Petroleomics: Chemistry of the underworld

Alan G. Marshall[a,b,1] and Ryan P. Rodgers[a,b,1]

[a]National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, FL 32310-4005; and [b]Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306

Edited by Fred W. McLafferty, Cornell University, Ithaca, NY, and approved August 14, 2008 (received for review May 24, 2008)

Each different molecular elemental composition—e.g., $C_cH_hN_nO_oS_s$—has a different exact mass. With sufficiently high mass resolving power ($m/\Delta m_{50\%} \approx 400{,}000$, in which m is molecular mass and $\Delta m_{50\%}$ is the mass spectral peak width at half-maximum peak height) and mass accuracy (<300 ppb) up to ≈800 Da, now routinely available from high-field (≥9.4 T) Fourier transform ion cyclotron resonance mass spectrometry, it is possible to resolve and identify uniquely and simultaneously each of the thousands of elemental compositions from the most complex natural organic mixtures, including petroleum crude oil. It is thus possible to separate and sort petroleum components according to their heteroatom class ($N_nO_oS_s$), double bond equivalents (DBE = number of rings plus double bonds involving carbon, because each ring or double bond results in a loss of two hydrogen atoms), and carbon number. "Petroleomics" is the characterization of petroleum at the molecular level. From sufficiently complete characterization of the organic composition of petroleum and its products, it should be possible to correlate (and ultimately predict) their properties and behavior. Examples include molecular mass distribution, distillation profile, characterization of specific fractions without prior extraction or wet chemical separation from the original bulk material, biodegradation, maturity, water solubility (and oil:water emulsion behavior), deposits in oil wells and refineries, efficiency and specificity of catalytic hydroprocessing, "heavy ends" (asphaltenes) analysis, corrosion, etc.

Fourier transform | ion cyclotron resonance | mass spectrometry | petroleum | fossil fuel

The rapidly ballooning interest in characterization of petroleum crude oil and its products derives from the confluence of three recent developments: (*i*) the rapidly increasing cost of crude oil (up to $120 per barrel at this writing), (*ii*) the global market shift toward heavier/more acidic/higher sulfur crude oil as the supplies of light "sweet" (low sulfur) crudes are depleted, and (*iii*) the introduction of ultrahigh-resolution mass analysis to separate and identify up to tens of thousands of crude oil components in a single step. Because the organic composition of petroleum is so complex, its characterization was until recently limited to bulk properties (e.g., density, viscosity, osmotic pressure, light scattering, UV-visible and infrared spectroscopy, NMR, x-ray scattering, and absorption-edge spectroscopy) and various wet chemical separations based on (e.g.) solubility, boiling point, gas chromatography, and liquid chromatography. The historical development of those applications, as well as prior low- and high-resolution mass spectrometry of petroleum, have been reviewed elsewhere (1).

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) offers the highest available broadband mass resolution, mass resolving power, and mass accuracy (2). It was first applied to petroleum via electron ionization of petroleum distillates (3, 4). However, the relatively low magnetic field (3.0 T) limited the m/z range, the need to volatilize the sample limited the molecular weight range (especially for heteroatom-containing species), and electron ionization produced extensive fragmentation (especially of alkyl chains).

Electrospray ionization (ESI) (5) is most efficient for polar molecules and typically generates positive ions by protonating (basic) neutrals and negative ions by deprotonating (acidic) neutrals. Because many heteroatom-containing components ($N_nO_oS_s$) of petroleum are highly polar, ESI is specific and especially efficient in generating their gas-phase ions. Although petroleum crude oils typically contain 90% hydrocarbons (C_cH_h), the $N_nO_oS_s$-containing molecules are typically the most problematic with respect to pollution, fouling of catalysts, formation of deposits during production and processing, corrosion, emulsions, and the highest-boiling fractions of lowest economic value. ESI coupled with low-resolution MS was first applied to petroleum by Zhan and Fenn (6). High-resolution ESI FT-ICR MS of petroleum has subsequently resolved and identified >17,000 different elemental compositions for organic bases and acid in crude oil (7).

Access to many of the remaining 90% of petroleum components is afforded by field desorption/ionization (FD) and atmospheric pressure photoionization (APPI). Continuous-flow FD FT-ICR MS yields abundant ions from several species not observed by ESI, including benzo- and dibenzothiophenes, furans, cycloalkanes, and polycyclic aromatic hydrocarbons (PAHs) (8). However, FD experiments are slow because of the need to ramp the current to the FD emitter over a period of a couple of minutes so as to volatilize/ionize species of successively increasing boiling point. APPI FT-ICR MS (9) can accumulate one mass spectral dataset in a few seconds and is thus better suited for signal averaging of a few hundred scans for increased dynamic range. APPI of petroleum requires the ultrahigh resolution of FT-ICR MS for two reasons: (*i*) APPI ionizes a broader range of compound classes, and thus the mass spectrum contains approximately five times as many peaks as an ESI mass spectrum of the same crude oil sample; and (*ii*) the same analyte molecule may produce both $M^{+\bullet}$ and $(M+H)^+$ ions, making it necessary to resolve $M^{+\bullet}$ containing one ^{13}C from $(M+H)^+$ containing all ^{12}C, a mass difference of only 4.5 mDa (see below). Laser desorption/ionization mass spectrometry, although highly useful for biomolecule analysis, does not provide a reliable representation of petroleum components because of significant aggregation and fragmentation at the laser power required to generate a useful number of gas-phase ions (10). However, recent introduction of a two-color laser method (11), in which the first laser desorbs the neutrals and the second laser ionizes them, appears to overcome the disadvantages of single-color laser desorption/ionization.

