

Development and Application of a Software Tool for the Interpretation of Organic Mixtures' Spectra – Hydrogen Deuterium Exchange (STORM-HDX) to Interpret APPI HDX MS Spectra[†]

Sunghyup Lee,[‡] Yunju Cho,[‡] and Sunghwan Kim^{‡,§,*}

[‡]Department of Chemistry, Kyungpook National University, Daegu 702-701, Korea

[§]Green-Nano Materials Research Center, Daegu 702-701, Korea. *E-mail: sunghwank@knu.ac.kr

Received August 16, 2013, Accepted October 14, 2013

New software was developed for the assignment of elemental formulae based on high-resolution mass spectra and subsequent hydrogen/deuterium exchange data. Entire peaks in high-resolution mass spectra were grouped by their Kendrick mass defect values, and the weighted RMS deviations between theoretical and experimental values were used to determine elemental formulae. After this initial assignment, formulae containing deuterium atoms were sorted in order to interpret hydrogen/deuterium exchange spectra. The software was successfully applied to hydrogen/deuterium exchange spectra of resins and aromatic fractions from heavy crude oil.

Key Words : Software, High resolution, Mass spectra, Deuterium exchange

Introduction

As oil resources are depleted, interest in the extraction and refining of unconventional crude oils, such as oil sands and shale oils, has been growing. However, the efficient use of these heavy, carbon-rich oils requires characterization of their chemical compositions.^{1,2} Petroleomics is defined as the characterization of the compounds in crude oils and the identification of correlational relationships between chemical composition and the physical properties and reactivity of crude oils.^{2,3} Due to its ultrahigh resolving power and mass accuracy, Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), which can be used to determine the elemental compositions of compounds in petroleum, has become one of the main tools used in petroleomics.⁴⁻¹⁰ A single high-resolution mass spectrum obtained from a crude oil sample routinely contains over 10,000 peaks, posing an enormous analytical challenge.^{11,12} Therefore, the development of algorithms and software that can help interpret these complex data is extremely important.^{11,13-16}

Recently, a new method that performs structural analyses of nitrogen compounds in crude oils by atmospheric pressure photo ionization hydrogen/deuterium exchange mass spectrometry (APPI HDX MS) was reported.^{17,18} In this APPI HDX MS method, a crude oil sample dissolved in toluene was mixed with deuterated methanol and sprayed into an APPI source. The resulting spectra contained a mixture of ions such as radical ions, hydrogen/deuterium exchanged radical ions, and hydrogen/deuterium exchanged plus deuterium ions. The list of peaks obtained by HDX MS spectra can be used for structural elucidation of nitrogen compounds in crude oils. However, the obtained HDX MS

spectra are even more complex than high-resolution spectra of pristine crude oils. Consequently, it is difficult to interpret HDX MS spectra for structural analysis by use of existing software.¹⁴ Therefore, new algorithms and software are required for the interpretation of HDX MS spectra.

This report describes software that was designed specifically for the interpretation of mass spectra obtained following hydrogen/deuterium exchange of crude oils. An algorithm was developed to assign elemental formulae and to sort the deuterium-exchanged peaks. The program was named Software Tool for the interpretation of ORganic Mixtures' spectra – hydrogen deuterium exchange (STORM-HDX).

Experimental

STORM-HDX. STORM-HDX was written in the C programming language (LabwindowsTM/CVITM 2010). STORM-HDX reads comma separated value (CSV) files in which mass numbers and abundances are listed and separated by commas. Commercial MS software, such as MassLynx and XMASS, can export CSV-formatted data. All output of STORM-HDX can be saved as CSV files so that it can be read and processed by commercial spreadsheet software such as Microsoft Excel or Origin.

Mass Spectrometry Analysis. All deuterated solvents used in this study were purchased from Sigma-Aldrich (St. Louis, MO, USA) and used without further purification. The shale oils used in this study were produced by Fischer assay pyrolysis of oil shales from the Eocene Green River Formation and collected from outcrops of the Piceance Basin Mahogany Zone oil shale at the Anvil Points Mine (APM) near Rifle, Colorado.⁴ The shale oil was separated into saturated, aromatic, resin, and asphaltene (SARA) fractions according to previously reported procedures.⁶ The aromatic and resin fractions were dissolved in perdeuterated toluene and deute-

[†]This paper is to commemorate Professor Myung Soo Kim's honourable retirement.

rated methanol to a final concentration of 0.5 mg/mL.

