Chemosphere 287 (2022) 132146

Contents lists available at ScienceDirect

Chemosphere

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

Assessment of oil refinery wastewater and effluent integrating bioassays, mechanistic modelling and bioavailability evaluation

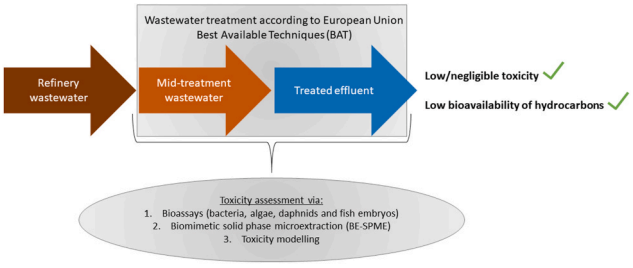
G.F. Whale [a], M. Hjort [b,*], C. Di Paolo [b,c], A.D. Redman [b,d], J.F. Postma [e], J. Legradi [f], P.E. G. Leonards [f]

[a] Whale Environmental Consultancy Limited, 55 Earlsway, Curzon Park, Chester, CH48AZ, United Kingdom
[b] Concawe, Boulevard du Souverain 165, 1160, Brussels, Belgium
[c] Shell International, Shell Health Risk Science Team, The Hague, the Netherlands
[d] ExxonMobil Petroleum and Chemical, Machelen, Belgium
[e] Ecofide, Singel 105, 1381 AT, Weesp, the Netherlands
[f] Department of Environment & Health, Vrije Universiteit Amsterdam, 1081 HV, Amsterdam, the Netherlands

HIGHLIGHTS

- Assessment of oil refinery wastewater treatment using toxicity bioassays.
- Comparison of hydrocarbon bioavailability by BE-SPME to observed toxicity.
- BE-SPME combined with simple (e.g. bacterial) bioassays are useful screening tools.
- All methods demonstrated significant decrease in toxicity after effluent treatment.
- All methods indicated negligible toxicity in the final treated effluents.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Willie Peijnenburg

Keywords:
Refinery
Petroleum hydrocarbons
Effluents
Effect based methods
SPME

ABSTRACT

Water is used in petroleum oil refineries in significant volumes for cooling, steam generation and processing of raw materials. Effective water management is required at refineries to ensure their efficient and responsible operation with respect to the water environment. However, ascertaining the potential environmental risks associated with discharge of refinery effluents to receiving waters is challenging because of their compositional complexity. Recent European research and regulatory initiatives propose a more holistic approach including biological effect methods to assess complex effluents and surface water quality. The study presented here investigated potential effects of effluent composition, particularly hydrocarbons, on aquatic toxicity and was a component of a larger study assessing contaminant removal during refinery wastewater treatment (Hjort *et al* 2021). The evaluation of effects utilised a novel combination of mechanistic toxicity modelling based on the exposure composition, measured bioavailable hydrocarbons using biomimetic solid phase microextraction (BE-SPME), and bioassays. The results indicate that in the refinery effluent assessments measured bioavailable hydrocarbons using BE-SPME was correlated with the responses in standard bioassays. It confirms that bioassays are providing relevant data and that BE-SPME measurement, combined with knowledge of other known non-hydrocarbon toxic constituents, provide key tools for toxicity identification. Overall, the results indicate that oil refinery effluents treated in accordance to the EU Industrial Emissions Directive requirements have low to

* Corresponding author. Concawe, Boulevard du Souverain, 165, B-1160, Brussels, Belgium.
E-mail address: markus.hjort@concawe.eu (M. Hjort).

https://doi.org/10.1016/j.chemosphere.2021.132146
Received 8 June 2021; Received in revised form 31 August 2021; Accepted 1 September 2021
Available online 4 September 2021

negligible toxicity to aquatic organisms and their receiving environments. Low-cost, animal-free BE-SPME represents a compelling tool for rapid effluent characterization.

1. Introduction

Water is a critical component of petroleum oil refinery operations. It is used in numerous applications including cooling, steam generation and hydrocarbon/chemical processing. The levels of contaminants found in wastewater vary due to a different factors including refinery complexity and the application from which the wastewater originates (Concawe, 1999). Contaminants and components of environmental concern which can be present in refinery wastewaters include oil, ammonia, phenols, substances responsible for oxygen depletion (as measured by their biological oxygen demand (BOD) and chemical oxygen demand (COD)), suspended solids and metals.

Refinery wastewater treatment systems are designed to remove inorganic and organic constituents to reduce contaminant loads. To assure treatment efficiency and environmental protection, since 1969, Concawe[1] has been gathering and compiling data on aqueous effluents from European oil refinery installations. Over this period technological advances and the installation of increasingly sophisticated treatment systems have led to significant reductions in the discharge of oil and of most other refinery pollutants (Concawe, 2020a). Techniques to investigate the potential impacts of effluents have also advanced since the first environmental and biological effects assessments of oil refinery effluents using ecological field monitoring programmes (Concawe, 1979) and effluent toxicity case studies (Concawe, 1982) were undertaken.

For example, whole effluent assessments (WEA) have included both evaluations using raw effluent and organic extracts and may also include evaluations of the effluent and in the receiving waters (Chapman, 2000; Comber et al., 2015; Hamers et al., 2018). It is recognised that WEA are complex and should be used in combination with other risk assessment tools. A range of bioassays including standard methods (de Vlaming et al., 2000; Smolders et al., 2003) and *in vitro* (Hamers et al., 2018) test systems have been developed to support chemical measurement-based assessments. It has been observed that often chemical measurements are not sufficient to explain observed toxicity, due to the likely under characterization of actual exposure (Hamers et al., 2018; Teodorović et al., 2009). For refinery effluents, this is further complicated by the complex nature of petroleum substances where the specific composition can affect the observed toxicity (Redman and Parkerton, 2015).

In response, Concawe has continued to develop methods and data to improve risk assessment using 1) sophisticated chemical analyses (e.g. 2-dimensional gas chromatography (Redman et al., 2012; Redman et al., 2017a), 2) biological based tools including whole-effluent toxicity (WET) assessments (Comber et al., 2015) and mesocosm studies (Cailleaud et al., 2019) and 3) application of biomimetic extraction to quantify the bioavailable fraction of organics in the effluent using solid phase microextraction (BE-SPME) as described by Leonards et al. (2011), Redman et al. (2017a), Redman et al. (2018a). The BE-SPME method has particular applicability to refinery WEA because it provides measurements of the total bioavailable hydrocarbons thereby providing broad analytical coverage of the relevant chemical exposure. Further, BE-SPME has been correlated with acute and chronic aquatic toxicity for hydrocarbons (Redman et al., 2018a), and therefore provides a mechanistic basis for evaluating potential toxicity observed in a WEA.

Concawe's recent activities have included assessing whether member company oil refinery discharges are meeting pollutant emissions criteria laid down in the Industrial Emissions Directive 2010/75/EU (IED) (EU, 2010), as well as providing data to support other risk assessment requirements and initiatives. The IED aims to achieve a high level of protection of human health and of the environment taken as a whole by reducing harmful industrial emissions across the EU, in particular through better application of Best Available Techniques (BAT). Under the IED, the BAT Reference document (BREF) for the refining of mineral oil and gas (REF BREF) describes the BATs (EU-JRC, 2015). Further, effluent parameters that refineries are obliged to monitor and control are described by BAT Associated Emission Levels (BAT-AELs) mentioned in the associated Commission Implementing Decision 2014/738/EU (EU, 2014).

However, in addition to the analytical focus on chemicals the implementation of the EU Water Framework Directive (WFD - 2000/60/EC) has led to increased attention being paid to biological effects. This is because the WFD regulates both the ecological and chemical status of surface water bodies. Good ecological status is defined in Annex V of the WFD, in terms of the quality of the biological community, the hydrological characteristics and the chemical characteristics (EC, 2000). Furthermore, there have been recommendations from the EU Solutions project to integrate effect-based methods (EBMs) to improve the monitoring of water quality (Brack et al., 2019).

Due to their compositional complexity, oil refineries are on occasions required to use effluent toxicity tests in addition to chemical analytical methods to monitor their effluents. Although the predominant reason for performing toxicity testing was to satisfy permit/regulatory conditions, a number of refineries voluntarily performed toxicity assessments (Comber et al., 2015; Concawe, 2018a) and mesocosm studies (Cailleaud et al., 2019). A summary of WEA using various types of biological-based effect methods and their value to provide additional relevant data to assess oil refinery effluent discharges and surface water monitoring have been reported (Concawe, 2012; Comber et al., 2015). The current study was undertaken to provide additional confidence that WEA tools might be used to complement traditional analytical approaches to improve the environmental impact assessment of oil refinery wastewaters and effluents.

