UC Berkeley

UC Berkeley Previously Published Works

Title

Comprehensive Chemical Characterization of Hydrocarbons in NIST Standard Reference Material 2779 Gulf of Mexico Crude Oil.

Permalink

<https://escholarship.org/uc/item/2zr0m0tv>

Journal

Environmental science & technology, 49(22)

ISSN

0013-936X

Authors

Worton, David R
Zhang, Haofei
Isaacman-VanWertz, Gabriel
et al.

Publication Date

2015-11-01

DOI

10.1021/acs.est.5b03472

Peer reviewed

1 **Comprehensive chemical characterization of hydrocarbons in NIST standard**
2 **reference material 2779 Gulf of Mexico crude oil**

3
4 David R. Worton^{1,2,&,*}, Haofei Zhang¹, Gabriel Isaacman-VanWertz^{1,§}, Arthur W. H.
5 Chan^{1,#}, Kevin R. Wilson³ and Allen H. Goldstein^{1,4}.

6
7 ¹ Department of Environmental Science, Policy and Management, University of
8 California, Berkeley, California, 94720, United States.

9 ² Aerosol Dynamics Inc., Berkeley, California, 94710, United States.

10 ³ Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley,
11 California, 94720, United States.

12 ⁴ Department of Civil and Environmental Engineering, University of California,
13 Berkeley, California, 94720, United States.

14 [&] Now at National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11
15 0LW, U.K.

16 [#] Now at Department of Chemical Engineering and Applied Chemistry, University of
17 Toronto, Ontario, M5S 3E5, Canada.

18 [§] Now at Department of Civil and Environmental Engineering, Massachusetts Institute of
19 Technology, Cambridge, MA, 02139, United States.

20
21 ^{*}Corresponding author: National Physical Laboratory, Hampton Road, Teddington,
22 Middlesex, TW11 0LW, U.K., Tel. +44 208 9436591, Fax. +44 208 6140446,
23 dave.worton@npl.co.uk.

24
25 **Keywords:** NIST SRM 2779, crude oil, *Deepwater Horizon* disaster, Surrogate,
26 GC/VUV-MS, chemical composition

27
28 **Abstract**

29 Comprehensive chemical information is needed to understand the environmental fate and
30 impact of hydrocarbons released during oil spills. However, chemical information
31 remains incomplete due to the limitations of current analytical techniques and the
32 inherent chemical complexity of crude oils. In this work, GC-amenable C₉ – C₃₃
33 hydrocarbons were comprehensively characterized in the National Institute of Standards
34 and Technology Standard Reference Material (NIST SRM) 2779 Gulf of Mexico crude
35 oil, by gas chromatography coupled to vacuum ultraviolet photoionization mass
36 spectrometry (GC/VUV-MS), with a mass balance of 68 ± 22 %. This technique
37 overcomes one important limitation faced by traditional GC and even comprehensive two
38 dimensional gas chromatography (GC×GC), the necessity for individual compounds to be
39 chromatographically resolved from one another in order to be characterized. VUV
40 photoionization minimizes fragmentation of the molecular ions facilitating the
41 characterization of the observed hydrocarbons as a function of molecular weight (carbon
42 number, *N_C*), structure (number of double bond equivalents, *N_{DBE}*) and mass fraction (mg
43 kg⁻¹), which represent important metrics for understanding their fate and environmental
44 impacts. Linear alkanes (8 ± 1 %), branched alkanes (11 ± 2 %) and cycloalkanes (37 ±
45 12 %) dominated the mass with the largest contribution from cycloalkanes containing one
46 or two rings and one or more alkyl side chains (27 ± 9 %). Linearity and good agreement

47 to previous work for a subset of > 100 components and for the sum of compound classes
48 provided confidence in our measurements and represented the first independent
49 assessment of our analytical approach and calibration methodology. Another crude oil
50 collected from the Marlin platform (35 km northeast of the Macondo well) was shown to
51 be chemically identical, within experimental errors, to NIST SRM 2779 demonstrating
52 that Marlin crude is an appropriate ‘surrogate’ oil for researchers conducting laboratory
53 research into impacts of the *DeepWater Horizon* disaster.

54 TOC artwork

55

56

Introduction

In the three months following the explosion and loss of the *Deepwater Horizon* (DWH) drilling platform on April 20th 2010 an estimated five million barrels of sweet light crude oil were released from the Macondo well, located in the Mississippi Canyon lease block 252 (MC252), into the Gulf of Mexico^{1, 2}. The oil ascended from 1.5km depth to the sea surface forming an oil slick that eventually impacted more than 1000 km of coastline³. The released oil was a complex mixture of many thousands of different organic compounds with different molecular structures. The environmental fate from different weathering pathways, i.e., evaporation, dissolution, photooxidation and biodegradation, are dependent on their chemical and physical properties, which are influenced by molecular structure^{4, 5}. A considerable number of oxygenated products were formed during some of these weathering processes, many of which are believed to be from saturated precursors although the exact species remain largely unknown^{6, 7}. Molecular structure also substantially influences the oxidation chemistry and secondary organic aerosol (SOA) yields of the volatile components that evaporate from the surface slicks⁸⁻¹¹. Therefore, comprehensive information on the chemical composition of the released hydrocarbons is essential for evaluating the fates and impacts of the emitted species in the environment. Additionally, without a better accounting of the oil composition, especially that of the persistent components of weathered oil, it is challenging to determine what oxygenated products could form and what their potential impacts on marine and coastal ecosystems might be.

Traditionally, gas chromatography coupled to flame ionization detection (GC-FID) or mass spectrometry (GC-MS) has been used to chemically characterize crude oils^{12, 13}. These chromatograms typically exhibited a large raised baseline often referred to as the 'unresolved complex mixture' (UCM)^{4, 14, 15}, comprised of many thousands of constitutional isomers of aliphatic hydrocarbons that are difficult or impossible to separate using conventional GC-based techniques¹⁶. Additionally, the significant fragmentation observed upon electron impact (EI) ionization yields many smaller fragments, which prevents the determination of molecular mass. The analytical challenge of deciphering the complexity of the UCM has meant that historically the majority of chemical analysis of crude oil have focused on components that can be distinguished from the bulk of the mass in the UCM, typically *n*-alkanes, polycyclic aromatic hydrocarbons (PAHs) and sterane and hopane biomarkers. These components comprised < 20 % of the total mass of the Macondo crude oil¹⁷.

More recently, comprehensive two-dimensional gas chromatography (GC × GC) has been successfully used to identify and quantify many more individual compounds and classes of compounds in fresh and weathered crude oils^{4, 5, 14, 18}. However, with the number of possible constitutional isomers increasing exponentially with increasing carbon number¹⁹ it becomes increasingly difficult to resolve individual species even with high resolution capillary GC × GC. Previously, this limitation has prevented a comprehensive understanding of the mass distribution of all compounds that are present. Determining the mass contribution of every constitutional isomer is extremely challenging and probably an unnecessary task since this information is likely to be too detailed for current models. However, being able to characterizing the total mass

distribution as a function of volatility and structure would provide useful constraints for environmental modeling of the released hydrocarbons.

