Environmental Science & Technology

Article

pubs.acs.org/est

Atmospheric Oxidation of Squalene: Molecular Study Using COBRA Modeling and High-Resolution Mass Spectrometry

David R. Fooshee,[†] Paige K. Aiona,[‡] Alexander Laskin,[§] Julia Laskin,[∥] Sergey A. Nizkorodov,[*,‡] and Pierre F. Baldi[*,†]

[†]School of Information and Computer Sciences, University of California, Irvine, California 92697, United States

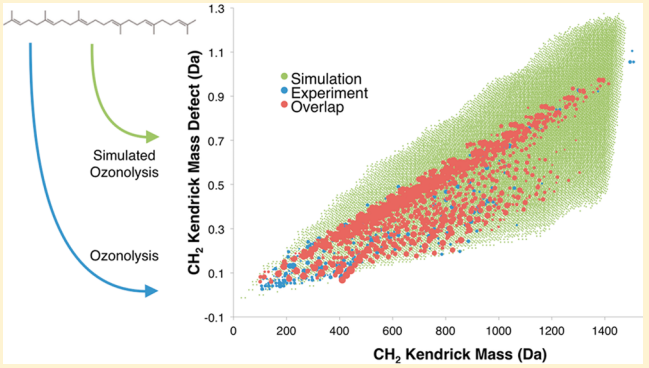
[‡]Department of Chemistry, University of California, Irvine, California 92697, United States

[§]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352, United States

[∥]Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States

Supporting Information

ABSTRACT: Squalene is a major component of skin and plant surface lipids and is known to be present at high concentrations in indoor dust. Its high reactivity toward ozone makes it an important ozone sink and a natural protectant against atmospheric oxidizing agents. While the volatile products of squalene ozonolysis are known, the condensed-phase products have not been characterized. We present an analysis of condensed-phase products resulting from an extensive oxidation of squalene by ozone probed by electrospray ionization (ESI) high-resolution mass spectrometry (HR−MS). A complex distribution of nearly 1300 peaks assignable to molecular formulas is observed in direct infusion positive ion mode ESI mass spectra. The distribution of peaks in the mass spectra suggests that there are extensive cross-coupling reactions between hydroxy-carbonyl products of squalene ozonolysis. To get additional insights into the mechanism, we apply a Computational Brewing Application (COBRA) to simulate the oxidation of squalene in the presence of ozone, and compare predicted results with those observed by the HR−MS experiments. The system predicts over one billion molecular structures between 0 and 1450 Da, which correspond to about 27 000 distinct elemental formulas. Over 83% of the squalene oxidation products inferred from the mass spectrometry data are matched by the simulation. The simulation indicates a prevalence of peroxy groups, with hydroxyl and ether groups being the second-most important O-containing functional groups formed during squalene oxidation. These highly oxidized products of squalene ozonolysis may accumulate on indoor dust and surfaces and contribute to their redox capacity.



1. INTRODUCTION

Squalene (see Scheme 1) is a naturally occurring product found in human and plant lipids.[1] This unsaturated, nonvolatile triterpene reacts readily with ozone, and is the most abundant ozone-reactive constituent of human sebum.[2,3] Squalene accounts for 10−12% of adult skin lipids, making it one of the major lipids found on the surface of the human skin.[4] This contributes to squalene's presence in human skin flakes, which are a major component of indoor dust. Lipids in skin and plant oils, such as squalene, are a natural form of protection against oxidizing agents found in air, with squalene shown to account for about 40% of ozone removal by human skin and hair.[4,5]

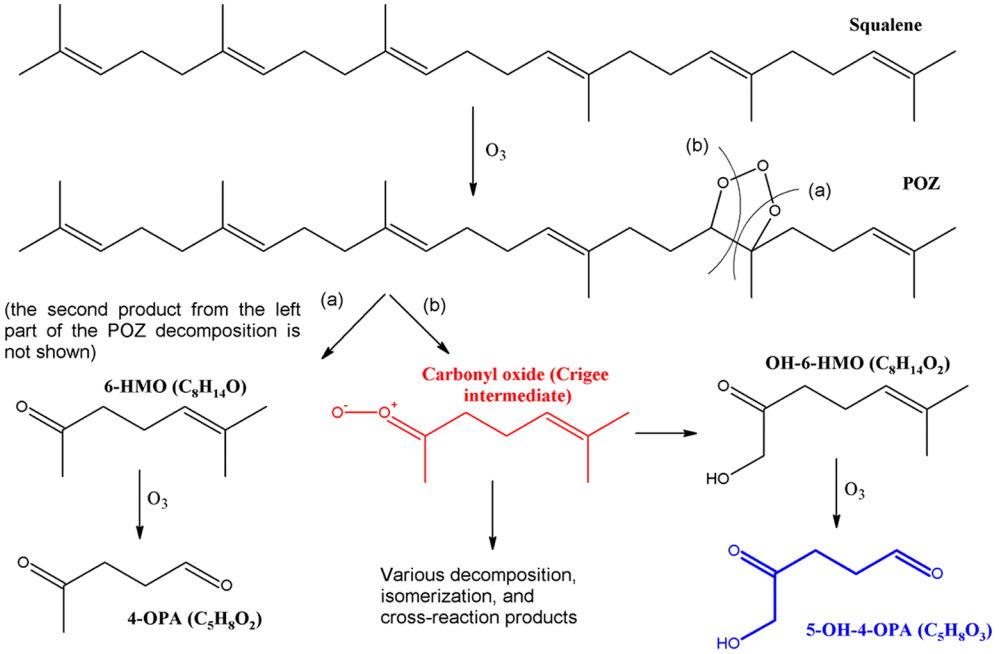
Squalene is highly reactive toward ozone due to its six unconjugated carbon−carbon double bonds.[1,6] It reacts with ozone based on the Criegee mechanism, whereby ozone attacks a double bond to form a primary ozonide (POZ), which then decomposes into a carbonyl and a carbonyl oxide, also known as a "Criegee intermediate".[5,7] Carbonyls, carboxyls, and α-hydroxy ketones are the main functional groups in squalene ozonolysis products, largely arising from the isomerization and decomposition reactions of its carbonyl oxides, R—CHOO and R—C(CH_3)OO.[5]

In view of its importance as an ozone sink in human occupied environments, the reaction between squalene and ozone has been increasingly studied in recent years.[2−6,8−12] Human skin lipids reacting with ozone have been examined in real and simulated environments, and results indicate that humans are a significant sink for ozone present in such environments, accounting for up to half of ozone removal.[2,5,11,12] For example, simulated office and aircraft cabin studies have examined the products of squalene ozonolysis due

