Hazard/Risk Assessment

Human and Aquatic Toxicity Potential of Petroleum Biodegradation Metabolite Mixtures in Groundwater from Fuel Release Sites

Timothy J. Patterson,¹,* Lauren Kristofco,² Asheesh K. Tiwary,³ Renae I. Magaw,³ Dawn A. Zemo,² Kirk T. O’Reilly,⁴ Rachel E. Mohler,² Sungwoo Ahn,⁵ Natasha Sihota,⁴ and Catalina Espino Devine⁴

¹Chevron Energy Technology, San Ramon, California, USA
²Zemo & Associates, Incline Village, Nevada, USA
³Exponent, Bellevue, Washington, USA
⁴Chevron Energy Technology, Richmond, California, USA

Abstract: The potential toxicity to human and aquatic receptors of petroleum fuel biodegradation metabolites (oxygen-containing organic compounds [OCOCs]) in groundwater has been investigated as part of a multi-year research program. Whole mixtures collected from locations upgradient and downgradient of multiple fuel release sites were tested using: 1) in vitro screening assays for human genotoxicity (the gamma-H2AX assay) and estrogenic effects (estrogen receptor transcriptional activation assay), and 2) chronic aquatic toxicity tests in 3 species (Ceriodaphnia dubia, Raphidocelis subcapitata, and Pimephales promelas). In vitro screening assay results demonstrated that the mixtures did not cause genotoxic or estrogenic effects. No OCOC-related aquatic toxicity was observed and when aquatic toxicity did occur, upgradient samples typically had the same response as samples downgradient of the release, indicating that background water quality was impacting the results. This information provides additional support for previous work that focused on the individual compounds and, taken together, indicates that OCOCs from petroleum degradation at fuel release sites are unlikely to cause toxicity to human or freshwater receptors at the concentrations present. Environmental Toxicology and Chemistry published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Biotransformation; Hazard/risk assessment; Mixture toxicology; Polar metabolites; Petroleum; Oxygen-containing organic compounds (OCOCs)

INTRODUCTION

Environmental assessment and management of petroleum release sites with hydrocarbon-impacted groundwater is a topic with ongoing interest from regulators, academia, and industry (Interstate Technology Regulatory Council 2018). The natural attenuation including biodegradation of petroleum fuels has been an accepted regulatory strategy for managing petroleum releases and is based on the concepts of decreasing mass and toxicity associated with these releases over time (Wiedemeier et al. 1995; US Environmental Protection Agency 1999, 2011a; Interstate Technology Regulatory Council 2018). Recently the appropriateness of this strategy has received renewed attention.

The most appropriate term to use for compounds found associated with hydrocarbon biodegradation is subject to debate. Our group initially used “polar” or “petroleum metabolites” to represent these compounds (Mohler et al. 2013; Zemo et al. 2013). Others introduced terms including petroleum hydrocarbon oxidation product (San Francisco Bay Regional Water Quality Control Board 2019) and petroleum-derived dissolved organic matter (Podgorski et al. 2018). Each of these terms suggests that the compounds are produced through the biodegradation of petroleum. Recent studies propose that these mixtures may contain both oxidized hydrocarbons and compounds synthesized as the result of in situ microbial activity (Mohler et al. 2013; O’Reilly et al. 2019). We now use the term oxygen-containing organic compounds (OCOCs) because it
describes their chemistry without making assumptions about their biological source (O’Reilly et al. 2019).

Oxygen-containing organic compounds are a complex mixture that potentially contains thousands of compounds. Regulatory bodies recommend 3 general approaches for evaluating the toxicity of mixtures: test the whole mixture, test similar mixtures, and/or evaluate the component toxicity (US Environmental Protection Agency 1986, 2000; European Commission 2012; Organisation for Economic Co-operation and Development 2018a). Previous studies from our research program used a component toxicity approach for evaluating the toxicity and potential health risks associated with metabolites of petroleum biodegradation found in groundwater at fuel release sites (Zemo et al. 2013, 2017). In summary, targeted and nontargeted chemical analyses were used to identify potential hydrocarbon degradation intermediates including alcohols, phenols, ketones, aldehydes, esters, and acids (Mohler et al. 2013). According to their chemical structure these were compiled into 5 chemical families and 22 structural groups. A toxicity ranking system consistent with those used by the US Environmental Protection Agency (USEPA) and the United Nations was developed for evaluating the potential chronic oral toxicity to humans of the different chemical classes (Zemo et al. 2013). Most of the identified OCOCs are expected to have a low chronic toxicity profile. In line with these findings, reference dose-based concentration targets for screening have been proposed for groundwater at various life-cycle stages of a biodegrading hydrocarbon plume at petroleum release sites (Zemo et al. 2017).

The present study uses a whole-mixture approach by evaluating the potential toxicity of the complex mixture of OCOCs present downgradient from biodegrading petroleum fuel releases, regardless of whether the OCOCs are biodegradation intermediates or background organic compounds in groundwater including those synthesized by microbes. Groundwater samples were collected from upgradient locations representing local background and downgradient locations (intended to contain petroleum metabolites but not the original petroleum fuel hydrocarbons) at 14 fuel-terminal sites where petroleum releases are undergoing biodegradation.

Traditional toxicological approaches to evaluate chronic human health toxicity endpoints involve animal testing and were not feasible for the present study. Also, no single regulatory acceptable in vitro toxicity study for comprehensive human health screening is currently available. Therefore, to evaluate the potential human health effects of the complex mixture of OCOCs in groundwater samples at fuel release sites and to support the literature-based toxicity-ranking approach presented previously (Zemo et al. 2013, 2017) 2 commercially available high-throughput in vitro assays were used to screen for estrogenic and genotoxic effects.

Genotoxicity was selected as an endpoint for additional study to complement the reference dose ranking that was based on repeat dose (chronic) systemic toxicity data (Zemo et al. 2013, 2017). Genotoxicity may indicate that a substance has a potential for carcinogenic effects. To evaluate genotoxicity potential, the gamma-H2AX assay was used. This is a sensitive and reliable high-throughput method for detecting double-strand DNA breaks (Mah et al. 2010; Tsamou et al. 2012; Ando et al. 2014; Nikolova et al. 2014) that utilizes the accumulation of gamma-H2AX in the cell as an endpoint for determining DNA damage with and without metabolic activation (Watters et al. 2009; Smart et al. 2011; Khoury et al. 2013). Cytotoxicity is also assessed to assist in interpreting whether a response is truly genotoxic (Khoury et al. 2013). It has a comparable sensitivity, specificity, and concordance to the Ames assay, as well as other in vitro genotoxicity tests such as the mouse lymphoma chromosomal aberration assay (Nikolova et al. 2014).