A 15-T FT-ICR MS (Bruker Daltonics, Billerica, MA) was used to analyze the shale oil samples. For positive-ion mode APPI analyses, the prepared samples were directly injected with a syringe pump (Harvard, Holliston, MA) at a flow rate of 500 $\mu\text{L/h}$. Three replicate spectra were acquired with a 4-MW dataset size and obtained in continuous accumulation of selected ions (CASI) mode with 325–530 windows. 200 time-domain transients were summed to enhance the signal-to-noise ratio. A nebulizing gas temperature of 450 $^{\circ}\text{C}$ and a flow rate of 2.0 L/min were used. The drying gas temperature was 210 $^{\circ}\text{C}$ at a flow rate of 2.3 L/min. The capillary voltage was 3,600 V. The collision cell radio-frequency (RF) voltage and energy were 1,500 V and -3.0 eV, Internal calibration was performed using the radical cations of the N_1 and N_1D_1 series in the (+) mode.¹⁸

Results and Discussion

STORM-HDX Algorithm.

Overall Scheme: The overall scheme of the STORM-HDX algorithm is presented in Figure 1. A more detailed explanation of selected parts is provided in a later section of this report. Briefly, a list of m/z values and relative abundances, saved in CSV format, is fed to the program. The list of peaks is sorted and grouped by Kendrick mass defect (KMD) analysis. Elemental formulae are calculated and assigned for each of the Kendrick mass groups. All of the results of the Kendrick mass sorting and elemental formulae assignments can be saved as CSV files. These can then be subjected to statistical analyses (paired t-tests and principle component analysis analyses) and compared with HDX data.

Assigning Elemental Formulae. The flow chart in Figure

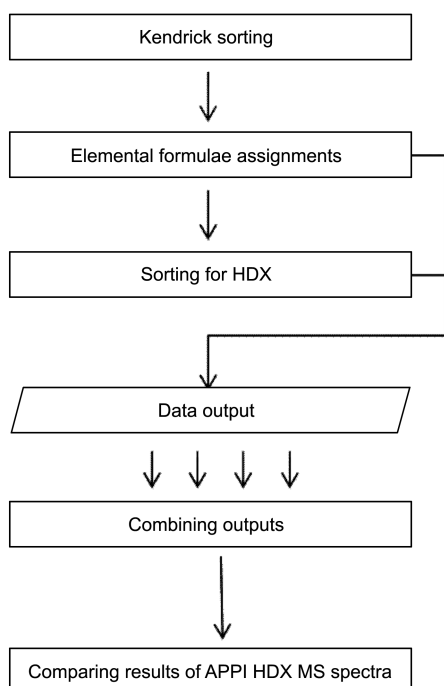


Figure 1. A flow chart of the STORM-HDX algorithm scheme.

1 shows the strategy employed for assigning elemental formulae with STORM-HDX. First, KMD values are calculated from m/z values, and peaks with the same KMD values are grouped together. After grouping, the peak with the highest relative abundance, and therefore the highest signal-to-noise ratio, in each group is chosen and possible elemental compositions for this peak are calculated within 2 ppm error. An error margin of 2 ppm was chosen instead of the more typical 1 ppm error in order to allow more candidate formulae. Usually, two candidates are calculated and assigned for each group. After assignment of the most abundant peak, the other peaks in the KMD group are assigned by adding or subtracting a series of (CH_2) groups. At this stage, all of the possible candidates are considered.

