1 Qualitative and quantitative assessment of unresolved complex mixture in PM_{2.5} of

2 Bakersfield, CA.

Punith Dev Nallathamby¹, Michael Lewandowski², Mohammed Jaoui³, John H. Offenberg², 3 Tadeusz E. Kleindienst², Caitlin Rubitschun⁴, Jason D. Surratt⁴, Sascha Usenko¹ and Rebecca J. 4 Sheeslev^{1*} 5 6 7 ¹ Department of Environmental Science, Baylor University, Waco, TX ² National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC, USA 8 ³ Alion Science and Technology, Research Triangle Park, NC, USA 9 ⁴ Department of Environmental Sciences and Engineering, Gillings School of Global Public 10 Health, University of North Carolina at Chapel Hill, NC, USA 11 12 13 Abstract 14 The 2010 CalNex (California Nexus) field experiment offered an opportunity for detailed characterization of atmospheric particulate carbon composition and sources in Bakersfield, CA. 15 In the current study, the authors describe and employ a new protocol for reporting unresolved 16 complex mixture (UCM) in over 30 daily samples. The Bakersfield, CA site has significant 17 contribution from UCM, $2.9 \pm 2.2\%$ of the daily OC, which makes it an ideal first application. The new 18 protocol reports two UCM peaks for Bakersfield with unique mean vapor pressure, retention 19 time, mass spectra and daily ambient concentration trends. The first UCM peak, UCM-A, was 20 comprised of semi-volatile compounds including alkanes, alkenes, and alkynes, with a mean 21 vapor pressure of 2E-04 Torr and medium to heavy-duty diesel exhaust as a likely source. The 22 second UCM peak, UCM-B, was comprised of linear, branched, and cyclic alkanes, with a mean 23 vapor pressure of 1E-08 Torr. UCM-B had strong similarities to UCM in the NIST Standard 24 25 Reference Material 1649b (urban dust) and to previously reported, detailed UCM for a representative Bakersfield sample, with possible sources including: motor vehicle exhaust, 26 27 agricultural activities, and construction activities.

- 28 *Corresponding author. Tel.: +1 2547103158; fax: +1 2547103409; email: rebecca_sheesley@baylor.edu,
- 29 Baylor University, OBP #97266, Waco, TX USA 76798

30

31 Keywords: Southern California, semi-volatile organic carbon, traffic emissions, PM2.5

32

33

34 **1.0 Introduction**

In 2010, the CalNex field experiment was conducted with two main ground sites to 35 36 investigate the intersection of climate and air quality in Southern and Central California (California Nexus, http://www.arb.ca.gov/research/calnex2010/calnex2010.htm). Bakersfield, 37 CA was chosen to improve understanding of the sources and atmospheric chemistry in the 38 Central Valley with a focus on particulate matter (PM). Located at the southern end of the 39 Central Valley, Bakersfield is bordered by coastal ranges in the west and by the Sierra Nevada 40 Mountains in the east. These topographic features can combine with shallow boundary layers to 41 trap and accumulate pollutants in the valley (Chow et al., 2006; Ahlm et al., 2012). $PM_{2.5}$ (PM \leq 42 2.5 μ m in aerodynamic diameter) concentrations in this region can reach 50 μ g m⁻³ in the winter 43 (well above the annual NAAQS standard of 12 μ g m⁻³ and the 24-hour standard of 35 μ g m⁻³) of 44 which ammonium nitrate and organic mass are the dominant components (Chow et al., 2006). 45 46 The Bakersfield-Delano metropolitan area has ~800,000 inhabitants and, in terms of known 47 emission sources, it is surrounded by extensive highway systems, agriculture, and industrial 48 areas including crude oil extraction and refinement which could contribute to the ambient PM_{25} 49 concentrations (Chan et al., 2013; Guzman-Morales et al., 2014). These sources can impact both 50 primary and secondary organic aerosols (SOA) concentrations. For a more detailed site description, see Liu et al. (2012), who indicated that motor vehicle exhaust, wood combustion, 51 and meat cooking comprise a significant fraction of the local PM (Liu et al., 2012). In addition to 52 53 local anthropogenic sources, the prevailing wind direction during the CalNex campaign was from the Northwest (Ahlm et al., 2012), which could potentially transport biogenic hydrocarbons 54 to the site (Liu et al., 2012). Thus, CalNex 2010 is ideally located to investigate PM composition 55 at the urban-rural boundary. 56

57 Recent research has targeted semi-volatile organic compounds (SVOC) for their potential importance in the production of SOA (Robinson et al., 2007; Presto et al., 2010; Tkacik et al., 58 2012). Gasoline and diesel-powered motor vehicle exhaust emit characteristic SVOCs and 59 thereby contribute to SOA formation, which include individual compounds and groups of 60 unresolved organic mass. Both composition and vapor pressure are relevant for modeling SOA 61 62 formation. Using one-dimensional gas chromatography mass spectrometry (GCMS), the majority of the SVOC mass cannot be speciated and is reported as an unresolved complex 63 mixture (UCM) (Schauer et al., 1999; Schauer et al., 2002). Usually the term UCM refers to a 64 65 raised hump in the total ion count (TIC) of a chromatogram, which comprises hundreds of unresolved peaks (Gough and Rowland, 1991; Fraser et al., 1997; Wang et al., 2012; White et 66 al., 2013). The presence of UCM was described in earlier studies (Blumer et al., 1970), but 67 Farrington and Quinn were the first to use the term UCM, in their paper describing sediments 68 from Narragansett Bay (Rhode Island, USA) (Farrington and Quinn, 1973; White et al., 2013). 69 70 Though first measured in sediments, UCM has been measured in atmospheric semi-volatile and particulate organic carbon as well (Fraser et al., 1997). Very recent advances in UCM analysis, 71 for example GCxGCMS, have been applied to characterize atmospheric particulate organic 72 73 carbon collected at CalNex sites on select days in Bakersfield and Pasadena, CA. This GCxGCMS technique offers improved resolution of the individual constituents which compose 74 the UCM (Chan et al., 2013). 75

