Acc. Chem. Res. **2004,** *37,* 53−59

Petroleomics: The Next Grand Challenge for Chemical Analysis

ALAN G. MARSHALL*[,†,‡] AND
RYAN P. RODGERS[†,‡]

Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32306, and Ion Cyclotron Resonance Program, National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, Florida 32310-4005

Received August 12, 2003

ABSTRACT

Ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry has recently revealed that petroleum crude oil contains heteroatom-containing (N,O,S) organic components having more than 20 000 distinct elemental compositions ($C_cH_hN_nO_oS_s$). It is therefore now possible to contemplate the ultimate characterization of all of the chemical constituents of petroleum, along with their interactions and reactivity, a concept we denote as "petroleomics". Such knowledge has already proved capable of distinguishing petroleum and its distillates according to their geochemical origin and maturity, distillation cut, extraction method, catalytic processing, etc. The key features that have opened up this new field have been (a) ultrahigh-resolution FT-ICR mass analysis, specifically, the capability to resolve species differing in elemental composition by C_3 vs SH_4 (i.e., 0.0034 Da); (b) higher magnetic field to cover the whole mass range at once; (c) dynamic range extension by external mass filtering; and (d) plots of Kendrick mass defect vs nominal Kendrick mass as a means for sorting different compound "classes" (i.e., numbers of N, O, and S atoms), "types" (rings plus double bonds), and alkylation ($(-CH_2)_n$) distributions, thereby extending to >900 Da the upper limit for unique assignment of elemental composition based on accurate mass measurement. The same methods are also being applied successfully to analysis of humic and fulvic acids, coals, and other complex natural mixtures, often without prior or on-line chromatographic separation.

Introduction

Detailed and quantitative measurement of compound types and carbon number distributions of petroleum feed and products is a cornerstone of molecular-based management of refining processes.[1,2] Knowledge of the molecular composition of a crude oil is critical to selection of refining ("downstream") processes and refining conditions, and ultimately to the economic value of the crude oil. The speciation of heteroatomic (particularly N and S) compounds is especially important due to their ultimate fate as NO_x and SO_x emissions upon combustion. At the "upstream" end, flow assurance requires understanding of solid deposits in pipelines, including wax, asphaltene, hydrates, and organic scale, all of which depend on oil composition.

Ultrahigh-resolution mass spectrometry has recently spawned a new field of "petroleomics", namely, the relationship between the chemical composition of a fossil fuel and its properties and reactivity. In this Account, we describe how to resolve and identify components of different elemental composition in a complex mixture, with experimental demonstrations from petroleum crude oil and its distillates. Although the principle has been understood for half a century, its experimental achievement has been realized only within the past few years, based on our recent developments detailed below.

Petroleum crude oil is arguably the world's most compositionally complex organic mixture, in terms of the number of chemically distinct constituents within a dynamic abundance range of 10 000−100 000. For such mixtures, chemists have traditionally relied on multiple wet chemical separations (extraction, precipitation, distillation, chromatography, etc.) to reduce the number of components and thus simplify the determination of the various molecular elemental compositions and structures. Fortunately, each chemical element (more precisely, each isotope of each element) has a different mass defect (i.e., the difference between its exact mass and the nearest integer ("nominal") mass).[3] Moreover, it is now possible to ionize virtually any organic molecule to produce a molecular (e.g., M^+) or quasimolecular (e.g., $[M + H]^+$, $[M - H]^-$) ion. Thus, with sufficiently high mass resolving power, $m/\Delta m_{50\%}$, in which m is ion mass and $\Delta m_{50\%}$ is mass spectral peak full width at half-maximum height, it is in principle possible to resolve and identify the elemental composition (chemical formula) of virtually any organic molecule, even in a complex mixture. In this Account, we describe recent advances culminating in the resolution and identification of crude oil components of more than 10 000 different elemental compositions ($C_cH_hN_nO_oS_s$), with analogous extension to other complex mixtures (coal, humic and fulvic acids). We shall also discuss the three analytically useful molecular characteristics that derive directly from elemental composition: "class" (i.e., numbers of heteratoms, $N_nO_oS_s$), "type" (i.e., hydrogen deficiency or number of rings plus double bonds), and carbon distribution (number of $-CH_2$ groups).

Alan G. Marshall, born in Bluffton, Ohio, on May 26, 1944, received a B.A. with honors in chemistry from Northwestern University and a Ph.D. in physical chemistry from Stanford University. He joined the faculty in chemistry at the University of British Columbia in 1969, and moved to The Ohio State University in 1980 as Professor of Chemistry and Biochemistry and Director of the Campus Chemical Instrument Center. In 1993, he moved to Florida State University, where he is now Kasha Professor of Chemistry and Biochemistry and Director of the Ion Cyclotron Resonance Program at the National High Magnetic Field Laboratory in Tallahassee, FL. His research focuses on development of new techniques and applications of Fourier transform ion cyclotron resonance mass spectrometry, a technique for which he is co-inventor.

Ryan P. Rodgers, born in Pensacola, Florida, on September 7, 1972, received a B.S. in chemistry from the University of Florida in 1995, and a Ph.D. in analytical chemistry from Florida State University in 1999. Following a postdoctoral appointment in aerosol mass spectrometry at Oak Ridge National Laboratory, he joined the Ion Cyclotron Resonance Program at the National High Magnetic Field Laboratory as an Assistant Scholar-Scientist and a courtesy faculty member of the Department of Chemistry and Biochemistry at Florida State University. He directs environmental and petrochemical applications of FT-ICR mass spectrometry at NHMFL.

---

* Corresponding author. Phone: (850) 644-0529. Fax: (850) 644-1366. E-mail: marshall@magnet.fsu.edu.

