Environmental Science & Technology

pubs.acs.org/est

Article

High-Resolution Tandem Mass Spectrometry-Based Analysis of Model Lignin–Iron Complexes: Novel Pipeline and Complex Structures

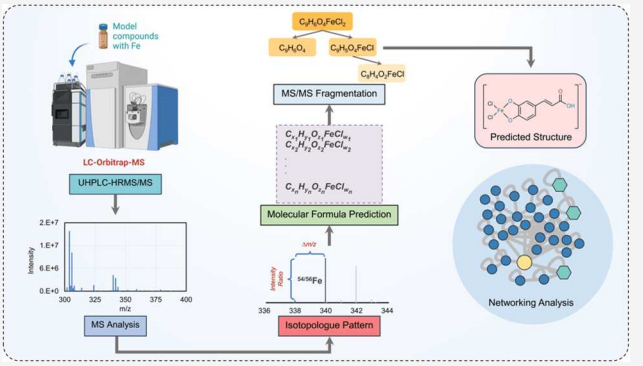
Abrar Shahriar, Srinidhi Lokesh, Anil Timilsina, Travis Numan, Tilman Schramm, Paolo Stincone, Laurinda Nyarko, Christian Dewey, Daniel Petras, Rene Boiteau, and Yu Yang*

Cite This: *Environ. Sci. Technol.* 2024, 58, 15090–15099

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Understanding the chemical nature of soil organic carbon (SOC) with great potential to bind iron (Fe) minerals is critical for predicting the stability of SOC. Organic ligands of Fe are among the top candidates for SOCs able to strongly sorb on Fe minerals, but most of them are still molecularly uncharacterized. To shed insights into the chemical nature of organic ligands in soil and their fate, this study developed a protocol for identifying organic ligands using ultrahigh-performance liquid chromatography-high-resolution tandem mass spectrometry (UHPLC-HRMS/MS) and metabolomic tools. The protocol was used for investigating the Fe complexes formed by model compounds of lignin-derived organic ligands, namely, caffeic acid (CA), *p*-coumaric acid (CMA), vanillin (VNL), and cinnamic acid (CNA). Isotopologue analysis of $^{54/56}$Fe was used to screen out the potential UHPLC-HRMS (m/z) features for complexes formed between organic ligands and Fe, with multiple features captured for CA, CMA, VNL, and CNA when $^{35/37}$Cl isotopologue analysis was used as supplementary evidence for the complexes with Cl. MS/MS spectra, fragment analysis, and structure prediction with SIRIUS were used to annotate the structures of mono/bidentate mono/biligand complexes. The analysis determined the structures of monodentate and bidentate complexes of FeL_xCl_y (L: organic ligand, $x = 1-4$, $y = 0-3$) formed by model compounds. The protocol developed in this study can be used to identify unknown organic ligands occurring in complex environmental samples and shed light on the molecular-level processes governing the stability of the SOC.



KEYWORDS: *high-resolution mass spectrometry, lignin, iron complexes, organic ligands, molecular networking*

INTRODUCTION

Association with iron (Fe) minerals is perceived as one of the important processes that stabilizes soil organic carbon (SOC) in soils and mediates its biogeochemical cycles. Association with Fe minerals contributes to up to 37.8% of SOC in forest soils across a range of temperate ecosystems, grasslands, wetlands, and Arctic tundra and permafrost soils.[1−4] Experimental studies documented the slower metabolism and respiration of Fe-bound SOC compared to other pools of SOC.[1,5,6] In particular, recent investigations have uncovered the important roles of Fe minerals and their redox reactions in the response of permafrost soil to climate change and processes of SOC during the melting process of permafrost soils.[4,7,8]

However, to date, the chemical nature of Fe-bound SOC is still largely unknown. Synchrotron-based X-ray absorption spectral analyses have determined that Fe-bound SOC is more aromatic and phenolic than bulk soils.[9,10] High-resolution mass spectrometry (HRMS) has been used in analyzing the chemical properties of Fe-bound SOC and its release upon the redox reactions of Fe.[11,12] Results indicate that Fe-bound OC in general has higher O/C and double bond equivalence than other SOCs.[1−3,13] Upon the reduction of Fe, a more hydrophobic SOC with more lignin-like components was released. More investigations into the chemical nature of Fe-bound SOC can provide in-depth insights into the Fe-mediated stabilization and cycling of the SOC.

Organic Fe ligands can be top candidates for SOC that strongly associate with Fe mineral surfaces. Ultrahigh-performance liquid chromatography-high-resolution mass spectrometry (UHPLC-HRMS) analysis as well as metabolomic studies were

Received: April 11, 2024
Revised: July 31, 2024
Accepted: August 1, 2024
Published: August 8, 2024

used to study the unknown Fe-binding compounds, mostly siderophores.[11,14,15] Together with HPLC-inductively coupled plasma mass spectrometry (ICP-MS) analysis, HPLC-HRMS identified the complexes with Fe formed by microbial-originated siderophores, and further MS/MS analysis determined the chemical structures for important unknown Fe-binding OC.[11,16,17] Beyond microbial siderophores, there is limited research about other naturally occurring organic Fe ligands in soil environments and their complexes with Fe.[18,19] Compared to other organic compounds, analysis for organic ligands and their complexes with Fe are even more challenging, with additional difficulties, including but not limited to the absence of standards for complexes, and discrepancies between MS analysis results and aqueous chemistry studies.[20,21]

Lignin is one of the major plant macromolecules and precursors for SOC and can be potentially an important source for Fe-binding SOC given its phenolic/catechol-dominated degradation products. Among litter materials, cellulosic tissue is relatively easily depolymerized into small saccharides or monosaccharides that are efficiently assimilated by microorganisms, whereas lignin is broken down into small compounds via an energy-intensive process that oxidatively cleaves chemical bonds.[22−25] The resulting degradation products of lignin are a complex mixture with various oxidation states and functional groups that may be stabilized by association with Fe minerals.[22,26] Phenolic, catechol, ether, and other O-containing functional groups are widely occurring among the lignin degradation products, which makes them good candidates for Fe-binding ligands, for which there are limited MS studies about their complexation with Fe.

This study aimed to (1) identify the Fe complexes formed by model compounds of lignin degradation products; (2) develop a protocol that can be used to identify molecularly unknown lignin-derived Fe ligands and corresponding complexes with Fe. Lignin-derived model compounds, i.e., caffeic acid (CA), *p*-coumaric acid (CMA), vanillin (VNL), and cinnamic acid (CNA), were selected based on four criteria: (a) identification as lignin degradation products,[27,28] (b) presence of functional groups potentially binding with Fe, (c) molecular weight in the range of 100−1000 Da (suitable for electrospray-ionization (ESI) MS), and (d) authentic standard commercially available.

MATERIALS AND METHODS

Materials and Sample Preparation. Four model compounds, i.e., caffeic acid (CA) (>98%, Sigma-Aldrich), *trans-p*-coumaric acid (CMA) (>98%, TCL), vanillin (VNL) (99%, Acros), and *trans*-cinnamic acid (CNA) (99%, Thermo Fisher), were used for this study. Details of model compounds and their chemical properties are compiled in the SI (Table S1). All of the model compounds were prepared in HPLC-grade methanol. $FeCl_3{\cdot}6H_2O$ (>97%, Fisher Chemical) salt solution was prepared with double-deionized (DDI) water (18.2 $M\Omega{\cdot}cm^{-1}$). Model compounds were reacted with Fe(III) at different concentrations (0−1000 μM). The concentration range of the samples was selected based on their environmental relevance and instrument limits. The typical Fe(III) concentration in soil ranges from 50 mM to 200 mM; however, this concentration can be within the range of 20−1600 μM in soil porewater.[8,29] In the case of Cl^-, the typical average concentration in soil is 2.85 mM and it can be as high as 8.67 mM in sediment porewater.[30,31] The complex samples were equilibrated overnight at room temperature. A pH range of 3−7 was chosen for the analysis. All of the samples were prepared in duplicates.