Endocrine disruption is an ongoing environmental and human health concern and screening assays are increasingly being used to detect and monitor potential estrogenic activity in surface water and wastewater effluents (Colborn et al. 1993; Purdom et al. 1994; Filby et al. 2007; Wehmas et al. 2011; US Environmental Protection Agency 2011b; Barber et al. 2012; Rottrof et al. 2013; Dreier et al. 2015). Endocrine disruption was a data gap in the previous reference-dose approach evaluating OCOCs (Zemo et al. 2013, 2017). Therefore, to evaluate the potential for endocrine activity, the estrogen receptor transcriptional activation (ERTA) assay in the human HeLa-9903 cell line was used. This is an Organisation for Economic Co-operation and Development guideline in vitro screening assay for identifying compounds that activate or inhibit the estrogen receptor (Organisation for Economic Co-operation and Development 2018b). It is considered a Level 2 study defined as "in vitro assays providing data about selected endocrine mechanism(s)/pathway(s).” Level 2 is the highest tier before in vivo studies. This assay is designed to evaluate the molecular initiating event associated with activation of the estrogen receptor and is suitable for screening chemicals; it is considered a Tier 1 assay in the estrogen-activation adverse outcome pathway and used for prioritizing chemicals for additional study (US Environmental Protection Agency 2014). Furthermore, it is recommended as a screening tool for water quality by the California State Water Resources Control Board (Drewes et al. 2018).

Because metabolite plumes containing OCOCs may reach downgradient surface water, understanding their risk to target organisms is critical. The aquatic toxicity of groundwater containing OCOCs has been questioned by some regulatory agencies (San Francisco Bay Regional Water Quality Control Board 2016; Interstate Technology Regulatory Council 2018). In addition, ecological screening and target clean-up levels for OCOCs have been or are being developed (San Francisco Bay Regional Water Quality Control Board 2019; Washington State Department of Ecology 2019). To evaluate aquatic toxicity, standard USEPA whole-effluent toxicity tests were conducted with 3 freshwater species (Ceriodaphnia dubia, Raphidocelis subcapitata, and Pimephales promelas). These species were selected because the fuel release sites were not upgradient of marine waters. Such tests are often required as permitting conditions for managing industrial discharges to water bodies (Norberg-King et al. 2018). These studies are characterized as short-term methods for estimating chronic toxicity (US Environmental Protection Agency 2002).

A better understanding of the human and ecological toxicity of OCOCs from petroleum biodegradation is critical for determining how to manage fuel release sites so that informed risk-based decisions can be made. The data obtained in the
present study provides novel information on the potential toxicity of whole groundwater from a large sample size of fuel release sites that can be used with existing evidence to inform such decisions.

**MATERIALS AND METHODS**

**Groundwater sample collection and analytical chemistry**

The samples for the present study were collected from 14 fuel-marketing terminals in California. The terminals were a combination of active and closed facilities and distributed both gasoline and diesel fuels (see Zemo et al. 2013, 2017 for additional details). Release dates and source area mass are unknown. Most plumes have been monitored for at least 10 yr and the plumes are stable or shrinking. The plumes had detections of diesel-range organics but no detections of gasoline-range organics. Monitoring wells at these sites were previously identified as either being upgradient of the release within the hydrocarbon plume, or downgradient of the hydrocarbon plume but within the metabolite plume. Wells within the hydrocarbon plume had detectable diesel-range organics following silica gel clean-up of the solvent extract (Zemo et al. 2013, 2017). Downgradient samples had detectable diesel-range organics without silica gel clean-up but were nondetect after silica gel clean-up indicating that the groundwater contained OCOCs but not dissolved hydrocarbons. Samples with dissolved hydrocarbon were not tested because the purpose of the present study was to evaluate the toxicity of OCOCs. Upgradient and downgradient locations were selected based on hydraulic gradient (i.e., groundwater flow direction) and analytical results from years of routine groundwater monitoring at these sites.

Groundwater samples from each well were collected for analytical testing into unpreserved 1-L amber bottles and shipped directly to the commercial testing laboratories. A total of 39 samples were collected from 12 terminals in 2015 for testing in in vitro screening assays for genotoxic and estrogenic effects; and a total of 36 samples were collected in 2013 and 2014 from 14 terminals for aquatic toxicity testing. Sample information is summarized in Supplemental Data, Table S1.

Analytical characterization was conducted using targeted and nontargeted methods: diesel-range organics with and without silica gel clean-up by gas chromatography–flame ionization detection (USEPA Methods 8015B and 3630C); 76 target compounds by gas chromatography–mass spectrometry (GC–MS; Supplemental Data, Table S3; modified USEPA Method 8270C); nontargeted GC–MS library search; and nontargeted comprehensive two-dimensional GC with time-of-flight mass spectrometry. Target chemicals were selected based on availability of regulatory-derived reference doses and existing data in the scientific literature as described in previous work (Zemo et al. 2017). Chemical analyses were performed on methylene chloride extracts (USEPA Method 3510) of each groundwater sample. Methanol eluates from the silica column were also analyzed by all methods except for diesel-range organics with and without silica gel clean-up. All analytical methods are described in Zemo et al. 2013, Mohler et al. 2013, and Zemo et al. 2017. For reporting the two-dimensional GC results for each sample, the tentatively identified compounds in the extracts and silica column eluates were combined and any replicate tentatively identified compounds were removed.

For the purpose of this discussion the extractable metabolite concentration containing OCOCs in each groundwater sample is defined as the difference in the concentration between diesel-range organics with silica gel clean-up and diesel-range organics without silica gel clean-up.

**In vitro screening assays for genotoxic and estrogenic effects**

All in vitro assays for the present study were performed by Cyprotex Laboratories (Kalamazoo, MI, USA) using their in-house developed assay protocols. Groundwater samples were collected from upgradient (1 L) and downgradient (1 L) locations at 12 fuel terminal sites and sent directly to Cyprotex for testing. Seven of the 12 sites were re-sampled approximately 6 to 7 mo after first sampling and tested to demonstrate reproducibility. Sample dates are in Supplemental Data, Table S1.