† Department of Chemistry and Biochemistry.

‡ National High Magnetic Field Laboratory.

10.1021/ar020177t CCC: $27.50

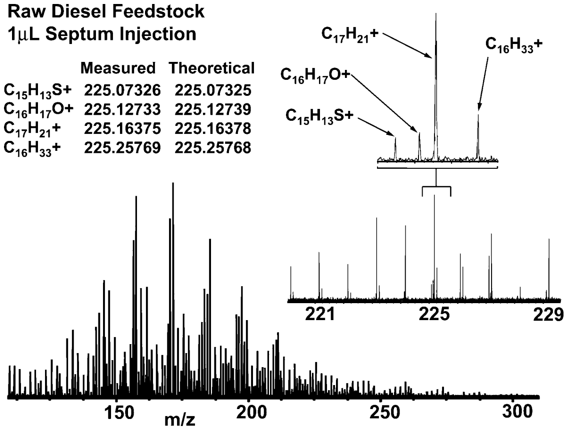
Petroleum and Mass Spectrometry: Divergent and Convergent Evolution

The unique connection between molecular mass and chemical formula was recognized early in the development of mass spectrometry.[4] In fact, the earliest high-resolution commercial mass spectrometers were developed for analysis of petroleum distillates. Unfortunately, those mass analyzers relied on successive magnetic and electric deflection of an ion beam through narrow slits. Mass resolving power depended inversely on slit width, whereas signal-to-noise ratio varied directly with slit width—thus, resolution could be increased only at the expense of lower signal-to-noise ratio. Moreover, scanning the mass range required scanning the magnetic field strength, and acquisition of a single high-resolution mass spectrum could take hours. Mass calibration was also problematic. Finally, the principal ionization method was bombardment with high-energy (~70 eV) electrons, resulting in extensive fragmentation of molecular radical cations ($M^{+\bullet}$). Fragmentation is reduced by lowering the electron beam energy (e.g., to ~15 eV), but at the cost of vastly reduced sensitivity and increased discrimination between different compound types. As a result, petroleum analysis by mass spectrometry has typically been limited to components sufficiently volatile (i.e., low mass and nonpolar) for gas chromatography/mass spectrometry.

The advent of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS)[5,6] ultimately made it possible to obtain an ultrahigh-resolution ($m/\Delta m_{50\%} >$ 100 000) mass spectrum in ~1 s.[7] Nevertheless, it has taken ~25 years to realize that potential for analysis of complex mixtures. To see why, we need to digress briefly to understand how to calibrate and assign mass spectral peaks to particular elemental compositions.

Neutrals, Ions, and Electrons. As reviewed elsewhere,[8,9] it is possible to generate odd-electron molecular ions ($M^{+\bullet}$) from low-voltage electron impact (LVEI) or field ionization/field desorption (FI/FD) and quasimolecular even-electron ions (e.g., $(M + H)^+$, $(M - H)^-$) by chemical ionization (CI), matrix-assisted laser desorption/ionization (MALDI), or electrospray ionization (ESI), without significant fragmentation. (The obvious advantage is that there is ~1 mass spectral peak per analyte composition—compare, for example, to NMR or optical spectroscopy, for which a single molecule can yield thousands of spectral peaks.[10]) It can be shown that even-electron ions of odd (even) mass must contain an even (odd) number of nitrogen atoms.[11] Conversely, odd-electron ions of odd (even) mass contain an odd (even) number of nitrogens. Moreover, every additional ring or double bond decreases the number of hydrogens by 2, without changing the numbers of other atoms in a molecule. For a molecule containing c carbons, h hydrogens, and n nitrogens, one can show that[11]

$$\text{rings} + \text{double bonds} = c - (h/2) + n/2 + 1 \quad (1)$$


Raw Diesel Feedstock
1μL Septum Injection

	Measured	Theoretical
$C_{15}H_{13}S+$	225.07326	225.07325
$C_{16}H_{17}O+$	225.12733	225.12739
$C_{17}H_{21}+$	225.16375	225.16378
$C_{16}H_{33}+$	225.25769	225.25768

FIGURE 1. Low-energy electron impact ionization 5.6-T FT-ICR mass spectrum, showing ~500 resolved peaks, for a 1-μL septum injection of raw diesel feedstock. The inset mass window shows resolution of four ions of different elemental composition at the same nominal mass, each identified with mass accuracy <0.3 ppm at a mass resolving power $m/\Delta m_{50\%} \approx$ 100 000.[19]

Compositional Nominal Mass Isobars. The primary problem in assigning correct elemental composition to ions of a given mass in a complex mixture is to distinguish two compositions of the same nominal mass but different exact mass. Equation 1 immediately reveals which such nominal mass "isobars" are allowed by chemical bonding rules. For example, replacement of CH_4 by O (both nominal mass 16 Da, but differing by 0.0364 Da in exact mass) is possible, provided that another ring or double bond is added (e.g., isobutane vs acetone). Similarly, replacement of SH_4 by C_3 (both nominal mass 32 Da, but differing by 0.0034 Da in exact mass) is possible, if accompanied by addition of five rings plus double bonds. Conversely, various other nominal mass isobars are impossible: e.g., molecules differing by CH_2 vs N, NH_2 vs O, etc.

