Non-target profiling of bitumen-influenced waters for the identification of tracers unique to oil sands processed-affected water (OSPW) in the Athabasca watershed of Alberta, Canada

Craig B. Milestone | Chenxing Sun | Jonathan W. Martin | Greg Bickerton | James W. Roy | Richard A. Frank | L. Mark Hewitt

1Sheridan College, School of Chemical and Environmental Sciences, 7899 McLaughlin Road, Brampton, ON, L6Y 5H9, Canada
2Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, AB, T6G 2G3, Canada
3Department of Environmental Sciences and Analytical Chemistry, Stockholm University, Stockholm, 10691, Sweden
4Environment and Climate Canada, Water Science and Technology Directorate, 867 Lakeshore Road, Burlington, ON, L7S 1A1, Canada

Correspondence
L. M. Hewitt, WSTD, Environment and Climate Canada, 867 Lakeshore Road, Burlington, ON, L7S 1A1, Canada. Email: mark.hewitt@canada.ca

Funding information
Canada-Alberta Oil Sands Monitoring Program; Environment and Climate Change Canada

Rationale: The objective of this study was to identify unique chemical tracers of oil sands process-affected water (OSPW) to enable definitive discrimination of tailings pond seepage from natural bitumen-influenced waters from the Canadian Alberta McMurray formation.

Methods: The approach involved comparing unknowns from an unprecedented sample set of OSPW (n = 4) and OSPW-affected groundwaters (n = 15) with natural bitumen-influenced groundwaters (n = 20), using high-performance liquid chromatography/electrospray ionisation high-resolution mass spectrometry (HPLC/ESI-HRMS) operated in both polarities.

Results: Four unknown chemical entities were identified as potential tracers of OSPW seepage and subsequently subjected to structural elucidation. One potential tracer, tentatively identified as a thiophene-containing carboxylic acid [C_{15}H_{23}O_{3}S]^{-}, was only detected in OSPW and OSPW-affected samples, thereby showing the greatest diagnostic potential. The remaining three unknowns, postulated to be two thiochroman isomers [C_{17}H_{25}O_{3}S]^{+} and an ethyl-naphthalene isomer [C_{16}H_{21}]^{+}, were detected in one and two background groundwaters, respectively.

Conclusions: We advanced the state of knowledge for tracers of tailings seepage beyond heteroatomic classes, to identifying diagnostic substances, with structures postulated. Synthesis of the four proposed structures is recommended to enable structural confirmations. This research will guide and inform the Oil Sands Monitoring Program in its efforts to assess potential influences of oil sands development on the Athabasca River watershed.

1 | INTRODUCTION

The oil sands deposits of northern Alberta, Canada, cover an area of 140,200 km², the third largest petroleum reserve behind Saudi Arabia and Venezuela. They are a compacted deposit of sand, silt, clay, water and bitumen, containing about 6–14 wt % bitumen and 80–85 wt % mineral solids, with the balance being water. Bitumen is a viscous hydrocarbon that is not recoverable in its natural state using...
conventional oil-well production methods. In oil sands mining operations, an alkaline hot water process based on flotation principles liberates the bitumen, creating a complex mixture of oil sands process-affected water (OSPW, 70–80%), solids (20–30%) and residual bitumen (1–3%).

OSPW is of environmental concern due to its potential toxicity to aquatic and mammalian species. To date, the industry has not treated or released OSPW, but worked under a zero-discharge practice. Currently, OSPW is contained in tailing ponds to allow clarification, and recycling back to the extraction process. As of 2016, the tailings ponds covered about 257 km², and contained 1.21 billion m³ of fine fluid tailings and OSPW. Within these containments are features, including clay substrates at the base and perimeters, and perimeter interceptor ditches to return seepage from the containment walls back into the ponds. However, there are numerous studies that have documented or indicated OSPW seepage into groundwater, thus raising concerns about a vector to surface water.

The OSPW recycling process, along with the hot alkaline extraction process, results in an enrichment of natural bitumen-derived substances in OSPW. This mixture of organics has been termed supercomplex and is similar in chemical profile to natural derived substances in OSPW. This mixture of organics has been extracted process, results in an enrichment of natural bitumen-aqueous. Conventional and chemometric approaches may identify the presence/absence of specific chemicals or chemical classes, and are useful for both known and unknown chemicals to show differences between groups of natural and contaminated wastewaters. In the Alberta oil sands region, non-target profiling and analysis of AEOs has successfully differentiated tailings ponds from each other and also has been used to differentiate surface water (e.g., lakes or upstream Athabasca) from bitumen-influenced water.

This study applied a non-target approach to identify tracer compounds unique to OSPW and OSPW-affect groundwater, so that definitive identification of OSPW-affect groundwater at sites influenced by natural bitumen could be realized. To progress beyond previous AEO-centric studies, we investigated both acidic and neutral bitumen-derived organics to further increase the compound classes evaluated within an unprecedented sampling of OSPW (n = 4), OSPW-influenced groundwater (n = 15), and non-OSPW-influenced background groundwater (n = 20) as recently defined in a companion paper by Hewitt et al using a weight of evidence approach.

2 METHODS

2.1 Sample preparation

A total of 35 samples were analyzed. With the exception of sample volumes of OSPW (2 L) vs non-OSPW (~19 L), all aspects of sample collections, storage, handling, extraction, and analysis were identical to minimize variability from these factors. All aspects of Syncrude OSPW, Suncrude Mildred Lake Settling Basin (MLSB) plume and Suncor Tar Island Dyke (TID) drive-point groundwater samples, background groundwater and the method blank for this study were previously described in the companion paper by Hewitt et al. In brief, OSPW samples were collected with stainless dippers into pre-cleaned stainless-steel containers fitted with Viton seals or glass bottles, with assistance and access provided by the respective oil sands operators’ personnel. Shallow riparian groundwater samples were collected using a stainless-steel drive-point system at depths of 30–120 cm below the streambed of the Athabasca River and associated tributaries. The locations (e.g., edge of river, outside of meanders) and timing (e.g., low river flow periods) of sampling were all chosen to best ensure collection from groundwater discharge zones. Furthermore, the drive point was pushed deeper prior to sampling if the water quality parameters indicated a similarity to surface water, which might indicate hyporheic exchange. OSPW from MLSB and groundwater from six sets of nested wells (each with wells at 1–3 depths) along the known MLSB seepage plume were collected. All groundwater sampling commenced following purging of the well or mini-profiler system and equilibration of field-parameters (temperature, electrical conductivity, pH, dissolved oxygen) measured with hand-held meters. No surface water sampling was conducted.
TABLE 1  Sample groupings, dilution factors, total naphthenic acids (Total NAs) and total entities for each ionization mode (ESI− and ESI+). Samples in bold were further analyzed using HRMS as examples of each grouping. Total NAs previously determined in Hewitt et al20 prior to dilution