Gas chromatography or comprehensive two dimensional gas chromatography with vacuum ultraviolet ionization mass spectrometry (GC or GC×GC/VUV-MS) overcomes the limitations of isomer resolution in GC for complex mixtures. VUV photoionization results in substantially reduced fragmentation of the molecular ion, which facilitates the classification of hydrocarbon compounds by carbon number (N_C), number of double bond equivalents in their structure (N_{DBE}) and degree of branching²⁰ determined from their molecular weight and GC retention times. The advantage of this technique relative to GC×GC is that individual compounds do not necessarily need to be chromatographically resolved from one another to be characterized. Additionally, compound classes, e.g., alkanes and cycloalkanes are unambiguously separated from one another as a result of molecular mass differences and data processing times are reduced because individual peak assignments are not necessary. As a result, this technique provides a more comprehensive methodology than was previously possible, especially for species with more than 10 carbon atoms. This analytical methodology has previously been shown to provide a more comprehensive approach for the characterization of complex organic mixtures of petrochemical products derived from crude oil^{10, 20, 21}. The main objectives of this work are to use these advanced techniques and classification schemes to provide a more comprehensive characterization of the insoluble semi-volatile hydrocarbons released during the DWH disaster and to provide a comprehensive assessment of the chemical composition of National Institute of Science and Technology (NIST) Standard Reference Material (SRM) 2779 Gulf of Mexico crude oil for the research community to utilize.

Experimental Methodology

Description of Oil Samples

Analysis of two crude oil samples are presented in this work; NIST SRM 2779 Gulf of Mexico Crude Oil and ‘surrogate’ oil (SURR). Both these crude oils are being widely used by the research community. The SURR oil is being provided to supplement the high demand and limited supply of Macondo well crude oil. The NIST SRM 2779 crude oil was collected on May 21, 2010 on the drillship *Discoverer Enterprise* from the insertion tube receiving oil directly from the Macondo well during response operations. The water was separated from the oil and the resulting oil was homogenized before being transferring into 2 mL amber glass ampoules. Certified mass fractions of 21 PAHs and reference mass fractions for an additional 22 PAHs and 31 alkylated PAHs were reported by NIST (available online at: <https://www-s.nist.gov/srmors/certificates/2779.pdf>).

Certified mass fractions were weighted means obtained from two to five different analytical methods at NIST and represent values with the highest level of confidence in their accuracy in that all known or suspected sources of bias have been investigated or taken into account. Reference values for the additional 22 PAHs were weighted means obtained from two to five different analytical methods at NIST. The reference mass fractions of the 31 alkylated PAHs were the means of results obtained using one analytical technique from an interlaboratory intercomparison exercise of 24 laboratories although not all returned data for each analyte. The reference mass fractions are

noncertified values that are estimates of the true value. However, these values do not meet NIST's criteria for certification and were provided with associated uncertainties that may reflect only measurement precision, may not include all sources of uncertainty or may reflect a lack of sufficient statistical agreement among multiple analytical methods. The SURR oil, classified as a 'dead' oil as it does not contain any dissolved gases, was collected from the Marlin Platform of the Dorado field, 35 km northeast of the Macondo well. The oil was stored in 350 gallons tanks that were filled on site. The SURR oil was made available to Gulf of Mexico Research Initiative (GoMRI) funded researchers by BP and was obtained from BP's Knox Storage Facility in Denver, Colorado by submitting a sample request to Architecture, Engineering, Consulting, Operations, and Maintenance (AECOM), the contracted managers of the facility.

Analytical Methods

Samples were analyzed using GC/VUV-HRTOFMS. Oil samples were diluted (50:1) in chloroform (HPLC grade, Sigma Aldrich), directly injected via a septumless inlet into a liquid nitrogen cooled inlet (CIS4, Gerstel Inc.), and cyrofocused at -25 °C on a quartz wool packed inlet liner. Analytes were transferred to the gas chromatograph (GC, Agilent 7890) by rapid heating of the CIS to 300 °C under a helium flow. Analytes were separated on a non-polar primary column (60 m × 0.25 mm × 250 µm Rxi-5Sil-MS; Restek) with a carrier gas flow rate of 2 ml min⁻¹ of helium. The GC temperature program was 40 °C hold for 2 min, 3.5 °C min⁻¹ until 320 °C and hold for 10 mins. Analytes were ionized with 10.5 electron volts (eV) vacuum ultraviolet photoionization (VUV) using a high resolution ($m/\Delta m \sim 4000$) time of flight mass spectrometer (HTOF, Tofwerk). The frequency of data collection on the mass spectrometer was 200 Hz, averaged to 0.5 Hz to improve identification and deconvolution of high resolution mass spectral peaks. The transfer line from the GC to the TOF was maintained at 270 °C. To minimize fragmentation in VUV the ion source was operated at 170 °C. The VUV beam (tunable from 8 – 30 eV but typically used at 10.5 ± 0.2 eV or 9.0 ± 0.2 eV, photon flux ~10¹⁵ photons s⁻¹) was generated by the Chemical Dynamics Beamline 9.0.2 of the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory. In our group, we have been pioneering the use of synchrotron radiation as the source of VUV photons due to the high photon flux. However, we acknowledge that this limits the accessibility of this technique beyond our research group. However, it should be noted that the first application of VUV ionization utilized a novel excimer VUV light source, in which an electron beam is used to form rare gas excimer species²² while alternative soft ionization technologies are also now emerging, e.g., a variable eV ion source (Markes International Ltd, UK) and field ionization mass spectrometry (JEOL USA Inc.)²³, that will likely lead to much wider use of our analytical methodology within the research community.

As the ionization energies of most organic compounds are between 8 and 11 eV, the minimal excess energy (compared to EI at 70 eV) limits fragmentation of ionized molecules, allowing for significant detection of the molecular ions²⁰. In this work, as in previous work, the molecular ion mass and chromatographic retention time were used to classify compounds comprising the UCM by N_C , N_{DBE} and degree of branching^{20, 21}. High mass resolution data processing was employed to distinguish hydrocarbons from any GC-amenable oxygen containing molecular ions¹¹. All data processing and visualization of

data were performed using custom code written in Igor 6.3.6 (Wavemetrics) adapted from high-resolution aerosol mass spectrometer data analysis code²⁴.