In a companion study to the research reported here, investigations were undertaken to provide information regarding contaminant removals from refinery wastewaters as these pass through various treatment steps (Hjort et al., 2021). The refineries used in that study were shown to operate according to BAT expectations, as almost all values from spot sample analysis were observed to be within BAT-AEL ranges. Removal factors of total petroleum hydrocarbons (TPH) in the wastewater treatment plants (WWTP) of the 13 refineries varied from 97% to >99.8% (Hjort et al., 2021). These same samples were used in the present study for WEA. By undertaking both studies the effectiveness of wastewater treatment could be assessed using both chemical and biologically based parameters. To our knowledge this is the first time such a combined systematic approach has been applied to assess petroleum refinery effluents.

The specific objectives of the biological assessments were to perform WEAs on refinery effluents using both raw effluent and organic extracts from the effluent. The aquatic toxicity assessment of organic extracts enabled the focus to be primarily on the organic contaminants present in the wastewaters. This was justified because the concentrations of metals and confounding factors (e.g. nitrite, ammonium and conductivity) were

[1] Concawe was established as CONCAWE (CONservation of Clean Air and Water in Europe) in 1963 by a small group of leading oil companies to carry out research on environmental issues relevant to the petroleum refining industry. Its membership has broadened and currently includes most oil companies operating in EU-27, UK, Norway and Switzerland, representing approximately 99% of petroleum refining capacity in those countries. In 2014, it became the Scientific Division of the European Petroleum Refiners Association.

relatively low (Hjort et al., 2021).

For each toxicity bioassay, responses of the organisms were compared to TPH (i.e. oil in water), BE-SPME measurements and predicted toxicity. The predicted toxicity, expressed as Toxic Units (TU), was calculated by PetroTox (Redman et al 2012, 2014b; Concawe, 2020b). The latter required advanced analytical characterization to support mechanistic modelling, used within PetroTox, to explain the observed toxicity with respect to the specific composition of the organic material found in the refinery effluents.

In addition, BE-SPME measurements were taken to provide a direct measurement of bioavailable hydrocarbons for comparison to the observed toxicity. These comparisons were used to evaluate the potential of using BE-SPME to support WEA due to its potential to simplify WEA by providing a measurement of total bioavailable hydrocarbons. The relationship between aquatic toxicity and BE-SPME was specifically investigated to assess if this could be used as a suitable indicator of effluent toxicity because this technique has already proven to be useful in assessing the hazard of petroleum substances (Leonards et al., 2011; Redman et al., 2014b, 2017b, 2018a; Concawe, 2020c). Furthermore, the data generated using EBMs, including *in vivo* bioassays, provided an opportunity to evaluate 1) whether these could help elucidate factors responsible for toxicity in refinery effluents and 2) improve assessment of potential for refinery effluents to impact on receiving surface water quality.

2. Material and methods

2.1. Background

The investigations were undertaken as a companion study to an analytical investigation to assess the efficacy of oil refinery wastewater treatment systems. Toxicity tests were performed on a combination of whole (raw) effluents and reconstituted samples derived on XAD®-extracts which effectively removed metals and other confounding factors. Chemical analyses, including Total Petroleum Hydrocarbons (TPH) and biomimetic solid phase micro extraction (BE-SPME) were performed to estimate exposure levels and measure bioavailable concentrations as described by Letinski et al. (2014). BE-SPME is a specific application of SPME and is considered a proxy for toxicity as has been described by Parkerton et al. (2000); Redman et al. (2018a). This has previously been used to assess the toxicity of both oil refinery effluents (Comber et al., 2015), oil sand process waters (Redman et al., 2018b), produced waters from oil and gas operations (Worden et al., 2021) and can be used to assess potential toxicity of oil spills (Letinski et al., 2014).

In essence, traditional analytical methods fail to characterize the bioavailability of individual hydrocarbon components because TPH are very often not related to bioavailable fractions due to variable compositions (Letinski et al., 2014; Redman and Parkerton, 2015; Redman et al., 2017a; Hedgpeth et al., 2019). Also, hydrocarbons act by a common narcotic mode of action and ecotoxicity occurs when their corresponding molar concentration in an organism's lipid exceeds a critical threshold known as the target lipid model (Di Toro et al., 2000). Since the ecotoxicity of narcotic mixtures appears to be additive (McGrath et al., 2005; McGrath and Di Toro, 2009; Redman et al., 2017a), ecotoxicity will depend upon the partitioning of individual hydrocarbons from the environment to lipids and the total molar sum of individual hydrocarbons in lipids. This has led to the development of the concept of 'biomimetic' extraction as a novel analytical tool for assessing narcosis-type or 'baseline" toxicity. The BE-SPME method is an operational measurement of bioavailable hydrocarbons and is intended to represent the accumulation of constituents at the sensitive tissues, e.g., target lipid. See Letinski et al. (2014) for additional mechanistic discussion.

The BE-SPME measurements were collected on the 100% effluent samples (or equivalent 100% loading with the XAD® extracts), and measurements were not collected for dilutions in the bioassays described below. Based on the past several years of research there are several BE-SPME-based critical thresholds, e.g. when the exposure is lethal to 50% of the studied population (LC_{50}), for a variety of aquatic organisms, for a variety of test substances and effluents (Redman et al., 2014a, 2017b, 2018a, 2018b; Redman, 2015; Redman and Parkerton, 2015 and Hansen et al., 2019). Therefore, the thresholds for acute and chronic toxicity are relatively well understood thus providing a strong basis for interpreting the results of the bioassays.

2.2. Sample collection

To align with the companion analytical project (see Hjort et al., 2021) this study was also undertaken in two phases. The first phase assessed mid-treatment (wastewater) as well as final effluent spot samples from four refineries. The second phase encompassed nine additional refineries in which final effluents were assessed. A fuller description regarding treatment trains and sample collection can be found in Hjort et al. (2021) with a brief description of the samples, together with toxicity tests undertaken and recipient environment, summarised in Table 1. Note that due to competition law restrictions, the refineries are coded and not represented by their names and locations.

To reduce errors sampling kits for sample collection were prepared and provided with a sampling protocol to the refineries. The approach was for single spot samples, which were taken in 2015 and 2016, was based on previous Concawe experience of collecting and ensuring integrity of oil refinery effluent samples (Concawe, 2010). Operators were asked to ensure that the sample line or collection device was flushed so that the effluent quality (visual, olfactory) was stable prior to sample collection. Sampling could not be aligned with the water retention in the full wastewater treatment train. Therefore, some results are not directly comparable. Nevertheless, a snapshot of refinery wastewater treatment is obtained.

Samples for bioassays were collected in 20 L stainless steel containers, which were completely filled with sample, leaving no headspace and then securely capped.

Samples for other analyses were effectively conserved using 500 mL of 4% nitric acid added to 20L stainless steel containers prior to the addition of the sample. The samples were shipped as soon as possible after collection using express couriers to the Vrije Universiteit Amsterdam (the Netherlands) where they were stored at 4 °C prior to XAD® extraction or use in toxicity tests.

2.3. Sample preparation

2.3.1. Raw (untreated) samples

These samples were taken directly from the sample container and gently warmed up to the appropriate test temperature prior to use. Where appropriate adjustments were made (e.g. by the addition of salts/nutrients) to the untreated effluent to ensure the media were appropriate for the test organisms. For example, in the Microtox® and algae tests, salts and nutrients were added corresponding to the standard test medium. However, *D. magna* were tested in the raw effluent.

2.3.2. XAD®-extracts

XAD®-extracts were prepared according to Struijs and van de Kamp (2001). The XAD®-extraction procedure was undertaken to ensure that the organic compounds could be extracted from the sample to remove these from other components of the sample, such as metals, ammonium and nitrate, therefore excluded possible confounding toxicity factors. The XAD®-extraction method has been validated for 27 chemicals by Struijs and van de Kamp (2001). Struijs and van de Kamp showed that the average recovery of 27 chemicals, with a wide range of physico-chemical and biological properties, was 62% including surfactants. The average recovery of 11 pesticides was for instance 70%, and for anionic LAS and nonionic glycol ethers they varied between 40% and 80%, respectively. Volatile compounds were also extracted with

Table 1
Overview of sampling codes and toxicity tests performed (Phase 1 samples are shaded - For more details on refinery treatment steps and codes see Hjort et al., 2021).

