Title

Interlaboratory comparison of in vitro bioassays for screening of endocrine active chemicals in recycled water

Authors

Mehinto, Alvine C
Jia, Ai
Snyder, Shane A
et al.

Publication Date

2015-10-01

DOI

10.1016/j.watres.2015.06.050

Peer reviewed

Water Research 83 (2015) 303–309

Contents lists available at ScienceDirect

Water Research

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

Interlaboratory comparison of *in vitro* bioassays for screening of endocrine active chemicals in recycled water

Alvine C. Mehinto [a, *], Ai Jia [b], Shane A. Snyder [b], B. Sumith Jayasinghe [c], Nancy D. Denslow [c], Jordan Crago [d], Daniel Schlenk [d], Christopher Menzie [e], Sandy D. Westerheide [e], Frederic D.L. Leusch [f], Keith A. Maruya [a]

[a] Southern California Coastal Water Research Project Authority, Costa Mesa, CA, 92626, United States
[b] Department of Chemical and Environmental Engineering, University of Arizona, Tuscon, AZ, 85721, United States
[c] Center for Environmental and Human Toxicology, Department of Physiological Sciences, University of Florida, Gainesville, FL, 32611, United States
[d] Department of Environmental Sciences, University of California Riverside, Riverside, CA, 92521, United States
[e] Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, Tampa, FL, 33620, United States
[f] Smart Water Research Centre, School of Environment, Griffith University, Southport, QLD, 4222, Australia

ARTICLE INFO

Article history:
Received 24 February 2015
Received in revised form
20 June 2015
Accepted 29 June 2015
Available online 6 July 2015

Keywords:
Bioanalytical screening
Standardization
Recycled water
Water quality
Endocrine disrupting chemicals

ABSTRACT

In vitro bioassays have shown promise as water quality monitoring tools. In this study, four commercially available *in vitro* bioassays (GeneBLAzer® androgen receptor (AR), estrogen receptor-alpha (ER), glucocorticoid receptor (GR) and progesterone receptor (PR) assays) were adapted to screen for endocrine active chemicals in samples from two recycled water plants. The standardized protocols were used in an interlaboratory comparison exercise to evaluate the reproducibility of *in vitro* bioassay results. Key performance criteria were successfully achieved, including low background response, standardized calibration parameters and high intra-laboratory precision. Only two datasets were excluded due to poor calibration performance. Good interlaboratory reproducibility was observed for GR bioassay, with 16–26% variability among the laboratories. ER and PR bioactivity was measured near the bioassay limit of detection and showed more variability (21–54%), although interlaboratory agreement remained comparable to that of conventional analytical methods. AR bioassay showed no activity for any of the samples analyzed. Our results indicate that ER, GR and PR, were capable of screening for different water quality, *i.e.*, the highest bioactivity was observed in the plant influent, which also contained the highest concentrations of endocrine active chemicals measured by LC–MS/MS. After advanced treatment (*e.g.*, reverse osmosis), bioactivity and target chemical concentrations were both below limits of detection. Comparison of bioassay and chemical equivalent concentrations revealed that targeted chemicals accounted for ≤5% of bioassay activity, suggesting that detection limits by LC–MS/MS for some chemicals were insufficient and/or other bioactive compounds were present in these samples. Our study demonstrated that *in vitro* bioassays responses were reproducible, and can provide information to complement conventional analytical methods for a more comprehensive water quality assessment.

1. Introduction

The scarcity of potable water is a growing issue worldwide, particularly for urban centers located in arid regions. To fulfill water supply needs, impacted entities are pursuing policies to increase the supply and use of recycled water (SWRCB, 2013). Herein we define recycled water as highly treated municipal wastewater that is available for indirect potable reuse (NRMMC, 2009), and ultimately in the future for direct potable reuse (WRRF, 2011). Because treated wastewater effluents that serve as source water for recycled water facilities typically contain chemical residues (Ternes et al., 2004), purification is needed to attenuate these contaminants. Prior to widespread public acceptance of potable reuse, recycled water utilities are faced with the challenge of demonstrating that chemicals present in product water are not harmful to environmental and human health (WRRF, 2011; SWRCB, 2013).

