

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231272517>

Correlation of FT-ICR Mass Spectra with the Chemical and Physical Properties of Associated Crude Oils

ARTICLE in ENERGY & FUELS · SEPTEMBER 2010

Impact Factor: 2.79 · DOI: 10.1021/ef1007165

CITATIONS

27

READS

164

8 AUTHORS, INCLUDING:



[Manhoi Hur](#)

Iowa State University

22 PUBLICATIONS 161 CITATIONS

[SEE PROFILE](#)



[Eunkyoung Kim](#)

SK Biopharmaceuticals

11 PUBLICATIONS 169 CITATIONS

[SEE PROFILE](#)



[Jae Won Lee](#)

Korea University

106 PUBLICATIONS 2,404 CITATIONS

[SEE PROFILE](#)



[Sunghwan Kim](#)

Kyungpook National University

66 PUBLICATIONS 2,232 CITATIONS

[SEE PROFILE](#)

Correlation of FT-ICR Mass Spectra with the Chemical and Physical Properties of Associated Crude Oils

Manhoi Hur,^{†,‡} Injoon Yeo,^{§,‡} Eunkyoung Kim,^{||} Myoung-han No,^{||} Jaesuk Koh,^{||} Yun Ju Cho,[§] Jae Won Lee,[‡] and Sunghwan Kim^{*,§}

[†]Department of Bioinformatics, [‡]Department of Statistics, Korea University, Seoul, Korea, Republic of Korea, [§]Department of chemistry, Kyungpook National University, Daegu, Korea, and ^{||}SK energy Institute of Technology, Daejeon, Korea. [†]Authors have contributed equally to this work.

Received June 8, 2010. Revised Manuscript Received August 11, 2010

In this study, the peaks observed using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) were correlated with properties of crude oils. The correlations were statistically analyzed and graphically presented using Circos diagrams. Numerous peaks with statistical significance ($p < 0.05$) correlated strongly with elemental sulfur, nitrogen, nickel, and vanadium contents. In addition, a number of peaks correlated with properties such as acidity, gravity, and weight percent of residue after atmospheric residue distillation of crude oils. The correlation agreed well with generally accepted ideas, thereby validating this approach. For example, sulfur-containing classes such as S₁, S₂, and NS correlated positively with sulfur content. Positive correlation denotes that the relative abundance of the peaks containing S₁, S₂, and NS heteroatoms increased as bulk concentrations of sulfur in the samples increased. The O₂ and O₄ classes of compounds, presumably with COOH functional groups, had a strong correlation with total acid number. Subsequent analyses showed some correlations had carbon number and double-bond equivalence dependence. This study clearly shows the correlation between the chemical compositions determined using FT-ICR MS and the chemical and physical properties of crude oils.

Introduction

As the world's remaining deposits of crude oil become heavier and sourer, the demand to know the exact composition of crude oil is increasing, and investigations of the chemical composition of crude oil are ongoing. The combination of gas chromatography and mass spectrometry (GC-MS) has been successfully applied to the study of crude oils, and GC-MS has greatly improved our knowledge of detailed chemical compositions of crude oils. However, GC-MS has limitations. For example, polar compounds, which are important components for characterizing heavy crude oil, are not compatible with the method without chemical derivatization. Even with derivatization, the types of compounds that can be observed with GC-MS are limited and studying high-mass molecular components ($m/z > 400$) is difficult.

Recently, a new field of study, petroleomics,^{1,2} has emerged to broaden our knowledge of petroleum composition. In petroleomics, polar and heavy components of crude oil, which have been difficult to detect using conventional techniques, are studied using ultrahigh-resolution mass spectrometry, particularly Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR MS).³ FT-ICR MS is well-known for its high resolution and accurate mass measurement capabilities. Broadband FT-ICR MS spectra of petroleum are usually very complex and routinely contain thousands of peaks over a

wide dynamic range. Many studies have developed new instrumentation,⁴ ionization methods,^{5–10} and data interpretation methods^{11–14} to improve the application of FT-ICR MS to petroleomics. FT-ICR MS has been combined with ion mobility separation for structural elucidation of crude oil.¹⁵ A petroleomic approach with FT-ICR MS has been shown to be very effective for studying the chemical composition of crude oils at the molecular level.^{16–26}

- (4) Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. *Anal. Chem.* **2008**, *80*, 3985–3990.
- (5) Schaub, T. M.; Hendrickson, C. L.; Qian, K. N.; Quinn, J. P.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 2172.
- (6) Constapel, M.; Schellentrager, M.; Schmitz, O. J.; Gab, S.; Brockmann, K. J.; Giese, R.; Benter, T. *Rapid Comm. Mass Spec.* **2005**, *19*, 326–336.
- (7) Crawford, K. E.; Campbell, J. L.; Fiddler, M. N.; Duan, P.; Qian, K.; Gorbaty, M. L.; Kenttamaa, H. I. *Anal. Chem.* **2005**, *77*, 7916–7923.
- (8) Purcell, J. M.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 1265–1273.
- (9) Schmitt-Kopplin, P.; Englmann, M.; Rossello-Mora, R.; Schiewek, R.; Brockmann, K. J.; Benter, T.; Schmitz, O. J. *Anal. Biol. Chem.* **2008**, *391*, 2803–2809.
- (10) Schrader, W.; Panda, S. K.; Brockmann, K. J.; Benter, T. *Analyst* **2008**, *133*, 867–869.
- (11) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 4676–4681.
- (12) Kim, S.; Kramer, R. W.; Hatcher, P. G. *Anal. Chem.* **2003**, *75*, 5336–5344.
- (13) Barrow, M. P.; Headley, J. V.; Peru, K. M.; Derrick, P. J. *Energy Fuels* **2009**, *23*, 2592–2599.
- (14) Hur, M.; Yeo, I.; Park, E.; Kim, Y.; Yoo, J.; Kim, E.; No, M.-h.; Koh, J.; Kim, S. *Anal. Chem.* **2009**, *82*, 211–218.
- (15) Fernandez-Lima, F. A.; Becker, C.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G.; Russell, D. H. *Anal. Chem.* **2009**, *81*, 9941–9947.
- (16) Qian, K.; Rodgers, R. P.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Energy Fuels* **2001**, *15*, 492–498.
- (17) Hughey, C. A.; Rodgers, R. P.; Marshall, A. G. *Anal. Chem.* **2002**, *36*, 4145–4149.