Sample refinery	Effluent type	Treatment step	Microtox®	Algae	Daphnia	Zebrafish	Receiving environment
D(23)	Raw	Final effluent	X	X	X	X	River
A(4)	Raw	Final effluent	X	X	X	X	River
C(16)	Raw	Final effluent	X	X	X	X	River
B1(7)	Raw	Mid treatment wastewater before polishing step	X	X	X	X	uWWTP
Control	-	XAD	X	X	X	X	na
D(18)	XAD	Mid treatment wastewater after floatation	X	X	X	X	River
D(22)	XAD	Mid treatment wastewater after biological treatment	X	X	X	X	River
A(2)	XAD	Mid treatment wastewater after floatation	X	X	X	X	River
A(3)	XAD	Mid treatment wastewater after biological treatment	X	X	X	X	River
D(23)	XAD	Final effluent	X	X	X	X	River
A(4)	XAD	Final effluent	X	X	X	X	River
C(16)	XAD	Final effluent	X	X	X	X	River
B1(7)	XAD	Mid treatment wastewater before polishing step	X	X	X	X	uWWTP
E(26)	XAD	Final effluent	X	X	X	nt	River
F(28)	XAD	Final effluent	X	X	X	nt	uWWTP
G(31)	XAD	Final effluent	X	X	X	nt	River
H(33)	XAD	Final effluent	X	X	X	nt	Estuary
I1(37)	XAD	Final effluent	X	X	X	nt	River
K(45)	XAD	Final effluent	X	X	X	nt	River
I2(39)	XAD	Final effluent	X	X	X	nt	River
L(49)	XAD	Final effluent	X	X	X	nt	Sea
J(42)	XAD	Final effluent	X	X	X	nt	River
M(51)	XAD	Final effluent	X	X	X	nt	Sea

Notes: Final effluents are those from the oil refinery wastewater treatment plant (WWTP) and may be subject to further off-site treatment. uWWTP = Urban wastewater treatment plant na = not applicable, nt = not tested (the Zebrafish test was not conducted due to lack of observed toxicity in initial tests).

efficiencies from 28% (dichlorobenzene) to 64% for pentachlorobenzene.

In this study, organic compounds from the samples were extracted with a mixture of resins XAD®-4 and XAD®-8. After extraction the resins were dried and eluted with acetone. The bulk of the acetone was subsequently removed by Kuderna-Danish distillation. The distillation residue was then stored and added to demineralised water just before toxicity testing was initiated. The final concentration of substances in the demineralised water was equal the concentration (based on the initial amount of water extracted with the XAD®) in the original sample and acetone concentration was on a non-toxic level. Each XAD®-extract was subsequently divided into separate sample fractions for each toxicity test. Salt solutions were then added according to the standardized test medium specified by the test guidelines. As a procedure blank (XAD® control), demineralised water was extracted in a comparable manner and simultaneously tested with the XAD®-extracts of the effluent samples.

2.4. Toxicity assessments

Selection of the bioassays (toxicity tests) was based on past experience with the testing of similar samples, the types of tests which oil

refineries had been required to undertake; feasibility (i.e. resource and sample volumes) and types of tests being currently considered under the WFD. For the latter, this considered the outcomes of a review (Concawe, 2018b) of potential tests being considered by the Sub Group working on the "Proposal for Effect-Based Monitoring and assessment in the Water Framework Directive" (Carere et al., 2019).

An overview of the types of tests which have been undertaken as elucidated from the 2010 and 2013 surveys indicate that the predominant ecotoxicity tests were with *Daphnia* and fish embryos (fertilised fish eggs and unspecified fish tests). Bacterial bioluminescence tests were also used as these have been shown to be a useful toxicity screening tool (Concawe, 2012). Apart from the fish tests (where fish cell lines had been proposed due to ethical considerations) all of these tests were deemed to be appropriate in the EBT review (Concawe, 2018b).

Whole effluent toxicity tests were therefore carried out using the bioluminescent bacteria *Aliivibrio fischeri* (referred to by its trade name Microtox®), algae (*Raphidocelis subcapitata*), daphnids (*Daphnia magna*) and zebrafish embryos (*Danio rerio*). Toxicity tests were performed on samples from mid-treatment wastewater, final effluents or XAD®-extracts for refineries as summarised in Table 1.

As such these biological effect methods included a rapid acute screening test (Microtox® test), an acute (5 day) fish embryo test (zebrafish) and two chronic (algae and daphnids, 3 and 21 day respectively) toxicity tests. Furthermore, these included representatives of several key trophic/functional levels notably degraders (bacteria), primary producers (algae), primary consumers (*Daphnia*) and secondary consumers (fish).

Toxicity tests were performed in open test vessels with the results expressed as lethal/adverse effect (effluent/sample) concentrations observed in x% of the test organisms (LC_x/EC_x values) after a defined exposure period.

All tests were started within one week after preparing the XAD®-extracts, but typically within four days. For semi-static test procedures (e.g. chronic test with *D. magna* and tests with early life-stages of zebrafish), test solutions were stored in the refrigerator at 4 °C. To determine if changes in toxicity occurred during storage, Microtox®-tests were conducted on three final effluents and one mid-treatment wastewater sample on arrival at the laboratory and after these had been stored for 4 weeks.

2.4.1. Microtox® assay with marine bacteria Aliivibrio fischeri

In the Microtox®-test, bioluminescence of the bacterium *A. fischeri* (formerly known as *Vibrio fischeri* and *Photobacterium phosphoreum*) is used to assess possible acutely toxic effects. The test was performed according to the ISO 11348-3 guideline (ISO, 2007) in which the bioluminescence is determined after a short exposure period (typically 5, 15 and 30 min) to a range of test concentrations (in this study effluent dilutions). As the bacterium is a marine species, in line with the ISO guidance, salt (molarity adjustment) solutions were added to each test solution which therefore lowered the maximum concentration to be tested to 45 vol% of the original sample. Each effluent concentration was tested at 15 °C in duplicate. All toxicity tests were valid based on the validation criteria described in the luminescent bacteria test ISO 11348-3 guidelines. The EC_{20} and EC_{50} values, including confidence limits, were subsequently calculated using the Microtox Omni Software by the decrease in bioluminescence in the samples relative to the controls.

2.4.2. Algae Raphidocelis subcapitata

Chronic growth inhibition was determined with the freshwater algae species *Raphidocelis subcapitata* (formerly known as *Selenastrum capricornutum* and *Pseudokirchneriella subcapitata*) based on the test procedures specified in the international guideline ISO 8692:2012 (ISO, 2012). All samples were tested in triplicate for 72 h (with six control replicates) in unsealed glass Erlenmeyer flasks. Additional blank controls (i.e. without the addition of algae) were prepared for all concentrations assessed to check for background noise (i.e. influence of particulates in raw samples).

Algal growth was measured on an approximately 24 h basis with tests conducted at 23 °C (±1 °C) in open glass vials, under continuous light and stirring. The observed growth rates were compared to the growth rates of unexposed (control) algae. Validity of the toxicity test was confirmed by comparing the growth rate (1.5–1.7 d^{-1}) and variation coefficients (0.3–2.4%) in the controls with the criteria mentioned in the ISO 8692 guideline (growth rate >1.4 d^{-1}; variation coefficient <5%). Before the start of the toxicity tests, pH, ammonium concentration and conductivity were measured to check for possible confounding factors in the raw wastewater and effluent samples only. Lowest effect (EC_{10}) and fifty percent effect concentrations (EC_{50}), including confidence limits, were determined for all toxicity tests using the intrinsic growth rates after 72 h. These concentration-response relationships were calculated with the ToxCalc software for Excel.

2.4.3. Daphnia magna reproduction test

Daphnia magna chronic toxicity assessments were undertaken in accordance to the OECD 211 guideline (OECD, 2012). The primary objective of this test was to assess the effect of toxicants in the effluents on the reproductive output of daphnids. Juvenile daphnids, aged less than 24 h at the start of the test, were individually exposed for 21 d. At the end of the test, the total number of living offspring produced per parent animal alive at the end of the test was assessed. Juveniles produced by adults that died during the test were excluded from the calculations. In addition, survival and the time to production of first brood were determined. Tests were performed with ten replicates at 20 °C and a light-cycle of 16 h/8 h (light/dark). All daphnids were fed daily with algae (*Chlorella*) and observed for any abnormalities, mortality, or the presence of juveniles. The amount of food increased with the age of the organisms. Test medium were refreshed twice a week. Before the start of the toxicity tests, pH, ammonium concentration, oxygen saturation level and conductivity were measured to check for possible confounding factors (in the raw wastewater and effluent samples). Validity of the toxicity tests was confirmed by comparing mortality (0–10%) and reproductive output (101–153 juveniles) in the controls with the criteria mentioned in the OECD 211 guideline (mortality ≤20%; reproductive output ≥60 juveniles/female). Lowest effect ($L(E)C_{10}$) and fifty percent effect concentrations ($L(E)C_{50}$), including confidence limits, were determined for all toxicity tests using the survival and reproductive output after 21 days. These concentration-response relationships were calculated with the ToxCalc software for Excel.