**Gamma-H2AX screening assay.** Groundwater samples were filter sterilized using 0.22-micron polytetrafluoroethylene (PTFE) filters into 2 separate vials. Ten doses were used starting at 50% groundwater (maximum concentration that allows for required cell culture media concentrations) and serial diluted with sterile reverse osmosis water by 0.5 for a final dosing regimen of 50, 25, 12, 6.2, 3.1, 1.5, 0.78, 0.39, 0.19, and 0.097%. One of the vials was spiked with a 10-mM stock of camptothecin (CAS 7689-03-4) as a positive control in each groundwater sample as well as in the negative control (sterile reverse osmosis water) for a final top dosing concentration of 1 μM; the remaining dilutions followed the same progression as the groundwater dosing (10 doses following a progression of 0.5-fold each). Each groundwater sample was divided and tested in triplicate. Both positive and vehicle controls responded as expected, indicating the assay was valid and that the groundwater was not altering the assay sensitivity or reliability. To avoid false-positive genotoxic signals caused by cell apoptosis or necrosis, Cyprotex followed their standard practice of qualifying a compound or mixture as genotoxic only if the measured cell viability was above 70% and the induction of H2AX phosphorylation was at least 20% higher than the control.

The minimum effective concentration (MEC) was calculated when DNA damage occurred and is defined as the point where the curves in the cell loss and the H2AX phosphorylation cross the vehicle control threshold (significance line). The test is considered positive when the response points continue to increase or decrease (or at least stay consistently above or below); and cytotoxicity is not observed at or below the MEC for the H2AX phosphorylation (see example for determining
ERTA screening assay. The ERTA assay using the HeLa-9903 cell line (USEPA Office of Prevention, Pesticides and Toxic Substances 890.1300) was run to assess potential estrogenic effects. Groundwater samples were sterilized using 0.22-micron PTFE filtration. Dilutions of the filtrate were prepared in sterile water for a total of 7 half-log concentrations beginning with a 1:2 dilution in 2x culture media for a final dosing of 50, 16, 5.0, 1.6, 0.5, and 0.05%. A top dose of 50% groundwater was the result of assay conditions requiring ample cell culture media.

Each exposure concentration was performed in replicates of 6. Several control groups were included on each plate: vehicle control (0.1% dimethyl sulfoxide), strong agonist control (1-nM estradiol [E2]), weak agonist control (4-cumylphenol), strong antagonist (ICI-182, 780), and negative control (corticosterone). For all plates, the amount of stock solution solvent (dimethyl sulfoxide) was held constant at 0.1% for the 24-h assay period. Luminescence was measured with a Packard TopCount NXT luminescence counter. All processed data were examined to determine whether negative and positive induction controls within each plate were within acceptable limits. Cell viability was monitored by propidium iodide uptake after a 24-h incubation with the test material. A 20% drop below vehicle-treated controls was considered evidence of cytotoxicity. Solubility and precipitation were determined by a light-scattering procedure using nephelometry with a signal ≥80% of the transmittance standard as the criterion.

The acceptance criteria used were: background value ratio of vehicle control to antagonist control should be less than 10-fold, and the ratio of positive control to vehicle control should be greater than 3-fold. Each data point was normalized to the average of the vehicle-only-treated control with the average 1-nM E2 response set at 100% (percentage of induction). Groundwater samples were considered positive for agonism if the maximum differential from the vehicle control was found to be above 20% with the target 1-nM E2 response of approximately 100% in the absence of confounding factors such as cytotoxicity and/or insolubility. Additional methodological details including example figures of the dose–response gene expression results and data for the control are provided in the Supplemental Data, Section 4.

Freshwater aquatic toxicity testing

An in vivo freshwater aquatic testing program was conducted by Pacific EcoRisk (Fairfield, CA, USA) using 3 freshwater species as described in Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Water to Freshwater Organisms (US Environmental Protection Agency 2002). The test program included: 1) a 96-h growth test using green algae R. subcapitata (formerly Selenastrum capricornutum; USEPA Method 1003), which is considered a chronic study because it covers multiple life cycles; 2) a 3-brood (6–8 d) survival and reproduction test using the crustacean C. dubia (USEPA Method 1002), also considered a chronic study because it evaluates multiple life cycles from neonate through reproduction; and 3) a 7-d survival and growth test using larval fathead minnows P. promelas (USEPA Method 1000), which is suitable for predicting chronic toxicity per the US Environmental Protection Agency (2002).

Groundwater samples were collected from monitoring wells at 14 fuel-terminal sites. Two sites were resampled twice (2 mo and 12 mo after first sampling) to investigate repeatability in the aquatic toxicity assays. For each site, testing was conducted using undiluted groundwater from an upgradient location and a downgradient location. Twelve liters of samples from each well were sent to the testing laboratory. All sample handling and testing was conducted consistent with USEPA guidance, including initiating aquatic testing within 15 h of collection (US Environmental Protection Agency 2002). On receipt at the aquatic testing laboratory, multiple subsamples were separated for quantitative analysis of water quality characteristics; the samples were stored at 0 to 6°C until used for testing.

The groundwater ionic matrix of each sample was analyzed to predict the impact on toxicity using the Gas Research Institute freshwater salinity toxicity relationship model. Anions and cations measured were sodium, potassium, calcium, magnesium, chloride, sulfate, and bicarbonate. Aqueous anions were analyzed using USEPA Method 300.0; alkalinity (as bicarbonate) was analyzed using Standard Method 2320B (alkalinity by titration); and all analyses were performed by a state-certified commercial analytical laboratory (Calsci Environmental Laboratories, Garden Grove, CA, USA). The freshwater salinity toxicity relationship model predicts the C. dubia 48-h% survival and the P. promelas 96-h% survival (Tietge et al. 1994). The predicted toxicity responses were compared with the observed mean 48-h and 96-h survival responses from the chronic C. dubia and P. promelas tests, respectively.

