SUPPLEMENTARY INFORMATION

Mapping and quantifying isomer sets of hydrocarbons (≥C\textsubscript{12}) in diesel fuel, lubricating oil and diesel exhaust samples using GC×GC-ToFMS

Mohammed S. Alam, Zhirong Liang, Soheil Zeraati-Rezaei, Christopher Stark, Hongming Xu, A. Rob MacKenzie and Roy M. Harrison

CONTENTS

Section S1 – Chromatographic Methodology
Section S2 – Computer language for identification of chemicals (CLIC qualifiers)
Section S3 – Quantification of compounds with no authentic standards
Section S4 – Scaling ion current to molar quantity for <C\textsubscript{25} and mass for >C\textsubscript{25}
Section S5 – Representative Mass Spectra for diesel and gas phase hydrocarbon identification
Section S6 – Unresolved peaks
Figure S7
Section S8 – Representative mass spectra for compound classes identified in the lubricating oil
Figure S9
Section S1: Chromatography Methodology

Exposed adsorption tubes were spiked with 1 ng of deuterated internal standard mix for quantification, and desorbed onto the cold trap at 350 °C for 15 min (trap held at 20 °C). The trap was then purged onto the column in a split ratio of 102:1 at 350 °C and held for 4 min. An initial temperature of 90 °C of the primary oven was held for 2 min and then increased at 2 °C min\(^{-1}\) to 240 °C, followed by 3 °C min\(^{-1}\) to 310 °C and held for 5 min. The initial temperature of the secondary oven of 40 °C was held for 2 min, and then increased to 250 °C by 3 °C min\(^{-1}\), followed by an increase of 1.5 °C min\(^{-1}\) to 315 °C and held for 5 min. The modulation period was 10 s. The complete run time was approximately 105 min.

Diluted lubricating oil samples and particulate phase filters were spiked with 50 µL of 1 ng µL\(^{-1}\) deuterated internal standard mix for quantification. Filters were immersed in dichloromethane (DCM) and ultrasonicated at 20 °C for 20 min. The extract was then concentrated to 50 µL under a gentle flow of N\(_2\) for analysis on the GC\(\times\)GC-ToF-MS. 1 µL of the extracted sample was injected in a split ratio 100:1 at 300 °C. An initial temperature of 120 °C of the primary oven was held for 2 min and then increased at 2 °C min\(^{-1}\) to 210 °C, followed by 1.5 °C min\(^{-1}\) to 325 °C and held for 5 min. The initial temperature of the secondary oven of 120 °C was held for 2 min, and then increased to 200 °C by 3 °C min\(^{-1}\), followed by an increase of 2 °C min\(^{-1}\) to 300 °C and a final increase of 1 °C min\(^{-1}\) to 330 °C, ensuring all species pass through the column efficiently. The modulation period was 13 s.

1 µL of diluted (1:1000) lubricating oil was injected in a split ratio 100:1 at 300 °C. An initial temperature of the primary and secondary ovens were kept the same (175 °C) held for 5 min. The
primary oven temperature was increased by 1°C min\(^{-1}\) to 325°C, while the secondary oven temperature was increased by 1°C min\(^{-1}\) to 330°C. A modulation time of 8 s was used, while a total run time of each sample was 120 min. The transfer line and ion source temperatures were 325 °C and 320 °C, respectively and were kept consistent for all sample analyses. Helium was used as the carrier gas at a constant flow rate of 1 mL min\(^{-1}\).

Section S2: Computer Language for Identification of Chemicals (CLIC qualifiers)

Language functions: See Reference 2 (below) for more details on selection language functions.

**Ordinal** – Returns the ordinal position of the indicated channel (m/z in a mass spectrum) in the intensity-ordered multi-channel array of the current object (blob)

**Retention** – Returns the retention time of the current object (blob) with respect to the chromatographic column indicated by the dimension parameter. Retention time for dimension 1 is expressed in minutes and dimension 2 is expressed in seconds.

**Relative** – Returns the intensity value of the indicated channel (m/z in a mass spectrum) in the multi-channel intensity array of the current object (blob) as a relative percentage of the largest intensity value of the array.