Accurate Mass Measurement To Yield Elemental Composition. The generation of chemical formulas from accurate mass measurements of petroleum-related samples by non-FT-ICR mass analyzers has been reviewed recently.[12] The first applications of FT-ICR MS (at 3 T magnetic field strength) to petroleum distillates were based on low-energy electron ionization of volatile components.[13,14] At 3 T, it was not possible to obtain sufficient mass resolution to span the full mass range in a single spectrum, and so it was necessary to "stitch" 10–20 separately acquired mass-selected,[15,16] heterodyne-detected[17] mass spectral segments together to obtain the final composite spectrum. Increasing the magnetic field to 5.6 T[18] allowed for acquisition of the full mass range at once, as shown for raw diesel feedstock in Figure 1.[19] The use of electron ionization limited that analysis to relatively low-mass species (<300 Da), and the analysis yielded ~500 different elemental compositions (note the four resolved elemental compositions at 225 Da nominal mass). In each case, it was possible to assign a unique elemental composition on the basis of sub-ppm mass measurement accuracy. That accuracy in turn depends on mass calibration.

Mass Calibration. An ion moving in a spatially uniform magnetic field rotates about the field at a "cyclotron" frequency proportional to z/m (in which m is the ion mass and z is the number of elementary charges per ion). The addition of other ions and an axially quadrupolar electric potential (to keep the ions from escaping along the magnetic field direction) shifts the cyclotron frequency of ions of all m/z values by an approximately constant amount.[20] These effects lead to a simple quadratic equation that relates ion cyclotron frequency to ion m/z.[21,22] Thus, if ions of at least two known m/z values are present, one can convert the ICR frequencies of all other ions to their m/z values. In an EI FT-ICR mass spectrum (as in Figure 1), it is easy to recognize $C_cH_h^+$ ions as those with the highest exact mass at a given nominal mass (e.g., $C_{16}H_{33}^+$ at 255 Da nominal mass). The hydrocarbon ions thus serve as "internal" mass calibrants, allowing for very precise interpolation to yield the masses of the remaining species. For electrospray ionization FT-ICR mass spectra, one may add known "external" mass calibrants, either to the liquid sample itself, or by external accumulation[23] of calibrant and analyte ions electrosprayed alternately from two different needles.[24]

Electrospray Ionization of Petroleum. John Fenn first demonstrated that electrospray ionization could generate abundant ions from the more polar molecules (such as N, S, and O compounds) in petroleum distillates.[25] Although such species make up a small fraction (less than 15%) of most crude oils,[26] they are important for petroleum exploration, production, and refining.[27] Because electrospray is a continuous ionization source, whereas FT-ICR MS is an inherently pulsed detector, ions are accumulated continuously in an external multipole ion trap, and then pulsed periodically into the ICR ion trap for mass analysis.[23]

Ultrahigh Mass Resolution Is Not Enough. The most direct (and most expensive) way to improve FT-ICR performance is to increase the magnetic field strength, B. Mass resolution and mass resolving power increase linearly with B, whereas upper mass limit and dynamic range increase as B^2, and the tendency for closely spaced resonances to coalesce varies as $1/B^2$.[28] Positive- and negative-ion ESI 9.4-T FT-ICR mass spectra are shown in Figure 2. The high magnetic field allows for resolution of more than 17 000 different compounds. Because ESI generates positive ions by protonation (i.e., of the more basic compounds) and negative ions by deprotonation (i.e., of the more acidic compounds), positive- and negative-ion ESI mass spectra typically originate from different neutrals (even of the same elemental composition). However, even at a mass resolving power, $m/\Delta m_{50\%} >$ 350 000, the number of possible elemental compositions increases so rapidly with increasing mass that it is typically not possible to assign a unique elemental composition for ions greater than ~300 Da in mass (see below).[29]

The dynamic range (ratio of highest- to lowest-magnitude peak) in FT-ICR MS is limited to ~10 000−50 000 for data from a single acquisition. Thus, because precision in measurement of mass (i.e., the horizontal axis in a mass spectrum) is proportional to the peak height-to-noise ratio (i.e., the vertical axis in a mass spectrum),[30] it is clear that mass measurement accuracy will be poorest for ions of low relative abundance. One solution is to co-add time-domain signals from hundreds of individual acquisitions,[8,9] because signal-to-noise ratio accumulates as the square root of the number of co-added acquisitions.[17] Alternatively, it is possible to select ions spanning a specified m/z range, either by mass-selective ejection inside the ICR trap,[31] or by passage through an external quadrupole mass filter—thus, the ICR trap may be filled with ions only within a particular m/z range, thereby increasing the signal-to-noise ratio within that range. Of course, one must then "stitch" together many such segments to reconstitute a full m/z range spectrum. Finally, since mass measurement precision also scales as the square root of the number of data points per peak width,[30] it is advisable to extend the time-domain data acquisition period (to yield up to 4 Mword data sets).

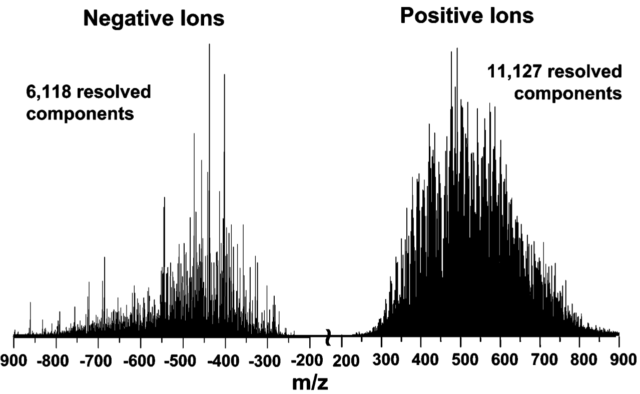


FIGURE 2. Electrospray ionization 9.4-T Fourier transform ion cyclotron resonance composite mass spectrum of a crude oil. Positive-ion and negative-ion mass spectra are combined in the display. Average mass resolving power, $m/\Delta m_{50\%}$, is ~350 000, allowing for resolution of thousands of different elemental compositions.[41]

