Environ Sci Pollut Res (2018) 25:4037–4050
DOI 10.1007/s11356-017-0094-z

In vitro tools for the toxicological evaluation of sediments and dredged materials: intra- and inter-laboratory comparisons of chemical and bioanalytical methods

Kathrin Eichbaum[1] · Markus Brinkmann[1,2] · Leonie Nuesser[1] · Carolin Gembé[1] · Marina Ohlig[3] · Sebastian Buchinger[3] · Georg Reifferscheid[3] · Markus Hecker[2] · John P. Giesy[4,5,6,7] · Henner Hollert[1,8,9,10]

Received: 20 December 2016 / Accepted: 4 September 2017 / Published online: 14 September 2017

Abstract The implementation of in vitro bioassays for the screening of dioxin-like compounds (DLCs) into management guidelines of dredged material is of increasing interest to regulators and risk assessors. This study reports on an intra- and inter-laboratory comparison study between four independent laboratories. A bioassay battery consisting of RTL-W1 (7-ethoxy-resorufin-*O*-deethylase; EROD), H4IIE (micro-EROD), and H4IIE-luc cells was used to assess aryl hydrocarbon receptor-mediated effects of sediments from two major European rivers, differently contaminated with DLCs. Each assay was validated by characterization of its limit of detection (LOD) and quantification (LOQ), *z*-factor, reproducibility, and repeatability. DLC concentrations were measured using high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS) and compared to bioassay-specific responses via toxicity equivalents (TEQs) on intra- and inter-laboratory levels. The micro-EROD assay exhibited the best overall performance among the bioassays. It was ranked excellent (*z*-factor = 0.54), reached a repeatability > 75%, was highly comparable (r^2 = 0.87) and reproducible (83%) between two laboratories, and was well correlated (r^2 = 0.803) with TEQs. Its LOD and LOQ of 0.5 and 0.7 pM 2,3,7,8-TCDD, respectively, approached LOQs of HRGC/HRMS measurements. In contrast, cell lines RTL-W1 and H4IIE-luc produced LODs > 0.7 pM 2,3,7,8-TCDD, LOQs > 1.7 pM 2,3,7,8-TCDD, and repeatability < 70%. Based on the data obtained, the micro-EROD assay is the most favorable bioanalytical tool, and via a micro-EROD-based limit value, it would allow for the assessment of sediment DLC concentrations; thus, it could be considered for the implementation into testing and management guidelines for dredged materials.

Keywords Biological effect-based assessment · BEQ · Dredged material management · Micro-EROD · TEQ

Responsible editor: Philippe Garrigues

✉ Henner Hollert
henner.hollert@bio5.rwth-aachen.de

[1] Department of Ecosystem Analysis, Institute for Environmental Research, ABBt–Aachen Biology and Biotechnology, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany

[2] School of the Environment and Sustainability and Toxicology Centre, University of Saskatchewan, 44 Campus Drive, Saskatoon, SK S7N 5B3, Canada

[3] Department G3: Biochemistry, Ecotoxicology, Federal Institute of Hydrology (BFG), Am Mainzer Tor 1, 56068 Koblenz, Germany

[4] Department of Veterinary Biomedical Sciences and Toxicology Centre, University of Saskatchewan, 44 Campus Drive, Saskatoon, SK S7N 5B3, Canada

[5] Department of Zoology and Center for Integrative Toxicology, Michigan State University, East Lansing, MI, USA

[6] Department of Biology, Hong Kong Baptist University, Kowloon, Hong Kong, SAR, China

[7] School of Biological Sciences, University of Hong Kong, Hong Kong, SAR, China

[8] Key Laboratory of Yangtze River Environment of Education Ministry of China, College of Environmental Science and Engineering, Tongji University, Shanghai, China

[9] College of Resources and Environmental Science, Chongqing University, Chongqing, China

[10] School of Environment, Nanjing University, Nanjing, China

Introduction

Industrial and municipal emissions to rivers have been reduced considerably during the last decades due to regulations such as the European Water Framework Directive (WFD) (Besselink et al. 2004). However, as sediments serve as sinks for persistent and bioaccumulative contaminants, they remain important sources by reintroducing particulate-bound organic pollutants back into the water phase through events such as dredging or flood events (Burton 1992, Schüttrumpf et al. 2011). Among sediment-bound pollutants, the group of dioxin-like compounds (DLCs) represents one of the most relevant groups of legacy contaminants. This group includes polychlorinated dibenzo-*p*-dioxins and dibenzofurans (PCDD/Fs) as well as dioxin-like polychlorinated biphenyls (DL-PCBs), which have the potential to cause adverse effects to wildlife and humans (as reviewed by White and Birnbaum 2009). Through frequently conducted sediment dredging activities, which are required and unavoidable for the maintenance of navigable waterways, fisheries, hot spot controlling, and flood defense (Breitung and Keller 2010), sediment-bound pollutants can be re-mobilized and transferred into surface waters where they become bioavailable again (Burton 1992, Otte et al. 2013). Thus, dredging may be contradictive to the aim of the WFD (management plan, 2000/60/EG 2009) and most likely prevent achieving the framework's goal of a "good ecological status" (Barceló and Petrovic 2007; Förstner et al. 2008; Hallare et al. 2011).

Recently, several environmental risk assessment (ERA) approaches have been developed to progressively enhance the water quality of rivers. For instance, the WFD daughter directive requested the concentrations of 33 priority pollutants (annex, 2000/60/EC 2006) not to increase in water (Hollert et al. 2009). For these substances, environmental quality standards (EQSs) were set (2008/105/EC 2008), which are used to identify effects or no effects of sediment-borne contaminants and to define measures (e.g., disposal or habitat construction) to be undertaken with dredged materials (Apitz and Power 2002, Manz et al. 2007, Wenning and Ingersoll 2002). A proposal for a directive amending 2000/60/EC and 2008/105/EC (COM(2011)876 2011) states 15 additional priority substances and introduces EQS for biota (additional to those stated in 2008/105/EC).

The management and handling of dredged materials in Germany follows guidelines that have been compiled by the Ministry of Transport and Digital Infrastructure (BMVI) under the coordination of the German Federal Institute of Hydrology (BfG) (Breitung and Keller 2010). Waterways located outside the jurisdiction of the Water and Shipping Administration (WSV) are subjected to the regulations of the respective German federal states (den Besten et al. 2003). Two directives of the management of dredged material on both federal inland ("Handlungsanweisung für den Umgang mit Baggergut im Binnenland," HABAB; HABAB 2000) and coastal waterways ("Handlungsanweisung für den Umgang mit Baggergut Küste," HABAK; HABAK 1999) have been established and reestablished within joint transitional arrangements ("Gemeinsame Übergangsbestimmungen zum Umgang mit Baggergut in den Küstengewässern," GÜBAK; GÜBAK 2009). For disposal of dredged materials, characteristics of dredging and relocation sites have to be comparable and evaluated according to economic and ecological aspects (Breitung and Keller 2010; HABAK 1999).

According to the German GÜBAK (GÜBAK 2009), the assessment of sediment quality is based on a stepwise chemical analysis and ecotoxicological test methods. Chemical analysis in this context focuses on pollutants such as heavy metals and hydrophobic organic pollutants such as polycyclic aromatic hydrocarbons (PAHs) and PCBs, which all are known to be of high relevance for sediments and suspended particulate matter. In case of hydrophobic organic pollutants, seven so-called indicator PCBs (i-PCBs; IUPAC No. 28, 52, 101, 118, 138, 153, und 180) are required to be analyzed, and of which mono-*ortho* PCB 118 belongs to the 12 DL-PCBs.

Ecotoxicological tools such as tests with algae, bacteria, and microcrustaceans do not provide information about causative compounds responsible for the observed biological effects and are not comparable to chemical analytical results (den Besten et al. 2003, Manz et al. 2007). Bioassays, which belong to the most important lines of evidence in support of integrated sediment assessment schemes (sediment quality triad (SQT); Chapman and Hollert 2006; Hollert et al. 2002), overcome such issues by being complementary to chemical results and by taking synergetic and antagonistic effects into account (Ahlf et al. 2002). This is why their role in a biological effect-based assessment in decision making frameworks is increasingly discussed (Ahlf et al. 2002; Besselink et al. 2004; den Besten et al. 2003; Förstner et al. 2008).

Since the year 2004, successful implementations of in vitro assays for the screening of DLCs in form of the DR-CALUX assay can be found in the Dutch dredging guideline for coastal sediments, which formerly only included chemical analysis. Here, a biological equivalent quotient (BEQ) signal value of 50 ng BEQ g^{-1} dry weight (dw) sediment has been set, which—if exceeded—involves further, detailed investigations (Manz et al. 2007). In German legislation, in vitro assays as semiquantitative methods prior to quantitative instrumental analysis have only been established in the field of food analysis, where BEQs allow for a simple yes/no decision (2012/252/EU 2012).

The present study addresses the question if specific in vitro assays for the detection of dioxin-like effects, especially those based on freely available non-proprietary cell lines, could be of added value for the assessment of sediment quality in the context of dredging activities in German waterways. Based on the results of the most reliable bioassay, possibilities of a threshold definition are presented and discussed.

Materials and methods

Study design

Participating laboratories included (1) the Institute for Environmental Research, RWTH University, Aachen, Germany, (2) the Federal Institute for Hydrology (BfG), Koblenz, Germany and (3) the BfG contract laboratory GBA (Gesellschaft für Bioanalytik mbH), and (4) Münster Analytical Solutions (mas), Münster, Germany, which in the following are abbreviated as (1) lab 1, (2) lab 2, (3) lab 2*, and (4) lab 3. The Institute for Environmental Research, RWTH University, Aachen, Germany, was the main laboratory, where most of the present study's work has been conducted.