Calibration Methodology

The calibration methodology used in this work was based on the method described in Isaacman et al.,²⁰ and Chan et al.,²⁵. In brief, the molecular ion signals for linear, branched, cyclic and aromatic hydrocarbons under VUV ionization are used as the basis for quantification. The sensitivity of the molecular ion of any given compound is a function of its thermal transfer efficiency, ionization efficiency and degree of fragmentation. Owing to ionization efficiency and extent of fragmentation, the relative molecular ion signal increases with increasing N_{DBE} , due to reduced fragmentation. Authentic standards of more than 80 compounds were used for calibration, these included *n*-alkanes, branched alkanes, *n*-alkyl cyclohexanes, *n*-alkyl benzenes, hopanes and steranes, PAHs and alkylated PAHs. These species were selected to span both the N_C and N_{DBE} ranges of the crude oil samples. As compounds with very similar structures fragment almost identically after undergoing photoionization by VUV they yield essentially the same mass-response calibration curves meaning that mass-response is strongly a function of structure and less dependent on carbon number. This resulted in N_{DBE} specific calibration curves for $N_{DBE} = 0, 1, 4$ and $7+$, due to the limited availability of authentic standards for the other N_{DBE} 's. The dependence of fragmentation on N_{DBE} was previously characterized using a standard mixture consisting of isomers of the same carbon number but different N_{DBE} ²⁵ and this was used to interpolate for N_{DBE} 's for which very limited authentic standards are available.

The thermal transfer efficiency is not linear with carbon number as early eluting components and late eluting components show lower responses relative to components eluting between them as a result of challenges efficiently transferring all the high and low volatility material. A series of perdeuterated *n*-alkanes (*n*-octane, *n*-decane, *n*-dodecane, *n*-tetradecane, *n*-hexadecane, *n*-octadecane, *n*-eicosane, *n*-docosane, *n*-tetracosane, *n*-hexacosane, *n*-octacosane, *n*-triacontane, *n*-dotriacontane and *n*-tetratriacontane; C/D/N Isotopes) added as an internal standard to all samples were used to generate a relationship between thermal transfer and retention time. This relationship was used for all species because the nature of the transfer efficiency is related to volatility and is not chemically specific. These were selected to span the volatility range of the analytes of interest. As aliphatics including the perdeuterated *n*-alkanes were not ionized at 9 eV, perdeuterated polycyclic aromatic hydrocarbons (Phenanthrene, Pyrene, Chrysene, Perylene, Dibenzo[*a,h*]anthracene; C/D/N Isotopes) were used to correct for thermal transfer efficiency in the 9eV runs. The perdeuterated *n*-alkanes were also used to post-calibrate the high resolution mass spectral data. However, high resolution peak fitting was not conducted at 9 eV because mass calibration was not possible due to an insufficient number of ions spanning a wide enough mass range to generate a robust relationship between exact mass and time of flight. This is a result of the much lower fragmentation of the PAHs relative to the *n*-alkanes. For these runs the nominal mass data was used. A comparison between the nominal and high resolution mass data for the samples collected at 10.5 eV yielded very similar results.

Total analytical uncertainty includes contributions from transfer efficiency, structural differences in fragmentation within a N_{DBE} class and uncertainties in calibration curves. The uncertainty in calibrating response to mass, determined from calibration curves of authentic standards, was structurally and mass dependent with larger uncertainties for lower N_{DBE} species and smaller mass fractions (Figure S1, supporting information). In general, the total analytical uncertainty was $< 40\%$ for all species at mass fractions above 1000 mg kg^{-1} , increasing to $< 75\%$ at mass fractions of 100 mg kg^{-1} (Table S1, supporting information). Periodic system blanks were analysed and these showed that background levels were negligible relative to observed levels of analytes in the samples. Estimated detection limits were < 10 and $< 1\text{ mg kg}^{-1}$ for aliphatic and aromatic species, respectively, and analytical precision was $< 10\%$ for all species as determined from replicate analysis of the NIST SRM 2779 crude oil.

Results

Comprehensive Mass Fraction Distributions of NIST SRM 2779 Oil Components

Figure 1a shows the comprehensive mass fraction (mg kg^{-1}) distribution of the NIST SRM 2779 Gulf of Mexico crude oil for components with 9 to 33 carbon atoms ($N_C = C_9 - C_{33}$) separated by the N_{DBE} . Contributions of branched and straight compounds to $N_{DBE} = 0$ and 1 are also shown as they are clearly distinguishable in the chromatograms (Figure S2, supporting information). The total contribution of each N_{DBE} class to the total mass divide by the total mass injected (mass balance) is also shown in Figure 1a as a percentage for all $C_9 - C_{33}$ compounds observed. For clarity, all branched compounds for $N_{DBE} = 0$ and 1 are shown as a sum. The mass distribution of the branched components of $N_{DBE} = 0$ separated by the number of methyl groups observed on the carbon chain are shown in Figure 1b. Again, for clarity, all components with $N_{DBE} \geq 7$ ($N_{DBE} = 7+$) are presented as a sum and the contribution of each N_{DBE} class from 7 to 15 is shown in Figure 1c. The data used to construct Figure 1 are shown in Tables 1 and S2 (supporting information), respectively, along with the total analytical uncertainty.

The lower N_{DBE} components dominate with contributions from $N_{DBE} = 0, 1, 2$ and 3 making up approximately 50% of the mass of the oil (Figure 1a). These N_{DBE} classes represent straight and branched acyclic alkanes and mono-, bi- and tricyclic alkanes. From the complexity of the observed parent ion chromatograms and knowledge of the N_C and N_{DBE} we infer that the vast majority of the cycloalkane mass is dominated by those with one or more alkyl side chain on the ring or rings. This is supported by the GC retention time data, which are consistent with the literature for where branched alkyl cyclohexanes elute relative to the n -alkyl cyclohexane isomers²¹. The majority of these compounds have not previously been reported. The straight chain alkanes and the branched alkanes both contribute approximately similar fractions of $\sim 10\%$. There are substantially fewer alkane isomers than are statistically possible and the observed isomer distributions are consistent with previous work on diesel fuel and crude oil that suggest branched alkanes in these fuels are primarily comprised of methylated isomers^{5, 26, 27}. On average, methyl alkanes (B1) comprised about half of the branched alkane mass with the majority of the remaining mass being comprised of the most branched isoprenoid alkanes, e.g., for $N_C = 18$ this is 3 methyl groups (B3) but for $N_C = 19$ this is 4 methyl groups (B4) (Figure S3, supporting information). This is consistent with previous work

on diesel fuel¹⁰ but the observation in crude oil implies that it is not a result of the refining process and is related to the formation from the concatenation of isoprene (C_5H_8) subunits to form more complex molecules in living organisms prior to crude oil formation²⁸. It is therefore highly likely that the alkylated side chains of the cyclic alkanes are also restricted to being only methyl branched isomers derived from isoprene subunits, though this cannot be confirmed from their mass spectra. However, many more isomers are observed for the cycloalkanes relative to acyclic alkanes (Figure S2, supporting information) because there are a larger number of possible configurations resulting from the different positions for alkyl substitution on the ring. The most likely cycloalkane ring structures are cyclopentane and cyclohexane, due to the minimal ring strain of these configurations. This is supported by the retention time data which is consistent with these structures, similar to previous work on lubricating oil²¹, and not with larger cycloalkane rings that elute later. Alkene, diene and triene moieties would also have N_{DBE} of 1, 2 and 3. However, unsaturated chains are relatively unstable and are thus very unusual in crude oils. As a result, we expect contributions from these species to be negligible, consistent with previous work on refined petrochemical products^{21, 29}.

