Jun Han[1]
Yi Yi[2]
Karen Lin[1]
S. Jean Birks[3]
John J. Gibson[2]
Christoph H. Borchers[1,4]

[1]University of Victoria—Genome British Columbia Proteomics Centre, Vancouver Island Technology Park, Victoria, British Columbia, Canada
[2]Alberta Innovates—Technology Futures, Vancouver Island Technology Park, Victoria, British Columbia, Canada
[3]Alberta Innovates—Technology Futures, Calgary, Alberta, Canada
[4]Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, Canada

Received May 29, 2016
Revised September 4, 2016
Accepted September 17, 2016

Research Article

Molecular profiling of naphthenic acids in technical mixtures and oil sands process-affected water using polar reversed-phase liquid chromatography–mass spectrometry

In this work, a reversed-phase ultra-HPLC (UHPLC) ultrahigh resolution MS (UHRMS) method was evaluated for the comprehensive profiling of NAs containing two O atoms in each molecule (O2NAs; general formula $C_nH_{2n+z}O_2$, where n is the number of carbon atoms and z represents hydrogen deficiency). Using a polar cyanopropyl-bonded phase column and negative-ion electrospray ionization mass spectrometric detection at 120,000 FWHM (m/z 400), 187 and 226 O2NA species were found in two naphthenic acid technical mixtures, and 424 and 198 species with molecular formulas corresponding to O2NAs were found in two oil sands process-affected water samples (one from a surface mining operation and the other from a steam-assisted gravity drainage operation), respectively. To our knowledge, these are the highest numbers of molecular compositions of O2NAs that have been profiled thus far in environmental samples. Assignments were based on accurate mass measurements (≤3 ppm) combined with rational molecular formula generation, correlation of chromatographic behavior of O2NA homologues with their elemental compositions, and confirmation with carboxyl group-specific chemical derivatization using 3-nitrophenylhydrazine. Application of this UHPLC–UHRMS method to the quantitation of O2NAs in the surface mining operation-derived water sample showed excellent linearity ($R^2 = 0.9999$) with external calibration, a linear range of 256-fold in concentration, and quantitation accuracies of 64.9 and 69.4% at two "standard substance" spiking levels.

Keywords:
Naphthenic acid / Oil sands process-affected water / Polar cyanopropyl-bonded phase column / Ultra-HPLC ultrahigh resolution mass spectrometry
DOI 10.1002/elps.201600250

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Naphthenic acids (NAs) are complex mixtures of large numbers of saturated aliphatic and polycyclic carboxylic acids, found in naturally occurring hydrocarbon deposits such as petroleum, oil sands, bitumen, and crude oils [1]. NAs with elemental compositions conforming to a general molecular formula of $C_nH_{2n+z}O_2$ are known as "classical" O_2-specific naphthenic acids (O2NAs) [2]. In this molecular formula, n is the number of carbon atoms, and z is zero or a negative even integer that represents the hydrogen deficiency resulting from formation of cyclic rings in the structures [2]. The definition of NAs has recently been expanded to include not

Correspondence: Christoph H. Borchers, University of Victoria-Genome British Columbia Proteomics Centre, #3101-4464 Markham St., Vancouver Island Technology Park, Victoria, Canada BC V8Z 7×8
E-mail: christoph@proteincentre.com
Fax: (250) 483–3238

Abbreviations: **3-NPH**, 3-nitrophenylhydrazine; **aLLE**, acidic liquid–liquid extraction; **BEH**, ethylene-bridged hybrid; **HWE**, hot-water extraction; **MW**, molecular weight; **NA**, naphthenic acids; **O2NA**, O_2-specific naphthenic acid; **OSPW**, oil sands process-affected water; **RT**, retention time; **SAGD**, steam-assisted gravity drainage; **UHRMS**, ultrahigh resolution MS; **XIC**, extracted ion chromatogram

Colour Online: See the article online to view Figs. 4 and 5 in colour.

only O2NAs, but also the unsaturated and aromatic NA-like compounds, the compounds with increased numbers of oxygen atoms in the formula (also referred as oxygenated NAs; O*x*NAs, *x* is an integer of >2), as well as the compounds that contain heterogeneous atoms such as nitrogen (N) and/or sulfur (S) [3, 4].

In northern Alberta, Canada, extractions of bitumen from the oil sands reserves lead to the storage of over 720 billion liters of oil sands process-affected water (OSPW), which contains bitumen-derived NAs and other uncharacterized organic compounds. There is a serious concern about the ecological impact of OSPW from oil sands extractions, given that the amount of OSPW stored in the tailing ponds is expected to continuously increase, while there is still no solution for safe release and reintegration of the water back to the aquatic environment [5, 6]. More specifically, among the thousands or even tens of thousands of compounds dissolved in tailing ponds, O2NAs are recognized as the major component contributing to the toxicity of OSPW [7–9]. Therefore, unambiguous determination of the molecular compositions of NAs and concentrations of individual species in OSPW and in aquatic systems would assist in evaluating their potential environmental impacts. However, the development of analytical methods to characterize and quantify individual NAs has long been a challenge. FTIR spectroscopy has been used to quantitate *total* NAs in OSPW, but overestimates the concentrations and cannot differentiate between the individual NA compounds such as O2NAs and OxNAs [10]. GC coupled to MS (GC-MS) has also been used for the analysis of the molecular compositions of some O2NAs following chemical derivatization [1, 9].

Over the past 10 years, direct infusion MS with ESI or atmospheric pressure photoionization (APPI) has been used for profiling NAs and the compounds containing N and S atoms in OSPW and other complex environmental water samples. Lo et al. [11] used low-resolution MS to characterize model O2NAs from the tailing ponds of the Athabasca oil sands, without the use of chromatographic separation prior to MS analysis. Martin et al. [12] compared the MS profile of O2NAs from low- and high-resolution MS instruments. Direct infusion ultrahigh resolution MS with ESI or APPI, in combination with front-end LC fractionation [13] or without LC fractionation [14], has been used by some researchers for the profiling of the organic components in OSPW. However, even with ultrahigh resolution MS (UHRMS), the comprehensive and quantitative analysis of O2NAs in OSPW without front-end chromatographic separation is difficult to achieve because of the complexity of the organic components in OSPW, severe ionization suppression in ESI due to the matrix, and the lack of separation of isomeric compounds.

LC coupled to ESI MS has become the preferred analytical technique for qualitative and quantitative analysis of O2NAs in OSPW and surface waters over the past years because of its potential to resolve individual O2NAs and other organic compounds and the reduced matrix effects during ionization. Smith and Rowland [15] synthesized amide derivatives of some O2NAs and used LC-(+)ESI-MS/MS to characterize the structures of the O2NA molecules. Hindle et al. [16] reported an HPLC-QTOF MS method for quantitation of several short-chain NAs ($n = 2$–4) in water samples on a C_{18} LC column. Ross et al. [17] used HPLC-QTOFMS with a Cosmosil C_{18}-MS column for the selective analysis of O2NAs and performed qualitative and quantitative analysis of some O2NAs in surface and pore waters in the lower Athabasca River region. More recently, Huang et al. reported the use of ultra-HPLC (UHPLC) with ion mobility spectroscopy QTOFMS for characterization of some O2NAs in OSPW [18]. Most of these LC-ESI/MS methods used RP columns packed with C_8- or C_{18}-bonded phases. A major disadvantage of these commonly used RP columns is that hydrophobic and high molecular weight (MW) O2NAs usually produce broad and tailing peaks due to the strong interactions between the nonpolar aliphatic or cyclic part of the O2NA molecules and the nonpolar covalently bonded stationary phases of the LC columns [19, 20]. Chemical derivatization with oxalyl chloride [15] and N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide (EDC), combined with LC-ESI/MS detection, has been also used for the analysis of model O2NA compounds [21].

