Water Research 150 (2019) 140–152

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Evaluating emerging organic contaminant removal in an engineered hyporheic zone using high resolution mass spectrometry

Katherine T. Peter [a, b, *], Skuyler Herzog [c], Zhenyu Tian [a, b], Christopher Wu [a, b], John E. McCray [c], Katherine Lynch [d], Edward P. Kolodziej [a, b, e]

[a] Interdisciplinary Arts and Science, University of Washington Tacoma, Tacoma, WA, 98421, USA
[b] Center for Urban Waters, Tacoma, WA, 98421, USA
[c] Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO, 80401, USA
[d] Seattle Public Utilities, Seattle, WA, 98104, USA
[e] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, 98195, USA

ARTICLE INFO

Article history:
Received 5 September 2018
Received in revised form
16 November 2018
Accepted 17 November 2018
Available online 23 November 2018

Keywords:
Non-target analysis
Suspect screening
Urban stream restoration
Stormwater treatment
Hyporheic zone
Sorption

ABSTRACT

The hyporheic zone (HZ), located at the interface of surface and groundwater, is a natural bioreactor for attenuation of chemical contaminants. Engineered HZs can be incorporated into stream restoration projects to enhance hyporheic exchange, with flowpaths optimized to promote biological habitat, water quantity, and water quality improvements. Designing HZs for in-stream treatment of stormwater, a significant source of flow and contaminant loads to urban creeks, requires assessment of both the hydrology and biogeochemical capacity for water quality improvement. Here, we applied tracer tests and high resolution mass spectrometry (HRMS) to characterize an engineered hyporheic zone unit process, called a hyporheic design element (HDE), in the Thornton Creek Watershed in Seattle, WA. Dye, NaCl, and bromide were used to hydrologically link downwelling and upwelling zones and estimate the hydraulic retention time (HRT) of hyporheic flowpaths. We then compared water quality improvements across hydrologically-linked surface and hyporheic flowpaths (3–5 m length; ~30 min to >3 h) during baseflow and stormflow conditions. We evaluated fate outcomes for 83 identified contaminants during stormflow, including those correlated with an urban runoff mortality syndrome in coho salmon. Non-target HRMS analysis was used to assess holistic water quality improvements and evaluate attenuation mechanisms. The data indicated substantial water quality improvement in hyporheic flowpaths relative to surface flow and improved contaminant removal with longer hyporheic HRT (for ~1900 non-target compounds detected during stormflow, <17% were attenuated >50% via surface flow *vs.* 59% and 78% via short and long hyporheic residence times, respectively), and strong contributions of hydrophobic sorption towards observed contaminant attenuation.

1. Introduction

The hyporheic zone (HZ), the interface of surface water and groundwater exchange in a lotic system, has been described as the "river's liver" (Fischer et al., 2005) because it is a natural bioreactor for chemical contaminants (Huntscha et al., 2013; Lewandowski et al., 2011). However, hydromodification of urban streams increases scour, reducing HZ exchange and attenuation potential (Lawrence et al., 2013). Including engineered HZs in stream restoration projects can promote ecosystem function by improving biological health, habitat, water quantity and quality (Hester and Gooseff, 2010; Lawrence et al., 2013). In terms of organic contaminants, the potential for water quality improvement in natural HZs has been assessed by evaluating the fate of wastewater-derived organic contaminants, such as pharmaceuticals (Kunkel and Radke, 2008; Lewandowski et al., 2011; Lin et al., 2006), perfluorinated chemicals (Hoehn et al., 2007), pesticides (Huntscha et al., 2013), and alkylphenols (Lin et al., 2006). However, urban stormwater contributes substantial flow and contaminant loads to urban creeks, including many contaminants distinct from those above (Grebel et al., 2013). Importantly, engineered HZs can treat non-point source pollution in-stream, eliminating the need for

* Corresponding author. Center for Urban Waters, 326 E. D Street, Tacoma, WA 98421, USA.
E-mail address: ktpeter@uw.edu (K.T. Peter).

https://doi.org/10.1016/j.watres.2018.11.050
0043-1354/

side-stream treatment options and land in space-limited urban environments.

In the HZ, hydraulic residence time (HRT) distributions, mass transfer rates, and microbial abundance govern hyporheic contaminant attenuation (Kunkel and Radke, 2011; Mendoza-Lera and Datry, 2017). Generally, both laboratory and field experiments indicate that longer contact times with HZ sediments (i.e., longer HRTs) increase contaminant attenuation, although half-life estimates for specific chemicals vary widely (e.g., 2.7 h *vs.* 10.5 days for gemfibrozil) (Kunkel and Radke, 2008; Lin et al., 2006). Notably, studies by Lewandowski et al. (2011) and Hoehn et al. (2007) indicated that little reach-scale contaminant removal was attributable to HZ flow because hyporheic exchange volumes were small relative to total surface flow and/or HZ HRTs were too short relative to attenuation timescales, respectively. Likewise, "traditional" restoration drop structures such as cross-vanes and low-head dams are often ineffective for water quality improvement, in part because they promote short HRTs (Azinheira et al., 2014; Hester et al., 2016). However, engineered HZs can be designed as unit processes (known as hyporheic design elements, or HDEs) to increase hyporheic exchange, lengthen flowpaths, and provide longer residence times (Harvey et al., 2013; Herzog et al., 2018, 2016), potentially promoting water quality improvements.

Quantifying contaminant attenuation within the HZ is complicated by spatiotemporally dynamic hyporheic flow patterns. Reach-scale approaches that use injections (or ambient fluctuations) of tracers to model solute retention (Runkel, 1998) assume uniform stream and hyporheic zone characteristics (e.g., velocity, hyporheic cross-sectional area) in the direction of flow, and are not appropriate to characterize engineered restoration structures with disparate hydraulic properties. Further, reach-scale models do not separate surface *vs.* hyporheic HRT (Ward et al., 2017), so they cannot differentiate between surface eddies above *vs.* hyporheic flow beneath a restored drop structure. In contrast, current state-of-the-art design to assess flowpath-scale solute retention in the HZ relies on pore water samplers and the assumption of 1D vertical flow (Knapp et al., 2017; Lewandowski et al., 2011; Posselt et al., 2018; Schaper et al., 2018). However, this approach cannot reliably assess water quality changes if flow fields are 2D or 3D, where downwelling water may travel both downstream and laterally prior to upwelling. Instead, the start and end of individual flowpaths must be identified to provide hydrologically-linked downwelling (influent) and upwelling (effluent) samples. Although the value of monitoring individual hyporheic flowpaths is clear (Abbott et al., 2016; Krause et al., 2017; Lewandowski et al., 2011), only one prior study monitored water quality changes along specific HZ flowpaths by injecting and tracking several tracers using a piezometer array (Zarnetske et al., 2011). However, this method relied on fortuitous flow connections rather than installing piezometers after identifying linked downwelling and upwelling zones, and has not been applied in restored or natural step-pool type structures (Zarnetske et al., 2011). Thus, a reliable and generalizable sampling approach to 3D hyporheic flow fields for water treatment evaluation is still needed.

With a suitable flowpath-based sampling design, water treatment performance assessments in engineered HZs should evaluate both the occurrence and fate of novel and emerging urban stormwater contaminants, as well as incorporate holistic metrics of water quality. Accordingly, the use of suspect and non-target screening analyses via high resolution mass spectrometry (HRMS) supports assessment of known "targeted" contaminants, enables detection of unknown and unanticipated chemicals (Bade et al., 2015a), and provides a powerful, statistically robust approach for more comprehensive evaluation of process performance (Bader et al., 2016; Nürenberg et al., 2015). Such assessments can reduce biases among smaller numbers of targeted chemicals, where similar chemical properties skew data outputs and are less representative of overall contaminant loads (Parry and Young, 2016). Insight into treatment performance can be developed by evaluating the number of unique molecular formulas (McEachran et al., 2018), using mass defect and spectral similarity screening to identify transformation products and homologous series (Hollender et al., 2017; Schollée et al., 2018), or applying data clustering techniques to group chemicals by fate and identify indicators (Merel et al., 2015). More importantly, linking treatment outcomes to molecular characteristics (e.g., polarity) can assess the role of common removal mechanisms (e.g., sorption) on performance. Existing applications of HRMS to evaluate treatment processes have used tentative formula and structure assignments to evaluate removal for specific contaminant classes (Parry and Young, 2016) and applied average *m/z* and retention time to determine overall changes in contaminant characteristics (Schollée et al., 2018). However, existing approaches have not further incorporated molecular characteristics to delineate and characterize attenuation mechanisms.

In this study, we characterized both hydrology and water quality improvement in an engineered hyporheic zone in the Thornton Creek watershed in Seattle, WA. To the best of our knowledge, the Thornton Creek HDEs are among the first direct hyporheic manipulations installed in any urban stream restoration in the U.S.; restoration efforts were intended to address both water quantity and quality. To guide water quality sampling, non-reactive tracers (dye, NaCl, and bromide) were used to hydrologically link downwelling and upwelling zones and estimate flowpath HRTs. We then evaluated the fate of a suite of urban stormwater contaminants, including those correlated with an urban runoff mortality syndrome in coho salmon (*Oncorhynchus kisutch*) (Peter et al., 2018), by sampling surface and hyporheic flowpaths during both baseflow and stormflow conditions. Expanding on previous data analysis strategies (e.g., Bader et al., 2017; Parry and Young, 2016), we used suspect and non-target HRMS data to evaluate the potential for water quality improvements in the HDE and the role of HRT and sorption in contaminant attenuation.

2. Materials and methods

2.1. Study site

Thornton Creek is the largest (3000 ha) and most urbanized watershed in Seattle, WA. The "Kingfisher Reach", located high on the South Fork of Thornton Creek (47°42'5"N, 122°18'27"W; Fig. 1), has a flashy hydrograph due to the watershed's high fraction of impervious surfaces. Discharge data for January 2014–December 2016 (from a gauge 1.2 km upstream) and for water years 2017–2018 (estimated via regression to 2014–2016 discharge at the creek mouth) are provided in Figures S2-S4. Summer baseflow in 2016 and 2017 was 30–40 L/s, whereas multiple winter storms generated peak flows >600 L/s.

