Temporal dynamics of halogenated organic compounds in Marcellus Shale flowback

Jenna L. Luek  
*University of New Hampshire, Durham, Jenna.Luek@unh.edu*

Mourad Harir  
*Helmholtz Zentrum Muenchen*

Philippe Schmitt-Kopplin  
*Helmholtz Zentrum Muenchen*

Paula J. Mouser  
*University of New Hampshire, Durham*

Michael Gonsior  
*University of Maryland*

Follow this and additional works at: https://scholars.unh.edu/civeng_facpub

Recommended Citation


This Article is brought to you for free and open access by the Civil Engineering at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Civil Engineering Scholarship by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
Temporal Dynamics of halogenated organic compounds in Marcellus Shale flowback

Jenna L. Luek¹, Mourad Harir²,³, Philippe Schmitt-Kopplin²,³, Paula J. Mouser⁴, Michael Gonsior¹

¹University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, USA.
²Helmholtz Zentrum Muenchen, Research Unit Analytical BioGeoChemistry, Neuherberg, Germany
³Technische Universität München, Chair of Analytical Food Chemistry, Freising-Weihenstephan, Germany
⁴University of New Hampshire, Department of Civil and Environmental Engineering, USA

*Corresponding author contact: Jenna L. Luek, Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, PO Box 38 Solomons, MD 20688 USA +14126054898, Email address: jluek88@gmail.com

Key words: high volume hydraulic fracturing, Marcellus Shale, flowback fluid, produced water, halogenated organic compounds
Graphical Abstract

Highlights

• FT-ICR-MS reveals temporal changes in shale gas well organic chemical composition
• Many organohalogens in fluid were unique to the first 3 months of well operation
• Iodinated organic ions in fluid remain abundant over ten months of well operation
• Abiotic and biotic reactions may drive iodinated organic compound formation
Abstract

The chemistry of hydraulic fracturing fluids and wastewaters is complex and is known to vary by operator, geologic formation, and fluid age. A time series of hydraulic fracturing fluids, flowback fluids, and produced waters was collected from two adjacent Marcellus Shale gas wells for organic chemical composition analyses using ultrahigh resolution mass spectrometry. Hierarchical clustering was used to compare and extract ions related to different fluid ages and many halogenated organic molecular ions were identified in flowback fluids and early produced waters based on exact mass. Iodinated organic compounds were the dominant halogen class in these clusters and were nearly undetectable in hydraulic fracturing fluid prior to injection. The iodinated ions increased in flowback and remained elevated after ten months of well production. We suggest that these trends are mainly driven by dissolved organic matter reacting with reactive halogen species formed abiotically through oxidizing chemical additives applied to the well and biotically via iodide-oxidizing bacteria. Understanding the implications of these identified halogenated organic compounds will require future investigation in to their structures and environmental fate.

1. Introduction

Halogenated organic compounds (HOCs) have been recently identified in shale gas wastewaters using both targeted and non-targeted analytical approaches (Hoelzer et al., 2016; Luek et al., 2017; Maguire-Boyle and Barron, 2014). Furthermore, HOCs are of particular interest because they are not known additives used in the hydraulic fracturing process and although a handful of mechanisms have been proposed, their origin remains unknown (Hoelzer...
et al., 2016; Luek et al., 2017; Maguire-Boyle and Barron, 2014). The environmental fate and toxicity of these compounds also remains unknown.

Non-targeted ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is widely used to investigate the chemical composition of diverse organic matter based on its high mass accuracy and resolution (Dvorski et al., 2016; Gonsior et al., 2011; Kellerman et al., 2014; Roullier-Gall et al., 2015; Walker et al., 2014). This approach has been applied to complex natural and engineered systems to describe compositional changes not understood a priori and without the use of hundreds or thousands of costly analytical standards (Chen et al., 2016; Gonsior et al., 2016; Lavonen et al., 2013; Shakeri Yekta et al., 2012; Sleighter et al., 2014). Ultrahigh resolution FT-ICR-MS operated in negative mode is an appropriate approach for identifying unknown deprotonated HOCs including diverse disinfection by-products (DBPs) (Gonsior et al., 2015; Lavonen et al., 2013; Luek et al., 2017; Xu et al., 2013). Paired with solid phase extraction (Dittmar et al., 2008), FT-ICR-MS is uniquely suited for describing the temporal dynamics of diverse HOCs found in high salinity shale gas wastewater.

