Molecular composition of organic aerosols in central Amazonia: an ultra-high-resolution mass spectrometry study

Ivan Kourtchev1,2, Ricardo H. M. Godoi3, Sarah Connors1, James G. Levine4, Alex T. Archibald1–5, Ana F. L. Godoi3, Sarah L. Paralovo3, Cybelli G. G. Barbosa3, Rodrigo A. F. Souza6, Antonio O. Manzi7, Roger Seco8, Steve Sjostedt9, Jeong-Hoo Park10, Alex Guenther8,11, Saewung Kim8, James Smith12,13, Scot T. Martin14,15, and Markus Kalberer1

1Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK
2Department of Chemistry and Environmental Research Institute, University College Cork, Cork, Ireland
3Environmental Engineering Department, Federal University of Parana, Curitiba, Brazil
4School of Geography Earth & Environmental Sciences, University of Birmingham, Birmingham, B15 2TT, UK
5NCAS climate, University of Cambridge, Cambridge, CB2 1EW, UK
6State University of Amazonas, Av. Darcy Vergas, 1200, 69065-020, Manaus-AM, Brazil
7Instituto Nacional de Pesquisas da Amazônia (INPA), Clima e Ambiente (CLIAMB), Manaus-AM, Brazil
8Department of Earth System Science, University of California, Irvine, CA 92697, USA
9NOAA ESRL Chemical Sciences Division, Boulder, CO, USA
10Air Quality Forecasting Center, National Institute of Environmental Research, Republic of Korea
11Pacific Northwest National Laboratory, Richland, WA, USA
12Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, CO, USA
13Department of Chemistry, University of California, Irvine, CA, USA
14School of Engineering and Applied Science, Harvard University, Cambridge, MA 02138, USA
15Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA

Correspondence to: Ivan Kourtchev (i.kourtchev@ucc.ie) and Markus Kalberer (markus.kalberer@atm.ch.cam.ac.uk)

Received: 14 May 2016 – Published in Atmos. Chem. Phys. Discuss.: 25 May 2016
Revised: 12 August 2016 – Accepted: 31 August 2016 – Published: 23 September 2016

Abstract. The Amazon Basin plays key role in atmospheric chemistry, biodiversity and climate change. In this study we applied nanoelectrospray (nanoESI) ultra-high-resolution mass spectrometry (UHRMS) for the analysis of the organic fraction of PM2.5 aerosol samples collected during dry and wet seasons at a site in central Amazonia receiving background air masses, biomass burning and urban pollution. Comprehensive mass spectral data evaluation methods (e.g. Kendrick mass defect, Van Krevelen diagrams, carbon oxidation state and aromaticity equivalent) were used to identify compound classes and mass distributions of the detected species. Nitrogen- and/or sulfur-containing organic species contributed up to 60% of the total identified number of formulae. A large number of molecular formulae in organic aerosol (OA) were attributed to later-generation nitrogen- and sulfur-containing oxidation products, suggesting that OA composition is affected by biomass burning and other, potentially anthropogenic, sources. Isoprene-derived organosulfate (IEPOX-OS) was found to be the most dominant ion in most of the analysed samples and strongly followed the concentration trends of the gas-phase anthropogenic tracers confirming its mixed anthropogenic–biogenic origin. The presence of oxidised aromatic and nitro-aromatic compounds in the samples suggested a strong influence from biomass burning especially during the dry period. Aerosol samples from the dry period and under enhanced biomass burning conditions contained a large number of molecules with high carbon oxidation state and an increased number of aromatic compounds compared to that from the wet period. The results of this work demonstrate that the studied site is influenced not only by biogenic emissions from the forest but also by biomass burning and potentially other anthropogenic emissions from the neighbouring urban environments.
1 Introduction

The Amazon Basin plays key role in atmospheric chemistry, biodiversity and climate change (Keller et al., 2009; Andreae et al., 2015). The Amazon rainforest is an important source of biogenic volatile organic compound (BVOC) emissions to the atmosphere (Greenberg et al., 2004; Alves et al., 2016), which give rise to secondary organic aerosol (SOA) through reaction with atmospheric oxidants (i.e. O_3, OH• and NO_2•) (e.g. Martin et al., 2010). SOA particles scatter and absorb solar and terrestrial radiation, influence cloud formation, and participate in chemical reactions in the atmosphere and are thus suggested to play an important role in climate change (Andreae and Crutzen, 1997; Haywood and Boucher, 2000; Hallquist et al., 2009; Pöschl et al., 2010). Aerosol optical properties, which govern the ability to absorb solar radiation, strongly depend on SOA composition (Laskin et al., 2015). It has been shown that organic nitrates, nitrooxy organosulfates and organic sulfates may contribute to light absorption by SOA (e.g. Song et al., 2013; Jacobson, 1999; Lu et al., 2011; Laskin et al., 2015). Chemical interactions between anthropogenic and biogenic aerosol precursors can play a significant role in the formation of SOA (Goldstein et al., 2009; Hoyle et al., 2011; Kleinman et al., 2015). For example, anthropogenic nitrogen oxides (NO_x) and sulfur dioxide (SO_2) are shown to react with a range of BVOCs, leading to formation of organic nitrates (e.g. Roberts, 1990; Day et al., 2010; Fry et al., 2014), nitroxy organosulfates and organosulfates (Surratt et al., 2008; Budisulistiorini et al., 2015). Much remains to be explored in terms of the molecular diversity of these compounds in the atmosphere.

