Available online at www.sciencedirect.com
ScienceDirect
Fuel 86 (2007) 758–768
FUEL
www.fuelfirst.com

Detailed compositional comparison of acidic NSO compounds in biodegraded reservoir and surface crude oils by negative ion electrospray Fourier transform ion cyclotron resonance mass spectrometry

Christine A. Hughey [a,*], Samantha A. Galasso [a], John E. Zumberge [b]

[a] *Department of Physical Sciences, Chapman University, 1 University Drive, Orange, CA 92866, USA*
[b] *GeoMark, Ltd., 9748 Whithorn Drive, Houston, TX 77095, USA*

Received 18 March 2006; received in revised form 30 August 2006; accepted 31 August 2006
Available online 27 September 2006

Abstract

Traditional hydrocarbon biomarker analyses determined the degree of biodegradation in two reservoir and two surface oils. These data were then correlated to the distribution and type (rings plus double bonds) of acidic NSO species selectively ionized and mass resolved by negative ion electrospray Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS) to determine if oxygenated polars could be used to estimate the degree of biodegradation in crude-oil contaminated soil (surface) samples. Since the histories of the surface samples were unknown (as it often the case for environmental samples), non-degraded and severely degraded reservoir oils were used as compositional benchmarks. The biodegraded reservoir crude oil exhibited an increase in relative abundance of O_2-species, a decrease in acyclic fatty acids, an increase in multi-ring naphthenic acids and a decrease in C_{32} hopanoic acids compared to the non-biodegraded reservoir crude oil. The surface oils exhibited trends similar to the biodegraded reservoir oil, indicating that the surface oils were biodegraded in the reservoir, the environment or both. However, one surface sample also exhibited biomarker signatures indicative of a non-degraded oil. This evidence, in combination with the ESI FT-ICR MS data, suggests that the sample was contaminated with a mixture of biodegraded and non-degraded oils and demonstrates how analysis of hydrocarbon and polar fractions can reveal complex histories of surface contamination. This work is the first detailed compositional analysis of acidic NSO species in crude oil soil extracts and lays the foundation for our understanding of how surface biodegradation effects the composition of polars.

Keywords: Microbial alteration; Environment; Petroleum acids

1. Introduction

Years of research has shown that weathering and biodegradation changes crude oil. Volatile, low molecular weight and water soluble hydrocarbons are first lost during weathering. Microorganisms may then catabolize petroleum compounds over a much longer period resulting in the selective removal of specific hydrocarbons. For example, linear alkanes are biodegraded faster than branched alkanes, cyloalkanes or aromatics. Little knowledge exists concerning the recalcitrance or modification of the polar components by these same processes. The data that are available come from geochemical research of crude oil reservoirs—not from surface seeps, spills or contamination. Polar compounds such as phenols, acids, and carbazoles have been used as indicators of oil's source, thermal maturity, migration and microbial alteration. It is reasonable to assume that these compounds also may provide clues to processes that influence compositional change in surface environments. Carboxylic acids are of particular interest as they result from the microbial oxidation of hydrocarbons

* Corresponding author. Tel.: +1 714 628 7346; fax: +1 714 532 6048.
E-mail address: hughey@chapman.edu (C.A. Hughey).

0016-2361/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.fuel.2006.08.029

[1–3]. Watson et al. [2] found in laboratory biodegradation studies that once formed, lower molecular weight acids (C_{10}–C_{20}) were rapidly biodegraded, while higher molecular weight acids (>C_{20}) were recalcitrant. These recalcitrant acids could potentially be used as molecular monitors of the biodegradation process in environmental spills. In addition, once polars have been identified in environmental samples, more information can be gained regarding their fate, persistence and potential toxicity in the environment.

Unfortunately, the characterization of polar compounds is challenging because (1) they typically make-up a small fraction of the petroleum sample (<15% by mass for crude oils), (2) they require time-consuming chromatographic isolation prior to analysis, (3) they are, once isolated, too complex to be resolved by traditional chromatographic or mass spectrometric techniques, and (4) their non-volatile nature excludes many from analysis by GC and ionization by traditional mass spectrometric ionization sources. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry overcomes many of the above-described challenges. Negative ion electrospray selectively ionizes acidic polar heteroatomic compounds (e.g., acids, phenols and carbazoles) from the predominately hydrocarbon matrix, which eliminates the need for chromatographic isolation prior to analysis [4–8]. In addition, FT-ICR MS routinely affords ultrahigh mass resolving powers $m/\Delta m_{50\%} > 400{,}000$ (in which $\Delta m_{50\%}$ denotes mass spectral full width at half-maximum height) and mass accuracy <0.5 ppm (with internal calibration), allowing the molecular formula assignment of 10,000+ ions in a single mass spectrum (250–800 Da) [9]. From the assignment of elemental formulas comes the component's class (heteroatom content), degree of alkylation, and type (rings plus double bonds). Furthermore, peak assignment is facilitated by conversion of mass spectrometric data from the IUPAC mass scale ($CH_2 = 14.0156$ Da) to the Kendrick mass scale [10] ($CH_2 = 14.000$) and subsequent sorting by Kendrick mass defect (difference between exact and nearest-integer Kendrick mass). Such data treatment allows for easy identification of homologous series (e.g., series in which compounds have the same number of heteroatoms and rings plus double bonds, but differ by the number of CH_2 groups) [11,12]. Once data are sorted and peaks are identified, compositional differences can be displayed easily in Kendrick mass defect plots [11] and van Krevelen diagrams [13–15] to allow rapid compositional comparisons between samples. Such detailed compositional analyses have been conducted by ESI FT-ICR MS for petroleum distillates [16,17], crude oils [5,6,9,18,19], and coal extracts [20,21]. This information has proven valuable to the petroleum industry in both upstream exploration by improving predictions of oil quality and in downstream refining where the goal is to increase the yield of marketable products and reduce the negative impacts of the polar components (e.g., liquid phase corrosion from acids, inhibition of catalysts by basic nitrogen compounds, etc.). ESI FT-ICR MS also has been used for environmental analysis, mainly for the compositional analysis of natural organic matter (e.g., humic and fulvic acids) [22–26]. Prior to the use of ESI FT-ICR MS, electron ionization FT-ICR MS was used to determine compositional changes of hydrocarbons in jet fuel as a result of weathering and biodegradation [27]. To date, detailed compositional analyses have not been reported on the polar components of crude oil from soil extracts.

Here we present the first detailed compositional analysis of acidic polars in crude oil contaminated soils. Two soil samples were collected from sites in Southern California and exposed to environmental conditions for different time periods. Traditional terpane and sterane biomarker analyses were conducted to determine the degree of biodegradation in the surface samples. The acidic polar compositions were then correlated to the biomarker data to determine if polars alone provide insight into surface biodegradation processes. Since the history of these crude oils is unknown, their hydrocarbon and acidic polar compositions were compared to two reservoir crude oils of known history. One reservoir crude oil is a light, non-biodegraded crude oil from the Urals (Russia). The other reservoir crude oil is a severely biodegraded on-shore crude oil from California. The acidic polar composition of these oils is representative of non-biodegraded and biodegraded oils previously analyzed by negative ion ESI FT-ICR MS [5,6,18]. Acidic polar composition comparisons are made by class, type and alkylation distributions. In addition, we investigate the robustness of a new biodegradation index recently proposed by Kim et al. [28]. This index, based on the ratio of acyclic to cyclic acids, indicates that the relative abundance of acyclic acids decreases as the degree of microbial degradation increases. This work lays the foundation for our understanding of the compositional changes that occur in the polar fraction when oil is exposed to surface environments and will direct future work in this area.