Conventional motor oil (lubricating oil), which consists of mineral base oil and other polymers, has been particularly challenging to resolve using 1-D GCMS techniques. The mineral base oil in this blend consists of petroleum hydrocarbons in the C_{20} to $>C_{40}$ range, and appears as a UCM hump in a chromatogram (Frysinger et al., 2003; Wang et al., 2004).

| 80 | Lubricating oils are known to be a major source of environmental hydrocarbon contamination |
|----|--|
| 81 | (Gough and Rowland, 1990; Gough and Rowland, 1991). In environmental samples, the |
| 82 | presence of UCM is considered to be an indicator for petroleum pollution (White et al., 2013); |
| 83 | other than crude oil and petroleum products, recently photosynthesized carbon or naturally |
| 84 | derived lipids may also contribute to UCM in aerosols and sediments (Venkatesan and Kaplan, |
| 85 | 1982; Laureillard et al., 1997; Silva et al., 2012). In earlier studies, a ratio of the UCM to |
| 86 | resolved components (U:R) in atmospheric organic aerosols, observed during GCMS analysis, |
| 87 | has been used as a qualitative indicator for the degree of anthropogenic contribution to |
| 88 | atmospheric organic carbon (Cox et al., 1982; Simoneit and Mazurek, 1982; Mazurek et al., |
| 89 | 1989; Hildemann et al., 1991). Aerosol samples dominated by vegetation derived organics |
| 90 | showed U:R ratios less than 2.0; whereas urban aerosol samples had U:R ratios of 2.6-25 |
| 91 | (Mazurek et al., 1989; Hildemann et al., 1991). Characterization of UCM in ambient $PM_{2.5}$ is |
| 92 | relevant for understanding potential SOA formation and emission impacts on organic aerosol |
| 93 | concentrations. Because of the potential variability in UCM methods, it is also vital to present |
| 94 | physicochemical properties and method validation using standard reference materials. |
| 95 | To understand the daily trends in UCM composition and concentrations observed at |
| 96 | Bakersfield, CA during the CalNex 2010 campaign, a new method was developed for |
| 97 | deconvolution and quantification of multiple UCM peaks. This method was applied to the full |

98 CalNex Bakersfield campaign (May 19- June 26, 2010) and will be compared with the UCM
99 previously characterized by Chan et al (2013) which used GCxGCMS for the June 23, 2010
100 Bakersfield sample.

101 **2.0 Materials and methods**

102 2.1 Field Sampling

103 The CalNex Bakersfield sampling campaign ran from May 19 to June 26, 2010 (Liu et al., 2012; Chan et al., 2013). The sampling site was located in Kern County, near Bakersfield in 104 the San Joaquin Valley, 35.35°N, 118.97°W (Liu et al., 2012). During that time 38 samples were 105 106 collected on a midnight to 11 pm schedule local time (23 h, PDT). PM_{2.5} samples were collected on 86 cm² quartz fiber filters (Pall Life Sciences, NY, USA), with a medium-volume sampler 107 (Tisch, OH, USA; flow rate of 226 L/min). The filters used for sampling were preconditioned by 108 baking at 550 °C for 6 h to remove any potential organic contaminants. Field blanks were 109 110 collected approximately every 10 days by placing a preconditioned filter on the sampler for ~ 15 min. The field blanks were handled and stored in the same manner as the sampled filters. During 111 the sampling period, the temperature and relative humidity were consistent with averages of $24 \pm$ 112 7 °C and $38 \pm 17\%$, respectively (Liu et al., 2012). 113

114 **2.2 Chemicals**

Phosphate-free biodegradable detergent (Liqui-Nox by Alconox, NY, USA) was used for 115 116 washing the glassware and extraction cells (34 mL). The cleaned and baked glassware and extraction cells were rinsed with methanol followed by dichloromethane (DCM; \geq 99.8% purity, 117 J.T. Baker, NJ, USA) immediately prior to usage. The samples were spiked with a known 118 quantity of an isotopically-labelled internal standard, IS#6 (Wisconsin State Laboratory of 119 Hygiene, WI, USA) before extraction. The IS#6 standard includes: Pyrene-d₁₀, 120 Benz[a]anthracene-d₁₂, Coronene-d₁₂, Cholestane-d₄, n-C₁₅-d₃₂, n-C₂₀-d₄₂, n-C₂₄-d₅₀, n-C₃₀-d₅₈, n-121 C₃₂-d₆₆ and n-C₃₆-d₇₄. A standard mixture (PMSTD#12, Wisconsin State Laboratory of Hygiene, 122

WI, USA) was used for quantification during GC-MS analysis. The PMSTD#12 includes nalkanes from C_8 to C_{40} , cyclic and branched alkanes.

125 **2.3 Extraction and Analysis**

126 **2.3.1 Pressurized Liquid Extraction**.

127 The sampled filters were extracted using the Accelerated Solvent Extraction system (ASE; ASE 350, Thermo Scientific Dionex, CA, USA). The method developed for 128 129 deconvoluting and quantifying the UCM was tested using the National Institute of Standards and 130 Technology standard reference material for urban dust (NIST SRM 1649b, Gaithersburg, MD, 131 USA). Each batch of CalNex filters (batch = 10), included one SRM (0.008 - 0.01 g) and one lab blank. Cells were preconditioned using the following ASE parameters: mixture of 1:1 132 methanol:DCM, temperature at 100 °C, 5 min heat time, 5 min static time, 120% rinse volume, 133 134 three cycles and purged with Nitrogen for 100 sec. The sampled filter was then placed inside the 135 conditioned ASE cell and spiked with IS#6 and allowed to come to equilibrium for 60 min. Filter 136 samples were extracted using the same protocol as the preconditioning except a sequential 137 extraction of methanol followed DCM was used. The methanol and DCM extracts were combined and then concentrated under a gentle stream of nitrogen using a Turbovap to a volume 138 139 of 1 ml (Zymark, Hopkinton, MA). Samples were further concentrated using a Techne Sample concentrator (Bibby Scientific Limited, United Kingdom). The final volume of the extract was 140 141 150 µl.