Firstly, a comparison of three in vitro cell bioassays, including the rainbow trout liver RTL-W1 EROD assay, the rat hepatoma H4IIE micro-EROD assay, and the H4IIE-luc assays, was conducted by lab 1 (main laboratory). Bioassays were validated by means of a set of criteria (listed in Table 1) such as repeatability, limits of detection (LOD) and quantification (LOQ), different levels of effect concentrations (EC) of the 2,3,7,8-TCDD standard, sample induction strengths relative to positive control as well as *z*-factors. Raw extracts, multilayer fractions (containing the sum of dioxin-like PCBs and PCDD/Fs), DL-PCB and PCDD/F fractions of differently contaminated sediments as well as positive control 2,3,7,8-TCDD served as bases for bioassay validation.

It turned out that those criteria disqualified the RTL-W1 EROD assay as a reliable screening tool, while indicating the need for further investigations of the H4IIE-luc assay. To uncover possible operator-related variations in the H4IIE-luc assay, additional intra-laboratory comparisons were performed in lab 1 using this assay, which in turn disqualified (repeatability < 75%; 2012/278/EU 2012) this assay, too.

For this and due to its overall best performance (Table 1), only the H4IIE micro-EROD assay was chosen for an inter-laboratory comparison between two operators of labs 1 and 2.

Finally, high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS) results for selected DLCs were compared between labs 2* and 3, and results of bioassay and HRGC/HRMS measurements were compared via toxicity equivalents (TEQs) and biological equivalents (BEQs), respectively.

Sediments

Sediment sampling was conducted by lab 2 in April 2012. The three sampling locations comprised Ehrenbreitstein (EBR) at the Rhine river as well as Prossen (PR) and Zollelbe (ZE) located along the Elbe river. EBR has been used in former studies, as it represents a slightly contaminated sediment (Feiler et al. 2013; Höss et al. 2010). Sediment PR possibly reflects the toxicological burden coming from the Czech part of the Elbe river (Stachel et al. 2011).

All sediments were sampled to a depth of 15 cm using Van-Veen grabs, filled in polyethylene buckets and immediately transferred to lab 1, where they were thoroughly homogenized, and an additional 1:10 mixture (EBR/ZE), consisting of nine parts dry weight (dw) EBR and one part dw ZE, was prepared in the laboratory to present a fine particulate sediment contaminated with highly persistent DLCs. For simplicity, the mixture of EBR/ZE will be subsequently named *sediment* as well. All sediments were stored at 4 °C until further use.

Sample preparation, extraction, and clean-up

The sample treatment procedures described in this section were all conducted in lab 1. Sediments were freeze-dried for 72 h (Alpha 1–4 LD plus, Martin Christ GmbH, Osterode, Germany), sieved to < 2 mm, and homogenized by using a mortar and pestle. Sediments were extracted for 48 h according to the methodology described by Umlauf et al. (2004), using Soxhlet extraction (behr Labor Technik, Düsseldorf, Germany) and a solvent mixture of n-hexane/acetone (352/48; *v*/*v*). An amount of 20 g sediment (dw) was mixed with an amount of 5 g muffled sodium sulfate (99% anhydrous powder, Sigma Aldrich, Germany). A process control sample, only containing 5 g sodium sulfate, was extracted in parallel. Samples for chemical analysis were spiked with $^{13}C_{12}$-labeled PCDD/F standards (EPA 1613 LCS, Wellington Laboratories, Campro Scientific GmbH, Germany) and a $^{13}C_{12}$-labeled PCB standards (EPA 68C LCS, Wellington Laboratories, Germany).

Clean-up of extracts included the following steps in chronological order: desulfurization with activated copper (24 h), sulfuric acid treatment (24 h), multilayer silica column clean-up, and activated carbon column clean-up. Each step was performed in accordance with US EPA method 8290 (US-EPA 1994) with the following modifications: Multilayer silica columns were equipped in the bottom-to-top order: glass wool, 1 g of activated silica gel, 2 g of basic silica gel (30 g of sodium hydroxide dissolved in 750 mL methanol, combined together with 100 g of silica gel that was then rotary evaporated until dryness for approximately 90 min in a 55 °C water bath), 1 g of activated silica gel, 4 g of acidic silica gel, 1 g of activated silica gel, and 1 g of sodium sulfate. Supelclean™ ENVI-Carb™ (Sigma Aldrich) was chosen as carbon adsorbent in the activated carbon columns. Extracts of fractions containing DL-PCBs and PCDD/Fs were aliquoted volumetrically and stored in 4-mL vials (amber glass, 45 × 14.7 mm with Butyl/PTFE septum and screw cap, VWR International) until further use in bioassays or chemical analysis (HRGC/HRMS analysis). For

Table 1 Bioanalytical quality criteria achieved for three different in vitro bioassays (RTL-W1 EROD, H4IIE micro-EROD, and H4IIE-luc assays) including effect concentrations (EC) levels of positive control 2,3,7,8-TCDD, limits of detection (LOD), and quantification (LOQ), *z*-factors, repeatability, and reproducibility

	In vitro bioassay				
	EROD[a]	Micro-EROD[a]	Micro-EROD[b]	H4IIE-luc[a]	H4IIE-luc[c]
Cell line	RTL-W1	H4IIE	H4IIE	H4IIE-luc	H4IIE-luc
Passages used	60–77	26–50	26–30	10–26	19–27
Number of tests	273	75	16	71	48
EC_{10}TCDD [pM]					
Mean ± SD	2.53 ± 0.82	2.39 ± 0.76	1.80 ± 0.20	1.21 ± 0.87	1.03 ± 0.60
Min/max	0.72/6.25	1.09/5.99	1.49/2.20	0.26/6.00	0.11/2.73
EC_{25}TCDD [pM]					
Mean ± SD	4.72 ± 1.32	3.59 ± 1.28	3.08 ± 0.35	2.52 ± 0.85	2.71 ± 1.40
Min/max	1.81/12.08	1.61/10.35	2.58/3.69	0.72/4.22	0.77/7.67
LOD [pM]					
Mean ± SD	0.94 ± 0.61	0.45 ± 0.32	0.43 ± 0.16	0.78 ± 0.63	0.73 ± 0.91
Min/max	0.01/5.41	0.07/2.05	0.06/0.61	0.06/2.96	0.01/4.72
LOQ [pM]					
Mean ± SD	1.72 ± 1.28	0.69 ± 0.36	0.72 ± 0.18	2.32 ± 2.05	2.12 ± 2.27
Min/max	0.02/13.05	0.17/2.47	0.20/1.01	0.16/9.93	0.06/10.86
z-factor					
Mean ± SD	0.36 ± 0.46	0.54 ± 0.22	0.72 ± 0.10	0.67 ± 0.14	0.64 ± 0.25
Min/max	− 5.13/0.99	− 0.06/1.00	0.58/0.90	0.27/0.90	− 0.88/0.91
Repeatability [%]					
Extract/TCDD	64/70	76/77	88/90	69/61	43/62
Reproducibility [%]					
Extract/TCDD	n.a.	$83_{(bl)}/98_{(bl)}$	$83_{(bl)}/98_{(bl)}$	$83_{(wl)}/78_{(wl)}$	$83_{(wl)}/78_{(wl)}$

bl between-laboratory reproducibility, *wl* within-laboratory reproducibility, *n.a.* not analyzed

[a] Operator 1, lab 1, method comparison

[b] Operator 2, lab 2, inter-laboratory comparison

[c] Operator 2, lab 1, intra-laboratory comparison

bioassay purposes, extracts were reduced close to dryness under a gentle stream of nitrogen and redissolved in dimethyl sulfoxide (DMSO ≥ 99.5% p.a., Carl ROTH).

For intra-laboratory comparison purposes, 20 g of each freeze-dried sediment was sent to lab 2* where an appropriate extraction and clean-up were performed.

Bioanalytical analysis

RTL-W1 EROD (7-Ethoxyresorufin-O-deethylase) assay

EROD activity was measured using the permanent fish cell line RTL-W1 (*Oncorhynchus mykiss*, rainbow trout liver-Waterloo 1) (Lee et al. 1993), donated by Dr. Niels C. Bols, University of Waterloo, Canada. Passage numbers 73 to 76 were used to obtain the here presented results. Cell culture and assay were performed according to the methods described by Wölz et al. (2009) with the exception that two samples were tested in triplicate per plate and each well of a plate was adapted to a final concentration of 0.5% DMSO. Briefly, cells were seeded in 96-well plates (TPP, Trasadingen, Switzerland) and incubated 72 h until confluence. Thereafter, medium was removed and cells were exposed to serial diluted concentrations of extracts and positive control 2,3,7,8-TCDD (3.1 to 100.0 pM; Promochem, Wesel, Germany).

Following a 72-h incubation time, exposure medium was removed and cells were lysed by freezing them at − 80 °C for at least 1 h. First, an enzyme-substrate complex consisting of EROD present in the cells and added substrate 7-ethoxyresorufin was allowed to develop within a reaction time of 10 min. Thereafter, addition of reduction equivalent NADPH caused the deethylation of the substrate, which was stopped after another 10 min through the addition of fluorescamine dissolved in acetonitrile. After 15 min, specific EROD activity was determined by measuring the fluorescence of reaction product resorufin (extinction 544 nm, emission 590 nm) and the absorbance of fluorescamine-amine-complexes (extinction of 360 nm, emission of 460 nm), according to a method of Lorenzen et al. (1997) with a multiwell plate reader (Tecan infinite M200).