Higher N_{DBE} 's contribute a smaller fraction of the mass ($N_{DBE} \geq 4$, ~15 %) and include contributions from tetra-, penta- and hexacyclic alkanes, benzenes, tetralins, indanes and PAHs including the alkylated varieties. Naphthalene and alkylated naphthalenes ($N_{DBE} = 7$) and phenanthrenes and its alkylated homologues ($N_{DBE} = 10$) and $N_{DBE} = 8$ were the dominant PAH components present (Figure 1c). It should be noted that at 10.5 eV $N_{DBE} = 8$ was heavily impacted by fragment ions of aliphatic compounds (Figure S4, supporting information) that prevented reliable quantitation. At 9 eV, the interference from these fragments was minimal as aliphatic compounds have ionization energies higher than 9 eV³⁰. A calibration curve at 9 eV was not available, so we estimate the mass of $N_{DBE} = 8$ by converting the raw signal at 9 eV to mass at 10.5 eV through comparison to $N_{DBE} = 7$ and 9 (Figure S4, supporting information). The $N_{DBE} = 8$ values listed in Table S2 (supporting information) therefore include an additional uncertainty of 10 %.

In this work, we do not explicitly report the aromatic and aliphatic contributions to $N_{DBE} = 4, 5$ and 6 due to difficulties in distinguishing them using our methodology because the IE of the species is similar and thus cannot be differentiated in 9 versus 10.5 eV runs (see text and Figures S5, S6 and S7 in the supporting information). However, we can rationalize that a greater proportion of the $N_{DBE} = 4, 5$ and 6 mass at lower N_C (9 – 20) is likely aromatic. This is because the complexity (many co-eluting peaks) observed in the parent ion chromatograms for these N_{DBE} 's imply that there are a large number of structural isomers present. For example, consider compounds with $N_C = 15$, for a monoaromatic compound this represents a benzene ring with a nonyl side chain that could be distributed in hundreds of different configurations. Whereas for a tetracyclic compound with two cyclohexane rings and two cyclopentane rings only one configuration is possible. This suggests a larger contribution from aromatic compounds at lower carbon numbers ($\leq C_{20}$) because statistically there are more possible isomer structures.

Table 2 compares the mass fractions eluting between different *n*-alkane windows observed in this work with values reported by Reddy et al.,¹⁷ for oil sampled directly from the Macondo well by a remotely operated vehicle on June 21st, 2010 during response efforts to the DWH disaster. The very good agreement within each of these elution ‘bins’ provides confidence in our results. Small compounds are not quantitatively captured by our inlet, and elution of large compounds is limited by temperatures within the mass spectrometer. Within the range of compounds studied in this work (C₉ – C₃₃), we account for 68 ± 22 % of the mass.

Comparison with other published data

Figure 2 shows the comparison between this work and published mass fraction (mg kg⁻¹) data for selected species reported by Woods Hole Oceanographic Institution (WHOI)¹⁷ (Figure 2a) and NIST (Figure 2b). Figure 2c shows the comparison between NIST and WHOI. All compounds included in these comparisons are shown in Table S3 (supporting information) along with their measured mass fractions and uncertainties, if available. As part of the comparison the elution order of specific isomers was confirmed through comparison of reported Kovats retention indices (KI) for temperature programmed GC³¹ (or Lee retention indices (LI)³² converted to KI), from the literature to those observed in this work (Table S3, supporting information). In general, there is very good agreement and linearity ($r^2 = 0.84$, $n = 67$) between this work and WHOI (Figure 2a) with all data agreeing to within a factor of two, except the C₉ aromatics (Figure S8, supporting information), with a relative mean standard deviation (RMSD) of 18 %. There is also good agreement between this work and NIST ($r^2 = 0.64$, $n = 55$, RMSD = 32 %; Figure 2b) although the mass fractions observed in this work were typically higher (on average a factor of 2, see Figure S8 supporting information) for most compounds, though most were less than a factor of four with the exception of the dimethyl decalins (Figure S8, supporting information). A comparison between NIST and WHOI also agreed very well ($r^2 = 0.95$, $n = 30$, RMSD = 32 %; Figure 2c). The mass fractions reported by WHOI were also higher than NIST with a similar magnitude to that observed in this work. This intercomparison represents the first independent assessment of our calibration methodology and the good agreement provides confidence in our analytical approach.

Table 3 shows the total mass fraction (%) and total mass released from the Macondo well during the DWH disaster for all *N*_{DBE} classes reported in this work for compounds with *N*_C = 9 – 33 (C₉ – C₃₃). Table 3 also shows the GC-amenable percent mass (%) data reported by Reddy et al.,¹⁷ separated by broad chemical class. There is excellent agreement for *n*-alkanes, sum of the branched alkanes and cycloalkanes, monoaromatics and PAHs with both datasets reporting the same total mass fractions, within the reported uncertainties. However, while the sum of the branched and cycloalkanes agree well the contribution of those compound classes differ. In this work, we observe approximately a factor of two less branched alkanes and almost three times more cycloalkanes, respectively, than Reddy et al.,¹⁷. Some of these discrepancies might be from contributions from branched and/or cycloalkanes that were not characterizable using their method and that were reported in a category labeled as ‘other’ (18 %). As such, we suggest that the results presented here provide a more accurate representation of the total emission of branched and cyclic alkanes released during the DWH disaster (Table 3).

Intercomparison of NIST SRM 2779 and Surrogate Oil from the Marlin Platform

With many scientific studies into the impacts of the DWH disaster planned for at least the next 10 years there is expected to be continued demand for Macondo source oil. Demand may exceed the supply of the available source oil such that surrogate crude oils will be necessary. Oil from the BP owned and operated Marlin platform was selected as surrogate oil based on its similar geography, oil family, composition, crude oil chemistry and aquatic toxicity. Previous chemical analyses of the composition were limited to total petroleum hydrocarbons (TPH), *n*-alkanes, selected aromatics including benzene, toluene, xylenes (BTX), polycyclic aromatic hydrocarbons (PAH) and geochemical biomarkers. This data was presented at the 32nd annual meeting of the Society of Environmental Toxicology and Chemistry (SETAC) in 2011 but to the best of our knowledge this work has not been published in the peer reviewed literature.