142 2.3.2 Gas Chromatography-Mass Spectrometery analysis.

The final extract was analyzed using an Agilent Technologies 7890A gas chromatograph
with a DB-5MS column (30 m x 0.25 mm x 0.25 μm), coupled with an Agilent Technologies

145 5975C mass spectrometer equipped with an electron ionization source (GC-MS). The following parameters were used: the injector and GC-MS interface were held at a temperature of 300 °C, 146 initial oven temperature was 65 °C for 10 min and then ramped at 10 °C min⁻¹ to a final oven 147 148 temperature of 300 °C for 26.5 min. Ultra-high purity helium was used as the carrier gas with a flow rate of 1 ml min⁻¹, purge flow rate of 66.7 ml min⁻¹ and a total flow rate of 70.1 ml min⁻¹. A 149 solvent delay of 5 min was used prior to data collection by full scan with a range of 50 to 650 150 m/z. For each batch of samples (5-12 samples per batch), several quality control measures were 151 included: a five point calibration curve, solvent blanks and check standards. Compounds were 152 153 identified based on retention time and target ions; after identification the peaks were quantified against the internal standard. Targeted analysis of additional organic tracers will be reported in a 154 forthcoming manuscript. 155

156 2.4 Unresolved Complex Mixture quantification and method development

In general, UCM represents a large area of unresolved compounds; it does not include 157 158 areas associated with specific peaks nor does it include areas associated with a solvent or instrument blank. UCM has been quantified by dividing the unresolved area into sub-areas 159 which elute between two consecutive n-alkanes, C_n and C_{n+1} (Hildemann et al., 1991). The 160 relation between C_n area counts and C_n mass was used to quantify the sub-region between the 161 two n-alkanes. UCM area has also been quantified using a response factor of the most central n-162 alkane within the UCM area (St'avova et al., 2012). In another method, the area of the UCM was 163 determined using the average response factor of n-alkanes (Wang et al., 2012). 164 The Hildemann et al. method was modified for the quantification of UCM used in the 165

166 current study (Hildemann et al., 1991). Modifications in the method were included to allow for167 deconvolution of multiple UCM peaks and for improved blank correction (see Section 2.3.1).

Ambient UCM concentrations are reported for 33 filters collected during the Bakersfield campaign, which includes the filter collected on June 23, 2010 for comparison to the work conducted by Chan et al (2013). The filter from May 25, 2010, which has roughly equal concentrations of both UCM-A and –B, has been used as an example in the UCM protocol description (Figures 1 and 2).

173 **2.4.1 Blank correction**

The total ion chromatograph, TIC, for the Bakersfield samples included three distinct 174 areas of UCM centered at 22, 29, and 34 min (Figure 1). The UCM centered at 29 min was also 175 176 present in the field blank, but neither the solvent nor instrument blanks (Figure 1). The average spectrum from the field blank and the sampled filters were comparable for the area of UCM 177 centered at 29 min (Figure S1 in the supplemental materials). Therefore, based on retention time 178 179 and average spectrum, the UCM that peaked at 29 min was identified as field blank contamination. All samples were corrected for this field blank contamination using the 180 181 deconvolution method described below.

182 **2.4.2** Chromatogram smoothing and peak deconvolution

The raw data was exported from Agilent Chemstation Data Analysis software as 3D data 183 184 to Matlab (Mathworks, Natick, Massachusetts, USA) prior to the UCM quantification. Prior to performing peak deconvolution, individual peaks and solvent blank contributions were first 185 removed. The first step was subtraction of a solvent blank chromatogram from the sample 186 187 chromatogram using Matlab. For the purposes of this paper, the process of plotting a new chromatogram with only the identified minimum values after solvent blank subtraction will be 188 189 called smoothing of the chromatogram. Details of the MatLab method are included in the 190 Supplemental Materials.

191 The smoothed UCM areas centered at 22, 29 and 34 min overlap and require 192 deconvolution, including field blank correction, prior to quantification (Figure 2). The area centered at 22 min is referred to as UCM-A and the area centered at 34 min is referred to as 193 UCM-B. Deconvolution was performed by using an iterative method, in which parameters of a 194 proposed model were adjusted systematically until the model fit the data. The UCM peaks were 195 assumed to be Gaussian and were quantified using overlapping Gaussian functions. This enabled 196 separate quantification of the UCM-A, UCM-B and the field blank UCM. The field blank UCM 197 area was composed of two overlapping Gaussian curves. The smaller curve of the field blank 198 199 contaminant co-eluted with the UCM-B area in the samples. Therefore, for the field blank correction, the value of the co-eluting Gaussian field blank curve was subtracted from the UCM-200 B curve. 201

The three deconvoluted peaks were then quantified (MatLab). The deconvoluted UCM peaks were divided into sub-regions using the included n-alkanes. For example if C_n , C_{n+1} , and C_{n+2} eluted at times t_n , t_{n+1} , and t_{n+2} , then the response factor for C_{n+1} was used to quantify the deconvoluted UCM that eluted between $(t_n+t_{n+1})/2$ and $(t_{n+1}+t_n+2)/2$. The quantified sub-regions were then summed to give the deconvoluted UCM.

207 **2.5 Method test**

The method developed for deconvoluting and quantifying the UCM was tested using NIST SRM 1649b. Figure 1 shows the smoothed chromatogram from one of the SRM extract runs with a UCM peak at 32 min. The UCM at 32 min aligns with the UCM B from the Bakersfield campaign. The Gaussian curve fitting is easily modified to fit the UCM distribution in the sample, i.e. number of curves required to match the chromatogram (Figure 2). The nalkanes C_{21} to C_{31} were used to quantify the UCM, which corresponds with UCM-B from the samples. The SRM UCM had an average value of $201 \pm 33 \ \mu g$ UCM g⁻¹ (± standard deviation).