Understanding the behavior of HOCs within an individual hydraulic fracturing well is essential for narrowing down their possible origins and environmental fate. HOCs have been hypothesized to be transformation products (Hoelzer et al., 2016; Luek et al., 2017), but time series data have not been used previously to investigate this possibility. Therefore, the aim of this study was to track changes to the dissolved organic matter (DOM) pool of Marcellus Shale gas well fluids over the first ten months of well operation. Specifically, we combined solid phase extraction with FT-ICR-MS and used hierarchical clustering analysis to identify key shifts in the distribution of HOCs and suggest plausible formation mechanisms.
2. Methods

Hydraulic fracturing fluid, flowback fluid, and produced water samples were collected from two adjacent hydraulic fracturing wells at the Marcellus Shale Energy and Environment Laboratory (MSEEL) field site in Morgantown, WV (Figure S1) between November 2015 and September 2016 (Carr, 2017). DOM was solid phase extracted from water samples and analyzed using FT-ICR-MS. Hierarchical clustering analyses were performed on the resulting ions and assigned molecular formulas to track the temporal trends of HOCs.

2.1 Sample Collection

MSEEL wells MIP-3H and MIP-5H (herein referred to as 3H and 5H) were sampled from a gas-fluid separator in autoclaved high-density polyethylene carboys from the separator outlet. Fluid was then transferred in to 1L base-washed low-density polyethylene containers using a peristaltic pump, minimizing headspace. Samples were collected approximately daily during the initial week of flowback (December 2015, fluid production rates as high as 1000 barrels d⁻¹), bi-weekly for the following 3 months (early production water, fluid production rates 10s of barrels d⁻¹), monthly for 3 months, and then bimonthly (late production water, fluid production rates very low, ~1 barrel d⁻¹) (mseel.org). On certain dates, the 5H well was not producing fluid so no sample was collected. Samples were refrigerated at 4 °C and shipped on ice within two weeks of collection. One field blank was collected by taking MilliQ water in to the field and then processed in the same manner as samples.

2.2 Organics extraction

Upon receipt, samples were filtered over a 0.7 μm glass fiber filter (Whatman GF/F) in to glass bottles previously baked at 500 °C. Filtered samples (200 mL) for solid phase extraction
were immediately acidified to pH 2 with concentrated hydrochloric acid and extracted over 1g/6mL Bond Elut PPL solid phase extraction cartridges (preconditioned with 2 cartridge volumes of methanol followed by 2 volumes of 0.1% formic acid solution) (Dittmar et al., 2008). Loaded cartridges were desalted using a 200 mL dilute hydrochloric acid rinse (pH = 2) followed by a 30 mL 0.1% formic acid solution rinse to avoid halide contamination of the methanolic extract. Large volume washing of cartridges reduces the likelihood of iodo-adducts (Luek et al., 2017; Xu et al., 2013) and prior investigations with high and ultrahigh resolution electrospray mass spectrometry have confirmed covalently-bound iodine in complex mixtures (Luek et al., 2017; Moulin et al., 2001; Xu et al., 2013). Cartridges were dried under vacuum and eluted with 10 mL ultrapure methanol. Methanolic extracts were stored at -20 °C prior to FT-ICR-MS analysis.