*Corresponding author phone: 82-53-950-5333; fax: 82-53-950-6330; e-mail: sunghwank@knu.ac.kr.

(1) Marshall, A. G.; Rodgers, R. P. *Acc. Chem. Res.* **2004**, *37*, 53–59.
(2) Marshall, A. G.; Rodgers, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18090–18095.

(3) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *26*, 489–490.

Table 1. Bulk Properties of Samples Used in This Study

| oil samples | sulfur (%) | N (ppm) | TAN (mg KOH/g) | API | V (ppm) | Ni (ppm) | AR (weight %) |
|-------------|------------|--------------|----------------|-------|--------------|--------------|---------------|
| crude 01 | 0.61 | 3861 | 2.05 | 20.1 | 17.5 | 9.5 | 65.46 |
| crude 02 | 2.87 | 1645 | 0.29 | 27.6 | 56.3 | 18.5 | 50.75 |
| crude 03 | 4.5 | ^a | 3.5 | 8 | 105.6 | ^a | ^a |
| crude 04 | 0.13 | 949 | 0.58 | 22.3 | 0 | 18.4 | 45.18 |
| crude 05 | 0.09 | ^a | 0.54 | 39.57 | ^a | ^a | 29.12 |
| crude 06 | 0.08 | ^a | 0.87 | 38.61 | ^a | ^a | 31.06 |
| crude 07 | 0.08 | 2865 | 2.34 | 27.2 | 0 | 42.0 | 73.33 |
| crude 08 | 0.11 | 1884 | 4.26 | 22 | 0 | 8.3 | 78.31 |
| crude 09 | 0.2 | 3231 | 1.46 | 20.7 | 1.3 | 49.1 | 77.17 |
| crude 10 | 0.13 | 1654 | 0.25 | 32.6 | 1.7 | 2.9 | 51.11 |
| crude 11 | 4.79 | 2136 | 0.27 | 18.3 | 54.6 | 21.4 | 65.66 |
| crude 12 | 0.18 | 3532 | 0.79 | 23.1 | 3.7 | 3.9 | 68.34 |
| crude 13 | 2.71 | 1350 | 0.18 | 30.4 | 28.2 | 9.5 | 45.86 |
| crude 14 | 1.85 | 3536 | 0.25 | 19.6 | 308.7 | 128.6 | 62.61 |
| crude 15 | 2.01 | 4011 | 0.18 | 19.5 | 329.8 | 129.0 | 64.88 |
| crude 16 | 0.25 | 4405 | 3.15 | 16.1 | 0 | 12.7 | 81.17 |
| crude 17 | 3.77 | 1620 | 0.3 | 24.1 | 43.8 | 21.7 | 56.26 |
| crude 18 | 3.57 | 4019 | 0.45 | 19.3 | 93.5 | 71.0 | 64.32 |
| crude 19 | 3.53 | 3123 | 0.47 | 19.5 | 108.3 | 38.7 | 64.48 |
| crude 20 | 1.91 | 900 | 0.38 | 34 | 11.5 | 10.6 | 41.18 |

^aData not available

One fundamental goal of petroleomics is to understand and predict the properties of crude oils.² To achieve this goal, it must be determined whether there is a correlation between the spectral information obtained using FT-ICR MS and the properties of crude oils. There have been many petroleomic studies using FT-ICR MS,^{13–23} and it was assumed that the spectral findings were related to crude oil properties. However, to our knowledge, no systematic investigation of this assumption has been published.

In the current study, the validity of the assumption was evaluated by seeking a correlation between high-resolution mass spectra and the chemical and physical properties of crude oil samples. The results are graphically presented using Circos diagrams.²⁷

Experimental Section

Mass Spectrometry Analysis. Twenty crude oil samples and their bulk properties were provided by the SK Energy Corporation. Twenty oil samples with relatively high and low sulfur content were selected for this study. The properties of the crude oils are listed in Table 1. Samples were prepared by diluting the crude oils to 1 mg/mL with a 50:50 v/v solution of toluene/methanol. HPLC grade methanol and toluene were purchased from Merck (Gibbstown, NJ, USA) and were used without further purification. The prepared samples were directly injected with a syringe pump (Harvard, Holliston, MA, USA) at a flow

rate of 200 $\mu\text{L}/\text{h}$. Analyses were performed with a 15 T FT-ICR mass spectrometer at the Korean Basic Science Institute (KBSI, Ochang-eup, Korea). Positive and negative mode atmospheric pressure photo ionization (APPI) was used as an ionization method. The APPI source was obtained from Bruker Daltonics (Billerica, MA, USA). The Apex hybrid Qq-FT instrument was equipped with a Bruker Apollo II dual source. Nitrogen was used as the drying and nebulizing gas. The operating parameters for APPI analyses were set at a nebulizing temperature of 200 °C with a 3.0 L/min flow rate and a drying gas temperature of 200 °C with a 2.0 L/min flow rate. The skimmer voltage was set to 13.0 V to minimize in-source fragmentation. Ionized samples were stored in an argon-filled collision cell for 1 s and transferred to the ICR cell with a 2 ms time-of-flight window. Both sidekick and gated trapping approaches were used. A sidekick voltage of 20 V was used to initially trap the ions. After transferring the ions to the ICR trap, the trap voltage was raised to 3 V and ramped down to 1.5 V for detection. At least 100 scans were accumulated and averaged to improve the signal-to-noise ratio of the resulting spectra. For each spectrum, at least 2×10^6 data points were recorded. A resolving power greater than 300 000 at $\sim 400 \text{ m/z}$ was routinely achieved for the 40 spectra collected (20 crude oils negative and positive ion data).

Mass Calibration. Forty mass spectra collected in this study were calibrated in two steps. In the first step, an internal calibrant material (G2421A electrospray “tuning mix” from Agilent, Santa Clara, CA, USA) was used as a mass internal calibration standard. The internal standard was added to the samples just before analysis. In the second step, the samples were analyzed without the internal calibrant to eliminate possible peak contributions from the calibrant material. The m/z scale of the resulting spectra was first calibrated using an external calibrant material (G2421A electrospray “tuning mix” from Agilent) and second, using the exact m/z numbers of major sample peaks obtained from the first spectrum.