2. Experimental

2.1. Surface crude oils

Crude oil contaminated soil was collected from two sites in Los Angeles County, CA, one at the L.A. port in Long Beach and the second in El Segundo. The Long Beach (LB) sample is the result of subsurface contamination from a pipeline that transported crude oil. At the time of sample collection, a retention pond existed on the site. Oil seeped through the water column of the pond to form a sheen on the water's surface. Soil samples were collected at the edge of the pond in July 2003. The E.P.A. estimates that the site had been contaminated with petroleum for 10+ years [29]. The El Segundo (ES) contamination is the result of an oil well blowout in May 2003. The well, originally drilled in 1937, produced low-grade oil with a high water content until 1996. The blowout released approximately 11,000 barrels of oil, which contaminated the top 5' of

the soil. Samples were collected in July 2003, prior to site clean-up in August 2003 [30].

Crude oil was extracted from the soil as described by Wang et al. [31]. Briefly, a 30 mL aliquot of hexanes and a 30 mL aliquot of dichloromethane (DCM) were added to a flask containing approximately 10 g of contaminated soil and 10 g anhydrous Na_2SO_4. The mixture was sonicated for 30 min. and was then vacuum filtered with a Büchner funnel. Sonication was continued with 60 mL aliquots of DCM until the filtrate was colorless. Fractions were consolidated, evaporated with a rotary evaporator, then further evaporated with dry N_2 to constant mass. The LB sample contained approximately 64.8 ± 7.0 mg oil/g soil; the ES sample contained approximately 39.6 ± 1.8 mg oil/g soil.

2.2. Reservoir crude oils

The non-degraded Urals (UR) and severely degraded Midway Sunset (MWS) reservoir crude oils were obtained from ExxonMobil Research and Engineering (Annandale, NJ) and analyzed as received.

2.3. Biomarker analyses

Oils were analyzed by conventional geochemical methods. Briefly, the C_{15+} deasphalted fractions were separated into saturate, aromatic and resin fractions by silica gel column chromatography. GC/MS in SIM mode was used to identify and quantitate biomarkers (e.g., steranes and terpanes) in the C_{15+} branched/cyclic and aromatic hydrocarbon fractions. Stable carbon isotopic compositions of the fractions and whole crude oils were also determined. Analyses are described in more detail by Zumberge et al. [32].

2.4. Sample preparation and analysis by ESI FT-ICR MS

Crude oils were dissolved in toluene at a concentration of 1 mg/mL. A 1 mL aliquot of this solution was diluted with 1 mL of methanol, then spiked with 10 μL of 35% v/v conc. NH_4OH. The base was added to facilitate deprotonation of acidic compounds by negative ion electrospray.

The crude oil was analyzed using a homebuilt 9.4 T FT-ICR mass spectrometer, housed at the National High Magnetic Field Laboratory in Tallahassee, FL. Data were collected and processed with a modular ICR data acquisition system (MIDAS) [33,34]. Ions were generated by negative ion electrospray equipped with a 50 μm i.d. fused silica ESI needle. Samples were infused at a flow rate of 400 nL/min. Typical ESI conditions were as follows: needle voltage, −2.5 kV, tube lens, −350 V, and heated capillary current, 4.8 A. Ions were accumulated externally in the first linear octapole ion trap for 10 s and transferred through the quadrupole and octapoles, both operated in rf-only mode. Multipoles [35] were operated at 1.5 MHz (front and middle octapole) and 2.0 MHz (last octapole). One hundred co-added time domain data sets were converted to mass-to-charge ratio to generate the spectra shown in Figs. 1 and 3.

2.5. LTQ mass analysis

The transmission of ions from an external ion source to the ICR cell imposes constraints on m/z range arising from ion stability in multipole ion guides and traps as well as time-of-flight dispersion. As a result, mass spectra of the oil extracts were collected on a ThermoFinnigan LTQ, which is not limited by m/z range constraints, to confirm the mass range of the samples. Both LTQ spectra exhibited a mass range from 100 to 1000 Da (data not shown). The ICR mass range spanned 250–800 Da. Therefore, the lower

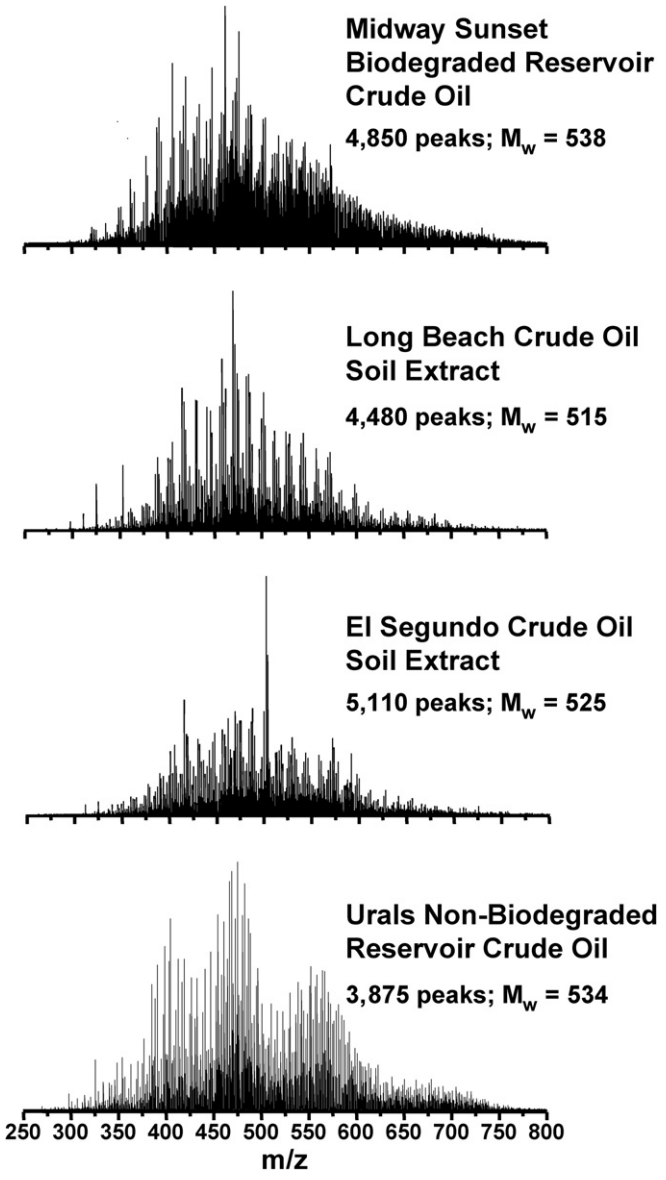

Fig. 1. Broadband negative ion ESI FT-ICR mass spectra of the reservoir (MWS and UR) and surface (LB and ES) crude oils. Peaks identified have a threshold greater than 5σ the baseline noise. M_w represents the weight-average molecular weight.

(<250 Da) and higher mass range (>800 Da) components are excluded from the FT-ICR analyses discussed below. However, the species excluded in the FT-ICR spectra were present at low relative abundance in the LTQ spectra. The most abundant species in the LTQ span the same mass range as observed in the FT-ICR spectra.

