Open Access Articles

Predicting Polycyclic Aromatic Hydrocarbon Concentrations in Resident Aquatic Organisms Using Passive Samplers and Partial Least-Squares Calibration

Oregon State
UNIVERSITY

Predicting Polycyclic Aromatic Hydrocarbon Concentrations in Resident Aquatic Organisms Using Passive Samplers and Partial Least-Squares Calibration

Norman D. Forsberg, Brian W. Smith, Greg J. Sower, and Kim A. Anderson*

Environmental and Molecular Toxicology Department, Oregon State University, ALS 1007, Corvallis, Oregon 97331, United States

S Supporting Information

ABSTRACT: The current work sought to develop predictive models between time-weighted average polycyclic aromatic hydrocarbon (PAH) concentrations in the freely dissolved phase and those present in resident aquatic organisms. We deployed semipermeable membrane passive sampling devices (SPMDs) and collected resident crayfish (*Pacifastacus leniusculus*) at nine locations within and outside of the Portland Harbor Superfund Mega-site in Portland, OR. Study results show that crayfish and aqueous phase samples collected within the Mega-site had PAH profiles enriched in high molecular weight PAHs and that freely dissolved PAH profiles tended to be more populated by low molecular weight PAHs compared to crayfish tissues. Results also show that of several modeling approaches, a two-factor partial least-squares (PLS) calibration model using detection limit substitution provided the best predictive power for estimating PAH concentrations in crayfish, where the model explained ≥72% of the variation in the data set and provided predictions within ~3× of measured values. Importantly, PLS calibration provided a means to estimate PAH concentrations in tissues when concentrations were below detection in the freely dissolved phase. The impact of measurements below detection limits is discussed.

SPMDs + PLS calibration predicts crayfish PAHs
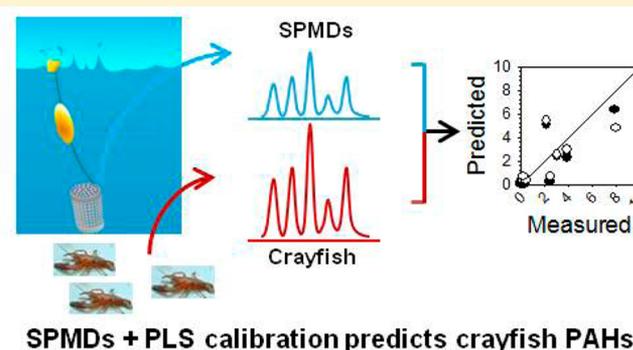
INTRODUCTION

Health officials are tasked with identifying and mitigating uncontrolled releases of chemicals that pose unacceptable levels of risk to public health, such as exposure to polycyclic aromatic hydrocarbons (PAHs) at Superfund sites or following an oil spill. In cases where the public can be exposed to contaminated fish and/or shellfish, it is necessary to collect and evaluate tissue levels of hazardous chemicals in order to protect consumers.[1,2] However, obtaining necessary amounts of tissue data is logistically challenging, time-consuming, expensive, and can further stress compromised ecological communities.[3−6] Furthermore, careful consideration must be given to organism home range, sample availability, and spatial and temporal contamination trends in order to obtain samples that capture site-representative variability in tissue concentrations.[3,7] Having a predictive approach to estimate concentrations of hazardous chemicals in aquatic organism tissues would be useful to risk assessors in instances when tissue data are limited.

One approach to estimating chemical concentrations in aquatic biota is through the application of empirical steady-state bioaccumulation factors (BAFs), typically calculated as the quotient of tissue concentrations, expressed on a wet weight, dry weight, or lipid normalized basis, and total aqueous phase concentrations.[8,9] However, applying published BAFs to predict site-specific contaminant concentrations in aquatic organisms can produce estimates that vary by orders of magnitude for contaminants of a given log K_{ow}.[7,9] This is due in part to variable chemical bioavailability,[8,9] physiological differences between organisms used to derive BAFs and those of interest in risk assessments, and the error incurred from applying average regression-based BAFs to site-specific data.[9] Additionally, using BAFs to estimate chemical concentrations in tissues is particularly problematic when concentrations in the aqueous phase are below detection limits.

Passive samplers, such as semipermeable membrane devices (SPMDs), uptake and concentrate the bioavailable fraction of lipophilic contaminants in a time-integrated and site-specific fashion, where the rate of chemical uptake is related, in part, to the log K_{ow} of the contaminant and physical properties of the sampler.[4,7] It is recognized that physical processes that control chemical uptake into SPMDs, namely chemical partitioning and diffusion, are similar to those that control chemical uptake by aquatic organisms. As a result, a large effort has been made to compare SPMDs to aquatic biota in side-by-side controlled experiments using chemical uptake mechanisms,[4,10] exchange kinetics,[7,11] and accumulation signatures[8] to determine if SPMDs could be used as surrogates for biomonitoring

organisms in water quality assessments. These types of studies have been conducted for fish[10,12,13] and several species of bivalve.[4,5,8,12,14−16] However, few studies have tried to mathematically relate the chemical concentrations and signatures captured by SPMDs directly to tissue levels;[8,17] only one has tried to predict concentrations in resident, or wild, aquatic organisms,[16] and none have combined concentrations measured in the freely dissolved phase, converted from SPMDs ($C_{w,SPMD}$), with multivariate regression techniques.

In the present study, we sampled PAHs in the freely dissolved phase using SPMDs and collected colocated resident crayfish (*Pacifastacus leniusculus*) at several locations within and outside the Portland Harbor Superfund Mega-site. The sampling design was chosen to capture $C_{w,SPMD}$ of PAHs that ranged 100× or more and PAH sources that are common to Superfund sites.[18] Crayfish were targeted because they have less efficient CYP450 systems, lower rates of chemical excretion compared to aquatic vertebrates, and relatively small home ranges;[19−21] they also represent a direct route of PAH exposure to the public.[1] Differences between freely dissolved and crayfish tissue PAH profiles are described based on results from principal components analysis (PCA). The predictive performance of several calibration models that relate PAH concentrations in the freely dissolved phase to crayfish concentrations are presented, along with recommendations for handling measurements that are below method detection limits.

■ MATERIALS AND METHODS

Study Area, Sample Collection, and Preparation. The study was conducted on the lower 18.5 miles of the Willamette River, just prior to its confluence with the Columbia River. In 2000, a portion of the river (river mile 3.5 to 9.2) was designated the Portland Harbor Superfund Mega-site and was later extended to include river mile 2 through 11.8.[1] This stretch of the Willamette and its shorelines has seen nearly 100 years of heavy industrial, commercial, and urban modification. Areal and satellite images showing approximate sampling locations are provided in the Supporting Information, Figure 1.