During a 2014 restoration of the Kingfisher Reach, the creek was re-meandered and reconnected to its floodplain. A homogeneous streambed mixture (coarse gravel to medium sand, 1–2.5 m depth; Figure S1) was installed over the native substrate (a relatively impermeable compacted till) (Bakke, 2018). Many partially-buried logs were added to the reach, but no additional organic matter was incorporated into the streambed aggregate. Additionally, six engineered HDEs were constructed to optimize hyporheic exchange fluxes and residence times. HDEs use a channel-spanning log to create a plunge pool, appearing similar to low head dams at the surface, but incorporate several novel subsurface features: 1) an impermeable barrier extends ~1 m below the channel-spanning log, capping the streambed for the first several meters of the plunge

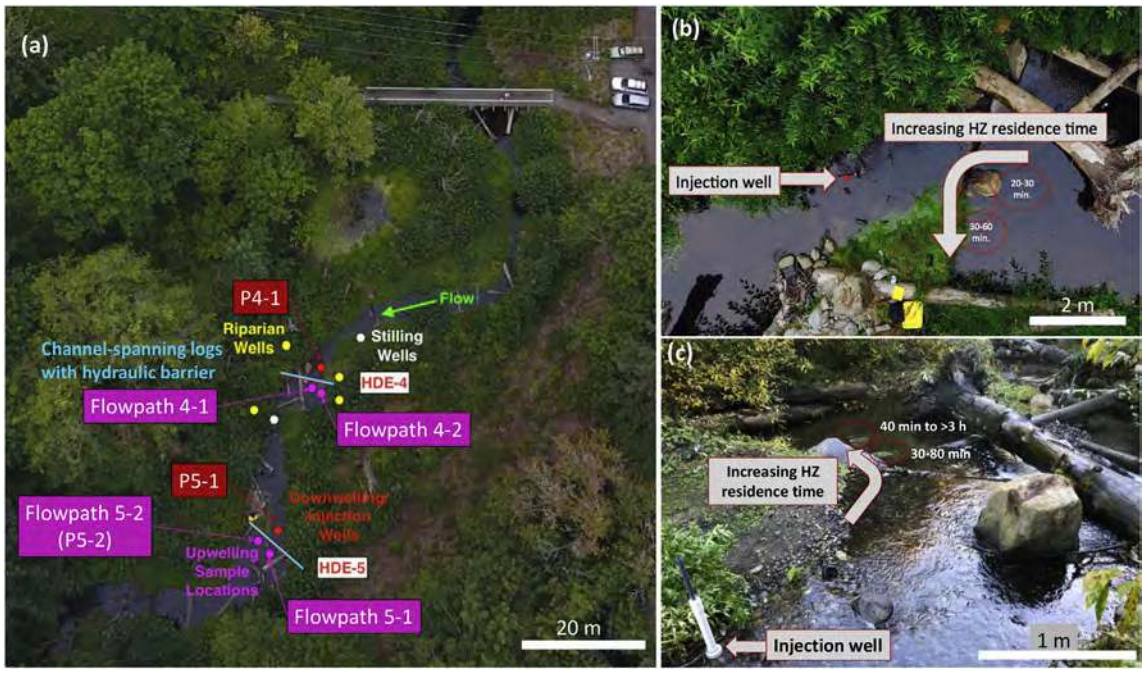

Fig. 1. (a) Map of study site within Kingfisher Reach on the South Fork of Thornton Creek showing hyporheic design elements (HDEs) and downwelling/upwelling sampling sites, and photographs of (b) HDE-5 and (c) HDE-4, showing locations of dye injection and upwelling, and approximate residence time distributions.

pool; and 2) the hyporheic zone is excavated deeper than the rest of the reach and backfilled with high permeability coarse gravel (Fig. 2). The first design feature induces longer HZ flowpaths by preventing short-circuiting flows beneath the log, and the second increases the HZ exchange capacity across the structure to increase the proportion of streamflow treated by the HZ.

Although fine sediments accumulate and scour seasonally in the Kingfisher Reach plunge pools, the surface sediment remained coarse after restoration (75% of samples had >90% pebbles and cobbles), while some finer sands accumulated in the subsurface (75–85% pebbles and cobbles) (Figure S1). The HDEs are each 4.6 m wide and 1.1 m deep across the channel, but the same coarse fill material extends into both stream banks, increasing the HDE cross-sectional area to 8.4 m^2. We performed initial flowpath delineation and sampling in HDE-5 in September 2017 (Fig. 1). However, the upwelling zone at HDE-5 later shifted, reducing the HRT. Therefore, subsequent flowpath delineation and sampling efforts were moved to HDE-4 (~40 m upstream of HDE-5) in October 2017 to complement an existing piezometer array (HOBO U20s and U20Ls, Onset Computer Corp, Bourne, MA). This array included 4 riparian piezometers that bracketed HDE #4 (upstream/downstream and left/right bank) and in-stream stilling wells upstream and downstream of the structure. All wells were 3.8 cm diameter stainless steel, with the stilling wells slotted throughout their length and the riparian wells screened at depths of 75–125 cm.

2.2. Flowpath delineation

We characterized individual hyporheic flowpaths by a multistep approach. To hydrologically link hyporheic flowpaths, non-toxic red food dye was injected into downwelling zones, with exfiltration monitored visually to identify upwelling zones. Sampling

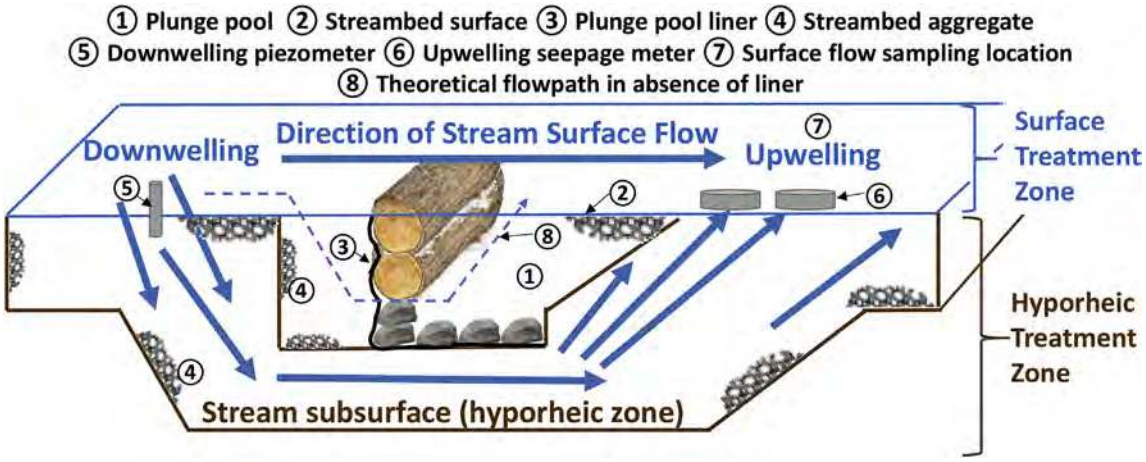

Fig. 2. Schematic of an engineered hyporheic design element (HDE), showing surface flow and hyporheic flowpaths and our conceptual approach to sampling hydraulically paired samples.

infrastructure was then installed at upwelling locations, and quantitative tracers (NaCl, bromide) were then used to connect sampling points and estimate HRT. Where possible, electrical conductivity loggers were deployed over longer timescales to assess flowpath hydrology during multiple storm events.

At HDE-5, dye was injected into piezometer P5-1 in the downwelling zone (2.5 m upstream from the channel-spanning log; screened 10–30 cm below the streambed). At ~15 min after injection, dye emerged in a focused location 1 m downstream from the log, in the center of the pool, then developed in an arc shape toward the right bank over the next hour. Two sampling locations were selected from this upwelling region to provide relatively fast (5–1) and slow (5–2) flowpaths. Minimum flowpath lengths of 3.6 m for Flowpath 5–1 and 3.1 m for Flowpath 5–2 (despite the longer HRT) were calculated by assuming a direct line from the injection point (average 20 cm depth) to the bottom of the HDE impermeable barrier and then to the upwelling location. At HDE-4, dye was injected into downwelling piezometer P4-1 (2.75 m upstream from channel-spanning log), and subsequent upwelling spread from the center of the pool to the left bank. Two flowpaths representing different HRTs were selected: 4–1 was faster (4.4 m long), 4–2 was slower (5 m long). Seepage meters (47 cm diameter, 110 L HDPE barrel sections; fitted with sampling ports and volumetric bags) were installed at each upwelling location (5-2 did not seal well and was replaced with a piezometer, P5-2). Seepage meter 5–1 (20 cm tall) was installed to a depth of 15 cm (internal volume ~16.5 L, assuming 30% sediment porosity). P5-2 (1 cm radius) had an internal volume of 0.340 L. To achieve faster equilibration rates in HDE-4, seepage meters were cut to 4 cm tall and installed flush with the sediment (internal volume ~2 L). Dye testing was performed on September 13, 2017 in HDE-5 (one day prior to baseflow sampling; flow 28 L/s) and on October 22, 2017 in HDE-4 (2 weeks prior to storm sampling; flow 47 L/s).

Bromide tracer tests informed by preliminary dye testing were used to confirm flowpath connectivity and quantify HRT. Seepage meters were purged at 40 mL/min (twice the planned sampling rate) for 20 min to reduce dead volume. To introduce a compact Br- pulse into the hyporheic flowpath, a 1 L bromide solution (0.1 M Br-) was rapidly injected into the downwelling well (in HDE-5, poured via funnel over ~30 s; in HDE-4, pumped at 100 mL/min). Then, during continuous 20 mL/min pumping from the downwelling and upwelling locations, samples for Br-analysis were collected at 10 min intervals over 1–2 h in 50 mL polypropylene tubes. The slow sampling pump rate was selected based on approximate HZ flow rates (measured via volumetric bag fill rates) to ensure that pumping did not artificially impact HZ transport rates or hydraulics. Bromide testing at HDE-5 was performed over 1 h after dye tests on September 13, 2017. At HDE-4, bromide testing was performed over 2 h immediately after water sample collection (November 12, 2017). Bromide concentrations were measured with an HI4102 bromide combination ion selective electrode (Hanna Instruments, Woonsocket, RI), calibrated at 10^{-6}–0.1 M Br immediately before sample analysis (a calibration standard within the observed concentration range was re-analyzed every 4 samples, with any shift in the voltage used to correct measurements in real samples).

In HDE-4, a NaCl tracer was also injected into the downwelling well (monitored as electrical conductivity) at ~1,000x background to evaluate turnover rates in upwelling seepage meters on October 23 and 24, 2017. The downwelling well sample represented the injection solution rather than the resulting diluted concentration of the tracer-labeled hyporheic water. Given the uniformity of the streambed aggregate and relative impermeability of the underlying native sediment, we considered tracer presence/absence appropriate to establish hydrologic connectivity, rather than fitting the breakthrough curves with an advection-dispersion model. Our approach assumed that any dispersion along the hyporheic flowpath represented mixing with comparable parcels of hyporheic water, rather than dilution with clean groundwater. Electrical conductivity loggers deployed in the HDE-4 seepage meters also logged ambient specific conductivity every 3 min from October 23 to December 2, 2017, capturing several distinct specific conductivity peaks attributable to storms.

2.3. Water sample collection

Samples for HRMS analysis were collected in HDE-5 on September 14, 2017 (baseflow conditions, 28 L/s), and in HDE-4 on November 12, 2017 (stormflow conditions; flows turbid and elevated, maximum 110 L/s). The sampling design was based on dye tracer-based flowpath delineation to permit direct comparison between spatially paired flowpaths: surface flow immediately above the HZ versus transport through the HZ at different HRTs (Fig. 2).