Received: July 22, 2015
Revised: October 18, 2015
Accepted: October 22, 2015
Published: October 22, 2015

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

Scheme 1. Structures of Squalene, One of the Six Possible Primary Ozonides (POZ) Squalene Can Form upon Reaction with Ozone, and Selected Volatile Products Resulting from the Decomposition of the Unstable POZ[a]



[a]To keep the diagram simple, the carbonyl oxides and stable products produced from the left part of the POZ are not shown. The product labeled in blue, 5-hydroxy-4-oxopentanal (5-OH-4-OPA) may participate in generating the oligomers of the type $M(C_5H_8O_3)_n$, which are prominent in the mass spectrum of non-volatile squalene ozonolysis products observed in this work.

to human presence.[5,11,12] In one simulated aircraft cabin study, it was found that ozone reacting with squalene on human skin and clothes was responsible for a 40−60 ppb increase in the concentration of certain acids, aldehydes, and ketones and >55% ozone removal.[11]

The reaction of squalene with ozone forms several primary products, including acetone, hydroxyacetone, 6-methyl-5-hepten-2-one (6-MHO), 1-hydroxy-6-methyl-5-hepten-1-one, and geranyl acetone, as well as a number of secondary products, including 4-oxopentanal (4-OPA), 5-hydroxy-4-oxopentanal (5-OH-4-OPA), 1,4-butanedial (succinic dialdehyde), 4-methyl-8-oxo-4-nonenal (4-MON), 4-methyl-4-octene-1,8-dial (4-MOD), 4-oxopentanoic acid, and 4-oxobutanoic acid.[1,4,5,13,14] In studies where squalene directly from human skin was scrubbed onto glass wool and exposed to 100 ppb ozone, acetone (∼17 ppb) and 6-MHO (∼14 ppb) represented the most abundant first-generation products, while 4-MON (∼4 ppb) and 4-OPA (∼3.5 ppb) represented the most abundant second-generation products.[5] Some of these products are represented in Scheme 1.

Many of these products have also been detected in forest air and are believed to result from reactions of ozone with plant surfaces containing squalene and related terpenes.[13,14] Squalene's high degree of unsaturation allows for the formation of tertiary and higher-order reaction products, such as ethanedial (glyoxal).[5,6] The volatility of these poorly characterized oxidation products is low, and they may be present on surfaces of indoor dust particles contributing to their redox activity and potential for irritancy.[15]

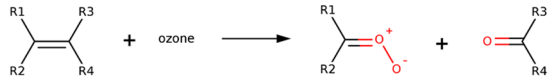
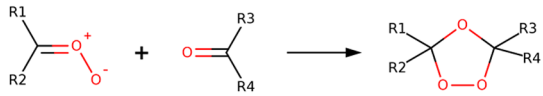
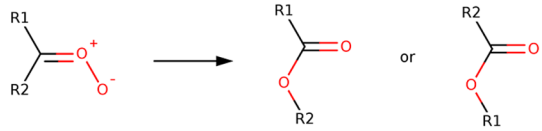
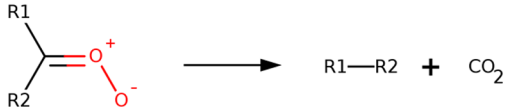
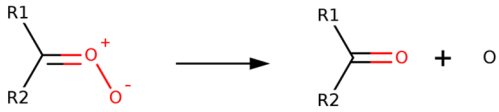
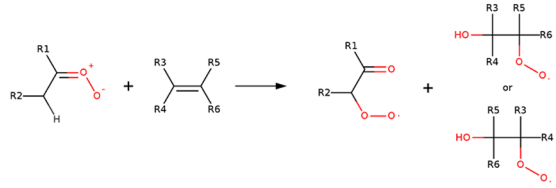
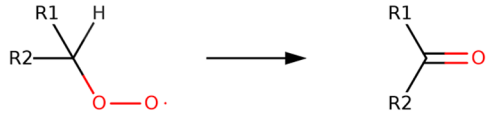
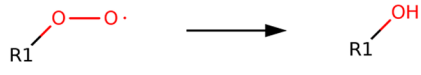
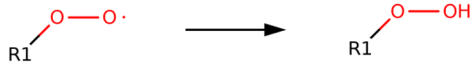
In this study, squalene is exposed to relatively high levels of ozone to simulate its prolonged oxidation, yielding a highly complex mixture of condensed-phase oxidized products. We utilize high-resolution mass spectrometry (HR−MS) to examine these oxidation products in detail and discuss the underlying mechanism. To further aid in understanding this complex system, we use a Computational Brewing Application (COBRA) to model the products of squalene oxidation by ozone.[16] COBRA is optimized to model the evolution of a complex chemical mixture by simulating millions of reactions between compounds in the mixture, after sets of starting products and chemical reaction mechanisms are specified. By studying these computationally predicted products and comparing them with the experimental data, we can better understand the molecular and structural characteristics of squalene oxidation products present on aged indoor surfaces and dust.

2. METHODS

2.1. Experimental Methods. Approximately 50 μL of pure squalene (Sigma-Aldrich, > 98% purity) was uniformly distributed over the surface of a 47 mm polyvinylidene fluoride (PVDF) filter (Millipore HVLP04700, 0.45 μm pore size). The filter was sealed inside a stainless steel filter holder (Millipore XX4404700) and exposed to a 2 SLM (standard liters per minute) flow of oxygen containing 50 ppm ozone for a period of time, typically 1 h. Such a high concentration was chosen to simulate the effects of long-term exposure of squalene to ambient levels of ozone. In terms of integrated exposure, it is roughly equivalent to 1000 h of exposure to 0.050 ppm ozone; however, we cannot exclude the possibility of nonlinear effects on the mechanism associated with the very high ozone mixing ratio used in this study. In the first set of control experiments, a blank filter was exposed to the same flow of ozone. In the second set of control experiments, squalene was applied to the filter but not exposed to ozone. All experiments were done under dry conditions and in the absence of other copollutants, such as NO_x. In actual indoor environments, water and NO_x could potentially affect the reaction mechanism and product

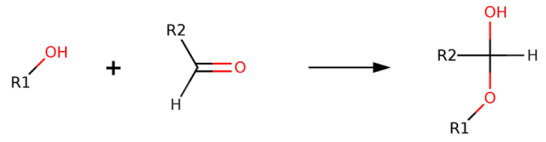
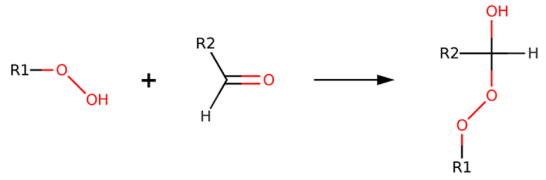
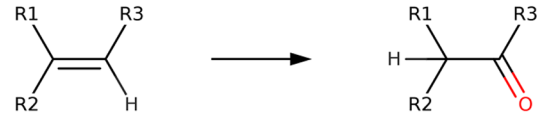
DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