215

216 **3.0 Results and Discussion**

217 The deconvolution and quantification method was used to quantify the UCM present in the Bakersfield samples collected during the CalNex 2010 field campaign. Figure 3 displays the 218 concentration of UCM-A and UCM-B in ng m⁻³ for the entire campaign. The n-alkanes for the 219 campaign that eluted at the same time as the UCM-A, and UCM-B ranged from C_{13} to C_{19} and 220 C₂₂ to C₃₂, respectively (Table 1). Table 1 also provides a detailed assessment of May 25 (a 221 sample of equal UCM-A and UCM-B) and June 23 (for comparability between this study and a 222 223 previous CalNex study) (Chan et al., 2013). The values reported in Table 1 for the Bakersfield samples and the NIST SRM provide a means of intercomparison across GCMS methods and 224 225 sample campaigns. It is via these physicochemical properties and proxy compounds (i.e. vapor 226 pressure; n-alkane at peak height and corresponding n-alkane range) that the Bakersfield UCM analysis reported here will be compared the UCM analysis conducted by Chan et al. (2013). 227 228 Chan et al. (2013) identified a single UCM peak which corresponds most closely to UCM-B using GCxGCMS and thermal desorption for sample introduction to the instrument (see section 229 3.2). Unlike the Bakersfield samples which had a bi-modal UCM, the SRM only contained a 230 single, later eluting, UCM peak (Figure 1). 231

Over the course of the campaign, the ambient concentration of the UCM-A had a much higher relative standard deviation (90.8 \pm 100.5 ng m⁻³) than that of UCM-B (74.9 \pm 37.9 ng m⁻³). Differences in gas-particle partitioning and vapor pressure in combination with variable emission rates could be impacting the particulate collection of UCM-A and UCM-B. The combined UCMs accounted for 4 to 92 ng μ g⁻¹ OC during the campaign, or 2.9 ± 2.2% of the daily OC. There were six days where the UCM contributed over 5% of the OC, however the daily UCM A and/or B did not correlate with the daily OC. Figure S2 in the Supplemental Materials displays the daily trends in the OC normalized ambient concentration for UCM-A and UCM-B.

Guzman-Morales et al reported three fossil fuel combustion (FFC) factors impacting 241 Bakersfield which had alkane functional groups and different degrees of oxidation (Guzman-242 Morales et al., 2014): an FFC source factor for Bakersfield which was identified as low 243 oxidation, high alkane primary/first generation secondary organic matter tentatively identified as 244 petroleum operation SOA, a more oxidized FFC which was identified as alkane SOA and a 245 nighttime FFC factor which had similar characteristics to diesel fuel. However, all these FFC 246 included high alkane functional group contribution and comprised more than half of the ambient 247 248 organic matter concentrations. It is likely that the UCM measured in the current study is represented by one of more of these FFC source factors. 249

250 3.1 Characterization of Unresolved Complex Mixture-A (UCM-A)

The normal distribution for UCM-A for the campaign peaked at ~21.7 min near the retention time for n-hexadecane; the n-alkanes that eluted within a single standard deviation from this mean ranged from C_{13} to C_{19} (Table 1). The presence of this range of molecular weight hydrocarbon in the samples indicates the potential sorption of SVOC onto the collected particulate matter. This lower molecular weight SVOC is often considered a sampling artifact, because it is potentially affected by ambient temperature and filter loading. Therefore, the concentration of this UCM fraction may be semi-quantitative with regards to the total atmospheric concentration. Figure 1 compares the chromatogram for the May 25 Bakersfield sample and SRM 1649b. UCM-A is
only present in the Bakersfield samples, potentially indicating that this is representative of a nonurban source impacting Bakersfield (i.e. crude oil extraction and refinement). Schauer et al
reported separate, semi-volatile and particle-phase UCM emissions from medium duty diesel
trucks collected on denuder-filter-polyurethane foam sampling trains (Schauer et al., 1999). The
Bakersfield sampling did not include a denuder prior to the filter, which potentially allowed this
semi-volatile UCM to be collected by the quartz fiber filters.

Figure 4 contains day-of-the-week (DoTW) averages and standard deviations for UCM-A 265 and UCM-B (n = 5 full weeks). As discussed previously, there is more variability in the daily 266 267 UCM-A concentrations, which translates into more variability in the UCM-A DoTW 268 concentrations among days and also within each DoTW average. UCM-A demonstrates a steady increase in concentration towards Friday, and a drop on Saturday. This is similar to national 269 270 trends seen for select particulate pollutants including elemental carbon, organic carbon and crustal elements (Murphy et al., 2008). Previous FTIR analysis of the Bakersfield PM_{2.5} samples 271 showed the presence of Si, Al, Ca, and Mg, suggesting a significant presence of organic 272 273 components associated with dust (Liu et al., 2012). This trend in UCM-A may indicate input 274 from medium to heavy duty diesel emissions associated with strong weekend reductions in emissions in California (Marr et al., 2002; Harley et al., 2005). 275

Figure 5a displays an average spectrum for UCM-A; from the major identified qualifier ions, the presence of branched alkanes (57, 71, and 85), alkenes (55, 83, and 97), and alkynes (55, 67, and 81) can be observed (based on the NIST Mass Spectral Library). A study conducted by Schauer et al (1999) reports diesel exhaust composition including n-alkanes from C_{12} to C_{29} , branched and cyclic alkanes. Laboratory-based spectral analysis of oil pump vapors by a thermal desorption aerosol gas chromatograph – aerosol mass spectrometer (TAG-AMS) did
show low oxygen to carbon ratio in the UCM (Williams et al., 2014). However, it cannot be
assumed that ambient UCM is only composed of hydrocarbons, as oxidation of branched and
cyclic alkanes is likely during aging of the aerosol.

