

## Electronic Supplementary Information (ESI) Section

### Brown carbon formation from ketoaldehydes of biogenic monoterpenes

**Tran B. Nguyen,<sup>\*a‡</sup> Alexander Laskin,<sup>b</sup> Julia Laskin,<sup>c</sup> and Sergey A. Nizkorodov<sup>a</sup>**

<sup>a</sup> Department of Chemistry, University of California, Irvine, California, USA 92697. E-mail: tbn@caltech.edu; Tel: +1-626-395-8650

<sup>b</sup> Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, USA 99352

<sup>c</sup> Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA 99352

<sup>†</sup>Electronic Supplementary Information (ESI) available: Table S1 – NMR data for identification of PA, LA, and KLA; Figure S1 – proton NMR spectra of LA, KLA, and PA confirming their purity; Figure S2 – LC-UVIS separation of the C<sub>500</sub> fraction; Figure S3 – control experiment with evaporation of KLA alone with reduced nitrogen compounds added; Figure S4 – comparison of the rates of aging of KLA and limonene SOA by gaseous ammonia; Figure S5 – proton NMR spectra of the C<sub>500</sub> fraction in D<sub>2</sub>O and in H<sub>2</sub>O; Figure S6 – absorption spectra recorded by the PDA detector at different elution times; Figure S7 – graphical representation of the compounds detected by HR-MS; Scheme S1 – reactions of KLA leading to low DBE compounds and supporting references.

<sup>‡</sup>Currently at: Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125

**Table S1.** <sup>1</sup>H resonance chemical shifts (δ), multiplicity (M; s = singlet, d = doublet, m = multiplet) and corresponding number of H atoms (from integrated peak areas) for identification of pinonaldehyde (PA), limononaldehyde (LA) and keto-limononaldehyde (KLA). <sup>1</sup>H NMR spectra and structures are shown in the main text of the manuscript. NMR shift data reported here are identical to those reported elsewhere for PA,<sup>1</sup> LA,<sup>2</sup> and KLA.<sup>3,4</sup>

<i>Pinonaldehyde</i>			<i>Limononaldehyde</i>			<i>Ketolimononaldehyde</i>		
δ (ppm)	M	#H	δ (ppm)	M	#H	δ (ppm)	M	#H
0.85	s	3	1.68 – 1.71	m	5	1.64 – 1.82	m	2
1.35	s	3	2.14	s	4	1.86 – 2.02	m	1
1.90 – 2.05	m	2	2.30 – 2.52	m	4	2.16	s	3
2.07	s	3	2.68	s	1	2.27	s	3
2.36 – 2.53	m	3	4.78, 4.84	d	2	2.38 – 2.60	m	2
2.88 – 2.96	m	1	9.68	s	1	2.90 – 3.18	m	2
9.75	s	1				9.73	s	1

### Identification of synthesized compounds

Generic name: Limononaldehyde (LA)

Molecular formula: C<sub>10</sub>H<sub>16</sub>O<sub>2</sub>

IUPAC name: 3-isopropenyl-6-oxoheptanal

SMILES (canonical): CC(=O)CCC(CC=O)C(=C)C

CAS RN#: 7086-79-5

Generic name: Ketolimononaldehyde (KLA)

Molecular formula: C<sub>9</sub>H<sub>14</sub>O<sub>3</sub>

IUPAC name: 3-Acetyl-6-oxoheptanal

SMILES (canonical): CC(=O)C(CC=O)CCC(=O)C

CAS RN#: 71155-94-7

Generic name: Pinonaldehyde (PA)