Upwelling sites were purged at 40 mL/min for 40 min, then samples (downwelling well, upwelling seepage meters/piezometer, downstream surface water immediately above the seepage meter) were pumped at 20 mL/min over ~3.5 h to collect a 4 L integrated sample [Masterflex peristaltic pump (Masterflex 07528–10 pump, Masterflex 7519–05 4-channel pump head); 3/16" C-FLEX® tubing (Masterflex L/S 25); 4-L amber glass bottles]. Using a Lagrangian sampling design, sample collections at the upwelling locations were delayed (20–40 min, HRT estimated from dye tests) relative to the downwelling well. All tubing was rinsed with deionized (DI) water prior to sampling. At each HDE, a field blank (DI water) was pumped through the sampling system for >30 min. To ensure that the downwelling well sample provided a hydrologically representative influent to the HZ, we compared a second possible influent during each sampling event: surface water adjacent to the downwelling well (composited ~200 mL grabs over the 3.5 h sampling window). Based on the total number of non-target compounds, distribution of non-target peak area, and distribution of compound polarity, the two samples were very similar in terms of their chemical composition (RSD <16% for each metric) (Figure S5), and the downwelling well was used as a representative influent to both surface and hyporheic flowpaths for all subsequent analyses. All samples were transported on ice to the Center for Urban Waters (Tacoma, WA) and extracted within 24 h.

2.4. Sample processing and analysis

A complete list of chemicals and reagents is provided in the supplementary material. Sample processing and data analysis followed methods described previously (Du et al., 2017). SPE cartridges (3 mL, 100 mg Infinity SPE cartridges; ABS Materials, Wooster, OH, USA) were preconditioned with 3 mL 50% (v/v) methanol in DI water, then 25 mL DI water. Water samples (1 L) were loaded (without pre-filtration) at 5–10 mL/min, then cartridges were rinsed with DI water (10 mL), nitrogen-dried (15 min), and eluted with methanol (2x, 2.5 mL). Extracts were concentrated with nitrogen to 1 mL and spiked with a mixture of 12 isotope-labeled internal standards (to evaluate sample-specific matrix effects) (Table S1). Extracts were analyzed on an Agilent 1290 Infinity UHPLC system (Santa Clara, CA, USA) for separation and an Agilent 6530 Quadrupole Time-of-Flight LC-MS system with electrospray Jet Stream Technology for detection. Chromatography used a reverse-phase C18 analytical column (Agilent ZORBAX Eclipse Plus 2.1×100 mm, 1.8 μm particle size) with a C18 guard column at 45 °C, injection volume 5 μL, flow rate 0.4 mL/min, and a binary gradient of 5 mM ammonium acetate plus 0.1% acetic acid in water (A) and 5 mM ammonium acetate plus 0.1% acetic acid in MeOH (B) [5% B at 0–1 min, 50% B at 4 min, 100% B at 17–20 min, 5% B at 20.1 min;

stop time 22.5 min; post-time 2 min]. HRMS spectra were acquired across 100–1700 *m/z* (MS only) and 50–1700 *m/z* (MS/MS) in 2 GHz Extended Dynamic Range mode (collision-induced dissociation; data-dependent acquisition).

For quality assurance and quality control (QA/QC), we monitored detector performance via check tunes before each analytical run, re-tuning if mass error was >2 ppm. Every 8–12 samples, we analyzed solvent blanks (no column carryover detected), a mixture of external reference standards (cotinine (retention time (RT) 3.4 min, 120 ng/mL), carbamazepine (RT 6.5 min, 50 ng/mL), and prometryn (RT 9.5 min, 50 ng/mL)), and an internal standard mixture control (Table S1). The instrument was re-tuned and samples re-analyzed if mass accuracy was >5 ppm or area counts were >20% of initial sensitivity during the batch. Relative standard deviation (RSD) of area counts and RT were <16% and <1.5%, respectively, in both reference standard and internal standard blanks across all analytical batches. Method (DI water) and field blanks (DI water through sampling equipment) were extracted via SPE and analyzed alongside samples, with fold change analyses (see below) used to exclude any signals detected in blanks.

2.5. Data reduction and analysis

HRMS analyses detect 100s–1000s of non-target compounds (exact mass-retention time pairs) in each sample. We used Agilent software (Profinder B.08.00; Mass Profiler Professional B.13.00) to: 1) extract and align detections across samples, 2) filter the data to avoid false positives, and 3) prioritize detections for identification efforts (Du et al., 2017). Compounds with peak area >5000, occurring in ≥3 of 4 samples, and present at peak area ≥5-fold that of a solvent, method, or field blank were retained. This fold change criteria was selected to limit false negatives, as true detections can occur in blanks at low levels (e.g., pervasive contaminants like surfactants) (Bader et al., 2017). Methods to establish peak area-based (*vs.* concentration-based) limits of detection for non-target compounds still require reference standards (Li et al., 2018), limiting their applicability when the majority of detections remain unidentified. Thus, if a non-target compound did not meet the detection criteria, its peak area was set to a fixed value (zero) for data processing. Removal was estimated as reduction in peak area using Equation (1). (Nürenberg et al., 2015; Parry and Young, 2016).

$$Removal\ (R) = \frac{Peak\ Area_{Influent} - Peak\ Area_{Effluent}}{Peak\ Area_{Influent}} * 100\% \qquad (1)$$

Non-target compounds prioritized for identification were assigned molecular formulas and screened (via accurate-mass, isotope-pattern, and isotope abundance) against in-house stormwater and vehicle/roadway-related contaminant databases (~1200 total compounds). Identifications were scored against criteria proposed by Schymanski et al. to communicate detection confidence (Schymanski et al., 2014a). The highest confidence levels indicate that analyte retention time and fragmentation patterns match pure reference standards (S1), MS/MS spectra match library databases (S2a) or diagnostic information excludes other possible structures (S2b). For compounds identified at level S1, semi-quantitative concentration (as peak areas were not corrected by internal standard response) was estimated via 3 to 6-point calibration.

3. Results and discussion

3.1. Hyporheic flowpath delineation

3.1.1. HDE-5

Results of tracer testing are summarized in Table S2. Dye tracer tests in HDE-5 indicated a ~20–30 min HRT in the HZ for the bulk of the flow (Flowpath 5–1; Fig. 1b), and detected a slower secondary Flowpath 5–2 with an average HRT of ~30–60 min, likely due to finer sediments (deposited sand/silt) at the 5–2 upwelling location. Repeated dye tests consistently produced similar outcomes, indicating simple dye tests can quickly delineate HZ flowpath connectivity. The Flowpath 5-1 seepage meter produced an upwelling flow rate of ~1.2 L in 30 min (areal flux = 240 mL m^{-2} min^{-1}). The Flowpath 5–2 piezometer (P5-2) produced ~5 mL/min (areal flux = 17 mL m^{-2} min^{-1}). These distinct flux rates indicate considerable spatial heterogeneity across the HDE. The observed HRTs and flux rates indicated that, without pumping, seepage meter 5–1 and P5-2 required ~6 and ~1 h, respectively, to fully equilibrate with upwelling hyporheic water.

Subsequent bromide tracer tests at HDE-5 failed to improve the accuracy of our HRT estimates for these flowpaths. Injection concentrations may have been too low or breakthrough not monitored over an appropriate time interval, given the slow turnover rate of the seepage meter. Unfortunately, the bromide electrode used for on-line monitoring malfunctioned, preventing on-site correction. However, the baseflow samples are linked qualitatively, with HRTs estimated *via* dye tracer, and the close alignment (1 day) of dye testing and water quality sampling provided substantial confidence in the flowpath linkage, given steady state conditions.

Dye tests in HDE-5 one month later (October 20, 2017, in preparation for storm sampling) revealed extensive lateral shifts in the hyporheic flow field, as dye exited the HZ >3 m from seepage meter 5–1 at HRT <10 min, consistent with prior observations of focused lateral inflows at HDE-5 during the wet season. Hyporheic flowpaths are expected to shift seasonally as the hyporheic zone is compressed by rising riparian water tables. By visual observation, 20–30 cm of fine sediment had deposited over the HDE-5 seepage meters, also contributing to altered hydraulics. These conditions, especially the shorter HRT, redirected our sampling efforts to HDE-4.

3.1.2. HDE-4

For the surface flowpath in HDE-4, from the downwelling well to the surface above the upwelling zone, we visually estimated a ~1 min HRT. Two hyporheic flowpaths with HRTs of ~30 and ~40 min (Flowpaths 4–1 and 4–2, respectively) were delineated by dye testing (Fig. 1c). The observed hyporheic flow direction (generally from river right above HDE-4 towards river left below) was reasonable in light of similar head gradients observed in the lateral (right to left) and downstream directions (Fig. S2). In Flowpath 4–1, after seepage meter installation, the NaCl tracer injection (HRT 32 min; Fig. S6) and ambient specific conductivity from two late October storms (HRT ~30–60 min; Fig. S7) validated this estimate. During bromide testing (November 12, 2017; three weeks after dye testing), bromide breakthrough (>10x background) first occurred in Flowpath 4–1 at ~80 min (Fig. S8). Ambient conductivity data from six early November storms also reflected this longer HRT, with an 86 ± 13 min lag-time between the start of the reach and seepage meter 4–1 (Fig. S7).

Similar changes occurred in Flowpath 4–2, although bromide tracer data were complicated by the too-short 2 h tracer sampling design. NaCl tracer and ambient storm conductivity data indicated the 4–2 flowpath was still hydrologically linked to downwelling surface water during water sampling. The NaCl tracer data (Fig. S6) in Flowpath 4-2 showed an attenuated specific conductivity spike (10% of max in 4–1) at 195 min after the max peak in 4–1, indicating an HRT of ~3.75 h. This offset relative to Flowpath 4–1 aligned with the 140 ± 42 min offset observed during ambient specific conductivity storm peaks during late October and early November.

The increased HRT estimates in subsurface flowpaths suggested slower HZ water velocities during fall storm events (e.g., from reduced head drop across the HDE, clogging of upwelling zones by fine sediments mobilized during storms). Water elevations from stilling wells bracketing HDE-4 showed 21% decreases in downstream head drop (and similar decreases across lateral gradients) during the November 12, 2017 storm (Fig. S2), which would reduce hyporheic flow velocities correspondingly. In addition to shifting hydraulics, seepage meter installation may partially "cap" the HZ, forcing hyporheic flow to slow and/or partially re-direct around the seepage meters. Given the dramatic increase in HRT for Flowpath 4–2, we hypothesize that seepage meter 4-2 had slower rates of upwelling and thus equilibration than 4–1, artificially increasing HRT. To mitigate this possibility, we recommend the use of sampling equipment that minimizes internal volumes and maximizes flow (e.g., piezometers). Effectively capturing area- and time-integrated volumetric samples from linked upwelling or downwelling zones remains a substantial technical challenge in these dynamic systems. While the multi-tracer hydraulic sampling approach was effective, experiments should be carefully adapted to the site conditions and updated in close to real-time to ensure that collected data are accurate and fit to purpose.

Overall, the tracer detections in Flowpath 4-1 delineated hydrologic connectivity between downwelling and upwelling locations for subsequent water quality sampling, while the Flowpath 4-2 linkage and HRT estimate are less certain. Many contaminant peaks exhibited conservative behavior (see Section 3.3), supporting our assumption that upwelling samples were not meaningfully diluted by groundwater. Instead, the unlabeled upwelling water is assumed to be other hyporheic water of comparable residence time and influent quality. This assumption is reasonable in the homogeneous coarse fill in these HDEs (*vs.* a natural, heterogeneous HZ), and the relatively impermeable underlying sediment (compacted till). Thus, despite uncertainty in exact HRTs, comparison between short and long HRT in HDE-4 derived from dye and NaCl studies (and distance: Flowpath 4–2 was further downstream) was appropriate.

