Refinement of compound aromaticity in complex organic mixtures by stable isotope label assisted ultra-high resolution mass spectrometry

Alexander Zherebker*[a], Oliver J. Lechtenfeld[b], Anastasia Sarycheva[a], Yury Kostyukevich[a], Oleg Kharybin[a], Elena I. Fedoros[c,d], Evgeny N. Nikolaev*[a]

[a] Skolkovo Institute of Science and Technology, 143025 Skolkovo, Moscow region, Russia. E-mail: a.zherebker@skoltech.ru, e.nikolaev@skoltech.ru

[b] Helmholtz Centre for Environmental Research - UFZ, Department of Analytical Chemistry, Leipzig, Germany, DE-04318

[c] N.N. Petrov National Medical Research Center of Oncology, Saint-Petersburg, Russia, 197758

[d] Nobel LTD, Saint-Petersburg, Russia, 192012

Abstract

Fourier transform ion cyclotron resonance mass-spectrometry (FTICR MS) provides a unique opportunity for molecular analysis of natural complex mixtures. In many geochemical and environmental studies structure-properties relations are based solely on the elemental compositional information. Several calculated parameters were proposed to increase reliability of structural attribution, among which aromaticity indices (AI and AI_{mod}) are widely used. Herein, we applied a combination of selective labeling reactions in order to obtain direct structural information on the individual components of lignin-derived polyphenolic material. Carboxylic (COOH), carbonyl (C=O) and hydroxyl (OH) groups were enumerated by esterification, reducing and acetylation reactions, respectively, followed by FTICR MS analyses. Obtained information enabled to constrain aromaticity accounting for carbon skeleton only. We found that actual aromaticity of components may be both higher or lower than approximated values depending on the abundance of COOH, C=O and OH groups. The results are of importance for geochemical community studying terrestrial NOM with structural gradients.

Keywords isotopic labeling, FTICR MS, molecular structure, aromaticity index, complex mixtures, polyphenols

INTRODUCTION

Nowadays, Fourier transform ion cyclotron resonance mass-spectrometry (FTICR MS) is widely used in environmental and chemical studies of natural polyphenolic and oxy-acids mixtures such as humic substances (HS)[1] or natural organic matter.[2–4] FTICR MS routinely resolves thousands of molecular compositions in a single sample.[5] The necessity of thorough molecular analysis of such mixtures is justified by the scientific request for a description of their biological activity, transformation pathways, fate and overall environmental role. The major limitation of FTICR MS to study NOM is a lack of direct structural information. Due to the extreme structural complexity of NOM tandem mass-spectrometry analysis are challenging and not routinely applied to date.[6] Most studies instead focus on direct infusion (DI) analysis; and individual mass-to-charge ratios detected by DI-FTICR MS may correspond to a large number of structural isomers.[7] In case of known boundary conditions for component structures it is possible to perform *in silico* algorithms to search for molecular components in publicly available databases of natural metabolites[8] or biological active compounds.[9] However, in case of HS or NOM samples, which have undergone extensive microbial or geochemical transformations for tens to thousands of years, this approach is highly questionable. Therefore, geochemical researches report interconnection of environmental properties with particular molecular ions in NOM with probability-based suggestions for structural attribution.[10]

Several molecular formula-based approaches were proposed to relate the elemental composition of molecular formulas to structural features. Such approaches can be divided into compositional correlation ("*Similia similibus*") and structural plausibility constraints. In the first case, components of NOM or HS are attributed to major (macromolecular) precursor units based only on similar H/C and O/C atomic ratios (e.g. lignins, tannins, carbohydrates, etc.).[11–13] Changes in the molecular composition and chemical properties of NOM or HS are then

explained by variations and reactivity of the respective structural precursor species, although it is known that biogenic precursors are subject to substantial diagenetic alteration.[14,15]

Chemical constraints of valences and oxidation state of contributing elements, on the other hand, also limit the plausibility of the molecular formula space. A particularly useful example is the estimation of component aromaticity, which is related to the unsaturation state of molecules. For highly aromatic samples (e.g. coal, petroleum, etc.), aromaticity is connected to double bond equivalent (DBE) and DBE/C ratio where e.g. condensed aromatic structures require DBE/C $\geq$ 0.7.[16] The introduction of aromaticity index (AI) and modified aromaticity index (AI_{mod}) accounted for functional groups found in HS and NOM that contribute to DBE, but not to the aromaticity of the carbon skeleton.[17,18] AI and AI_{mod} assume that all oxygen atom and half of oxygen atoms form double bonds with sp^2-hybridized carbon, respectively. Hence AI is the most conservative aromatic system approximation. Despite known limitations of these indices (especially AI_{mod}), they currently serve as major structural parameters in many geochemical studies of NOM/HS[19–21]. The use of AI may thus lead to false conclusions about aromatic character of molecules for structurally constrained samples. For example, we have recently shown that molecules with the same low AI value (e.g. 0.3) may correspond to isomers with and without aromatic rings depending on the geochemical origin.[22]

In order to obtain direct structural information, we have previously proposed a combination of FTICR MS with isotopic labeling reactions, which enable to enumerate particular classes of oxygenated functional groups[23] or carbon skeleton fragments[24] of individual components in complex mixtures. Selective labelling techniques for the reduction of ketones and quinones in humic substances of dissolved organic matter (DOM) samples were also applied by Baluha et al.[25] However, there is a lack of studies devoted to a combination of labeling techniques coupled to FTICR MS.

The objective of this study was to explore major structural moieties and to use this information to constrain aromaticity of individual components of lignin-derived polyphenol mixture – BP-Cx-1,

which has been described in our previous study: this material possesses similar optical properties to natural organic matter and exhibits high inhibitory activity against a wide range of ferments[26]. To reach the main goal, enumeration of different oxygen functional groups was performed for the first time using selective incorporation of deuterium by esterification, reducing and acetylation reactions followed by FTICR MS analysis of labeled material.