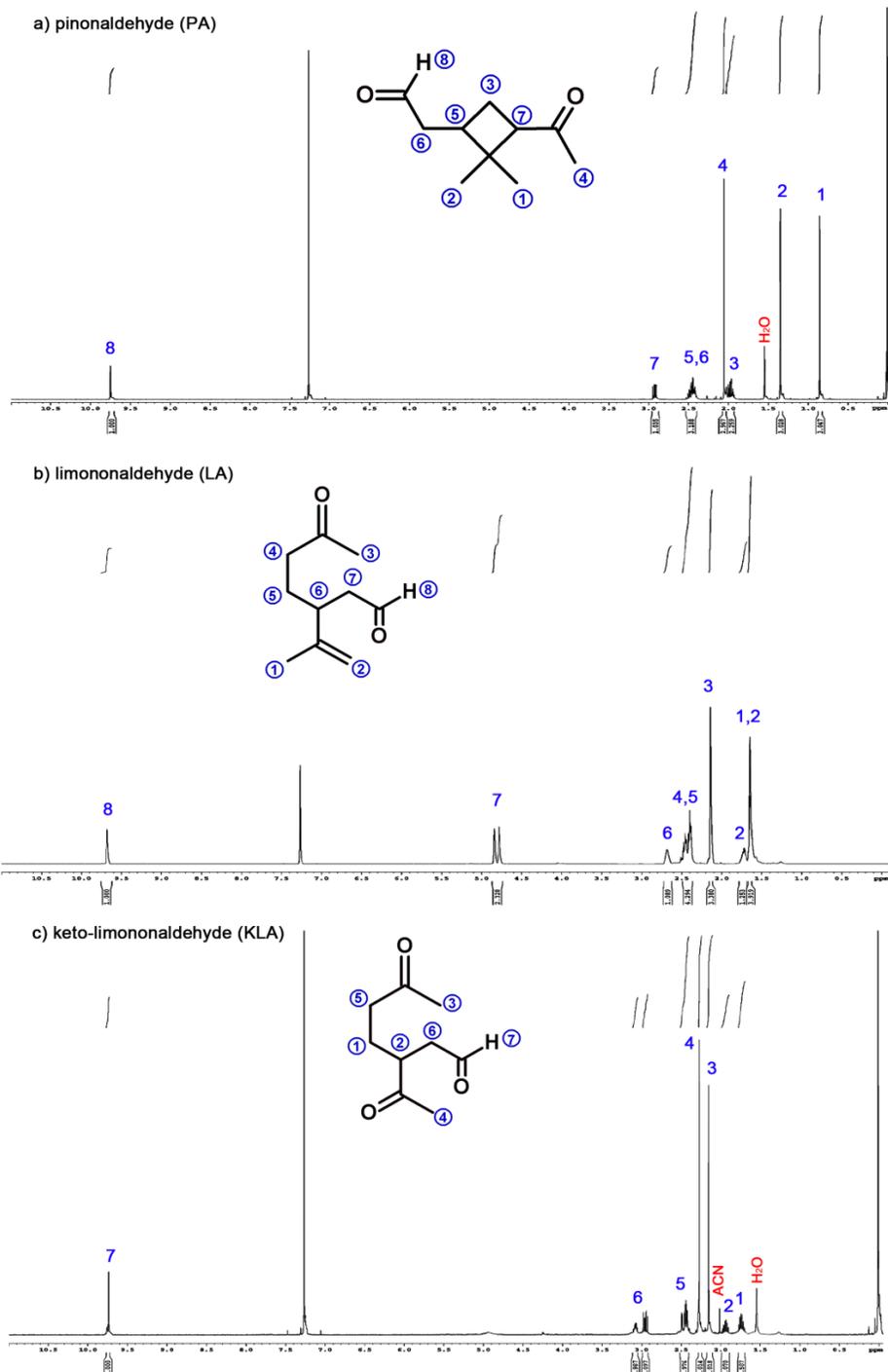
Molecular formula: C<sub>10</sub>H<sub>16</sub>O<sub>2</sub>

IUPAC name: 2-(3-Acetyl-2,2-dimethylcyclobutyl)ethanal

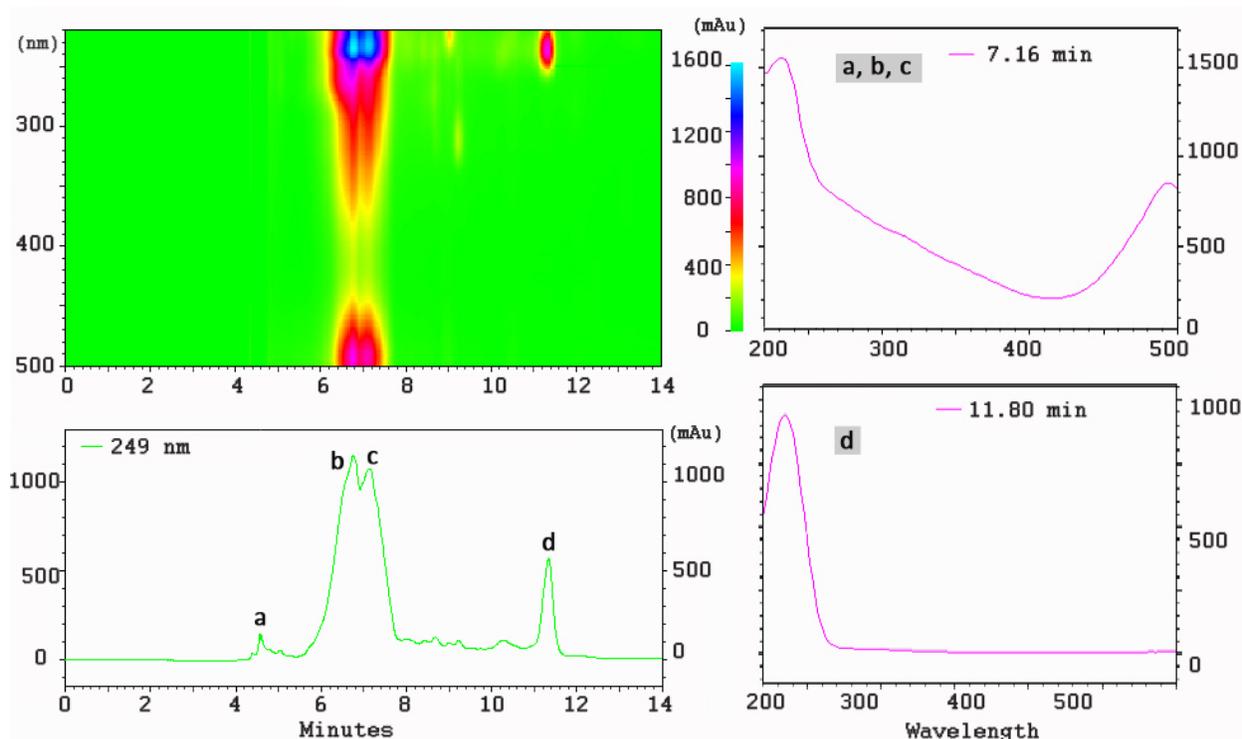
SMILES (canonical): O=C(C)C1CC(CC=O)C1(C)C

