Molecular composition of biogenic secondary organic aerosols using ultrahigh-resolution mass spectrometry: comparing laboratory and field studies

I. Kourtchev
S. J. Fuller
C. Giorio
R. M. Healy
E. Wilson
I. O’Connor
J. C. Wenger
M. McLeod
J. Aalto
T. M. Ruuskanen
W. Maenhaut
R. Jones
D. S. Venables
J. R. Sodeau
M. Kulmala
M. Kalberer

1Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK
2Department of Chemistry and Environmental Research Institute, University College Cork, Cork, Ireland
3Department of Forest Sciences, University of Helsinki, P.O. Box 27, Helsinki, Finland
4Hyytiälä Forestry Field Station, Hyytiäläntie 124, Korkeakoski, 35500, Finland
5Department of Physics, University of Helsinki, P.O. Box 64, Helsinki, Finland
6Department of Analytical Chemistry, Ghent University, Krijgslaan 281, S12, 9000 Ghent, Belgium
7Department of Pharmaceutical Sciences, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium

Correspondence to: I. Kourtchev (ink22@cam.ac.uk) and M. Kalberer (mk594@cam.ac.uk)

Received: 25 October 2013 – Published in Atmos. Chem. Phys. Discuss.: 12 November 2013
Revised: 17 January 2014 – Accepted: 21 January 2014 – Published: 26 February 2014

Abstract. Numerous laboratory experiments have been performed in an attempt to mimic atmospheric secondary organic aerosol (SOA) formation. However, it is still unclear how close the aerosol particles generated in laboratory experiments resemble atmospheric SOA with respect to their detailed chemical composition. In this study, we generated SOA in a simulation chamber from the ozonolysis of α-pinene and a biogenic volatile organic compound (BVOC) mixture containing α- and β-pinene, Δ3-carene, and isoprene. The detailed molecular composition of laboratory-generated SOA was compared with that of background ambient aerosol collected at a boreal forest site (Hyytiälä, Finland) and an urban location (Cork, Ireland) using direct infusion nanoelectrospray ultrahigh-resolution mass spectrometry. Kendrick mass defect and van Krevelen approaches were used to identify and compare compound classes and distributions of the detected species. The laboratory-generated SOA contained a distinguishable group of dimers that was not observed in the ambient samples. The presence of dimers was found to be less pronounced in the SOA from the BVOC mixtures when compared to the one component precursor system. The molecular composition of SOA from both the BVOC mixture and α-pinene represented the overall composition of the ambient sample from the boreal forest site reasonably well, with 72.3 ± 2.5 % (n = 3) and 69.1 ± 3.0 % (n = 3) common ions, respectively. In contrast, large differences were found between the laboratory-generated BVOC samples and the ambient urban sample. To our knowledge this is the first direct comparison of molecular composition of laboratory-generated SOA from BVOC mixtures and ambient samples.

1 Introduction

Biogenic volatile organic compounds (BVOCs) play an important role in atmospheric chemistry and give rise to secondary organic aerosols (SOA), which have effects on climate (Hallquist et al., 2009) and human health (Pope and Dockery, 2006). SOA is formed within the atmosphere from gaseous precursors and gas-to-particle conversion processes. Laboratory chamber experiments have been performed for decades in an attempt to mimic atmospheric SOA formation. However, it is still unclear how close the aerosol particles generated in laboratory experiments resemble atmospheric SOA with respect to their detailed chemical composition.
One of the major challenges is the identification of the organic composition of the SOA, which is composed of thousands of organic compounds (Kanakidou et al., 2005). These compounds generally cover a wide range of polarities, volatilities and masses (Goldstein and Galbally, 2007) and therefore it is difficult to find a single analytical technique for their detailed chemical analysis at the molecular level. Conventional chromatographic methods (gas chromatography (GC) and liquid chromatography (LC)) are not capable of resolving the highly complex mixtures with a wide variety of physicochemical properties. Moreover, commonly used mass spectrometers, which are often used as detectors following chromatographic separation, do not have sufficient mass-resolving power to distinguish and differentiate all the compounds present in the complex mixture of organic aerosol. Ultra-high resolution mass spectrometers (UHR-MS) (i.e. Fourier transform ion cyclotron resonance MS and Orbitrap MS) have a mass resolution power that is at least 1 order of magnitude higher (≥100 000) than conventional MS and thus have the potential for solving this problem. Direct infusion electrospray ionisation (ESI)-UHR-MS has been successfully applied for the analysis of both ambient and laboratory-generated SOA and facilitated the characterisation of hundreds of species with individual molecular formulae (Nizkorodov et al., 2011). Despite the high analytical throughput of direct infusion MS, this method is prone to artefacts such as changes in the ionisation efficiency of an analyte due to the presence of “matrix” compounds in the complex organic mixtures (Pöschl, 2005). For instance, sulfates, nitrates and ammonium salts are important constituents of atmospheric aerosols (Pöschl, 2005) and once injected into the ESI source can cause ion suppression, adduct formation and a rapid deterioration of instrument performance (Dettmer et al., 2007). NanoESI-MS, which generally produces smaller droplet sizes and analyte flow in the electrospray (Schmidt et al., 2003), can substantially reduce interference effects from inorganic salts. Moreover, it provides better sensitivity towards a variety of analytes in samples containing relatively high levels of salts (Jurasek et al., 1999; Schmidt et al., 2003) and decreases source contamination (Schmidt et al., 2003) compared to conventional ESI sources.