| Sample grouping                        | Sample within grouping | ESI− dilution factor | ESI+ dilution factor | Total NAs (mg/L) | ESI− # entities | ESI+ # entities |
|----------------------------------------|------------------------|----------------------|----------------------|------------------|----------------|----------------|
| OSPW                                   | Syncrude OSPW          | 4                    | 4                    | 117              | 148            | 6              |
|                                        | Suncor OSPW 1          | 4                    | 4                    | a                | 106            | 10             |
|                                        | Suncor OSPW 2          | 4                    | 4                    | a                | 129            | 10             |
|                                        | Suncor OSPW 3          | 6.7                  | 6.7                  | a                | 119            | 10             |
| OSPW−affected groundwater - Syncrude plume | Syncrude plume 1A | 8                    | 8                    | 168              | 195            | 12             |
|                                        | Syncrude plume 1B      | 4                    | 4                    | 191              | 187            | 12             |
|                                        | Syncrude plume 2B      | 4                    | 4                    | 170              | 180            | 12             |
|                                        | Syncrude plume 2C      | 4                    | 16                   | 148              | 186            | 12             |
|                                        | Syncrude plume 3A      | 1                    | 1                    | 7                | 74             | 10             |
|                                        | Syncrude plume 3B      | 13.3                 | 26.7                 | 144              | 164            | 11             |
|                                        | Syncrude plume 3C      | 13.3                 | 13.3                 | 122              | 151            | 12             |
|                                        | Syncrude plume 4B      | 13.3                 | 13.3                 | 93               | 121            | 12             |
| OSPW−affected groundwater - Suncor Drive point | Suncor DP-4     | 6.7                  | 6.7                  | 39               | 123            | 12             |
|                                        | Suncor DP-5            | 20                   | 26.7                 | 41               | 110            | 12             |
|                                        | Suncor DP-6            | 4                    | 14.8                 | 3                | 93             | 5              |
| Background groundwater (non-OSPW affected) | BG-1                   | 4                    | 4                    | 10               | 34             | 1              |
|                                        | BG-2                   | 2.7                  | 2.7                  | 11               | 74             | 3              |
|                                        | BG-3                   | 2                    | 2                    | 6                | 94             | 4              |
|                                        | BG-4                   | 1.3                  | 1.3                  | 8                | 78             | 1              |
|                                        | BG-5                   | 20                   | 20                   | 153              | 49             | 3              |
|                                        | BG-6                   | 2                    | 2                    | 7                | 73             | 3              |
|                                        | BG-7                   | 2                    | 2                    | 8                | 42             | 2              |
|                                        | BG-8                   | 2.7                  | 2.7                  | 35               | 68             | 6              |
|                                        | BG-9                   | 8                    | 8                    | 47               | 102            | 8              |
|                                        | BG-10                  | 8                    | 8                    | 54               | 92             | 9              |
|                                        | BG-11                  | 8                    | 8                    | 72               | 50             | 6              |
|                                        | BG-12                  | 2.7                  | 2.7                  | 29               | 32             | 0              |
|                                        | BG-13                  | 2                    | 2                    | 9                | 92             | 9              |
|                                        | Suncor DP-1            | 13.3                 | 26.7                 | 48               | 87             | 4              |
|                                        | Suncor DP-2            | 13.3                 | 13.3                 | 37               | 110            | 6              |
|                                        | Suncor DP-3            | 1                    | 1                    | 1                | 58             | 2              |
|                                        | Syncrude plume 4A      | 3.3                  | 3.3                  | 3                | 63             | 5              |
|                                        | Syncrude plume 5B      | 3.3                  | 3.3                  | 1                | 15             | 1              |
|                                        | Syncrude plume 5C      | 3.3                  | 3.3                  | 1                | 17             | 1              |
|                                        | Syncrude plume 6A      | 3.3                  | 3.3                  | 3                | 14             | 1              |

*Total NA was not determined.
Three additional OSPW samples collected from three different sites within Suncor Pond 8A in May 2013 were also included and prepared as described above. Aliquots (0.5 mL) of all extracts (including a method blank) were evaporated and reconstituted in methanol (0.5 mL), with any further dilution also using methanol (dilution factors, Table 1). A quality control (QC) sample was generated by pooling 5 μL from each sample to create one representative QC sample.

Samples were placed into four distinct groupings, based on the conclusions of Hewitt et al. OSPW; OSPW-affected groundwater within an OSPW plume at MLSB; OSPW-affected groundwater sampled by drive point adjacent to Suncor TID; and natural background groundwater (Table 1). The OSPW and OSPW-affected groundwater samples included four OSPW samples collected from tailings ponds (1 Syncrude, 3 Suncor), three OSPW-affected drive-point samples (DP) collected adjacent to a tailings pond (Suncor), and eight samples collected from wells within the MLSB OSPW plume which were previously determined to be influenced by OSPW (Syncrude). This overall group of OSPW-affected samples were compared with samples from 20 background groundwater sites (collected by drive point >1 km from any mining operations), most affected by natural bitumen, and including MLSB plume well samples previously determined to be free of OSPW influence.

2.2 | Non-target analyses

2.2.1 | HPLC/QTOF-MS data acquisition

Analyses by high performance liquid chromatography/quadrupole time-of-flight mass spectrometry (HPLC/QTOF-MS, Agilent 1200 LC system interfaced to an Agilent QTOF 6520 mass spectrometer, resolution = 4700 at m/z 118; Agilent Technologies, Santa Clara, CA, USA) were performed using a Poroshell 120 EC-C18 column (Agilent Technologies; 3.0 x 50 mm, particle size 2.7 μm), temperature controlled at 40° C. A binary solvent system of water (Solvent A) and methanol (Solvent B), both containing 0.1% formic acid, were used in a gradient programme with a constant flow rate of 0.4 mL min⁻¹. The initial mobile phase was 5% B for 2 min, followed by a linear gradient to 95% B over 20 min, held for 10 min, before returning to equilibrate at initial conditions (16 min). A sandwich injection of sample (2 μL) and isotopically labelled internal standards (δ₁₂₀-decanoic acid (mass 191.38); δ₇-9-anthracene carboxylic acid (mass 231.30); C/D/N Isotopes, Pointe-Claire, QC, Canada; 1 μL), with a needle rinse for 10 s, was used for each injection.