After completing the series, weighted root-mean-square (RMS) mass errors between experimental (m/z_{exp}) and theoretical (m/z_{theo}) values were calculated for each group of candidate elemental formulae using Eq. (1), where 'RA' represents relative abundance, 'n' is the total number of peaks in a group, and 'w' is a weighting coefficient.

$$\sqrt{\frac{\sum_i^n \left[RA_i^w \times \left(\frac{m/z_{\text{exp}_i} - m/z_{\text{theo}_i}}{m/z_{\text{theo}_i}} \right)^2 \right]}{n \times \sum_i^n RA_i^w}} \quad (1)$$

A weighting factor based on relative abundance was used because peaks with higher abundance were less affected by noise and therefore deemed more reliable. Typically, $w = 1$ was applied in these calculations. The calculated RMS errors were compared, and the group of candidate elemental formulae with the lowest RMS errors was chosen. When a candidate elemental formula contained a less abundant isotope, the existence of a peak with an m/z value corresponding to the same formula but with the most abundant isotope was checked. For example, when a candidate elemental composition for m/z x contained a ^{13}C atom, the existence of a peak at $(x - 1.0033)$ m/z was checked. The mass difference between ^{13}C and ^{12}C is 1.0033 amu. If this peak was not found, the elemental formula containing the less abundant isotope was discarded and the next candidate with the second lowest RMS error was chosen. This procedure was repeated until all of the KMD groups were assigned.

Sorting and Comparing HDX Data in STORM-HDX.

In previous studies, it was shown that nitrogen-containing compounds (M) with a primary amine functional group produce a combination of radical ions with two H/D exchanges ($[\text{M}-2\text{H}+2\text{D}]^{+\bullet}$) and an ion with two H/D exchanges and D^+ ($[\text{M}-2\text{H}+2\text{D}]+\text{D}^+$) ions. Compounds with a secondary amine yield a combination of a radical ion with one H/D exchange ($[\text{M}-\text{H}+\text{D}]^{+\bullet}$) and an ion with one H/D exchange and D^+ ($[\text{M}-\text{H}+\text{D}]+\text{D}^+$). In addition, compounds with a tertiary amine produce a combination of a radical ion without H/D exchanges ($[\text{M}]^{+\bullet}$ and $[\text{M}]+\text{D}^+$), and finally molecules with a pyridine functional group produce an $[\text{M}]+\text{D}^+$ ion.

These ion combinations observed in APPI HDX MS spectra can be used for structural elucidation. As explained above, high-resolution mass spectra of crude oil contain numerous

peaks, and it would be impractical to identify all of the deuterium-exchanged peaks manually. Therefore, STORM HDX searches a list of assigned peaks for peaks resulting from H/D exchange reaction and groups them together automatically. This grouped information is then tabulated and can be exported as a CSV file. In the table, the different types of exchanged ions generated from the same elemental formula are listed together so that combinations of ions can be easily identified. This summarized information enables the assignment of elemental composition. In addition, the abundances of different types of ions can be summed separately and listed by STORM HDX. This summed abundance can be used to show overall functional group distributions.

STORM HDX Demonstration. APPI HDX MS spectra were obtained from resin and aromatic fractions of the shale oil sample. The resulting spectra were analyzed by STORM HDX. As described above, the program generated a table of exchanged ions generated from the same elemental formulae. An example of such a table is provided in Table 1, which allows for easy comparison of relative abundance.

Furthermore, STORM HDX provides the ability to compare the various types of exchanged ions in different samples. To achieve this, the data in Table 1, obtained from different samples, have to be merged into a single table. An example

of a combined table is presented in Table 2.

A bar graph showing the summed abundances of different types of ions is shown in Figure 2. In the aromatic fraction (red bars in Figure 2), both $[M]^+$ and $[M-H+D]^+$ ions were the most abundant. However, $[M]+D^+$ ions were most abundant in the resin fraction. According to the results from a previous publication,¹⁸ $[M]^+$ and $[M-H+D]^+$ ions were observed from compounds with secondary and tertiary amine functional groups. Also, $[M]+D^+$ ions were mostly observed