LC-Orbitrap MS/MS Analysis. A Vanquish UHPLC (Thermo Fisher Scientific) coupled to a Q Exactive HF quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) was used for the analysis of the prepared model compound-$FeCl_3$ samples. The injection volume of the sample was set to be 10 μL and eluted through a C18 core−shell column (Kinetex C18, 150 mm × 2.1 mm, 1.7 m particle size, 100 Å pore size, Phenomenex, Torrance). The samples were eluted with a mobile phase comprised of solvent A (LC-MS grade H_2O + 0.1% formic acid) and solvent B (LC-MS grade methanol (MeOH) + 0.1% formic acid) with a flow rate of 500 μL/min for a total method run of 16 min. The solvent gradient of the mobile phase was set as 40% of solvent B from 0 to 6 min of the run, a linear gradient from 40 to 50% of B up to 8 min, then 50−99% of B from 8 to 10 min, followed by the washout phase for the next 3 min at 99% of B, and then the last 3 min of re-equilibration phase at 5% of B. Data-dependent acquisition (DDA) of MS spectra was performed primarily in negative mode. The parameters for ESI were set as a sheath gas flow of 50 arbitrary units (arb. units), the auxiliary gas flow of 12 arb, and the auxiliary gas temperature of 400 °C. The temperature of the inlet capillary was 320 °C, while the spray current and voltage were set as 100 A and 2400 V, respectively, for the negative mode. The scan range for MS was set to 150−1500 m/z with a resolution at m/z 200 ($R_{m/z\ 200}$) of 120,000 with one microscan. The maximum ion injection time was set to 100 ms with an automated gain control (AGC) target of 1.0×10^{06}. Up to five MS/MS spectra per MS^1 survey scan were recorded in DDA mode with $R_{m/z\ 200}$ of 15,000 with one microscan. For the DDA MS/MS, the maximum ion injection time was set to 100 ms with an AGC target of 5.0×10^{05}. The MS/MS precursor isolation window was set to m/z 1. The normalized collision energy was stepped from 25 to 35% to 45% with z = 1 as the default charge state. MS/MS scans were triggered at the apex of chromatographic peaks within 2−15 s from their first occurrence. The intensity threshold was set as 8.0×10^{04}, and the dynamic exclusion was 5.0 s. Both isotope peaks and ions with unspecified charge states were omitted from the MS/MS acquisition. The selected resolutions for both MS^1 and MS/MS and other parameters for data-dependent MS/MS acquisition are sufficient for analyzing complex samples.[14,32] A subset of samples were analyzed using an Orbitrap IQ-X quadrupole mass spectrometer coupled with a Thermo Scientific Vanquish Horizon UHPLC system using a Phenomenex BioZen XBC18 (2.1 mm × 100 mm with guard column) column. Details of this analysis can be found in the SI (SI, Text S1.1). Description of the LC-Orbitrap MS/MS analysis in positive mode can also be found in the SI (SI, Text S1.2).

Data Processing. *Isotopologue Pattern Matching.* The acquired LC-Orbitrap-MS data (in .raw format) was converted to mzML format using the MSconvert application (ProteoWizard) before the data were analyzed by an in-house algorithm written in R using the xcms library (code details in SI, Text S1.3).[16,33,34] The developed algorithm searches all of the MS scans and compiled all of the peaks that have exhibited Fe isotope pattern ($^{54/56}$Fe, $\Delta m/z$ = 1.9953 and intensity ratio = 0.058 in case of only one Fe present in the complex) along with extracted ion chromatogram (XIC) and defined apo-ligand XIC. Mass tolerance between the isotopologue was set to 0.001 ($\Delta m/z$ of $^{54/56}$Fe), whereas the intensity ratio

https://doi.org/10.1021/acs.est.4c03608
Environ. Sci. Technol. 2024, 58, 15090−15099

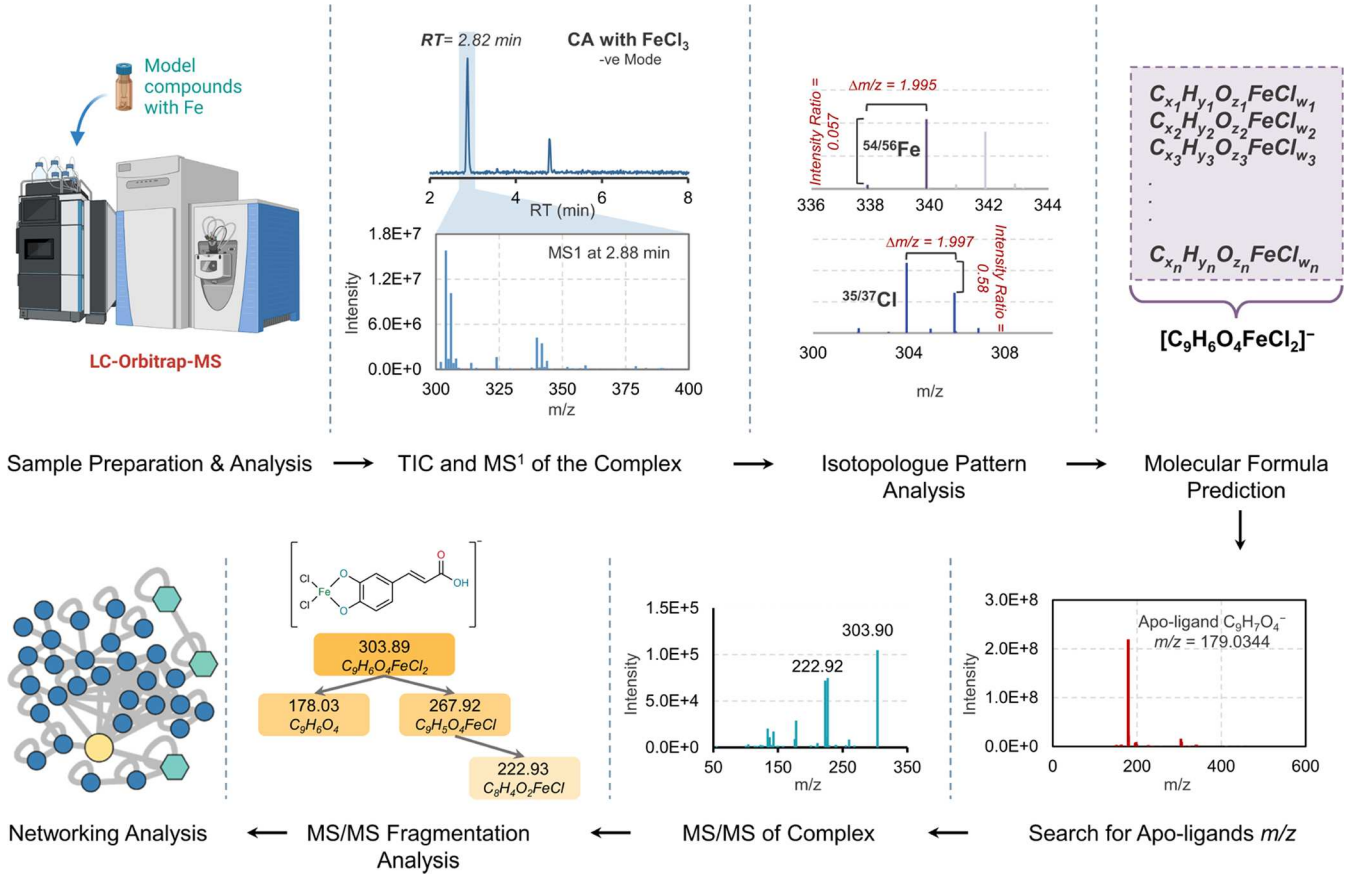


Figure 1. Overview of the data analysis process to identify Fe-model compound complexes. For better illustration purposes, a complex of CA (m/z = 303.899) identified in this study was used as an example throughout the figure. Each step elucidates step-by-step processes that were followed to figure out the chemical formula and structure of the complex. Searching for apo-ligands is not critical for the model compound but is important for searching for and identifying unknown ligands in complex environmental samples.

tolerance was considered within a factor of 1.3 ($^{54/56}$Fe intensity). The parameters and criteria for removing mass sets, which are the result of background and instrumental noises, can be found in the previous study.[16] Two additional criteria were included for the current study in the case of consideration of Fe isotopologue: (1) the XIC of the potential isotopologue was captured as a feature in MZmine3 (details of MZmine3 can be found in a later section) and (2) the retention time (RT) of the isotopologue is not too early (<1 min) or late (>10 min, washout or re-equilibration phase). In addition, the Cl isotopologue ($^{35/37}$Cl, $\Delta m/z$ = 1.997) pattern was also searched, and based on the intensity ratio, the number of Cl present in the potential compound was predicted. The maximum number of Cl present in the singly charged complex was assumed to be 3. The intensity ratios between the major pair of Cl isotopologues were 0.97, 0.64, and 0.32 for $^{35}Cl_3/^{35}Cl_2{}^{37}Cl$, $^{35}Cl_2/^{35}Cl^{37}Cl$, and $^{35}Cl/^{37}Cl$, respectively. The mass difference uncertainty was chosen more stringently with a mass error of 5 ppm when the uncertainty regarding the intensity was higher (intensity ratio $^{54}Fe/^{56}Fe = 0.058 \pm 0.02$).