H4IIE micro-EROD assay

H4IIE cells were provided by the Lower Saxony State Office for Consumer Protection and Food Safety (LaVes) and were cultivated using Dulbecco's modified eagle medium (DMEM; low glucose, Life Technologies GmbH) supplemented with 10% FBS (Biochrom AG) and 2% L-glutamine (2 mM, GIBCO® GlutaMAX™, Life Technologies GmbH, Darmstadt, Germany). Passage numbers 26 to 50 were used for the assay, which was performed according to a protocol provided by LaVes (LAVES 2013). For the detailed protocol, please refer to Schiwy et al. (2015). Briefly, confluent cells were trypsinized and 50 μL cell suspension (200,000 cells mL^{-1} DMEM without phenol red) was seeded in a 96-well plate (96-Well, Growing surface, Sarststaedt) and incubated for 2 h in a humidified 95:5air/CO_2 atmosphere at 37 °C in darkness.

Thereafter, cells were exposed to triplicates of serially (1:2) pre-diluted concentrations of extracts and positive control 2,3,7,8-TCDD using DMSO as a carrier solvent (0.58 to 18.64 pM; Promochem, Wesel, Germany). DMSO concentration was 0.5% in all wells. Following a 72-h incubation, medium was removed and 100 μL of 8 μM ethoxyresorufin (ETX) solution containing 10 μM dicumarol was added to all cell-containing wells. After 30 min, the reaction was stopped by adding 75 μL methanol (p.a.; ROTH). Plates were shaken horizontally (300 rpm) for 10 min, and resorufin production was determined fluorometrically (excitation 530 nm, emission 590 nm) by using a multiwell plate reader (Tecan infinite M200; Tecan Germany GmbH, Crailsheim, Germany). For the calculation of the specific EROD activity, protein was determined by using a bicinchonic acid (BCA) protein assay kit (Sigma Aldrich). A protein standard curve was prepared in the remaining ETX solution and added to the plate in a 1:2 serial dilution (3.9–500.0 μg mL^{-1}). Absorption was measured at 550 nm following the addition of 100 μL/well BCA solution (Tecan infinite M200).

For intra-laboratory comparison purposes, personnel of lab 2 were trained in lab 1. Frozen cells (passage number 22), aliquots of extract fractions, 96-well plates, and stock solutions required for test performance were sent to lab 2. Plates were measured using the same protocol (Tecan Infinite M200 Pro) and data was then sent to and evaluated by lab 1 personnel.

The H4IIE-luc assay

The H4IIE-luc cell line was donated by Prof. Dr. John P. Giesy (University of Saskatchewan, Saskatoon, Canada) and cultivated with Dulbecco's modified eagle medium (DMEM; low glucose, Life Technologies GmbH) supplemented with 10% FBS (Biochrom AG) and 2% L-glutamine (2 mM, GIBCO® GlutaMAX™, Life Technologies GmbH, Darmstadt, Germany). The assay was performed according to a method developed by Sanderson and co-workers (1996), with modifications (Khim et al. 1999). Confluent cells of passage numbers 10 to 27 were trypsinized, seeded in 96-well plates (ViewPlate™-96, Perkin Elmer, Rodgau-Jügesheim, Germany) at a density of 80,000 cells mL^{-1} in DMEM, and incubated for 24 h in a humidified 95:5air/CO_2 atmosphere at 37 °C in darkness.

Thereafter, medium was removed (liquid handling device, IBS INTEGRA Bioscience, Landquart-Davos, Switzerland) and cells were exposed to triplicates of 250 μL serially (1:3) pre-diluted concentrations of extracts and positive control (0.5% DMSO per well) for 72 h. It should be noted that according to Lee et al. (2013), the prolonged 72-h incubation time compared to 24 h of the original protocol by Sanderson and co-workers (1996) may result in slightly poorer analytical quality control criteria, without affecting relative potencies in general. Thereafter, plates were washed twice with phosphate buffered saline (PBS; 10 ×; with 1.33 g calcium L^{-1} and 1.0 g magnesium chloride L^{-1}, Sigma). The bottoms of the plates were closed with backing tape (white; for ViewPlate™-96, Unifilter™-96, PerkinElmer) and each well was equipped with 100 μL PBS and 50 μL LucLite (LucLite®, Constant Quanta™, Perkin Elmer). After 10 min, luminescence was determined using a multiwell plate reader (TECAN infinite M200).

Calculation of biological equivalent quotients (BEQs)

Bioassay concentration-response curves were plotted via GraphPad Prism 5 (GraphPad Prism 5 Software Inc., La Jolla, CA, USA) using a four-parameter logistic regression model (dose response stimulation; log agonist vs. response). BEQs were calculated based on a 25% effect concentration (EC) level (EC_{25}), with BEQ [pg/g dw] representing the quotient of the EC_{25}TCDD [pg mL^{-1}] and the extract EC_{25} [g sediment mL^{-1} medium] (Eq. 1). Each BEQ represents the mean value of three independent replicates.

$$\mathrm{BEQ}\left[\mathrm{pg\ g^{-1}}\right] = \frac{TCDD\ \mathrm{EC_{25}}\left[\mathrm{pg\ mL^{-1}}\right]}{extract\ \mathrm{EC_{25}}\left[\mathrm{g\ mL^{-1}}\right]} \quad (1)$$

All bioanalytical data from intra- and inter-laboratory comparisons were evaluated by the same operator.

Calculation of bioanalytical quality criteria

High-throughput screening assays require adequate sensitivity, reproducibility, and accuracy in order to be used as high throughput assays for the identification of samples of highest concern (2012/278/EU 2012).

Repeatability is defined as the precision under conditions, where independent test results are obtained with the same

method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time (ISO/5752 2002). It is calculated as the coefficient of variation of n measurements ($n = 3$ in the present study).

Reproducibility can be divided into (1) *within-laboratory reproducibility* and (2) *between-laboratory reproducibility*. *Within-laboratory reproducibility* is defined as precision under conditions, where test results are obtained with the same method on identical test items in the same laboratory with different operators over a long period of time. *Between-laboratory reproducibility* is defined as precision under conditions, where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment (ISO/5752 2002). Both (1) and (2) are calculated as the coefficient of variation of two mean values, each consisting of three independent replicates.

The z-factor was calculated according to Zhang et al. (1999) (Eq. 2):

$$z\text{–}factor = 1 - \frac{3(\sigma_s + \sigma_c)}{|\mu_s - \mu_c|} \quad (2)$$

with standard deviation σ and arithmetic mean μ of the sample s (here TCDD maximum induction) and the solvent control c. The factor represents the assays' dynamic range and data variation of both sample and reference compound measurement. A z-factor of 1 indicates an *ideal assay*, whereas z-factors in the ranges $1 > z \geq 0.5$ and $0.5 > z > 0$ indicate *excellent* and *marginal assay*, respectively. *Marginal assay* signifies that the separation of positive and negative control is small. The classification is based on the general assumption that the better an assay, the higher its dynamic range and the smaller its variability (Zhang et al. 1999). *Limit of detection (LOD)* (Eq. 3) and *quantification (LOQ)* (Eq. 4) were determined according to MacDougall and Crummett (1980),

$$\mathrm{LOD} = \mu_c + 3\sigma_c \quad (3)$$

$$\mathrm{LOQ} = \mu_c + 10\sigma_c \quad (4)$$

with μ and σ being the arithmetic mean and standard deviation, respectively, of a negative control c, which in the case of this study is represented by solvent control DMSO.

Chemical analysis

HRGC/HRMS analyses

HRGC/HRMS analyses of extracts prepared by lab 1 were performed by lab 3. For instrumental analysis, a capillary gas chromatograph coupled to a high-resolution mass spectrometry was used (Thermo Scientific Trace Ultra GC with Thermos scientific DFS HRMS, Thermo Fisher Scientific, Bremen, Germany). The GC was equipped with a 60-m DB-5MS capillary column of 0.25 mm inner diameter and 0.25 μm film thickness (Agilent J&W, Santa Clara, CA, USA). The capillary column was used for both PCDD/F and PCB analyses. Separate HRGC/HRMS runs at different instrumental conditions were applied for the analysis of the two compound classes.

Since the pre-cleaned sediment extracts provided by lab 1 partly showed insufficient extract purification, the PCDD/F and PCB fractions of the initial clean-up were recombined and reprocessed for chemical analysis by lab 3. The HRGC/HRMS analyses also revealed hepta-, octa-, and partly HexaCDD/Fs to be retained within the initial clean-up. Hence, quantified congeners comprised 2,3,7,8-tetraCDD/F, 1,2,3,7,8-pentaCDD/F, 2,3,4,7,8-pentaCDF, and most 2,3,7,8-HexaCDD/Fs, as well as the 12 WHO-DL-PCBs comprising the non-*ortho* PCBs 77, 81, 126, and 169 and the mono-*ortho* PCBs 105, 114, 118, 123, 156, 157, 167, and 189. However, the comparably low TEF values of these retained congeners make them negligible for TEQ calculation.

