Supporting Information

Molecular characterisation of organophosphorus compounds in wildfire smoke using 21-Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Amna Ijaz[1*], William Kew[2], Swarup China[2], Simeon K Schum[1], Lynn R Mazzoleni[1*]

[1]*Department of Chemistry, Michigan Technological University, Houghton, Michigan 49931, USA*

[2]*Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, USA*

Table of Contents

Supplementary Text

S1. Molecular compositions from the 15-T and the 21-T FT-ICR MS

Overall, 63% of the total ion current could be assigned a formula in the 15-T FT-ICR MS (Fourier transform - Ion Cyclotron Resonance mass spectrometer) and 75% in the 21-T FT-ICR MS. Formula error was centred at zero for both mass spectrometers (Figure S2). The largest fraction of assigned formulae belonged to CHNO species (~50%) in both FT-ICR instruments with CHO species being a close second.

Previously, for this aerosol mixture, using the FT-OTE (FT – Orbitrap Elite) MS, Brege et al.[1] observed a continuum of C in the detected species that extended from C_3 to C_{45} and encompassed aliphatic, aromatic, and condensed aromatic species with a continuous range of O atoms and DBE values using (+/-)APPI and (+/-)ESI. Based on AI_{mod} values, a majority of species were identified to be olefinic in nature by Brege et al.[1], which was corroborated by both FT-ICR instruments employed here. As per the 21-T FT-ICR analysis, 14.68% of the species identified here were aromatic or condensed aromatic, 18.41% were aliphatic, while ~47.56 and 19.33% (66.89% altogether) were low- and high-O unsaturated, i.e., olefinic species. Similar distributions were seen with 15-T FT-ICR analysis: 16.68% species were aromatic or condensed aromatic, 23.01% were aliphatic, while ~30% each (60.3% altogether) were either low- or high-O unsaturated species. Generally, organic emissions from anthropogenic and biogenic sources are rich in aliphatic or olefinic species, including alkanes, alkenes, alkanoic and alkenoic acids, etc.[2, 3] that are effective precursors for the formation of secondary organic aerosol. The substantial fraction of aromatic species observed for the BBOA mixture analysed here may be because of a strong interaction of this aerosol with wildfires in the 48-hours preceding its collection.

Average values of O/C, H/C, and DBE were 0.56 ± 0.2, 1.19 ± 0.3, and 9.38 ± 4.4 from 15-T (-)ESI/FT-ICR MS analysis and 0.46 ± 0.2, 1.20 ± 0.3, and 10.38 ± 4.8 from 21-T (-)ESI/FT-ICR MS analysis. As described in the main text of the report, the benefits of exceptional resolving power and sensitivity afforded by the 21-T FT-ICR MS became obvious only when species exclusively detected by it above the signal-to-noise ratio were considered.

CHO species. Species constituting only C, H, and O formed the second largest fraction of all molecular formulae assigned to the BBOA sample by both instruments, but they accounted for the most ion current as the tallest peaks throughout the scan range belonged to the CHO group (Figure 1). By 21-T analysis, the formulae assigned exhibited a continuous range of oxygenation: $O\text{-}O_{19}$ with O_{avg} of 9.46 ± 3.9 (Figure S2). The highest C_{avg} was noted for this molecular group at 25.15 ± 7.6 with the number of C atoms per molecule ranging from 6 to 41 (Figure S2). The O/C_{avg} and H/C_{avg} of all CHO formulae were 0.40 ± 0.2 and 1.22 ± 0.3. Their DBE also covered a wide range from 1 to 23 with the most ion intensity originating from species with DBE of 6-12 (Figure S2) and DBE_{avg} of 10.80 ± 5.0. The intermediate DBE values were supported by $AI_{mod,avg}$ of 0.30 ± 0.2, which indicated most species to be olefinic (Figure S9).

Molecular species observed at *m/z* values >400 are considered to be an indication of oxidative ageing and/or oligomerisation of atmospheric aerosol.[2] However, owing to the isomeric complexity of BBOA, the formulae assigned cannot be used to guarantee the presence of oligomers/chemical species commonly present in BBOA and only putative deductions can be made. For instance, most direct products of cellulose pyrolysis are CHO compounds, such as levoglucosan ($C_6H_{10}O_5$, 162.052824 Da)[4, 5] and cellobiosan ($C_{12}H_{20}O_{10}$, 324.105649 Da)[6]. Levoglucosan is a particle-phase marker of biomass burning, and the formula corresponding to it was observed only in 15-T FT-ICR MS analysis with 3% relative intensity. It was outside the *m/z* range covered by the 21-T FT-ICR MS, where the smallest analyte ion was observed at 182 Da. The molecular formula denoting levoglucosan's dimer, cellobiosan, was detected only by the 21-T FT-ICR MS with 4.0% relative intensity. Other pyrolysis products of lignin include syringol ($C_8H_{10}O_3$, 154.062995 Da), guaiacol ($C_7H_8O_2$, 124.052430 Da)[5], vanillic acid ($C_8H_8O_4$, 168.042260 Da), syringaldehyde ($C_9H_{10}O_4$, 182.057910 Da), and syringic acid ($C_9H_{10}O_5$, 198.052824 Da)[2]. In the 15-T FT-ICR MS, these formulae were assigned to peaks with relative intensities of 0.0 (not detected), 0.0, 9.0, 7.6, and 6.6%, while in the 21-T FT-ICR MS, we recorded only the peak corresponding potentially to syringic acid with 9.7% relative intensity. Although molecular formulae for syringol and guaiacol were not detected by either instrument, molecular formulae corresponding to their dimers ($C_{16}H_{18}O_6$; 306.110340 Da and $C_{14}H_{14}O_4$; 246.089210 Da), trimers ($C_{24}H_{26}O_9$; 458.157685 Da and $C_{21}H_{20}O_6$; 368.125990 Da), and tetramers ($C_{32}H_{34}O_{12}$; 610.205030 Da and $C_{28}H_{26}O_8$; 490.162770 Da) were detected by both instruments. Relative intensities recorded by the 21-T FT-ICR MS for the aforementioned species were quite high at 37.7, 27.7, and 6.1%. Relative intensities for peaks that were assigned formulae corresponding to guaiacol oligomers were similarly substantial at 21.4, 22.9, and 11.67%. A pentamer was potentially detected only for guaiacol ($C_{35}H_{32}O_{10}$; 612.199550 Da) exclusively by the 21-T FT-ICR MS at 2.0% relative intensity.

