Quantification of isomerically summed hydrocarbon contributions to crude oil by carbon number, double bond equivalent, and aromaticity using gas chromatography with tunable vacuum ultraviolet ionization†

Jeremy A. Nowak,*a Robert J. Weverb and Allen H. Goldsteinb,c

The ability to structurally characterize and isomerically quantify crude oil hydrocarbons relevant to refined fuels such as motor oil, diesel, and gasoline represents an extreme challenge for chromatographic and mass spectrometric techniques. This work incorporates two-dimensional gas chromatography coupled to a tunable vacuum ultraviolet soft photoionization source, the Chemical Dynamics Beamline 9.0.2 of the Advanced Light Source at the Lawrence Berkeley National Laboratory, with a time-of-flight mass spectrometer (GC × GC-VUV-TOF) to directly characterize and isomerically sum the contributions of aromatic and aliphatic species to hydrocarbon classes of four crude oils. When the VUV beam is tuned to 10.5 ± 0.2 eV, both aromatic and aliphatic crude oil hydrocarbons are ionized to reveal the complete chemical abundance of C9–C30 hydrocarbons. When the VUV beam is tuned to 9.0 ± 0.2 eV only aromatic hydrocarbons are ionized, allowing separation of the aliphatic and aromatic fractions of the crude oil hydrocarbon chemical classes in an efficient manner while maintaining isomeric quantification. This technique provides an effective tool to determine the isomerically summed aromatic and aliphatic hydrocarbon compositions of crude oil, providing information that goes beyond typical GC × GC separations of the most dominant hydrocarbon isomers.

Introduction

Crude oil is a complex mixture containing tens of thousands of distinct organic species,1,2 dominated primarily by hydrocarbons (~90%).3 The composition of a crude oil varies depending on its origin and degree of weathering,4,5 as hydrocarbons become photooxidized, biodegraded, or subjected to abiotic environmental processes. While the range of crude oil hydrocarbons has been observed up to C80 with a double bond equivalent (NDBE)6 chemical class up to 30,2 refined fuels such as diesel, gasoline, and motor oil mostly consist of hydrocarbons below C30 in a NDBE range of 0–10.7–9 Hydrocarbons in these refined fuels increase in aromatic and aliphatic compound complexity as the carbon number (NC) and the potential number of isomers of the compound both increase, particularly as NC ≥ 9. This study therefore focuses on the characterization of those C9–C30 hydrocarbons of crude oils pertaining to the aforementioned refined fuels that are amendable to separation in our GC system.

There have been many chromatographic and mass spectrometric advancements to characterize both crude oils and their fuel derivatives as a function of hydrocarbon content. Liquid chromatographic (LC) separations distinguish resins, aromatic compounds, saturated compounds, and acidic compounds in crude oil,10,11 but are insufficient for separating compounds of increased alkylation in a higher boiling range. LC also does not isomerically quantify compounds of a specific NC and the analysis processes are often laborious, time-consuming, and prone to sample loss.

Ultrahigh resolution Fourier transform ion cyclotron mass spectrometry (FT-ICR MS) allows the identification of unique elemental compositions in crude oil between C10–C80 and NDBE = 0–30.1,2,4,5 FT-ICR MS analysis of crude oil is typically performed using electrospray ionization (ESI), which targets polar compounds,5,12 or atmospheric pressure photoionization (APPI), which targets nonpolar compounds and sulfur contai-
ing species. While these soft ionization methods retain molecular ions necessary for molecular formula identification, they have difficulty ionizing pure hydrocarbons without significant ion suppression or reduced ionization efficiency. Even when pure hydrocarbons are ionized in FT-ICR MS,12 this technique is unable to offer structural characterizations of hydrocarbons or differentiate hydrocarbon isomers in a single NDBE chemical class at a particular NC. Attempts at combining ion mobility/mass spectrometry (IM/MS) with FT-ICR MS have provided limited isomeric separation of predominantly heteroatom containing crude oil components rather than separation of pure hydrocarbons.14,15 Furthermore, FT-ICR MS relies on characterizing molecular formulas in terms of relative abundances in a mass spectrum rather than quantifying them based on mass fractions of a particular NC or NDBE.1,2,4,5

Gas chromatographic (GC) techniques are traditionally used in gasoline, diesel, and motor oil analyses when there is no necessity to characterize compounds up to C30 and NDBE = 30, as these heavier non-GC amendable species are rarely found in the aforementioned fuels.7-9 Traditional GC techniques are often limited to amendable hydrocarbons up to C30, due to the temperature limitations of columns and transfer lines (∼320 °C). GC typically couples a volatility separation of the hydrocarbons with a flame ionization detector (FID) or electron ionization mass spectrometer (EI MS).16,17 Due to the crude oil’s or fuel’s chemical complexity, much of the organic mass co-elutes as an unresolved complex mixture (UCM) in the resulting chromatogram,18 which complicates mass spectra interpretation and compound identification. EI, a hard ionization technique (∼70 eV), imparts a large excess of energy to organic molecules, resulting in fragmentation that complicates compound identification on a molecular level.16-18 GC has been combined with FT-ICR MS,19 but EI fragmented compounds >155 amu, making molecular identification of heavier compounds difficult. GC has also been previously combined with atmospheric pressure chemical ionization (APCI) and FT-ICR MS,20 but the results illustrated limited isomeric separation of hydrocarbons with the same elemental composition and retention time as well as limited characterization of structural contributions to a particular NDBE class of one NC.