Routine quantitation of O2NAs in OSPW and other environmental samples is still challenging, however, due not only to some technical limitations of direct infusion MS and LC-MS mentioned above, and the complicated compositions of the organic components in OSPW, but also because of the limitations of commercially available NA technical mixtures that are often used as the standard substances in quantitative analysis. The use of appropriate NA standard substances that match the varying O2NA components in OSPW is a key factor for reliable quantitation of O2NAs. New analytical methods are also needed for the accurate and comprehensive analysis of O2NAs, including the determination of their molecular compositions and concentrations, in OSPW.

In this study, we demonstrate a reversed-phase UHPLC–UHRMS method using a polar cyanopropyl-bonded phase for the chromatographic separation of O2NAs in two commercial NA technical mixtures and two representative OSPW samples, one produced from the surface mining operation using the Clark caustic hot-water extraction (HWE) process, and the other from the in situ mining operation utilizing the steam-assisted gravity drainage (SAGD) process. By combining two different sample preparation approaches for extraction of O2NAs from the two different OSPW samples and the use of a polar RP–UHPLC–UHRMS method, the O2NA profiles of the NA technical mixtures and the OSPW samples were determined. The UHPLC–UHRMS method was also used for the preliminary quantitative analysis of O2NAs in the HWE-derived OSPW sample.

2 Materials and methods

2.1 Chemicals and reagents

LC-MS grade acetonitrile (ACN), methanol, ethyl acetate, water, a 25% ammonia hydroxide solution, formic acid,

and HPLC grade phosphoric acid (≥85%), and pyridine were purchased from Sigma-Aldrich (St Louis, MO, USA). 3-Nitrophenylhydrazine (3-NPH)•HCl, EDC·HCl, and palmitic acid were also obtained from Sigma-Aldrich. Oasis mixed-mode anion exchange (MAX)/RP-SPE cartridges (150 mg sorbent per cartridge, 6 mL, and 30 μm particle size) were obtained from Waters (Milford, MA, USA). A Merichem (Houston, TX, USA) refined NA technical mixture provided by Environment Canada as part of a 2014 Naphthenic Acid Inter-Laboratory Study (NAILS)—a similar study to the one described in [4], and a Sigma-Aldrich NA technical mixture (cat no. 70340; 250 mL) were used in this study as commercially available technical mixtures. Two OSPW samples were collected by Alberta Innovates-Technology Futures from two different operators in northern Alberta, Canada—one from HWE processing (OSPW-1) and the other produced by SAGD (OSPW-2). Both of the water samples were analyzed to determine the O2NA compositions. The two NA technical mixtures and the two OSPW samples were stored at 4°C until used.

2.2 Standard solution

Solutions (0.5 mg/mL each) of the two NA technical mixtures were prepared in 75% aqueous methanol and were used for determining the molecular compositions of their O2NAs. The Merichem NA technical mixture was used as the "standard substance" to make standard solutions in 75% aqueous methanol for generating a calibration curve with nominal concentrations ranging from 0.010 to 1000 μg/mL.

2.3 Sample preparation

2.3.1 Acidic liquid–liquid extraction (aLLE)

aLLE extraction was performed using a similar approach to that described in [22], with ethyl acetate as the organic solvent. One milliliter aliquots of each of the two OSPW samples were mixed with 40 μL of phosphoric acid and 3 mL of ethyl acetate in 5 mL borosilicate glass test tubes. The tubes were capped and vortex mixed for 60 s followed by centrifugation at 3000 × *g* for 10 min in a Beckman X-22R centrifuge. The upper organic phases were carefully removed using 1 mL pipette tips without agitating the lower aqueous phases, and were transferred to 15 mL borosilicate glass test tubes. The aqueous phases were then extracted with 3 mL of ethyl acetate two more times. After vortex mixing and centrifugation, the organic phases of each extraction for a same sample were pooled and then dried under a gentle nitrogen gas flow at 30°C in a 24-position N-EVAP nitrogen evaporator (Organomation Associates, Berlin, MA, USA). The residues were dissolved in 0.5 mL of 75% aqueous methanol; 10 μL aliquots were injected for LC-MS.

2.3.2 Mixed-mode SPE

One milliliter aliquots of the OSPW samples were mixed with 40 μL of phosphoric acid and the mixtures were loaded onto the MAX/RP-SPE cartridges that had been activated with 3 mL of isopropanol and reconditioned with 6 mL of 5% ammonium hydroxide. Under a 5 mm Hg negative pressure, the cartridges were washed with 3 mL × 3 of 5% ammonium hydroxide solution followed by elution with 3 mL × 3 of isopropanol containing 5% ammonium hydroxide. The eluents were dried down under a gentle flow of nitrogen gas at 30°C. The dried residues were dissolved in 0.5 mL of 75% aqueous methanol. Ten microliters aliquots were injected for LC-MS.

2.3.3 ACN precipitation

One milliliter aliquots of the OSPW samples were mixed with 9 mL of ACN in 15 mL borosilicate glass test tubes. After vortex mixing for 60 s, the tubes were centrifuged at 3000 × *g* for 10 min. Five milliliters aliquots of the clear supernatants were carefully transferred to 9 mL borosilicate glass test tubes and then dried down under a gentle nitrogen gas flow at 30°C. The dried residues were dissolved in 0.25 mL of 75% aqueous methanol. Ten microliters aliquots were injected for LC-MS.

2.4 Column comparison experiment

Four UHPLC columns packed with four different bonded phases (C_{18}, C_8, C_4, and cyanopropyl [CN]) were compared to determine the chromatographic behavior of different homologues of O2NAs detected in both commercially available NA technical mixtures and the two OSPW samples involved in this study. The UHPLC columns were a high-strength silica CN column (2.1 × 100 mm, 1.8 μm, 100 Å), an ethylene-bridged hybrid (BEH) C_{18} column (2.1 × 50 mm, 1.7 μm, 130 Å), a BEH C_8 column (2.1 × 50 mm, 1.7 μm, 130 Å), and a BEH C_4 column (2.1 × 50 mm, 1.7 μm, 300 Å). For a parallel comparison, the same mobile phase (i.e., 0.01% formic acid in water as solvent A and 0.01% formic acid in ACN as solvent B) was used at the same flow rate (0.3 mL/min) and at the same column temperature (40°C) for binary solvent gradient elution. The elution gradient was 15–100% of solvent B in 15 min and 100% B for 2 min before column equilibration at 15% for each of the columns. Other LC and MS operating parameters for the comparison were kept the same. The parameters of these columns and the corresponding MS operating conditions are tabulated in Supporting Information Table S1.

2.5 LC-MS and MS/MS

A Waters Acquity UPLC system (Milford, MA, USA) coupled to an LTQ-Orbitrap Fusion MS (Thermo Scientific, CA, USA) was used. After the preliminary column comparison

experiment (Section 2.4), the high-strength silica CN column was used for the molecular profiling and quantitative analysis of O2NAs, with an optimized solvent gradient consisting of 15–90% B in 15 min followed by 2-min column wash at 100% B and 3-min column re-equilibrium at 15% B. The column flow was 0.3 mL/min and the column temperature was maintained at 40°C for all LC-MS runs.

The MS instrument was equipped with an atmospheric pressure ESI source and was operated in the negative-ion survey scan mode with FTMS detection and a mass resolution of 120 000 FWHM (at m/z 400). The scan range was m/z 80–1200, and all of the LC-MS data were recorded in the "profile" mode. The automatic gain control target was set at 5×10^4 ion counts and the maximum microinjection time was 100 ms for the Orbitrap. A background ion at m/z 112.98563 (the $[2M + Na\text{-}2H]^-$ ion of formic acid) was used in the survey scans as the reference mass to provide real-time internal mass calibration so as to ensure mass accuracy throughout all LC-MS runs. These LC-MS operating parameters are included in Supporting Information Table S1.