To date, most laboratory experiments reproducing atmospheric SOA formation have been performed using a single organic precursor (e.g. α- or β-pinene or isoprene) while in the atmosphere a wide range of precursors contribute to SOA, which results in a more complex SOA composition compared to the one-precursor laboratory systems. Although there are a few studies where oxidation of volatile organic compound (VOC) mixtures were performed, their main goal was to investigate SOA formation, yields (VanReken et al., 2006; Jaoui et al., 2008; Hao et al., 2009, 2011; Kiendler-Schr et al., 2009; Mentel et al., 2009; Hatfield and Huff Hartz, 2011; Waring et al., 2011) and specific products (Jaou et al., 2003; Amin et al., 2013) rather than detailed molecular composition.

The main objective of this work is to compare the detailed molecular composition of laboratory-generated SOA from oxidation of a single BVOC (α-pinene) and from a mixture of four BVOCs, with samples of ambient aerosol from urban and remote locations using chip-based direct infusion nanoESI-UHR-MS. In a preceding study, we examined aerosol samples from the boreal forest site Hyytiälä, Finland, and determined that a dominant fraction of the detected compounds are reaction products of a multi-component mixture of BVOCs (Kourtchev et al., 2013). In the present study, we compare the composition of these field samples with SOA generated in chamber experiments from the ozonolysis of α-pinene and of BVOC mixtures containing four species (α- and β-pinene, Δ3-carene, and isoprene) that are most abundant in Hyytiälä’s environment. The laboratory experiments were performed under conditions (e.g. relative humidity (RH), aerosol seed, and BVOC ratios) resembling those at the boreal sampling site during the summer 2011 period. To the best of our knowledge this is the first direct comparison of the molecular composition of laboratory-generated SOA from the BVOC mixtures and ambient samples.

2 Materials and methods

2.1 Atmospheric simulation chamber

Experiments were carried out in an atmospheric simulation chamber described in detail elsewhere (Thüner et al., 2004). Briefly, the chamber is a cylinder made of fluorine–ethene–propene (FEP) Teflon® foil with a volume of 3.91 m3. It was operated at 296 ± 2 K using purified air at 0.1–1 mbar above atmospheric pressure. The experiments were performed at 60–68 % relative humidity produced from bubbling purified air through heated water. The humidity and temperature were measured using a dew point meter (DRYCAP® DM70 Vaisala). Fans installed at both ends of the chamber were used during the first 5 min of the reaction to provide rapid and uniform mixing of the reactants and products. Between experiments the chamber was cleaned by introducing about 1 ppm of ozone into the chamber and flushing with purified air at a flow rate of 0.15 m3 min−1. The experiments were performed with neutral ammonium sulfate ((NH4)2SO4, Sigma-Aldrich, 99.99 %) seed particles, produced using an atomiser and dried before introduction into the chamber. Aerosol seed particles were subjected to Krypton-85 (Kr-85) charge neutraliser before introduction to the chamber. Seed particle concentrations for each experiment are shown in Table 1. Cyclohexane at a molar concentration 1000 times higher than the BVOC precursors was used to scavenge OH radicals produced from ozonolysis of the reactants. BVOCs (i.e. α-pinene, β-pinene, Δ3-carene and isoprene) were introduced into the chamber by flowing purified air over known amounts of the compounds in a gently heated Pyrex impinger. The BVOC concentrations...
Table 1. Summary of α-pinene and BVOC mixture ozonolysis experiments.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Total BVOC conc., ppb</th>
<th>Ozone conc., ppb</th>
<th>RH, %</th>
<th>Seed conc., µg m⁻³</th>
<th>ΔM, µg m⁻³</th>
<th>SOA yield (γ)ᵃᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-pinene</td>
<td>49.5</td>
<td>200</td>
<td>60</td>
<td>3.1</td>
<td>42.3</td>
<td>0.15</td>
</tr>
<tr>
<td>α-pinene</td>
<td>50.5</td>
<td>200</td>
<td>64</td>
<td>3.0</td>
<td>50.0</td>
<td>0.16</td>
</tr>
<tr>
<td>α-pinene</td>
<td>55.5</td>
<td>200</td>
<td>62</td>
<td>4.3</td>
<td>55.9</td>
<td>0.16</td>
</tr>
<tr>
<td>BVOC mixture</td>
<td>62.9</td>
<td>200</td>
<td>63</td>
<td>5.2</td>
<td>34.8</td>
<td>0.11</td>
</tr>
<tr>
<td>BVOC mixture</td>
<td>67.7</td>
<td>200</td>
<td>68</td>
<td>3.8</td>
<td>43.6</td>
<td>0.12</td>
</tr>
<tr>
<td>BVOC mixture</td>
<td>62.9</td>
<td>200</td>
<td>65</td>
<td>3.6</td>
<td>41.8</td>
<td>0.12</td>
</tr>
</tbody>
</table>

ᵃYields were calculated from the equation \( \gamma = \Delta M / \Delta HC \), where \( \Delta M (\mu g \text{ m}^{-3}) \) is the particle mass loading and \( \Delta HC (\mu g \text{ m}^{-3}) \) is the reacted concentration of BVOCs;ᵇthe data is wall loss corrected.