* Corresponding author.
E-mail address: alvinam@sccwrp.org (A.C. Mehinto).

http://dx.doi.org/10.1016/j.watres.2015.06.050
0043-1354/

Bioanalytical techniques such as *in vitro* bioassays have been shown to be a suitable screening alternative for water quality applications. Many of these assays are used to quantify chemical bioactivity based on mode of action (MOA), *e.g.*, as part of the U.S. Environmental Protection Agency (EPA) ToxCast™ and Endocrine Disruptor Screening Programs (Dix et al., 2007; Reif et al., 2010). Over the last decade, a number of studies have applied *in vitro* bioassays to ascertain the endocrine activity of surface water and wastewater (van der Linden et al., 2008; Leusch et al., 2010; Jarosova et al., 2014). Moreover, bioassays that target molecular initiating events (*e.g.* gene transactivation) can be linked to higher order adverse outcomes via toxicity pathway analyses (Piersma et al., 2013; Sonneveld et al., 2006), providing additional biological relevance for these tools in the screening mode. Whereas much of the groundwork has been laid for endocrine disrupting endpoints, researchers are currently broadening the scope of bioanalytical tools to include other relevant MOAs, *e.g.*, genotoxicity, immunotoxicity and oxidative stress (Escher et al., 2014; Leusch et al., 2014; van der Linden et al., 2014).

A recent study by Escher et al. (2014) evaluating 103 different *in vitro* bioassays to screen for chemicals in wastewater, recycled water and drinking water, concluded that some of the bioassay endpoints were capable of discriminating among samples of different quality. Other studies have reached a similar conclusion, showing that a few bioassays are capable of relatively good measurement precision (within laboratories) while demonstrating the ability to differentiate among water qualities (Leusch et al., 2010; Jarosova et al., 2014). However, these studies have employed bioassays that are not widely available, and as a result, lack standardization across multiple laboratories. To successfully transfer and implement this technology for water quality monitoring, it is critical to demonstrate that commercially available bioassays can be standardized, and that measurements using these standardized assays agree well across multiple laboratories (Andersen et al., 1999).

The goals of this study were to 1) develop standardized protocols using commercially available bioassays for screening water samples from recycled water treatment process units; 2) evaluate the reproducibility of the bioassay responses among participating laboratories; and 3) assess the ability of the bioassays to screen for water quality by comparing bioassay responses with targeted chemical occurrence. To accomplish our goals, four "off-the-shelf" *in vitro* transactivation bioassays representing different pathways of the endocrine system were adapted to screen water samples from two water recycling plants. Replicate aliquots of water extracts were analyzed by five research laboratories using the standardized protocols, with bioassay responses translated into bioanalytical equivalent concentrations and compared with conventional analytical measurements.

2. Materials and methods

2.1. Materials

GeneBLAzer® androgen receptor (AR), estrogen receptor-alpha (ER), glucocorticoid receptor (GR) and progesterone receptor (PR) cell assay kits and media components were purchased from Life Technologies (Carlsbad, CA). Bioassay kits contained division arrested cells stably transfected with the beta-lactamase reporter gene, a LiveBLAzer FRET B/G loading kit, and CCF4-AM substrate. Black wall, clear-bottom 96-well plates were purchased from Corning (Corning, NY).

Chemicals known or suspected to activate AR (testosterone, trenbolone), ER (17α-ethinylestradiol, 17α-estradiol, 17β-estradiol, bisphenol A, estriol, estrone), GR (dexamethasone, hydrocortisone, prednisolone, and triamcinolone), and PR (levonorgestrel, norethisterone, norgestrel) were purchased at the highest purity available from Sigma–Aldrich (St. Louis, MO). The AR active chemical methyltrienolone (R1881) was purchased from Perkin–Elmer. Isotopically labeled 17α-ethinylestradiol-$^{13}C_2$, bisphenol A-$^{13}C_{12}$, estriol-$^{13}C_3$, estrone-$^{13}C_6$ were purchased from Cambridge Isotope Laboratories (Andover, MA), dexamethasone-d_4, norethindrone-d_6 and norgestrel-d_6 from C/D/N Isotopes (Pointe-Claire, Canada), and 17β-estradiol-$13C_3$ and hydrocortisone-d_2 from Sigma–Aldrich (St. Louis, MO). The purity of all the isotope standards was ≥98%.

Molecular grade dimethyl sulfoxide (DMSO, 99.5% purity) was obtained from Sigma–Aldrich (St. Louis, MO). HPLC grade methanol, acetone, acetonitrile, ethyl acetate, hexane and formic acid were purchased from Fisher Scientific (Pittsburgh, PA).