Saturated hydrocarbons are especially problematic, because virtually all methods for ionizing neutral saturated hydrocarbons produce extensive fragmentation, thereby making it hard to identify the neutral precursors in the original sample as well as vitiating quantitation. Recently, laser-induced acoustic desorption (LIAD) of neutrals, followed by chemical ionization with appropriate reagents, has shown great promise in achieving

Author contributions: A.G.M. and R.P.R. designed research; R.P.R. performed research; R.P.R. contributed new reagents/analytic tools; R.P.R. analyzed data; and A.G.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. E-mail: marshall@magnet.fsu.edu or rodgers@magnet.fsu.edu.

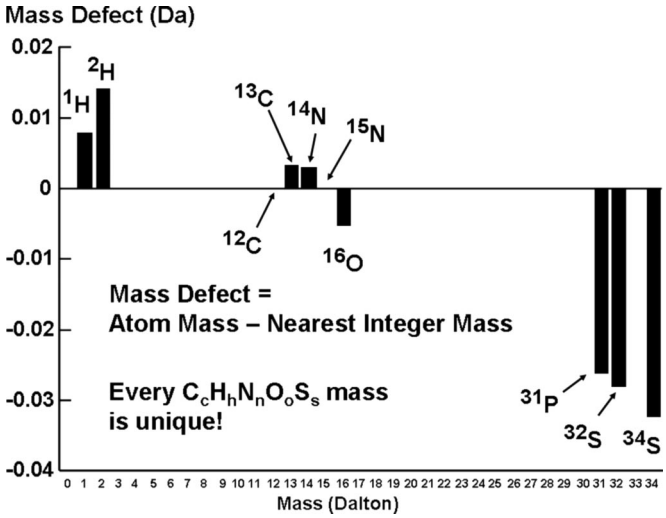

Fig. 1. Atomic mass defects for selected isotopes of some common chemical elements. Because no two have the same mass defect, it is possible to determine a unique elemental composition for any molecule from a sufficiently accurate mass measurement.

efficient (and uniformly efficient) ionization of saturated hydrocarbons over a wide molecular mass range (12, 13).

"Petroleomics" is the principle that from sufficiently complete characterization of the organic composition of petroleum and its relatives and products, it should be possible to correlate (and ultimately predict) their properties and behavior (7, 14). In this article, we describe the basis for extraction of elemental compositions from sufficiently accurate mass measurements and present various methods to sort the thousands of resulting chemical formulas in chemically and practically important ways.

Mass Defects: The Key to Unlocking Chemical Formulas

Ultrahigh mass resolving power ($m/\Delta m_{50\%} \geq 400{,}000$) up to ≈800 Da is needed to resolve the thousands of peaks in a petroleum mass spectrum (15). Ultrahigh mass accuracy is needed for a different reason. Fig. 1 shows the mass defects (defined in the figure) for the most abundant isotope of each of several common chemical elements. Every isotope of every element has a different mass defect. (By definition, the mass defect of ^{12}C is zero.) Thus, if the mass of a molecule can be measured with sufficiently high accuracy (in practice, ≈0.0003 Da for molecules up to ≈1,000 Da in mass), then its elemental composition can usually be determined uniquely from its mass alone (16, 17). Fig. 2 provides examples of some of the closest mass "splits" encountered in mass spectrometry of petroleum. Two of the most important close doublets are molecules whose elemental compositions differ by $^{12}C_3$ vs. $^{32}SH_4$ (to identify sulfur-containing components, see Fig. 2) and ^{13}C vs. $^{12}C^{1}H$ (for APPI MS, see above).

Results and Discussion

Molecular Mass Distribution. The first step in characterization of organics in petroleum is to determine the molecular mass distribution, particularly for "heavy ends" high-boiling components and residues. It is now well established that the average molecular mass of asphaltenes is <1,000 Da (18). Higher average molecular masses inferred from prior vapor-phase osmometry and size-exclusion chromatography may be attributed to formation of noncovalent aggregates because of too-high concentration and/or use of solvents that promote aggregation. Aggregation has been demonstrated by mass-selective isolation of singly charged dimer ions in the range 840–860 Da, followed by collisional activation at collision energies too low to break covalent bonds. The heterodimers dissociate to yield a distribution of monomeric ions, and those ions do not dissociate when subjected to the same collision conditions (1). In this way, aggregates up to tetramers have been demonstrated. Moreover, the abundance of dimers relative to monomers increases with increasing sample concentration—additional evidence for noncovalent complexes.