CAS RN#: 2704-78-1

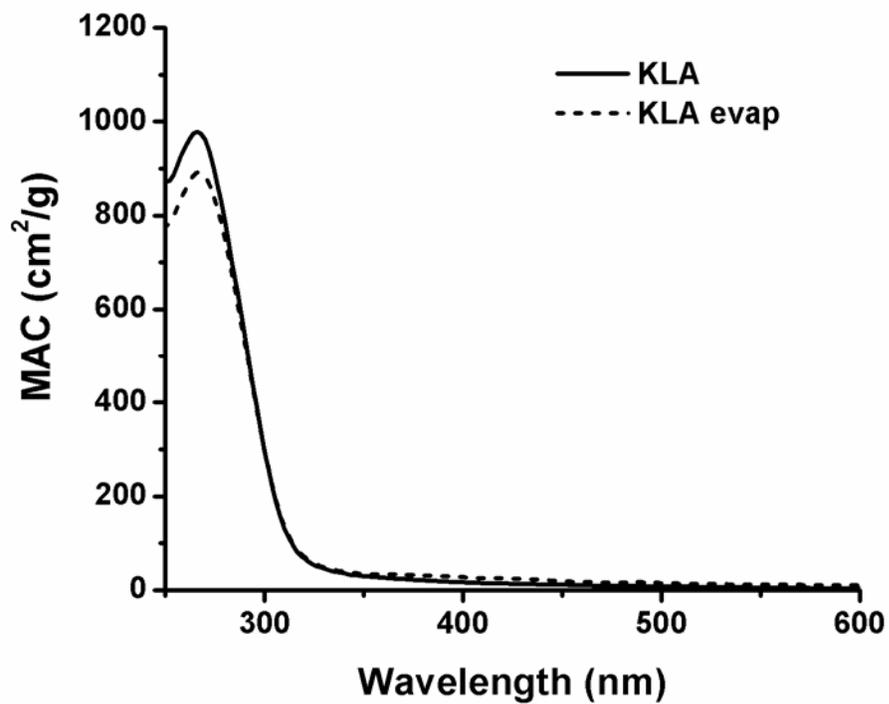
**Figure S1:**  $^1\text{H}$  NMR spectra taken in  $\text{CDCl}_3$  of pinonaldehyde (PA), limononaldehyde (LA) and keto-limononaldehyde (KLA) in the range of -0.5 - 11 ppm. No peaks were observed above 11 ppm, which indicates an absence of carboxylic acids. Chemical shift and integration data are presented in Table S1. Resonances at 7.27 ppm and 0 ppm are due to  $\text{CHCl}_3$  and TMS, respectively. Note LA spectrum was not spiked with TMS. Trace impurities from solvents are noted based on expected positions reported in Ref 5. All synthesized compounds are > 95% pure and are identical to spectra reported in the literature for PA<sup>1</sup>, LA<sup>2</sup>, and KLA.<sup>3,4</sup>