The QTOF mass spectrometer used an electrospray ionization (ESI) source in both negative and positive ion mode. Parameters for both modes were as follows: nebulizer gas pressure 35 psig; drying gas flow rate 10 L min⁻¹; gas temperature 350°C; and skimmer voltage 65 V. Further parameters for negative ion mode were as follows: capillary voltage (VCAP) 3000 V, fragmentor voltage 130 V; all other conditions were set by the data system autotune. Parameters for positive ion mode were as follows: VCAP 3500 V and fragmentor voltage 180 V. Initial instrument calibration was performed using the Agilent tune mixture G1969-85000 at low mass range (ESI⁺ m/z 118–1521), using six calibration points. Accurate mass spectra were acquired at 3 scans s⁻¹ in the range of m/z 100–1050 (ESI–) and m/z 100–980 (ESI⁺). Continuous internal mass calibration was performed using two internal reference ions for constant alignment (Agilent reference mixture, purine (m/z 119.0362 (ESI–) or 121.0508 (ESI⁺)), HP-0921 (m/z 966.0007 (ESI–) or 922.0097 (ESI⁺)).

A randomized injection order was utilized, with QC injections (QC sample and instrument blank) after every five samples. The instrument blank confirmed no carryover at any point, while the QC sample showed no change throughout the run. Due to considerable ion suppression, including suppression of internal reference ions, samples were diluted (see Table 1 for dilution factors) based on the relative amount of these internal reference ions at 17.8 min compared with the counts observed in the instrument blank. The selection of the 17.8-min mark was based on the peak of the second observable hump in sample baselines which was interpreted as representing the maximal suppression effect of matrix.

2.2.2 | Non-target data analysis workflow

Data were acquired using Agilent MassHunter Workstation Software -Data Acquisition software (LC/MS version B5.00), with further analysis using Agilent MassHunter Profinder B.06.00. Data files from HPLC/QTOF-MS spectra were processed using Profinder to generate a list of potential entities, where an entity is defined as a potential unknown compound - described by its retention time, calculated mass and abundance. The files were processed using the Batch Molecular Feature Extraction algorithm, and the following parameters that were not default: peak height >600 counts; a charge state ≤3, a quality score of 50.00, and with all features requiring two or more ions (i.e. deprotonated molecule with at least one ¹³C isotope), isotope model of “Common organic (no halogens)”, and in 25% of at least one sample group (sample groupings, Table 1). These entities/compounds were then manually screened (examined one by one) within Profinder to remove any false positives, poor matches or duplication, reducing the number of entities from 1220 to 210 for ESI– and 1923 to 14 for ESI⁺ (Table 2).

2.3 | HRMS analysis for entity confirmation and structural elucidation

2.3.1 | Chemicals and materials

Ammonium acetate and methanol (LC/MS grade; Fisher Scientific, Fair Lawn, NJ, USA), and compressed CO₂ gas (grade 4.5; Praxair Canada, Mississauga, ON, Canada) were used. Analogues of postulated structures 5-(2-thienyl)pentanoic acid, thiochroman-4-one, 1-methylnaphthalene, 1-ethylmethylmethoxyphenylmethane...
and 2,6-diisopropynaphthalene were obtained from Sigma-Aldrich (Oakville, ON, Canada).

### 2.3.2 Analytical methods

Analyses by HPLC/Orbitrap-ESI-HRMS used an ARIA MX transcend HPLC system coupled with an Orbitrap Elite hybrid mass spectrometer (both from Thermo Fisher Scientific, San Jose, CA, USA). The HPLC conditions were the same as used for the non-target approach. The ionization potential was set at ±4 kV, with the sheath, aux, and sweep gas (N₂) flow rates at 40, 25, and 2 (arbitrary units), respectively. The vaporizer and capillary temperatures were 350°C and 325°C, respectively. Acquisition was performed in full scan mode (m/z 100 to 500) at 1.2 Hz with the resolving power set to a nominal full width half-maximum value of 240,000 at m/z 400. The tandem mass spectrometry (MS/MS) analysis was performed on these entity precursor ions using collision-induced dissociation (CID) and higher-energy collision dissociation (HCD). Xcalibur 2.2 software (Thermo Fisher Scientific) was used for data acquisition. For formula assignment, elements were restricted to C (2–40), H (4–80), O (0–10), and N (0–2), S (0–2), with the double-bond equivalent between 0 and 40 and mass error <5 ppm. Element count heuristics were applied according to the Seven Golden Rules to exclude certain proposed empirical formulae. After applying these restrictions, ions < m/z 400 were all assigned a single formula. Isotope patterns were then verified for these entities.

### 2.3.3 Supercritical fluid chromatography (SFC)/Orbitrap HRMS analysis

SFC separation used an 1260 Infinity Hybrid SFC/UHPLC system (Agilent Technologies) in SFC mode, adapting the method of Pereira et al. Three Zorbax Rx-Sil columns (250 × 4.6 mm, 5 μm particle size; Agilent Technologies) were connected in series (total column length 750 mm) and kept at 40°C. Compressed CO₂ gas (grade 4.5) was the feed gas for generation of supercritical CO₂ as the mobile phase (A). To promote separation and ionization in the SFC system the mobile phase modifier (B) was 95% methanol containing 5% 20 mM ammonium acetate in water. The flow rate was 2.5 mL/min with the following gradient: 0–3 min, 2% B; 3–40 min, linear gradient to 5% B; 40–60 min, linear gradient to 15% B; 60–80 min, linear gradient to 30% B, hold for 5 min; 85–90 min, linear gradient return to 2% B, hold for 20 min until the next injection. The system back pressure was set at 120 bar with the backpressure regulator at 60°C. The samples were injected through a fixed-volume (5-μL) sample loop. A makeup flow of methanol at 0.3 mL/min was first combined with the analytical flow, then split between the backpressure regulator and then via a stainless-steel restriction capillary (50 μm ID × 0.5 m) to the ion source of the mass spectrometer. The Orbitrap ESI-HRMS and MS/MS parameters were the same as described above for the HPLC separations.