CHNO species. We assigned 5022 N-containing formulae in total, i.e., the largest fraction of all formulae assigned, by 21-T FT-ICR MS analysis. These formulae belonged to 41 heteroatom classes: $NO_2\text{-}NO_{18}$, $N_2O_3\text{-}N_2O_{16}$, and $N_3O_4\text{-}N_3O_{13}$, where the number of formulae assigned to each class decreased in the order listed (Figure S2). The majority of formulae assigned were concentrated between DBE values of 7 to 16. The average O/C was 0.45 ± 0.2 and H/C_{avg} was 1.17 ± 0.2 for the CHNO molecular group. Previously, BBOA has been reported to be rich in CHNO species[7, 8].

The CHNO species identified in this BBOA mixture with the 21-T FT-ICR MS was predominantly low-O olefinic compounds (Figure S9). The species with N_{1-2} comprised a considerable fraction of aromatic species, which were likely nitroaromatic compounds previously reported to be present in tar balls[9]. Aliphatic and condensed aromatic species were present, but only sparsely. Average O/N values of 7.91 ± 4.1 and a minimum of at least two O atoms per CHNO molecule, along with a high average number of O atoms per molecule (9.76 ± 3.3), indicated these compounds to comprise organic species with up to three nitro ($-NO_2$) or nitrate ($-ONO_2$) groups along with other O functional groups. It must be noted that the elemental ratios alone are not sufficient to draw conclusions about the presence of these functional groups. However, reasonable inferences can be drawn based on findings of fragmentation studies on atmospheric organic aerosol that have reported a loss of HNO_3[10, 11], HNO_2, and NH_3 from BBOA [11]. There were 43 CHNO species with three N atoms, whose O/N values were <2, i.e., they had an insufficient number of O to have $-NO_2$ or $-ONO_2$ functionalities. Considering that the N_3O_x class was detected exclusively by the 21-T FT-ICR MS, these compounds may likely have low ionisability or were minor constituents of the sample, which caused them to remain undetected in the 15-T FT-ICR and FT-OTE instruments. Their DBE values ranged from 7 to 12 and they were mostly classified as olefinic species with a small proportion of aromatic species. They could be unsaturated species containing reduced N, for instance, amine ($-NH_3$) or azo ($-N_2$) functional groups, with at least one other polar functional group as speculated previously in water-soluble organic aerosol[3]. They could also be aromatic species, such as phenols, naphthalenes, or benzoic acids, substituted with nitroso (-NO) groups.

CHOS species. The 21-T FT-ICR MS detected 190 CHOS species with much lower intensities than CHNO and CHO species. The number of O atoms in these compounds ranged from 5-12 with O/S_{avg} of 7.61 ± 1.5. Figure S3 shows the distribution of formulae belonging to these molecular groups across O_oS_s classes. These species had very low overall $AI_{mod,avg}$ of 0.06 ± 0.5 and DBE_{avg} of 3.2 ± 1.8 suggesting them to be very saturated and highly oxygenated. This was supported by the O/C_{avg} and H/C_{avg} of this molecular group being 0.70 ± 0.3 and 1.64 ± 0.2, both of which are much higher than average elemental ratios for CHO and CHNO groups (Table 1). Willoughby et al.[3] attributed aliphatic, oxygenated CHOS species in aerosol to the co-emission and reactions of sulphate (SO_x) aerosol with carbonaceous, hydrogen-saturated aerosol. It must be noted here that 15-T FT-ICR MS detected almost twice the number of CHOS species (n = 355) with the same overall elemental ratios, but relatively higher unsaturation (DBE_{avg} = 4.0 ± 2.6 and $AI_{mod,avg}$ = 0.12 ± 0.6). These CHOS species were present from *m/z* 200-450 (Figure S7). In both FT-ICR instruments, low levels of unsaturation, O/S of ≥4, and a minimum of five O atoms per CHOS molecule imply S to be present as organic sulfonates ($-SO_3$) or sulphates ($-OSO_3$) with other O-containing functional groups. The neutral loss of these functional groups as SO_3, SO_4, and H_2SO_4 has been reported in literature[10, 11].

References

1. Brege, M.; China, S.; Zelenyuk-Imre, A.; Mazzoleni, L., Extreme molecular complexity resulting in a continuum of carbonaceous species in biomass burning tar balls. *ACS Earth Space Chem*. **2021**, 5 (10), 2729-2739

2. Dzepina, K.; Mazzoleni, C.; Fialho, P.; China, S.; Zhang, B.; Owen, R. C.; Helmig, D.; Hueber, J.; Kumar, S.; Perlinger, J. A., Molecular characterization of free tropospheric aerosol collected at the Pico Mountain Observatory: a case study with a long-range transported biomass burning plume. *Atmos. Chem. Phys.* **2015**, 15 (9), 5047-5068.

3. Willoughby, A.; Wozniak, A.; Hatcher, P. G., A molecular-level approach for characterizing water-insoluble components of ambient organic aerosol particulates using ultrahigh-resolution mass spectrometry. *Atmos. Chem. Phys.* **2014**, 14 (18).

4. Simoneit, B. R.; Elias, V. O., Detecting organic tracers from biomass burning in the atmosphere. *Mar. Pollut. Bull.* **2001**, 42 (10), 805-810.

5. Simoneit, B. R., Biomass burning—a review of organic tracers for smoke from incomplete combustion. *J. Appl. Geochem.* **2002**, 17 (3), 129-162.

6. Bai, X.; Johnston, P.; Sadula, S.; Brown, R. C., Role of levoglucosan physiochemistry in cellulose pyrolysis. *J. Anal. Appl. Pyrolysis.* **2013**, 99, 58-65.

7. Laskin, A.; Smith, J. S.; Laskin, J., Molecular characterization of nitrogen-containing organic compounds in biomass burning aerosols using high-resolution mass spectrometry. *Environ. Sci. Technol.* **2009**, 43 (10), 3764-3771.

8. Schmitt-Kopplin, P.; Gelencser, A.; Dabek-Zlotorzynska, E.; Kiss, G.; Hertkorn, N.; Harir, M.; Hong, Y.; Gebefügi, I., Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy: organosulfates as photochemical smog constituents. *Anal. Chem.* **2010**, 82 (19), 8017-8026.

9. Adachi, K.; Sedlacek, A. J.; Kleinman, L.; Springston, S. R.; Wang, J.; Chand, D.; Hubbe, J. M.; Shilling, J. E.; Onasch, T. B.; Kinase, T., Spherical tarball particles form through rapid chemical and physical changes of organic matter in biomass-burning smoke. *Proc. Natl. Acad. Sci.* **2019**, 116 (39), 19336-19341.