Sampling was performed during the fall of 2003 at nine sites that were known to vary in their degree of PAH contamination. Samples within the Mega-site were collected from river miles 3.5, 7 west (7w), 7 east (7e), and 8, whereas those upstream of the Mega-site were from river miles 13, 17, and 18.5. A more focused sampling event was carried out near the former McCormick and Baxter creosoting company site at river mile 7e prior to sediment capping in 2004. SPMDs and crayfish cages were placed upriver (7e-S), in close proximity to (7e-C), and downriver (7e-N) of known sediment contamination hot spots and nonaqueous phase liquid plumes that resulted from creosoting operations. In addition, sites within the Mega-site receive inputs from coal tar contaminated industrial sites, petroleum spills, urban runoff, combined sewer overflows, and atmospheric deposition. Upriver sites receive inputs primarily associated with residential and commercial land use.

Standard triolein containing (1 mL and ≥95% pure) SPMDs were purchased from Environmental Sampling Technologies (St. Joseph, MO) and consisted of 91−106 cm strips of 2.5 cm wide low-density polyethylene lay-flat tubing with a wall thickness of 70−95 μm, surface area of 450 cm^2, and total weight of 4.5 g. SPMD samples were collected in October and early November, 2003 in 14- to 19-day deployments. Stainless steel cages containing five SPMDs were colocated with crayfish traps at each site and deployed 3 m from the river bed. Crayfish

(*P. leniusculus*) samples were collected over three sampling campaigns from September to October, 2003 using standard operating procedures adapted from EPA's *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories*.[3] Crayfish were captured using crayfish traps fashioned with chicken breast as bait; traps were retrieved within 24 h of deployment. Upon retrieval, external crayfish surfaces were cleaned of foreign material using ambient and 18 MΩ cm$^{−1}$ water and inspected for physical damage. Samples were subsequently euthanized using liquid nitrogen, individually wrapped in aluminum foil, and stored on frozen ice packs in labeled zip-lock bags for transport. Crayfish PAH concentrations were assumed to be at steady state because crayfish were likely exposed to site-specific PAH mixtures for the entirety of their lives.

In the laboratory, frozen crayfish were brought to room temperature, sexed, sized for body and carapace lengths, weighed, and dissected to isolate visceral tissue. Whole visceral samples were frozen individually with liquid nitrogen and homogenized to a fine powder using a prechilled stainless steel mortar and pestle. Crayfish extraction proceeded using a previously described method with modifications (described in the Supporting Information).[22] Samples were stored at −20 °C until analysis.

Chemicals and Instrumental Analysis. SPMD extracts and crayfish tissue samples were analyzed for the following 16 PAHs: naphthalene (NAP), acenaphthylene (ACY), acenaphthene (ACE), fluorene (FLO), phenanthrene (PHE), anthracene (ANT), fluoranthene (FLA), pyrene (PYR), benz[*a*]anthracene (BAA), chrysene (CHR), benzo[*b*]fluoranthene (BBF), benzo[*k*]fluoranthene (BKF), benzo[*a*]pyrene (BAP), indeno[1,2,3-*cd*]pyrene (IPY), dibenz[*ah*]anthracene (DBA), and benzo[*ghi*]perylene (BPY). SPMDs were extracted and chemically analyzed using high performance liquid chromatography (HPLC) with diode-array and florescence detectors as described previously,[18] whereas crayfish samples were analyzed using a previously described GC−MS method.[22] Method detection limits (MDLs) for crayfish were calculated as 10*SD of replicate low-level standard measurements or as the lowest standard that reproducibly produced signal-to-noise ratios >3 and ranged from 0.05 to 1.2 ng/g wet weight. Quality control samples met data quality objectives and are further described in the Supporting Information.

Freely Dissolved Concentration Calculation. PAH concentrations in SPMDs (C_{SPMD}, ng/g) were converted to freely dissolved water concentrations ($C_{w,SPMD}$, ng/L) using the following established equation:

$$C_{w,SPMD} = \frac{C_{SPMD} x V_{SPMD}}{R_s x t} \tag{eq 1}$$

where R_s are laboratory determined PAH sampling rates (L/d), V_{SPMD} is the sampler volume (mL), and t is the deployment time in days (d).[4,18] Laboratory-based sampling rates were temperature adjusted using methods described by Sower and Anderson.[18] All statistical analyses and modeling experiments were performed using $C_{w,SPMD}$ in order to account for deployment time, chemical specific sampling rates, and the effect of site-to-site differences in particulate and dissolved organic matter conditions on PAH bioavailability.

Statistical Methods. Crayfish PAH concentrations, $C_{w,SPMD}$, and crayfish morphological data were checked for normality using Shapiro−Wilk tests. Differences between

upriver and Superfund site PAH concentrations and crayfish morphology were evaluated using Mann–Whitney rank sum tests for data that failed Shapiro–Wilk tests or two-sample t-tests for normally distributed data. All pairwise spatial differences in PAH concentrations and crayfish morphology, and at individual superfund sites compared to upriver control sites, were evaluated using Kruskal–Wallis one way ANOVAs on ranks incorporating Dunn's procedure for multiple comparisons when there were unequal treatment group sizes, or standard one way ANOVAs when site data were normally distributed. Statistical analyses were considered significant at $p \leq 0.05$. Crayfish PAH concentrations are expressed on a wet tissue weight basis. Summary statistics were calculated for individual PAHs, the summed concentration of PAHs not including FLO (ΣPAHs), and summed levels of carcinogenic PAHs (ΣC-PAHs), which included FLA, CHR, BAA, BBF, BKF, BAP, IPY, and DBA. Descriptive statistics, graphing, and simple linear regression were performed using SigmaPlot 11.0 (Systat Software Inc.).

Principal components analysis (PCA) was used to investigate similarities and differences between freely dissolved and crayfish PAH profiles. The goal of PCA is to identify a few uncorrelated linear combinations of original data (i.e., principal components) that capture as much multivariate response variation as possible. Freely dissolved and crayfish tissue PAH concentrations were standardized to total sample PAH concentration (ΣPAH) and 4th root transformed prior to PCA because preliminary tests indicated that this transformation improved overall data visualization and reduced the impact of large scaling differences between PAH proportions. Values <MDLs were inputted as zeros, whereas FLO and BAP were excluded from PCA due to chromatographic interferences in SPMD extracts.[18] In total, PCA was performed on a mathematical matrix composed of 88 rows (60 crayfish samples, 26 $C_{w,SPMD}$ samples) and 14 columns (PAHs). Score and loading plots were used to help visually interpret patterns in the data.