Spectral Interpretation. Spectral interpretation was performed using STORMS 1.0 (statistical tool for organic mixtures’ spectra).²⁸ An automated peak-picking algorithm was later implemented for more reliable and faster results.²⁹ Elemental formulas were calculated and assigned based on m/z values within 1 ppm error range. Normal conditions for petroleum data ($\text{C}_n\text{H}_m\text{N}_o\text{O}_p\text{S}_q$, c unlimited, h unlimited, $0 \leq n \leq 5$, $0 \leq$

- (18) Hughey, C. A.; Rodgers, R. P.; Marshall, A. G.; Walters, C. C.; Qian, K.; Mankiewicz, P. *Org. Geochem.* **2004**, 35, 863–880.
- (19) Fu, J. M.; Kim, S.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G.; Qian, K. N. *Energy Fuels* **2006**, 20, 661–667.
- (20) Klein, G. C.; Kim, S.; Rodgers, R. P.; Marshall, A. G.; Yen, A. *Energy Fuels* **2006**, 20, 1973–1979.
- (21) Headley, J. V.; Peru, K. M.; Barrow, M. P.; Derrick, P. J. *Anal. Chem.* **2007**, 79, 6222–6229.
- (22) Schaub, T. M.; Jennings, D. W.; Kim, S.; Rodgers, R. P.; Marshall, A. G. *Energy Fuels* **2007**, 21, 185–194.
- (23) Smith, D. F.; Rahimi, P.; Teclamariam, A.; Rodgers, R. P.; Marshall, A. G. *Energy Fuels* **2008**, 22, 3118–3125.
- (24) Panda, S. K.; Andersson, J. T.; Schrader, W. *Angew. Chem., Int. Ed.* **2009**, 48, 1788–1791.
- (25) Pinkston, D. S.; Duan, P.; Gallardo, V. A.; Habicht, S. C.; Tan, X. L.; Qian, K. N.; Gray, M.; Mullen, K.; Kenttamaa, H. I. *Energy Fuels* **2009**, 23, 5564–5570.
- (26) Smith, D. F.; Rodgers, R. P.; Rahimi, P.; Teclamariam, A.; Marshall, A. G. *Energy Fuels* **2009**, 23, 314–319.
- (27) Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S. J.; Marra, M. A. *Genome Res.* **2009**, 19, 1639–1645.

- (28) Hur, M.; Lim, D.; Ye, I.; Yoo, J.; Kim, Y.; Kim, S. In Statistical Interpretation of Crude Oil High Resolution Spectra by STORMS; 56th American Society for Mass Spectrometry Conference, Denver, 2008.

- (29) Hur, M.; Oh, H. B.; Kim, S. *Bull. Korean Chem. Soc.* **2009**, 30, 2665–2668.

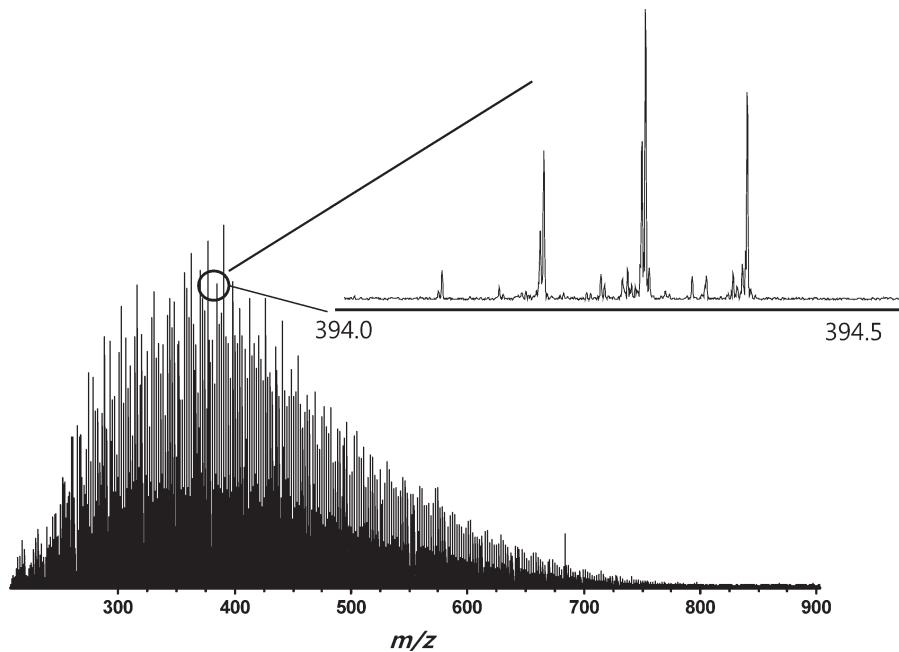


Figure 1. (+) Mode APPI spectrum of crude oil sample 3 and the expanded view of the 394.0–394.5 m/z region.

Table 2. List of Assigned Elemental Formulae and ppm Error for the Peaks in Figure 1

| observed mass | assigned formula | theoretical mass | ppm error |
|---------------|---------------------------|------------------|-----------|
| 394.0845 | $C_{26}H_{18}S_2$ | 394.0844 | 0.0 |
| 394.0878 | $C_{26}H_{18}N_3O_9$ | 394.0881 | 0.8 |
| 394.1387 | $C_{27}H_{22}OS$ | 394.1386 | -0.2 |
| 394.1625 | $C_{27}H_{24}NS$ | 394.1624 | 0.0 |
| 394.1719 | $C_{31}H_{22}$ | 394.1716 | -0.7 |
| 394.1751 | $C_{25}H_{26}S$ | 394.1750 | -0.2 |
| 394.1784 | $C_{25}H_{30}S_2$ | 394.1783 | -0.1 |
| 394.2291 | $C_{29}H_{30}O$ | 394.2291 | -0.0 |
| 394.2325 | $C_{26}H_{34}OS$ | 394.2325 | 0.0 |
| 394.2485 | $C_{28}H_{31}N_1^{13}C_1$ | 394.2485 | 0.0 |
| 394.2529 | $C_{29}H_{32}N$ | 394.2529 | 0.2 |
| 394.2564 | $C_{28}H_{32}^{13}C_2$ | 394.2566 | 0.4 |
| 394.2598 | $C_{25}H_{36}S^{13}C_2$ | 394.2599 | 0.3 |
| 394.2613 | $C_{25}H_{34}N_2O_2$ | 394.2615 | 0.4 |
| 394.2656 | $C_{30}H_{34}$ | 394.2655 | -0.2 |
| 394.2689 | $C_{27}H_{38}S$ | 394.2689 | -0.1 |
| 394.2722 | $C_{24}H_{42}S_2$ | 394.2722 | 0.1 |
| 394.2867 | $C_{27}H_{38}O_2$ | 394.2866 | -0.1 |
| 394.3105 | $C_{27}H_{40}NO$ | 394.3104 | -0.1 |
| 394.3231 | $C_{28}H_{42}O$ | 394.3230 | -0.2 |
| 394.3469 | $C_{28}H_{44}N$ | 394.3468 | -0.1 |
| 394.3504 | $C_{27}H_{44}^{13}C_2$ | 394.3505 | 0.1 |
| 394.3550 | $C_{28}H_{45}^{13}C_1$ | 394.3550 | -0.1 |