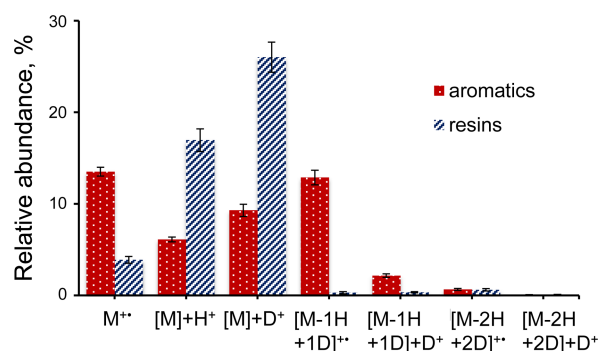


Figure 2. The bar graph shows the summed abundance of deuterium-exchanged ions observed in aromatic and resin fractions of shale oil.

Table 1. A list of peaks sorted by STORM-HDX

Formula	DBE	M**	[M]+H ⁺	[M-H+D]**	[M]+D ⁺	[M-2H+2D]**	[M-H+D]+D ⁺	[M-3H+3D]**	[M-2H+2D]+D ⁺
C ₂₆ H ₃₇ N	9	0.0319	0.0201	0.0912	0.0268	0.0000	0.0089	0.0000	0.0000
C ₂₇ H ₃₉ N	9	0.0312	0.0203	0.0983	0.0269	0.0000	0.0122	0.0000	0.0000
C ₂₈ H ₄₁ N	9	0.0346	0.0216	0.1100	0.0309	0.0000	0.0124	0.0000	0.0000
C ₂₉ H ₄₃ N	9	0.0451	0.0258	0.1359	0.0342	0.0000	0.0164	0.0000	0.0000
C ₃₀ H ₃₇ N	13	0.0462	0.0220	0.1114	0.0268	0.0037	0.0148	0.0000	0.0000
C ₃₁ H ₃₉ N	13	0.0548	0.0251	0.1202	0.0243	0.0030	0.0119	0.0000	0.0000
C ₃₂ H ₄₁ N	13	0.0576	0.0238	0.1223	0.0248	0.0031	0.0139	0.0026	0.0000
C ₃₃ H ₄₃ N	13	0.0674	0.0259	0.1267	0.0302	0.0036	0.0153	0.0000	0.0000
C ₂₇ H ₂₅ N	16	0.0429	0.0108	0.0678	0.0161	0.0000	0.0047	0.0000	0.0036
C ₂₈ H ₂₇ N	16	0.0718	0.0177	0.1432	0.0268	0.0000	0.0154	0.0000	0.0042
C ₂₉ H ₂₉ N	16	0.0470	0.0145	0.0759	0.0146	0.0000	0.0062	0.0000	0.0040
C ₃₀ H ₃₁ N	16	0.0816	0.0195	0.2093	0.0303	0.0000	0.0236	0.0000	0.0048

Table 2. Merged data for comparing H/D exchanged ions from different samples

Formula	DBE	M**	[M-H+D]**	[M]+D ⁺	[M-2H+2D]**	[MH+D]+D ⁺	[M-2H+2D]+D ⁺
C ₂₄ H ₃₃ N	9	0.0124	0.0366	0.0117	0.0075	0.0052	0.0000
	9	0.0077	0.0221	0.0293	0.0054	0.0160	0.0000
C ₂₅ H ₃₅ N	9	0.0193	0.0561	0.0163	0.0000	0.0067	0.0000
	9	0.0095	0.0315	0.0368	0.0000	0.0217	0.0000
C ₃₀ H ₃₇ N	13	0.0345	0.0831	0.0206	0.0037	0.0109	0.0000
	13	0.0134	0.0680	0.0487	0.0021	0.0415	0.0000
C ₃₁ H ₃₉ N	13	0.0360	0.0929	0.0175	0.0030	0.0086	0.0000
	13	0.0149	0.0745	0.0478	0.0000	0.0441	0.0000
C ₂₅ H ₂₃ N	15	0.0201	0.0382	0.0095	0.0000	0.0053	0.0000
	15	0.0086	0.0376	0.0243	0.0000	0.0234	0.0045
C ₂₆ H ₂₅ N	15	0.0281	0.0456	0.0108	0.0000	0.0042	0.0000
	15	0.0108	0.0415	0.0298	0.0000	0.0227	0.0059

with pyridine compounds. Based on the information provided in Figure 2, it was concluded that compounds with secondary and tertiary amines were abundant in the aromatic fraction and that pyridine type compounds were abundant in the resin fraction. The examples shown in Tables 1 and 2 and Figure 2 show how the information provided by STORM-HDX facilitates the interpretation of HDX spectra.