are shown in Table 1. Ozone (ca. 200 ppbv) was introduced at the beginning of the reaction over a period of 1 min using an electric discharge generator. Ozone decay was monitored with an automated analyser (Thermo Model 49i). A scanning mobility particle sizer (TSI model 3081) was used to measure particle number-size distributions in the range 11–478 nm (mobility diameter) every 3 min. The particle mass concentrations were calculated assuming a density of 1 g cm⁻³. Dilution and wall-loss corrections were calculated by considering the particle loss as a first order rate process. The background NOₓ concentration measured using a NOₓ analyser (Thermo Model 42i) was below 2 ppbv for all experiments. The aerosol samples were collected on quartz fibre filters (Pallflex® Tissuquartz 2500QAT-UP, 47 mm diameter) 20–30 min after the maximum SOA concentration was observed. Before use, the quartz fibre filters were preheated at 650°C for 12 h to remove any possible organic impurities. A charcoal denuder was used in front of the filter pack to remove gas phase species formed during the ozonolysis reaction. The aerosol sampling flow rate was approximately 12 L min⁻¹ and the sampling time was 40 min. In addition, “blank” chamber samples were collected by drawing “clean” air from the chamber for 40 min. The filter samples were wrapped in baked aluminum foil and stored at −20°C prior to analysis.

2.2 Ambient samples

Biogenic ambient samples were collected at the boreal forest site SMEAR II in Hyytiälä, southern Finland (61°51′N, 24°17′E) as previously described in detail (Kourtchev et al., 2013). The forest around the station is dominated by conifers (mainly Scots pine and Norway spruce) with some deciduous trees, such as aspen and birch, with a tree density of about 2500 ha⁻¹. Detailed descriptions of the site, instrumentation, meteorological data collection and sampling are given elsewhere (Kulmala et al., 2001; Hari and Kulmala, 2005). In total, 10 separate day and night atmospheric aerosol PM₁ samples, each representing 12 h of sampling, were collected over the period 16 to 25 August 2011.

Anthropogenic ambient samples were collected during 9–17 September 2011 at the Tivoli Industrial Estate and Docks (TIED), Cork, Ireland (51°54′5″ N, 8°24′38″ W). A detailed description of the TIED site is given elsewhere (Healy et al., 2009; Hellebust et al., 2010; Kourtchev et al., 2011). The site is located approximately 3 km east of Cork city centre with a population of about 120 000 inhabitants. A shipping berth is located 400–600 m to the south-west, while the main road carrying traffic east out of the city lies 200 m to the north of the site. Residential areas surround the site on all sides except the north and north-east. The vegetation that surrounds the site mainly consists of shrubs and some native deciduous trees, which include oak (Quercus robur L. and Quercus petraea), ash (Fraxinus excelsior), birch (Betula pendula), sycamore (Acer pseudoplatanus), poplar (Populus tremula) and beech (Fagus sylvatica). It should be noted that there are not many conifer species present around the sampling area and if they do exist, they are likely to be spruce (Picea sitchensis). PM₂.₅ aerosol samples were collected on quartz fibre filters (Pallflex® Tissuquartz 2500QAT-UP, 150 mm diameter, pre-fired for 24 h at 650°C) using a high volume (Digitel DHA-80, Switzerland) sampler with a flow rate of 500 L min⁻¹.

2.3 Aerosol sample analysis

All ambient filters were analysed for organic carbon (OC) and elemental carbon (EC) using a thermal-optical transmission (TOT) technique (Birch and Cary, 1996). For each sample, a part of the quartz fibre filter (6–30 cm², depending on OC or total aerosol loading for ambient and laboratory samples, respectively) was extracted three times with 5 mL of methanol (Optima® grade, Fisher Scientific) under ultrasonic agitation for 30 min. The three extracts were combined, filtered through a Teflon® filter (0.2 µm) and reduced by volume to approximately 200 µL under a gentle stream of nitrogen.

The final extracts were analysed using an ultrahigh-resolution LTQ Orbitrap Velos mass spectrometer (Thermo Fisher, Bremen, Germany) equipped with a TriVersa NanoMate robotic nanoflow chip-based ESI source (Advion...
Biosciences, Ithaca NY, USA). The Orbitrap MS instrument was calibrated using an Ultramark 1621 solution (Sigma-Aldrich, UK). The mass accuracy of the instrument was below 1.5 ppm and was routinely checked before the analysis. The instrument mass resolution was 100 000 at m/z 400. The negative ionisation mass spectra were collected in three replicates at ranges m/z 100–650 and m/z 200–900 and processed using Xcalibur 2.1 software (Thermo Scientific). A mixture of camphor, sulfonic acid (20 ng µL⁻¹) and a T3 Atlantis C18 column (3 µm; 2.1 × 150 mm; Waters, Milford, USA). 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 kept for 12 min at 3 % B (total run time 60 min). MS spectra were collected in full scan using the lock mass for the deprotonated dimer of formic acid at m/z 91.00368 with a resolution of 100 000 and the mass ranges of m/z 50–650 and m/z 150–900. Based on pre-scan information from the full scan MS, a parallel data-dependent collision induced dissociation (CID) multi-stage mass spectrometry (MSn) (n = 1, 2, 3 and 4) was performed on the most intense precursor ion in three scans at the resolution of 30 000.