A comprehensive knowledge of aerosol molecular composition, which in turn leads to a better understanding of aerosol sources, is required for the development of effective air pollution mitigation strategies. However, identification of the organic aerosol composition remains a major analytical challenge (Nozière et al., 2015). Organic aerosol is composed of thousands of organic compounds, which cover a wide range of physical and chemical properties (Goldstein and Galbally, 2007), making it difficult to find a single analytical technique for a detailed chemical analysis at the molecular level. Methods based on ultra-high-resolution mass spectrometry (UHRMS) have shown great potential in solving this longstanding problem. UHR mass spectrometers (e.g. Fourier transform ion cyclotron resonance mass spectrometer and Orbitrap mass spectrometer) have mass resolution power that is at least 1 order of magnitude higher (≥ 100 000) than conventional MS and high mass accuracy (< 5 ppm) and thus, when coupled with soft ionisation techniques (e.g. electrospray ionisation, ESI), can provide a detailed molecular composition of the organic aerosol (Nizkorodov et al., 2011; Nozière et al., 2015). Direct infusion ESI-UHRMS has been applied successfully for the analysis of aerosol samples from remote (e.g. boreal forest in Finland, Pico Island of the Azores archipelago), rural (e.g. Millbrook, USA; Harcum, USA; K-Puszta, Hungary) and urban (e.g. Cambridge, UK; Birmingham, UK; Cork, Ireland; Shanghai, China; and Los Angeles, USA) locations (Wozniak et al., 2008; Schmitt-Kopplin et al., 2010; Kourtchev et al., 2013, 2014; Tao et al., 2014; Dzepina et al., 2015). UHRMS has proven to be extremely useful in assessing chemical properties of the SOA.

The aim of this study was to investigate the detailed molecular composition of organic aerosol from a site that received air masses from a wide range of origins, including the background atmosphere of Amazonia, biomass burning and urban pollution plumes. The measurements were performed as a part of the Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5) campaign (Martin et al., 2016). The location of the research site where aerosol was collected for this study is ~69 km downwind of Manaus (population 2 million), intersecting background and polluted air with day-to-day variability in the position of the Manaus plume. The study designed served as a laboratory for investigating anthropogenic perturbations to biogenic processes and atmospheric chemistry.

2 Methods

2.1 Sampling site

Aerosol sampling was conducted at site “T3” of GoAmazon2014/5 located at ~3.2133° and ~60.5987°. The T3 site is located in the pasture area, ~2.5 km from the rainforest. The air masses arriving to the sampling site often passed over the single large city (Manaus) in the region. Detailed descriptions of the site and instrumentation are provided in Martin et al. (2016).

PM_{2.5} aerosol samples were collected on 47 mm polycarbonate filters (Nuclepore) using a Harvard impactor (Air Diagnostics, Harrison, ME, USA) with flow rate of 10 L min^{-1} from 5 to 26 March 2014 and 5 September to 4 October 2014, which were during intensive observation periods 1 and 2 (IOP1 and IOP2) of GoAmazon2014/5, respectively, corresponding to the traditional periods of wet and dry seasons of Amazonia. The sampling durations are shown in Table S1 in the Supplement. The airflow through the sampler was approximately 10 L min^{-1}. After collection, the aerosol samples were transferred into Petri dishes and stored in the freezer at ~4 °C until analysis.

2.2 Aerosol sample analysis

Fifteen samples, 5 from IOP2 and 10 from IOP1, were extracted and analysed using a procedure described elsewhere (Kourtchev et al., 2014, 2015). Depending on the aerosol loading of the analysed samples, a part (half to whole) of the filter was extracted in methanol (Optima TM LC/MS grade, Fisher Scientific) in a chilled ice slurry, filtered through a Teflon filter (0.2 μm, ISODiscTM, Supelco) and reduced by
volume using a nitrogen line to achieve approximately 0.3 µg of aerosol per microlitre of methanol. Several samples with the highest aerosol loading were divided into two parts for both direct infusion and LC/MS analyses, while the samples with the lowest loading were only analysed using direct infusion analysis. The LC/MS portion was further evaporated to 20 µL and diluted to 100 µL by aqueous solution of formic acid (0.1%). The final extracts were analysed as described in Kourtchev et al. (2013) using a high-resolution LTQ Orbitrap Velos mass spectrometer (Thermo Fisher, Bremen, Germany) equipped with an ESI and a TriVersa NanoMate robotic nanoflow chip-based ESI (Advion Biosciences, Ithaca NY, USA) source. The Orbitrap MS was calibrated using an Ultramark 1621 solution (Sigma-Aldrich, UK). The mass accuracy of the instrument was below 1 ppm. The instrument mass resolution was 100,000 at m/z 400. The ion transmission settings were optimised using a mixture of camphor sulfonic acid (20 ng µL⁻¹), glutaric acid (30 ng µL⁻¹), and cis-pinnonic acid (30 ng µL⁻¹) in methanol and Ultramark 1621 solution.

2.3 Direct infusion UHRMS analysis

The ionisation voltage and back pressure of the nanoESI direct infusion source were set at −1.4 kV and 0.8 psi, respectively. The inlet temperature was 200 °C and the sample flow rate was approximately 200–300 nL min⁻¹. The negative ionisation mass spectra were collected in three replicates at two mass ranges (m/z 100–650 and m/z 150–900) and processed using Xcalibur 3.1 software (Thermo Fisher Scientific Inc.). Similar to our preceding studies (Kourtchev et al., 2015), the average percentage of common peaks between analytical replicates was ~80%. This is also in agreement with literature reports for similar data analysis (Sleigher et al., 2012). The identification of IEPOX organosulfates was performed by comparing MS fragmentation patterns and chromatographic elution time with a synthesised IEPOX-OS standard which was provided by Dr Surratt from the University of North Carolina. It must be noted that, due to competitive ionisation of analytes in the direct infusion ESI analysis of the samples with a very complex matrix (i.e., aerosol extracts), the ion intensities do not directly reflect the concentration of the molecules in the sample (Oss et al., 2010); therefore, data shown in this work are semi-quantitative.