Select groundwater samples for which toxicity (both upgradient and downgradient) was observed were analyzed for dissolved metals to determine their potential impact on the observed toxicity. Dissolved metals were analyzed by a commercial laboratory (Eurofins Calscience, Garden Grove, CA, USA; formerly Calscience Environmental) using USEPA Method 200.7, and included antimony, arsenic, barium, beryllium, cadmium, chromium, cobalt, copper, iron, lead, molybdenum, nickel, selenium, silver, thallium, vanadium, and zinc. The reported metals’ concentrations were compared with USEPA ambient water quality criteria (US Environmental Protection Agency 2020), National Oceanic and Atmospheric Administration SQuaRT cards (2019), and other literature-based screening values (Suter and Tsao 1996) to assess the potential for the detected metals to cause toxicity.

RESULTS

Groundwater analytical chemistry results

Analytical results are summarized in the Supplemental Data, Tables S2 and S3. As per the study design, the majority of upgradient samples were nondetect for diesel-range organics
without silica gel clean-up (<100 µg/L), meaning that the OCOCs were not measurable using USEPA Method 8015B. Most downgradient samples had detections of diesel-range organics without silica gel clean-up (120–4400 µg/L) but were nondetect for diesel-range organics with silica gel clean-up. This means that the OCOC complex mixture was present and petroleum hydrocarbons were absent or nondetect using USEPA Method 8015B. In some cases, downgradient samples were nondetect for diesel-range organics without silica gel clean-up, meaning that the downgradient edge of the plume had shifted since it was last monitored, most likely caused by slight changes in hydraulic gradient. Results of the targeted OCOC analyses using GC–MS (76 analytes) show that a few acids and one alcohol were sometimes detected at low concentrations (Supplemental Data, Table S4, Section 2, contains a list of the target analytes). No compounds with a regulatory-derived reference dose were detected.

Tentatively identified compounds identified by nontargeted two-dimensional gas chromatography analysis were found in both upgradient and downgradient samples. Many of the tentatively identified compounds identified in these samples are of the same types of compounds known to either occur naturally in groundwater systems (e.g., organic acids) or as a result of ambient anthropogenic sources (e.g., phthalates); see O’Reilly et al. (2015, 2019) for detailed discussion. Structural classes identified included alcohols, ketones, aldehydes, phenols, and organic acids/esters (Supplemental Data, Tables S2 and S3). The most frequently identified family for both upgradient and downgradient wells was acids/esters (Supplemental Data, Tables S2 and S3).

These results are consistent with a larger set of samples collected from 22 fuel marketing sites (Zemo et al. 2017; O’Reilly et al. 2019).

**In vitro screening assays**

**Gamma-H2AX screening assay.** No evidence of DNA damage was seen in any sample, up to and including the maximum tested concentration of 2200 µg/L. Activity was noted at one site, in one sample (T4, MW-31) at the maximum concentration tested. However, at this concentration cytotoxicity was evident, as seen by a decrease in cell count by approximately 50%, well above the 30% acceptance criterion. The study report stated that this response indicated a loss of total cells due to necrosis, apoptosis, or a reduction in cellular proliferation. In addition, cytotoxicity was observed at the next 3 sample concentrations, which further demonstrates that cytotoxic effects occurred; these 3 dose levels did not result in DNA damage (Supplemental Data, Figure S3, Section 3). Per the assay protocol, cytotoxicity is a confounding factor and does not indicate that genotoxicity occurred. Finally, negative results were obtained at essentially the same diesel-range organics concentration for a sample collected from the same well in a previous sampling round (Supplemental Data, Figure S4, Section 3). Representative test data and figures for the assay controls as well as for the MW-31 sample are provided in the Supplemental Data, Section 3.

**ERTA screening assay.** None of the groundwater samples produced a positive estrogenic response at any OCOC concentration up to and including the maximum tested concentration (2200 µg/L). In addition, no unusual dose–response curve was obtained because it was consistently flat across doses and ranged from a minimum of 0.37% to a maximum of 5.45% expression relative to 1 nM estradiol. All samples were negative for cytotoxicity and there were no issues with solubility, which was measured in parallel for each assay as precipitation (see Supplemental Data, Figure S5, Section 4). The data obtained for all reference controls fell within the test laboratory’s historical range and corresponded well to the results of others in the literature. An example groundwater result is presented in Supplemental Data, Figure S5, and the control data are in Supplemental Data, Table S5.

**Chronic freshwater aquatic toxicity testing**

A total of 36 groundwater samples (17 upgradient, 18 downgradient, and 1 transgradient) from 14 sites were tested with varying extractable OCOC concentrations up to 1800 µg/L. The toxicity results are found in Table 1 and the results for water quality parameters are presented in the Supplemental Data, Table S6, Section 5. Figures 1 and 2 present the toxicity testing results for *C. dubia* (survival) and *P. promelas* (survival and growth) compared with the OCOC concentrations (diesel-range organics without silica gel clean-up) present in each sample and demonstrate that there is no correlation between OCOC concentration and toxicity.

Algal toxicity results with *R. subcapitata* were confounded by algal plating, which is when the algal cells adhere to the inside surface of the test flasks. This was frequently observed in both upgradient and downgradient samples. The testing laboratory reports that algal cell plating has been observed in a variety of wastewater treatment plant effluents that have ionic matrices that, although typical for those types of effluents, differ from laboratory control water (documented in aquatic toxicity study reports, Pacific EcoRisk Laboratory, Fairfield, CA, USA). There was also no consistent correlation between OCOC concentration and algal toxicity.

*Ceriodaphnia dubia* survival was not impacted relative to the laboratory control in either the upgradient or downgradient samples at 12 of the 14 sites. Survival was impacted in both the upgradient and downgradient samples at 2 sites (T2 and T1). *Ceriodaphnia dubia* reproduction was impacted in all but one of the upgradient and downgradient groundwater samples, regardless of OCOC presence. The alkalinity was above the reported *C. dubia* alkalinity chronic median lethal concentration and median inhibitory concentration for 19 of 36 site samples (Supplemental Data, Table S6, Section 5), which may have been a reason for the decreased reproduction.