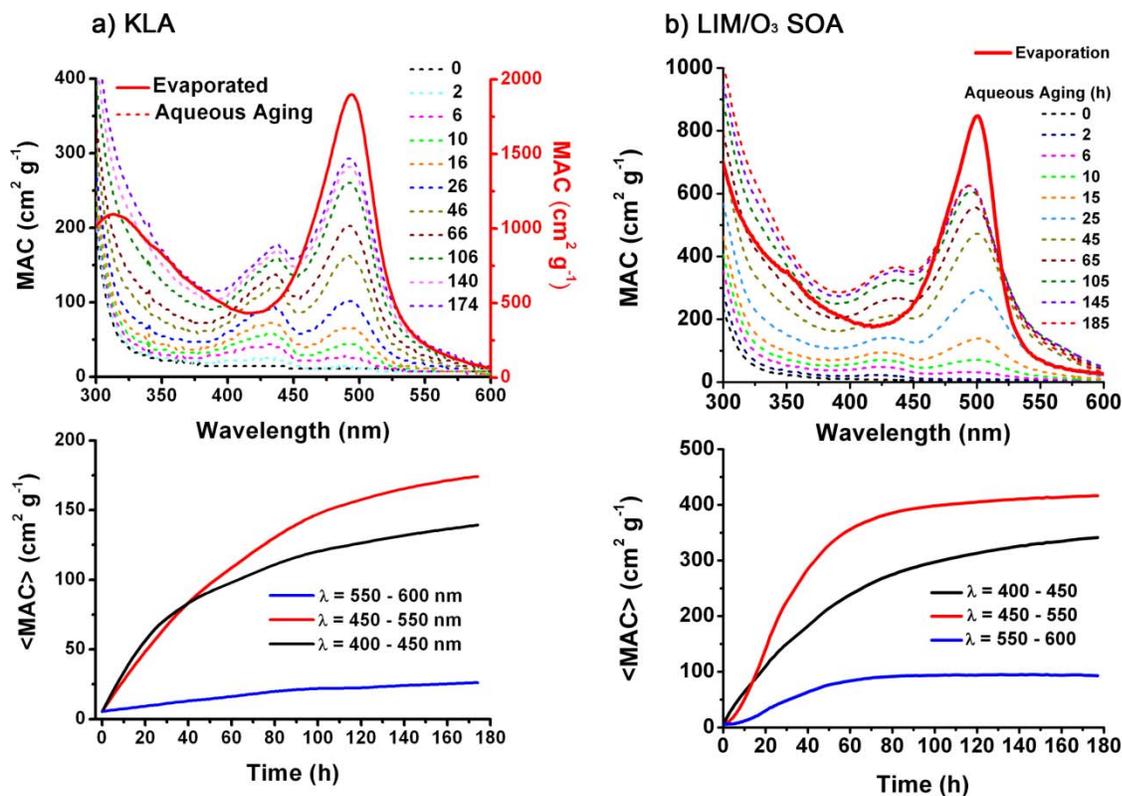
**Figure S2:** A sample LC-UVVIS chromatogram and spectra for the evaporated KLA+AS mixture using (1:1) ACN:H<sub>2</sub>O mixture as the eluent on a C18 semi-prep column. The upper left panel shows the 2D absorption spectrum vs. retention time plot of the mixture. The lower left panel shows the chromatogram recorded at 249 nm. The right panels show characteristic absorption spectra corresponding to peaks (a – d) in the chromatogram. The C<sub>500</sub> fraction discussed in the text corresponds to combined peaks b and c.



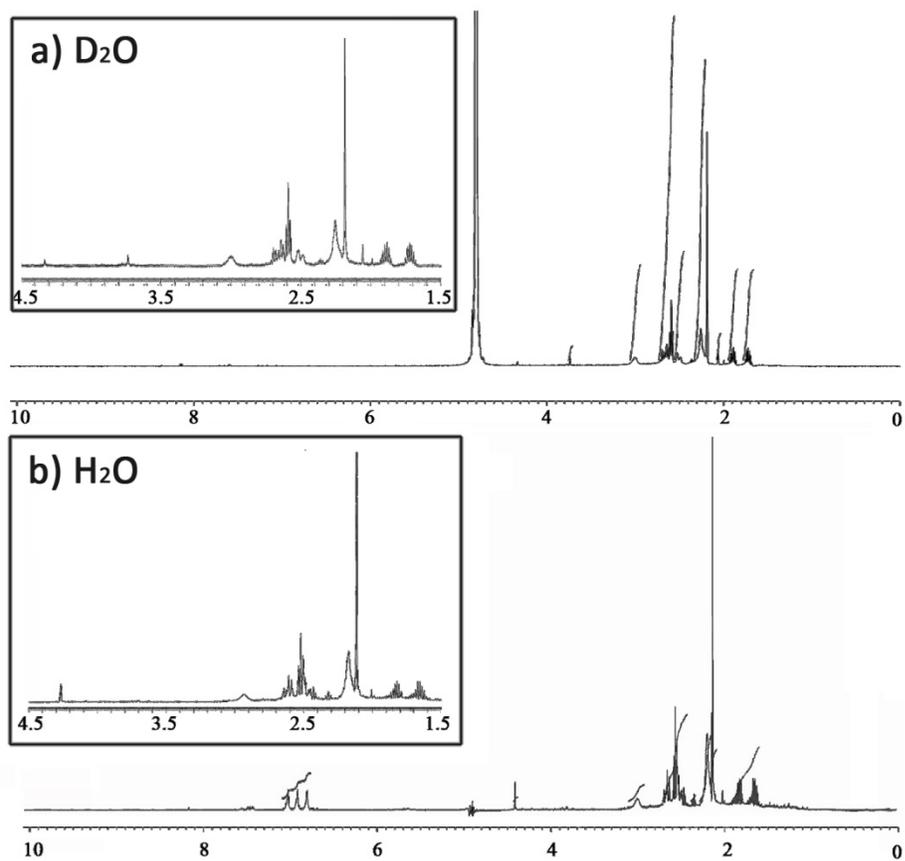
**Figure S3:** Control for the evaporation experiment of KLA. Although KLA efficiently browns in aqueous or evaporative experiments when  $\text{NH}_4^+$  or GLY is added, negligible changes in light absorbing properties (e.g., positive MAC in visible wavelengths) occur when KLA stays in solution or is evaporated without additives. PA and LA do not brown in any experiments, with or without  $\text{NH}_4^+$  or GLY.