2.4 LC/MS analysis

LC/MS ESI parameters were as follows: spray voltage, −3.6 kV; capillary temperature, 300 °C; sheath gas flow, 10 arbitrary units, auxiliary gas flow, 10 arbitrary units; sweep gas flow rate, 5 arbitrary units; and S-lens RF level, 58%. LC/(−)ESI-MS analysis was performed using an Accela system (Thermo Scientific, San Jose, USA) coupled with LTQ Orbitrap Velos MS and a T3 Atlantis C18 column (3 µm; 2.1 × 150 mm; Waters, Milford, USA). The sample extracts were injected at a flow rate of 200 µL min⁻¹. The mobile phases consisted of 0.1% formic acid (v/v) (A) and methanol (B). The applied gradient was as follows: 0–3 min 3 % B, 3–25 min from 3 to 50 % B (linear), 25–43 min from 50 to 90 % B (linear), 43–48 min from 90 to 3 % B (linear), and then kept for 12 min at 3 % B. The collision-induced dissociation (CID) settings for MS/MS analysis are reported in Kourtchev et al. (2015).

2.5 High-resolution MS data analysis

The direct infusion data analysis was performed using procedures described in detail by Kourtchev et al. (2013). The calibration (as described in Kourtchev et al., 2013) and ion transmission checks that include monitoring of ion signal intensity were routinely performed. Briefly, for each sample analysis, 60–90 mass spectral scans were averaged into one mass spectrum. Molecular formulae assignments were made using Xcalibur 3.1 software using the following constraints: 12C ≤ 100, 13C ≤ 1, 1H ≤ 200, 16O ≤ 50, 14N ≤ 5, 32S ≤ 2, 34S ≤ 1. The data processing was performed using a Mathematica 8.0 (Wolfram Research Inc., UK) code developed in-house that utilises a number of additional constraints described in previous studies (Kourtchev et al., 2013, 2015). Only ions that appeared in all three replicates were kept for evaluation. The background spectra obtained from the procedural blanks were also processed using the rules mentioned above. The formulae lists of the background spectra were subtracted from those of the ambient (or chamber) sample and only formulae with a sample/blank peak intensity ratio ≥10 were retained.

The Kendrick mass defect (KMD) is calculated from the difference between the nominal mass of the molecule and the exact KM (Kendrick, 1963). Kendrick mass of the CH2 unit is calculated by renormalising the exact IUPAC mass of CH2 (14.01565) to 14.00000.

2.6 Benzene and isoprene measurements

For benzene and isoprene analysis we used a high-resolution selective-reagent-ionisation proton transfer reaction time-of-flight mass spectrometer (SRI-PTR-TOF-MS 8000, Ionicon Analytik, Austria). The data reduction process used and a description of the PTR-TOF-MS instrument are provided elsewhere (Graus et al., 2010; Müller et al., 2013). Background of the instrument was measured regularly by passing ambient air through a platinum catalyst heated to 380 °C. Sensitivity calibrations were performed by dynamic dilution of VOCs using several multi-component gas standards (Apel Riemer Environmental Inc., Scott-Marrin, and Air Liquide, USA). The calibration cylinders contained acetaldehyde, acetone, benzene, isoprene, α-pinene, toluene and trichlorobenzene, among other species. During IOP1, the instrument was operated with the H3O⁺ reagent ion and at a drift tube pressure of 2.3 mbar, voltage of 600 V, and temperature of 60 °C.
corresponding to a field density ratio $E/N$ ratio of 130 Td ($E$ being the electric field strength and $N$ the gas number density; 1 Td = $10^{-17}$ V cm$^{-2}$). During IOP2, the reagent ion was NO$^+$ and the drift tube settings were 2.3 mbar, 350 V, and 60 °C, resulting in an $E/N$ ratio of 76 Td. The sampling was done with 1 min time resolution and the instrument detection limit for benzene and isoprene were below 0.02 and 0.04 ppbv, respectively.

### 2.7 Air mass history analysis

Air mass history analysis was done for the sampling period using the Numerical Atmospheric-dispersion Modelling Environment (NAME) model, developed by the UK Met Office (Maryon et al., 1991). NAME is a Lagrangian model in which particles are released into 3-D wind fields from the operational output of the UK Met Office Unified Model meteorology data (Davies et al., 2005). These winds have a horizontal resolution of 17 km and 70 vertical levels, which reach ~80 km. In addition, a random walk technique was used to model the effects of turbulence on the trajectories (Ryall and Maryon, 1998). To allow the calculation of air mass history for the average sampling time (which varied between samples, 24, 36 or 48 h), $10^4$ particles per hour were released continuously from the T3 site. The trajectories travelled back in time for 3 days with the position of the particles in the lowest 100 m of the model atmosphere recorded every 15 min. The particle mass below 100 m was integrated over the 72 h travel time. The air mass history (“footprints”) for the periods of the analysed filters are shown in Fig. S1 in the Supplement. The majority of the 3-day air mass footprints originated from the east, although wind direction showed variability nearer to the sampling site on some occasions, e.g. sample MP14-17 (Fig. S1). Almost all air masses pass over Manaus and therefore highlight this city as a potential source. Some air masses also pass over Manacapuru, but this is rare and the corresponding time-integrated concentrations are lower than the equivalent Manaus values.