Two-dimensional gas chromatography with electron ionization mass spectrometry (GC × GC-EI-MS) enhances compound separation and peak capacity by over 10-fold compared to GC-MS.31-33 The addition of a second polar column in GC × GC-EI-MS allows aromatic compounds of the same volatility as aliphatic compounds to separate in the second dimension of the chromatogram, facilitating the acquisition of cleaner mass spectra and more easily quantifiable peak areas. However, even the most advanced GC × GC separation has its limitations in the characterization of oils and fuels. As the number of carbon atoms in molecules increases, the number of possible hydrocarbon constitutional isomers becomes exponentially larger (>1045 isomers of alkanes with NC = 15 and >105 isomers of alkanes with NC = 20),24 making it difficult to resolve these individual species even with high resolution GC × GC-EI-MS. While cycloalkanes and alkylbenzenes are more easily identifiable in the second dimension of a GC × GC chromatogram than they are in a chromatogram using traditional GC-MS with EI, only the most concentrated species are quantifiable and it remains difficult to quantify the total range of isomers of a hydrocarbon class. It can therefore be challenging to use GC × GC-EI-MS to achieve a comprehensive quantitative summary of hydrocarbon distributions that takes into account the contribution of structural isomers based on volatility, polarity, and chemical structure.

Further complicating advanced GC × GC separation is the fact that at NDBE ≥ 4, both pure aromatic, pure aliphatic, and heavily alkylated aromatic hydrocarbons are present with the same chemical formulas in crude oil and fuels. It becomes increasingly difficult to separate and quantify how much material is aliphatic versus aromatic as NC and NDBE increase due to the aforementioned exponential increase in isomers and the similar polarities of heavier aliphatic compounds and heavily alkylated aromatic compounds.25

In order to quantify both the isomeric hydrocarbon content and the aromatic distributions of crude oil hydrocarbons relevant to refined fuels, this proof-of-concept study incorporates GC × GC coupled with tunable synchrotron vacuum ultraviolet photoionization and time-of-flight mass spectrometry (GC × GC-VUV-TOF).25,26 The VUV beam at the Advanced Light Source at the Lawrence Berkeley National Laboratory provides a soft ionization source (8-30 eV, less than the 70 eV ionization energy of typical “hard” electron ionization) with high photon flux (∼1016 photons per s)25 to reduce hydrocarbon fragmentation and enhance the signal of the molecular ion compared to that of EI. This technique allows complete classes of GC-amendable hydrocarbons to be integrated, summed, and characterized based on molecular weight, NC, degree of branching, and NDBE,25,26 which is represented for only hydrocarbons in this study by

\[ N_{DBE}(C_nH_y) = X - Y/2 + 1. \] (1)

Unlike other soft photoionization methods combined with GC that rely on a rare gas excimer lamp to ionize petroleum alkanes at only one ionization energy,27,28 the VUV beam in this study can be tuned to different photoionization energies. While previous studies have also incorporated tunable synchrotron VUV photoionization mass spectrometry to analyze heavy oils and aromatic crude oil hydrocarbons,29-31 those investigations did not combine chromatographic resolution with soft ionization to isomerically sum both aliphatic and aromatic hydrocarbon fractions of those C9-C30 compounds relevant to fuels. This proof-of-concept investigation combines GC × GC-MS with tunable synchrotron VUV photoionization energies to determine the NC dependent and isomerically summed aromatic and aliphatic fractions, as a function of mass loading, of GC-amendable hydrocarbons of four crude oils from different locations. VUV radiation at 10.5 eV was chosen as the ionization energy to ionize all aliphatic and aromatic hydrocarbons, thereby simultaneously maximizing the signal from the molecular ion needed for quantification and
minimizing signal interferences from fragment ions. Ionization energies greater than 10.5 eV result in greater compound fragmentation and reduced molecular ion signal, making it more difficult to accurately and isomerically sum both the aromatic and aliphatic hydrocarbon fractions. VUV radiation at 9.0 eV ionizes aromatic species without significantly ionizing aliphatic hydrocarbons, which allows the molecular ion signal of aromatic hydrocarbons to be maximized while minimizing both fragmentation of aromatic hydrocarbons and interfering ions from aliphatic compounds. Aromatic fractions, representing those compounds that contain at least one aromatic functional group, of isomerically summed hydrocarbon molecular formulas can then be determined based on the response ratios of hydrocarbon signals obtained at 9.0 eV versus 10.5 eV ionization energy ($R_{9.0/10.5}$). This tunable soft ionization method allows immediate quantification, isomeric summarization, and aromatic characterization of those crude oil hydrocarbons relevant to refined fuels. The ability to efficiently quantify and distinguish isomerically summed aliphatic from aromatic fractions of one hydrocarbon mass and $N_{DBE}$ is useful to characterize crude oil hydrocarbons for environmental modeling, fingerprinting of the hydrocarbon component of fuels for industrial applications, examining environmental weathering pathways, and understanding the impacts of oil spills in the environment.