MATERIALS AND METHODS

Solvents and other reagents used in this study were commercially available. Methanol of HPLC grade (Lab-Scan) was used for elution and dissolution of BP-Cx-1 components. High-purity distilled water (18.2 MΩ) was prepared using a Millipore Simplicity 185 system. D-enrichment of deuterated methanol (CD_3OD), sodium borodeuteride ($NaBD_4$) and acetyl chloride-d_3 (CD_3COCl) were 99.8%, 98% and 99%, respectively (Sigma). Bond Elut PPL (Priority PolLutant, Agilent Technologies) cartridges (50 mg, 3 mL) were used for isolation and purification of the parent and the labeled samples. PPL represents a modified styrene-divinylbenzene polymer designed for polar organic compounds extraction. Parent BP-Cx-1 was provided by Nobel Ltd as a sterile 0.42% ammonia solution (batch X112K14A1) described elsewhere.[26] The carbon distribution obtained by qualitative ^{13}C NMR is provided in Supporting Information (Table S1). All labeled samples were purified using solid-phase extraction (SPE) from aqueous solution on PPL cartridges according to the procedure described for DOM samples.[27]

Labeling procedures

Carboxylic groups in the parent sample were selectively deuteromethylated following the previously developed method.[28] Briefly: $SOCl_2$ (60 μL) was added dropwise to the solution of 0.5 mg of BP-Cx-1 fractions in 1.5 mL of CD_3OD under continued stirring and ice-cooling. The reaction mixture was then refluxed for 4 h and dried under vacuum.

Acetylation was performed by a modified procedure[29]: a drop of conc. H_2SO_4 was added to a mixture of dried BP-Cx-1 (0.5 mg) and 1.5 ml of CD_3COCl in 10-ml round-bottomed flask. Further mixture was refluxed under stirring on a boiling water bath for 4 hours. The content was poured into 5 ml of ice-water and vigorously stirred to ensure the hydrolysis of unreacted acetyl chloride. The organic material was extracted on PPL followed by methylation using $CH_3OH/SOCl_2$ procedure. As acetyl chloride leads also to carboxylic and alcohol groups acetylation, a second step – methylation with CH_3OH – was performed, which led to trans-esterification of carboxylic groups and acidic hydrolysis of non-conjugated and aliphatic esters due to *in situ* generation of HCl, while the phenyl esters remain intact.[30]

Reducing was performed by a modified procedure for DOM.[31] Under ice-cooling 0.1M NaOD solution in D_2O was added to $NaBD_4$ (5 mg) until full dissolving and the solution was purged by Ar for oxygen removal. Similarly, the solution of BP-Cx-1 sample (1 mg) in D_2O was prepared. Next, sample solution was added drop-wise to $NaBD_4$ solution under intense stirring. The stirring was maintained for 3 hours under continuous Ar flow. Further, mixture was transferred to a beaker and acidified with 1M HCl under ice-cooling until gassing stopped.

Determination of labeling series by FTICR mass spectrometry and data treatment

Detailed information on FTICR MS analyzes is provided in the Supporting Information. Functional groups were enumerated by juxtaposition of FTICR mass-spectra of labeled and parent samples and by extraction of peak series with mass differences corresponding to the particular labeling procedure as it was previously described by us for the H/D exchange[32]. Here, these series are produced by peaks with the m/z difference of 17.03448, 3.02193 and 45.029396 Da, respectively. In case of $NaBD_4$ reducing [HD] series were examined instead of $[D]_2$ because the produced –OD alcohol groups are back exchanged to -OH during subsequent sample treatment. Labeling series were manually determined for 200 most abundant ions in the parent BP-Cx-1 material from the extracted mass spectra

fragments with designated peak series as it has been described previously for H/D exchange of DOM[32]. The examples of the corresponding peaks series are represented in Fig. S1. Error constraint was set to 0.0005 m/z.

For van Krevelen and Kendrick diagrams, molecular formulas were divided into 5 molecular classes according to atomic ratios and modified aromaticity index (AI_{mod}) as described in Kellerman et al.[33] AI_{mod}, AI, and conventional double bond equivalent (DBE) indices were calculated according to Koch et al.[17] Constrained AI (AI_{cor}) and DBE (DBE_{cor}) values were calculated by subtracting carbon atoms in carboxyl and carbonyl groups from the total carbon number. AI_{mod}, AI, DBE, AI_{cor} and DBE_{cor} were calculated according to equations 1-5, respectively.

$$AI = \frac{1+C-O-0.5H}{C-O} \quad (1) \qquad AI_{mod} = \frac{1+C-0.5O-0.5H}{C-0.5O} \quad (2); \; AI_{cor} = \frac{1+C-(COOH)_n-(C=O)_m-0.5H}{C-(COOH)_n-(C=O)_m} \quad (3)$$

$$DBE_{cor} = 1 + C - 0.5H - (COOH)_n - (C = O)_m \quad (4) \qquad DBE = 1 + C - 0.5H \quad (5)$$

RESULTS AND DISCUSSION

Preliminary sample characterization

The mass spectrum of the parent material was characterized by singly charged ions in mass-window of 200-1000 Da. In total, 1659 formulae were resolved in which 1156 contained only CHO atoms accounting for 68% of the total intensity. The resulting molecular formula data were plotted in the commonly used van Krevelen diagram displaying the H/C vs O/C chemical space (Fig. 1A). As it is seen, nominally aromatic and unsaturated compounds are most abundant in BP-Cx-1, which is typical for lignin derived materials.[34] Kendrick mass defect plots for CH_2-, CO_2- and O-base masses (Fig. 1B and Fig. S2) justify the necessity for deeper structural study.