Elemental Composition Assignment from Accurate Mass Measurement. Once the thousands of peaks in an FT-ICR mass spectrum of a petroleum sample have been resolved and converted from

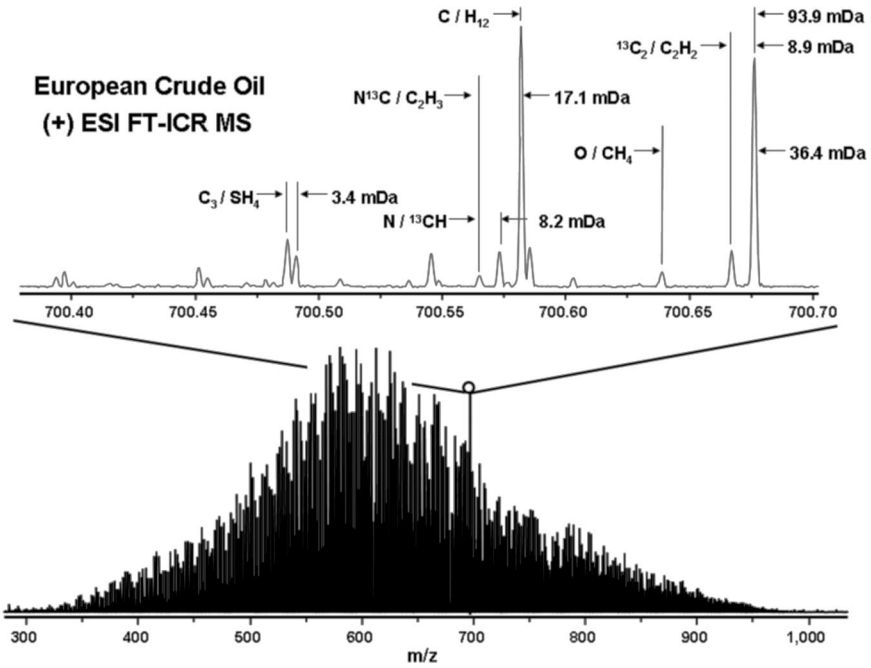

Fig. 2. Positive-ion electrospray 9.4-T FT-ICR mass spectrum of a European crude oil, containing ≈8,000 resolved and identified peaks. Several of the mass splits commonly encountered in crude oil are shown in the mass-scale-expanded 300-mDa *Inset*. Data were provided by A. M. McKenna.

ion cyclotron frequency to mass, the next step is to assign each mass to a unique elemental composition. Even with ultrahigh mass resolving power (e.g., $m/\Delta m_{50\%}$ = 400,000 up to 800 Da), such assignments may not be possible based on mass measurement alone. Fortunately, petroleum and its derivatives typically consist primarily of homologous series. Each homologous series is categorized by heteroatom class ($N_nO_oS_s$), double bond equivalents (DBE) (see Eq. **1**), and carbon number. The DBE is equal to the number of rings plus double bonds involving carbon (19):

$$\text{Double bond equivalents } (C_cH_hN_nO_oS_s) = c - h/2 + n/2 + 1. \quad [1]$$

Within each heteroatom class, many DBE values are possible. For example, $C_{42}H_{59}N_1$ belongs to the N_1 class with DBE = 14, whereas $C_{42}H_{53}N_1$ belongs to the same heteroatom class, N_1, but its DBE is 17. DBE thus affords a direct measure of aromaticity of petroleum components. [In the petroleum industry, a more common index is the "hydrogen deficiency," Z, defined from the elemental composition, C_cH_{2c+Z}. We prefer DBE, because a saturated hydrocarbon with zero rings or double bonds has a DBE of 0 but a Z-value of −2. In any case, it is easy to convert between the two: $Z = -2(\text{DBE}) + n + 2$, in which n is the number of nitrogens in the chemical formula.]

Consider a homologous series of molecules of the same heteroatom class and DBE, but a varying degree of alkylation. Members of that series will vary in composition by multiples of $-CH_2$, corresponding to multiples in mass of 14.0565 Da. A particularly useful way to identify such a series is to convert the Système international d'unités (SI) mass to "Kendrick" mass (20):

$$\text{Kendrick mass} = \text{SI mass} \times (14.00000/14.0565). \quad [2]$$

Thus, successive members of an alkylation series (i.e., same heteroatom class and same number of rings plus double bonds involving carbon) will differ by 14.00000 in Kendrick mass and will therefore each have the same Kendrick mass defect:

$$\text{Kendrick mass defect} = \text{nominal Kendrick mass} - \text{Kendrick mass}, \quad [3]$$

in which nominal Kendrick mass is the Kendrick mass rounded to the nearest integer.

Members of a given alkylation series are thereby readily identified. Because elemental compositions can be assigned with high confidence for the lower mass members of each series, the assignments can be extended to 800–1,000 Da by extrapolation. That procedure was the key to "cracking" the problem of mass assignments for petroleum components (21).

Compositional Sorting: Graphical Images Derived from Heteroatom Class, DBE, and Carbon Number. Having proceeded from thousands of ICR frequencies to thousands of accurate masses to thousands of elemental compositions, the next issue is how to represent those compositions in convenient graphical form rather than in a table with thousands of entries. In fact, the elemental composition yields three independent properties of a molecule: heteroatom class ($N_nO_oS_s$), double bond equivalents, and number of carbons (for a given class and DBE, the carbon number is a measure of the degree of alkylation). Those properties may be sorted hierarchically, as shown in Fig. 3. First, compositions are sorted according to heteroatom class. For a given heteroatom class (in this case, O_2, typically consisting mainly of carboxylic acids), there is a distribution of DBE values. Finally, for a given class (in this case, O_2) and DBE (in this case, 2), one can display the carbon distribution.