LC-MS/MS Feature Detection. Before the molecular formula prediction of the potential complexes, acquired MS data were processed using MZmine3 (version 3.4.16) (http://mzmine.github.io/).[35] The intensity thresholds of MS^1 and MS/MS spectra were set to 1.0×10^{05} and 1.0×10^{03}, respectively. ADAP chromatogram module was used for the chromatogram building of MS^1 with a scan accuracy (m/z) of 10 ppm.[36] Deconvolution of the XIC was performed using the local minima resolver (LMR) with a retention time tolerance of 0.2 min, a chromatographic threshold of 0.85, and at least 3 data points were considered for resolving the features. An isotope filter module was also incorporated to filter out the same features containing a ^{13}C isotope with an m/z tolerance of 10 ppm. Join aligner is also performed after all of those processes. The predicted complex m/z features were manually searched in the feature list, and their XIC and MS data were further analyzed. For the MS^1 peaks, the m/z values were reported to four decimal places, whereas for the MS/MS peaks, they were reported to two decimal places. Molecular formulas for all of the features were then annotated considering the presence of C, H, O, Fe, and Cl in the molecule at different numbers. A simple complementary molecular formula prediction was also performed using the ChemCal web application (https://www.chemcalc.org/). Based on the monoisotopic mass of the predicted complexes and the potential number of Cl present in the complex (if any based on the $^{35/37}$Cl isotopologue), the molecular formula was predicted. The apo-ligands were also manually searched in the feature list, and their XICs were compared with the potential complex XICs. The feature tables were exported from MZmine3 for creating fragment analysis in SIRIUS and molecular networking in Global Natural Product Social Molecular Networking (GNPS).

Although the current study is a targeted analysis, the spectral library search was done within GNPS. GNPS was incorporated

https://doi.org/10.1021/acs.est.4c03608
Environ. Sci. Technol. 2024, 58, 15090−15099

for library search with a minimum matched peak for the library search of 4 as well as a minimum cosine value of 0.70.

In Silico Annotation and Molecular Networking. MS/MS data were further processed, and *in silico* annotation was done using SIRIUS (version 5.6.3).[37] Fragmentation trees of the list of potential complexes were created based on the MS/MS data with an m/z error of 5 ppm. CSI:FingerID module of SIRIUS was also incorporated during the data analysis process as dimers or other products originating from parent compounds may be observed in the prepared samples. SIRIUS can also predict the structure of the features with MS/MS that may not be found based on library search, but it should be noted that complex structures cannot be resolved by SIRIUS. Instead, SIRIUS was incorporated to predict all of the potential apo-ligand structure. As stated earlier, exported files from MZmine3 (feature table file as .csv along with additional pair data and corresponding MS/MS spectra files as .mgf) were submitted to GNPS for ion-identity molecular networking (IIMN) analysis.[38] metaCorrelate module in MZmine3 was utilized with a retention time tolerance of 0.1 min. For the feature shape correlation, a Pearson correlation value of 0.85 was selected for the minimum shape correlation, and the feature height correlation was set as 0.70. In addition to the frequently found adducts, additional adducts of $[M+FeCl_3]^-$ and $[M+FeCl_2]^-$ were considered with an m/z tolerance of 10 ppm. The generated files were submitted to GNPS. Cytoscape (version 3.9.1) was used for visualizing the molecular networks.[39]

QA/QC Analysis. For the QA/QC purpose, a mix of sulfamethazine, sulfamethizole, sulfachloropyridazine, sulfadimethoxine, and amitryptiline was prepared to a final concentration of 10 μg/mL and the final concentration of coumarin-314 was 20 μg/mL. The instrument was calibrated to <3 ppm mass accuracy prior to the analysis of the samples. The detailed table of the compounds and the QA/QC MS figures can be found in the SI (SI, Table S2 and Figure S1). Negative control or reference background sample was taken with an average signal intensity of 1.0×10^{05}. Feature list blank subtraction was also performed after an aligned feature list was generated by MZmine3 as described in the previous section. Features present in the blank were subtracted from the final feature list based on intensity.

RESULTS AND DISCUSSION

Overview of the Protocol for Identifying the Complexes Formed by Model Compounds. A protocol was developed and used for identifying the Fe complexes formed by model lignin degradation products with four major steps (Figure 1): (1) screening of $^{54/56}$Fe isotopologue followed by $^{35/37}$Cl isotopologue for potential metal complexes: $^{54/56}$Fe isotopologue will be screened as a primary benchmark for the complexes with Fe when $^{35/37}$Cl isotopologue is used as an additional evidence as hybrid complexes can often occur; (2) identification of features from MS^1 data using MZmine3 and annotation of molecular formula: based on the m/z value of the potential complex feature combined with Fe and an additional atom's isotope pattern, a list of molecular formulas were predicted for potential complexes; (3) fragmentation analysis using SIRIUS based on the MS/MS spectra: the structure of the complexes can be resolved and explained by scrutinizing fragmentation trees generated by SIRIUS; and (4) ion-identity molecular networking analysis (IIMN) and complementary library search using HRMS/MS database (e.g., GNPS): IIMN and library search can reveal additional potential complexes and related apo-ligands.

Results reported in the main text are based on the (−) MS analysis results, whereas complementary (+) MS results can be found in the SI. The complexes and apo-ligands reported in both modes were singly charged. The major analyses were focused on results for samples with 250 μM model compounds and 500 μM $FeCl_3$ at pH of 3.14−3.31 when the results for other concentrations and pH setups were used for analyzing the impact of analyte concentrations and solution pH (SI, Text S1.4 and 1.5). This analysis protocol was developed with the aim of identifying unknown organic ligands in complex environmental samples and validating using the model compounds (Figure 1). All of the mass spectral data are publicly available in Mass spectrometry Interactive Virtual Environment (MassIVE) (https://massive.ucsd.edu/) (SI, Table S3).

Isotopologue Pattern Screening Analysis. The results of $^{54/56}$Fe isotopologue analysis revealed 25−106 unique features (MS^1 signal intensity $>1 \times 10^{05}$) at different retention times, likely with associated Fe for all four model compounds (SI, Figure S2). After cleaning up, i.e., background removal (S/N > 10 at various concentration ratios), and validation in replicates, there were 5, 4, 10, and 6 features with $^{54/56}$Fe isotope patterns for samples of CA, CMA, VNL, and CNA, respectively, showing the coelution of monoisotopic peaks containing ^{54}Fe and ^{56}Fe. As the next step, molecular formulas were successfully assigned for 2, 2, 3, and 1 features of complexes for CA, CMA, VNL, and CNA, respectively. CA had 2 peaks with m/z values of 303.8987 and 411.9872 at a retention time of 2.82 and 2.88 min, respectively, with $^{54/56}$Fe isotope patterns (SI, Figure S3), while CMA also had 2 peaks with m/z of 451.9531 and 708.0937 at a retention time of 3.67 min.

In parallel to compilation of the potential m/z matching of the natural abundance Fe isotope pattern, MS^1 peaks were searched independently for the $^{35/37}$Cl isotope for possible complexes with Cl. The most likely number of Cl present in the complex was determined by comparing the isotopologue pattern of molecules with those with different numbers of Cl using an in-house R script (SI, Figure S4). In the case of CA, 18−38, 19−36, and 24−42 unique features with $^{35}Cl_3/^{35}Cl_2{}^{37}Cl$, $^{35}Cl_2/^{35}Cl^{37}Cl$, and $^{35}Cl/^{37}Cl$ isotopic patterns were observed, respectively, for samples with different concentration ratios of Fe and CA. As an example, for CNA, m/z = 307.8867, 309.8838, and 311.8809 triads showed $\Delta m/z$ = 1.997 among them and the intensity ratios were 0.93 (peak intensity of 309.8838/307.8867) and 0.27 (peak intensity of 311.8809/307.8867), matching the isotopic pattern of a molecule with 3 Cl (SI, Figure S5). Such kinds of Fe complexes containing Cl were also observed in previous studies.[20,40] Beck et al.[20] found complexation between organic acids (e.g., oxalic acid, malic acid) and $FeCl_3$ resulted in $[M-H+FeCl_3]^-$, $[M-2H+FeCl_2]^-$, and $[M-3H+FeCl]^-$ complex species.