### Experimental section

#### Oil samples

Crude oil samples were obtained from four sources: the North Sea, the Gulf of Mexico, Texas, and Azerbaijan. The oils were dissolved in chloroform (HPLC grade, Sigma-Aldrich, St Louis, MO, USA) prior to analysis. Densities were determined by weighing a known volume of the original crude oil (Mettler AB104 analytical balance, Toledo, Columbus, OH, USA). API gravities (in degrees) of the oils were calculated using the equation below, where SG is the specific gravity of crude oil:

$$\text{API Gravity} = 141.5/\text{SG-131.5}. \quad (2)$$

#### GC × GC-VUV-TOF

Oil dissolved in chloroform was directly injected, without prior fractionation or chromatographic separation, via a septumless head into a liquid nitrogen cooled inlet (−40 °C) with a fused silica liner (CIS4, Gerstel Inc, Linthicum, MD, USA). The sample was transferred to the gas chromatograph (GC, Agilent 7890, Santa Clara, CA, USA) by heating the liner to 320 °C at 10 °C s$^{-1}$. Analytes were separated based on volatility using a nonpolar column (60 m × 0.25 mm × 0.25 μm Rxi-5Sil-MS; Restek, Bellefonte, PA, USA) followed by polarity using a medium-polarity second dimension column (1 m × 0.25 mm × 0.25 μm Rtx-200MS; Restek, Bellefonte, PA, USA), with a helium gas flow rate of 2 mL min$^{-1}$. A dual-stage thermal modulator (Zoex, Houston, TX, USA) consisting of a guard column (1 m × 0.25 mm Rxi; Restek, Bellefonte, PA, USA) cryogenically focused the first column effluent and rapidly heated the analyte every 2.4 s via a 400 ms hot jet pulse into the second column. The GC program began at 40 °C, ramped at 3.5 °C min$^{-1}$ to 320 °C, and was held isothermal for a final 10 minutes. These temperatures provide an upper limit of amendable hydrocarbons of C$_{30}$. The second column was placed in a secondary oven maintained 15 °C above the main oven temperature. This system created an orthogonal separation of the analyte based on volatility in the 1st-dimension and polarity in the 2nd-dimension.

The analyte from the second column was transferred from the GC to a 170 °C ion source via a 270 °C transfer line. These temperatures reduce the thermal energy imparted to the hydrocarbons and subsequently reduce compound fragmentation. Analytes were ionized with 10.5 ± 0.2 eV or 9.0 ± 0.2 eV VUV photoionization with a photon flux of ~10$^{16}$ photons per s and detected using a time-of-flight mass spectrometer (TOF, Tofwerk, Thun, Switzerland) with $m/Δm_{50%} ≈ 4000$. The VUV beam was from the Chemical Dynamics Beamline 9.0.2 of the Advanced Light Source at the Lawrence Berkeley National Laboratory. Mass spectra were collected at 100 Hz and data processing was performed using custom code written in Igor 6.3.7 (Wavemetrics). To correct for thermal transfer efficiency of high- and low-volatility analytes through the GC column, a series of perdeuterated alkanes (Sigma-Aldrich, St Louis, MO, USA) from C$_{10}$-C$_{30}$ were used as an internal standard in 10.5 eV analyses to establish a relationship between transfer efficiency and retention time. Perdeuterated polycyclic aromatic hydrocarbons (PAHs) (Sigma-Aldrich, St Louis, MO, USA) were used to correct for thermal transfer efficiency in the 9.0 eV analyses (when alkane ionization is negligible).