### 3 Results and discussion

Figure 1 shows mass spectra from two typical samples collected during IOP1 and IOP2. The majority of the ions were associated with molecules below 500 Da, although the measured mass goes up to 900 Da. Although ESI is a “soft” ionisation technique resulting in minimal fragmentation, we cannot exclude the possibility that some of the detected ions correspond to fragments, also in light of the many relative fragile compounds (e.g. thermally labile compounds) that constitute OA. The largest group of identified molecular formulae in all samples were attributed to molecules containing CHO atoms only ($1051 \pm 141$ formulae during IOP2 and $820 \pm 139$ during IOP1), followed by CHON ($537 \pm 71$ during IOP2 and $329 \pm 71$ during IOP1), CHOS ($183 \pm 34$ during IOP2 and $137 \pm 31$ during IOP1) and CHONS ($37 \pm 11$ during IOP2 and $28 \pm 10$ during IOP1) (Fig. 2). The number of molecular formulae containing CHO and CHON subgroups increased by ~20% from IOP1 to IOP2; however, a rather insignificant increase was observed for CHOS and CHONS subgroups. The Student’s $t$ test showed that the observed difference for CHO ($p = 0.0092$) and CHON ($p = 0.00007$) subgroups between two seasons is statistically significant. This is consistent with the observed increase in odd reactive nitrogen species (NO$_3^-$) from IOP1 to IOP2 (Table S1). Organic nitrates are believed to form in polluted air through reaction with nitrogen oxides during daytime and from reaction of NO$_3^-$ with BVOCs during nighttime (Day et al., 2010; Ayres et al., 2015). The average concentration of NO$_3^-$ during IOP1 was found to be on almost 2 times higher, which is possibly reflected in the increased number of organonitrates in the aerosol samples from IOP2. Moreover, the increase in the number of organonitrates during IOP2 is consistent with recent studies which demonstrated that organonitrates groups in aerosol particles may hydrolyse under high-RH conditions (Liu et al., 2012). In this respect, while night-time maximum RH during both filter sampling periods was very similar (~90%), daytime RH during IOP1 was higher (89%) compared to that from IOP2 (66%) (Fig. S2).

Carbon oxidation state (OS$_C$) introduced by Kroll et al. (2011) can be used to describe the composition of a complex mixture of organics undergoing oxidation processes. OS$_C$ was calculated for each molecular formula identified in the mass spectra using the following equation:

$$OS_C = -\sum_i OS_i \frac{n_i}{n_C},$$

where OS$_i$ is the oxidation state associated with element $i$ and $n_i/n_C$ is the molar ratio of element $i$ to carbon within the molecule (Kroll et al., 2011).

Figure 3 shows overlaid OS$_C$ plots for two samples from IOP1 and IOP2. Consistent with previous studies, the majority of molecules in the sampled organic aerosol had OS$_C$ between −1.5 and +1 with up to 30 (nC) carbon atoms throughout the selected mass range (m/z 100–650) (Kroll et al., 2011, and the references therein). The molecules with OS$_C$ between −1 and +1 with 13 or fewer carbon atoms (nC) are suggested to be associated with semivolatile and low-volatility oxidised organic aerosol (SV-OOA and LV-OOA) produced by multistep oxidation reactions. The molecules with OS$_C$ between −0.5 and −1.5 with seven or more carbon atoms are associated with primary biomass burning organic aerosol (BBOA) directly emitted into the atmosphere (Kroll et al., 2011). The cluster of molecules with OS$_C$ between −1 and −1.5 and nC fewer than 10 could be possibly associated with OH radical oxidation products of isoprene (Kourtchev et al., 2015), which is an abundant VOC in Amazon rainforest (Rasmussen and Khalil, 1988; Chen et al., 2015). The isoprene daytime average was above 1.5 ppbv during both seasons, with hourly campaign averages reach-
The line colours in the mass spectra correspond to the CHO (black), blue (CHON), CHOS (red) and CHONS (green) formulae assignments. The relative-intensity axis was split to make a large number of ions with low intensities visible.

In general, aerosol samples from IOP1 contained fewer oxidised molecules compared to those from IOP2. Wet deposition of aged or processed aerosol during the wet (i.e. IOP 2) sampling period cannot be the only reason for the observed differences in OSC. It has been shown that different oxidation regimes to generate SOA (e.g. OH radical vs. ozonolysis) can result in significantly different OSC of SOA (Kourtchev et al., 2015). For example, the SOA component from the OH radical-initiated oxidation of α-pinene as well as BVOC mixtures had a molecular composition with higher OSC throughout the entire molecular mass range compared to that obtained from the ozonolysis reaction (Kourtchev et al., 2015).

Figure 4 shows the distribution of ion signal intensities for selected tentatively identified tracer compounds for anthropogenic, biogenic and mixed sources in all 15 samples. The structural or isomeric information is not directly obtained from the direct infusion analysis; therefore, the identification of the tracer compounds was achieved by comparing MS/MS fragmentation patterns from authentic standards and published literature. The tracer compounds include anhydrosugars, structural isomers with a molecular formula of C_6H_10O_5 at m/z 161.0456 corresponding to levoglucosan, mannosan, galactosan and 1,6-anhydro-β-D-glucofuranose, which are regarded as marker compounds for biomass burning (Simoneit et al., 1999; Pashynska et al., 2002; Kourtchev et al., 2011). Nitrocatechols, with a molecular formula of C_6H_5NO_4 (m/z 154.01458), are attributed to mixed anthropogenic sources, e.g. biomass and vehicular emissions, and methyl-nitrocatechols (C_7H_7NO_4, m/z 168.03023) are important markers for biomass burning OA, formed from m-cresol emitted during biomass burning (Inumaa et al., 2010) as well as diesel exhaust. 3-Methyl-1,2,3-butanetricarboxylic acid (3-MBTCA), with a molecular formula of C_8H_12O_6 at m/z 203.05611, is an OH radical-initiated oxidation product of α- and β-pinene (Szmigielski et al., 2007), and regarded as a tracer for processed or biogenic SOA. Finally, isoprene epoxydiol organosulfate ester (IEPOX-OS), with a molecular formula of C_5H_12O_7S at m/z 215.0231, has been suggested to be formed through reactions between SO_4 and isoprene oxidation products (Pye et al., 2013; Budisulistiori-
Figure 3. Carbon oxidation state plot for CHO-containing formulae in organic aerosol from IOP1 (red squares) and IOP2 (blue diamonds).