Quantification of PCDD/Fs and PCBs was performed via isotope dilution and based on the labeled internal PCDD/F and PCB standards added by lab 1 prior to the initial clean-up. Overall recoveries of the internal standards through both clean-up procedures were determined by means of labeled recovery standards added prior to the instrumental analysis. Based on blank analyses and the dry weight of the sediment samples, analytical limits of quantification (LOQs) were determined to be below 1 pg WHO2005TEQ/g for both PCDD/Fs and PCBs, respectively. Recoveries of the $^{13}C_{12}$-labeled Tetra- through HexaCDD/F quantification standards were in the range of 15–106%, whereas recoveries of the DL-PCBs ranged from 43 to 120%. Recoveries of PCDD/Fs and DL-PCBs in the three process control ranged from 11 to 102% and from 52 to 108%, respectively. Inter-laboratorial conducted HRGC/HRMS results were evaluated by different operators of laboratories 2* and 3.

Calculation of toxicity equivalent quotients

TEQs for DL-PCB and PCDD/F fractions were calculated based on the relative potency (REP) of individual congeners described previously for the different assays used in this study (Behnisch et al. 2002, Clemons et al. 1997, Lee et al. 2013, Van den Berg et al. 2006) and which allowed for assay and cell line specific conversion of instrumentally derived results.

In the case of RTL-W1 cells, REPs were obtained from 72 h-EC_{50} values previously described by Clemons et al. (1997), whereas 72 h-EC_{20} values for H4IIE and H4IIE-luc cells were obtained from Behnisch et al. (2002) and Lee et al. (2013), respectively. Although BEQs were calculated for 72 h of exposure, the chosen effect levels for BEQ calculation slightly deviated from those available for REP-based TEQ calculation. For bioassay comparisons, TEQs were calculated

based on WHO_{2005} mammalian toxicity equivalent factors (Van den Berg et al. 2006) (Eq. 5):

$$TEQ_x\ [\mathrm{pg\ g^{-1}}] = \sum\left(conc_i * REP_x\right) \quad (5)$$

Connecting the present study with German dredged material directives

To connect the present study's results with German dredged material directives, we conducted an attempt to derive a bioanalytical threshold value (H4IIE micro-EROD assay) from chemical analytical data. TEQ values (sum of WHO_{2005} DL-PCB and PCDD/F TEQs) from an Elbe length profile study, conducted by Stachel and co-workers (2011), were collected and a 75% percentile of all data points, including TEQs of the present studies' Elbe sampling locations Prossen and Magdeburg (ZE), was determined and served as a limit value for stronger contaminated locations.

Data analysis and presentation

All graphs and correlation analyses (Pearson correlation; $p < 0.05$) were calculated using GraphPad Prism 5. Illustrations were created using the vector graphic program Inkscape 0.91. A one-tailed Student's t test ($p < 0.05$) was performed using Sigma Stat 12.0 to statistically analyze differences between intra- and inter-laboratorial derived bioassay results. In case of uneven variances, Welch's correction ($p < 0.005$) was applied for Student's t test adjustment.

Results

Physical chemical characterization of sediments

Generally, Elbe sediments PR and ZE, which on average exhibited higher percentages of sand (21%) and smaller percentages of silt (69%) and clay (10%) clearly differed from Rhine sediment EBR (4, 79, and 17% of sand, silt, and clay, respectively). In contrast to EBR (49.6 g kg^{-1} TOC), they had higher amounts of TOC (63.7 g kg^{-1}), confirmed by their slightly higher losses on ignition (13.6 and 10.6%, respectively).

WHO-PCDD/F concentrations (data not shown) increased in the order EBR < PR < EBR/ZE < ZE (0.03, 0.24, 1.22 and 3.70 ng g^{-1} dw, respectively), whereas WHO-DL-PCB concentrations increased in the order EBR < EBR/ZE < PR < ZE (3.50, 4.18, 5.22 and 9.72 ng g^{-1} dw, respectively). When expressed as toxicity equivalents (TEQs, Van den Berg et al. 2006) both PCDD/F and DL-PCB TEQs increased in the order EBR < PR < EBR/ZE < ZE. PCDD/F TEQs were 5.13, 5.49, 17.53, and 84.22 ng g^{-1} dw and DL-PCB TEQs were 2.66, 2.95, 4.04, and 5.85 ng g^{-1} dw for EBR, PR, EBR/ZE, and ZE, respectively.

Inter- and intra-laboratory comparison study

Comparison of three in vitro bioassays

This section focuses on the general handling and performance of the EROD, micro-EROD, and H4IIE-luc assays (refer to assays marked with "a", Table 1) as well as a comparison between bioassay and HRGC/HRMS-derived BEQs and TEQs, respectively.

The H4IIE micro-EROD assay showed an average z-factor of 0.54, an average repeatability (coefficient of variation (CV) of three independent measurements) of > 75% for both sediment extracts and TCDD (Table 1) with an LOD and LOQ of 0.5 and 0.7 pM 2,3,7,8-TCDD, respectively.

The average LOD and LOQ of the H4IIE-luc assay was 0.7 and 2.1 pM 2,3,7,8-TCDD, respectively. Repeatability for sediment extracts, which was independent of the different fractions (raw, multilayer, DL-PCB, and PCDD/F) and 2,3,7,8-TCDD, was high and averaged at 31 and 39%, respectively. The quotient of the average EC_{25}TCDD and the average EC_{10}TCDD levels was the highest (2.6) for the H4IIE-luc assay (Table 1).

RTL-W1 EROD assay average LODs and LOQs of 0.94 and 1.72 pM 2,3,7,8-TCDD (Table 1), respectively, were high compared to the remaining assays. The assay's repeatability was the lowest (64%), but independent of the different fractions tested (raw, multilayer, DL-PCB, and PCDD/F). The CV among replicate experiments with the single substance 2,3,7,8-TCDD was lesser with a value of 30% (Table 1). The assays' overall z-factor was 0.36 and was accompanied by the overall highest standard deviation among the three assays.

The comparability between TEQs and BEQs increased from the H4IIE-luc ($r^2 = 0.642$) to the EROD ($r^2 = 0.779$) to the micro-EROD assay ($r^2 = 0.803$). The percentage of relative potency (REP)-based TEQs in micro-EROD BEQs amounted up to 49% (Fig. 1). Thereby, it was proven that REP-based TEQs explained a greater part (26.5%) of BEQs than it was the case for routinely used WHO TEQs of year 2005 (16.0%, data not shown).

Intra-laboratory comparison

Results obtained by the different operators with the H4IIE-luc assay were highly similar. The intra-laboratory validation study achieved LODs that did not significantly differ ($p = 0.339$) from each other following Welch's correction ($p < 0.0001$). However, repeatability differed considerably between operator 1 (69%) and 2 (43%), respectively. The within-laboratory reproducibility was independent of the

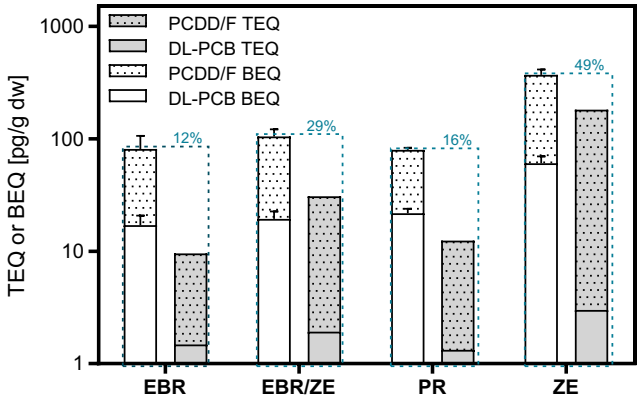


Fig. 1 Comparison of bioanalytical (BEQs) and instrumentally derived toxicity equivalents (TEQs) of DL-PCB and PCDD/F extract fractions of four sediments. EBR = Ehrenbreitstein (Rhine), PR = Prossen (Elbe), ZE = Zollelbe (Elbe) and a 1:10 mixture (EBR/ZE) consisting of one dry weight part EBR and nine dry weight parts ZE. BEQs were determined on an EC_{25} level via the H4IIE micro-EROD assay, while TEQs were calculated using H4IIE micro-EROD assay-specific relative potencies (REP). The overall share of TEQs in respective BEQs is given in percentages