$\rho \leq 10, 0 \leq s \leq 2$ ³⁰ were used in these calculations. An example of a (+) mode APPI spectrum and its expanded region is shown in Figure 1. The elemental formulas found and assigned in the expanded region of Figure 1 are listed in Table 2. All the peaks were successfully assigned within an error range of 1 ppm.

Statistical Analysis. Prior to statistical analysis, peak lists with elemental formula assignments were merged into a single table for each ion modes' data. Further information can be found in a previous report.¹⁴ Briefly, 20 petroleum samples were analyzed in triplicate using FTICR MS with both positive and negative mode APPI. As a result, 60 mass spectra from 20 oil samples were obtained for each ionization mode. Triplicate spectra were later combined into a single spectrum, resulting in 20 mass spectra and an equal number of peak lists. The

relative abundance of individual peaks was normalized to the summed relative abundance of each peak list. Elemental formulas were assigned to each peak list using the method described above. The 20 resulting peak lists with elemental formula assignments were merged into a single table. The information in the merged table was used for statistical analysis.

Statistical computations were performed using R statistical package, version 2.6.2, and STORMS 1.0.²⁸ STORMS 1.0 was specifically developed for this study and used a R-(D)COM interface³¹ between STORMS 1.0 and the R calculation engine. Spearman's rank correlation³² was used in this study because the Spearman correlation does not require prior knowledge of a probability distribution. Further information on Spearman's rank correlation can be found in the literature.^{32,33} Briefly, the original two variables were converted into ranks. The ranks were determined based on size; variables were sorted by size and ranked according to their order. A variable with a smaller value had a lower rank, and a variable with a larger value had a higher rank. If variables were of equal size, they had an equal ranking. After ranking, a correlation analysis was performed on the ranks. In this study, the relative abundance of peaks in the spectra was the first original variable and properties of the crude oil were the second original variable. The correlation coefficient (ρ , spearman's rho) was calculated using the following equation when there were no tied ranks, and it represents the measure of correlation between two variables,³³

$$\rho = 1 - \frac{6 \sum_i (x_i - y_i)^2}{n(n^2 - 1)} \quad (1)$$

where x_i , and y_i are the ranks of relative abundance and properties of crude oil, respectively, for each set. If there were tied ranks, the following equation was used for the calculation:³³

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (2)$$

(31) Thomas, B.; Erich, N. *Comp. Stat.* **2007**, 22, 91–108.

(32) Spearman, C. *Am. J. Psychol.* **1904**, 15, 72–101.

(33) Maritz, J. S., *Distribution-Free Statistical Methods*; Chapman & Hall: New York, 1981; p 217.