Figure 4. Ion intensity distributions (left axis) of selected tentatively identified markers in individual samples using UHRMS analysis and averaged benzene concentration (right axis) from PTR-TOF-MS analysis. Benzene concentration was averaged for the aerosol filter sampling intervals. The UHRMS data were corrected for organic carbon load in each individual filter sample (see Methods section).

ini et al., 2015) and thus can be used to observe the extent of SO₂ ageing effects on the biogenic SOA. Direct infusion analysis suffers from competitive ionisation in the complex matrices and thus comparing ion intensities across samples must be done with caution. Moreover, other compounds with similar molecular composition present in the aerosol matrix may also contribute to the ion signal intensities of the above-discussed compounds. All selected tracers showed very similar variations with benzene concentration that was measured in the gas phase using PTR-MS (Fig. 3). Benzene, generally regarded as an anthropogenic species, has various sources, including industrial solvent production, vehicular emissions and biomass burning (Hsieh et al., 1999; Seco et al., 2013; Friedli et al., 2001). Recent studies have indicated that vegetation (leaves, flowers, and phytoplankton) emits a wide variety of benzenoid compounds to the atmosphere at substantial rates (Misztal et al., 2015). However, considering that benzene concentration correlated very well with another anthropogenic tracer CO ($R^2 = 0.77$, Fig. S3) during IOP1 and IOP2, it is likely that the observed benzene concentrations were mainly due to anthropogenic emissions. During the sampling period, irrespective of the season, air masses passed over the large city Manaus and small municipalities located near the T3 site (Fig. S1). It must be noted that, due to rather low sampling resolution time ($\geq 24$ h), the molecular composition of all analysed samples is likely to be influenced by clean air masses and anthropogenic plumes from these urban locations which usually last only a few hours per day, and thus individual urban plume events cannot be identified with the data analysed here. In Manaus natural gas is mainly used for heating and cooking and therefore the contribution from these activities to biomass burning OA at our site is highly unlikely. During IOP1 much lower incidents of forest fires were observed compared to that during IOP2 (Martin et al., 2016). For example, the number of forest fires in the radius of 200 km from the sampling site varied between 0 and 340 fires (http://www.dpi.inpe.br/proarco/bdqueimadas/). This is reflected in the ion signal intensities of the particle-phase biomass burning markers, i.e. anhydrosugars (C₆H₁₀O₅) and nitrocatechols (C₆H₅NO₄) and gas-phase benzene concentrations, which were significantly lower during IOP1 compared to that from IOP2, when on average more fires are observed.

It should be noted that ion signal intensities for anhydrosugars (C₆H₁₀O₅) and nitrocatechols (C₆H₅NO₄) showed a very good correlation ($R^2 > 0.7$), suggesting that nitrocatechols, observed at the sampling site, are mainly associated with biomass burning sources. The highest ion signal in-
tensities of these tracer compounds were observed during two periods: 7–9 September 2014 (sample MP14-128) and 27–28 September 2014 (sample MP14-148) with the latter one coinciding with highest incident of fires (340 fires). Although during 7–9 September (sample MP14-128) a significantly lower number (22 fires) of fires was observed compared to the period of 27–28 September 2014, lower wind speed occurring during 7–9 September suggests that a high intensity of the biomass burning markers could be due to the biomass burning emissions from nearby sources. Between the T3 sampling site and Manaus (about 20 km east of the site), there are a number of small brick factories, which use wood to fire the kilns (Martin et al., 2016) and are thus an additional local wood-burning source besides the forest and pasture fires.

Interestingly the sample MP14-148 had the highest ion intensity corresponding to IEPOX-OS (Fig. 4), which also coincided with the strong increase in the ion intensity at m/z 96.95987 corresponding to [HSO\(^4\)]\(^-\). This is consistent with the organosulfate formation mechanism through reactive uptake of isoprene epoxydiols (IEPOX) in the presence of acidic sulfate (Surratt et al., 2010; Lin et al., 2012, 2013). A similar relationship between sulfate and organosulfates concentrations has been observed previously in field studies in the southeastern US (Surratt et al., 2007, 2008, 2010; Lin et al., 2012, 2013). This is also in agreement with previous studies from the Amazon, where the highest levels of 2-methyltetrols were observed during the dry period, which was characterised by biomass burning (and higher particle concentrations of sulfuric acid) (Claeys et al., 2010). Considering that Claeyts et al. (2010) employed an alternative GC/MS procedure with prior trimethylsilylation, 2-methyltetrol sulfates were converted to 2-methyltetrols and not detectable as separate organosulfate compounds. It should be noted that the 27–28 September period (sample MP14-148) was marked by a very strong increase in the CO concentration (Fig. S4). In mid-latitude environments it has been suggested that the production of anthropogenic SOA in an air mass, as it travels from an urban source region, can be estimated by using a relatively inert pollution tracer, such as CO occurring in the air mass (De Gouw et al., 2005; Hoyle et al., 2011). At the T3 sampling site, the highest CO concentrations are observed in air masses affected by biomass burning. Therefore, it is possible that organic aerosol in the sample MP14-148 has experienced the highest contribution from biomass burning as well as other anthropogenic activities.

To investigate the influence of anthropogenic activities (i.e. biomass burning) on a detailed molecular composition of organic aerosol at the T3 site, we compared samples from the periods with the lowest (9 fires), moderately high (254 fires) and the highest (340 fires) incidents of fires occurring within 200 km around the site.