10. LeClair, J. P.; Collett, J. L.; Mazzoleni, L. R., Fragmentation analysis of water-soluble atmospheric organic matter using ultrahigh-resolution FT-ICR mass spectrometry. *Environ. Sci. Technol.* **2012**, 46 (8), 4312-4322.

11. Schum, S. K. Molecular characterization of free tropospheric organic aerosol and the development of computational tools for molecular formula assignment. Dissertation, Michigan Technological University, Houghton, MI, USA, **2019**.

Supplementary Figures

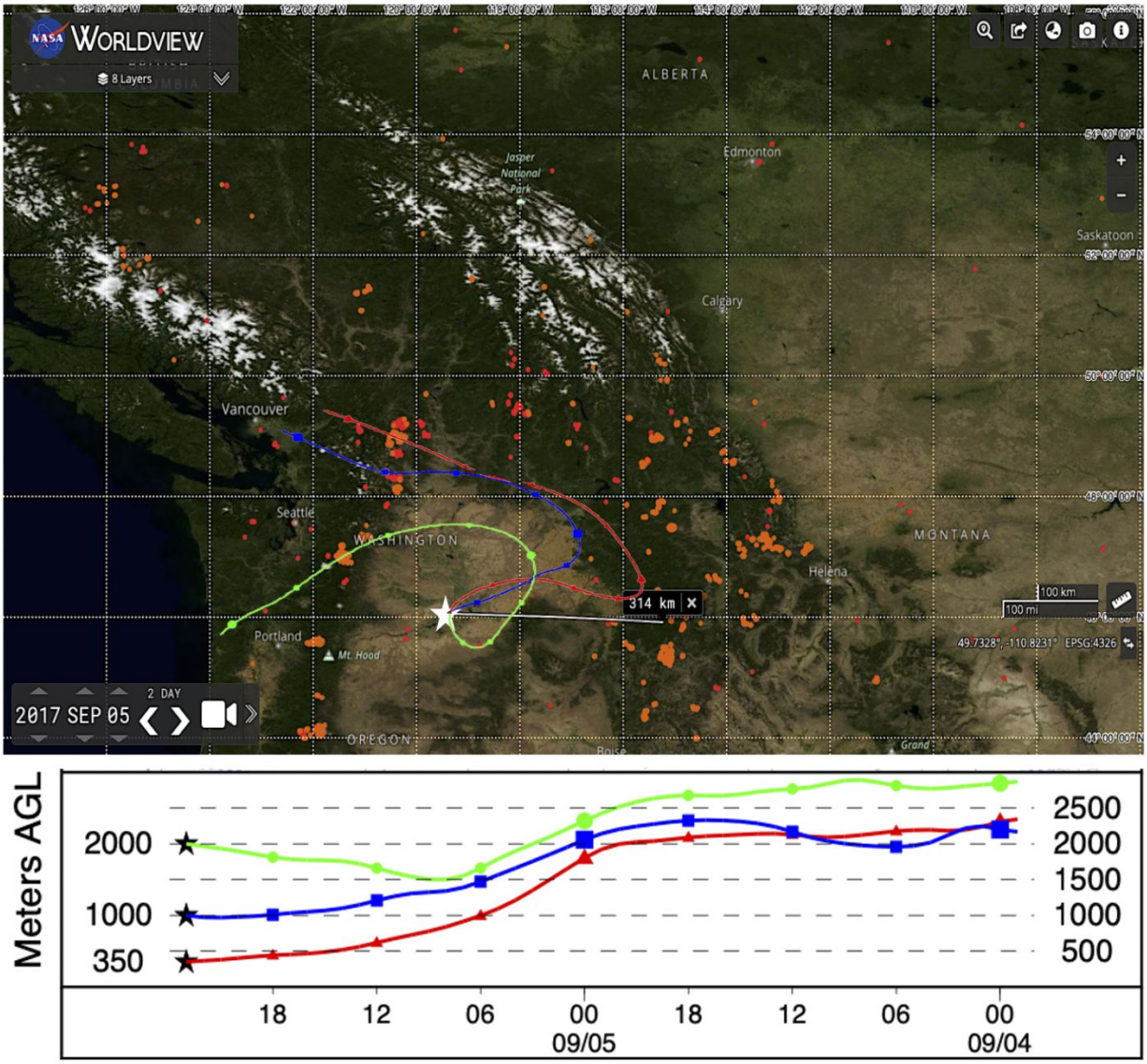

Figure S1 Backward trajectory of wind plumes extending 48-hours from the time of sample collection. The trajectory was modelled using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) at 350, 1000, and 2000 m above ground level and overlaid on fires and thermal anomalies (day and night) curated by NASA Worldview from Aqua/MODIS, Terra/MODIS, Suomi NPP/VIIRS for the same period.

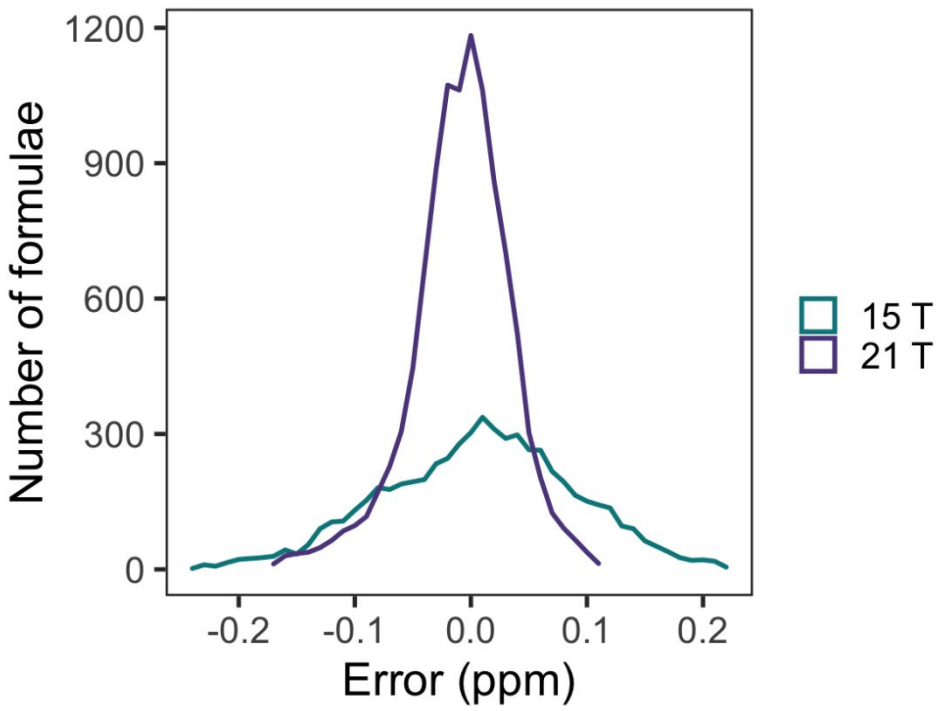

Figure S2 Error distribution of molecular formulae assigned to species detected by (-)ESI analysis with 15-T and 21-T FT-ICR mass spectrometers.

