Faraday Discussions

RSCPublishing

PAPER

Average chemical properties and potential formation pathways of highly oxidized organic aerosol

Kelly E. Daumit,[a] Sean H. Kessler[b] and Jesse H. Kroll*[ab]

Measurements of ambient organic aerosol indicate that a substantial fraction is highly oxidized and low in volatility, but this fraction is generally not reproduced well in either laboratory studies or models. Here we describe a new approach for constraining the viable precursors and formation pathways of highly oxidized organic aerosol, by starting with the oxidized product and considering the possible reverse reactions, using a set of simple chemical rules. The focus of this work is low-volatility oxidized organic aerosol (LV-OOA), determined from factor analysis of aerosol mass spectrometer data. The elemental composition and volatility of the aerosol enable the determination of its position in a three-dimensional chemical space (defined by H/C, O/C, and carbon number) and thus its average chemical formula. Consideration of possible back-reactions then defines the movement taken through this chemical space, constraining potential reaction pathways and precursors. This approach is taken for two highly oxidized aerosol types, an average of LV-OOA factors from ten field campaigns (average formula $C_{10.5}H_{13.4}O_{7.3}$), and extremely oxidized LV-OOA (from Mexico City, average formula $C_{10}H_{12.1}O_{8.4}$). Results suggest that potential formation pathways include functionalization reactions that add multiple functional groups per oxidation step, oligomerization of highly oxidized precursors, and, in some cases, fragmentation reactions that involve the loss of small, reduced fragments.

1 Introduction

Organic aerosol (OA) constitutes a substantial fraction (20%–90% by mass[1]) of atmospheric fine particulate matter, and thus a thorough understanding of its formation and evolution is necessary in order to better evaluate the effects of particulate matter on climate and human health. However, despite substantial research efforts, there remains a great deal unknown about the chemical mechanisms important to OA. Especially uncertain are the sources and chemistry of the

[a]MIT Department of Civil and Environmental Engineering, Cambridge, MA, United States. E-mail: jhkroll@mit.edu; Tel: +617 253 2409

[b]MIT Department of Chemical Engineering, Cambridge, MA, United States
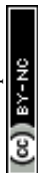
most oxidized fraction of OA, sometimes termed "humic-like substances" (HULIS) (when measured using offline techniques of fractionated filter samples),[2] or low-volatility oxidized organic aerosol (LV-OOA) (when determined from factor analysis of online aerosol mass spectrometer data).[3] This class of atmospheric organic matter is ubiquitous in the atmosphere,[1,3] highly oxidized (with an average carbon oxidation state of approximately 0–1),[4] exceedingly low in volatility (saturation concentration $c^* \sim 0.1$ to $<10^{-7}$ µg m^{-3}),[5] and generated relatively quickly (over time scales of $\sim$1–3 days).[3,6]

While this highly oxidized OA is presumed to be secondary in nature (formed from chemical transformations of gas-phase organics that subsequently condense as particulate matter), its detailed formation pathways are poorly understood and have not been adequately reproduced in the laboratory. Laboratory chamber experiments of secondary organic aerosol (SOA) formed from various classes of precursors tend to produce aerosol that is relatively unoxidized (average carbon oxidation state $(OS_C) \sim -1.1$ to 0.1).[4,7] Several efforts to "age" OA in laboratory chambers have also fallen short of reaching oxidation levels as high as those measured for ambient aerosol.[8,9] Those studies that have succeeded in generating highly oxidized SOA have generally been able to do so only with a select few precursors[10,11] or at OH exposures substantially higher than those relevant to the formation of oxidized aerosol in the atmosphere.[12–14] Alternative routes for aging aerosol, such as aqueous-phase oxidation[15] or photosensitized reactions[16] have received recent attention, but there is still much unknown about their chemistry and role in OA formation.

Recent modeling efforts have also attempted to reproduce the amount and degree of oxidation of organic aerosol as measured in field studies, by introducing additional aging chemistry to simulate oxidation beyond the initial SOA formation.[17–19] However, such models, whether simulating a specific environment (*e.g.*, Mexico City[17,20,21]), or SOA formation from a specific class of precursors,[18,22] still tend to be unsuccessful in reproducing the formation of highly oxidized aerosol. They either underpredict OS_C at a given loading[17,20,21] or predict the formation of highly oxidized OA over longer time scales than what is observed.[23] These model-measurement discrepancies suggest major gaps in our understanding of the chemistry underlying the formation of highly oxidized organic aerosol. Unfortunately at present we have very little information as to whether these gaps relate to uncertainties in reaction mechanisms, SOA precursors, or both.

Here we present a new approach for constraining the formation of highly oxidized aerosol, involving the use of chemical properties of the aerosol to better understand its possible formation pathways and precursors. In contrast to most modeling approaches that start with potential precursors and simulate their forward evolution in an attempt to generate products chemically similar to those measured in the atmosphere, we start with the known products (highly oxidized OA components) and work backwards toward reactants (SOA precursors). This approach is similar to the retrosynthetic approach of Pun *et al.* for assessing formation pathways of atmospheric water-soluble organic compounds.[24] That approach was highly molecule- and reaction-specific, involving the identification of detailed chemical reactions that link individual precursors to individual aerosol components. Here we take a more generalized view of the chemistry, assessing classes of reactions that could ultimately lead to the formation of highly oxidized, low-volatility organics. Such an approach requires knowledge of the
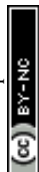
chemical properties of the "targets" (molecules representing the oxidized fraction of the aerosol); we define these in terms of the average chemical formulas of their constituent molecules, as determined from properties measured in field studies as well as structure–activity relationships. From these targets, a set of simple, general rules governing possible atmospheric reactions lets us work backwards, allowing for identification of viable reaction pathways and aerosol precursors (which are also defined in terms of their average chemical properties). By focusing on generic processes rather than detailed chemical structures and mechanisms, we can draw general conclusions about the types of processes and precursors that might yield highly oxidized organic species. These can be used to provide guidance in identifying relevant reaction conditions and precursors in future studies.

In the following section we describe this methodology in detail, including the chemical space we use to describe the organics, the characterization of our targets, and the types and characteristics of reverse pathways considered. Results – the chemical characteristics of our targets (constrained by AMS measurements of LV-OOA components) and the possible range of aerosol formation pathways – are then described. We conclude with discussions of implications of this work for our understanding of highly oxidized organic aerosol, as well as potential areas of future work.

2 Methodology

We describe molecules in terms of their position in a three-dimensional chemical space, defined by their carbon number (n_C), oxygen-to-carbon ratio (O/C), and hydrogen-to-carbon ratio (H/C). O/C and H/C have already seen substantial use in describing the average properties of atmospheric organic aerosol using the two-dimensional van Krevelen diagram;[7,25] however, this provides no information on the size of the molecule (a crucial factor governing volatility). By considering n_C as a third dimension to the van Krevelen space, we can represent any molecular formula $C_xH_yO_z$ as a single, unique point, and can relate its position in chemical space to volatility.[12,13] This space is equivalent to the three-dimensional space defined by carbon number, hydrogen number, and oxygen number (n_C vs. n_H vs. n_O); we choose to use O/C and H/C here since they are also used in the van Krevelen diagram, and are more amenable to visualization and comparison of elemental composition (hydrogen and oxygen content) of molecules of different sizes. A previously proposed alternative three-dimensional space uses molecular weight, heteroatom mass, and double bond equivalents.[26] While this has advantages for describing certain atmospheric systems, especially those with nitrogen or sulfur containing species, the space proposed here is more conducive to the description of atmospheric reactions in terms of changes to carbon, hydrogen, and oxygen content of the molecules.

The three-dimensional space we use here is closely related to many of the two-dimensional frameworks that have recently been used to describe and/or model organic aerosol. These include the van Krevelen plot,[7] polarity vs. n_C,[27] OS_C vs. saturation vapor concentration (c^*),[3,19] the f44 vs. f43 space for plotting AMS data,[25,28] OS_C vs. n_C,[4] and n_O vs. n_C.[18] Although adding a third dimension introduces complexity to any of these descriptions, it allows for more chemical information to be represented. It explicitly includes the effects of carbon number, which are missing from functional-group-based frameworks (e.g., van Krevelen
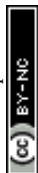
and f44–f43), and it requires fewer assumptions about the identity or distributions of functional groups than frameworks that reduce the descriptions of functional groups to a single dimension (polarity–n_C, OS_C–n_C, OS_C–c^*, n_O–n_C). We note that isomeric species of a given formula $C_xH_yO_z$ are not distinguished even in this 3-D space, which can lead to errors in the calculation of c^* when the identity of the functional groups changes, as discussed below. Nonetheless, chemical transformations and properties (*e.g.*, c^*) can be represented in this space in a straightforward manner.

I. Characterization of the target

Our general approach is to start with the product (molecules representative of highly oxidized organic aerosol) and work backwards. We refer to this species as the "target" and the backwards reactions as "transforms" following the conventions of retrosynthetic analysis.[29] Key to this approach is the accurate determination of the chemical formula (position in O/C–H/C–n_C space) of the target.

Since the target in this case (highly oxidized organic aerosol) is not a single molecule, but rather a complex mixture, we define it in terms of its average properties and average chemical formula. In the present study we determine this from field studies using the Aerodyne High Resolution-Time of Flight-Aerosol Mass Spectrometer (HR-ToF-AMS, or simply AMS),[30,31] which provides measurements of the elemental ratios of OA.[32,33] Positive matrix factorization (PMF) of AMS data[34] typically yields several factors, of which LV-OOA is the most highly oxidized. We take this to be representative of the most oxidized organic aerosol, and use measurements of LV-OOA elemental ratios to define the location of our target in van Krevelen space.

Because the AMS provides no information regarding carbon number (the third dimension of our chemical space), we determine n_C from field measurements of aerosol volatility (c^*).[5,36,38] The carbon number can be estimated from c^* using structure–activity relationships (SARs) for determining vapor pressure. In this work we use SIMPOL,[39] allowing for the direct determination of c^* from n_C and the functional group abundance, as described below. We focus on partitioning only into the condensed organic phase, as described by c^*. Partitioning into liquid water could also be included using this general approach, but it would require use of a SAR for estimating the volatility over water (Henry's law constant), which is beyond the scope of this work.