2.4 Ultrahigh MS resolution data analysis

The ultrahigh-resolution mass spectral data interpretation was carried out using a procedure as described elsewhere (Kourtchev et al., 2013). For each direct infusion sample analysis 70–80 mass spectral scans were averaged into one mass spectrum. Molecular assignments were performed using Xcalibur 2.1 software applying the following constraints ¹²C ≤ 100, ¹³C ≤ 1, ¹⁹H ≤ 200, ¹⁶O ≤ 50, ¹⁴N ≤ 5, ³²S ≤ 2, ³⁴S ≤ 1. All mathematically possible elemental formulae with a mass tolerance of ±5 ppm were calculated. The data filtering was performed using a Mathematica 8.0 (Wolfram Research Inc., UK) code developed in-house that employed several conservative rules and constraints similar to those used in previous studies (Koch et al., 2005; Wozniak et al., 2008; Lin et al., 2012). Only ions with intensities ten times above the noise level were kept for the data analysis. The mass tolerance range for keeping mathematically assigned elemental formulae was set to approximately ±0.5 ppm and varied within the ±5 ppm tolerance window. This range was determined by establishing the average difference between the theoretical and the experimental mass for nine compounds with known elemental composition determined by LC/MS analyses (Kourtchev et al., 2013). All molecular formulae where O/C ≥ 1.2, N/C ≥ 0.5, S/C ≥ 0.2 ≤ 0.3 H/C and H/C ≥ 2.5 were eliminated with the aim of removing molecules that are not likely to be observed in nature. Moreover, neutral formulae that had either a non-integer or a negative value of the double bond equivalent (DBE) were removed from the list of possible molecules. Double bond equivalents were calculated using Xcalibur 2.1 software. The assigned formulae were additionally checked for the “nitrogen-rule” and isotopic pattern as described elsewhere (Kourtchev et al., 2013). 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. All molar ratios, DBE factors and chemical formulae presented in this paper refer to neutral molecules.

2.5 Hierarchical agglomerative cluster analysis

Aerosol samples were classified by hierarchical agglomerative cluster analysis (Lukasová, 1979). The data was organised in a two-way table Xnm, where n is the number of samples (six smog chamber samples and two ambient samples) and m is the number of molecules analysed by UHR-MS in the mass range 100–300 Da (451 molecular formulae). The Xnm is a binary value indicating the presence/absence of the molecule m in sample n. The cluster analysis was performed using Statistica 10 (StatSoft Inc., Tulsa, OK, USA), based on the unweighted pair-group average linkage method (or average linkage method) and the percent disagreement (Georgieva et al., 2005) distance measure. The metric used in this study is analogous to the Jaccard’s dissimilarity distance measure that is commonly applied for the analyses of binary patterns (Sneath and Sokal, 1973; Anthony et al., 2002; Cordeiro et al., 2003; Kosman and Leonard, 2005). The percent disagreement or simple mismatch metric considers as a match the absence of molecules in all compared spectra. This might give misleading results because two samples could be considered close to each other just because they shared many absences of molecules (Kosman and Leonard,
SOA concentrations and yields from the ozonolysis of \( \alpha \)-pinene and BVOC mixtures are shown in Table 1. The average SOA yields (corrected for wall losses) for \( \alpha \)-pinene and BVOC mixtures were \( 0.16 \pm 0.01 \) \((n = 3)\) and \( 0.11 \pm 0.01 \) \((n = 3)\), respectively. The obtained yields for the \( \alpha \)-pinene-only experiments are in reasonable agreement with those reported in the literature for similar BVOC concentration ranges (Pathak et al., 2007; Hatfield and Huff Hartz, 2011).

Surprisingly, in the present study, SOA yields for the BVOC mixture were significantly lower than for the single BVOC (\( \alpha \)-pinene) system. \( \alpha \)-pinene and \( \Delta^3 \)-carene accounted for a major fraction \((\sim 70\%)\) of the total BVOC mixture and thus are expected to make the major contributions to the SOA mass. \( \Delta^3 \)-carene is reported to have a similar SOA yield to \( \alpha \)-pinene (Jonsson et al., 2006; Hatfield and Huff Hartz, 2011) and therefore cannot be responsible for the observed low yield from the precursor mixture. \( \beta \)-pinene and isoprene account for 20\% and 10\% of the total BVOC mixture, respectively. The ozonolysis of these BVOCs generally results in a lower yield for \( \beta \)-pinene compared to \( \alpha \)-pinene (Jonsson et al., 2006) and a very low yield for isoprene \((\sim 0.014)\) (Kleindienst et al., 2007). However, considering that the latter BVOCs account for a small fraction of the reaction mixture their contribution to the total SOA mass is expected to be rather low. It has been shown that the addition of \( \beta \)-pinene to the \( \alpha \)-pinene/O\(_3\) system did not affect the SOA yield significantly (Hatfield and Huff Hartz, 2011). Therefore, the possibility that isoprene is responsible for suppressing SOA formation from the precursor mixture cannot be ruled out. Furthermore, as shown in Table 1, ozone was present in all experiments in excess and thus the differences in yield are not expected to be due to the limited availability of oxidants.