3.1.3. Hyporheic flow as a fraction of streamflow

Based on the most constrained HRT estimate (32 min in Flowpath 4–1), a head drop of 0.33 m over 4.25 m in the longitudinal direction ($dh/dx = 0.078$), a 4.4 m 3D hyporheic flowpath ($v = 0.0023$ m/s), and assumed 30% porosity, Darcy's Law yields a hydraulic conductivity (K) of 0.009 m/s, in line with typical rates for gravel (Fetter, 2001). Applying these data uniformly across the HDE cross-sectional area, a 5.9 L/s total hyporheic flux was estimated in each HDE (35 L/s across all six HDEs). Although the streambed aggregate is uniform along the study reach, this flux estimate is based on a single, relatively fast flowpath and does not account for seasonal fine sediment deposits or reductions in head drop; thus, it is an upper limit most relevant to baseflow conditions. Notably, the total hyporheic flux is quite similar to the 2016–2017 baseflow, indicating near equivalent volumes of surface water and HZ flow through the six HDEs. This estimate highlights the relevance of the engineered HDEs for reach-scale water quality, especially for seasonal first flush events. In 2017, the first four fall storms (prior to the October 2017 tracer test) generated discharges of 57–172 L/s. Therefore, we estimate 20–60% of these first flush events were likely to have passed through the engineered HDEs. Of course, the HDE flow capacity would represent a much smaller fraction of overall stream flow during larger winter storms due to increases in surface discharge and lateral groundwater inflows. Treating the entirety of such flows may not be practical with in-stream methods, although this also depends on the length of restored reach. The above flow fraction calculation does not include the several hundred meters of additional restored hyporheic zone in the reach beyond the six HDEs.

3.2. Chemical identification

We detected ~1900 individual non-target compounds during stormflow in HDE-4, while only 320 non-target compounds were detected during baseflow conditions in HDE-5. In the HDE-4 stormwater, 83 were identified at a confidence of ≥S2b as contaminants that are: 1) commonly detected in urban waters; 2) linked to specific sources (e.g., roadway runoff, residential areas); and/or 3) structurally related to chemicals strongly correlated with an urban stormwater-linked coho salmon acute mortality syndrome (McIntyre et al., 2018; Peter et al., 2018; Scholz et al., 2011). Of those, 27 were detected during baseflow conditions, although at peak areas consistently lower (~2–90-fold) than those observed during stormflow. The identified contaminants were grouped by structure and/or sources (Table S3), and peak area (indicating relative abundance) in the HDE-4 downwelling well (Fig. 3a).

Several organophosphates and pesticides that are pervasive anthropogenic indicators in urban watersheds (Bradley et al., 2017) were detected. In particular, tris(1-chloro-2-propyl)phosphate (TCPP, S2a) and tris(2-butoxyethyl)phosphate (TBOEP, S1; ~190 ng/L), used as flame retardants and paint additives, are known endocrine disruptors and reproductive toxicants (Meeker and Stapleton, 2010; van der Veen and de Boer, 2012), and were quite abundant (peak areas $>10^6$) during stormflow. During baseflow conditions, TCPP and TBOEP were again among the largest peak area detections, albeit at >8-fold lower abundance ($\sim 3 \times 10^5$). Cotinine (a nicotine metabolite), diuron (an herbicide and booster biocide), and DEET (an insect repellent) were each detected (S1) in stormflow at ~10 ng/L, while only DEET was detected during baseflow conditions (~5 ng/L).

Several contaminants from vehicles and roadway runoff were detected, including the commonly observed corrosion inhibitor 4(5)-methyl-1H-benzotriazole (S1; ~160 ng/L stormflow, ~70 ng/L baseflow) (LeFevre et al., 2015). Other vehicle-derived contaminants (all <9 ng/L during baseflow) included 1,3-diphenylguanidine (S1, ~125 ng/L in stormflow), 1,3-dicyclohexylurea (S1, ~60 ng/L), 1-cyclohexyl-3-phenylurea (S1, ~65 ng/L), N-methyl-dicyclohexylamine (S1, ~2 ng/L), N-cyclohexyl-benzothiazolamine (S1, ~3 ng/L), dicyclohexylamine (S2a), and 1,3-diphenylurea (S2a), which are used in tire rubber manufacturing or are reaction by-products (Wik and Dave, 2009). Hexa(methoxymethyl)melamine (HMMM, S1; ~150 ng/L stormflow, ~4 ng/L baseflow) and its structural family leach from tire rubbers, and are one of the most abundant (high peak area and detection frequency) organic contaminants detected in road runoff and urban receiving waters such as Thornton Creek (Peter et al., 2018). The significantly elevated peak areas observed during storm events suggest these roadway-related contaminants are potentially useful as stormwater-specific indicators.

Eight homologous series of poly(propylene glycols) (PPGs), poly(ethylene glycols) (PEGs), and octylphenol ethoxylates (OPEOs) (60 contaminants in total) were identified in the HDE-4 stormflow. Glycols are widely used in antifreeze, personal care products, and polymer formulations, and are degradation products of non-ionic surfactants (such as OPEOs), which are pervasive in aquatic environments (Ferguson et al., 2001; González et al., 2007). The PEG series consisted of saturated PEGs ($C_{2n\text{-}2}H_{4n\text{-}2}O_n$, n = 17–18, 21–23; S1), PEG dimethyl ethers ($C_{2n}H_{4n+2}O_n$, n = 4–7; S1), PEG monomethyl ethers ($C_{2n\text{-}1}H_{4n}O_n$, n = 4–15; S1), and PEG monobutyl ethers ($C_{2n+2}H_{4n+6}O_n$, n = 5–6; S2b). Detected OPEOs had formulas $C_{14}H_{22}O + C_{2n}H_{4n}O_n$ (n = 8–16, S1). The PPG series consisted of saturated PPGs ($C_{3n\text{-}3}H_{6n\text{-}4}O_n$, n = 5–18; S1), and two monounsaturated PPGs ($C_{3n\text{-}3}H_{6n\text{-}6}O_n$, n = 6–16, S2b and $C_{3n\text{-}6}H_{6n\text{-}12}O_n$,

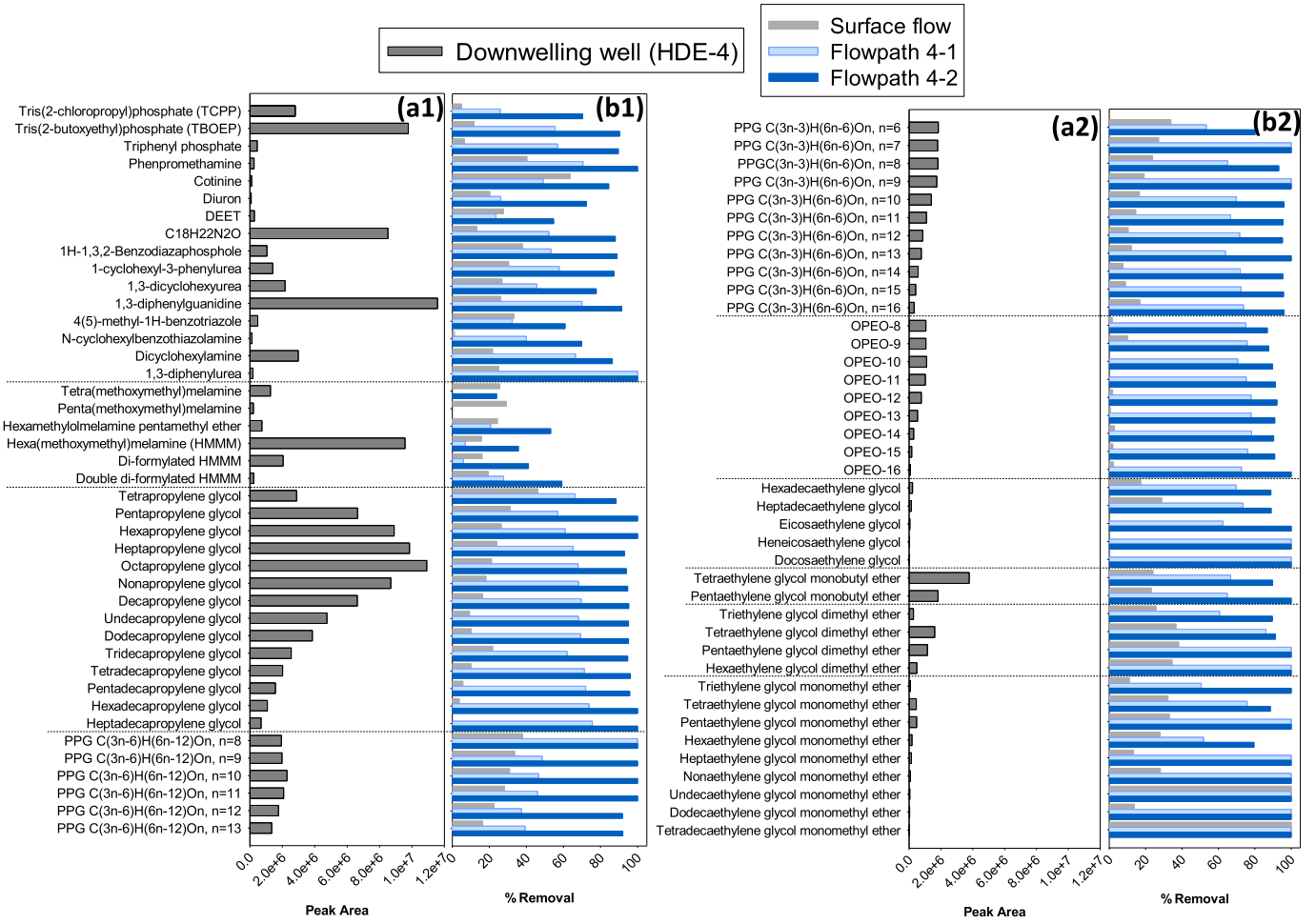

Fig. 3. Contaminants identified in Thornton Creek during stormflow in HDE-4, with (a) peak area in the HDE-4 downwelling well and (b) percent removal (based on change in peak area relative to the downwelling well) in the surface flowpath, hyporheic Flowpath 4–1 (shorter HRT), and hyporheic Flowpath 4–2 (longer HRT). Contaminants without removal data are those for which no removal (or an increase in peak area) was observed in that flowpath.

n = 8–13, S2b). While long-chain surfactants are commonly detected in wastewater and hydraulic fracking waters (Schymanski et al., 2014b; Thurman et al., 2017), this is among the first reports of long-chain PPGs in surface waters linked to urban runoff. Notably, the saturated PPGs were among the most abundant compounds present in the HDE-4 surface water, comprising 14% of total peak area and 42% of peak area for identified compounds (with an additional 5% and 14% of total and identified peak area, respectively, attributed to unsaturated PPGs). By comparison to available technical standards (using total peak area), we estimate ~40 ng/L OPEO, ~70 ng/L PEG, ~45 ng/L PEG dimethyl ether, ~35 ng/L PEG monomethyl ether, and ~700 ng/L PPG in HDE-4 during stormflow. In contrast, only 7 glycols/ethoxylates were detected (≤4 members in any series) during baseflow conditions, at ≥8-fold lower peak areas relative to stormflow. Thus, the glycol and ethoxylate families also represent potential surrogates/indicators for urban pollution.