285 **3.2 Characterization of Unresolved Complex Mixture-B (UCM-B)**

The UCM-B consisted of heavier compounds with low vapor pressures as compared to 286 UCM-A (i.e. later eluting compounds). The normal distribution of UCM-B for the campaign 287 peaked at 31.9 min, the retention time of n-heptacosane (Table 1). The n-alkanes C_{22} to C_{32} 288 eluted within a single standard deviation from the mean and the vapor pressure ranged from 2E-289 290 10 to 6E-7 torr. Chan et al. (2013) reported UCM with a similar range of n-alkanes (20-25 291 carbon atoms) and vapor pressure (9E-8 to 6E-6 torr) for two urban sites in California: Pasadena 292 and Bakersfield. Comparison with the SRM 1649b chromatogram reveals that the Bakersfield UMC-B is similar in retention time range and peak shape of the SRM UCM (Figure 1 and Table 293 294 1); this potentially indicates that UCM-B traces the urban component of the Bakersfield UCM (i.e. traffic emissions). Previous studies in the San Joaquin Valley have reported UCM 295 296 associated with paved road dust (Rogge et al., 2012). However, a qualitative assessment of the relative retention times for select compounds displayed in chromatograms reported by Rogge et 297 298 al. (2012) for urban and rural paved roads indicate that the road dust UCM peaks have a later retention time maxima than the Bakersfield both the UCM-A and UCM-B reported here. 299

UCM-B concentrations had no discernable day-of-the-week trend (Figure 4). It has been
reported that passenger car traffic has no day-of-the-week cycles in California (Marr et al., 2002;
Harley et al., 2005) nor do national electric power generation emissions (Chinkin et al., 2003;

303 Motallebi et al., 2003). This consistent contribution combined with the similarity to the urban 304 dust SRM UCM indicates that the UCM-B tracks the urban contribution, including traffic. Figure 5b displays the average spectrum of UCM-B; it is dominated by branched alkanes 305 (57, 71, and 85 ions, based on the NIST Mass Spectral Library), and cyclic alkanes (55, 69, 83, 306 and 97 ions)). This spectral average is similar to compounds identified in Bakersfield UCM by 307 308 Chan et al. (2013). The retention time and composition UCM-B are similar to UCM found in samples which have had crude oil or motor oil contributions (White et al., 2013). Both Chan et 309 al and Guzman-Morales implicate local petroleum operations for the UCM and low oxidation 310 311 hydrocarbon, respectively, reported in each publication for the Bakersfield campaign (Chan et al., 2013; Guzman-Morales et al., 2014). However, based on the high-molecular weight n-alkane 312 concentrations, the carbon preference indices (CPI) for this study averaged above 2. Chan et al. 313 also reported CPIs consistently above 2 for Bakersfield, suggesting contribution from plant wax 314 or vegetative detritus (Chan et al., 2013) and Guzman-Morales also reported a vegetative detritus 315 factor (Guzman-Morales et al., 2014). This suggests the possibility of mixed biogenic and 316 317 anthropogenic emission sources impacting the low vapor pressure UCM at Bakersfield. 318

319

3.3 Unresolved to Resolved ratio (U:R)

In earlier studies a simple ratio of the concentrations of Unresolved to Resolved alkanes (U:R) was utilized to semi-quantitatively assess biogenic and anthropogenic contributions (Hildemann et al., 1991). A U:R greater than 2.0 indicates a predominance of anthropogenic sources, while a U:R ratio less than 2.0 indicates biogenic sources including vegetative detritus. In Figure 3, the U:R ratios are displayed for each day with a line of demarcation at the U:R ratio of 2.0. The values of the U:R ratio ranged from 0.8 to 16 with a mean value of 6.0 for the 33

| 326 | days analyzed. The variability in the U:R ratios suggest that the relative contributions of |
|-----|---|
| 327 | anthropogenic and biogenic PM sources fluctuates at least on a daily time scale, likely less. |
| 328 | Based on their U:R ratio (for A+B), the 33 daily samples were separated into two bins: |
| 329 | anthropogenic-dominated days (26 at U:R $>$ 2.0), and biogenic-dominated days (7 at U:R $<$ 2.0). |
| 330 | The Figure 3 inset displays the mean values of UCM-A and UCM-B for each bin of sample days. |
| 331 | This analysis reveals that the anthropogenic-dominated days have overall higher concentrations |
| 332 | of UCM-A and UCM-B than the biogenic-dominated days. As with the daily trends, UCM-A has |
| 333 | a much larger difference between bins with very little contribution towards the biogenic- |
| 334 | dominated days. Additional investigations are needed to clarify the strengths of the |
| 335 | anthropogenic and biogenic contributions on a daily basis. |

336

4.0 Conclusions

The deconvolution and quantification protocol developed herein was successfully applied to daily CalNex 2010 Bakersfield samples to quantify two UCM peaks (A & B). The separate quantification of two UCM peaks within a single sample, which were then further characterized by vapor pressure and ranges of representative n-alkanes, is a unique contribution to UCM characterization in atmospheric aerosols. This new deconvolution method enables quick, reproducible, quantification of multiple components of the UCM as analyzed by GCMS which is an improvement over previously reported manual quantification of UCM (Schauer et al., 1999).

By reporting the daily concentration and also simple physical characteristics, the UCM measured at Bakersfield in the current study can be compared to UCM from disparate campaigns using disparate analytical techniques. Likewise, the inclusion of the UCM characterization of the urban dust SRM can be used to facilitate intercomparison across these same studies. By
including the vapor pressure range and representative compounds, these UCM concentrations
can also be included more readily in photochemical models.

For the CalNex Bakersfield campaign, the UCM-A was comprised of semi-volatile 351 compounds including alkanes, alkenes, and alkynes. Based on the mass spectrum, daily trends 352 and comparison with published emission inventory studies, the major source was assumed to be 353 medium to heavy-duty diesel exhaust. UCM-B, which eluted later, was comprised of heavier 354 compounds, and based on their average spectrum, its composition was most likely to be linear, 355 branched, and cyclic alkanes. Both the vapor pressure range and composition of branched and 356 cyclic alkanes are similar to those reported by Chan et al (2013) in a detailed description of the 357 358 Jun 23, 2010 sample. UCM-B showed indicators for heavy motor oil, as well as vegetative detritus (based on the U:R and the CPI). On comparison with other literature, the SRM 1649b 359 360 and U:R ratios, possible sources of UCM-B were identified to be urban sources, including motor 361 vehicle exhaust, combined with agricultural activities.