2.4.4. Fish embryo acute toxicity (FET) test with Danio rerio

Fish toxicity tests were performed using the zebrafish (*Danio rerio*) embryo acute toxicity test (ZFET) according to the OECD 236 guideline (OECD, 2013). Measurements of pH, ammonium concentration and conductivity were done before starting the tests. Zebrafish embryos were exposed to the water samples and mortality, hatching, and abnormalities (e.g. swim bladder malformation or oedema, curved column, heart oedema) were scored daily until 5 days post fertilisation. Lowest effect ($L(E)C_{10}$) and fifty percent effect concentrations ($L(E)C_{50}$) values were determined for all toxicity tests using the survival and frequency of malformations after 5 days. Lethal and sublethal effect levels, including confidence limits, were based on concentration-response relationships calculated with GraphPad Prism software.

2.5. Biomimetic solid phase micro extraction (BE-SPME)

The BE-SPME represents an organism-free assay to measure bioavailable organic material in the effluent, which includes hydrocarbons, polar organics, and ionic organics (e.g., naphthenic acids). Assessments of potentially bioavailable hydrocarbons were undertaken in the effluent samples using BE-SPME combined with analysis by gas chromatography coupled with a flame ionization detector (GC-FID)

according to the protocol of Leslie and Leonards (2005). The BE-SPME method represents the manual adaptation from what was previously referred to as potentially bioaccumulating substances (PBS) and is considered consistent with the automated method (Letinski et al., 2014; Redman et al., 2018a, 2018b). In this approach SPME fibres (100 μm poly-dimethylsiloxane (PDMS) fibres) were exposed to 250 ml of effluents, with agitation, for 24 h in a closed glass bottle without head space. After 24 h exposure, the fibres were removed from the effluent solution, dried with a tissue and then the injection was directly done into a GC equipped with a FID. A DB1 (210 m × 0.25 mm x 0.1 μm) GC column was used. For quantification 2,3-dimethylnaphthalene was used as external standard. The temperature programme of the GC was designed to sum the peaks as much as possible by using a fast temperature ramp. The total peak area of the chromatogram was integrated (between C9 and C38) and the molar concentration was calculated by applying the average molar response factor of 2,3-dimethylnaphthalene for converting the measured GC-FID response (total integrated area) to millimolar (mM) of organic constituents on the PDMS fibre (Redman et al., 2018a).

The BE-SPME assessments were undertaken primarily on final effluents although some mid-treatment wastewaters were also assessed.

2.6. Chemical analysis

The chemical analyses included TPH, benzene, toluene, ethyl benzene, xylenes (BTEX), phenol index, polycyclic aromatic hydrocarbons (PAHs), (bio) chemical oxygen demand, total nitrogen, total suspended solids and selected metals (Cd, Hg, Ni, Pb, V) before, and after, wastewater treatment steps. In addition, detailed hydrocarbon characterization of the wastewater samples was performed according to a GCxGC protocol previously developed (Concawe, 2010). These analyses were all undertaken as part of the companion study with further information provided by Hjort et al. (2021).

3. Results and discussion

The efficiency of the XAD®-extraction was studied by analysing both BE-SPME and toxicity before and after XAD®-extraction. Both the toxicity (Table SD6) and BE-SPME data (Fig. SD1) showed comparable results before ('raw' sample) and after XAD®-extraction indicating that the extraction procedure did not lead to a significant loss of hydrocarbons. Although slightly greater toxicity was observed in some of the raw samples this is considered to be related to confounding factors as described later in this section.

The analytical assessment of the effluents, as reported by Hjort et al. (2021), demonstrated that concentrations of TPH (or Oil in Water) were reduced substantially by treatment. This was typically in the order of 90% on average across the biodegradation steps in the treatment systems and the final effluent TPH concentrations were often less than 1 mg/L. The only two exceptions were for refinery I1 and refinery M, in which TPH was higher in the effluent than in the influent to the biological treatment step. This could be the consequence of using spot samples as sampling could not be aligned with the water retention in the full water treatment train. Consequently, unless steady state conditions prevail, the influent and effluent results are not always directly comparable. For refinery I1 this may be due to the very low influent concentrations (0.06 mg-TPH/L), whereas for refinery M there is no obvious explanation for the observed TPH increase and we postulate that much of the TPH was associated with the total suspended solids (TSS) that also increased over the biological treatment step.

A similar reduction in the predicted toxicity was observed before and after the refinery biological treatment steps (Table 2). The notable exceptions being for refineries I1 and M in which the predicted toxicity decreased even though TPH appeared to increase. This supports our interpretation that the increased TPH of the effluent was associated with the increased TSS. In Table 2 the predicted toxicity results are based on PetroTox calculations for aquatic organisms with median sensitivity using the GCxGC and TPH measurements from the analytical studies. These results demonstrate that the biological treatment steps were effective at reducing both concentration and predicted toxicity. These results will be compared to the observed toxicity in the bioassays in the present study.

3.1. Toxicity (bioassay) assessments

An overview of the results expressed as LC_{50}/EC_{50} values for all of the toxicity assessments is presented in Table 3 and discussed below. Further results are provided in the supplementary data (Table SD1 (Microtox®), Table SD2 (algae), Table SD3 (*D. magna*) and Table SD4 (*D. rerio*). No toxicity was observed in the procedural blank (XAD® control) in any of the tests.

The criteria used to assess potentially confounding factors of nitrite, ammonium and conductivity in the raw effluents were based on those derived by Postma et al. (2002). These data, and where exceedance potentially occurred in three final raw effluents and a wastewater, are presented in Table SD5.

Table 2
Effects of effluent treatment on TPH through biological treatment and predicted aquatic toxicity. (TPH = total petroleum hydrocarbons; TU = toxic units as predicted using PetroTox).

Refinery	Influent TPH (mg/L)	Effluent TPH (mg/L)	TPH Removal (%)	Influent TU [−]	Effluent TU [−]	TU Removal [%]
Refinery A	18	0.08	>99	1.98	0.059	97
Refinery B1	8.8	3	66	2.41	0.16	93
Refinery B2	31	3	90	2.2	0.28	87
Refinery C	5	0.06	99	0.54	0.056	90
Refinery D	1.5	0.22	85	2.01	0.049	98
Refinery E	1.9	0.41	78	0.048	0.005	89
Refinery F	6.6	<0.05	>99	1.81	0.034	98.
Refinery G	10	<0.05	>99	1.19	0.009	99
Refinery I2	13	<0.05	>99	1.61	0.051	97
Refinery I1	0.06	0.11	nc*	0.16	0.015	91
Refinery K	18	0.79	96	1.25	0.30	76
Refinery L	140	2.8	98	0.26	0.15	43
Refinery M	8.4	10	nc*	3.45	0.22	93
Refinery J	25	<0.05	>99	1.11	0.014	99
Refinery H	4.4	0.2	95	2.48	0.031	99

Notes: i) Influent is to the biological treatment step and effluent is the final refinery WWTP effluent; ii) Refinery WWTP effluent do not always represent final discharge since some transfer to final polishing steps and others to off site (uWWTPs); iii) nc* not calculated as TPH higher in effluent than influent which could be consequence of spot sample and/or organic suspended particles in final effluent; iv) Analytical data from Hjort et al. (2021).

Table 3
Summary of results of all the toxicity tests (all values as % volume of effluent). Raw (unextracted) samples are shaded grey.

Sample code and description	*D. magna* (21 days)		Microtox	Algal growth	Fish embryo (120 h)	
	EC_{50} Reproduction	LC_{50} Mortality	15 min EC_{50}	72 h EC_{50}	LC_{50} Mortality	EC_{50} Malformation
XAD Control	>100	>100	>45	>100	>100	>100
D(23) Final effluent	>100	>100	> 45	58.9	>100	>100
A(4) Final effluent	>100	>100	>45	>100	>100	>100
C(16) Final effluent	98.6	69.6	>45	>100	>100	>100
B1(7) Mid treatment	13.4	17.9	30	53.3	50	17
D(18) XAD Mid treatment	56.2	62.7	11.1	>100	>100	>100
D(22) XAD Mid treatment	>100	>100	35.9	>100	>100	>100
A(2) XAD Mid treatment	46.5	34.9	13.1	>100	>100	>100
A(3) XAD Mid treatment	>100	>100	>45	>100	>100	>100
D(23) XAD Final effluent	>100	>100	>45	>100	>100	>100
A(4) XAD Final effluent	>100	>100	>45	>100	>100	>100
C(16) XAD Final effluent	>100	>100	>45	>100	>100	>100
B1(7) XAD Mid treatment	>100	>100	>45	>100	>100	>100
E(26) XAD Final effluent	>100	>100	42.2	>100	-	-
F(28) XAD Final effluent	>100	>100	>45	>100	-	-
G(31) XAD Final effluent	80.1	>100	>45	>100	-	-
H(33) XAD Final effluent	>100	>100	>45	>100	-	-
I1(37) XAD Final effluent	84.2	>100	>45	>100	-	-
K(45) XAD Final effluent	>100	>100	>45	>100	-	-
I2(39) XAD Final effluent	>100	>100	>45	>100	-	-
L(49) XAD Final effluent	>100	>100	>45	>100	-	-
J(42) XAD Final effluent	>100	>100	>45	>100	-	-
M51(46) XAD Final treatment	>100	>100	38.3	>100	-	-