Table 1. List of Chemical Rules Used to Define the Squalene Oxidation System

No.	Description	Chemical Rule
1	Ozone + alkene	
2	CI + carbonyl	
3	CI isomerization into acid or ester	
4	CO_2 elimination from CI	
5	O atom elimination from CI	
6	OH loss from CI + OH addition to an alkene + addition of O_2 to the resulting radicals	
7	Carbonyl formation from RO_2	
8	Alcohol formation from RO_2	
9	Peroxide formation from RO_2	

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

Table 1. continued

No.	Description	Chemical Rule
10	Hemiacetal formation	
11	Peroxyhemiacetal formation	
12	Double bond oxidation	

distribution. Therefore, comparisons of the observed and predicted products of oxidation of squalene with compounds found indoors should be done with care.

The reaction products were extracted from the filter with 5 mL acetonitrile containing 50 μM NaCl to help improve ionization efficiency in the positive ion mode. The resulting solutions were analyzed with a high-resolution ($m/\Delta m = 10^5$ at m/z 450) linear-ion-trap (LTQ) Orbitrap mass spectrometer (Thermo Corp.) using an electrospray ionization (ESI) source in the positive ion mode. The compounds were detected as sodiated $[M + Na]^+$ and/or protonated $[M + H]^+$ species, with the strong predominance of the former. The resulting mass spectra were calibrated with respect to peaks in an Ultramark LTQ ESI positive ion calibration solution (Thermo Scientific) and with respect to the observed prominent peaks in the oxidized squalene mass spectra, such as $[C_{30}H_{50}O_nNa]^+$ with n = 0–6, and $[C_{20}H_{34}O_{12}(C_5H_8O_3)_nNa]^+$ with n = 0–7. The data analysis was carried out as discussed in Nizkorodov et al.[17] Specifically, peaks corresponding to ^{13}C isotopes were removed; peaks that appeared in the oxidized filter control sample were discarded; formula assignments were limited to $C_{1-80}H_{2-140}O_{0-50}Na_{0-1}^+$ with 0.0012 m/z tolerance; the H/C and O/C ratios were constrained to 0.5–2.2 and 0–1.2; and only closed-shell ions (no ion-radicals) were considered. Multiple formula assignments were sometimes possible for peaks with high m/z values; in these cases, preference was given to formulas that continued prominent CH_2 Kendrick series of unambiguously assigned peaks. All formulas discussed in the remainder of this paper correspond to the neutral compounds for the convenience of comparison with the COBRA simulation.

2.2. Computational Methods (COBRA). COBRA is a computational brewing application that can be customized to simulate highly complex chemical systems.[16] It uses a bottom-up approach to simulation, wherein the starting compounds are selected and reaction rules describing the chemical transformations allowed within the system are defined based on experiments or prior chemical knowledge. Starting compounds are input using the commonly used SMILES language, and reaction rules are written using the related SMIRKS language.[18−20] SMIRKS is a reaction transform language used to define chemical reactions to an arbitrary degree of specificity. Each SMIRKS reaction rule represents one possible chemical transformation. Once the set of starting molecules and list of reaction rules is defined, COBRA simulates the evolution of the chemical system by exhaustively applying reaction rules to the starting pool of reactants, thereby generating a set of first-iteration products, which is mixed back into the pool of reactants for future iterations. The simulation continues in this way for an arbitrary number of iterations, resulting in a comprehensive list of all possible products of reactions.

In this study, COBRA was applied to modeling the oxidation of squalene in the presence of ozone. Two starting compounds were used for the simulation: squalene and ozone. Twelve chemical rules based on the current understanding of the mechanism of ozonolysis of alkenes were written to model the chemistry of the system (Table 1).[21] Some of the rules produce Criegee intermediates (CIs) or other free radicals, which react further to form more stable products. To avoid modeling these unstable intermediates explicitly within the product pool, we combined CI- and radical-forming rules with later rules which consume CIs and radicals to form the final set of 12 SMIRKS-encoded rules used by the COBRA simulation (Table S1).

Specifically, Rule 1 describes the formation of a Criegee intermediate (CI) and carbonyl from an alkene, as shown in Scheme 1. Rule 2 corresponds to the formation of a secondary ozonide (SOZ) from the CI and a carbonyl, which is known to be efficient in condensed phases.[22] Rules 3–5 correspond to possible decomposition pathways of CI into stable products.[21] Rule 6 encompasses a series of reactions starting from an OH loss from CI, followed by the addition of OH to a double bond in a neighboring molecule, and an addition of molecular oxygen to all the alkyl radical sites to form peroxyl radicals, RO_2.[23] The extent of OH loss from the CI in the condensed phase is

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

uncertain, and may be low based on pressure-dependence studies, but we include it since it is one of the most important reactions of CIs in the gas phase.[24,25] The RO_2 radicals are relatively stable, and are expected to decay by cross-reactions resulting in carbonyl, alcohol, and peroxide products (Rules 7−9). Our first set of simulations included only rules 1−9 in Table 1, but they could not reproduce some of the major peaks in the experimental mass spectrum. Thus, the rules were amended to include three additional pertinent rules: hemiacetal formation (Rule 10), which appeared to occur in the products based on the experimental data; peroxyhemiacetal formation (Rule 11); and autoxidation of double bonds to carbonyls (Rule 12), which appeared to occur readily in squalene exposed to air (see Figure S1).[9]

We applied a filter such that the maximum mass of the generated structures was limited to 100 heavy atoms (C and O), or approximately 1450 Da, because this was expected to cover the range of peaks detected by mass spectrometry. No other filters were applied. The simulation was allowed to proceed for four iterations, producing over 1 billion unique molecular structures after running for one month on an AMD Opteron 6274 2.2 GHz system. Ideally, the simulations would have run for six iterations to allow all six of squalene's double bonds to be oxidized, but extending the simulation was prohibitively expensive in terms of computer resources.

