Supporting Information

Response to “Comment on the Molecular Composition of Natural Organic Matter Analyzed by Electrospray FTICR MS”

Will Kew, John W. T. Blackburn, Dušan Uhrín*

EaStCHEM School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ, UK

*E-mail: dusan.uhrin@ed.ac.uk

Table of Contents

Experimental Parameters

Further discussion on the range of *m/z* values observed

Investigating Potential Oligomerization

Figure S1. ESI-(-) FTICR mass spectra of SRFA when tuned for low masses

Figure S2. Comparison of spectra acquired at 1 mg/ml, 2.5 mg/ml and 5 mg/ml

Figure S3. ESI-(-) FTICR mass spectra of SRFA. Expansions of *m/z* 653.0 to 653.3 showing the effect of varying dry gas temperature in the ESI source from 150 to 220°C.

Figure S4. CID ESI-(-) FTICR mass spectra of SRFA from 200 to 660 *m/z* acquired with an isolation window of 10 *m/z* at *m/z* 650 while increasing the CID voltage from 0 to 12.5V.

Figure S5. CID SI-(-) FTICR mass spectrum of SRFA. Expansion at *m/z* 600 to 660 of the bottom spectrum from Figure S4 acquired with a 12.5 V CID voltage.

Figure S6. CID ESI-(-) FTICR mass spectra of SRFA from 200 to 760 *m/z* acquired with an isolation window of 10 *m/z* at *m/z* 750 while increasing the CID voltage from 0 to 10 V.

Figure S7. CID SI-(-) FTICR mass spectrum of SRFA. Expansion at *m/z* 700 to 760 of the bottom spectrum from Figure S6 acquired with a 10 V CID voltage.

Experimental Parameters

All data were acquired on a SolariX 12 T (Bruker Daltonics) FTICR MS using *ftmscontrol* 2.1 and processed with *DataAnalysis* 4.4. Data were all acquired in negative mode ESI. Spectra were calibrated using a quadratic function against previously known assignments. Molecular formula assignments performed using *PetroOrg* 12.0 (Florida State University). Unless otherwise noted, spectra were acquired with 0.1 mg/ml in 50:50 methanol:water. Solvents were LC-MS grade from Sigma Aldrich.

All experiments were acquired with the following settings unless otherwise noted. All FIDs were processed, including apodization and zero-filling, with default parameters of full sine apodization and zero-filling once prior to fast Fourier transform. 200 scans were summed into one FID. The ESI parameters were a flow rate of 200 µL/hr, nebuliser pressure of 2 bar, dry gas flow of 6 L/min and temperature of 180°C. Source optics were as follows; capillary exit was set to -200 V, deflector plate was set to -180 V, funnel 1 was set to -100 V and skimmer 1 was set to -45 V. Funnel RF amplitude was set to 120 Vpp. The octopole was set to a frequency of 5 MHz with RF amplitude of 300 Vpp. The quadrupole Q1 mass was set to 100 *m/z*. The ICR cell setting were as follows; a sidekick potential of 2 V, sidekick offset of 2 V, analyser entrance was set to 4 V, front trapping potential was -0.5 V and back trapping potential was -0.6 V.

For complete metadata on acquisition, please see the raw data available online at http://dx.doi.org/10.7488/ds/2061

Published bimodal dataset

The bimodal dataset was acquired with a 2 MW FID and low mass limit of 147.44 *m/z* and high mass limit of 1000 *m/z.* The syringe flow rate was 500 µL/hr. Ions were accumulated in the collision cell for 1.2 seconds with a voltage of 0.5 V and DC extract bias of -0.7 V, gas control was set to 35%, RF frequency of 2 MHz and collision RF amplitude of 400 Vpp. Transfer optics were set to a time of flight of 0.7 ms, frequency of 4 MHz and RF amplitude of 440 Vpp. The ICR cell setting were as follows; a sidekick potential of 2 V, sidekick offset of 2 V, analyser entrance was set to 4 V, front trapping potential was -0.4 V and back trapping potential was -0.5 V.

New "Normal" distributed dataset

The normal dataset was acquired with a 2 MW FID and low mass limit of 147.44 *m/z* and high mass limit of 1000 *m/z*. Ions were accumulated in the collision cell for 0.15 seconds with a voltage of 0.5 V and DC extract bias of -0.7 V, gas control was set to 35%, RF frequency of 2 MHz and collision RF amplitude of 400 Vpp. Transfer optics were set to a time of flight of 0.7 ms, frequency of 4 MHz and RF amplitude of 440 Vpp.