#### Data visualization for isomeric summarization

Fig. S-1a† illustrates a chromatogram of molecular ion mass versus first dimension retention time of the Gulf of Mexico oil obtained using GC × GC-VUV-TOF at 10.5 eV ionization energy. In this figure the second-dimension separation is not visible. Increasing molecular mass is observed as the GC retention time increases, as is to be expected with heavier compounds of decreasing volatility. Fig. S-1b† illustrates the mass spectrum as a function of chromatographic retention time for the isomers of hydrocarbons with 20 carbon atoms in the Gulf of Mexico oil, showing their classification by mass and $N_{DBE}$ under 10.5 eV VUV ionization. $N_{DBE} = 0$ represents alkanes and $N_{DBE} = 6$ represents a mixture of hexacycloalkanes and aromatic species. PAHs of $N_{DBE} \geq 7$ were also observed, but are not shown in Fig. S-1b† for simplicity. The volatility separation based on retention time successfully distinguishes aliphatic compounds and PAHs of equal molecular mass. Additionally, hydrocarbon volatility of $N_{DBE} = 0$ compounds increases with branching of the straight-chain alkane, therefore branched isomers elute earlier in the chromatogram than the straight-chain hydrocarbon. The enhancement of the molecular ion under VUV ionization, combined with chromatographic separation by volatility, distinguishes the branched isomers from the straight-chain alkane of the same $N_C$. Fig. S-1c† demon-
strates this concept by showing the single ion chromatogram (SIC) of pentadecane \((m/Q = 212, C_{15}H_{32})\) of the Gulf of Mexico oil, where the straight-chain alkane and its branched isomers are clearly distinguished.

In addition to retaining molecular ions necessary for hydrocarbon classification and quantification, Fig. S-2† illustrates the advantages of GC × GC-VUV-TOF over traditional GC × GC-EI-MS. Fig. S-2a and S-2b† illustrate the complete GC × GC-EI-TOF chromatogram at 70 eV ionization energy and the complete GC × GC-VUV-TOF chromatogram at 10.5 eV ionization energy, respectively, of mass versus first dimension retention time of the Gulf of Mexico oil. While the isomers of heavier molecular weight and lower volatility compounds are observable and quantifiable at 10.5 eV, they are not quantifiable at 70 eV. Fragmentation of hydrocarbon isomers is also increased in the chromatogram at 70 eV, as illustrated by the increase in signal between \(m/Q = 100–200\). Fig. S-2c† further illustrates the advantages of GC × GC-VUV-TOF over traditional GC × GC-EI-MS by depicting \(m/Q = 326\) (\(C_{24}H_{38}, N_{DBE} = 6\)) hydrocarbon isomers using the two techniques. In the GC × GC-EI-MS total ion chromatogram in Fig. S-2c† only a few most concentrated isomers are identifiable as individual peaks corresponding to peaks in the single ion GC × GC-VUV-TOF chromatogram of \(m/Q = 326\). Most of the signal for \(C_{24}H_{38}\) is from the extremely large number of unresolved isomers eluting between 55 and 70 minutes, which are not separable using GC × GC-EI-MS. It is extremely difficult to separate or resolve the numerous isomers present at such lower concentrations (>10⁷ isomers could be present with \(N_C = 20\)). It is difficult, time consuming, and perhaps impossible to isomerically separate and quantify all of these hydrocarbons, especially as \(N_C\) further increases, to achieve complete chemical characterization of a compound class. GC × GC-VUV-TOF represents these \(C_{24}H_{38}\) compounds in Fig. S-2c† as a single chromatographic trace using the molecular ion. By integrating across the trace and summing the signal of the complete range of isomers of a particular hydrocarbon molecular ion, we achieve the ability to quantify the complete mass of these isomers. This methodology extends across all GC-amendable molecular weights and compound classes, allowing us to isomerically summed compounds more easily and to a greater analytical extent than even in the most advanced GC × GC-EI-MS separation.

**Differences in chromatograms at 9.0 eV and 10.5 eV**

All aliphatic and aromatic compounds are ionized using GC × GC-VUV-TOF with 10.5 eV VUV energy to reveal the chemical composition of crude oil hydrocarbons from \(C_9-C_{30}\) with \(N_{DBE} = 0–10\). Under 9.0 eV ionization energy, only the aromatic species are ionized. Fig. 1 shows the mass spectrum as a function of chromatographic retention time for the isomers of hydrocarbons with 20 carbons atoms (\(C_{20}\)) in the Gulf of Mexico oil, showing their classification by mass and \(N_{DBE}\) under 10.5 and 9.0 eV VUV ionization (1a and 1b, respectively). All \(N_{DBE} = 0–10\) hydrocarbons are ionized in the 10.5 eV chromatogram, while only \(N_{DBE} = 4–6\) aromatic compounds and PAHs (\(N_{DBE} = 7–10\), not shown here for simplicity) are ionized in the 9.0 eV chromatogram. There is some detection of aliphatic compounds at 9.0 eV, but the ionization efficiency is <5% of that at 10.5 eV and is therefore considered negligible in our analyses. Fig. S-3† shows two-dimensional chromatograms of a known standard hydrocarbon mixture to further illustrate the use of tunable VUV to selectively ionize hydrocarbons depending on molecular structure.

