# Identification of Water-Soluble Heavy Crude Oil Organic-Acids, Bases, and Neutrals by Electrospray Ionization and Field Desorption Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

LATEEFAH A. STANFORD,§
SUNGHWAN KIM,‡ GEOFFREY C. KLEIN,§
DONALD F. SMITH,§
RYAN P. RODGERS,§,† AND
ALAN G. MARSHALL*,†,§

*National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, FL 32310-4005, Korean Basic Science Institute, 52 YeoEun-Dong, Yuseong-Gu, Daejeon, Korea, and Department of Chemistry, Florida State University, Tallahassee, FL, 32306-4590.*

We identify water-soluble (23 °C) crude oil NSO nonvolatile acidic, basic, and neutral crude oil hydrocarbons by negative-ion ESI and continuous flow FD FT-ICR MS at an average mass resolving power, $m/\Delta m_{50\%} = 550\,000$. Of the 7000+ singly charged acidic species identified in South American crude oil, surprisingly, many are water-soluble, and much more so in pure water than in seawater. The truncated $m/z$ distributions for water-soluble components exhibit preferential molecular weight, size, and heteroatom class influences on hydrocarbon solubility. Acidic water-soluble heteroatomic classes detected at >1% relative abundance include O, $O_2$, $O_3$, $O_4$, OS, $O_2S$, $O_3S$, $O_4S$, $NO_2$, $NO_3$, and $NO_4$. Parent oil class abundance does not directly relate to abundance in the water-soluble fraction. Acidic oxygen-containing classes are most prevalent in the water-solubles, whereas acidic nitrogen-containing species are least soluble. In contrast to acidic nitrogen-containing heteroatomic classes, basic nitrogen classes are water-soluble. Water-soluble heteroatomic basic classes detected at >1% relative abundance include N, NO, $NO_2$, NS, $NS_2$, NOS, $NO_2S$, $N_2$, $N_2O$, $N_2O_2$, OS, $O_2S$, and $O_2S_2$.

## Introduction

Anthropogenic sources and natural seepage from the ocean floor contribute to the millions of gallons of oil entering the sea (*1*). The amount and composition of spilled oil water-soluble components determine the subsequent aquatic toxic risk and response. Successful aqueous partitioning requires overcoming the inability of water molecules to direct their electric dipole moments toward hydrophobic surfaces.

* Corresponding author phone: 1-850-644-0529; fax: 1-850-644-1366; e-mail: marshall@magnet.fsu.edu.
† National High Magnetic Field Laboratory, Florida State University.
‡ Korean Basic Science Institute.
§ Department of Chemistry, Florida State University.

Despite the overall nonpolar nature of crude oil, some heteroatom-containing classes nevertheless exhibit water solubility. Therefore, it is reasonable that heterocyclic aromatic NSO-containing crude oil species should display similar water-solubility.

Research on polar water solubles primarily focuses on naphthenic acids (surfactant-like acyclic, monocyclic, and polycyclic acids) and pure hydrocarbons (*2*). Naphthenic acids are soluble at neutral pH and ionize under slightly basic solutions such as seawater (*2*). Up to ~3% w/w of crude oil may be acids, of which naphthenic acids are most abundant (*3*). Crude oil heterocyclic/aromatic NSO compound characterization and monitoring is fundamentally necessary for environmental protection, because NSO-containing compound solubilities are orders of magnitude greater than those of purely aromatic hydrocarbons. For example, the aqueous solubility of phenanthrene, a neutral tri-aromatic hydrocarbon, is 1.1 mg L$^{-1}$, whereas acridine (a nitrogen-containing polar basic tri-aromatic) has a solubility of 46.6 mg L$^{-1}$ (*4*). Identification of solubility patterns for organic NSO heteroatomic basic water-soluble species should not go ignored, because basic nitrogen species are generally more toxic than acidic nitrogen compounds (*5*).

Gas chromatography mass spectrometry (GC-MS) and low voltage electron ionization Fourier transform ion cyclotron resonance mass spectrometry (EI FT-ICR MS) provide detailed compositional information for volatile (typically low molecular weight) and nonvolatile (typically higher molecular weight) polar crude oil compounds (*6–8*). However, less is known about the water solubility of petroleum hydrocarbons. Furthermore, solubility differs between pure water and seawater environments, because high molar volume or small or multivalent seawater ions "salt out" hydrophobic species (*9*). Here, we apply high-resolution electrospray ionization (ESI) and field desorption (FD) FT-ICR MS for nonvolatile ($m/z > 225$) polar and nonpolar heteroatomic species compositional characterization of geographically distinct crude oils and their water-soluble acids, bases, and neutrals. 9.4 T FT-ICR MS (*10*) attains mass resolving power, $m/\Delta m_{50\%} > 200\,000$ ($\Delta m_{50\%}$ is the mass spectral full peak width at half-height) throughout the range, $225 < m/z < 1000$, suitable to differentiate compounds of the same nominal mass, but different exact mass, at an accuracy to within better than 0.3 ppm (*7, 11*). Negative-ion ESI selectively generates [M−H]$^-$ quasimolecular ions at atmospheric pressure by deprotonation without fragmentation (*12, 13*). ESI selectively ionizes nonvolatile acidic and basic species, and low emitter current FD ionizes nonvolatile neutrals in addition to acids and bases without extensive ion fragmentation or matrix interference. FD ionization generates primarily M$^{+\cdot}$ ions and, to a lesser degree, [M−2H]$^{+\cdot}$ and [M−H]$^+$ ions from samples deposited on a field ionization emitter and desorbed by an ionizing electric field and thermal assistance (*14, 15*). Traditional FD is a pulsed ionization technique that requires multiple emitter dosages to accommodate signal averaging. Continuous-flow FD enables sustained FD ion formation for improved mass resolution, dynamic range, and signal-to-noise ratio (*15, 16*). Positive-ion ESI and continuous-flow FD FT-ICR MS have proved successful for complex mixture hydrocarbon analysis (*6, 15, 16*). In an effort to project general water solubility profiles of oils for the future, we systematically map heteroatom class, aromaticity, and alkylation trend solubility for heavy North and South American and Middle Eastern oils in pure water and seawater.
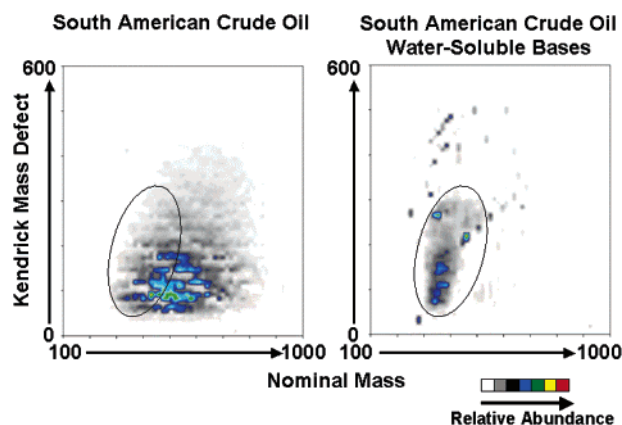
**FIGURE 1. Isoabundance-contoured Kendrick mass defect vs Kendrick nominal mass plot for positive ions from a South American crude oil and its water-soluble bases. An oval highlights the water-soluble bases and is overlaid on the parent oil Kendrick plot. "Bases and neutrals" is an operational definition denoting components obtained by liquid/liquid solvent extraction at pH 12.**