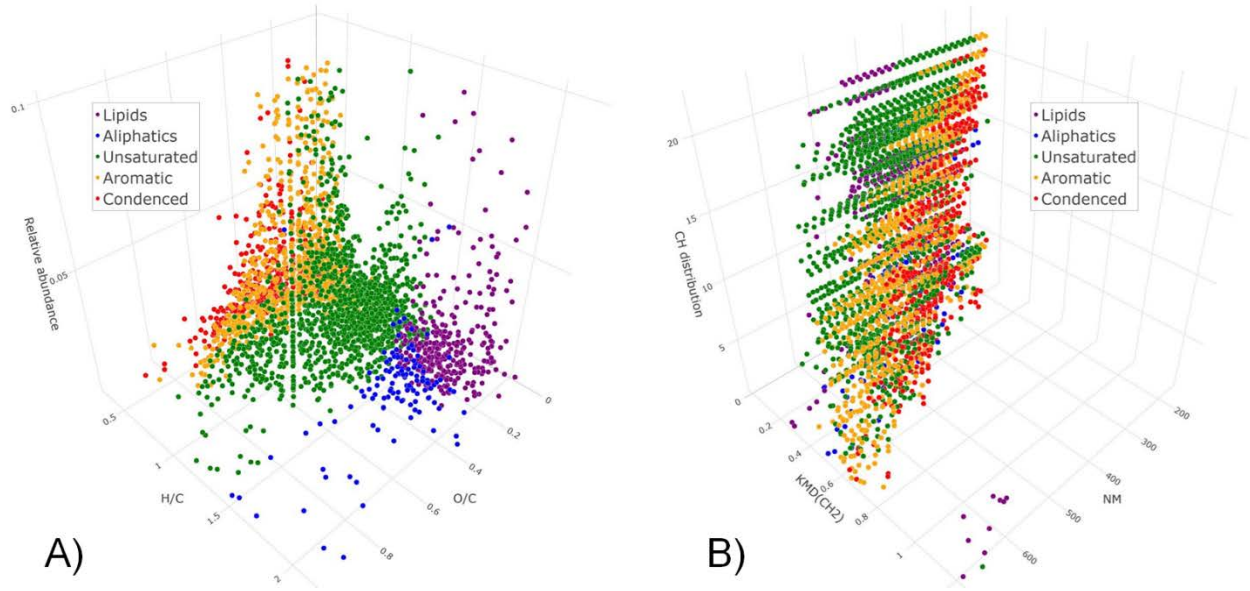

Figure 1. 3D van Krevelen and Kendrick diagrams for BP-Cx-1 with highlighted molecular classes. Z-axis corresponds to the relative intensity (A) and homologous series lengths (B).

Typically, CH_2-homologous series are the most abundant in BP-Cx-1 with a maximum length of 20. These series were observed for all molecular classes including aliphatics and lipids. When using KMD analysis, addition of a methylene fragment (CH_2) is interpreted as chain elongation of a carbon skeleton and should not influence aromatic systems present in a molecule. However, addition of CH_2 decreases the mean compound aromaticity described e.g. by AI_{mod} due to an increase of denominator in the eq. 2. As a result, molecular formulas from the same homologous series can be attributed to different (AI_{mod}-based) molecular classes. Expectedly, this was also observed for O-homologous series, but not for CO_2 series (Fig. S2). These series were abundant mostly in aromatic and unsaturated compounds, which is in agreement with our previous findings that these compounds are enriched with oxygenated functional groups, such as carboxyls and phenols.[35] As it is seen there are two ambiguities coming from the AI and AI_{mod} calculations: the approximation that all oxygen atoms are represented by carbonyl or carboxylic groups, respectively, and by manipulating with mean aromaticity values, which may be confusing for structural interpretation. Therefore, an enumeration of oxygenated functional groups would help to decrease structural ambiguity and add an important

information about functional moieties of individual molecular components. For example, oxygen series may correspond to the appearance of methoxyl and alcohol hydroxyl groups or oxidation of aldehydes. These groups would drastically change chemical properties of KMD-based determined homologous.

Labeling experiments

The reaction scheme for a determination of carboxylic, carbonyl and phenolic groups is summarized in Fig. 2 showing a model structure, which contains the functional groups of interest for this study. The sum of ketone and aldehyde groups were determined by selective reducing with $NaBD_4$ (1) by a modified efficient procedure for DOM leading to deuterium incorporation in C-H bonds in newly formed alcohols.[31]. Carboxylic groups were esterified with CD_3OD (2) following the previously developed method[28], which ensures selectivity of the reaction and tolerates the presence of alcohol groups.[36] Phenolic groups were determined by acetylation with D-labeled acetyl chloride (3) according to the modified procedure[29]. As it is seen from Fig. S1 in cases of all reactions high yields of labeling were achieved.