See Reference 2 (below) for more details on selection language functions.

Some examples of CLIC qualifiers are shown below:
C2-Alkyl Benzenes
=(ordinal(91)=1)&(ordinal(106)=2);

C3-Alkyl Benzenes
=(ordinal(91)=1)&(ordinal(120)<=3)&(retention(1)<30))|(ordinal(105)=1)&(ordinal(120)<=3)&(retention(1)<30));

C4-Alkyl
Benzenes=((ordinal(91)=1)&(ordinal(134)<=3))|(ordinal(105)=1)&(ordinal(134)<=3)&(retention(1)<80))|(ordinal(119)=1)&(ordinal(134)<=3));

C5-Alkyl
Benzenes=((ordinal(92)=1)&(ordinal(148)<=3))|(ordinal(105)=1)&(ordinal(148)<=4))|(ordinal(119)=1)&(ordinal(148)<=4))|(ordinal(133)=1)&(ordinal(148)<=3));

C6-Alkyl
Benzenes=((ordinal(92)=1)&(ordinal(162)<=4))|(ordinal(105)=1)&(ordinal(162)<=5))|(ordinal(119)=1)&(ordinal(162)<=5))|(ordinal(133)=1)&(ordinal(162)<=5))|(ordinal(147)<=2)&&(ordinal(162)<=5)&(retention(2)<6))|(ordinal(106)=1)&(ordinal(162)<=4))|(ordinal(105)=1)&(ordinal(162)<=4));

C7-Alkyl
Benzenes=((ordinal(105)=1)&(ordinal(176)<=4))|((ordinal(133)=1)&(ordinal(176)<=4))|((ordinal(161)=1)&(ordinal(176)<=4)&(Retention(2)<5.5))|((ordinal(91)=1)&(ordinal(176)<=3))|((ordinal(147)=1)&(ordinal(119)<=4)&(ordinal(176)<=4))|((ordinal(106)=1)&(ordinal(176)<=4))|((ordinal(119)=1)&(ordinal(176)<=5));

C1-Alkyl NAP=(ordinal(141)<=3)&(ordinal(115)<=3)&(ordinal(142)<=3);

C2-Alkyl
NAP=((ordinal(141)=1)&(ordinal(156)=2)&(ordinal(115)=3)&(relative(141)>90)&(relative(156)>50))|(ordinal(156)=1)&(ordinal(141)=2)&(ordinal(155)=3)&(relative(156)>90)&(relative(141)>50));

C3-Alkyl NAP
=|(ordinal(155)=1)&(ordinal(170)=2)&(Relative(155)>80)&(Relative(170)>25))|((ordinal(170)=1)&(ordinal(155)=2)&(Relative(170)>80)&(Relative(155)>60))|((ordinal(141)=1)&(ordinal(170)=2)&(ordinal(115)=3)&(Relative(141)>80)&(Relative(170)>25)&(Relative(115)>15));

C4-Alkyl
NAP=((ordinal(155)=1)&(ordinal(184)=2)&(Relative(155)>80)&(Relative(184)>15))|((ordinal(141)=1)&(ordinal(184)=2)&(Relative(141)>80)&(Relative(184)>15))|((ordinal(169)=1)&(ordinal(184)=2)&(Relative(169)>80)&(Relative(184)>30))|((ordinal(184)=1)&(ordinal(169)=2)&(Relative(184)>80)&(Relative(169)>50));
Section S3: Quantification of compounds with no authentic standards

The authentic standard mixture contains 74 standard compounds including 9 internal standards (see Section 2.3 for more details). These standards were chosen in order to cover as much of the overall chromatogram as possible and are used for obtaining a calibration for quantifying groups of isomers with the same molecular ion and functionality. For example; the response for C\textsubscript{11} (undecane, m/z 156) was used to quantify all C\textsubscript{11} alkane isomers which were positively identified in the analysed samples (and have the same molecular mass and retention times in all sample runs). Using retention times for C\textsubscript{11} alkane isomers as well as mass spectra, only those C\textsubscript{11} isomer peaks were selected that were observed using the SIC for m/z 156 (or the respective CLIC expressions). Isomer sets were comprehensively identified in the analysed samples using mass fragmentation at 70eV and molecular masses at lower ionisation energies (10-15eV); the methodology of identifying individual isomers has been published elsewhere (Alam et al. 2016). The total ion current within each polygon (C\textsubscript{11} isomer sets in this case) was then integrated and the isomer set abundance was estimated using the response ratio of the closest structurally related deuterated standard (dodecane-d\textsubscript{26}) to the corresponding compound class natural standard with the same carbon number (C\textsubscript{11} in this case).

