Organic sulfur fingerprint indicates continued injection fluid signature 10 months after hydraulic fracturing

Journal:	*Environmental Science: Processes & Impacts*
Manuscript ID	EM-ART-07-2018-000331.R2
Article Type:	Paper
Date Submitted by the Author:	24-Sep-2018
Complete List of Authors:	Luek, Jenna; University of New Hampshire College of Engineering and Physical Sciences, Civil and Environmental Engineering Harir, Mourad; Helmholtz Zentrum Munchen Deutsches Forschungszentrum fur Umwelt und Gesundheit Schmitt-Kopplin, Philippe; Helmholtz Zentrum Munchen Deutsches Forschungszentrum fur Umwelt und Gesundheit Mouser, Paula; University of New Hampshire College of Engineering and Physical Sciences, Civil and Environmental Engineering Gonsior, Michael ; University of Maryland Center for Environmental Science

Organic sulfur fingerprint indicates continued injection fluid signature 10 months after hydraulic fracturing

Luek, J.L.[1,2], Harir, M.[3], Schmitt-Kopplin, P.[3,4], Mouser, P.J.[1], Gonsior, M.[2]

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

*Corresponding author contact:
Jenna L. Luek, University of New Hampshire, Durham, NH 03825
+14126054898, Email address: jenna.luek@unh.edu

For submission to Environmental Science: Processes and Impacts

Environmental Significance Statement

A substantial quantity of wastewater is produced through hydraulic fracturing activities to obtain oil and natural gas from unconventional geological shale formations. In this study, we investigate organic sulfur substances present in two shale gas wastewater time series using ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry and show distinct signatures of aliphatic organic sulfur compounds in the fluids 10 months after hydraulic fracturing. These findings suggest that these sulfur compounds, likely alcohol ethoxysulfates used as fracturing additives, are relatively stable under the deep shale well conditions.

Abstract

Hydraulic fracturing requires the injection of large volumes of fluid to extract oil and gas from low permeability unconventional resources (e.g., shale, coalbed methane), resulting in the production of large volumes of highly complex and variable waste fluids. Shale gas fluid samples were collected from two hydraulically fractured wells in Morgantown, WV, USA at the Marcellus Shale Energy and Environment Laboratory (MSEEL) and analyzed using ultrahigh resolution mass spectrometry to investigate the dissolved organic sulfur (DOS) pool. Using a non-targeted approach, ions assigned DOS formulas were analyzed to identify dominant DOS classes, describe their temporal trends and their implications, and describe the molecular characteristics of the larger DOS pool. The average molecular weight of organic sulfur compounds in flowback decreased and was lowest in produced waters. The dominant DOS classes were putatively assigned to alcohol sulfate and alcohol ethoxysulfate surfactants, likely injected as fracturing fluid additives, on the basis of exact mass and homolog distribution matching. This DOS signature was identifiable 10 months after the initial injection of hydraulic

fracturing fluid, and an absence of genes that code for alcohol ethoxysulfate degrading proteins (e.g., sulfatases) in the shale well genomes and metagenomes support that these additives are not readily degraded biologically and may continue to act as a chemical signature of the injected fluid. Understanding the diversity, lability, and fate of organic sulfur compounds in shale wells is important for engineering productive wells and preventing gas souring as well as understanding the consequences of unintended fluid release to the environment. The diversity of DOS, particularly more polar compounds, needs further investigation to determine if the identified characteristics and temporal patterns are unique to the analyzed wells or represent broader patterns found in other formations and under other operating conditions.

1. Introduction

More than 100 billion gallons of wastewater are produced annually from unconventional oil and gas extraction using high volume hydraulic fracturing methods[1]. The high complexity and spatial-temporal variability of these wastewaters have limited characterization of these fluids, particularly the organic components, and a limited understanding of the factors controlling the observed variability[2–4]. The diversity of compounds found in hydraulic fracturing fluids and wastewaters and their mechanistic controls are of broad interest in understanding the natural and engineered processes occurring in deep shales following hydraulic fracturing, as well as understanding the fate of these compounds in the environment and during treatment. Sulfur plays a dynamic biogeochemical role in these fluid systems[5,6], yet the distribution of dissolved organic sulfur (DOS) and its stability and contribution to biological activity at depth has yet to be assessed.

Persulfate is the most frequently used inorganic sulfur hydraulic fracturing additive in a survey of the FracFocus 1.0 database (ammonium persulfate applied to 27 percent of gas/oil wells and 60 percent of oil wells; sodium persulfate applied to 11 percent of oil wells)[7], followed by sulfates and to a lesser degree, thiosulfate, bisulfate, and metabisulfate salts[8,9]. Inorganic sulfur trends in gas wells are commonly measured to understand scaling potential and gas souring due to sulfides[5,10–12]. Temporal flowback and produced fluid sampling in Colorado and Pennsylvania for sulfate generally showed the same trends, peaking during early flowback (2-430 mg L^{-1}) and returning to lower levels in produced waters after 90 days (8 – 100 mg L^{-1})[2,3,13]. However, only a small number of organic sulfur (OS) compounds have been previously targeted in flowback and produced waters, although both OS additives (e.g., sulfate and sulfonate surfactants, heterocyclic biocides[8,9,14], thiourea polymers[7]) and natural/shale derived OS compounds (e.g., alkanethiols and thioheterocycles[15,16]) may be present. The disclosure frequency varies by time period and location, with thiourea polymers (>3900 disclosures), thioheterocyclic biocides (>400 disclosures) and thioglycolic acid (>100 disclosures) most frequently disclosed[7].

Tracking DOS in oil and gas wastewaters can provide information on the natural and engineered processes occurring in deep wells. For example, identification of known hydraulic fracturing additives such as cocamidopropyl hydroxysultaine[14] in flowback fluid helps understand their use usage frequency and stability within the hydraulically fractured formation. OS additives such as sodium dodecyl sulfate (friction reducer, emulsion inhibitor) have also been suggested as incidental long-term sources of unwanted sulfide in high temperature wells[17,18]. Additionally, the identification of an OS compound used during natural gas cleaning near an

unconventional oil and gas wastewater treatment facility advanced understanding of the potential environmental impacts during waste fluid management[19].

Although OS compounds may be involved in many biotic and abiotic processes occurring in hydraulic fracturing wells and wastewaters, selecting targeted analytes among the diverse possible OS additives and naturally occurring compounds limits the scope of any given study. Many known OS additives are not amenable to gas chromatography-based methods, requiring liquid chromatography mass spectrometry for analysis[14,20]. Another analytical challenge limiting targeted quantitation is the presence of many possible OS homologues and co-products for a given additive, such as seen with sulfonated surfactants[21,22]. Ultrahigh resolution mass spectrometry is capable of analyzing ionizable OS compounds in hydraulic fracturing wastewaters including not only known additives such as OS surfactants[23,24] but also potential degradation products and naturally occurring OS compounds[25].