Several empirical univariate and multivariate regression models were investigated for calibrating concentrations of PAHs in the dissolved phase to concentrations in crayfish tissues. Prior to modeling, predictor variables (X, $C_{w,SPMD}$) were 4th root transformed and averaged by site, while response variables (Y, crayfish tissue concentrations) were first averaged by site and then 4th root transformed. We determined empirically that this combination of transformations maximizes model performance, while providing a simple procedure for converting predicted PAH tissue concentrations from modeled units into original wet weight concentration units using a 4th power transformation. PAH concentrations that were less <MDLs were input as either "zero" or the method detection limit before calculating site averages. Thus, regression analyses were performed using $\mathbf{X}$ and $\mathbf{Y}$ matrices that were each composed of 9 rows (sampling sites) and 13 or 15 columns (average site concentration of individual PAHs), for univariate regression and PLS, respectively.

Partial least-squares calibration (PLS) is a multivariate regression method, that seeks to predict a set of response variables Y, from a set of predictor variables, X. PLS overcomes overfitting problems often encountered during multiple linear regression analyses when predictor variables are highly multicollinear. To avoid overfitting, the PLS algorithm constructs predictive models indirectly, extracting new variables called *factors* from the X data, and expresses the predictive models as linear combinations of these factors. Like principal

components extracted by PCA, PLS factors are labeled sequentially, explaining decreasing variation in the X variables. However, unlike PCA, the PLS algorithm seeks factors which also contribute to the prediction of the Y variables. Good PLS models use the fewest factors possible while retaining good predictive capabilities.

The optimal number of PLS factors to include in the models was determined using leave-one-out-at-a-time-cross-validation because our data set was relatively small. Here, each sample was withheld in turn as a "new" sample and predicted using a model parametrized with the remainder of the predictor data ($X - 1$ row). This process was repeated with models set to systematically select one additional factor up until the PLS model composed of all possible factors was evaluated. Predictive performance was assessed using the predicted residual sum of squares (PRESS) statistic and Hotelling's T^2 statistic, where the predicted residual is calculated as the difference between the predicted and measured value and the T^2 statistic is the multivariate equivalent of the Student's t statistic. Ideally, as factors are added, both the PRESS and T^2 statistics will decrease and then begin to increase indicating the onset of overfitting. The number of PLS factors associated with the smallest PRESS and T^2 statistic represents the best number of factors to include in the final model. Predictive performance was evaluated by inputting the transformed $C_{w,SPMD}$ data into the final cross-validated models and comparing residuals between PAH concentrations observed in crayfish tissues and those predicted by PLS. Multivariate analyses and graphing were performed using a combination of PRIMER (version 6, PRIMER-E, United Kingdom), SAS (version 9.2, USA), and TIBCO Spotfire S+ (version 8.1, USA) software.

■ RESULTS AND DISCUSSION

Crayfish Morphology. A total of 60 wild-caught resident crayfish, composed of 28 males and 32 females, were evaluated during the course of this study (Supporting Information, Table 1). It has been reported that organism age, body size, and physiological state can influence contaminant uptake and resulting tissue concentrations.[8] The morphological data are consistent with the hypothesis of no differences in crayfish body length, carapace length, and body weight between individual sites or between crayfish captured within and outside (controls) the designated Superfund Mega-site (Kruskal–Wallis one way ANOVA on ranks, $p > 0.1$ across all tests). Crayfish carapace lengths observed in this study ranged from 3 to 6.4 cm with a mean of 4.5 ± 0.8 cm, suggesting that organisms were generally adults greater than 2 years of age.[23,24] Studies conducted in British rivers reported that *P. leniusculus* generally migrates less than 225 m over a 2 year period and may have a more limited home range in high water flow environments.[21,25] As a result, PAH concentrations in both crayfish and SPMDs are taken to be site-specific and represent time-weighted average contaminant concentrations.[4,26] Taken together, these results suggest that crayfish evaluated in the present study were similar in age and physical attributes regardless of sampling location, and display site-specific contamination patterns. It is not readily apparent how well crayfish analyzed in the present study represent other wild crayfish populations or if chronic exposure to environmental contamination has affected their development, physiology, behavior, or ecology. However, it is likely that fishermen using this section of river capture and consume crayfish with morphological attributes similar to those reported here.

PAH Concentrations in Crayfish and the Feely Dissolved Phase. The range of PAH concentrations measured in crayfish and the freely dissolved phase across all sites is presented in Figure 1A,B, respectively, and further described in
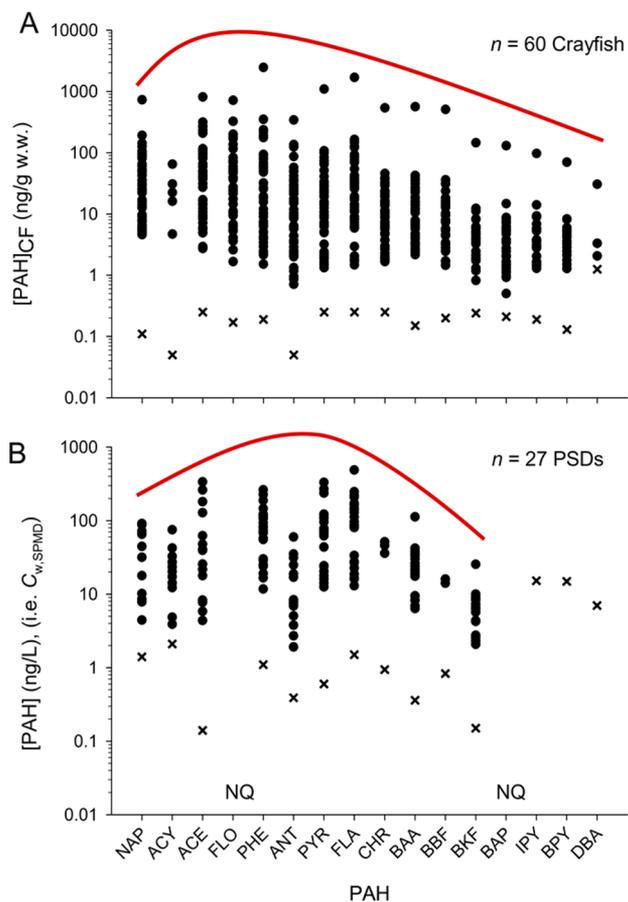