### 3.1 UHR-MS analysis

Representative (-) nanoESI high resolution mass spectra for ambient summer aerosol from the boreal forest site Hyytiälä, Finland, laboratory-generated SOA from ozonolysis of \( \alpha \)-pinene, and the BVOC mixture are shown in Fig. 1. The molecular composition of the organic aerosol at Hyytiälä is found to be strongly affected by air mass origin. Depending on the sampling day, 460–730 molecular formulae were identified in the 10 ambient samples (Kourtchev et al., 2013). The NanoESI mass spectra of the ambient samples is mainly composed of low molecular mass compounds (i.e. peaks below \( m/z \) 350, Fig. 1a) irrespective of the sampling period, which is consistent with previous studies of ambient organic aerosol (OA) from urban (e.g. Riverside, USA, Reemtsma et al., 2006), Cambridge, UK (Rincón et al., 2012), suburban (e.g. Pearl River Delta Region, China, Lin et al., 2012), and rural (e.g. Mount Werner, north-western Colorado, USA, Mazzoleni et al., 2012) environments. This is in contrast to laboratory-generated SOA from both \( \alpha \)-pinene (Fig. 1c) and the BVOC mixture (Fig. 1d) which contain high molecular weight compounds with distinguishable groups.
of dimers. Similar observations were reported in the literature for laboratory-generated SOA from biogenic or anthropogenic BVOCs where UHR mass spectra often contain a large number of oligomers (Reinhardt et al., 2007; Walser et al., 2007; Putman et al., 2012).

Figure 1b shows a mass spectrum containing only those ions that were observed in all 10 Hyytiälä samples (referred to hereafter as “common ions”). Considering that “common ions” exclude all species that occurred during the individual days, they are potentially characteristic of locally formed and emitted OA because their presence is independent of the air mass origin. The BVOC mixture samples have a fewer number of peaks in the dimeric region than the $\alpha$-pinene samples. The total number of assigned formulae in the $\alpha$-pinene and BVOC mixture mass spectra were on average 632 ± 84 and 501 ± 54, respectively (where ± describes the variability between three replicate chamber experiments). A higher number of formulae (about 900) were identified from the negative electrospray ultrahigh-resolution FT-ICR mass spectra of SOA from $\alpha$-pinene ozonolysis in the previous study of Putman et al. (2012). However, the latter study identified formulae in the range $100 < m/z < 850$, whereas we only considered ions below $m/z$ 650. The number of possible empirical formulae assignments increases significantly with higher masses, especially above 400 Da. Because no common ions $> m/z$ 300 are present in the ambient samples, only ions from the monomeric region of the laboratory-generated SOA were used for further comparison with the ambient sample.

In the monomeric region (below $m/z$ 300), the number of formulae in SOA from $\alpha$-pinene and the BVOC mixture were comparable, on average 199 ± 29 and 215 ± 17, respectively (Fig. 1a and b). At first, such a small difference was somewhat puzzling. However, considering that three of the four BVOC-precursor compounds (i.e. $\alpha$-pinene, $\beta$-pinene and $\Delta^3$-carene) used in the mixture are structural isomers, their oxidation with $O_3$ is expected to yield products with similar elemental composition but different structures, which cannot be separated using the analytical technique employed here. For instance, the mass spectra from both chamber experiments and ambient OA were dominated by an ion at $m/z$ 185.0818. While in $\alpha$-pinene experiments this ion corresponded to cis-pinic acid, in the BVOC mixture experiments and Hyytiälä ambient samples this ion was related to three (i.e. cis-pinic acid, homoterpenylic acid, and cis-caric acid) and five (i.e. cis-pinic acid, homoterpenylic acid, limonic acid, ketolimononic acid and cis-caric acid) different compounds, respectively. The separation and identification of these compounds was achieved using LC/MS analysis. The mass spectral differences observed in the monomeric region of SOA produced from the single-precursor and BVOC mixture were mainly associated with variation of the ion intensities caused by differences in the number of products with the same mass in the single component and BVOC mixture experiments as described above.