Key inputs for SIMPOL (and most other SARs for estimating vapor pressures) are the abundances of different functional groups in the molecule.[39] In order to determine these from our values for n_C, H/C, and O/C, we make two assumptions. The first is that functional groups in the molecule are limited to carbonyls, hydroxyls, or some combination of the two (*e.g.*, carboxylic acids). Since each functional group contains only one oxygen atom, their abundances can be related to the oxygen number of the molecule:

$$n_O = n_C O/C = n_{-OH} + n_{=O} \qquad (1)$$

where n_O, n_C, and O/C are as defined above, and n_{-OH} and $n_{=O}$ are the number of hydroxyl groups and carbonyl groups in the molecule, respectively. While other functional groups are also likely to be present in organic aerosol, several groups can be approximated using this simple treatment, as discussed below.
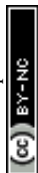
The second assumption is that all sites of unsaturation (double bond equivalents, DBE) in the molecule arise from carbonyl groups:

$$n_{=O} = DBE = 1 + n_C\left(1 - \frac{1}{2}H/C\right) \tag{2}$$

We therefore assume that our target contains no carbon–carbon double bonds or rings. While this is reasonable for aliphatic C=C bonds, which are highly susceptible to oxidation and unlikely to survive significant atmospheric processing, cyclic (or aromatic) structures may be present in highly oxidized OA. The effect of neglecting any rings present is to overestimate $n_{=O}$ and underestimate n_{-OH}, thereby overestimating c^* somewhat.

The two assumptions above (eqn (1) and (2)) allow for the straightforward derivation of expressions relating elemental ratios, functional groups, volatility, and carbon number. The number of hydroxyl groups (n_{-OH}) can be determined by combining eqn (1) and (2):

$$n_{-OH} = n_O - n_{=O} = n_C O/C - \left(1 + n_C\left(1 - \frac{1}{2}H/C\right)\right)$$

$$= -1 + n_C\left(O/C + \frac{1}{2}H/C - 1\right) \tag{3}$$

The carbon number (n_C) and functional group abundances (n_{-OH} and $n_{=O}$) allow for the calculation of c^* ($\mu g\ m^{-3}$) using SIMPOL:

$$\log_{10} c^* = \log_{10}(\alpha P_{sat}) = \log_{10}\alpha + b_0 + b_C n_C + b_{=O} n_{=O} + b_{-OH} n_{-OH}$$

$$= \log_{10}\alpha + b_0 + b_C n_C + b_{=O}\left(1 + n_C\left(1 - \frac{1}{2}H/C\right)\right) \tag{4}$$

$$+ b_{-OH}\left(-1 + n_C\left(O/C + \frac{1}{2}H/C - 1\right)\right)$$

where P_{sat} is the vapor pressure in atm, α is the conversion from P_{sat} to c^*, and the b terms are the different group contribution terms for quantifying the contribution of each chemical moiety to vapor pressure: b_0 is the zero order term, b_C is the carbon number term, $b_{=O}$ is the carbonyl group term, and b_{-OH} is the hydroxyl group term (for MW = 200 g mol^{-1} and T = 293 K, equal to 1.79, −0.438, −0.935, and −2.23, respectively).[39] The conversion factor α is given by $\alpha = 10^6(MW)/RT$, or 8.314×10^9 at 293 K, where MW is the average molecular weight (g mol^{-1}) of the molecules making up the absorbing phase, R is the ideal gas constant (8.21×10^{-5} atm m^3 mol^{-1} K^{-1}), and T is the temperature in K. For these calculations we use an assumed MW value of 200 g mol^{-1}; the actual value used has little effect on the results. Rearranging eqn (4) to solve for carbon number, and substituting in values of $n_{=O}$ and n_{-OH} from eqn (2) and (3), we obtain

$$n_C = \frac{\log_{10} c^* - \log_{10}\alpha - b_0 - b_{=O} + b_{-OH}}{b_C + b_{=O}\left(1 - \frac{1}{2}H/C\right) + b_{-OH}\left(O/C + \frac{1}{2}H/C - 1\right)} \tag{5}$$

Eqn (5) allows for the determination of carbon number from elemental (H/C, O/C) and volatility (c^*) data. This approach, used in previous descriptions of heterogeneous oxidation systems,[12,13] is similar to that of Donahue *et al.*[19] for
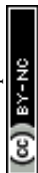
estimating carbon number from OS_C (or O/C) and $c*$ alone; however that approach required making assumptions about functional group distribution (specifically that $n_{-OH} = n_{=O}$). With the explicit inclusion of H/C, we can directly calculate n_{-OH} and $n_{=O}$, which allows for a more accurate estimation of n_C.

Eqn (5) relies critically on the assumption that target molecules contain only hydroxyl and carbonyl groups. While molecules containing other functional groups will be less accurately represented, many such functional groups are reasonably approximated as carbonyls, alcohols, or some combination of the two. Because the vapor pressure effect of an acid group ($b_{(=O)OH} = -3.58$) is similar to that of a hydroxyl plus a carbonyl ($b_{-OH} + b_{=O} = -3.165$), treating an acid as the sum of the two introduces only a minor error. Likewise, epoxides are very similar to carbonyls in terms of both bonding and effect on volatility: like carbonyls, they contribute one DBE, and the vapor pressure effect of an ether ($b_{-O-} = -0.718$) is similar to that of a carbonyl. Similarly, nitrate and hydroperoxyl groups are connected to the carbon skeleton *via* a single C–O bond and so have similar bonding as a hydroxyl, and also have roughly the same vapor pressure effect ($b_{-NO_3} = -2.23$, $b_{-OOH} = -2.48$). These functional groups have additional oxygen atoms, but in the AMS these are likely lost as NO_2 or OH fragments[40] and so it is likely that only one oxygen atom is measured (as is the case for a hydroxyl group). Since these moieties contribute one fewer hydrogen than a hydroxyl group, the measured H/C may not be exactly the same as for the corresponding alcohol, but this will have only a minor effect on results. Unfortunately, not all functional groups are as well-represented in the above treatment. Acyclic ethers have the same bonding (and contribution to O/C and H/C) as hydroxyl groups, but have a substantially smaller effect on vapor pressure ($b_{-O-} = -0.718$ *vs.* $b_{-OH} = -2.23$). Similarly, esters, like acids, will be treated as a carbonyl plus a hydroxyl group,
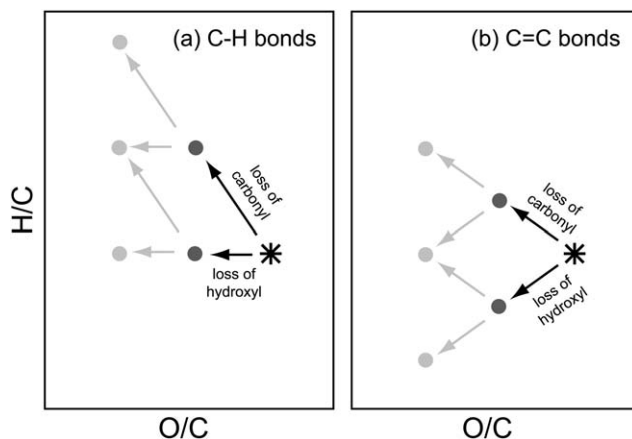

Fig. 1 Possible transforms (back-reactions) available to a target molecule, assuming functionalization reactions only (with no change to carbon number). Loss of functional groups is illustrated in terms of movement in van Krevelen space, depicting changes in H/C and O/C. The exact trajectory depends on the identity both of the functional group and the bond being replaced. Panel a: functionalization of C–H bonds (saturated case), with slopes of −2 for carbonyl groups and 0 for hydroxyl groups. Panel b: functionalization of C=C bonds or rings (unsaturated case), with slopes of −1 for carbonyl groups and +1 for hydroxyl groups. In either case, changes to more than one functional group (light grey arrows and points) can be described by simple vector addition.
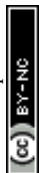
but their actual effect on vapor pressure is far smaller ($b_{(=O)O-} = -1.20$ *vs.* $b_{-OH} + b_{=O} = -3.165$). There is no good way to account for these errors using elemental ratios alone; direct measurements of functional group abundances in OA (*e.g.*, using FTIR[41]) would thus greatly aid in this general approach for determining the chemical formula and properties of the target.

We note there are ways to determine the chemical formula (position in a given chemical space) other than the elemental- and volatility-based approach described above. One widely-used technique is ultrahigh-resolution electrospray ionization mass spectrometry (UHR-ESI-MS) of filter samples to directly determine exact chemical formulas.[42,43] There are currently far fewer measurements of ambient OA by UHR-ESI-MS than by AMS, but this is another promising technique for characterizing the target. While concerns such as variations in sensitivity toward different compounds, and the possibility of fractionation arising from dissolution in a given solvent, still need to be addressed, estimates of OS_C and n_C from AMS and ESI measurements are generally in good agreement.[4]

II. Transforms

Once the target is well-defined, transforms (the reverse of atmospheric reactions) allow for the formation chemistry to be traced backwards toward potential precursors. Again, the aim is not to identify specific formation pathways and precursor molecules, but rather to examine the key features of viable ones. Here, we consider three reaction classes—functionalization, fragmentation, and oligomerization—each of which is discussed below. Simple chemical rules for each reaction class help constrain their potential role and possible precursors.

a) Functionalization. Functionalization reactions are those that add (or interconvert) functional groups without other changes to the carbon skeleton. They are easily represented in a 2-D van Krevelen plot since they involve no change in n_C. Movement through this space has been used previously to describe different functional group additions.[7,25] Reverse functionalization reactions connecting target molecules back to their precursors involve the loss of carbonyl or hydroxyl groups, as shown in Fig. 1. The exact direction in van Krevelen space depends not only on the functional group involved but also the nature of the precursor, specifically the bond (C–H *vs.* C=C) being replaced by the functional group. Fig. 1a shows the "saturated" case, in which the functional groups replace C–H bonds, which has been considered previously.[7] Removing a single functional group decreases O/C by $1/n_C$, with the effect on H/C depending on the identity of the group. Loss of a carbonyl decreases H/C by $2/n_C$ (following a slope of -2 on the van Krevelen plot), while removing a hydroxyl has no effect on the H/C ratio (following a slope of 0). These steps are fully additive, such that the loss of multiple functional groups can be described simply by the sum of the vectors associated with the individual groups. For example, removing a hydroxycarbonyl or carboxylic acid (the equivalent of removing one hydroxyl and one carbonyl) decreases O/C by $2/n_C$ and increases H/C by $2/n_C$ (thereby following a slope of -1). Fig. 1b shows the "unsaturated" case, in which the functional group replaces a carbon-carbon double bond (or C–C single bond within a ring). Loss of a functional group still decreases O/C by $1/n_C$, but now removing a carbonyl increases H/C by only $1/n_C$ (slope = -1), whereas removing a hydroxyl actually decreases H/C, by $1/n_C$ (slope = $+1$). Note that the oxidation of C=C bonds typically involves
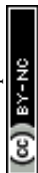
changes to two functional groups, so movement through this space can be quite rapid. In either case (saturated or unsaturated), for a given n_C, each point in van Krevelen space corresponds to a value of c^*, allowing the change in volatility associated with any transform to be easily computed.