LC_{50}/EC_{50} = Lethal/adverse effect concentrations which are seen in 50% of the organisms in the defined exposure period

3.1.1. *Bacteria A. fischeri (Microtox®)*

Toxic effects as determined by EC_{20} luminescence inhibition values as shown in the supplementary data (SD) Table SD1 were observed in 64% of the samples (14 out of 22). However, evident toxicity identified by EC_{20} values $\leq$ 10 vol% was only found in four samples. These were all mid-treatment wastewaters and therefore not discharged to the environment without further treatment. Greater toxicity would be anticipated in these samples as contaminant concentrations were typically higher in comparison to final effluents as confirmed by Hjort et al. (2021). Since no obvious confounding factor such as pH and ammonium was identified and no toxicity was observed in the corresponding extracts, the toxicity in the raw effluents could have been caused by possible differential bioavailability of the material in the XAD® extract relative to the raw effluent, or possible constituents that did not partition to the XAD® resins. Identification of these potential toxicants would require a more comprehensive assessment (e.g. full suite of metals analysis or full toxicity identification evaluation) which was outside the scope of this project. Furthermore, this was not considered to be an important omission as the BE-SPME measurements were associated with the observed toxicity so further analysis was not considered a high priority.

The Microtox® assessments of stored samples (Table SD1) indicated that toxicity was not significantly affected by a three week storage period under 4 °C as EC_{20} and EC_{50} values did not differ significantly before and after storage. This provides reassurance that the sample integrity and associated toxicity was not compromised during storage and provides input for planning of future studies.

3.1.2. *Algae R. subcapitata*

Overall, results indicate that hydrocarbons in refinery final effluents will cause either no or only minor effects on the algal growth in open test systems. Clear growth inhibition (EC_{50} < 100 vol% Table 3) was found in one raw final effluent sample (D23) as well as in the one raw mid-treatment wastewater sample (B1). In the latter the toxicity could be partly due to relatively high ammonium concentrations (range <10–47 mg/L upper limit of which exceeds the maximum recommended value of 25 mg/L; Table SD 5).

Toxicity to algae among the XAD® extracts was only observed at low level (EC_{10} value = 81 vol%) and for a mid-treatment wastewater sample (A(2)).

3.1.3. *Daphnid D. magna*

The toxicity of XAD®-extracts (Table 3, Table SD3) was generally lower compared to respective raw samples which could potentially be explained by the removal of confounding factors such as high nitrite, ammonium or conductivity levels (Table SD 5). Toxic effects causing >50% decrease in survival (LC_{50}) or reproduction (EC_{50}) were only observed in two raw samples i.e. C(16) and B1(7) and four XAD®-extracts (D(18), A(2), G(31) and I1(37)). However, smaller but still significant effects on reproduction were observed in 89% of the XAD®-extracts (Table SD3) for which many EC_{10}-values fell below the lowest concentration tested (i.e. <10 vol%).

For some tests, confidence limits could not be calculated for all samples. This typically occurred when effects were only found in one concentration or in the case that no partial effects (i.e. effects either 0 or 100%) were observed. In these instances ideally additional tests with an extended number of concentrations should be performed but these were not undertaken due to limitations on sample volume and concerns regarding the age of samples. Such aspects might be considered and minimized in future studies for example by the application of low-volume bioassays whenever possible.

3.1.4. *Zebrafish (D. rerio)*

Acute toxicity to fish as assessed using the zebrafish embryo acute toxicity assay indicate samples did not show any mortality or significant adverse effects after 120 h exposure (Table SD 4). The only exception was sample B1(7), a raw mid-treatment wastewater sample, for which malformations were observed after 72 h exposure around the hatching

period. The toxicity in this sample was not considered to be due to hydrocarbon components as no toxicity was observed in the corresponding XAD® extract. Ammonium might have influenced the test results as the preliminary criteria set for this species (Table SD 6) were exceeded in sample B1(7).

3.1.5. Summary of toxicity (bioassay) data

Raw final effluent samples showed variable, toxic effects as observed on Microtox®, algae and daphnids, while no effects were found on fish larvae. Overall, the three final effluent raw samples did not present high levels of toxicity in any of the bioassays with a dilution factor of 1.7 required to reduce the most toxic sample (D23) to below its lowest EC_{50} (Table 3).

The criteria for confounding factors were potentially exceeded for some of the raw sample bioassays. However, on the basis of the limited observed toxicity and comparing the results for raw samples compared to XAD® extracts (Tables SD6a and SD6b) only the mid water treatment sample (B1(7)) indicated that these may be contributing to the toxicity observed in the raw sample.

Also, an alternative explanation for a decreased toxicity observed in the mid treatment sample (B1(7)) XAD® extract compared to respective raw samples could be that the extracts did not exhibit the same level of bioavailability as the raw effluents. This may have been due to the change in the physical state of the test material (e.g., dispersed in the effluents vs concentrated material in the extract). More investigations would be required to establish if this was the case.

Microtox® was the most sensitive of the toxicity tests, even though this is a rapid screening bioassay. The fish embryo toxicity test was the least sensitive. This is broadly in line with previous refinery effluent toxicity assessments where *D. magna* and Microtox® have been shown to be the most sensitive assays compared to fish and algae (Comber et al., 2015). However, care has to be taken in making assumptions about the sensitivity of the test organism as many other factors need to be considered. For example, the nature of the test system may in itself lead to loss of toxicants (and hence sensitivity) due to absorptive losses of toxicants to the surfaces of the test exposure vessels. Considering this aspect, the apparent sensitivity of the Microtox® test may be a consequence of its rapid nature and that test dilutions are prepared in glass cuvettes immediately prior to exposure. This minimises loss of potential toxic components e.g. via adsorption to test vessels. Consequently, future investigations with other species could attempt to avoid adsorption losses by e.g. using glass vessels or applying other dosing regimens (e.g. flow through, passive dosing) and ideally whenever possible confirming exposure concentrations (i.e. BE-SPME measurements in test vessels during the tests).

3.2. Influence of wastewater treatment on toxicity

The toxicity data are summarised in Fig. 1 showing the distribution of results before and after treatment for the bioassays, together with the BE-SPME measurements and predicted toxicity, expressed as TUs. These demonstrate that toxicity decreased during wastewater treatment as shown by the shifts in the plots although this shift is more obvious (due to sensitivity) for some of the bioassays. The data from this current study demonstrate apparent species differences where the *D. rerio* embryo and algae bioassays appear to be relatively insensitive whilst the *D. magna* and Microtox® bioassays appear to be the most sensitive.

There were few instances of toxicity (e.g., EC_{50}) in the bioassays in the final effluent where the few observed EC_{50} values occur at high dilutions, or are not observed even in the 100% effluent. In contrast, the

