Supplementary Material

Molecular-scale investigation of soil fulvic acid and water-extractable organic matter by high-resolution mass spectrometry and $^1$H NMR spectroscopy

Ruixia Han$^A$, B Jitao Lv$^A$, C Lei Luo$^A$, Bei Wen$^A$ and Shuzhen Zhang$^A$, B

$^A$State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, P.O. Box 2871, Beijing 100085, China

$^B$University of Chinese Academy of Sciences, Beijing 100049, China

$^C$Corresponding author. Email: jtlv@rcees.ac.cn
**Text S1.** Detailed extraction procedures.

**Table S1.** Properties and elemental compositions of standard IHSS soils.

**Fig. S1.** Extraction and desalination approaches of three methods.

**Fig. S2.** DOC content in the extracts by the three methods in the two soils.

**Fig. S3.** Venn diagrams of comparing the unique and common peaks extracted by three different protocols for Elliott silt loam soil (a) and Pahokee peat soil (b) investigated with negative ion mode ESI FT-ICR-MS.

**Fig. S4.** Van Krevelen diagrams of the negative ion mode DOM molecules extracted by the three methods in the two soils.

**Fig. S5.** The percentage of each molecular group in FA, FAPPL and WEOM extracted from the two soils (Elliott silt loam soil (ESLS) and Pahokee peat soil (PP)) investigated with negative ion mode ESI FT-ICR-MS.

**Fig. S6.** Number of molecules of all the groups extracted by the three methods (FA(a, d), FAPPL(b, e) and WEOM(c, f)) in the two soils (Elliott silt loam soil (a, b, c), Pahokee peat soil (d, e, f)) investigated with negative ion mode ESI FT-ICR-MS.

**Fig. S7.** Quantitative distribution of O/C (a, b), H/C (c, d), DBE (e, f), NOSC (g, h) and Almod (i, j) of DOM molecules extracted from Elliott silt loam soil (a, c, e, g, i) and Pahokee peat soil (b, d, f, h, j) investigated with negative ion mode ESI FT-ICR-MS.

**Fig. S8.** Broadband ESI-FT-ICR-MS spectrum of the positive ion mode DOM components extracted by WEOM from Elliott silt loam soil (non Na-adduct molecules (black); Na-adduct molecules (red)).

**Fig. S9.** van Krevelen diagrams of the positive ion mode DOM components extracted by the
three methods: FA (a, d), FAPPL (b, e) and WEOM (c, f) from Elliott silt loam soil (a, b, c) and Pahokee peat soil (d, e, f).

Fig. S10. Number of molecules of all the groups extracted by the three methods (FA(a, d), FAPPL(b, e) and WEOM(c, f)) in the two soils (Elliott silt loam soil (a, b, c), Pahokee peat soil (d, e, f)) investigated with positive ion mode ESI FT-ICR-MS.
Text S1. Detailed extraction procedures.

Extraction of FA: soil sample was acidified to pH = 1-2 and dissolved in 0.1 M HCl solution, followed by shaking, centrifuging and filtering with 0.45 μm filters to get FA extract 1. Then 0.1 M NaOH was added to the neutralized soil residue under N₂ atmosphere and stored overnight in a sealed flask. After centrifugation, the supernatant was acidified to pH = 1 and settled for 12 to 16 h followed by centrifugation to obtain the extracts (FA extract 2). FA extracts 1 and 2 were combined and passed through XAD-8 column following the elution program recommended by IHSS. Passed the eluate through H⁺ saturated cation exchange resin (Rohm & Haas) to desalt. Eluates were freeze-dried and stored at -20°C.

FAPPL was extracted with alkaline-acid method as same as FA. The extracts were collected and filtered with 0.22 μm filters and purified and concentrated with PPL based SPE. The PPL cartridges were rinsed with 2 cartridge filling of methanol before loading samples. Then samples were acidified to pH = 2 with 1 M HCl and passed through the cartridges, followed by rinsing cartridges with 3 cartridge volumes of pH = 2 ultrapure water. After sorbents were dried, 3 cartridge volumes of methanol were used to elute the cartridge. Eluates were dried under nitrogen and stored at -20°C until analysis.