2.2. Sample collection and processing

Grab samples of water (4 L each) were collected in June 2013 from various treatment processes of a fully operational (Plant 1) and a pilot water recycling plant (Plant 2) located in southwestern U.S (Table 1). Upon collection, samples containing chlorine were immediately quenched with sodium thiosulfate (50 mg/L). All water samples and a field blank sample consisting of milli-Q water were treated with sodium azide (1 g/L) to inhibit microbial activity (Vanderford et al., 2011). Samples were stored in methanol-rinsed amber glass bottles at 4 °C and extracted within one week of collection.

Solid phase extraction (SPE) was performed according to the methods described by Macova et al. (2011). Briefly, 1 L samples were filtered and passed through two pre-conditioned cartridges: a 6 cc Oasis HLB cartridge (Waters, Milford, MA) stacked on top of a Supelclean coconut charcoal cartridge (Sigma Aldrich). After rinsing with milli-Q water and vacuum drying for 2 h, the cartridges were eluted separately with 10 mL methanol and 10 mL acetone:hexane (1:1, v/v). For each 4-L water sample, all eluates were combined and evaporated under a gentle stream of nitrogen. Each water extract was then reconstituted in 4 mL methanol and aliquots of 1.5 mL were stored at −20 °C for chemical analyses. The remaining extract was solvent exchanged to 2.5 mL DMSO and kept in amber glass vials at −20 °C. Aliquot samples of 500 μL were shipped in plastic Eppendorf tubes on ice overnight to the participating laboratories for blind bioassay analysis, where they were transferred into amber glass vials and stored at −20 °C.

Table 1
Samples collected from two U.S. water recycling treatment plants. Plant 1 was a fully operational facility; Plant 2 was operating as a pilot plant at the time of collection.

Sample no.	Description
1	Field blank (milli-Q water)
2	Plant 1 – influent (final secondary effluent from WWTP#1)
3	Plant 1 – influent subject to ozonation (Oz)
4	Plant 1 – product water subject to microfiltration (MF)
5	Plant 1 – product water subject to reverse osmosis (RO)
6	Plant 1 – product water subject to ultraviolet (UV)
7	Plant 2 – influent (final secondary effluent from WWTP#2)
8	Plant 2 – influent subject to UV
9	Plant 2 – influent subject to UV/hydrogen peroxide (H_2O_2)
10	Plant 2 – influent subject to Oz
11	Plant 2 – influent subject to Oz/UV
12	Plant 2 – influent subject to chlorination

WWTP – wastewater treatment plant.

2.3. Standardization of bioassay protocols

To standardize *in vitro* bioassays protocols to screen water samples, the following parameters were optimized to: 1) identify appropriate reference chemicals; 2) determine cell plating density for a simplified 96-well plate format; and 3) determine suitable DMSO content per well (see Supplementary Information SI-1). Optimized bioassay parameters were used to establish standard operating procedures, including sample dilution protocols and standardized plate layout that all participating laboratories would follow.

Standardized bioassay protocols are described in detail in Supplementary Information SI-1. Briefly, standardized protocols were developed for a 96-well plate format using GeneBLAzer® AR, ER, GR and PR division arrested cells. Cell density was set at 40,000 to 50,000 cells per well and solvent vehicle content was set at 0.5% DMSO per well. For each water extract, four dilutions were prepared in assay-specific media and evaluated for cytotoxic effects using PrestoBlue (Life Technologies). The diluted water extracts were then analyzed for AR, ER, GR and PR bioactivity. The REF (relative enrichment factor), defined as the product of SPE concentration factor and bioassay dilution factor, was 5 × for the initial dilution. Raw bioassay fluorescence data, expressed as blue (460 nm) to green (530 nm) absorbance ratio, were obtained using a bottom-reading fluorescence plate reader. Bioassay responses were calibrated using a 9-point assay-specific standard curve, and the calibrated response was verified on successive plates by running five concentrations of the reference chemical. Assay media (cell free), cells and media, and solvent vehicle controls were analyzed on each plate to correct for bioassay artifacts. All samples including water extracts, reference chemicals and controls were run in triplicate. Quality assurance/quality control (QA/QC) criteria were established to assess the quality of bioassay results obtained (Tables SI-2 and SI-3).