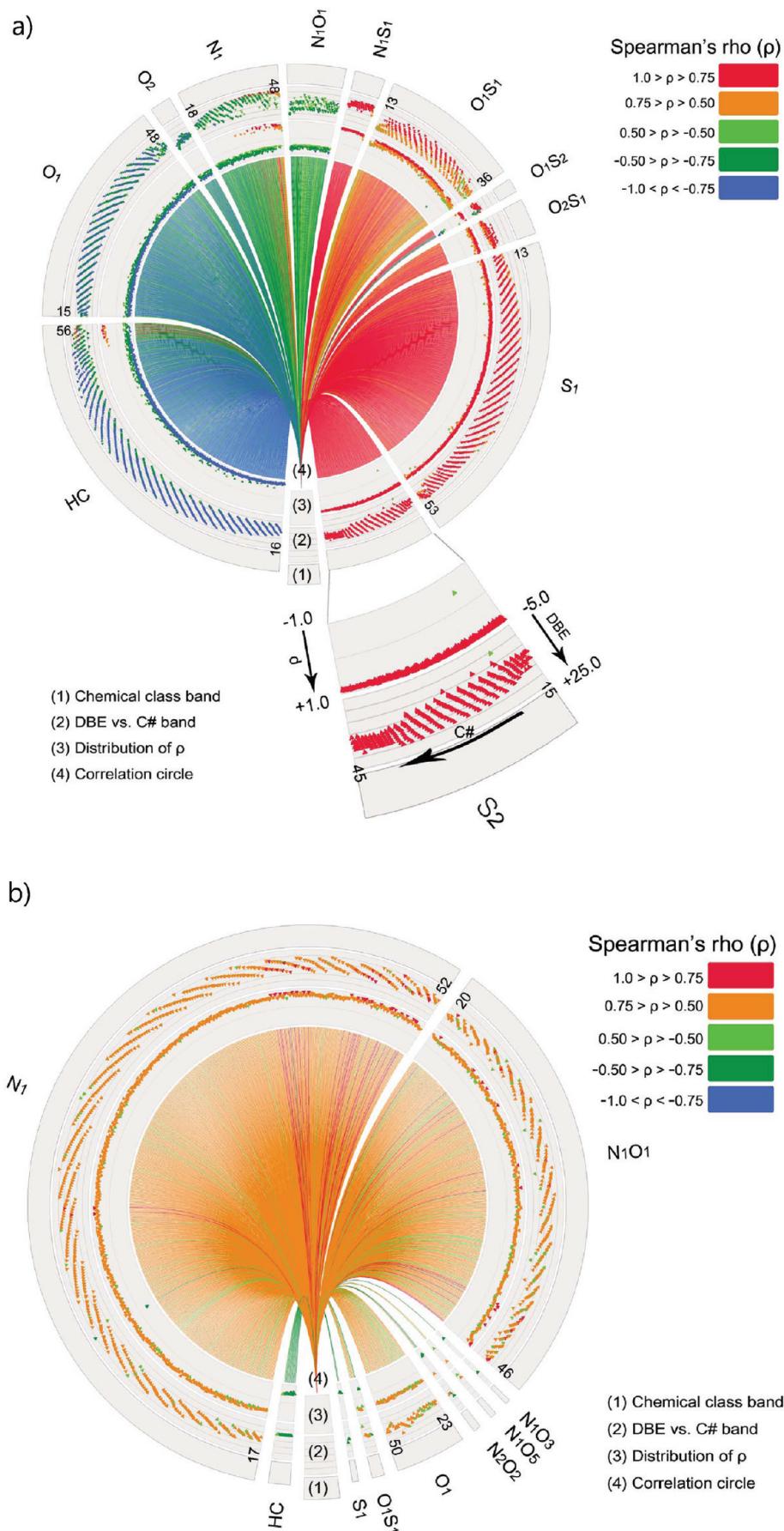


Figure 2. Circos diagrams showing correlations between oil compounds observed in (+) APPI FT-ICR mass spectra and bulk (a) sulfur and (b) nitrogen content.

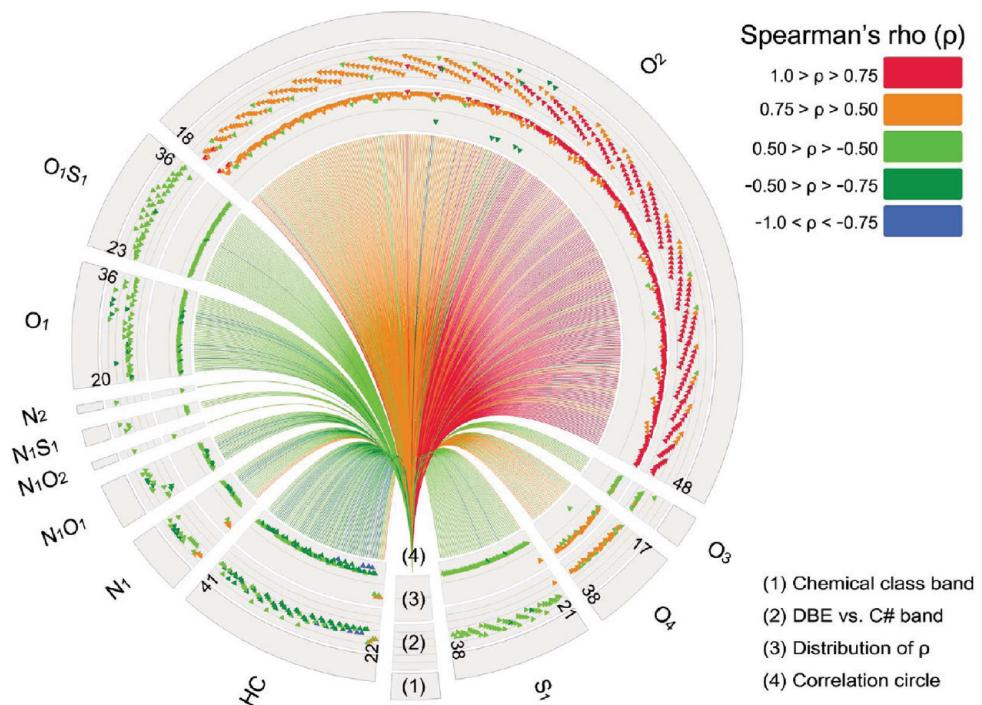


Figure 3. Correlation of oil compounds observed in (−) APPI FT-ICR mass spectra and TAN.

The sign of ρ indicates the direction of the correlational relationship between variables. The value of ρ is positive when one variable tends to increase as the other variable increases. The value of ρ is negative when one variable tends to decrease as the other variable increases.

Data Visualization Using Circos Diagrams. The Circos diagram²⁷ was originally developed for comparative genomics and is a very effective visualization tool for large amounts of data. The Circos diagram was used in this study to visualize results from statistical analysis. The Circos diagrams were downloaded from the original website (<http://mkweb.bcgsc.ca/circos/>) and were modified to accommodate correlational results. Interpretation of the diagrams is further explained in the following section.

Results and Discussion

Interpretation of the Circos Diagram. The Circos diagrams were constructed in the same manner throughout the manuscript, and Figure 2 is used as an example to explain how to interpret the diagrams. For each ionization mode, spectral information from 20 spectra was combined into a table, and the combined information was used for the correlation analysis. The p -value was calculated using the Hmisc library,³⁴ which is available in the R statistical package.³¹ Among the compounds in the table, only the ones with statistical significance (p -values < 0.05) were used to construct the Circos diagrams. The shell located farthest from the center designates heteroatom classes. The length of the shell is proportional to the number of peaks. The layer just inside the class designator shows double-bond equivalences (DBE = number of rings plus double bonds to carbon) and carbon number distribution for each class [DBE and carbon number (C#) band]. DBE values were calculated from chemical formulas using the following equation:

$$\text{DBE} = c - h/2 + n/2 + 1 \quad (3)$$

(for the ions, $\text{C}_c\text{H}_h\text{N}_n\text{O}_o\text{S}_s^+$ or $\text{C}_c\text{H}_h\text{N}_n\text{O}_o\text{S}_s^-$)

In this band, the peaks with the same C# appeared on a line. Carbon numbers for the lines increased in a clockwise direction as noted in Figure 2. The minimum and maximum C#s are marked for each class. Dots residing closer to the center have lower DBE values than outer dots. The DBE value increased according to the direction marked in the figure. For example, the lowest DBE observed for the S_1 class compounds was 0.5, and DBE values increased in the directions designated with arrows. This applies to the other classes as well.

The circle at the center is the correlation circle, which indicates the overall distribution of the correlation coefficient (ρ). The lines connecting a property with observed peaks were color-coded according to their ρ values. Red and orange indicate a strong positive correlation, light green a weak or no correlation, and green and blue show a strong negative correlation. The color code for ρ is identical to that used in other parts of the diagram. A more detailed distribution of ρ is displayed in the band just outside of the circle. In this part of the diagram, the dots residing near the correlation circle have ρ values of −1, and the dots located near the “DBE & C# Band” have ρ values of 1.