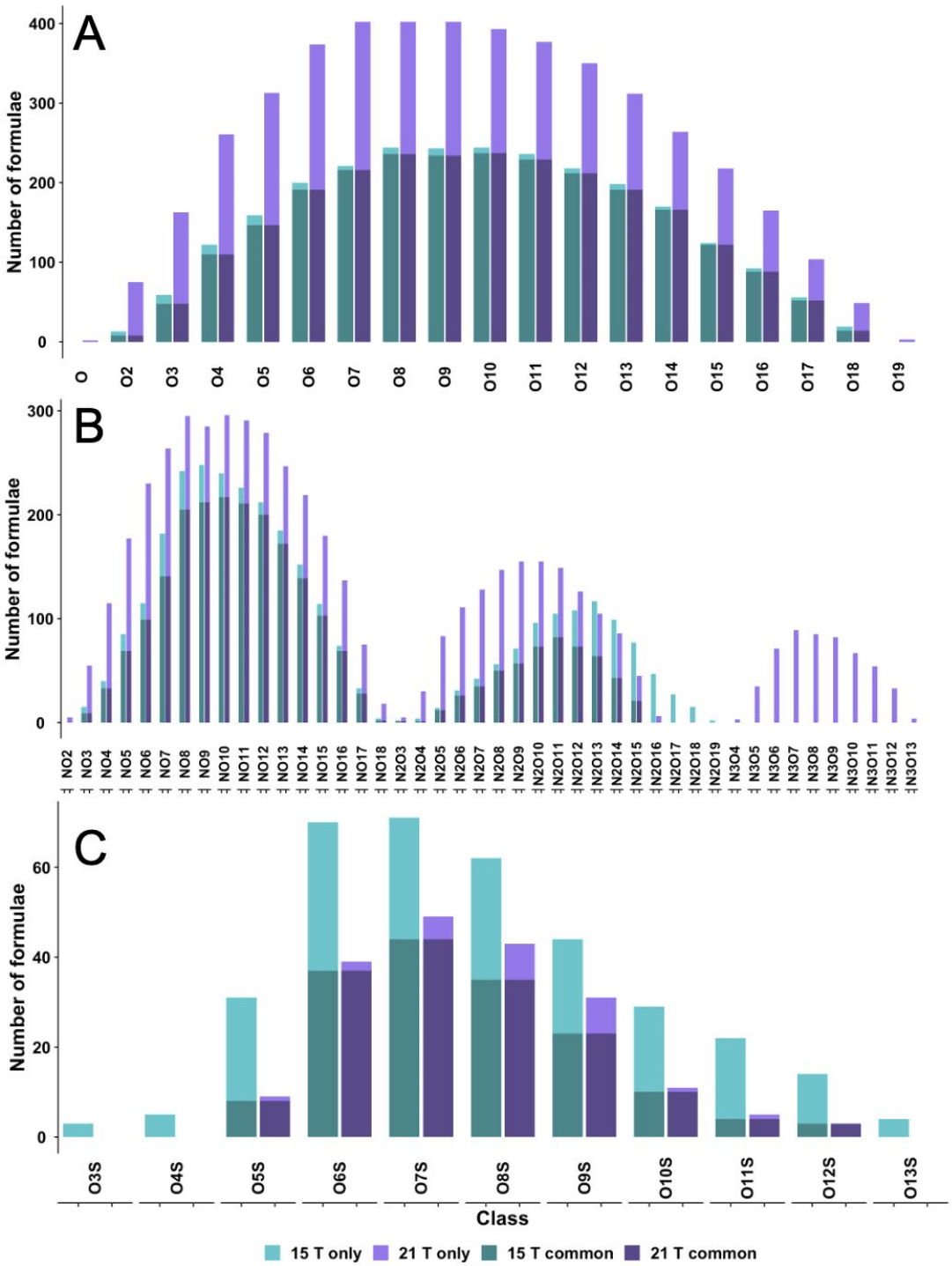

Figure S3 Class distributions of species belonging to (A) CHO, (B) CHNO, and (C) CHOS molecular groups in organic aerosol as observed by (-)ESI analysis with 15-T and 21-T FT-ICR mass spectrometers. Darker shades in the bars denote the number of molecular formulae that were commonly assigned in both analyses, while the lighter shades denote formulae that were uniquely observed in either the 15-T (turquoise) or the 21-T (violet). Class distributions for CHOP and CHNOP detected by the 21-T FT-ICR MS are shown in Figure 3.

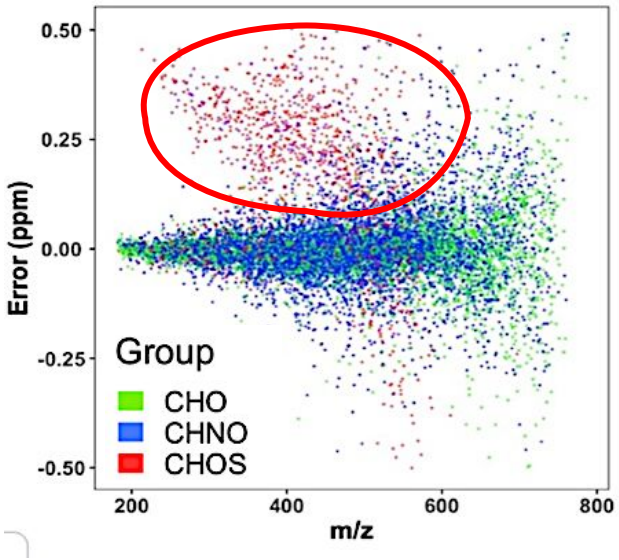

Figure S4 Error distribution of 'potential' CHO^{32}S assignments (with respect to the remaining dataset) that were considered as potential alternatives for CHOP and CHNOP formulae

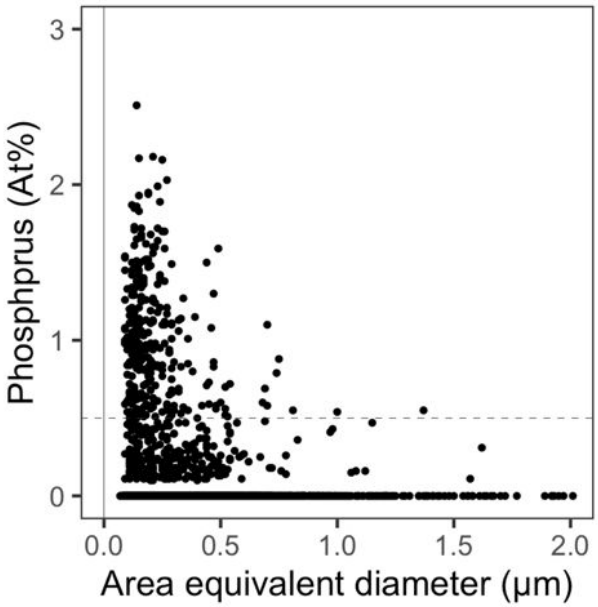

Figure S5 Phosphorus content by atomic percentage in organic aerosol sample detected from spectromicroscopy analyses.