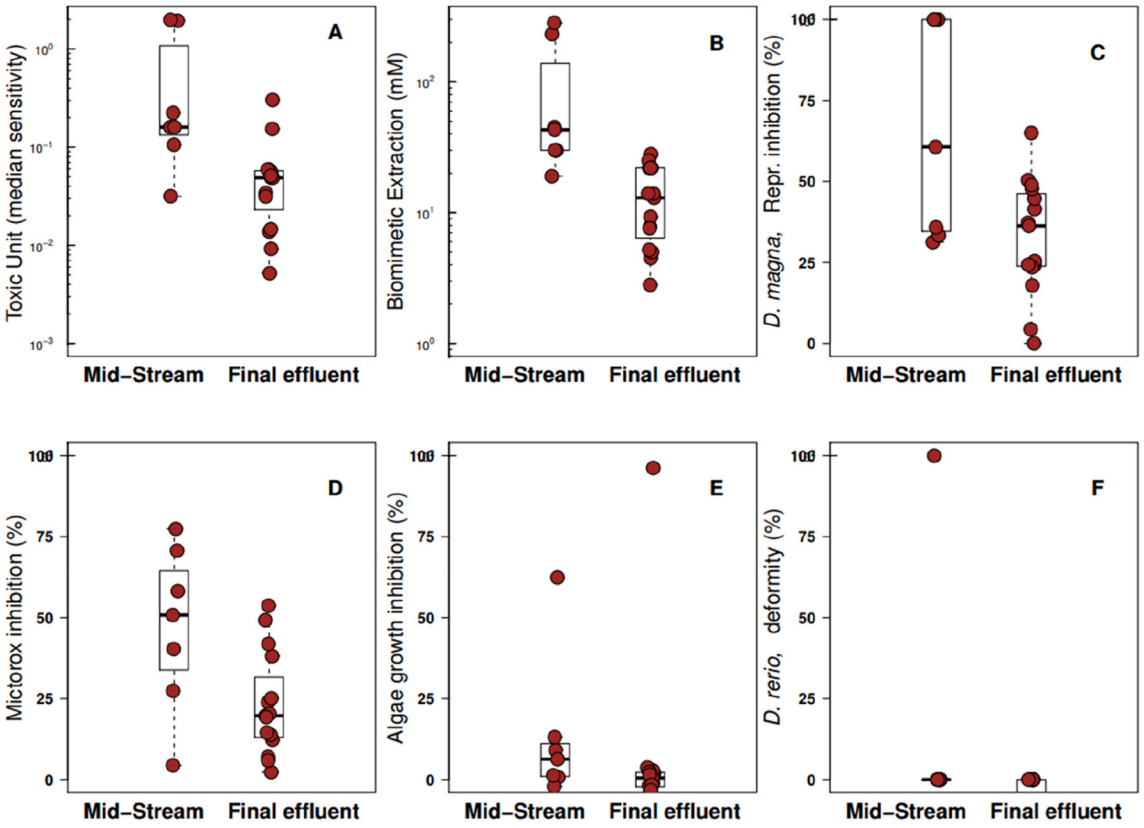


Fig. 1. Comparison of mid-stream (influent to biological treatment step) and final effluents from oil refinery WWTPs. Toxicity (toxic units) predicted using PetroTox (Panel A), BE-SPME measurements (Panel B), % effects on 21 day *D. magna* reproduction (Panel C), 15 min % light inhibition Microtox® (Panel D), 72 h % algal growth inhibition (Panel E) and 120 h % malformation observed in fish (*D. rerio*) FET test (Panel F). Actual data are presented as circles, and are shown along with summary box plots (median, inner quartile, and range).

mid-stream effluent samples often exhibit toxicity with multiple observed EC_{50}s. This empirical decline in toxicity was consistent with the decline in the theoretical predicted TUs and measured BE-SPME measurements (Fig. 1, Panel A and B, respectively). On average BE-SPME and TUs declined by approximately a factor of 2–3 after final treatment.

This factor is generally consistent between the different toxicity metrics including predicted TU, measured BE-SPME, and the observed toxicity in the 100% effluents for the different bioassays. There is notable variability in the various parameters, spanning up to an order of magnitude in all cases, including before and after treatment. This reflects the complex nature of a petrochemical refinery which handles different types of crude oils, and manufactures different products through different processes. However, it is clear that effluent toxicity is reduced by treatment in a WWTP regardless of the assessment technique used to measure it, and that aquatic toxicity of the final effluent is negligible.

3.3. BE-SPME as measurement of bioavailability

This section builds on the observed association between toxicity determined in the bioassays (section 3.1) and BE-SPME by performing direct comparison to the observed toxicity as illustrated in Fig. 2 and Table SD7. In the present work, BE-SPME measurements were collected only on 100% effluent samples. The dilutions of the effluent clearly result in reduced exposure, and toxicity, however the relationship between toxicity and exposure concentrations for petroleum substances is often not linear (Redman et al., 2014b; Redman, 2015 and Redman et al., 2017a, 2017b). Therefore, linear extrapolation of the BE-SPME measurements to the lower loadings could result in errors or misleading trends. For this reason, direct comparison of the BE-SPME on the 100% effluent to the EC_{50}s is likely not very informative. Despite this limitation, the 100% effluent represents the worst case exposure. Therefore, the comparisons were limited to effect data at the 100% dilution, which were compared further to other existing BE-SPME-toxicity data (Redman et al., 2018a, b).

There appears to be a reasonable association between the chronic 21 day *D. magna* survival data from the present study and the 2-d LC_{50} values from literature data (Redman et al., 2018a) with a critical effect concentration (LC_{50}) near 10 mM. The chronic data found in literature were collected at lower exposures and show generally low effects and the addition of the chronic aquatic toxicity data for the present study show higher effects at higher BE-SPME measurements. These data appear consistent with each other and appear to form a reasonable dose response with a chronic EC_{50} around 10 mM. This implies an acute to chronic ratio (ACR) around 1, which is consistent with relatively low ACRs (median ~5) for these types of organic chemicals (McGrath et al., 2018).

The Microtox® data in the present study (BE-SPME LC_{50} near 20–60 mM) appear to be offset from the literature study (LC_{50} 150 mM). However, the BE-SPME LC_{50} is consistent with the range of other critical BE-SPME values of 15–90 mM (Redman et al., 2018a,b) so it is considered to represent typical variability observed between laboratories. It is often necessary to understand variability so that sufficiently conservative and meaningful monitoring thresholds can be determined.

For the toxicity to algae and zebrafish embryos as summarised in Fig. 2 plots D to F there is an apparent discrepancy at the high loading when the results of this study are compared to those in the literature.

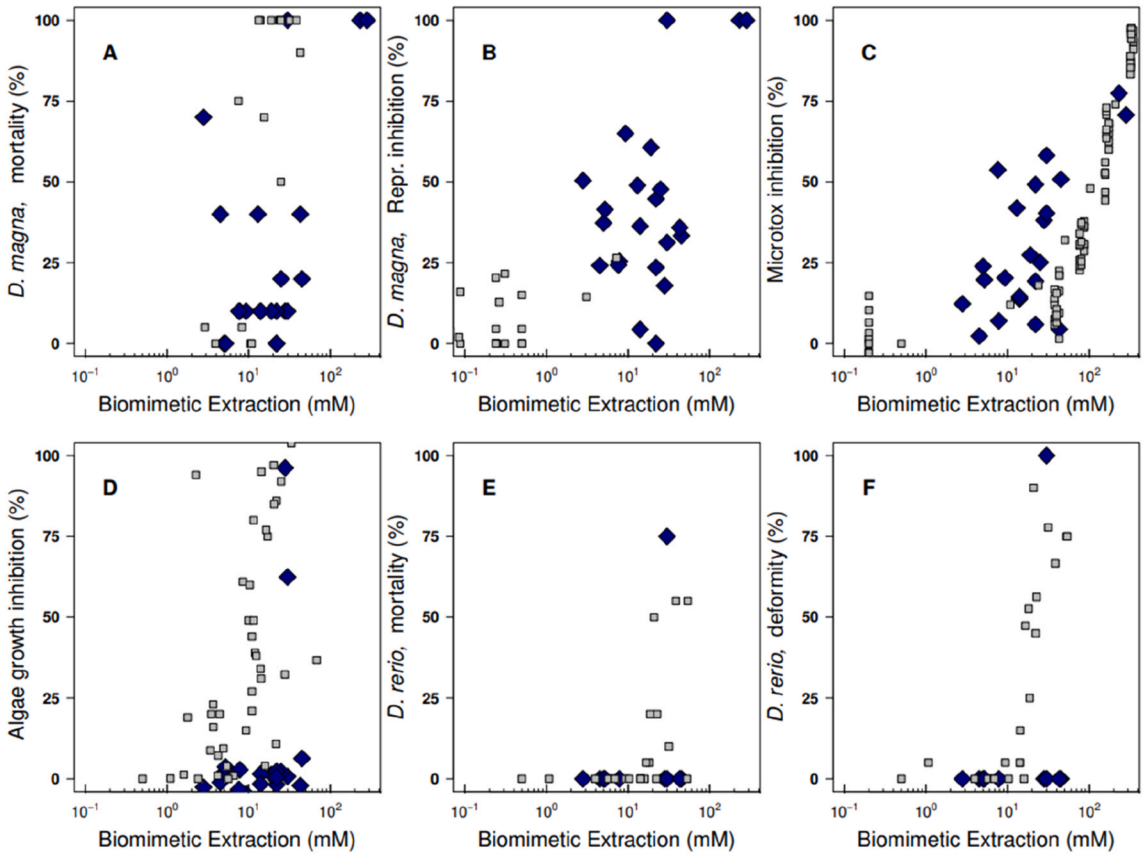


Fig. 2. Direct comparison of toxicity data from bioassays to BE-SPME measurements *D. magna* 2-d mortality (literature data) or 21-d mortality (present study) (Panel A), *D. magna* 21-d reproduction (Panel B), and Microtox (Panel C) inhibition, 72 h algal growth inhibition (Panel D), 120 h FET (*D. rerio*) mortality and deformity (Panels E and F, respectively). Grey square points represent literature data (Redman et al., 2018a,b) and blue diamonds are data from the present study. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

One possible explanation is that this is exposure related since the BE and toxicity data were taken from separate sources and it is possible that in some studies, the test material did not fully equilibrate with the test system. It is clear that these same samples produced toxicity in the Microtox® and daphnid tests. For example, Microtox® is a rapid bioassay (15 min) compared to the other bioassays which take several days. The Daphnid testing employed static renewals of the test substance in larger test vessels than the fish embryo tests. The algal test did not employ renewals.