Correlation with Sulfur and Nitrogen Contents. The correlational results between peaks observed using positive mode APPI and sulfur and nitrogen contents listed in Table 1 are displayed in Figure 2, panels a and b. In Table 1, the sulfur and nitrogen contents were obtained by elemental analysis.³⁵ In Figure 2, a positive correlation (red or orange color) means that the relative abundance of a certain peak increased as bulk concentrations of sulfur or nitrogen in the samples increased. A negative correlation (blue or dark green color) denotes a decrease in relative abundance of a peak as the bulk concentration of sulfur or nitrogen decreased.

In the sulfur correlation diagram (Figure 2a), sulfur-containing classes (e.g., S, S_2 , SO, and NS) were dominant, and the correlation circle was primarily either orange or red,

(35) ASTM D7455 - 08 Standard Practice for Sample Preparation of Petroleum and Lubricant Products for Elemental Analysis; American Society for Testing and Materials (ASTM) International: West Conshohocken, PA.

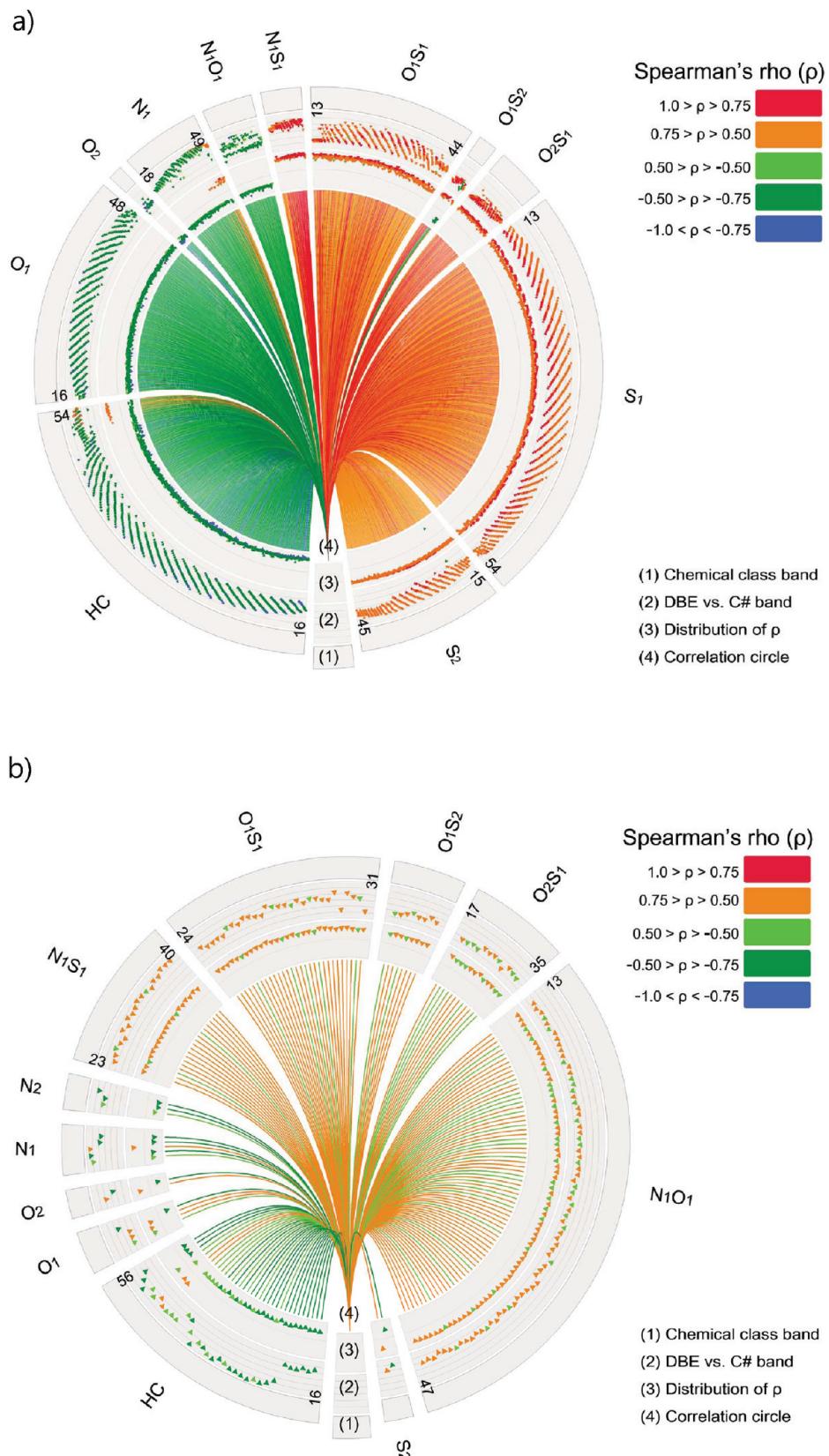
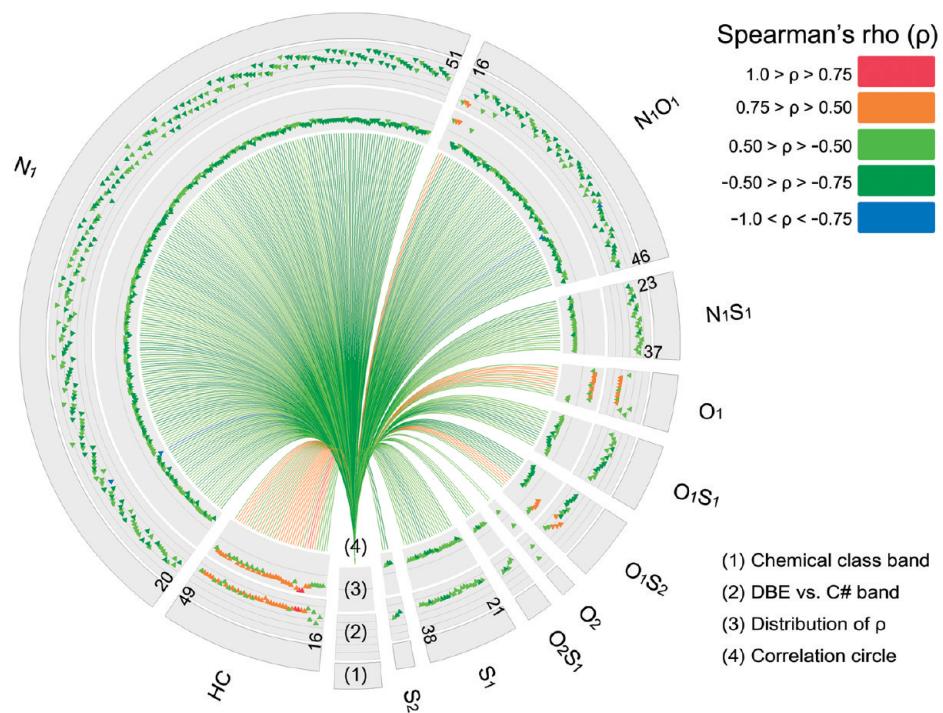