2.3 FT-ICR-MS analysis

Methanolic extracts were diluted 1:5 in ultrapure methanol and injected at 120 uL hr⁻¹ using a Bruker Solarix 12T electrospray ionization FT-ICR-MS located at the Helmholtz Zentrum Munich, Germany. The instrument was operated in negative mode to target solid phase extracted compounds and target HOCs. Complementary positive ionization was not performed although this mode could have ionized additional organic compounds, including nitrogen containing HOCs. 500 scans were averaged for each sample and a post calibration was performed using a list of known DOM internal calibrants to obtain a mass accuracy of less than 0.1 ppm (Table S1). The obtained full scan mass resolution was better than 400,000 at m/z 400, allowing for precise formula assignments (Hertkorn et al., 2008). All m/z ions identified in the field blank were removed prior to further processing. Following the methods of Sleighter et al., (2012) replicate sample mass spectra were compared to confirm that variability in the mass
spectral analysis across samples was different from variability among extraction replicates as a
function of the % of overlapping m/z ions and a regression of peak magnitudes from two
replicates (Figure S2, S3).

Because the ions of interest were not understood *a priori*, no surrogate or internal
standards were added to samples prior to extraction or analysis, and hence why FT-ICR-MS is
used as a semi-quantitative approach. Ion suppression issues caused by changes in the matrix
were limited by diluting samples sufficiently as determined by the transient spectra. However,
remaining extraction and ionization efficiency issues are not addressed using this non-targeted
approach. For this reason, changes that would only be consistent with the expected changes in
the matrix itself (i.e., consistently present in only unbroken fracturing fluids but absent in all
flowback and produced waters, the largest contrast in the fluid matrices) were not discussed to
limit false pattern identification.

2.4 Hierarchical Cluster Analysis

Hierarchical cluster analysis was performed using Gene Cluster 3.0 and TreeView on log
transformed ion abundances to limit clustering driven only by high intensity ions. Clustering
analysis using average linkages was performed on uncentered m/z ions identified in each well on
ions present in 2 or more samples between m/z 150 – 400 (3H, n= 6613; 5H, n=5296). Clusters
were selected with ions unique to flowback and early produced waters (first three months of well
operation) and assigned formulas.

2.5 Formula Assignments

Formulas were assigned to individual ions from the entire spectra including selected
hierarchical clusters using in-house software (Hertkorn et al., 2008). A range of 150 - 700 m/z
was selected to encompass the majority of ions in the spectra and where the calibration is reliable
to 0.1 ppm. Formulas were assigned with a maximum value per assignment of

$$C_{100}H_{280}O_{80}N_{3}S_{2}Cl_{3}Br_{3}I_{3}$$ and a maximum error of 0.2 ppm. The mass error associated with ions identified below the lowest molecular weight calibrant was sufficient for formula assignment in this 0.2 ppm window. Nonsensical formula assignments were removed using a number of criteria in favor of alternative plausible formula assignments. Formulas not passing the nitrogen rule (McLafferty and Turecek, 1993) were removed within this software and remaining assignments were further reduced to remove invalid formulas by removing those with an oxygen to carbon ratio (O/C) greater than one or a negative double bond equivalent (DBE). Raw values of formulas containing only carbon, hydrogen, and oxygen (CHO), as well as nitrogen (CHON) or sulfur (CHOS) number between m/z 150-700 are reported in Figure S4.

Additional filtering of assigned formulas identified during the cluster analysis involved removing assignments with more than 3 heteroatoms (e.g., CHOI₃ kept, CHOI₃S₁ removed) and preferentially removing duplicate assignments with very low O/C ratios and higher heteroatoms based on consistencies found when checking many assigned formulas against isotopic pattern matching and consistently confirming their alternatives. For example, of the observed m/z ions assigned to an iodinated formula, approximately 50% had a duplicate formula assignment containing S and Cl, but were determined false assignments because the distinctive $^{35}$Cl to $^{37}$Cl isotopic ratios were not observed in the mass spectra. It is possible that this stringent filtering criteria may have removed a small number of correct formula assignments but was necessary for management of the large dataset.

