



#### Supplement of

#### Molecular characterization of free tropospheric aerosol collected at the Pico Mountain Observatory: a case study with a long-range transported biomass burning plume

K. Dzepina et al.

Correspondence to: K. Dzepina (kdzepina@mtu.edu) and L. R. Mazzoleni (lrmazzol@mtu.edu)

#### **1. Details of the ultrahigh-resolution ESI FT-ICR MS analysis of WSOC**

#### 32 **1.1 Aerosol sample preparation**

33 The preparation of aerosol samples for FT-ICR MS analysis was done using Strata-X 34 (Phenomenex, Torrance, CA, USA) reversed-phase solid-phase extraction (SPE) to concentrate 35 the analytes and remove inorganic ions (Mazzoleni et al., 2010, 2012; Zhao et al., 2013). All 36 solvents used to prepare samples for ESI FT-ICR MS analysis were of HPLC grade purity or 37 higher. Strata-X SPE cartridges were conditioned with an application of three solvents in the 38 following order: 3 mL of methanol, 3 mL of acetonitrile and 3 mL of optima grade HPLC water. 39 Exactly half of 9/24 and 9/25 filters and a laboratory filter blank were cut into smaller pieces and 40 sonicated for 30 minutes twice, first time in 30 mL and second time in 20 mL water. The pH was 41 not adjusted. The 50 mL aqueous extracts were applied at a slow rate (approx. 1 mL min<sup>-1</sup>) to 42 enable aerosol organic species to partition to the SPE stationary phase. SPE cartridges were rinsed 43 with water and the WSOC species retained on the SPE stationary phase were eluted with a 2 mL 44 solution composed of water and acetonitrile in a 1:9 ratio. Prepared extracts were stored in a freezer 45 until cold overnight shipment on freezer packs to the Woods Hole Oceanographic Institution 46 (Woods Hole, MA) Fourier Transform Mass Spectrometry facility for analysis.

47

#### **1.2 ESI FT-ICR MS instrumental parameters**

Aerosol WSOC samples and blanks were analyzed at the Woods Hole Oceanographic Institution Fourier Transform Mass Spectrometry facility using the ultrahigh-resolution 7 tesla FT-ICR mass spectrometer (LTQ FT Ultra, Thermo Scientific, San Jose, CA) equipped with an ESI source. Samples were infused directly into the ESI interface and the ESI parameters were adjusted to obtain a stable ion current with minimum ion injection time into the mass analyzer. After optimization, the infusion flow rate was 4  $\mu$ L min<sup>-1</sup> and the ESI needle voltage was -3.9 kV. 54 Between measurements, the sample delivery apparatus was flushed with at least 500 µL of pure 55 water and acetonitrile and their solution in 1:9 ratio until background noise levels were reached. 56 Three replicate measurements of negative ion mass spectra were collected using full-scan analysis 57 for the range of m/z values 100 to 1000 with the mass resolving power set at 400,000 (at m/z 400). 58 Automatic gain control was used to consistently fill the linear ion trap quadropole mass analyzer 59 (LTQ) with the same number of ions  $(n = 10^6)$  for each acquisition and to avoid space charge 60 effects from overfilling the mass analyzer. The instrument was externally calibrated in negative 61 ion mode with a standard solution of sodium dodecyl sulfate and taurocholic acid. Mass accuracy 62 for the spectra was < 2 ppm. For each sample, more than 200 individual mass spectra recorded in 63 the time domain were collected and stored as transients using Thermo Xcalibur software.

#### 64 **1.3 Data processing and assignment of molecular formulas**

The analysis of ambient aerosol by FT-ICR MS method yields thousands of elemental 65 66 compositions per sample (e.g., Mazzoleni et al., 2010; 2012). The aim of data processing is to 67 interpret and represent data in an understandable and clear form. Transient co-addition and 68 molecular formula assignment of ultrahigh-resolution FT-ICR MS data was performed with 69 Composer software (Sierra Analytics, Modesto, CA version 1.0.5) as described in detail in 70 previous publications (Putman et al., 2012; Mazzoleni et al., 2012). Briefly, ~200 recorded 71 transients were first co-added for improved analyte detection (Kujawinski et al., 2002; Stenson et 72 al., 2003). Internal recalibration of the mass spectra was not done and the mass accuracy was 73 sufficient to assign the majority of the detected ions. The molecular formula calculator was set to 74 allow up to 100 carbon (C), 400 hydrogen (H), 100 oxygen (O), 3 nitrogen (N) and 1 sulfur (S) 75 atoms per molecular formula composition. The molecular formula calculator is based on the 76 PREDATOR algorithm (Blakney et al., 2011) and uses a Kendrick mass defect (KMD) analysis