## Experimental Section

See Supporting Information (6, 7, 15−28).

## Results and Discussion

**Molecular Size and Electrostriction.** Figure SI-1 (see Supporting Information) shows isoabundance-contoured Kendrick mass defect (KMD) plots based on mass spectra of the parent oil (left), pure water-soluble acids (middle), and seawater water-soluble acids (right). KMD plots best illustrate the extent to which aromaticity (number of rings plus double bonds, or double bond equivalents (DBE)) influences water partitioning. KMD vs nominal Kendrick mass (NKM, i.e., the nearest integer value of the Kendrick mass) for all ions of a single mass spectrum) concisely displays type (DBE) dependence for complex mass spectra typical for FT-ICR MS analysis of petroleum derived materials.

$$\text{Kendrick mass} = \text{IUPAC mass} \times (14.00000/14.01565) \tag{1}$$

$$\text{KMD} = (\text{NKM} - \text{Kendrick mass}) \times 1000 \tag{2}$$

The utility of the Kendrick plot is that members of a given homologous alkylation series (i.e., same heteroatom class and type (DBE value) but different number of -CH₂ groups) have the same KMD. Thus, members of such a series fall on a horizontal line, with successive members separated by exactly 14.0000 Da (Kendrick exact mass of $CH_2$) along that line (29, 30). Also, Kendrick mass defect increases with DBE, so that the Kendrick plot gives a simultaneous graphical picture of the degree of unsaturation (*y*-axis) and degree of alkylation (*x*-axis). In Figure 1, we highlight inside an oval the Kendrick plot entries for species from the positive-ion ESI South American crude oil water-soluble bases (The outlying peaks are from internal calibrants of high mass defect). The overlay clearly confirms the lower water solubility with increasing alkylation (i.e., higher NKM for a given Kendrick mass defect (KMD)). The three parent oils exhibit widely variable spectral shape, distribution, and composition. In contrast, the water-soluble fractions share many common spectral and compositional features, starting with similar mass range. Despite the large number of highly abundant high molecular weight species in the parent oils, the water-soluble species populate a much lower *m/z* range (upper mass limit of ∼600 Da in pure water and ∼550 Da in 30 ppt salinity seawater), presumably due to salt ion water compression (electrostriction) and salt−water complexation. The

most abundant species in each water-soluble fraction, independent of geographical origin, fall between 300 < NKM < 400 with a second abundance maximum from 450 < NKM < 500. $[C_{18}H_{29}O_3S]^-$ (exact mass 325.1843 Da, DBE 4 SO₃) is the most abundant anion in South American pure water, South American seawater, and Middle Eastern seawater soluble acids. Consequently, O₃S is the most abundant class. An additional commonality between the water-solubles is DBE 11 O₂ $[C_{32}H_{43}O_2]^-$ (exact mass 459.3269 Da). It is the most abundant ion in the 450 < NKM < 500 segment of the South American pure water-soluble, South American seawater-soluble, and the North American pure water-solubles. Of the >7000 South American crude oil singly charged negative ions, we identify 1441 pure water-soluble, and 768 seawater-soluble species; 1800+ pure water-soluble ions of the 8000 parent oil positive-ion ESI species; and 1205 pure water-soluble FD ions of the 2171 parent oil ions. The greatly reduced compositional complexity for the water-soluble spectra relative to that of the parent oil implies that heteroatom class, molecular weight, and/or structure significantly influence solubility.

Aqueous solubility is inversely related to hydrophobic surface area (31). Polar functional groups, alkyl branching, and aromatization decrease hydrophobic surface area, whereas alkyl chains increase hydrophobic surface area. The effect is more exaggerated for the seawater sample because electrostriction decreases the area into which hydrophobes may solvate between ordered water molecules. The red highlighted area defining the main seawater-soluble components (Figure SI-1 right) is overlaid on the Kendrick plots for the parent oil (left) and pure water-solubles (middle) for comparison.

**Heteroatomic Class Distribution.** Table SI-1 (see Supporting Information) lists the parent crude oil and water-soluble acidic, basic, and neutral class composition for classes of ≥1% relative abundance. Abundance is scaled relative to the highest-magnitude peak in each mass spectrum, so that even if the *absolute* abundance of a given class is the same for two samples, its *relative* abundance depends on differences in the abundances of other species. $O_x$ and $O_xS$ compounds are the most abundant water-soluble classes; therefore, oxygenated functional groups appear to promote aqueous solubility. In comparing the NSO classes for the water-solubles and parent oils, we do not find a direct relationship between heteroatomic class abundance in the parent oil and its corresponding water-soluble fractions. Species of high abundance in the three parent oils are not necessarily high in abundance for their respective water-soluble fractions. For example, the positive-ion NO₂ class is only 2% abundant in the North American crude oil but is the most abundant (26%) in the water-soluble base fraction. We observe NO₂ species in the North American crude oil water-solubles sample by positive and negative-ion ESI at similar abundance and the same DBE range of 5−17. The simplest explanation is that we detect the hydrophilic carboxylic O₂ functional group by negative-ion ESI and the pyridinic functional group in the same molecule by positive-ion ESI. The high abundance of the NO₂ class may relate to the presence of the highly hydrophilic carboxylic and pyridinic functional groups in the same molecule.