To assess the uncertainty of this method, we estimated known concentrations of compounds for which authentic standards were available. Table S3-1 shows the difference between concentrations estimated with the generic standard and the authentic standard. The overall uncertainty is difficult to estimate. This depends upon both the uncertainty associated with quantifying an individual compound (U), and the number of compounds in a polygon (n). Then

\[ U \text{ (polygon)} = (U_1^2 + U_2^2 + U_3^2 \ldots \ldots U_n^2)^{1/2} \]
where $U_1$, $U_2$ …… $U_n$ are the uncertainties associated with individual compounds. There were three polygons for which calibration standards were available for all compounds. These gave collective uncertainties calculated as above of 13.9%, 18.9% and 39.3% (mean = 24%). Although statistically rigorous, we feel that this overestimates the uncertainty as the mass closure figures for the samples as a whole appear realistic and none deviates appreciably from 75-100%, including samples not described in this paper.

Table S3-1. Comparison of true calibrated concentrations and estimated concentrations using this methodology.

<table>
<thead>
<tr>
<th>Compound</th>
<th>True Calibrated Concentrations in sample (ng/µL)</th>
<th>Concentration using TIC and n-alkane calibration (ng/µL)</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pristane</td>
<td>9.9</td>
<td>9.3</td>
<td>-6%</td>
</tr>
<tr>
<td>Phytane</td>
<td>10.0</td>
<td>10.4</td>
<td>4%</td>
</tr>
<tr>
<td>Cyclohexane, pentyl-</td>
<td>3.1</td>
<td>3.3</td>
<td>6%</td>
</tr>
<tr>
<td>Cyclohexane, hexyl-</td>
<td>15.2</td>
<td>18.1</td>
<td>19%</td>
</tr>
<tr>
<td>Cyclohexane, heptyl-</td>
<td>14.0</td>
<td>16.5</td>
<td>18%</td>
</tr>
<tr>
<td>Cyclohexane, octyl-</td>
<td>11.1</td>
<td>12.5</td>
<td>13%</td>
</tr>
<tr>
<td>Cyclohexane, nonyl-</td>
<td>8.9</td>
<td>10.0</td>
<td>12%</td>
</tr>
<tr>
<td>Cyclohexane, decyl-</td>
<td>10.0</td>
<td>9.5</td>
<td>-5%</td>
</tr>
<tr>
<td>Cyclohexane, undecyl-</td>
<td>3.8</td>
<td>4.0</td>
<td>5%</td>
</tr>
<tr>
<td>Cyclohexane, dodecyl-</td>
<td>4.1</td>
<td>3.8</td>
<td>-7%</td>
</tr>
<tr>
<td>Cyclohexane, tridecyl-</td>
<td>3.7</td>
<td>3.7</td>
<td>0%</td>
</tr>
<tr>
<td>Cyclohexane, tetradecyl-</td>
<td>3.1</td>
<td>2.5</td>
<td>-19%</td>
</tr>
<tr>
<td>Cyclohexane, pentadecyl-</td>
<td>3.0</td>
<td>2.8</td>
<td>-7%</td>
</tr>
<tr>
<td>naphthalene, 1-methyl-</td>
<td>20.0</td>
<td>24.0</td>
<td>20%</td>
</tr>
<tr>
<td>naphthalene, 1-ethyl-</td>
<td>15.0</td>
<td>13.0</td>
<td>-13%</td>
</tr>
<tr>
<td>naphthalene, 1-propyl-</td>
<td>12.0</td>
<td>10.0</td>
<td>-17%</td>
</tr>
<tr>
<td>naphthalene , 1-hexyl-</td>
<td>32.0</td>
<td>35.0</td>
<td>9%</td>
</tr>
</tbody>
</table>
Section S4: Scaling ion current to molar quantity for \(<C_{25}\) and mass for \(>C_{25}\)