2.4. Targeted chemical analyses

Sample aliquots preserved for chemical analysis were further processed to reduce matrix effects using the extraction method of Chang et al. (2009). Target analytes (Table SI-4) were analyzed using an Agilent 1290 UHPLC system combined with 6460 triple quadrupole mass spectrometer (LC-MS/MS) in positive and negative electrospray ionization (ESI) mode. ER active chemicals were analyzed using an Agilent Zorbax Eclipse Plus C8 Rapid Resolution HD column (50 mm × 2.1 mm, 1.8 μm), while AR, GR and PR active chemicals were analyzed using a C18 Rapid Resolution HD column (100 mm × 2.1 mm, 1.8 μm). Both columns were maintained at 30 °C at a flow rate of 0.4 mL/min. Gas temperature was set at 300 °C, gas flow was maintained at 11 L/min and the nebulizer was set at 45 psi. An injection volume of 5 μL was used for analysis of all samples. For ER active chemicals, water (A) and methanol (B) were used as mobile phases with the following gradient: 10% B held for 2.5 min, increase to 40% B in 0.5 min, linear increase to 70% B in 6 min, then to 100% B in 0.1 min and held for 1.0 min. For AR, GR and PR active chemicals, 0.1% (v/v) formic acid in water (A) and acetonitrile (B) were used with the following gradient: 5% B held for 1.5 min, linear increase to 20% B in 1.5 min, to 45%B in 1 min, to 65% B in 3 min, up to 100% in 1 min and held at 100% B for 1 min.

Detection parameters used to monitor the thirteen target analytes are provided in Table SI-4. Instrumental limits of detection (LODs), method detection limits (MDLs) and sample specific method reporting limits (MRLs) were determined according to Anumol and Snyder (2015) (Tables SI-5 and SI-6). For quality control purposes, at least one laboratory blank and one laboratory blank sample fortified with the target analytes were analyzed for every 10 samples.

2.5. Data analyses

Raw bioassay fluorescence data were background corrected by subtracting the average response of the media-only control wells. Datasets that did not meet the established QA/QC criteria were excluded from further analyses. The 9-point calibration curve was used to determine the bioassay specific LOD, calculated as the minimum calibration response plus two standard deviations of the mean of that response. Background-corrected fluorescence data were converted to bioanalytical equivalent concentrations (BEQ, ng/L) using the equation:

$$\text{BEQ} = \text{EC}_{50} \text{ of reference chemical} / \text{EC}_{50} \text{ of water sample}$$

where EC_{50} is the 50% effect concentration of the reference chemical or water sample, and EC of water sample is the product of

Table 2
Interlaboratory comparison of calibration performance for GeneBLAzer® estrogen receptor-alpha (ER), glucocorticoid receptor (GR) and progesterone receptor (PR) bioassays. Results deemed not acceptable were excluded. The calibration performance of the GeneBLAzer® androgen receptor (AR) bioassay was not evaluated as no bioactivity was observed in any of the samples analyzed.

	Lab A	Lab B	Lab C	Lab D	Lab E
ER (referenced to 17β-estradiol)					
Hill slope	1.6	1.6	1.4	1.5	1.0
R^2	0.99	0.99	0.99	0.97	0.99
Log EC_{50} (M)	−9.8	−9.9	−9.6	−10.2	−9.6
LOD	11	7	10	4	7
Data accepted	yes	yes	yes	no	yes
GR (referenced to dexamethasone)					
Hill slope	2.3	1.9	2.5	2.4	2.2
R^2	0.99	0.99	0.99	0.98	1.00
Log EC_{50} (M)	−8.6	−8.5	−8.4	−8.7	−8.5
LOD	5	4	2	9	4
Data accepted	yes	yes	yes	yes	yes
PR (referenced to levonorgestrel)					
Hill slope	2.1	1.8	1.4	1.5	1.3
R^2	0.99	1.00	0.96	0.98	0.99
Log EC_{50} (M)	−9.1	−9.9	−9.5	−9.9	−9.5
LOD	3	2	3	2	8
Data accepted	no	yes	yes	yes	yes

LOD – limit of detection, expressed as a percent effect concentration (EC) relative to the referenced chemical.

average response of the water extract and the REF. The BEQ for a given sample was calculated and included for intercomparison if: 1) extract dilutions exhibited a concentration-dependent response above the LOD; and 2) the bioassay response fell within the linearized portion of the calibration curve (estimated between EC_{10} and EC_{50} of the reference chemical).

