ion structures are given in **SI Figures 7** and **8**. Formula assignments, assignment error, and neutral losses are in **Tables 3** and **4**.

Table 3. Fragment ions of CuL **G** with m/z 648.16394 and molecular formula $[C_{32}H_{32}O_7N_4Cu + H]^+$.

Fragment Ion Formula	*m/z*	Error (ppb)	Neutral Loss	Relative Abundance (%)
$[C_{31}H_{30}O_7N_4Cu]^+$	633.14055	8	CH_3•	100.00
$[C_{30}H_{33}O_4N_4Cu]^+$	576.17925	4	C_2O_3	39.96
$[C_{25}H_{26}O_4N_3Cu]^+$	495.12138	2	$C_7H_7O_3N$	2.04
$[C_{24}H_{26}O_4N_3Cu]^+$	483.12138	-0.1	$C_8H_7O_3N$	3.02
$[C_{23}H_{23}O_4N_3Cu]^+$	468.09790	1	$C_9H_{10}O_3N$	4.55

Therefore, the corresponding metal-free ligands were targeted for fragmentation. The loss of $C_7H_7O_3N$ and $C_8H_7O_3N$ is also observed in the MS^2 spectrum of the metal free ligand **G** (**Fig. 4**) as fragment ions at *m/z* 434.20745, $[C_{25}H_{28}O_4N_3]^+$ and 422.20745, $[C_{24}H_{28}O_4N_3]^+$ (**Table 4**). In contrast to the MS^2 spectrum of the CuL **G**, the MS^2 spectrum for the metal-free ligand was dominated by fragment ions arising from the loss of N-containing moieties, supporting the hypothesis that coordination with Cu stabilizes the structure. The N:C ratio of the fragment ions (0.12 to 0.13) is the same or similar to the parent CuL ion (0.13), which indicates that the nitrogen is dispersed across the molecule rather than localized, which is consistent with a tetrapyrrole structure.

Table 4. Fragment ions of metal-free ligand **G** with m/z 587 and molecular formula $[C_{32}H_{34}O_7N_4 + H]^+$.

Fragment Ion Formula	*m/z*	Error (ppb)	Neutral Loss	Relative Abundance (%)
$[C_{16}H_{17}O_2N_2]^+$	269.12845	-2	$C_{16}H_{18}O_5N_2$	100.00
$[C_{17}H_{15}O_5N_2]^+$	327.09755	0.6	$C_{15}H_{20}O_2N_2$	35.91
$[C_{24}H_{28}O_4N_3]^+$	422.20745	4	$C_8H_7O_3N$	23.71
$[C_{16}H_{19}O_3N_2]^+$	287.13901	3	$C_{16}H_{16}O_4N_2$	17.88
$[C_{16}H_{17}O_4N_2]^+$	301.11828	-1	$C_{16}H_{18}O_3N_2$	14.34
$[C_{31}H_{35}O_5N_4]^+$	543.26024	8	CO_2	13.23
$[C_{25}H_{28}O_4N_3]^+$	434.20745	4	$C_7H_7O_3N$	2.57

A major fragmentation pathway for metal-free ligand **G** is the sequential loss of CO_2 to produce the fragment ion with nominal *m/z* 543, followed by the loss of a terminal pyrrole (C_7H_7ON) to produce nominal *m/z* 422, followed by the loss of a second pyrrole ($C_8H_{11}O_2N$) to produce *m/z* 269 (**Fig. 4**). The loss of CO_2 with the loss of pyrroles has been seen when using higher fragmentation energies (like that used in this study) on bilin tetrapyrroles, [41] heme catabolites that are structurally similar to chlorophyll catabolites, though they lack ring E. In order to confirm the sequential loss of pyrroles from nominal *m/z* 422, this ion was targeted for further fragmentation. The MS^3 spectrum *of m/z* 422 (**SI Fig. 3a**) yielded a fragment ion at *m/z* 269, confirming the sequential loss of pyrroles from metal-free ligand **G**. Fragmentation of ion *m/z* 269 (**SI Fig. 3b**), yielded fragment ions at *m/z* 241.0 ($[M - CO]^+$) and 226.0 ($[M - CHON]^+$), indicating the presence of a carbonyl and a cyclic amide. [42] Based on the fragmentation patterns and the similarities to chlorophyll catabolites, a draft structure for metal-free and Cu-bound ligand **G** is proposed (**Fig. 4**). The extracted ion chromatograms (EICs) for CuLs **A**-**K** yielded unique peaks for each CuL. In contrast, the corresponding EICs of the metal-free ligands exhibited multiple peaks (**SI Fig. 4**). This further supports the stabilization of the ligand structure by complexation to Cu and suggests that there may be different isomers of each metal-free ligand.[43,44]