LC- MS/MS was performed to analyze the 3-NPH derivatives of the two NA technical mixtures, the two OSPW samples, and palmitic acid (a representative O2NA compound) following 3-NPH derivatization to see if any useful information about the chemical structures of individual O2NA compounds could be obtained. For LC-MS/MS, the survey scan was carried out within a mass range of m/z 80 to 1200 at a mass resolution of 30 000 FWHM (at m/z 400) with FTMS detection. The top five most intensive peaks were then picked up from the survey scan for MS/MS scans by CID in the linear ion trap. The precursor ion isolation width was 1 Da and the normalized collision energy was 30%.

2.6 Chemical derivatization

Derivatization of the carboxylic compounds in both the technical mixtures and the OSPW extracts was carried out using 3-NPH as the derivatizing reagent according to a protocol for the derivatization of short-chain fatty acids that has been described elsewhere [23]. The reaction follows a mechanism of nucleophilic addition/elimination as depicted in Supporting Information Fig. S1A. Briefly, 100 μL of 1 mg/mL of the NA technical mixtures or the OSPW extracts was mixed with 50 μL of 200 mM of 3-NPH•HCl in 75% aqueous methanol and 50 μL of 150-mM EDC·HCl in methanol–water–pyridine (70.5:23.5:6, v/v/v). The mixtures were allowed to react at 30°C for 30 min. After reaction, the mixture was mixed with 200 μL of 75% aqueous methanol and 10 μL aliquots were injected into the LC-MS instrument for analysis.

2.7 Assignment of O2NAs

The assignment of the O2NAs was performed in three steps. First, the theoretical m/z values of $(M\text{-}H)^-$ of O2NAs were calculated according to the general formula $C_nH_{2n+z}O_2$. These were used to extract the ion current chromatograms of putative O2NAs from the recorded UHPLC–UHRMS profiles using *QualiBrower* of the Thermo *Xcaliber* software, with an allowable mass error of ±3 ppm. Only peaks with ion intensities of $\geq 1 \times 10^3$ on their extracted ion chromatograms (XICs) were included for further processing. For each extracted peak, its measured m/z value was averaged from the scans across the entire XIC peak. The averaged m/z's of individual peaks were then recorded and exported to the "*MolWeightToFormula*" utility of the *Compass* 1.3 software (Bruker Daltonics, Bremen, Germany) to generate their chemically rational elemental compositions. During the molecular formula generation, unlimited number of C, H, and O atoms and a maximum of six N atoms were considered, and the allowable mass error was 3 ppm (the experimentally determined mass accuracy on the instrument during LC-MS runs). The charge state was set as −1. In addition, only the even electron configuration for the ESI-generated ions and the C/H ratios of between 0.2 and 2 were considered. No limit was applied to the total number of cyclic rings plus double bonds. Only those compounds whose generated molecular formulas included $C_nH_{2n+z}O_2$ were recorded, along with their retention times (RTs) and peak areas. For all the detected O2NAs, only those elemental compositions with the same z value in the $C_nH_{2n+z}O_2$ formula (i.e., homologues) and which contained at least three consecutive carbon atom numbers were retained. Next, quadratic regression analyses were performed on the chromatographic RTs (min) and the carbon atom numbers (n) for each of the assigned O2NA homologues. The correlation efficient (R^2) cutoff of the regressions was ≥0.990—that is, any detected compound would not be counted if inclusion of the compound's RT and its carbon atom number (n) for the regression induced an R^2 value of <0.990. Finally, for the assigned O2NA compounds, the XICs using the m/z values of the $(M\text{-}H)^-$ ions for both of the underivatized form and the 3-NPH derivative (a mass increment of 135.04326 Da) were checked in the LC-MS profiles recorded from the solutions at a same concentration with and without chemical derivatization. The XIC peak for any *underivatized* O2NA should not appear in the LC-MS profiles of the samples analyzed *after* 3-NPH derivatization. Any compounds that did not pass the above three steps were excluded from the list of assigned O2NAs.

2.8 Quantitation of O2NAs

The sensitivity of the UHPLC–UHRMS method for analysis of O2NAs was determined from the lower LOQ and the upper LOQ, defined as the lowest and the highest concentrations of the linear calibration curves with external calibration using $R^2 \geq 0.999$ as the linearity cutoff value. Calibration was done using linear regression on the total peak areas of all of the assigned O2NAs (measured from each of the working standard solutions made with the Merichem NA technical mixture) versus their nominal total O2NA concentrations.

Quantitation accuracy was estimated from "standard substance" spiking tests. Briefly, 1 mL aliquots of OSPW-1 were spiked with the Merichem technical mixture at two spiking levels, equivalent to 20 and 40 mg/L. "Total O2NAs" was then determined by ACN precipitation as described in Section 2.3.3, followed by UHPLC–UHRMS quantitation with external calibration. The percent recoveries were calculated as [(the measured total concentration minus the endogenous concentration)/the spiked-in amount] × 100. Triplicate samples for each spiking level were analyzed. During quantitative data processing, if doublet peaks for a single O2NA elemental composition were detected (resulting from structural isomers), and the assignment of the compounds as O2NAs passed the three-step criteria as described in Section 2.6, then the total peak area for this particular O2NA formula was used.

2.9 Data analysis

The contour plots of the O2NA molecular compositions were generated with *SigmaPlot* 12.0 (Systat Software, San Jose, CA, USA) after normalization of the total peak areas for all of the detected O2NAs from each of the two technical mixtures and each of the two OSPW samples to the same value. Other data processing and statistics was performed within Microsoft Excel.

3 Results and discussion

3.1 Polar CN-bonded phase for UHPLC of O2NA separation

Figure 1A shows the XICs of four homologous O2NA compounds ($z = -16$; $n = 25$–28), detected in the Merichem technical mixture. The chromatographic behavior of O2NAs detected in this mixture was compared on four UHPLC columns packed with four different bonded phases. As expected, the O2NA compounds showed the strongest LC retention and the broadest peak shapes (2–2.5 min in peak width) on the 5-cm long C_{18} UHPLC column. They showed weaker retention and narrower peak shapes on the medium-chain C_8 column with a pore size of 130 Å and the short-chain C_4 column with a wide pore size (300 Å), which was also as expected. The sharpest peak shapes (30–35 s in peak width) were observed from the CN column, even though it had twice the column length (10 cm) of the other three columns (5-cm long) and the smallest pore size (100 Å).

For both of NA technical mixtures ($n = 2$) and OSPW samples ($n = 2$) involved in this study, the CN column gave not only the sharpest peaks but also a narrow RT elution window (2.2 to 13.6 min) for all of the detected O2NAs. The use of the CN column had two additional advantages: first, all of the detected O2NAs were found to be eluted at ≤90% ACN on the 10-cm long column. In contrast, many higher-MW O2NAs not only showed much broader peak shapes (up to 4 min in peak width) on the short 5-cm C_{18} column but also some were not even eluted at 100% ACN, though these compounds did elute from the C_4 column at <100% ACN, due to the reduced non-polar interactions of these O2NAs with the C_4 acyl chain and the reduced resistance to liquid mass transfer on the surface of the wide-pore (300 Å) stationary phase.