3. Results

3.1. Evaluation of individual laboratory performance

All five laboratories successfully employed the standardized bioassay protocols to conduct a blind analysis of water samples. Most datasets (12 out of 15) generated for ER, GR and PR bioassays were deemed satisfactory. The AR assay was not responsive to the water samples. There was no evidence of cytotoxicity except for samples 11 and 12 that showed mild cytotoxic effects but at levels that did not exceed the acceptance threshold (<20% cell mortality, Table SI-2). All bioassay responses reported for assay media only and DMSO-exposed cells were in agreement with the QA/QC criteria shown in Table SI-2. The majority of assay-specific calibration curves also met acceptance thresholds (Table 2). Only two laboratories, lab A and lab D, reported log EC_{50} values below the expected range for PR and ER assays, respectively. These two datasets were excluded from the calculation of sample BEQs. Variability in triplicate bioassay responses within a given laboratory was sufficiently low (relative standard deviation (RSD) between 7 and 18%) except for one laboratory that exceeded the 20% RSD threshold for samples analyzed using the PR bioassay.

3.2. Interlaboratory agreement

Bioactivity was consistently detected in samples 2, 3 and 4 from Plant 1 (Fig. SI-2). Measured GR-BEQ (expressed as ng dexamethasone/L) was highly reproducible, with a coefficient of variation less than 30% among the five laboratories (Table 3). However the measured ER bioactivity (expressed as ng 17β-estradiol/L) showed a higher degree of interlaboratory variation (greater than 30%), especially for samples 3 and 4 where the measured response was close to the bioassay LOD (Table 3). Due to unacceptable calibration performance from one laboratory (Table 2) and poor intra-laboratory precision from another (above 20% RSD), PR-BEQ (expressed as ng levonorgestrel/L) was only reported in samples 2 and 3 by three out of the five participating laboratories. The reproducibility of PR bioassay responses among these laboratories was intermediate compared to that achieved for GR and ER (21 and 36% RSD, Table 3). Bioassay responses were uniformly below detection levels for all samples collected from Plant 2 (samples 7–12). None of the participating laboratories detected AR activity above LOD in the water samples. Thus, AR bioassay results were not used for this interlaboratory comparison exercise.

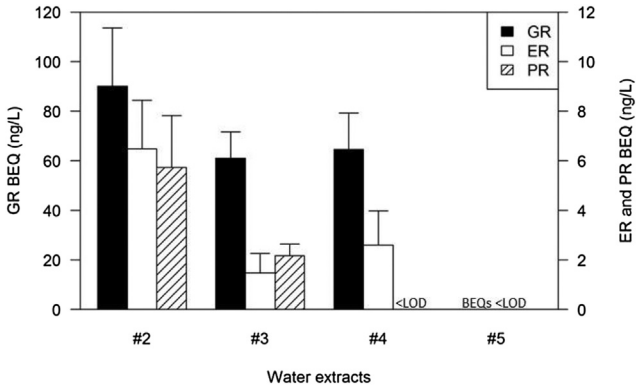

Fig. 1. Mean bioassay equivalent concentrations (BEQ) for ER (expressed as ng 17β-estradiol/L ± SD), GR (expressed as ng dexamethasone/L ± SD) and PR (expressed as ng levonorgestrel/L ± SD) for samples from an operational recycled water facility (Plant 1). Samples 2, 3, 4 and 5 correspond to Plant 1 influent, OZ – ozonation, MF – microfiltration and RO – reverse osmosis, respectively. The median limit of detection (LOD) was <1.7 ng/L for ER-BEQ, <52 ng/L for GR-BEQ, and <1.4 ng/L for PR-BEQ.

3.3. Bioscreening of water quality

Patterns of bioassay responses were in agreement with the quality of the water samples assessed by LC-MS/MS. The highest level of ER, GR and PR bioactivity was detected in the Plant 1 influent extract (sample 2), which contained the highest concentrations of target analytes and was thus considered to have the lowest water quality (Fig. 1, Table SI-6). The mean GR-, ER- and PR-BEQ for sample 2 were estimated at 90 ng dexamethasone/L, 6.5 ng 17β-estradiol/L and 5.7 ng levonorgestrel/L, respectively. Ozonation (sample 3) led to a reduction of ER, GR and PR bioactivity. Following microfiltration (sample 4), the mean PR-BEQ was below detection limits. However, ER and GR bioactivities were detected in this sample at concentrations equivalent to 2.6 ng 17β-estradiol/L and 65 ng dexamethasone/L, respectively. At the highest level of treatment (samples 5 and 6), bioassay responses were no longer detected. None of the laboratories detected a bioassay response in the field blank (sample 1).