Figure 3 shows an intercomparison of the comprehensive chemical composition of the Marlin Surrogate and NIST SRM 2779 (Macondo well) crude oils and clearly demonstrates that they are chemically very similar ($r^2 = 0.86$, $n = 236$, RMSD = 10 %) with almost all N_C and N_{DBE} classes reported in this work being within error of unity. This is perhaps not too surprising given that BP made a concerted effort to identify a similar crude oil that would make a suitable surrogate. However, what is interesting is how similar these two crude oils are considering they are located 35 km apart, which underscores the challenges of studying in the Gulf of Mexico. As a result of the similarity in these crude oils, the comprehensive chemical composition data reported in this work (Table 1 and Table S2, supporting information) will be of interest to those researchers conducting research using authentic Macondo well source oil, NIST SRM 2779 oil and/or the Surrogate oil from the Marlin Platform.

Environmental Implications

In this work, the chemical composition of hydrocarbons present in Macondo well crude oil was characterized using GC/VUV-MS, with a mass balance of 68 ± 22 % for $C_9 - C_{33}$ hydrocarbon compounds. The chemical classification scheme used in this work provides data in a format that can be utilized by environmental models to better understand the comprehensive chemical evolution of crude oil during weathering processes. One such example is to estimate the speciated mass fluxes of evaporating components from surface oil slicks and to predict the impact of these evaporating hydrocarbons on air pollutant concentrations, namely particulate matter and ozone, with implications for air quality in the vicinity of an oil spill^{10, 33}. Analysis by GC/VUV-MS is particularly well suited for the determination of saturated semi-volatile compounds where the presence of a large numbers of constitutional isomers inhibits the ability to chromatographically resolve individual peaks, which would otherwise prevent their detection. This is especially true for cycloalkanes where, in addition to skeletal isomerism of the alkyl side chain or chains there are many positional isomers as a result of different arrangements of those alkyl side chains on the ring or rings. Recent work has identified these hydrocarbons as being important precursors to observed oxygenated hydrocarbons, formed via photooxidation and biodegradation, although their identities have not been categorically determined^{6, 7}. The application of GC/VUV-MS to crude oil samples with varying degrees of

environmental weathering could provide invaluable insights into the identities of the key precursors and therefore to the likely identity of the oxygenated products formed during the weathering process.

Supporting Information

Supporting information includes additional information as noted in the text and supplementary Figures S1 – S8 and Tables S1 – S3. This material is available free of charge via the internet at <http://pubs.acs.org>. The data presented in Table 1 and Table S2 are also available from the GRIIDC data archive, accessible via the internet at <http://ezid.cdlib.org/id/doi:10.7266/N7K64G1H>.

Acknowledgements

This work was funded by the Gulf of Mexico Research Initiative (GoMRI) as part of the Gulf Integrated Spill Response (GISR) consortium under contract SA12-09/GoMRI-006. GIVW was supported by the National Science Foundation (NSF) Graduate Research Fellowship (NSF grant: DGE 1106400). The Advanced Light Source, the Chemical Dynamics Beamline and KRW were supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

1. McNutt, M. K.; Camilli, R.; Crone, T. J.; Guthrie, G. D.; Hsieh, P. A.; Ryerson, T. B.; Savas, O.; Shaffer, F. Review of flow rate estimates of the Deepwater Horizon oil spill. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, (50), 20260-20267. doi:10.1073/pnas.1112139108.
2. Ryerson, T. B.; Camilli, R.; Kessler, J. D.; Kujawinski, E. B.; Reddy, C. M.; Valentine, D. L.; Atlas, E.; Blake, D. R.; de Gouw, J.; Meinardi, S.; Parrish, D. D.; Peischl, J.; Seewald, J. S.; Warneke, C. Chemical data quantify Deepwater Horizon hydrocarbon flow rate and environmental distribution. *Proceedings of the National Academy of Sciences* **2012**, *109*, (50), 20246-20253. doi:10.1073/pnas.1110564109.
3. Graham, B.; Reilly, W. K.; Beinecke, F.; Boesch, D. F.; Garcia, T. D.; Murray, C. A.; Ulmer, F. *Deep water: The Gulf oil disaster and the future of offshore drilling*; US Government Printing Office, Washington, DC: Report to the President: national commission on the BP *Deepwater Horizon* oil spill offshore drilling, 2011.
4. Nelson, R. K.; Kile, B. M.; Plata, D. L.; Sylva, S. P.; Xu, L.; Reddy, C. M.; Gaines, R. B.; Frysinger, G. S.; Reichenbach, S. E. Tracking the weathering of an oil spill with comprehensive two-dimensional gas chromatography. *Environ. Forensics* **2006**, *7*, (1), 33-44. doi:10.1080/15275920500506758.
5. Gros, J.; Reddy, C. M.; Aeppli, C.; Nelson, R. K.; Carmichael, C. A.; Arey, J. S. Resolving biodegradation patterns of persistent saturated hydrocarbons in weathered oil samples from the *Deepwater Horizon* disaster. *Environ. Sci. Technol.* **2014**, *48*, (3), 1628-1637. doi:10.1021/es4042836.
6. Hall, G. J.; Frysinger, G. S.; Aeppli, C.; Carmichael, C. A.; Gros, J.; Lemkau, K. L.; Nelson, R. K.; Reddy, C. M. Oxygenated weathering products of *Deepwater Horizon* oil come from surprising precursors. *Mar. Pollut. Bull.* **2013**, *75*, (1-2), 140-149. doi:10.1016/j.marpolbul.2013.07.048.
7. Aeppli, C.; Carmichael, C. A.; Nelson, R. K.; Lemkau, K. L.; Graham, W. M.; Redmond, M. C.; Valentine, D. L.; Reddy, C. M. Oil weathering after the *Deepwater Horizon* disaster led to the formation of oxygenated residues. *Environ. Sci. Technol.* **2012**, *46*, (16), 8799-8807. doi:10.1012/es3015138.
8. Lim, Y. B.; Ziemann, P. J. Effects of molecular structure on aerosol yields from OH radical-initiated reactions of linear, branched, and cyclic alkanes in the presence of NO_x. *Environ. Sci. Technol.* **2009**, *43*, (7), 2328-2334. doi:10.1021/es803389s.
9. Tkacik, D. S.; Presto, A. A.; Donahue, N. M.; Robinson, A. L. Secondary organic aerosol formation from intermediate-volatility organic compounds: Cyclic, linear, and branched alkanes. *Environ. Sci. Technol.* **2012**, *46*, (16), 8773-8781. doi:10.1021/es301112c.
10. Gentner, D. R.; Isaacman, G.; Worton, D. R.; Chan, A. W. H.; Dallmann, T. R.; Davis, L.; Liu, S.; Day, D. A.; Russell, L. M.; Wilson, K. R.; Weber, R.; Guha, A.; Harley, R. A.; Goldstein, A. H. Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, (45), 18318-23. doi:10.1073/pnas.1212272109.
11. Isaacman, G.; Chan, A. W. H.; Nah, T.; Worton, D. R.; Ruehl, C. R.; Wilson, K. R.; Goldstein, A. H. Heterogeneous OH oxidation of motor oil particles causes selective