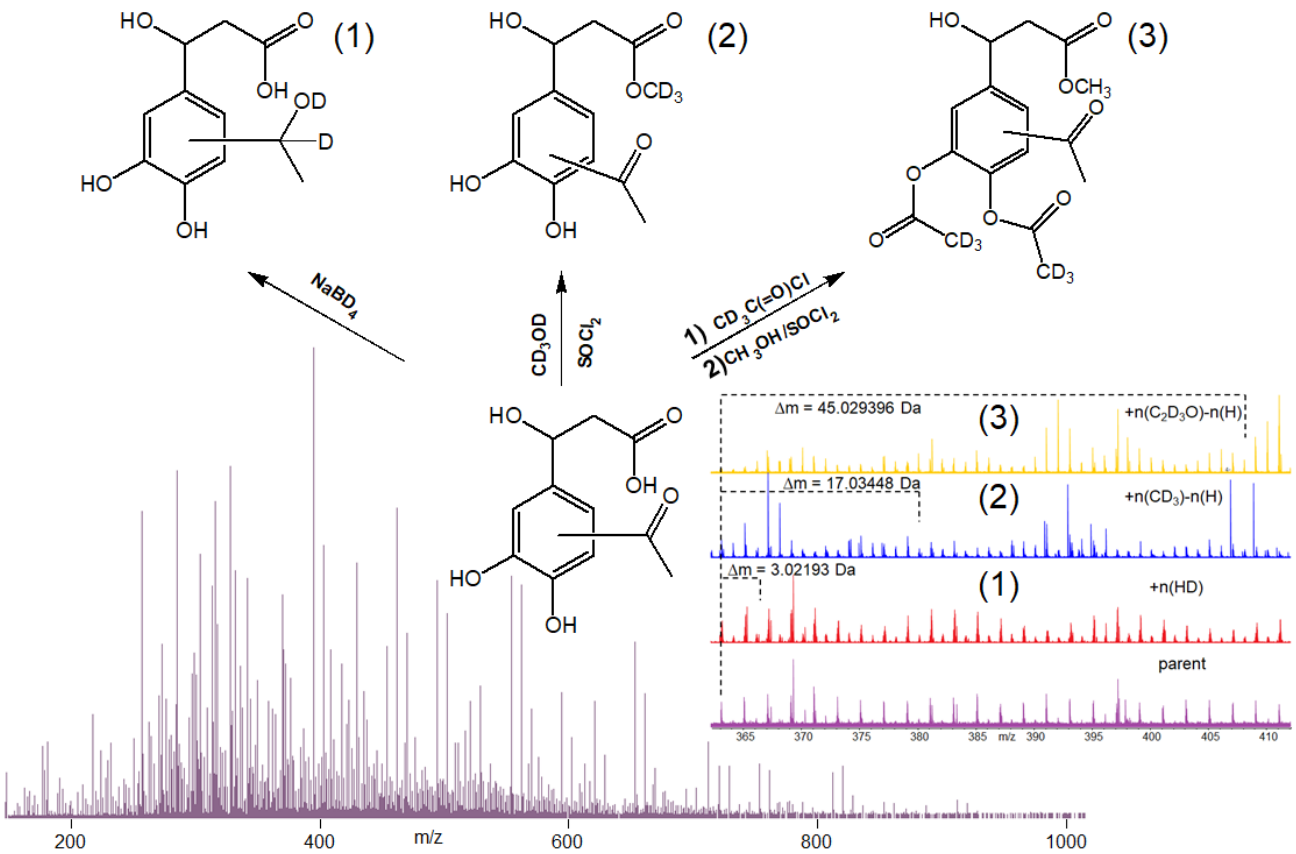

Figure 2. General scheme of functional groups enumeration on the model structure including reaction pathways, FTICR mass-spectra acquisition and extraction of labeling peak series.

Functional groups of interest were enumerated for the 200 most abundant ions in the FTICR mass-spectrum of the parent material and plotted in van Krevelen diagrams. Correspondingly, the number of carboxylic groups increased with the O/C ratio (Fig. 3A). Number of carboxylic groups varied from 0 for most reduced unsaturated and aromatic compounds to 4 for the relatively oxidized compounds with O/C > 0.4 and H/C < 1. Interestingly, the number of carbonyl groups did not exhibit a clear distribution in the van Krevelen space. For all molecular classes the number of each functional group varied from 0 to 3 per molecular formula (Fig. 3B). Previously it was shown the presence of up to two carbonyl groups per molecular component of riverine NOM.[25,37] However, typically riverine NOM contains a lack of low-oxidized aromatic and condensed compounds with O/C < 0.5 and H/C < 1. Moreover, based on the synthetic nature of BP-Cx-1 obtained by oxidation of lignin, carbonyl

groups detected in our work likely corresponded to the quinone structures, which are typical products for the oxidative condensation of lignin residues.[38,39] Further, the number of phenolic groups were mirroring the carboxylic group distribution. Their number decreased from 3 to 0 with an increase of O/C ratio. Similar to carbonyl compounds, species with the maximum numbers of phenolic groups occupied an area in van Krevelen diagram with O/C < 0.4 and H/C < 1, which was characterized by the absence of carboxylic group. Collectively, the absence of carboxyls and the abundance of phenols indicate the presence of flavonoid-like and quercetin-like structures typical for the plant-derived materials.[40]

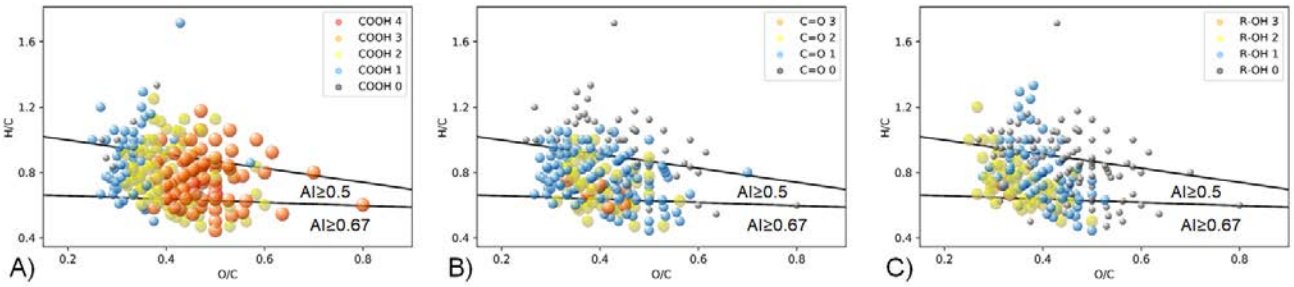

Figure 3. Van Krevelen diagrams for 213 compounds in BP-Cx-1 with highlighted numbers of A) carboxylic, B) carbonyl and C) phenolic groups as determined by isotopic labeling and FTICR MS

Constraining Aromaticity index

Information on the carboxyl and carbonyl groups was further used to constrain compound aromaticity. Determination of functional groups with sp^2-carbon enables to correct DBE and AI values to accurately account for carbon skeleton only. For this purpose, DBE_{cor} and AI_{cor} were calculated taking into account only carbon atoms untied to carboxylic or carbonyl functionality. Fig. 4 (A, B) shows the comparison of conventional and constrained DBE value distributions for BP-Cx-1 components with assigned functional groups. It is clearly seen that the maximum of DBE distribution is 12, which corresponds to up to 3 aromatic rings. At the same time maximum of DBE_{cor} distribution is 8, which corresponds to up to only 2 aromatic rings per molecule. Therefore, carbon unsaturation of compounds based on MS1 analysis may be significantly overestimated. Unlike petroleum in which