3. RESULTS AND DISCUSSION

3.1. Experimental Mass Spectrum. The background-corrected ESI mass spectrum of oxidized squalene contained

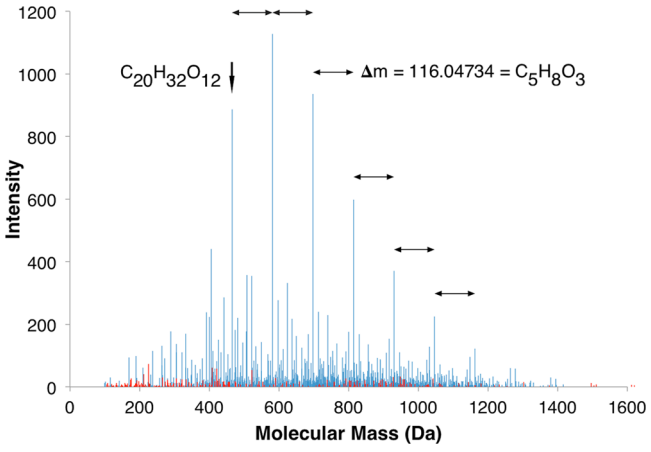


Figure 1. HR−MS experimental data and COBRA predictions of squalene ozonolysis products. HR−MS spectrum peaks that were predicted by the simulation are shown in blue. The remaining experimental peaks that were not predicted are shown in red. Within the simulation mass range of 0−1450 Da, over 83% of HR−MS peaks were predicted, including all major peaks as shown. Also noted is the series of strong peaks with $\Delta m = C_5H_8O_3$.

more than 1500 peaks, and nearly 1300 of them could be unambiguously assigned to elemental formulas within the constraints described above. The reconstructed spectrum of the assigned neutral compounds is shown in Figure 1. ESI is regarded as a "soft" ionization technique, so it is assumed that each neutral compound should produce just one peak in the mass spectrum. In practice, some amount of ion fragmentation or cluster formation may occur even under soft ESI conditions. But even considering this complication, it is clear that

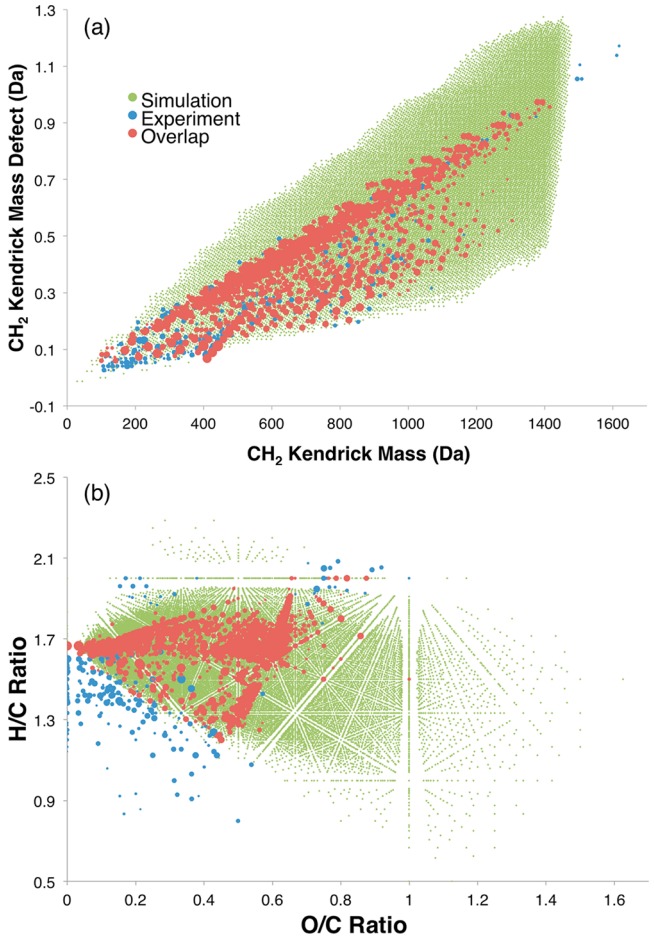


Figure 2. (a) Kendrick plot and (b) Van Krevelen diagram comparing the COBRA results (small green dots) with experimental results (blue circles with diameters proportional to the log of the peak abundance). Red circles are those for which simulated and experimental data overlap. The COBRA simulation produces products that are more oxidized on average (have higher O/C ratios).

ozonolysis of squalene generates a complex mixture of products. This can be contrasted with ozonolysis of lipids that have just one double bond, such as oleic acid or undecylenic acid, which produce a simpler product distribution.[26−28] For example, the undecylenic acid ozonolysis in ref 28 produced a factor of 20 fewer peaks compared to the number of assigned peaks in the squalene system in this work.

In direct-infusion ESI−MS of complex mixtures, the relative intensities of different peaks are not necessarily indicative of the corresponding concentrations because of the selectivity in the ionization process. For example, the spectrum of unoxidized squalene shown in Figure S1, in addition to the major $C_{30}H_{50}$ squalene peak, contains peaks of $C_{30}H_{50}O_n$ with n = 1−6 corresponding to the products of squalene oxidation (all of them detected as Na adducts). These species are likely present at small levels in squalene, but they are enhanced in the mass spectrum because sodium ion affinity increases with the polarity of the analyte. Despite the selectivity in ionization, we can reasonably expect that all highly oxidized products should show up in the mass spectrum. Therefore, we can draw qualitative conclusions from the occurrence and relative intensity distribution of structurally related peaks.

One series of formulas that especially stands out in the experimental spectrum is a progression of

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

Table 2. Comparison of Experimental and Simulated Results from This Work with Experimental Results from Several Prior Studies[a]

system	number of assigned peaks in system	percentage of assigned peaks matched by squalene + O_3 simulation	percentage of assigned peaks matched by squalene + O_3 experimental system
squalene + O_3 products	1258	83%	100%
limonene + O_3 SOA	924	90%	22%
isoprene + OH/NO_x SOA	463	62%	10%
naphthalene + OH/NO_x SOA	242	23%	7%

[a]Percentage of overlapping peaks between experimental and simulated results for ozonized squalene and several SOA systems.