Therefore, the goal of this study was to investigate DOS compounds in hydraulic fracturing fluids and wastewaters using ultrahigh resolution mass spectrometry and identify dominant DOS classes and the molecular characteristics of these compounds. By analyzing ten months of flowback and produced water from two adjacent MSEEL shale gas wells, we determined that the DOS pool was dominated by putative alcohol ethoxysulfates likely applied as hydraulic fracturing fluid additives that remained in the DOS fluid signature ten months after injection. Determining the chemical makeup of these fluids allows us to begin assessing the fate of these fluids and the roles they play in biogeochemical cycling at depth, well productivity, wastewater engineering, and impacting the broader environment.

2. Methods

2.1 Sample Collection and Extraction

Fluid samples were collected from two shale gas wells at the MSEEL site in Morgantown, WV between November 2015 (hydraulic fracturing fluid) and September 2016 (produced waters) (n=24). Detailed information regarding MSEEL fluid sample collection and processing has been reported elsewhere[4]. Briefly, 1L fluid samples were collected from the gas-fluid separator tank and stored at 4°C or on ice until processing. Hydraulic fracturing fluids, flowback fluids, and produced waters (200 mL) were filtered through 0.7 μm glass fiber filters (Whatman GF/F, 47mm) and acidified to pH 2 using hydrochloric acid prior to solid phase extraction with activated Bond Elut PPL cartridges (Agilent, 1g/6 ml). Solid phase extraction was performed under gravity at a flow rate of <10 mL min^{-1}; loaded cartridges were subsequently rinsed to remove salts using 200 mL dilute hydrochloric acid (pH 2) followed by 30 mL of 0.1% formic acid solution. Methanolic extracts (10 mL) were stored at -20°C until analysis.

2.2 FT-ICR-MS Analysis

Methanolic extracts were analyzed using a Bruker Solarix 12 Tesla (Bruker Daltonics, Bremen, Germany) Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR-MS) with electrospray ionization in negative mode located at the Helmholtz Center for Environmental Health in Munich, Germany as reported elsewhere[26]. Briefly, the mass spectrometer was calibrated using arginine clusters and individual samples were post-calibrated using known reference mass lists to obtain a mass accuracy of less than 0.2 ppm and ions identified in an MSEEL field blank were subtracted from the spectra. Formulas were assigned using in-house software (maximum $C_{100}H_{\infty}O_{80}N_3S_2$) to ions between *m/z* 150 and 1000 and invalid formulas were removed by eliminating formula assignments with an oxygen to carbon

ratio greater than 1 and/or a negative double bond equivalency. Further reduction of possible false formula assignments was performed by removing formulas with a double bond equivalent minus the number of oxygen [DBE-O] less than -10 and greater than +10 [27].

3. Results & Discussion

3.1 Distribution of CHOS in MSEEL flowback

The number of ions assigned formulas containing carbon, hydrogen, oxygen and sulfur (CHOS) in the MSEEL wells varied over the first nine months of well flowback.[4] The number of CHOS ions was lowest in the fracturing fluid (90 – 345 ions), consistent with the low total number of ions in these samples, a likely artifact of the gelled matrix of the fracturing fluid. At all other time points, fluids were a high salinity aqueous matrix and relatively consistent across samples. The highest number of CHOS ions was found in early flowback, with more than 900 CHOS ions identified in each well. The percent of CHOS ions relative to CHO ions (unique formulas detected) was also lowest in the fracturing fluid (13 – 28 percent); in flowback and produced water, the number of CHOS ions ranged from 43% -76% of CHO ions. 1354 unique ions containing one sulfur atom ($CHOS_1$), and 203 unique ions containing two sulfur atoms ($CHOS_2$) were identified.

The average and weighted average molecular weight (MW_a, MW_{wa}) and carbon oxidation state[28] (COx_a, COx_{wa}) were calculated for neutral molecules containing 1 or 2 sulfur atoms (**Figure 1**). COx_a and COx_{wa} were calculated by

(Eq. 1) $$COx_a = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{H-2O-2S}{C}\right)_i$$

(Eq. 2) $$COx_{wa} = \frac{\sum_{i=1}^{n}\left(\frac{H-2O-2S}{C}\right)_i \times w_i}{\sum_{i=1}^{n} w_i}$$

where C, H, O, and S correspond to the number of each individual element in a given molecular formula, *i*, containing only C, H, O, and S; *n* is the number of CHOS formulas, and *w* is the relative abundance (peak intensity) of the molecular formula *i* among CHOS formulas. MW_a and MW_{wa} were determined in the same manner, replacing the carbon oxidation state calculation by molecular weight.

The MW_a and MW_{wa} were both highest in fracturing fluid and early flowback (maximum 380 Da) and lower in later produced waters (minimum 305 Da) (**Tables S1, S2**), with significant regressions in all but the 5H MW_a. The MW_{wa} was generally lower than the MW_a, indicating that higher molecular weight ions were generally of lower intensity, as typically observed in FT-ICR mass spectra of DOM[29]. The MW decrease over time likely indicates a switch from the characteristics of the fluid additives in early samples to fluids dominated by the Marcellus shale connate fluids. Active microbial communities are also likely consuming a portion of the CHOS, contributing to the reduction in molecular weight. However, a truly quantitative assessment is not possible due to possible changes in the matrix and hence likely changes in ion suppression characteristics and the lack of internal standards.

The averaged COx values ranged from -0.12 to -0.77, with lower COx values observed in the weighted average than the unweighted average. This deviation from the unweighted average indicates an abundance of high intensity ions with a more negative (reduced) COx. No clear temporal trend was observed in the 5H well, which may be a function of the limited number of samples analyzed or an actual lack of a temporal trend. The 3H well COx_a decreased over time (slope = -0.0003 d^{-1}, $R^2 = 0.50$, p=0.003), and although the temporal trends observed in the 3H well COx varied substantially, higher initial COx values could be indicative of persulfate application as an oxidative breaker in the 3H well. Persulfate acts as a non-specific oxidizer,

likely resulting in the oxidation of molecules and potentially driving an increase in the COx of organic compounds including those containing sulfur. At *in situ* shale well pH (neutral – slightly acidic), activated persulfate will primarily produce the highly electrophilic sulfate radical which prefers to react with unsaturated molecules and molecules containing amino, hydroxyl, and alkoxy groups[30–33]. Persulfate was applied at higher rates in the 3H well (75 times higher)[4], where a more oxidized COx_a was observed. Although many other production chemicals applied to these wells could contribute to the observed shift, the quantity of persulfate applied prior to flowback is one of the few differences between the 3H and 5H well fracturing fluid chemistry that could contribute to the observed difference in COx_a between the two wells. Alternatively, a decrease in COx could indicate a shift from the more oxidized COx of compounds applied with the hydraulic fracturing fluid to compounds expected in the more reduced fossil fuel environment, or microbial reduction via fermentation[34].