(Hughey et al., 2001) to sort ions into user-defined homologous series. Kendrick transformation (Kendrick, 1963) helps in identifying homologous series of compounds that differ in number of base units (e.g., CH<sub>2</sub>, O<sub>x</sub> where x is an integer > 0, and CH<sub>2</sub>O) only. KMD for an elemental composition of a compound *C* observed in a mass spectrum containing base unit *B* (e.g., CH<sub>2</sub>) can be calculated from Kendrick Mass (KM) as:

82 
$$KM = C(observed mass) * \frac{B(nominal mass)}{B(exact mass)}$$
 (1)

$$83 KMD = Nominal Kendrick Mass - KM (2)$$

All molecular formulas presented here were assigned using the CH<sub>2</sub> homologous series. Note that accurate formula assignments can be achieved with molecular formula extensions other than CH<sub>2</sub> (e.g., Kujawinski and Behn, 2006). Homologous compounds that differ by number of base units only will have the same KMD. KMD for the entire mass spectrum and selected base unit can be visualized in a Kendrick mass plot, in which KMD is plotted *vs.* nominal Kendrick mass (NKM). This visual representation aids in recognition of elemental compositions of the same type and class.

A user-defined *de novo* cutoff for the homologous series is used in the molecular formula 91 92 assignments. De novo cutoff is the m/z value above which the molecular formula assignments are 93 made only for the ions that belong to homologous series starting below the cutoff. We used a two-94 step approach here following the work of Koch et al. (2007), who reported that unequivocal 95 molecular formula assignments can be determined for species having CHO elemental composition 96 for masses up to 1000 Da with small measurement error (<1 ppm). However, this is not possible 97 for species with elemental compositions including N, S and other heteroatoms. Therefore, we 98 performed the following two methods of molecular formula assignments. Method A: elemental 99 composition including C, H, O, N and S with a *de novo* cutoff at m/z 500; and Method B: elemental

100 composition including only C, H and O with *de novo* cut off at m/z 1000. The assigned molecular 101 formulas presented here are the combination of both methods without double counting. An 102 assignment threshold of six times the root mean square values of the signal-to-noise ratio between 103 m/z 988 - 990 was applied to the data. These assignment thresholds were 0.9%, 0.2% and 0.1% 104 (relative to the molecular ion peak at 100%) for 9/24, 9/25 and laboratory filter blank, respectively. 105 For the three replicates of each sample measured by FT-ICR MS, only the molecular formula 106 assignments found in all three replicate analyses were retained. Molecular formulas with 107 measurement errors larger than 3 ppm were discarded. Chemical rules and assumptions for 108 additional quality assurance of the assigned formulas were adopted from Koch et al. (2005) and 109 are described in detail in Putman et al. (2012).

110 Structural properties of molecular formulas can be inferred from the hydrogen deficiency, 111 which is defined as the difference between the number of H atoms in a molecule and its completely 112 saturated acyclic analogue. This difference corresponds to the number of double bonds and rings 113 in a molecule (McLafferty and Turecek, 1993; Lobodin et al., 2012) and is referred to as the double 114 bond equivalents (DBE) for a molecule in the general form of  $C_cH_hN_nO_oS_s$  as:

115 
$$DBE(C_c H_h N_n O_o S_s) = c - \frac{h}{2} + \frac{n}{2} + 1$$
 (3)

Another indicator of molecular saturation is the carbon-normalized DBE, defined as the
ratio of DBE to the number of C atoms (DBE/C) for each molecular assignment, i.e.:

118 
$$DBE/C = \frac{DBE(C_cH_hN_nO_oS_s)}{c(C_cH_hN_nO_oS_s)}$$
(4)