DBE values are widely used to indicate the number and type of aromatic rings,[41,42] in case of polyphenols and NOM widely used indices for characterization of compounds aromaticity are AI and AI_{mod}. According to the definition by Koch et al. (2006), AI ≥ 0.5 and AI ≥ 0.67 conservatively indicate the presence of aromatic and condensed aromatic rings in NOM molecules, respectively.[17] Exploration of constrained AI based on actual number of oxygenated functional groups revealed discrepancy between estimated and actual functional group-accounted compound aromaticity. Fig. 4(C) shows van Krevelen diagram with molecular formulas, for which the attribution to molecular classes changed depending on the used aromaticity calculation: AI, AI_{mod} and AI_{cor}. AI considers all oxygen atoms bound in carbonyl groups, and it's value is always smaller than AI_{mod} and AI_{cor}. Number-averaged values for the formulas shown in Fig. 4C were 0.40, 0.55 and 0.57 for AI, AI_{mod} and AI_{cor}, respectively. Nevertheless, only 31% of molecules were always attributed to the same molecular classes in case of all indices. Further, we found that AI_{cor} may be both higher or lower than AI_{mod}. In 53% of cases wherein AI_{mod} differed from AI_{cor} the class attribution remained the same. At the same time 16% of molecules were attributed to different classes, in which for 3% of compounds AI_{cor} was smaller than AI_{mod}, while for 13% of molecules AI_{cor} was larger than AI_{mod}.

Applying AI_{cor} demonstrates that not only calculated aromaticity values were different but the attribution of some compounds needs to be corrected. For example, compounds with H/C < 1.5 and AI < 0.5 are frequently assigned to alicyclic compounds like carboxyl-rich alicyclic molecules CRAM[43] or CCAM[44]. Application of carboxyl and carbonyl groups enumeration enabled to re-assigned several species to the aromatic-ring containing compounds. Moreover, relative contributions of aromatic and condensed aromatic systems were also different for aromaticity indices Fig. 4(D). Using of the most conservative AI for the BP-Cx-1 sample resulted in only few compounds attributed to condensed aromatics. Further, carboxyl-approximated AI_{mod} showed similar results to AI_{cor},

however, using AI_{cor} enabled to reassign 10% of "non-aromatic" compounds to aromatic and condensed molecules (Fig. 4D, S3).

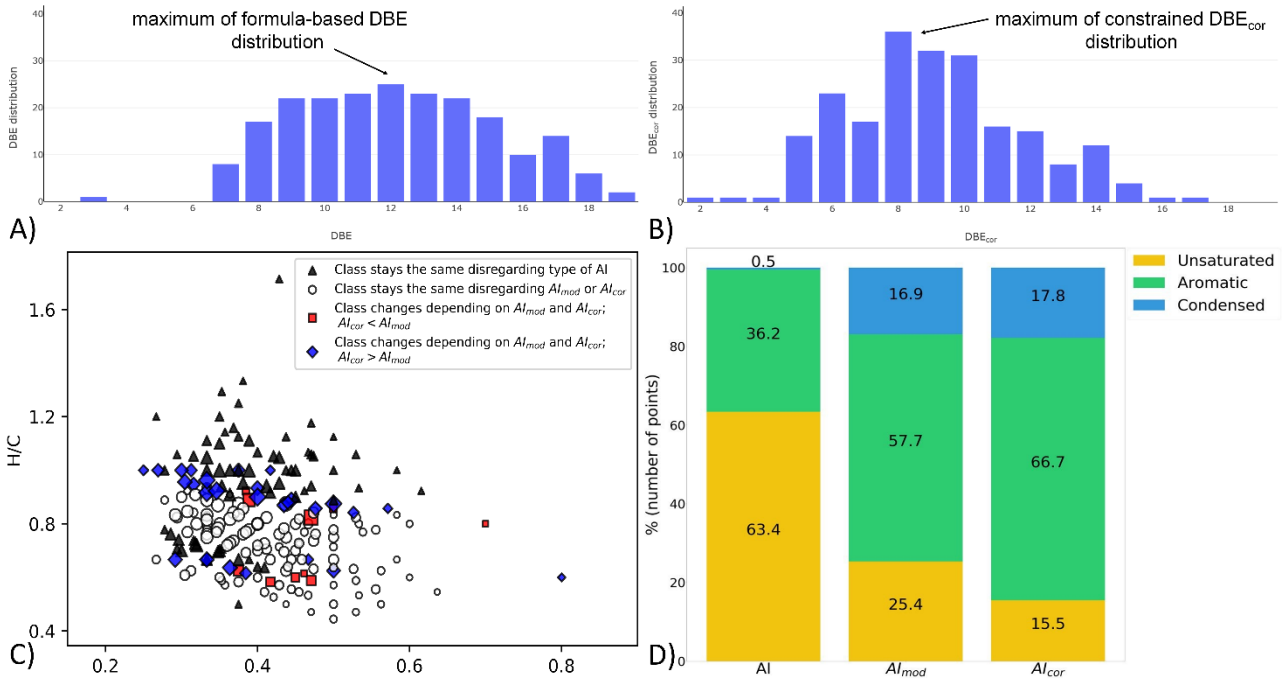

Figure 4. Color-coded A, B) DBE and DBE_{mod} distributions and C) Van Krevelen diagram for BP-Cx-1 components with the assigned components regarding the consistency of molecular class attribution depending of the aromaticity indices: AI, AI_{mod} and AI_{cor}. D) Relative contribution of three molecular classes calculated according to different aromaticity indices.