**Calculating \(N_C\)-dependent aromatic fractions of \(N_{DBE}\) classes**

In this proof-of-concept study, aromatic fraction is defined as the fraction of isomers of a hydrocarbon molecular formula, based on \(N_C\) and \(N_{DBE}\), that contain an aromatic functional group. An aromatic fraction of 1 represents isomerically summed species that contain no aliphatic character, such as methyl naphthalenes or anthracene. We assumed that compounds of \(N_{DBE} = 4\) with fewer than 14 carbon atoms are entirely aromatic (aromatic fraction = 1). As the aromatic fraction decreases from 1, aliphatic contributions to an isomerically summed compound increase in the form of alkyl side chains or pure cycloalkanes. Isomers of a hydrocarbon molecular formula with an aromatic fraction <1 therefore can consist of pure aromatic compounds, aliphatic compounds, mixtures of compounds containing fused aromatic and cyclic rings, etc. An aromatic fraction of 0 represents an isomerically summed molecular formula that contains no aromaticity, and...
The response ratio of hydrocarbon signal responses obtained by integrating across a molecular formula at 9.0 eV versus 10.5 eV ionization energy (\(R_{9.0/10.5}\)) is indicative of the aromatic fraction in a specific \(N_{DBE}\) as a function of \(N_C\). It should be noted that \(R_{9.0/10.5} \neq 1\) even if all isomers of a molecular formula are completely aromatic. For instance, \(C_{10}H_{14}\) represents isomers of alkylated benzenes with ten carbon atoms. While a combination of double bonded cyclic compounds could theoretically possess a chemical formula of \(C_{10}H_{14}\), the likelihood of finding such compounds in crude oils or refined fuels is low, as they do not generally contain alkenes.\(^7\,^{35}\) \(R_{9.0/10.5}\) of these \(C_{10}H_{14}\) alkylated benzenes is still less than 1 due to the decreased photoionization efficiencies of aromatic hydrocarbons at 9.0 versus 10.5 eV ionization energies.\(^33\) As the length of the alkyl side chain group in heavier \(N_{DBE} = 4–6\) compounds increases, there is an increasing trend in ionization energy that leads to a lower response at 9.0 eV and thus a lower \(R_{9.0/10.5}\).\(^32\,^{33}\) Therefore, there is a \(N_C\) dependence in \(R_{9.0/10.5}\) for alkylated aromatic compounds.

To calibrate for \(R_{9.0/10.5}\) of the hydrocarbons as a function of \(N_C\) and \(N_{DBE}\) in the oils, authentic standards (Sigma-Aldrich, St Louis, MO, USA), including alkanes, branched alkanes, cycloalkanes, alkylbenzenes, hopanes, steranes, PAHs, and alkylated PAHs were analyzed at 9.0 eV and 10.5 eV. The ratio of these hydrocarbon signals is referred to as \(R_{standard}\). Fig. 2 shows \(R_{9.0/10.5}\) for a standard mixture of hydrocarbons consisting of the aforementioned species. Compounds with \(N_{DBE} < 4\) in this standard mixture are not aromatic and have a \(R_{9.0/10.5} < 0.05\), corroborating the fact that the photoinization probability of aliphatic compounds is extremely low at 9.0 eV. \(R_{9.0/10.5}\) for alkylbenzenes decreases with increasing \(N_C\) and becomes indistinguishable from \(R_{9.0/10.5}\) for tetracyclic steranes above \(C_{25}\). PAHs have a \(R_{9.0/10.5}\) of 0.45–0.50 independent of molecular weight. To estimate the response ratios between 9.0 eV and 10.5 eV energies for purely aliphatic components of \(N_{DBE} = 4, 5, \text{ and } 6\) (\(R_{aliphatic}\)), we extrapolated the trend in response ratios observed in \(N_{DBE} = 0, 1, 2, \text{ and } 3\), which cannot include aromatic rings. Using the \(N_C\) dependent values for the aromatic and aliphatic hydrocarbon signal responses, the aromatic fraction of an isomerically summed molecular formula is calculated as

\[
\text{Aromatic Fraction} = \frac{R_{9.0/10.5} - R_{aliphatic}}{R_{standard}}
\]

when \(N_C\) in an isomerically summed \(N_{DBE} = 4\) molecular formula exceeds 25, the aromatic hydrocarbon isomers have a similar \(R_{9.0/10.5}\) to that of tetracyclic steranes and therefore the aromatic fraction cannot be discerned by this method.