Helpful Patterns in Mass Spectra. Further progress in assigning the mass spectral peaks thus depends on identifying spacings (not just individual ion masses) in the spectrum. For example, Figure 3 shows two mass scale-expanded segments of the ESI FT-ICR mass spectrum of a Chinese crude oil. Even at low "magnification" (bottom spectrum), families of peaks are separated by multiples of 14.01565 Da (i.e., the mass of CH_2). Those species must therefore be ions of the same "class" (i.e., same numbers of heteroatoms ($N_nS_sO_o$) and the same "type" (i.e., same number of rings plus double bonds, but differing numbers of $-CH_2$ groups (i.e., a "carbon number distribution"). At higher "magnification" (top spectrum), one finds series of peaks separated by multiples of 2.01565 Da (i.e., the mass of two hydrogens). Those species are thus ions with the same class and carbon number, but different "type" (i.e., different numbers of rings plus double bonds). Compound "type" is also commonly reported as the value of Z in the elemental composition, expressed as $C_cH_{2c+Z}N_nO_oS_sP_p$. For example, a carbazole molecule, $C_{12}H_9N$, contains three rings and six double bonds, so that $2c + Z = 9$, and $Z = -15$. Thus, an alkylated carbazole has elemental composition, $C_cH_{2c-15}N$, and a shorthand notation for its compound class and type is $-15N$.

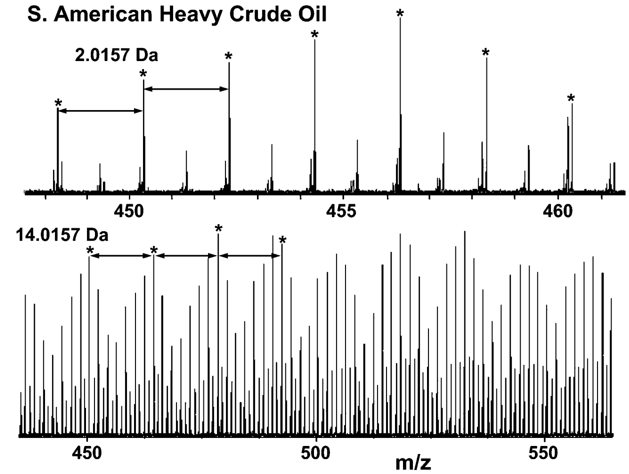


FIGURE 3. Mass spacings in an ESI FT-ICR mass spectrum of a crude oil. Each added ring or double bond lowers the mass by 2.01565 Da, and each additional $-CH_2$ group increases the mass by 14.01565 Da. Identification of such spacings is essential to correct assignment of elemental composition (see text).

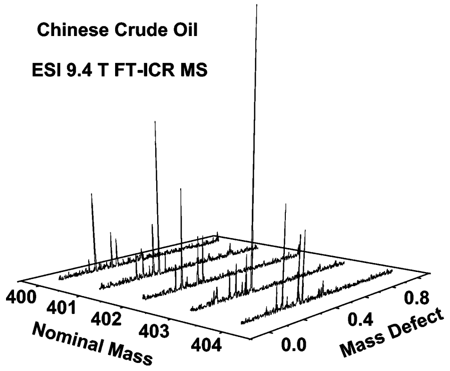


FIGURE 4. Three-dimensional display of a 5-Da segment of a broadband electrospray ionization FT-ICR mass spectrum of a heavy crude oil. The average resolving power across the broadband spectrum is $m/\Delta m_{50\%} = 390\,000$, and the signal-to-noise ratio of the base peak is 1480:1. The spectrum has been sliced into 1-Da segments, and each segment is then rotated by 90° and scaled according to its "mass defect" (i.e., difference between exact and nominal mass). The segments are then stacked according to their nominal masses as shown.[32]

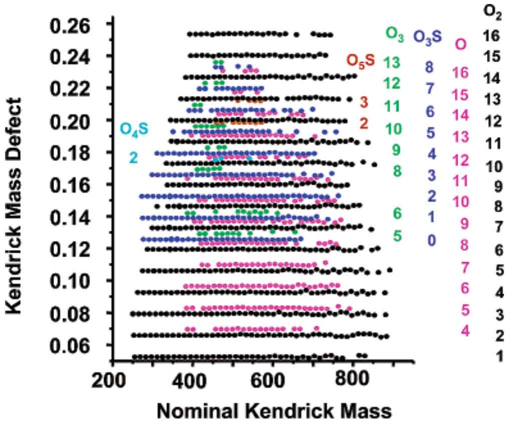


FIGURE 5. Plot of Kendrick mass defect vs nominal Kendrick mass for ions of *odd* nominal mass from 200 to 900 Da in a broadband ESI FT-ICR mass spectrum of a Chinese crude oil. Note the periodicities of 2 and 14 Da on the nominal mass axis, as well as 0.01340 Da (i.e., Kendrick mass defect for two hydrogens) on the mass defect axis. These spacings make it possible to determine molecular "class" and "type" simultaneously over a wide mass range from a single display (see text).

Species containing ^{13}C provide additional information. Because mass analyzers distinguish ions according to their m/z ratio, a singly charged ion containing one ^{13}C atom will be found 1.00335 higher in m/z than ions of the same elemental composition but with all ^{12}C atoms. However, for an ion with z elementary charges, the corresponding ^{13}C-containing ion will be shifted upward by $1.003/z$ on the m/z axis. In this way, we find that virtually all electrosprayed petroleum ions of <1000 Da are singly charged.