Low-mass optimised dataset

For acquisition of a low-mass optimised dataset (**Figure S1**), the transient was increased to 4 MW. Ion accumulation time was 300 ms. The low mass was set to 73.5 *m/z* and high mass 1000 *m/z*.

Further discussion on the range of *m/z* values observed

He *et al*. pointed out that SRFA spectra are typically presented as a normal distribution between 200-700 *m/z*. However, SRFA is expected to contain compounds smaller and larger than this. Furthermore, He *et al*. used a Q1 mass filter of 300 m/z, and mass spectral limits of 200-800 *m/z*, biasing their results to this mass range. Typically, we acquire spectra between 100 or 150 and 1000 or 2000 *m/z*, with the Q1 mass set to the low mass limit. **Figure S1** shows that formulae can be assigned at low masses approaching 100 *m/z.*

Investigating Potential Oligomerization

High-concentration dataset

Spectra were acquired at 1 mg/ml, 2.5 mg/ml and 5 mg/ml (**Figure S2**). At 5 mg/ml spray stability began to suffer. Above this concentration we believe the sprayer would block. These data were acquired at 4 MW with a low mass limit of 98.3 *m/z* and high mass limit of 2000 *m/z*.

At these higher concentrations, we had to severely reduce the ion accumulation time (IAT) to 0.01s to prevent overfilling the ICR cell and to continue acquiring high quality spectra. IAT has been shown to impact on relative peak intensity, with longer IATs allowing for observation of more minor components.[1] Interestingly, the very short IAT presented by Cao *et al*.[1] had an apparent bimodal distribution for ESI of SRNOM. Our original bimodal dataset, however, had been acquired with a longer IAT of 1.2s.

Capillary Temperature

One means to determine if the higher mass ions were non-covalent aggregates is to vary the temperature of the ESI capillary. At higher temperatures, aggregates should be less likely to form. The ESI capillary temperature is controlled by the dry gas temperature. In our standard method, this is 180 °C. To see if the ESI capillary temperature would have an impact, several spectra of SRFA were acquired at varying dry gas temperatures from 150°C to 220°C.

The overlaid results at *m/z* 653 are shown in **Figure S3**. There are only marginal changes in signal intensities for each spectrum, as may be caused for random variation and varying dry gas temperature. The mean standard deviation at *m/z* 653 in peak area was only 0.06%. If these ions had been non-covalently aggregates, their intensity would likely decrease significantly as capillary temperature increased. This was not observed, and further supports our conclusions that we do not produce a substantial levels of aggregates in our experiments.

Collision Cell Voltages

We experimented with a variety of collision cell voltages for the 2.5 mg/ml sample concentration – including 0 V (not recorded), 0.5 V (typical acquisition voltage), 2 V, 4 V, and 6 V. If our data at 2.5 mg/ml (and thus at all concentrations) is dimerising, increasing the collision voltage should break up these aggregates before causing further fragmentation. These collision cell voltages were applied without collision induced dissociation (CID). Increasing the collision voltage to 4 V made no substantial difference to the profile, though it did increase the relative intensities of most peaks. At 6 V, the profile of the spectrum shifted marginally to lower masses, from around 380 m/z to 320 m/z, suggesting some degree of fragmentation.

To further investigate the possibility that the higher mass ions are non-covalently bonded aggregates, MS/MS experiments were performed on two higher mass regions – *m/z* 650 and *m/z* 750. Spectra were acquired with isolation of a 10 *m/z* window around these masses and increasing collision cell CID fragmentation energies applied, from 0 to 12.5 V (650 *m/z*) or 0 to 10 V (750 *m/z*). A non-selective, 10 *m/z* was chosen to test multiple ions at once and to simplify the experiment, as in-cell isolation and fragmentation can be non-trivial and very insensitive.[2] The spectra were acquired with a slightly increased transient length (1.67 s) to increase resolving power at these higher masses.

The results of these experiments are shown in **Figures S4-S7**, including stacked mass spectra plots for the increasing voltages (**Figures S4** and **S6**, 650 and 750 *m/z,* respectively), and expansions highlighting neutral losses of CO_2 and H_2O (**Figures S5** and **S7**, 650 and 750 *m/z,* respectively). The mass accuracy of these losses is sub-ppm. The loss of these fragments is typical for small molecule fragmentation,

especially of NOM samples.[2] If these ions at 650 and 750 m/z had been non-covalently bonded aggregates, it would be expected for their 'monomer' units to be lost before covalently bonded fragments appear. Thus, these experiment further confirm that we are not observing any significant level of aggregates in our mass spectra.