- depletion of branched and less cyclic hydrocarbons. *Environ. Sci. Technol.* **2012**, *46*, (19), 10632-10640. doi:10.1021/es302768a.
12. Wang, Z.; Fingas, M. Differentiation of the source of spilled oil and monitoring of the oil weathering process using gas chromatography-mass spectrometry. *Journal of Chromatography A* **1995**, *712*, (2), 321-343. doi:10.1016/0021-9673(95)00546-Y.
13. Wang, Z. D.; Fingas, M. F. Development of oil hydrocarbon fingerprinting and identification techniques. *Mar. Pollut. Bull.* **2003**, *47*, (9-12), 423-452. doi:10.1016/s0025-326x(03)00215-7.
14. Frysinger, G. S.; Gaines, R. B.; Xu, L.; Reddy, C. M. Resolving the unresolved complex mixture in petroleum-contaminated sediments. *Environ. Sci. Technol.* **2003**, *37*, (8), 1653-1662. doi:10.1021/es020742n.
15. Ventura, G. T.; Hall, G. J.; Nelson, R. K.; Frysinger, G. S.; Raghuraman, B.; Pomerantz, A. E.; Mullins, O. C.; Reddy, C. M. Analysis of petroleum compositional similarity using multiway principal components analysis (MPCA) with comprehensive two-dimensional gas chromatographic data. *Journal of Chromatography A* **2011**, *1218*, (18), 2584-2592. doi:10.1016/j.chroma.2011.03.004.
16. Gough, M. A.; Rowland, S. J. Characterization of unresolved complex mixtures of hydrocarbons in petroleum. *Nature* **1990**, *344*, (6267), 648-650. doi:10.1038/344648a0.
17. Reddy, C. M.; Arey, J. S.; Seewald, J. S.; Sylva, S. P.; Lemkau, K. L.; Nelson, R. K.; Carmichael, C. A.; McIntyre, C. P.; Fenwick, J.; Ventura, G. T.; Van Mooy, B. A. S.; Camilli, R. Composition and fate of gas and oil released to the water column during the Deepwater Horizon oil spill. *Proceedings of the National Academy of Sciences* **2012**, *109*, (50), 20303-20308. doi:doi: 10.1072/pnas.1101242108.
18. Reddy, C. M.; Eglinton, T. I.; Hounshell, A.; White, H. K.; Xu, L.; Gaines, R. B.; Frysinger, G. S. The west falmouth oil spill after thirty years: The persistence of petroleum hydrocarbons in marsh Sediments. *Environ. Sci. Technol.* **2002**, *36*, (22), 4754-4760. doi:10.1021/es020656n.
19. Goldstein, A. H.; Galbally, I. E. Known and unexplored organic constituents in the earth's atmosphere. *Environ. Sci. Technol.* **2007**, *41*, (5), 1514-1521. doi:10.1021/es072476p.
20. Isaacman, G.; Wilson, K. R.; Chan, A. W. H.; Worton, D. R.; Kimmel, J. R.; Nah, T.; Hohaus, T.; Gonin, M.; Kroll, J. H.; Worsnop, D. R.; Goldstein, A. H. Improved resolution of hydrocarbon structures and constitutional isomers in complex mixtures using gas chromatography-vacuum ultraviolet-mass spectrometry. *Anal. Chem.* **2012**, *84*, (5), 2335-2342. doi:10.1021/ac2030464.
21. Worton, D. R.; Isaacman, G.; Gentner, D. R.; Dallmann, T. R.; Chan, A. W. H.; Ruehl, C.; Kirchstetter, T. W.; Wilson, K. R.; Harley, R. A.; Goldstein, A. H. Lubricating oil dominates primary organic aerosol emissions from motor vehicles. *Environ. Sci. Technol.* **2014**, *48*, (7), 3698-3706. doi:10.1021/es405375j.
22. Muhlberger, F.; Wieser, J.; Ulrich, A.; Zimmermann, R. Single photon ionization (SPI) via incoherent VUV-excimer light: Robust and compact time-of-flight mass spectrometer for on-line, real-time process gas analysis. *Anal. Chem.* **2002**, *74*, (15), 3790-3801. doi:10.1021/ac0200825.

23. Hejazi, L.; Guilhaus, M.; Hibbert, D. B.; Ebrahimi, D. Gas chromatography with parallel hard and soft ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **2015**, 29, (1), 91-99. doi:10.1002/rcm.7091.
24. DeCarlo, P. F.; Kimmel, J. R.; Trimborn, A.; Northway, M. J.; Jayne, J. T.; Aiken, A. C.; Gonin, M.; Fuhrer, K.; Horvath, T.; Docherty, K. S.; Worsnop, D. R.; Jimenez, J. L. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. *Anal. Chem.* **2006**, 78, (24), 8281-8289. doi:10.1021/ac061249n.
25. Chan, A. W. H.; Isaacman, G.; Wilson, K. R.; Worton, D. R.; Ruehl, C. R.; Nah, T.; Gentner, D. R.; Dallmann, T. R.; Kirchstetter, T. W.; Harley, R. A.; Gilman, J. B.; Kuster, W. C.; de Gouw, J. A.; Offenberg, J. H.; Kleindienst, T. E.; Lin, Y. H.; Rubitschun, C. L.; Surratt, J. D.; Hayes, P. L.; Jimenez, J. L.; Goldstein, A. H. Detailed chemical characterization of unresolved complex mixtures in atmospheric organics: Insights into emission sources, atmospheric processing, and secondary organic aerosol formation. *Journal of Geophysical Research: Atmospheres* **2013**, 118, (12), 6783-6796. doi:10.1002/jgrd.50533.
26. Schauer, J. J.; Kleeman, M. J.; Cass, G. R.; Simoneit, B. R. T. Measurement of emissions from air pollution sources. 2. C₁ through C₃₀ organic compounds from medium duty diesel trucks. *Environ. Sci. Technol.* **1999**, 33, (10), 1578-1587. doi:10.1021/es980081n.
27. Schauer, J. J.; Kleeman, M. J.; Cass, G. R.; Simoneit, B. R. T. Measurement of emissions from air pollution sources. 5. C₁–C₃₂ organic compounds from gasoline-powered motor vehicles. *Environ. Sci. Technol.* **2002**, 36, (6), 1169-1180. doi:10.1021/es0108077.
28. Kuzuyama, T.; Seto, H. Diversity of the biosynthesis of the isoprene units. *Natural Product Reports* **2003**, 20, (2), 171-183. doi:10.1039/B109860H.
29. Mao, D.; Van De Weghe, H.; Lookman, R.; Vanermen, G.; De Brucker, N.; Diels, L. Resolving the unresolved complex mixture in motor oils using high-performance liquid chromatography followed by comprehensive two-dimensional gas chromatography. *Fuel* **2009**, 88, (2), 312-318. doi:10.1016/j.fuel.2008.08.021.
30. Adam, T.; Zimmermann, R. Determination of single photon ionization cross sections for quantitative analysis of complex organic mixtures. *Analytical and Bioanalytical Chemistry* **2007**, 389, (6), 1941-1951. doi:10.1007/s00216-007-1571-x.
31. Kovats, E. Gas-chromatographische charakterisierung organischer verbindungen .1. Retentionsindices aliphatischer haolgenide, alkohole, aldehyde und ketone. *Helvetica Chimica Acta* **1958**, 41, (7), 1915-1932. doi:10.1002/hlca.19580410703.
32. Lee, M. L. Retention indices for programmed-temperature capillary-column gas chromatography of polycyclic aromatic hydrocarbons. *Anal. Chem.* **1979**, 51, (6), 768-774. doi:10.1021/ac50042a043.
33. Drozd, G. T.; Worton, D. R.; Aeppli, A.; Reddy, C. M.; Zhang, H.; Variano, E.; Goldstein, A. H. Modeling comprehensive chemical composition of weathered oil following a marine spill to predict ozone and potential secondary aerosol formation. *in review, Journal of Geophysical Research - Oceans* **2015**.