At currently achievable Fourier transform ion cyclotron resonance broadband mass spectrometry resolving power ($m/\Delta m_{50\%} > 350\,000$ for $200 < m/z < 1000$), it would be necessary to spread out a mass spectrum (e.g., Figure 2) over ~200 m to provide visual resolution of the most closely resolved peaks (e.g., top inset in Figure 1). Fortunately, there are natural gaps in a typical mass spectrum, spaced 1 Da apart, because virtually no commonly encountered elemental compositions yield masses at those values (see Figure 5, below—the vertical scale spans a mass defect range of only 0.05–0.26 (rather than 1.00), but includes all observed chemical compositions). Thus, it is possible to break a broadband mass spectrum into 1-Da segments, rotate each segment by 90°, scale each segment according to its mass defect (i.e., difference between exact and nominal mass), and then compress the spacing between the segments to yield a compact display (see Figure 4).[32] Such a display turns out to be the key that unlocks mass assignments up to ~1000 Da.

Kendrick Mass and Kendrick Plot. We previously noted that the repeating mass (CH_2) for a carbon distribution series is 14.01565 Da. It is particularly useful to rescale the mass spectrum from the usual IUPAC mass scale (based on the ^{12}C atomic mass as exactly 12 Da) to the "Kendrick" mass scale:[33]

$$\text{Kendrick mass} = \text{IUPAC mass} \times (14/14.01565) \quad (2)$$

The Kendrick scale effectively converts the mass of CH_2 from 14.01565 to exactly 14.00000. Thus, homologous series (namely, compounds with the same constitution of heteroatoms and number of rings plus double bonds, but different numbers of CH_2 groups) will have identical Kendrick mass defect:

$$\text{Kendrick mass defect} = (\text{exact Kendrick mass} - \text{nominal Kendrick mass}) \quad (3)$$

Figure 5 is two-dimensional slice (i.e., peaks with magnitude greater than 3σ of baseline noise) of the three-dimensional display of Figure 4, rescaled from IUPAC to Kendrick mass. For simplicity, only data from odd nominal masses are shown. First, compounds of the same class and type but different number of CH_2 units are separated by 14 Da in nominal Kendrick mass, but zero difference in Kendrick mass defect, and thus fall on a single horizontal line. Second, compounds of different class are

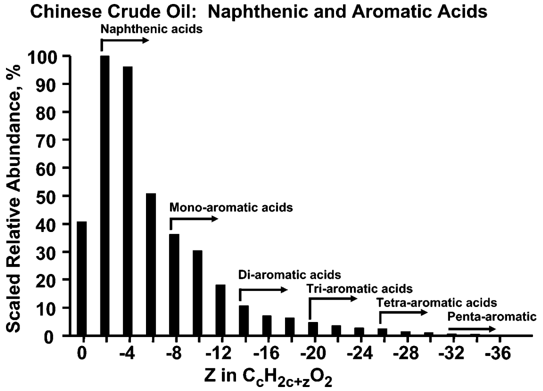


FIGURE 6. Type distribution for $C_cH_{2c+Z}O_2$ acids identified in an ESI FT-ICR mass spectrum of Chinese crude oil. The Chinese crude is highly naphthenic with a predominance of one- and two-ring acids (i.e., $0 > Z > -8$, for species containing two oxygen atoms). The acids may contain up to five aromatic rings.[27]

now readily identified, because the Kendrick mass defect for compounds of a given class is displaced vertically from those of other classes (see the horizontal carbon distributions for the O and O_2 classes in Figure 5). Third, every additional ring or double bond lowers the nominal Kendrick mass by 2 Da, but also increases the Kendrick mass defect. For example, Figure 5 shows clear vertical separation between carbon distributions (horizontal rows of black data points) for the O_2-compounds containing 1−16 rings plus double bonds.

The primary value of the Kendrick plot[32] is that once a few related compounds have been identified, extension of that pattern to higher mass allows for confident elemental composition assignment of ions whose mass would otherwise be too high to allow for unique assignment based only on the measured mass of that peak. Thus, despite increasing vertical uncertainty with increasing nominal mass in Figure 5, it is still possible to assign high-mass peaks with confidence by extrapolation from lower-mass members of an existing carbon series uniquely assigned at low mass.

If an elemental composition for a compound whose carbons are all ^{12}C is assigned at a given *even* mass, then one would expect to find the same elemental composition, except with one ^{13}C instead of ^{12}C, at an *odd* mass, 1.00335 Da (IUPAC) mass units higher than the corresponding even-mass peak. Thus, Kendrick plots at odd and even masses provide a convenient double-check on mass assignments. Finally, the data in the horizontal rows of Figure 5 begin to scatter noticeably above ~600 Da for two reasons. First, the FT-ICR mass resolving power decreases with increasing mass.[34] Second, above ~600 Da, the ICR signal from ions of a given listed elemental composition may no longer be fully resolved from that for ions of other elemental composition(s), so that the observed signal appears to shift slightly from its expected ICR frequency—evidence for the need for even higher magnetic field.

What To Do with the Data? The next question is how to make use of such a vast array of thousands of elemental compositions. As one example, Figure 6 shows a type distribution (scaled as hydrogen deficiency, or Z, values) for all species containing two oxygen atoms in a Chinese crude oil. (All members of the carbon distribution for each class and type are combined in each vertical bar.) One can thus gain an understanding of the distribution of aromatic ring sizes in the mixture. In contrast, Figure 7 shows the class distribution for each of three crude oils of different geographic origin, revealing major differences in the relative abundances of species containing different number of N, S, and O atoms. It is also possible to compare carbon distributions (not shown) for different crude oils—the resulting differences in alkylation patterns should prove useful for modeling of the flow properties of such liquids.