The ionisation of organic compounds can be affected by the presence of inorganic salts in the analyte solutions, potentially leading to a decrease in MS signal intensity when using direct infusion mass spectrometry methods. Thus, we tested whether the presence of atmospherically abundant salts (e.g. ammonium sulfate) in our filter extracts could cause such a matrix effect and whether this could be responsible for the lack of dimers observed in the ambient samples. Laboratory-generated samples were spiked with ammonium sulfate at atmospherically realistic proportions (30% of the total aerosol mass). The addition of salts suppressed the intensities of all ions in the entire mass range but did not selectively decrease the intensity of ions in the dimeric region (Fig. S1). However, due to competitive ionisation of analytes in the ESI direct infusion analysis of the aerosol samples that are known to have a very complex matrix, the ion intensities do not directly reflect the concentration of the molecules in the sample. Therefore, signal intensities should be interpreted with caution and thus were not considered for the mass spectral comparison in this study. In contrast, LC/MS which is a quantitative technique showed significant difference in the abundances of peaks associated with higher-molecular weight (HMW) compounds between ambient and laboratory-generated samples.
dimer with m/z is plotted as a function of the O/C ratio for each formula in our ambient and laboratory-generated samples. We can thus conclude that anthropogenic emissions. As indicated above, the elemental composition from the boreal forest site only included “common ions” because they are potentially characteristic for locally emitted OA as their presence is independent of the air mass origin. The elemental ratios from the TIED site included “common ions” from 3–4 September 2011, associated with westerly air masses. The composition of the latter samples is discussed in detail in a separate article (Kourtchev et al., 2014). It can be seen from Fig. 3, that the distribution of elemental ratios of laboratory-generated SOA from α-pinene is very similar to that of the BVOC mixture. Moreover, the elemental distribution of laboratory SOA generated from a single precursor and a mixture of BVOCs represent fairly well that of the ambient SOA from Hyytiälä, except that the latter sample contained an additional cluster of molecules as displayed at the upper left part of the diagram. In general, this region is associated with the most reduced/saturated and biogenic origin. Aromatic compounds are typically associated with anthropogenic sources (Henze et al., 2008) and could therefore possibly be fatty acids emitted from a local biogenic source (Kourtchev et al., 2013). In contrast, the VK diagrams of the laboratory-generated SOA were very different from that of the anthropogenic aerosol from the TIED site, which contained a large cluster of ions with low H/C (≤1.0) and O/C (≤0.5) ratios, possibly corresponding to oxidised aromatic hydrocarbons (Mazzoleni et al., 2012). These differences were also apparent when the data was expressed as DBE vs. mass-to-charge ratio (m/z) (Fig. 4). The samples from α-pinene and the BVOC mixtures had very similar DBE distribution with values in the range 1–7. A small number of species observed in the laboratory samples with DBE values of 5–7 were possibly associated with dimers that were formed through accretion reactions (Putman et al., 2012). The DBE distribution of molecules from the Hyytiälä samples clearly resembled those of the α-pinene and BVOC mixtures, except that the ambient sample contained an additional cluster of ions with DBE 0–1. As determined by MS/MS analysis they are attributed to unsaturated and saturated fatty acids (Kourtchev et al., 2013). The DBE plot for OA from TIED was very different from the rest of the compared samples and contained an additional large cluster of molecules with DBE between 7 and 13 (Fig. 4b), once more indicating the presence of oxidised aromatic species. Aromatic compounds are typically associated with anthropogenic sources (Henze et al., 2008) whereas aliphatic compounds can be of both anthropogenic and biogenic origin.

The average O/C and H/C ratios for SOA from α-pinene (0.55 and 1.46, respectively) and the BVOC mixture (0.58 and 1.40) were fairly comparable to the ratios for a sample, is often used to describe the evolution of organic mixtures. Moreover, VK diagrams can also be used to visualise the differences in the elemental composition of different samples. Figure 3 shows an overlaid VK diagram for SOA from (a) α-pinene, the BVOC mixture and a boreal forest sample from Hyytiälä and (b) α-pinene, the BVOC mixture and a sample from the TIED site, which is heavily influenced by anthropogenic emissions. As indicated above, the elemental composition from the boreal forest site only included “common ions” because they are potentially characteristic for locally emitted OA as their presence is independent of the air mass origin. The elemental ratios from the TIED site included “common ions” from 3–4 September 2011, associated with westerly air masses. The composition of the latter samples is discussed in detail in a separate article (Kourtchev et al., 2014). It can be seen from Fig. 3, that the distribution of elemental ratios of laboratory-generated SOA from α-pinene is very similar to that of the BVOC mixture. Moreover, the elemental distribution of laboratory SOA generated from a single precursor and a mixture of BVOCs represent fairly well that of the ambient SOA from Hyytiälä, except that the latter sample contained an additional cluster of molecules as displayed at the upper left part of the diagram. In general, this region is associated with the most reduced/saturated and biogenic origin. Aromatic compounds are typically associated with anthropogenic sources (Henze et al., 2008) and could therefore possibly be fatty acids emitted from a local biogenic source (Kourtchev et al., 2013). In contrast, the VK diagrams of the laboratory-generated SOA were very different from that of the anthropogenic aerosol from the TIED site, which contained a large cluster of ions with low H/C (≤1.0) and O/C (≤0.5) ratios, possibly corresponding to oxidised aromatic hydrocarbons (Mazzoleni et al., 2012). These differences were also apparent when the data was expressed as DBE vs. mass-to-charge ratio (m/z) (Fig. 4). The samples from α-pinene and the BVOC mixtures had very similar DBE distribution with values in the range 1–7. A small number of species observed in the laboratory samples with DBE values of 5–7 were possibly associated with dimers that were formed through accretion reactions (Putman et al., 2012). The DBE distribution of molecules from the Hyytiälä samples clearly resembled those of the α-pinene and BVOC mixtures, except that the ambient sample contained an additional cluster of ions with DBE 0–1. As determined by MS/MS analysis they are attributed to unsaturated and saturated fatty acids (Kourtchev et al., 2013). The DBE plot for OA from TIED was very different from the rest of the compared samples and contained an additional large cluster of molecules with DBE between 7 and 13 (Fig. 4b), once more indicating the presence of oxidised aromatic species. Aromatic compounds are typically associated with anthropogenic sources (Henze et al., 2008) whereas aliphatic compounds can be of both anthropogenic and biogenic origin.

The average O/C and H/C ratios for SOA from α-pinene (0.55 and 1.46, respectively) and the BVOC mixture (0.58 and 1.40) were fairly comparable to the ratios for
Fig. 3. Van Krevelen diagrams for compounds containing only carbon, hydrogen and oxygen in the samples from (a) \( \alpha \)-pinene/O\(_3\), BVOC mixture/O\(_3\) and the boreal forest, Hyytiälä, Finland; and (b) \( \alpha \)-pinene/O\(_3\), BVOC mixture/O\(_3\) and Tivoli Industrial Estate and Docks, Cork, Ireland. Filled blue circles correspond to species from \( \alpha \)-pinene/O\(_3\) experiment, red diamonds to BVOC mixture/O\(_3\) experiments, black hashes to the Hyytiälä samples and black crosses to the Tivoli Industrial Estate and Docks, Cork samples.