Halogenated formula assignments were compared to their theoretical isotopic patterns (Figure S5) and those ions not matching their theoretical isotopic distribution were removed (<0.1 ppm error with error consistent across isotopes and maximum 10% error in magnitude).
HOCs found using the cluster analysis with insufficient intensity to confirm isotopically were not removed (43% of HOC formulas). However, about half of these low intensity ions were members of a homologous series for which at least one member was confirmed isotopically. Degree of confidence for halogenated formula assignments is discussed in Section 3.1 and reported in Table S4. Assigned formulas matching known compounds are putatively named as such based on their plausible presence in these fluids and likelihood to ionize under the methods used, but have not been confirmed structurally.

3 Results and Discussion

3.1 Cluster analysis reveals halogenated ions unique to early produced waters

Due to the obvious differences in the injected fluid (unbroken gel) and the flowback and produced waters (broken gels & shale derived fluids), hierarchical clusters that only described differences between identified ions the injected fluid and flowback and produced waters were not compared to avoid improper comparisons based on possible extraction differences between these two fluid types. Instead, ions unique to a certain period of flowback were selected, absent both before and after the selected period and therefore both the broken and unbroken fluid types. Three clusters could be identified in the 3H well series that contained ions unique to different flowback and early produced waters, representing three fluid age groups (3S-1, 3S-2, and 3S-3) (Figure S6). The 5H well series contained fewer samples and only one cluster of ions unique to flowback and early produced waters could be identified (5S-2) (Figure S7). All m/z ions identified in these clusters are given in supplemental Table S2 regardless of whether or not they could be assigned an unambiguous molecular formula.
The molecular formulas identified in the four selected clusters are presented as Van Krevelen diagrams to visualize differences in the formula assignments as a function of heteroatom type (Figure 1) (van Krevelen, 1950) and oxidation and saturation status. The selected clusters contained a large number of heteroatom formula assignments rather than those containing only carbon, hydrogen, and oxygen typically dominant in natural organic matter (Table S3). The ions in clusters 3S-1, 3S-2, and 5S-2 contain similar DOM heteroatom classes and distributions, with 74 ions shared between the 5H well cluster and either the 3S-1 or 3S-2 cluster. In contrast, only three of the 5H well cluster ions overlapped with those in cluster 3-S3. A large number of ions containing both nitrogen and sulfur were observed.

**Figure 1.** Oxygen to carbon (O/C) ratio versus hydrogen to carbon (H/C) ratio of formulas assigned to molecular ions (van Krevelen, 1950) for hierarchical clusters unique to flowback and...
early produced waters a) 3S-2 b) 5S-2 c) 3S-1 d) 3S-3. Fluid ages for most ions in each cluster are given.

Cluster 3S-3, representing ions present only in early produced waters (32-80 d), was distinct from the other clusters, containing 155 ions with halogenated formula assignments that could be validated to varying degrees (Table S4). The majority of ions were iodinated (52%), while 20% were brominated, 9% were chlorinated, and 19% contained two different halogens. Sixty-five of these halogenated formula assignments were supported with secondary peaks matching their theoretical stable isotopic spectra (e.g., Figure S5). Five iodinated ions had been previously identified in a North Dakota flowback fluid where their assignment was supported by the appearance of a 126.9045 m/z peak (iodine) during fragmentation (Luek et al., 2017). Of the remaining ions assigned plausible halogenated formulas, many were members of homologous series [separated by \( \text{CH}_2 \) groups determined using \( \text{kmd/z}^* \) values (Shakeri Yekta et al., 2012)] where at least one member of the series had been confirmed isotopically. Thirty-eight of the remaining halogenated formula assignments had intensities too low to rely on isotopic pattern matching (particularly iodinated assignments which rely solely on the \(^{13}\text{C} \) peak). Among the other three clusters, only five plausible halogenated formula assignments were identified and three confirmed using isotopic pattern matching.