MS^2 spectra were obtained for all metal-free ligands except for **C** and **I**, and all Cu-bound ligands except **A**, **B**, and **J**. However, structures of these ligands can be inferred based on systematic differences stemming from their molecular formulae. Ligand **C** differs from ligand **B** by the addition of a CH_2 and elutes slightly later in the LC-ESI MS chromatogram (**Fig. 1b**). Therefore, we infer that ligand **C** is a homologue of ligand **B**. Similarly, ligand **I** differs from

ligand **H** by an H_2, and from ligand **J** by an O. Given this, we infer that ligands **H**, **I**, and **J** have the same parent structure with ligand **I** having an additional double bond compared to **H** and an alcohol functionality compared to **J**. This is supported by the MS^2 spectra of metal-free ligands **H** and **J** which have nine fragment ions in common (**SI Fig. 5**). While the MS^2 spectrum of CuL **J** was not acquired, spectra for CuLs **H** and **I** show five fragment ions in common and four fragment ions that differ by an H_2 (**SI Fig. 6**).

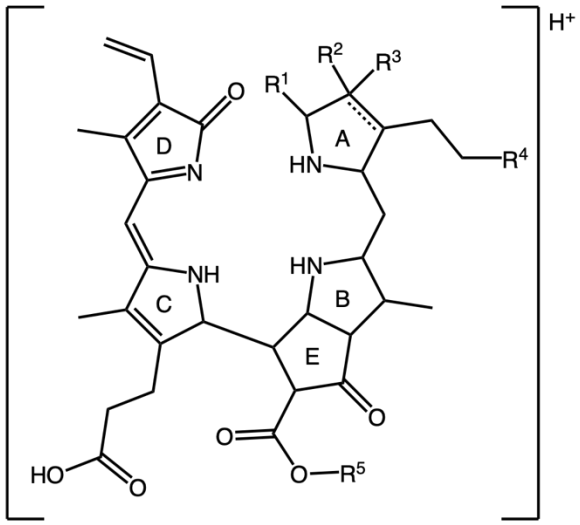
The MS^2 spectra show that the suite of ligands produced by *P. tricornutum* are structurally related (**Table 5**). Major fragment ions *m/z* 422 and 269 observed in the MS^2 spectrum of metal-free ligand **G** are also observed in the MS^2 spectra of metal-free ligands **D**, **F**, **H**, and **J**. While the MS^2 spectrum of metal-free ligand **E** does not contain *m/z* 422 or 269, it does contain *m/z* 420 and 267 indicating an additional unsaturation.

Table 5. Fragments found in metal-free ligands that match losses characteristic to chlorophyll catabolites. Only one fragment ion is shown for the loss of pyrrole(s), though in many cases there were multiple fragment ions from the loss of pyrrole(s). MS^2 spectra were not acquired for metal-free ligands **C** and **I** due to low abundance of the precursor.