Fig. 4. Double bond equivalents (DBE) vs. mass to charge (\( m/z \)) ratios for all compounds containing only carbon, hydrogen and oxygen in the samples from (a) \( \alpha \)-pinene/O\(_3\), BVOC mixture/O\(_3\) experiments and boreal forest, Hyytiälä, Finland, and (b) \( \alpha \)-pinene/O\(_3\), BVOC mixture/O\(_3\) experiments and Tivoli Industrial Estate and Docks, Cork, Ireland. Filled blue circles correspond to species from \( \alpha \)-pinene/O\(_3\) experiment, red diamonds to BVOC mixture/O\(_3\) experiments, black hashes to the Hyytiälä samples and black crosses to the Tivoli Industrial Estate and Docks, Cork samples.

OA from Hyytiälä (0.52 and 1.48) (Kourtchev et al., 2013) but higher than those from TIED (0.36 and 1.1). The H/C value for laboratory-generated SOA indicated that the identified SOA molecules are of aliphatic and alicyclic nature (Putman et al., 2012). The elemental O/C ratios found in this study are within the range obtained for SOA generated in laboratory experiments from a number of BVOCs, for example, \( \alpha \)-pinene/O\(_3\) (0.42–0.55) (Putman et al., 2012), and photo-oxidation of isoprene under low-NO\(_x\) conditions (0.54) (Nguyen et al., 2011). These values are slightly higher than the average values (0.43–0.46) reported for the limonene/O\(_3\) system (Bateman et al., 2009).

It has been demonstrated that O/C ratio, as measured by the Aerodyne High Resolution Time-of-Flight Aerosol Mass Spectrometer (HR-ToF-AMS), is positively correlated with the hygroscopic parameter of the organic fraction (Wu et al., 2013), which in turn is related to the cloud condensation nucleus (CCN) activity of aerosol particles (Petters and Kreidenweis, 2007). Therefore, considering that the O/C ratio for SOA generated from the \( \alpha \)-pinene-only system is very similar to that from the BVOC mixture and OA from Hyytiälä, we suggest that the simplified BVOC system can possibly be used for parameterisation of OA at the boreal site.

O/C ratios may not accurately describe the degree of oxidation of organics, because other non-oxidative processes (e.g. hydration and dehydration) can also affect these parameters (Kroll et al., 2011). Carbon oxidation state (OS\(_C\)), on the other hand, may change differently upon oxidation, but the average OS\(_C\) will increase. Therefore, OS\(_C\) could be a useful metric for the degree of oxidation of organic species in the atmosphere and can serve as a key variable.
to describe organic mixtures. OSc is shown to be strongly linked to aerosol volatility and thus is a useful parameter for the classification of SOA (Hao et al., 2011). Carbon oxidation state can be calculated from the following equation:

\[ \text{OS}_C = - \sum_i \text{OS}_i \frac{n_i}{n_C}, \]

where \( \text{OS}_i \) is the oxidation state associated with element \( i \) and \( n_i/n_C \) is the molar ratio of element \( i \) to carbon (Kroll et al., 2011). Figure 5 shows the overlaid carbon oxidation state vs. the number of carbon atoms for molecules from laboratory-generated SOA and the ambient boreal samples. The OSc distribution of laboratory-generated SOA generally resembles that of the ambient samples ranging from \(-1.7\) to \(+1\) and only a few species have an oxidation state greater than \(+1\). Interestingly, molecules with OSc greater than \(+1\) were only observed in SOA from the BVOC mixture experiments and ambient aerosol from Hyytiälä. Such compounds are expected to contain several carbonyl groups. However, literature data suggests that the average oxidation state of organic aerosol rarely exceeds this value because species with several carbonyl groups are highly unstable and will rapidly decompose to smaller molecules (Kroll et al., 2011). Considering that the studied BVOC mixture mainly contained monoterpenes, which are structural isomers and the fact that highly oxidised molecules were not observed in the SOA generated from the ozonolysis of \( \alpha \)-pinene, it is likely that species with OSc > \(+1\) were produced from the ozonolysis of isoprene. Moreover, in the BVOC mixture experiments cross reactions between radicals and oxidation products of the different BVOCs are expected to occur which may lead to formation of a complex range of species. These highly oxidised species are worthy of further investigation. It should be mentioned that highly oxidised multifunctional molecules have been previously observed in both ambient air (Hyytiälä) and in chamber experiments of \( \alpha \)- and \( \beta \)-pinene ozonolysis (Ehn et al., 2012).

The majority of the species exhibited OSc values between \(-1\) and \(+1\) with 15 or less carbon atoms, suggesting that they are semi- and low-volatile organic compounds corresponding to “fresh” and “aged” SOA produced by multistep oxidation reactions (Jimenez et al., 2009; Kroll et al., 2011). Compared to chamber samples, the Hyytiälä samples additionally contained ions with OSc < \(-1\) and more than 7 carbon atoms which is characteristic of primary biomass burning aerosol (Kourtchev et al., 2013).