Table 3
Interlaboratory agreement for analysis of recycled water samples exhibiting activity using GeneBLAzer® bioassays. Results are expressed as bioanalytical equivalent concentrations (BEQ, ng/L).

Sample no.	ER[a]			GR[b]			PR[c]		
	N	Mean BEQ	% RSD	N	Mean BEQ	% RSD	N	Mean BEQ	% RSD
2	4	6.5	30	5	90	26	3	5.7	36
3	4	1.5	54	5	61	16	3	2.2	21
4	4	2.6	53	5	65	23	3	<1.4	N/A
5	4	<1.7	N/A	5	<52	N/A	3	<1.4	N/A
6	4	<1.7	N/A	5	<52	N/A	3	<1.4	N/A

N – number of laboratories meeting QA/QC criteria.
RSD – relative standard deviation.
N/A – not applicable.
Below detection is represented by "<median limit of detection" (*e.g.* "<1.7").
[a] ER-BEQ expressed as ng 17β-estradiol/L.
[b] GR-BEQ expressed as ng dexamethasone/L.
[c] PR-BEQ expressed as ng levonorgestrel/L.

Table 4
Comparison of bioanalytical and chemical equivalent concentrations (BEQs and CEQs, respectively), both expressed in ng/L.

Sample no.	CEQs for estrogens[a]						Sum CEQ	Percent BEQ explained[c]	CEQs for glucocorticoids[b]				Sum CEQ	Percent BEQ explained[c]
	αE2	βE2	E1	EE2	E3	BPA			DEX	HYD	PRED	TRIM		
2	<0.51	<1.3	0.13	<12.9	<0.058	0.00046	0.13	2	<0.18	0.24	0.09	<0.04	0.52	0.4
3	<0.47	<1.2	0.05	<7.0	<0.073	0.00003	0.05	4	<0.17	0.22	0.07	<0.04	0.49	0.5
4	<0.49	<1.2	0.13	<8.6	<0.08	0.00006	0.13	5	<0.16	0.2	0.06	<0.04	0.46	0.4
REP[d]	0.011	1	0.027	5.03	0.017	0.000016			1.00	0.26	0.21	0.15		

[a] αE2 = 17α-estradiol, βE2 = 17β-estradiol, E1 = estrone; EE2 = 17α-ethinylestradiol, E3 = estriol, BPA = bisphenol A.
[b] DEX = dexamethasone; HYD = hydrocortisone, PRED = prednisolone, TRI = triamcinolone.
[c] (sum CEQ/mean BEQ)* 100; mean BEQs are from Table 3.
[d] REP (unpublished data) – relative effect potency factor, calculated as EC_{50} of the reference compound/EC_{50} of the test compound.

3.4. Comparison of in vitro bioassay and targeted chemical analyses

Concentrations of ER and GR active chemicals frequently detected in wastewater as determined by LC–MS/MS are shown in Table SI-6. Many of the chemicals detected had low relative effect potency (REP, Table 4). Chemical equivalent concentrations (CEQs), defined as the sum of the product of agonist concentrations determined by LC-MS/MS and their REP, were calculated for each water sample. Targeted chemistry data explained less than 10% of the ER bioassay responses and less than 1% of the GR bioassay responses (Table 4). PR bioassay responses could not be explained, as the two known PR active chemicals analyzed in this study were below analytical detection limits.

4. Discussion

Standardization is a critical step in establishing robust methodologies that allow multiple entities to produce comparable results. In the present study, performance-based QA/QC criteria (Table SI-2) were established to control for poor method performance due to (1) cytotoxicity; (2) interferences from media and solvents ("background response"); and (3) variability in calibration response. All participating laboratories achieved results within the acceptance thresholds for cytotoxicity and background response, indicating that the likelihood of matrix interference was small. During optimization of our protocols, we observed first-hand the negative impact that both parameters can exert on the bioassay response if not properly controlled. Cytotoxicity is particularly important to implement for cell bioassays to confirm that non-responses are an accurate representation of non-bioactivity, and are not caused by sample exposure issues with cell viability.