Figure 5a–c show H/C ratios of CHO-containing formulae as a function of their molecular mass and double-bond equivalent (DBE), which shows a degree of unsaturation of the molecule, for a sample with the lowest (a) moderately high (b) and highest incidents (c) of fires. One of the obvious differences between these samples is the abundance of ions with low H/C ratios (\(<1\)). The majority of these ions have a DBE value above 7, indicating that they likely correspond to oxidised aromatic compounds, which are mainly of anthropogenic origin (Kourtchev et al., 2014; Tong et al., 2016). For example, the smallest polycyclic aromatic hydrocarbon (PAH), naphthalene, with a molecular formulae of C\(_{10}\)H\(_8\), has H/C = 0.8 and DBE = 7. The number of CHO-containing formulae with high DBE equivalent and low H/C increased dramatically during the days with moderately high and high incidents of fires (Fig. 5a–c), suggesting that they are mainly associated with biomass burning. The largest grey circles in Fig. 5a–c correspond to the ions at m/z 133.01425 (with a neutral molecular formula of C\(_4\)H\(_8\)O\(_3\)), m/z 187.0612 (C\(_8\)H\(_{12}\)O\(_3\)), m/z 201.07685 (C\(_9\)H\(_{14}\)O\(_3\)), m/z 203.05611 (C\(_9\)H\(_{12}\)O\(_6\)), and m/z 215.05611 (C\(_9\)H\(_{12}\)O\(_6\)) with DBE < 6.

Recent studies have indicated that different families of compounds with heteroatoms (e.g. O, S) overlap in terms of DBE and thus may not accurately indicate the level of unsaturation of organic compounds. For example, the valent atoms, such as oxygen and sulfur, do not influence the value of DBE, yet they may contribute to the potential double bonds of that molecule (Reemtsma, 2009; Yassine et al., 2014). Yassine et al. (2014) suggested using aromaticity equivalent (X\(_c\)) to improve the identification and characterisation of aromatic and condensed aromatic compounds in water-soluble organic carbon. The aromaticity equivalent can be calculated as follows:

\[
X_c = \frac{3(DBE - (mN_O + nN_S)) - 2}{(DBE - (mN_O + nN_S))},
\]

where “m” and “n” correspond to a fraction of oxygen and sulfur atoms involved in π-bond structures of a compound which varies depending on the compound class. For example, carboxylic acids, esters, and nitro functional groups have m = n = 0.5. For compounds containing functional groups such as aldehydes, ketones, nitroso, cyanate, alcohol, or ethers “m” and “n” are 1 or 0. Considering that ESI, in negative mode, is most sensitive to compounds containing carboxylic groups we therefore used m = n = 0.5 for the calculation of the X\(_c\). For molecular formulae with an odd number of oxygen or sulfur, the sum (mN\(_O\) + nN\(_S\)) in Eq. (2) was rounded down to the closest integer as detailed in Yassine et al. (2014). The authors proposed that aromaticity equivalent with X\(_c\) ≥ 2.50 and X\(_c\) ≥ 2.71 as unambiguous minimum criteria for the presence of aromatics and condensed aromatics.

Expressing our data using aromaticity equivalents confirmed that the increase in the number of molecules with high DBE from the sample with the lowest to the highest incidents of fires was due to the increase in the number of
Figure 5. H/C vs. m/z plot for CHO-containing formulae in the samples from the periods with (a) low, (b) moderately high and (c) very high incidents of fires. The marker size reflects relative ion abundance in the sample. The colour code shows double-bond equivalent (DBE) in the individual molecular formula. Molecular formulae with DBE < 6 are shown as grey markers. The largest grey circles correspond to the ions at m/z 133.01425 (with a neutral molecular formula of C₄H₆O₅), m/z 187.0612 (C₈H₁₂O₅), m/z 201.07685 (C₉H₁₄O₅), m/z 203.05611 (C₈H₁₂O₆), and m/z 215.05611 (C₉H₁₂O₆).

aromatic and condensed aromatic compounds in the aerosol samples (Fig. S5). Considering the Yassine et al. (2014) assignment criteria for the aromatic-reach matrices, the highest number of the aromatic compounds in the Amazon samples was observed for formulae with a benzene core structure (X_c = 2.50), followed by formulae with a pyrene core structure (X_c = 2.83), an ovalene core structure (X_c = 2.92), and highly condensed aromatic structures or highly unsaturated compounds (X_c > 2.93). The largest grey circles in Fig. S5a correspond to the ions at m/z 187.11357 with a neutral molecular formula of C₉H₁₇NO₃ and m/z 281.26459 with a neutral molecular formula of C₁₈H₁₅NO. The largest grey circles in Fig. S5b and c correspond to the ions at m/z 154.0146, m/z 168.03023 and m/z 152.03532 with neutral molecular formulae of C₆H₅NO₄, C₇H₇NO₄ and C₇H₇NO₃, respectively.