362

363 Acknowledgements

The U.S. Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research described here under Contract EP-D-10-070 to Alion Science and Technology. The manuscript has not been subjected to external peer review and has not been cleared for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. This study was supported in part by Baylor

- 369 University funds including the University Research Committee and the C. Gus Glasscock, Jr.
- 370 Endowed Fund for Excellence in Environmental Sciences.

371

372 **References**

373 Ahlm, L., Liu S., Day D.A., Russell L.M., Weber R., Gentner D.R., Goldstein A.H., DiGangi J.P., et al., 374 2012. Formation and growth of ultrafine particles from secondary sources in Bakersfield, California. Journal of Geophysical Research-Atmospheres 117 DOI: 10.1029/2011JD017144. 375 376 Blumer, M., Souza G. and Sass J., 1970. Hydrocarbon Pollution of Edible Shellfish by an Oil Spill. 377 Marine Biology 5, 195-202. 378 Chan, A.W.H., Isaacman G., Wilson K.R., Worton D.R., Ruehl C.R., Nah T., Gentner D.R., Dallmann 379 T.R., et al., 2013. Detailed chemical characterization of unresolved complex mixtures in 380 atmospheric organics: Insights into emission sources, atmospheric processing, and secondary 381 organic aerosol formation. Journal of Geophysical Research: Atmospheres 118, 6783-6796 382 10.1002/jgrd.50533. 383 Chickos, J.S. and Hanshaw W., 2004. Vapor Pressures and Vaporization Enthalpies of the n-Alkanes 384 from C31 to C38 at T = 298.15 K by Correlation Gas Chromatography. Journal of Chemical & Engineering Data 49, 620-630 10.1021/je030236t. 385 386 Chinkin, L.R., Coe D.L., Funk T.H., Hafner H.R., Roberts P.T., Ryan P.A. and Lawson D.R., 2003. Weekday versus weekend activity patterns for ozone precursor emissions in California's South 387 Coast Air Basin. Journal of the Air & Waste Management Association 53, 829-843. 388 389 Chow, J.C., Chen L.W.A., Watson J.G., Lowenthal D.H., Magliano K.A., Turkiewicz K. and Lehrman D.E., 2006. PM2.5 chemical composition and spatiotemporal variability during the California 390 391 Regional PM10/PM2.5 Air Quality Study (CRPAQS). Journal of Geophysical Research-392 Atmospheres 111 10.1029/2005JD006457. 393 Cox, R.E., Mazurek M.A. and Simoneit B.R.T., 1982. Lipids in Harmattan Aerosols of Nigeria. Nature 394 296, 848-849. Farrington, J.W. and Quinn J.G., 1973. Petroleum hydrocarbons in Narragansett Bay: I. Survey of 395 396 hydrocarbons in sediments and clams (Mercenaria mercenaria). Estuarine and Coastal Marine 397 Science 1, 71-79 10.1016/0302-3524(73)90059-5. 398 Fraser, M.P., Cass G.R., Simoneit B.R.T. and Rasmussen R.A., 1997. Air quality model evaluation data 399 for organics .4. C-2-C-36 non-aromatic hydrocarbons. Environmental Science & Technology 31, 400 2356-2367. 401 Frysinger, G.S., Gaines R.B., Xu L. and Reddy C.M., 2003. Resolving the unresolved complex mixture in 402 petroleum-contaminated sediments. Environmental Science & Technology 37, 1653-1662. 403 Gough, M. and Rowland S., 1991. Characterization of Unresolved Complex-Mixtures of Hydrocarbons 404 from Lubricating Oil Feedstocks. Energy & Fuels 5, 869-874. 405 Gough, M.A. and Rowland S.J., 1990. Characterization of Unresolved Complex-Mixtures of Hydrocarbons in Petroleum. Nature 344, 648-650. 406 Goss, K.-U. and Schwarzenbach R.P., 1999. Empirical Prediction of Heats of Vaporization and Heats of 407 408 Adsorption of Organic Compounds. Environmental Science & Technology 33, 3390-3393 409 10.1021/es980812j. 410 Guzman-Morales, J., Frossard A.A., Corrigan A.L., Russell L.M., Liu S., Takahama S., Taylor J.W., Allan J., et al., 2014. Estimated contributions of primary and secondary organic aerosol from 411 412 fossil fuel combustion during the CalNex and Cal-Mex campaigns. Atmospheric Environment 88, 413 330-340 10.1016/j.atmosenv.2013.08.047. Harley, R.A., Marr L.C., Lehner J.K. and Giddings S.N., 2005. Changes in motor vehicle emissions on 414 diurnal to decadal time scales and effects on atmospheric composition. Environmental Science & 415 416 Technology 39, 5356-5362. 417 Hildemann, L.M., Mazurek M.A., Cass G.R. and Simoneit B.R., 1991. Quantitative characterization of 418 urban sources of organic aerosol by high-resolution gas chromatography. Environmental Science 419 & Technology 25, 1311-1325.