References

(1) Cao, D.; Lv, J.; Geng, F.; Rao, Z.; Niu, H.; Shi, Y.; Cai, Y.; Kang, Y. *Anal. Chem.* **2016**, *88* (24), 12210–12218.

(2) Witt, M.; Fuchser, J.; Koch, B. P. *Anal. Chem.* **2009**, *81* (7), 2688–2694.

Figures

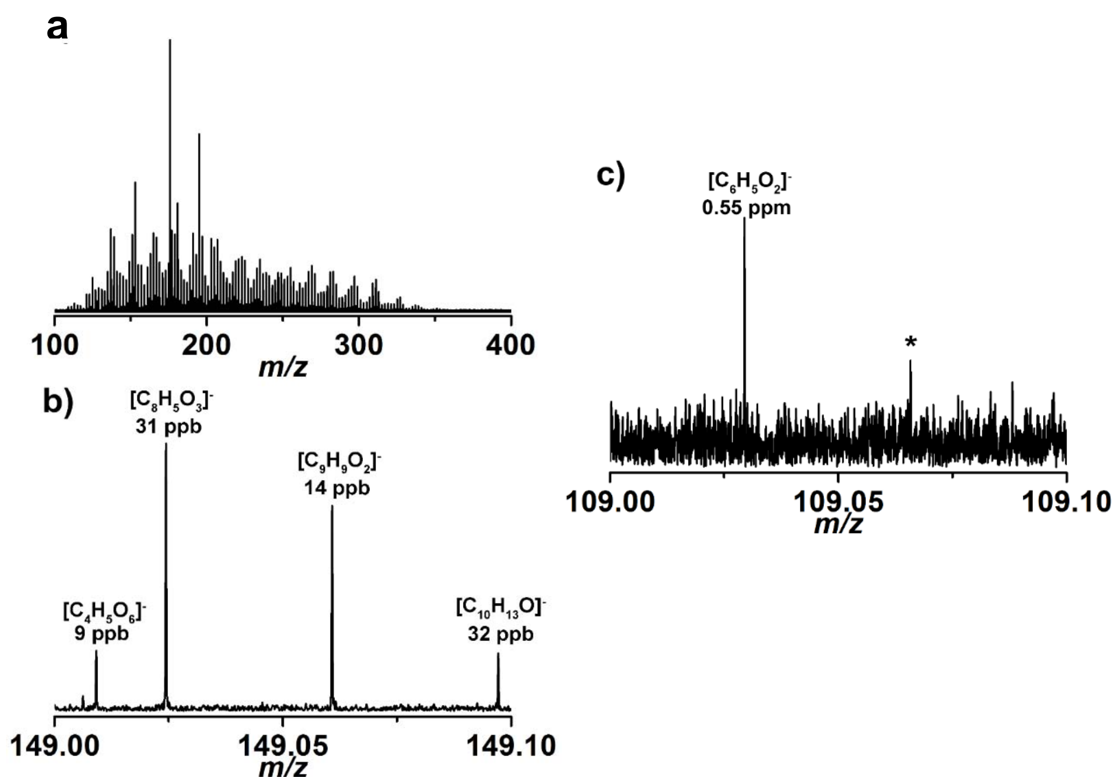

Figure S1 - ESI-(-) FTICR mass spectra of SRFA when tuned for low masses. (a) the 100-400 *m/z* range, (b) a single nominal mass at 149 *m/z*, and (c) the lowest assigned mass observed at 109 *m/z*. The peak labelled by an asterisk (*) was not assigned as it was below our peak picking thresholds.