Pure hydrocarbons compose >80% of the parent oils, and 12−15% O, O₂, O₄, and S heteroatomic classes. In contrast, the water-soluble base and neutral fraction FD ions consist of ∼34−45% pure hydrocarbons and ∼53−59% O, O₂, and O₄ classes. Although class abundances do not differ drastically from one parent oil to the other, DBE and carbon number distributions are unique to each oil. The near 5:1 HC:NSO parent oil abundance ratio vs 1:1 HC:NSO water-soluble fraction abundance ratio indicate preferential class, DBE, and carbon number water-solubility among neutrals.

As revealed by FD FT-ICR MS, the parent oils are highly aromatic: DBE = 0.5−11 for the South American crude and DBE = 0.5−13 for the North American crude oils. We find an aromaticity limit of DBE ≤ 9 for HC water-solubility with DBE 2 and 6 $C_{16-30}$ ions most abundant. For the present extraction at pH 12, carboxylic and di-carboxylic acids are completely ionized and not extracted in the base fraction. Therefore, water-soluble bases and neutral fraction $O_2$ and $O_4$ ions may include neutral $O_2$ and $O_4$ classes common to crude oils including di- and tetra alcohols, furans, ketones, aldehydes, esters, and ethers un-ionized at pH 12. DBE 0.5−6 $C_{14-40}$ $O_2$ and DBE 1.5−5.5 $C_{15-35}$ $O_4$ class neutrals are enriched in the South American crude oil water-soluble fraction, and DBE 0.5−5 $C_{14-40}$ $O_2$ and DBE 1.5−5.5 $C_{15-36}$ $O_4$ class neutrals are enriched in the North American crude oil water-soluble fraction. DBE 1.5 $O_2$ and DBE 5.5 $O_4$ class ions are most abundant in each water-soluble base and neutral fraction. FD ionization produces ions of integer DBE values by loss of an electron from a neutral, whereas half-integer DBE ions are produced by FD as $[M-H]^{+\cdot}$ ions.

**Pyrrolic vs Pyridinic Solubility.** Our negative-ion ESI FT-ICR MS water-solubility experiments indicate poor water solubility of pyrrolics (acidic nitrogen) independent of their abundance in the parent oil (Table SI-1). Moreover, nitrogen-containing species at >1% abundance are water-soluble only if an oxygenated functional group is present in combination with a pyrrolic group (i.e., $NO_x$ classes): not surprising, because carbazole (DBE 9, N class) solute−solvent hydrogen bonding with water is inefficient because of strong water molecule self-associations that repel carbazole from the bulk solvent. Thus, the parent oil heteroatomic classes that survive in the water-soluble fractions must contain functional groups that strongly hydrogen bond with neighboring water molecules (32, 33). However, positive-ion ESI FT-ICR MS reveals high water solubility for basic nitrogen-containing (pyridinic) hydrocarbons. Pyridinic vs pyrrolic nitrogen water-solubility is nicely illustrated in 0.5 Da segments of the negative- and positive-ion ESI mass spectra for the South American crude oil and its water-soluble acids and bases in Figure SI-2. Negative-ion ESI spectra are on the left in Figure SI-2, and positive-ion ESI spectra are to the right. Parent oil spectra are plotted normally, and water-soluble spectra are inverted for comparison. DBE and class labels are shown for peaks with magnitude higher than 3σ of baseline noise. For the crude oil, pyrrolic species are highly abundant as negative ions, whereas pyridinic species are highly abundant as positive ions. South American crude oil negative ions include N, NS, and NO classes that are not seen in the corresponding water-solubles spectral segment. Even at 37% relative abundance in the parent oil, the negative 15 N class ions are absent in the acidic water-soluble fraction. Acidic $^{13}$C oxygenated species were most abundant in the acidic water-solubles 364.0−364.5 Da segment. Pyridinic species dominate the positive-ion ESI crude oil and water-soluble 364.0−364.5 Da inset. Peaks noted by an asterisk are oxygenated pyridines. It is reasonable that oxygenated pyridines are low in abundance in the parent oil spectrum, but enriched in the water-soluble fraction due to hydrogen bonding with water molecules.

Among several explanations for differences in pyrrolic vs pyridinic solubility, specific hydrogen bond interactions are likely the main factor. Interactions between organic molecule hydrogen bond acceptors with hydrogen atoms of water molecules mainly influence partitioning. Pyridinic molecules contain an easily accessible lone pair of electrons on the nitrogen atom to interact with hydrogen atoms of water (34). Pyrollic molecules contain a hydrogen atom bonded to the nitrogen atom and a lone pair of electrons on the nitrogen atom to interact with oxygen atoms of water. However, because of steric hindrance, it is believed that the lone pair
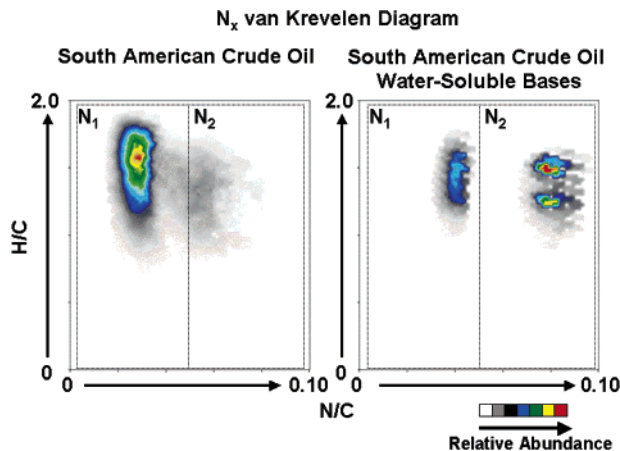


$N_x$ van Krevelen Diagram

South American Crude Oil | South American Crude Oil Water-Solubles

**FIGURE 2.** Relative isoabundance-contoured van Krevelen plots for $N_x$ positive ions from a heavy South American crude oil and its water-solubles.

of electrons belonging to the nitrogen atom of pyrollic molecules is not easily accessible to hydrogen atoms of water, and pyrrollics are thus less soluble than pyridines (34).