Figure 1. PAH concentrations measured in colocated (A) resident crayfish and (B) freely dissolved phase ($C_{w,SPMD}$) samples. Samples were collected along a gradient of PAH contamination within and outside of the Portland Harbor Superfund Mega-site. Measurements below detection limits are represented by "x", whereas PAHs that were not quantified due to chromatographic interferences are indicated by "NQ". $C_{w,SPMD}$ samples were collected in tandem with a larger study by Sower and Anderson.[18] Curved red lines highlight differences in the range of high molecular weight PAHs measured in crayfish tissues and the freely dissolve phase.

theSupporting Information, Tables 2 and 3. The median summed concentration of PAHs in crayfish tissue collected within the Mega-site (214 ng/g w.w.) was similar to previous reports[1] and significantly larger compared to upriver sites (12 ng/g w.w.). This was true for roughly 90% of the individual PAHs (Supporting Information, Table 4). ACY and DBA were often below detection in crayfish regardless of sampling location. Comparison of individual and ΣPAH crayfish concentrations between individual sites within the Mega-site and upriver "control" sites revealed significant differences between river miles 7eN, 7eC and in a few instances 3.5 and 7eS. Median ΣPAH concentrations were generally 13 to 30 times larger (P generally <0.001) within the Mega-site than upriver. Tissue concentrations of carcinogenic PAHs were high in crayfish collected within the Superfund, while they were frequently below detection at upriver sites (52 to73 ng/g w.w.

compared to 5 ng/g w.w. upriver). Crayfish PAH profiles at river mile 7e were comparable to source profiles for creosote contaminated sediments reported for the Wyckoff/Eagle Harbor Superfund site[27] and support earlier findings by Sower and Anderson.[18] ΣPAH concentrations measured in tissues were not strongly correlated ($R^2 < 0.25$) with crayfish body length, carapace length, or whole-body weight. However, concentrations were similar in magnitude to those measured in crayfish exposed to effluent from a U.S. Department of Energy Superfund site in Oakridge, TN.[28] The $C_{w,SPMD}$ data paired to crayfish tissue concentrations in the present study were reported in detail by Sower and Anderson[18] and are summarized in the Supporting Information, Table 3.

Chemical Profile Comparisons. PCA was used to explore similarities and differences between crayfish and $C_{w,SPMD}$ PAH profiles across the study area. Several key trends were revealed when data were standardized to total PAH content and plotted against PC1 and PC2 (Figure 2). The scores of $C_{w,SPMD}$ PAHs
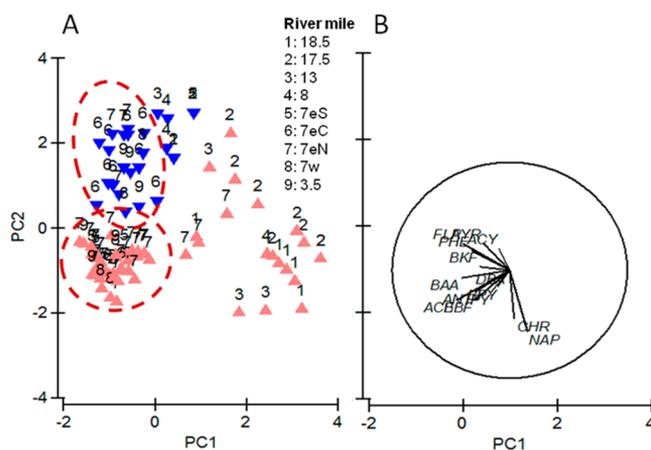

Figure 2. PCA on 4th root transformed PAH profiles using the first two principal components, PC1 and PC2, which together account for 55.4% of data set variation. The PCA scores plot for sample profiles (A) is provided with numbers indicating sampling location and colored triangles differentiating freely dissolved (▼) and crayfish (▲) samples. Red circles enclose samples collected within the Superfund Mega-site. The PCA loadings plot (B) shows directions of increasing weight for each indicated PAH variable in terms of PC1 and PC2. The length of the vectors corresponds to the percentage of that PAH's variation explained jointly by PC1 and PC2, with the perimeter of the reference circle indicating 100%. FLO and BAP were not included in the analysis.

tended to cluster together, indicating that PAH profiles in the dissolved phase were more similar across sites compared to crayfish (Figure 2A). In contrast, the large spread of crayfish PAH scores indicates that crayfish PAH profiles were more variable between sites. Crayfish collected from within the Mega-site were clearly unique compared to those collected upriver. Differences are attributed to the influence of high concentration point-source contamination within the Harbor and low, variable, nonpoint source inputs at upriver sites.[18] The PCA also revealed that PAH profiles in both crayfish and the aqueous phase varied according to collection site, where moving from upriver through the Superfund tended to be negatively correlated with PC1 and PC2.