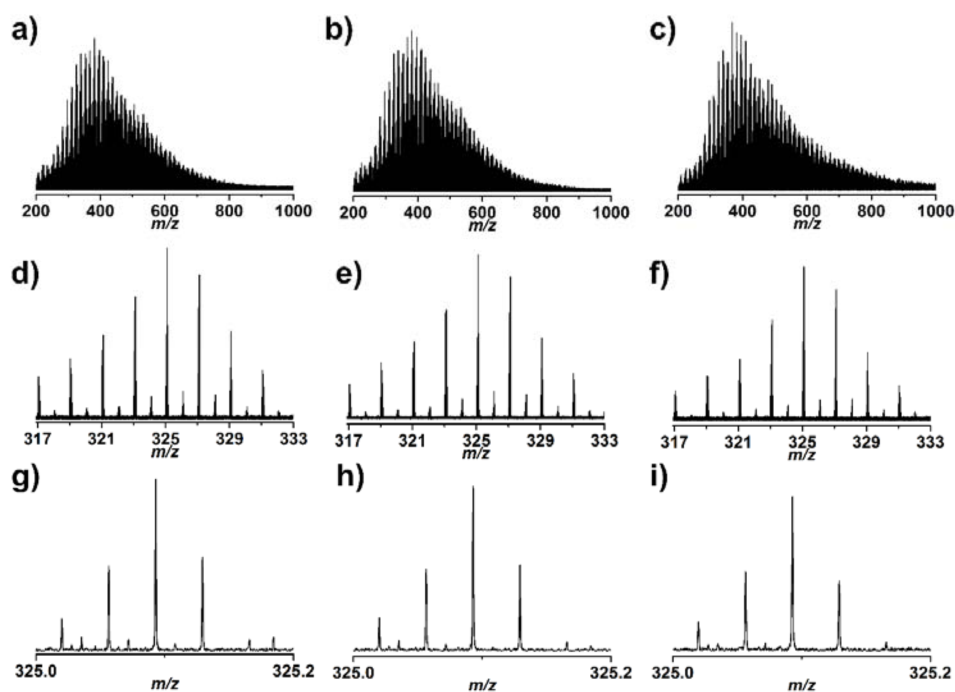

Figure S2 – ESI-(-) FTICR mass spectra of SRFA acquired at different concentrations; left column (a, d, g) at 1 mg/ml, middle column, (b, e, h) at 2.5 mg/ml, and right column (c, f) i) at 5 mg/ml. Top row shows the 200-1000 *m/z* range, the middle row shows a per-14 *m/z* distribution, and the bottom row shows a single *m/z* at 325.

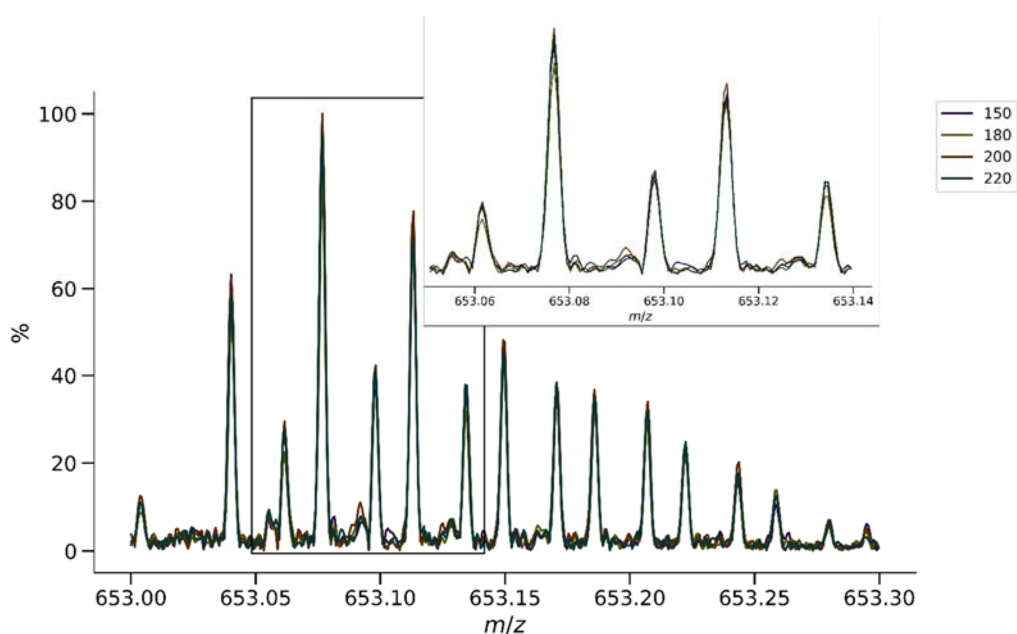

Figure S3 – ESI-(-) FTICR mass spectra of SRFA. Expansions of *m/z* 653.0 to 653.3 showing the effect of varying dry gas temperature in the ESI source from 150 to 220°C. Plots are overlaid and colour coded. Inset shows additional expansion at *m/z* 653.05-653.14. Data are uncalibrated. Peak intensities have been scaled to largest peak across all spectra within shown region. Spectra are a sum of 20 scans.