Among the identified contaminants in stormflow, 27 overlap with a coho mortality chemical signature reflecting urban pollution (Peter et al., 2018). In HDE-5 during baseflow, 39% (22 compounds, total peak area ~1.7×10^6) of the mortality signature (both identified and unidentified compounds) was detected in the downwelling water. In HDE-4 during stormflow, and consistent with our expectation of increased coho mortality risk during storms in urbanized creeks, 86% of the mortality signature (50 non-target compounds, total peak area ~9×10^7) was detected.

3.3. Water quality assessment

To assess potential water quality improvement, we evaluated the HDEs as a set of paired in-stream "reactors" with two possible flowpaths that reflect the two potential treatment outcomes that may impact water quality in this reach: 1) surface attenuation, represented by surface flow immediately above the hyporheic paths; or 2) HZ treatment via flow through the subsurface HDE at varying HRTs (Fig. 2). We note that due to mixing of surface water and upwelling hyporheic water in the plunge pool (turnover time of several to tens of minutes) somewhat more attenuation in the sampled surface flowpath, relative to a hydraulically isolated 1-min surface flowpath, can be expected, although attenuation via mechanisms such as photolysis or sorption were not anticipated to be significant over such short timescales. A visual comparison between the HDE-4 influent and effluents is shown in Fig. S9, where a substantial reduction in turbidity was evident after transport through the hyporheic flowpaths.

We assessed water quality using both a "traditional analysis" focused on targeted analytes (within the HRMS data), and a complementary analysis focused on non-target data. Both approaches evaluated the fate of stormwater-derived contaminants along hydraulically paired surface and hyporheic flowpaths; the non-target HRMS data were used to assess treatment performance more broadly and evaluate the role of sorption-based attenuation

mechanisms in contaminant attenuation. On a practical note, while the "benefit" of HRMS analysis is the simultaneous detection of many contaminants, it is not typically feasible to acquire all relevant standards (even when pure standards are available) and quantify all detected contaminants. Therefore, water quality is typically compared on the basis of peak area reduction (Equation (1)) (Nürenberg et al., 2015; Parry and Young, 2016). Peak area comparison may over- or underestimate removal relative to the use of concentration (e.g., due to matrix enhancement or suppression), but provides relative comparisons across contaminants and flowpaths.

3.3.1. Removal of identified compounds

We first evaluated the fate of the 83 identified contaminants in Thornton Creek (Fig. 3b). Overall, consistently higher removal occurred in hyporheic Flowpath 4–1 relative to the paired surface flowpath. For example, the long-chain glycols and OPEOs were removed <38% during surface flow. In hyporheic Flowpath 4–1, removals increased by 47 ± 20% for PPGs to 37–100% removal, by 75 ± 5% for OPEOs (71–78% removal), and by 55 ± 29% for PEGs (51–100% removal). During baseflow conditions, PPGs were better removed via hyporheic Flowpath 5–1 (65–100%) relative to surface flow (14–23%) (Table S3). Similarly, the organophosphates persisted during surface flow (<15% removal), while moderate removal (26–57%) occurred along Flowpath 4–1. For vehicle and roadway-derived contaminants, removals <35% occurred during surface flow. With the exception of the HMMM family and 4(5)-methyl-1H-benzotriazole, removals in Flowpath 4–1 (40–100%) were 40 ± 18% higher than during surface flow.

Increased removals at the longer HRT of Flowpath 4–2 relative to Flowpath 4–1 were evident, indicating that longer HZ residence time contributed to contaminant attenuation and improved water quality (Fig. 3). Additional removal of the glycols and ethoxylates occurred with longer hyporheic flow, with removals exceeding 80% (and often >95%) for all PPGs, PEGs, and OPEOs in Flowpath 4–2, representing an increase of 22 ± 16% relative to Flowpath 4–1. Removal of the organophosphates also was 37 ± 6% higher in Flowpath 4–2 (70–90% removal) than in Flowpath 4–1. Removal of 4(5)-methyl-1H-benzotriazole and the HMMM family compounds (excepting the most polar and persistent members, tetra- and penta-MMM) was 31 ± 3% higher in Flowpath 4–2 than in Flowpath 4–1. For all other roadway-associated contaminants, removals exceeded 70% in Flowpath 4–2, an increase of 23 ± 11% relative to Flowpath 4–1. Based on the substantial peak areas of the identified contaminants in the HDE-4 stormflow, their effective attenuation in the HZ contributes significantly to reduction in anthropogenic contaminant loads to downstream reaches: assuming 20–60% of stormflow passes through the series of HDEs (see Section 3.1.3), the observed removals for all identified contaminants corresponds to a net ~15–50% reduction in the total load of these contaminants.

Observed removal of the coho mortality signature mirrored overall trends in the removal of identified contaminants. In HDE-4 (stormflow) and HDE-5 (baseflow), only 20% of the signature peak area detected in the downwelling well was removed during surface flow (one compound removed completely). In contrast, flow through hyporheic Flowpath 5–1 yielded a 51% reduction in signature peak area (two compounds removed completely). Likewise, signature peak area reductions of 55% and 84% occurred in Flowpaths 4–1 and 4–2, respectively (with 6 and 15 compounds removed completely). Thus, hyporheic flow (particularly at longer HRT) reduced the number and peak area of compounds correlated to the mortality phenomenon, although ongoing studies are assessing the potential for mortality signature outcomes to reflect actual reductions in risk to aquatic organisms.

3.3.2. Assessment of treatment performance with non-target compounds

After evaluating individual chemicals, the entire non-target dataset was used to assess treatment performance. We compared non-target compound peak area (Equation (1)) to estimate removal outcomes in paired samples, and applied predicted logK_{ow} characteristics to evaluate the potential contribution of sorption to attenuation in the HZ.

One metric of treatment performance is the total number of non-target compounds (Fig. S10). In HDE-4 during stormflow, 1198 compounds were detected. While 4% fewer compounds were

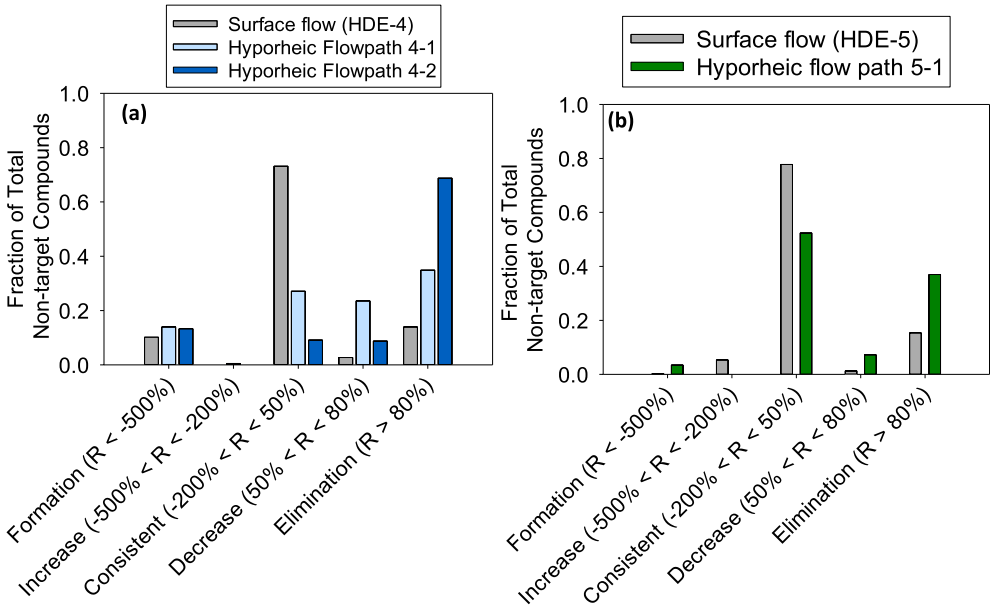

Fig. 4. Observed removal of non-target compounds (based on change in peak area relative to the corresponding downwelling well) in (a) HDE-4 during stormflow and (b) HDE-5 during baseflow, based on five categories for chemical fate: formation (R < −500%), increase (−500% < R < −200%), consistent (−200% < R < 50%), decrease (50% < R < 80%), and elimination (R > 80%). To facilitate comparison across the different flowpaths and HDEs, the fraction of total detections in each removal category is reported.

detected in downstream surface flow, 23% and 52% fewer compounds were detected after hyporheic flow at short (Flowpath 4–1) and long HRT (Flowpath 4–2), respectively. In HDE-5 during baseflow, 320 non-target compounds were detected in the downwelling well. In the downstream sample, ~15% fewer compounds were observed *vs.* 35% fewer compounds in Flowpath 5–1. Residual dye detected in Flowpath 5-2 during this event precluded non-target analysis of Flowpath 5-2 data. Overall, enhanced removal was evident for hyporheic *vs.* surface flow, and fewer compounds were detected as hyporheic HRT increased.

Per Bader et al. (2016), we grouped removal outcomes into defined intervals to categorize non-target compound fate. Notably, Bader et al. (2016) used peak height as a surrogate for concentration; our use of peak area is conceptually similar, but accounts for peak shape variability across different matrices and analytical runs. The intervals describe possible outcomes for extent of peak area removal: formation ($R < -500\%$), increase ($-500\% < R < -200\%$), consistent ($-200\% < R < 50\%$), decrease ($50\% < R < 80\%$), and elimination ($R > 80\%$). 'Formation' and 'increase' represent compounds that are newly detected during treatment (e.g., leaching, transformation products), while 'decrease' and 'elimination' represent compounds removed from the aqueous phase. The broad range for 'consistent' is intended to compensate for possible peak area variations derived from matrix effects (e.g., matrix suppression and enhancement), as well as small changes in compound concentration (Nürenberg et al., 2015). The reported range for 'consistent' was validated by spiking 12 isotopically-labeled internal standards into process influent and effluent waters (Fig. S11; Bader et al., 2016). The ratio of effluent/influent peak area ranged from 0.80 to 1.32 (75% of pairs within 0.85–1.15), indicating: 1) peak area comparisons to assess removal are appropriate, with consistent matrix bias expected throughout the treatment train (Li et al., 2018); and 2) the reported interval categories should be broadly applicable across surface waters and similar matrices.

Results of this analysis are shown for stormflow (Fig. 4a) and baseflow (Fig. 4b). For both, most non-target compounds in the downstream surface flowpath (shown in gray) were 'consistent' (i.e., no clear attenuation). Relative to surface flow, hyporheic flow (shown in blue/green) drove more compounds to the 'decrease' and 'elimination' categories. Notably, in HDE-4 stormflow, the impact of HRT on removal was clear. In Flowpath 4–1 (shorter HRT), 24% of non-target compounds decreased in peak area ($50\% < R < 80\%$) and 35% were eliminated ($R > 80\%$). In Flowpath 4–2 (longer HRT), 9% of compounds decreased while 69% were eliminated – longer hyporheic HRT clearly enhanced contaminant attenuation.