No effects on *P. promelas* survival or growth were seen at 13 of the 14 sites (33 of 36 samples). Effects were limited to downgradient samples collected from a single site (T1) tested at 3 different times. However, this is attributed to the groundwater based on the Gas Research Institute freshwater salinity toxicity relationship modeling.
TABLE 1: Aquatic toxicity testing results for groundwater samples

| Site | Sample receipt date | Sample ID | Sample type | DRO without silica gel | Raphidocelis subcapitata mean algal cell density | Ceriodaphnia dubia | Pimephales promelas |
|------|---------------------|-----------|-------------|------------------------|-----------------------------------------------|------------------|-------------------|
|      |                     |           |             |                        | Mean % survival | Mean reproduction (neonates/female) | Mean % survival | Mean biomass (mg) |
| T4 7/10/13 laboratory water control | 7/10/2013 | MW-11 | U | ND | 3.73 | 100 | 34 | 92.5 | 0.55 |
| T4 7/10/13 laboratory water control | 7/10/2013 | MW-31 | D | 610 | 0.00* | 100 | 18* | 82.5 | 0.49 |
| T4 9/30/13 laboratory water control | 9/30/2013 | MW-31 | D | 250 | 3.03* | 100 | 17* | 85 | 0.51 |
| T4 7/10/13 laboratory water control | 7/10/2013 | MW-31 | U | ND | 3.73 | 100 | 59* | 95 | 0.57 |
| T4 9/30/13 laboratory water control | 9/30/2013 | MW-31 | D | 610 | 0.00* | 100 | 17* | 85 | 0.63 |
| T4 7/9/14 laboratory water control | 7/9/2014 | MW-11 | U | ND | 3.98 | 100 | 32.7 | 95 | 0.68 |
| T4 7/9/14 laboratory water control | 7/9/2014 | MW-31 | D | 230 | 0.86* | 100 | 24.8* | 95 | 0.54 |
| T4 7/11/13 laboratory water control | 7/11/2013 | MW-3 | U | ND | 2.44* | 100 | 0* | 97.5 | 0.59 |
| T4 7/11/13 laboratory water control | 7/11/2013 | MW-7 | D | 7.62* | 100 | 36.5 | 97.5 | 0.76 |
| T1 7/23/14 laboratory water control | 7/23/2013 | MW-65A | U | ND | 5.16 | 80* | 1.6* | 95 | 0.91 |
| T1 7/23/14 laboratory water control | 7/23/2013 | MW-50A | D | 680 | 4.66 | 100 | 3.3* | 37.5* | 0.25* |
| T1 9/30/13 laboratory water control | 9/30/2013 | MW-65A | U | ND | 4.4 | 40* | 0.3* | 94.7 | 0.69 |
| T1 9/30/13 laboratory water control | 9/30/2013 | MW-50A | D | 540 | 2.36* | 50* | 1.3* | 95 | 0.46 |
| T1 7/15/14 laboratory water control | 7/15/2014 | MW-65A | U | ND | 6.23* | 100 | 0* | 95 | 0.5 |
| T1 7/15/14 laboratory water control | 7/15/2014 | MW-50A | D | 780 | 3.27* | 100 | 15* | 0.11* |
| WG1 9/30/13 laboratory water control | 9/30/2013 | MW-6 | U | ND | 4.94 | 100 | 18.9* | 95 | 0.6 |
| WG1 9/30/13 laboratory water control | 9/30/2013 | MW-7 | D | 1800 | 8.15 | 100 | 15.4* | 97.5 | 0.63 |
| UKSS 10/16/13 laboratory water control | 10/16/2013 | MW-15 | U | ND | 5.7 | 100 | 12.8* | 95 | 0.46 |
| UKSS 10/16/13 laboratory water control | 10/16/2013 | MW-18 | D | 640 | 0.17* | 100 | 18.4* | 92.5 | 0.45 |
| RB 7/7/14 laboratory water control | 7/7/2014 | MW-12 | U | ND | 7.36* | 90 | 6.5* | 95 | 0.57 |
| RB 7/7/14 laboratory water control | 7/7/2014 | MW-13 | D | 120 | 6.82* | 100 | 11.1* | 97.5 | 0.59 |
| SacWB 7/9/14 laboratory water control | 7/9/2014 | MK-14 | U | ND | 4.01 | 100 | 33.1 | 95 | 0.52 |
| SacWB 7/9/14 laboratory water control | 7/9/2014 | MK-20 | D | 1500 | 0.07* | 100 | 23.2* | 92.5 | 0.6 |
| GV 7/10/14 laboratory water control | 7/10/2014 | MW-7 | U | ND | 3.64 | 100 | 27.1 | 100 | 0.44 |
| GV 7/10/14 laboratory water control | 7/10/2014 | MW-5 | D | 484* | 100 | 8.1* | 100 | 0.43 |
| SacE 7/10/14 laboratory water control | 7/10/2014 | C-23 | U | ND | 3.64 | 100 | 27.1 | 100 | 0.44 |
| SacE 7/10/14 laboratory water control | 7/10/2014 | C-33 | D | 120 | 6.61* | 100 | 8.2* | 95 | 0.44 |
| Dun 7/10/14 laboratory water control | 7/10/2014 | C-7 | U | ND | 2.00* | 70 | 0.1* | 95 | 0.48 |
| Dun 7/10/14 laboratory water control | 7/10/2014 | C-2 | D | 400 | 4.77* | 100 | 9.5* | 100 | 0.45 |
| Arc 7/21/14 laboratory water control | 7/21/2014 | MW-9 | T | 620* | 5.63* | 60* | 0.6* | 95 | 0.36 |
| Arc 7/21/14 laboratory water control | 7/21/2014 | MW-11 | D | 740 | 4.64* | 100 | 6.9* | 95 | 0.41 |
| CrescC 7/21/14 laboratory water control | 7/21/2014 | MW-8 | U | ND | 4.75 | 100 | 45 | 95 | 0.39 |
| CrescC 7/21/14 laboratory water control | 7/21/2014 | MW-5 | D | 480 | 5.39* | 100 | 29.6* | 100 | 0.45 |
| Son 7/22/14 laboratory water control | 7/22/2014 | MW-6 | U | ND | 5.75* | 100 | 28.5* | 97.5 | 0.41 |
| Son 7/22/14 laboratory water control | 7/22/2014 | MW-10 | D | ND | 7.27* | 100 | 26.8* | 100 | 0.46 |
| WG2 7/23/14 laboratory water control | 7/23/2014 | MW-11 | U | ND | 4.04 | 100 | 36.3 | 97.5 | 0.47 |
| WG2 7/23/14 laboratory water control | 7/23/2014 | MW-5 | D | 410 | 0.59* | 100 | 0.2* | 95 | 0.43 |

*Laboratory report (P) notes that visible algal plating was observed; indicating potential confounding results.