Interestingly, a similar trend was observed for the molecules containing CHON subgroups (Fig. S6). A number of CHON molecules with low H/C (< 1) and high DBE (≥ 5) almost doubled from the days with 9 to 340 fires (Fig. S7). Nitro-aromatic compounds such as nitrophenols (DBE = 5) and N-heterocyclic compounds, including 4-nitrocatechol and isomeric methyl-nitrocatechols, are often observed in the OA from biomass burning sources (Kitanovski et al., 2012a, b; Iinuma et al., 2010) and have been suggested to be potential contributors to light absorption by brown carbon (Laskin et al., 2015). It is worth mentioning that aerosol samples affected by biomass burning contained another interesting ion at m/z 182.04588 with a neutral molecular formula of C₈H₇NO₄, possibly corresponding to other biomass burning OA markers, i.e. isomeric dimethyl-nitrocatechols (Kahnt et al., 2013). The differences in the in-
increased number of nitro-aromatic compounds in aerosol samples affected by biomass burning are also apparent in overlaid Van Krevelen diagrams (Fig. 6), which show H/C and O/C ratios for each formula in a sample. Van Krevelen diagrams can be used to describe the overall composition or evolution of organic mixtures (Van Krevelen, 1993; Nizkorodov et al., 2011; Nozière et al., 2015). Organic aerosol affected by biomass burning contained a significantly larger number of CHON formulae with O/C < 0.5 and H/C < 1 (Fig. 6a and b, area B) but a smaller number of formulae with O/C < 0.5 and H/C > 1 (Fig. 6a and b, area A). While molecules with H/C ratios (< 1.0) and O/C ratios (< 0.5) (area A in Fig. 3) are generally associated with aliphatic compounds, molecules with high H/C ratios (> 1.5) and low O/C ratios (< 0.5) typically belong to oxidised aromatic hydrocarbons (area B in Fig. 3) (Mazzoleni et al., 2010, 2012). Although the smaller number of nitro-aliphatic compounds in the samples affected by biomass burning requires further investigation, it is possible that they were oxidised in the polluted air by NO₃ and O₃ (Zahardis et al., 2008; Malloy et al., 2009), whose production is significantly enhanced during fire events (e.g. Galanter et al., 2000). The majority (up to 80 %) of the CHON molecules in the analysed samples have O/C ratios < 0.7 (Fig. 6). The relatively low oxygen content suggests that these molecules include decreased nitrogen-containing compounds (Zhao et al., 2013). Although biomass burning material type is expected to result in a different molecular composition, the presence of a large number of molecules with low O/C ratio is consistent with the literature. For example, most of the CHON molecules in OA from wheat straw burning in K-Puszta in the Great Hungarian Plain of Hungary and biomass burning at Canadian rural sites (Saint Anicet, Quebec, and Canterbury, New Brunswick) had O/C ratios below 0.7 (Schmitt-Kopplin et al., 2010). In addition, the CHON molecules identified by LC/MS in biomass burning OA from Amazonia showed O/C ratios below 0.7, i.e. 4-nitrocatechol (C₆H₄NO₄; O/C = 0.67), isomeric methyl-nitrocatechols (C₇H₇NO₄; O/C = 0.57), and isomeric dimethyl-nitrocatechols (C₈H₉NO₄; O/C = 0.50) (Claeys et al., 2012).

Figure 7 shows overlaid OSC plots for OA from the days with low, moderately high and high incidents of fires. During the days affected by high and moderately high number of fires, OSC was shifted towards a more oxidised state for the CHO molecules containing more than seven carbon atoms. The difference in OSC becomes even more pronounced with the increased number of carbons (e.g. > 7 carbon atoms) in the detected molecular formulae. Interestingly, the affected ions with high OSC do not fall into the category of the BBOA (encircled area in Fig. 7) which are associated with primary particulate matter directly emitted into the atmosphere as defined in Kroll et al. (2011).

At a first glance, biomass burning seems to influence the number and intensity of the CHOS-containing formulae; however, the effect was much lower compared to that for the CHO and CHON molecules (see discussion above). A higher number of CHOS-containing molecules was observed in the sample (MP14-148) corresponding to the highest incident of fires (Fig. 8a). Interestingly, IEPOX-OS was found to be very abundant in the sample that experienced the highest incidents of fires (Fig. 8a). The significant IEPOX-OS mass was previously observed during low-altitude flight campaigns in northern California and southern Oregon under high-NO conditions (> 500 pptv) (Liao et al., 2015). The authors explained this observation by the transport or formation of IEPOX from isoprene hydroxynitrate oxidation (Jacobs et al., 2014) and higher sulfate aerosol concentrations occurring during their sampling period (Nguyen et al., 2014). This explanation is also consistent with our results. The ion at m/z 96.95987 corresponding [HSO₄]⁻ in UHR mass spectra of the sample MP14-148 was 3 times more abundant compared to that in the sample MP14-129, suggesting that particle acidity may be one of the reasons for the high abundance of the IEPOX-OS in this sample. Considering that the main sources of sulfate at T3 site are industrial pollution (e.g. power plants), natural and long-range sources, they could also be responsible for the high abundance of the sulfate and IEPOX-OS in the samples besides the overlapping biomass burning event. Noticeably, these samples contained not only
Figure 7. Overlaid carbon oxidation state (OS\(_C\)) plots for CHO subgroups in the samples from the periods with low (blue markers) and very high (red markers) incidents of fires. The marker areas reflect relative ion abundance in the sample. The area marked as SV-OOA, LV-OOA, BBOA and HOA correspond to the molecules associated with semivolatile and low-volatility oxidised organic aerosol, biomass burning organic aerosol and hydrocarbon-like organic aerosol as outlined by Kroll et al. (2011).

Figure 8. Overlaid Van Krevelen diagram (a) and Kendrick mass defect plot (b) for CHOS-containing formulae in the samples from the periods with low (blue markers) and very high incidents of fires (red markers). The marker areas reflect relative ion abundance in the sample. Red markers correspond to the ions from the period with the lowest incidents of fires. Note that IEPOX-OS is not a part of any homologous series in the sample with very low incident of fires and it has only one additional homologue in the sample that experienced very high incident of fires (see enlarged area of Fig. 8a). Area “A” in Kendrick mass defect plot shows formulae with KMD > 0.33 that are mainly present in the sample with high incident of fires.