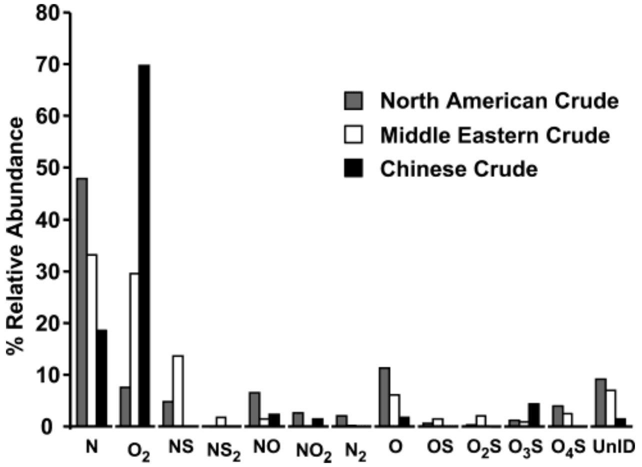


FIGURE 7. Comparison of compound classes identified in each of three crude oils. Relative abundance for acids (O_2) mirror the total acid number (TAN) values: Chinese crude > Middle Eastern crude > North American crude. The Middle Eastern and North American crude oils contain a high relative abundance of sulfur-containing classes, as expected from their high sulfur content. Only classes with a relative abundance >1% are shown.[27]

Noncovalent Aggregates. Virtually all of the ions in the crude oil ESI FT-ICR mass spectrum of Figure 2 are below ~1000 Da in mass. That result contrasts strikingly with prior estimates of crude oil molecular weight distributions based on size exclusion chromatography, light scattering, etc., but is corroborated by recent fluorescence depolarization results,[35] as explained in a recent review.[36] At the high concentrations typically used for non-MS analysis, noncovalent adducts and aggregates are likely present. Noncovalent adducts are *not* present in our ESI FT-ICR mass spectra, because we see no significant change in the mass spectra on CO_2 infrared laser irradiation. For mixtures of porphyrins, on the other hand, noncovalent dimers dominate the ESI FT-ICR mass spectrum and may be dissociated by IR irradiation.[37]

Caveats

Because ionization efficiency (for any ionization method) for one species can be greatly affected by the presence of other species, it is not easy to relate the observed ion relative abundances to the relative abundances of their precursor neutrals in the original sample. For example, positive-ion electrospray ionization favors the most basic compounds, whose presence can reduce the relative abundance of species of lower basicity. Ultimately, it will be necessary to calibrate the relative ionization efficiencies by spiking the mixture with species of known ionization efficiency.

Moreover, electrospray selectively ionizes only the relatively minor polar constituents of crude oil, and it is important to gain access to less polar species (e.g., paraffins, cycloparaffins, aromatics, thiophenes, etc.) Field desorption ionization coupled with FT-ICR mass analysis promises to extend compositional analysis to hydrocarbons (at least the aromatics) and other nonpolar components.[38] Further extensions to field ionization, atmospheric pressure chemical ionization, and atmospheric pressure photoionization promise to broaden the range of compounds accessible to ultrahigh-resolution FT-ICR mass analysis.

Finally, additional structural information is, in principle, available by measuring at high resolution the mass of ion fragments produced by dissociation of precursor molecular or quasimolecular ions, i.e., tandem mass spectrometry (MS/MS).[39] However, the best present resolution in mass selection of precursor ions is about 100 ppm,[40] so such experiments can currently select ions only to within ∼1 Da. Since we already know that such a range can include at least 29 different elemental compositions,[41] MS/MS experiments are presently at the edge of feasibility.

Future Projections

ESI FT-ICR MS alone exposes characteristic differences between crude oils of different origin and maturity.[27] Thousands of N-, S-, and O-containing components in diesel fuel may be monitored simultaneously at various stages of processing[19,42]—information critical to the removal of such species as catalysis poisons and pollutants resulting from combustion. A single ESI FT-ICR mass spectrum of crude oil can resolve and identify ∼20 000 different elemental compositions, a number that should double when combined with FT-ICR mass spectra generated by other ionization methods. Thus, the “petroleome” for a single crude oil may contain as many or more compositionally distinct components as there are open reading frames in the human genome![43]

An important class of future applications is the forensic identification of complex mixtures. Human fingerprints or DNA are useful for identification because their patterns are so complex. Similarly substances which themselves are compositionally complex may be identified precisely by virtue of that complexity. For example, arson “accelerants” (kerosene, gasoline, diesel fuel, lighter fluid, etc.) contain components whose elemental compositions are characteristic of each mixture. A small amount of uncombusted accelerant may thus be compared by FT-ICR MS to a library of such accelerants for forensic identification.[44] The same idea applies to explosives.[45]

Other complex organic mixtures also lend themselves to compositional analysis by FT-ICR MS: coal (more than 10 000 resolved elemental compositions),[46] humic and fulvic acids,[47−52] organic phosphates,[53] etc. Environmental weathering[54,55] and biodegradation of fossil fuels[56] may now be understood in compositional detail.

Finally, chemical composition requires both identity and quantity of each component. This paper mainly addresses identification. The ultimate success of “petroleomics” will require advances in quantification, modeling, and informatics. For example, since MS cannot distinguish isomers by mass alone, other tools such as NMR, IR, XANES, and chromatography will be needed to enable full “petroleomics”.