We can define the range of potential precursors of our target by providing two fundamental limits on what reverse reactions are possible: (1) a molecule cannot lose more of a given functional group than it actually has; and (2) going backwards, the average carbon oxidation state cannot increase, since reduction reactions are thermodynamically unfavorable. This can significantly narrow the area in chemical space that defines the potential precursors, as discussed in the Results section.

b) Fragmentation. Fragmentation reactions decrease the carbon number of the molecule *via* the cleavage of carbon–carbon bonds. Since these reactions are oxidative, they will involve the addition of functional groups at the site of the bond breakage. (The scission of a single bond within a ring is therefore not considered a fragmentation reaction, and is instead a functionalization reaction, involving the addition of multiple functional groups.) Fragmentation reactions can result in rapid oxidation of the carbon, since they can lead to a high abundance of functional groups on relatively small molecules. However, the decrease in carbon number can also lead to a substantial increase in volatility. This tradeoff likely limits the importance of fragmentation pathways for the formation of very oxidized, low-volatility species.

Because n_C changes during fragmentation, simple 2-D van Krevelen plots are not sufficient to accurately capture these transformations. Furthermore, because of the range of possible precursors (with any number of carbon, oxygen, or hydrogen atoms), there are no simple generic transforms to describe this change. Nonetheless, we can apply constraints that considerably narrow the chemical space associated with the possible transforms and precursors. One approach is to assume that functional groups are evenly distributed across all the carbon in a molecule, and that this distribution is conserved during the fragmentation process (with the exception of any functional groups added during fragmentation). This statistical treatment of fragmentation is similar to that used in the recently-developed Statistical Oxidation Model.[18] For the present study, we assume that two functional groups (one hydroxyl and one carbonyl) are added to a given fragment, but this distribution can be altered as necessary. Such assumptions (plus the requirement that n_C of the precursor must exceed that of the target), give the following constraints: $n_{C,p} > n_{C,t}$, $n_{-OH,p} = (n_{-OH} - 1)(n_{C,p}/n_{C,t})$, and $n_{=O,p} = (n_{=O,t} - 1)(n_{C,p}/n_{C,t})$, where subscript "t" refers to the target and subscript "p" refers to the precursor. The subtraction of 1 from $n_{-OH,t}$ and $n_{=O,t}$ accounts for the addition of the two functional groups upon fragmentation.

An alternative, more exact approach is to allow the functional groups to be localized somewhere on a molecule, rather than distributed evenly across it. This allows for a broader range of potential precursors, since fragments can have varying degrees of oxidation (*i.e.*, one that is more oxidized than the other). In this case, the number of functional groups in the target provides constraints on precursors: after taking into account any new functional groups, the precursor cannot have fewer hydroxyl or carbonyl groups than the target. Again, if we assume that two groups (one hydroxyl and one carbonyl) are added to the fragment, thereby increasing $n_{=O}$ and n_{-OH} each by one, we obtain the following constraints: $n_{C,p} > n_{C,t}$, $n_{-OH,p} \geq (n_{-OH,t} - 1)$, and $n_{=O,p} \geq (n_{=O,t} - 1)$.
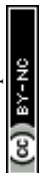
Table 1 HR-ToF-AMS data from measurements at 10 field locations. Elemental ratios are reported for the average LV-OOA PMF factor at each site and are used to calculate each average LV-OOA OS_C ($\sim 2O/C - H/C$)

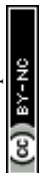| Measurement Location	O/C	H/C	OS_C	Reference
Riverside, CA (2005)	0.72	1.27	0.17	Docherty 2011[45]
Mexico City Aircraft (2006)	1.02	1.12	0.92	DeCarlo 2010[35]
Mexico City Ground (2006)	0.84	1.21	0.47	Aiken 2008[32]
Kaiping, China (2008)	0.64	1.30	−0.02	Huang 2011[37]
Barcelona, Spain (2009)	0.75	1.18	0.32	Mohr 2012[46]
Paris, France (2009)	0.73	1.33	0.13	Crippa 2013[47]
New York City, NY (2009)	0.63	1.29	−0.03	Sun 2011[48]
Hong Kong, China (2009)	0.59	1.26	−0.08	He 2011[49]
Shanghai, China (2010)	0.65	1.49	−0.19	Huang 2012[50]
Sacramento, CA (2010)	0.54	1.32	−0.24	Setyan 2012[51]
Heshan, China (2010)	0.55	1.30	−0.20	Gong 2012[52]
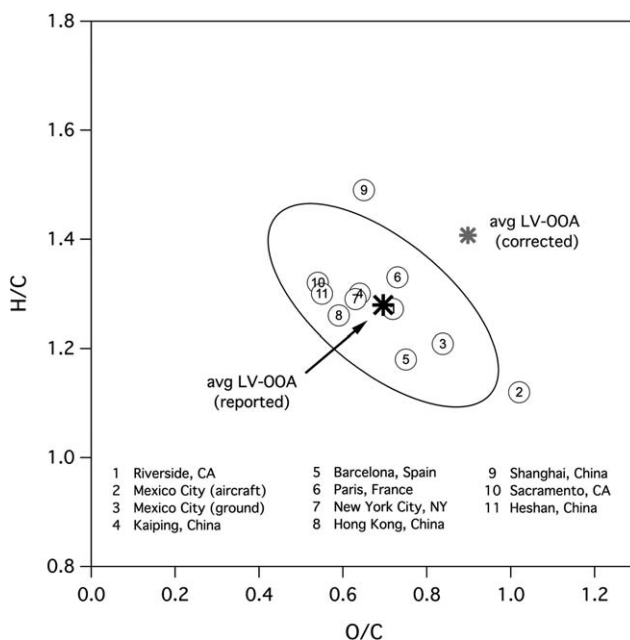
Fig. 2 Van Krevelen diagram showing the reported elemental ratios of the LV-OOA factors from HR-AMS field campaigns (see Table 1). The black star is the average reported O/C and H/C from all measurements, with the ellipse indicating the covariance of the data within 1.96 standard deviations of the uncorrected mean value. The grey star is the average elemental composition after applying the recommended correction of 1.3 and 1.1 to O/C and H/C, respectively.[53]

c) Oligomerization. Oligomerization reactions occur when two molecular species combine to form a larger one. Because of the associated increase in carbon number, oligomerization reactions are an efficient way to lower vapor pressure. However, this type of reaction is typically non-oxidative[44] so is not accompanied by an overall change in OS_C. Therefore the carbon atoms in the precursors must, on average, be as oxidized as those in the target.

As with fragmentation, oligomerization transforms are more difficult to describe than those for functionalization, since any number of changes to n_C, n_H, and n_O are possible. However, we can still apply basic constraints related to the allowable number of carbons and functional groups to narrow the range of potential precursors. The same statistical method described for fragmentation reactions can be used for oligomerization to constrain the precursor molecules. Here, we require the n_C of the precursor to be smaller than that of the target, and we assume that the precursor molecules have the same functional group distribution as the target, giving the following constraints: $n_{C,p} < n_{C,t}$, $n_{-OH,p} = n_{C,p}(n_{-OH,t}/n_{C,t})$, and $n_{=O,p} = n_{C,p}(n_{=O,t}/n_{C,t})$.

As with fragmentation, we can expand the range of precursors further by allowing the functional groups to be unevenly distributed between the two precursors. This allows one of the oligomerization reactants to be more functionalized that the other. While a large number of different oligomerization transformations are possible, we assume that for any oligomerization reaction each precursor must have at least one functional group. We also allow the total oxygen and hydrogen content to change somewhat by the gain or loss of a water molecule, or by the conversion of a carbonyl to an alcohol (as in a hemiacetal formation). These assumptions, plus the requirement that the precursor n_C be lower than the target n_C, yield the following simple rules for the number of oxygen atoms and carbonyl and hydroxyl groups: $n_{O,p} \leq n_{O,t}$, $n_{=O,p} \leq (n_{=O,t} + 1)$, and $n_{-OH,p} \leq (n_{-OH,t} + 1)$. It should be noted that if ester or ether linkages are formed during the oligomerization reaction, they will introduce an error in the c^* calculation. However, these effects are likely to be relatively small since the change in c^* is governed mostly by changes to the carbon number, rather than by interconversions of functional groups.