Figure 1B shows the chromatographic behavior of four homologous O2NA compounds ($z = -4$, $n = 28$ to 32) detected in OSPW-1 on the four different columns. These compounds were not eluted from the C_{18} column, even at 100% ACN. It was also observed that, when the ACN was at >90%, the ESI-MS detection sensitivity for O2NAs dramatically decreased. For this reason, the four compounds eluted from the C_8 column as four narrow but weak peaks during the column wash step at 100% ACN. Thus, the cyanopropyl (CN)-bonded phase UHPLC column provided an improved LC/ESI-MS method by overcoming these technical limitations and enabling more comprehensive profiling of O2NAs in OSPW.

O2NAs are complicated mixtures containing a large collection of organic acids that have both a non-polar acyl chain ($z = 0$) or polycyclic rings ($z = -2, -4, -6 \ldots$) and a polar carboxyl group in their heterogeneous structures. CN columns can be used in both RP and normal-phase LC modes due to the dipolar nature of the bonded phase. When used in RPLC mode, the nitrile group of the bonded phase strongly interacts with the polar moiety of dipolar compounds, while the propyl ligand provides moderate hydrophobic interactions with the non-polar moiety [24]. CN columns are often used for fast RPLC separations of mixtures of different compounds with a broad range of polarities [24]. In this study, the CN column provided an alternative selectivity of separation to the more commonly used RPLC columns, and was better than the other three UHPLC columns in this study for molecular profiling of O2NAs, though we expect its ability to resolve isomeric compounds may be weaker than the C_{18} or C_8 columns.

3.2 Assignment of O2NAs and determination of the molecular compositions of O2NAs detected in the NA technical mixtures

With the use of the CN column and the LTQ-Orbitrap MS instrument with FTMS detection at a high mass resolution of 120 000 HWFM (at m/z 400), we first analyzed the two NA technical mixtures to profile the O2NAs under an elution gradient optimized for the CN column (15% to 90% B in 15 min). Figure 2A shows the UHPLC–UHRMS base peak chromatograms for the two technical mixtures. Using the polar CN column and the UHPLC system for chromatographic separation, the same O2NA compounds detected in the two NA technical mixtures showed very consistent retention between different LC-MS runs. Supporting Information Fig. S1A and B display the RTs of eight O2NA compounds detected in the two technical mixtures. These two technical mixtures displayed significantly different profiles of organic compounds.

We followed the three steps, as described in Section 2.6, to assign the detected O2NAs. In RPLC, the RTs of

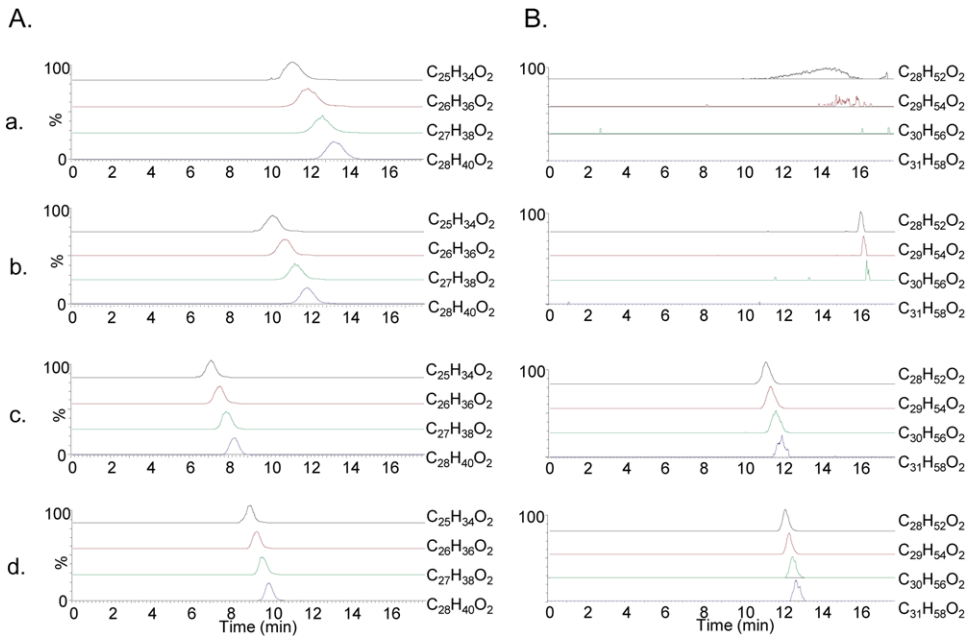

Figure 1. Comparison of chromatographic behavior of O2NAs on four UHPLC columns with different bonded phases. The four columns are (**a**) BEH C_{18} (2.1×50 mm, 1.8 μm, 130 Å); (**b**) BEH C_8 (2.1×50 mm, 1.7 μm, 130 Å); (**c**) BEH C_4 (2.1×50 mm, 1.7 μm, 300 Å) and (**d**) high-strength silica CN (2.1×100 mm, 1.8 μm, 100 Å). (**A**) Homologous O2NAs (z = −16 and n = 25 to 28 in $C_nH_{2n+z}O_2$) detected in the Merichem NA technical mixture. (**B**) Homologous O2NAs (z = −4 and n = 28 to 31 in $C_nH_{2n+Z}O_2$) detected in OSPW-1.

homologous O2NAs increased with increasing n, while the RTs of O2NAs with the same n decreased with an increase in the absolute z value [25]. Figure 2B shows the quadratic regression lines for RT vs. carbon number for four O2NA homologues (z = −4, −8, −12, and −16) detected in the Merichem technical mixture. Good correlations were observed for all of these homologues with R^2 values between 0.9950 and 0.9986. Figure 2C shows good correlations ($R^2 = 0.9934$ for n = 12 to 0.9985 for $n = 24$) between z and RT for detected O2NAs with the same n for each group. Compounds were excluded from the list of assigned O2NAs if inclusion of any of them resulted in R^2 values of <0.990. This avoided some mis-assignments based solely on the measured masses for molecular generation. In this way, many non-O2NA compounds detected in the OSPW samples and shown with their XIC peaks based on the m/z's ± 3 ppm were excluded from the assigned O2NA list. The close correlations of n and z versus RT provided an additional level of confidence in the assignments of O2NAs, particularly for those high-MW species (m/z >500) in different homologues of O2NAs which were detected in the OSPW samples in this work. For some higher-MW O2NAs, the measured m/z values generated more than one molecular formula. Confident assignment of these compounds to O2NAs was achieved by extrapolating the elemental compositions of the homologous low-MW species, which were assigned with unique molecular formulas. The relationships between z and n versus RT aided in the assignments.

Chemical derivatization using 3-NPH as the derivatizing reagent was then used as a final step to confirm the existence of carboxyl groups in the assigned O2NA compounds. The derivatization was conducted at 30 °C for 30 min. The reaction follows a mechanism of nucleophilic addition/elimination as depicted in Supporting Information Fig. S2A. Figure 3 shows the XICs of eight homologous O2NA compounds (z = −16, n = 20 to 27) from the UHPLC–UHRMS of two aliquots of a same solution prepared from OSPW-1 before and after 3-NPH derivatization. As shown, all the peaks representing the underivatized O2NAs disappeared after the derivatization, except two very weak peaks for $C_{26}H_{36}O_2$ and $C_{27}O_{38}O_2$, probably because of the incomplete derivatization. Moreover, all of the peaks corresponding to the 3-NPH derivatives of individual O2NAs (i.e., having a mass increment of 135.0433 Da due to the derivatization with one molecule of 3-NPH) appeared in the XICs (Fig. 3C). These peaks were not detectable in the solutions before chemical derivatization. It was noted that the 3-NPH derivatives of the assigned O2NAs showed broader peaks than their underivatized molecules in Fig. 3, which resulted from the strong polar-polar interactions between the polar CN bonded phase and the polar 3-NPH moiety of the derivatives. Comparison of the XICs extracted with the m/z values of underivatized O2NAs before and after 3-NPH derivatization strongly indicated that most of the assigned O2NAs were authentic carboxyl group-containing compounds. However, there were also some compounds detected in the two OSWP samples, which passed the first two steps of the assignments but did not react with 3-NPH, indicating that there was no carboxyl group in the compounds. These compounds were removed from the assigned O2NA lists. Supporting Information Fig. S2B and C show the CID mass spectra of the 3-NPH derivatives of two O2NA compounds, one acquired from the standard substance of an aliphatic long-chain fatty acid — palmitic acid ($C_{16}H_{32}O_2$), and the other from an assigned O2NA compound ($C_{20}H_{24}O_2$, z = −16) detected in OSPW-1. As shown, MS/MS by CID generated common fragment ions at *m/z* 122, 137, 152, and 178 from the two derivatives. These fragment ions were from CID fragmentation of the 3-NPH moiety of the derivatives, as the fragmentation mechanism being depicted in Supporting