In HDE-4, ~13% of total non-target compounds in each flowpath were only detected downstream (i.e., "formation" interval; Fig. 4a). These new compounds could arise from transformation of influent compounds or leaching. A few new detections are probably analytical artifacts not captured by controls. For example, several poly-caprolactam compounds detected only in the HDE-4 upwelling samples likely leached from seepage meter materials (caprolactam is a monomer of nylon-6). While we did not conclusively detect a parent-product pair, we anticipate that many other new compounds are transformation products; for example, detection (S2b) of short-chain PEGs di- and triethylene glycol nonyl ether ($C_{13}H_{28}O_3$, RT 11.93 min; $C_{15}H_{32}O_4$, RT 12.1 min) (Fig. S12).

3.3.3. Assessing sorption potential with non-target data

Based on the relatively short HRT of the HZ flowpaths and previously observed organic micropollutant fate in HZs (Ding et al., 1999; Lin et al., 2006), hydrophobic sorption (e.g., to accumulated organic matter and subsurface biofilms) is expected to be a dominant removal mechanism. Without conclusively identifying each non-target detection, the contribution of sorption to attenuation was evaluated by using instrument RT to predict octanol-water partitioning coefficients ($\log K_{ow}$) (Bade et al., 2015b; Nurmi et al., 2012), *via* a previously established correlation built with 260 pure standards (Fig. S13, from Du et al., 2017). This approach relies on the statistical power of the hundreds-thousands of detected compounds to overcome expected uncertainties and errors arising from assignation of chemical characteristics to unidentified compounds. While some will be outliers from the $\log K_{ow}$ prediction (e.g., ionizable compounds), this approach supports evaluation of broader system trends and attenuation mechanisms using the extensive non-target dataset. Although more advanced modeling approaches (e.g., artificial neural networks, quantitative structure-retention time relationships) can improve the accuracy of $\log K_{ow}$ prediction (Bade et al., 2015a; Barron and McEneff, 2016), the simple linear correlation remains an effective predictor. For

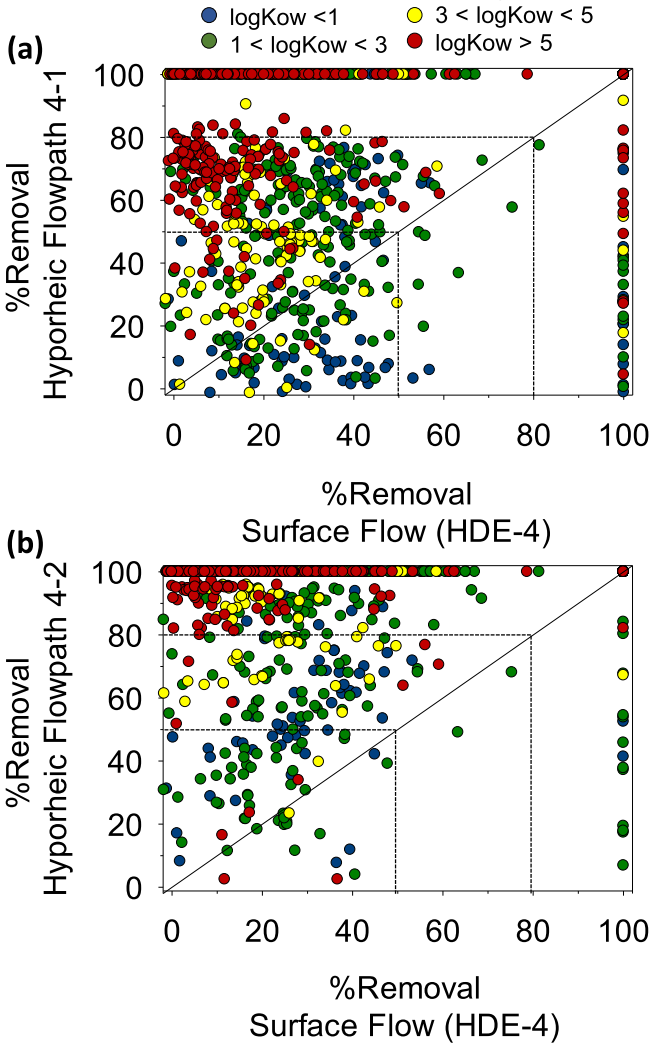

Fig. 5. Scatterplots showing relative removal of non-target compounds during storm flow in HDE-4 for (a) hyporheic Flowpath 4–1 (shorter HRT) *vs.* surface flow and (b) hyporheic Flowpath 4–2 (longer HRT) *vs.* surface flow. The plots only include compounds detected in the influent and with observed $R \geq 0\%$ (i.e., 'increase' and 'formation' categories are excluded). Points are colored according to estimated $\log K_{ow}$ range, with $\log K_{ow}$ increasing from blue to red, to evaluate the effect of sorption on observed removal. $\log K_{ow}$ ranges are at least 2 log units wide (i.e., at least half the 95% prediction interval of the linear relationship between $\log K_{ow}$ and instrument RT). Dashed lines at 50% and 80% removal delineate boundaries for the 'consistent', 'decrease', and 'elimination' categories described in text. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

example, Bade et al. (2015b) demonstrated that for 95% of 595 pesticides and pharmaceuticals, actual RT was within 4 min of the prediction (and for a 345-point, pesticide-only regression, 79% were within 2 min of the prediction). Similarly, for our regression, actual and predicted RT were within 2 and 4 min for 78% and 96% of compounds, respectively.

To depict removal trends as a function of predicted sorption potential for the different flowpaths, we compared relative removal performance using scatterplots (HDE-4 stormflow data in Fig. 5, HDE-5 baseflow data in Fig. S14) *via* Parry and Young (2016). The 1:1 line indicates identical removal in the two flowpaths, while compounds above the 1:1 line were more effectively removed during hyporheic flow than surface flow (and *vice versa*). A second visualization approach, not described previously, supports a statistical comparison across logK_{ow} ranges and treatment types. This approach used box-plots to evaluate removal as a function of estimated logK_{ow}, with median removals extracted as a single-point comparison of removal performance across logK_{ow} ranges (HDE-4 in Fig. 6, HDE-5 in Fig. S15).

Both approaches indicated substantial water quality improvement in hyporheic flowpaths, relative to limited removal during surface flow, and strong contributions of hydrophobic sorption towards observed attenuation. In the scatterplots (Figs. 5 and S16), more polar compounds (estimated logK_{ow} <3, blue and green) primarily lay along the 1:1 line at removals <50%, while more hydrophobic compounds (estimated logK_{ow} >3, yellow and red) clearly exhibited much higher removal along hyporheic flowpaths. In hyporheic Flowpath 4–1, as logK_{ow} increased, the distribution visibly shifted towards 'elimination' (Fig. 5a), indicating the HZ was increasingly effective for removal of more hydrophobic compounds, particularly during stormflow.

This trend was even more apparent in the longer-HRT hyporheic Flowpath 4–2 (Fig. 5b). Non-target detections in all logK_{ow} ranges exhibited greater removal in HZ flow than surface flow, and nearly all detections with logK_{ow} >5 (in red) exhibited near 100% removal in hyporheic flow (clusters along the top edge of the plot). Thus, by comparison to the shorter hyporheic flowpath, increasing HRT in the HZ promoted enhanced removal of more hydrophobic compounds and was especially important for increased removal of polar compounds. These trends were reflected by the removal of identified compounds described above (scatterplots for the 83 identified compounds provided in Fig. S16). For example, removal of the polar 4(5)-methyl-1H-benzotriazole (logK_{ow} 1.8) was similar in surface flow and Flowpath 4–1 (33% and 32%, respectively), but increased to 61% in Flowpath 4–2. The organophosphates exemplify trends observed in HDE-4 for non-polar compounds: removals of TBOEP (logK_{ow} 3.75) and triphenyl phosphate (logK_{ow} 4.59) increased from 12% and 7%, respectively, during surface flow to 55% and 57%, respectively, in Flowpath 4–1, with even higher removal (90% for both) in Flowpath 4–2.

In the boxplots (Fig. 6, S15), two statistically significant trends were apparent in the data: (1) increasing removal with compound hydrophobicity in Flowpaths 4–1 and 4–2 (Dunn's test, $p < 0.01$ for all comparisons except logK_{ow} 1–3 *vs.* 3–5 in Flowpath 4–1 ($p = 0.41$) and logK_{ow} <1 *vs.* 1–3 in Flowpath 4–2 ($p = 0.36$)), and (2) increasing removal with hyporheic HRT (Dunn's test, $p < 0.01$

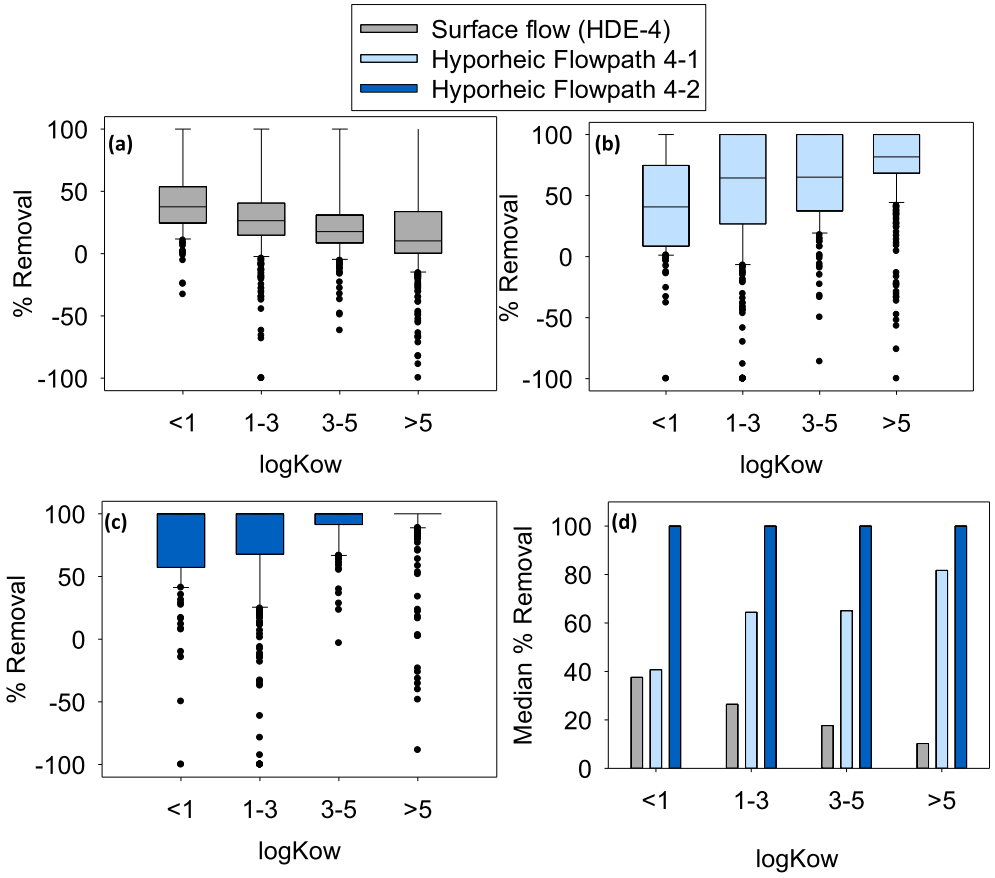

Fig. 6. Box-plots showing peak area removal of non-target compounds in HDE-4 during stormflow for (a) surface flow, (b) the shorter hyporheic flowpath, and (c) the longer hyporheic flowpath. Median removal for each range (extracted from the box-plots) is shown in panel (d). The box-plots only include compounds present in the downwelling well (i.e., detections completely unique to a surface or hyporheic flow "effluent" are excluded). For any compound with R < −100%, the value was set to −100% to avoid skewing the data. Box: 25th to 75th percentile; whiskers: 10th to 90th percentile; black dots: outliers; line within box: median (when the median line is not visible within a box, the value is 100%).

for all comparisons except removal in surface flow *vs.* Flowpath 4–1 at logK_{ow} <1). Thus, sorption to immobile geomedia, biofilms, and organic matter phases was a dominant removal mechanism in the HZ. However, we observed more variability in removal performance within the lower logK_{ow} ranges. This is consistent with the potential for additional molecular characteristics (e.g., functional groups, ionizability) and removal mechanisms (e.g., electrostatic interactions) to govern the fate outcomes of more polar compounds (Schaffer et al., 2012; Tülp et al., 2009).