CONCLUSIONS

Application of selective labeling reaction and FTICR MS to the complex polyphenolic mixture enabled to directly enumerate carboxyl, carbonyl and hydroxyl functional groups in individual components and to map them in van Krevelen diagram. Added value of the obtained data is the possibility to exclude sp^2-hybridized carbon atoms bonded to oxygen from the consideration of DBE or AI parameters. We found that actual aromaticity of components may be both higher or lower than approximated by widely used AI_{mod} values. This resulted from the approximation of AI and AI_{mod}, that all oxygen atoms are represented by carbonyl and carboxyl groups, respectively, while oxygen in

polyphenols may be represented by alcohols and ether groups only. Therefore, the use of AI (as most conservative aromaticity descriptor) may result in significant underestimation of a sample's mean aromaticity and false attribution to molecular classes. While AI_{mod} is a good approximation for carboxyl-rich compounds predominantly found in marine NOM,[39] it underrepresents widely distributed polyphenols when applied to terrestrial NOM. Our results are thus of high importance for geochemical studies of e.g. soil organic matter formation[45], mineral sorption[46], or photochemical transformation[47], where aromatic structures of the NOM are used to explain molecular reactivity.

ASSOCIATED CONTENT

Supporting Information

Details of experimental procedures and FTICR MS analysis. Table S2 with samples molecular compositions with calculated aromaticity indices and assigned numbers of functional groups. This material is available free of charge via the Internet at https://pubs.acs.org/journal/ancham

Acknowledgments

This work was supported by Russian Science Foundation grant 19-75-00092. The authors are grateful to Jan Kaesler for FT-ICR MS measurements and for using the analytical facilities of the Centre for Chemical Microscopy (ProVIS) at the Helmholtz Centre for Environmental Research, Leipzig which is supported by the European Regional Development Funds (EFRE - Europe funds Saxony) and the Helmholtz Association.

Competing financial interest

The authors declare no competing financial interests.

References.

(1) Zherebker, A. Y.; Kostyukevich, Y. I.; Kononikhin, A. S.; Nikolaev, E. N.; Perminova, I. V. Molecular Compositions of Humic Acids Extracted from Leonardite and Lignite as Determined by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Mendeleev*

Commun. **2016**, *26* (5), 446–448.

(2) Hertkorn, N.; Harir, M.; Cawley, K. M.; Schmitt-Kopplin, P.; Jaffé, R. Molecular Characterization of Dissolved Organic Matter from Subtropical Wetlands: A Comparative Study through the Analysis of Optical Properties, NMR and FTICR/MS. *Biogeosciences* **2016**, *13* (8), 2257–2277.

(3) Li, Y.; Harir, M.; Lucio, M.; Kanawati, B.; Smirnov, K.; Flerus, R.; Koch, B. P.; Schmitt-Kopplin, P.; Hertkorn, N. Proposed Guidelines for Solid Phase Extraction of Suwannee River Dissolved Organic Matter. *Anal. Chem.* **2016**, *88* (13), 6680–6688.

(4) Raeke, J.; Lechtenfeld, O. J.; Tittel, J.; Oosterwoud, M. R.; Bornmann, K.; Reemtsma, T. Linking the Mobilization of Dissolved Organic Matter in Catchments and Its Removal in Drinking Water Treatment to Its Molecular Characteristics. *Water Res.* **2017**, *113*, 149–159.

(5) Hertkorn, N.; Ruecker, C.; Meringer, M.; Gugisch, R.; Frommberger, M.; Perdue, E. M.; Witt, M.; Schmitt-Kopplin, P. High-Precision Frequency Measurements: Indispensable Tools at the Core of the Molecular-Level Analysis of Complex Systems. *Anal. Bioanal. Chem.* **2007**, *389* (5), 1311–1327.

(6) Witt, M.; Fuchser, J.; Koch, B. P. Fragmentation Studies of Fulvic Acids Using Collision Induced Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2009**, *81* (7), 2688–2694.

(7) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M. Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008**, *80* (23), 8908–8919.

(8) Forcisi, S.; Moritz, F.; Lucio, M.; Lehmann, R.; Stefan, N.; Schmitt-Kopplin, P. Solutions for Low and High Accuracy Mass Spectrometric Data Matching: A Data-Driven Annotation Strategy in Nontargeted Metabolomics. *Anal. Chem.* **2015**, *87* (17), 8917–8924.

(9) Orlov, A. A.; Zherebker, A.; Eletskaya, A. A.; Chernikov, V. S.; Kozlovskaya, L. I.; Zhernov, Y. V.; Kostyukevich, Y.; Palyulin, V. A.; Nikolaev, E. N.; Osolodkin, D. I.; et al. Examination of Molecular Space and Feasible Structures of Bioactive Components of Humic Substances by FTICR MS Data Mining in ChEMBL Database. *Sci. Rep.* **2019**, *9* (1), 1–12.

(10) Spencer, R. G. M.; Mann, P. J.; Dittmar, T.; Eglinton, T. I.; McIntyre, C.; Holmes, R. M.; Zimov, N.; Stubbins, A. Detecting the Signature of Permafrost Thaw in Arctic Rivers. *Geophys. Res. Lett.* **2015**, *42* (8), 2830–2835.

(11) Hockaday, W. C.; Purcell, J. M.; Marshall, A. G.; Baldock, J. A.; Hatcher, P. G. Electrospray and Photoionization Mass Spectrometry for the Characterization of Organic Matter in Natural Waters: A Qualitative Assessment. *Limnol. Oceanogr.* **2009**, *7*, 81–95.

(12) Sleighter, R. L.; Hatcher, P. G. The Application of Electrospray Ionization Coupled to Ultrahigh Resolution Mass Spectrometry for the Molecular Characterization of Natural Organic Matter. *J. mass Spectrom.* **2007**, *42* (5), 559–574.

(13) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the van Krevelen Diagram. *Anal. Chem.* **2003**, *75* (20), 5336–5344.

(14) Polyakov, A. Y.; Lebedev, V. A.; Shirshin, E. A.; Rumyantsev, A. M.; Volikov, A. B.; Zherebker, A.; Garshev, A. V.; Goodilin, E. A.; Perminova, I. V. Non-Classical Growth of Water-Redispersible Spheroidal Gold Nanoparticles Assisted by Leonardite Humate. *CrystEngComm* **2017**, *19* (5), 876–886.