**Figure S4:** Aging time profile of (a) KLA +  $\text{NH}_4^+$  and (b) d-limonene/ $\text{O}_3$  SOA +  $\text{NH}_4^+$  reaction in solution. Panels in (b) were reproduced with permission from *Nguyen et al (2012)*.<sup>6</sup> The red trace shows the MAC value of the same solution after one evaporation and re-dissolution. The bottom panel shows time dependence of MAC averaged over three different wavelength ranges surrounding the 430 nm band (400-450 nm, black trace), the 500 nm band (450-550 nm, red trace), and the position of the spectrum above 550 nm (blue trace).

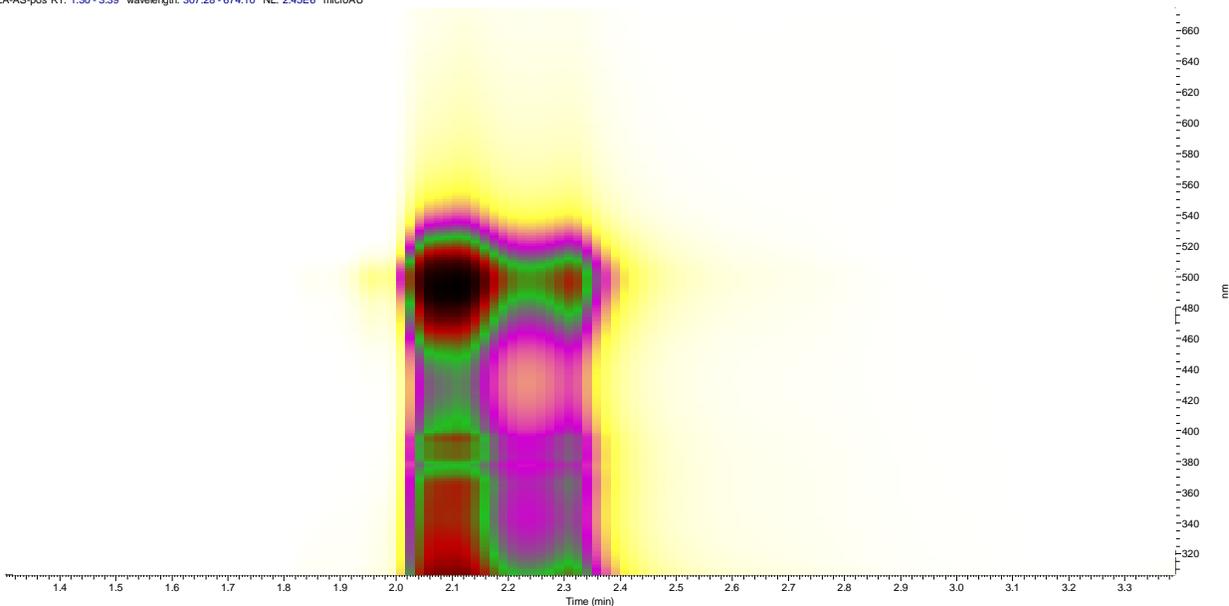


**Figure S5:**  $^1\text{H}$  NMR spectrum of  $\text{C}_{500}$  taken in primarily  $\text{D}_2\text{O}$  or  $\text{H}_2\text{O}$ . Protons that exchange with the solvent (NH, OH, COOH) are not visible in the  $\text{D}_2\text{O}$  spectrum.