Ligand	Theoretical *m/z*	$[M - CH_4O]^+$	$[M - H_2O]^+$	$[M - CO]^+$	$[M - CO_2]^+$	[M - pyrrole ring]$^+$	[M - 2 pyrrole rings]$^+$
A	522.187091	490.16088	504.17663	494.19212	478.19735	n.d.	n.d.
B	629.260591	597.23441	611.25009	n.d.	n.d.	406.17610	271.10772
D	659.271155	627.24501	641.26054	631.27619	615.28132	422.20740	269.12846
E	643.276241	611.25009	625.26580	615.28132	599.28623	420.19175	267.11281
F	645.255505	613.22938	627.24500	617.26059	601.26573	422.20745	269.12845
G	587.250026	n.d.	569.23951	559.25518	543.26024	422.20745	269.12845
H	601.265676	n.d.	583.25513	573.27079	557.27591	422.20740	269.12846
J	583.255111	n.d.	565.24470	555.26036	539.26542	422.20740	269.12846
K	460.186697	n.d.	442.17612	432.19174	416.19686	341.11317	n.d.

n.d. = not detected

Given these similarities, it is likely that the tetrapyrrole ligands **B-J** share the same parent structure but differ in the number, type, and arrangement of functional group substitutions (**Fig. 5**). The other two ligands detected, **A** and **K**, are likely tripyrroles as the same fragmentation pathways are observed (loss of methoxycarbonyl, H_2O, CO, pyrrole ring, decarboxylation). Interestingly, the tripyrroles were present as both the metal-free and Cu-bound ligands, indicating that the ligands produced by *P. tricornutum* are tri-dentate ligands, which is consistent with metal-bound complexes of chlorophyll catabolites.[45,46]



ID	R^1	R^2	R^3	R^4	R^5
B	=O	CH_3	=	OH	CH_3
C	=O	H	=	OH	CH_3
D	=O	CH_3	CH_2OH	OH	CH_3
E	=O	CH_3	CH_2OH	H	CH_3
F	=O	=	CH_2OH	OH	CH_3
G	=O	H	=	H	H
H	=O	CH_3	=	H	H
I	=O	CH_3	=	H	H
J	H	CH_3	=	H	H

Figure 5. Draft structure for ligands **B – J** with differences represented with R groups. When a double bond is indicated for R^2, the 2,3 carbons are unsaturated. Given that these ligands have similarities with chlorophyll catabolite in composition and fragmentation pathways, the positions of functional groups is inferred from known chlorophyll catabolite structures. The position of double bonds in rings A and B are not indicated as the number and position of double bonds varies among the different ligands.

Comparison with Existing Cu Ligand Compositional Data

Although nearly all Cu dissolved in seawater is bound to organic ligands, the structures of these ligands remain uncharacterized. To the best of our knowledge there is only one published high resolution MS^2 spectrum for a CuL detected in the Eastern Tropical South Pacific (ETSP).[23] Compositionally, the ETSP ligand, $[C_{20}H_{21}O_8S_2N_4]^+$, is similar to the ligands we characterized from *P. tricornutum* in that it has four N atoms and an H:C ratio of 1.0, which falls within the range of H:C ratios found in this study (0.96 to 1.09). However, there are two significant differences: the ETSP ligand contains two S heteroatoms and was found to bind Ni as well as Cu, while the ligands from *P. tricornutum* did not contain S and were specific to Cu. Additionally, the O:C and N:C ratios of the ETSP ligand, 0.4 and 0.19, respectively, were above the ranges found in this study (0.18 to 0.30 and 0.11 to 0.13, respectively). These compositional differences indicate that the ETSP ligand is more protein-like than the ligands in this study.[24]

Chlorophyll catabolites of higher plants readily form complexes with metals such as Zn^{2+}, Ni^{2+}, and Cu^{2+} and it has been suggested that this may point to a biological role of heavy

metal detoxification.[45,46] Given the structural similarities between the CuLs in *P. tricornutum* media extracts and chlorophyll catabolites, it is possible that CuLs may likewise play a role in Cu detoxification. In order to determine this, the strength of the CuL complexes needs to be determined and further experiments with cultures grown under different Cu concentrations up to the toxicity threshold of *P. tricornutum* need to be made. The structural information provided by our study may help with these studies and in further characterizing CuLs in environmental samples.