Molecular Formula Prediction and Capturing of Apo-Ligands. For predicting the molecular formula of the potential complexes, C, H, and O atoms were considered along with Fe, and the number of Cl atoms in the complexes was taken into account. A mass error below 3 ppm was observed compared to the theoretical exact mass (SI, Figure S6). When molecular formulas were assigned for complexes, a lower mass error was observed when considering Fe and Cl, as opposed to when

https://doi.org/10.1021/acs.est.4c03608
Environ. Sci. Technol. 2024, 58, 15090−15099

Table 1. List of High-Performance Liquid Chromatography (HPLC)-High-Resolution Tandem Mass Spectrometry (HRMS/MS) Features Exhibiting $^{56/54}$Fe (and $^{35/37}$Cl Isotope Pattern if Present) and Their Predicted Molecular Formula with Mass Error in Negative Mode[a]

model compounds	retention time (min)	m/z of potential complex	possible formula	intensity	exact mass of potential complex	mass error (ppm)	m/z of apo-ligands	mass error for apo-ligands	MS/MS for complex and apo-ligands?
CA	2.82	303.8987	$[(C_9H_6O_4)FeCl_2]^-$	8.99×10^{05}	303.8992	1.61	179.0339	3.28	Yes
	2.88	411.9872	$[(C_9H_6O_4)_2Fe]^-$	1.78×10^{07}	411.9882	2.47	179.0339	2.95	Yes
CMA	3.67	451.9531	$[(C_9H_7O_3)_2FeCl_2]^-$	4.72×10^{05}	451.9517	3.11	163.0402	4.34	Yes
	3.67	708.0937	$[(C_9H_7O_3)_4Fe]^-$	2.27×10^{06}	708.0931	0.86	163.0396	0.96	Yes
VNL	3.38	427.9522	$[(C_8H_7O_3)_2FeCl_2]^-$	2.19×10^{06}	427.9517	1.22	151.0397	1.15	Yes
	3.38	544.0231	$[(C_8H_7O_3)_3FeCl]^-$	7.99×10^{06}	544.0224	1.39	151.0397	1.26	Yes
	3.38	660.0939	$[(C_8H_7O_3)_4Fe]^-$	7.30×10^{06}	660.0930	1.34	151.0397	1.34	Yes
CNA	5.66	307.8867	$[(C_9H_7O_2)FeCl_3]^-$	1.70×10^{06}	307.8861	1.91	147.0452	3.93	Yes

[a]All of the possible formulas contain a single negative charge. The intensity and mass error of the complexes in the table compiled are based on 250 μM model compounds: 500 μM $FeCl_3$ concentration samples. Mass errors are shown as absolute values.

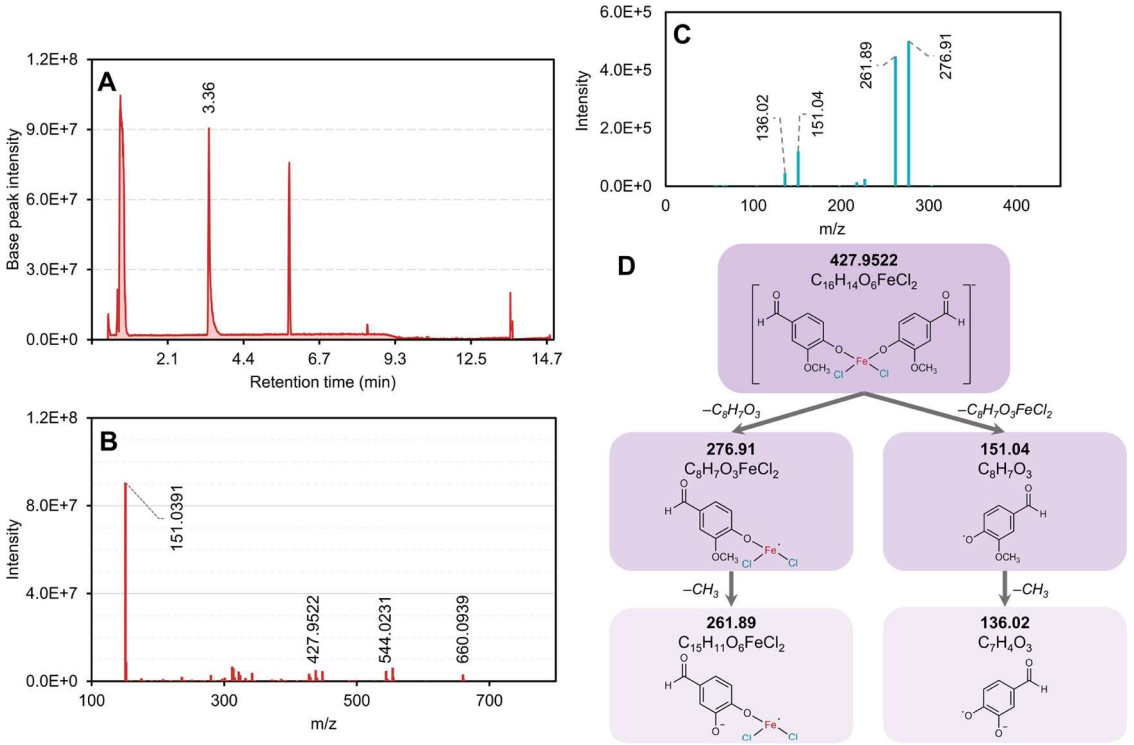


Figure 2. (A) LC-Orbitrap MS chromatogram of VNL in the −ve mode; the parent peak and the potential complex peaks were found at marked 3.36 min RT, (B) MS^1 spectrum at an RT of 3.36 min; 151.0392 is the VNL peak in the -ve mode, whereas the other three marked peaks are potential complex peaks, (C) MS/MS spectra of m/z = 427.9522, and (D) MS/MS fragment analysis of VNL complex (m/z = 427.9522) based on SIRIUS and its potential molecular structure.

they were disregarded. It reinforced their presence within the molecule, complementing their isotopic pattern as evidence. For instance, the predicted molecular formula for the feature with m/z = 544.0231 (a potential complex of VNL, which exhibited isotopologue of $^{54/56}$Fe and $^{35/37}$Cl) presuming the presence of Fe and Cl in the complex was $[C_{24}H_{21}O_9FeCl]^-$ with an m/z error of 1 ppm, whereas without considering Fe and Cl, the predicted formula was $[C_{27}H_{12}O_{13}]^-$ with 9 ppm error assuming common annotation rules.[41,42] Similar outcomes for all of the annotated molecular formulas were observed, with the absolute mass error ranging from 0.86 to 3.1 ppm for complex features when considering both Fe and Cl when the error was 3.6 ppm or higher when calculated with only C/H/O (**SI**, Figure S7).

A range of complexes ($[FeL_xCl_y]^-$, L indicates organic ligands, x = 1−4, y = 0−3) were observed for model compounds (Table 1). CA formed monoligand bidentate complexes of $[(C_9H_6O_4)FeCl_2]^-$ and biligand bidentate complexes of $[(C_9H_6O_4)_2Fe]^-$, which dominated with an intensity of 1.78×10^{07}. It is reasonable given the configuration is stereometrically favored as bidentate catechol ligands and Fe(III) can form very stable complexes.[43,44] CMA formed biligand monodentate complex ($[(C_9H_7O_3)_2FeCl_2]^-$) and quadra-ligand complexes ($[(C_9H_7O_3)_4Fe]^-$). VNL formed biligand monodentate complex $[(C_8H_7O_3)_2FeCl_2]^-$, triligand monodentate complex ($[(C_8H_7O_3)_3FeCl]^-$), and quadra-ligand complex ($[(C_8H_7O_3)_4Fe]^-$). Only the monoligand monodentate complex ($[(C_9H_7O_2)FeCl_3]^-$) was observed for CNA. Bidentate complexes were only observed for CA, and for

https://doi.org/10.1021/acs.est.4c03608
Environ. Sci. Technol. 2024, 58, 15090−15099