Among the various possible graphical images derived from elemental compositions, we find that the most useful is a plot of

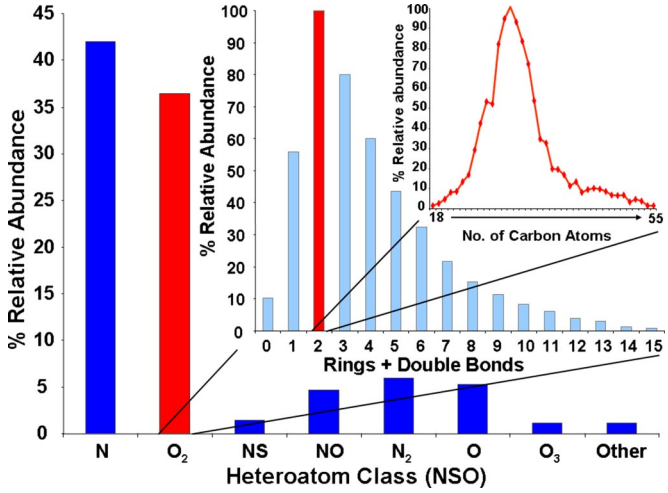

Fig. 3. Sorting of compounds based on their elemental compositions. (*Bottom*) Heteroatom class (all species with the same $N_nO_oS_s$ composition). (*Middle*) Distribution of DBE (double bond equivalents = rings plus double bonds to carbon) distribution for members of the O_2 heteroatom class. (*Top*) Carbon number distribution for O_2 species with DBE = 2. Graphical combinations of these distributions furnish characteristic images of various petroleum materials.

DBE vs. carbon number for a given heteroatom class. For example, Fig. 4 shows such a plot for a distillation cut from Athabasca bitumen. Interestingly, the O_2 class (presumably carboxylic acids) are not aromatic because their DBE values are typically 4 or less (and the simplest aromatic carboxylic acid would have DBE = 5). Plots of DBE vs. carbon number have proved especially useful in characterizing hydrotreatment in petroleum processing (22–24), vacuum gas oil distillation cuts (25), water-soluble acids and bases (26), saturates vs. aromatics (27), heat exchanger deposits (28), oil:water emulsion interfacial material (29–31), (lack of) matrix effects in saturates/aromatics/resins/asphaltenes (SARA) (32) fractionation (33), naphthenic acids (1, 34, 35), sulfur-containing polycyclic aromatics (27, 36–38), distillates (39, 40), and other applications discussed in this article.

Another graphical compositional image is a "van Krevelen" plot,

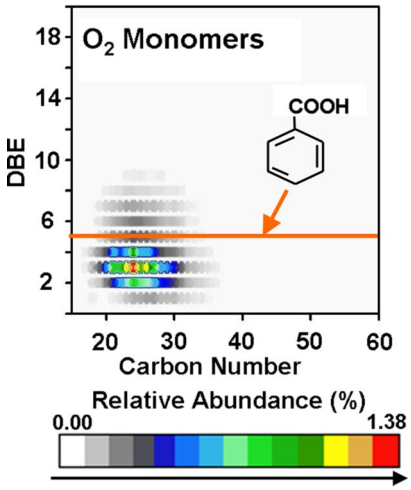

Fig. 4. Plots of double bond equivalents (DBE) vs. carbon number for the 375–400°C distillation cut from the negative-ion ESI 9.4 T FT-ICR mass spectrum of Athabasca bitumen (1 mg/ml) in toluene/methanol, spiked with 2% (by volume) ammonium hydroxide. Data were provided by D. F. Smith.

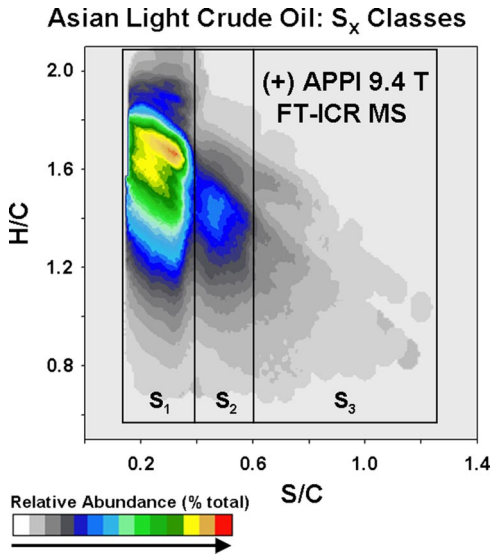

Fig. 5. van Krevelen plot of H/C ratio vs. S/C ratio for all sulfur-containing ions in an APPI positive-ion 9.4 T FT-ICR mass spectrum of an Asian light crude oil. To avoid overlap between S_x classes, only ions with $300 < m/z < 750$ are included. Data were provided by A. M. McKenna.

typically of H/C ratio vs. O/C ratio. The van Krevelen plot (like Kendrick mass) was originally devised to represent bulk elemental compositions (41). The advent of ultrahigh-resolution FT-ICR MS made it possible to extend both of those ideas to distinguish all of the individual elemental compositions simultaneously. The van Krevelen plot was initially applied to dissolved organic matter (42), and later to petroleum (43, 44) and coal (45). van Krevelen plots have been used to characterize, e.g., asphaltene vs. its parent oil (46), biodegradation (14), and crude oil vs. its deposit (47). Fig. 5 is a van Krevelen plot for an Asian light crude oil but plotted here as H/C ratio vs. S/C ratio. The van Krevelen plot nicely separates the sulfur classes (S1, S2, and S3, each of which would require a separate plot of DBE vs. carbon number) on the abscissa, and the ordinate provides a measure of unsaturation. (Note that lower H/C ratio corresponds to higher DBE.)