Hockaday et al. (2006) introduced DBE/C as a criterion for identifying molecular species with condensed aromatic ring structures. The value of DBE/C  $\geq 0.7$  is a threshold for molecular assignments with fused-ring structures, and thus a proxy for combustion derived species and their degradation products (Hockaday et al., 2006). 123 Another metric to estimate the degree of molecular saturation and structure is the 124 aromaticity index (AI). AI is defined as the minimum number of C-C double bonds and rings 125 required to rationalize these molecular features based on elemental composition (Koch and 126 Dittmar, 2006). AI is calculated for  $C_cH_hN_nO_oS_s$  molecular compositions by taking into account 127 the C bonds with heteroatoms as:

128 
$$AI = DBE_{AI} / C_{AI} = (1 + c - o - s - 0.5 * h) / (c - o - s - n)$$

The AI calculation of Koch and Dittmar (2006) assumes that all oxygen atoms are in the form of carbonyl C=O bonds and provides a conservative lower boundary for the aromaticity of the molecular formulas. This is not true for aerosol WSOM where carboxyl and alcohol (as well as nitrate and sulfate) functional groups are abundant. Thus, the molecular formulas may be more aromatic than indicated using this index.

134 The elemental composition from each of the molecular assignments, accessible with 135 ultrahigh-resolution of FT-ICR MS, enables the calculation of their O/C, H/C and OM/OC ratios. 136 O/C (H/C) elemental ratio is calculated as the ratio of the number of O (H) vs. C atoms directly 137 from detected elemental composition. Elemental ratios determined from the assigned molecular 138 formulas of the ESI FT-ICR mass spectrum are visualized using van Krevelen diagrams (Kim et 139 al., 2003) with the H/C ratio plotted against the O/C, or N/H or S/H ratio (Wu et al., 2004). OM/OC 140 ratio is calculated as the ratio between measured or estimated mass and calculated mass of carbon 141 in assigned molecular formulas. Properties of samples can also be evaluated by relative abundance 142 weighting, which is determined by Equations (6)-(10) for each elemental composition 143  $C_c H_h N_n O_o S_s$ :

144 
$$O/C_w = \frac{\sum w_i * o_i}{\sum w_i * c_i}$$
(6)

145 
$$H/C_w = \frac{\sum w_i * o_i}{\sum w_i * h_i}$$
(7)

6

(5)

146 
$$DBE_{w} = \frac{\sum w_{i} * DBE_{i}}{\sum w_{i}}$$
(8)

147 
$$OM/OC_w = \frac{\sum w_i * OM/OC_i}{\sum w_i}$$
(9)

148 
$$DBE/C_w = \frac{\sum w_i * DBE_i}{\sum w_i * c_i}$$
(10)

where  $w_i$  is the relative abundance for each individual molecular formula *i*. The sum includes all molecular formula assignments in a sample or its subset. Error of the weighted values is calculated by propagating the standard deviation (1 $\sigma$ ) of the relative abundance used to weight the values for the group (Mazzoleni et al., 2012; Zhao et al., 2013).

153 Due to the presence of thousands of molecular assignments, it is difficult to see the details 154 of the graphical representations of data available with ultrahigh resolution mass spectrometry. 155 Therefore, Fig. S9 presents the examples of details in graphical representations for monoisotopic 156 molecular assignments detected in WSOM for 9/24. Fig. S9a shows an excerpt of the Kendrick 157 plot to illustrate the molecular complexity. Numerous molecular assignments are apparent and the 158 separation in homologous series is evident with the same KMD differing by number of CH<sub>2</sub> base 159 units only. As examples, we highlighted random three homologous series belonging to the CHO, 160 CHNO and CHOS groups: (CH<sub>2</sub>)<sub>1-19</sub>C<sub>3</sub>H<sub>4</sub>O<sub>4</sub>, (CH<sub>2</sub>)<sub>1-5</sub>C<sub>6</sub>H<sub>3</sub>NO<sub>2</sub>, and (CH<sub>2</sub>)<sub>1-2,5,7</sub>C<sub>7</sub>H<sub>16</sub>SO<sub>4</sub>, 161 respectively. Next, molecular assignments in the van Krevelen diagram are aligned to lines 162 reflecting different chemical trends and reactions. Examples of those lines are visualized in Fig. 163 S9b, including: i) saturation or unsaturation; ii) oxidation or reduction; and iii) functionalization 164 or fragmentation. Finally, numerous homologous series of species overlap at each DBE value. 165 Figure S9c highlights random examples of homologous series with DBE values of 10-12 for CHO 166 (C14H10O10(CH2)n, C15H10O11(CH2)n and C14H6O7(CH2)n,) and CHNO groups (C14H11NO10(CH2)n, 167  $C_{16}H_{13}NO_{11}(CH_2)_n$  and  $C_{16}H_{11}NO_{11}(CH_2)_n$ ).