3 Results

I. LV-OOA target

Properties of the highly oxidized OA target are determined from HR-ToF-AMS field measurements. Table 1 summarizes AMS elemental ratios for studies that report H/C and O/C for the LV-OOA PMF factor (or, more generally, the most oxidized of multiple OOA factors, which includes factors such as OOA-I[32] or MO-OOA[51]). Fig. 2 shows the location of these factors in van Krevelen space. The degree of oxidation varies widely from study to study, indicating that LV-OOA is not one class of compounds but rather an aged form of SOA with a composition that varies as a function of location, emissions profile, and/or oxidant exposure. For example, LV-OOA from the MILAGRO field campaign in the Mexico City Metropolitan Area (MCMA)[32,35] is substantially more oxidized than that measured at the other field sites. Nonetheless, we use the average values of LV-OOA H/C and O/C (1.28 and 0.696, respectively) from these studies as our generic highly-oxidized OA target. This is shown as a black star in Fig. 2. The ellipse represents the covariance of the data within 1.96 standard deviations of the uncorrected mean value. AMS elemental ratios for ambient organic aerosol are typically corrected using empirical factors to account for biases in ion fragmentation.[32] Recent calibration studies indicate that the recommended factors result in an underestimation of O/C and H/C values for a range of multi-functional oxidized organic standards, indicating that AMS elemental ratios reported for ambient environments will
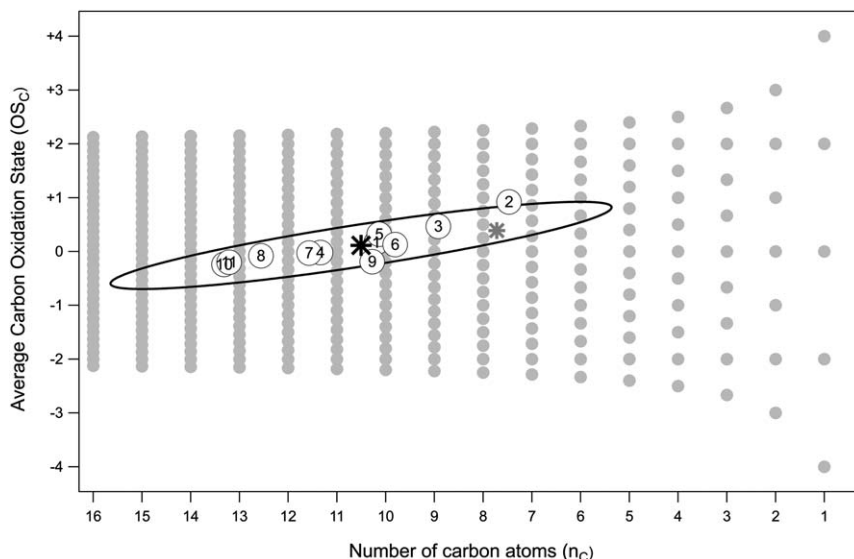
Paper
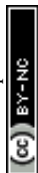
Fig. 3 Locations of the different LV-OOA factors in OS_C–n_C space. Grey circles indicate the possible combinations of average carbon oxidation state and carbon number.[4] Numbered points denote LV-OOA from all ten field campaigns (numbers are the same as in Fig. 2); OS_C is determined from elemental ratios, and n_C from eqn (5), using a c^* of 10^{-3} μg m^{-3}. The black star shows the average values for all the factors, with the ellipse depicting the uncertainty in its n_C and OS_C values (see text for details). The grey star shows the position of the average LV-OOA after applying the recommended correction of 1.3 and 1.1 to O/C and H/C, respectively.[53]

need to be modified by new correction factors (effectively 1.3 for O/C and 1.1 for H/C).[53] The average ratios calculated after using these new proposed correction factors are also shown in Fig. 2.

Volatility of the aerosol, as measured with thermodenuders, has to our knowledge been reported for only a limited number of field studies,[38] and values for c^* have been determined only for the MCMA ground-based data.[5] The decrease in the LV-OOA factor upon heating in the thermodenuder is consistent with LV-OOA volatilities spanning a range of c^* values between 10^{-7} and 10^{-1} μg m^{-3}.[5] However, since the LV-OOA measured in MCMA appears not to be representative of LV-OOA everywhere (Fig. 2), these volatility ranges may not be widely applicable to other locations. On the other hand, the elemental composition of Riverside LV-OOA is close to that of the overall LV-OOA average (Fig. 2), suggesting its volatility, which is substantially higher than that of LV-OOA in MCMA,[38] may be more representative of typical LV-OOA. We choose c^* in the range of 10^{-5} to 10^{-1} μg m^{-3} as reasonable values based on the reported c^* distribution.[5] From these c^* values, carbon numbers can be calculated (eqn (5)): for $\log_{10}(c^*)$ of -1, -3, and -5, the corresponding chemical formulas are $C_{9.2}H_{11.8}O_{6.4}$, $C_{10.5}H_{13.4}O_{7.3}$, and $C_{11.8}H_{15.1}O_{8.2}$, respectively, for the reported average, and $C_{6.8}H_{9.5}O_{6.1}$, $C_{7.7}H_{10.9}O_7$, and $C_{8.7}H_{12.2}O_{7.9}$, respectively, for the corrected average.[53] These formulas are not meant to define exact chemical species or specific molecules, but instead represent averages of all the compounds in LV-OOA.

The sources of uncertainty considered in evaluating the average number of carbon atoms in the LV-OOA mixture arise from estimated errors in the elemental
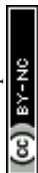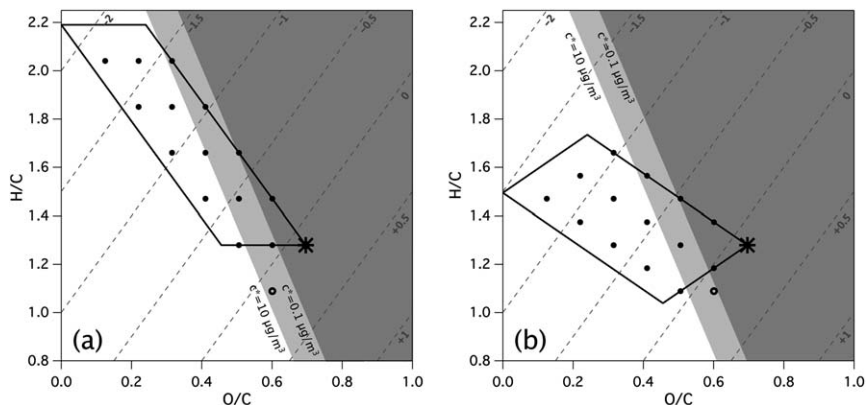
Fig. 4 Functionalization transforms available to the LV-OOA target (black star). These are calculated for $n_C = 10.5$ and $c^* = 10^{-3}$ μg m^{-3}, for the fully saturated case (panel a) and the fully unsaturated case (panel b). Potential oxidative transforms and precursors are bounded by the parallelograms, whose dimensions are defined by the number of hydroxyl and carbonyl functional groups in the target. Each black circle depicts an intermediate precursor, formed from the removal of an integer number of functional groups; the leftmost corner of each parallelogram (y-intercept) gives the limiting H/C value of the unfunctionalized hydrocarbon precursor. OS_C contours are shown as dashed grey lines, and shading indicates the predominant phase of the species: gas phase (white, $c^* > 10$ μg m^{-3}), semi-volatile (light grey, c^* 0.1 to 10 μg m^{-3}), or condensed-phase (dark grey, $c^* < 0.1$ μg m^{-3}). Multiple functional groups (including at least one hydroxyl group) must be lost for the immediate precursor of the target to be in the gas phase. Nonoxidative processes (e.g. hydrolysis) may also be important in SOA formation; the open circles indicate immediate hydrolysis precursors.

analysis calculation, the distribution of vapor pressures of the compounds within the mixture, and variation in the elemental ratios measured in the different field studies. The measurement error arising from elemental analysis is estimated to be ±30% of the reported O/C and ±10% for H/C.[32,33] Because these errors were reported for individual compounds, we expect the measurement error of a large mixture of compounds to be substantially smaller, so that the values given represent the most conservative case. The uncertainty in c^* for the average of all OOA components is taken to be $10^{-3\pm2}$ μg m^{-3}, as approximated from the distribution of values reported.[5] The variance in elemental ratios for the mixtures is simply the variance of the data presented in Table 1, multiplied by 1.96^2 in order to reflect the range in which 95% of the data is expected to fall. These errors are propagated forward using the sensitivity of n_C to O/C, H/C, and $\log_{10}(c^*)$ as described by eqn (5), to give three separate values of the uncertainty in the carbon number, δ_{NC}. The contributions from elemental analysis measurement error, variation in the volatility, and variance are 3.3, 1.3, and 3.8, respectively. An overall uncertainty, $\delta_{NC,tot}$, is calculated as the square root of sum of squares of each separate uncertainty, to yield a final estimate of 10.5 ± 5.2 for the LV-OOA carbon number, with the greatest source of uncertainty being the variance in elemental ratios reported for each field study discussed here.

Fig. 3 (OS_C vs. n_C) shows the location of the different LV-OOA factors in carbon oxidation state vs. carbon number space.[4] Also shown is the average LV-OOA target, with the ellipse depicting the uncertainty associated with its n_C and OS_C values. The rotated ellipse arises from the covariance between the outputs and reflects the fact that more highly oxidized compounds require a smaller carbon
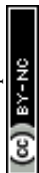
number for a given value of c^*. The carbon numbers determined here ($n_C \sim 6$–15) are slightly larger than those estimated by Donahue et al. ($n_C \sim 5$–10) using a similar technique and the same general c^* range.[54] This difference in part results from the lower value of OS_C that we use, and our direct determination of the abundance of hydroxyl and carbonyl functional groups individually, rather than assuming the two are equally abundant.

Functionalization. Taking the LV-OOA target determined using $c^* = 10^{-3}\ \mu g\ m^{-3}$ ($C_{10.5}H_{13.4}O_{7.3}$), and assuming it was formed via functionalization reactions only (with changes to H/C and O/C but not carbon number), allows us to work backwards toward precursors in van Krevelen space, using the simple rules described in the Methodology section. Fig. 4 shows the potential formation pathways and precursors, but is different than most other van Krevelen plots used to represent atmospheric OA since it includes gas-phase as well as particulate organics. The black star denotes the target, and the black circles depict the possible transforms associated with functionalization (as depicted in Fig. 1), with each point corresponding to the loss of an integer number of functional groups. Fig. 4a shows the fully saturated case in which functional groups replace C–H bonds only, and Fig. 4b shows the unsaturated case in which they replace only C=C bonds.