Nitrogen Speciation. For an APPI FT-ICR mass spectrum, additional information emerges if DBE in Eq. **1** is calculated from the elemental composition of the ion rather than its precursor neutral in the original sample. Thus, odd-electron radical cations have integer DBE values (because loss of an electron does not affect the DBE calculated from Eq. **1**), whereas protonated or deprotonated (even-electron) ions have half-integer DBE values (because changing the number of hydrogens by one changes DBE by ± 0.5 in Eq. **1**). That property makes it possible to distinguish five-membered from six-membered nitrogen ring species as follows (48).

The primary ionization pathways for aromatic N_1 class compounds (acridine, with nitrogen in a six-membered pyridine ring, and carbazole, with nitrogen in a five-membered pyrrolic ring) by APPI (positive-ion) and ESI (negative-ion and positive-ion) differ. Proton transfer yields even-electron $[M+H]^+$ or $[M-H]^-$ ions with half-integer calculated DBE values. Positive-ion APPI can also form odd-electron radical cations with integer DBE values. Thus, from elemental compositions derived from APPI FT-ICR MS, it is possible to distinguish between heterocyclic nitrogen species in six-membered rings [positive-ion ESI (half-integer ion DBE value)] or APPI (half-integer ion DBE value) from five-membered nitrogen ring species [negative-ion ESI (half-integer ion DBE value)] or positive-ion APPI (integer ion DBE value). APPI is more convenient because both types of nitrogen rings can be found and identified in a single (positive-ion) mass spectrum, whereas ESI requires separate positive- and negative-ion mass spectra to reach the same result.

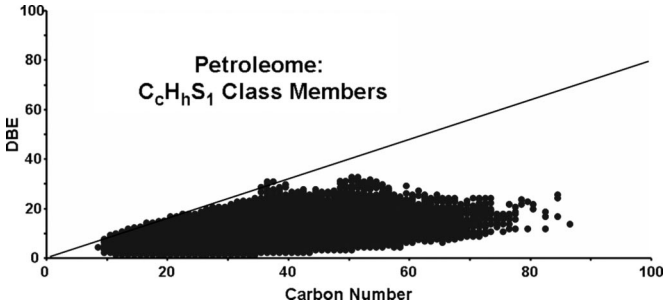

Fig. 6. Plot of double bond equivalents vs. carbon number for all monoisotopic $C_cH_hS_1$ members of the petroleome database, compiled from dozens of positive- and negative-ion ESI and APPI FT-ICR mass spectra. The diagonal line represents the maximum possible DBE for each carbon number. The entries are not abundance-weighted; i.e., each different elemental composition is shown by a single dot. Data were provided by I. Stroe and J. E. Velasquez.

The Petroleome. We have, to date, resolved and identified >60,000 distinct elemental monoisotopic elemental compositions, $^{12}C_c{}^{1}H_h{}^{14}N_n{}^{16}O_o{}^{32}S_s$, from ions in positive- and negative-ion ESI and APPI FT-ICR mass spectra of dozens of petroleum crude oil samples. [Species with one or more ^{13}C, ^{15}N, ^{34}S, etc. are not counted separately because they are chemically essentially identical. Moreover, a neutral composition based on observation of both $M^{+\bullet}$ and $(M+H)^+$ ions is counted only once.] Fig. 6 is a distribution of DBE vs. carbon number for just the $C_cH_hS_1$ members of the "petroleome." The diagonal line represents the maximum possible DBE for each carbon number in a planar polycyclic aromatic molecule. Interestingly, over the tens of millions of years during which petroleum is formed, virtually all possible combinations of DBE and carbon number are generated up to DBE $\approx$ 25 and carbon number $\approx$ 40, and, with decreasing aromaticity, up to carbon number $\approx$ 80. Similar trends are found for other heteroatom classes.

As the petroleome database expands, we envision several uses, by analogy to proteomic and other such databases. First, once the database is sufficiently complete, it should be possible to search only among its members (rather than all chemically possible elemental compositions), thereby simplifying and speeding the assignment of a unique elemental composition to each mass spectral peak. Second, once the monoisotopic species have been identified, one can search for the same species containing one or more ^{13}C, ^{15}N, and/or ^{34}S in place of ^{12}C, ^{14}N, ^{32}S to improve the reliability of the original assignment. Third, by including additional descriptors, it becomes possible to characterize crude oils (and their processed products) with respect to geographic origin (e.g., for oil spills), density, total acid number (milligrams of KOH to neutralize 1 g of sample), corrosivity, tendency to form sodium or calcium naphthenate deposits, emulsions, distillation profiles sorted by heteroatom compound class, fractions separated by solubility and/or chromatographic elution [e.g., saturates/aromatics/resins/asphaltenes, SARA (32)], etc. Finally, similar databases may be compiled and exploited for coal (49), biofuels, humic (50) and fulvic (51) acids, dissolved organic matter (52–54), and other complex organic mixtures.