Our standardized protocols and QA/QC criteria are consistent with those established for ER transactivation bioassays employed in the Endocrine Disruptor Screening Program (EDSP) (USEPA, 2012). For example, the 9-point calibration curve over four logs of concentration of reference chemical in our protocols compares favorably to those employed in the EDSP, which vary between single point (*i.e.*, one concentration) to 16-point calibration curves stretching over 10 orders of magnitude. In the present study, the laboratories were able to routinely operate within the established QA/QC criteria for calibration. Verification of calibration response performed to assess plate-to-plate and day-to-day variation revealed that calibration was adequately stable. There were only 2 out of 15 incidences where a laboratory failed to meet the QA/QC criteria for the calibration curve (Table 2), confirming that bioassay methodologies have matured in reliability and emphasizing the importance of routinely monitoring QA/QC parameters to ensure reliable bioassay results (Leusch et al., 2010; Macova et al., 2011). In both cases, the estimated EC_{50} for the reference chemical and/or the Hill slope fell outside of the acceptance range by a slight margin. Further refinement of the protocols to confirm accurate preparation of reference chemical stock solutions and dilutions could improve calibration performance.

Bioassay response patterns for ER and GR were similar across laboratories, with samples 2–4 showing clear bioactivity and the remaining samples showing little evidence of activity above LODs (Fig. SI-2). The response patterns for AR and PR were more variable as bioassay responses were often close to or below the LOD (Table 3). Interlaboratory agreement was highest in estimating mean GR-BEQs (16–26% RSD; Table 3). This is to be expected since the GR responses for samples 2–4 were the highest measured. Mean ER-BEQs for four laboratories (data from the fifth laboratory did not meet QA/QC criteria) agreed to within a factor of 2 (54% RSD), and mean PR-BEQs estimated for three laboratories agreed to within 36%. The level of interlaboratory agreement achieved was similar in both magnitude and range (10%–~50%) to the variability noted in previous studies. van der Linden et al. (2008) reported an intra-laboratory variability for triplicate analysis of WWTP effluent samples ranging from 10% for ER to 35% RSD for GR and PR assays. In the same study, samples' variability increased to approximately 40% for surface and drinking water samples, which is lower but consistent with our results for highly treated samples. In another study that measured estrogenicity of single chemicals in human breast cancer (MCF-7) cells, the interlaboratory agreement ranged from 10 to 50% (Andersen et al., 1999).

Our results also indicate that the level of interlaboratory reproducibility attained using the standardized bioassays is approaching that achievable using conventional analytical chemistry. Interlaboratory agreement for targeted chemical analyses using environmental samples can range from single digit to more than 50% RSD, depending on the analyte, matrix and measured concentration (Vanderford et al., 2014). It should be noted that the present study addresses only the reproducibility of bioassay responses since a single laboratory collected and processed all the samples. Other factors including extraction method, collection, handling and preservation can have a significant impact on reproducibility especially when dealing with exceedingly low concentrations of chemicals. In our study, aliquots of sample extracts in DMSO were shipped in plastic Eppendorf tubes. The timing of subsequent transfer into glass vials varied among laboratories. This may have resulted in a variable loss of extractable residues, thus affecting the interlaboratory agreement in bioactivity measurements.

The *in vitro* bioassays evaluated were capable of qualitatively ranking samples from Plant 1 by water quality (Fig. 1). For example, the reduction in mean bioactivity measured for Plant 1 influent after microfiltration (sample 4) is in agreement with the reduction in bioassay response observed by Escher et al. (2014) for recycled water subjected to ultrafiltration. Bioactivity of all three endpoints was below LODs in water subjected to RO and UV (samples 5 and 6), consistent with bioassay screening results published earlier (Escher et al., 2014; Leusch et al., 2014). In contrast, AR revealed no pattern

among the 12 samples analyzed. It should be noted that anti-AR activity has often been observed in the environment and this endpoint may be a more relevant MOA for bioscreening of chemicals (Escher et al., 2014; USEPA, 2014).