The parallelograms represent the range of possible transforms and precursors, and are bounded by the requirement that the target cannot lose more hydroxyl or carbonyl groups than it has (so that the dimensions of the parallelogram are given by the number of hydroxyl and carbonyl groups). The left corner of the parallelogram (y-intercept in the van Krevelen diagram) corresponds to the fully de-functionalized hydrocarbon precursor. For the saturated case, this gives a precursor with a formula consistent with an acyclic alkane. For the unsaturated case, it gives an alkene containing $n_O/2$ C=C bonds (or 3.7 for our target). This corresponds to a precursor of formula $C_{10.5}H_{15.7}$, which interestingly is similar to that of monoterpenes, a well-studied class of SOA precursors. These two limiting cases bound the H/C of the hydrocarbon precursor to between 1.5 and 2.2.

The areas defined by the parallelograms in Fig. 4, and movement within them, are not necessarily tied to specific chemistry, but rather describe the range of possible precursors. The dashed grey lines are carbon oxidation state contours ($OS_C \approx 2O/C - H/C$). From these, it can be seen that replacing C–H bonds with functional groups (Fig. 4a) moves more directly along the OS_C gradient than replacing C=C bonds (Fig. 4b) does, and results in a larger increase in OS_C per functional group. At the same time, the oxidation of C=C bonds usually involves the addition of two functional groups per oxidation step. It should be noted that the areas bounded by the parallelograms are for oxidation reactions only; non-oxidative transformations (involving movement along lines of constant OS_C) could lead to a wider range of possible precursors. One example is the dehydration of the target molecule (open circles in Fig. 4), whose corresponding forward reaction is the hydrolysis of a carbonyl or epoxide group to form two hydroxyl groups, believed to be an important reaction in the formation of isoprene SOA.[55]

The shading in Fig. 4 indicates the volatility (and therefore predominant phase) as a function of location in van Krevelen space: volatile/gas-phase (white, $c^* > 10\ \mu g\ m^{-3}$), semi-volatile (light grey, $0.1 < c^* < 10\ \mu g\ m^{-3}$), or low-volatility/particle-phase (dark grey, $c^* < 0.1\ \mu g\ m^{-3}$). This is similar to the work of Kessler et al. in which c^* cutoffs were calculated as a function of carbon number.[12,13] While the exact cutoffs will vary with organic aerosol loading, these are reasonable
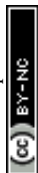
values for most ambient conditions. Values for c^* were calculated using eqn (4), though for Fig. 4b it was modified to allow for one DBE to be from a C=C bond:

$$\log_{10} c^* = \log_{10}\alpha + b_0 + b_C n_C + b_{=O}\left(n_C\left(1 - \frac{1}{2}H/C\right)\right) + b_{-OH}\left(n_C\left(O/C + \frac{1}{2}H/C - 1\right)\right) \tag{6}$$

While the timescales of formation of highly oxidized OA are highly uncertain, there is evidence (based on Mexico City data) that the process is relatively fast, on the order of a day.[3,6] This is faster than the rate of heterogeneous oxidation (which occurs on the order of several days),[12,13,56,57] suggesting that the immediate precursor to highly oxidized OA is likely not in the particle phase. The production of low-volatility products by gas-phase oxidation leads to a "trapping" in the particle phase where the rate of oxidation slows dramatically.[58,59] Thus the rapid formation of highly oxidized OA *via* functionalization reactions would seem to require gas-phase precursors. Semi-volatile compounds (light grey area in Fig. 4) may serve as gas-phase precursors, but their oxidation reactions tend to be slow, since they spend some fraction of their time in the particle phase; in addition they are highly susceptible to depositional losses to other environmental surfaces. Thus, in the absence of fast particle phase oxidation reactions, the most likely
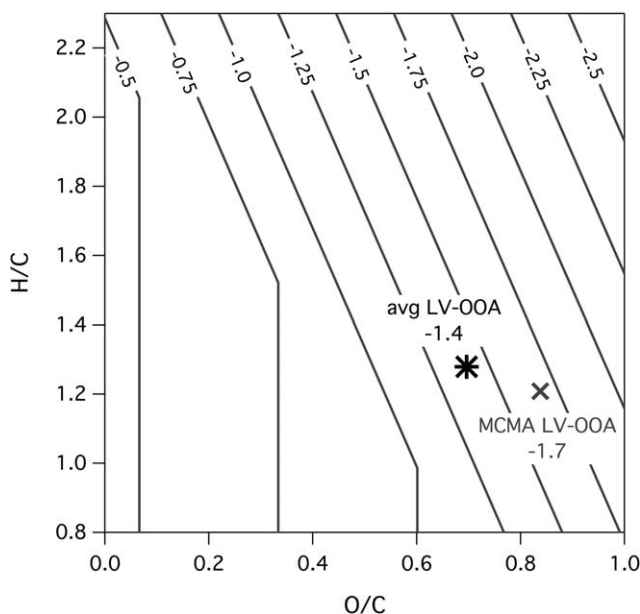

Fig. 5 Contour plot showing the influence of a single carbon atom in a molecule, including its associated functional groups, on the molecule's volatility ($\partial(\log_{10} c^*)/\partial n_c$), as calculated from eqn (4). These are calculated for $n_C = 10.5$, but do not vary significantly with carbon number above $n_C = 7$. Instead H/C and O/C have the strongest influence on "per-carbon volatility" which is greatest when both ratios are high (top right corner, when hydroxyl groups are abundant).[12,13] The discontinuities near the left of the figure reflect the fact that eqn (4) does not always hold at low O/C, where unsaturation arises from C=C double bonds or rings rather than carbonyl groups; thus when DBE > n_O, eqn (4) is modified to allow for C=C double bonds. LV-OOA targets used in this study are denoted by markers.
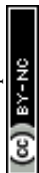
immediate gas-phase precursors to highly oxidized aerosol are those that are fully in the gas phase; as shown in Fig. 4, this requires the precursor to have substantially fewer functional groups than the target (at least two fewer hydroxyl groups or three fewer mixed hydroxyl/carbonyl groups). Known routes for the addition of multiple functional groups in a single oxidation step include the oxidation of carbon–carbon double bonds or oxidative ring-opening reactions, both of which add at least two (and possibly more) functional groups to the molecule. In addition, some alkoxy isomerization reactions can add three or more functional groups at a time;[60] plus there is evidence for other, poorly-understood mechanisms that involve the rapid addition of multiple functional groups.[61]

The above discussion assumes that the formation of low-volatility organics is rapid compared to the rate of heterogeneous oxidation, and also that gas-particle partitioning is essentially instantaneous. In general, the formation kinetics of the most oxidized fraction of OA are not well constrained; if they are substantially slower in most environments than in MCMA, heterogeneous oxidation could play an important role, allowing the immediate oxidative precursors to be in the condensed phase. This uncertainty underscores the need for additional ambient studies of OA at a variety of photochemical ages. Additionally, recent studies of partitioning kinetics suggest that condensable organics may remain in the gas phase longer than predicted by assumptions of thermodynamic equilibrium.[62] This could potentially allow for additional gas-phase oxidation even when the precursors have volatilities below the semi-volatile/condensed-phase cutoff. On the other hand, several experimental studies suggest a mass transfer limitation to
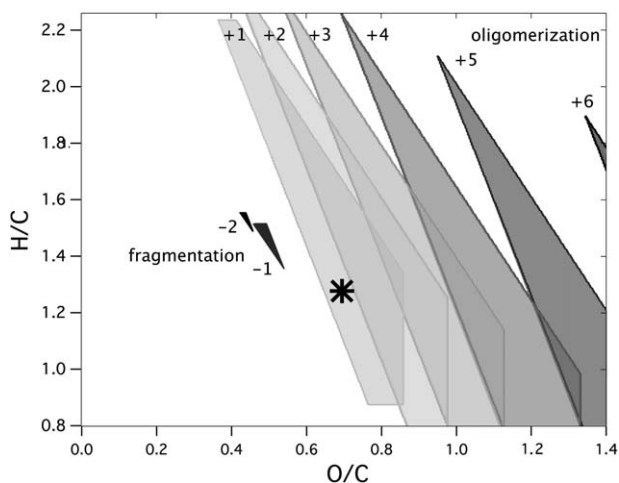

Fig. 6 Ranges of possible precursors of fragmentation (solid-shaded areas) and oligomerization (lightly-shaded areas) reactions that can form the LV-OOA target (black star). Three-dimensional chemical space is depicted by showing the range in van Krevelen space available to precursors of different carbon numbers. The numbers correspond to the carbon number change of the reaction; for example "−2" refers to a fragmentation reaction in which the target has two fewer carbon atoms than the precursor. The target can be formed via fragmentation of a gas-phase precursor only if a small (C_1–C_2) reduced fragment is lost; oligomerization is a more viable channel, though it requires the precursors to be low in volatility and thus extremely oxidized. In both cases, the range of possible precursors is considerably narrower if it is assumed that functional groups are evenly distributed on the target and precursor(s).
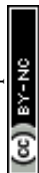
the evaporation of particulate organics;[63–65] this would have the opposite effect, enhancing the "trapping" effect and causing the semi-volatile precursors to behave more like the least-volatile species.

The potential formation pathways of highly oxidized aerosol also depend on our knowledge of the physicochemical properties (elemental ratios and c^*), which can be uncertain. While the volatility (c^*) chosen for our target has a relatively minor effect on calculated carbon number, it strongly influences the potential formation processes, most importantly by affecting the phase of the precursors. Decreasing the c^* of the target (for example, from 10^{-3} to 10^{-5} µg m^{-3}, increasing the calculated n_C from 10.5 to 11.8), increases the number of functional groups that must be removed in order for the immediate precursor to be in the gas phase. Therefore targets with low c^* are less likely to be formed from functionalization reactions, whereas targets with higher c^* can be easily formed from reactions that add only 1–2 functional groups.