It is important to recognize that the present analysis is limited to compositional possibilities. For many elemental compositions, there can be multiple isomers (e.g., positional isomers, such as butane vs. isobutane; alcohol vs. ketone, stereoisomers, etc.). Additional (typically spectroscopic) information is needed to distinguish some of those possibilities. For example, the number

of isomers for the single elemental composition, $C_{14}H_{10}$, is 5.3×10^6 (55)!

The biggest remaining problem for mass-based petroleomics is quantitation. No single ionization method produces ions from different analyte neutrals with equal efficiency. For example, ESI typically generates positive ions by protonation (of basic analytes) and negative ions by deprotonation (of acidic analytes). Because ESI is typically achieved by addition of a weak acid (e.g., formic acid) or weak base (e.g., ammonium hydroxide), the species accessible by ESI are typically limited to relatively strong acids (e.g., carboxylic acids) or relatively strong bases (e.g., pyridinic nitrogen). Thus, ESI is "blind" to aromatic hydrocarbons, thiophenes, furans, etc. Efforts are currently underway to extend the range of acidity/basicity for ESI. For example, incubation of petroleum with methyl iodide can result in methylation of benzo- and dibenzothiophenes and their alkylated congeners, thereby producing positive ions in solution for high yield from positive-ion electrospray (37). Unfortunately, the reaction is slow (overnight), and its efficiency diminishes with increasing molecular mass. Much more uniform ionization efficiency across widely different functional groups is currently provided by field desorption or atmospheric pressure photoionization, as noted above.

At this writing, petroleomics has been the subject of two review articles (7, 14), one monograph (47), two petroleomics symposia (*Petroleomics: The Next Grand Challenge for Chemical Analysis*, PittCon 2003, Orlando, FL, March, 2003; *Symp. VI. Petroleomics: From Petroleum Composition to Commercial Reality*, II. International Applied Statistical Physics Molecular Engineering Conference, Puerto Vallarta, Mexico, August, 2003), and 65 journal publications. Petroleomics is rapidly taking its place as the most comprehensive tool for characterization of crude oil and its derivatives.

Materials and Methods

Petroleum samples were typically dissolved at 100 mg/mL in toluene. Two serial dilutions yielded a 2 mg/mL solution that may be further diluted in toluene to 0.5 mg/mL for APPI analysis and 1 mg/mL in toluene:methanol (1:1 vol/vol) for ESI analysis. Conditions for microelectrospray (1) and APPI (9) analyses are described elsewhere. Our FT-ICR mass spectra were acquired with custom-built 9.4 T[56] and 14.5 T[57] instruments. Multiple (100–200) time-domain acquisitions were averaged for each sample, Hanning-apodized, and zero-filled once before fast Fourier transform and magnitude calculation. The quadrupolar electric trapping potential approximation (58) converted ICR frequency to mass-to-charge ratio. Instrument control, data acquisition, and data analysis were handled by a modular ICR data station.[§] Negative ion data were collected with similar parameters and appropriate polarity changes. Petroleum ions were typically singly charged, based on the unit *m/z* separation between $^{12}C_n$ and $^{13}C_1{}^{12}C_{n-1}$ isotopic variants of the same elemental composition (59). FT-ICR mass spectra were internally calibrated with respect to a homologous series of ions present in high abundance in each sample. Masses for singly charged ions with relative abundance of $>6\sigma$ of baseline rms noise were exported to a spreadsheet. Measured masses were then converted from the SI mass scale to the Kendrick mass scale for identification of homologous series. Kendrick mass defect analysis was used for peak assignments as described above (21).

[§]Blakney GT, Hendrickson CL, Marshall AG, Improved data acquisition system for Fourier transform ion cyclotron resonance mass spectrometry, 55th American Society for Mass Spectrometry Annual Conference on Mass Spectrometry, June 3–7, 2007, Indianapolis, IN, MPD067 (abstr).

ACKNOWLEDGMENTS. We thank the various coauthors in our listed references, as well as Priyanka Juyal, Brandie M. Ehrmann, Amy M. McKenna, Mmilili Mapolelo, Jade E. Velasquez, Brian S. Bingham, and Ioana Stroe. This work was supported by the National Science Foundation National High Field Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Facility (Grant DMR-06-54118), Florida State University, and the National High Magnetic Field Laboratory.