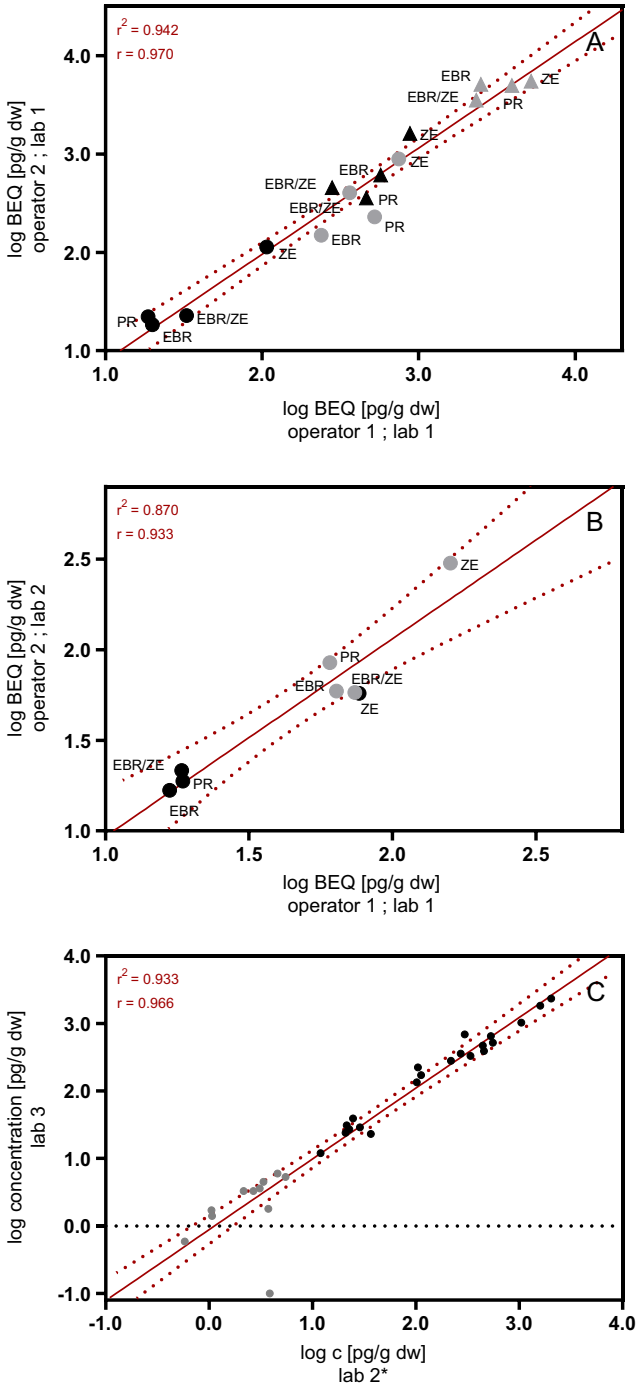
different fractions and amounted to 78%**.** There was a highly significant (r^2 = 0.942) correlation between H4IIE-luc BEQs obtained by two different operators (Fig. 2a). Among the different sediments, both operators characterized sediment ZE and its fractions to possess the highest overall AhR activating potential (Fig. 2a).

Inter-laboratory comparisons

Results obtained for the H4IIE micro-EROD assays were highly similar between both operators and laboratories. A one-tailed Student's *t* test, which due to unequal variances was adapted by Welch's correction ($p < 0.005$), indicated inter-laboratorial achieved LODs to not significantly differ from each other ($p = 0.752$).

Repeatability achieved by both operators was lower for DL-PCB than for PCDD/F fractions and averaged to 76 and 88% for lab 1 and 2, respectively. The between-laboratory reproducibility for sediment extracts and single substance 2,3,7,8-TCDD averaged to 83 and 98%, respectively. Average LOD and LOQ values of the micro-EROD assay were 0.4 and 0.7 pM 2,3,7,8-TCDD, respectively (Table 1). Furthermore, the correlation of micro-EROD assay results obtained by different laboratories was significant (r^2 = 0.87, Fig. 2b). Both operators and laboratories found highest AhR activating potential for fractions of sediment ZE.

Concerning the inter-laboratory comparison of HRGC/HRMS measurements, a highly significant correlation was found between DL-PCB and PCDD/F congener concentrations determined in sediments PR and EBR (r^2 = 0.933; Fig. 2c and Table 2), even though extracts were differently extracted and cleaned.



Summarized consideration of bioanalytical results

Multilayer and raw fractions were only analyzed using the EROD and H4IIE-luc assays (data not shown). In both assays, the sum of activities of DL-PCB and PCDD/F fractions was lower than total activity measured in multilayer extracts (by 73 and 82% for DL-PCB and PCDD/F fractions, respectively). Regarding intra- and inter-laboratorial achieved BEQs of the three different assays, they all indicated sediment ZE and

◀ **Fig. 2** Pearson correlations of an intra-laboratory comparison via H4IIE-luc EC_{25}BEQs (**a**) as well as inter-laboratory comparisons via H4IIE micro-EROD EC_{25}BEQ (**b**) and instrumental determined concentrations (**c**) in different sediments extracts and fractions thereof. EBR = Ehrenbreitstein (Rhine), PR = Prossen (Elbe), ZE = Zollelbe (Elbe), and a 1:10 mixture (EBR/ZE) consisting of one dry weight part EBR and nine dry weight parts ZE. Black and gray circles show BEQs (**a**, **b**) and single congener concentrations (c**)** determined for sediment DL-PCB and PCDD/F fractions, respectively. Black and gray triangles show BEQs determined for multilayer fractions and raw extracts, respectively. Linear regression line is depicted with its 95% confidence interval (dashed red line). Lab 1 = RWTH Aachen University, lab 2 = German Federal Institute for Hydrology (BfG), lab 2* = BfG contract laboratory, lab 3 = Münster Analytical Solutions (mas)

fractions thereof to possess the overall highest toxicity. The fish cell line RTL-W1 produced the highest and the mammalian wild-type cell line H4IIE the lowest BEQs (Fig. 3a, b). Logarithmic BEQs of the EROD assay correlated well with those determined via the H4IIE-luc (r^2 = 0.930) and micro-EROD assays (r^2 = 0.910). Logarithmic BEQs of the H4IIE-luc and micro-EROD assays are correlated well, too (r^2 = 0.900).

Intra- (H4IIE-luc) and inter-laboratory (micro-EROD) BEQs were comparable (p = 0.008). Only the micro-EROD BEQs obtained for ZE PCDD/F fraction significantly differed between the two operators and laboratories (Fig. 3b).

Table 2 Concentrations of PCDD/F and DL-PCB congeners in pg g^{-1} in sediments from Prossen/Schmilka (PR) and Ehrenbreitstein (EBR), measured by two independent laboratories (lab 2* = BfG contract laboratory; lab 3 = Münster Analytical Solutions); (n.d. = not determinable), data built the basis for the correlation analysis shown in Fig. 2c

	Sediment PR		Sediment EBR	
	Lab 3	Lab 2*	Lab 3	Lab 2*
2,3,7,8-PCDDs				
2,3,7,8-TetraCDD	0.6	0.5	0.6	0.6
1,2,3,7,8-PentaCDD	0.7	0.7	1.1	1.4
1,2,3,4,7,8-HexaCDD	0.8	0.7	1.1	1.7
1,2,3,6,7,8-HexaCDD	1.4	1.5	3.4	4.5
2,3,7,8-PCDFs				
2,3,7,8-TetraCDF	6.8	8.8	4.6	6.0
1,2,3,7,8-PentaCDF	2.5	3.7	2.2	3.3
2,3,4,7,8-PentaCDF	4.0	4.5	3.1	3.6
1,2,3,4,7,8-HexaCDF	3.7	5.2	5.5	5.3
1,2,3,6,7,8-HexaCDF	2.3	3.1	2.7	3.3
1,2,3,7,8,9-HexaCDF	1.4	0.4	n.d.	n.d.
2,3,4,6,7,8-HexaCDF	3.1	2.1	3.7	1.8
Non-ortho PCB				
PCB 77	272.2	357.0	297.2	686.0
PCB 81	12.0	12.0	n.d.	5.0
PCB 126	20.9	24.0	24.7	39.0
PCB 169	n.d.	10.0	n.d.	10.0
Mono-ortho PCB				
PCB 105	457.8	391.0	446.5	470.0
PCB 114	28.9	29.0	22.7	27.0
PCB 118	2041.7	2330.0	1600.1	1830.0
PCB 123	36.8	23.0	21.5	31.0
PCB 156	1052.2	1030.0	532.8	650.0
PCB 157	105.0	223.0	102.1	134.0
PCB 167	556.2	518.0	339.0	331.0
PCB 189	219.4	279.0	112.3	172.0

Connecting the present study with German dredged material directives

When combining the results obtained for Elbe sediments analyzed by this study with TEQ values from a more extensive data set that includes multiple locations analyzed along the Elbe river (Stachel et al. 2011), the top 25% most contaminated sediments were clearly separated from the remaining samples by a TEQ limit value (LV) of 35 pg g^{-1} dw (Fig. 4a). Linear regression analysis of micro-EROD BEQs and TEQs determined for Rhine and Elbe sediment DL-PCB and PCDD/F fractions obtained during the present study (Fig. 4b) resulted in a respective BEQ_{LV} of 145 pg BEQ g^{-1} dw sediment.

Discussion

Inter- and intra-laboratory comparison study

Comparison of three in vitro bioassays

Each of the performed assays was able to detect dioxin-like activity in complex samples. Altogether, according to our findings, the micro-EROD assay with the cell line H4IIE constitutes the preferable bioanalytical screening tool among the examined assays. The micro-EROD assay according to its average z-factor of 0.54 could be classified as *excellent*. Its average repeatability > 75% corresponds to aforementioned regulatory recommendations (2012/278/EU 2012), and the here applied two sample-plate layout (i.e., testing of two separate samples per plate) allowed for the simultaneous testing of 16 samples per cycle. The assays' most promising criteria were its remarkably low LOD and LOQ of 0.5 and 0.7 pM 2,3,7,8-TCDD (Table 1), respectively, which approach the limits achieved by instrumental analysis such as HRGC/HRMS.