### 3 Results and discussion

#### 3.1 Non-target screening for potential tracer compounds using HPLC/QTOF-MS

Representative total-ion chromatograms (TICs) for each of the four sample groupings are depicted in Figure 1, and are consistent with previous analyses illustrating the high complexity inherent in oil sands mixtures of organics. Visually, the TICs of each group type are characterized by unresolved humps spanning large retention time ranges (9–22 min) with few individual peaks evident other than the internal standards at 13.93 and 16.97 min. These internal standards were consistent, with very little retention time shift (±0.02 min) among all 35 samples. Tight alignments enabled comparisons of samples based on retention times (Figure 1). A secondary hump, or shoulder, was observed at later retention times (~16–19 min) in ESI–mode for OSPW-affected samples, but was less pronounced in ESI+ mode (Figure 1). Under reversed-phase chromatographic conditions, this second hump is suggestive of less polar or higher molecular weight compounds, possibly indicating a source difference because

| TABLE 2 Workflow summary | ESI- | ESI+ |
|---------------------------|------|------|
| Entity generation – Profinder | 154 | 13 |
| Initial Total entities | 1220 | 1923 |
| Manual screening | 210 | 14 |
| Total entities | 1220 | 1923 |

- Entity included for HRMS structural characterizations (mass = 284.1433)
- Entity excluded from HRMS structural characterization due to low abundance (mass = 499.9372)
concentrations, as an indicator of potential matrix effects.\textsuperscript{20} Ion in Table 1 and, interestingly, were not correlated with NA

Dilution factors required to ameliorate ion suppression are provided are not normalized based on dilution during sample preparation.

impacted background groundwater samples.

this hump was not present in either natural bitumen- or non-bitumen-impacted background groundwater samples.

The TICs in Figure 1 are presented for comparison purposes and are not normalized based on dilution during sample preparation. Dilution factors required to ameliorate ion suppression are provided in Table 1 and, interestingly, were not correlated with NA concentrations, as an indicator of potential matrix effects.\textsuperscript{20} Ion suppression was not previously observed in temporal and spatial comparisons of AEO profiles between two OSPW containments, which were obtained using solid-phase extraction (SPE).\textsuperscript{46} It is possible that the signal suppressions observed in this study are derived from non-polar compounds that would have been co-extracted during the acidified liquid–liquid extraction, and retained, since further sample cleanup was undertaken. Evidence for the presence of non-polar compounds is presented below for one of the postulated structures.

Profinder was next applied to generate a list of total entities from all samples (Table 2, Initial Total Entities), with the intention of using chemometric statistical software (Mass Profiler Professional, Agilent) to interrogate the data. Two problems necessitated a workflow change from that normally prescribed.\textsuperscript{48} First, upon inspection many of the entities were determined to be either false positives, as the generated extracted ion chromatograms (EICs) were electronic noise (data not shown), or misaligned due to wide and non-normal peak shapes (e.g. Figure S1, supporting information). Manually removing false positives and realigning misaligned entities, together with removing entities associated with method blanks, produced a revised and reduced list (Table 2, Total Entities), with entities reduced dramatically (~80% ESI−, 99% ESI+). The high number of false positives is probably due to a combination of the complex mixtures involved, which were not fully resolved, either chromatographically (Figure 1) or by mass (Figures 2–4), and the low threshold for entity detection used in Profinder (>600 counts). These complexities, while expected, resulted in a high number of ions present at any given chromatographic time point, thereby impeding deconvolution using Profinder. Nevertheless, sufficient numbers of screened entities remained for further interrogation, where generally more entities were detected in ESI− mode (210 vs 14 for ESI+). It is important to note that the total number of entities reflect the number of entities that could be extracted by the software, and is not the total number of compounds in each sample.

At this point in the planned workflow, we imported the screened total entities from each polarity into MPP, whereupon a second workflow obstacle, broad peaks (>1 min, Figure S1, supporting information), was encountered. Similar to generating the initial total entity lists from Profinder, the broadness and shape of most of these peaks also prevented alignment in MPP, against internal standards and between the 35 samples, with the software often classifying them as different entities entirely. This limitation of the software for these complex mixtures led to a second manual screening of the Total Entities (Table 2) using Profinder. The presence/absence of entities within the various sample groupings was then manually itemized (e.g. entities unique to OSPW, Table 2). For this study, an entity was considered as present in a sample grouping if it was present in at least one sample from that group.

In general, for both polarities, the total number of entities in the OSPW-affected groupings was much higher than at natural background groundwater sites (Table 1). In non-target analysis, we utilized entity patterns to discriminate and reconfirm placement of samples with questionable OSPW influence: Syncrude Plume 3A, 4A, 5B, 5C, 6A and Suncor DP-6.\textsuperscript{20} As a result of these inspections, we utilized entity patterns to discriminate and reconfirm placement of samples with questionable OSPW influence: Syncrude Plume 3A, 4A, 5B, 5C, 6A and Suncor DP-6.\textsuperscript{20} A new grouping of samples was undertaken. Evidence for the presence of non-polar compounds is presented below for one of the postulated structures.

In general, for both polarities, the total number of entities in the OSPW-affected groupings was much higher than at natural background groundwater sites (Table 1). In non-target analysis, we utilized entity patterns to discriminate and reconfirm placement of samples with questionable OSPW influence: Syncrude Plume 3A, 4A, 5B, 5C, 6A and Suncor DP-6.\textsuperscript{20} A new grouping of samples was undertaken. Evidence for the presence of non-polar compounds is presented below for one of the postulated structures.

In general, for both polarities, the total number of entities in the OSPW-affected groupings was much higher than at natural background groundwater sites (Table 1). In non-target analysis, we utilized entity patterns to discriminate and reconfirm placement of samples with questionable OSPW influence: Syncrude Plume 3A, 4A, 5B, 5C, 6A and Suncor DP-6.\textsuperscript{20} A new grouping of samples was undertaken. Evidence for the presence of non-polar compounds is presented below for one of the postulated structures.