The overlap of many nitrogen and/or sulfur containing compounds between the 3H well and the 5H well clusters suggests similar processes are occurring in both wells resulting in these ions unique to flowback and early produced waters. This is expected due to their similar hydraulic fracturing fluids and underlying geology. In contrast, the absence of a 5H well cluster containing HOCs suggests a possible differences between the two wells, possibly related to the smaller quantity of ammonium persulfate breaker added to 5H well (see Section 3.3).
3.2 Iodinated organic ions high in flowback and produced water

Due to the large number of iodinated organic compounds identified during the hierarchical cluster analysis, we further investigated the temporal dynamics of this specific class of compounds. Iodinated organic compounds are of particular interest due to their limited known natural occurrence (Dembitsky, 2006) and the toxicity of known iodinated disinfection by-products (Duirk et al., 2011; Plewa et al., 2004; Richardson et al., 2008). Iodinated organic compounds were tracked across all fluids in the time series to investigate this class of HOCs.

Prior to injection, the fracturing fluids were nearly devoid of iodinated ions, with three or fewer identified at relatively low abundances in individual samples (Figure 2a). Beginning with the first week of flowback, the number of iodinated ions increased, and remained high in all produced water samples out to 276 days. The cumulative abundance of all iodinated ions also followed this trend, with higher intensities but more variability observed in the 3H well. A small volume (20 mL) quality control extraction was performed in August 2016 on all previously collected samples and showed the same trends over the time series. The small volume extractions had slightly lower numbers of ions and intensities, likely related to either the smaller sample volumes or the aging of the fluids prior to extraction (3-7 months storage unfiltered at 4 °C).

Because of the expected variability in the make-up of the injected fluid and wastewaters over time, the changes in iodinated ion number and abundance were also computed relative to the number and cumulative abundance of CHO ions present in each fluid sample (Figure 2b). The same temporal trends are observed on the basis of raw number and abundance as are seen relative to CHO ions. The observed pattern is in contrast to the temporal trends for CHO, CHON, and CHOS ions over the time series (Figure S4). This supports that the observed trends are
indeed a function of actual changes in the fluid mixture rather than as a function of analytical
differences due to ion suppression or extraction efficiency.

The majority of the iodinated ions confirmed with $^{13}$C isotopic pattern matching
contained only carbon, hydrogen, oxygen, and one iodine atom, but seven ions contained
nitrogen or sulfur (Table 1), and five ions contained two iodine atoms. The largest iodinated ion
confirmed with $^{13}$C was $m/z$ 433.1206. Above this value, other peaks could be assigned iodinated
formulas, but none had sufficient intensity to allow for confirmation with the $^{13}$C peak. All ions
identified in the 5H well were also identified in the 3H well, but 11 iodinated ions present in the
3H well were not identified in the 5H well. FT-ICR-MS provides no structural information, but
some structures can be inferred based on the limited number of structural isomers for small
compounds and their ability to be extracted and ionize under the experimental conditions. For
example, the corresponding neutral formula $C_2H_2O_2I_2$ is expected to be diiodoacetic acid, a
known disinfection by-product (DBP) (Plewa et al., 2004) and naturally occurring compound
(Dembitsky, 2006). However, most identified ions are large and therefore cannot be structurally
determined without additional analyses.

The observed distribution of O/C and H/C ratios of these iodinated compounds (Figure
S8) was consistent with the distribution of aromatic, particularly oxygen-rich and phenolic
compounds (e.g., fulvic acids) susceptible to reaction with reactive iodine (Moulin et al., 2001).
Halogenation of unsaturated compounds, aromatics, substituted aromatics (i.e., phenols) and
natural organic matter can occur rapidly via electrophilic aromatic substitution (Criquet et al.,
2015; Westerhoff et al., 2004). A similar distribution (based on O/C vs. H/C) of iodinated DBPs
was formed during the chloramination of drinking water containing iodide, indicating the
preferential formation of iodinated compounds from these aromatic DOM precursors (Wang et al., 2016).