In contrast to cell line H4IIE, a two sample-plate layout turned out to be inappropriate using H4IIE-luc cells due to cross-talk of adjacent wells during the luminescence measurements (Puga et al. 2009). This limited the number of samples/cycle to six. Although luminescence is known to be one of the

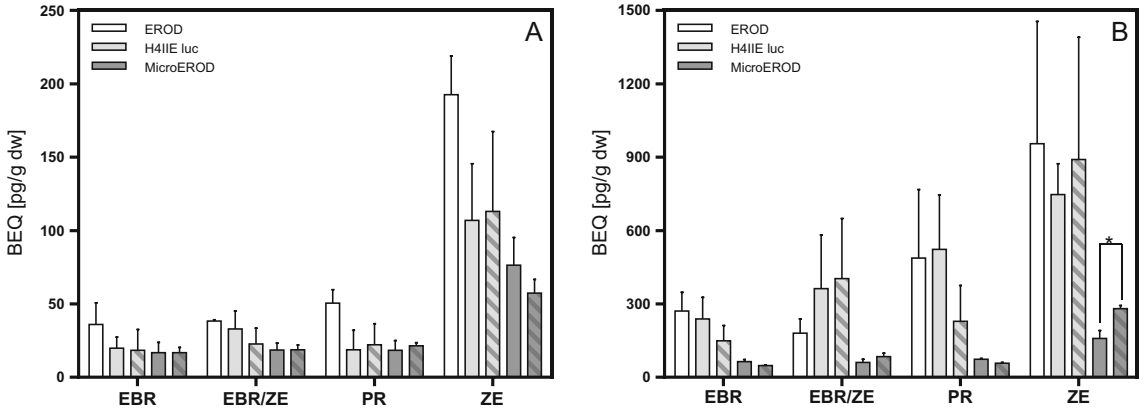


Fig. 3 Biological toxicity equivalents (BEQs) obtained for DL-PCB (**a**) and PCDD/F (**b**) fractions of three sediments from Ehrenbreitstein (EBR), Prossen (PR), and Zollelbe (ZE) as well as a mixture (EBR/ZE) consisting of nine dry weight parts EBR and one dry weight part ZE. BEQs were obtained using the RTL-W1 EROD, H4IIE-luc, and H4IIE micro-EROD assays. Dashed bars show results of intra- (H4IIE-luc) and inter-laboratory (H4IIE micro-EROD) comparisons. Bars show mean values of three independent replicates with standard deviations. Asterisks show significant differences between results obtained by different operators, which was analyzed using a Student's *t* test ($p < 0.05$)

most sensitive endpoints (Sanderson et al. 1996, Willett et al. 1997), the average LOD of 0.8 pM 23,7,8-TCDD was higher than expected and the average LOQ of 2.3 pM 2,3,7,8-TCDD could not compete with those of the remaining assays. While the assay's average repeatability was satisfactory for complex samples (69%), it was unexpectedly low (61%) for the standard 2,3,7,8-TCDD. It should be emphasized that a better performance of the assay can be expected for the 24-h incubation time described in the original protocol (Sanderson et al. 1996) as indicated by Lee et al. (2013). The distance between the average EC_{25}TCDD and the average EC_{10}TCDD showed that the H4IIE-luc assay covered the widest concentration range among the analyzed assays (Table 1), hence might compensate for time-consuming range finding tests prior to the actual assay.

The RTL-W1 EROD assay allows for the testing of up to 36 samples per cycle, which constitutes the assays' most promising feature and equates three to five times the testing capacity of mammalian cells such as H4IIE and H4IIE-luc. High sample numbers require long test periods, which due to differing culture conditions (e.g., lower temperature, no need for culture in CO_2 atmosphere) are much better tolerated by the fish cell RTL-W1 compared to their mammalian relatives. RTL-W1 cells are very slow growing (one doubling after 72 h) and have stable cytochrome concentrations even at high passage numbers. Hence, they allow for the testing of large numbers of samples using a single subculture (Lee et al. 1993), which decreases subculture-related variability. Nevertheless, repeatability of the RTL-W1 EROD assay was higher than formerly determined (2012/278/EU 2012; Besselink et al. 2004; Engwall and Van Bavel 2004) and showed values of 64 and 70% for extracts and single substance 2,3,7,8-TCDD, respectively. Hence, repeatability decreased with increasing sample

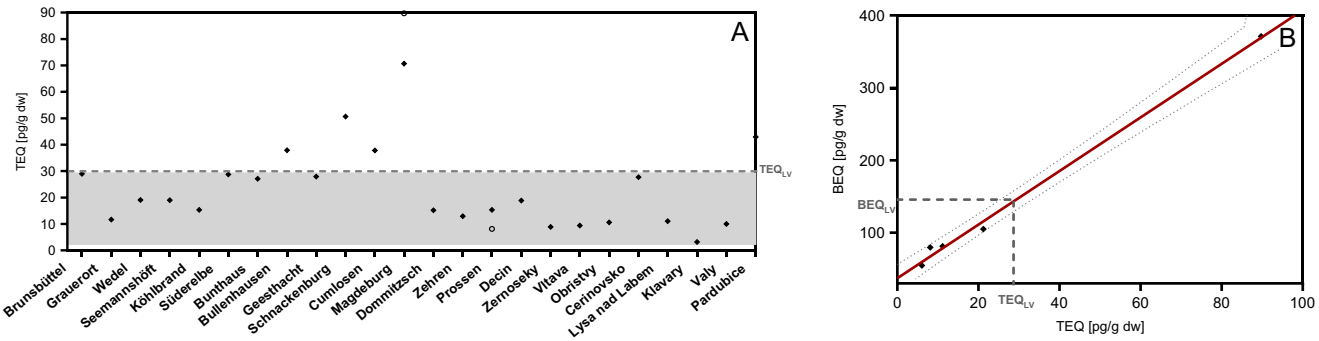


Fig. 4 **a** Sum of DL-PCB and PCDD/F WHO_{2005} toxicity equivalents (TEQs) measured via HRGC/HRMS in sediment extracts of an Elbe length profile sampling campaign conducted by Stachel et al. (2011). Sampling locations follow the rivers course. Dashed line represents a TEQ limit value (LV) derived from the 75% percentile (gray area) of all data points; blank dots at locations Magdeburg and Prossen represent TEQs measured in the present study. **b** Linear correlation of TEQs [pg/g dw] and biological equivalents (BEQs); TEQs represent the sum of DL-PCB and PCDD/F measured in sediments of rivers Elbe and Rhine of the present study, whereas BEQs show the AhR inducing potential of the respective DL-PCB and PCDD/F fractions analyzed using the micro-EROD assay with H4IIE cells. Through the TEQ_{LV} derived in **b**, a limit value for BEQ calculation was determined (dashed line, BEQ_{LV})

complexity, which corresponds to previous observations of Besselink et al. (2004).

Moreover, care has to be taken when samples with low EROD-inducing potential have to be evaluated because of the assay's relatively high average LOD and LOQ of 0.9 and 1.7 pM 2,3,7,8-TCDD (Table 1), respectively. The fact that the *z*-factor of 0.36 was accompanied by the overall highest standard deviation most likely indicates high intra-assay fluctuations of positive and negative control (Zhang et al. 1999) due to the *z*-factors' high sensibility towards variability. In conclusion, the EROD assay (RTL-W1) constitutes an assay particularly suitable for the pre-screening of large sampling sets, and despite its comparably high detection limits, promptly can be used to detect samples of highest concern.

Intra-laboratory comparison

Intra-laboratory results obtained for the H4IIE-luc assay were highly similar (r^2 = 0.942) between two operators of the same laboratory (Fig. 2a), but repeatability for sediment extracts of 69 and 43% achieved by operator 1 and 2, respectively, was different and indicated the assays' high variability. The within-laboratory reproducibility for sediment extracts of 83% (in contrast to the percentage of repeatability) corresponded to recommendations set by European guidelines (2012/278/EU 2012) and most likely was due to the fact that test results of both operators and laboratories were evaluated by the same operator (Table 1). Results generated by both operators indicated that sediment ZE and its fractions possessed the overall highest AhR activating potential (Fig. 2a), which demonstrated the suitability of the H4IIE-luc assay to be used as prioritizing tool in sediment evaluations.

Inter-laboratory comparisons

The inter-laboratory comparison of the H4IIE micro-EROD assay showed comparable results (r^2 = 0.87) between both operators and laboratories, but showed slightly higher deviations than the intra-laboratory comparison (Fig. 2a, b). The between-laboratory reproducibility averaged to 83%, indicating the H4IIE micro-EROD as a useful cross-laboratory method. A standard (here 2,3,7,8-TCDD), which delivers highly comparable repeated measures, is one basic requirement for implementing an assay as a regulatory tool. For the standard, between-laboratory reproducibility and repeatability on an EC_{25}TCDD level were 98 and 84% (calculated from respective standard deviations depicted in Table 1), respectively, and thus are greater than the recommendations of regulatory guidelines (2012/278/EU 2012) and performance of reported by other studies (Engwall and Van Bavel 2004).

In general, the variability observed for this assay was below the inter-laboratory variability that was achieved using the CALUX assay. The reproducibility for 2,3,7,8-TCDD and sediment extracts reached values of 86 and 80%, respectively, in case of the CALUX assay (Besselink et al. 2004). However, the fact that results of both laboratories were calculated by the same operator most likely lowered the intra-assay variation. For instance, Engwall and Van Bavel (2004) concluded for their inter-laboratory bioassay comparison study that different evaluation methods conducted by the participating laboratories influenced inter-laboratory variance. To lower this influencing factor, the authors strongly recommended standardized evaluation methods (Engwall and Van Bavel 2004).

Average LOD and LOQ values of 0.4 and 0.7 pM 2,3,7,8-TCDD were comparable to former studies (as reviewed by Eichbaum et al. 2014) and indicate that the H4IIE micro-EROD is a highly suitable screening tool for complex samples with low dioxin-like activity. The inter-assay variability of EC_{10}TCDD values was very high compared to EC_{25}TCDD levels, which possibly indicates EC_{25} values to be the more reliable effect level for sample evaluation (Table 1). Finally, both operators found the highest overall induction potential for fractions of sediment ZE, again showing the suitability as prioritization tool and supporting the results of the remaining assays (Fig. 2b).