The distribution of $CHOS_1$ ions in early flowback was dominated by *m/z* ions and assigned formulas with high hydrogen to carbon ratios (H/C) (**Figure 2**), and appeared to be unique from the typical distribution of natural organic matter (NOM) which generally centers around an oxygen to carbon ratio (O/C) of ~0.5 and a hydrogen to carbon ratio of ~1 with regard to the number and relative abundance of *m/z* ions [29,35]. $CHOS_1$ ions with a H/C ratio greater than 2 remained high in relative abundance in later samples as well. The distribution of $CHOS_2$ ions was shifted towards a higher O/C ratio than $CHOS_1$ ions, with most ions falling between an O/C ratio of 0.5 and 0.8.

3.2 MSEEL CHOS abundance dominated by saturated organic compounds

To further investigate the observed abundance of $CHOS_1$ ions with a high H/C ratio, the distribution of ions was plotted as a function of double bond equivalency (DBE). The majority of

MSEEL CHOS flowback and produced water *m/z* ions containing carbon, hydrogen, oxygen, and one sulfur atoms fell between a DBE of 0 and 12 (**Figure 3**). A handful of formula assignments were confirmed up to a DBE of 17; formulas with a DBE >17 were removed as probable incorrect formula assignments due to their very high or very low DBE-O value [<-10 or >+10][27]. At each fluid stage, a bell curve of ion abundances was observed between a DBE of 1 and 12. However, at most time points, the dominant OS DBE class was 0, indicating its aliphatic nature. This proportionally high abundance of DBE=0 is not generally observed in diverse types of NOM analyzed using FT-ICR-MS, regardless of ionization source[36–39]. This is also true of NOM associated with shales[40] and a water soluble Utica shale drill cutting extract (**Figure 3a**). Several high intensity OS *m/z* ions were checked for ^{13}C and ^{34}S isotopic patterns and confirmed by precise mass and relative isotopic abundance (e.g., **Table S3**).

These high intensity aliphatic DOS *m/z* ions were present in the hydraulic fracturing fluid samples and were particularly dominant in early flowback, indicating that the aliphatic DOS may have been associated with the hydraulic fracturing fluid additives. Sulfur-containing additives applied to the MSEEL wells included inorganic compounds (ammonium sulfate, diammonium peroxidisulfate, copper sulfate, sodium sulfate) and sulfur containing copolymers (acrylamide with methylpropanesulfonic acid and a thiourea-containing polymer). However, the molecular formulas for these additives or the individual monomers do not match the identified formulas in this study. The high intensity of aliphatic peaks indicates an exogenous source that is likely associated with unlisted additives/additive mixtures.

3.3 Dominant OS species are likely alcohol ethoxysulfates

Both MSEEL wells use ethoxylated alcohol surfactants, some of which are listed as "trade secret" on the FracFocus report. Ethoxylated alcohols (AEO) themselves do not contain

sulfur, but alcohol ethoxysulfates (AES) are a common class of industrial ethoxylated surfactants. Indeed, the measured masses of many high intensity OS ions matched the exact masses of known AES surfactants (<0.2 ppm error). The most abundant classes of the putative AES were those containing one to four ethoxylate groups ($AE_{1-4}S$) (**Figure 4**, **Figures S1-S8**), consistent with a reported European industrial production average of $AE_{2.4}S$[41]. AES in the 5H well had a slightly longer chain length distribution than AES in the 3H well at all time points, likely a function of slightly different chemical batch mixtures injected in to the two wells. These ethoxylated groups also had the longest homologous series chains (separated by CH_2 groups), consistent with the range of alkane change lengths typical of these industrial mixtures[41]. Many AES peaks were among the highest peaks in the overall spectra, particularly in the early flowback fluids. Exact masses matching alkane sulfates (AS) were also identified in substantial abundances, consistent with their known prevalence in AES mixtures (15-45%)[41]. Exact masses matching secondary alkane sulfonates (SAS) were not identified consistently above the baseline nor were these present in consistent homologous series, indicating these surfactants were not likely applied to the wells. The lack of *m/z* ions in DOM that match these DOS signatures and the appearance of plausible homologous series strongly suggest that these high abundance *m/z* ions are likely related to AES. Similar to our findings, other unlisted polyethoxylated alcohols (amino-polyethylene glycols) have also recently been identified in hydraulic fracturing flowback fluids[42].

3.4 Alcohol ethoxysulfate stability

Although AES and AS are readily biodegraded in surface and wastewaters[41,43,44], the prevalence of these probable additives in fluids 10 months after fluid injection suggests that they are not being degraded in the well environment or rapidly degraded in the waste fluids when

returned to the surface. The distribution of ethoxylate chain length in the identified AES decreased slightly over time (**Table S4**), which may indicate some biological degradation occurring in the well. Further inferences on the biological stability of organic sulfur including AES were drawn from metagenomic data collected from the MSEEL wells during the sample time period and made available through the Joint Genome Institute (JGI)[45].

As a key component of many proteins, sulfur is essential for growth in all microorganisms. It can be acquired from the cell's surrounding environment in a variety of organic and inorganic (e.g., sulfate) forms. When present as organic sulfate, extracellular enzymes such as alkyl sulfatase and aryl sulfatase, or via Fe^{2+}-dependent sulfatases can be released from the cell to cleave inorganic sulfate via hydrolase or dioxygenase reactions, respectively[46]. Sulfonates can also be imported across the membrane using specialized ABC transporters[47]. In order to investigate the relationship between organic sulfur compounds and microbial enzymes associated with sulfur cleavage and/or uptake, we queried metagenome and genomic data from the sampled MSEEL wells. Five produced water metagenomes from the 5H well, two produced water metagenomes from the 3H well (spanning Days 70 – 280), and nine bacterial isolate draft genomes (*Halanaerobium spp.*, dominant shale gas late produced water taxa) were identified in the Joint Genome Institute (JGI) database and analyzed using the integrated microbial genomes & microbiomes (IMG/M) system[48] (**Table S5**). Genes encoding for sulfatases, organic sulfur ABC transporters, and sulfate uptake were queried within the JGI IMG/M system usage of either enzyme name or KEGG orthology/E.C. number. A full list of genes queried and raw gene counts are listed in **Tables S6 and S7**. A comparison of raw counts and relative gene counts (value divided by assembled metagenome count $x10^5$) showed similar temporal trends.