Table 1. Mass fractions (mg kg⁻¹) of components present in the NIST SRM 2779 Gulf of Mexico oil. Components are separated by carbon number (N_C) for C₉ – C₃₃ species, number of double bond equivalents (N_{DBE}) and where possible degree of branching (B0 = straight; B = all branched, B1 = one methyl group; B2; two methyl groups etc). The contributions of higher N_{DBE} components ($N_{DBE} = 7 - 15$) to $N_{DBE} = 7+$ total are given in Table S1 (Supporting Information).

N_C	N_{DBE}														
	0							1		2	3	4	5	6	7+
	B0	B1	B2	B3	B4	B5	B6	B0	B						
9	8600 ± 1200	13200 ± 1830	9910 ± 1377					4240 ± 590	25700 ± 8490	3070 ± 1020		17300 ± 5210	144 ± 53		21 ± 20
10	6790 ± 944	7920 ± 1100	7300 ± 1014					2820 ± 391	17600 ± 5810	7580 ± 1740	177 ± 113	7030 ± 2320	9270 ± 305		942 ± 160
11	6390 ± 888	4930 ± 685	5280 ± 733					2000 ± 277	13700 ± 4520	10300 ± 3420	690 ± 292	3660 ± 1210	1740 ± 656		2650 ± 895
12	5700 ± 792	3930 ± 545	3050 ± 424					1550 ± 252	10820 ± 3580	6810 ± 2250	1210 ± 435	2420 ± 800	1960 ± 649	119 ± 60	4730 ± 1510
13	5090 ± 707	2960 ± 411	3360 ± 466					1300 ± 245	9640 ± 3190	6070 ± 2010	1610 ± 544	1660 ± 619	2060 ± 680	493 ± 164	4720 ± 1620
14	4620 ± 642	2550 ± 354	2050 ± 284	2260 ± 314				1100 ± 237	8510 ± 2810	5470 ± 1810	1990 ± 656	1560 ± 573	1800 ± 595	840 ± 273	3310 ± 1890
15	4530 ± 629	2290 ± 318	1610 ± 253	1910 ± 265				996 ± 231	7720 ± 2550	5110 ± 1690	2150 ± 710	1600 ± 588	1490 ± 538	979 ± 324	4310 ± 1810
16	4290 ± 596	2090 ± 291	1440 ± 249	2170 ± 302				871 ± 223	7260 ± 2400	4640 ± 1530	2190 ± 724	1650 ± 613	1280 ± 445	1020 ± 338	3840 ± 1600
17	4010 ± 557	1730 ± 257	1030 ± 233	964 ± 229				753 ± 213	6690 ± 2210	4320 ± 1430	2150 ± 710	1640 ± 608	1220 ± 422	992 ± 329	3510 ± 1450
18	3490 ± 485	1510 ± 251	645 ± 201	1890 ± 262				658 ± 202	6350 ± 2100	4310 ± 1430	2150 ± 709	1620 ± 601	1240 ± 428	933 ± 307	3580 ± 1560
19	3380 ± 469	1370 ± 247	654 ± 202	695 ± 206	2550 ± 354			622 ± 198	5940 ± 1960	4260 ± 1410	2180 ± 720	1750 ± 662	1240 ± 428	905 ± 297	3610 ± 1600
20	3240 ± 451	1490 ± 250	712 ± 208	516 ± 182	1790 ± 258			578 ± 191	5630 ± 1860	4190 ± 1380	2120 ± 701	1640 ± 610	1250 ± 431	863 ± 282	3190 ± 1400
21	3050 ± 424	1130 ± 238	376 ± 155	307 ± 138	752 ± 212			534 ± 185	5350 ± 1770	4110 ± 1360	2070 ± 684	1770 ± 674	1240 ± 429	870 ± 284	2830 ± 1170
22	2830 ± 394	1040 ± 234	241 ± 121	205 ± 109	535 ± 185			432 ± 167	5080 ± 1680	3610 ± 1190	1960 ± 648	1720 ± 649	1160 ± 395	818 ± 266	2460 ± 968
23	2570 ± 357	887 ± 224	178 ± 100		633 ± 199			379 ± 156	4620 ± 1530	3260 ± 1080	1880 ± 620	1510 ± 549	1080 ± 365	795 ± 258	2140 ± 866
24	2450 ± 340	699 ± 207	77 ± 54		285 ± 133	467 ± 174		319 ± 142	4150 ± 1370	3030 ± 1000	1810 ± 598	1410 ± 501	1050 ± 353	778 ± 252	1420 ± 620
25	2220 ± 308	693 ± 206	89 ± 62		142 ± 86	481 ± 176		292 ± 135	3810 ± 1260	2810 ± 929	1770 ± 590	1350 ± 474	1020 ± 338	772 ± 250	1080 ± 487
26	2010 ± 280	585 ± 192				305 ± 139		251 ± 124	3460 ± 1150	2640 ± 872	1680 ± 565	1340 ± 473	982 ± 325	784 ± 254	869 ± 346
27	1690 ± 255	434 ± 168				221 ± 114		205 ± 109	3070 ± 1020	2370 ± 784	1660 ± 559	2080 ± 688	1130 ± 383	796 ± 258	133 ± 65
28	1430 ± 248	298 ± 137				326 ± 144		154 ± 91	2630 ± 869	2130 ± 703	1550 ± 528	2110 ± 697	1030 ± 345	780 ± 253	57 ± 36
29	1350 ± 246	372 ± 155				136 ± 84	252 ± 124	203 ± 108	2380 ± 785	1940 ± 640	1510 ± 518	2270 ± 751	1180 ± 382	718 ± 232	
30	1180 ± 240	325 ± 144				34 ± 30	254 ± 124	43 ± 36	2260 ± 746	1860 ± 616	1450 ± 501	1690 ± 633	1500 ± 543	748 ± 242	
31	1060 ± 235	260 ± 126					197 ± 106	69 ± 52	2160 ± 714	1770 ± 589	1250 ± 446	1260 ± 438	1160 ± 396	767 ± 248	
32	830 ± 219	196 ± 106					136 ± 83		2280 ± 753	1820 ± 601	1560 ± 532	1300 ± 454	1190 ± 409	811 ± 263	
33	660 ± 202	108 ± 71					107 ± 71		2490 ± 821	2310 ± 764	1670 ± 563	1170 ± 399	1190 ± 407	746 ± 242	