**Figure 2.** MSEEL 3H and 5H iodinated ion number and cumulative iodinated ion abundance as raw values (a) and as a percent of CHO number and cumulative abundance (b). Injected fluids are shown as a time point prior to Day 1 of flowback.
Table 1. Iodinated molecular formulas supported with $^{13}$C peak identified in 2 or more samples.

<table>
<thead>
<tr>
<th>Measured Mass</th>
<th>No. Times identified</th>
<th>H</th>
<th>C</th>
<th>O</th>
<th>N</th>
<th>S</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>236.9054$^p$</td>
<td>16</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>276.9367$^S$</td>
<td>16</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>282.9473</td>
<td>11</td>
<td>8</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>292.8952</td>
<td>8</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>292.9316$^p$</td>
<td>18</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>305.9633</td>
<td>$^3$</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>306.9109</td>
<td>18</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>306.9473$^S$</td>
<td>17</td>
<td>8</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>308.9266</td>
<td>21</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>310.8072$^p$</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>320.963</td>
<td>17</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>322.8072$^S$</td>
<td>$^3$</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>322.9422</td>
<td>19</td>
<td>8</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>322.9786</td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>327.0099$^S$</td>
<td>21</td>
<td>16</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>334.9422$^S$</td>
<td>18</td>
<td>8</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>334.9786</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>336.9579</td>
<td>19</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>338.9194</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>338.9735</td>
<td>15</td>
<td>12</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>339.0099</td>
<td>2</td>
<td>16</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>342.9143</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>343.0048</td>
<td>14</td>
<td>16</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>345.0205</td>
<td>14</td>
<td>18</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>350.9372</td>
<td>19</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>350.9735</td>
<td>16</td>
<td>12</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>351.0099$^p$</td>
<td>$^3$</td>
<td>16</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>352.9528</td>
<td>16</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>352.9892</td>
<td>18</td>
<td>14</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>355.0048$^p$</td>
<td>12</td>
<td>16</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>355.0412$^p$</td>
<td>$^3$</td>
<td>20</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>357.0205</td>
<td>4</td>
<td>18</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>364.9528$^S$</td>
<td>17</td>
<td>10</td>
<td>11</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>367.0049</td>
<td>7</td>
<td>16</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>369.0205$^S$</td>
<td>$^3$</td>
<td>18</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>370.0521</td>
<td>2</td>
<td>21</td>
<td>12</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
3.3 Possible origins of halogenated organic compounds

Four sources have been proposed for the origin of HOCs in flowback fluids: a) chemical additives in the hydraulic fracturing fluid b) leached shale molecules c) biotic reactions between additives and/or shale compounds and d) abiotic reactions between additives and/or shale compounds (Hoelzer et al., 2016; Luek et al., 2017; Maguire-Boyle and Barron, 2014). The specific additives reported for MSEEL wells on the FracFocus database (fracfocus.org) do not contain HOCs, so this is an unlikely source.

Thousands of naturally occurring HOCs do exist (Gribble, 2010), but the number of known iodinated compounds is limited, with just over 110 compounds identified in the published literature (Dembitsky, 2006). We searched all samples for the exact masses of all known biogenic iodinated organic compounds (Dembitsky, 2006) and identified four. Three of these