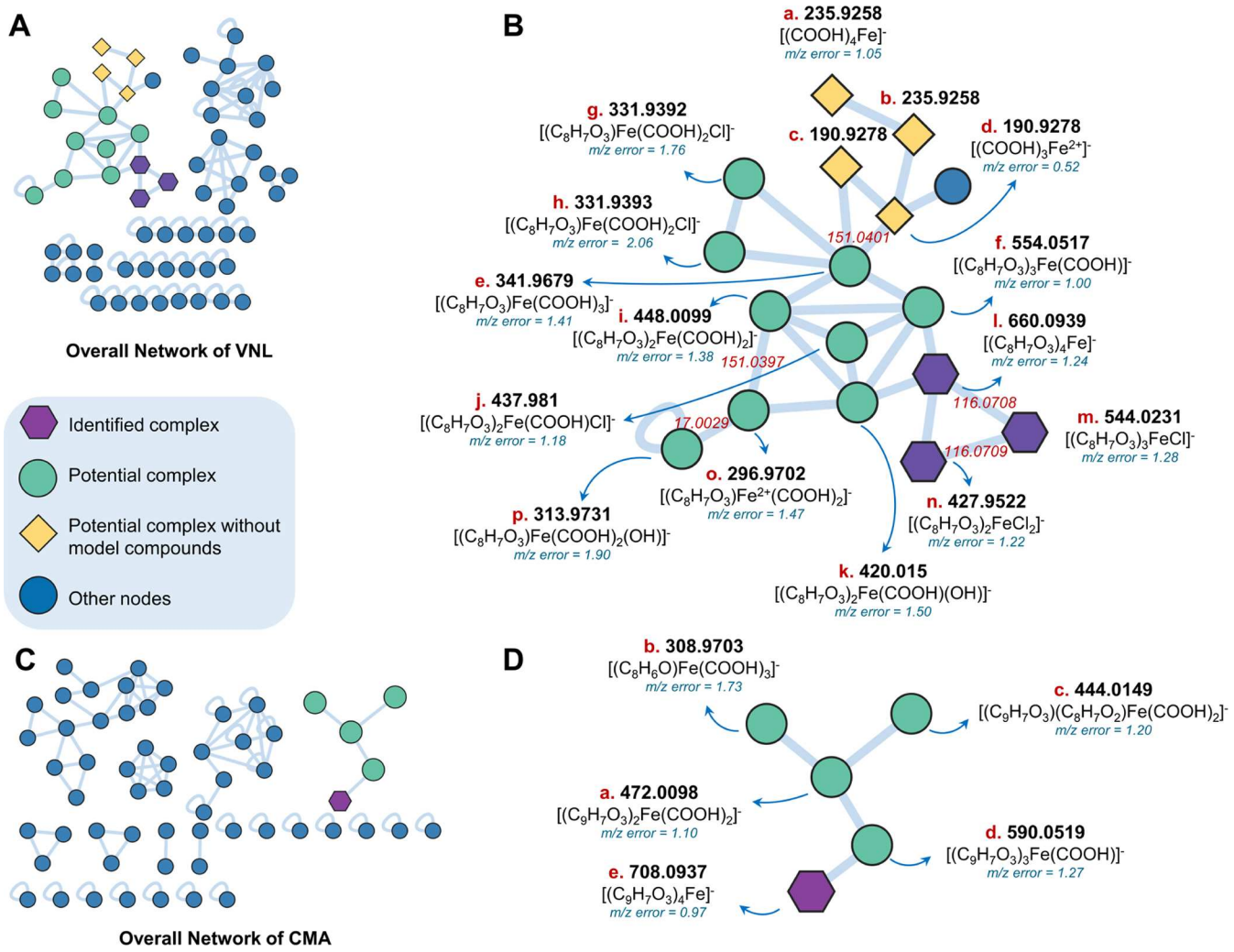


Figure 3. (A) Overall ion-identity molecular network (IIMN) results for VNL in the −ve mode, (B) a cluster where three of the identified VNL complexes were found along with potential complexes, (C) IIMN for CMA, and (D) cluster with identified CMA complexes. The black bold numbers represent the m/z value of each node, whereas the red numbers in between the nodes represent the mass difference between the highlighted nodes. The predicted molecular formula and the absolute value of m/z error in ppm are also shown for each node.

CA only bidentate complexes were captured. Hybrid complexes with organic ligands and Cl^- were observed, with the signal intensity determined by the stability constants, which are unavailable for these organic ligands at this point.

The prediction of the apo-ligand molecular formula based on MZmine3 and ChemCal matched the molecular formula of model compounds or their derivatives. For example, the calculated m/z errors of VNL as apo-ligands from $[2M-2H+FeCl_2]^-$, $[3M-3H+FeCl]^-$, and $[4M-4H+Fe]^-$ were 1.05−1.55, 1.26−1.66, and 1.24−1.55 ppm, respectively. All of these complex species eluted at retention times similar to those of their parent ligands (Figure 2A,B): apo-ligands and complexes of CA, CMA, VNL, and CNA were eluted at 2.82−2.89, 3.67−3.71, 3.36, and 5.66 min, respectively. Such a trend was also observed in previous studies for Fe-binding ligands.[14] TIC and XICs of the potential complexes of CMA and CNA as well as MS^1 peaks at that retention time are illustrated in the SI (SI, Figure S8).

In the +ve mode, however, Cl was not observed as a ligand in the predicted complexes. Instead, OH was observed to be present as a ligand in the complexes. The general configuration for the complexes in the + ve mode was determined as $[FeL_x(OH)_y]^+$, where $x = 1-4$, $y = 0-3$ (SI, Table S4). Similar to the -ve mode, the complexes and apo-ligands coeluted as the similar retention time. Most of the complexes exhibited $^{54/56}Fe$ isotope patterns and have MS/MS fragmentation data except CA (SI, Figure S9 and Table S4). Although CA complexes listed in the table did not always have a clear isotope pattern, their XICs at the same retention time as their parent peak and accurate formula assignment (<2 ppm for most of the species) strengthen the fact of the existence of such complexes.

Fragmentation Analysis and Prediction of Molecular Structure for Complexes and Apo-Ligands. The predicted formulas of the complexes were corroborated by the analysis of MS/MS data using SIRIUS. MS/MS data of the complexes were utilized for making fragmentation trees with predicted formulas for all of the major fragment peaks. VNL complexes with Fe were illustrated where fragmentation trees of $[(C_8H_7O_3)_2FeCl_2]^-$ (Figure 2C,D) and $[(C_8H_7O_3)_3FeCl]^-$ (SI, Figure S10) were built based on the major MS/MS fragment peaks. SIRIUS can corroborate the predicted structures of the detected complexes by explaining the major MS/MS fragments and neutral losses of complexes. For CA complexes with Fe, the fragmentation tree is more complicated

https://doi.org/10.1021/acs.est.4c03608
Environ. Sci. Technol. 2024, 58, 15090−15099

as depicted for $[(C_9H_6O_4)_2Fe]^-$ with more branches and fragment species (SI, Figure S11). SIRIUS also successfully interpreted the MS/MS and fragmentation for apo-ligands (SI, Figure S11). As a demonstration for identifying unknown ligands with standard available ("unknown (for their generation in lignin degradation or binding capacity for Fe) known (MS library for standard available)"), apo-ligands of CA, CMA, CNA, and VNL were matched with the standard through GNPS annotation (cosine value = 0.94−0.98, m/z error <1 ppm). Such success laid the foundation for annotating the MS^1 and MS/MS data to be collected for the unknown ligands and complexes detected in the environmental samples.