Figure 4. Correlation of oil compounds observed in (+) APPI FT-ICR mass spectra and bulk (a) vanadium and (b) nickel content.

indicating a strong positive correlation. The fact that sulfur-containing compounds showed a strong positive correlation with the sulfur content of the samples agrees very well with what is expected from crude oils. Note also that a correlation with SO class compounds showed a DBE dependence. In

the “DBE & C# Band,” the SO compounds with relatively high DBE values had a stronger correlation than those with lower DBE values. The SO compounds with lower DBE values were assigned to sulfoxide compounds, whereas the SO compounds with higher DBE values were assigned to

a)



b)

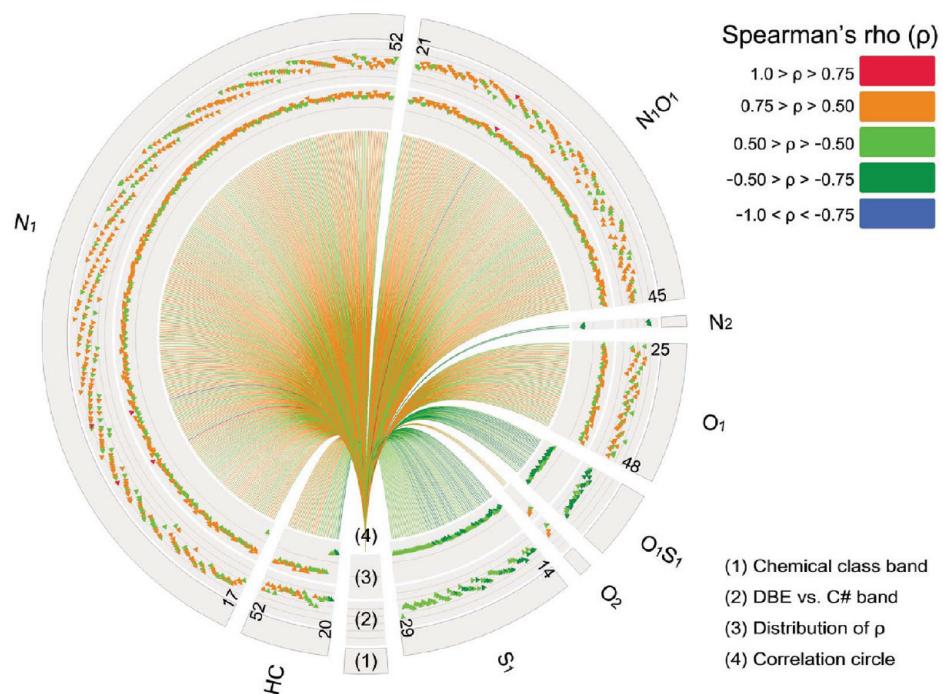


Figure 5. Correlation of oil compounds observed in (+) APPI FT-ICR mass spectra and (a) API and (b) AR.

compounds containing furanyl and thiophenyl groups.³⁶ The difference in correlation may have originated from their structural differences.

In Figure 2, the classes that did not contain sulfur had a negative correlation with sulfur content, which could have

been due to differences in ionization efficiency or ionization suppression caused by a matrix effect. More studies with synthetic standard compounds are currently being carried out in our laboratory to address this point.

Correlation with nitrogen content was investigated (Figure 2b). The nitrogen-containing classes (e.g., N₁ and N₁O₁) showed a strong positive correlation with the nitrogen

(36) Kim, Y. H.; Kim, S. J. *J. Am. Soc. Mass. Spectrom.* 2010, 21, 386–392.

Table 3. Average DBE Values for S₁ Class Compounds in Different Ranges of R Values

| correlation coefficient | $0.5 < R < 0.75$ | $0.75 < R < 1$ |
|---|-------------------|--------------------|
| average DBE (\pm standard deviation) | 4.0 (± 2.7) | 10.8 (± 3.6) |

content of the bulk samples, which agrees with the general conception that petroleum samples with high nitrogen content contain abundant nitrogen compounds. Thus, this result, combined with the result described in the above paragraph, strongly support the conclusion that information obtained from the high-resolution mass spectra correlates well with the nitrogen and sulfur contents of crude oils.

Correlation with Total Acid Number. Correlation analyses of the total acid number (TAN) for peaks observed using (−) mode APPI FT-ICR MS are presented in Figure 3. TAN values were determined by the commonly used titration method.³⁷ TAN is a commonly used parameter in the oil industry to indicate acidity of the crude oils. The number is defined by the amount of KOH (in mg) needed to neutralize 1 g of crude oil. Spectra obtained with negative mode APPI were used in this analysis because negative mode analysis is more suitable for acidic compounds. Figure 3 shows a strong correlation between TAN and O₂ and O₄ class compounds. Because typical organic acids contain one or more COOH functional groups, one may reasonably expect that the amount of O₂ and O₄ compounds would be related to the TAN of a given sample. The DBE distribution of the O₂ group correlating with the TAN ranged widely from 1 to 17. This means that both naphthenic compounds and acids with aromatic structures could contribute to the acidity of crude oils. In addition, the peaks with a higher carbon number generally correlated more strongly with the TAN. In the case of the O₄ class, the peaks correlating positively with TAN had relatively low DBE distributions, ranging from 3 to 5. The COOH functional group had a DBE value of 1, and thus, the O₄ class compounds minus two COOH functional groups would have DBE values between 1 and 3. Since an aromatic ring has a DBE value of 4, O₄ class compounds are likely to have dicarboxylic naphthenic acid structures.