In general, for both polarities, the total number of entities in the OSPW-affected groupings was much higher than at natural background groundwater sites (Table 1). In non-target analysis, we utilized entity patterns to discriminate and reconfirm placement of samples with questionable OSPW influence: Syncrude Plume 3A, 4A, 5B, 5C, 6A and Suncor DP-6.\textsuperscript{20} A new grouping of samples was undertaken. Evidence for the presence of non-polar compounds is presented below for one of the postulated structures.
influence, which they were unable to distinguish clearly from non-OSPW-affected reference samples. Our data indicate that the OSPW influence is at the limit of detectability for these sites, consistent with the Hewitt et al.20 data for Family A/B monoaromatic acids, SFS, fluoride, sweeteners and perfluorinated surfactants. Likewise, the decision to maintain Syncrude plume samples 4A, 5B, 5C and 6A as natural background groundwater samples was supported by the low number of entities detected in these samples (Table 1), that aligned with the natural bitumen-influenced background.

For entities detected by ESI− (210, Table 2), the majority (154) were present in all sample groups. Eight entities were found in both OSPW-affected Syncrude plume samples (44) and the source pond Syncrude OSPW sample (Table 2), suggesting that these could be potential tracers for this site. A further 32 entities were only found in
OSPW-affected Syncrude plume samples (≥4). These 40 entities were broad peaks (e.g. Figure S1, supporting information), spanning >1 min, and were not considered further because we also sought to identify compounds diagnostic for all OSPW sources. When comparing all OSPW samples from both Syncrude and Suncor with the OSPW-affected groundwater samples of the MLSB Plume from Syncrude and DPs 4–6 adjacent to Tar Island Dyke (Suncor), eight common entities were detected that were not in the background samples. A further nine entities were present in these groupings, but also in only one natural background groundwater sample (Table 2, ESI–). The broadness and shape of 15 of these peaks are probably due to insufficient chromatographic resolution of multiple isomers, thus negating them for further investigation in the present study.

Applying a peak width criterion of ≤0.25 min to the remaining entities yielded two sharply defined peaks, indicative of single compounds (Table 2). The first entity (mass = 284.1433, Table S1, supporting information) was present in two Syncrude plume samples, all three Suncor OSPW-affected DP samples, and all four OSPW samples. Low intensity peaks for this entity were tentatively identified in two additional Syncrude plume samples (3B and 4B), but the intensities were such that confirmation of these entities using their isotopic ratios was not possible. As these samples were amongst the most diluted, we have included these as tentatively present (Table S1, supporting information). It is unclear at this point why this entity was not detected in other Syncrude plume samples (1A, 1B, 2B, 3A or 3C), although it is possible that biodegradation, sample matrix effects, or simply detection limits are involved. The second entity (mass = 499.9372, Table S1, supporting information) was present in five plume samples (3B, 3C and 4B tentatively), one OSPW-affected DP sample, and four OSPW (three of which are tentative) samples. As the general abundance (data not shown) of the latter (mass = 499.9372) was considerably lower, it was not selected for further HRMS work as this necessitated sufficient abundance for MS/MS structural characterizations, but nevertheless it is noted as a potential tracer. Thus, the entity at 12.1 min with mass = 284.1433 (Figure 2) was the lone entity detected by ESI– that was selected for structural elucidation as a potential tracer for OSPW seepage. It should be noted that further sample preparation to better remove matrix effects (e.g. separate neutrals from polar compounds) and concentrate samples may improve detection of these specific two entities.

For entities detected by ESI+ (n = 14; Table 2), all but one were found in all groups, with the lone unique entity being detected only in the background groundwater grouping. Of the 14 entities, three met the peak width criterion described above. Two of these entities were in only one reference site (DP-2), while the third entity was in three reference sites (DP-2, BG-3 and Plume 4A) (Table S1, supporting information). As noted in Hewitt et al., DP-2 did contain a high level of natural bitumen, with visible oil globules. All three chosen entities were detected in all four OSPW samples, all Syncrude OSPW-affected plume samples, and the two primary Suncor OSPW-affected DP samples adjacent to Tar Island Dyke. While still being detected at a limited number of reference sites, they were still included for structural elucidation work because they were highly abundant, sharply defined peaks with greater diagnostic potential than the previously identified Family A and B mono-aromatic acids identified by Hewitt et al.

Following the redesigned workflow (Table 2), four entities were ultimately selected for further evaluation. These are depicted below (Figures 2–4) and described by retention time, full mass spectra detected at that time, and the extracted entity 13C isotopic mass used by Profinder. Of these four entities, two are apparent isomers detected by ESI+ at m/z 309.1496 at retention times of 12.1 and 12.4 min. These four entities were subjected to detailed HRMS/MS experiments to generate data for structure assignment.

3.2 | Structural characterizations of potential OSPW tracers by HRMS

Representative samples from each grouping were selected for further Orbitrap HRMS experiments: Syncrude and Suncor-1 OSPW, Plume 2C, Suncor DP-5, BG-3, BG-5, and method blank (Table 1, in bold). Using these same samples, the peaks and ions associated with the possible tracers identified by HPLC/QTOF-MS were first independently confirmed using HPLC/ESI-Orbitrap HRMS full scanning mode (both ESI– and ES+), thus validating the alternative workflow.

Accurate mass measurements for the four entities were performed using enhanced separations by SFC on long columns (750 mm) to investigate the presence of other isomers that may exist in a single m/z value attributed to an entity derived from the QTOF data, and to provide further confirmation that the entities were consistent (same retention times) between samples (example EICs are provided in the supporting information). The SFC system was coupled to an Orbitrap HR mass spectrometer (Resolution = 240,000) to facilitate formula confirmation and structural elucidation by MS6.