**MW-9 at the Arc site is approximately transparent of the residual source area and MW-11. Water levels were very low when these 2 were sampled. It is possible that drought conditions caused a shift in the gradient direction, and that MW-9 was temporarily downgradient of the residual source area (and thus within the OCOC "plume") at the time the wells were sampled for the present study.

*OCOC concentration in MK-14 at the SacWB site was the result of a single compound identified by traditional GC-MS as dodecaneoic acid.

*p < 0.05. The replications at these test control treatments were significantly less than the lab control treatment response.

ND = non-detect; U = upgradient; D = downgradient; T = transgradient; DRO = diesel-range organics; OCOC = oxygen-containing organic compound; GC-MS = gas chromatography–mass spectrometry.
The Gas Research Institute freshwater salinity toxicity relationship modeling conducted for the T1 site predicted the occurrence of *P. promelas* and *C. dubia* toxicity based on the ionic composition of the groundwater (Table 2). The model only predicts acute toxicity (48 h for *C. dubia* and 96 h for *P. promelas*); thus the survival at these time points was also evaluated during the present study for comparison to predicted values. The freshwater salinity toxicity relationship model predictions were a good indicator of toxicity for most of the downgradient samples, including predicting the trend of increasing toxicity at the longer time point of 6 to 8 d for the chronic study. Toxicity is consistent with the elevated alkalinity, hardness, conductivity, and salinity in the samples from the T1 site compared with other sites (Supplemental Data, Table S6). In addition, the freshwater salinity toxicity relationship modeling predicted that higher levels of toxicity would occur for *Daphnia* rather than for fish,

**FIGURE 1**: *Ceriodaphnia dubia* survival comparison with oxygen-containing organic compound (OCOC) concentration as diesel-range organics. Data points on x axis with OCOC concentration = 0 µg/L, representing the upgradient/background samples. GRI = Gas Research Institute; STR = salinity toxicity relationship.

**FIGURE 2**: *Pimephales promelas* survival and growth comparison with oxygen-containing organic compound (OCOC) concentration as diesel-range organics. Data points on x axis with OCOC concentration = 0 µg/L, representing the upgradient/background samples. GRI = Gas Research Institute; STR = salinity toxicity relationship.
Toxicity of petroleum biodegradation metabolites—Environmental Toxicology and Chemistry, 2020;39:1634–1645

which is consistent with the observed data. Finally, it is not expected that the chemical composition and OCOCs present were the source of toxicity because the OCOC concentrations for the downgradient T1 samples were 540, 680, and 780 µg/L, which falls approximately mid-range with downgradient samples from other sites. In addition, no unique targeted compounds via USEPA modified Method 8270C were observed and the distribution among the structural classes was similar among sites, indicating that it is unlikely an unidentified compound was present and causing toxicity. Salinity toxicity relationship modeling predictions are not provided for other sampling sites because there were no other predictions of toxicity from the ionic matrix.

The presence of dissolved metals did not appear to cause toxicity to any species; however, several metals exceeded the screening criteria at a number of sites (both upgradient and downgradient). Supplemental Data, Table S7, Section 5 contains the measured metal concentrations.

**DISCUSSION**

In our previous work, the potential toxicity of OCOCs in groundwater to humans was investigated from an individual component perspective (Zemo et al. 2017). A reference dose-based relative toxicity ranking system was developed (low, with reference dose ≥ 0.1; low-to-moderate, with reference dose < 0.1–0.01; and moderate, with reference dose < 0.01–0.001 mg/kg/d) and assigned to 22 structural classes of OCOCs that could potentially be present based on sampling results and known biodegradation pathways (Zemo et al. 2013, 2017). The overall conclusion of that work was that most of the OCOCs identified in groundwater could be ranked as low (reference dose ≥ 0.1 mg/kg/d) toxicity to humans and that they would be unlikely to present a significant health risk. Furthermore, the toxicity ranking profile of the tentatively identified compounds in the complex mixture decreased with biodegradation of the plumes.

Since publication of Zemo et al. (2017), two regulatory agencies have reviewed and used these results to develop their own weighted-average reference doses to be used for screening OCOCs in groundwater. The Hawaii Department of Health (2018) developed reference doses for various stages of the plume from least to most biodegraded. The Australia Cooperative Research Centre for Contamination Assessment and Remediation of the Environment (2018; similar to the Interstate Technology Regulatory Council in the United States) independently reviewed and applied a chemical category component approach to develop reference doses for each biodegradation stage.

To complement the component approach, the present study presents a second line of evidence by conducting whole-mixture testing of groundwater samples using in vitro screening assays for human genotoxic and estrogentic effects, as well as chronic toxicity assays for freshwater aquatic receptors. Toxicological approaches for characterizing the potential health impacts of environmental concentrations of complex mixtures are
challenging (Krishnan et al. 1997; Cassee et al. 1998; European Commission 2012). This is especially true for mixtures where the individual components are transient in nature and cannot be concentrated for testing without altering their physical or chemical properties. Each approach has benefits and limitations, and the appropriate choice depends on the circumstances and the data needs. A marked benefit of following the whole-mixture testing approach is that all chemicals present in the groundwater are tested at the concentrations at which they exist and are bioavailable in the environment. The testing does not rely on analytical measurement of the components of the mixture and thus provides a comprehensive evaluation of the potential toxicity.

The in vitro testing results supplement the previous reference dose-based component work and further demonstrate the low human health hazard potential from OCOCs and that the screening levels of 1500 to 15 000 µg/L are protective of human health (Zemo et al. 2017). Specifically, genotoxicity and endocrine disruption via estrogenicity are unlikely to occur.