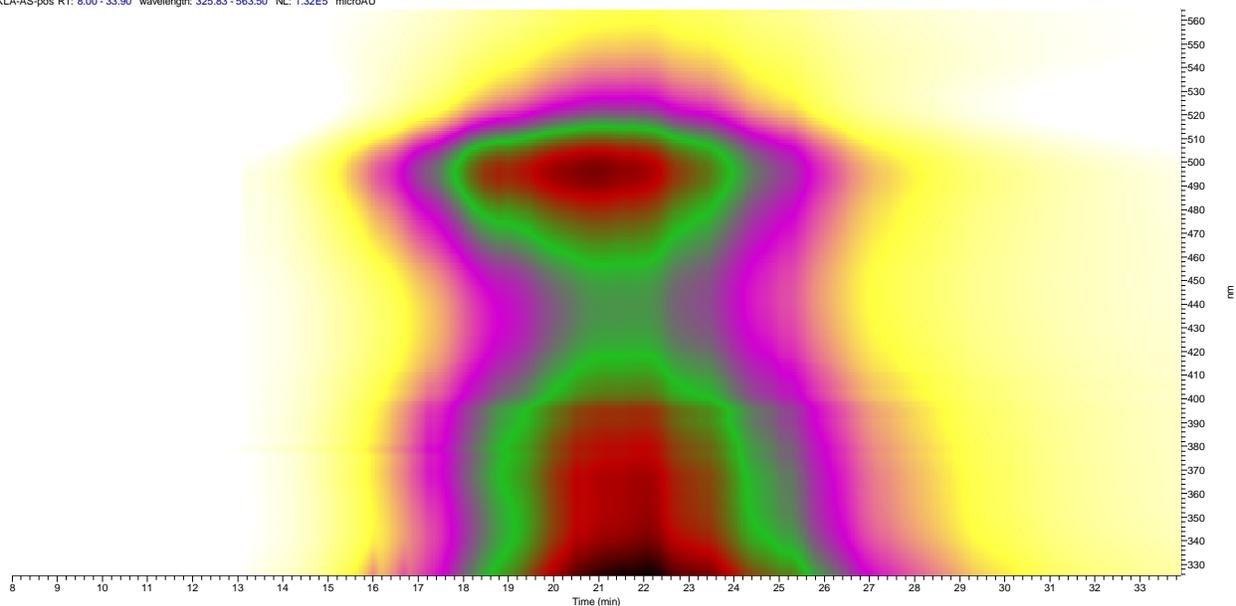


**Figure S6:** PDA absorption spectra of the compounds eluting from the column at early (top) and late (bottom) retention times following injection of the KLA+  $\text{NH}_4^+$  reaction products in the LC-UVVIS-HRMS instrument. The 500 nm absorption band is clearly observed for both early-eluting charged species and late-eluting neutral species.

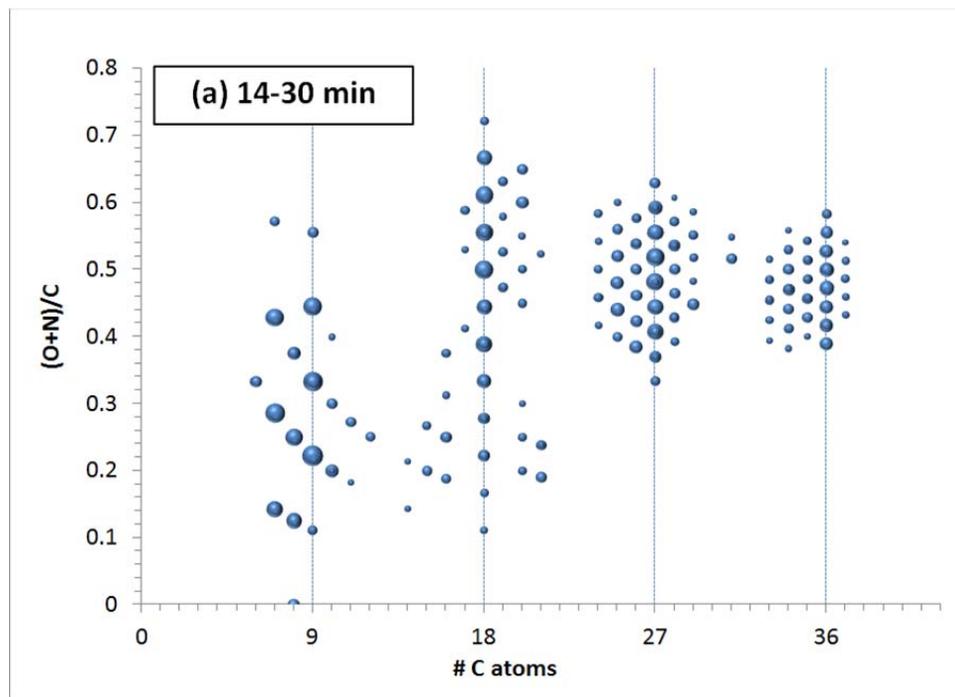
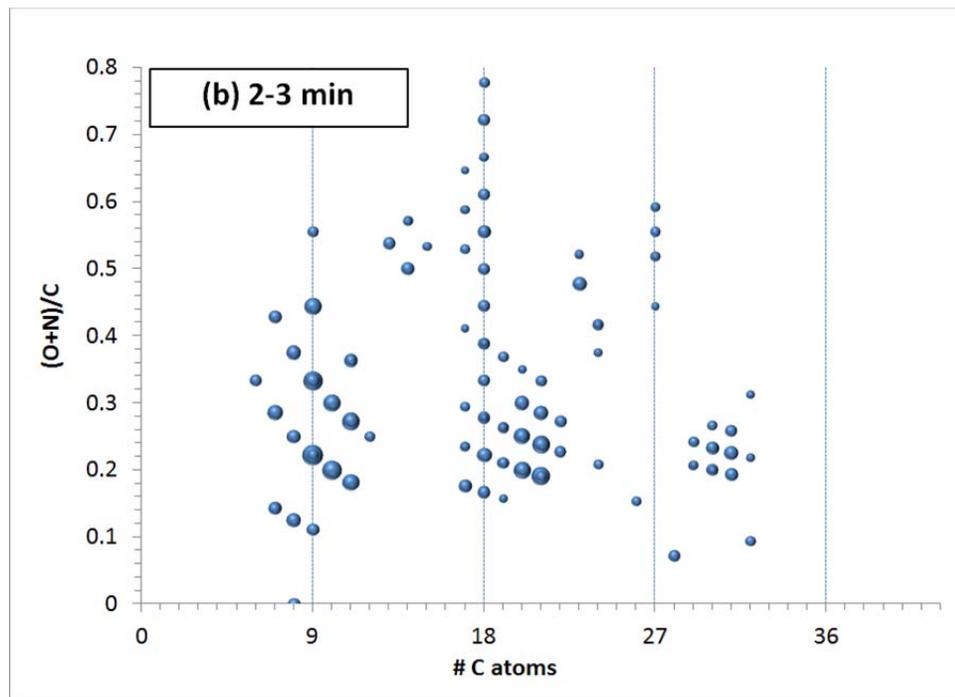
KLA-AS-pos RT: 1.30 - 3.39 wavelength: 307.28 - 674.10 NL: 2.45E6 microAU