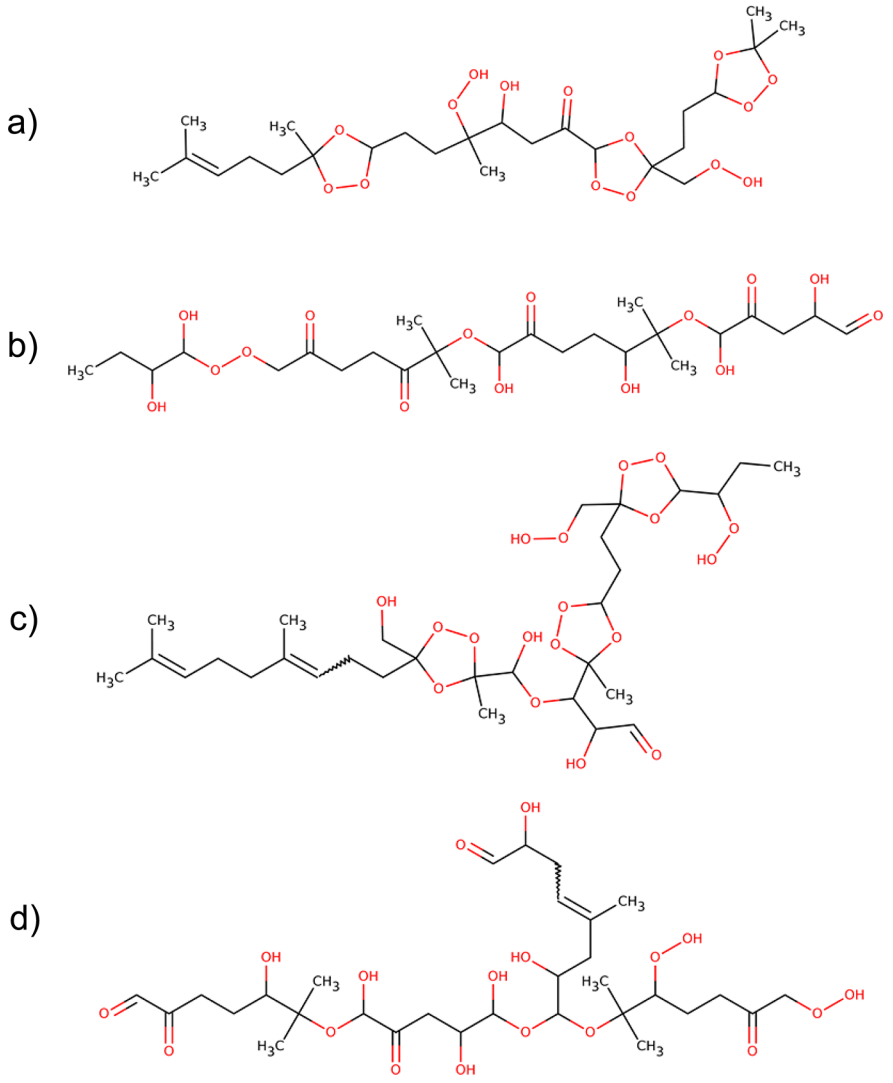


Figure 3. Examples of predicted structures from the two most abundant HR−MS peaks. Parts (a) and (b) are from the most abundant peak, with the formula $C_{25}H_{42}O_{15}$. Parts (c) and (d) are from the second most abundant peak, with the formula $C_{30}H_{50}O_{18}$. COBRA generated 9804 and 32 527 structures for these two peaks, respectively, and the examples above were randomly selected from those sets.

Table 3. Percentage of O Atoms by Functional Group for All Predicted Products

functional group	percentage of O atoms
peroxy	42.5%
hydroxyl	20.9%
ether	18.0%
ketone	9.7%
aldehyde	5.8%
carboxyl	2.0%
ester	1.1%

$C_{20}H_{34}O_{12}(C_5H_8O_3)_n$ with n = 0−7 (Figure 1). The oligomer building block $C_5H_8O_3$ could represent a combination of isoprene (C_5H_8), from which squalene is built, and ozone (O_3). However, it is also possible that $C_5H_8O_3$ corresponds to an actual product of squalene ozonolysis, such as 5-hydroxy-4-oxopentanal (5-OH-4-OPA) shown in Scheme 1. The molecules possessing both hydroxyl and carbonyl groups can cross-couple by means of hemiacetal formation reactions, and 5-OH-4-OPA should be especially active in such a process because aldehydes are more reactive than ketones. In fact, the COBRA simulations described below did not predict some

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

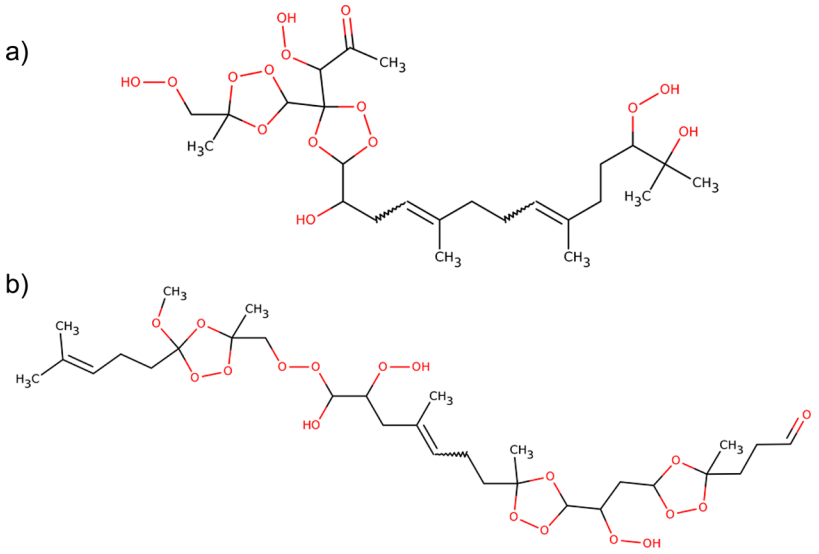


Figure 4. Examples of structures with a high number of peroxy functional groups. Peroxy groups were the most common functional group for O atoms in the predicted products (Table 3). The structures above were predicted for experimental peaks $C_{25}H_{42}O_{15}$ (a) and $C_{30}H_{50}O_{18}$ (b). These structures represent the highest fraction of O atoms in peroxy groups for structures predicted with these molecular formulas.

members of the $C_{20}H_{34}O_{12}(C_5H_8O_3)_n$ series when the hemiacetal formation rule (Rule 10) was disabled, further supporting that reaction's importance.

To help identify other series of related compounds in the mass spectrum, we carried out an analysis of the frequency of occurrence of mass differences between the peaks, similar to the one described in ref 29. According to the results summarized in Table S2, the most common mass differences correspond to $C_5H_8O_3$ (possibly 5-OH-4-OPA), $C_3H_6O_2$ (possibly hydroxyacetone), C_5H_8 (isoprene), O_3 (ozone), oxygen atom, and various combinations of these species. It is significant that all of these differences are chemically meaningful. For example, hydroxyacetone is a known product of squalene oxidation and an easily oligomerizable molecule.[5] Isoprene is a basic building block of squalene, so it is a naturally repeating unit for the squalene oxidation products. The reaction of ozone with alkenes produces SOZ compounds in high yields (Rule 2 in Table 1), with the formula of SOZ derived from the formula of the initial alkene with the addition of 3 oxygen atoms. Finally, the oxygen atom can be added to the molecule via the autoxidation reaction (Rule 12).