2.6. Mass calibration of FT-ICR MS data

Mass spectra were initially externally calibrated with respect to a negative ion electrospray "tuning mix" (Agilent # G2421A, Palo Alto, CA). The mass spectra were then internally calibrated (with an error of less than 0.2 ppm) with respect to a homologous series within each crude oil spanning a mass range of 250–800 Da. All species present in the mass spectra are singly charged as evidenced by the ~1 Da spacing between each monoisotopic species and its corresponding nuclide containing one ^{13}C atom.

2.7. Data analysis

The mass values for (singly charged) ions of 250–800 Da and peak height greater than five times the standard deviation of the noise were collected. Measured masses were converted from the IUPAC mass scale ($CH_2 = 14.01565$ Da) to the Kendrick mass scale ($CH_2 = 14.00000$ Kendrick mass units) [10], sorted by even and odd nominal Kendrick masses (nearest integer) and the Kendrick mass defect (i.e., difference between exact and nominal Kendrick mass) within the MIDAS program [33]. Even and odd nominal Kendrick masses were then sorted into homologous series based on their identical Kendrick mass defects [12]. Molecular formulas of species of less than ~500 Da could be unambiguously assigned based solely on mass measurement to ±0.5 ppm. Elemental compositions were assigned by use of a molecular formula calculator program limited to molecular formulas consisting of up to 100 ^{12}C atoms, 2 ^{13}C, 200 ^{1}H, 5 ^{14}N, 8 ^{16}O, 5 ^{32}S, and 1 ^{34}S. Molecular formulas could be confirmed/eliminated by the presence/absence of the corresponding nuclide containing one ^{13}C. Because members of a homologous series differ by only integer multiples of CH_2, assignment of a single member of a series accurately assigns formulas to higher mass members.

Once a homologous series is identified, it is described by its class (heteroatom content) and the number of rings plus double bonds. The number of rings plus double bonds is commonly described by either a double bond equivalent value (DBE) or a Z value. The DBE value increases by one unit for each additional ring plus double bond. The Z value decreases by two units for each additional ring plus double bond. Both conventions are used here. For a compound with the molecular formula C_cH_hX (where X refers to the heteroatom present) the Z value is determined by

$$h - 2c \quad (1)$$

and DBE is determined by

$$c - \frac{h}{2} + \frac{n}{2} + 1 \quad (2)$$

For example, the $C_nH_{2n-10}O_2$ molecule has a Z value of −10 and a DBE value of 6 and, therefore, has 6 rings plus double bonds. Ring plus double bond designations correspond to the neutral molecule—not the $[M-H]^-$ ion formed by negative ion ESI.

3. Results and discussion

3.1. Traditional biomarker analyses

Biomarker analyses were conducted to determine the oils' projected source rock type, level of thermal maturity, age and degree of biodegradation. All oils are of marine origin. The Midway Sunset (MWS, reservoir) and El Segundo (ES, surface) oils were both generated from the Miocene Monterey formation (CA, USA). The Biomarker Biodegradation Index (PM Index) described by Peters [36] classifies the MWS oil as severely degraded (PM Index = 8). The oil has a lower API gravity, a higher sulfur content and a high TAN value (Table 1). In addition, the regular steranes are degraded and the demethylated hopane/hopane ratio is high: 1.37. The ES oil (PM Index = 4) is less degraded than the MWS as the *n*-alkane and isoprenoids are absent, but the sterane and hopane biomarkers are in tact. The Urals (UR, reservoir) oil from Russia is a light, non-degraded oil (PM = 0) as evidenced by the presence of *n*-alkanes, its Ph/*n*-C_{17} ratio, higher API gravity and low TAN value (Table 1). The Long Beach (LB, surface) oil, believed to originate from Eocene marine shales (CA, USA), is also non-degraded (PM = 0) due to the presence of *n*-alkanes. However, as subsequently discussed, the LB sample may actually be a mixture of oils with varying degrees of biodegradation. From here on, the hydrocarbon biomarker data will be compared to the ESI FT-ICR MS data to determine if analysis of acidic NSO components supports and/or provides further insight into surface biodegradation processes.

3.2. Broadband negative ESI FT-ICR MS spectra

Previous studies with model compounds demonstrated the selective ionization of acidic species, such as carboxylic acids, phenols and neutral nitrogen compounds (e.g., pyrrole benzologues) by negative ion ESI [5,7,37]. These studies demonstrate that quasimolecular ions, $[M-H]^-$, are formed with minimal fragmentation. The generation of one ionic species for each neutral analyte present is desirable given the complexity of petroleum polars. As shown in Fig. 1, each mass spectrum contains ~4000 peaks ($\geqslant 5\sigma$ RMS noise) between the mass range $250 < m/z < 800$. The non-degraded oil (UR) contains fewer peaks than the reservoir degraded oil (MWS) and the soil extracts (LB and ES). A mass resolving power $m/\Delta m_{50\%} > 400{,}000$ (in which $\Delta m_{50\%}$ denotes mass spectral

Table 1
Geochemical properties of oil samples

Sample	PM Index	Degree of biodegradation	API[a]	S (wt%)	N (wt%)	V (ppm)	Ph/n-C_{17}	TAN	C_{15+} group-type composition			
									Sat (%)	Aro (%)	NSO (%)	Asph (%)
UR-reservoir	0	Non-degraded	27.6	1.71%	0.100%	72	1.11	0.05	36.6	46.7	11.8	5.1
MWS-reservoir	8	Severe	12.9	1.33%	0.816%	31	–	4.05	12.9	44.1	34.6	8.5
ES-surface	4	Severe	ND	2.21%	ND	138	–	ND	5.7	24.9	32.6	36.8
LB-surface	0[b]	Non-degraded	ND	0.21%	ND	0	0.86	ND	43.2	16.0	32.7	8.2

[a] API gravity = (141.5/s.g.) – 131.5, where s.g. is specific gravity; TAN, total acid number, mg KOH/g oil; PM Index, Biodegradation Index from Peters et al. and degree of biodegradation from Wenger et al. [36]. ND, not determined due to limited sample size.
[b] The LB oil is likely a mixture of oils that vary in degree of biodegradation.

full width at half-maximum height) was obtained for all oils. The weight-average molecular weight (M_w) for each oil is also shown in Fig. 1. Kim et al. recently demonstrated with six genetically related oils that M_w decreases with increasing levels of biodegradation [28]. This trend is not observed in Fig. 1 as the oils are not genetically related.

3.3. Acidic NSO compound class distributions

In Fig. 2, the relative abundance for each compound class is summed over all nominal masses, divided by the total relative abundance for all peaks, and depicted in three bar graphs—one for oxygen-only containing compounds, one for nitrogen-containing compounds and one for sulfur-containing compounds. (Classes with relative abundances less than 2% are not shown.) For all oils, >98% of peaks were unambiguously assigned a molecular formula.