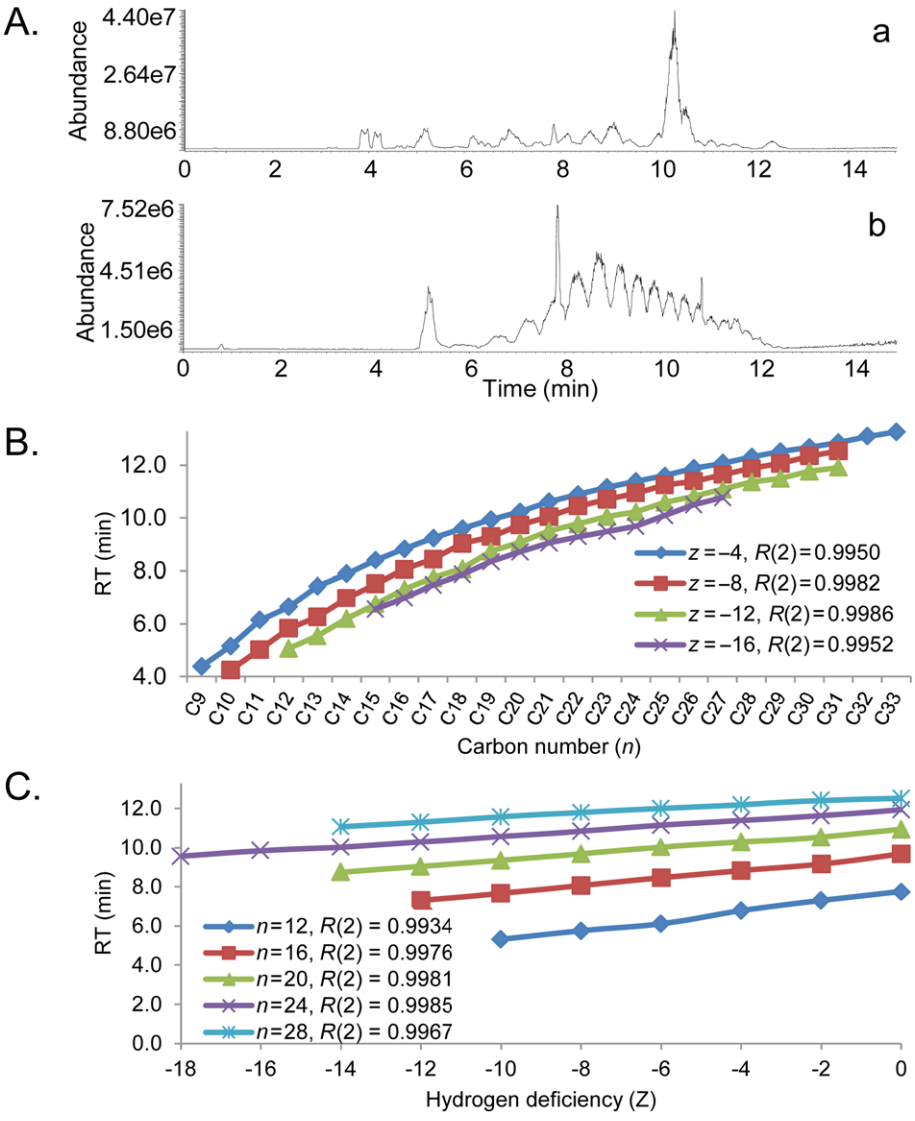

Figure 2. **(A)** Base peak chromatograms of organic compounds detected by UHPLC–UHRMS in the Sigma (**a**) and Merichem (**b**) NA technical mixtures and the relationships of chromatographic retention time (RT, min) of O2NAs with n (**B**) and z (**C**) of $C_nH_{2n+z}O_2$ using quadratic regression.

Information Fig. S2D. However, no O2NA-specific structural information can be deduced from the acquired MS/MS spectra. In the CID mass spectrum from the assigned O2NA compound ($C_{20}H_{24}O_2$) detected in OSPW-1, there were also some additional fragment ions which we attributed to the CID fragmentation of isobaric or nearly-isobaric compounds which co-eluted with this O2NA compound ($C_{20}H_{24}O_2$). The $(M\text{-}H)^-$ precursor ions of these co-eluted compounds could not be separated from that of the target compound because of the 1-Da width of the precursor ion isolation window.

After data processing, 187 and 226 molecular formulas conforming to $C_nH_{2n+z}O_2$ were determined from UHPLC–UHRMS of the Sigma and Merichem NA technical mixtures, respectively. Figure 4A and B illustrates the molecular compositions of O2NAs detected in these two mixtures. The O2NAs detected in the Sigma technical mixture are characterized by hydrogen deficiency (z) between 0 and −16 and carbon numbers ranging from n = 5 to 38. Based on the detected peak areas, the most abundant single O2NA species in this technical mixture was $C_{19}H_{38}O_2$, a saturated fatty acid with an odd number of carbon atoms. This observation is in agreement with a previous study in which GC-MS and infusion-ESI/MS were used for profiling O2NAs in two NA technical mixtures including the one from Sigma [26]. For the Merichem NA technical mixture analyzed in this work, the hydrogen deficiency ranged from $z = 0$ to -24 and the n ranged from 5 to 38 for $z = 0$ and from 23 to 26 for $z = -24$. In addition to the different numbers of the molecular formulas of O2NAs, the two technical mixtures had significantly different distributions of MWs and their relative abundances. Based on peak areas, the most abundant class of homologous O2NAs detected in the Sigma technical mixture were saturated fatty acids (z = 0), followed by those with $z = -2$, -4, and -6. In contrast, the most abundant homologous O2NAs detected in the Merichem NA technical mixture were the species with $z = -4$, followed by those with $z = -6$