3.2. COBRA Simulations. COBRA simulations generated over 1 billion unique structures representing 26 899 elemental formulas after four iterations of computational brewing (where the first iteration is the initial reaction of ozone with one of the double bonds in squalene). Overlap with experimental HR−MS results was good, with the majority of high-intensity peaks being predicted (Figure 1). Specifically, when considering HR−MS observed peaks within the COBRA simulation range, which was limited to a maximum of 100 heavy atoms, or roughly 1450 Da, we predicted 1045 of the 1258 experimentally observed peaks (83.1%). As Figure 1 demonstrates, the peaks predicted by COBRA represent the most abundant compounds detected within this mass range. This suggests that the reaction rules defined in Table 1 and used by COBRA to generate predicted products reasonably capture the essential chemistry of the squalene ozone oxidation system.

Despite the good overlap between the observed and predicted formulas, the fact that 17% of the observed formulas were not reproduced by the simulation suggests that some reaction mechanisms are still missing. Specifically, the simulation appears to miss experimentally observed compounds with low H/C and low O/C ratios, which tend to be outliers on the Van Krevelen diagram in Figure 2b. To further compare the group of observed compounds that is predicted and the group that is not predicted, we computed the average molecular formula and average double bond equivalent (DBE) for the two groups. For observed compounds that are predicted, the average formula is $C_{37.3}H_{61.4}O_{15.2}$ (DBE 7.6), while the average formula for observed compounds that are not predicted is $C_{32.6}H_{49.3}O_{8.3}$ (DBE 9.0). These formulas indicate that the simulation predicts observed compounds that are, on average, larger and more oxidized compared to the ones it does not predict. Interestingly, observed compounds that are not predicted tend to be less saturated on average (have a higher DBE).

We should note that the overlap of the predicted formulas and observed ones is not a definite proof that the chemistry rules we have chosen are correct and comprehensive. We tested the overlap between the formulas predicted for squalene ozonolysis in this work, and formulas observed in ozonolysis of limonene by Bateman et al. (ref 30), and found a high degree of overlap between them (90% of the observed formulas appeared in the simulation output). Limonene is a monoterpene, which, like squalene, is built from isoprene units. Therefore, limonene undergoes ozonolysis by a similar mechanism, so the overlap in this case is not too surprising. We also tested the overlap between formulas predicted in this work and those observed in two other experiments: an isoprene photooxidation SOA (ref 16), and a naphthalene photooxidation SOA (ref 31). Overlap between our squalene simulation from this work and the observed formulas in isoprene and naphthalene SOA was 62% and 23%, respectively. The lower degree of overlap with these two systems is consistent with the mechanistic differences between OH-driven oxidation for isoprene and naphthalene cases, and O_3-driven oxidation in the squalene and limonene cases.

To provide further context for the aforementioned comparisons, we also tested the overlap between experimental results from this work and experimental results from the other

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

works. When comparing experimentally observed formulas for squalene ozonolysis in this work and formulas observed in the other works, the overlap (the percentage of formulas from the other studies that were also observed experimentally in this work) was as follows: 22% for the limonene ozonolysis system (ref 30), 10% for the isoprene photooxidation system (ref 16), and 7% for the naphthalene photooxidation system (ref 31). All of these comparison results are summarized in Table 2.

Figure 2 offers another perspective on the overlap between HR−MS and COBRA results. The top panel shows a CH_2 Kendrick plot for the experimentally observed and simulated peaks. The predicted structures have similar values of the Kendrick mass defects compared to the experimental ones. The Kendrick mass defect is calculated as the difference between the nominal mass and Kendrick mass with respect to the $^{12}CH_2$ group (eqs 1 and 2 in ref 17). The second panel shows a Van Krevelen diagram, wherein the H/C and O/C ratios for all the compounds are plotted against each other.[32] This diagram has a wide range on the O/C axis, which is expected since the chemistry strongly favors oxygen atom addition to the molecules. The simulated compounds are on average more oxidized than the experimentally observed ones. This is not surprising, as the simulation can proceed without kinetic constraints, which is of course not true in reality. First of all, as the products build up, the product mixture likely becomes more viscous, constraining access of ozone and oxygen to the mixture components.[33] Second, it has been shown for the OH-driven oxidation of lipids that oxidation reactions adding oxygen atoms to the products are efficient at the early stages of oxidation, but at later stages they are replaced by fragmentation reactions, which our simulation does not take into account. For OH oxidation of squalane (the fully saturated version of squalene) the fragmentation processes took over when O/C ratio reached about 0.4.[34] On the basis of Figure 2b, the fragmentation may become significant in the ozone−squalene system once the O/C ratio reaches ∼0.7 since not many products are observed with O/C in excess of this value, while there is clearly a potential for accommodating larger amounts of oxygen according to the simulation.

We performed 13 "exclusion simulations" to examine which reaction rules were contributing most to the prediction of experimentally observed products. For each exclusion simulation, a single reaction rule from Table S1 was omitted, and the simulation was run until 500 000 reactions were computed. Then the number of experimental molecular formulas recovered was compared against an analogous control simulation in which all rules were included. The "exclusion percent" column in Table S1 is the percent reduction in the number of correctly predicted formulas when that rule is omitted. That is, a higher exclusion percent indicates a more important rule in terms of contribution to the product distribution. With an exclusion percent of 16.3%, the most important rule in recovering experimentally observed formulas was the sequence of reactions 1 + 6 + 7 from Table 1, which involved OH loss from a Criegee intermediate (CI) with subsequent addition of O_2 to the resulting radical, followed by carbonyl formation from the resulting RO_2 group. The second most important rule, with an exclusion percent of 15%, was the sequence of reactions 1 + 6 + 9, analogous to the previous rule, but with peroxide formation from RO_2 as the last step. The third most important, at 12.2%, appeared to be reaction 2 from Table 1, the formation of a trioxolane ring (SOZ) from CI and a carbonyl. Finally, formation of hemiacetals and peroxyhemiacetals (reactions 10 and 11 from Table 1) had an exclusion percent of 5.8% and appeared to be required to reproduce the most abundant peaks in the experimental spectrum. Some rules had an exclusion percent of 0%, implying that they either produce strictly a subset of products generated by other rules, or that they do not produce experimentally observed structures within the number of reactions computed in an exclusion simulation.