The loadings diagram (Figure 2B) shows that PAH profiles in the freely dissolved phase (scores in Figure 2A) were enriched in ACY, PHE, PYR, and FLA (log $K_{ow}s > 4$ and < 5.2)
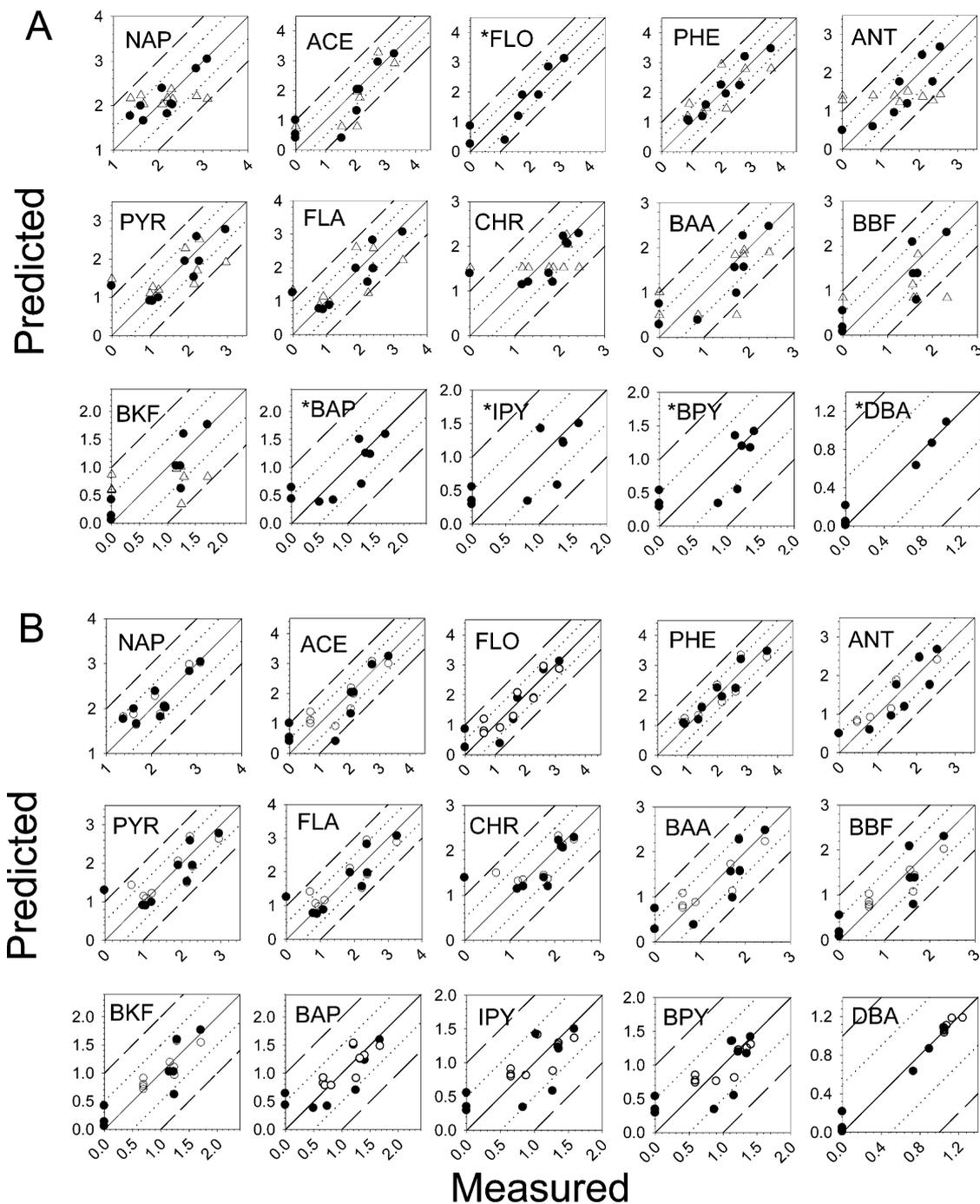
6294

dx.doi.org/10.1021/es5000534 | *Environ. Sci. Technol.* 2014, 48, 6291−6299

Figure 3. Predicted and measured PAH levels in crayfish tissues determined using PLS on averages calculated by inputting "zero" for values below MDLs (●) compared to (A) ordinary least-squares regression (Δ) or (B) PLS on averages calculated by inputting MDLs for nondetects (○). Asterisks indicate that FLO, BAP, IPY, BPY, and DBA were unable to be modeled using ordinary least-squares regression. The solid 1:1 diagonal line indicates perfect prediction, whereas the dotted (±0.5 from 1:1) and dashed (±1 from 1:1) lines are included simply to assist the reader in comparing general trends between the predictive models. Data points are in modeled units.

compared to crayfish profiles. The relatively diminished proportion of high molecular weight PAHs (log $K_{ow}s > 5.5$) in the dissolved phase is likely related to solubility limitations and competitive sorption to dissolved- and particulate-bound organic carbon. Similar findings were reported by Axelman et al. for water samples and caged blue mussels (*Mytilus edulis* L.) impacted by aluminum smelting operations.[8] But crayfish PAH profiles in the present study had larger relative proportions of CHR than mussels. This may be because crayfish were site residents with complex PAH exposure scenarios that occurred

at the sediment−water interface where high molecular weight PAHs predominate, whereas the data reported by Axelman et al. were for caged mussels that were exposed to PAHs in the water column for a relatively short period of time.[8] In addition to high relative proportions of CHR, crayfish PAH profiles were also enriched in NAP compared to the freely dissolved phase. This trend was more pronounced for samples collected within the Superfund. This could be because sediments that crayfish interacted with in the Mega-site have been impacted by point-source inputs from creosoting and manufactured gas
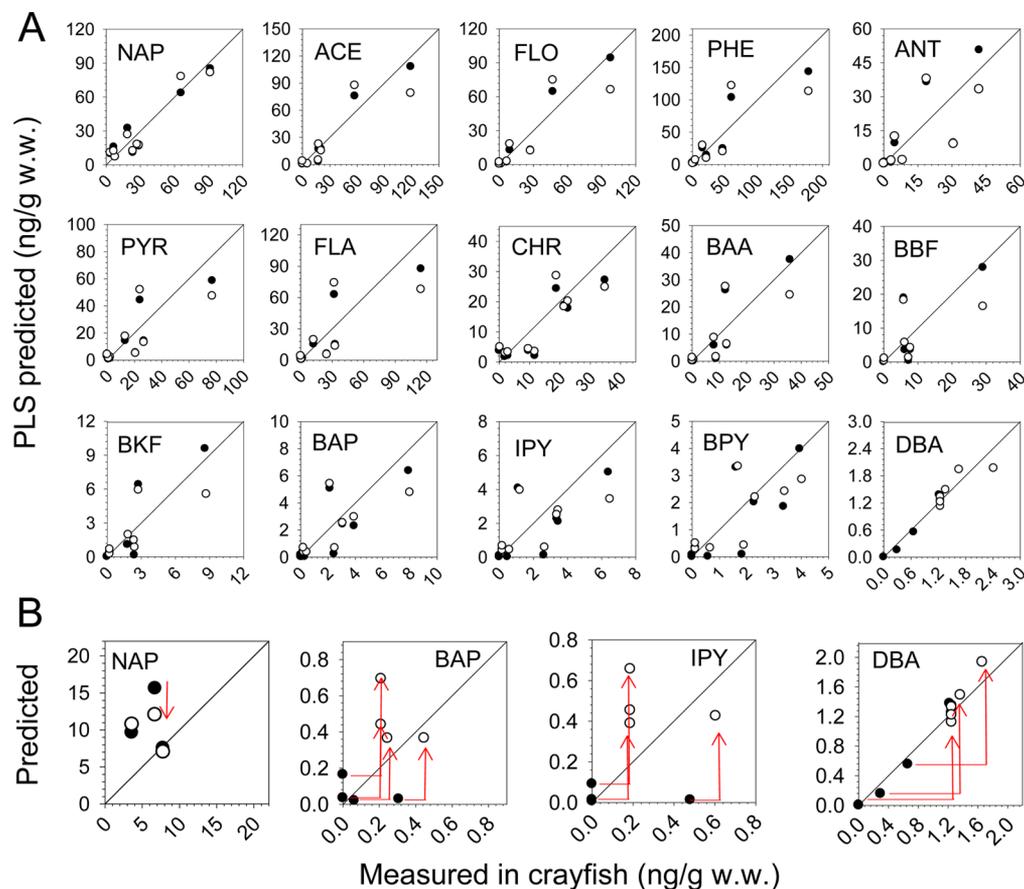
Figure 4. (A) Predicted crayfish tissue PAH concentrations compared to measured concentrations expressed in wet tissue weight and (B) close-up view of measurements near method detection limits for four representative PAHs. Data points represent (●) average site concentrations calculated by inputting "zero" for measurements ≤ MDLs and (○) average site concentrations calculated after inputting MDLs for measurements that were below detection. The diagonal 1:1 reference line indicates perfect prediction.