Conclusions

Spent culture media from the marine diatom *P. tricornutum* was analyzed by LC-FT-ICR MS at 21 T to characterize CuLs. Eleven CuLs were detected and assigned molecular formulae, which were confirmed through an analysis of their ^{15}N, ^{13}C, and ^{18}O isotopologues. The ligands are highly oxidized and rich in nitrogen, with O:C and N:C ratios of 0.18-0.30 and 0.11-0.13, respectively, and the Cu ion was in the +2 oxidation state. Tandem MS^2 and MS^3 spectra of CuLs and their metal-free analogues yielded diagnostic ions from fragmentation pathways characteristic of tri- and tetrapyrroles. These ligands are tri-dentate and bind Cu through coordination with nitrogen, similar to a CuL previously characterized from seawater. However, unlike the single CuL previously characterized from seawater, the CuLs isolated from *P. tricornutum* media were specific for Cu and did not bind Ni. Draft structures of *P. tricornutum* CuLs show strong similarities to known chlorophyll catabolites, suggesting that the CuLs characterized here maybe widely produced by marine photoautotrophs.

Acknowledgements

We thank Dr. Riss Kell and Dr. Mak Saito for providing the culture material and assistance with the culture work. We also thank the Ion Cyclotron Resonance group at the National High Magnetic Field Laboratory, especially Dr. Chad Weisbrod, Dr. Huan Chen, and Dr. Zeljka Popovic for their assistance with the 21T FT-ICR MS. We also thank Dr. Jeffrey Shabanowitz for helpful conversations on mass spectrometry and fragmentation. This work was funded by the National Science Foundation (NSF) program in chemical oceanography (awards OCE-1736280 and OCE-2045223) to DJR, the NSF Graduate Research Fellowship Program (award 1122374) to LBA, and the Simons Foundation Ocean Processes and Ecology program (SCOPE) award 721227 to DJR. This work was also supported through the National High Magnetic Field Laboratory User Collaboration Grants program, which is supported by NSF Division of Materials Research and Division of Chemistry through DMR-1644779, DMR-2128556, and the State of Florida.

Author Contributions

L.B.-A. and D.J.R. designed the project. L.B.-A. grew the cultures, collected, and processed the samples. L.B.-A. and J.L. performed copper ligand identification of orbitrap data. L.B.-A., A.M.M., and C.L.H. designed the LC-FT-ICR MS analyses. L.B.-A. conducted tandem mass spectrometry experiments and spectra interpretation. L.B.-A. and D.J.R. performed data analysis and wrote the first draft of the paper. All authors contributed to writing the manuscript.

Supporting Information

Additional experimental details and spectra (DOC).

Data Availability

All raw LC-MS data files are publicly available as a MassIVE dataset (https://massive.ucsd.edu, accession MSV000095816).

References

1. Sandmann, G., Reck, H., Kessler, E. & Böger, P. Distributions of plastocyanin and soluble plastidic cytochrome c in various classes of algae. *Arch Microbiol* **134**, 23–27 (1983).
2. Stryer, L. *Biochemistry*. (W.H. Freeman, New York, 1995).
3. Groussman, R. D., Parker, M. S. & Armbrust, E. V. Diversity and Evolutionary History of Iron Metabolism Genes in Diatoms. *PLoS One* **10**, 1–25 (2015).
4. Parker, M. S., Mock, T. & Armbrust, E. V. Genomic insights into marine microalgae. *Annu Rev Genet* **42**, 619–645 (2008).
5. Maldonado, M. T. *et al.* Copper-dependent iron transport in coastal and oceanic diatoms. *Limnol Oceanogr* **51**, 1729–1743 (2006).
6. Brand, L. E., Sunda, W. G. & Guillard, R. R. L. Reduction of marine phytoplankton reproduction rates by copper and cadmium. *J Exp Mar Biol Ecol* **96**, 225–250 (1986).
7. Moffett, J. W., Brand, L. E., Croot, P. L. & Barbeau, K. A. Cu speciation and cyanobacterial distribution in harbors subject to anthropogenic Cu inputs. *Limnol Oceanogr* **42**, 789–799 (1997).
8. Halliwell, B. & Gutteridge, J. M. C. Role of free radicals and catalytic metal ions in human disease: An overview. *Methods Enzymol* **186**, 1–85 (1990).
9. Imlay, J. & Linn, S. DNA Damage and Oxygen Radical Toxicity. *Science* **240**, 1302–1309 (1988).
10. Annett, A. L., Lapi, S., Ruth, T. J. & Maldonado, M. T. The effects of Cu and Fe availability on the growth and Cu : C ratios of marine diatoms. *Limnol Oceanogr* **53**, 2451–2461 (2008).
11. Lopez, J. S., Lee, L. & Mackey, K. R. M. The toxicity of copper to Crocosphaera watsonii and other marine phytoplankton: A systematic review. *Front Mar Sci* **6**, 1–13 (2019).
12. Heller, M. I. & Croot, P. L. Copper speciation and distribution in the Atlantic sector of the Southern Ocean. *Mar Chem* **173**, 253–268 (2014).
13. Jacquot, J. E. & Moffett, J. W. Copper distribution and speciation across the International GEOTRACES Section GA03. *Deep Sea Res 2 Top Stud Oceanogr* **116**, 187–207 (2015).
14. Sunda, W. G. Trace Metal Interactions with Marine Phytoplankton. *Biological Oceanography* **6**, 411–442 (1989).
15. Sunda, W. G. & Huntsman, S. A. Processes regulating cellular metal accumulation and physiological effects: Phytoplankton as model systems. *Sci Total Environ* **219**, 165–181 (1998).
16. Ahner, B. A. & Morel, F. M. M. Phytochelatin production in marine algae. 2. Induction by various metals. *Limnol Oceanogr* **40**, 658–665 (1995).
17. Ahner, B. A., Morel, F. M. M. & Moffett, J. W. Trace metal control of phytochelatin production in coastal waters. *Limnol Oceanogr* **42**, 601–608 (1997).