We also compared the table of major primary and secondary products of squalene ozonolysis listed in ref 3. to both our experimental HR−MS data, and the simulated data. All of these products were predicted by the simulation, while half of them were detected in the HR−MS data (Table S3). We note that experiments in ref 3 predominantly detected early products of oxidation generated at low ozone exposures, most of which are volatile species, whereas the experiments reported here focused on later-generation products remaining in the condensed phase. It is therefore expected that some of the readily oxidizable compounds observed under the milder oxidation conditions in ref 3 were not observed here under stronger oxidation conditions. The simulation captures both early- and late-generation products of oxidation, regardless of whether they are volatile or not, hence it agrees with both sets of experiments.

One of the benefits of the simulation is that it has information on molecular structures of every product produced by the simulation. Figure 3 shows examples of COBRA-predicted structures with masses corresponding to the two most abundant HR−MS peaks. The structures in Figure 3A and 3B correspond to the highest intensity peak, $C_{25}H_{42}O_{15}$, and Figure 3C and 3D are structures selected from the second highest peak, $C_{30}H_{50}O_{18}$. We should emphasize that these structures are just randomly selected structural isomers out of thousands of possibilities. For example, there were as many as 9804 structures generated for the aforementioned C_{25} peak, and 32 527 structures generated for the C_{30} peak. These numbers indicate the broad range of structural isomers generated by the system for a given molecular mass. For some molecular formulas, particularly at higher masses, there were hundreds of thousands of structures predicted. Of course, in reality, not all of these isomers will be accessible because of various kinetic and energetic constraints, which are not included in the simulation. Nevertheless, the implication of these results for the interpretation of the experimental mass spectrum (which cannot resolve isobaric compounds) is that the mixture of products is likely even more complex than the number of distinct peaks in the direct-infusion ESI mass spectrum would suggest.

Another advantage of the simulation is that it can provide statistical information on the distribution of important functional groups in the products (once again, we are neglecting possible kinetic and energetic constraints in the actual reaction). Therefore, we calculated the percentage of oxygen atoms in simulated products by functional group, to better understand the chemical composition of the complete set of predicted structures (Table 3). On the basis of the simulation results, peroxy groups are the most prevalent functional group containing oxygen atoms, accounting for 42.5% of oxygen atoms in the products. The next most prevalent oxygen-containing functional groups were hydroxyl, ether, and ketone groups, representing 20.9%, 18.0%, and 9.7% of all oxygen atoms in the simulated formulas, respectively. Figure 4 shows examples of predicted structures containing a high fraction of O atoms in peroxy groups. The two structures

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

were selected from the pool of predicted products for the two most abundant HR−MS peaks, as in Figure 3. They represent the maximum fraction of oxygen atoms in peroxy groups among structures predicted at these masses and are shown for illustrative purposes. We note that while the simulation does not consider possible kinetic and energetic constraints, these statistics and examples are still of use for understanding the composition and characteristics of the simulated product pool.

Simulating the oxidation of squalene in the presence of ozone using COBRA is a good example of the utility of computational tools for modeling and understanding complex natural systems. COBRA can aid in validating experimental data, generating proposed structures for compounds observed in mass spectrometry experiments when combined with electronic structure calculations, and generating possible structures for other experimentally observed chemical formulas. By defining a relatively simple system of chemical transformation rules and starting compounds, this computational approach can be a powerful tool for better understanding the complexity of a natural chemical system composed of tens or hundreds of thousands of compounds. These results are atmospherically relevant because they may help understand the nature of compounds found in highly aged indoor dust resulting from skin flakes and highly soiled indoor surfaces, such as airplane cabins and heavily populated office spaces.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.5b03552.

(Figure S1) An ESI mass spectrum of squalene observed before the ozone exposure; (Table S1) a graphical representation of the SMIRKS-encoded reaction rules used to define the squalene oxidization simulation; (Table S2) prominent mass differences observed in oxidized squalene, which correspond to $C_cH_hO_o$ compounds; (Table S3) a list of major primary and secondary products generated by ozonolysis of squalene (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: (949) 824-1262; e-mail: nizkorod@uci.edu (S.A.N.).
*Phone: (949) 824-5809; e-mail: pfbaldi@ics.uci.edu (P.F.B.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Partial initial support for this project was provided by a seed grant from the UCI Environmental Institute. The work of D.F. and P.B. has been supported in part by NSF grant IIS 0513376, a Google Faculty Research Award, and an unrestricted gift from ExxonMobil to P.B. S.N. acknowledges support by the NSF grants AGS-1227579. P.A. thanks the Ford Foundation Predoctoral Fellowship Program of the National Academy of Science for their support. The PNNL group acknowledges support from the Chemical Sciences Division (J.L.), Office of Basic Energy Sciences of the U.S. DOE, and Laboratory Directed Research and Development program (A.L.) of the W.R. Wiley Environmental Molecular Sciences Laboratory (EMSL)—a national scientific user facility located at PNNL, and sponsored by the Office of Biological and Environmental Research of the U.S. PNNL is operated for US DOE by Battelle Memorial Institute under Contract No. DE-AC06-76RL0 1830.

■ REFERENCES

(1) Fruekilde, P.; Hjorth, J.; Jensen, N. R.; Kotzias, D.; Larsen, B. Ozonolysis at vegetation surfaces: a source of acetone, 4-oxopentanal, 6-methyl-5-hepten-2-one, and geranyl acetone in the troposphere. *Atmos. Environ.* **1998**, *32* (11), 1893−1902.

(2) Wang, C.; Waring, M. S. Secondary organic aerosol formation initiated from reactions between ozone and surface-sorbed squalene. *Atmos. Environ.* **2014**, *84*, 222−229.

(3) Weschler, C. J. Roles of the human occupant in indoor chemistry. *Indoor Air* **2015**, DOI: 10.1111/ina.12185.

(4) Weschler, C. J.; Langer, S.; Fischer, A.; Bekö, G.; Toftum, J.; Clausen, G. Squalene and cholesterol in dust from Danish homes and daycare centers. *Environ. Sci. Technol.* **2011**, *45* (9), 3872−3879.

(5) Wisthaler, A.; Weschler, C. J. Reactions of ozone with human skin lipids: Sources of carbonyls, dicarbonyls, and hydroxycarbonyls in indoor air. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (15), 6568−6575.

(6) Wells, J. R.; Morrison, G. C.; Coleman, B. K.; Spicer, C.; Dean, S. W. Kinetics and reaction products of ozone and surface-bound squalene. *J. ASTM Int.* **2008**, *5* (7), 1−12.

(7) Criegee, R. Mechanism of ozonolysis. *Angew. Chem., Int. Ed. Engl.* **1975**, *14* (11), 745−752.

(8) Petrick, L.; Dubowski, Y. Heterogeneous oxidation of squalene film by ozone under various indoor conditions. *Indoor Air* **2009**, *19* (5), 381−391.