production, both of which can have relatively large proportions of NAP depending on the type of process residue and degree of material weathering.[27,29] Not surprisingly, the loadings diagram shows that in both crayfish and water samples traveling from upriver through the Superfund is associated with an increased proportion of PAHs, such as ACE, BAA, BBF, IPY, BPY, and DBA, which are constituents of known contamination sources to the Harbor.[18,27] Although crayfish within the Superfund were not readily distinguishable from one another by PCA, PAH profiles in resident crayfish and the dissolved phase both indicate a dominant contamination region that is unique to the Mega-site.

Predictive Modeling. Several approaches were investigated for mathematically linking $C_{w,SPMD}$ to crayfish PAH concentrations, including ordinary least-squares regression (OLS), principal components regression (PCR), and partial least-squares regression (PLS). Though the predictive capability of PCR was analyzed, the focus of further discussion will be on PLS because it consistently outperformed PCR in terms of cross-validation and R^2 values (data not shown). Results for OLS and PLS are presented in Figure 3 for average crayfish PAH concentrations by site, calculated by inputting zeros for measurements that were below detection (PLS-"zero"). Plots of predicted versus measured values were used to evaluate modeling performance graphically, where values that fell closer to the solid diagonal line indicated strong predictive capacity.

Ideally, if the model explained the data perfectly, all values would fall on the diagonal line.

As shown Figure 3A, FLO, BAP, and other high-molecular weight carcinogenic PAHs were not calibrated using OLS. This is because levels in SPMD extracts were either unable to be quantified due to chromatographic matrix interferences or because levels were <MDLs. As a result, predictions could not be made because $C_{w,SPMD}$ were unable to be calculated for these PAHs. Of the remaining PAHs, 50% appeared not to be linearly related to concentrations in the freely dissolved phase and would be better modeled as the average concentration across all sampling sites. This was the case for NAP, ANT, CHR, BBF, and BKF. Although it does appear that OLS provided some predictive capacity for ACE, PHE, PYR, FLA, and BAA, R^2 values were consistently low with values ranging from <0.001 to 0.69, where most were below 0.4 and half were below 0.2.

For PLS calibration, leave-one-out-at-a-time cross-validation was used to identify the appropriate number of PLS factors to include in the final model. Root mean PRESS and T^2 values plotted against increasing number of model factors indicated that a two-factor PLS model provided good predictions (overall R^2 value of 72.5%), while simultaneously guarding against overfitting (Supporting Information, Figure 2). Correlation-loading plots of the two-factor PLS model revealed that PAHs frequently above detection limits in both crayfish and the freely dissolved phase provided the strongest predictions, where 70 to 95% of the variation was explained for these PAHs (Supporting

Information, Figure 3). PLS grouped sites similarly to PCA, but also provided improved resolution between sites within the Superfund that vary in terms of PAH profiles, such as river mile 7 west and 7 east.[18] Compared to low molecular weight PAHs that were frequently detected, the model accounted for a smaller amount of variation for several high molecular weight PAHs that were detected infrequently and at low levels in aqueous phase samples, such as CHR, BPY, IPY, and BAP. This is likely because data for these residues were more variable as a result of their lower relative solubility and increased interaction with organic-rich solid phases.

PAH concentrations in crayfish tissues predicted using PLS-"zero" showed remarkable improvements compared to OLS regression. PLS overcame challenges associated with data that were below detection in the aqueous phase (i.e., predictor block), at low to sub part-per-trillion levels. This is because PLS generates a multidimensional calibration curve composed of linear combinations (factors) of predictor variables using all of the predictor PAHs. These so-called factors are mathematically orthogonal to one another, which allow them to simultaneously overcome challenges associated with multicollinearity. For example, though BAP was not quantified in the aqueous phase due to matrix interferences during instrumental analysis, crayfish tissue concentrations were estimated using PLS models composed of two factors, each of which was composed of a unique linear combination of predictor PAHs for which data in the aqueous phase did exist. Tissue concentrations of FLO, IPY, BPY, and DBA were estimated in the same way and agreed well with measured tissue concentrations. R^2 values for PAHs included in the present study ranged from 0.52 to 0.77 and were generally greater than 0.65 (Supporting Information, Figure 3).

The influence of measurements that were at, or below, detection on calibration performance was evaluated using a second PLS model, developed by substituting PAH method detection limits for zeros in the data set (PLS-"MDL"). Measurements were pretreated as described previously, subjected to PLS with cross-validation, and modeled using two-factors, which accounted for 71.8% of the variation in response variables. Consistency between predicted and measured values are presented in comparison to the PLS-"zero" model in Figure 3B. It is important to keep in mind that average measured concentrations were increased in the PLS-"MDL" model compared to the PLS-"zero" model. In terms of performance, both models described >70% of response variation, provided a mechanism for estimating PAHs in crayfish tissues for PAHs not measured in the freely dissolved phase, and provided balanced model fits for tissue concentrations that ranged ~200× on average across all PAHs and sampling locations (Supporting Information, Figure 4). However, substituting method detection limits for zero values led to systematically larger averages for sites that had high frequencies of PAH measurements below detection. This was the case for PAHs with molecular weights greater than 228 g/mol (i.e., > BAA) and log K_{ow}s > 5.5, where measurements in the aqueous phase and crayfish for these contaminants were more variable between replicates.