### Results and discussion

#### Crude oil chemical composition

The calculated densities and API gravities of the four crude oils used in this study are given in Table 1. Based on the API gravities, the Azerbaijan and Texas oils are classified as light and the Gulf of Mexico and North Sea oils are medium-heavy. All four oils were initially analyzed using GC × GC-VUV-TOF with 10.5 eV photoionization energy. Fig. 3 illustrates the distribution of each oil’s GC-amendable isomerically summed hydrocarbon mass fractions (measured in milligram per kilogram injected oil) as a function of both \(N_C\) and \(N_{DBE}\).\(^{~60–65\%}\) of the measured mass fractions of the hydrocarbons in the Texas and Azerbaijan oils are between the \(C_9–C_{15}\) hydrocarbons. The mass fraction distributions of the hydrocarbons in the North Sea and Gulf of Mexico oils are similar, with \(~50\%\) of the mass fractions split between \(C_9–C_{15}\) and \(C_{20–C_{30}}\) species. The Texas and Azerbaijan oils contain \(~70\%\) of the total measured hydrocarbon mass fraction within completely aliphatic compound classes, \(N_{DBE} = 0–3\). Only \(~60\%\) of the total measured hydrocarbon mass fractions of the North Sea and Gulf of Mexico oils belongs to \(N_{DBE} = 0–3\).

\(\text{GC} \times \text{GC-VUV-TOF} \text{ at } 10.5 \text{ eV ionization energy also provides separation of crude oil alkanes based on their degree of branching. Alkanes can be classified by retention time based on the degree of branching } B_x, \text{ where } x \text{ represents the number of alkyl branches from methyl (B1) to hexamethyl (B6). For simplicity, each branch is assumed to represent a methyl group, rather than ethyl, propyl, etc. Fig. S-4† illustrates the isomerically summed mass fractions of the branched alkanes in each of the oils as a function of } N_C \text{ and degree of branching.}\)

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Density (g cm⁻³)</th>
<th>API gravity (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf of Mexico oil</td>
<td>∼0.92</td>
<td>∼22.0</td>
</tr>
<tr>
<td>North Sea oil</td>
<td>∼0.91</td>
<td>∼22.5</td>
</tr>
<tr>
<td>Azerbaijan oil</td>
<td>∼0.85</td>
<td>∼35.0</td>
</tr>
<tr>
<td>Texas oil</td>
<td>∼0.83</td>
<td>∼40.0</td>
</tr>
</tbody>
</table>
B1 branched alkanes dominate the branched alkanes in the Texas oil. The branched signatures of both the Gulf of Mexico and North Sea oils exhibit similar mass distributions, corroborating the similarities in the hydrocarbon compositions of these two oils seen in Fig. 1. The Azerbaijan oil has a similar branched alkane distribution to that of the North Sea and Gulf of Mexico oils, but the Azerbaijan oil contains substantially higher mass fractions of B4–B6 branched alkanes between C19–C30.

Response ratios at 9.0 eV and 10.5 eV ionization energies

Fig. S-5, S-6, and S-7† display overlayed SICs with arbitrary units of intensity at 9.0 eV and 10.5 eV ionization energies for C18H30 (NDBE = 4), C20H32 (NDBE = 5), and C20H30 (NDBE = 6), respectively, for all four oils. The signal peaks that are present in the 10.5 eV SICs, but not in the 9.0 eV SICs, represent aliphatic isomers that are not ionized at 9.0 eV. R9.0/10.5 of the integrated signal of the summed isomers of NDBE = 4, 5, and 6 hydrocarbons was calculated for each oil and plotted as a function of NC in Fig. S-8† R9.0/10.5 decreases with increasing molecular weight independent of an oil’s source, indicating that aliphatic contributions in the form of alkyl groups, cyclic rings, etc. to these isomerically summed compound classes increase with NC.

As previously discussed and illustrated, traditional GC × GC-EI-MS separates aliphatic and aromatic compounds based on both volatility and polarity, but the exponential increase in the number of constitutional isomers as NC increases makes quantification of the whole chemical class by NC and NDBE challenging and time-consuming. Only concentrated compounds are identifiable and quantifiable in GC × GC space as NC increases, therefore it becomes increasingly difficult to isomerically sum and distinguish aliphatic compounds from aromatic compounds as molecular weight increases. In order to test our structural separation method based on ionization energies, Table S-1† displays the ratio of aromatic components to the total aromatic plus aliphatic components of nine isomerically summed hydrocarbon molecular formulas in the North Sea oil, spanning a range of molecular weights. This ratio was calculated using the second-dimension separation of GC × GC-EI-MS and tunable GC × GC-VUV-TOF. As NC grows from NC = 16 to NC = 24 in NDBE = 4, 5, and 6, it becomes difficult to separate and quantify isomers of a given hydrocarbon. The calculated ratios using the GC × GC-EI-MS separation differ from those calculated using tunable GC × GC-VUV-TOF by ~35–65% as NC increases. This comparison of experimental techniques illustrates the increased fraction of isomers which are not measured using GC × GC-EI-MS as the oil’s chemical complexity increases. The mass closure afforded by summing the isomers of hydrocarbons of higher molecular weights, as well as the structural separations afforded by the tunable ionization energies, present tunable GC × GC-VUV-TOF as an analytical tool that can efficiently and quantitatively characterize sums of hydrocarbon isomers found in crude oil to a greater extent than GC × GC-EI-MS separation can alone.