1. Smith DF, *et al*. (2007) Self-association of organic acids in petroleum and Canadian bitumen characterized by low- and high-resolution mass spectrometry. *Energy Fuels* 21:1309–1316.
2. Marshall AG, Hendrickson CL, Jackson GS (1998) Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom Rev* 17:1–35.
3. Hsu CS, Liang Z, Campana JE (1994) Hydrocarbon characterization by ultrahigh resolution FTICRMS. *Anal Chem* 66:850–855.
4. Guan S, Marshall AG, Scheppele SE (1996) Resolution and chemical formula identification of aromatic hydrocarbons containing sulfur, nitrogen, and/or oxygen in crude oil distillates. *Anal Chem* 68:46–71.
5. Fenn JB, Mann M, Meng CK, Wong SF (1990) Electrospray ionization–principles and practice. *Electrospray Mass Spectrom Rev* 9:37–70.
6. Zhan DL, Fenn JB (2000) Electrospray mass spectrometry of fossil fuels. *Int J Mass Spectrom*, 194:197–208.
7. Marshall AG, Rodgers RP (2004) Petroleomics: The next grand challenge for chemical analysis. *Acc Chem Res* 37:53–59.
8. Schaub TM, Hendrickson CL, Quinn JP, Rodgers RP, Marshall AG (2005) Instrumentation and method for ultrahigh resolution field desorption ionization Fourier transform ion cyclotron resonance mass spectrometry of non-polar species. *Anal Chem* 77:1317–1324.
9. Purcell JM, Hendrickson CL, Rodgers RP, Marshall AG (2006) Atmospheric pressure photoionization Fourier transform ion cyclotron resonance mass spectrometry for complex mixture analysis. *Anal Chem* 78:5906–5912.
10. Hortal AR, Hurtado P, Martinex-Haya B, Mullins OC (2007) Molecular weight distributions of coal and petroleum asphaltenes from laser desorption/ionization experiments. *Energy Fuels* 21:2863–2868.
11. Pomerantz AE, Hammond MR, Morrow AL, Mullins OC, Zare RN (2008) Two-step laser mass spectrometry of asphaltenes. *J Am Chem Soc* 130:7216–7217.
12. Crawford KE, *et al*. (2005) Laser-induced acoustic desorption/Fourier transform ion cyclotron resonance mass spectrometry for petroleum distillate analysis. *Anal Chem* 77:7916–7923.
13. Duan P, *et al*., (2008) Analysis of base oil fractions by $ClM_n(H_2O)^+$ chemical ionization combined with laser-induced acoustic desorption/Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 80:1847–1853.
14. Rodgers RP, Schaub TM, Marshall AG (2005) Petroleomics: Mass spectrometry returns to its roots. *Anal Chem* 77:20A–27A.
15. Panda SK, Andersson JT, Schrader W (2007) Mass-spectrometric analysis of complex volatile and nonvolatile crude oil components: A challenge. *Anal Bioanal Chem* 389:1329–1339.
16. Beynon JH (1954) Qualitative analysis of organic compounds by mass spectrometry. *Nature* 174:735–737.
17. Kim S, Rodgers RP, Marshall AG (2006) Truly 'exact' mass: Elemental composition can be determined uniquely from molecular mass measurement at ~0.1 mDa accuracy for molecules up to ~500 Da. *Int J Mass Spectrom* 251:260–265.
18. Mullins OC (2006) Rebuttal to comment by professors Herod, Kandiyoti, and Bartle on "Molecular size and weight of asphaltene and asphaltene solubility fractions from coals, crude oils and bitumen." *Energy Fuels* 86:309–312.
19. McLafferty FW, Turecek F (1993) *Interpretation of Mass Spectra* (University Science Books, Mill Valley, CA), 4th Ed, p 371.
20. Kendrick E (1963) A mass scale based on CH_2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Anal Chem* 35:2146–2154.
21. Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K (2001) Kendrick mass defect spectroscopy: A compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal Chem* 73:4676–4681.
22. Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG (2001) Elemental composition analysis of processed and unprocessed diesel fuel by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 15:1186–1193.
23. Klein GC, Rodgers RP, Marshall AG (2006) Identification of hydrotreatment-resistant heteroatomic species in a crude oil distillation cut by electrospray ionization FT-ICR mass spectrometry. *Fuel* 85:2071–2080.
24. Miyabayashi K, *et al*. (2008) Compositional analysis of deasphalted oil before and after hydrocracking over zeolite catalyst by Fourier transform ion cyclotron resonance mass spectrometry. *Fuel Process Technol* 89:397–405.
25. Stanford LA, Kim S, Rodgers RP, Marshall AG (2006) Characterization of compositional changes in vacuum gas oil distillation cuts by electrospray ionization FT-ICR mass spectrometry. *Energy Fuels* 20:1664–1673.
26. Stanford LA, *et al*. (2007) Identification of water-soluble heavy crude oil organics. Acidic and basic NSO compounds in fresh water and sea water by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Environ Sci Technol* 41:2696–2702.
27. Purcell J, *et al*. (2007) Sulfur speciation in petroleum: Atmospheric pressure photoionization or chemical derivatization and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 21:2869–2874.
28. Schaub TM, Jennings DW, Kim S, Rodgers RP, Marshall AG (2007) Heat exchanger deposits in an inverted steam assisted gravity drainage operation. Part 2. Organic acid analysis by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 21:185–194.
29. Hemmingson PV, *et al*. (2006) Structural characterization and interfacial behavior of acidic compounds extracted from a North Sea oil. *Energy Fuels* 20:1980–1987.