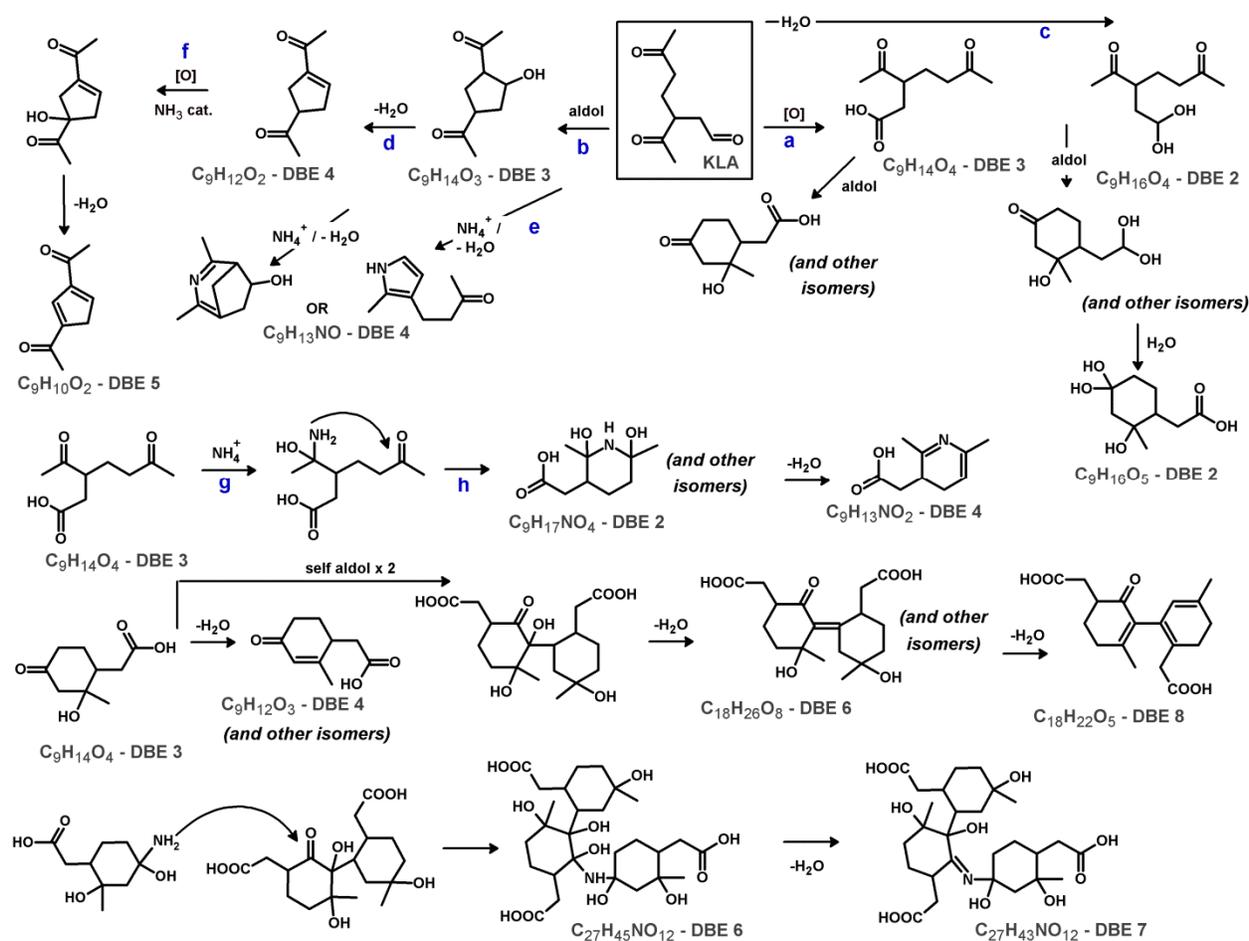
KLA-AS-pos RT: 8.00 - 33.90 wavelength: 325.83 - 563.50 NL: 1.32E5 microAU