Aquatic toxicity was not assessed in the previous work by Zemo et al. (2013, 2017). The present study results demonstrate that OCOCs up to 1800 µg/L diesel-range organics do not cause chronic aquatic toxicity to P. promelas or C. dubia (survival). Whether OCOCs in groundwater caused reproductive toxicity to C. dubia could not be assessed because a decrease in reproduction occurred at both upgradient and downgradient wells, which was likely caused by unsuitable groundwater matrix conditions. Indeed, water quality characteristics such as ion concentration have been shown to influence or directly cause toxicity to invertebrates used in aquatic assays (Cowgill and Milazzo 1991; Douglas and Horne 1997; Pillard et al. 2000, 2002; Kline and Stekoll 2002; Mount et al. 2016). In addition, the R. subcapitata data were confounded by algal cell plating leading to inconsistent results among both upgradient and downgradient samples. Based on the structural classes or families comprising OCOCs, it is unlikely that selective toxicity to algae would occur; this is being further confirmed in ongoing work.

Based on the extensive amount of testing conducted in the present study, it is recommended that before embarking on a testing program to assess aquatic toxicity at a petroleum-release site, a representative sample well should first be examined to ensure compatibility with the assay; this may include measuring general water quality characteristics and/or pilot aquatic toxicity testing. In addition, selecting appropriate background well controls that have water quality similar to the target wells is critical to avoid false-positive results. Additional research is likely needed to optimize the C. dubia reproductive toxicity assay to determine which water quality parameter(s) are most critical.

A goal of the present study is to establish concentrations of OCOCs not hazardous to human and aquatic receptors that may be applied to all petroleum fuel release sites because testing every mixture at every site is often not feasible (based on the amount of mixtures that would need to be tested) or ethical (if conducting animal studies). This leads to the second recommended approach, testing similar mixtures and extrapolating the results. However, this requires understanding certain parameters that may contribute to or influence toxicity, such as the mechanism of action and whether there are any interactions among the constituents present (e.g., leading to synergistic effects), as well as whether there are OCOCs that have a greater hazard potential than others. Whereas a limitation to the present study is that the concentration of each individual OCOC was not quantified, the large sample size across multiple sites reveals that the OCOCs in groundwater from fuel release sites are unlikely to cause aquatic toxicity at OCOC diesel-range organic concentrations up to 1800 µg/L.

The limitations with the mixture testing approaches highlight the value of the third approach: evaluating the toxicity of individual components of a mixture. Future work could further extend the present component work (Zemo et al. 2017) to evaluate the toxicity based on the functional groups present for each molecular structural class of compounds identified by two-dimensional GC. This could be accomplished by using computational toxicology tools to apply the wealth of toxicity data available in the public literature, and/or structural activity relationships to make in silico predictions. The value of this additional work is to: 1) cover all relevant toxicological endpoints, 2) determine toxicity trends within the chemical categories of OCOCs present (e.g., impact of branching, alkyl chain length), 3) identify ecological receptors that may be most sensitive for potential future screening studies, and 4) evaluate mechanism of action and whether there are any synergies among OCOC chemical categories. Not only will this directly provide information on the potential toxicity of OCOCs but it may also supply data on targeted testing approaches for petroleum release sites.

We are aware of 2 other research programs evaluating the aquatic toxicity of OCOCs in groundwater with the intent of developing screening or action levels. The first developed an environmental screening level of 510 µg/L for petroleum hydrocarbon oxidation products (same as OCOCs; San Francisco Bay Regional Water Quality Control Board 2019) for saltwater aquatic species based on a study with Americamysis bahia. The study (Terraplace 2018) tested 4 wells at one site. However, toxicity was incorrectly attributed to OCOCs even though: 1) toxicity occurred in the background control wells, 2) water quality characteristics of the background and target wells differed, 3) there was no dose–response relationship between OCOC concentration and toxicity, and 4) bicarbonate and calcium concentrations were near or exceeding levels that are known to cause toxicity to A. bahia (Douglas and Horne 1997; Pillard et al. 2000, 2002; Kline and Stekoll 2002).

The second program is ongoing and conducted by the Washington State Department of Ecology (2019). The aim of this project was to study the effects of OCOCs in groundwater to both freshwater and saltwater aquatic species. Additional studies primarily addressing the aquatic toxicity of OCOCs are found in the scientific literature. These include research conducted on: 1) weathered crude oils (Wolfe et al. 1995; Neff et al. 2000; McGuire et al. 2018), 2) crude oil-produced water (Thomas et al. 2009; Scarlett et al. 2012; Hughes et al. 2017), 3) water-accommodated fractions (WAFs)
or water-soluble fractions of fresh or artificially weathered crude oils (Barron et al. 1999 [WAFs]; Melbye et al. 2009 [water-soluble fractions]), 4) WAFs of refined fuels (Neff et al. 2000), and 5) oil-sands process water to assess the toxicity of naphthenic acids (Rogers et al. 2002; McKee et al. 2014).

In contrast, the present study differs in that: 1) groundwater samples were collected from naturally biodegrading (i.e., stable or shrinking plumes) fuel releases, thus they represent typical conditions expected at biodegrading fuel release sites, 2) groundwater tested was derived from refined fuels and not from crude oil, 3) OCOC mixtures were tested directly as they exist in site groundwater without any laboratory manipulation except for dilution that was required for the in vitro assays, 4) only dissolved polar metabolites (OCOCs) were tested, whereas dissolved or nondissolved hydrocarbons and other fuel components were carefully avoided to preclude confounding toxicity test results, and 5) a field control sample was incorporated by testing groundwater samples from upgradient of the fuel sources representing background groundwater conditions at each site.

Finally, the hydrocarbon and nitrogen-sulfur-oxygen-containing compounds present in refined fuels differ significantly from those present in crude oil, oil-tar sands, or oil-sands produced water; and thus the composition of the soluble biodegradation metabolites from these different sources could differ significantly from the mixtures tested for the present study. For example, studies have focused on the presence of naphthenic acids as potentially active agents producing endocrine disruption in aquatic receptors. Whereas naphthenic acids can be present at high concentrations in oil-sands process waters, they were infrequently identified by the targeted and nontargeted analytical approaches used in the present study.

CONCLUSIONS

A multi-year research program has applied multiple lines of investigation to evaluate the hazard and risk of groundwater at fuel release sites including: 1) determining the presence of compounds with potential toxicological concern, 2) assessing the toxicity of identified compounds based on existing data in the literature, and 3) targeted toxicity testing of whole mixtures. These approaches are similar to those taken by multiple regulatory agencies (Australia Cooperative Research Centre for Contamination Assessment and Remediation of the Environment 2018; Drewes et al. 2018; Hawaii Department of Health 2018).