Identification of the entity at m/z 283.14 (ESI–): The Orbitrap HRMS full scan of the peak at 12.1 min showed a base peak at m/z 283.1371, assigned as [C15H23O3S]– with high mass accuracy (±0.71 ppm). The isotopic pattern also matched the theoretical isotopic pattern of [C15H23O3S]–, with the signal abundance of [C15H23O3S]–, 4.4% higher than the intensity of [C15H23O3S]–. Using an HCD energy of 120 eV, MS/MS analysis was performed on this ion (Figure 5A) and product ions at m/z 265.1270 and 221.1374 were observed from neutral loss of H2O and CO2, suggesting the presence of hydroxyl and carboxylic functional groups, respectively. The series of product ions separated by 14 units (CH2)n were observed from neutral loss of O2 and CO2, supporting the presence of hydroxyl and carboxylic functional groups, respectively. The series of product ions separated by 14 m/z units (CH2)n at m/z 193, 179, 165 and 151 provides strong evidence for the presence of a linear alkyl chain core structure. A sulfur atom was present in all these product ions, suggesting that this heteroatom was part of a relatively stable ring structure, such as a thiophene group. MS/MS analysis, using a HCD energy of 80 eV, of a reference standard of 5-(2-thienyl)pentanoic acid also showed characteristic neutral losses of H2O and CO2 (Figure S2, supporting information), supporting the possible structure for this entity proposed in Figure 5A, although there is uncertainty regarding the position of the hydroxyl group and of the
alkyl chain substitution. Thiophenes occur widely in petroleum products, therefore, our suggestion of this entity as a thiophene-containing carboxylic acid is consistent with the sample type. Further analysis using SFC/Orbitrap HRMS of the ion at m/z 283.1371 did not improve the separation, and did not help us achieve better structural elucidation in this case.

Identification of the entity at m/z 309.15 (ESI+): Using HPLC/Orbitrap HRMS accurate mass measurements, the formula of the ion at m/z 309.1523 from the peaks at 12.1 and 12.4 min was assigned as [C_{17}H_{25}O_{3}S]^{+} (Δm = 1.29 ppm). Similar to the aforementioned entity at m/z 283.1371, the presence of the isotopic pattern of sulfur further supported this formula assignment. The HPLC/Orbitrap-MS/MS analysis (CID 30 eV) of the two ions at 12.08 and 12.4 min generated extensive fragmentation patterns that were very similar (Figure S3, supporting information). Better chromatographic separation was achieved using SFC, with 12 major peaks observed eluting between 50 and 80 min, as shown in the extracted ion chromatogram of the ion at m/z 309.1523 (±5 ppm) (Figure S4, supporting information). However, we could not confirm which peaks separated by SFC corresponded to the two peaks separated by HPLC due to different column chemistry. Thus, MS/MS analysis was performed on the m/z 309.1523 ions (50–80 min). The MS/MS spectra of these ions at 69.3 and 69.5 min were similar (Figures S5B and S5C, supporting information), and were quite different from that of the m/z 309.1523 ion at 55.6 min (Figure S5A, supporting information). The major product ions at m/z 291.1418, 245.1367 and 212.1560 generated under HPLC/Orbitrap MS/MS analysis (Figure S3, supporting information) were also observed in the

FIGURE 5 A, HPLC/MS/MS (ESI−, HCD 120 eV) of the ion at m/z 283.1371 at 12.1 min; B, SFC/MS/MS (ESI+, CID 30 eV) of the ion at m/z 309.1523 at 69.3 min; C, SFC/MS/MS (ESI+, CID 30 eV) of ion at m/z 213.1644 at 8.2 min. The proposed structure is shown in shade for each corresponding MS/MS spectrum.
SFC/Orbitrap MS/MS spectra of the m/z 309.1523 ion at 69.3 and 69.5 min (Figures S5B and S5C, supporting information). Therefore, we focused the interpretation of MS/MS spectra on the two peaks at 69.3 and 69.5 min.

The neutral loss of H2O from the precursor ion at m/z 309.1523 yielded the product ion at m/z 291.1418, which then underwent further loss of either CO (to yield the ion at m/z 263.1472) or H2O (ion at m/z 273.1314) (Figure 5B). These neutral losses supported the presence of a carbonyl and two hydroxyl groups in the precursor ion at m/z 309.1523. Neutral loss of H2O from m/z 263.1472 generated a product ion at m/z 245.1367. This product ion could then undergo either loss of SH (33 Da) resulting in the product ion at m/z 212.1560 [C16H20S]+ or of C2H4 (28 Da) to yield the ion at m/z 217.1045 [C14H17S]+. The radical ion at m/z 212.1560 then underwent extensive fragmentation, resulting in the formation of the tropylium cation at m/z 91.0540 corresponding to a benzene ring core-structure. Thus, we proposed that this entity could be a thiochroman with a ketone functional group on the S-containing cyclic moiety, with two hydroxyl groups probably located on an alkyl side chain. This proposal was also supported by the MS/MS analysis (CID 30 eV) of a reference standard of thiochroman-4-one, which showed the loss of CO and H5, and the formation of a tropylium cation that are characteristic to the cyclic ketone, thiane and benzene structure, respectively (Figure S6, supporting information). The two m/z 309.1523 ions at 69.3 and 69.5 min appear to be isomers with minor structural differences, as their MS/MS spectra were very similar (Figure S5B and S5C, supporting information).

Identification of the entity at m/z 213.1644 (ES+): The accurate mass measurement of the ion at m/z 213.1644 at 18.7 min allowed assignment of the formula as [C16H21]+ with good mass accuracy (Δm = −1.88 ppm). The chromatographic separation was improved by using SFC (Figure S7, supporting information), and the SFC peak at 8.2 min was determined to correspond to the peak at 18.7 min detected by HPLC/Orbitrap HRMS, based on the consistency of this ion being detected in the OSPW-A, OSPW-B-1, GW-2C and GW-DP5 samples, and being absent in the method blank, BG-3 and BG-6 samples. SFC/Orbitrap MS/MS analysis was performed on the m/z 213.1644 ion at 8.2 min, and the spectrum showed a series of product ions separated by 14 m/z units (CH2) indicative of a linear alkyl chain.53 The double-bond equivalent number (DBE = 7) of this entity (C16H20S) indicated the probable existence of a naphthalene structure. Therefore, MS/MS analysis was performed on 1-methylnaphthalene, 1-ethyl-naphthalene and 2,6-disopropynaphthalene reference standards in order to further elucidate and confirm the structure of the entity (Figure S8, supporting information). Consistent among all standards was the fragmentation of the alkyl side chains (methyl, ethyl and isopropyl) from the respective naphthalene core structures. For the unknown entity, loss of C2H4 from the precursor ion at m/z 213.1644 suggested an ethyl side chain (Figure S5C). We therefore propose that this entity is a naphthalene with one ethyl branch; however, the structure of ~C4H8 is not clear and we cannot rule out the possibility that ~C4H8 exists as two ethyl branches, or that a longer side chain is present, given the stability of the C11H11 ion and previous observations of alkylated aromatics.54