WEOM was extracted by adding ultrapure water to liquid: soil = 10:1 and shaking at 150 rpm at 25°C overnight. The supernatant was collected after centrifuging at 4000 rpm for 20 min, and then filtered with 0.22 μm membrane. Extracts were processed through PPL column to purify and desalt. Eluates were dried under nitrogen and stored at -20°C.
Table S1. Properties and elemental compositions of standard IHSS soils.\(^1\)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Texture</th>
<th>H(_2)O</th>
<th>Ash</th>
<th>C</th>
<th>H</th>
<th>O</th>
<th>N</th>
<th>S</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliott Silt Loam Soil (ESLS)</td>
<td>loam</td>
<td>1.52</td>
<td>nd*</td>
<td>2.9</td>
<td>nd*</td>
<td>nd*</td>
<td>0.25</td>
<td>nd*</td>
<td>nd*</td>
</tr>
<tr>
<td>Pahokee Peat Soil (PP)</td>
<td>muck</td>
<td>6.2</td>
<td>12.7</td>
<td>46.9</td>
<td>3.90</td>
<td>30.3</td>
<td>3.42</td>
<td>0.58</td>
<td>nd*</td>
</tr>
</tbody>
</table>

Note: *nd* means that an item was not determined.

Elemental compositions of IHSS samples are available at

Fig. S1. Extraction and desalination approaches of the three methods (FA, FA\textsubscript{PPL}, WEOM).

Fig. S2. DOC content in the extracts by the three methods (FA, FA\textsubscript{PPL} and WEOM) in the two soils (ESLS, Elliott silt loam soil; PP, Pahokee peat soil).
Fig. S3. Venn diagrams comparing the unique and common peaks extracted by three different protocols for Elliott silt loam soil (a) and Pahokee peat soil (b) investigated with negative ion mode ESI FT-ICR-MS.
Fig. S4. Van Krevelen diagrams of the negative ion mode DOM molecules extracted by the three methods in the two soils. (CHO molecules: a, b, c, g, h, i; CHON molecules: d, e, f, j, k, l; soil ESLS: a ~ f; soil PP: g ~ l; FA: a, d, g, j; FAPPL: b, e, h, k; WEOM: c, f, i, l).
**Fig. S5.** The percentage of each molecular group in FA, FAPPL and WEOM extracted from the two soils (Elliott silt loam soil (ESLS) and Pahokee peat soil (PP)) investigated with negative ion mode ESI FT-ICR-MS.

**Fig. S6.** Number of molecules of all the groups extracted by the three methods (FA(a, d), FAPPL(b, e) and WEOM(c, f)) in the two soils (Elliott silt loam soil (a, b, c), Pahokee peat soil (d, e, f)) investigated with negative ion mode ESI FT-ICR-MS.
Figure S7. Quantitative distribution of O/C (a, b), H/C (c, d), DBE (e, f), NOSC (g, h) and AImod (i, j) of DOM molecules extracted from Elliott silt loam soil (a, c, e, g, i) and Pahokee peat soil (b, d, f, h, j) investigated with negative ion mode ESI FT-ICR-MS.

Figure S8. Broadband ESI-FT-ICR-MS spectrum of the positive ion mode DOM components extracted by WEOM from Elliott silt loam soil (non Na-adduct molecules (black); Na-adduct molecules (red)).
**Fig. S9.** van Krevelen diagrams of the positive ion mode DOM components (CHO molecules (red); CHON molecules (blue); digits are number of molecules) extracted by the three methods: FA (a, d), FA<sub>PPL</sub> (b, e) and WEOM (c, f) from Elliott silt loam soil (a, b, c) and Pahokee peat soil (d, e, f).

**Fig. S10.** Number of molecules extracted by the three methods (FA(a, d), FA<sub>PPL</sub>(b, e) and WEOM(c, f)) in the two soils (Elliott silt loam soil (a, b, c), Pahokee peat soil (d, e, f)) investigated with positive ion mode ESI FT-ICR-MS.