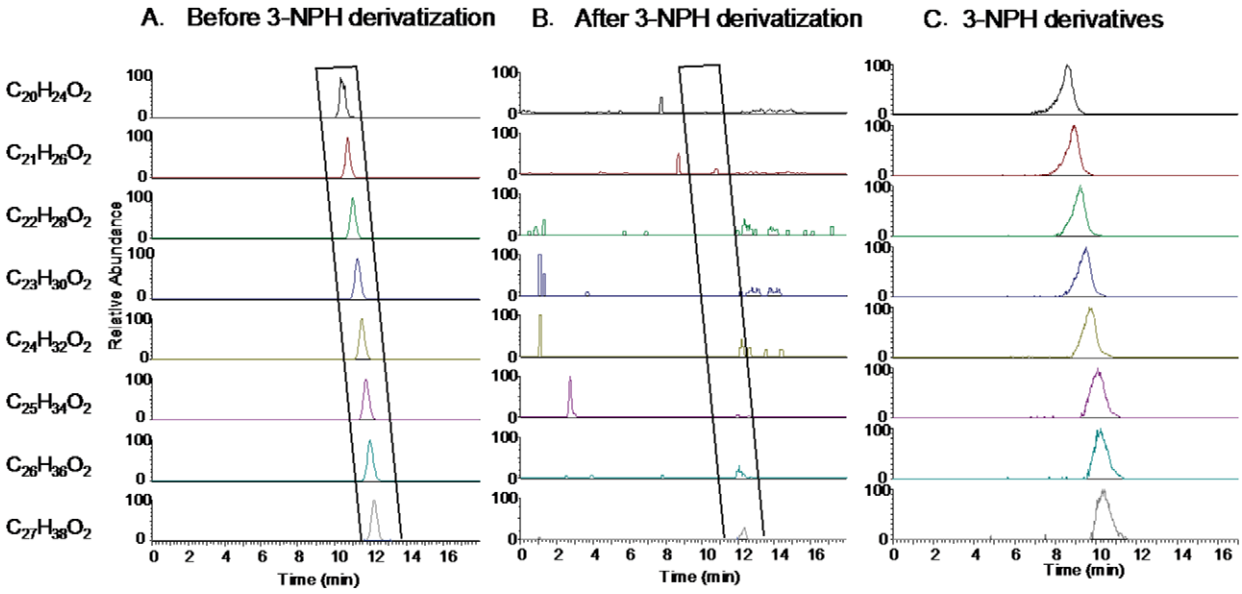

Figure 3. Confirmation of the presence of a carboxylic group in the tentatively assigned O2NAs by chemical derivatization UHPLC–UHRMS. (A) XICs of free O2NAs ($z = -16$ and $n = 20$–27 in $C_nH_{2n+z}O_2$) without derivatization; (B) XICs of the same O2NAs after 3-NPH derivatization, and (C) XICs of 3-NPH derivatives of the same O2NAs following reaction with 3-NPH.

and −2. Also, the Merichem technical mixture had a more uniform compositional distribution of O2NAs with a much lower abundance of saturated acyl acids, as compared to the O2NA compositions of the Sigma technical mixture.

The significantly different NA compositions, as detected in the two NA technical mixtures, demonstrated the complicated compositions of O2NAs in these commercial mixtures and highlights a need for a standardized NA technical mixture that could be used for quantitative analysis of O2NAs in various water samples, including OSPW.

3.3 Molecular compositions of O2NAs in two OSPW samples

With the improved chromatographic behavior of O2NAs on the polar CN column and their more complete elution at <90% ACN, which assisted in their detection by ESI/MS, we next profiled the O2NAs in OSPW, particularly those high-MW and highly hydrophobic species that have not previously been characterized or assigned. We used this new UHPLC–UHRMS method to determine the compositions of bitumen-derived O2NAs in OSPW-1 from the surface mining HWE operation and OSPW-2 from the SAGD operation.

3.3.1 Extraction of O2NAs in OSPW

To efficiently extract O2NAs from the two OSPW samples, we compared three sample preparation approaches, i.e., aLLE using ethyl acetate as the organic solvent, MAX/RP-SPE, and ACN precipitation. It was found that the three sample preparation approaches showed different extraction efficiencies for O2NAs detected in the two different OSPW samples. Supporting Information Fig. S3 shows the bar graphs of the average peak areas + standard deviations (SDs) of representative O2NAs from the analyses of triplicate aliquots of each of the two OSPW samples. These detected O2NA species have different hydrogen deficiencies with z values of −2, −8, −14 and −20, and the n values are 20, 24 and 28 for OSPW-1. For OSPW-2, the z values are −2, −8, −14 and −18, with a same set of the n values. Based on the comparison, ACN precipitation was more effective for extraction of O2NAs from OSPW-1 than other two procedures, while aLLE and MAX/RP-SPE showed very similar extractions of O2NAs from OSPW-2. For consistency, aLLE was used to extract O2NAs from OSPW-2 in the subsequent experiments.

3.3.2 O2NA molecular compositions in OSPW

The molecular compositions of O2NAs were determined in the extracts from OSPW-1 and OSPW-2 using the two different sample preparation approaches. As a result, 424 and 198 molecular formulas corresponding to $C_nH_{2n+z}O_2$ O2NAs were assigned. Figures 5A and 5B are the contour plots and bar graphs showing the molecular compositions of O2NAs detected in the two OSPW samples. OSPW-1 from the surface mining process using HWE contained many more species of O2NAs, with molecular compositions ranging from z = 0 to −24 and n = 8 to 50 in the $C_nH_{2n+z}O_2$ formula. The major O2NA components in this sample were those with z = −2, −4 and −6, and, to a lesser extent, with z = 0 and −8 to −16. In contrast, the O2NAs detected in OSPW-2 from the SADG mining process had O2NA compositions with a

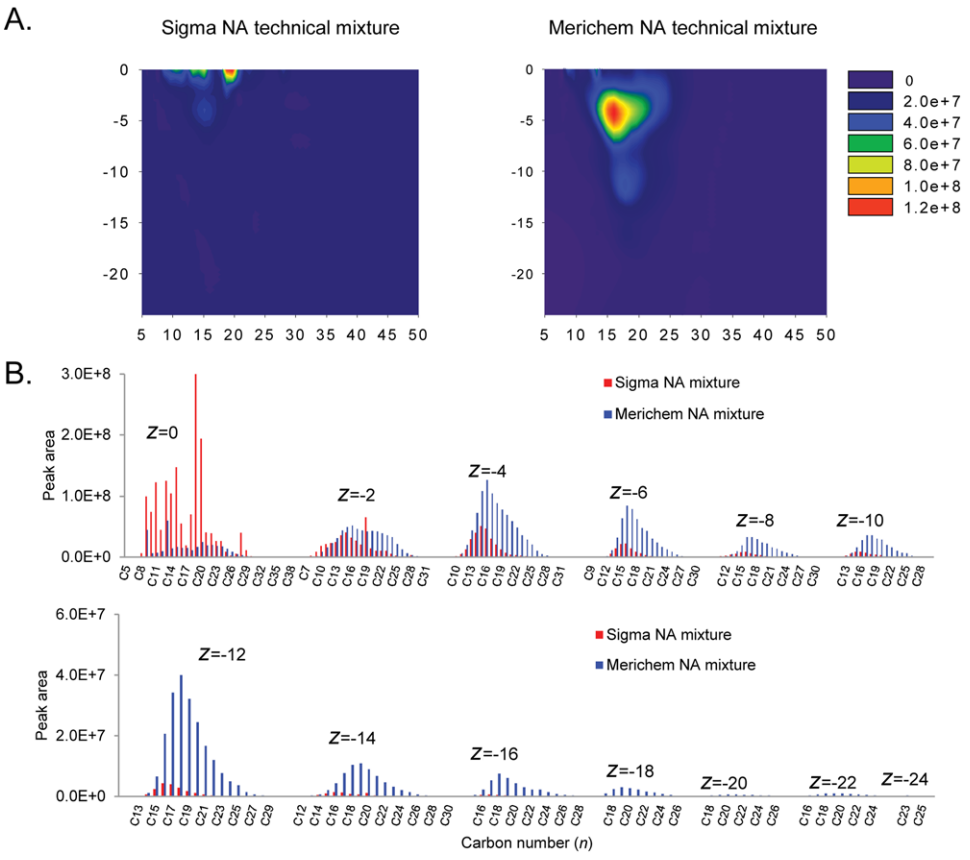

Figure 4. Molecular compositions of the O2NAs detected in the two technical NA mixtures by UHPLC–UHRMS. (A) Contour plots; (B) bar graphs.

narrower range of z values, from 0 to −20. The major O2NAs detected in OSPW-2 are those with z = −4 to −18, with the most abundant species being those with n = −12.