To compare and contrast the utility of the PLS-"zero" and PLS-"MDL" calibration models, PLS-predicted and measured crayfish PAH concentrations were 4th power transformed to wet weight tissue concentrations (ng/g w.w.), and plotted together (Figure 4A). Predictions from both PLS models successfully recreated the range of PAH concentrations measured in resident crayfish across the sampling area. For sites with mid-to-high levels of contamination, downriver of RM 7e-C, both models predicted tissue concentrations on average to within 2× of measured values across all 15 PAHs evaluated. Predicted crayfish PAH concentrations for samples collected at upriver sites, where PAH concentrations were lower and more variable, differed more substantially from measured values, but were generally within 19× and 3× for the PLS-"zero" and "MDL" models, respectively. Improved estimates of site-specific tissue concentrations were observed using the PLS-"MDL" model compared to a previous approach developed by Thomann and Komlos[30] that relied on equilibrium partitioning-based crayfish biota-sediment accumulation factors (BSAFs), especially for PAHs with log K_{ow}s > 6. This is significant because PAHs with log K_{ow}s in this range are typically drivers in human health risk assessments that include consumption of contaminated seafood as an exposure pathway.[1,6]

Although both PLS models provided good estimates for PAHs that were frequently above MDLs, differences were apparent for PAHs below or near MDLs. For instance, NAP was detected in every crayfish sample (>6.5 ng/g w.w. → 60× above the MDL) and a majority of SPMD extracts. As a result, differences between PLS-"zero" and -"MDL" predictions were relatively small, with the largest differences found at sites that had large frequencies of measurements below detection in the aqueous phase (Figure 4B). This was found to be the case for other low molecular weight PAHs as well (ACE, FLO, PHE, PYR). However, because these PAHs are classified as noncarcinogenic and their human health screening levels in tissues are relatively large, the differences between the two models do not have significant impacts on their interpretation in a risk assessment framework.[1] In contrast, high molecular weight PAHs, such as BAP, IPY, and DBA, were frequently below or near MDLs in crayfish, the freely dissolved phase, or both. As expected, substituting MDLs into the dataset produced average crayfish concentration estimates that were larger in magnitude than observed using zeros. From a screening level risk assessment stand point, the PLS-"MDL" modeling approach will provide tissue concentrations that are more health protective and guard against false negatives. This is especially important for high molecular weight PAHs that often exhibit carcinogenic activity because these compounds have low tissue concentration screening levels and are often drivers of risk assessments at Superfund sites.[1] However, it is important to recognize that actual PAH concentrations fall somewhere between zero and MDLs, which affords risk assessors some discretion in the treatment of measurements that fall below detection limits. In practice, measurements below detection should be handled in a way that supports study objectives.

Human health and ecological risk assessments are often challenged by a limited amount of site-specific data for contaminant concentrations in resident aquatic biota. Combining passive sampling with PLS calibration appears to provide a means to fill this knowledge gap. Additionally, using passive samplers would provide assessors dissolved phase concentration estimates, extracts suitable for detecting site-specific toxicity to aquatic systems,[31,32] and a means for identifying contamination sources based on chemical ratios and profiles,[18,33,34] thus increasing the cost effectiveness of each field sampling event. The approach may be especially useful during early project planning and screening-level assessments, when resources for performing field sampling may be in short

6297

dx.doi.org/10.1021/es5000534 | *Environ. Sci. Technol.* 2014, 48, 6291–6299

supply. Though our models were developed using PAH concentrations in visceral tissue, the approach may perform comparably for tissues more commonly consumed by the public. Future research should focus on validating the current model using a larger data set, comparing lipid filled SPMDs to nonlipid filled samplers that employ performance reference compounds, evaluating the predictive capacity of the approach in other aquatic biota and tissue types, and extending the approach to additional contaminants that drive risk assessments.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Figures and tables include additional method and result details, areal/satellite images, cross-validation diagnostic plots, PLS correlation loading plots, a summary of crayfish morphology, and detailed statistical summaries and evaluations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*K. A. Anderson. E-mail: kim.anderson@oregonstate.edu. Phone: 541-737-8501. Fax: 541-737-0497.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) ATSDR. Public Health Assessment: Portland Harbor, Multnomah County, Oregon, EPA Facility ID: OR0001297969; Oregon Department of Human Services Superfund Health Investigation and Education Program: Atlanta, GA, 2006; http://www.atsdr.cdc.gov/HAC/pha/PortlandHarbor/PortlandHarborPHA032206.pdf.

(2) FDA. Protocol for interpretation and use of sensory testing and analytical chemistry results for re-opening oil-impacted areas closed to seafood harvesting due to the Deepwater Horizon Oil Spill; U.S. Food and Drug Administration: Washington, DC, 2010; http://www.fda.gov/downloads/Food/RecallsOutbreaksEmergencies/Emergencies/UCM233818.

(3) EPA. Guidance for assessing chemical contaminant data for use in fish advisories; EPA 823-B-00-007; U.S. Environmental Protection Agency Office of Water: Washington, DC, 2000.

(4) Huckins, J. N.; Petty, J. D.; Booij, K. Monitors of organic chemicals in the environment: Semipermeable membrane devices; Springer: New York, NY, 2006; pp 139−167.

(5) Lance, E. W.; Matz, A. C.; Reeves, M. K.; Verbrugge, L. A. Petroleum hydrocarbon contamination in Nelson Lagoon, Alaska, sampling three different matrices. Mar. Pollut. Bull. 2012, 64 (10), 2129−2134.

(6) Allan, S. E.; Sower, G. J.; Anderson, K. A. Estimating risk at a Superfund site using passive sampling devices as biological surrogates in human health risk models. Chemosphere 2011, 85 (6), 920−927.

(7) Booij, K.; Smedes, F.; Van Weerlee, E. M.; Honkoop, P. J. C. Environmental monitoring of hydrophobic organic contaminants: The case of mussels versus semipermeable membrane devices. Environ. Sci. Technol. 2006, 40 (12), 3893−3900.

(8) Axelman, J.; Næs, K.; Näf, C.; Broman, D. Accumulation of polycyclic aromatic hydrocarbons in semipermeable membrane devices and caged mussels (Mytilus edulis L.) in relation to water column phase distribution. Environ. Toxicol. Chem. 1999, 18 (11), 2454−2461.

(9) Arnot, J. A.; Gobas, F. A. P. C. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environ. Rev. 2006, 14 (4), 257−297.

(10) Ke, R.; Xu, Y.; Huang, S.; Wang, Z.; Huckins, J. N. Comparison of the uptake of polycyclic aromatic hydrocarbons and organochlorine pesticides by semipermeable membrane devices and caged fish (Carassius carassius) in Taihu Lake, China. Environ. Toxicol. Chem. 2007, 26 (6), 1258−1264.