In general, most genes associated with using organic sulfur were abundant in early produced waters and less abundant or absent in later produced waters (**Figure S9**). No genes coding for alkyl sulfatases, the expected sulfate cleavage enzymes for the putatively identified AES surfactants, were identified in the metagenomes using E.C. number. However, when searching by name rather than E.C. number, alkyl sulfatase-encoding genes were identified in the metagenomes of two produced water samples during early production. Similarly, genes coding for arylsulfatases were identified in metagenomes collected in the first few months of flowback when searching by E.C. number or gene name. The absence of these extracellular cleavage pathways is consistent with the continued abundance of AES in later produced water, although it is possible that the sulfatases are not well characterized in the prevalent microbes and therefore not readily identified using our search approach. Externally cleaved sulfate produced by sulfatases (and other available sulfate) can subsequently be imported into the cell through active transporters. Genes coding for sulfate ABC transporters were also most prevalent in earlier produced waters, when sulfate concentrations are the highest[49].

During times of sulfate starvation, microorganisms increase production of proteins associated with organic sulfur uptake, including taurine and alkanesulfonate transporters[50,51]. We identified four different organic sulfur transporters with relative abundances decreasing from D-methionine>taurine~alkanesulfonate>cystine. Genes coding for methionine ABC transporters were abundant at all analyzed time points while the other transporters decreased after the first few months of flowback.

Once across the membrane, genes for cleaving sulfite from alkanesulfonates were identified in metagenomes sampled from earlier produced waters but were nearly absent at later dates. In the dominant taxa *Halanaerobium* genomes, only genes for D-methionine and taurine

transport were detected while sulfate transporters and sulfatases were absent. These results indicate that microbial cleavage and uptake of organic sulfur is higher and more diverse in earlier produced waters, consistent with a wide variety of exogenous inputs from fracturing fluid present in the earlier flowback. Higher relative abundance of organic sulfur uptake proteins for methionine and taurine in later produced waters suggests inorganic sulfate limitation and macronutrient recycling within the microbial community as a source of sulfur.

Although AES and AS appear to be somewhat stable at depth, these compounds are known to be readily biodegraded in surface and wastewaters[41,43,44]. The rate of photodegradation of AES is not reported in the literature, but a photo-Fenton treatment strategy for AEO removal in wastewaters indicated rapid mineralization[52]. Preliminary photochemistry and anaerobic biodegradation experiments of hydraulic fracturing fluid components were performed and analyzed using non-target ultrahigh resolution mass spectrometry to determine the possible fate of AES in these complex mixtures. The methods and results of these experiments are detailed in the supplemental materials and the observed results are consistent with the expected rapid photochemical degradation and slower biological degradation under anaerobic conditions. Although AES persist at depth, when flowback and produced fluids are returned to the surface, they are less stable and would not be conserved tracers of the shale well fluids in the event of a surficial spill.

4. Conclusions

A large number of diverse DOS compounds were identified in shale gas wastewaters using an ultrahigh resolution FT-ICR-MS approach intended to capture ionizable and solid-phase extractable molecules. The observed changes in OS substances including the decrease in average

molecular weight and the declining signal attributed to AES surfactants indicated that the produced waters 10 months after the initial fracturing continue to be influenced by the injected hydraulic fracturing fluid although they are increasingly dominated by the connate fluids. AES and other OS surfactants are readily ionized under electrospray ionization[23], resulting in their detection at very low concentrations and in highly complex matrices. AES surfactants may therefore be ideal sensitive tracers of the injected hydraulic fracturing fluid even with substantial dilution. In the subsurface, metagenomic and genomic data indicates AES are not likely being readily degraded as a source of sulfur by dominant taxa. Understanding the molecular characteristics of DOS and their physiochemical and biological fates in shale wells is critical in addressing questions of effective hydraulic fracturing engineering and biogeochemical cycling within the well. However, the diversity of DOS needs further investigation to determine if the identified characteristics and trends are unique to the analyzed wells or represent broader patterns found in other formations and under other operating conditions. Additional analyses of volatile compounds (e.g., GCxGC-TOF-MS) and the hydrophilic organic fraction of these fluids are important for providing a more complete understanding of the dynamics of OS in hydraulic fracturing systems.

Conflicts of Interest

The authors have no conflicts to declare.

Acknowledgements

We thank the Marcellus Shale Energy and Environment Laboratory for providing site access and samples and Mary Evert, Andrea Hanson, and Katie Heyob for assistance with obtaining

samples. We thank the Analytical Biogeochemistry research group at the Helmholtz Zentrum Muenchen for assistance with FT-ICR-MS analyses. The genome and metagenome sequence data were produced by the US Department of Energy Joint Genome Institute http://www.jgi.doe.gov/ in collaboration with the user community. We thank Kelly Wrighton and Rebecca Daly for providing MSEEL metagenome samples through the JGI database, and are grateful for the feedback from our reviewers which improved this manuscript. Funding was provided by NSF CBET Award #1604475/1823069 to MG and PJM and NSF MRI Award #1039768. This is contribution XXXX of the University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory.