Previous work shows that the total ion current (TIC) signal can be scaled to molar quantity (depending on saturation and effective vapour pressure). The TIC intensity of the individual authentic n-alkane standards \((C_{12}-C_{34})\), with seven injected concentrations covering 0.5ng/µL, 1ng/µL, 1.5ng/µL, 2ng/µL, 2.5ng/µL, 4ng/µL and 5ng/µL were investigated. Below \(C_{25}\), the TIC intensity of alkanes increased linearly as the molar quantity (Figure S1, red curves). The TIC molar quantity of alkyl-cyclohexane \((C_{12}-C_{34})\) are also plotted here to compare with the curves of n-alkanes (Figure A, black curves). The curves of the n-alkyl-cyclohexanes overlap with those of the n-alkanes in the same injection, illustrating that n-alkyl-cyclohexane and n-alkanes have a very similar response per mole. This is also found to be the same for other species with increasing alkyl chain length (e.g. alkyl benzenes and alkyl naphthalenes). However, as the carbon number increases \(>C_{25}\), the TIC response is linearly related to the compound mass, which is plotted in Figure B as an illustration. The perfect overlap of the curves also demonstrates that the TIC intensity was almost unaffected by the individual chemical structures under the same injection mass. (For the high molecular weight alkyl-cyclohexanes, only \(C_{25}\) and \(C_{33}\) were investigated due to the limitation of the available authentic standards).

The TIC molar quantity responses of deuterated standards are also shown in Figure S4-3 to further support the above view. Additionally, the TIC versus molar quantity for the alkane and alkyl-cyclohexane series are shown in Figure S4-2. This has been explored in more detail in Liang et al. (2017, in prep).
Figure S4-1. TIC intensity response to the molar quantity of individual n-alkanes and n-alkyl-cyclohexanes (C12-C24).
**Figure S4-2.** TIC intensity response to the molar quantity (left) and mass(right) of individual n-alkanes and n-alkyl-cyclohexanes (C$_{25}$-C$_{34}$).

**Figure S4-3.** TIC intensity response to the normalized molar quantity of individual deuterated standards (C$_{12}$, C$_{15}$, C$_{20}$, C$_{25}$, C$_{30}$).
**Figure S4-4.** Plots of total ion current vs. molar quantity for the individual alkanes with carbon number from C12-C34.
Section S5: Representative Mass Spectra for diesel and gas phase hydrocarbon identification

Representative mass spectra at 70eV and 12eV ionisation are shown below. These mass spectra were identical at their respective retention times in different samples. Mass spectra at the bottom panel (blue peaks) are from 12 eV ionisation, while the top panels (red peaks) are from 70eV ionisation.

**Figure S5-1.** Representative mass spectra of n-alkanes, n-alkylcyclohexanes and n-alkylbenzenes. Top graphs with red peaks are 70eV ionisation mass spectra, while bottom graphs with blue peaks are 12eV spectra.
**Figure S5-2.** Representative mass spectra of n-alkynaphthalenes, cis and trans-decalin, tetralin, methyl-, dimethyl-, trimethyl-, and tetramethyl tetralins. Top graphs with red peaks are 70eV ionisation mass spectra, while bottom graphs with blue peaks are 12eV spectra.
**Figure S5-3.** Representative mass spectra of four PAH. Top graphs with red peaks are 70eV ionisation mass spectra, while bottom graphs with blue peaks are 12eV spectra.

When two compounds co-elute, it is often possible to quantify each by using the low ionisation energy mass spectrum. The molecular ion of the higher molecular weight component is used to provide an ion current which is scaled upward to include the other peaks in the mass spectrum to give a total ion current due to the component. The second component can be determined by difference. The molecular ions for a C\textsubscript{11} and a C\textsubscript{12} cycloalkane are at m/z 154 and 168, respectively. These peaks would be present in the mass spectrum of co-eluting compounds, but can be separated by isolating the ion current attributed to a peak by assessing the mass spectrum that contains only the molecular ion of one of the compounds. This avoids the issue of co-elution of some compounds.
Section S6: Unresolved peaks

For the diesel fuel 90% of the total mass injected was accounted for. We suspect that a significant amount of the mass that was unaccounted may be <C_{10} and/or any unresolved peaks that we were unable to measure and/or identify using our technique. There were 123 unresolved peaks in the 2D chromatogram and their positioning is shown in Figure S6-1. The intensities of the peaks, however, are relatively small and thus suggest a small contribution to the overall mass. Representative mass spectra of some of these unresolved peaks are shown in Figure S6-2.