A larger number of oxygenated CHOS-containing molecules with O/C > 1.2 but also molecules with O/C < 0.6 and H/C ranging from 0.4 to 2.2. Recent laboratory and field studies indicated the presence of a large number of aromatic and aliphatic organosulfates and sulfonates in OA and linked them to anthropogenic precursors (Tao et al., 2014; Wang et al., 2016; Riva et al., 2015, 2016; Kuang et al., 2016). Riva et al. (2015, 2016) demonstrated formation of organosulfates and sulfonates in the laboratory smog chamber experiments from photooxidation of alkanes and PAHs, respectively. The authors indicated enhancement of organosulfates yields in the presence of acidified ammonium sulfate seed and suggested that these organosulfates are mainly formed through reactive uptake of gas-phase epoxides. It must be noted that the above-cited field studies are based on measurements in the Northern Hemisphere and thus organosulfate formation pathways and sources may differ from that of Amazonia.

KMD plots are a useful visualisation technique for identification of homologous series of compounds differing only by the number of a specific base unit (e.g. a CH\(_2\) group). Anthropogenically affected aerosol samples have longer homologous series of molecules containing CHOS subgroups.
(Fig. 8b). One of these longer series includes a second most abundant ion at \( m/z \) 213.0075 (C\(_5\)H\(_{10}\)O\(_3\)S). The compound with a molecular formula of C\(_5\)H\(_{10}\)O\(_3\)S has been previously observed in the laboratory and field studies and attributed to isoprene-derived organosulfates (Surratt et al., 2008; Gómez-González, 2008; Kristensen and Glassius, 2011; Nguyen et al., 2014; Hettiyyadura et al., 2015). This molecular formula could also be associated with organosulfates (e.g. isomeric 3-sulfooxy-2-hydroxypentanoic acid and 2-sulfooxy-3-hydroxypentanoic acid) formed from the green leaf volatiles 2-E-pentenal, 2-E-hexenal, and 3-hexenal (Shalamzari et al., 2016). The KMD plot (Fig. 8b) shows that OA from the anthropogenically affected samples contained an additional series of CHOS molecules with high KMD > 0.33 that were not present in the sample from the less polluted period. Most of these ions are highly oxygenated (containing > 10 oxygens) and are likely to be associated with molecules produced through photochemical ageing reactions (Hildebrandt et al., 2010).

It is worth noting that, in most of the samples, IEPOX-OS was not a part of any homologous series in KMD plot (e.g. Fig. 8b). This observation confirms that atmospheric oxidation reactions resulting in the incorporation of S and N functional groups do not always conserve homologous series but could also lead to a wide range of possible reaction products (Rincón et al., 2012; Kourtchev et al., 2013).

4 Conclusions

In this study we applied direct infusion nanoESI UHRRMS for the analysis of the organic fraction of PM\(_{2.5}\) samples collected IOP1 and IOP2 of GoAmazon2014/5 in central Amazonia, which is influenced by both background and polluted air masses. Up to 2100 elemental formulae were identified in the samples, with the largest number of formulae found during IOP2. The distribution of several tracer compounds along with the comprehensive mass spectral data evaluation methods (e.g. Kendrick mass defect, Van Krevelen diagrams, carbon oxidation state and aromaticity equivalent) applied to the large UHRRMS datasets were used to identify various sources of organic aerosol components, including natural biogenic sources, biomass burning and anthropogenic emissions. The distinguishable homologous series in the KMD diagram contained nitrogen-containing series included NACs, e.g. nitrocatechols, nitrophenols, nitroguaianclos and nitrosaliclycic acids derived from biomass burning material. Isoprene-derived IEPOX-OS was found to be the most dominant ion in most of the analysed samples and strongly followed the concentration trends of the gas-phase anthropogenic tracer benzene and CO (with biomass burning as the dominant tracer at the T3 site), supporting its mixed biomass burning–anthropogenic–biogenic origin. Van Krevelen, DBE and X\(_e\) distributions along with relatively low elemental O / C and H / C ratios indicated the presence of a large number of oxidised aromatic compounds in the samples. A significant number of CHO-containing formulae in aerosol samples from IOP2 had a higher oxidation state compared to that from IOP1 and became even more important during the days with the highest incidents of fires. Although our results suggest that the studied site is significantly influenced by not only biogenic emissions and biomass burning but also anthropogenic emissions from the neighbouring urban activities, future work is needed to better understand the quantitative contributions of the various factors to the aerosol composition at the T3 site. The analysis of aerosol samples with higher sampling resolution or quantifying specific marker compounds and applying a receptor modelling techniques (Alves et al., 2015) would allow for separating these sources in more detail and would thus significantly improve understanding of the aerosol formation sources at the site.

5 Data availability

The data are available upon request.

The Supplement related to this article is available online at doi:10.5194/acp-16-11899-2016-supplement.

Acknowledgements. Research at the University of Cambridge was supported by ERC grant no. 279405. The authors would like to thank Jason Surratt (University of North Carolina) for providing a synthesised IEPOX-OS standard. O\(_x\), CO, NO\(_x\), RH and rain data were obtained from the Atmospheric Radiation Measurement (ARM) Climate Research Facility, a U.S. Department of Energy (grant DE-SC0011122) Office of Science user facility sponsored by the Office of Biological and Environmental Research. We acknowledge the support from the Central Office of the Large Scale Biosphere Atmosphere Experiment in Amazonia (LBA), the Instituto Nacional de Pesquisas da Amazonia (INPA), and the Universidade do Estado do Amazonia (UEA). The work was conducted under 001030/2012-4 of the Brazilian National Council for Scientific and Technological Development (CNPq).

Edited by: J. Allan
Reviewed by: two anonymous referees

References


I. Kourtchev et al.: Molecular composition of organic aerosols in central Amazonia

Kendrick, E.: A Mass Scale Based on CH
Keller, M., Bustamante, M., Gash, J., and Dias, P.: Amazonia and
Kitanovski, Z., Grgic, I., Yasmeen, F., Claeys, M., and Cusak,
(iii) The formation, properties and impact
of secondary organic aerosol: current and emerging issues,


I. Kourtchev et al.: Molecular composition of organic aerosols in central Amazonia