18. Calvo, J., Jung, H. & Meloni, G. Copper metallothioneins. *IUBMB Life* **69**, 236–245 (2017).
19. Campos, M. L. A. M. & van den Berg, C. M. G. Determination of copper complexation in sea water by cathodic stripping voltammetry and ligand competition with salicylaldoxime. *Anal Chim Acta* **284**, 481–496 (1994).
20. Ruacho, A., Richon, C., Whitby, H. & Bundy, R. M. Sources, sinks, and cycling of dissolved organic copper binding ligands in the ocean. *Commun Earth Environ* **3**, 263-281 (2022).
21. Waska, H., Koschinsky, A., Ruiz Chancho, M. J. & Dittmar, T. Investigating the potential of solid-phase extraction and Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) for the isolation and identification of dissolved metal-organic complexes from natural waters. *Mar Chem* **173**, 78–92 (2015).
22. Waska, H. *et al.* Inorganic and organic iron and copper species of the subterranean estuary: Origins and fate. *Geochim Cosmochim Acta* **259**, 211–232 (2019).
23. Boiteau, R. M. *et al.* Structural Characterization of Natural Nickel and Copper Binding Ligands along the US GEOTRACES Eastern Pacific Zonal Transect. *Front Mar Sci* **3**, 1–16 (2016).
24. Rivas-Ubach, A. *et al.* Moving beyond the van Krevelen Diagram: A New Stoichiometric Approach for Compound Classification in Organisms. *Anal Chem* **90**, 6152–6160 (2018).
25. Biswas, H., Bandyopadhyay, D. & Waite, A. Copper addition helps alleviate iron stress in a coastal diatom: Response of Chaetoceros gracilis from the Bay of Bengal to experimental Cu and Fe addition. *Mar Chem* **157**, 224–232 (2013).
26. Peers, G., Quesnel, S.-A. & Price, N. M. Copper requirements for iron acquisition and growth of coastal and oceanic diatoms. *Limnol Oceanogr* **50**, 1149–1158 (2005).
27. Peers, G. & Price, N. M. N. Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature* **441**, 341–4 (2006).
28. Zhou, X., Slauenwhite, D. E., Pett, R. J. & Wangersky, P. J. Production of copper-complexing organic ligands during a diatom bloom: Tower tank and batch-culture experiments. *Mar Chem* **27**, 19–30 (1989).
29. Kong, L. & Price, N. M. A reduction-dependent copper uptake pathway in an oceanic diatom. *Limnol Oceanogr* **65**, 601–611 (2020).
30. González-Dávila, M. *et al.* Cu transport and complexation by the marine diatom Phaeodactylum tricornutum: Implications for trace metal complexation kinetics in the surface ocean. *Sci Total Environ* **919**, 170752-170771 (2024).
31. Wei, Y. *et al.* Copper toxicity to Phaeodactylum tricornutum: A survey of the sensitivity of various toxicity endpoints at the physiological, biochemical, molecular and structural levels. *BioMetals* **27**, 527–537 (2014).
32. Morelli, E. & Pratesi, E. Production of Phytochelatins in the Marine Diatom Phaeodactylum Tricornutum in Response to Copper and Cadmium Exposure. *Bull Environ Contam Toxicol* **59**, 657-664 (1997).
33. Hendrickson, C. L. *et al.* 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer: A National Resource for Ultrahigh Resolution Mass Analysis. *J Am Soc Mass Spectrom* **26**, 1626–1632 (2015).
34. Bahureksa, W. *et al.* Improved Dynamic Range, Resolving Power, and Sensitivity Achievable with FT-ICR Mass Spectrometry at 21 T Reveals the Hidden Complexity of Natural Organic Matter. *Anal Chem* **94**, 11382–11389 (2022).