#### 168 2. Additional References to the Supplement

- Kendrick, E.: A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of
   organic compounds, Anal. Chem., 35, 2146–2154, 1963.
- 171 Kim, S., Kramer, R.W., and Hatcher, P.G.: Graphical Method for Analysis of Ultrahigh-Resolution
  172 Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram, Anal. Chem.,
  173 75, 5336-5344, 2003.
- Koch, B. P., Witt, M. R., Engbrodt, R., Dittmar, T., and Kattner, G.: Molecular Formulas of Marine
  and Terrigenous Dissolved Organic Matter Detected by Electrospray Ionization Fourier
  Transform Ion Cyclotron Resonance Mass Spectrometry, Geochim. Cosmochim. Acta, 69,
  3299-3308, 2005.
- Koch, B.P.and Dittmar, T.: From mass to structure: an aromaticity index for high-resolution mass
  data of natural organic matter, Rapid Commun. Mass Spectrom., 20, 926-932,
  doi:10.1002/RCM.2386, 2006.
- 181 Koch, B. P., Dittmar, T., Witt, M., and Kattner, G.: Fundamentals of Molecular Formula
  182 Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter, Anal. Chem., 79,
  183 1758-1763, 2007.
- Kujawinski, E. B., Hatcher, P. G., and Freitas, M. A.: High-Resolution Fourier Transform Ion
  Cyclotron Resonance Mass Spectrometry of Humic and Fulvic Acids, Improvements and
  Comparisons, Anal. Chem., 74, 413–419, 2002.
- Kujawinski, E.B., and Behn, M.D.: Automated Analysis of Electrospray Ionization Fourier
   Transform Ion Cyclotron Resonance Mass Spectra of Natural Organic Matter, Anal. Chem., 78,
   4363-4373, 2006.
- Lobodin, V. V., Rodgers, R. P., and Marshall, A. G.: Petroleomics and the analysis of complex
  organic mixtures with Fourier transform ion cyclotrone resonance, in: Comprehensive
  Environmental Mass Spectrometry, edited by: Lebedev, A. T., ILM Publications, a trading
  division of International Labmate Ltd., St. Albans, Hertfordshire, United Kingdom, 415–442,
  2012.
- McLafferty, F.W., and Turecek, F.: Interpretation of Mass Spectra, 4th Edn., University Science
   Books, Sausalito, CA, 1993.
- Putman, A. L., Offenberg, J. H., Fisseha, R., Kundu, S., Rahn, T. A., and Mazzoleni, L. R.:
  Ultrahigh-resolution FT-ICR mass spectrometry characterization of a-pinene ozonolysis SOA,
  Atmos. Environ., 46, 164–172, 2012.
- Wu, Z., Rodgers, R.P. and Marshall, A.G.: Two- and Three-Dimensional van Krevelen Diagrams:
   A Graphical Analysis Complementary to the Kendrick Mass Plot for Sorting Elemental
   Compositions of Complex Organic Mixtures Based on Ultrahigh-Resolution Broadband
   Fourier Transform Ion Cyclotron Resonance Mass Measurements, Anal. Chem., 76, 2511-2516,
   204
- 205

206

### Molecular characterization of free tropospheric aerosol collected at the Pico Mountain Observatory: A case study with a long-range transported biomass burning plume

K. Dzepina, C. Mazzoleni, P. Fialho, S. China, B. Zhang, R.C. Owen, D. Helmig, J. Hueber, S. Kumar, J.A.
Perlinger, L.J. Kramer, M.P. Dziobak, M.T. Ampadu, S. Olsen, D.J. Wuebbles and L. R. Mazzoleni

> Correspondence to: kdzepina@mtu.edu lrmazzol@mtu.edu

# **Figures to the Supplement**



**Figure S1.** (a): Location of the Pico Island of the Azores Archipelago in the North Atlantic Ocean (photo credit: Google maps). Volcano on the Pico Island (b) (photo credit: Christer Johansson, Wikipedia) on the top of which is a crater where the Pico Mountain Observatory is located (c) (photo credit: Claudio Mazzoleni).