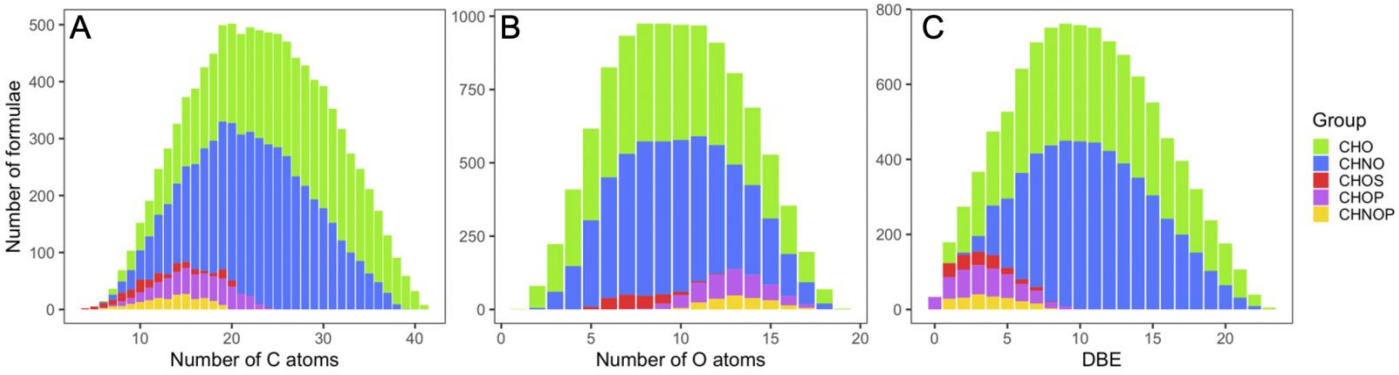

Figure S6 Distribution of formulae assigned to organic aerosol sample across integer values of (A-B) carbon and oxygen atoms and (C) double bond equivalence. Bars are coloured by molecular groups.

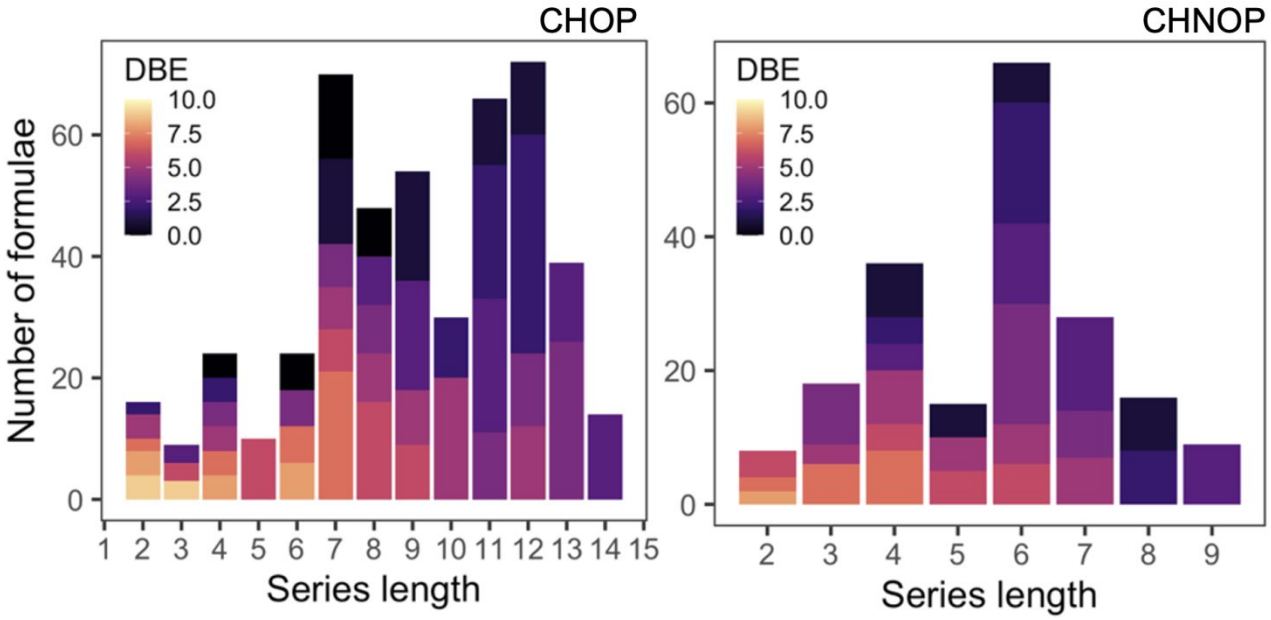

Figure S7 Number and length of CH_2-homologous series formed by all CHOP and CHNOP species. Corresponding Kendrick mass defect plots are drawn in Figures 3A-B.