The authors thank the following colleagues, who have participated as coauthors in the body of work underlying this review: Karin V. Andersen, Erin N. Blumer, Helen J. Cooper, William T. Cooper, Mark R. Emmett, Anne Fievre, Michael A. Freitas, Shenheng Guan, Mark A. Greaney, Christopher L. Hendrickson, Christine A. Hughey, Sara Jernström, William M. Landing, Daniel G. McIntosh, Kuangnan Qian, John P. Quinn, Winston K. Robbins, Tanner M. Schaub, Stuart E. Scheppele, Michael V. Senko, Touradj Solouki, Alexandra C. Stenson, Clifford C. Walters, Forest M. White, and Zhigang Wu. Finally, we thank Carol L. Nilsson for suggesting the term, “petroleomics”. The work cited here has been supported by Amoco, ExxonMobil Research and Engineering, NJ, The National Science Foundation (currently CHE-99-09502), Florida State University, The Ohio State University, and the National High Magnetic Field Laboratory in Tallahassee, FL.

References

(1) Quann, R. J.; Jaffe, S. B. Structure-Oriented Lumping: Describing the Chemistry of Complex Hydrocarbon Mixtures. *Ind. Eng. Chem. Res.* **1992**, *31*, 2483−2497.

(2) Jaffe, S. B. Compositional Based Modeling. *Abstracts of Papers*, 222nd National Meeting of the American Chemical Society, Chicago, IL, Fall 2001; American Chemical Society: Washington, DC, 2001; PETR-032.

(3) Audi, G.; Wapstra, A. H. The 1995 Update to the Atomic Mass Evaluation. *Nucl. Phys. A* **1995**, *595*, 409−425.

(4) Beynon, J. H.; Williams, A. E. *Mass and Abundance Tables for Use in Mass Spectrometry*; Elsevier: New York, 1963.

(5) Comisarow, M. B.; Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *25*, 282−283.

(6) Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Acc. Chem. Res.* **1985**, *18*, 316−322.

(7) 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.

(8) 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.

(9) Qian, K.; Robbins, W. K.; Hughey, C. A.; Cooper, H. J.; Rodgers, R. P.; Marshall, A. G. Resolution and Identification of 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.

(10) Marshall, A. G.; Hendrickson, C. L.; Shi, S. D.-H. Scaling MS Plateaus with High-Resolution FTICR MS. *Anal. Chem.* **2002**, *74*, 252A−259A.

(11) McLafferty, F. W.; Turecek, F. *Interpretation of Mass Spectra*; University Science Books: Sausalito, CA, 1993.

(12) Roussis, S. G. Exhaustive Determination of Hydrocarbon Compound Type Distribution by High-Resolution Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1031−1051.

(13) Hsu, C. S.; Liang, Z.; Campana, J. E. Hydrocarbon Characterization by Ultrahigh-Resolution FTICRMS. *Anal. Chem.* **1994**, *66*, 850−855.

(14) Guan, S.; Marshall, A. G.; Scheppele, S. E. Resolution and Chemical Formula Identification of Aromatic Hydrocarbons Containing Sulfur, Nitrogen, and/or Oxygen in Crude Oil Distillates. *Anal. Chem.* **1996**, *68*, 46−71.

(15) Marshall, A. G.; Wang, T.-C. L.; Ricca, T. L. Tailored Excitation for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Chem. Soc.* **1985**, *107*, 7893−7897.

(16) Guan, S.; Marshall, A. G. Stored Waveform Inverse Fourier Transform (SWIFT) Ion Excitation in Trapped-ion Mass Spectrometry: Theory and Applications. *Int. J. Mass Spectrom. Ion Processes* **1996**, *157/158*, 5−37.

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

(18) Rodgers, R. P. I.; White, F. M.; McIntosh, D. G.; Marshall, A. G. 5.6 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer for Analysis of Volatile Complex Mixtures. *Rev. Sci. Instrum.* **1998**, *69*, 2278−2284.

(19) Rodgers, R. P.; Andersen, K. V.; White, F. M.; Hendrickson, C. L.; Marshall, A. G. Resolution, Elemental Composition, and Simultaneous Monitoring by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Organosulfur Species Before and After Diesel Fuel Processing. *Anal. Chem.* **1998**, *70*, 4743−4750.

(20) Brown, L. S.; Gabrielse, G. Geonium Theory. Physics of a Single Electron or Ion in a Penning Trap. *Rev. Mod. Phys.* **1986**, *58*, 233−311.

(21) Ledford, E. B., Jr.; Rempel, D. L.; Gross, M. L. Space Charge Effects in Fourier Transform Mass Spectrometry. Mass Calibration. *Anal. Chem.* **1984**, *56*, 2744−2748.

(22) 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.* **2000**, *195/196*, 591−598.

(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) Flora, J. W.; Harris, J. C.; Muddiman, D. C. High-Mass Accuracy of Product Ions Produced by SORI−CID Using a Dual Electrospray Ionization Source Coupled with FTICR Mass Spectrometry. *Anal. Chem.* **2001**, *73*, 1247−1251.

(25) Zhan, D. L.; Fenn, J. B. Electrospray Mass Spectrometry of Fossil Fuels. *Int. J. Mass Spectrom.* **2000**, *194*, 197−208.

(26) Altgelt, K. H.; Boduszynski, M. M. *Composition and Analysis of Heavy Petroleum Fractions*; Marcel Dekker: New York, 1994.

(27) Hughey, C. A.; Rodgers, R. P.; Marshall, A. G.; Qian, K.; Robbins, W. R. 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.

(28) Marshall, A. G.; Guan, S. Advantages of High Magnetic Field for FT-ICR Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1819−1823.

(29) Rodgers, R. P.; Blumer, E. N.; Hendrickson, C. L.; Marshall, A. G. Stable Isotope Incorporation Triples the Upper Mass Limit for Determination of Elemental Composition by Accurate Mass Measurement. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 835−840.