3.2 Comparison using statistical tools

Figure 6 shows the fraction of molecular formulae below 300 Da found in both the laboratory-generated SOA and the ambient samples relative to the total number of formulae in the ambient samples. Evidently, the molecular composition of SOA from both the BVOC mixture and \( \alpha \)-pinene represented the overall composition of the ambient sample from the boreal forest site reasonably well, with 72.3 ± 2.5 % (\( n = 3 \)) and 69.1 ± 3.0 % (\( n = 3 \)) common ions, respectively. Although, the BVOC mixture resulted in a slightly higher number of common formulae than that from the boreal forest aerosol compared to the pure \( \alpha \)-pinene-SOA, the difference in the mean values among the treatment groups is not large enough to exclude the possibility that the difference is due to random sampling variability; according to ANOVA test, the difference was not statistically significant (\( p = 0.348 \)). In contrast, the molecular composition of laboratory-generated SOA was substantially different from that of the anthropogenically affected TIED site. The fraction of common molecular formulae from \( \alpha \)-pinene and the BVOC mixture relative to the total number of ions from the TIED sample was only 16.1 ± 1.7 % and 16.9 ± 1.2 %, respectively, indicating the very different sources of organic compounds in these samples. The inverse comparison of the fraction of common formulae relative to the total number of formulae in the laboratory-generated samples in the monomeric region indicated that laboratory-generated SOA contained 20–25 % formulae that were not observed in the boreal samples. We suggest that these molecules are first generation products that are possibly oxidised with time in the atmosphere resulting in aged oxidation products.

Laboratory-generated and ambient samples were also compared by hierarchical cluster analysis (HCA) that divides samples into groups (clusters) of similar molecular composition. HCA separated the samples into three clusters (Fig. 7): (1) \( \alpha \)-pinene (replicates from three different experiments); (2) BVOC mixture (three replicates) together with the common ions of the ambient samples from Hyytiälä; and (3) common ions of the ambient samples from TIED. The branches
Fig. 6. Fraction of molecular formulae below 300 Da found in SOA from both laboratory-generated and ambient samples relative to the total number of formulae in the ambient samples. The error bars represent the standard deviation of three replicate chamber experiments.

in the tree diagram (dendrogram) represent the average distance between the connected samples. It is evident from the dendrogram that all replicate samples from the α-pinene and the BVOC mixture experiments cluster together, implying very good reproducibility of the applied technique (i.e. smog chamber experimental and MS analysis) to separate two experimental conditions relative to each other. Although the α-pinene data is separated from the BVOC mixture and Hyytiälä cluster, the linkage distance is not large enough to conclude that their chemical composition is very different. On the other hand, the data from TIED was classified into a separate cluster confirming that its molecular composition is very different from the rest of the samples. The results from HCA clearly support the findings obtained from the statistical analysis and other visualisation methods (van Krevelen diagrams, carbon oxidation state, DBE).

4 Conclusions

The detailed molecular composition of background ambient aerosol from a boreal forest site (Hyytiälä, Finland), an urban location (Cork, Ireland), laboratory-generated SOA from α-pinene and a mixture of four BVOCs were compared using nanoESI UHRMS. Our results demonstrate that the molecular composition of SOA in the monomeric mass range up to m/z 300 from both the ozonolysis of the BVOC mixture and α-pinene represented the overall composition of the ambient sample from the boreal forest site fairly well, with 72.3 ± 2.5 % (n = 3) and 69.1 ± 3.0 % (n = 3) common ions, respectively. Other atmospheric oxidants (e.g. OH radicals and NOx) will certainly influence the composition of SOA and their reaction products are likely to explain some of the remaining molecules that were not observed in our laboratory-generated SOA.

The elemental (O/C and H/C) ratios of SOA from the α-pinene-only system were very similar to those from the

BVOC mixtures and ambient aerosol from boreal forest. Considering that the O/C ratio is positively correlated with hygroscopicity of the organic fraction, the simplified α-pinene-only system can potentially be useful for parameterisation of boreal OA. A specific class of CHO compounds identified as fatty acids was present exclusively in the ambient samples suggesting that the composition at the boreal forest OA is also influenced by primary emissions. In contrast, the overall molecular composition of the urban samples is dominated by a high number of oxidised aromatic hydrocarbons and is very different from the boreal and laboratory-generated OA. The presence of dimers observed exclusively in the laboratory samples requires further investigation. Thus it appears that while the laboratory experiments performed here and in other studies do a fairly good job of simulating biogenic SOA formation (as indicated by the high level of agreement between the laboratory SOA and the boreal forest samples in the monomeric mass region of the mass spectra), they do not completely replicate the processes occurring in the ambient atmosphere.

Supplementary material related to this article is available online at http://www.atmos-chem-phys.net/14/2155/2014/acp-14-2155-2014-supplement.pdf.
Acknowledgements. Research at the University of Cambridge was supported by a M. Curie Intra-European fellowship (project no. 254319) and the European Research Council (ERC starting grant 279405). The work at University College Cork was funded by the European Commission (project EUROCHAMP-2, contract no. 228335). The work at University of Helsinki was supported by Academy of Finland Center of Excellence program (project no. 1118615). The authors would like to acknowledge the use of NAME, developed and provided by the Atmospheric Dispersion Group, UK Met Office.

Edited by: S. A. Nizkorodov

References


I. Kourtchev et al.: Molecular composition of biogenic secondary organic aerosols


Nizkorodov, S. A., Laskin, J., and Laskin, A.: Molecular chemistry of organic aerosols through the application of high resolution