Table 2. Comparison of the eluting mass fractions (%) from this work and high temperature stimulated distillation of the gas chromatography amenable fraction of the Macondo well oil reported by Reddy et al.,¹⁷.

Eluting Fraction	Reddy et al.	This Work
before <i>n</i> -nonane (<C ₉)	15	-
<i>n</i> -nonane to <i>n</i> -undecane (C ₉ – C ₁₁)	10	13 ± 4
<i>n</i> -dodecane to <i>n</i> -octadecane (C ₁₂ – C ₁₈)	25	27 ± 8
<i>n</i> -nonadecane to <i>n</i> -triacontane (C ₁₉ – C ₃₀)	25	24 ± 8
after <i>n</i> -triacontane(>C ₃₀)	25	4 ± 2
Total	100	68 ± 22

Table 3. Comparison of the mass fractions for analyte classes from this work with those reported by Reddy et al.,¹⁷. Also shown are the calculated total releases from the Macondo well for each analyte class reported here.

Analyte		N_{DBE}	Mass Fraction (%)		Total released This Work ($\times 10^{10}$ g) ^b	
			Reddy et al.,			This Work $C_9 - C_{33}$
			$C_5 - C_{38}$	$C_9 - C_{38}$ ^a		
<i>n</i> -alkanes (B0)		0	15	$\begin{matrix} 10 \\ 9\ (C_9 - C_{33}) \end{matrix}$	8.0 ± 1.0	5.11 ± 0.64
branched alkanes		0	26	22	11 ± 2.0	7.03 ± 1.28
<i>methyl</i> (B1)					5.3 ± 0.9	3.39 ± 0.58
<i>dimethyl</i> (B2)					3.8 ± 0.6	2.43 ± 0.38
<i>trimethyl</i> (B3)					1.1 ± 0.2	0.70 ± 0.13
<i>tetramethyl</i> (B4)					0.7 ± 0.1	0.45 ± 0.06
<i>pentamethyl</i> (B5)					0.2 ± 0.1	0.13 ± 0.06
<i>hexamethyl</i> (B6)					0.1 ± 0.1	0.06 ± 0.06
cycloalkanes		0 – 6	18	14	37 ± 12	23.87 ± 7.86
<i>1 ring</i>	<i>straight alkyl chain</i> (B0)	1			2.1 ± 0.5	1.34 ± 0.32
	<i>branched alkyl chain</i> (B)	1			17 ± 5.6	10.90 ± 3.58
2 rings		2			10 ± 3.2	6.39 ± 2.04
3 rings		3			4.0 ± 1.3	2.56 ± 0.83
4 rings ^c		4			2.4 ± 1.0	1.53 ± 0.64
5 rings ^c		5			1.0 ± 0.4	0.64 ± 0.26
6 rings		6			0.8 ± 0.3	0.51 ± 0.19
monoaromatics		4 – 6	9	7	6.8 ± 2.7	4.37 ± 1.72
		4			4.0 ± 1.5	2.56 ± 0.97
		5			2.1 ± 0.8	1.35 ± 0.50
		6			0.72 ± 0.39	0.46 ± 0.25
PAHs		7+	3.9		4.7 ± 1.6	3.00 ± 1.02
		7			2.1 ± 0.68	1.34 ± 0.43
		8			0.74 ± 0.26	0.47 ± 0.17
		9			0.53 ± 0.18	0.34 ± 0.12
		10			0.62 ± 0.21	0.40 ± 0.13
		11			0.27 ± 0.11	0.17 ± 0.07
		12			0.26 ± 0.10	0.17 ± 0.06
		13			0.16 ± 0.07	0.10 ± 0.04
		14			0.04 ± 0.02	0.03 ± 0.01
		15			0.03 ± 0.02	0.02 ± 0.01

^a Data from Table S1 (supporting information) of Reddy et al.,¹⁷, corrected to remove contributions from C₅ – C₈ components listed in Table S2 (supporting information) of Reddy et al.,¹⁷.

^b Calculated with a net oil emission of 4.9 x 10⁶ barrels^{1, 2}.

^c Including the traditional sterane (N_{DBE} = 4) and hopane biomarkers (N_{DBE} = 5).

Figure 1. (a) Mass fraction distribution of the NIST SRM 2779 Gulf of Mexico crude oil as a function of carbon number (N_C) and the number of double bond equivalents (N_{DBE}). $N_{DBE} = 0$ and 1 are separated into straight and branched. For simplicity, all branched alkanes are colored red and all compounds with $N_{DBE} \geq 7$ are colored black. The data is provided in Table 1. The contribution of each compound class to the total mass for all observed carbon numbers ($C_9 - C_{33}$) is also shown. (b) Mass fraction distribution of the branched components of $N_{DBE} = 0$ (red bars from (a)) separated by number of methyl groups (B1-6) observed on the carbon chain²⁰. (c) Mass fraction distribution of the polycyclic aromatic hydrocarbon (PAH) and thiophene components with $N_{DBE} \geq 7$ (black bars from (a)) separated by N_{DBE} . This data is provided in Table S1 (supporting information).

Figure 2. Comparison of mass fractions (mg kg^{-1}) for individual compounds or groups of isomers observed in this work (listed in Table S2, supporting information) with reported values from WHOI (a)¹⁷ and NIST (b). A comparison between NIST and WHOI is also shown in (c). Data are color coded by N_{DBE} class, consistent with the color scheme used in Figure 1. Solid lines indicate 1:1 and dashed lines indicate 2:1 and 1:2.

Figure 3. Comparison of mass fractions in the NIST SRM 2779 Gulf of Mexico crude oil and Surrogate oil for $N_{DBE} = 0$ (left panel) colored by number of methyl groups (n-alkane, B0; methyl, B1; dimethyl, B2; trimethyl, B3; tetramethyl, B4, pentamethyl, B5 and hexamethyl, B6) and $N_{DBE} > 0$ colored by N_{DBE} (right panel). For simplicity, all components with $N_{DBE} \geq 7$ are colored black.