3.3.1. Distribution of oxygen-only containing compounds

Among acidic NSO species in crude oil, the distribution, type (number of rings plus double bonds) and abundance of O_2 species (e.g., carboxylic acids) provides the most compelling evidence for biodegradation. Numerous analyses of reservoir crude oils by negative ion ESI FT-ICR MS have shown that (1) the abundance of O_2-species is greater in biodegraded crude oils than non-biodegraded oils and (2) the abundance of O_2-species increases as the degree of biodegradation increases. These O_2 species are generated during the aerobic and anaerobic metabolism of the hydrocarbon fraction [1,3,36,38–40]. In Fig. 2, the LB soil extract has the highest relative abundance in O_2 species (45%), followed by the MWS biodegraded reservoir oil (30%), the ES soil extract (29%), and the UR non-degraded reservoir oil (4%). Overall, 83% of the species ionized in the LB oil contain only oxygen: O (2%), O_2 (45%), O_3 (18%), O_4 (11%), O_5 (5%), O_6 (2%), and O_7 (in trace amounts). Upon the addition of an oxygen (e.g., from O_2 to O_3) the relative abundance decreases by approximately half. In the MWS, ES, and UR oils, oxygen-only species comprise 45%, 42%, and 18% of the ionized species, respectively. The trend for total relative abundance of oxygen-only species reflects the trend observed for O_2 species: LB $\gg$ MWS $>$ ES $\gg$ UR. The dramatic compositional differences between the LB and UR oils are shown in the zoomed mass inset between 469.1 and 469.5 Da in Fig. 3.

The higher relative abundances of O_2 species in the LB and ES oils, suggest that, like the MWS oil, they too have undergone biodegradation in the reservoir, the environment or both. This conclusion correlates well to the ES hydrocarbon biomarker data, but not the LB data. As mentioned previously, the LB oil is classified as a non-degraded oil because of the distinct presence of *n*-alkane peaks in the whole crude gas chromatogram (not shown). However, a closer look indicates that the LB oil may actually be a mixture of non-degraded and biodegraded oils. Since the LB soil was collected from a site with a long

history of oil contamination from a subsurface pipeline, it is highly likely that the site was contaminated by multiple oils at various time periods. The large unresolved hump in the gas chromatogram, the large percentage of NSO compounds (32.7%) determined by LC, and the trace amount of demethylated hopanes (demethylated hopane/hopane ratio = 0.04) observed by GC/MS provide clues that a biodegraded oil may be present. A detailed investigation of the type and carbon number distribution of oxygenated species (see Section 3.4) by ESI FT-ICR MS provides an even more convincing argument that a mixture of oils exists in the LB sample.

3.3.2. Distribution of nitrogen- and sulfur-containing compounds

Nitrogen- and sulfur-containing heteroatomic classes also are present in the oils. Nitrogen-containing classes are made up of pyrroles, carbazoles and indoles (e.g., N, N_2, NS). Compounds with nitrogen and oxygen may be furollic, phenolic, or contain carboxylic functional groups (e.g., NO, NO_2, NO_3). All have been previously identified [5,18,28,41]. The summed contribution of nitrogen-containing classes for the oils is: UR (70%) ≫ MWS (46%) > ES (20%) > LB (13%). In the ESI FT-ICR mass spectrum the UR oil appears to have considerably more nitrogen-containing species than the MWS. In fact, that is incorrect. Elemental analysis indicates that nitrogen compounds comprise 0.100 wt% of the UR oil and 0.816 wt% of the MWS (Table 1). This disparity is due to different ionization efficiencies in the ESI source. In negative ion ESI, ionization efficiency tends to increase as the acidity of the species increases [5]. Therefore, carboxylic acids ionize more efficiently than carbazoles. As evidenced by the TAN values in Table 1, the MWS oil has a higher acid content than the UR oil. The acids in the MWS oil more efficiently compete for charge than the neutral nitrogen compounds, thus causing the nitrogen content to appear less. The UR oil has a low acid content and, subsequently, higher ionization efficiency for the nitrogen-containing compounds. For sulfur containing species, summed relative abundance rankings from Fig. 2 roughly correlate to % sulfur values in Table 1: ES > UR > MWS > LB. However, one must caution that only a fraction of sulfur-containing compounds are ionized by negative ion ESI. Neutral and basic sulfur compounds are not observed. Fortunately, it is the acid content (e.g., the oxygen-only containing species) that is of greatest molecular interest when comparing crude oils with varying degrees of biodegradation. For these compounds, relative abundances in the mass spectra correlate with the TAN values of the oils. The remainder of this manuscript will, therefore, focus on the oxygen-containing species.

3.4. Type (ring plus double bonds) and carbon number distributions for O_2 species

The type and carbon number distribution for O_2 species can provide further evidence of biodegradation. Kim et al. recently reported that with increasing biodegradation (1) the abundance of acyclic fatty acids ($Z = 0$ or 1 DBE) decreases, (2) the abundance of multi-ring naphthenic acids ($Z = -4$ to -8 or 3–5 DBE) increases and (3) the abundance of $C_{31,32}H_{n-10}O_2$ hopanoic acids (or 6 DBE) initially increases (PM Index of 1–2) then decreases (PM Index 3+)

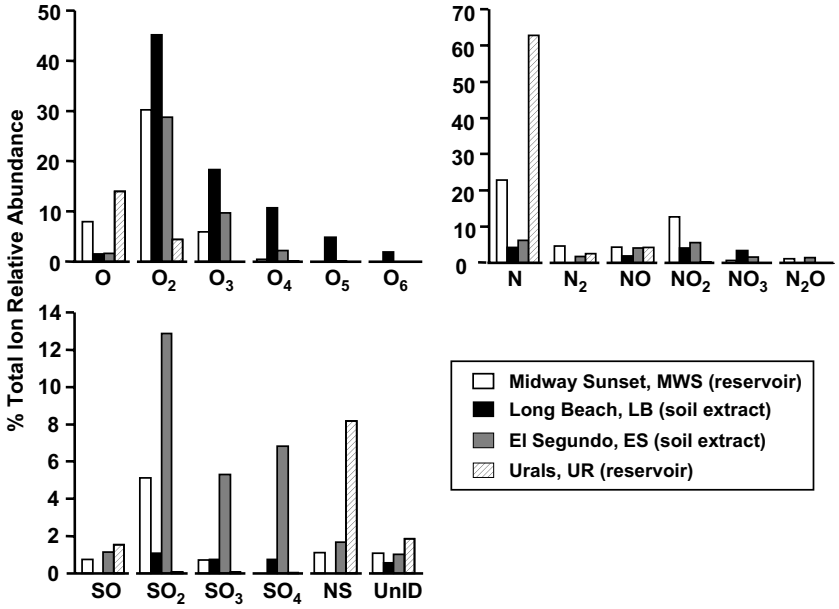

Fig. 2. Compound class distribution of the crude oils, normalized to 100%. The biodegraded oils (MWS, LB* and ES) contain a higher relative abundance in oxygen-only species than the non-biodegraded reservoir oil (UR). Of the biodegraded oils, the LB oil contains the highest relative abundance of highly oxygenated (>O_5) species—species not detected in the other oils. The non-degraded oil contains predominately nitrogen-containing compounds. *The LB oil is likely a mixture of biodegraded and non-degraded oils.