**Figure S2.** Scatter plots of total filter-collected aerosol mass ( $OM + EC + SO_4^{2-} + NO_3^{-} + Cl^{-}$ ) shown in Figure 1a *vs.* on-line aerosol measurements of BC mass concentration (**a**), aerosol light scattering (**b**) and number concentration (**c**) averaged over filter-collection periods. Linear regression best fit intercepts, slopes and correlation coefficients as well as their standard errors are reported.



**Figure S3.** MODIS satellite fire counts for the period of September 10-17. MODIS fire counts are shown with the size of symbols denoting fire radiative properties in MW (**a**) and brightness temperature in K (**b**).



**Figure S4.** Simulated FLEXPART retroplumes for 9/24 sample with 3-hrs time resolution; first simulation is for Sep 24 at 15:00 UTC (a) and the last one for Sep 25 at noon (h). Residence times of the retroplumes are normalized by each vertical intervals in the lower plots for better visualization of the distribution along an irregular height scale. The upwind time in days at each location of the plume is labeled by white numbers. Filter-collection period for this sample is Sep 24, 2012 at 15:00 - Sep 25, 2012 at 15:00 (Table 1).



**Figure S5.** Simulated FLEXPART retroplumes for 9/25 sample with 3-hrs time resolution; first simulation is for Sep 25 at 15:00 UTC (a) and the last one for Sep 26 at noon (h). Residence times of the retroplumes are normalized by each vertical intervals in the lower plots for better visualization of the distribution along an irregular height scale. The upwind time in days at each location of the plume is labeled by white numbers. Filter-collection period for this sample is Sep 25, 2012 at 15:00 - Sep 26, 2012 at 15:00 (Table 1).



**Figure S6.** FLEXPART CO spectra for September 2012, as determined by multiplying the FLEXPART retroplumes with emission inventories. Fire CO includes contributions from the whole north hemisphere while the anthropogenic CO is further divided into the continent regions. Please refer to the text in Section 2.4 for the details of this calculation.



**Figure S7.** Time series of non-methane hydrocarbons ethane and propane and their ln(ethane/propane) ratio (**f**), as well as ozone (**g**). Also shown are measured mass concentrations of the filter-collected species as a percentage of the total (**b**) and all aerosol measurements shown in Fig. 1 for comparison. On-line measurements of gas-phase species are averaged for the sampling periods overlapping with filter measurements (**a**).



**Figure S8.** Mass resolved H/C elemental ratio diagram for all of the monoisotopic molecular assignments in the Pico WSOM 9/24 (**a**) and 9/25 (**b**) samples. Mass resolved H/C elemental ratio diagram is also shown for common and unique formula assignments between the two Pico samples (**c**).



**Figure S9.** Illustration of chemical composition complexity available with FT-ICR MS analysis for molecular assignments detected in WSOM of 9/24 and shown in Figs 6a-c. Kendrick mass defect *vs.* nominal Kendrick mass for homologous series of CHO, CHNO and CHOS groups, which are highlighted with arrows (**a**). van Krevelen diagram with different chemical trends highlighted with lines and marked with letters (i), (ii) and (iii) (**b**). Double bond equivalents *vs.* number of carbon atoms with examples for CHO and CHNO groups given above and below, respectively, the DBE value (**c**). The size of markers in (**c**) denotes relative abundance. See text in the Supplement for the details.



**Figure S10.** Sum of the relative abundance *vs.* number of oxygen atoms for CHNO compounds with one and two N atoms ( $N_1$  and  $N_2$ ) in the 9/24 (**a**) and 9/25 (**b**) WSOM samples. Sum of relative abundances is also normalized to the number of compounds with a certain oxygen number (x axis).



**Figure S11.** Sum of the relative abundance *vs.* number of oxygen atoms for CHOS compounds with one S atoms (S<sub>1</sub>) in the 9/24 (**a**) and 9/25 (**b**) WSOM samples. Sum of relative abundances is also normalized to the number of compounds with a certain oxygen number (x axis).