Figure 2 shows van Krevelen plots for South American crude oil and its water-soluble bases, constructed from positive-ion ESI FT-ICR MS $N_x$ species. From the assigned chemical formula for the $N_x$ ions, we plot the H/C ratio vs the N/C ratio for each observed elemental composition. Thus, the N/C ratio separates compositions horizontally by compound class (N, $N_2$, $N_3$, etc.) and the H/C ratio separates the same class but with different type (DBE value) and carbon number distribution. Ions of the same homologous series fall along the same diagonal plane with increasing alkylation. The data shift vertically downward with increasing DBE (i.e., increasing aromaticity or a decrease in H/C ratio) (35). Among all basic $N_x$ classes, $N_1$ species are more prevalent in the whole oil, whereas $N_2$ species are most prevalent for the water solubles.

Isoabundance-contoured DBE vs carbon number plots for the South American crude oil and water-soluble N class ions are displayed in Figure SI-3. We emphasize DBE and carbon number solubility limits, by highlighting the oval region enclosing the water-soluble N class ions and overlaying it on the parent oil plot in Figure SI-3 (left). The positive-ion ESI water-soluble N class species span the range, 6 < DBE < 13 and $C_{22-28}$, i.e., only a small fraction of that for their parent oil N class ions, 4 < DBE < 20 and $C_{25-55}$. Similarly (not shown), the water-soluble North American crude oil N class positive ions range from DBE 6 < DBE < 15 and $C_{20-36}$, i.e., a fraction of the range for their parent crude oil, 5 < DBE < 20 at $C_{25-48}$. We do not expect the DBE range for the water-soluble bases to extend as high as for basic species in the parent crude oils because solubility decreases as aromatization increases (36). South American crude oil $C_{32-40}$ DBE 9 N class ions are most abundant, whereas $C_{22-30}$ DBE 9 N class water-solubles are most abundant. For the nitrogen-containing heteroatomic classes, the South American crude oil and water-solubles most abundant species share the same DBE values; however, the water-solubles span only about one-quarter of the carbon number distribution range of the parent oil. The North American crude oil and its water-soluble fraction follow the same pattern.

**$O_x$ Classes.** The South American crude oil has a carbon number distribution of $C_{19-45}$ and DBE range 4−27 for the $O_1$ class. We observe DBE 4−13 $C_{23-40}$ pure water-soluble, and DBE 12−13 $C_{22-33}$ seawater-soluble $O_1$ ions at low ion abundance. $O_1$ class ions are presumably phenols because electrospray ionization selectively ionizes phenolic species (37). Furthermore, phenolic species exhibit DBE values ≥4,
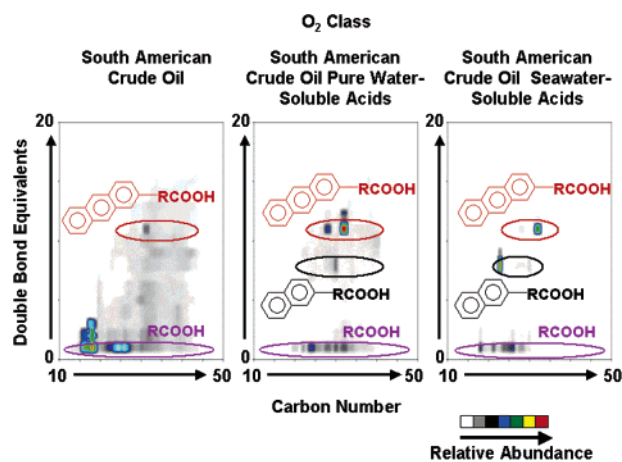
**O₂ Class**

South American Crude Oil — South American Crude Oil Pure Water-Soluble Acids — South American Crude Oil Seawater-Soluble Acids

**FIGURE 3.** Relative isoabundance-contoured plots of double bond equivalents vs carbon number for O$_2$ class compounds in a heavy South American crude oil, pure water-solubles, and seawater-solubles. Highly alkylated O$_2$ species soluble in pure water preferentially "salt out" in seawater.



**O₂ Class**

North American Crude Oil — Medium Arabian Crude Oil — North American Pure Water-Solubles — Medium Arabian Seawater-Solubles
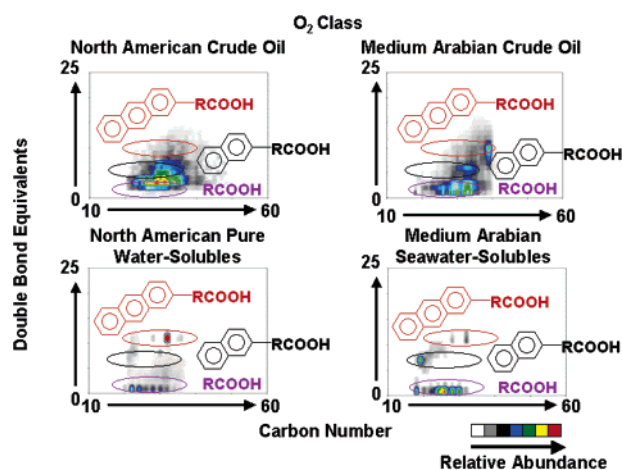
**FIGURE 4.** Relative isoabundance-contoured plots of double bond equivalents vs carbon number for O$_2$ ions from a North American and a Medium Arabian crude oils and their water-solubles. Polyaromatic O$_2$ are less soluble in seawater than in pure water, indicating the "salting out" of those organic species in seawater.

consistent with the lowest DBE range for O$_1$ ions from each sample. Low molecular weight short chain (C$_0$–C$_3$) alkyl phenols are extremely soluble in water, as evident in petroleum formation waters (38). However, a C$_{13}$ alkyl phenol is the lowest molecular weight (275.23796 Da) O$_1$ class ion we observed in the South American crude oil. C$_{13}$ alkyl phenols fall outside the typical water solubility range observed by Bennett et al. (39). Indeed, we observed only slight water-solubility for phenolics (Table SI-1).