References

1 A. J. Kondash, E. Albright and A. Vengosh, *Sci. Total Environ.*, 2017, **574**, 314–321.
2 K. Oetjen, K. E. Chan, K. Gulmark, J. H. Christensen, J. Blotevogel, T. Borch, J. R. Spear, T. Y. Cath and C. P. Higgins, *Sci. Total Environ.*, 2018, **619-620**, 654–664.
3 J. Rosenblum, A. W. Nelson, B. Ruyle, M. K. Schultz, J. N. Ryan and K. G. Linden, *Sci. Total Environ.*, 2017, **596-597**, 369–377.
4 J. L. Luek, M. Harir, P. Schmitt-Kopplin, P. J. Mouser and M. Gonsior, *Water Res.*, 2018, **136**, 200–206.
5 A. E. Booker, M. A. Borton, R. A. Daly, S. A. Welch, C. D. Nicora, D. W. Hoyt, T. Wilson, S. O. Purvine, R. A. Wolfe, S. Sharma, P. J. Mouser, D. R. Cole, M. S. Lipton, K. C. Wrighton and M. J. Wilkins, *mSphere*, 2017, **2**, e00257–17.
6 A. N. Paukert Vankeuren, J. A. Hakala, K. Jarvis and J. E. Moore, *Environ. Sci. Technol.*, 2017, **51**, 9391–9402.
7 U.S. EPA, *Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Final Report).*, U.S. Environmental Protection Agency, Washington, DC, 2016.
8 W. T. Stringfellow, J. K. Domen, M. K. Camarillo, W. L. Sandelin and S. Borglin, *J. Hazard. Mater.*, 2014, **275**, 37–54.
9 M. Elsner and K. Hoelzer, *Environ. Sci. Technol.*, 2016, **50**, 3290–3314.
10 R. A. Marriott, P. Pirzadeh, J. J. Marrugo-Hernandez and S. Raval, *Can. J. Chem.*, 2016, **94**, 406–413.
11 C. He, T. Zhang and R. D. Vidic, *Disruptive Sci. Technol.*, 2013, **1**, 169–176.
12 G. P. Thiel and J. H. Lienhard, *Desalination*, 2014, **346**, 54–69.
13 T. Hayes, *Sampling and Analysis of Water Streams Associated with the Development of Marcellus Shale Gas*, Gas Technology Institute, 2009.
14 I. Ferrer and E. M. Thurman, *Anal. Bioanal. Chem.*, 2015, **407**, 6417–6428.
15 W. Orem, C. Tatu, M. Varonka, H. Lerch, A. Bates, M. Engle, L. Crosby and J. McIntosh, *Int. J. Coal Geol.*, 2014, **126**, 20–31.
16 S. J. Maguire-Boyle and A. R. Barron, *Env. Sci Process. Impacts*, 2014, **16**, 2237–2248.
17 P. Pirzadeh, S. Raval and R. A. Marriott, *Org. Geochem.*, 2015, **83-84**, 94–100.
18 P. Pirzadeh, K. L. Lesage and R. A. Marriott, *Energy Fuels*, 2014, **28**, 4993–5001.
19 W. Orem, M. Varonka, L. Crosby, K. Haase, K. Loftin, M. Hladik, D. M. Akob, C. Tatu, A. Mumford, J. Jaeschke, A. Bates, T. Schell and I. Cozzarelli, *Appl. Geochem.*, 2017, **80**, 155–167.
20 G. J. Getzinger, M. P. O'Connor, K. Hoelzer, B. D. Drollette, O. Karatum, M. A. Deshusses, P. L. Ferguson, M. Elsner and D. L. Plata, *Environ. Sci. Technol.*, 2015, **49**, 8347–8355.
21 M. L. Trehy, W. E. Gledhill, J. P. Mieure, J. E. Adamove, A. M. Nielsen, H. O. Perkins and W. S. Eckhoff, *Environ. Toxicol. Chem.*, 1996, **15**, 233–240.
22 A. Di Corcia, L. Capuani, F. Casassa, A. Marcomini and R. Samperi, *Environ. Sci. Technol.*, 1999, **33**, 4119–4125.
23 M. Gonsior, M. Zwartjes, W. J. Cooper, W. Song, K. P. Ishida, L. Y. Tseng, M. K. Jeung, D. Rosso, N. Hertkorn and P. Schmitt-Kopplin, *Water Res.*, 2011, **45**, 2943–2953.
24 G. Morales-Cid, I. Gebefugi, B. Kanawati, M. Harir, N. Hertkorn, R. Rosselló-Mora and P. Schmitt-Kopplin, *Anal. Bioanal. Chem.*, 2009, **395**, 797–807.
25 S. E.-M. Dvorski, M. Gonsior, N. Hertkorn, J. Uhl, H. Müller, C. Griebler and P. Schmitt-Kopplin, *Environ. Sci. Technol.*, 2016, **50**, 5536–5546.

26 J. L. Luek, P. Schmitt-Kopplin, P. J. Mouser, W. T. Petty, S. D. Richardson and M. Gonsior, *Environ. Sci. Technol.*, 2017, **51**, 5377–5385.
27 P. Herzsprung, N. Hertkorn, W. von Tümpling, M. Harir, K. Friese and P. Schmitt-Kopplin, *Anal. Bioanal. Chem.*, 2014, **406**, 7977–7987.
28 J. H. Kroll, N. M. Donahue, J. L. Jimenez, S. H. Kessler, M. R. Canagaratna, K. R. Wilson, K. E. Altieri, L. R. Mazzoleni, A. S. Wozniak, H. Bluhm, E. R. Mysak, J. D. Smith, C. E. Kolb and D. R. Worsnop, *Nat. Chem.*, 2011, **3**, 133–139.
29 N. Hertkorn, M. Frommberger, M. Witt, B. P. Koch, P. Schmitt-Kopplin and E. M. Perdue, *Anal. Chem.*, 2008, **80**, 8908–8919.
30 C. Liang and H.-W. Su, *Ind. Eng. Chem. Res.*, 2009, **48**, 5558–5562.
31 L. W. Matzek and K. E. Carter, *Chemosphere*, 2016, **151**, 178–188.
32 P. Neta, V. Madhavan, H. Zemel and R. W. Fessenden, *J. Am. Chem. Soc.*, 1977, **99**, 163–164.
33 W.-D. Oh, Z. Dong and T.-T. Lim, *Appl. Catal. B Environ.*, 2016, **194**, 169–201.
34 J. Valle, M. Gonsior, M. Harir, A. Enrich-Prast, P. Schmitt-Kopplin, D. Bastviken, R. Conrad and N. Hertkorn, *Water Res.*, 2018, **129**, 252–263.
35 M. Gonsior, J. Valle, P. Schmitt-Kopplin, N. Hertkorn, D. Bastviken, J. Luek, M. Harir, W. Bastos and A. Enrich-Prast, *Biogeosciences*, 2016, **13**, 4279–4290.
36 H. P. Dias, T. M. C. Pereira, G. Vanini, P. V. Dixini, V. G. Celante, E. V. R. Castro, B. G. Vaz, F. P. Fleming, A. O. Gomes, G. M. F. V. Aquije and W. Romão, *Fuel*, 2014, **126**, 85–95.
37 V. Carré, S. Schramm and F. Aubriet, *AIMS Environ. Sci.*, 2015, **2**, 547–564.
38 E. Bae, I. J. Yeo, B. Jeong, Y. Shin, K.-H. Shin and S. Kim, *Anal. Chem.*, 2011, **83**, 4193–4199.
39 T. M. C. Pereira, G. Vanini, E. C. S. Oliveira, F. M. R. Cardoso, F. P. Fleming, A. C. Neto, V. Lacerda, E. V. R. Castro, B. G. Vaz and W. Romão, *Fuel*, 2014, **118**, 348–357.
40 E. Bae, J.-G. Na, S. H. Chung, H. S. Kim and S. Kim, *Energy Fuels*, 2010, **24**, 2563–2569.
41 *Alcohol Ethoxysulphates (AES) Environmental Risk Assessment*, HERA, 2004.
42 K. A. Sitterley, K. G. Linden, I. Ferrer and E. M. Thurman, *Anal. Chem.*, , DOI:10.1021/acs.analchem.8b02439.
43 G.-G. Ying, *Environ. Int.*, 2006, **32**, 417–431.
44 D. C. McAvoy, S. D. Dyer, N. J. Fendinger, W. S. Eckhoff, D. L. Lawrence and W. M. Begley, *Environ. Toxicol. Chem.*, 1998, **17**, 1705–1711.
45 H. Nordberg, M. Cantor, S. Dusheyko, S. Hua, A. Poliakov, I. Shabalov, T. Smirnova, I. V. Grigoriev and I. Dubchak, *Nucleic Acids Res.*, 2014, **42**, D26–D31.
46 M. Toesch, M. Schober and K. Faber, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 1485–1496.
47 M. A. Kertesz, *Res. Microbiol.*, 2001, **152**, 279–290.
48 I.-M. A. Chen, V. M. Markowitz, K. Chu, K. Palaniappan, E. Szeto, M. Pillay, A. Ratner, J. Huang, E. Andersen, M. Huntemann, N. Varghese, M. Hadjithomas, K. Tennessen, T. Nielsen, N. N. Ivanova and N. C. Kyrpides, *Nucleic Acids Res.*, 2017, **45**, D507–D516.
49 M. V. Evans, J. Panescu, A. J. Hanson, J. Sheets, S. A. Welch, N. Natasi, R. A. Daly, D. R. Cole, T. H. Darrah, M. J. Wilkins, K. C. Wrighton and P. J. Mouser, *Rev.*
50 J. van der Ploeg, E. Eichhorn and T. Leisinger, *Arch. Microbiol.*, 2001, **176**, 1–8.
51 J. R. van der Ploeg, M. A. Weiss, E. Saller, H. Nashimoto, N. Saito, M. A. Kertesz and T. Leisinger, *J. Bacteriol.*, 1996, **178**, 5438–5446.
52 S. S. da Silva, O. Chiavone-Filho, E. L. de Barros Neto, A. L. N. Mota, E. L. Foletto and C. A. O. Nascimento, *Environ. Technol.*, 2014, **35**, 1556–1564.