(15) Herzsprung, P.; Von Tümpling, W.; Hertkorn, N.; Harir, M.; Büttner, O.; Bravidor, J.; Friese, K.; Schmitt-Kopplin, P. Variations of DOM Quality in Inflows of a Drinking Water Reservoir: Linking of van Krevelen Diagrams with EEMF Spectra by Rank Correlation. *Environ. Sci. Technol.* **2012**, *46* (10), 5511–5518.

(16) Hockaday, W. C.; Grannas, A. M.; Kim, S.; Hatcher, P. G. Direct Molecular Evidence for the Degradation and Mobility of Black Carbon in Soils from Ultrahigh-Resolution Mass Spectral Analysis of Dissolved Organic Matter from a Fire-Impacted Forest Soil. *Org. Geochem.* **2006**, *37* (4), 501–510.

(17) Koch, B. P.; Dittmar, T. From Mass to Structure: An Aromaticity Index for High-Resolution Mass Data of Natural Organic Matter. *Rapid Commun. Mass Spectrom.* **2006**, *20* (5), 926–932.

(18) Koch, B. P.; Dittmar, T. Erratum: From Mass to Structure: An Aromaticity Index for High-Resolution Mass Data of Natural Organic Matter (Rapid Communications in Mass Spectrometry (2006) 20 (926-932) DOI: 10.1002/Rcm.2386). *Rapid Communications in Mass Spectrometry*. 2016, p 250.

(19) Gonsior, M.; Peake, B. M.; Cooper, W. T.; Podgorski, D.; D'Andrilli, J.; Cooper, W. J. Photochemically Induced Changes in Dissolved Organic Matter Identified by Ultrahigh Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environ. Sci. Technol.* **2009**, *43* (3), 698–703.

(20) Zherebker, A.; Podgorski, D. C.; Kholodov, V. A.; Orlov, A. A.; Yaroslavtseva, N. V.; Kharybin, O.; Kholodov, A.; Spector, V.; Spencer, R. G. M.; Nikolaev, E.; et al. The Molecular Composition of Humic Substances Isolated From Yedoma Permafrost and Alas Cores in the Eastern Siberian Arctic as Measured by Ultrahigh Resolution Mass Spectrometry. *J. Geophys. Res. Biogeosciences* **2019**, *124* (8), 2432–2445.

(21) Stubbins, A.; Mann, P. J.; Powers, L.; Bittar, T. B.; Dittmar, T.; McIntyre, C. P.; Eglinton, T. I.; Zimov, N.; Spencer, R. G. M. Low Photolability of Yedoma Permafrost Dissolved Organic Carbon. *J. Geophys. Res. Biogeosciences* **2017**, *122* (1), 200–211.

(22) Zherebker, A.; Shirshin, E.; Rubekina, A.; Kharybin, O.; Kononikhin, A.; Kulikova, N. A.;

Zaitsev, K. V.; Roznyatovsky, V. A.; Grishin, Y. K.; Perminova, I. V.; et al. Optical Properties of Soil Dissolved Organic Matter Are Related to Acidic Functions of Its Components as Revealed by Fractionation, Selective Deuteromethylation, and Ultrahigh Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2020**, *54* (5), 2667–2677.

(23) Kostyukevich, Y.; Acter, T.; Zherebker, A.; Ahmed, A.; Kim, S.; Nikolaev, E. Hydrogen/Deuterium Exchange in Mass Spectrometry. *Mass Spectrom. Rev.* **2018**, *37* (6), 811–853.

(24) Zherebker, A. Y.; Airapetyan, D.; Konstantinov, A. I.; Kostyukevich, Y. I.; Kononikhin, A. S.; Popov, I. A.; Zaitsev, K. V.; Nikolaev, E. N.; Perminova, I. V. Synthesis of Model Humic Substances: A Mechanistic Study Using Controllable H/D Exchange and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analyst* **2015**, *140* (13), 4708–4719.

(25) Baluha, D. R.; Blough, N. V.; Del Vecchio, R. Selective Mass Labeling for Linking the Optical Properties of Chromophoric Dissolved Organic Matter to Structure and Composition via Ultrahigh Resolution Electrospray Ionization Mass Spectrometry. *Environ. Sci. Technol.* **2013**, *47* (17), 9891–9897.

(26) Fedoros, E. I.; Orlov, A. A.; Zherebker, A.; Gubareva, E. A.; Maydin, M. A.; Konstantinov, A. I.; Krasnov, K. A.; Karapetian, R. N.; Izotova, E. I.; Pigarev, S. E.; et al. Novel Water-Soluble Lignin Derivative BP-Cx-1: Identification of Components and Screening of Potential Targets in Silico and in Vitro. *Oncotarget* **2018**, *9* (26), 18578.

(27) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A Simple and Efficient Method for the Solid-Phase Extraction of Dissolved Organic Matter (SPE-DOM) from Seawater. *Limnol. Ocean. Methods* **2008**, *6*, 230–235.

(28) Zherebker, A.; Kostyukevich, Y.; Kononikhin, A.; Kharybin, O.; Konstantinov, A. I.; Zaitsev, K. V.; Nikolaev, E.; Perminova, I. V. Enumeration of Carboxyl Groups Carried on Individual

Components of Humic Systems Using Deuteromethylation and Fourier Transform Mass Spectrometry. *Anal. Bioanal. Chem.* **2017**, *409* (9), 2477–2488.

(29) Andjelkovic, T.; Perovic, J.; Purenovic, M.; Blagojevic, S.; Nikolic, R.; Andjelkovic, D.; Bojic, A. A Direct Potentiometric Titration Study of the Dissociation of Humic Acid with Selectively Blocked Functional Groups. *Eclética Química* **2006**, *31* (3), 39–46.