<table>
<thead>
<tr>
<th>Mass</th>
<th>17</th>
<th>10</th>
<th>12</th>
<th>6</th>
<th>0</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>376.9528S</td>
<td>17</td>
<td>10</td>
<td>12</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>376.9891</td>
<td>11</td>
<td>14</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>378.9684S</td>
<td>18</td>
<td>12</td>
<td>12</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>379.0048</td>
<td>9</td>
<td>16</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>381.0205</td>
<td>10</td>
<td>18</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>383.0361</td>
<td>3</td>
<td>20</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>387.031</td>
<td>4</td>
<td>20</td>
<td>12</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>392.9476S</td>
<td>18</td>
<td>10</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>397.0154</td>
<td>11</td>
<td>18</td>
<td>13</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>399.0311S</td>
<td>7</td>
<td>20</td>
<td>13</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>401.0467</td>
<td>2</td>
<td>22</td>
<td>13</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>402.8334</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>404.00</td>
<td>3</td>
<td>15</td>
<td>14</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>406.9633S</td>
<td>16</td>
<td>12</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>420.8076</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>421.0153</td>
<td>9</td>
<td>18</td>
<td>15</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>423.031</td>
<td>10</td>
<td>20</td>
<td>15</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>425.0467</td>
<td>11</td>
<td>22</td>
<td>15</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>427.026</td>
<td>8</td>
<td>20</td>
<td>14</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>427.0623</td>
<td>6</td>
<td>24</td>
<td>15</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>428.7796</td>
<td>9</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>433.1206</td>
<td>20</td>
<td>30</td>
<td>14</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
four were supported with their $^{13}$C peak, diidoacetic acid ($m/z$ 310.80717), iodotyrosine ($m/z$ 305.96326) and iodophloroglucinol ($m/z$ 376.95276, also identified in fracturing fluid prior to injection), while diiodomethane ($m/z$ 266.81732) was present at low abundances so the expected $^{13}$C peak would be below the baseline and therefore could not be confirmed. To this end, the Marcellus shale could contain HOCs derived from ancient biogenic processes that were not remineralized during diagenesis and slowly leached from the shale as the wells matured. However, many biogenic halogenated formulas, such as methyl halides and halogenated phenols can be microbially degraded (Gribble, 2010), and likely would not persist in the environment over geologic timescales. Although the HOC content of shales is poorly characterized, their probable degradation indicates that an ancient biogenic origin for the identified HOCs is unlikely.

In addition to their possible sourcing as natural biogenic products, diiodomethane and diidoacetic acid can also be formed as disinfection by-products (DBPs). These two compounds were not the only putatively identified DBPs in this dataset; several other exact masses also matched known DBPs. The exact masses of DBPs detected only in flowback and produced waters included those matching halogenated acetic acids, iodomethylbutenedioic acid, and several halogenated aromatic structures (halogenated benzaldehydes, benzoic acids, phenols, and benzoquinones). Halogenated DBPs can be formed when oxidizing chemicals such as chloramines and persulfates react with halides to form reactive halogen species, which subsequently react with DOM and xenobiotic compounds (Gong and Zhang, 2015; Plewa et al., 2004; Postigo et al., 2016; Wang et al., 2016; Xie et al., 2015). Of the identified iodinated molecular formulas (Table 1), all but 7 had their non-iodinated counterparts (replacing I with H) present in MSEEL samples and were also present at high intensity ($>10^8$).
Ammonium persulfate (listed as diammonium peroxidisulphate on the FracFocus report), a strong oxidizing agent, was used in both the 3H and 5H well hydraulic fracturing fluids (fracfocus.org). Persulfate oxidation has been used as an advanced oxidation process in wastewater treatment, relying on the activation of persulfate (via heat, UV light, ultrasound, or an electron) to form two sulfate radicals (Matzek and Carter, 2016). High temperatures found in the Marcellus Shale at depth and many potential electron donors (e.g., transition metals, additives) could activate the added persulfate. The resulting sulfate radicals are highly reactive, and can propagate a number of diverse reactions beyond their intended role of breaking polymers. Persulfate oxidation can form reactive halogens including iodine, hypiodite, radical iodine, and others that can ultimately react with organic compounds including DOM and phenols to form DBPs (Lu et al., 2015; Wang et al., 2017). The observation of more diverse HOCs present in the 3H well than the 5H well is consistent with a persulfate source for these compounds, as ammonium persulfate was applied at a concentration 75 times higher in the 3H well than the 5H well (0.00074% vs. 0.00001% by mass of hydraulic fracturing fluid).