30. Stanford LA, Rodgers RP, Marshall AG, Czarnecki J, Wu XA (2007) Compositional characterization of bitumen/water emulsion films by negative- and positive-ion electrospray ionization and field desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 21:973–981.
31. Stanford LA, *et al*. (2007) Detailed elemental compositions of emulsion interfacial material vs. parent oil for nine geographically distinct light, medium, and heavy crude oils, detected by negative- and positive-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 21:963–972.
32. Rudzinski WE, Aminabhavi TM, Sassman S, Watkins LM (2000) Isolation and characterization of the saturate and aromatic fractions of a Maya crude oil. *Energy Fuels* 14:839–844.
33. Klein GC, Angström A, Rodgers RP, Marshall AG (2006) Use of saturates/aromatics/resins/asphaltenes (SARA) fractionation to determine matrix effects in crude oil analysis by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 20:668–672.
34. Barrow MP, Headley JV, Peru KM, Derrick PJ (2004) Fourier transform ion cyclotron resonance mass spectrometry of principal components in oilsands naphthenic acids. *J Chromatogr A* 2058:51–59.
35. Headley JV, Peru KM (2007) Characterization of naphthenic acids from Athabasca oil sands using electrospray ionization: The significant influence of solvents. *Anal Chem* 79:6222–6229.
36. Rodgers RP, Andersen KV, White FM, Hendrickson CL, Marshall AG (1998) 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* 70:4743–4750.
37. Müller H, Andersson JT, Schrader W (2005) Characterization of high-molecular-weight sulfur-containing aromatics in vacuum residues using Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 77:2536–2543.
38. Panda SK, Schrader W, Al-Hajji A, Andersson JT (2007) Distribution of polycyclic aromatic sulfur heterocycles in three Saudi Arabian crude oils as determined by Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 21:1071–1077.
39. Pakarinen JMH, Teräväinen MJ, Pirskanen A, Wickström K, Vainiotalo P (2007) A positive-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry study of Russian and North Sea crude oils and their six distillation fractions. *Energy Fuels* 21:3369–3374.
40. Teräväinen MJ, Pakarinen JMH, Wickström K, Vainiotalo P (2007) Comparison of the composition of Russian and North Sea crude oils and their eight distillation fractions studied by negative-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry: The effect of suppression. *Energy Fuels* 21:266–273.
41. van Krevelen DW (1950) Graphical-statistical method for the study of structure and reaction processes of coal. *Fuel* 29:269–284.
42. Kim S, Kramer RW, Hatcher PG (2003) Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Anal Chem* 75:5336–5344.
43. Wu Z, Strohm JJ, Song C, Rodgers RP, Marshall AG (2005) Comparative compositional analysis of untreated and hydrotreated oil by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 19:1072–1077.
44. Kim S, *et al*. (2005) Microbial alteration of the acidic and neutral polar NSO compounds in petroleum revealed by Fourier transform ion cyclotron resonance mass spectrometry. *Org Geochem* 36:1117–1134.
45. Wu Z, Rodgers RP, Marshall AG (2004) Compositional determination of acidic species in Illinois #6 coal extracts by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 18:1424–1428.
46. Klein GC, Kim S, Rodgers RP, Marshall AG, Yen A (2006) Mass spectral analysis of asphaltenes. II. Detailed compositional comparison of asphaltenes deposit to its crude oil counterpart for two geographically different crude oils by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels* 20:1973–1979.
47. Rodgers RP, Marshall AG (2006) *Chapter 3: Petroleomics: Advanced Characterization of Petroleum Derived Materials by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS)*. Asphaltenes, heavy oils and petroleomics, eds Mullins OC, Sheu EY, Hammami A, Marshall AG (Springer, New York), pp 63–93.
48. Purcell JM, Rodgers RP, Hendrickson CL, Marshall AG (2007) Speciation of nitrogen containing aromatics by atmospheric pressure photoionization or electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* 18:1265–1273.
49. Wu Z, Jernström S, Hughey CA, Rodgers RP, Marshall AG (2003) 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* 17:946–953.
50. Stenson AC, Landing WM, Marshall AG, Cooper WT (2002) Ionization and fragmentation of humic substances in electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 74:4397–4409.
51. Stenson AC, Marshall AG, Cooper WT (2003) 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* 75:1275–1284.
52. Tremblay LB, Dittmar T, Marshall AG, Cooper WJ, Cooper WT (2007) Molecular characterization of dissolved organic matter in a north Brazilian mangrove porewater and mangrove-fringed estuary by ultrahigh resolution Fourier transform-ion cyclotron resonance mass spectrometry and excitation/emission spectroscopy. *Mar Chem* 105:15–29.
53. Kujawinski EB, Behn M (2006) Automated analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra of natural organic matter. *Anal Chem* 78:4363–4373.
54. Hertkorn N, *et al*. (2007) An integrated NMR and FTCIR mass spectroscopic study to characterize a new and major refractory component of (marine) natural organic matter (NOM) at the molecular level, CRAM: Carboxyl-rich alicyclic molecules. *Geochim Cosmochim Acta* 71:A399–A399.
55. Hertkorn N, *et al*. (2007) High-precision frequency measurements: Indispensable tools at the core of the molecular-level analysis of complex systems. *Anal Bioanal Chem* 389:1311–1327.
56. Senko MW, Hendrickson CL, Emmett MR, Shi SDH, Marshall AG (1997) External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* 8:970–976.
57. Schaub TM, *et al*. (2008) High-performance mass spectrometry: Fourier transform ion cyclotron resonance at 14.5 tesla. *Anal Chem* 80:3985–3990.
58. Ledford EB, Jr, Rempel DL, Gross ML (1984) Space charge effects in Fourier transform mass spectrometry. Mass calibration. *Anal Chem* 56:2744–2748.
59. Senko MW, Beu SC, McLafferty FW (1995) Automated assignment of charge states from resolved isotopic peaks for multiply charged ions. *J Am Soc Mass Spectrom* 6:52–56.