SIRIUS was also valuable for annotating the structures of the dimers or other derivative compounds that originated from the model compounds. For example, one of the CA derivatives (dehydrodicaffeic acid dilactone, $C_{18}H_{13}O_8^-$, m/z = 357.0603, retention time = 4.80 min) was found in $FeCl_3$-spiked CA samples based on SIRIUS prediction (similarity score = 95.3%). It was reported that dehydrodicaffeic acid dilactone can be generated when CA is oxidized. Huber et al.[45] validated the structure of dilactone dimer of CA with an authentic standard and showed that this compound can be a biotransformation product of CA when secretome generated from *Botrytis cinerea* was introduced in CA samples. Dehydrodicaffeic acid dilactone was likely produced due to the oxidation of CA by $FeCl_3$ as it has been reported that Fe^{3+} ion can stimulate the process.[46]

Additional library matchings were found for samples of complexes for CA (6), CMA (1), VNL (2), and CNA (1) based on a GNPS library search (SI, Table S5). For instance, based on annotation using GNPS, dehydrodicaffeic acid dilactone (intensity of 6.86×10^{07}, retention time = 4.80 min) and 4-(3,4-dihydroxyphenyl)-6,7-dihydroxynaphthalene-2-carboxylic acid (intensity of 1.36×10^{07}, retention time = 4.55 min) were captured with higher cosine value (>0.75) (SI, Figure S12). The intensity of dehydrodicaffeic acid dilactone (m/z = 357.0603, retention time = 4.80 min) was 6.95×10^{07} in spiked samples, whereas it was 2.95×10^{05} (closer to background level) in the CA only sample suggesting the oxidation of CA by Fe(III). Although it has been reported that these compounds can originate from CA, the exact generation pathways from the reactions between CA and $FeCl_3$ are unknown.[47,48] A similar case was observed for 4-(3,4-dihydroxyphenyl)-6,7-dihydroxynaphthalene-2-carboxylic acid (intensity of CA only sample is 3.88×10^{05}). In the case of the other three model compounds, library hits were very limited. CMA was found in CNA samples as a library hit with a retention time of similar length (3.37 min). The generation of CMA as a result of hydroxylation of CNA is well-known.[49] For additional compounds captured from the library search, their possible Fe complexes based on the configuration determined for the known ligands were also searched. The complex of dehydrodicaffeic acid dilactone as a configuration of $[(C_{18}H_{12}O_8)_2Fe]^-$ was observed, eluting at the same retention time (4.80 min) as its parent compound (SI, Figure S13). ^{54}Fe isotope of the complex was not observed, likely due to its potential low intensity (4.55×10^{05}), which is a limiting factor when using isotopologues for screening. Although 4-(3,4-dihydroxyphenyl)-6,7-dihydroxynaphthalene-2-carboxylic acid also complexed with Fe as $[(C_{17}H_{10}O_6)_2Fe]^-$, no complexes were detected for the library-matched compound of 1,5-dihydroxy-2,3-dimethoxyxanthone in the $FeCl_3$-spiked VNL sample. For complex environmental samples with unknown ligands, such annotation and library matching will enable the identification of the presence of compounds with reported MS.

Networking Analysis for Additional Complexes and Ligands. Networking analysis performed using GNPS uncovered additional complexes and ligands as well as related compounds or their transformation products. Clusters were built based on the MS^1 characteristics and MS/MS similarities among the apo-ligands and their corresponding complexes.[38] All of the networks in both + ve and -ve modes can be accessed using the links provided in the SI (SI, Table S6).

The overall IIMN (Figure 3A,B), with linkage based on the mass difference of nodes, of VNL showed multiple clusters, with one including the identified complexes ($[(C_8H_7O_3)_2FeCl_2]^-$, $[(C_8H_7O_3)_3FeCl]^-$, and $[(C_8H_7O_3)_4Fe]^-$). Based on the mass difference, the molecular formulas of connected nodes were annotated. For instance, the mass difference of 116.0709 ($\Delta m/z$ between VNL^- and Cl^-) between node n with m/z = 427.9522 and m with m/z = 544.0231 nodes indicated the introduction of another VNL in the complex, and another VNL could be added for the node l with m/z = 660.0939. Using such rules with Cl, COOH (from formic acid), and OH as possible ligands, all of the nodes were annotated with the molecular formula of $[(C_8H_7O_3)_x\text{Fe}(COOH)_y(OH)_zCl_w]^-$, where w, x, y, and z can be 0−4, with a mass error of <2 ppm. The feature for all nodes was eluted at the same retention time of 3.36 min, as the parent ligand of VNL and identified complexes, when the MS^1 intensities of the annotated nodes were relatively low for many peaks; therefore, $^{54/56}Fe$ and/or $^{35/37}Cl$ isotopologue screening did not identify these features. The retention time of these features, their absence in the sample of ligand only, and only presence for the corresponding ligand mixture with $FeCl_3$, as well as the accurate determination of their molecular formula, served as solid evidence for the presence of these complexes with VNL. Similarly, additional complexes of $[L_x\text{Fe}(COOH)_y(OH)_zCl_w]^-$ (L indicates ligands) were also captured for CA, CMA, and CNA (SI, Table S7). As portrayed in Figure 3C,D, additional complexes of CMA were also unveiled (nodes a−d). The clusters containing complexes can be more complicated with a copious number of nodes, as shown in the network of CA in SI (**SI**, Figure S14). In addition, for CA, dimer and complexes with dimer were also detected. For environmental samples with unknown ligands, such networking analysis can provide supplemental evidence for the detection of target complexes, when the isotopologue analysis can identify the top species of complexes and ligands. For the next step, native metabolomics making uses of postcolumn pH adjustment and metal infusion can be a very useful tool to identify and connect apo-ligands and their metal-bond species in complex environmental samples, where most of the nodes may not be explained based on known available standards. Aron et al.[14] have successfully demonstrated that IIMN can be applied to identify apo-ligand and complexes in ultracomplex surface ocean samples. They identified and verified apo-ligand domoic acid with authentic standards and found its dimer complex $[2M + Cu^{2+} - H]^+$ using native metabolomics and IIMN. Overall molecular networking has been proven to be an effective way to predict from the association of siderophores with metal to annotate the structure of unknown metabolites.[50−52]

Environmental Implication. This study developed a protocol for identifying unknown complexes and ligands, as well as predicting their molecular structures. It has been evident that the isotopologue pattern is a useful tool for

https://doi.org/10.1021/acs.est.4c03608
Environ. Sci. Technol. 2024, 58, 15090−15099

screening the Fe-binding compounds and can be utilized to search for organic ligands in complex environmental samples when the structures of complexes can be well annotated based on MS[1] and MS/MS data (SI, Figure S15). When the undeniable compound validation is limited to the available authentic standards, IIMN analysis can recognize a range of additional complexes and ligands, including derivatives, dimers, and possible transformation products of identified ligands and complexes. In addition to the successful development of this protocol, the molecular formula and structure determined for the model compounds representing degradation products of lignin built up a crucial database and a roadmap for searching for important lignin-derived compounds in complex environmental samples.

The developed methods can be used to search for organic ligands in complex soil environments and comprehensively understand the impact of climate change. As association with minerals plays an essential role in the stability of organic carbon and its response to climate change, organic ligands, with the potential to strongly bind with minerals, are key for the stability of organic carbon. Our study aimed to develop a protocol that can be applied to complex environmental samples, including but not limited to actual degraded lignin samples and various soil samples such as permafrost soil and wildfire-affected soil. Additionally, the molecular formula and stoichiometry of the model complexes laid a solid foundation for identifying potential ligands and corresponding complexes in real complex environmental samples.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c03608.

Additional information regarding LC-Orbitrap IQ-X MS/MS analysis, LC-Orbitrap MS/MS analysis in positive mode, code availability, effect of aqueous chemistry and pH, QA/QC details, isotope pattern figures for $^{54/56}$Fe and $^{35/37}$Cl, m/z error of potential complexes, XIC and fragmentation analysis of additional complexes, mirror plot based on library matching compounds, detailed results of CA dimer compound and complex, IIMN analysis of CA, list of all possible complex configurations, effect of model compound vs $FeCl_3$ concentration, physicochemical properties of model compounds, links to all raw MS data, list of potential complexes in positive mode, list of library-matched compounds, links for IIMN and library searching, and list of additional complexes found in negative mode (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Yu Yang − *Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada 89557, United States;* orcid.org/0000-0002-7568-0202; Email: yuy@unr.edu

Authors

Abrar Shahriar − *Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada 89557, United States; Nuclear and Chemical Sciences Division, Physical and Life Sciences, Lawrence Livermore National Laboratory, Livermore, California 94550, United States*

Srinidhi Lokesh − *Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada 89557, United States*

Anil Timilsina − *Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada 89557, United States;* orcid.org/0000-0001-7053-0662

Travis Numan − *Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada 89557, United States*

Tilman Schramm − *CMFI Cluster of Excellence, University of Tuebingen, 72076 Tuebingen, Germany; Department of Biochemistry, University of California Riverside, Riverside, California 92521, United States*

Paolo Stincone − *CMFI Cluster of Excellence, University of Tuebingen, 72076 Tuebingen, Germany;* orcid.org/0000-0002-2214-6655

Laurinda Nyarko − *College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, United States; Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431, United States*

Christian Dewey − *College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, United States; Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431, United States*

Daniel Petras − *CMFI Cluster of Excellence, University of Tuebingen, 72076 Tuebingen, Germany; Department of Biochemistry, University of California Riverside, Riverside, California 92521, United States*

Rene Boiteau − *College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, United States; Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431, United States;* orcid.org/0000-0002-4127-4417

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.4c03608

Notes

The authors declare no competing financial interest.