Concerning HRGC/HRMS measurements conducted by different operators and laboratories, a highly significant correlation (r^2 = 0.933; Fig. 2c, d Table 2) revealed that extraction, clean-up, and analytical methods applied by labs 1 and 2* were comparable and robust. Findings of an intra- and inter-laboratory comparison study conducted by Besselink et al. (2004) revealed that different extraction and clean-up methodologies distinctly influenced TEQs and BEQs and most likely increased reproducibility. Consequently, it is strongly recommended to use extracts of the same origin for future cross-method comparisons.

Although DL-PCBs and PCDD/Fs show high and low concentrations, respectively, (Fig. 2a–c) bioanalytical results show low and high induction levels for DL-PCBs and PCDD/Fs, respectively, which reveals their high sensitivity towards dioxins. PCDDF/s, so far, are not present among the target compounds of guidelines for dredged material (HABAB 2000; HABAK 1999). However, this compound class essentially influences the overall induction potential in the group of DLCs; thus, the future implementation of in vitro bioassays for the screening of environmental trace contaminations with the highly relevant PCDD/Fs should be considered.

Summarized consideration of bioanalytical results

Only the H4IIE micro-EROD assay showed the same order of contamination levels (Fig. 3a, b) for DL-PCBs (EBR < EBR/ZE < PR < ZE) and PCDD/Fs (EBR < PR < EBR/ZE < ZE) as the results from the chemical analysis ("Physical chemical characterization of sediments" section). Furthermore, this

assay exhibited the smallest amount of unexplained percentages of 26.5% when compared to chemical analytical results (Fig. 1). Discrepancies between TEQs and BEQs were considerably higher than guideline recommendations of ± 20% (2012/278/EU 2012), which may be explained through antagonistic or synergistic effects, contradicting the additive character of the TEQ approach (Safe 1998a, b). Moreover, complex environmental mixtures are known to contain a certain fraction of dioxin-like inducers non-targeted by chemical analysis (Engwall and Van Bavel 2004). Moreover, adsorption of DLCs to the test vessel material might have influenced the actual test concentrations during the exposure of cells. This was observed by scientists, who compared adsorption behavior of organic chemicals to glass and plastic materials, showing that the latter one crucially lowers actual test concentrations (Palmgren et al. 2006; Zielke 2011).

EROD and H4IIE-luc assays, by which multilayer fractions containing both DL-PCBs and PCDD/Fs were investigated, showed that the sum of the activity of the two fractions (DL-PCB, PCDD/F) was lower than the total activity measured in multilayer extracts (73 and 82%, respectively; data not shown). This observation, although opposite to former findings (Manz et al. 2007), may indicate the presence of antagonistic substances in the multilayer fractions or possible compound losses during the fractionation process.

In general, BEQs of the different assays were comparable when DL-PCB fractions were investigated, while they differed for PCDD/Fs (Fig. 3a, b), indicating the bioassays differing sensitivities towards PCDD/Fs.

Bioanalytical threshold value derivation from chemical data

The present study's sampling location ZE (Magdeburg) was found to be among the top 25% of the most contaminated sediments of the Elbe river. Regarding the TEQ_{LV} of 35 pg g^{-1} dw, sediment ZE with TEQs of 70 and 90 pg g^{-1} dw as determined in the previous and in the present study, respectively, clearly separated from other sampling locations along the river such as sediment PR (Prossen/Schmilka). The TEQ_{LV} was chosen in an arbitrary manner, and thus could also be based on any other percentile of choice (Fig. 4a). The BEQ_{LV} of 145 pg BEQ g^{-1} dw sediment, deduced from a linear correlation of BEQs and TEQs determined in the present study (Fig. 4b), has to be considered as preliminary due to the limited data available for the bioassay to date (Fig. 4b). Hence, further biochemical and instrumental sediment evaluations with DL-PCB and PCDD/F fractions would have to be conducted to strengthen the basis for such in vitro assay sediment evaluations. Based on this, sediments could be evaluated and ranked using simple, rapid, and low-cost intensive in vitro bioassays such as the H4IIE micro-EROD assay.

Conclusions

The H4IIE micro-EROD assay showed the best performance within the investigated bioassays. It was ranked *excellent* (*z*-factor = 0.54), possessed a satisfactory sample/cycle-number, and its repeatability > 75% was independent of sample complexity and its remarkably low LOD and LOQ of 0.5 and 0.7 pM 2,3,7,8-TCDD, respectively, approached the limits achieved by instrumental analysis, with which the assay was highly comparable (r^2 = 0.803). Bioassay results were highly reproducible (83%) and comparable (r^2 = 0.87) between operators and laboratories. While all bioassays reliably indicated sediment ZE and its fractions to possess the overall highest dioxin-like potential among the four chosen sediments, only cell line H4IIE showed the same sequence in sediment contamination as it was determined by HRGC/HRMS.

In contrast, the RTL-W1 EROD assay due to its high sample/cycle-number was more suitable for the prescreening and prioritization of large sampling sets, and the H4IIE-luc assay due to its widest concentration range may compensate for time-consuming range finding tests.

With an exemplarily set limit value, derived from former determined Elbe sediment TEQs, we could moreover deduce a H4IIE micro-EROD-based limit value that in future regulatory decisions might be used as an additional quality measure for the assessment of sediments.

Acknowledgements The present work forms a part of the dioRAMA project ("dioxin Risk Assessment for sediment Management Approaches"), which received funds from the "Title Group 05" of the German Federal Government. M.B. received a personal stipend from the German National Academic Foundation ("Studienstiftung des Deutschen Volkes") and is currently a Banting postdoctoral fellow of the Natural Sciences and Engineering Research Council of Canada (NSERC). Profs. Dr. Giesy and Hecker were supported by the Canada Research Chair Program. Furthermore, Dr. Giesy was supported by a visiting distinguished professorship in the Department of Biology and Chemistry and State Key Laboratory in Marine Pollution, City University of Hong Kong, the 2012 "High Level Foreign Experts" (#GDW20123200120) program, funded by the State Administration of Foreign Experts Affairs, the People's Republic of China to Nanjing University, and the Einstein Professor Program of the Chinese Academy of Sciences. Prof. Dr. Hollert was supported by the Chinese 111 Program (College of Environmental Science and Engineering and Key Laboratory of Yangtze Water environment, Ministry of Education, Tongji University). Special thanks go to Dr. Stephan Hamm and Dr. Armin Maulshagen from mas (Münster Analytical Solutions GmbH), Münster, Germany for the HRGC/HRMS measurements and their kind contribution to this work.

References

2000/60/EC (2006) Directive of the European Parliament and Council—on environmental quality standards in the field of water policy and amending Directive 2000/60/EC

2000/60/EG (2009) Bewirtschaftungsplan nach Artikel 13 der Richtlinie 2000/60/EG für den deutschen Teil der Flussgebietseinheit Elbe (FGG Elbe)

2008/105/EC (2008) Directive of the European Parliament and Council—on environmental quality standards in the field of water policy and amending and subsequently repealing Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/EEC, 86/280/EEC and amending Directive 2000/60/EC

2012/252/EU (2012) Commission Regulation (EU) no 252/2012 of 21 March 2012 laying down methods of sampling and analysis for the official control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation EC1883/2006, Off J Eur Union

2012/278/EU (2012) VERORDNUNG (EU) Nr. 278/2012 DER KOMMISSION vom 28. März 2012 zur Änderung der Verordnung EG152/2009 hinsichtlich der Bestimmung der Gehalte an Dioxinen und polychlorierten Biphenylen

Ahlf W, Hollert H, Neumann-Hensel H, Ricking M (2002) A guidance for the assessment and evaluation of sediment quality a German approach based on ecotoxicological and chemical measurements. J Soils Sediments 2:37–42

Apitz S, Power E (2002) From risk assessment to sediment management an international perspective. J Soils Sediments 2:61–66

Barceló D, Petrovic M (eds) (2007) Sustainable management of sediment resources. Sediment quality and impact assessment of pollutants, Vol 1:, Amsterdam

Behnisch PA, Hosoe K, Brouwer A, Sakai S (2002) Screening of dioxin-like toxicity equivalents for various matrices with wildtype and recombinant rat hepatoma H4IIE cells. Toxicol Sci 69:125–130

Besselink HT, Schipper C, Klamer H, Leonards P, Verhaar H, Felzel E, Murk AJ, Thain J, Hosoe K, Schoeters G, Legler J, Brouwer B (2004) Intra- and interlaboratory calibration of the DR CALUX® bioassay for the analysis of dioxins and dioxin-like chemicals in sediments. Environ Toxicol Chem 23:2781–2789

Breitung V, Keller M (2010) Management von schadstoffbelasteten Feinsedimenten in Bundeswasserstraßen. Umweltwiss Schadst Forsch 22:645–650

Burton A (ed) (1992) Sediment toxicity assessment. Lewis Publishers, Inc., Chelsea, Chapter 1–3, 15 pp

Chapman PM, Hollert H (2006) Should the sediment quality triad become a tetrad, a pentad, or possibly even a hexad? J Soils Sediments 6:4–8