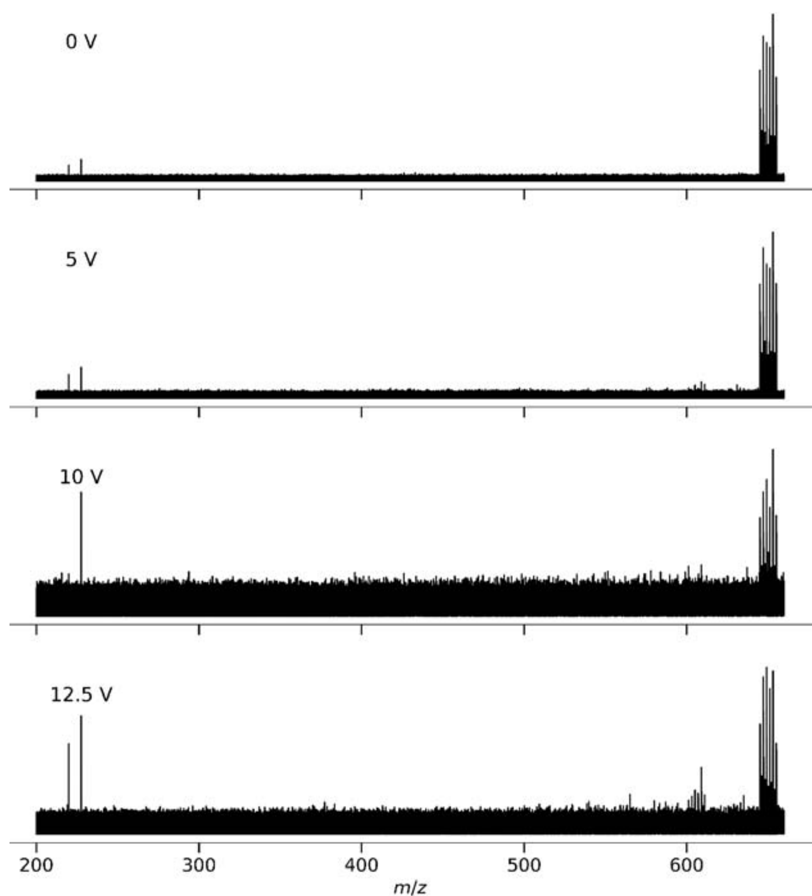

Figure S4 - CID ESI-(-) FTICR mass spectra of SRFA from 200 to 660 *m/z* acquired with an isolation window of 10 *m/z* at *m/z* 650 while increasing the CID voltage from 0 to 12.5V. Spectra show sum of 20 transients. At 12.5V ion accumulation time was increased to increase sensitivity. Note that peaks around 220 m/z are noise signals and not real ion signals. A chemical formula cannot be assigned to them, and they have no ^{13}C peak at +1 *m/z*, (or at +½ m/z if a doubly charged species, etc); they regularly appear in our CID spectra.

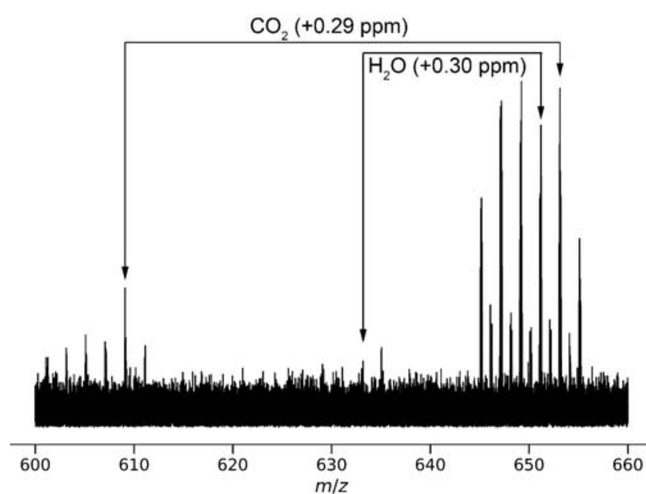

Figure S5 – CID ESI-(-) FTICR mass spectrum of SRFA. Expansion at *m/z* 600 to 660 of the bottom spectrum from Figure S4 acquired with a 12.5 V CID voltage. Annotations of two neutral losses are shown, along with mass accuracy errors for these losses.