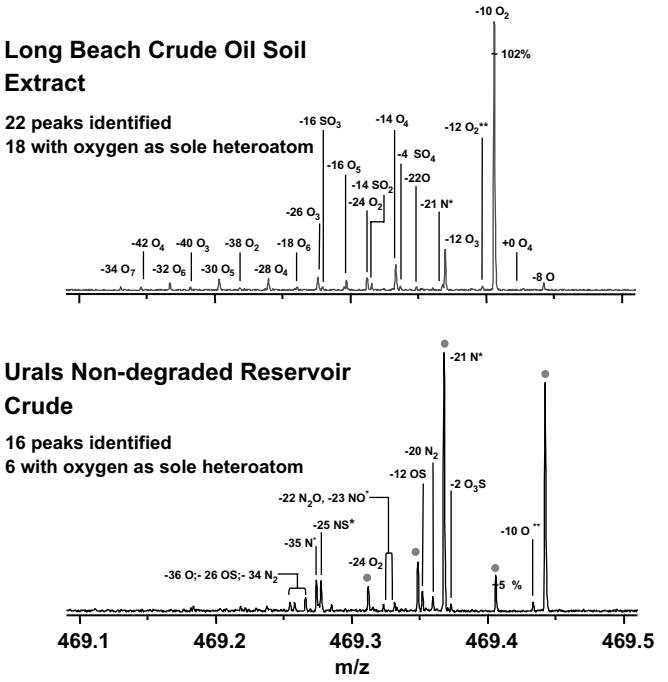

Fig. 3. The zoomed mass inset at 469 Da demonstrates the high mass resolving power/mass accuracy of FT-ICR MS and the difference in composition of the LB and UR oils. Mass resolving powers > 400,000 allows the unambiguous assignment of peaks with an error less than 0.5 ppm. Eighteen of 22 species identified in the LB sample contain 1–7 oxygens as the sole heteroatom. Also seen in the LB spectrum is an equal mass spacing of 36 mDa, which corresponds to the loss of an oxygen and addition of one double bond equivalent (DBE). This pattern is most evident by peaks labeled $-34\,O_7$, $-32\,O_6$, $-30\,O_5$, $-28\,O_4$, $-26\,O_3$, $-24\,O_2$ and $-22\,O$. Such chemical order in the mass spectrum provides an internal check on peak assignment. 6 of the 18 species identified in the UR oil contain only oxygen. Most species contain nitrogen or oxygen with sulfur or nitrogen. Asterisks denote nuclides containing one (*) or two (**) ^{13}C atoms. Species are identified by their Z-value and class. Dots (●) indicate species present in both oils.

[28]. We can use the distribution of O_2 species to qualitatively estimate the degree of biodegradation in the surface oils. As seen in Fig. 4 (top), the non-degraded reservoir UR oil exhibits the highest relative abundance of acyclic fatty acids (23.9%, scaled relative abundance for all O_2 species) and a relatively low abundance of multi-ring acids. The MWS, ES and LB oils, in contrast, contain lower abundances of acyclic acids and higher abundances of naphthenic acids with DBEs between 3 and 5 ($Z = -4$ to -8). The MWS, or most biodegraded oil, contains the lowest relative abundance of acyclic acids at 1.5%, followed by the ES and LB oils at 5.8% and 7.3%, respectively. Overall, the MWS, ES and LB oils have very similar O_2 type distributions. Furthermore, the relative abundance of $C_{31,32}H_{n-10}O_2$ hopanoic acids is less for the MWS, LB and ES oils than the non-degraded UR oil (Fig. 5). The summed relative abundance of hopanoic acids between C_{31} and C_{32} indicates that indeed the less degraded oils have a higher abundance of hopanoic acids: UR (32%) > LB (25%) > ES (19%) > MWS (15%). Past studies report that biodegradation not only reduces the abundance of C_{31} and C_{32} acids, but also increases the relative abundance of $\leqslant C_{30}$ acids [3]. This shift has been attributed to changes in microbial pathways. As seen in Fig. 5, all oils exhibit a maximum abundance at C_{32}. However, a summed relative abundance of hopanoic acids between C_{26} and C_{30} does show an increase in $\leqslant C_{30}$ acids with increasing levels of biodegradation: MWS (25%) > ES (20%) > LB (18%) > UR (16%). Thus, from the relative abundance (Fig. 2), type distribution (Fig. 4, top), and carbon number distribution (Fig. 5) data for O_2-species, we can conclude that, like the MWS oil, the surface oils are biodegraded. However, the higher abundance of acyclic fatty acids and hopanic acids in the LB oil suggests, as does the biomarker data, that a non-degraded oil is also present in the sample.

3.5. Characterization of O_3+ species

A van Krevelen diagram (Fig. 6) allows the simultaneous visual comparison of all oxygen-only species present in the four samples [13–15]. The diagram is constructed with the molar ratio of hydrogen-to-carbon (H/C ratio) on the *y*-axis and the molar oxygen-to-carbon ratio (O/C ratio) on the *x*-axis. Since FT-ICR MS analysis allows for elemental composition assignment of all resolved peaks, construction of van Krevelen diagrams is straight-forward. The region occupied by each class is governed by the number of oxygens and the DBEs. The addition of an oxygen increases the O/C ratio, therefore, the class containing a single oxygen is located on the far left of the diagram and the class containing 6 (LB) oxygens is on the far right. As previously seen in Figs. 2 and 3, the biodegraded oils contain a greater diversity of oxygen-only species.

The van Krevelen diagram also sorts data by saturation/aromaticity. For a given class, the slope of homologous series (individual lines that intersect at an H/C ratio value of 2) increases with an increase in aromaticity. Also, a vertical line (same O/C values, different H/C values) within a class represents differences in saturation. As seen in Fig. 6, all oxygen species in the UR, MWS and ES oils have an H/C ratio between 1 and ~2. O and O_2 species in the LB oil follow a similar trend, but classes O_3+ have a large number of species with an H/C ratio less than 1—indicating that these species are highly aromatic. Since portions the LB sample have likely been exposed to environmental conditions for long periods of time (10+ years), perhaps these highly oxygenated, highly aromatic species are the product of aerobic biodegradation in a surface environment. Tomcsyk and Winans previously identified highly oxygenated compounds (up to O_6) in a San Joaquin Valley (CA) oil and attributed their presence to aerobic biodegradation conditions [41]. However, the authors also reported that unsaturation/aromaticity decreased as the number of oxygen atoms increased. For example, the most abundant O_4 species had DBEs between 6 and 9 ($Z = -10$ and -16); and the most abundant O_6 species had DBEs between 2 and 6 ($Z = -2$ and -10). While the O_4 species