Figure S6-1. A 2D chromatogram showing the positioning of unresolved peaks (red oval).
Figure S6-2. Representative mass spectra of unresolved peaks by this methodology. These unresolved peaks were present in the chromatographic space outlined in Figure S6-1.
**Figure S7.** A chromatogram of diesel exhaust emissions in the gas phase collected on an adsorption tube and analysed via thermal desorption.

**Section S8 – Representative mass spectra for compound classes identified in the lubricating oil**

n + i-Alkanes were identified by the SIC of m/z 57 and selecting the corresponding molecular ion of the alkane. Figure S8-1 shows two representative alkanes with the molecular ions 282 (C20) and 296 (C21).

**Figure S8-1.** n-Alkane mass spectra at 12eV for m/z 57 and the molecular ion 282 and 296 for C20 n-alkane (top mass spectrum, red) and C21 n-alkane (bottom mass spectrum, blue), respectively.
Monocyclic alkanes were identified by the SIC of m/z 82 and selecting the corresponding molecular ion of the specific monocyclic alkane. For example, Figure S8-2 shows two representative monocyclic alkanes with the molecular ions 280 (C20) and 294 (C21).

Figure S8-2. Monocyclic alkane mass spectra at 12eV for m/z 82 and the molecular ions 280 and 294 for C20 monocyclic alkane (top mass spectrum, red) and C21 monocyclic alkane (bottom mass spectrum, blue), respectively.

Bicyclic alkanes were identified by the SIC of m/z 137 and selecting the corresponding molecular ion of the specific bicyclic alkane. For example, Figure S8-3 shows two representative bicyclic alkanes with the molecular ions 292 (C21) and 306 (C22).
**Figure S8-3.** Bicyclic alkane mass spectra at 12eV for m/z 137 and the molecular ions 292 and 306 for C21 bicyclic alkane (top mass spectrum, red) and C22 bicyclic alkane (bottom mass spectrum, blue), respectively.

Tricyclic terpanes were identified by the SIC of m/z 191 and selecting the corresponding molecular ion of the specific tricyclic terpane carbon number. For example, Figure S8-4 shows two representative tricyclic terpanes with the molecular ions 276 (C20) and 290 (C22).

**Figure S8-4.** Tricyclic terpanes mass spectra at 12eV for m/z 191 and the molecular ions 276 and 290 for C20 tricyclic terpane (top mass spectum, red) and C21 tricyclic terpane (bottom mass spectrum, blue), respectively.

This analysis was repeated for all the compounds shown in Table 1 and representative mass spectra are shown in Figures S8-5 to S8-7 below.
**Figure S8-5.** Monocyclic aromatics mass spectra at 12eV for m/z 92 and the molecular ions 288 and 302 for C21 (top mass spectrum, red) and C22 (bottom mass spectrum, blue), respectively.

**Figure S8-6.** Pentacyclic alkanes mass spectra at 12eV for the molecular ions 258 (top mass spectrum, red) and methyl pentacyclic alkane with molecular ion 272 and m/z 257 mass fragment (bottom mass spectrum, blue).

**Figure S8-7.** Hexacyclic alkanes mass spectra at 12eV for the molecular ions 298 (top mass spectrum, red) and methyl hexacyclic alkane with molecular ion 312 and m/z 297 mass fragment (bottom mass spectrum, blue).
Figure S9. A selected ion chromatogram of lubricating oil (5W30) with mass fragments (A) m/z 92 and 119 signifying monocyclic and methyl-monocyclic aromatics and (B) m/z 302 signifying the M⁺ for C₂₂ monocyclic aromatic isomers.

References