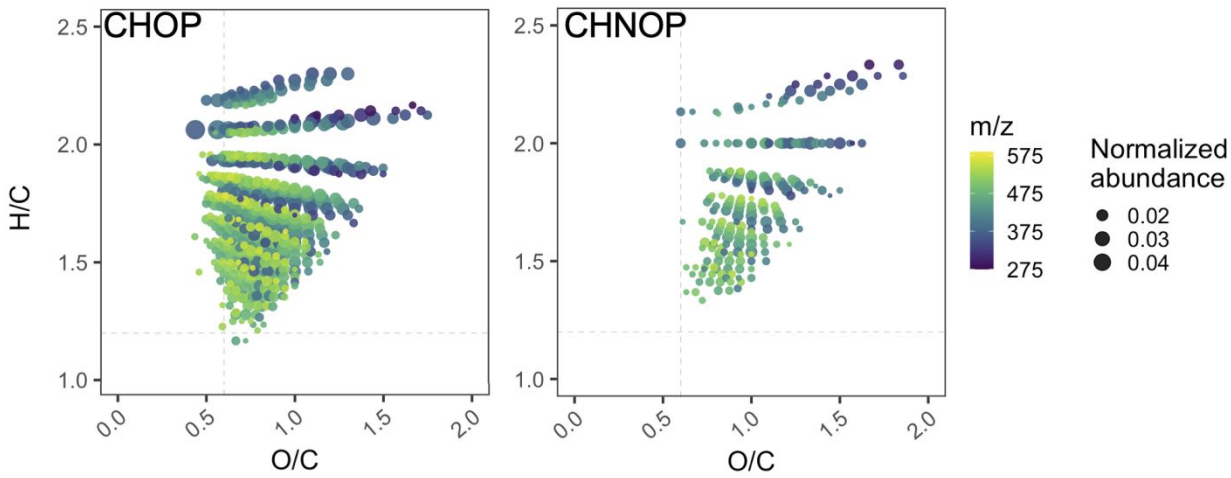

Figure S8 Arrangement of CHOP and CHNOP species in the van Krevelen space with data points coloured by m/z and sized to normalised abundance.

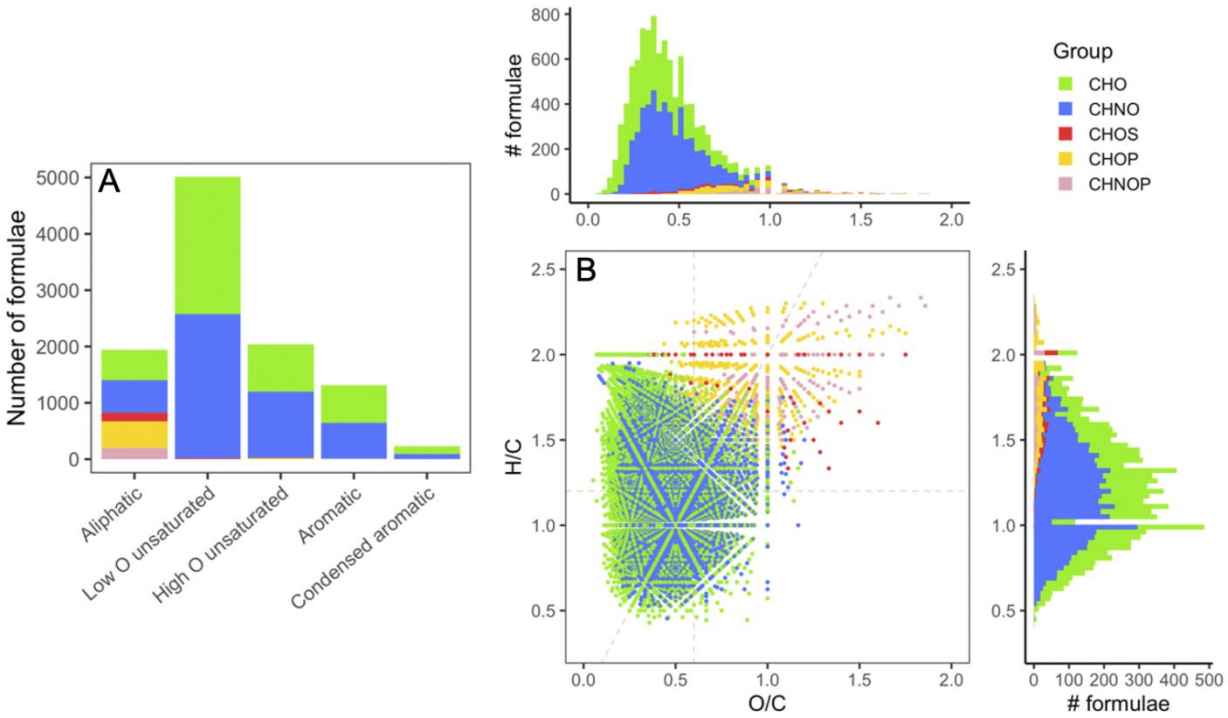

Figure S9 Distribution of formulae assigned to organic aerosol sample across estimated structural classes and elemental ratios: (A) Classification into five structural classes based on modified aromaticity index. Bars are coloured by molecular groups, and (B) van Krevelen diagram showing all molecular groups across H/C and O/C. Histograms are drawn for both axes to better represent data points that are overlaid on one another.