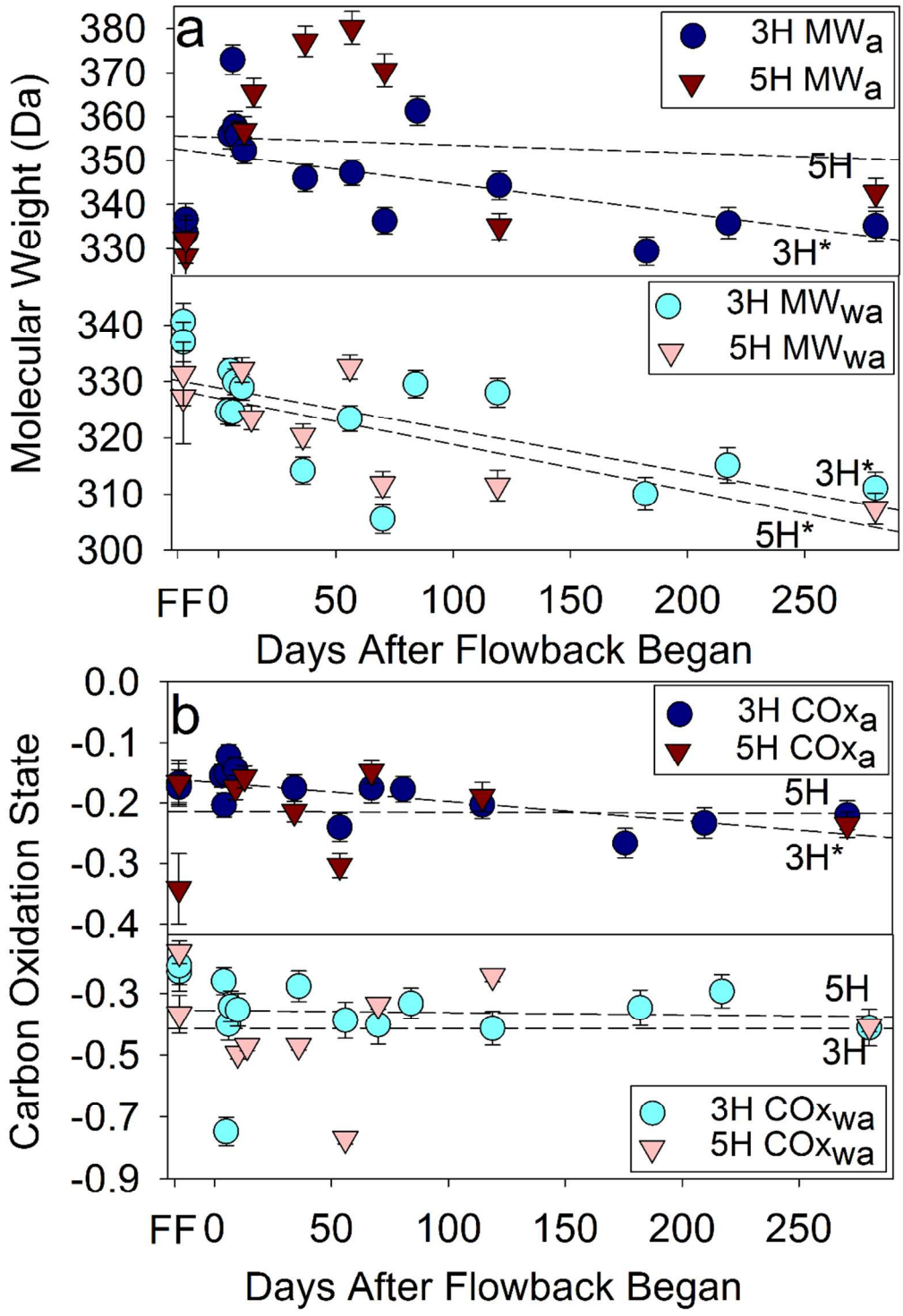


Figure 1. Average molecular weight (a) and carbon oxidation state (b) of fracturing fluid (FF), flowback, and produced waters from the 3H and 5H MSEEL wells with linear regressions applied. Error bars represent standard error of the mean. * indicates significant slope ($p<0.05$), FF included in calculation as day 0.