**Figure S7:** Pictorial representation of the compounds eluting early (2-3 min) and late (14-30 min) in the LC-UVVIS-ESI/HR-MS measurements of the  $\text{NH}_3(\text{g}) + \text{KLA}$  reaction products. The sizes of the points are proportional to the logarithms of the corresponding ion abundances. As expected, most products have carbon skeletons traceable to KLA, e.g., monomers ( $\text{C}_9$ ), dimers ( $\text{C}_{18}$ ), trimers ( $\text{C}_{27}$ ), and tetramers ( $\text{C}_{36}$ ).



**Scheme S1:** Possible reactions of KLA, leading to products that (likely) do not absorb visible light. Structures drawn are representative only, based on the KLA structure and expected chemistry. Likely, many isomers exist. Molecular formulas shown in the scheme, with their proposed corresponding structures, have been detected with HR-ESIMS (Table S2). Reactions shown: **(a)** oxidation in water of the aldehyde to an acid (spontaneous but slow, enhanced by evaporation); **(b)** aldol formation from a ketone enolate nucleophilic attack on the aldehyde; **(c)** hydration of the aldehyde to a geminal diol; **(d)** dehydration, in this case of the aldol to an  $\alpha,\beta$ -unsaturated ketone (also slow in water, but enhanced with evaporation); **(e)** Paal-Knorr pyrrole formation<sup>7,8</sup> from protonation of the aldehyde (by  $\text{NH}_4^+$ ) to form a hemiaminal (unstable, see g) which nucleophilically adds to a ketone (see h), subsequent dehydration forms a pyrrole; **(f)**  $\alpha$ -hydroxylation of ketones,<sup>9,10</sup> that may be catalyzed by ammonia and amino acids, dehydration of which may be the source of higher-than-expected unsaturation in some compounds (e.g.,  $\text{C}_9\text{H}_{10}\text{O}_2$ ); **(g)** hemiaminal formation from the reaction of  $\text{NH}_4^+$  and a ketone; **(h)** the fast nucleophilic attack by the hemiaminal to close the ring, subsequent dehydration forms heterocyclic nitrogen compounds.



## References to Scheme S1

1. C. W. Dicus, D. Willenbring and M. H. Nantz, *Journal of Labelled Compounds and Radiopharmaceuticals*, 2005, **48**, 223-229.
2. S.-T. Liu, K. V. Reddy and R.-Y. Lai, *Tetrahedron*, 2007, **63**, 1821-1825.
3. D. Yang and C. Zhang, *The Journal of Organic Chemistry*, 2001, **66**, 4814-4818.
4. K. Griesbaum, M. Hilb and J. Bosch, *Tetrahedron*, 1996, **52**, 14813-14826.
5. H. E. Gottlieb, V. Kotlyar and A. Nudelman, *The Journal of Organic Chemistry*, 1997, **62**, 7512-7515.
6. T. B. Nguyen, P. B. Lee, K. M. Updyke, D. L. Bones, J. Laskin, A. Laskin and S. A. Nizkorodov, *J. Geophys. Res.*, 2012, **117**, D01207; doi:01210.01029/02011JD016944
7. C. Paal, *Berichte der deutschen chemischen Gesellschaft*, 1884, **17**, 2756-2767.
8. L. Knorr, *Berichte der deutschen chemischen Gesellschaft*, 1884, **17**, 2863-2870.
9. R. Howe and F. J. McQuillin, *Journal of the Chemical Society (Resumed)*, 1958, 1513-1518.
10. B. Plietker, *Tetrahedron: Asymmetry*, 2005, **16**, 3453-3459.