- Laureillard, J., Pinturier L., Fillaux J. and Saliot A., 1997. Organic geochemistry of marine sediments of 420 the Subantarctic Indian Ocean sector: Lipid classes - Sources and fate. Deep-Sea Research Part Ii-421 422 Topical Studies in Oceanography 44, 1085-1108. 423 Liu, S., Ahlm L., Day D.A., Russell L.M., Zhao Y., Gentner D.R., Weber R.J., Goldstein A.H., et al., 2012. Secondary organic aerosol formation from fossil fuel sources contribute majority of 424 425 summertime organic mass at Bakersfield. Journal of Geophysical Research-Atmospheres 117 426 10.1029/2012JD018170. 427 Marr, L.C., Black D.R. and Harley R.A., 2002. Formation of photochemical air pollution in central 428 California - 1. Development of a revised motor vehicle emission inventory. Journal of 429 Geophysical Research-Atmospheres 107 10.1029/2001jd000689. 430 Mazurek, M.A., Cass G.R. and Simoneit B.R., 1989. Interpretation of high-resolution gas 431 chromatography and high-resolution gas chromatography/mass spectrometry data acquired from 432 atmospheric organic aerosol samples. Aerosol Science and Technology 10, 408-420. Motallebi, N., Tran H., Croes B.E. and Larsen L.C., 2003. Day-of-week patterns of particulate matter and 433 434 its chemical components at selected sites in California. Journal of the Air & Waste Management 435 Association 53, 876-888. Murphy, D., Capps S., Daniel J., Frost G. and White W., 2008. Weekly patterns of aerosol in the United 436
- Murphy, D., Capps S., Daniel J., Frost G. and White W., 2008. Weekly patterns of aerosol in the United
 States. Atmospheric Chemistry and Physics 8, 2729-2739.
- Presto, A.A., Miracolo M.A., Donahue N.M. and Robinson A.L., 2010. Secondary Organic Aerosol
 Formation from High-NOx Photo-Oxidation of Low Volatility Precursors: n-Alkanes.
 Environmental Science & Technology 44, 2029-2034.
- Robinson, A.L., Donahue N.M., Shrivastava M.K., Weitkamp E.A., Sage A.M., Grieshop A.P., Lane
 T.E., Pierce J.R., et al., 2007. Rethinking organic aerosols: Semivolatile emissions and
 photochemical aging. Science 315, 1259-1262 10.1126/science.1133061.
- Rogge, W.F., Medeiros P.M. and Simoneit B.R.T., 2012. Organic Compounds in Dust from Rural and
 Urban Paved and Unpaved Roads Taken During the San Joaquin Valley Fugitive Dust
 Characterization Study. Environmental Engineering Science 29, 1-13.
- Schauer, J.J., Kleeman M.J., Cass G.R. and Simoneit B.R.T., 1999. Measurement of emissions from air
 pollution sources. 2. C₁ through C₃₀ organic compounds from medium duty diesel trucks.
 Environmental Science & Technology 33, 1578-1587.
- Schauer, J.J., Kleeman M.J., Cass G.R. and Simoneit B.R.T., 2002. Measurement of emissions from air
 pollution sources. 5. C-1-C- 32 organic compounds from gasoline-powered motor vehicles.
 Environmental Science & Technology 36, 1169-1180.
- Silva, T.R., Lopes S.R.P., Sporl G., Knoppers B.A. and Azevedo D.A., 2012. Source characterization
 using molecular distribution and stable carbon isotopic composition of n-alkanes in sediment
 cores from the tropical Mundau-Manguaba estuarine-lagoon system, Brazil. Organic
 Geochemistry 53, 25-33.
- 457 Simoneit, B.R. and Mazurek M.A., 1982. Organic matter of the troposphere—II. Natural background of
 458 biogenic lipid matter in aerosols over the rural western United States. Atmospheric Environment
 459 (1967) 16, 2139-2159.
- St'avova, J., Stahl D.C., Seames W.S. and Kubatova A., 2012. Method development for the
 characterization of biofuel intermediate products using gas chromatography with simultaneous
 mass spectrometric and flame ionization detections. Journal of Chromatography A 1224, 79-88.
- 463 Tkacik, D.S., Presto A.A., Donahue N.M. and Robinson A.L., 2012. Secondary Organic Aerosol
 464 Formation from Intermediate-Volatility Organic Compounds: Cyclic, Linear, and Branched
 465 Alkanes. Environmental Science & Technology 46, 8773-8781.
- Venkatesan, M.I. and Kaplan I.R., 1982. Distribution and Transport of Hydrocarbons in Surface
 Sediments of the Alaskan Outer Continental-Shelf. Geochimica Et Cosmochimica Acta 46, 21352149.

- Wang, J.Z., Yang Z.Y. and Chen T.H., 2012. Source apportionment of sediment-associated aliphatic
 hydrocarbon in a eutrophicated shallow lake, China. Environmental Science and Pollution
 Research 19, 4006-4015.
- Wang, Z., Fingas M., Lambert P., Zeng G., Yang C. and Hollebone B., 2004. Characterization and
 identification of the Detroit River mystery oil spill (2002). Journal of Chromatography A 1038,
 201-214 10.1016/j.chroma.2004.03.004.
- White, H.K., Xu L., Hartmann P., Quinn J.G. and Reddy C.M., 2013. Unresolved Complex Mixture (UCM) in Coastal Environments Is Derived from Fossil Sources. Environmental Science & Technology 47, 726-731.
- Williams, B.J., Jayne J.T., Lambe A.T., Hohaus T., Kimmel J.R., Sueper D., Brooks W., Williams L.R.,
 et al., 2014. The First Combined Thermal Desorption Aerosol Gas Chromatograph-Aerosol Mass
- 480 Spectrometer (TAG-AMS). Aerosol Science and Technology 48, 358-370
- 481 10.1080/02786826.2013.875114.

482



Figure 1: Chromatogram of field blank (black), sample collected on May 25, 2010 (red), lab blank (purple) and standard reference material (blue).



Figure 2: Deconvoluted and smoothed chromatogram for May 25, 2010.



Figure 3: Stacked vertical bar chart showing UCM-A and UCM-B, along with the U:R ratio as a black dot for that sample. The horizontal black line represents the U:R ratio value of 2.0. Inset shows the average UCM-A and UCM-B concentration, with standard deviation, for anthropogenic-dominated days and biogenic-dominated sources as defined by U:R ratios.



Figure 4: Day of the week (DotW) averages with standard deviation for the UCM-A, and UCM-B ambient concentrations.