The elemental ratios used for the target in Fig. 4 represent the average from a number of studies (Table 1), yielding an average chemical formula of $C_{10.5}H_{13.4}O_{7.3}$ (with 2.5 hydroxyl groups and 4.8 carbonyl groups). If we use the corrected values for the elemental ratios[53] (grey star in Fig. 2), but the same c^* as in the base case (10^{-3} µg m^{-3}), we obtain a target with a chemical formula of $C_{7.7}H_{10.9}O_7$ (3.7 hydroxyl groups and 3.3 carbonyl groups). The higher $n_{-OH}/n_{=O}$ ratio implies a lower carbon number at a given c^*, since hydroxyl groups have a greater effect on c^* than do carbonyl groups. While this has little effect on the number of functional groups separating the target and a gas-phase precursor (since the separation is governed primarily by the volatility of the target), it can affect the overall pathways connecting the oxidized aerosol to hydrocarbon precursors; additional oxidation steps will be required to form a target with a higher number of functional groups.

Fragmentation reactions. In addition to the chemical constraints listed in the Methodology section, the phase of the fragmentation precursor also limits the possible transforms associated with this reaction. As in the functionalization case, we require the precursor to be in the gas phase ($c^* > 0.1$ µg m^{-3}) since heterogeneous oxidation is assumed to be too slow to account for the rate of formation of highly oxidized aerosol.

The overall, average effect of fragmentation reactions on volatility is illustrated in Fig. 5, which shows the contribution of a single carbon atom (and its associated functional groups) to the volatility of a molecule of $n_C = 10.5$. (Results are only weakly dependent on carbon number above $n_C = 7$.) This is calculated from eqn (4), assuming that the functional groups are evenly distributed throughout the molecule (i.e., the number of groups on the carbon atom is equal to the number of groups on the entire molecule divided by n_C). For our target (black star in Fig. 5), each carbon atom decreases the vapor pressure by 1.4 orders of magnitude. This is a full order of magnitude larger than the effect of a non-functionalized carbon (e.g., a CH_2 group). This simple view of "per-carbon volatility" helps guide assessments of the viability of fragmentation reactions as steps in the formation of highly oxidized organic aerosol; however, as discussed in the Methodology section, it does not accurately represent all fragmentation reactions, since functional groups generally are not evenly distributed throughout a given molecule. Here we consider potential fragmentation precursors both with and without evenly-distributed functional groups.
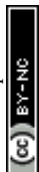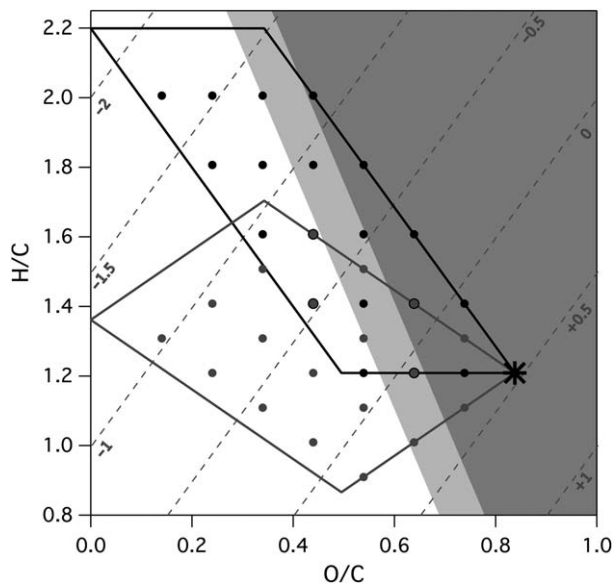
Fig. 7 Functionalization transforms available to the LV-OOA target from the MILAGRO ground site,[31,37] using $n_C = 10$ and $c* = 10^{-5}$ µg m^{-3}. (These are different than the values in Fig. 3–4 which used $c*$ of 10^{-3} µg m^{-3}.) This plot is similar to Fig. 4, except that the saturated and unsaturated cases are shown together, as black and grey parallelograms, respectively (the $c*$ cutoffs are given for the saturated case only). Because of the exceedingly low volatility of these organics,[5,37] their formation from gas-phase precursors can occur only if a large number of functional groups (three or more) are added to the precursor in a single oxidation step.

Assuming the functional groups are distributed evenly, fragmentation of a gas-phase molecule cannot form highly oxidized, low-volatility OA. This is because the decrease in volatility from the additional functional groups (in this case, one hydroxyl group and one carbonyl group, lowering $c*$ by 3.2 orders of magnitude) is largely if not completely offset by the volatility increase from the carbon loss (1.4 orders of magnitude per carbon atom). As a result, the large decrease in volatility needed to form a low-volatility target (2 orders of magnitude) is not possible, a conclusion broadly consistent with the findings of Cappa and Wilson.[18] However, when functional groups are not evenly distributed, and reduced carbon can be lost during fragmentation, the formation of the target *via* fragmentation is possible in some cases. The carbon numbers and elemental ratios of viable precursors are depicted in Fig. 6 (solid-shaded areas). The allowed precursors are only slightly larger than the target, and are significantly less oxidized, indicating that fragmentation can produce highly oxidized organic aerosol only when the precursor loses a small (C_1 or C_2) and essentially unfunctionalized fragment. (This exact requirement depends on the identity and number of functional groups added to the fragment). However, molecules that are more volatile than our target are more likely to form *via* fragmentation, since they may be more easily formed from gas-phase precursors. Similarly, less-oxidized species can also be formed from fragmentation reactions, since the increase in volatility from losing carbon atoms is not as great (top left corner of Fig. 5). Thus, while fragmentation reactions probably do not lead to the immediate formation of highly oxidized OA, they
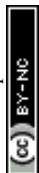
might occur as part of the overall reaction mechanism, primarily in the earlier oxidation steps.

Oligomerization reactions. Because oligomerization reactions decrease volatility, they are generally more viable steps in the formation of oxidized OA components than fragmentation reactions. As with fragmentation reactions, the phase of the precursor introduces constraints beyond those described in the Methodology section: since the probability of two gas-phase precursors reacting is quite low under most conditions, at least one precursor must be present in the particle phase ($c^* < 10$ µg m^{-3}).

Since oligomerization reactions involve changes to n_C, their effects can be visualized similarly to the effects of fragmentation reactions, by treating all carbon as evenly functionalized (Fig. 5). Assuming that functional groups are evenly-distributed in both the precursors and the targets, the condensed-phase precursor can be smaller than the target by one or two carbon atoms. Precursors with even fewer carbon atoms will be so volatile that they will be present in the gas phase only, and therefore cannot participate in oligomerization reactions. However, a wider range of precursors is possible if the functional groups are not evenly distributed, and instead the two precursor molecules have different levels of functionalization. The range of possible oligomerization precursors are shown as lightly-shaded regions in Fig. 6; lower carbon numbers are possible when the majority of the functional groups are on the condensed-phase precursor. Thus oxidized organic aerosol could form from the oligomerization of smaller, highly-oxidized molecular species, which themselves may be formed by previous functionalization or fragmentation reactions. This is broadly consistent with the view of formation of SOA from aqueous chemistry;[15] even though we do not explicitly model partitioning into the aqueous phase, this work provides some constraints into the chemical nature (H/C and O/C ratios) of these smaller species. On the other hand, the target itself (of formula $C_{10.5}H_{13.4}O_{7.3}$) is not consistent with the oligomeric species measured in SOA, such as those from monoterpene oxidation (which typically have twenty or more carbon atoms and relatively low OS_C values[66]), isoprene photooxidation, or glyoxal oligomerization (which typically have higher H/C values[11]). Thus, even though these oligomerization reactions can occur under atmospheric conditions, they do not appear to dominate formation of the most oxidized fraction of the aerosol.

MCMA case study. Although LV-OOA in Mexico City is significantly more oxidized than that measured in other locations (Table 1 and Fig. 2), it has received substantial study and to our knowledge is the only LV-OOA factor whose volatility and elemental ratios have both been reported. Here, we examine LV-OOA at the T0 ground site at MILAGRO[32,38] as a case study, in order to gain insight into potential formation pathways for this exceedingly highly oxidized organic aerosol. Using a c^* of $10^{-5 \pm 2}$ µg m^{-3} gives a target carbon number estimate of 10.0 ± 3.4 with the chemical formula $C_{10}H_{12.1}O_{8.4}$, (3.4 hydroxyl groups and 5 carbonyl groups). The sources of uncertainty studied in this case for the carbon number are the error in computing elemental ratios and the deviations in c^*, which give values of δ_{NC} of 3.2 and 1.1, respectively. As with the average LV-OOA, the overall error is computed from the sum of the square of the errors. The potential functionalization precursors are shown in Fig. 7 for the saturated and unsaturated cases. Because the target in this case is so low in volatility, the functionalization transforms involve the loss of at least two hydroxyl groups, or all five carbonyl
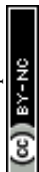
groups, to form precursors that are semi-volatile, and even more for the precursor to be present exclusively in the gas phase. This suggests that gas-phase functionalization reactions would have to be exceedingly rapid to lead to the formation of organic material that is so high in oxidation state and low in volatility. Similarly, this class of organics cannot be formed directly from fragmentation reactions (even if the loss of small, reduced fragments are allowed). Again, some fragmentation reactions might occur earlier in the overall oxidation mechanism, but they do not play a role in the last step that forms this target. As in our average LV-OOA target (Fig. 6), oligomerization could be a viable step in the formation of these compounds, provided highly oxidized organic precursor species are present in the condensed phase.

4   Conclusions

We have presented a new general approach for assessing the potential pathways and precursors involved in the formation of highly oxidized organic aerosol, which involves starting from the product or "target" and working backwards. We define this target based on measurements of volatility and average elemental ratios of the AMS LV-OOA factor, yielding an average approximate chemical formula of $C_{10.5}H_{13.4}O_{7.3}$. The utility of our retrosynthetic method for constraining its formation chemistry has been demonstrated for this average LV-OOA target, as well as an exceedingly oxidized case, LV-OOA measured at the MILAGRO ground site. Our results suggest that the most probable LV-OOA-forming reactions are functionalization reactions that add multiple functional groups (including at least one hydroxyl group) at a time, as well as oligomerization reactions of smaller, highly oxidized precursors. Fragmentation reactions are only possible as the final step if they involve the loss of a small number of relatively reduced carbon atoms. While the specific aim of this study was to apply this technique to LV-OOA, this approach could be useful for understanding a wider range of targets, including other fractions of the aerosol (e.g., SV-OOA[3]), as well as SOA generated in laboratory studies. However, this requires some information about the volatility and/or carbon number of the aerosol of interest.