Comparing the profiles of the O2NAs detected in the two technical mixtures with those in the two OSPW samples, OSPW-1 had a profile of O2NAs similar to the Merichem NA technical mixture although it contained many higher-MW O2NAs with n up to 50 (e.g., z = −18 to −24). Although OSPW-2 had a smaller number of O2NA species than OSPW-1, the XICs of O2NAs indicated that there were two peaks for 88 of the 198 molecular formulas determined from OSPW-2, indicating the existence of structural isomers for each of these 88 determined O2NA compositions. These 88 isomeric peaks were not detected in OSPW-1 or either of the two technical mixtures. Figure 6 shows the XICs for eight homologous O2NAs (n = 20 to 27) belonging to z = −8, determined in the two technical mixtures and the two OSPW samples. In addition to the peaks with same RTs which were detected in all four of the samples, there were a series of peaks with the measured mass errors of 0.8 to 1.6 ppm, which were only detected in OSPW-2 for each of these eight O2NAs but with shorter LC RTs. It is worth mentioning that even though 88 pairs of O2NA structural isomers detected in OSPW-2 were resolved on the CN column without 3-NPH derivatization, it was found that the structural isomers for each pair showed only a singlet peak for their 3-NPH derivatives (data not shown). This indicated that the introduction of the polar 3-NPH moiety to the O2NA structures changed the separation selectivity for these isomeric pairs on the CN column, in addition to the broadening the peak shapes of the 3-NPH derivatives as mentioned above.

In summary, these results indicated that the OSPW samples not only had more complicated compositions of O2NAs than those of the NA technical mixtures, but also had O2NAs profiles that were significantly different from each other. To our knowledge, this is the first time that more than 400 molecular formulas corresponding to O2-specific NAs have been determined in a single OSPW sample. This can be attributed to the use of the polar reversed-phase CN column which resulted in improved chromatographic behavior of O2NAs, including the sharper peaks which increased the detection sensitivity and the reduced ionization suppression of ESI during LC-MS runs. Use of the CN column also resulted in a more complete elution of high-MW O2NAs from the LC column at <90% aqueous ACN which favored ESI analysis of these hydrophobic molecular species.

3.4 Quantitation of O2NAs in OSPW

After determining the profiles of O2NAs in the two technical mixtures and the two OSPW samples, we next evaluated

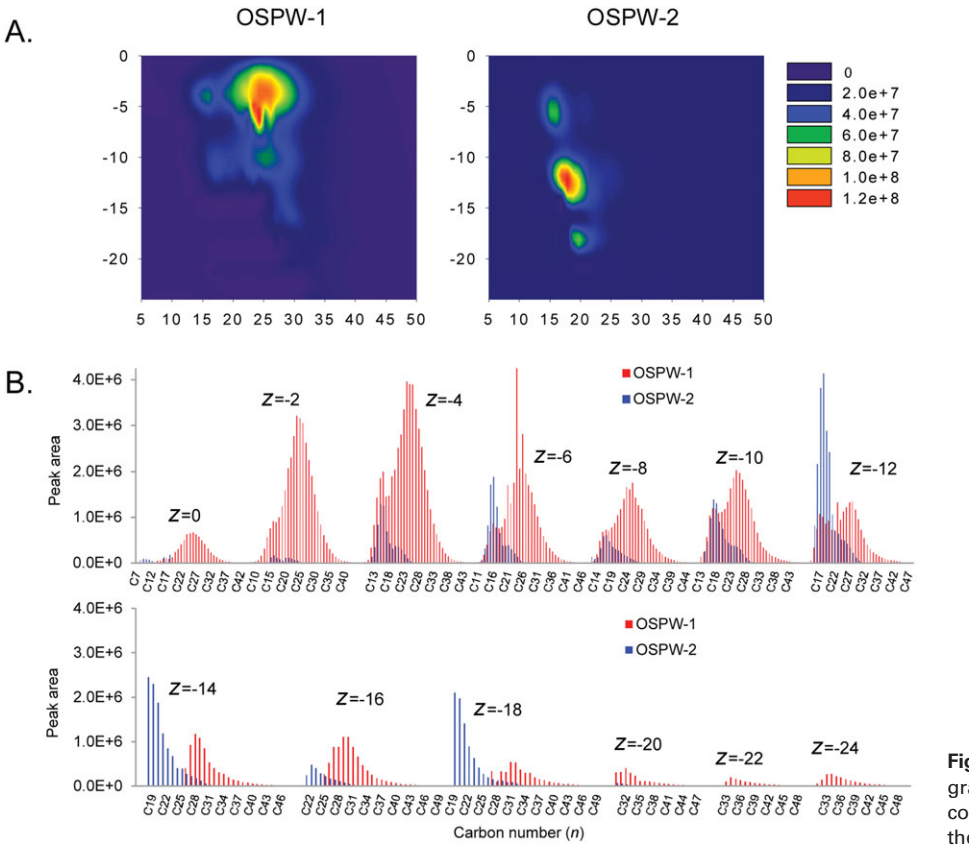

Figure 5. Contour plots (A) and bar graphs (B) showing the molecular compositions of O2NAs detected in the two OSPW samples.

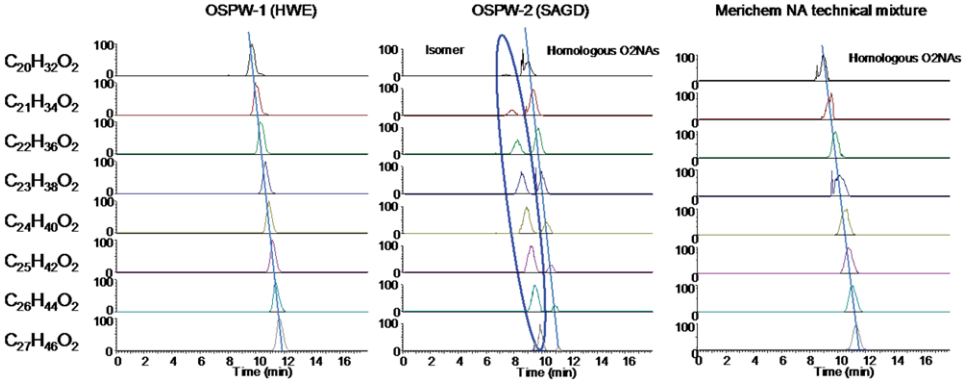

Figure 6. XICs of a same homologous O2NA homologue ($z = -8$; $n = 20$–27) detected in two OSPW samples and the Merichem technical mixture, indicating that an isomeric homologue was detected in the SAGD OSPW (OSPW-2) sample.

this method for quantitative analysis. Because of the need for two different sample preparation procedures for extraction of O2NAs from the two OSPW samples and the time-consuming post-acquisition data processing, only the quantitation of OSPW-1 is shown here. Because ACN precipitation showed higher efficiency than the other two sample preparation procedures, including aLLE, for extraction of O2NAs from OSPW-1, as demonstrated in the "3.3.1 *Extraction of O2NAs in OSPW"* section, it was used as the optimized sample preparation approach for quantitative analysis of O2NAs in OSPW-1. Also, the Merichem NA technical mixture used in this study was more similar in molecular composition of O2NAs to OSPW-1, it was used as the "standard substance" to make the solutions for preparation of the calibration curves. Supporting Information Fig. S4 shows the external calibration curve for the quantitation. This calibration curve was plotted using the *total* peak areas (As) versus the nominal concentrations (C) of all of the O2NAs, not selected O2NAs, that were detected in a step-by-step dilution of the "standard substance" solutions – a 1:4 (v/v) dilution of a stock solution (1 mg/mL) of the Merichem mixture. Good linearity (R^2 = 0.9999) that spanned a 256-fold concentration range was observed. The lower LOQ and upper LOQ were 2.7 and 667 μg/mL, respectively, with 10-μL injections. Also included in Supporting Information Fig. S4 are the calibration curves for four individual O2NAs (n = 20, z = 0, −6, −12 and −18) based on linear regression of their measured peak areas versus the nominal concentrations of total O2NAs. As shown,

good linearity was also observed for the individual O2NAs. After sample preparation using ACN precipitation, the total concentration of O2NAs in OSPW-1 was determined to be 25.8 mg/L (CV% = 4.5, n = 3).