3.3.4. Relevance for hyporheic restoration

Identifying and sampling hydrologically paired downwelling and upwelling locations in the HZ is challenging, but is needed to quantify the treatment capabilities of the restored HZ. Variability in the hyporheic residence times of individual flowpaths (observed during timescales of days to weeks) were attributed to both the changing water table during storm events, siltation, and artificial flowpath alterations caused by seepage meter installation, although longer-term studies are needed to fully evaluate these phenomena. In the Kingfisher HDEs (a few meters in length, 0.5–~3 h HRT), we observed >50% attenuation of hundreds of organic stormwater contaminants. The data also indicated increasing contaminant removal efficiency with longer hyporheic HRT, and that hydrophobic sorption processes to geomedia and organic matter phases are key contributors to rapid water quality improvement in HDEs. Although contaminants may degrade microbially over longer timescales after sorption, this requires further investigation. While these results are specific to a subset of emerging organic contaminants, we anticipate that HDEs also attenuate other, regulated pollutants (e.g., nutrients, pathogens, metals). At the reach scale, we estimated that the six Kingfisher HDEs can treat a water volume similar to the total surface discharge during summer baseflow and early hydrograph, first-flush fall storm events. The Kingfisher HDEs have several novel design features to lengthen HRT, but further improvements (e.g., longer caps, lateral groundwater seals) could improve treatment performance and reliability of water quality monitoring. While the potential for water quality improvement within individual HDEs is clear, the primary technical challenge for these systems is hydraulic: promoting the passage of sufficient volumes through HDEs (while maintaining adequate contact times with biofilms and media) to meaningfully improve reach-scale water quality, particularly during storm events that reduce the head drop across the HDE while introducing significant contaminant loads. Future work should include longer-term studies to assess the repeatability of HZ performance over time and should optimize hyporheic exchange to meet reach scale water quality goals, with HRT matched to contaminant attenuation timescales. Further, the HDE subsurface media could be amended with specific sorbents (e.g., biochar) to improve attenuation of high priority contaminants, thereby allowing higher throughput rates to provide proportionally larger reductions in the contaminant load. Continued research is needed to develop effective and reliable options for in-stream treatment of surface waters via HZ engineering to promote high quality and healthy urban freshwater ecosystems.

4. Conclusions

- Linked hyporheic flowpaths were identified for subsequent water quality sampling in an engineered hyporheic design element using dye, NaCl, and bromide tracer tests. Hydraulic retention times in the sampled HZ ranged from ~30 min to >3 h over 3–5 m hyporheic travel.
- Removal data for non-target compounds across surface and hyporheic flowpaths, including identified compounds linked to vehicle/roadway sources and a coho mortality chemical signature, demonstrated that hyporheic flow attenuates hundreds of organic compounds, particularly at longer HRT, thus improving water quality relative to surface transport.
- Non-target HRMS data can assess holistic water quality improvements using metrics such as total number and total peak area to describe overall system trends. Additionally, non-target compound hydrophobicity, estimated via RT-logK_{ow} relationships, can be used to evaluate the potential contributions of hydrophobic sorption to attenuation.
- During both baseflow and stormflow, the HZ is increasingly effective for removal of increasingly hydrophobic compounds, indicating that sorption is a dominant removal mechanism. Longer hydraulic retention times promote compound removal, particularly for more polar compounds.

Declarations of interest

None.

Acknowledgements

This work was supported by funding from Seattle Public Utilities and the University of Washington (Tacoma-Seattle). The authors gratefully acknowledge Paul Bakke (US Fish and Wildlife) and Steve Damm (Seattle Public Utilities) for their insight and assistance at the Thornton Creek field site, Sally Landefeld for her help with water sampling, and Nina Zhao for her help with bromide measurements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watres.2018.11.050.

References

Abbott, B.W., Baranov, V., Mendoza-Lera, C., Nikolakopoulou, M., Harjung, A., Kolbe, T., Balasubramanian, M.N., Vaessen, T.N., Ciocca, F., Campeau, A., Wallin, M.B., Romeijn, P., Antonelli, M., Gonçalves, J., Datry, T., Laverman, A.M., de Dreuzy, J.-R., Hannah, D.M., Krause, S., Oldham, C., Pinay, G., 2016. Using multi-tracer inference to move beyond single-catchment ecohydrology. Earth Sci. Rev. 160, 19–42. https://doi.org/10.1016/j.earscirev.2016.06.014.

Azinheira, D.L., Scott, D.T., Hession, W., Hester, E.T., 2014. Comparison of effects of inset floodplains and hyporheic exchange induced by in-stream structures on solute retention. Water Resour. Res. 50, 6168–6190. https://doi.org/10.1002/2013WR014400.

Bade, R., Bijlsma, L., Miller, T.H., Barron, L.P., Sancho, J.V., Hernández, F., 2015a. Suspect screening of large numbers of emerging contaminants in environmental waters using artificial neural networks for chromatographic retention time prediction and high resolution mass spectrometry data analysis. Sci. Total Environ. 538, 934–941. https://doi.org/10.1016/j.scitotenv.2015.08.078.

Bade, R., Bijlsma, L., Sancho, J.V., Hernández, F., 2015b. Critical evaluation of a simple retention time predictor based on LogKow as a complementary tool in the identification of emerging contaminants in water. Talanta 139, 143–149. https://doi.org/10.1016/j.talanta.2015.02.055.

Bader, T., Schulz, W., Kümmerer, K., Winzenbacher, R., 2017. LC-HRMS data processing strategy for reliable sample comparison exemplified by the assessment of water treatment processes. Anal. Chem. 89, 13219–13226. https://doi.org/10.1021/acs.analchem.7b03037.

Bader, T., Schulz, W., Kümmerer, K., Winzenbacher, R., 2016. General strategies to increase the repeatability in non-target screening by liquid chromatography-high resolution mass spectrometry. Anal. Chim. Acta 935, 173–186. https://doi.org/10.1016/j.aca.2016.06.030.

Bakke, P.D., 2018. Thornton Creek Floodplain Restoration: Effectiveness Monitoring of Physical Processes. Draft Report to City of Seattle, Seattle Public Utilities.

Barron, L.P., McEneff, G.L., 2016. Gradient liquid chromatographic retention time prediction for suspect screening applications: a critical assessment of a generalised artificial neural network-based approach across 10 multi-residue reversed-phase analytical methods. Talanta 147, 261–270. https://doi.org/10.1016/j.talanta.2015.09.065.

Bradley, P.M., Journey, C.A., Romanok, K.M., Barber, L.B., Buxton, H.T., Foreman, W.T.,

Furlong, E.T., Glassmeyer, S.T., Hladik, M.L., Iwanowicz, L.R., Jones, D.K., Kolpin, D.W., Kuivila, K.M., Loftin, K.A., Mills, M.A., Meyer, M.T., Orlando, J.L., Reilly, T.J., Smalling, K.L., Villeneuve, D.L., 2017. Expanded target-chemical analysis reveals extensive mixed-organic-contaminant exposure in U.S. streams. Environ. Sci. Technol. 51, 4792–4802. https://doi.org/10.1021/acs.est.7b00012.

Ding, W.-H., Wu, J., Semadeni, M., Reinhard, M., 1999. Occurrence and behavior of wastewater indicators in the Santa Ana River and the underlying aquifers. Chemosphere 39, 1781–1794. https://doi.org/10.1016/S0045-6535(99)00072-7.

Du, B., Lofton, J.M., Peter, K.T., Gipe, A.D., James, C.A., McIntyre, J.K., Scholz, N.L., Baker, J.E., Kolodziej, E.P., 2017. Development of suspect and non-target screening methods for detection of organic contaminants in highway runoff and fish tissue with high-resolution time-of-flight mass spectrometry. Env. Sci Process. Impacts 19, 1185. https://doi.org/10.1039/C7EM00243B.

Ferguson, P.L., Iden, C.R., Brownawell, B.J., 2001. Distribution and fate of neutral alkylphenol ethoxylate metabolites in a sewage-impacted urban estuary. Environ. Sci. Technol. 35, 2428–2435. https://doi.org/10.1021/es001871b.

Fetter, C.W., 2001. In: Applied Hydrogeology, fourth ed. Prentice Hall, Upper Saddle River, N.J.

Fischer, H., Kloep, F., Wilzcek, S., Pusch, M.T., 2005. A river's liver – microbial processes within the hyporheic zone of a large lowland river. Biogeochemistry 76, 349–371. https://doi.org/10.1007/s10533-005-6896-y.

González, S., Barceló, D., Petrovic, M., 2007. Advanced liquid chromatography-mass spectrometry (LC-MS) methods applied to wastewater removal and the fate of surfactants in the environment. TrAC Trends Anal. Chem. (Reference Ed.) 26, 116–124. https://doi.org/10.1016/j.trac.2006.12.003.

Grebel, J.E., Mohanty, S.K., Torkelson, A.A., Boehm, A.B., Higgins, C.P., Maxwell, R.M., Nelson, K.L., Sedlak, D.L., 2013. Engineered infiltration systems for urban stormwater reclamation. Environ. Eng. Sci. 30, 437–454. https://doi.org/10.1089/ees.2012.0312.

Harvey, J.W., Böhlke, J.K., Voytek, M.A., Scott, D., Tobias, C.R., 2013. Hyporheic zone denitrification: controls on effective reaction depth and contribution to whole-stream mass balance: scaling hyporheic flow controls on stream denitrification. Water Resour. Res. 49, 6298–6316. https://doi.org/10.1002/wrcr.20492.

Herzog, S., Higgins, C.P., Singha, K., McCray, J., 2018. Performance of engineered streambeds for inducing hyporheic transient storage and attenuation of resazurin. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.8b01145.

Herzog, S.P., Higgins, C.P., McCray, J.E., 2016. Engineered streambeds for induced hyporheic flow: enhanced removal of nutrients, pathogens, and metals from urban streams. J. Environ. Eng. 142, 04015053. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001012.