35. Smith, D. F., Podgorski, D. C., Rodgers, R. P., Blakney, G. T. & Hendrickson, C. L. 21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures. *Anal Chem* **90**, 2041–2047 (2018).
36. Boiteau, R. M. & Repeta, D. J. An extended siderophore suite from Synechococcus sp. PCC 7002 revealed by LC-ICPMS-ESIMS. *Metallomics* **7**, 877–884 (2015).
37. Karg, C. A., Taniguchi, M., Lindsey, J. S. & Moser, S. Phyllobilins– Bioactive Natural Products Derived from Chlorophyll– Plant Origins, Structures, Absorption Spectra, and Biomedical Properties. *Planta Med* ***89***, 637–662 (2022).
38. Muller, T., Vergeiner, S. & Krautler, B. Structure elucidation of chlorophyll catabolites (phyllobilins) by ESI-mass spectrometry—Pseudo-molecular ions and fragmentation analysis of a nonfluorescent chlorophyll catabolite (NCC). *Int J Mass Spectrom* **365**, 48–55 (2014).
39. Roca, M., Ríos, J. J. & Pérez-Gálvez, A. Mass spectrometry: the indispensable tool for plant metabolomics of colourless chlorophyll catabolites. *Phytochem Rev* **17**, 453–468 (2018).
40. Bale, N. J., Llewellyn, C. A. & Airs, R. L. Atmospheric pressure chemical ionisation liquid chromatography/mass spectrometry of type II chlorophyll-a transformation products: Diagnostic fragmentation patterns. *Org Geochem* **41**, 473–481 (2010).
41. Sekera, E. R. & Wood, T. D. Examination of the fragmentation behavior of hemin and bilin tetrapyrroles by electrospray ionization and collision-induced dissociation. *Mass Spectrom Lett* **9**, 91–94 (2018).
42. Niessen, W. M. A. & Correa C., R. A. Fragmentation of Even-Electron Ions in Interpretation of MS-MS Mass Spectra of Drugs and Pesticides. 71–128 (John Wiley & Sons, Inc., 2017).
43. Moser, S., Scherzer, G. & Kräutler, B. On the Nature of Isomeric Nonfluorescent Chlorophyll Catabolites in Leaves and Fruit – A Study with a Ubiquitous Phylloleucobilin and its Main Isomerization Product. *Chem Biodivers* **14**, (2017).
44. Hörtensteiner, S. Update on the biochemistry of chlorophyll breakdown. *Plant Mol Biol* **82**, 505–517 (2013).
45. Li, C. & Kräutler, B. Transition metal complexes of phyllobilins - a new realm of bioinorganic chemistry. *Dalton Trans* **44**, 10116–10127 (2015).
46. Li, C. *et al.* Blue transition metal complexes of a natural bilin-type chlorophyll catabolite. *Chem Sci* **5**, 3388–3395 (2014).

For Table of Contents Only