(9) Fu, D.; Leng, C.; Kelley, J.; Zeng, G.; Zhang, Y.; Liu, Y. ATR-IR Study of Ozone Initiated Heterogeneous Oxidation of Squalene in an Indoor Environment. *Environ. Sci. Technol.* **2013**, *47* (18), 10611−10618.

(10) Wang, H.; He, C.; Morawska, L.; McGarry, P.; Johnson, G. Ozone-initiated particle formation, particle aging, and precursors in a laser printer. *Environ. Sci. Technol.* **2012**, *46* (2), 704−712.

(11) Weschler, C. J.; Wisthaler, A.; Cowlin, S.; Tamás, G.; Strøm-Tejsen, P.; Hodgson, A. T.; Nazaroff, W. W. Ozone-initiated chemistry in an occupied simulated aircraft cabin. *Environ. Sci. Technol.* **2007**, *41* (17), 6177−6184.

(12) Wisthaler, A.; Tamas, G.; Wyon, D. P.; Strom-Tejsen, P.; Space, D.; Beauchamp, J.; Hansel, A.; Maerk, T. D.; Weschler, C. J. Products of Ozone-Initiated Chemistry in a Simulated Aircraft Environment. *Environ. Sci. Technol.* **2005**, *39* (13), 4823−4832.

(13) Matsunaga, S.; Mochida, M.; Kawamura, K. Variation on the atmospheric concentrations of biogenic carbonyl compounds and their removal processes in the northern forest at Moshiri, Hokkaido Island in Japan. *J. Geophys. Res.* **2004**, *109*, D04302 DOI: 10.1029/2003JD004100.

(14) Matsunaga, S.; Mochida, M.; Kawamura, K. High abundance of gaseous and particulate 4-oxopentanal in the forestal atmosphere. *Chemosphere* **2004**, *55*, 1143−1147.

(15) Anderson, S. E.; Franko, J.; Jackson, L. G.; Wells, J. R.; Ham, J. E.; Meade, B. J. Irritancy and allergic responses induced by exposure to the indoor air chemical 4-oxopentanal. *Toxicol. Sci.* **2012**, *127* (2), 371−381.

(16) Fooshee, D. R.; Nguyen, T. B.; Nizkorodov, S. A.; Laskin, J.; Laskin, A.; Baldi, P. COBRA: A Computational Brewing Application for Predicting the Molecular Composition of Organic Aerosols. *Environ. Sci. Technol.* **2012**, *46*, 6048−6055.

(17) Nizkorodov, S. A.; Laskin, J.; Laskin, A. Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3612−3629.

(18) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model.* **1988**, *28* (1), 31−36.

(19) *Daylight Theory Manual*; Daylight Chemical Information Systems, Inc., release date August 1, 2011. http://www.daylight.com/dayhtml/doc/theory/index.html.

(20) *OpenEye Scientific Software, I., OEChem*, version 1.7.4; In OpenEye Scientific Software: Santa Fe, NM, USA, 2010.

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313

(21) Finlayson-Pitts, B. J. and Pitts, J. N. *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*; Academic Press: San Diego, 2000; pp 1040.
(22) Bailey, P. S. *Organic Chemistry, Vol. 39, Pt. 1: Ozonation in Organic Chemistry, Vol. 1: Olefinic Compounds*; Academic Press: New York, 1978; 272 pp.
(23) Von Sonntag, C.; Schuchmann, H. P. The Elucidation of Peroxyl Radical Reactions in Aqueous Solution with the Help of Radiation-Chemical Methods. *Angew. Chem., Int. Ed. Engl.* **1991**, *30* (10), 1229−1253.
(24) Kroll, J. H.; Clarke, J. S.; Donahue, N. M.; Anderson, J. G.; Demerjian, K. L. Mechanism of HO_x formation in the gas-phase ozone-alkene reaction. 1. Direct, pressure-dependent measurements of prompt OH yields. *J. Phys. Chem. A* **2001**, *105* (9), 1554−1560.
(25) Drozd, G. T.; Donahue, N. M. Pressure Dependence of Stabilized Criegee Intermediate Formation from a Sequence of Alkenes. *J. Phys. Chem. A* **2011**, *115* (17), 4381−4387.
(26) Walser, M. L.; Desyaterik, Y.; Laskin, J.; Laskin, A.; Nizkorodov, S. A. High-resolution mass spectrometric analysis of secondary organic aerosol produced by ozonation of limonene. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1009−1022.
(27) Zahardis, J.; Petrucci, G. A. The oleic acid-ozone heterogeneous reaction system: products, kinetics, secondary chemistry, and atmospheric implications of a model system - a review. *Atmos. Chem. Phys.* **2007**, *7* (5), 1237−1274.
(28) Gomez, A. L.; Park, J.; Walser, M. L.; Lin, A.; Nizkorodov, S. A. UV photodissociation spectroscopy of oxidized undecylenic acid films. *J. Phys. Chem. A* **2006**, *110* (10), 3584−3592.
(29) Kunenkov, E. V.; Kononikhin, A. S.; Perminova, I. V.; Hertkorn, N.; Gaspar, A.; Schmitt-Kopplin, P.; Popov, I. A.; Garmash, A. V.; Nikolaev, E. N. Total mass difference statistics algorithm: a new approach to identification of high-mass building blocks in electrospray ionization Fourier transform ion cyclotron mass spectrometry data of natural organic matter. *Anal. Chem.* **2009**, *81* (24), 10106−10115.
(30) Bateman, A. P.; Nizkorodov, S. A.; Laskin, J.; Laskin, A. Time-resolved molecular characterization of limonene/ozone aerosol using high-resolution electrospray ionization mass spectrometry. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7931−7942.
(31) Lee, H. J.; Aiona, P. K.; Laskin, A.; Laskin, J.; Nizkorodov, S. A. Effect of solar radiation on the optical properties and molecular composition of laboratory proxies of atmospheric brown carbon. *Environ. Sci. Technol.* **2014**, *48*, 10217−10226.
(32) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Anal. Chem.* **2003**, *75* (20), 5336−5344.
(33) Turek, C.; Stintzing, F. C. Stability of Essential Oils: A Review. *Compr. Rev. Food Sci. Food Saf.* **2013**, *12*, 40−53.
(34) Kroll, J. H.; Smith, J. D.; Che, D. L.; Kessler, S. H.; Worsnop, D. R.; Wilson, K. R. Measurement of fragmentation and functionalization pathways in the heterogeneous oxidation of oxidized organic aerosol. *Phys. Chem. Chem. Phys.* **2009**, *11* (36), 8005−8014.

DOI: 10.1021/acs.est.5b03552
Environ. Sci. Technol. 2015, 49, 13304−13313