Correlation with Vanadium and Nickel Contents. Figure 4, panels a and b, shows the correlations between peaks observed in (+) mode APPI mass spectra and the vanadium and the nickel contents of the oil sample. The vanadium and nickel contents were determined using the ASTM method.³⁸ In the case of vanadium, a strong correlation with sulfur-containing compounds was observed. It was apparent that the Circos diagram for vanadium content (Figure 3a) is very similar to the one for sulfur content (Figure 2a) since sulfur-containing classes (e.g., S, S₂, SO, and NS) showed a strong positive correlation in both figures. The similarity suggests that sulfur and vanadium contents may be correlated with each other. This relationship has been previously reported.³⁹ The data presented in Figure 4 clearly show that the correlation between sulfur and vanadium is also valid at the molecular level.

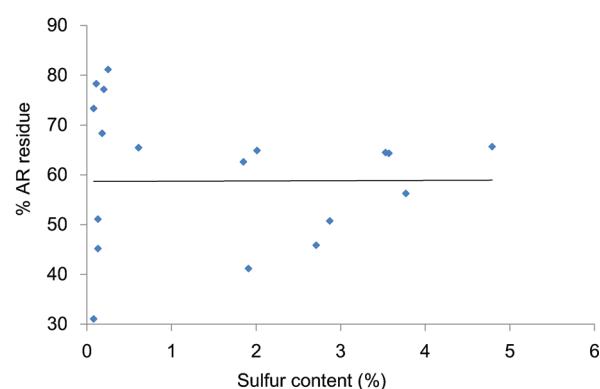
The S₁ class compounds with relatively low DBE numbers showed stronger correlations with vanadium content in the

(37) ASTM D974 - 08e1 Standard Test Method for Acid and Base Number by Color-Indicator Titration; American Society for Testing and Materials (ASTM) International: West Conshohocken, PA.

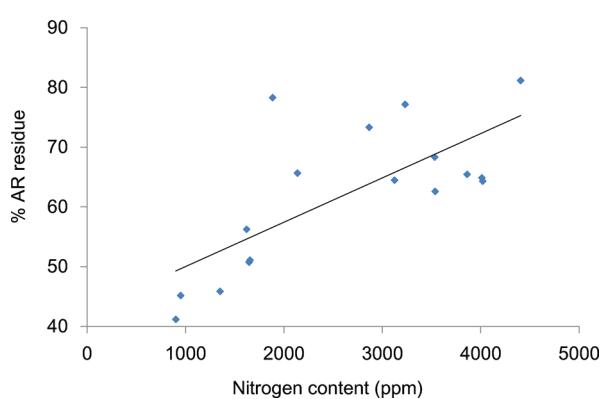
(38) ASTM D5708 - 05 Standard Test Methods for Determination of Nickel, Vanadium, and Iron in Crude Oils and Residual Fuels by Inductively Coupled Plasma (ICP) Atomic Emission Spectrometry; American Society for Testing and Materials (ASTM) International: West Conshohocken, PA.

(39) Barwise, A. J. G. Energy Fuels 1990, 4, 647–652.

a) Sulfur content vs AR



b) Nitrogen content vs AR

**Figure 6.** Plots of AR and (a) sulfur and (b) nitrogen content comparing correlational relationship between the bulk properties.

DBE vs C# Band of Figure 4a. The trend of a stronger correlation for lower DBE compounds was independent of carbon number. The average DBE values of S₁ class compounds with different ranges of Spearman's rho values are listed in Table 3. The average DBE with a ρ value between 0.75 and 1 was calculated to be 4.0, but those with ρ values between 0.5 and 0.75 had an average DBE of 10.8. This indicates that S₁ class compounds with a lower degree of aromaticity and/or unsaturation were more strongly correlated with vanadium content. The reason why the correlation exists is not fully understood and requires further study.

Fewer peaks with a significant correlation (p -value < 0.05) with nickel content were observed than those with vanadium content (compare Figure 4, panels a and b). However, the overall class distributions for vanadium and nickel were quite similar. The sulfur-containing compounds had a strong positive correlation with S₁ class compounds. S₁ class compounds with a strong positive correlation with nickel content showed DBE dependence. The average DBE of S₁ class compounds with a strong correlation was about 3, which indicates that most of the positively correlated S₁ compounds had structures similar to thiopene. The thiopene structure has a DBE value of 3.

Correlation with API and AR Values. Correlations between peaks observed in (+) mode APPI, the American Petroleum Institute (API) gravity value, and the percentage of atmospheric residue remaining after atmospheric pressure distillation processing (AR) are presented in Figure 5, panels a and b. Both API and AR are important physical parameters used to determine the economic value of crude

oils. In Figure 5, correlations with nitrogen-containing compounds, especially N_1 and N_1O_1 , were most significant. N_1 and N_1O_1 class compounds showed a negative correlation with API. This means that denser crude oils (i.e., those exhibiting smaller API) yielded more nitrogen-containing peaks. In contrast, N_1 and N_1O_1 class compounds showed a positive correlation with AR values, indicating that the oil samples with a greater amount of heavy components (i.e., those exhibiting larger AR) contained more N_1 and N_1O_1 class compounds. Both correlations with API and AR clearly showed that N_1 and N_1O_1 class compounds were closely related to the heaviness of crude oils.

The AR residue often includes a high concentration of nitrogen and/or sulfur compared to the light portion of crude oil. Figure 5b clearly shows that the AR value had a good correlational relationship with some nitrogen-containing compounds. However, the sulfur class compounds did not have a good correlation. The S_1 class had only a few statistically significant peaks, and those peaks even had a negative correlation. To determine whether the relationship between nitrogen and sulfur and AR could be extended to the bulk properties, two plots were generated (Figure 6, panels a and b). Very poor correlation was observed between sulfur

content and AR values (Figure 6a). However, nitrogen content and AR values showed a much better correlation (Figure 6b), which verifies that the correlational relationship using FT-ICR mass spectra agreed well with results predicted with commonly used parameters.

Conclusions

In conclusion, the results showed that the spectral information provided by FT-ICR MS correlated strongly with the chemical and physical properties of crude oils. The relationship was statistically analyzed and graphically presented by Circos diagrams. The relationship between mass spectrometric peaks and the composition of a given oil sample is paramount to the field of petroleomics, and this study opens the door for predicting the properties of crude oils based on petroleomics. Research is currently under way to develop this methodology.

Acknowledgment. The authors thank the Korea Basic Science Institute for providing FT-ICR MS instrument time to perform experiments. This work was supported by a Korea Science and Engineering Foundation (KOSEF) grant funded by the Korean government (MEST) (No. 2010-0017024), and in part by a grant of Korea Basic Science Institute (G30123) to S. H. Kim.