We observed significant enrichment (relative to the parent oil) of O$_2$, O$_3$, and O$_4$ classes in the pure water and seawater (Table SI-1) suggesting preferential solubility related to molecular structure. Compared to all crude oil O$_2$ compounds (i.e., diols, aldehydes, esters, etc.), carboxylic acids are most toxic because of their low volatility (40). ESI selectively ionizes these toxic acids, which occupy wide H/C distributions in the water-soluble fractions, an indication of preferential water solubility. We further examine the type and carbon number dependence of the parent oil and corresponding water-soluble O$_2$ compounds through isoabundance-contoured plots of DBE vs carbon number (Figures 3 and 4). South American crude oil (Figure 3, left) O$_2$ species include C$_{14}$–C$_{44}$ fatty acids (DBE 1) and acids with up to five aromatic cores (DBE 17). Low molecular weight fatty acids, monocyclic, and

dicyclic naphthenic acids (DBE 1–3) are the most abundant species in the parent oil. Interestingly, the parent oil and pure water-soluble fraction (Figure 3, middle) O$_2$ species span the same DBE range, but with markedly different relative abundances. Aromatic acid solubility is favored over naphthenic acid solubility in both the pure water and seawater solubles fractions. In fact, the 11 O$_2$ [C$_{32}$H$_{43}$O$_2$]$^-$ tri-aromatic acid is the most abundant ion of the O$_2$ class in the South American pure water-soluble and the North American pure water-soluble acids. Despite the high abundance of particular di- and tri-aromatic acid ions, overall the 1 O$_2$ fatty acids are the most abundant O$_2$ type overall for the South American seawater-soluble, South American pure water-soluble, and the Middle Eastern seawater-soluble acids. However, the 11 O$_2$ type is most abundant in the North American pure water-soluble acid fraction because of [C$_{32}$H$_{43}$O$_2$]$^-$ at unusually high 1.78% relative abundance. Finally, the 1 O$_2$ fatty acids type is the second most abundant type of the O$_2$ class in the North American pure water-soluble acid fraction. Note that the abundances of ions in the mass spectrum of water-soluble species are *not* necessarily directly proportional to their abundances in the parent oil.

At DBE 1–3, C$_{18-32}$ O$_2$ species are more soluble in seawater relative to C$_{20-41}$ for pure water. Salting out increases with the addition of naphthenic rings and increased alkyl carbon chain length. However, ~C$_{31}$ DBE 5–7 O$_2$ species are soluble in pure water. The truncated DBE and carbon number distributions for the O$_2$ class in seawater are not surprising, in view of the 5-fold increase in hydrophobicity of di-(2-ethylhexyl) phthalate (oxygenated polar hydrocarbon) in seawater relative to freshwater (31). However, seawater accommodates DBE 1 (fatty acids), short-chain DBE 8 (di-aromatic acids), and short-chain DBE 11 (tri-aromatic acids). For medium Arabian crude oil seawater-soluble and North American crude oil pure water-soluble negative-ion ESI O$_2$ species, the DBE vs C$_n$ profiles (see Figure 4) show similar solubility trends and electrostriction effects. In prior naphthenic acid extraction from freshwater via oil sands hot water extraction, only acids at DBE 1–7 C$_{14-33}$ were identified (2). Here, we expand the identification with high-resolution FT-ICR MS and find that di- and tri-aromatic acids are soluble as well.

The South American crude oil is low in O$_2$ species, whereas the O$_2$ class is highly abundant in its seawater soluble acid fraction. Likewise, the O$_2$ class is most abundant for North American and Middle Eastern Crude oils, but is greatly reduced in their water-soluble acid fractions. On the other hand, the O$_2$ class is established both here and elsewhere as a highly water soluble heteroatomic class. What then are the physical parameter(s) or abiotic conditions that determine solubility? The North American parent oil is most abundant in di-cyclic acids (DBE 3). However, aside from the peak at DBE 11, the most abundant species in its water-soluble fraction are DBE 1 fatty acids, even though they are less than 1% abundant in the parent oil. South American crude oil, lowest in the O$_2$ class abundance, is most abundant in low molecular weight ≤C$_{20}$ DBE 1 fatty acids (Figure 3, left). In contrast, the Middle Eastern and North American oils are low in abundance for low DBE O$_2$ species <C$_{22}$ (Figure 4), reflecting low solubility in the water-soluble fraction. Therefore, we infer that oils highly abundant in low carbon number fatty acids (i.e., South American crude in Figure SI-3) generate water-soluble fractions relatively enriched in the O$_2$ class.

**S and O$_x$S Classes.** We find negative-ion OS, O$_2$S, O$_3$S, and O$_4$S species ranging widely from long alkyl chain acyclic to polyaromatics in the parent oil. In contrast, both water-soluble fractions are limited mainly to O$_3$S species; OS, O$_2$S, and O$_4$S classes are very low in abundance and have short alkyl chain distributions relative to the parent oil. Figure 5 shows the parent oil and water-soluble O$_3$S and O$_4$S DBE vs

**South American Crude Oil** | **South American Crude Oil Pure Water-Soluble Acids** | **South American Crude Oil Seawater-Soluble Acids**

O₃S Class

O₄S Class

Carbon Number

Relative Abundance

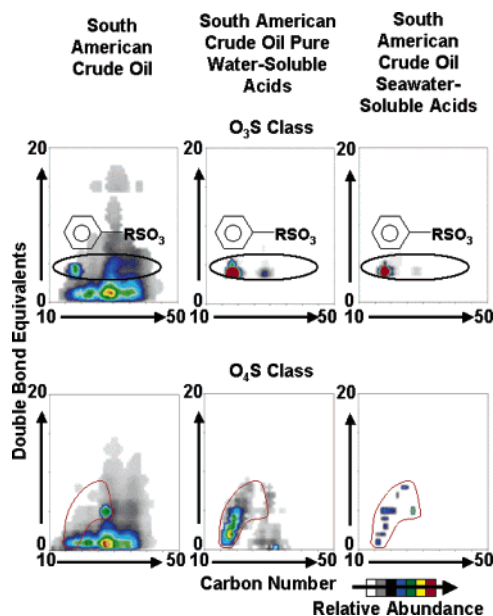**FIGURE 5. Relative isoabundance-contoured plots of double bond equivalents vs carbon number for O₃S and O₄S ions in a heavy South American crude oil, pure water-solubles, and seawater-solubles. The most saturated members of this class are highly abundant in crude oil but are negligibly water-soluble. As in Figure 1, the seawater-soluble region is outlined in red in all three plots.**

carbon number distributions. DBE 1 and DBE 4 $O_3S$ species are most abundant in the parent oil, whereas DBE 4 $O_3S$ species are most abundant in the water-soluble fractions and DBE <4 $O_3S$ species are absent. The $O_3S$ class structure is currently unclear. They may represent polyhydroxylated or carboxylated thiophenes, sulfoxides, or linear alkylbenzene sulfonates (LAS) (*34*, *41*). LAS (DBE = 4) are the commercially and domestically widest-used surfactants in the detergent industry and are highly soluble. Blanks collected from the extraction equipment and vials used for sample storage showed no detectable polar species. Thus, the suspected LAS species do not arise from sample contamination.