3.3 | Study implications

In this study, we successfully applied a non-target approach to tentatively identify specific compounds that were diagnostic of the influence of OSPW in highly complex mixtures of bitumen-influenced groundwaters. Previous non-target studies have applied two-dimensional chromatographic techniques and/or HRMS instrumentation to profile heteroatomic ion classes (O2S, etc.) that may have diagnostic potential for tracing OSPW.27,32,37 Here, we advanced the state of knowledge beyond heteroatomic classes to postulating structures for unknown chemical species that could be unique to OSPW seepage. This was possible because of the sufficiently large and unique sample set that included multiple OSPW samples and groundwater sampled from many natural bitumen-influenced background sites (>1 km from any tailings pond). Included with these were selected groundwater samples previously identified to have OSPW-influence, adjacent to tailings ponds.20

The four unknowns identified as having diagnostic potential in this study are more diagnostic of OSPW seepage than the Family A isomers identified in the weight of evidence approach used by Hewitt et al20 as the strongest single indicators of OSPW-aFFECTed groundwater. While being significantly enriched in OSPW and in most OSPW-aFFECTed groundwater samples, the Family A and B groups of monoaromatic acids were both routinely detected in natural bitumen-influenced samples. In this study, the entity tentatively identified as a thiophene-containing carboxylic acid was found to be unique to the MLSB Plume (one confirmed, and two possible occurrences), OSPW samples, and in groundwater adjacent to Suncor TID that was affected by OSPW, but not detected in any background sample, therefore showing the greatest diagnostic potential as a tracer. The other three entities, two isomers with thiocromhan-based structures, and an alkyl polycyclic aromatic hydrocarbon (PAH) (with an ethyl side chain), were similarly detected in all OSPW and OSPW-influenced groups. These entities were also detected in a small number of background samples (≤3) and are therefore of lesser diagnostic capability than the postulated thiophene. Despite this background level of incidence these entities still represent an increased diagnostic capability over the Family A isomers identified by Hewitt et al and, as such, would be useful as tracers in a weight of evidence approach, or within site-specific studies. Future structural confirmation could enable quantitative comparisons between background levels and those from industrial sources to enhance their utility.

The entity that was excluded from HRMS structural elucidation experiments (m/z 499.9372, Table S1, supporting information) also showed a high diagnostic potential, warranting further investigation beyond the scope of the present work. The fact that three of the four final tracer candidates were sulfur-containing is consistent with literature on sulfur-containing species as being diagnostic of industrial
sides of bitumen. In addition, the fact that all four entities were detected in DP4 and DPS samples provides further support to the conclusion of Hewitt et al. that OSPW-affected groundwaters is reaching the Athabasca River system adjacent to Tar Island Dyke.

We therefore recommend synthesis of the four proposed structures to enable structural confirmation. Upon structural confirmation, their presence and concentrations in OSPW-affected groundwaters and additional tailings ponds could be confirmed, allowing their use in a targeted analysis of plume tracking with much improved detection limits. In addition to the entity at m/z 499.9372 (Table S1, supporting information), we note that there are probably other unknowns with diagnostic potential not identified in this study. As noted above, the extraction methodology employed considered all water-soluble bitumen organics, while future work should utilize more selective extraction and fractionation approaches (e.g. Bauer et al.) to produce fractions and extracts of lesser complexity. By coupling these improvements together with other chromatographic systems (e.g. SFC/HRMS) and the chemometric approach, additional tracers may be identified. This research will continue to guide and inform the Oil Sands Monitoring Program in its efforts to assess potential influences of oil sands development on the Athabasca River watershed.