Figure 5 a: The average spectrum of UCM-A from the sample collected on May 25, 2010. b: The average spectrum of UCM-B from the sample collected on May 25, 2010

Table 1: Description of UCM including retention time, vapor pressure, and n-alkanes.

| | UCM-A | | | _ | <u> </u> | | | |
|-----------------------|-------|-------|-------|---|----------|-------|------------|--|
| Campaign (5/22-6-26) | -1σ | μ | +1σ | _ | -1σ | μ | +1σ | |
| Retention time (min) | 17.6 | 21.7 | 25.8 | | 27.7 | 31.9 | 36.2 | |
| Vapor pressure (Torr) | 9E-03 | 2E-04 | 4E-06 | | 6E-07 | 1E-08 | 2E-10 | |
| n-alkanes | C13 | C16 | C19 | | C22 | C27 | C32 | |
| 5/25/2010 | -1σ | μ | +1σ | _ | -1σ | μ | +1σ | |
| Retention time (min) | 18.8 | 21.9 | 25.1 | | 28.5 | 32.3 | 36.2 | |
| Vapor pressure (Torr) | 3E-03 | 1E-04 | 7E-06 | | 3E-07 | 7E-09 | 2E-10 | |
| n-alkanes | C14 | C16 | C19 | | C22 | C27 | C32 | |
| 6/23/2010 | -1σ | μ | +1σ | _ | -1σ | μ | $+1\sigma$ | |
| Retention time (min) | 18.0 | 22.4 | 26.7 | | 27.4 | 31.7 | 36.0 | |
| Vapor pressure (Torr) | 6E-03 | 1E-04 | 2E-06 | | 8E-07 | 1E-08 | 2E-10 | |
| n-alkanes | C13 | C16 | C20 | | C21 | C27 | C32 | |
| NIST SRM 1649b | -1σ | μ | +1σ | _ | -1σ | μ | +1σ | |
| Retention time (min) | - | na | - | | 31.1 | 33.3 | 35.4 | |
| Vapor pressure (Torr) | - | - | - | | 2E-08 | 3E-09 | 4E-10 | |
| n-alkanes | - | - | - | | C27 | C29 | C31 | |

* Vapor pressure were transcribed from previous publications (Goss and Schwarzenbach, 1999; Chickos and Hanshaw, 2004)

References

- Chickos, J.S. and Hanshaw W., 2004. Vapor Pressures and Vaporization Enthalpies of the n-Alkanes from C31 to C38 at T = 298.15 K by Correlation Gas Chromatography. Journal of Chemical & Engineering Data 49, 620-630 10.1021/je030236t.
- Goss, K.-U. and Schwarzenbach R.P., 1999. Empirical Prediction of Heats of Vaporization and Heats of Adsorption of Organic Compounds. Environmental Science & Technology 33, 3390-3393 10.1021/es980812j.

Supplemental Materials for:

Qualitative and quantitative assessment of unresolved complex mixture in PM_{2.5} of

Bakersfield, CA.

Punith Dev Nallathamby¹, Michael Lewandowski², Mohammed Jaoui³, John H. Offenberg², Tadeusz E. Kleindienst², Caitlin Rubitschun⁴, Jason D. Surratt⁴, Sascha Usenko¹ and Rebecca J. Sheesley^{1*}

¹ Department of Environmental Science, Baylor University, Waco, TX

² National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC, USA

³ Alion Science and Technology, Research Triangle Park, NC, USA

⁴ Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, NC, USA

*Corresponding author. Tel.: +1 2547103158; fax: +1 2547103409; email: <u>rebecca_sheesley@baylor.edu</u>, Baylor University, OBP #97266, Waco, TX USA 76798

Details of the MatLab method.

Individual peaks were identified and subtracted using the *findpeaks* function in Matlab. The *findpeaks* function identifies the local maximum and minimum values throughout the chromatogram. The local minima were then used to plot a new, smoothed chromatogram which included only the UCM(s), without any individual peaks (Figure 1).

The peak deconvolution was performed in Matlab by fitting Gaussian curves to the GCMS data. The Gaussian function for one dimension is described by the following equation:

$$Y_{i} = \frac{1}{\sigma\sqrt{2\pi}} e^{-(X_{i}-\mu)^{2}/2\sigma^{2}}$$
(1)

where the height of the curve Y_i at a given point X_i (retention time in min) can be calculated if the mean μ and standard deviation σ are known. For a normal distribution, the mean is the retention time of the curve where the height is at its maximum and the standard deviation is the half width of the curve. The *normpdf* function in Matlab can be used to generate Gaussian curves for a given set of X_i , μ (retention time at maximum height) and σ (half-width) (Equation 1). Given the sample UCM, three Gaussian curves were generated with an initial rough estimate of their position and half width. An individual regression analysis was done between each generated curve Y_i to the smoothed chromatogram, and a linear trendline was generated in each case, forcing the intercept to be zero. The slope of this trendline for each case was then calculated and multiplied with its respective generated curves. The magnitude of the generated curves then approximated the smoothed chromatogram. These new curves (slope x generated curve) are referred to here as redrawn curves. The three redrawn curves were then summed to give the model. The fit of the model was then checked by calculating the fitting error between the proposed model and smoothed chromatogram. By applying the *norm* function in Matlab to the difference between the model and smoothed chromatogram, the fitting error was estimated.

Iterations of this procedure were then performed using the *fminsearch* function in Matlab by varying the starting points of the μ and σ of each generated Gaussian curve until the least fitting error was obtained. Once the least fitting error was obtained, it was checked by overlapping the smoothed chromatogram, generated model, and generated curves. Additional iterations were performed as necessary to assure a close fit. Figure S1 in the Supplemental Materials is a smoothened chromatogram of the May 25, 2010 sample (same as Fig 1), the TIC broken into the deconvoluted peaks. The red dots represent the smoothed chromatogram, and the blue, green, and magenta lines represent the three deconvoluted peaks. The black line represents the fit model, which is the sum of the three deconvoluted peaks.



Figure S1: The average spectrum between the 26 and 31 min of (a) sample collected on May 25, 2010 and (b) field blank.



Sample date

Figure S2: Ambient daily concentration of UCM-A and UCM-B normalized by ambient organic carbon (OC).