(11) Huckins, J. N.; Prest, H. F.; Petty, J. D.; Lebo, J. A.; Hodgins, M. M.; Clark, R. C.; Alvarez, D. A.; Gala, W. R.; Steen, A.; Gale, R.; Ingersoll, C. G. Overview and comparison of lipid-containing semipermeable membrane devices and oysters (Crassostrea gigas) for assessing organic chemical exposure. Environ. Toxicol. Chem. 2004, 23 (7), 1617−1628.

(12) Baussant, T.; Sanni, S.; Jonsson, G.; Skadsheim, A.; Børseth, J. F. Bioaccumulation of polycyclic aromatic compounds: 1. Bioconcentration in two marine species and in semipermeable membrane devices during chronic exposure to dispersed crude oil. Environ. Toxicol. Chem. 2001, 20 (6), 1175−1184.

(13) Verweij, F.; Booij, K.; Satumalay, K.; van der Molen, N.; van der Oost, R. Assessment of bioavailable PAH, PCB and OCP concentrations in water, using semipermeable membrane devices (SPMDs), sediments and caged carp. Chemosphere 2004, 54 (11), 1675−1689.

(14) Peven, C. S.; Uhler, A. D.; Querzoli, F. J. Caged mussels and semipermeable membrane devices as indicators of organic contaminant uptake in Dorchester and Duxbury Bays, Massachusetts. Environ. Toxicol. Chem. 1996, 15 (2), 144−149.

(15) Moring, J. B.; Rose, D. R. Occurrence and concentrations of polycyclic aromatic hydrocarbons in semipermeable membrane devices and clams in three urban streams of the Dallas-Fort Worth Metropolitan Area, Texas. Chemosphere 1997, 34 (3), 551−566.

(16) Boehm, P. D.; Page, D. S.; Brown, J. S.; Neff, J. M.; Edward Bence, A. Comparison of mussels and semi-permeable membrane devices as intertidal monitors of polycyclic aromatic hydrocarbons at oil spill sites. Mar. Pollut. Bull. 2005, 50 (7), 740−750.

(17) Hofelt, C. S.; Shea, D. Accumulation of organochlorine pesticides and PCBs by semipermeable membrane devices and Mytilus edulis in New Bedford Harbor. Environ. Sci. Technol. 1996, 31 (1), 154−159.

(18) Sower, G. J.; Anderson, K. A. Spatial and temporal variation of freely dissolved polycyclic aromatic hydrocarbons in an urban river undergoing Superfund remediation. Environ. Sci. Technol. 2008, 42 (24), 9065−9071.

(19) Jewell, C. S. E.; Mayeaux, M. H.; Winston, G. W. Benzo[a]pyrene metabolism by the hepatopancreas and green gland of the red swamp crayfish, Procambarus clarkii, in vitro. Comp. Biochem. Physiol., Part C: Pharmacol., Toxicol. Endocrinol 1997, 118 (3), 369−374.

(20) James, M. O.; Boyle, S. M. Cytochromes P450 in crustacea. Comp. Biochem. Physiol., Part C: Pharmacol., Toxicol. Endocrinol. 1998, 121 (1), 157−172.

(21) Guan, R.; Wiles, P. R. The home range of the signal crayfish in a British lowland river. Freshwater forum 1997, 8, 45−54.

(22) Forsberg, N. D.; Wilson, G. R.; Anderson, K. A. Determination of parent and substituted polycyclic aromatic hydrocarbons in high-fat

salmon using a modified QuEChERS extraction, dispersive SPE and GC−MS. *J. Agric. Food Chem.* **2011**, *59* (15), 8108−8116.

(23) Guan, R.; Wiles, P. R. Feeding ecology of the signal crayfish *Pacifastacus leniusculus* in a British lowland river. *Aquaculture* **1998**, *169* (3), 177−193.

(24) Guan, R.; Roy Wiles, P. R. Growth, density and biomass of crayfish, *Pacifastacus leniusculus*, in a British lowland river. *Aquat. Living Resour.* **1996**, *9* (3), 265−272.

(25) Bubb, D. H.; Thom, T. J.; Lucas, M. C. Movement and dispersal of the invasive signal crayfish *Pacifastacus leniusculus* in upland rivers. *Freshwater Biol.* **2004**, *49* (3), 357−368.

(26) Schilderman, P.; Moonen, E.; Maas, L.; Welle, I.; Kleinjans, J. Use of crayfish in biomonitoring studies of environmental pollution of the river Meuse. *Ecotoxicol. Environ. Saf.* **1999**, *44* (3), 241−252.

(27) Brenner, R. C.; Magar, V. S.; Ickes, J. A.; Abbott, J. E.; Stout, S. A.; Crecelius, E. A.; Bingler, L. S. Characterization and fate of PAH-contaminated sediments at the Wyckoff/Eagle Harbor Superfund site. *Environ. Sci. Technol.* **2002**, *36* (12), 2605−2613.

(28) Rao, V. R.; Mitz, S. V.; Hadden, C. T.; Cornaby, B. W. Distribution of contaminants in aquatic organisms from East Fork Poplar Creek. *Ecotoxicol. Environ. Saf.* **1996**, *33* (1), 44−54.

(29) Hawthorne, S. B.; Poppendieck, D. G.; Grabanski, C. B.; Loehr, R. C. Comparing PAH availability from manufactured gas plant soils and sediments with chemical and biological tests. 1. PAH release during water desorption and supercritical carbon dioxide extraction. *Environ. Sci. Technol.* **2002**, *36* (22), 4795−4803.

(30) Thomann, R. V.; Komlos, J. Model of biota-sediment accumulation factor for polycyclic aromatic hydrocarbons. *Environ. Toxicol. Chem.* **1999**, *18* (5), 1060−1068.

(31) Hillwalker, W. E.; Allan, S. E.; Tanguay, R. L.; Anderson, K. A. Exploiting lipid-free tubing passive samplers and embryonic zebrafish to link site specific contaminant mixtures to biological responses. *Chemosphere* **2010**, *79* (1), 1−7.

(32) Allan, S. E.; Smith, B. W.; Tanguay, R. L.; Anderson, K. A. Bridging environmental mixtures and toxic effects. *Environ. Toxicol. Chem.* **2012**, *31* (12), 2877−2887.

(33) Luellen, D. R.; Shea, D. Semipermeable membrane devices accumulate conserved ratios of sterane and hopane petroleum biomarkers. *Chemosphere* **2003**, *53* (7), 705−713.

(34) Allan, S. E.; Smith, B. W.; Anderson, K. A. Impact of the Deepwater Horizon Oil Spill on bioavailable polycyclic aromatic hydrocarbons in Gulf of Mexico coastal waters. *Environ. Sci. Technol.* **2012**, *46* (4), 2033−2039.