Despite this discrepancy the observed relationships between toxicity and BE-SPME are consistent with other literature data (Redman et al., 2018a, 2018b). This confirms that toxicity in these tests were mainly from the bioavailable portion of the organic fraction in the raw effluents or extracts. This observation further highlights the utility of this non-animal test method for evaluating toxicity of refinery effluents. Coupled with the existence of BE-based critical toxicity thresholds for non-specific toxicity, BE-SPME has the potential to greatly simplify toxicity assessments of mixtures, and effluents.

4. Summary and conclusions

The study shows that the bioassays used displayed different levels of sensitivity but it is suspected that this may be an artefact of actual exposures achieved within the test systems rather than species sensitivity. Using XAD® resin extraction helped to remove uncertainty associated with potentially confounding factors that can be present in oil refinery effluents and thereby offered a viable approach for assessing toxicity of organic fractions of oil refinery effluents. It is therefore recommended that for future WEA undertaken on refinery effluents/wastewaters studies should 1) be designed to assess if exposure is a limiting factor and if so ensure bioassay procedures are modified to mitigate this problem and 2) that bioassay studies are conducted on both raw and extracted samples (e.g. XAD® resin) to improve understanding of the factors governing their toxicity.

The companion analytical study by Hjort et al. (2021) demonstrated that the refineries used in this study were typically shown to operate according to BAT expectations with removal factors of TPH for their whole WWTP varying from 97% to >99.8% between refineries. The observed toxicity data and other evaluation metrics (e.g., toxic units, BE-SPME) from the current study also show substantial reductions due to the refinery WWTP.

The results generated in these combined studies are in line with previous investigations providing evidence that oil refinery wastewater treatment steps remove both contaminants and toxicity. This was confirmed in all the assessments where predicted toxicity, BE-SPME and actual bioassay data all demonstrated a significant drop after treatment (i.e. when final effluents from the biological WWTPs were compared to mid-treatment wastewater samples). Furthermore, all methods indicated that the toxicity of final treated effluents can be considered to be negligible.

Observed toxicity was correlated with BE-SPME in the present study. In general the apparent BE-based critical toxicity thresholds were similar to the thresholds found in open literature (Redman et al., 2018a, 2018b).

BE-SPME measurements provide a suitable basis for evaluating aquatic toxicity of oil refinery effluents and are recommended to be used in WEA evaluations to support mechanistic interpretation of potential observed effects. Use of microbial toxicity screening tests such as Microtox® can be used to support the evaluations due to the simplicity of implementation and the rapid response. This has also been seen in investigations of toxicity of produced waters from oil and gas platforms (Worden et al., 2021). BE-SPME, combined where appropriate with simple bacterial bioassays, is therefore recommended as a suitable tool for screening water bodies and effluents to support effect-based monitoring and assessment of effluent discharge impact on the ecological status of surface water bodies under the WFD.

By utilising a range of biological effect methods, the results of this study provide additional reassurance that oil refinery effluents treated in accordance to EU BAT requirements are not toxic to aquatic organisms and considered unlikely to have significant impacts on the surface water bodies into which they are discharged.

CRediT author statement

G. Whale: Conceptualization, Investigation, Writing - Original Draft, Writing - Review & Editing; M. Hjort: Writing - Review & Editing, Project administration, Supervision; C. Di Paolo: Writing - Review & Editing; A.D. Redman: Conceptualization, Investigation, Visualization, Data Curation; J.F. Postma: Formal analysis, Data Curation; J. Legradi: Formal analysis; P.E.G. Leonards: Formal analysis, Investigation, Data Curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors declare no conflicts of interests. The costs of undertaking and reporting this research were funded by Concawe. We want to acknowledge the participating refineries and all the people who provided information on the wastewater treatment process and the collection of the samples, as the study would not have been possible without their cooperation. We acknowledge in particular Mike Spence (Concawe) for setting up the refinery contacts, preparation of the questionnaire used in sample collection, and discussions on the projects. We would like to thank Eleni Vaiopoulou (Concawe) for her critical review and comments as well as the members of Concawe STF32 and Concawe STF34 for the stimulating discussions, support and information. We also like to acknowledge Jacco Koekkoek, Martin van Velzen, Gerda Hopman, and Peter Cenijn of VU Amsterdam who performed the chemical and zebrafish analysis. Rineke Keijzers from Ecofide is acknowledged for the toxicity testing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2021.132146.

References

Brack, W., Aissa, S.A., Backhaus, T., Dulio, V., Escher, B.I., Faust, M., Hilscherova, K., Hollender, J., Hollert, H., Müller, C., Munthe, J., Posthuma, L., Seiler, T.-B., Slobodnik, J., Teodorovic, I., Tindall, A.J., de Aragão Umbuzeiro, G., Zhang, X., Altenburger, R., 2019. Effect-based methods are key. The European Collaborative project SOLUTIONS recommends integrating effect-based methods for diagnosis and monitoring of water quality. Environ. Sci. Eur. 31 (1) https://doi.org/10.1186/s12302-019-0192-2, 10 (6 pp.).

Cailleaud, K., Bassères, A., Gelber, C., Postma, J.F., ter Schure, A.T., Leonards, P.E., Redman, A.D., Whale, G.F., Spence, M.J., Hjort, M., 2019. Investigating predictive tools for refinery effluent hazard assessment using stream mesocosms. Environmen toxicol Chem 38, 650–659.

Carere, M., et al., 2019. Proposal for Effect-Based Monitoring and Assessment in the Water Framework Directive. Report to the WG Chemicals on the Outcome of the Work Performed in the Subgroup on Effect-Based Methods (EBMs). Mandate 2016-2018. Approved by the Strategic Coordination Group of the WFD in Feb 2019.

Concawe, 1999. Best available techniques to reduce emissions from refineries. Report No. 99/01. Brussels: Concawe.

Chapman, P.M., 2000. Whole effluent toxicity testing—usefulness, level of protection, and risk assessment. *Environmen toxicol Chem*: Int. J. 19, 3–13.

Comber, M.H.I., Girling, A., den Haan, K.H., Whale, G., 2015. Oil refinery experience with the assessment of refinery effluents and receiving waters using biologically based methods. Integrated Environ. Assess. Manag. 11 (4), 653–665.

Concawe, 1979. The Environmental Impact of Refinery Effluents. Report No. 5/79. Brussels: CONCAWE.