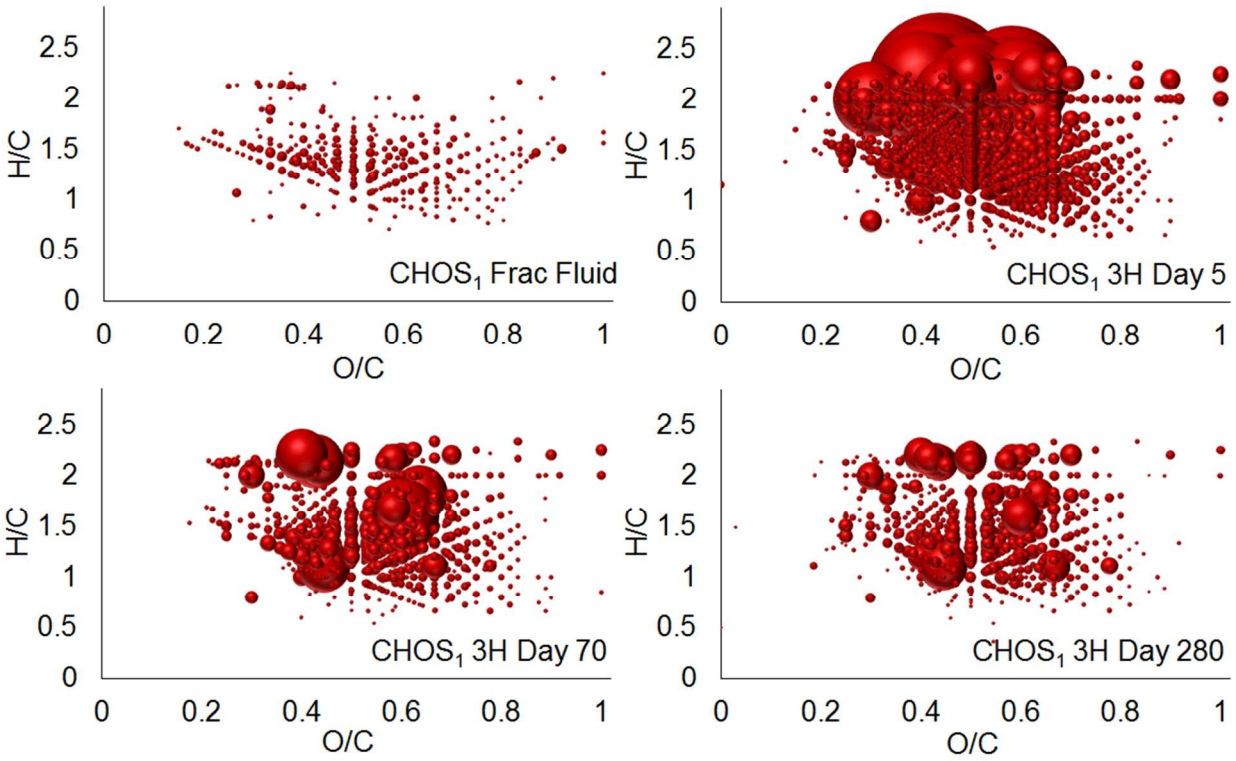


Figure 2. Distribution of OS ions observed in MSEEL 3H well fracturing fluid (a) and in flowback and produced waters obtained on Day 5 (b) Day 70 (c) and Day 280 (d). Bubble sizes are scaled by relative abundance of highest CHOS peak on Day 5 for comparison across plots.

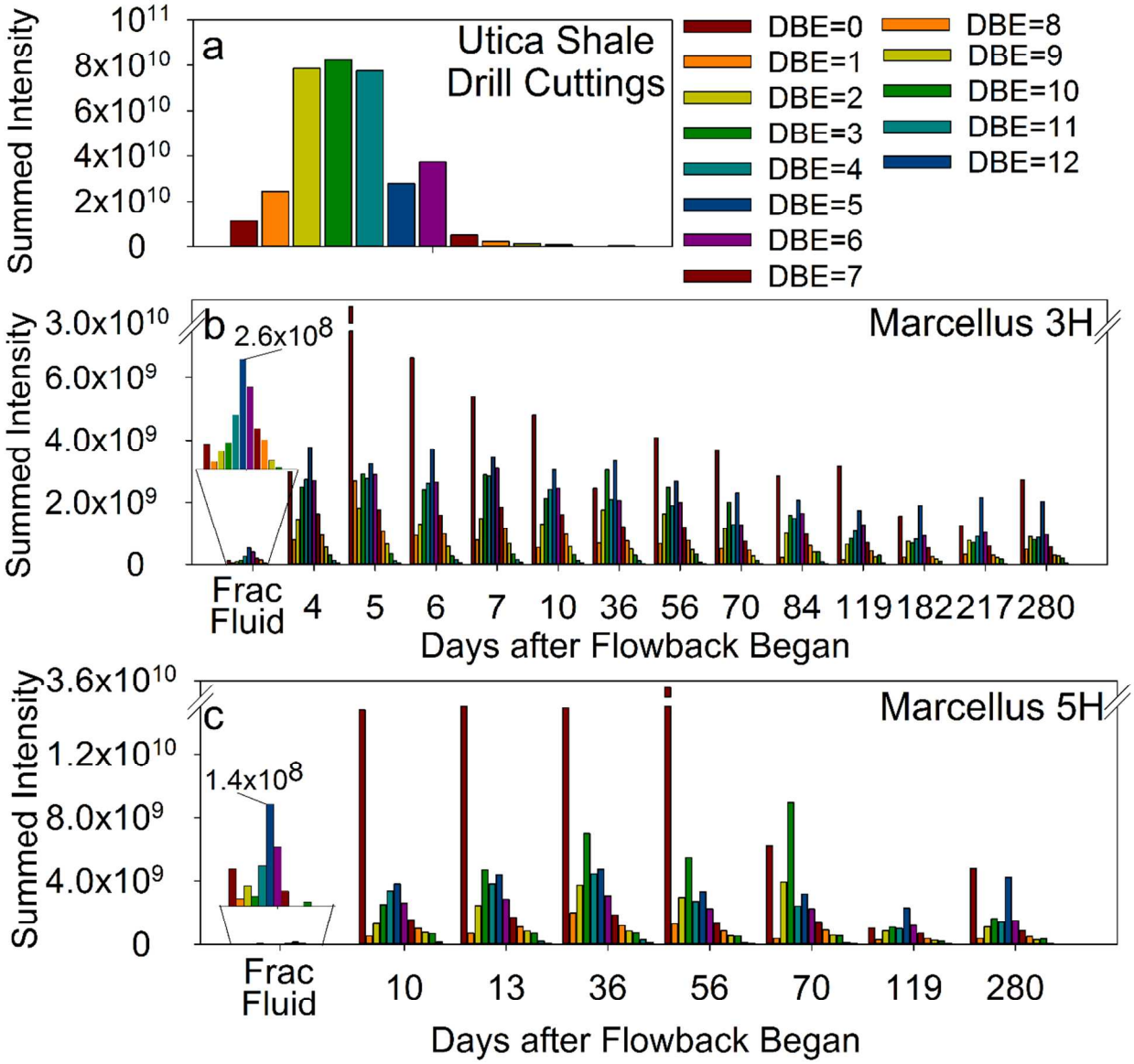


Figure 3. Summed abundance of $CHOS_1$ peaks corresponding to a DBE values between 0 and 12 in a) a water soluble Utica shale drill cutting extract b) MSEEL 3H well samples and c) MSEEL 5H well samples according to days after flowback began.