In summary, a new software to examine APPI hydrogen/deuterium exchange high-resolution mass spectra was developed. The software was named as STORM-HDX. For STORM-HDX, a new algorithm to assign elemental formulae was developed and applied. It was demonstrated that the software could be successfully applied to study and compare hydrogen/deuterium exchange spectra of resin and aromatic fractions from heavy crude oil.

Acknowledgments. This work was supported by NRF (National Research Foundation of Korea) Grant funded by the Korean Government (NRF-2011-Fostering Core Leaders of the Future Basic Science Program) and by grants-in-aid from the Ministry of Land, Transport and Maritime Affairs, Korea (PM56951).

References

1. Barrow, M. P. *Biofuels* **2010**, *1*, 651-655.
 2. Marshall, A. G.; Rodgers, R. P. *Proc. Natl. Acad. Sci.* **2008**, *105*.
 3. Hur, M.; Yeo, I.; Kim, E.; No, M.-h.; Koh, J.; Cho, Y. J.; Lee, J. W.; Kim, S. *Energy Fuels* **2010**, *24*, 5524-5532.
 4. Cho, Y.; Jin, J. M.; Witt, M.; Birdwell, J. E.; Na, J.-G.; Roh, N.-S.; Kim, S. *Energy Fuels* **2012**, *27*, 1830-1837.
 5. Cho, Y.; Witt, M.; Kim, Y. H.; Kim, S. *Anal. Chem.* **2012**, *84*, 8587-8594.
 6. Cho, Y.; Na, J.-G.; Nho, N.-S.; Kim, S.; Kim, S. *Energy Fuels* **2012**, *26*, 2558-2565.
 7. Bae, E.; Na, J.-G.; Chung, S. H.; Kim, H.; Kim, S. *Energy Fuels* **2010**, *24*, 2563-2569.
 8. Jin, J. M.; Kim, S.; Birdwell, J. E. *Energy Fuels* **2011**, *26*, 1054-1062.
 9. Eckert, P. A.; Roach, P. J.; Laskin, A.; Laskin, J. *Anal. Chem.* **2011**, *84*, 1517-1525.
 10. Corilo, Y. E.; Vaz, B. G.; Simas, R. C.; Lopes Nascimento, H. D.; Klitzke, C. C. F.; Pereira, R. C. L.; Bastos, W. L.; Santos Neto, E. V.; Rodgers, R. P.; Eberlin, M. N. *Anal. Chem.* **2010**, *82*, 3990-3996.
 11. Hur, M.; Yeo, I.; Park, E.; Kim, Y. H.; Yoo, J.; Kim, E.; No, M. H.; Koh, J.; Kim, S. *Anal. Chem.* **2010**, *82*, 211-218.
 12. Kim, Y. H.; Kim, S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 386-392.
 13. Kim, S.; Kramer, R. W.; Hatcher, P. G. *Anal. Chem.* **2003**, *75*, 5336-5344.
 14. Hur, M.; Oh, H. B.; Kim, S. *Bull. Korean Chem. Soc.* **2009**, *30*, 2665-2668.
 15. Hur, M.; Yeo, I.; Kim, E.; No, M. H.; Koh, J.; Cho, Y. J.; Lee, J. W.; Kim, S. *Energy Fuels* **2010**, *24*, 5524-5532.
 16. Islam, A.; Cho, Y.; Ahmed, A.; Kim, S. *Mass Spectrometry Letters* **2012**, *3*, 63-67.
 17. Ahmed, A.; Kim, S. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1900-1905.
 18. Cho, Y.; Ahmed, A.; Kim, S. *Anal. Chem.* **2013**, *85*, 9758-9763.
-