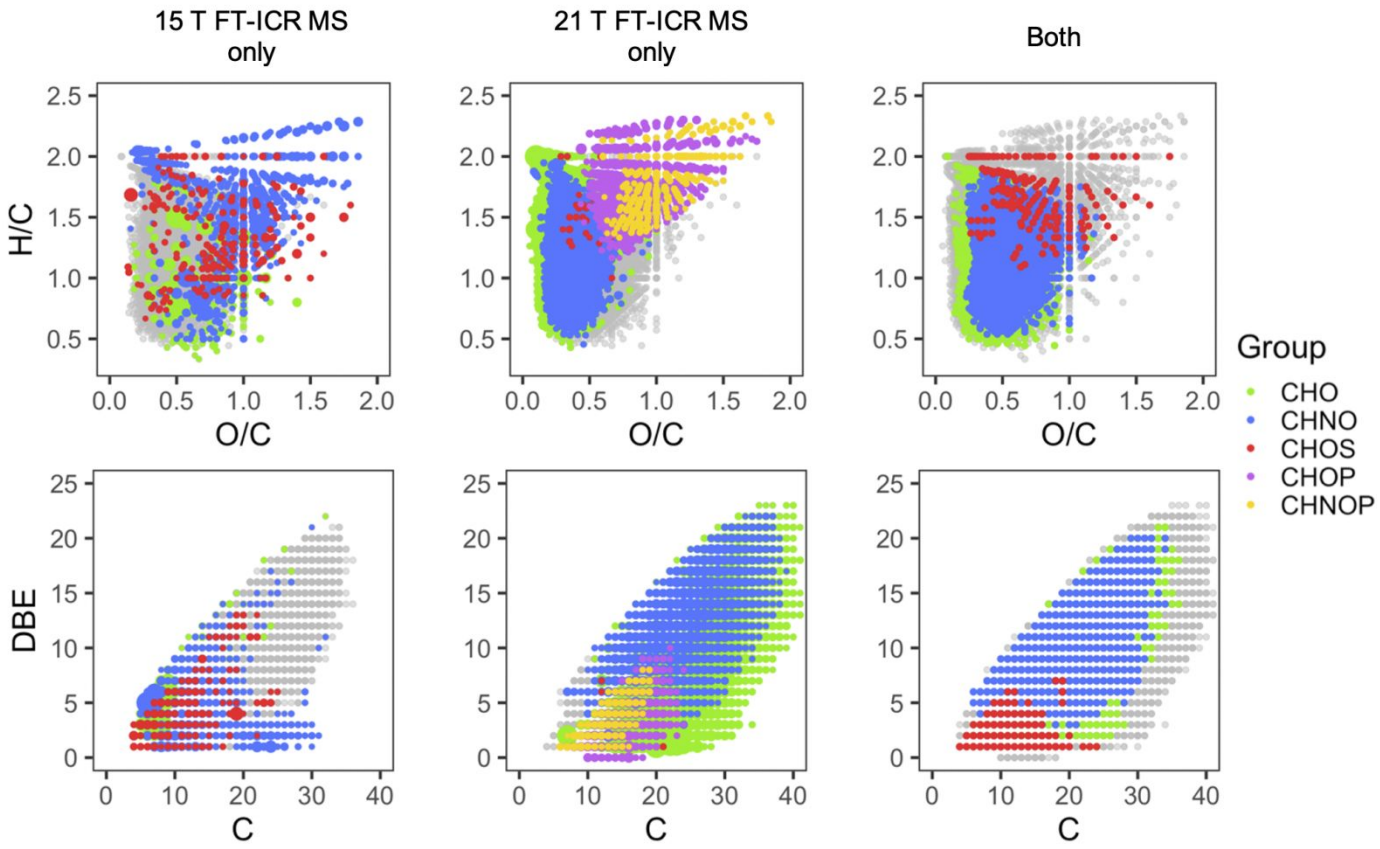

Figure S10 Molecular species identified exclusively or commonly by 15-T and 21-T FT-ICR mass spectrometers. The upper panel shows the distribution of such species in the van Krevelen space, and the bottom panel shows them in the DBE versus C plot. Grey data points in the background represent all species detected by a given mass spectrometer (the third column has both). Exclusive or common species are overlaid in colour, where the colours represent molecular groups. Data points are scaled to abundance in the first two columns.

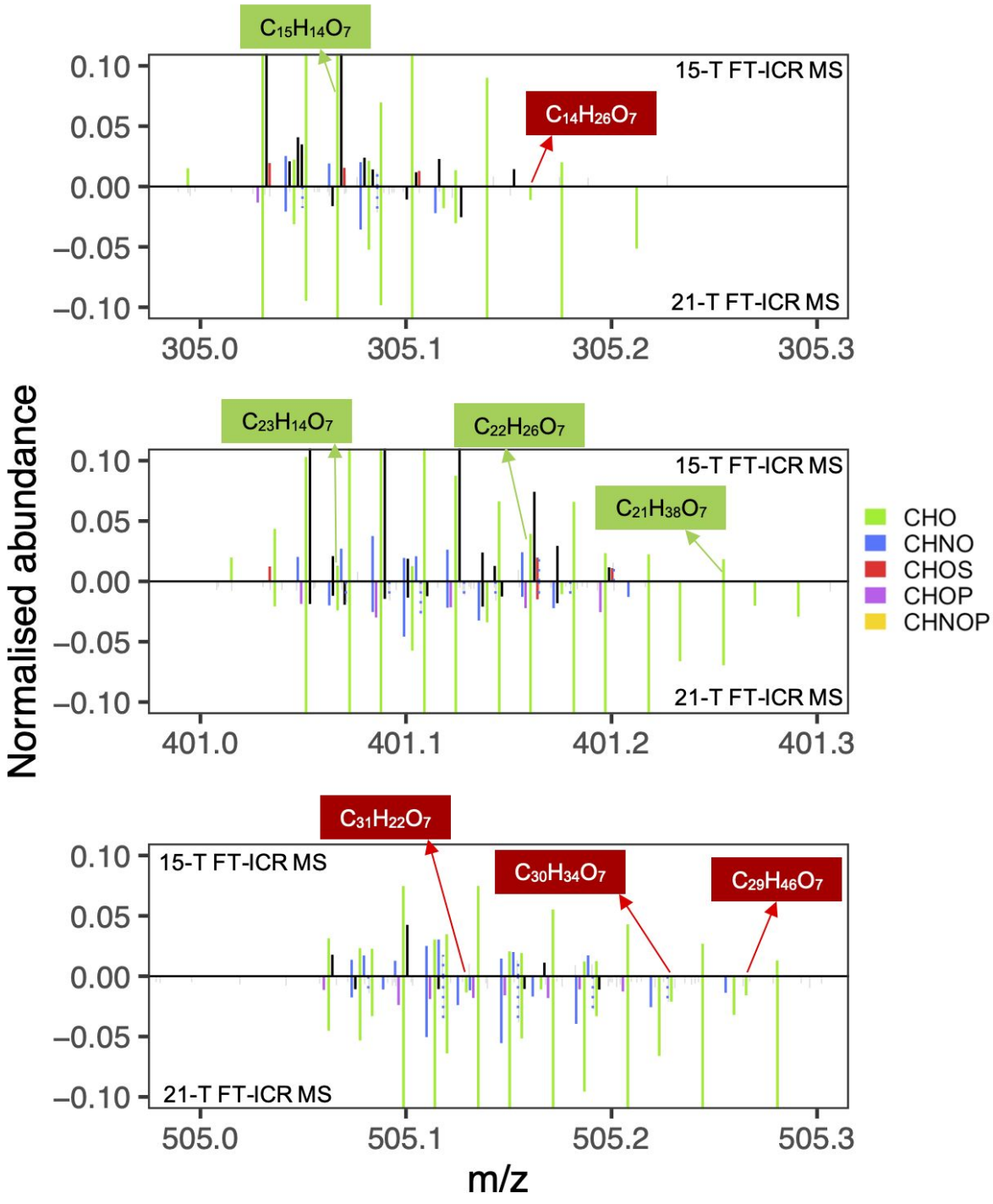

Figure S11 Reconstructed mass spectra at nominal masses: *m/z* 305, 401, and 505. In all panels, the top spectrum was acquired from 15-T FT-ICR MS and the bottom spectrum from 21-T FT-ICR MS. The spectra have been zoomed-in equally on both sides to aid visualisation, particularly for 15-T that had lower S/N at higher *m/z*. Species belonging to O_7 class are pinpointed with their neutral formulae. Formulae outlined in green were detected by both instruments, while those in red were detected only by the 21-T FT-ICR MS. Black peaks represent unassigned species.