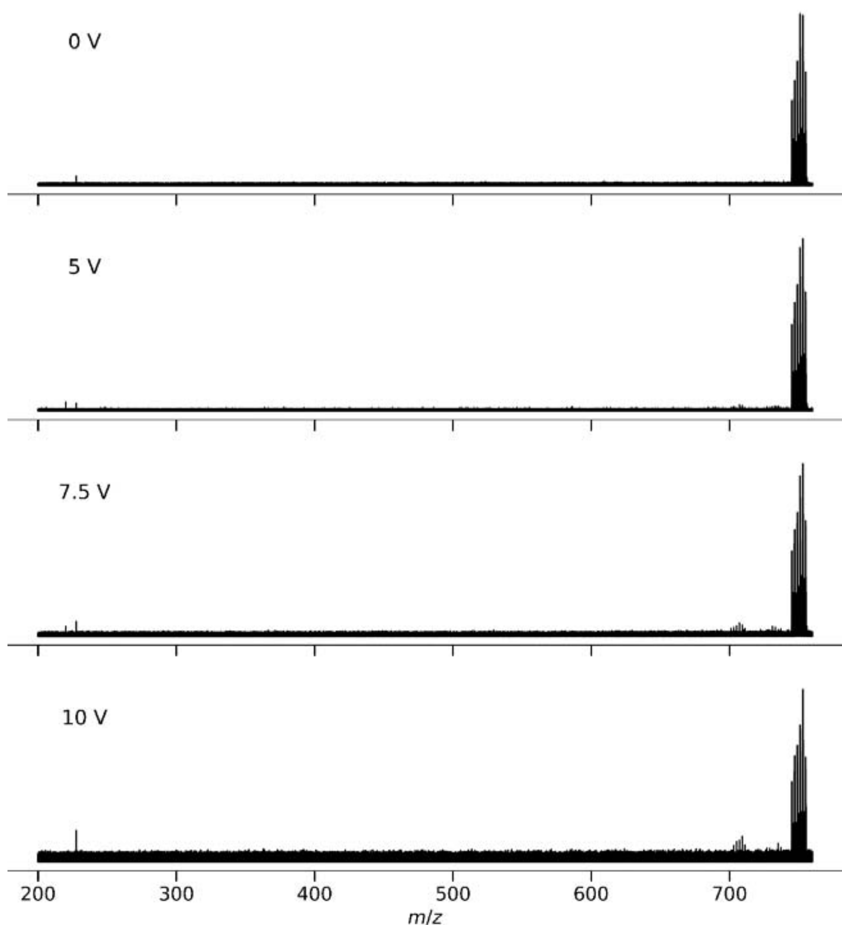

Figure S6 - CID ESI-(-) FTICR mass spectra of SRFA from 200 to 760 *m/z* acquired with an isolation window of 10 *m/z* at *m/z* 750 while increasing the CID voltage from 0 to 10 V. Spectra show sum of 20 transients. Note that peaks around 220 m/z are noise signals and not real ion signals. A chemical formula cannot be assigned to them, and they have no ^{13}C peak at +1 *m/z*, (or at +½ m/z if a doubly charged species, etc); they regularly appear in our CID spectra.

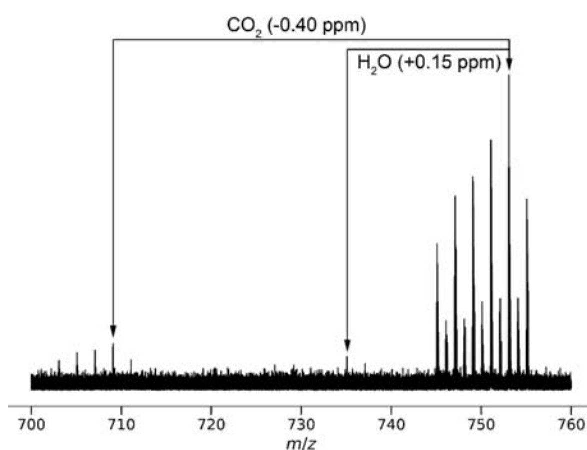

Figure S7 - CID ESI-(-) FTICR mass spectrum of SRFA. Expansion at *m/z* 700 to 760 of the bottom spectrum from Figure S6 acquired with a 10 V CID voltage. Annotations of two neutral losses are shown, along with mass accuracy errors for these losses.