To estimate the quantitation accuracy, spiking recovery tests were performed by adding the NA technical mixture to different aliquots of OSPW-1 at two spiked levels, 20 and 50 mg/L, respectively. The recoveries measured at the two spiking levels were 67.4% (CV% = 5.4, n = 3) and 64.2% (CV% = 6.2, n = 3), respectively. The accuracy decreased with the increased spiking levels, which can be explained by the increased matrix/suppression effects of ESI at the higher spiking level. Alternatively, the lower accuracy might be due to the fact that the exact composition of the O2NAs in the technical mixture did not match that of this OSPW sample.

4 Conclusion

The molecular compositions of the O2NAs in two commercial NA technical mixtures and two OSPW samples from two representative oil sands mining operations (i.e., HWE and SAGD) were qualitatively determined by reversed-phase UHPLC–UHRMS, with a polar cyanopropyl-bonded phase column for the chromatographic separation. Use of the CN column produced better chromatographic behavior, including sharper peak shapes, narrower elution time window, and more complete elution of the O2NAs from the LC column compared to the C_8 and C_{18} columns. 187 and 226 molecular formulas of O2NAs were found in the two commercial standard mixtures, and 428 and 198 molecular formulas of O2NAs were found in the two OSPW samples. It should be noted that each molecular formula could represent more than one O2NA compound. In addition, there were 88 isomeric peaks of O2NAs which were uniquely observed in the SAGD OSPW sample. These results show the molecular complexity of the O2NA compositions in the technical mixtures and the higher complexity of the O2NA compositions in OSPW. To our knowledge, these are the highest numbers of molecular compositions of O2NAs that have been profiled so far. The use of the polar CN column-based UHPLC–UHRMS method for quantitation of the O2NAs in the HWE OSPW sample has shown the potential of this method for the quantitation of O2NAs in OSPW.

This work was supported by funding to the Metabolomics Innovation Centre (TMIC) from Genome Canada, Genome BC, and Genome Alberta for operations (205MET and 7203) and technology development (215MET and MC3T). Research funding for sample collection was provided by Alberta Environment and OSRIN (Oil Sands Research Information Network). We thank Dr. Carol E. Parker for helpful discussions and careful review of this manuscript.

The authors have declared no conflict of interest.

5 References

[1] Headley, J. V., McMartin, D. W., *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.* 2004, *39*, 1989–2010.

[2] Headley, J. V., Peru, K. M., Barrow, M. P., *Mass Spectrom. Rev.* 2009, *28*, 121–134.

[3] Jones, D., West, C. E., Scarlett, A. G., Frank, R. A., Rowland, S. J., *J. Chromatogr. A* 2012, *1247*, 171–175.

[4] Headley, J. V., Peru, K. M., Mohamed, M. H., Frank, R. A., Martin, J. W., Hazewinkel, R. R., Humphries, D., Gurprasad, N. P., Hewitt, L. M., Muir, D. C., Lindeman, D., Strub, R., Young, R. F., Grewer, D. M., Whittal, R. M., Fedorak, P. M., Birkholz, D. A., Hindle, R., Reisdorph, R., Wang, X., Kasperski, K. L., Hamilton, C., Woudneh, M., Wang, G., Loescher, B., Farwell, A., Dixon, D. G., Ross, M., Pereira Ados, S., King, E., Barrow, M. P., Fahlman, B., Bailey, J., McMartin, D. W., Borchers, C. H., Ryan, C. H., Toor, N. S., Gillis, H. M., Zuin, L., Bickerton, G., McMaster, M., Sverko, E., Shang, D., Wilson, L. D., Wrona, F. J., *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.* 2013, *48*, 1145–1163.

[5] Martin, J. W., *Environ. Toxicol. Chem.* 2015, *34*, 2682.

[6] Quagraine, E. K., Peterson, H. G., Headley, J. V., *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.* 2005, *40*, 685–722.

[7] Clemente, J. S., Fedorak, P. M., *Chemosphere* 2005, *60*, 585–600.

[8] Frank, R. A., Kavanagh, R., Kent Burnison, B., Arsenault, G., Headley, J. V., Peru, K. M., Van Der Kraak, G., Solomon, K. R., *Chemosphere* 2008, *72*, 1309–1314.

[9] Holowenko, F. M., MacKinnon, M. D., Fedorak, P. M., *Water Res.* 2002, *36*, 2843–2855.

[10] Grewer, D. M., Young, R. F., Whittal, R. M., Fedorak, P. M., *Sci. Total Environ.* 2010, *408*, 5997–6010.

[11] Lo, C. C., Brownlee, B. G., Bunce, N. J., *Water Res.* 2006, *40*, 655–664.

[12] Martin, J. W., Han, X., Peru, K. M., Headley, J. V., *Rapid Commun. Mass Spectrom.* 2008, *22*, 1919–1924.

[13] Nyakas, A., Han, J., Peru, K. M., Headley, J. V., Borchers, C. H., *Environ. Sci. Technol.* 2013, *47*, 4471–4479.

[14] Barrow, M. P., Witt, M., Headley, J. V., Peru, K. M., *Anal. Chem.* 2010, *82*, 3727–3735.

[15] Smith, B. E., Rowland, S. J., *Rapid. Commun. Mass Spectrom.* 2008, *22*, 3909–3927.

[16] Hindle, R., Noestheden, M., Peru, K., Headley, J., *J. Chromatogr. A* 2013, *1286*, 166–174.

[17] Ross, M. S., Pereira Ados, S., Fennell, J., Davies, M., Johnson, J., Sliva, L., Martin, J. W., *Environ. Sci. Technol.* 2012, *46*, 12796–12805.

[18] Huang, R., McPhedran, K. N., Gamal El-Din, M., *Environ. Sci. Technol.* 2015, *49*, 11737–11745.

[19] Wang, X., Kasperski, K. L., *Anal. Methods* 2010, *2*, 1715–1722.

[20] Shang, D., Kim, M., Haberl, M., Legzdins, A., *J. Chromatog. A* 2013, *1278*, 98–107.

[21] Woudneh, M. B., Coreen Hamilton, M., Benskin, J. P., Wang, G. et al., *J. Chromatogr. A* 2013, *1293*, 36–43.

[22] Jie, W., Cao, X., Chai, L., Liao, J., Huang, Y., Tang, X., *Anal. Methods* 2015, *7*, 2149–2154.

[23] Han, J., Lin, K., Sequeira, C., Borchers, C. H., *Anal. Chim. Acta* 2015, *854*, 86–94.

[24] Euerby, M. R., Petersson, P., Campbell, W., Roe, W., *J. Chromatogr. A* 2007, *1154*, 138–151.

[25] Bataineh, M., Scott, A. C., Fedorak, P. M., Martin, J. W., *Anal. Chem.* 2006, *78*, 8354–8361.

[26] Damasceno, F. C., Gruber, L. D. A., Geller, A. M., Vaz de Campos, M. C., Gomes, A. O., Guimaraes, R. C. L., Peres, V. F., Jacques, R. A., Caramao, E. B., *Anal. Methods* 2014, *6*, 807–816.