Surfactants such as LAS are amphiphilic molecules that self-assemble into quasi-micelles, normal micelles, inverted micelles, and mixed micelles for the enhanced solubility of less polar compounds in aqueous systems, and below their critical micelle concentration by pseudomicelles (CMC) (*42*). Crude oils possess many surfactant species (e.g., metallic salts, organic acids, organic bases, and organometallics) capable of self-assembly so as to promote water-in-oil emulsions (inverted-micelles), and water solubility (normal, quasi-, mixed micelles) (*43*). In fact, at elevated pH, naphthenic acids and their soaps form micelles in the aqueous phase, thereby increasing the solubility of low-polarity compounds (*44*). Because of their high abundance, $O_3S$ ions in the water-soluble spectra are most likely LAS species, due to known high solubility of standard LAS compounds. Therefore, the suspected surfactant ability of DBE >4 $O_3S$ species may enhance and explain the solubility of the polyaromatic and long alkyl chain species seen here. Moreover, DBE <4 $O_3S$ species in the parent oil are most likely carboxylated and polyhydroxylated thiophenes.

The parent oil $O_4S$ class and $O_3S$ class (Figure 5) DBE vs carbon number distribution profiles are similar. We find DBE 0–14 $C_{12-42}$ $O_4S$ species in the parent oil, with DBE 1–5 $C_{15-37}$ $O_4S$ species most abundant. Molecular size and structure regulate $O_4S$ species solubility in pure water and seawater. The red border in Figure 5 highlights the limited DBE 1–8 $C_{16-27}$ $O_4S$ class solubility in seawater (right), for comparison to the DBE vs carbon number plots for the pure water-solubles

(middle), and the parent oil (left). The significant reduction in abundance of $O_4S$ species in the DBE 1–9 $C_{16-36}$ range for seawater relative to pure water attests to seawater salting out effects. However, some $O_4S$ species in the narrow range of DBE 1–5 $C_{16-22}$ are present in both the pure water and seawater-soluble fractions. $O_4S$ species may be aromatic sulfur-containing carboxylates (i.e., phenylthio and phenylsulfonyl carboxylates). Yin et al. state that the dominant factors for aqueous solubility of organosulfur compounds are the carboxyl group that hydrogen bonds with water molecules, the electron-donating ability to water molecules, and hydrophobic surface area (*45*). Their work illustrates increased hydrophobicity for molecules of larger volume and thus larger hydrophobic surface area (*45*). A similar effect is visible in Figure 5 by the absence of $O_4S$ water-soluble species at the higher carbon numbers and low DBE values seen in the whole crude oil.

Positive-ion multi-heteroatom sulfur-containing compounds are prevalent in the distilled water-soluble fraction of each crude oil (Table SI-1). The South American crude oil water-soluble bases include the following basic sulfur-containing species in order of increasing abundance: $NS_2$, OS, $O_2S$, $NO_2S$, NOS, NS, and $O_2S_2$. The South American crude oil and water-soluble fraction nitrogen-sulfur and oxygen-sulfur compounds exhibit similar carbon number distributions, but different DBE distributions. $N_xS_y$ compounds are polyaromatic, whereas $O_xS_y$ compounds are less aromatic. The $O_2S_2$ class (DBE 2–8 and $C_{11-31}$) is second most abundant of all classes in the South American crude oil water-soluble base fraction.

In positive-ion ESI FT-ICR MS, South and North American crude oil water-soluble base $O_2S_2$ compounds are protonated and sodiated $O_xS_y$ ions. The sodium adduct, potassium adduct, and proton-bound dimer $O_xS_y$ species have previously been reported as sulfoxides by positive-ion ESI FT-ICR MS in diesel fuel and petroporphyrins (*6*, *46*). The nature of the presently observed $O_2S_2$ compounds is not entirely clear. To test for the presence of $O_2S_2$ dimers, trapped ions were subjected to infrared multiphoton dissociation in the ICR cell. No fragmentation occurred, as in similar IRMPD experiments by Hughey on model compound sodiated, protonated, and potassium-bound sulfoxide dimers. Therefore, we infer that the present $O_xS_y$-containing species may also be sulfoxides. We find $O_2S_2$ species of DBE 2–7 in the North American and DBE 2–8 for the South American crude oil water-soluble bases. FT-ICR and other mass spectrometric studies suggest that positive-ion $O_xS_y$ ions are linear sulfoxides rather than aromatic sulfur species due to low DBE (*6*, *47*).

Benzothiophene (DBE 6), dibenzothiophene (DBE 9), and 5-carboxybenzothiophene exhibit water solubility (*17*, *48*). Carboxythiophenes are acidic, and therefore, not extracted into our water-soluble base and neutral fraction. Sulfur compounds observed in the oils and not in the water-soluble fractions presumably represent mercaptans, thiophenes, and polycyclic and polyaromatic thiophenes (based on their DBE 1–10 distribution in the parent oils). The ion optics of our FD instrument limits the transmission of ions below $m/z$ 200, such as benzothiophene and dibenzothiophene. However, alkylthiophenes and alkylpolyaromatic thiophenes of significant carbon chain length fall within the detectable $m/z$ range. We observe North American crude oil DBE 6 $C_{23-33}$ and DBE 9 $C_{25-28}$ ions, and South American crude oil DBE 6 $C_{20-45}$ and DBE 9 $C_{25-36}$ S class ions. For benzothiophenes ($C_8$) and dibenzothiophene ($C_{12}$) core structures, North American crude oil benzothiophenes have $C_{15-25}$ alkyl carbons and dibenzothiophenes have $C_{13-16}$ alkyl carbons. Similarly, South American crude oil benzothiophenes have $C_{12-37}$ alkyl carbons and dibenzothiophenes have $C_{13-24}$ alkyl carbons. Oil neutral organic water solubility experiments by Ali report alkylbenzothiophene and alkyldibenzothiophenes solubility

in trace amounts relative to alkylbenzenes, alkylphenols, and other neutral molecules solubility (*17*). The absence of alkyl monosulfur compounds of long carbon chain length in our results supports that finding. The South American crude oil water-soluble fraction is not completely devoid of sulfur species. We see the $O_2S_2$ class at 0.8% abundance. $O_2S_2$ class South American crude oil water-soluble base ions are the only positive-ion electrospray ionizable species detected by FD FT-ICR MS, yet second most abundant for positive-ion ESI classes (Figure SI-3). Moreover, at 0.2% abundance, North American crude oil water-soluble $NO_2$ ions are the only positive-ion electrospray ionizable species detected by FD FT-ICR MS, even though they are the most abundant for positive-ion ESI classes, suggesting a close direct relationship between positive-ion ESI ionization efficiency for the $O_2S_2$ and $NO_2$ classes and actual concentration within the basic polar fraction of petroleum samples.