Clemons JH, Dixon DG, Bols NC (1997) Derivation of 2,3,7,8-TCDD toxic equivalence factors (TEFs) for selected dioxins, furans and PCBs with rainbow trout and rat liver cell lines and the influence of exposure time. Chemosphere 34:1105–1119

COM(2011)875 (2011) Report from the Commission to the Eropean Parliament and the Council on the outcome of the review of Annex X to Directive 2000/60/EC of the European Parliament and of the Council on priority substances in the field of water policy

den Besten P, de Deckere E, Babut M, Power B, DelValls TA, Zago C, Oen AP, Heise S (2003) Biological effects-based sediment quality in ecological risk assessment for European waters. J Soils Sediments 3: 144–162

Eichbaum K, Brinkmann M, Buchinger S, Reifferscheid G, Hecker M, Giesy JP, Engwall M, van Bavel B, Hollert H (2014) In vitro bioassays for detecting dioxin-like activity—application potentials and limits of detection, a review. Sci Total Environ 487:37–48

Engwall M, Van Bavel B (2004) The second round of Interlaboratory comparison of dioxin-like compounds in food using bioassays. Man Technology Environment Research Centre, Department of Natural Sciences, Örebro University, Sweden, p 314

Feiler U, Höss S, Ahlf W, Gilberg D, Hammers-Wirtz M, Hollert H, Meller M, Neumann-Hensel H, Ottermanns R, Seiler T-B, Spira D, Heininger P (2013) Sediment contact tests as a tool for the assessment of sediment quality in German waters. Environ Toxicol Chem 32:144–155

Förstner U, Ahlf W, Calmano W (2008) Entwicklung von Qualitätskriterien für Gewässersedimente. Universitätsbibliothek der Technischen Universität Hamburg-Harburg

GÜBAK (2009) Gemeinsame Übergangsbestimmungen zum Umgang mit Baggergut in Küstengewässern GÜBAK-WSV

HABAB (2000) Handlungsanweisung für den Umgang mit Baggergut im Binnenland (HABAB-WSV), BfG-1251

HABAK (1999) Handlungsanweisung für den Umgang mit Baggergut im Küstenbereich (HABAK-WSV), BfG-1100

Hallare A, Seiler T-B, Hollert H (2011) The versatile, changing, and advancing roles of fish in sediment toxicity assessment—a review. J Soils Sediments 11:141–173

Hollert H, Dürr M, Olsman H, Halldin K, van Bavel B, Brack W, Tysklind M, Engwall M, Braunbeck T (2002) Biological and chemical determination of dioxin-like compounds in sediments by means of a sediment triad approach in the catchment area of the River Neckar. Ecotoxicology 11:323–336

Hollert H, Ernst M, Seiler TB, Wölz J, Braunbeck T, Kosmehl T, Keiter S, Grund S, Ahlf W, Erdinger L, Dürr M (2009) Strategien zur Sedimentbewertung—ein Überblick. Umweltwiss Schadst Forsch 21:160–176

Höss S, Ahlf W, Fahnenstich C, Gilberg D, Hollert H, Melbye K, Meller M, Hammers-Wirtz M, Heininger P, Neumann-Hensel H, Ottermanns R, Ratte HT, Seiler TB, Spira D, Weber J, Feiler U (2010) Variability of sediment-contact tests in freshwater sediments with low-level anthropogenic contamination—determination of toxicity thresholds. Environ Pollut 158:2999–3010

ISO/5752 (2002) Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation

Khim JS, Villeneuve DL, Kannan K, Lee KT, Snyder SA, Koh CH, Giesy JP (1999) Alkylphenols, polycyclic aromatic hydrocarbons, and organochlorines in sediment from Lake Shihwa, Korea: instrumental and bioanalytical characterization. Environ Toxicol Chem 18:2424–2432

LAVES (2013) Mikro-EROD-bioassay zur Bestimmung von Dioxinen und dioxinähnlichen Substanzen (gekürzte Fassung), 03-110-MAA-M-11EROD V4, LAVES (Lower Saxony State Office for Consumer Protection and Food Safety), pp 5

Lee LEJ, Clemons JH, Bechtel DG, Caldwell SJ, Han K-B, Pasitschniak-Arts M, Mosser DD, Bols NC (1993) Development and characterization of a rainbow trout liver cell line expressing cytochrome P450-dependent monooxygenase activity. Cell Biol Toxicol 9:279–294

Lee KT, Hong S, Lee JS, Chung KH, Hilscherová K, Giesy JP, Khim JS (2013) Revised relative potency values for PCDDs, PCDFs, and non-ortho-substituted PCBs for the optimized H4IIE-luc in vitro bioassay. Environ Sci Pollut Res 20:8590–8599

Lorenzen A, Kennedy SW, Bastien LJ, Hahn ME (1997) Halogenated aromatic hydrocarbon-mediated porphyrin accumulation and induction of cytochrome P4501A in chicken embryo hepatocytes, 53. Elsevier, Amsterdam

MacDougall D, Crummett WB (1980) Guidelines for data acquisition and data quality evaluation in environmental chemistry. Anal Chem 52: 2242–2249

Manz W, Krebs F, Schipper CA, Den Besten PJ (2007) Status of ecotoxicological assessment of sedient and dredged material in Germany and the Netherlands-Dutch-German exchange (DGE) on dredged material

Otte JC, Keiter S, Faßbender C, Higley EB, Rocha PS, Brinkmann M, Wahrendorf DS, Manz W, Wetzel MA, Braunbeck T, Giesy JP,

Hecker M, Hollert H (2013) Contribution of priority PAHs and POPs to Ah receptor-mediated activities in sediment samples from the River Elbe Estuary, Germany. PLoS One 8(10):e75596

Palmgren JJ, Monkkonen J, Korjamo T, Hassinen A, Auriola S (2006) Drug adsorption to plastic containers and retention of drugs in cultured cells under in vitro conditions. Eur J Pharm Biopharm 64:369–378

Puga A, Ma C, Marlowe JL (2009) The aryl hydrocarbon receptor cross-talks with multiple signal transduction pathways. Biochem Pharmacol 77:713–722

Safe SH (1998a) Development validation and problems with the toxic equivalency factor approach for risk assessment of dioxins and related compounds. J Anim Sci 76:134–141

Safe SH (1998b) Limitations of the toxic equivalency factor approach for risk assessment of TCDD and related compounds. Teratogen Carcin Mut 17:285–304

Sanderson JT, Aarts JMMJG, Brouwer A, Froese KL, Denison MS, Giesy JP (1996) Comparison of Ah receptor-mediated luciferase and ethoxyresorufin-O-deethylase induction in H4IIE cells: implications for their use as bioanalytical tools for the detection of polyhalogenated aromatic hydrocarbons. Toxicol Appl Pharmacol 137:316–325

Schiwy A, Brinkmann M, Thiem I, Guder G, Winkens K, Eichbaum K, Nuszer L, Thalmann B, Buchinger S, Reifferscheid G, Seiler T-B, Thoms B, Hollert H (2015) Determination of the CYP1A-inducing potential of single substances, mixtures and extracts of samples in the micro-EROD assay with H4IIE cells. Nat Protocols 10:1728–1741

Schüttrumpf H, Brinkmann M, Cofalla C, Frings RM, Gerbersdorf SU, Hecker M, Hudjetz S, Kammann U, Lennartz G, Roger S, Schäffer A, Hollert H (2011) A new approach to investigate the interactions between sediment transport and ecotoxicological processes during flood events. Env Sci Eur 23:39

Stachel B, Mariani G, Umlauf G, Götz R (2011) Dioxine und PCBs in Feststoffen aus der Elbe, ihren Nebenflüssen und der Nordsee (Längsprofilaufnahme 2008)

Umlauf G, Christoph EH, Bidoglio G (2004) Schadstoffuntersuchungen nach dem Hochwasser vom August 2002 – Ermittlung der Gefährdungspotenziale an Elbe und Mulde (Endbericht BMBF-FKZ PJT 0330492). In: Geller W, Ockenfeld K, Boehme M, Knoechel A (eds), ISBN: 3-00-013615-0

US-EPA (1994) Method 8290 - Analysis of polychlorinated dibenzodioxins and poly-chlorinated dibenzofurans by high resolution gas chromatography/high resolution mass spectrometry (HRGC/HRMS)

Van den Berg M, Birnbaum LS, Denison MS, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, Peterson RE (2006) The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol Sci 93:223–241

Wenning RJ, Ingersoll CG (2002) Summary of the SETAC Pellston Workshop on use of sediment quality guidelines and related tools for the assessment of contaminated sediments. 17–22 August 2002, Fairmont, Montana, USA. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola FL, USA

White SS, Birnbaum LS (2009) An overview of the effects of dioxins and dioxin-like compounds on vertebrates, as documented in human and ecological epidemiology. J Environ Sci Heal C 27:197–211

Willett KL, Gardinali PR, Sericano JL, Wade TL, Safe SH (1997) Characterization of the H4IIE rat hepatoma cell bioassay for evaluation of environmental samples containing polynuclear aromatic hydrocarbons (PAHs). Arch Environ Contam Toxicol 32:442–448

Wölz J, Borck D, Witt G, Hollert H (2009) Ecotoxicological characterization of sediment cores from the western Baltic Sea (Mecklenburg Bight) using GC–MS and biotests. J Soils Sediments 9:400–410

Zhang J-H, Chung TDY, Oldenburg KR (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen 4:67–73

Zielke H (2011) Time-related alterations and other confounding factors in direct sediment contact tests, PhD thesis RWTH-Aachen University: pp 216