The results from the present study demonstrate that the biodegrading mixtures containing OCOCs in downgradient groundwater samples from multiple fuel release sites—at concentrations up to the maximum tested of 2200 µg/L diesel-range organics—did not activate the estrogen receptor or cause genotoxicity. In addition, OCOCs up to the maximum concentrations tested at 1800 µg/L did not cause aquatic toxicity above background groundwater to 2 standardized freshwater aquatic species, P. promelas and C. dubia. Whole-mixture testing conducted in the present study adds to the weight of evidence, in combination with previous component-based work (Zeno et al. 2013, 2017), that once a plume from a fuel release is biodegraded to the point where no dissolved hydrocarbons are present, the toxicity to human and freshwater aquatic receptors is low.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at https://doi.org/10.1002/etc.4749.

Acknowledgment—The authors thank the Chevron Environmental Management Company for funding the present study under Chevron’s Remediation Technology Development Initiative.

Disclaimer—The authors declare that there are no conflicts of interest regarding this publication.

Data Availability Statement—Data, associated metadata, and calculation tools are available from the corresponding author (timothy.patterson@chevron.com).

REFERENCES

Ando M, Yoshikawa K, Iwase Y, Ishiura S. 2014. Usefulness of monitoring γH2AX and cell cycle arrest in HepG2 cells for estimating genotoxicity using a high-content analysis system. J Biomol Screen 19:1246–1254.

Australia Cooperative Research Centre for Contamination Assessment and Remediation of the Environment. 2018. Weathered petroleum hydrocarbons (silica gel clean-up). Technical Report No. 40. Callaghan, Australia.

Barber LB, Vajda AM, Douville C, Norris DO, Writer JH. 2012. Fish endocrine disruption responses to a major wastewater treatment facility upgrade. Environ Sci Technol 46:2121–2131.

Barron MG, Podrabsky T, Ogle S, Ricker RW. 1999. Are aromatic hydrocarbons the primary determinant of petroleum toxicity to aquatic organisms? Aquat Toxicol 46:253–268.

Cassee FR, Groten JP, Van Bladenen PJ, Feron VJ. 1998. Toxicological evaluation and risk assessment of chemical mixtures. Crit Rev Toxicol 28:73–101.

Colborn T, vom Saal FS, Soto AM. 1993. Developmental effects of endocrine-disrupting chemicals in wildlife and humans. Environ Health Perspect 101:378–384.

Cowgill UM, Milazzo DP. 1991. The sensitivity of two cladocerans to water quality variables: Salinity <467 mg NaCl/L and hardness <200 mg CaCO3/L. Arch Environ Contam Toxicol 21:218–223.

Douglas WS, Horne MT. 1997. The interactive effects of essential ions and salinity on the survival of Myisidopsis bahia in 96-H acute toxicity tests of effluents discharged to marine and estuarine receiving waters. Environ Toxicol Chem 16:1996–2001.

Dreier DA, Connors KA, Brooks BW. 2015. Comparative endpoint sensitivity of in vitro estrogen agonist assays. Regul Toxicol Pharmacol 72:185–193.

Drewes JE, Anderson P, Denslow N, Jakubowski W, Olivier A, Schlenk D, Snyder S. 2018. Monitoring strategies for constituents of emerging concern in recycled water. California State Water Resources Control Board, Sacramento, CA, USA.

European Commission. 2012. Toxicity and assessment of chemical mixtures. Brussels, Belgium. [cited 2020 April 11]. Available from: http://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_155.pdf

Filby AL, Neuparth T, Thorpe KL, Owen R, Galloway TS, Tyler CR. 2007. Health impacts of estrogens in the environment, considering complex mixture effects. Environ Health Perspect 115:1704–1710.

Hawaii Department of Health. 2018. TPH risk case studies. Honolulu, HI, USA. [cited 2020 March 30]. Available from: https://tphrisk-1.itrcweb.org/11-tph-risk-case-studies/
US Environmental Protection Agency. 2014. NICEATM Workshop on Adverse Outcome Pathways. An adverse outcome pathway framework for screening for potential endocrine disruption. Washington, DC.

US Environmental Protection Agency. 2020. National recommended water quality criteria—Aquatic life criteria table. Washington, DC. [cited 2020 April 29]. Available from: https://www.epa.gov/wqcf/national-recommended-water-quality-criteria-aquatic-life-criteria-table

Washington State Department of Ecology, 2019. Addendum to quality assurance project plan: Environmental effects-based concentrations for total petroleum hydrocarbons (TPHs) in marine and fresh water, Olympia, WA, USA. [cited 2020 April 29]. https://fortress.wa.gov/ecy/publications/documents/1903104.pdf

Watters GP, Smart DJ, Harvey JS, Austin CA. 2009. H2AX phosphorylation as a genotoxicity endpoint. Mutat Res 679:50–58.

Wehmas LC, Cavallin JE, Durhan EJ, Kahl MD, Martinovic D, Mayasich J, Tuominen T, Villeneuve DL, Ankley GT. 2011. Screening complex effluents for estrogenic activity with the T47D-KBlue cell bioassay: Assay optimization and comparison with in vivo responses in fish. Environ Toxicol Chem 30:439–445.

Wiedemeier TH, Wilson JT, Kampbel DH, Miller RN, Hansen JE. 1995. Technical protocol for implementing intrinsic remediation with long-term monitoring for natural attenuation of fuel contamination dissolved in groundwater, p. 323. US Air Force Center for Environmental Excellence. San Francisco, CA.

Wolfe DA, Scott KJ, Clayton JR, Lunz J, Payne JR, Thompson TA. 1995. Comparative toxicity of polar and non-polar organic fractions from sediments affected by the Exxon Valdez oil spill in Prince William Sound, Alaska. Chem Ecol 10:137–156.

Zemo DA, O’Reilly KT, Mohler RE, Magaw RI, Espino Devine C, Ahn S, Tiwary AK. 2017. Life cycle of petroleum biodegradation metabolite plumes, and implications for risk management at fuel release sites. Integr Environ Assess Manag 13:714–727.

Zemo DA, O’Reilly KT, Mohler RE, Tiwary AK, Magaw RI, Synowiec KA. 2013. Nature and estimated human toxicity of polar metabolite mixtures in groundwater quantified as TPHs/IDRO at biodegrading fuel release sites. Ground Water Monit Remediat 33:44–56.