[##]Department of Soil & Crop Sciences, Colorado State University, Fort Collins, Colorado 80523-1101, United States.

■ ACKNOWLEDGMENTS

This study was financially supported by the National Science Foundation (2108270). Yang was supported by Alexander von Humboldt Research Fellowship for Experienced Researchers. The authors are also grateful to the help of Dr. Ruprecht and Dr. Spain at the Shared Facility in the Department of Chemistry at University of Nevada, Reno. TOC artwork and Figure 1, include icons from BioRender (BioRender.com). The authors have the publication license from BioRender.

■ REFERENCES

(1) Lalonde, K.; Mucci, A.; Ouellet, A.; Gélinas, Y. Preservation of Organic Matter in Sediments Promoted by Iron. *Nature* **2012**, *483* (7388), 198−200.

(2) Wang, Y.; Wang, H.; He, J.-S.; Feng, X. Iron-Mediated Soil Carbon Response to Water-Table Decline in an Alpine Wetland. *Nat. Commun.* **2017**, *8* (1), No. 15972.

(3) Huang, X.; Liu, X.; Liu, J.; Chen, H. Iron-Bound Organic Carbon and Their Determinants in Peatlands of China. *Geoderma* **2021**, *391*, No. 114974.
(4) Patzner, M. S.; Mueller, C. W.; Malusova, M.; Baur, M.; Nikeleit, V.; Scholten, T.; Hoeschen, C.; Byrne, J. M.; Borch, T.; Kappler, A.; Bryce, C. Iron Mineral Dissolution Releases Iron and Associated Organic Carbon during Permafrost Thaw. *Nat. Commun.* **2020**, *11* (1), No. 6329.
(5) Riedel, T.; Biester, H.; Dittmar, T. Molecular Fractionation of Dissolved Organic Matter with Metal Salts. *Environ. Sci. Technol.* **2012**, *46* (8), 4419−4426.
(6) Jones, D. L.; Edwards, A. C. Influence of Sorption on the Biological Utilization of Two Simple Carbon Substrates. *Soil Biol. Biochem.* **1998**, *30* (14), 1895−1902.
(7) Sowers, T. D.; Wani, R. P.; Coward, E. K.; Fischel, M. H. H.; Betts, A. R.; Douglas, T. A.; Duckworth, O. W.; Sparks, D. L. Spatially Resolved Organomineral Interactions across a Permafrost Chronosequence. *Environ. Sci. Technol.* **2020**, *54* (5), 2951−2960.
(8) Patzner, M. S.; Kainz, N.; Lundin, E.; Barczok, M.; Smith, C.; Herndon, E.; Kinsman-Costello, L.; Fischer, S.; Straub, D.; Kleindienst, S.; Kappler, A.; Bryce, C. Seasonal Fluctuations in Iron Cycling in Thawing Permafrost Peatlands. *Environ. Sci. Technol.* **2022**, *56* (7), 4620−4631.
(9) Adhikari, D.; Poulson, S. R.; Sumaila, S.; Dynes, J. J.; McBeth, J. M.; Yang, Y. Asynchronous Reductive Release of Iron and Organic Carbon from Hematite−Humic Acid Complexes. *Chem. Geol.* **2016**, *430*, 13−20.
(10) Adhikari, D.; Zhao, Q.; Das, K.; Mejia, J.; Huang, R.; Wang, X.; Poulson, S. R.; Tang, Y.; Roden, E. E.; Yang, Y. Dynamics of Ferrihydrite-Bound Organic Carbon during Microbial Fe Reduction. *Geochim. Cosmochim. Acta* **2017**, *212*, 221−233.
(11) Boiteau, R. M.; Fansler, S. J.; Farris, Y.; Shaw, J. B.; Koppenaal, D. W.; Pasa-Tolic, L.; Jansson, J. K. Siderophore Profiling of Co-Habitating Soil Bacteria by Ultra-High Resolution Mass Spectrometry. *Metallomics* **2019**, *11* (1), 166−175.
(12) Bahureksa, W.; Tfaily, M. M.; Boiteau, R. M.; Young, R. B.; Logan, M. N.; McKenna, A. M.; Borch, T. Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation. *Environ. Sci. Technol.* **2021**, *55* (14), 9637−9656.
(13) Hodgkins, S. B.; Tfaily, M. M.; McCalley, C. K.; Logan, T. A.; Crill, P. M.; Saleska, S. R.; Rich, V. I.; Chanton, J. P. Changes in Peat Chemistry Associated with Permafrost Thaw Increase Greenhouse Gas Production. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (16), 5819−5824.
(14) Aron, A. T.; Petras, D.; Schmid, R.; Gauglitz, J. M.; Büttel, I.; Antelo, L.; Zhi, H.; Nuccio, S.-P.; Saak, C. C.; Malarney, K. P.; et al. Native Mass Spectrometry-Based Metabolomics Identifies Metal-Binding Compounds. *Nat. Chem.* **2022**, *14* (1), 100−109.
(15) Walker, L. R.; Tfaily, M. M.; Shaw, J. B.; Hess, N. J.; Paša-Tolić, L.; Koppenaal, D. W. Unambiguous Identification and Discovery of Bacterial Siderophores by Direct Injection 21 T Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Metallomics* **2017**, *9* (1), 82−92.
(16) Boiteau, R. M.; Repeta, D. J. An Extended Siderophore Suite from Synechococcus Sp. PCC 7002 Revealed by LC-ICPMS-ESIMS. *Metallomics* **2015**, *7* (5), 877−884.
(17) Boiteau, R. M.; Repeta, D. J. An Extended Siderophore Suite from Synechococcus Sp. PCC 7002 Revealed by LC-ICPMS-ESIMS. *Metallomics* **2015**, *7* (5), 877−884.
(18) Chen, H.; Yang, Z.; Chu, R. K.; Tolic, N.; Liang, L.; Graham, D. E.; Wullschleger, S. D.; Gu, B. Molecular Insights into Arctic Soil Organic Matter Degradation under Warming. *Environ. Sci. Technol.* **2018**, *52* (8), 4555−4564.
(19) Mann, B. F.; Chen, H.; Herndon, E. M.; Chu, R. K.; Tolic, N.; Portier, E. F.; Chowdhury, T. R.; Robinson, E. W.; Callister, S. J.; Wullschleger, S. D.; Graham, D. E.; Liang, L.; Gu, B. Indexing Permafrost Soil Organic Matter Degradation Using High-Resolution Mass Spectrometry. *PLoS One* **2015**, *10* (6), No. e0130557.
(20) Beck, J.; Brüggemann, M.; van Pinxteren, D.; Herrmann, H. Nontarget Approach to Identify Complexing Agents in Atmospheric Aerosol and Rainwater Samples. *Anal. Chem.* **2022**, *94* (25), 8966−8974.
(21) Tapparo, A.; Di Marco, V.; Badocco, D.; D'Aronco, S.; Soldà, L.; Pastore, P.; Mahon, B. M.; Kalberer, M.; Giorio, C. Formation of Metal-Organic Ligand Complexes Affects Solubility of Metals in Airborne Particles at an Urban Site in the Po Valley. *Chemosphere* **2020**, *241*, No. 125025.
(22) Bugg, T. D.; Ahmad, M.; Hardiman, E. M.; Singh, R. The Emerging Role for Bacteria in Lignin Degradation and Bio-Product Formation. *Curr. Opin. Biotechnol.* **2011**, *22* (3), 394−400.
(23) Filley, T. R.; Cody, G. D.; Goodell, B.; Jellison, J.; Noser, C.; Ostrofsky, A. Lignin Demethylation and Polysaccharide Decomposition in Spruce Sapwood Degraded by Brown Rot Fungi. *Org. Geochem.* **2002**, *33* (2), 111−124.
(24) Kögel-Knabner, I. The Macromolecular Organic Composition of Plant and Microbial Residues as Inputs to Soil Organic Matter: Fourteen Years On. *Soil Biol. Biochem.* **2017**, *105*, A3−A8.
(25) Thevenot, M.; Dignac, M.-F.; Rumpel, C. Fate of Lignins in Soils: A Review. *Soil Biol. Biochem.* **2010**, *42* (8), 1200−1211.
(26) Huang, W.; Hammel, K. E.; Hao, J.; Thompson, A.; Timokhin, V. I.; Hall, S. J. Enrichment of Lignin-Derived Carbon in Mineral-Associated Soil Organic Matter. *Environ. Sci. Technol.* **2019**, *53* (13), 7522−7531.
(27) Asina, F.; Brzonova, I.; Voeller, K.; Kozliak, E.; Kubátová, A.; Yao, B.; Ji, Y. Biodegradation of Lignin by Fungi, Bacteria and Laccases. *Bioresour. Technol.* **2016**, *220*, 414−424.
(28) Xu, Z.; Qin, L.; Cai, M.; Hua, W.; Jin, M. Biodegradation of Kraft Lignin by Newly Isolated Klebsiella Pneumoniae, Pseudomonas Putida, and Ochrobactrum Tritici Strains. *Environ. Sci. Pollut. Res.* **2018**, *25* (14), 14171−14181.
(29) Kappler, A.; Straub, K. L. Geomicrobiological Cycling of Iron. *Rev. Mineral. Geochem.* **2005**, *59* (1), 85−108.
(30) Geilfus, C.-M. Chloride in Soil: From Nutrient to Soil Pollutant. *Environ. Exp. Bot.* **2019**, *157*, 299−309.
(31) Poletaeva, V. I.; Tirskikh, E. N.; Pastukhov, M. V. Hydrochemistry of Sediment Pore Water in the Bratsk Reservoir (Baikal Region, Russia). *Sci. Rep.* **2021**, *11* (1), No. 11124.
(32) Stincone, P.; Pakkir Shah, A. K.; Schmid, R.; Graves, L. G.; Lambidis, S. P.; Torres, R. R.; Xia, S.-N.; Minda, V.; Aron, A. T.; Wang, M.; Hughes, C. C.; Petras, D. Evaluation of Data-Dependent MS/MS Acquisition Parameters for Non-Targeted Metabolomics and Molecular Networking of Environmental Samples: Focus on the Q Exactive Platform. *Anal. Chem.* **2023**, *95* (34), 12673−12682.
(33) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24* (21), 2534−2536.
(34) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78* (3), 779−787.
(35) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Dührkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdžija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L.-F.; Wang, M.; Orešič, M.; Weng, J.-K.; Böcker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T. Integrative Analysis of Multimodal Mass Spectrometry Data in MZmine 3. *Nat. Biotechnol.* **2023**, *41* (4), 447−449.
(36) Myers, O. D.; Sumner, S. J.; Li, S.; Barnes, S.; Du, X. One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and