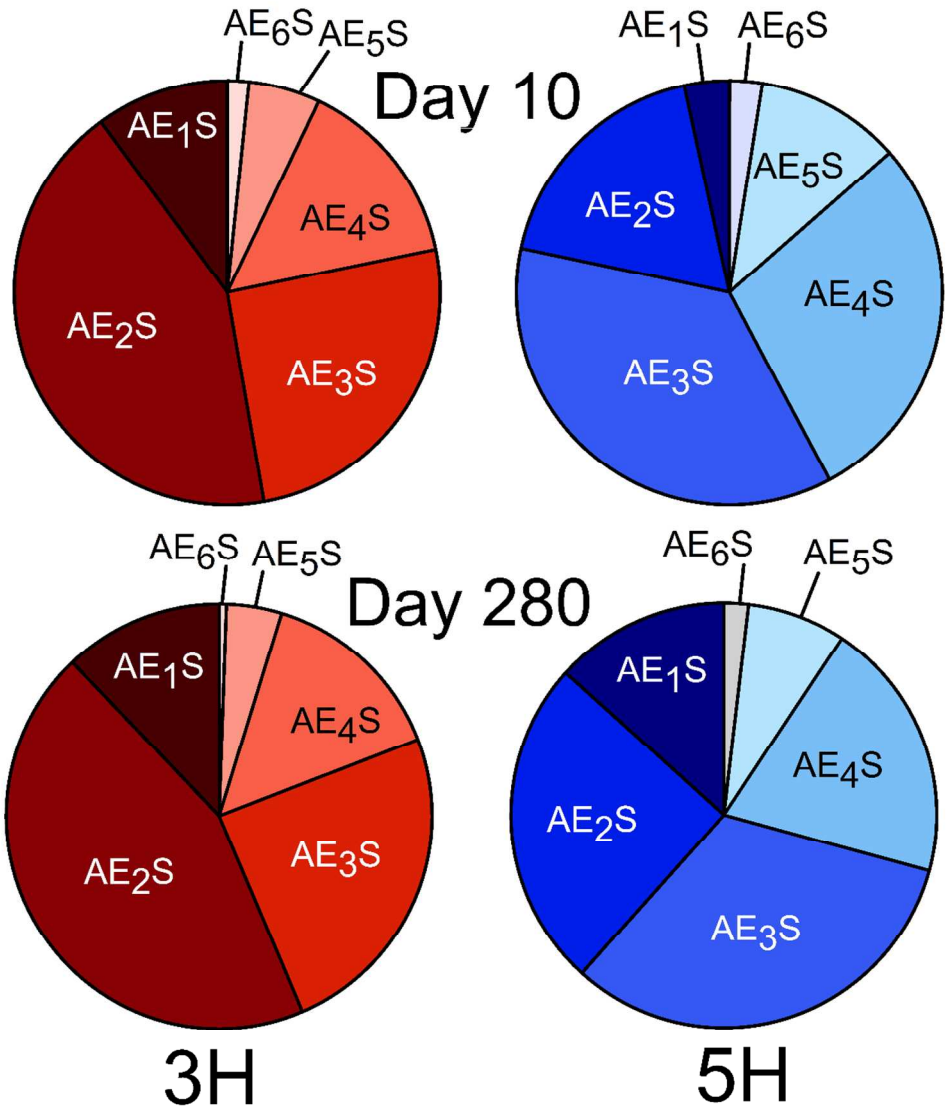
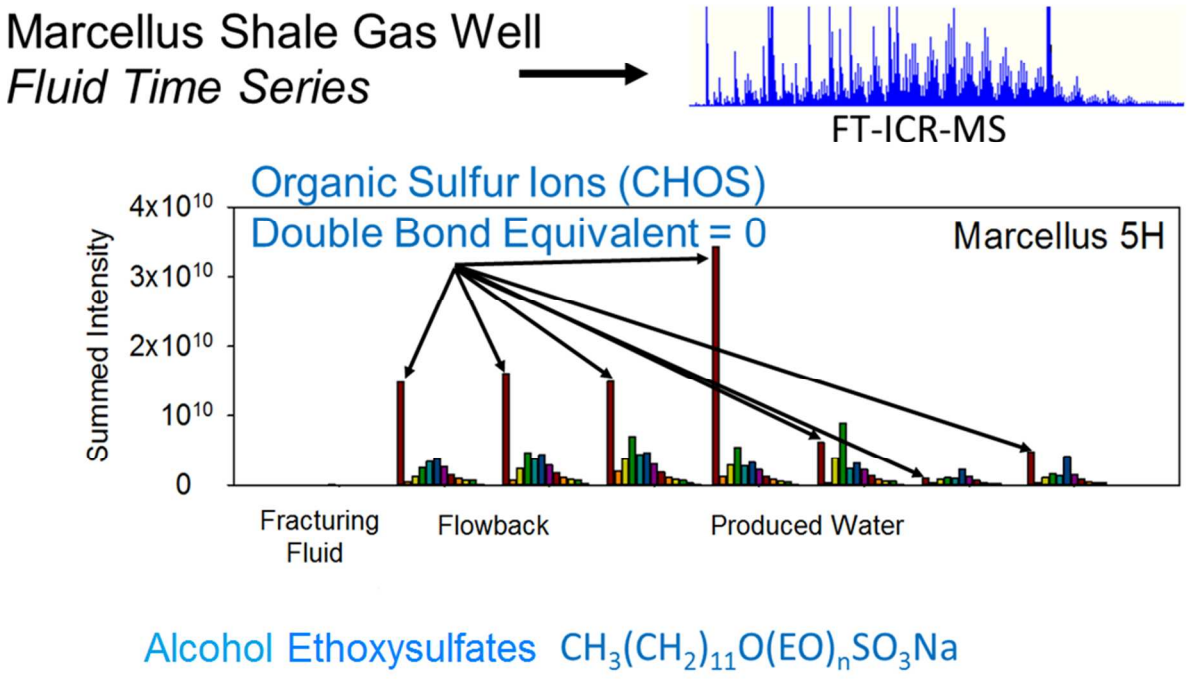


Figure 4. Distribution of summed abundances putative alcohol ethoxysulfate ion classes on Day 10 (top) and Day 280 (bottom) of flowback/produced water for the 3H (left) and 5H (right) wells. AE_7S and AE_8S were identified but contributed less than 0.01% and are therefore not shown.

Table of contents entry

Image:



Text: Ultrahigh resolution mass spectrometry used to identify unique organic sulfur signatures in hydraulic fracturing wastewaters likely associated with alcohol ethoxysulfate surfactants.