Concawe, 1982. Ecological Monitoring of Aqueous Effluents from Petroleum Refineries. Report No. 8/82. Brussels: CONCAWE.
Concawe, 2010. Concawe effluent speciation project. Report 3/10 Brussels, CONCAWE. https://www.concawe.eu/wp-content/uploads/2017/01/report-no.-3_10.pdf.
Concawe, 2012. Assessment of refinery effluents and receiving waters using biologically based effect methods. Report No. 1/12. Brussels: CONCAWE. https://www.concawe.eu/wp-content/uploads/2017/01/report-no.-1_12.pdf.
Concawe, 2018a. A review of toxicity testing conducted on European refinery effluents in 2010 and 2013. Report No. 2/18. Brussels: CONCAWE. https://www.concawe.eu/wp-content/uploads/2018/02/Rpt_18-2.pdf.
Concawe, 2018b. Review of options for refinery effluent assessment using effect-based tools in combination with passive samplers. Report no. 14/18 Brussels: CONCAWE. https://www.concawe.eu/wp-content/uploads/Rpt_18-14.pdf.
Concawe, 2020a. 2016 Survey of Effluent Quality and Water Use at European Refineries. Report No. 10/20. Brussels: CONCAWE. https://www.concawe.eu/wp-content/uploads/Rpt_20-10.pdf.
Concawe, 2020b. PetroTox Version 4.0 Report No. 10/20. Brussels: CONCAWE User Manual. https://www.concawe.eu/wp-content/uploads/Rpt_20-07.pdf.
Concawe, 2020c. Aquatic Toxicity of Petroleum Substances : Extending the Validation of the Biomimetic Extraction (BE) Method for Use in Hazard Assessments. Report No. 3/20. Brussels: CONCAWE. https://www.concawe.eu/wp-content/uploads/Rpt_20-3.pdf.
de Vlaming, V., Connor, V., DiGiorgio, C., Bailey, H.C., Deanovic, L.A., Hinton, D.E., 2000. Application of whole effluent toxicity test procedures to ambient water quality assessment. Environmen toxicol Chem: Int. J. 19, 42–62.
Di Toro, D.M., McGrath, J.A., Hansen, D.J., 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. I. Water and tissue. Environ. Toxicol. Chem. 19 (8), 1951–1970.
EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy. OJ L 327, 22.12.2000.
EU, 2010. Directive 2010/75/EU of the European Parliament and of the Council of 24 November 2010 on Industrial Emissions (Integrated Pollution Prevention and Control). OJ 2010, L 334: 17 – 119.
EU, 2014. Commission Implementing Decision (2014/738/EU) of 9 October 2014 Establishing Best Available Techniques (BAT) Conclusions, under Directive 2010/75/EU of the European Parliament and of the Council on Industrial Emissions, for the Refining of Mineral Oil and Gas. OJ 2014, L 307: 38 – 82. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014D0738&from=EN.
EU-JRC, 2015. Best Available Techniques (BAT) Reference Document for the Refining of Mineral Oil and Gas. EU-JRC 2015. http://eippcb.jrc.ec.europa.eu/reference/BREF/REF_BREF_2015.pdf.
Hamers, T., Legradi, J., Zwart, N., Smedes, F., de Weert, J., van den Brandhof, E.-J., van de Meent, D., de Zwart, D., 2018. Time-Integrative Passive sampling combined with TOxicity Profiling (TIPTOP): an effect-based strategy for cost-effective chemical water quality assessment. Environ. Toxicol. Pharmacol. 64, 48–59.
Hansen, B.H., Parkerton, T., Nordtug, T., Størseth, T.R., Redman, A., 2019. Modeling the toxicity of dissolved crude oil exposures to characterize the sensitivity of cod (Gadus morhua) larvae and role of individual and unresolved hydrocarbons. Mar. Pollut. Bull. 138, 286–294.
Hedgpeth, B., Redman, A., Alyea, R., Letinski, D., Connelly, M., Butler, J., Zhou, H., Lampi, M., 2019. Analysis of sublethal toxicity in developing zebrafish embryos exposed to a range of petroleum substances. Environmen toxicol Chem 38 (6), 1302–1312.
Hjort, M., den Haan, K.H., Whale, G.F., Koekkoek, J., Leonards, P.E.G., Redman, A.D., Vaiopoulou, E., 2021. Conventional and high resolution chemical characterization to assess refinery effluent treatment performance. Chemosphere 278 (2021), 130383. https://doi.org/10.1016/j.chemosphere.2021.130383.
ISO, 2007. International Standards Organisation Water Quality — Determination of the Inhibitory Effect of Water Samples on the Light Emission of Vibrio Fischeri (Luminescent Bacteria Test) — Part 3: Method Using Freeze-Dried Bacteria. ISO 11348-3:2007.
ISO, 2012. International Standards Organisation Water quality — fresh water algal growth inhibition test with unicellular green algae. ISO 8692:2012. https://www.iso.org/obp/ui/#iso:std:iso:8692:ed-3:v1:en.
Leonards, P.E., Postma, J.F., Comber, M., Whale, G., Stalter, G., 2011. Impact of biodegradation on the potential bioaccumulation and toxicity of refinery effluents. Environ. Toxicol. Chem. 30, 2175–2183.
Leslie, H.A., Leonards, P.E.G., 2005. Protocol, Determination of potentially bioaccumulatable substances (PBS) in whole effluents using biomimetic solidphase microextraction (SPME). In: OSPAR Intercessional Expert Group (IEG) on Whole Effluent Assessment (WEA), Interlaboratory Study. RIVO, IJmuiden, the Netherlands.
Letinski, D., Parkerton, T., Redman, A., Manning, R., Bragin, G., Febbo, E., Palandro, D., Nedwed, T., 2014. Use of passive samplers for improving oil toxicity and spill effects assessment. Mar. Pollut. Bull. 86, 274–282.
McGrath, J.A., Di Toro, D.M., 2009. Validation of the target lipid model for toxicity assessment of residual petroleum constituents: monocyclic and polycyclic aromatic hydrocarbons. Environ. Toxicol. Chem. 28, 1130–1148.
McGrath, J.A., Parkerton, T.F., Hellweger, F.L., Di Toro, D.M., 2005. Validation of the narcosis target lipid model for petroleum products: gasoline as a case study. Environ. Toxicol. Chem. 24, 2382–2394.
McGrath, J.A., Fanelli, C.J., Di Toro, D.M., Parkerton, T.F., Redman, A.D., Paumen, M.L., Comber, M., Eadsforth, C.V., den Haan, K., 2018. Re-evaluation of target lipid model–derived HC5 predictions for hydrocarbons. Environ. Toxicol. Chem. 37, 1579–1593.
OECD, 2012. Guidelines for the Testing of Chemicals, Section 2 Effects on Biotic Systems Test No. 211: Daphnia Magna Reproduction Test. Adopted: 2nd October 2012.
OECD, 2013. Guidelines for the Testing of Chemicals, Section 2 Effects on Biotic Systems Test No. 236: Fish Embryo Acute Toxicity (FET) Test. Adopted 26th July 2013.
Parkerton, T.F., Stone, M.A., Letinski, D.J., 2000. Assessing the aquatic toxicity of complex hydrocarbon mixtures using solid phase microextraction. Toxicol. Lett. (Shannon) 112–113, 273–282.
Postma, J.F., De Valk, S., Dubbeldam, M., Maas, J.L., Tonkes, M., Schipper, C.A., Kater, B.J., 2002. Confounding factors in bioassays with freshwater and marine organisms. Ecotoxicol. Environ. Saf. 53 (2), 226–237.
Redman, A., 2015. Role of entrained droplet oil on the bioavailability of petroleum substances in aqueous exposures. Mar. Pollut. Bull. 97, 342–348.
Redman, A.D., Parkerton, T.F., McGrath, J.A., Di Toro, D.M., 2012. PETROTOX: an aquatic toxicity model for petroleum substances. Environ. Toxicol. Chem. 31, 2498–2506.
Redman, A.D., Parkerton, T.F., Comber, M.H., Paumen, M.L., Eadsforth, C.V., Dmytrasz, B., King, D., Warren, C.S., den Haan, K., Djemel, N., 2014a. PETRORISK: a risk assessment framework for petroleum substances. Integrated Environ. Assess. Manag. 10, 437–448.
Redman, A.D., Parkerton, T.F., Letinski, D.J., Manning, R.G., Adams, J.E., Hodson, P.V., 2014b. Evaluating toxicity of heavy fuel oil fractions using complementary modeling and biomimetic extraction methods. Environ. Toxicol. Chem. 33, 2094–2104.
Redman, A.D., Parkerton, T.F., 2015. Guidance for improving comparability and relevance of oil toxicity tests. Mar. Pollut. Bull. 98, 156–170.
Redman, A.D., Parkerton, T.F., Leon Paumen, M., Butler, J.D., Letinski, D.J., den Haan, K., 2017a. A re-evaluation of PETROTOX for predicting acute and chronic toxicity of petroleum substances. Environ. Toxicol. Chem. 36, 2245–2252.
Redman, A.D., Butler, J.D., Letinski, D.J., Parkerton, T.F., 2017b. Investigating the role of dissolved and droplet oil in aquatic toxicity using dispersed and passive dosing systems. Environ. Toxicol. Chem. 36, 1020–1028.
Redman, A., Butler, J., Letinski, D., Di Toro, D., Paumen, M.L., Parkerton, T., 2018a. Technical basis for using passive sampling as a biomimetic extraction procedure to assess bioavailability and predict toxicity of petroleum substances. Chemosphere 199, 585–594.
Redman, A.D., Parkerton, T.F., Butler, J.D., Letinski, D.J., Frank, R.A., Hewitt, L.M., Bartlett, A.J., Gillis, P.L., Marentette, J.R., Parrott, J.L., 2018b. Application of the target lipid model and passive samplers to characterize the toxicity of bioavailable organics in oil sands process-affected water. Environ. Sci. Technol. 52, 8039–8049.
Smolders, R., Bervoets, L., Wepener, V., Blust, R., 2003. A conceptual framework for using mussels as biomonitors in whole effluent toxicity. Human and Ecological Risk Assessment 9, 741–760.
Struijs, J., van de Kamp, R.E., 2001. A revised procedure to concentrate organic micro-pollutants in water. In: Bilthoven : National Institute for Public Health and the Environment RIVM. Report No. 607501 001/2001.
Teodorović, I., Bečelić, M., Planojević, I., Ivančev-Tumbas, I., Dalmacija, B., 2009. The relationship between whole effluent toxicity (WET) and chemical-based effluent quality assessment in Vojvodina (Serbia). Environ. Monit. Assess. 158, 381.
Worden, J., Di Paolo, C., Whale, G., Eadsforth, C.V.E., Michie, E., Lindgren, A., Smit, M. G.D., 2021. A Toolbox for screening environmental hazard and risk of effluent discharges from facilities within the oil and gas industry. Integrated Environ. Assess. Manag. 2021, 1–12. https://doi.org/10.1002/ieam.4400.