Detecting Chromatographic Peaks. *Anal. Chem.* **2017**, *89* (17), 8696−8703.
(37) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299−302.
(38) Schmid, R.; Petras, D.; Nothias, L.-F.; Wang, M.; Aron, A. T.; Jagels, A.; Tsugawa, H.; Rainer, J.; Garcia-Aloy, M.; Dührkop, K.; et al. Ion Identity Molecular Networking for Mass Spectrometry-Based Metabolomics in the GNPS Environment. *Nat. Commun.* **2021**, *12* (1), No. 3832.
(39) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* (11), 2498−2504.
(40) Waska, H.; Koschinsky, A.; Dittmar, T. Fe-and Cu-Complex Formation with Artificial Ligands Investigated by Ultra-High Resolution Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS): Implications for Natural Metal-Organic Complex Studies. *Front. Mar. Sci.* **2016**, *3*, 119.
(41) Koch, B. P.; Dittmar, T.; Witt, M.; Kattner, G. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. *Anal. Chem.* **2007**, *79* (4), 1758−1763.
(42) Kind, T.; Fiehn, O. Seven Golden Rules for Heuristic Filtering of Molecular Formulas Obtained by Accurate Mass Spectrometry. *BMC Bioinf.* **2007**, *8* (1), 105.
(43) Kontoghiorghe, C. N.; Kolnagou, A.; Kontoghiorghes, G. J. Phytochelators Intended for Clinical Use in Iron Overload, Other Diseases of Iron Imbalance and Free Radical Pathology. *Molecules* **2015**, *20* (11), 20841−20872.
(44) Yuen, A. K. L.; A Hutton, G.; F Masters, A.; Maschmeyer, T. The Interplay of Catechol Ligands with Nanoparticulate Iron Oxides. *Dalton Trans.* **2012**, *41* (9), 2545−2559.
(45) Huber, R.; Marcourt, L.; Koval, A.; Schnee, S.; Righi, D.; Michellod, E.; Katanaev, V. L.; Wolfender, J.-L.; Gindro, K.; Queiroz, E. F. Chemoenzymatic Synthesis of Complex Phenylpropanoid Derivatives by the Botrytis Cinerea Secretome and Evaluation of Their Wnt Inhibition Activity *Front. Plant Sci.* 2022, *12*, DOI: 10.3389/fpls.2021.805610.
(46) Kumada, Y.; Takeuchi, T.; Umezawa, H. Characterization of the Dehydrodicaffeic Acid Dilactone-Forming Enzyme and the Enzymic and Chemical Synthesis of This Mushroom Product. *Agric. Biol. Chem.* **1977**, *41* (5), 877−885.
(47) Tinikul, R.; Chenprakhon, P.; Maenpuen, S.; Chaiyen, P. Biotransformation of Plant-Derived Phenolic Acids. *Biotechnol. J.* **2018**, *13* (6), No. 1700632.
(48) Alson, S. G.; Jansen, O.; Cieckiewicz, E.; Rakotoarimanana, H.; Rafatro, H.; Degotte, G.; Francotte, P.; Frederich, M. In-Vitro and in-Vivo Antimalarial Activity of Caffeic Acid and Some of Its Derivatives. *J. Pharm. Pharmacol.* **2018**, *70* (10), 1349−1356.
(49) Chen, H.-C.; Li, Q.; Shuford, C. M.; Liu, J.; Muddiman, D. C.; Sederoff, R. R.; Chiang, V. L. Membrane Protein Complexes Catalyze Both 4- and 3-Hydroxylation of Cinnamic Acid Derivatives in Monolignol Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (52), 21253−21258.
(50) Behnsen, J.; Zhi, H.; Aron, A. T.; Subramanian, V.; Santus, W.; Lee, M. H.; Gerner, R. R.; Petras, D.; Liu, J. Z.; Green, K. D.; Price, S. L.; Camacho, J.; Hillman, H.; Tjokrosurjo, J.; Montaldo, N. P.; Hoover, E. M.; Treacy-Abarca, S.; Gilston, B. A.; Skaar, E. P.; Chazin, W. J.; Garneau-Tsodikova, S.; Lawrenz, M. B.; Perry, R. D.; Nuccio, S.-P.; Dorrestein, P. C.; Raffatellu, M. Siderophore-Mediated Zinc Acquisition Enhances Enterobacterial Colonization of the Inflamed Gut. *Nat. Commun.* **2021**, *12* (1), No. 7016.
(51) Reher, R.; Aron, A. T.; Fajtová, P.; Stincone, P.; Wagner, B.; Pérez-Lorente, A. I.; Liu, C.; Shalom, I. Y. B.; Bittremieux, W.; Wang, M.; Jeong, K.; Matos-Hernandez, M. L.; Alexander, K. L.; Caro-Diaz, E. J.; Naman, C. B.; Scanlan, J. H. W.; Hochban, P. M. M.; Diederich, W. E.; Molina-Santiago, C.; Romero, D.; Selim, K. A.; Sass, P.; Brötz-Oesterhelt, H.; Hughes, C. C.; Dorrestein, P. C.; O'Donoghue, A. J.; Gerwick, W. H.; Petras, D. Native Metabolomics Identifies the Rivulariapeptolide Family of Protease Inhibitors. *Nat. Commun.* **2022**, *13* (1), No. 4619.
(52) Zhou, Z.; Luo, M.; Zhang, H.; Yin, Y.; Cai, Y.; Zhu, Z.-J. Metabolite Annotation from Knowns to Unknowns through Knowledge-Guided Multi-Layer Metabolic Networking. *Nat. Commun.* **2022**, *13* (1), No. 6656.