**PEER REVIEW**
The peer review history for this article is available at https://publons.com/publon/10.1002/rcm.8984.

**ORCID**
L. Mark Hewitt https://orcid.org/0000-0002-3683-562X

**REFERENCES**
1. Alberta Energy Regulator. Alberta’s energy reserves 2014 and supply/demand outlook 2015–2024: 2015 ST98–2015.
2. Canadian Association of Petroleum Producers. Crude oil: Forecast, markets and transportation; 2016.
3. CONRAD - Promoting Innovation in the Oil Sands. Canadian Oil Sands Network for Research & Development (CONRAD); Edmonton, Alberta: 2010.
4. Allen EW. Process water treatment in Canada’s oil sands industry: I. Target pollutants and treatment objectives. J Environ Eng Sci. 2008; 7(2):123-138.
5. Bartlett AJ, Frank R, Gillis PL, et al. Toxicity of naphthenic acids to invertebrates: Extracts from oil sands process-affected water versus commercial mixtures. Environ Pollut. 2017;227:271-279.
6. Kavanagh RJ, Frank RA, Oakes KD, et al. Fathead minnow (Pimephales promelas) reproduction is impaired in aged oil sands process-affected waters. Aquat Toxicol. 2011;101(1):214-220.
7. Leung SS, MacKinnon MD, Smith REH. The ecological effects of naphthenic acids and salts on phytoplankton from the Athabasca oil sands region. Aquat Toxicol. 2003;62(1):11-26.
8. Lister A, Nero V, Farwell A, Dixon DG, Van Der Kraak G. Reproductive and stress hormone levels in goldfish (Carassius auratus) exposed to oil sands process-affected water. Aquat Toxicol. 2008;87(3):170-177.
9. Nero V, Farwell A, Lee LEJ, Van Meer T, MacKinnon MD, Dixon DG. The effects of salinity on naphthenic acid toxicity to yellow perch: Gill and liver histopathology. Ecotoxicol Environ Saf. 2006;65(2):252-264.
10. Peters LE, MacKinnon MD, Van Meer T, van den Heuvel MR, Dixon DG. Effects of oil sands process-affected waters and naphthenic acids on yellow perch (Perca flavescens) and Japanese medaka (Oriizias latipes) embryonic development. Chemosphere. 2007; 67(11):2177-2183.
11. Li C, Li F, Stafford J, Belosevic M, Gamal E-DM. The toxicity of oil sands process-affected water (OSPW): A critical review. Sci Total Environ. 1785-1802:2017:601-602.
12. Rogers VV, Wickstrom M, Liber K, MacKinnon MD. Acute and subchronic mammalian toxicity of naphthenic acids from oil sands tailings. Toxicol Sci. 2002;66(2):347-355.
13. Alberta Government. Tailings ponds, oil sands tailings ponds locations. 2019. http://osip.alberta.ca/map/
14. Frank RA, Roy JW, Bickerton G, et al. Profiling oil sands mixtures from industrial developments and natural groundwaters for source identification. Environ Sci Technol. 2014;48(5):2660-2670.
15. Hunter GP. Investigation of groundwater flow within an oil sands tailings impoundment and environmental implications: Earth Sciences. University of Waterloo; 2001. http://scholar.google.com/scholar?hl=en&q=Hunter%2C+G.+P.+Investigation+of+groundwater+flow+within+an+oil+sands+tailings+impoundment+and+environmental+implications.+University+of+Warloo,+2001
16. MacKinnon MD, Kampala G, Marsh B, Federak PM. Guigard S. Indicators for assessing transport of oil sands process-affected waters. In: Thomson NR, ed. Bringing Groundwater Quality Research to the Watershed (Proceedings of GQ2004 International Conference of Groundwater Quality). Vol. 297 IAH Press; 2005:71-80.
17. Ross MS, Pereira AS, Fennell J, et al. Quantitative and qualitative analysis of naphthenic acids in natural waters surrounding the Canadian oil sands industry. Environ Sci Technol. 2012;46(23): 12796-12805.
18. Roy JW, Bickerton G, Frank RA, Grapentine L, Hewitt LM. Assessing risks of shallow riparian groundwater quality near an oil sands tailings pond. Groundwater. 2016;54(4):545-558.
19. Sun C, Shoiky W, Cuss CW, et al. Characterization of naphthenic acids and other dissolved organics in natural water from the Athabasca oil sands region. Canada Environ Sci Technol. 2017;51(17): 9524-9532.
20. Hewitt LM, Roy JW, Bickerton G, et al. Advances in distinguishing groundwater influenced by oil sands process-affected water (OSPW) from natural bitumen-influenced groundwaters. Environ Sci Technol. 2020;54(3):1522-1532.
21. Pereira AS, Martin JW. Exploring the complexity of oil sands process-affected water by high efficiency supercritical fluid chromatography/orbitrap mass spectrometry. Rapid Commun Mass Spectrom. 2015;29(8):735-744.
22. Ahad JME, Pakdel H, Savard MM, et al. Characterization and quantification of mining-related ‘naphthenic acids’ in groundwater near a major oil sands tailings pond. Environ Sci Technol. 2013;47(10): 5023-5030.
23. Grewer DM, Young RF, Whittal RM, Fedorak PM. Naphthenic acids and other acid-extractable in water samples from Alberta: What is being measured? Sci Total Environ. 2010;408(23):5997-6010.
24. Headley JV, Barrow MP, Peru KM, et al. Preliminary fingerprinting of Athabasca oil sands solar organics in environmental samples using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom. 2011;25(13): 1899-1909.
25. Ahad JME, Pakdel H, Savard MM, Simard M-C, Smirnoff A. Extraction, separation, and intramolecular carbon isotope characterization of Athabasca oil sands acids in environmental samples. Anal Chem. 2012;84(23):10419-10425.
26. Baker KM. Identification of process water in a surficial aquifer at Syncrude’s Mildred Lake Site: Earth Sciences. University of Waterloo; 1999. https://scholar.google.com/scholar?hl=en&q=Baker%2C+K.
41. Gobelius L, Hedlund J, Dürig W, et al. Per- and polyfluoroalkyl substances in Swedish groundwater and surface water: Implications for environmental quality standards and drinking water guidelines. Environ Sci Technol. 2018;52(7):4340-4349.

42. Anderson RH, Long GC, Porter RC, Anderson JK. Occurrence of select perfluoroalkyl substances at U.S. Air Force aqueous film-forming foam release sites other than fire-training areas: Field-validation of critical fate and transport properties. Chemosphere. 2016;150:678-685.

43. Ruff M, Mueller MS, Loos M, Singer HP. Quantitative target and systematic non-target analysis of polar organic micro-pollutants along the river Rhine using high-resolution mass-spectrometry - Identification of unknown sources and compounds. Water Res. 2015;87:145-154.

44. Schymanski EL, Singer HP, Slobodnik J, et al. Non-target screening with high-resolution mass spectrometry: Critical review using a collaborative trial on water analysis. Anal Bioanal Chem. 2015;407(21):6237-6255.

45. Hollender J, Schymanski EL, Singer HP, Ferguson PL. Nontarget screening with high resolution mass spectrometry in the environment: Ready to go? Environ Sci Technol. 2017;51(20):11505-11512.

46. Frank RA, Milestone CB, Rowland SJ, et al. Assessing spatial and temporal variability of acid–extractable organics in oil sands process-affected waters. Chemosphere. 2016;160:303-313.

47. Pereira AS, Martin JW. On-Line Solid Phase Extraction-HPLC-Orbitrap Mass Spectrometry for Screening and Quantifying Targeted and Non-Targeted Analytes in Oil Sands Process-Affected Water and Natural Waters in the Athabasca Oil Sands Region. School of Energy and Environment, University of Alberta; 2014. https://doi.org/10.7939/R3NZ280X5Z

48. Technologies A. Agilent G3835AA MassHunter mass profiler professional software – application guide. November, 2015 2015.

49. Brunswick P, Shang D, van Aggelen G, et al. Trace analysis of total naphthenic acids in aqueous environmental matrices by LC/MS-QToF direct injection. J Chromatogr A. 2015;1405:49-71.

50. Ahad JME, Paked H, Gammon PR, Siddique T, Kuznetsova A, Savard MM. Evaluating in situ biodegradation of 13C-labelled naphthenic acids in groundwater near oil sands tailings ponds. Sci Total Environ. 2018;64:392-399.

51. Bauer AE, Frank RA, Headley JV, et al. A preparative method for the isolation and fractionation of dissolved organics from bitumen-influenced waters. Sci Total Environ. 2019;671:587-597.

52. Speight JG. The application of spectroscopic techniques to the structural analysis of coal and petroleum. Appl Spectros Rev. 1972;5(1):211-263.

53. Ha J, Cho E, Kim S. Interpreting chemical structures of compounds in crude oil based on the tandem mass spectra of standard compounds obtained at the same normalized collision energy. Energy Fuels. 2017;31(7):6960-6967.

54. Headley JV, Peru KM, Fahlim B, Colodey A, McMartin DW. Selective solvent extraction and characterization of the acid extractable fraction of Athabasca oil sands process waters by Orbitrap mass spectrometry. Int J Mass Spectrom. 2013;345–347:104-108.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.