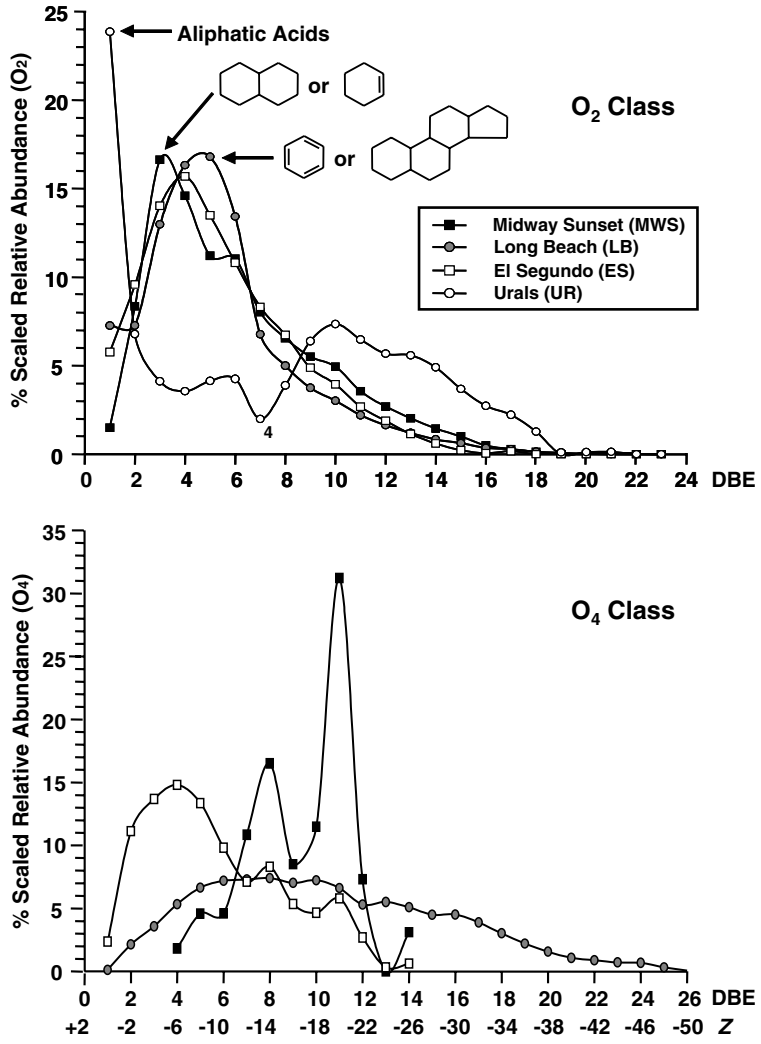

Fig. 4. O_2 type distribution in all four oils, normalized to the total relative abundance of the O_2 class (top) and O_4 type distribution in the MWS, LB and ES oils, normalized to the total relative abundance of the O_4 class (bottom). Type refers to the hydrogen deficiency as defined by ring plus double bond equivalents (DBE). Each additional ring/double bond makes the DBE more positive by 1 unit and Z more negative by two units.

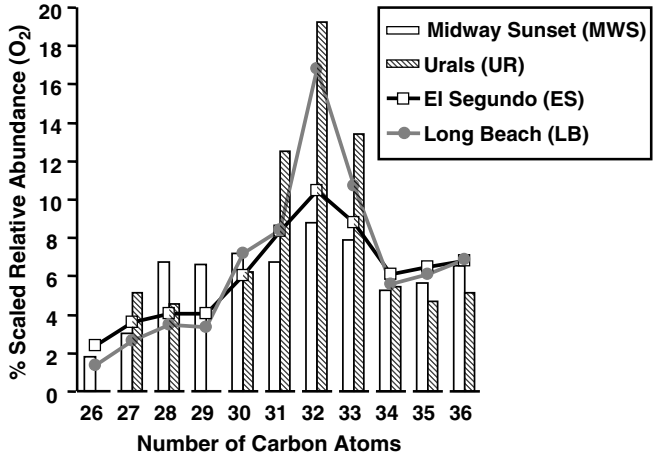

Fig. 5. Carbon number distribution for hopanoic acids, 6 DBE O_2 ($Z = -10$), normalized to the total relative abundance of the O_2 class. The relative abundance of hopanoic acids is greatest in the non-degraded reservoir oil (UR). Generally, the abundance of these acids decreases with increasing degrees of biodegradation.

in the LB oil exhibits a DBE distribution (Fig. 4, bottom) similar to that reported by Tomcsyk and Winans (with a maximum between 6 and 10 DBEs), the O_6 species in the LB oil (data not shown) exhibit a maximum DBE distribution at 18 ($Z = -34$)—indicating that unsaturation/aromaticity increases as the number of oxygen atoms increases. GC/MS analysis rules out the possibility that these highly oxygenated species originate from humic material as terrigenous markers, (e.g., large C19 tricyclic terpanes) are absent. The presence of oxygenated polymers also seems unlikely given that the data is sorted into homologous series in with repeating units of CH_2—a homologous series unique to crude oils and petroleum products. Also note that the DBE distributions for O_4 species (Fig. 4, bottom) in the ES, MWS and LB oils are different. The MWS reservoir oil exhibits two maxima at 8 and 11 DBE. This bimodal distribution likely results from two different core structures within the same class [6]. The ES surface oil exhibits a skewed Gaussian-like DBE distribution with a

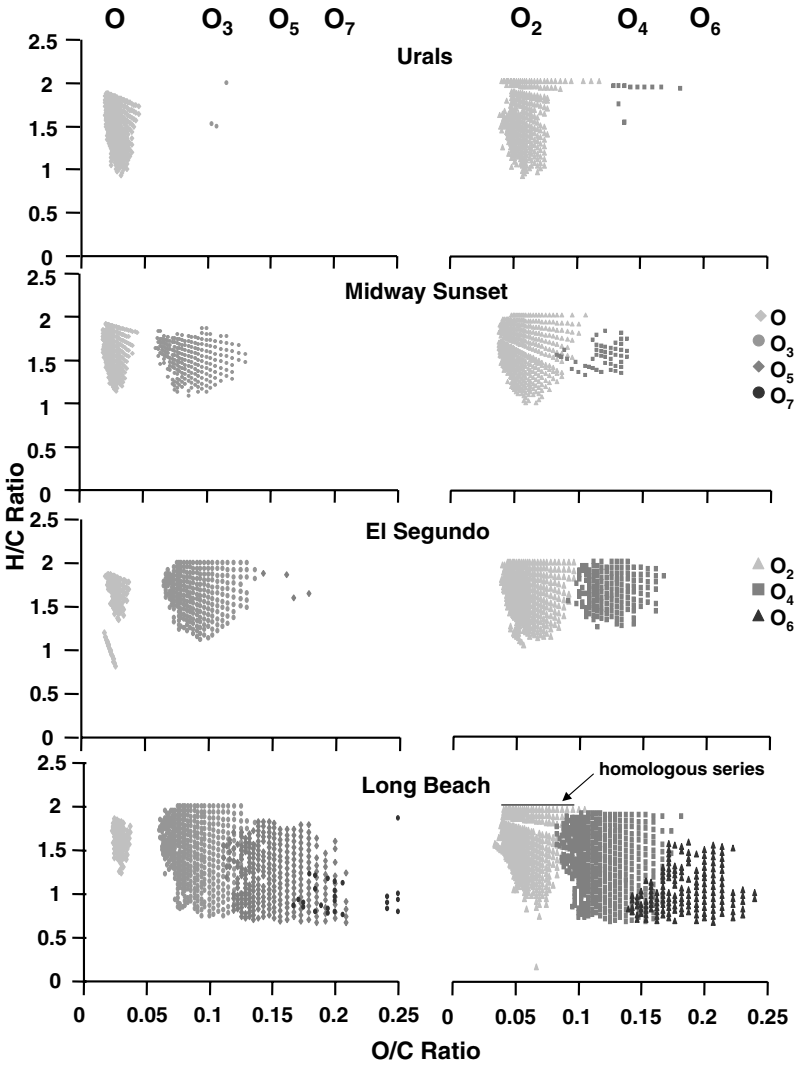

Fig. 6. van Krevelen diagrams for oxygen only-containing classes in all four oils. For clarity, two diagrams are presented for each oil—one for classes with an odd number of oxygen atoms (left) and one for even oxygen atoms (right). van Kevelen diagrams provide a visual comparison of the DBEs for different classes of oxygen species present. Angled horizontal lines that intersect the *y*-axis at 2 correspond to a homologous series within a given class. The biodegraded oils contain species with more oxygen atoms. Highly oxygenated species (O_5+) also have higher DBEs (or lower H/C ratios).

maximum between 3 and 5 DBEs—likely indicating the presence of one core structure. The LB surface oil DBE distribution is similar in shape to the ES oil, but is spread over a much wider DBE range. This broad distribution further adds to the curious nature of the oxygenated species in the LB oil. The O_6 DBE distribution is also broad (data not shown). No doubt, a greater understanding of surface vs. reservoir biodegradation processes is necessary to account for these differences in DBE distribution and unsaturation/aromaticity.