In the present work, we limited the partitioning of the organic species to only two phases, the gas phase and the condensed-organic phase. Recent work has suggested the importance of the liquid-water phase as a medium for oxidation reactions,[15] which could also be included in this treatment. This would require the determination of the volatilities of organics over water (Henry's Law constant) from chemical properties (elemental ratios, functional group distributions, etc.). Otherwise the general approach, including the transforms that describe the chemical reactions, is the same as described in this work.

This retrosynthetic approach relies critically on the accurate characterization of the target molecule and key characteristics of its formation. This present work highlights the need for accurate measurements of the elemental ratios and volatility of organic aerosol. Field measurements of $c*$ are limited, hindering accurate estimates of carbon number. Better constraints on $c*$ of the oxidized aerosol could significantly improve the accuracy of our target carbon number. There is also a need for better constraints on the kinetics of the formation of highly oxidized aerosol. This has been measured only in the outflow from MCMA, which might not be widely representative. This rate is a key constraint on the
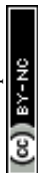
formation pathway, as it limits the number of oxidation steps that are possible as well as the phase in which those reactions can happen. Additionally, characterization of highly-oxidized OA using techniques other than AMS would be useful. The direct measurement of functional group distributions (*e.g.*, using FTIR[41]), can provide an alternative, and potentially more comprehensive, description of the relationships between chemical structure, volatility, and formation chemistry. Similarly, direct measurements of n_C (*e.g.*, using UHR-ESI-MS) would improve the accuracy of the target chemical formula and could provide more detailed information on the formation pathways of a wide range of molecular compounds.

We have made an effort to demonstrate the utility of this retrosynthetic approach for exploring the possible pathways by which highly oxidized aerosol is formed. By starting at the known endpoint/target, we improve our ability to identify pathways that can actually result in the formation of low-volatility, highly-oxidized organic compounds. This is in contrast to standard forward approaches that require some guesswork as to the starting molecules and pathways that may result in the formation of the target. This general approach could be particularly powerful when applied to rigorous models of oxidation mechanisms, gas-particle partitioning, and aerosol chemistry.[18,22,54,67] Running such models backwards from well-defined targets may help identify key classes (or even structures) of organic compounds, as well as reaction conditions, that are most likely to lead to the formation of this important class of organic aerosol.

Acknowledgements

References

1 Q. Zhang, J. L. Jimenez, M. R. Canagaratna, J. D. Allan, H. Coe, I. Ulbrich, M. R. Alfarra, A. Takami, A. M. Middlebrook, Y. L. Sun, K. Dzepina, E. Dunlea, K. Docherty, P. F. DeCarlo, D. Salcedo, T. Onasch, J. T. Jayne, T. Miyoshi, A. Shimono, S. Hatakeyama, N. Takegawa, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, P. Williams, K. Bower, R. Bahreini, L. Cottrell, R. J. Griffin, J. Rautiainen, J. Y. Sun, Y. M. Zhang and D. R. Worsnop, *Geophys. Res. Lett.*, 2007, **34**, L13801.
2 E. R. Graber and Y. Rudich, *Atmos. Chem. Phys.*, 2006, **6**, 729–753.
3 J. L. Jimenez, M. R. Canagaratna, N. M. Donahue, A. S. H. Prevot, Q. Zhang, J. H. Kroll, P. F. DeCarlo, J. D. Allan, H. Coe, N. L. Ng, A. C. Aiken, K. S. Docherty, I. M. Ulbrich, A. P. Grieshop, A. L. Robinson, J. Duplissy, J. D. Smith, K. R. Wilson, V. A. Lanz, C. Hueglin, Y. L. Sun, J. Tian, A. Laaksonen, T. Raatikainen, J. Rautiainen, P. Vaattovaara, M. Ehn, M. Kulmala, J. M. Tomlinson, D. R. Collins, M. J. Cubison, J. Dunlea, J. A. Huffman, T. B. Onasch, M. R. Alfarra, P. I. Williams, K. Bower, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, D. Salcedo, L. Cottrell, R. Griffin, A. Takami, T. Miyoshi, S. Hatakeyama, A. Shimono, J. Y. Sun, Y. M. Zhang, K. Dzepina, J. R. Kimmel, D. Sueper, J. T. Jayne, S. C. Herndon, A. M. Trimborn, L. R. Williams, E. C. Wood, A. M. Middlebrook, C. E. Kolb, U. Baltensperger and D. R. Worsnop, *Science*, 2009, **326**, 1525–1529.
4 J. H. Kroll, N. M. Donahue, J. L. Jimenez, S. H. Kessler, M. R. Canagaratna, K. R. Wilson, K. E. Altieri, L. R. Mazzoleni, A. S. Wozniak, H. Bluhm, E. R. Mysak, J. D. Smith, C. E. Kolb and D. R. Worsnop, *Nat. Chem.*, 2011, **3**, 133–139.
5 C. D. Cappa and J. L. Jimenez, *Atmos. Chem. Phys.*, 2010, **10**, 5409–5424.
6 R. Volkamer, J. L. Jimenez, F. San Martini, K. Dzepina, Q. Zhang, D. Salcedo, L. T. Molina, D. R. Worsnop and M. J. Molina, *Geophys. Res. Lett.*, 2006, **33**, L17811.

7 C. L. Heald, J. H. Kroll, J. L. Jimenez, K. S. Docherty, P. F. DeCarlo, A. C. Aiken, Q. Chen, S. T. Martin, D. K. Farmer and P. Artaxo, *Geophys Res. Lett.*, 2010, **37**, L08803.

8 N. M. Donahue, K. M. Henry, T. F. Mentel, A. Kiendler-Scharr, C. Spindler, B. Bohn, T. Brauers, H. P. Dorn, H. Fuchs, R. Tillmann, A. Wahner, H. Saathoff, K. H. Naumann, O. Mohler, T. Leisner, L. Muller, M. C. Reinnig, T. Hoffmann, K. Salo, M. Hallquist, M. Frosch, M. Bilde, T. Tritscher, P. Barmet, A. P. Praplan, P. F. DeCarlo, J. Dommen, A. S. H. Prevot and U. Baltensperger, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 13503–13508.

9 L. Qi, S. Nakao, Q. Malloy, B. Warren and D. R. Cocker III, *Atmos. Environ.*, 2010, **44**, 2990–2996.

10 R. Bahreini, M. D. Keywood, N. L. Ng, V. Varutbangkul, S. Gao, R. C. Flagan, J. H. Seinfeld, D. R. Worsnop and J. L. Jimenez, *Environ. Sci. Technol.*, 2005, **39**, 5674–5688.

11 P. S. Chhabra, N. L. Ng, M. R. Canagaratna, A. L. Corrigan, L. M. Russell, D. R. Worsnop, R. C. Flagan and J. H. Seinfeld, *Atmos. Chem. Phys.*, 2011, **11**, 8827–8845.

12 S. H. Kessler, J. D. Smith, D. L. Che, D. R. Worsnop, K. R. Wilson and J. H. Kroll, *Environ. Sci. Technol.*, 2010, **44**, 7005–7010.

13 S. H. Kessler, T. Nah, K. E. Daumit, J. D. Smith, S. R. Leone, C. E. Kolb, D. R. Worsnop, K. R. Wilson and J. H. Kroll, *J. Phys. Chem. A*, 2012, **116**, 6358–6365.

14 A. T. Lambe, T. B. Onasch, D. R. Croasdale, J. P. Wright, A. T. Martin, J. P. Franklin, P. Massoli, J. H. Kroll, M. R. Canagaratna, W. H. Brune, D. R. Worsnop and P. Davidovits, *Environ. Sci. Technol.*, 2012, **46**, 5430–5437.

15 B. Ervens, B. J. Turpin and R. J. Weber, *Atmos. Chem. Phys.*, 2011, **11**, 11069–11102.

16 E. Gómez Alvarez, H. Wortham, R. Strekowski, C. Zetzsch and S. Gligorovski, *Environ. Sci. Technol.*, 2012, **46**, 1955–1963.

17 K. Dzepina, C. D. Cappa, R. M. Volkamer, S. Madronich, P. F. DeCarlo, R. A. Zaveri and J. L. Jimenez, *Environ. Sci. Technol.*, 2011, **45**, 3496–3503.

18 C. D. Cappa and K. R. Wilson, *Atmos. Chem. Phys.*, 2012, **12**, 9505–9528.

19 N. M. Donahue, S. A. Epstein, S. N. Pandis and A. L. Robinson, *Atmos. Chem. Phys.*, 2011, **11**, 3303–3318.

20 K. Dzepina, R. M. Volkamer, S. Madronich, P. Tulet, I. M. Ulbrich, Q. Zhang, C. D. Cappa, P. J. Ziemann and J. L. Jimenez, *Atmos. Chem. Phys.*, 2009, **9**, 5681–5709.

21 A. Hodzic, J. L. Jimenez, S. Madronich, M. R. Canagaratna, P. F. DeCarlo, L. Kleinman and J. Fast, *Atmos. Chem. Phys.*, 2010, **10**, 5491–5514.

22 B. Aumont, R. Valorso, C. Mouchel-Vallon, M. Camredon, J. Lee-Taylor and S. Madronich, *Atmos. Chem. Phys.*, 2012, **12**, 7577–7589.

23 J. Lee-Taylor, S. Madronich, B. Aumont, A. Baker, M. Camredon, A. Hodzic, G. S. Tyndall, E. Apel and R. A. Zaveri, *Atmos. Chem. Phys.*, 2011, **11**, 13219–13241.

24 B. K. Pun, C. Seigneur, D. Grosjean and P. Saxena, *J. Atmos. Chem.*, 2000, **35**, 199–223.

25 N. L. Ng, M. R. Canagaratna, J. L. Jimenez, P. S. Chhabra, J. H. Seinfeld and D. R. Worsnop, *Atmos. Chem. Phys.*, 2011, **11**, 6465–6474.