## Acknowledgments

## Supporting Information Available

Experimental Procedures (including sample descriptions, sample preparation for ESI FT-ICR MS and continuous-flow FD FT-ICR MS, and mass spectrometry measurements), a Table, and three Figures whose contents are described in the main text of this paper. This material is available free of charge via the Internet at http://pubs.acs.org.

## Literature Cited

(1) Coleman, J. M.; Baker, J.; Cooper, C.; Fingas, M.; Hunt, G. L.; Kvenvolden, K. A.; McDowell, J. E.; Michel, J.; Michel, R. K.; Phinney, J.; Pond, R.; Rabalais, N. N.; Roesner, L. A.; Spies, R. B. Oil in the sea: Inputs, fates, and effects, 2002. National Research Council. National Academy Press, Washington, DC. *Spill Sci. Technol. Bull.* **2002**, *7*, 197−199.

(2) Holowenko, F. M.; MacKinnon, M. D.; Fedorak, P. M. Characterization of naphthenic acids in oil sands wastewaters by gas chromatography-mass spectrometry. *Water Res.* **2002**, *36*, 2843−2855.

(3) Jones, D. M.; Watson, J. S.; Meredith, W.; Chen, M.; Bennett, B. Determination of naphthenic acids in crude oils using non-aqueous ion exchange solid-phase extraction. *Anal. Chem.* **2001**, *73*, 703−707.

(4) Broholm, K.; Hansen, A. B.; Jorgensen, P. R.; Arvin, E.; Hansen, M. Transport and biodegradation of creosote compounds in a large, intact, fractured clayey till column. *J. Contam. Hydrol.* **1999**, *39*, 331−348.

(5) Oliveira, E. C.; de Campos, M. C. V.; Lopes, A. S.; Vale, M. G. R.; Caramao, E. B. Ion-exchange resins in the isolation of nitrogen compounds from petroleum residues. *J. Chromatogr. A* **2004**, *1027*, 171−177.

(6) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G. Elemental composition analysis of processed and unprocessed diesel fuel by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* **2001**, *15*, 1186−1193.

(7) Hughey, C. A.; Rodgers, R. P.; Marshall, A. G. Resolution of 11 000 compositionally distinct components in a single electrospray ionization Fourier transform ion cyclotron resonance mass spectrum of crude oil. *Anal. Chem.* **2002**, *74*, 4145−4149.

(8) Rodgers, R.; Blumer, E. N.; Freitas, M. A.; Marshall, A. G. Complete compositional monitoring of the weathering of transportation fuels based on elemental compositions from Fourier transform ion cyclotron resonance mass spectrometry. *Environ. Sci. Technol.* **2000**, *34*, 1671−1678.

(9) Cameron, I. L.; Kanal, K. M.; Keener, C. R.; Fullerton, G. D. A mechanistic view of the non-ideal osmotic and motional behavior of intracellular water. *Cell Biol. Int.* **1997**, *21*, 99−113.

(10) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998**, *17*, 1−35.

(11) Qian, K.; Rodgers, R. P.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. Reading chemical fine print: Resolution and identification of 3000 nitrogen-containing aromatic compounds from a single electrospray ionization Fourier transform ion cyclotron resonance mass spectrum of heavy petroleum crude oil. *Energy Fuels* **2001**, *15*, 492−498.

(12) Qian, K.; Robbins, W. K.; Hughey, C. A.; Cooper, H. J.; Rodgers, R. P.; Marshall, A. G. 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*, 1505−1511.

(13) Zhan, D. L.; Fenn, J. B. Electrospray mass spectrometry of fossil fuels. *Int. J. Mass Spectrom. Ion Process.* **2000**, *194*, 197−208.

(14) Beckey, H. D. Determination of structures of organic molecules and quantitative analyses with field ionization mass spectrometer. *Angew. Chem., Int. Ed. Eng.* **1969**, *8*, 623−639.

(15) Schaub, T. M.; Hendrickson, C. L.; Quinn, J. P.; Rodgers, R. P.; Marshall, A. G. Instrumentation and method for ultrahigh resolution field desorption ionization Fourier transform ion cyclotron resonance mass spectrometry of nonpolar species. *Anal. Chem.* **2005**, *77*, 1317−1324.

(16) Schaub, T. M.; Linden, H. B.; Hendrickson, C. L.; Marshall, A. G. Continuous-flow sample introduction for field desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1641−1644.

(17) Ali, L. N.; Mantoura, R. F. C.; Rowland, S. J. The dissolution and photodegradation of kuwaiti crude-oil in seawater.1. Quantitative dissolution and analysis of the seawater-soluble fraction. *Mar. Environ. Res.* **1995**, *40*, 1−17.

(18) Rogers, V. V.; Liber, K.; MacKinnon, M. D. Isolation and characterization of naphthenic acids from athabasca oil sands tailings pond water. *Chemosphere* **2002**, *48*, 519−527.

(19) Saeed, T.; Al-Mutairi, M. Comparative composition of volatile organic compounds in the water-soluble fraction of different crude oils produced in kuwait. *Water Air Soil Pollut.* **2000**, *120*, 107−119.

(20) Emmett, M. R.; White, F. M.; Hendrickson, C. L.; Shi, S. D. H.; Marshall, A. G. Application of micro-electrospray liquid chromatography techniques to ft-icr ms to enable high-sensitivity biological analysis. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 333−340.

(21) Senko, M. W.; Hendrickson, C. L.; Pasa-Tolic, L.; Marto, J. A.; White, F. M.; Guan, S.; Marshall, A. G. Electrospray ionization ft-icr mass spectrometry at 9.4 tesla. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1824−1828.