(30) Chen, L.; Cottrell, C. E.; Marshall, A. G. Effect of Signal-to-Noise Ratio and Number of Data Points upon Precision in Measurement of Peak Amplitude, Position, and Width in Fourier Transform Spectrometry. *Chemom. Intell. Lab. Syst.* **1986**, *1*, 51−58.

(31) Wang, T.-C. L.; Ricca, T. L.; Marshall, A. G. Extension of Dynamic Range in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry via Stored Waveform Inverse Fourier Transform Excitation. *Anal. Chem.* **1986**, *58*, 2935−2938.

(32) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. Kendrick Mass Defect Spectroscopy: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra. *Anal. Chem.* **2001**, *73*, 4676−4681.

(33) Kendrick, E. A Mass Scale Based on $CH_2 = 14.0000$ for High-Resolution Mass Spectrometry of Organic Compounds. *Anal. Chem.* **1963**, *35*, 2146−2154.

(34) Marshall, A. G.; Comisarow, M. B.; Parisod, G. Relaxation and Spectral Line Shape in Fourier Transform Ion Cyclotron Resonance Spectroscopy. *J. Chem. Phys.* **1979**, *71*, 4434−4444.

(35) Groenzin, H.; Mullins, O. C. Asphaltene Molecular Size and Structure. *J. Phys. Chem. A* **1999**, *103*, 11237−11245.

(36) Buenrostro-Gonzalez, E.; Groenain, H.; Lira-Galeana, C.; Mullins, O. C. The Overriding Chemical Principles that Define Asphaltenes. *Energy Fuels* **2001**, 972−978.

(37) 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.

(38) Schaub, T. M.; Hendrickson, C. L.; Qian, K.; Quinn, J. P.; Marshall, A. G. High-Resolution Field Desorption Ionization Fourier Transform Ion Cyclotron Resonance High-Resolution Mass Analysis of Non-Polar Molecules. *Anal. Chem.* **2003**, *75*, 2172−2176.

(39) McLafferty, F. W. High-Resolution Tandem FT Mass Spectrometry above 10 kDa. *Acc. Chem. Res.* **1994**, *27*, 379−386.

(40) Chen, L.; Marshall, A. G. Stored Waveform Mass-Selective Simultaneous Ejection/Excitation for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Int. J. Mass Spectrom. Ion Processes* **1987**, *79*, 115−125.

(41) 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.

(42) 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.

(43) Butler, D.; Smaglik, P. Draft Data Leave Geneticists with a Mountain To Climb. *Nature* **2000**, *405*, 984−985.

(44) Rodgers, R. P.; Blumer, E. N.; Freitas, M. A.; Marshall, A. G. Compositional Analysis for Identification of Arson Accelerants by Electron Ionization Fourier Transform Ion Cyclotron Resonance High-Resolution Mass Spectrometry. *J. Forensic Sci.* **2001**, *46*, 268−279.

(45) Wu, Z.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G. Compositional Analysis of Military Explosives by Electrospray Ionization Fourier Transform Negative Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2002**, *74*, 1879−1883.

(46) Wu, Z.; Jernström, S.; Hughey, C. A.; Rodgers, R. P.; Marshall, A. G. Resolution of 10,000 Compositionally Distinct Components in Polar Coal Extracts by Negative-Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Energy Fuels* **2002**, *17*, 946−953.

(47) Fievre, A.; Solouki, T.; Marshall, A. G.; Cooper, W. T. High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Humic and Fulvic Acids by Laser Desorption/Ionization and Electrospray Ionization. *Energy Fuels* **1997**, *11*, 554−560.

(48) Solouki, T.; Freitas, M. A.; Alomary, A. Gas-Phase Hydrogen/Deuterium Exchange Reactions of Fulvic Acids: An Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectral Study. *Anal. Chem.* **1999**, *71*, 4719−4726.

(49) Kujawinski, E. B.; Hatcher, P. G.; Freitas, M. A. High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Humic and Fulvic Acids: Improvements and Comparisons. *Anal. Chem.* **2002**, *74*, 413−419.

(50) Kujawinski, E. B.; Freitas, M. A.; Zang, X.; Hatcher, P. G.; Green-Church, K. B.; Jone, R. B. The Application of Electrospray Ionization Mass Spectrometry (ESI MS) to the Structural Characterization of Natarual Organic Matter. *Org. Geochem.* **2002**, *33*, 171−180.

(51) Stenson, A. C.; Landing, W. M.; Marshall, A. G.; Cooper, W. T. Ionization and Fragmentation of Humic Substances in Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2002**, *74*, 4397−4409.

(52) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. Exact Masses and Chemical Formulas of Individual Suwannee River Fulvic Acids from Ultrahigh-Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal. Chem.* **2003**, *75*, 1275−1284.

(53) Llewellyn, J. M.; Landing, W. M.; Marshall, A. G.; Cooper, W. T. Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Dissolved Organic Phosphorus Species in a Treatment Wetland after Selective Isolation and Concentration. *Anal. Chem.* **2002**, *74*, 600−606.

(54) Rodgers, R. P.; Blumer, E. N.; Freitas, M. A.; Marshall, A. G. 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*, 5171−5176.

(55) Rodgers, R. P.; 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.

(56) Rodgers, R. P.; Blumer, E. N.; Freitas, M. A.; Marshall, A. G. Efficacy of Bacterial Bioremediation: Demonstration of Complete Incorporation of Hydrocarbons into Membrane Phospholipids from Thodococcus Hydrocarbon Degrading Bacteria by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environ. Sci. Technol.* **2000**, *34*, 535.

AR020177T