Fig. 3 Distribution of GC-amendable isomerically summed hydrocarbon mass fractions as a function of both the number of carbon atoms and double bond equivalents (NDBE) for oils from (a) Azerbaijan, (b) North Sea, (c) Texas, and (d) Gulf of Mexico.
$N_{\text{DBE}} = 7$–10 response ratios

At 10.5 eV, the SICs of molecular weights pertaining to $N_{\text{DBE}} = 8$ hydrocarbons contain fragments of larger aliphatic hydrocarbons, making it difficult to discern the total pure $N_{\text{DBE}} = 8$ hydrocarbon content in the GC × GC-VUV-TOF chromatogram. Fig. S-9† illustrates this concept using the SIC of $m/Q = 196$ (C$_{15}$H$_{16} N_{\text{DBE}} = 8$) at 10.5 eV. Portions of the chromatographic trace align with retention times of heavier straight-chain and branched alkanes. The C$_{15}$H$_{16}$ molecular mass also corresponds with fragment ions from alkanes such as C$_{17}$H$_{36}$ and C$_{20}$H$_{42}$, following the loss of hydrogens and alkyl groups. Therefore at 10.5 eV, the chromatographic trace of C$_{15}$H$_{16}$ includes signal from C$_{15}$H$_{16}$ isomers and aliphatic fragments of higher molecular weight species. At 9.0 eV, these aliphatic compounds are not ionized and the interference from fragment ions is negligible. While the 9.0 eV data can detect the total aromatic contribution to a $N_{\text{DBE}} = 8$ molecular formula, an aromatic fraction of this isomerically summed molecular formula cannot be calculated due to the interfering fragment ions in the 10.5 eV signal.

Fig. 4b further illustrates this concept by overlaying the Gulf of Mexico oil’s C$_{15}$H$_{16}$ ($N_{\text{DBE}} = 8$) SICs (with arbitrary units of intensity) at both 10.5 eV and 9.0 eV ionization energies. Branched aliphatic hydrocarbons fragment more than straight-chain aliphatic compounds, thus the larger fragments in the 10.5 eV trace in Fig. 4b correspond to fragments of branched compounds of a higher molecular weight whose retention times overlap with those of the lower molecular weight C$_{15}$H$_{16}$ isomers. At 9.0 eV, the aliphatic compounds are not ionized and the $N_{\text{DBE}} = 8$ SIC consists purely of aromatic hydrocarbons with no observable interferences from fragmentation. This observation was consistent in all molecular formulas pertaining to $N_{\text{DBE}} = 8$ hydrocarbons in all four oils.

Fig. 4  Single ion chromatograms (SICs) with arbitrary units of intensity (A.U.) of Gulf of Mexico oil for (a) C$_{15}$H$_{18}$ ($N_{\text{DBE}} = 7$), (b) C$_{15}$H$_{16}$ ($N_{\text{DBE}} = 8$), (c) C$_{15}$H$_{14}$ ($N_{\text{DBE}} = 9$), and (d) C$_{15}$H$_{12}$ ($N_{\text{DBE}} = 10$) at both 10.5 and 9.0 eV ionization energies. The C$_{15}$H$_{16}$ $N_{\text{DBE}} = 8$ SIC at 10.5 eV includes signal from C$_{15}$H$_{16}$ isomers and aliphatic fragments of higher molecular weight species (e.g. C$_{17}$H$_{36}$, C$_{19}$H$_{40}$, C$_{20}$H$_{42}$) with overlapping retention times. The $N_{\text{DBE}} = 7$, 9, and 10 SICs at 10.5 eV contain no observable interferences from fragmentation.
Fig. 4a, c, and d display the overlapping 9.0 eV and 10.5 eV SICs (with arbitrary units of intensity) for C_{15}H_{18} (N\textsubscript{DBE} = 7), C_{13}H_{14} (N\textsubscript{DBE} = 9), and C_{15}H_{12} (N\textsubscript{DBE} = 10) of the Gulf of Mexico oil. Unlike in the SIC of C_{15}H_{16} at 10.5 eV, there were no observable interferences from fragmentation of heavier aliphatic compounds in these PAH SICs at 10.5 eV and R\textsubscript{9.0/10.5} values could be estimated. This observation was consistent in all molecular formulas pertaining to N\textsubscript{DBE} = 7, 9, and 10 compounds in all four oils. Fig. S-10† shows R\textsubscript{9.0/10.5} of the integrated signal of the summed isomers of N\textsubscript{DBE} = 7, 9, and 10 compounds for all four oils as a function of N\textsubscript{C}. All of the oils have a R\textsubscript{9.0/10.5} of ~0.4–0.6 in the aforementioned compound classes independent of N\textsubscript{C}. There is no decrease in R\textsubscript{9.0/10.5} as was observed in the N\textsubscript{DBE} = 4, 5, and 6 R\textsubscript{9.0/10.5}, implying that N\textsubscript{DBE} = 7, 9, and 10 hydrocarbons in these oils consist of pure PAHs with minimal aliphatic isomers.