26 Y. Wei, T. Cao and J. E. Thompson, *Atmos. Environ.*, 2012, **62**, 199–207.

27 J. F. Pankow and K. C. Barsanti, *Atmos. Environ.*, 2009, **43**, 2829–2835.

28 N. L. Ng, M. R. Canagaratna, Q. Zhang, J. L. Jimenez, J. Tian, I. M. Ulbrich, J. H. Kroll, K. S. Docherty, P. S. Chhabra, R. Bahreini, S. M. Murphy, J. H. Seinfeld, L. Hildebrandt, N. M. Donahue, P. F. DeCarlo, V. A. Lanz, A. S. H. Prévôt, E. Dinar, Y. Rudich and D. R. Worsnop, *Atmos. Chem. Phys.*, 2010, **10**, 4625–4641.

29 E. J. Corey and X.-M. Cheng, *The Logic of Chemical Synthesis*, Wiley, New York, 1995.

30 P. F. DeCarlo, J. R. Kimmel, A. Trimborn, M. J. Northway, J. T. Jayne, A. C. Aiken, M. Gonin, K. Fuhrer, T. Horvath, K. S. Docherty, D. R. Worsnop and J. L. Jimenez, *Anal. Chem.*, 2006, **78**, 8281–8289.

31 M. R. Canagaratna, J. T. Jayne, J. L. Jimenez, J. D. Allan, M. R. Alfarra, Q. Zhang, T. B. Onasch, F. Drewnick, H. Coe, A. Middlebrook, A. Delia, L. R. Williams, A. M. Trimborn, M. J. Northway, P. F. DeCarlo, C. E. Kolb, P. Davidovits and D. R. Worsnop, *Mass Spectrom. Rev.*, 2007, **26**, 185–222.

32 A. C. Aiken, P. F. DeCarlo, J. H. Kroll, D. R. Worsnop, J. A. Huffman, K. S. Docherty, I. M. Ulbrich, C. Mohr, J. R. Kimmel, D. Sueper, Y. Sun, Q. Zhang, A. Trimborn, M. Northway, P. J. Ziemann, M. R. Canagaratna, T. B. Onasch, M. R. Alfarra, A. S. H. Prevot, J. Dommen, J. Duplissy, A. Metzger, U. Baltensperger and J. L. Jimenez, *Environ. Sci. Technol.*, 2008, **42**, 4478–4485.

33 A. C. Aiken, P. F. DeCarlo and J. L. Jimenez, *Anal. Chem.*, 2007, **79**, 8350–8358.

34 I. M. Ulbrich, M. R. Canagaratna, Q. Zhang, D. R. Worsnop and J. L. Jimenez, *Atmos. Chem. Phys.*, 2009, **9**, 2891–2918.

35 P. F. DeCarlo, I. M. Ulbrich, J. Crounse, B. de Foy, E. J. Dunlea, A. C. Aiken, D. Knapp, A. J. Weinheimer, T. Campos, P. O. Wennberg and J. L. Jimenez, *Atmos. Chem. Phys.*, 2010, **10**, 5257–5280.
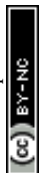
36 C. D. Cappa, *Atmos. Meas. Tech.*, 2010, **3**, 579–592.
37 X. F. Huang, L. Y. He, M. Hu, M. R. Canagaratna, J. H. Kroll, N. L. Ng, Y.-H. Zhang, Y. Lin, L. Xue, T. L. Sun, X. G. Liu, M. Shao, J. T. Jayne and D. R. Worsnop, *Atmos. Chem. Phys.*, 2011, **11**, 1865–1877.
38 J. A. Huffman, K. S. Docherty, A. C. Aiken, M. J. Cubison, I. M. Ulbrich, P. F. DeCarlo, D. Sueper, J. T. Jayne, D. R. Worsnop, P. J. Ziemann and J. L. Jimenez, *Atmos. Chem. Phys.*, 2009, **9**, 7161–7182.
39 J. F. Pankow and W. E. Asher, *Atmos. Chem. Phys.*, 2008, **8**, 2773–2796.
40 D. K. Farmer, A. Matsunaga, K. S. Docherty, J. D. Surratt, J. H. Seinfeld, P. J. Ziemann and J. L. Jimenez, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 6670–6675.
41 S. F. Maria, L. M. Russell, B. J. Turpin and R. J. Porcja, *Atmos. Environ.*, 2002, **36**, 5185–5196.
42 P. Lin, A. G. Rincon, M. Kalberer and J. Z. Yu, *Environ. Sci. Technol.*, 2012, **46**, 7454–7462.
43 A. S. Wozniak, J. E. Bauer, R. L. Sleighter, R. M. Dickhut and P. G. Hatcher, *Atmos. Chem. Phys.*, 2008, **8**, 5099–5111.
44 J. H. Kroll and J. H. Seinfeld, *Atmos. Environ.*, 2008, **42**, 3593–3624.
45 K. S. Docherty, A. C. Aiken, J. A. Huffman, I. M. Ulbrich, P. F. DeCarlo, D. Sueper, D. R. Worsnop, D. C. Snyder, R. E. Peltier, R. J. Weber, B. D. Grover, D. J. Eatough, B. J. Williams, A. H. Goldstein, P. J. Ziemann and J. L. Jimenez, *Atmos. Chem. Phys.*, 2011, **11**, 12387–12420.
46 C. Mohr, P. F. DeCarlo, M. F. Heringa, R. Chirico, J. G. Slowick, R. Richter, C. Reche, A. Alastuey, X. Querol, R. Seco, J. Penuelas, J. L. Jimenez, M. Crippa, R. Zimmermann, U. Baltensperger and A. S. H. Prevot, *Atmos. Chem. Phys.*, 2012, **12**, 1649–1665.
47 M. Crippa, I. El Haddad, J. G. Slowik, P. F. DeCarlo, C. Mohr, M. R. Heringa, R. Chirico, N. Marchand, J. Sciare, U. Baltensperger and A. S. H. Prevot, *J. Geophys. Res.: Atmos.*, 2013, **118**, 1950–1963.
48 Y.-L. Sun, Q. Zhang, J. J. Schwab, K. L. Demerjian, W.-N. Chen, M.-S. Bae, H.-M. Hung, O. Hogrefe, B. Frank, O. V. Rattigan and Y.-C. Lin, *Atmos. Chem. Phys.*, 2011, **11**, 1581–1602.
49 L.-Y. He, X.-F. Huang, L. Xue, M. Hu, Y. Lin, J. Zheng, R. Zhang and Y.-H. Zhang, *J. Geophys. Res.*, 2011, **116**, D12304.
50 X.-F. Huang, L.-Y. He, L. Xue, T.-L. Sun, L.-W. Zeng, Z.-H. Gong, M. Hu and T. Zhu, *Atmos. Chem. Phys.*, 2012, **12**, 4897–4907.
51 A. Setyan, Q. Zhang, M. Merkel, W. B. Knighton, Y. Sun, C. Song, J. E. Shilling, T. B. Onasch, S. C. Herndon, D. R. Worsnop, J. D. Fast, R. A. Zaveri, L. K. Berg, A. Wiedensohler, B. A. Flowers, M. K. Dubey and R. Subramanian, *Atmos. Chem. Phys.*, 2012, **12**, 8131–8156.
52 Z. Gong, Z. Lan, L. Xue, L. Zeng, L. He and X. Huang, *Front. Environ. Sci. Eng.*, 2012, **6**, 725–733.
53 M. R. Canagaratna, P. Massoli, J. L. Jimenez, S. Kessler, Q. Chen, L. Hildebrandt, E. Fortner, L. Williams, K. Wilson, J. Surratt, N. Donahue, J. Kroll, J. T. Jayne and D. R. Worsnop, *Improved calibration of O/C and H/C Ratios obtained by Aerosol Mass Spectrometry of Organic Species*, manuscript in preparation.
54 N. M. Donahue, J. H. Kroll, S. N. Pandis and A. L. Robinson, *Atmos. Chem. Phys.*, 2012, **12**, 615–634.
55 F. Paulot, J. D. Crounse, H. G. Kjaergaard, A. Kurten, J. M. St. Clair, J. H. Seinfeld and P. O. Wennberg, *Science*, 2009, **325**, 730–733.
56 I. J. George and J. P. D. Abbatt, *Nat. Chem.*, 2010, **2**, 713–722.
57 A. L. Robinson, N. M. Donahue and W. F. Rogge, *J. Geophys. Res.*, 2006, **111**, D03302.
58 J. D. Hearn and G. D. Smith, *Int. J. Mass Spectrom.*, 2006, **258**, 95–103.
59 E. A. Weitkamp, K. E. H. Hartz, A. M. Sage, N. M. Donahue and A. L. Robinson, *Environ. Sci. Technol.*, 2008, **42**, 5177–5182.
60 P. J. Ziemann and R. Atkinson, *Chem. Soc. Rev.*, 2012, **41**, 6582–6605.
61 M. Ehn, E. Kleist, H. Junninen, T. Petäjä, G. Lönn, S. Schobesberger, M. Dal Maso, A. Trimborn, M. Kulmala, D. R. Worsnop, A. Wahner, J. Wildt and T. F. Mentel, *Atmos. Chem. Phys.*, 2012, **12**, 5113–5127.
62 M. Shiraiwa and J. H. Seinfeld, *Geophys. Res. Lett.*, 2012, **39**, L24801.
63 T. D. Vaden, D. Imre, J. Beránek, M. Shrivastava and A. Zelenyuk, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 2190–2195.
64 V. Perraud, E. A. Bruns, M. J. Ezell, S. N. Johnson, Y. Yu, M. L. Alexander, A. Zelenyuk, D. Imre, W. L. Chang, D. Dabdub, J. F. Pankow and B. J. Finlayson-Pitts, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 2836–2841.
65 A. P. Grieshop, N. M. Donahue and A. L. Robinson, *Geophys. Res. Lett.*, 2007, **34**, L14810.
66 M. P. Tolocka, M. Jang, J. M. Ginter, F. J. Cox, R. M. Kamens and M. V. Johnston, *Environ. Sci.Technol.*, 2004, **38**, 1428–1434.
67 X. Zhang and J. H. Seinfeld, *Atmos. Chem. Phys.*, 2013, **13**, 5907–5926.