3.6. Use of a new biodegradation index to estimate the degree of biodegradation

From the detailed compositional analysis of biodegraded reservoir oils, Kim et al. devised a biodegradation index that allows one to estimate the degree of biodegradation in a crude oil [28]. This index is based on the ratio of acyclic to cyclic naphthenic acids (A/C ratio) and is calculated as follows

$$\text{A/C Ratio} = \left[\sum \mathrm{O}_{2\,Z=0} \Big/ \sum \mathrm{O}_{2\,Z=-4,-6,-8}\right] \tag{3}$$

in which the mass spectral peak magnitudes are summed for $Z = 0$ (1 DBE) and $Z = -4$, -6 and -8 (3, 4, and 5 DBE), respectively. This ratio was devised from data similar to that shown in Fig. 4, in which the A/C ratio decreases as the degree of biodegradation increases. To test the robustness of this newly proposed index, we obtained the trend line data from Kim et al., determined the slope of the line and plotted on the trend line the A/C ratio for each of our oils (Fig. 7). From this plot, the degree of biodegradation was qualitatively estimated by extrapolating to the *x*-axis, which correlates to the Biomarker Biodegradation Indices (designated as PM Index) described by Peters et al. [36]. A value of "0" corresponds to a non-degraded

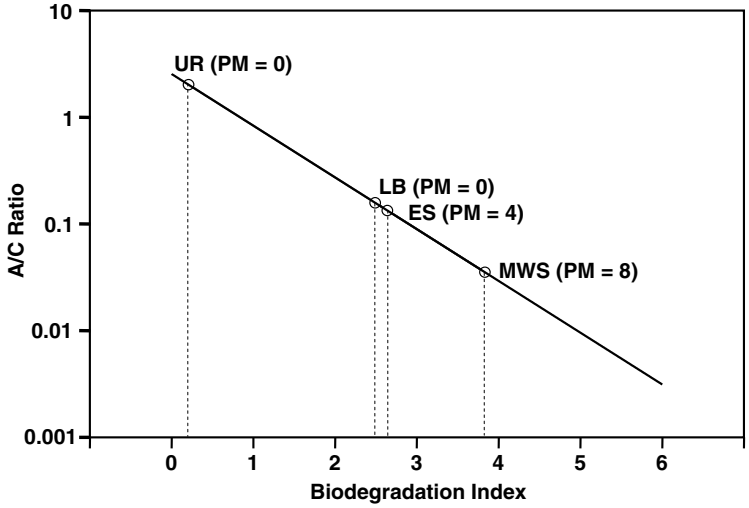

Fig. 7. Comparison of two biodegradation indices. The biodegradation index proposed by Kim et al. [28] (referred to as the Acid Biodegradation Index) is based on the ratio of acyclic to cyclic acids (*y*-axis). The Biomarker Biodegradation Index developed by Peters et al. [36] (referred to as PM Index) utilizes traditional hydrocarbon biomarkers (*x*-axis). For the MWS, ES and LB oils the two index values do not correlate well.

oil; a value of "4+" corresponds to a severely degraded oil. As expected, the non-degraded reservoir UR oil has an acid A/C Ratio of ~2 and a Acid Biodegradation and PM Index of ~0. The severely degraded MWS reservoir oil, with an estimated PM Index of 8 has an estimated Acid Biodegradation Index of ~4 in Fig. 7. Therefore, the A/C ratio used here may underestimate the degree of biodegradation. The surface oils, LB and ES, have an Acid Biodegradation Index of 2.5 and 2.6 (moderately degraded), respectively. Again, this value is an underestimation for the ES oil, which has a PM Index = 4. The Acid Biomarker Index for the LB oil, however, is greater than the PM Index = 0. Assuming that the LB sample is a mixture of non-degraded and biodegraded oils, the difference in PM and Acid Biomarker indices is logical given that the PM Index value is largely due to the presence of *n*-alkanes in the non-graded oil and the Acid Biodegradation Index value is due to the presence of naphthenic acids in the biodegraded oil. Overall, the Acid Biodegradation Index ranks the oils in the correct order of biodegradation: UR < LB < ES < MWS, but the calculated values do not match the PM Index values as well as the data presented by Kim et al. perhaps because our oils are not genetically related.

4. Conclusions and future work

Here we presented the first detailed compositional comparison of acidic NSO compounds in reservoir and surface crude oils by negative ion FT-ICR MS. The biodegraded reservoir crude oil (MWS) exhibited an increase in relative abundance of O_2-species, a decrease in acyclic fatty acids, an increase in multi-ring naphthenic acids and a decrease in C_{32} hopanoic acids compared to the non-biodegraded reservoir crude oil (UR). The surface oils (LB and ES) exhibited trends similar to the MWS, indicating that the oils have been microbially altered in the reservoir, the environment or both. Hydrocarbon biomarker analyses also indicated the presence of a non-degraded oil in the LB sample as evidenced by distinct *n*-alkane peaks in the whole gas chromatogram. This evidence, in combination with the ESI FT-ICR MS data, suggests that the LB soil was contaminated with a mixture of non-degraded and biodegraded (biodegraded in the reservoir, environment or both) oils. Here we see how traditional hydrocarbon analysis and ESI FT-ICR analysis of polars provide complimentary information to reveal a complex history of contamination. From biomarker data alone, the presence, type and distribution of oxygenated species would not have been known.

Future work will further investigate the effect of surface biodegradation on the composition of petroleum polars. Because samples of the LB and ES oils prior to their release in the environment were not available for this study (as is often the case for aged environmental samples), it was impossible to differentiate between microbial alteration that occurred in the environment and alteration that occurred in the reservoir. Such a differentiation is the focus of our current project. We also aim to determine the source of the highly oxygenated, highly aromatic compounds observed in the LB oil. It will be interesting to see if these species are unique to aerobic biodegradation in surface environments. Lastly, we will continue to examine the robustness of the biodegradation index proposed by Kim et al. [28]. We predict that this ratio will prove useful in monitoring the biodegradation of a single oil or a suite of genetically related oils in the environment.

Acknowledgements

We would like to thank Peter Reich and Robert Wise of the US EPA for helping us acquire the contaminated soil samples. We thank Dr. Cliff Walters from ExxonMobil Research and Engineering Co. for providing us with the

reservoir oils. The FT-ICR MS data was collected at the National High Magnetic Field Laboratory (NHMFL), Tallahassee, FL with the help of Dr. Ryan Rodgers and the continued support of Dr. Alan Marshall. Don Smith, also at the NHMFL, collected the mass spectra of the oil extracts with the ThermoFinnigan LTQ to confirm the mass spectral range. The NSF National High Field FT-ICR Facility is supported by CHE-99-09502 and Florida State University. Acknowledgement is also made to the Donors of the American Chemical Society Petroleum Research Fund and Chapman University for support of this research.