(30) Mase, T.; Arima, H.; Tomioka, K.; Yamada, T.; Murase, K. Imidazo[2,1-b]Benzothiazoles 3: Syntheses and Immunosuppressive Activities of 2-(m-Acyloxyphenyl)Imidazo[2,1-b]Benzothiazoles. *Eur. J. Med. Chem.* **1988**, *23* (4), 335–339.

(31) Ma, J.; Del Vecchio, R.; Golanoski, K. S.; Boyle, E. S.; Blough, N. V. Optical Properties of Humic Substances and CDOM: Effects of Borohydride Reduction. *Environ. Sci. Technol.* **2010**, *44* (14), 5395–5402.

(32) Kostyukevich, Y.; Kononikhin, A.; Zherebker, A.; Popov, I.; Perminova, I.; Nikolaev, E. Enumeration of Non-Labile Oxygen Atoms in Dissolved Organic Matter by Use of $^{16}O/^{18}O$ Exchange and Fourier Transform Ion-Cyclotron Resonance Mass Spectrometry. *Anal. Bioanal. Chem.* **2014**, *406* (26), 6655–6664.

(33) Kellerman, A. M.; Dittmar, T.; Kothawala, D. N.; Tranvik, L. J. Chemodiversity of Dissolved Organic Matter in Lakes Driven by Climate and Hydrology. *Nat. Commun.* **2014**, *5*, 3804.

(34) Waggoner, D. C.; Chen, H.; Willoughby, A. S.; Hatcher, P. G. Formation of Black Carbon-like and Alicyclic Aliphatic Compounds by Hydroxyl Radical Initiated Degradation of Lignin. *Org. Geochem.* **2015**, *82*, 69–76.

(35) Zherebker, A.; Shirshin, E.; Kharybin, O.; Kostyukevich, Y.; Kononikhin, A.; Konstantinov, A. I.; Volkov, D.; Roznyatovsky, V. A.; Grishin, Y. K.; Perminova, I. V; et al. Separation of Benzoic and Unconjugated Acidic Components of Leonardite Humic Material Using Sequential Solid-Phase Extraction at Different PH Values as Revealed by Fourier Transform

Ion Cyclotron Resonance Mass Spectrometry and Correlation Nuclear Magneti. *J. Agric. Food Chem.* **2018**, *66* (46), 12179–12187.

(36) Hosangadi, B. D.; Dave, R. H. An Efficient General Method for Esterification of Aromatic Carboxylic Acids. *Tetrahedron Lett.* **1996**, *37* (35), 6375–6378.

(37) Bianca, M. R.; Baluha, D. R.; Gonsior, M.; Schmitt-Kopplin, P.; Del Vecchio, R.; Blough, N. V. Contribution of Ketone/Aldehyde-Containing Compounds to the Composition and Optical Properties of Suwannee River Fulvic Acid Revealed by Ultrahigh Resolution Mass Spectrometry and Deuterium Labeling. *Anal. Bioanal. Chem.* **2020**, 1–11.

(38) Kinney, C. R.; Love, D. L. Quinone Character of Oxidation Products of a Bituminous Coal. *Anal. Chem.* **1957**, *29* (11), 1641–1645.

(39) Morreel, K.; Kim, H.; Lu, F.; Dima, O.; Akiyama, T.; Vanholme, R.; Niculaes, C.; Goeminne, G.; Inzé, D.; Messens, E.; et al. Mass Spectrometry-Based Fragmentation as an Identification Tool in Lignomics. *Anal. Chem.* **2010**, *82* (19), 8095–8105.

(40) Lei, Z.; Jing, L.; Qiu, F.; Zhang, H.; Huhman, D.; Zhou, Z.; Sumner, L. W. Construction of an Ultrahigh Pressure Liquid Chromatography-Tandem Mass Spectral Library of Plant Natural Products and Comparative Spectral Analyses. *Anal. Chem.* **2015**, *87* (14), 7373–7381.

(41) Schwemer, T.; Rüger, C. P.; Sklorz, M.; Zimmermann, R. Gas Chromatography Coupled to Atmospheric Pressure Chemical Ionization FT-ICR Mass Spectrometry for Improvement of Data Reliability. *Anal. Chem.* **2015**.

(42) Rüger, C. P.; Miersch, T.; Schwemer, T.; Sklorz, M.; Zimmermann, R. Hyphenation of Thermal Analysis to Ultrahigh-Resolution Mass Spectrometry (Fourier Transform Ion Cyclotron Resonance Mass Spectrometry) Using Atmospheric Pressure Chemical Ionization for Studying Composition and Thermal Degradation of Complex Materials. *Anal. Chem.* **2015**, *87* (13), 6493–6499.

(43) Hertkorn, N.; Benner, R.; Frommberger, M.; Schmitt-Kopplin, P.; Witt, M.; Kaiser, K.; Kettrup, A.; Hedges, J. I. Characterization of a Major Refractory Component of Marine Dissolved Organic Matter. *Geochim. Cosmochim. Acta* **2006**, *70* (12), 2990–3010.

(44) DiDonato, N.; Hatcher, P. G. Alicyclic Carboxylic Acids in Soil Humic Acid as Detected with Ultrahigh Resolution Mass Spectrometry and Multi-Dimensional NMR. *Org. Geochem.* **2017**, *112*, 33–46.

(45) Malik, A. A.; Roth, V. N.; Hébert, M.; Tremblay, L.; Dittmar, T.; Gleixner, G. Linking Molecular Size, Composition and Carbon Turnover of Extractable Soil Microbial Compounds. *Soil Biol. Biochem.* **2016**, *100*, 66–73.

(46) Coward, E. K.; Ohno, T.; Sparks, D. L. Direct Evidence for Temporal Molecular Fractionation of Dissolved Organic Matter at the Iron Oxyhydroxide Interface. *Environ. Sci. Technol.* **2019**, *53* (2), 642–650.

(47) Wilske; Herzsprung; Lechtenfeld; Kamjunke; von Tümpling. Photochemically Induced Changes of Dissolved Organic Matter in a Humic-Rich and Forested Stream. *Water* **2020**, *12* (2), 331.