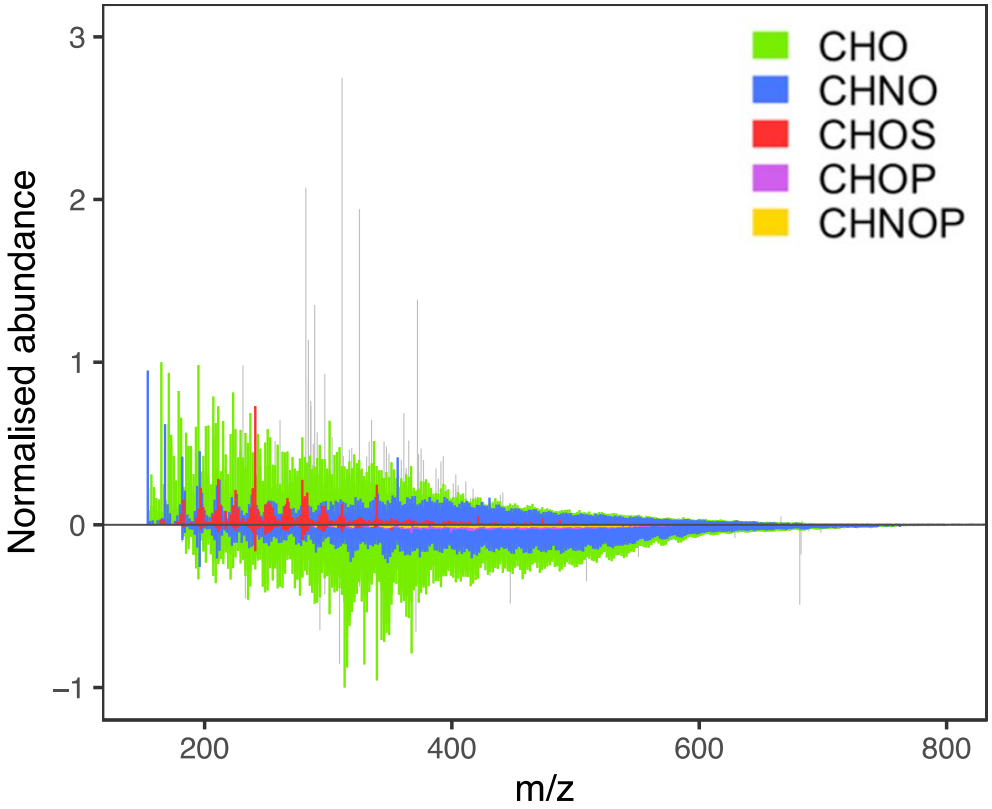

Figure S12 Reconstructed mass spectra of organic aerosol from (-)ESI analysis with 15-T (top) and 21-T (bottom) FT-ICR mass spectrometers. Peaks are coloured by the molecular group of the formula assigned to them. Grey peaks in the background are unassigned, below the S/N threshold, or contaminant (present in the blank with high relative abundances) peaks. Abundances in both mass spectra were normalised to the abundance of the tallest peak assigned a formula. As some very dominant peaks in the 15-T FT-ICR MS analysis were unassigned (shown in grey), their normalised abundances were >1.

Supplementary Tables

Table S1 Recalibrant series used to correct systematic bias in the exact masses of ions measured by (-)ESI FT-ICR MS analysis

Recalibrant series used			**Average formula error of preliminary C, H, and O-containing formulae (ppm)**	**Average formula error of preliminary C, H, and O-containing formulae after correction of mass (ppm)**	**Average formula error of final formula assignments, including C, H, O, N, S, and/or P-containing formulae (ppm)**
Adduct	Oxygen class	DBE			
-H (deprotonated)	O9	8	0.76 (15-T) 0.71 (21-T)	0.07 (15-T) 0.03 (21-T)	0.08 (15-T) 0.04 (21-T)
	O7	6			
	O6	7			
	O9	9			
	O7	7			
	O10	11			
	O5	5			
	O15	17			
	O16	15			

Table S2 Summary of bulk chemical metrics based on molecular formulae assigned to the organic aerosol sample. Here averages and standard deviations weighted to normalised abundance are presented, while the number average and standard deviations are presented in Table 1.

	Total number of peaks	**Total number of peaks above S/N**	**Unassigned peaks**	**Assigned monoisotopic formulae[a]**	**O/C_{avg}**	**H/C_{avg}**	**DBE_{avg}**	**m/z_{avg}**	**AE_{avg}**	**$AI_{mod,avg}$**
					Oxygen-to-carbon ratio	Hydrogen-to-carbon ratio	Double bond equivalence	Mass/charge	Absolute error (ppm)	Modified aromaticity index
15-T FT-ICR MS										
All MF	23701	9939	2447	6053	0.58 ± 0.2	1.13 ± 0.3	8.31 ± 3.6	363.17 ± 113.9	0.061 ± 0.04	0.33 ± 0.5
CHO				2618	0.54 ± 0.2	1.08 ± 0.3	8.52 ± 3.5	345.89 ± 114.4	0.061 ± 0.04	0.35 ± 0.2
CHNO				3080	0.62 ± 0.2	1.18 ± 0.3	8.60 ± 4.0	402.52 ± 105.2	0.062 ± 0.04	0.31 ± 0.8
CHOS				355	0.89 ± 0.3	1.58 ± 0.3	3.11 ± 1.9	282.87 ± 61.4	0.069 ± 0.04	0.12 ± 0.7
21-T FT-ICR MS										
All MF	34557	15312	1294	10533	0.44 ± 0.2	1.19 ± 0.3	9.80 ± 4.0	428.14 ± 112.0	0.025 ± 0.02	0.30 ± 0.2
CHO				4629	0.41 ± 0.2	1.20 ± 0.3	9.54 ± 4.0	416.48 ± 114.0	0.023 ± 0.02	0.31 ± 0.2
CHNO				5022	0.46 ± 0.1	1.13 ± 0.2	11.00 ± 3.5	454.03 ± 105.4	0.027 ± 0.02	0.32 ± 0.2
CHOS				190	0.75 ± 0.3	1.66 ± 0.2	2.93 ± 1.6	308.02 ± 63.8	0.022 ± 0.02	0.04 ± 0.4
CHOP				488	0.85 ± 0.2	1.77 ± 0.2	3.49 ± 2.1	450.63 ± 62.2	0.079 ± 0.04	0.00
CHNOP				204	1.03 ± 0.2	1.79 ± 0.2	3.66 ± 1.8	441.40 ± 56.6	0.078 ± 0.04	0.13 ± 1.3

[a]Number of formulae assigned to monoisotopic peaks with ^{12}C, ^{1}H, ^{16}O, ^{14}N, ^{32}S, and ^{31}P only.