References

[1] Thorn KA, Aiken GR. Biodegradation of crude oil into nonvolatile organic acids in a contaminated aquifer near Bemidji, Minnesota. Org Geochem 1998;29:909–31.
[2] Watson JS, Jones DM, Swannell RPJ, van Duin ACT. Formation of carboxylic acids during aerobic biodegradation of crude oil and evidence of microbial oxidation of hopanes. Org Geochem 2002;33:1153–69.
[3] Meredith W, Kelland S-J, Jones DM. Biodegradation and crude oil acidity. Org Geochem 2000;31:1059–73.
[4] Hsu CS, Dechert GJ, Robbins WK, Fukuda EK. Napthenic acids in crude oils characterized by mass spectrometry. Energy Fuels 2000;14:217–23.
[5] Hughey CA, Rodgers RP, Marshall AG, Qian K, Robbins WK. Identification of acidic NSO compounds in crude oils of different geochemical origins. Org Geochem 2002;33:743–59.
[6] Qian K, Robbins WK, Hughey CA, Cooper HJ, Rodgers RP, Marshall AG. Resolution and identification of elemental compositions for more than 3000 crude acids in heavy petroleum by negative-ion microelectrospray high-field Fourier transform ion cyclotron resonance mass spectrometry. Energy Fuels 2001;15:1505–11.
[7] Lo CC, Brownlee BG, Bunce NJ. Electrospray-mass spectrometric analysis of reference carboxylic acids and Athabasca Oil Sands naphthenic acids. Anal Chem 2003;75:6394–400.
[8] McMartin DW, Peru KM, Headley JV, Winkler M, Gillies JA. Evaluation of liquid chromatography-negative ion electrospray mass spectrometry for the determination of selected resin acids in river water. J Chromatogr A 2002;952:289–93.
[9] Hughey CA, Rodgers RP, Marshall AG. 11,000 components in a single (+) ESI spectrum. Anal Chem 2002;74:4145–9.
[10] Kendrick E. Anal Chem 1963;35:2146–54.
[11] Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K. Kendrick mass defect spectrum. Anal Chem 2001;73:4676–81.
[12] Hsu CS, Qian K, Chen YC. An innovative approach to data analysis in hydrocarbon characterization by on-line liquid chromatography-mass spectrometry. Anal Chim Acta 1992;264:79–89.
[13] van Krevelen DW. Fuel 1950;29:269–84.
[14] Kim S, Kramer R, Hatcher P. Graphical method of analysis of ultrahigh resolution broadband mass spectra of NOM, the van Krevelen diagram. Anal Chem 2003;75:5336–44.
[15] Wu ZG, Rodgers RP, Marshall AG. Two- and three-dimensional van Krevelen diagrams. Anal Chem 2004;76:2511–6.
[16] Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG. Diesel. Energy Fuels 2001;15:1186–93.
[17] Wu ZG, Rodgers RP, Marshall AG, Strohm JJ, Song C. Comparative compositional analysis of untreated and hydrotreated oil by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Energy Fuels 2005;19:1072–7.
[18] Hughey CA, Rodgers RP, Marshall AG, Walters CC, Qian K, Mankiewicz P. Acidic and neutral polar NSO compounds in Smackover oils of different thermal maturity. Org Geochem 2004;35:863–80.
[19] Qian K, Rodgers RP, Hendrickson CL, Emmett MR, Marshall AG. Resolution of 3000 N Bases. Energy Fuels 2001;15:492–8.
[20] Wu ZG, Jerstrom S, Hughey CA, Rodgers RP, Marshall AG. Coal. Energy Fuels 2003;17:946–53.
[21] Wu ZG, Rodgers RP, Marshall AG. Illinois coal. Energy Fuels 2004;18:1424–8.
[22] Grannas AM, Hockaday WC, Hatcher PG, Thomson LG, Mosley-Thompson E. New revelations on the nature of organic matter in ice cores. J Geophys Res 2006;111. Citation No. D04304.
[23] Hockaday WC, Grannas AM, Kim S, Hatcher PG. Direct molecular evidence for the degradation and mobility of black carbon in soils from ultrahigh-resolution mass spectral analysis of dissolved organic matter from a fire-impacted forest soil. Org Geochem 2006;37:501–10.
[24] Kim S, Kaplan LA, Hatcher PG. Biodegradable dissolved organic matter in a temperate and a tropical stream determined from ultra-high resolution mass spectrometry. Limnol Oceanog 2006;51:1054–63.
[25] Kujawinski EB. Electrospray ionization Foruier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex and environmental mixtures. Environ Forensics 2002;3:207–16.
[26] Stenson AC, Marshall AG, Cooper WT. Exact mass and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 2003;75:1275–84.
[27] Rodgers RP, Blumer EN, Freitas MA, Marshall AG. Jet fuel chemical composition, weathering, and identification as a contaminant at a remediation site, determined by Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 1999;71:5171–6.
[28] Kim S, Stanford LA, Rodgers RP, Marshall AG, Walters CC, Qian K, et al. Microbial alteration of the acidic and neutral polar NSO compounds revealed by Fourier transform ion cyclotron resonance mass spectrometry. Org Geochem 2005;36:1117–34.
[29] Reich P. Environmental Protection Specialist, Superfund Division, Oil Program; 2005: personal communication by e-mail.
[30] Nelson ED. Graner Oil Company spill response and oversight report, El Segundo, CA. Long Beach: START Ecology and Environment, Inc.; 2003.
[31] Wang Z, Fingas M, Sergy G. Study of 22-year-old arrow oil samples by GC/MS. Environ Sci Technol 1994;28:1733–46.
[32] Zumberge JE, Russell JA, Reid SA. Charging of Elk Hills reservoirs as determined by oil geochemistry. AAPG Bull 2005;89:1347–71.
[33] Blakney GT, van der Rest G, Johnson JR, Freitas MA, Drader JJ, Shi SDH, et al. Further improvements to the MIDAS data Station for FT-ICR mass spectrometry. Chicago, IL; 2001.
[34] Senko MW, Canterbury JD, Guan S, Marshall AG. A high-performance modular data system for FT-ICR mass spectrometry. Rapid Commun Mass Spectrom 1996;10:1839–44.
[35] Hendrickson CL, Quinn JP, Emmett MR, Marshall AG. Quadrupole mass filtered external accumulation for fourier transform ion cyclotron resonance mass spectrometry. Long Beach, CA; 2000.
[36] Peters KE, Walters CC, Moldowan JM. Biodegradation parameters. The biomarker guide. Cambridge: University Press; 2005. p. 661–702.
[37] Hendriksen T, Juhler RK, Svensmark B, Cech NB. The relative influences of acidity and polarity on resposiveness of small organic molecules to analysis with negative ion ESI MS. J Am Soc Mass Spectrom 2005;16:446–55.
[38] Behar FH, Albrecht P. Acids. Org Geochem 1984;6:597–604.
[39] Mackenzie AS, Patience RL, Yon DA, Maxwell JR. Geochim Cosmochim Acta 1982;46:783–92.
[40] Jaffé R, Gallardo MT. Application of carboxylic acid biomarkers as indicators of biodegradation and migration of crude oils from the Maracaibo Basin, western Venezuela. Organic Geochem 1993;20:973–84.
[41] Tomczyk NA, Winans RE, Shinn JH, Robinson RC. On the nature and origin of acidic species in petroleum. 1. Detailed acid type distribution in a California crude oil. Energy Fuels 2001;15:1498–504.