Additionally, reactive iodine species involved in these halogenation reactions can also be formed through natural processes (without the external addition of oxidant additives) through abiotic reactions with DOM (Li et al., 2012) and oxidation of iodide to reactive iodine by biotically produced hydrogen peroxide and organic acids (Steinberg et al., 2008).

Bacteria known to oxidize iodide have been identified in hydraulic fracturing wastewaters previously (Amachi et al., 2005; Murali Mohan et al., 2013), and may be an additional source of reactive iodine. These bacteria are capable of converting inorganic iodide to reactive iodine that can subsequently react with DOM along the same pathways as abiotically-produced reactive iodine. Although biocides are employed to limit bacterial growth, diverse and
active communities are found in these fluids downhole (Cluff et al., 2014; Daly et al., 2016; Mouser et al., 2016; Murali Mohan et al., 2013), with hydraulic fracturing increasing the shale poresize and removing this physical limitation for microbial life in the deep shales (Mouser et al., 2016). We searched for taxa phylogenetically associated to known iodide-oxidizing bacteria in 16S rRNA data from MSEEL samples. Taxa closely related to uncultured Roseovarius spp. were observed in MSEEL 3H drilling muds (2-4%) as well as flowback and early produced fluids (from 2 through 119 days) from both the 3H and 5H wells (<1%) (Figure S9, Table S5) (unpublished data from Kelly Wrighton, for methods see Cluff et al., 2014; Daly et al., 2016). Some Roseovarius spp. are capable of iodide oxidation in conjunction with production of iodinated organic compounds, including methyl halides (Amachi et al., 2005; Fuse et al., 2003). This reaction requires iodide, an oxidant (e.g., peroxide), and appropriate genes (i.e. halide peroxidases) that are poorly characterized in bacteria. As mineral iodides and oxidants (e.g. persulfate) are present in this system, these data suggest the potential exists for biotic production of iodinated organic compounds in conjunction with abiotic reactions in this system, albeit by low abundance microbial community members.

4. Conclusions

Iodinated organic ions were tracked through the first nine months of operation of two Marcellus Shale gas wells using FT-ICR-MS and revealed a steep increase in the number of ions assigned iodinated organic molecular formulas during the initial flowback period. The number and abundance of iodinated organic ions remained elevated in produced waters 276 days later. Hierarchical clustering analysis also revealed a large number of iodinated, brominated and chlorinated ions that were unique to fluids returning to the surface of the 3H well 1-3 months
We suggest that biotic and abiotic oxidation of halides subsequently reacting with diverse DOM contributes to the observed organohalogen diversity and temporal dynamics. Tracking changes in the chemical composition of shale gas fluids is essential for understanding fundamental changes occurring in hydraulic fracturing fluids, particularly those driven by known additives and microbial communities. Although many similarities were observed between the two wells, their differences raise questions as to why these differences exist: Do they reflect differences in geology, hydraulic fracturing fluid mixtures, hydraulic fracturing techniques, or some other undescribed variable? More work on this topic is needed to better understand how these results can be generalized to different wells. Understanding why differences are observed across hydraulic fracturing wastewaters is essential for understanding the fundamental functions occurring within a well, and for addressing more applied questions of which wastewaters are suitable for reuse or a given treatment technique.

Supplemental Materials

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.watres.XXXX. Additionally, calibrated mass lists for all samples with ions identified using FT-ICR-MS are available through the Dryad digital repository (http://datadryad.org/).

Acknowledgements

We thank Mary Evert and Andrea Hanson for assistance with obtaining samples, and the Analytical Biogeochemistry research group at the Helmholtz Zentrum Muenchen for assistance with FT-ICR-MS analyses. We thank Kelly Wrighton for supplying 16S rRNA data on the
abundance of *Roseovarius spp.* Financial assistance was obtained from NSF CBET Award #1604475/1604432 to MG and PJM, and NSF MRI Award #1039768. We are grateful for the insightful anonymous reviews that substantially improved this manuscript. This is contribution #XXXX of the University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory.
References