**Aromatic fraction calculations**

The measured R\textsubscript{9.0/10.5} of N\textsubscript{DBE} = 4, 5, and 6 hydrocarbons, along with R\textsubscript{aliphatic} and R\textsubscript{standard}, is used to calculate the N\textsubscript{C}-dependent isomerically summed aromatic fraction of each N\textsubscript{DBE} in each oil using eqn (3).

Fig. 5a–c present the isomerically summed aromatic fractions of N\textsubscript{DBE} = 4, 5, and 6 hydrocarbons of each oil as a function of N\textsubscript{C}. The Gulf of Mexico oil contains the largest fraction of isomerically summed aromatic hydrocarbons of the four oils used in this study across all N\textsubscript{C} in N\textsubscript{DBE} = 4–6. This oil contains ~1.5–2.0 times higher fractions of isomerically summed aromatic material across the N\textsubscript{DBE} = 4, 5, and 6 compound classes when compared to those of the other three oils. The Azerbaijan oil has the lowest isomerically summed aromatic fraction in the N\textsubscript{DBE} = 4 class, while the Texas oil has a ~1.5–2.0 times higher isomerically summed aromatic fraction than that of the North Sea oil in the C\textsubscript{15}–C\textsubscript{25} range of N\textsubscript{DBE} = 4. The North Sea, Texas, and Azerbaijan oils all possess similar isomerically summed aromatic fractions at N\textsubscript{DBE} = 5 across all N\textsubscript{C}. At N\textsubscript{DBE} = 6, the isomerically summed aromatic fraction from C\textsubscript{15}–C\textsubscript{25} in the North Sea oil is ~1.3–1.6 times that of the Texas and Azerbaijan oils, revealing that the North Sea oil is structurally similar to the Gulf of Mexico oil in the N\textsubscript{DBE} = 6 class and consists of more aromatic material than the Texas and Azerbaijan oils in the N\textsubscript{DBE} = 6 class.

This study shows that as molecular weight increases in N\textsubscript{DBE} = 4–6, the fraction of isomers of a hydrocarbon molecular formula that are purely aromatic decreases. This pattern implies that hydrocarbon structures consisting of benzenes or alkylated benzenes decrease in concentration as N\textsubscript{C} grows across N\textsubscript{DBE} = 4–6, and structures consisting of fused cyclic rings increase in concentration in the same scenario. It can therefore be concluded from this study that in N\textsubscript{DBE} = 4–6, isomerically summed heavier hydrocarbons are more likely to consist of fused cyclic rings and cycloalkanes than pure alkylated benzenes or a combination of aromatic functional groups and cyclic rings. Therefore, aliphatic contributions to the N\textsubscript{DBE} = 4–6 compound classes in all of these oils become more prevalent as molecular weight increases across one isomerically summed compound class.

**Conclusions**

Tunable soft VUV photoionization coupled to GC × GC-MS offers a fast, efficient, and unique analytical technique to further characterize and quantify the chemical complexity of crude oil hydrocarbons in the molecular range relevant to
refined fuels such as gasoline, diesel, and motor oil beyond that of traditional GC × GC-EI-MS separation. Soft ionization at 10.5 eV allows the full range of isomeric hydrocarbons to be summed as a function of $N_C$, $N_{DBE}$ chemical class, and the amount of branching. By altering the energy of the VUV beam between 10.5 eV and 9.0 eV, we successfully ionize hydrocarbon isomers with $N_{DBE} \geq 4$ into aliphatic and aromatic fractions as a function of $N_C$ while maintaining the ability to isomerically sum mass fractions. This study provides a novel chromatographic and mass spectrometric tool to constrain and strengthen existing models of crude oil hydrocarbon characterization with applications to oil spills, microbial transformations, and environmental impacts. Tunable soft ionization is therefore a powerful analytical tool which can now be applied to isomerically sum and further characterize crude oil hydrocarbons beyond that of traditional compound class and carbon number labels.

**Conflicts of interest**

There are no conflicts to declare.

**Acknowledgements**

We acknowledge the Energy Biosciences Institute (EBI) for funding this work under project 86217. We also acknowledge Kevin Wilson and Bruce Rude for their assistance at the Chemical Dynamics Beamline 9.0.2 at the Advanced Light Source, Coty Jen for GC × GC-VUV-TOF instrumental setup and transport, and Professor John Coates for access to the oils and advice.

**References**

18. H. K. White, L. Xu, P. Hartmann, J. G. Quinn and C. M. Reddy, Unresolved complex mixture (UCM) in coastal