(22) Blakney, G. T.; van der Rest, G.; Johnson, J. R.; Freitas, M. A.; Drader, J. J.; Shi, S. D.-H.; Hendrickson, C. L.; Kelleher, N. L.; Marshall, A. G. Further improvements to the midas data station for ft-icr mass spectrometry. 49th ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, IL, 2001.

(23) Senko, M. W.; Hendrickson, C. L.; Emmett, M. R.; Shi, S. D.-H.; Marshall, A. G. External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 970−976.

(24) Hendrickson, C. L.; Quinn, J. P.; Emmett, M. R.; Marshall, A. G. Quadrupole mass filtered external accumulation for Fourier transform ion cyclotron resonance mass spectrometry. 48th ASMS Conference on Mass Spectrometry and Allied Topics, Long Beach, CA, 2000.

(25) Marshall, A. G.; Verdun, F. R. *Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook*; Elsevier: Amsterdam, 1990.

(26) Ledford, E. B.; Rempel, D. L.; Gross, M. L. Space-charge effects in Fourier-transform mass-spectrometry. 1. Electrons. *Int. J. Mass Spectrom. Ion Process.* **1984**, *55*, 143−154.

(27) Shi, S. D. H.; Drader, J. J.; Freitas, M. A.; Hendrickson, C. L.; Marshall, A. G. Comparison and interconversion of the two most common frequency-to-mass calibration functions for Fourier transform ion cyclotron resonance mass spectrometry. *Int. J. Mass Spectrom. Ion Process.* **2000**, *196*, 591−598.

(28) Beu, S. C.; Laude, D. A. Elimination of axial ejection during excitation with a capacitively coupled open trapped-ion cell for Fourier-transform ion-cyclotron resonance mass-spectrometry. *Anal. Chem.* **1992**, *64*, 177−180.

(29) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. N. Kendrick mass defect spectrum: A compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal. Chem.* **2001**, *73*, 4676−4681.

(30) Kendrick, E. A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Anal. Chem.* **1963**, *35*, 2146−2154.

(31) Turner, A. Salting out of chemicals in estuaries: Implications for contaminant partitioning and modelling. *Sci. Total Environ.* **2003**, *314*, 599−612.

(32) Nobeli, I.; Price, S. L.; Lommerse, J. P. M.; Taylor, R. Hydrogen bonding properties of oxygen and nitrogen acceptors in aromatic heterocycles. *J. Comput. Chem.* **1997**, *18*, 2060−2074.

(33) Ruelle, P.; Sarraf, E.; Kesselring, U. W. Prediction of carbazole solubility and its dependence upon the solvent nature. *Int. J. Pharm.* **1994**, *104*, 125−133.

(34) Fileti, E. E.; Coutinho, K.; Malaspina, T.; Canuto, S. Electronic changes due to thermal disorder of hydrogen bonds in liquids: Pyridine in an aqueous environment. *Phys. Rev. E* **2003**, *67*, 061504.

(35) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal. Chem.* **2003**, *75*, 5336−5344.

(36) Ferreira, M. M. C. Polycyclic aromatic hydrocarbons: A qspr study. *Chemosphere* **2001**, *44*, 125−146.

(37) Hughey, C. A.; Rodgers, R. P.; Marshall, A. G.; Qian, K. N.; Robbins, W. K. Identification of acidic nso compounds in crude oils of different geochemical origins by negative ion electrospray Fourier transform ion cyclotron resonance mass spectrometry. *Org. Geochem.* **2002**, *33*, 743−759.

(38) Taylor, P.; Larter, S.; Jones, M.; Dale, J.; Horstad, I. The effect of oil-water-rock partitioning on the occurrence of alkylphenols in petroleum systems. *Geochim. Cosmochim. Acta* **1997**, *61*, 1899−1910.

(39) Bennett, B.; Aplin, A. C.; Larter, S. R. Measurement of partition coefficients of phenol and cresols in gas-charged crude oil/ water systems. *Org. Geochem.* **2003**, *34*, 1581−1590.

(40) Legal, L.; Moulin, B.; Jallon, J. M. The relation between structures and toxicity of oxygenated aliphatic compounds homologous to the insecticide octanoic acid and the chemotaxis of two species of drosophila. *Pest. Biochem. Physiol.* **1999**, *65*, 90−101.

(41) Krueger, C. J.; Barber, L. B.; Metge, D. W.; Field, J. A. Fate and transport a linear alkylbenzenesulfonate in a sewage-contaminated aquifer: A comparison of natural-gradient pulsed tracer tests. *Environ. Sci. Technol.* **1998**, *32*, 1134−1142.

(42) Chu, Y. P.; Gong, Y.; Tan, X. L.; Zhang, L.; Zhao, S.; An, J. Y.; Yu, J. Y. Studies of synergism for lowering dynamic interfacial tension in sodium alpha-(n-alkyl) naphthalene sulfonate/alkali/acidic oil systems. *J. Colloid Interface Sci.* **2004**, *276*, 182−187.

(43) Lee, R. F. Agents which promote and stabilize water-in-oil emulsions. *Spill Sci. Technol. Bull.* **1999**, *5*, 117−126.

(44) Havre, T. E.; Sjoblom, J.; Vindstad, J. E. Oil/water-partitioning and interfacial behavior of naphthenic acids. *J. Dispersion Sci. Technol.* **2003**, *24*, 789−801.

(45) Yin, C. S.; Liu, X. H.; Guo, W. M.; Lin, T.; Wang, X. D.; Wang, L. S. Prediction and application in qspr of aqueous solubility of sulfur-containing aromatic esters using ga-based mlr with quantum descriptors. *Water Res.* **2002**, *36*, 2975−2982.

(46) Rodgers, R. P.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G.; Greaney, M. A.; Qian, K. Molecular characterization of petroporphyrins in crude oil by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Can. J. Chem.* **2001**, *79*, 546−551.

(47) Tomczyk, N. A.; Winans, R. E.; Shinn, J. H.; Robinson, R. C. On the nature and origin of acidic species in petroleum. 1. Detailed acid type distribution in a california crude oil. *Energy Fuels* **2001**, *15*, 1498−1504.

(48) Mundt, M.; Hollender, J. Simultaneous determination of nso-heterocycles, homocycles and their metabolites in groundwater of tar oil contaminated sites using lc with diode array uv and fluorescence detection. *J. Chromatogr. A* **2005**, *1065*, 211−218.