Hester, E.T., Gooseff, M.N., 2010. Moving beyond the banks: hyporheic restoration is fundamental to restoring ecological services and functions of streams. Environ. Sci. Technol. 44, 1521–1525. https://doi.org/10.1021/es902988n.

Hester, E.T., Hammond, B., Scott, D.T., 2016. Effects of inset floodplains and hyporheic exchange induced by in-stream structures on nitrate removal in a headwater stream. Ecol. Eng. 97, 452–464. https://doi.org/10.1016/j.ecoleng.2016.10.036.

Hoehn, E., Plumlee, M.H., Reinhard, M., 2007. Natural attenuation potential of downwelling streams for perfluorochemicals and other emerging contaminants. Water Sci. Technol. 56, 59. https://doi.org/10.2166/wst.2007.804.

Hollender, J., Schymanski, E.L., Singer, H.P., Ferguson, P.L., 2017. Nontarget screening with high resolution mass spectrometry in the environment: ready to go? Environ. Sci. Technol. 51, 11505–11512. https://doi.org/10.1021/acs.est.7b02184.

Huntscha, S., Rodriguez Velosa, D.M., Schroth, M.H., Hollender, J., 2013. Degradation of Polar organic micropollutants during riverbank filtration: complementary results from spatiotemporal sampling and push-pull tests. Environ. Sci. Technol. 47, 11512–11521. https://doi.org/10.1021/es401802z.

Knapp, J.L.A., González-Pinzón, R., Drummond, J.D., Larsen, L.G., Cirpka, O.A., Harvey, J.W., 2017. Tracer-based characterization of hyporheic exchange and benthic biolayers in streams: HYPORHEIC EXCHANGE AND BENTHIC BIOLAYERS. Water Resour. Res. 53, 1575–1594. https://doi.org/10.1002/2016WR019393.

Krause, S., Lewandowski, J., Grimm, N.B., Hannah, D.M., Pinay, G., McDonald, K., Martí, E., Argerich, A., Pfister, L., Klaus, J., Battin, T., Larned, S.T., Schelker, J., Fleckenstein, J., Schmidt, C., Rivett, M.O., Watts, G., Sabater, F., Sorolla, A., Turk, V., 2017. Ecohydrological interfaces as hot spots of ecosystem processes: ecohydrological interfaces as hot spots. Water Resour. Res. 53, 6359–6376. https://doi.org/10.1002/2016WR019516.

Kunkel, U., Radke, M., 2011. Reactive tracer test to evaluate the fate of pharmaceuticals in rivers. Environ. Sci. Technol. 45, 6296–6302. https://doi.org/10.1021/es104320n.

Kunkel, U., Radke, M., 2008. Biodegradation of acidic pharmaceuticals in bed sediments: insight from a laboratory experiment. Environ. Sci. Technol. 42, 7273–7279. https://doi.org/10.1021/es801562j.

Lawrence, J.E., Skold, M.E., Hussain, F.A., Silverman, D.R., Resh, V.H., Sedlak, D.L., Luthy, R.G., McCray, J.E., 2013. Hyporheic zone in urban streams: a review and opportunities for enhancing water quality and improving aquatic habitat by active management. Environ. Eng. Sci. 30, 480–501. https://doi.org/10.1089/ees.2012.0235.

LeFevre, G.H., Paus, K.H., Natarajan, P., Gulliver, J.S., Novak, P.J., Hozalski, R.M., 2015. Review of dissolved pollutants in urban storm water and their removal and fate in bioretention cells. J. Environ. Eng. 141, 04014050. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000876.

Lewandowski, J., Putschew, A., Schwesig, D., Neumann, C., Radke, M., 2011. Fate of organic micropollutants in the hyporheic zone of a eutrophic lowland stream: results of a preliminary field study. Sci. Total Environ. 409, 1824–1835. https://doi.org/10.1016/j.scitotenv.2011.01.028.

Li, Z., Undeman, E., Papa, E., McLachlan, M.S., 2018. High-throughput evaluation of organic contaminant removal efficiency in a wastewater treatment plant using direct injection UHPLC-Orbitrap-MS/MS. Environ. Sci. Process. Impacts 20, 561–571. https://doi.org/10.1039/C7EM00552K.

Lin, A.Y.-C., Plumlee, M.H., Reinhard, M., 2006. Natural attenuation of pharmaceuticals and alkylphenol polyethyoxylate metabolites during river transport: photochemical and biological transformation. Environ. Toxicol. Chem. 25, 1458. https://doi.org/10.1897/05-412R.1.

McEachran, A.D., Hedgespeth, M.L., Newton, S.R., McMahen, R., Strynar, M., Shea, D., Nichols, E.G., 2018. Comparison of emerging contaminants in receiving waters downstream of a conventional wastewater treatment plant and a forest-water reuse system. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-018-1505-5.

McIntyre, J.K., Lundin, J.I., Cameron, J.R., Chow, M.I., Davis, J.W., Incardona, J.P., Scholz, N.L., 2018. Interspecies variation in the susceptibility of adult Pacific salmon to toxic urban stormwater runoff. Environ. Pollut. 238, 196–203. https://doi.org/10.1016/j.envpol.2018.03.012.

Meeker, J.D., Stapleton, H.M., 2010. House dust concentrations of organophosphate flame retardants in relation to hormone levels and semen quality parameters. Environ. Health Perspect. 118, 318–323. https://doi.org/10.1289/ehp.0901332.

Mendoza-Lera, C., Datry, T., 2017. Relating hydraulic conductivity and hyporheic zone biogeochemical processing to conserve and restore river ecosystem services. Sci. Total Environ. 579, 1815–1821. https://doi.org/10.1016/j.scitotenv.2016.11.166.

Merel, S., Anumol, T., Park, M., Snyder, S.A., 2015. Application of surrogates, indicators, and high-resolution mass spectrometry to evaluate the efficacy of UV processes for attenuation of emerging contaminants in water. J. Hazard Mater. 282, 75–85. https://doi.org/10.1016/j.jhazmat.2014.09.008.

Nürenberg, G., Schulz, M., Kunkel, U., Ternes, T.A., 2015. Development and validation of a generic nontarget method based on liquid chromatography – high resolution mass spectrometry analysis for the evaluation of different wastewater treatment options. J. Chromatogr. A 1426, 77–90. https://doi.org/10.1016/j.chroma.2015.11.014.

Nurmi, J., Pellinen, J., Rantalainen, A.-L., 2012. Critical evaluation of screening techniques for emerging environmental contaminants based on accurate mass measurements with time-of-flight mass spectrometry: critical evaluation of TOF-MS screening techniques. J. Mass Spectrom. 47, 303–312. https://doi.org/10.1002/jms.2964.

Parry, E., Young, T.M., 2016. Comparing targeted and non-targeted high-resolution mass spectrometric approaches for assessing advanced oxidation reactor performance. Water Res. 104, 72–81. https://doi.org/10.1016/j.watres.2016.07.056.

Peter, K.T., Tian, Z., Wu, C., Lin, P., White, S., Du, B., McIntyre, J.K., Scholz, N.L., Kolodziej, E.P., 2018. Using high-resolution mass spectrometry to identify organic contaminants linked to urban stormwater mortality syndrome in coho salmon. Environ. Sci. Technol. 52, 10317–10327. https://doi.org/10.1021/acs.est.8b03287.

Posselt, M., Jaeger, A., Schaper, J.L., Radke, M., Benskin, J.P., 2018. Determination of polar organic micropollutants in surface and pore water by high-resolution sampling-direct injection-ultra high performance liquid chromatography-tandem mass spectrometry. Environ. Sci. Process. Impacts. https://doi.org/10.1039/C8EM00390D.

Runkel, R.L., 1998. One-dimensional Transport with Inflow and Storage (OTIS): a Solute Transport Model for Streams and Rivers. https://doi.org/10.3133/wri984018.

Schaffer, M., Börnick, H., Nödler, K., Licha, T., Worch, E., 2012. Role of cation exchange processes on the sorption influenced transport of cationic β-blockers in aquifer sediments. Water Res. 46, 5472–5482. https://doi.org/10.1016/j.watres.2012.07.013.

Schaper, J.L., Seher, W., Nützmann, G., Putschew, A., Jekel, M., Lewandowski, J., 2018. The fate of polar trace organic compounds in the hyporheic zone. Water Res. 140, 158–166. https://doi.org/10.1016/j.watres.2018.04.040.

Schollée, J.E., Bourgin, M., von Gunten, U., McArdell, C.S., Hollender, J., 2018. Non-target screening to trace ozonation transformation products in a wastewater treatment train including different post-treatments. Water Res. 142, 267–278. https://doi.org/10.1016/j.watres.2018.05.045.

Scholz, N.L., Myers, M.S., McCarthy, S.G., Labenia, J.S., McIntyre, J.K., Ylitalo, G.M., Rhodes, L.D., Laetz, C.A., Stehr, C.M., French, B.L., McMillan, B., Wilson, D., Reed, L., Lynch, K.D., Damm, S., Davis, J.W., Collier, T.K., 2011. Recurrent die-offs of adult coho salmon returning to spawn in Puget Sound lowland urban streams. PloS One 6, e28013. https://doi.org/10.1371/journal.pone.0028013.

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014a. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ. Sci. Technol. 48, 2097–2098. https://doi.org/10.1021/es5002105.

Schymanski, E.L., Singer, H.P., Longrée, P., Loos, M., Ruff, M., Stravs, M.A., Ripollés Vidal, C., Hollender, J., 2014b. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. Environ. Sci. Technol. 48, 1811–1818. https://doi.org/10.1021/es4044374.

Thurman, E.M., Ferrer, I., Rosenblum, J., Linden, K., Ryan, J.N., 2017. Identification of polypropylene glycols and polyethylene glycol carboxylates in flowback and produced water from hydraulic fracturing. J. Hazard Mater. 323, 11–17. https://

doi.org/10.1016/j.jhazmat.2016.02.041.
Tülp, H.C., Fenner, K., Schwarzenbach, R.P., Goss, K.-U., 2009. pH-dependent sorption of acidic organic chemicals to soil organic matter. Environ. Sci. Technol. 43, 9189–9195. https://doi.org/10.1021/es902272j.
van der Veen, I., de Boer, J., 2012. Phosphorus flame retardants: properties, production, environmental occurrence, toxicity and analysis. Chemosphere 88, 1119–1153. https://doi.org/10.1016/j.chemosphere.2012.03.067.
Ward, A.S., Schmadel, N.M., Wondzell, S.M., Gooseff, M.N., Singha, K., 2017. Dynamic hyporheic and riparian flow path geometry through base flow recession in two headwater mountain stream corridors: hyporheic potential model. Water Resour. Res. 53, 3988–4003. https://doi.org/10.1002/2016WR019875.
Wik, A., Dave, G., 2009. Occurrence and effects of tire wear particles in the environment – a critical review and an initial risk assessment. Environ. Pollut. 157, 1–11. https://doi.org/10.1016/j.envpol.2008.09.028.
Zarnetske, J.P., Haggerty, R., Wondzell, S.M., Baker, M.A., 2011. Dynamics of nitrate production and removal as a function of residence time in the hyporheic zone. J. Geophys. Res. 116. https://doi.org/10.1029/2010JG001356.