The contribution of two known ER agonists detected by LC–MS/MS – estrone and bisphenol A – accounted for up to 5% of the measured BEQs (Table 4). The discrepancy between GR-BEQ and CEQ estimated for hydrocortisone and prednisolone was even greater, accounting for less than 1% of measured GR-BEQs (Table 4). Previous studies have shown good agreement between bioanalytical equivalent concentrations and the levels of natural and synthetic hormones present in water samples (Leusch et al., 2010; Tang et al., 2014). In the present study, relating the measured biological activity to individual chemicals was made difficult for two reasons: few chemicals were monitored by targeted chemical analyses (four GR agonists and two PR agonists, Table SI-4), and there were relatively high analytical method reporting limits (MRLs) for some of the potent estrogens. Previous studies have shown that a variety of synthetic glucocorticoids, including but not limited to those selected here, can contribute to overall GR activity in water samples (Schriks et al., 2010; Suzuki et al., 2014), and the same is likely to hold true for progestins and PR activity. For ER activity, the relatively high MRLs for 17β-estradiol and 17α-ethinylestradiol (0.4–1.3 and 0.5–2.4 ng/L, respectively) may have impacted the CEQ to BEQ comparison, as either of these compounds present just below their respective MRLs could account for the majority of the biological activity detected *in vitro*. Furthermore, assessing the mass balance between BEQ and CEQ relies on the accurate estimation of REPs of known agonists. Although the REP utilized in our study were within the same order of magnitude as those published previously (Leusch et al., 2014; Schriks et al., 2010), additional standardization of REP values is needed to establish consensus across cell lines for the same endpoint.

5. Conclusions

Standardization of bioassay protocols and QA/QC guidelines resulted in good interlaboratory agreement when applying commercially available *in vitro* transactivation bioassays for screening of recycled water samples. Bioassays that showed reproducible responses among the participating laboratories, and thus potential to assess water quality include ER, GR and PR in agonist mode. Androgenic activity was not observed; however, antagonist mode AR would represent an additional reasonable endpoint for screening of endocrine active chemicals in wastewater-derived sources.

The suite of estrogenic hormones and pervasive synthetic chemicals typically targeted in current monitoring efforts could not be quantitatively related to the bioactivity observed in this study. It is likely that inclusion of additional compounds and improvements in the detection of trace chemicals by targeted chemistry methods would improve the correlation between CEQ and BEQ. However, questions remain as to the identity of other bioactive substances in these samples. A strategy that incorporates bioanalytical, targeted and non-targeted chemical screening would constitute a more comprehensive and effects-directed monitoring effort (Snyder, 2014). Should screening using *in vitro* bioassays and targeted chemical analysis warrant further investigation, non-targeted chemical analysis could be implemented to identify the other chemicals contributing to the bioactivity.

Acknowledgments

Funding for this project was provided by State Water Resources Control Board Agreement No. 10-096-250. We acknowledge Shimin Wu (University of Arizona) for helping with the chemical analyses, Jens Scheideler (Xylem Corporation) for providing the Advanced Oxidation Processes (AOP) pilot system, and Gerry Pelanek, Kun Bi and John Printen (Life Technologies) for their technical assistance. We also acknowledge Beate Escher (UFZ) for helpful discussion and advice. The authors declare no conflict of interest, and reference to trade names does not imply endorsement.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.watres.2015.06.050.

References

Anumol, T., Snyder, S.A., 2015. Rapid analysis of trace organic compounds in water by automated online solid-phase extraction coupled to liquid chrmoatography-tandem mass spectrometry. Talanta 132, 77–86.

Andersen, H.R., Andersson, A.M., Arnold, S.F., Autrup, H., Barfoed, M., Beresford, N.A., Bjerregaard, P., Christiansen, L.B., Gissel, B., Hummel, R., Jørgensen, E.B., Korsgaard, B., Le Guevel, R., Leffers, H., McLachlan, J., Møller, A., Nielsen, J.B., Olea, N., Oles-Karasko, A., Pakdel, F., Pedersen, K.L., Perez, P., Skakkebœk, N.E., Sonnenschein, C., Soto, A.M., Sumpter, J.P., Thorpe, S.M., Grandjean, P., 1999. Comparison of short-term estrogenicity tests for identification of hormone-disrupting chemicals. Environ. Health. Perspect. 107 (Suppl. 1), 89–108.

Chang, H., Wan, Y., Hu, J., 2009. Determination and source apportionment of five classes of steroid hormones in urban rivers. Environ. Sci. Technol. 43 (20), 7691–7698.

Dix, D.J., Houck, K.A., Martin, M.T., Richard, M.A., Setzer, R.W., Kavlock, R.J., 2007. The ToxCast program for prioritizing toxicity testing of environmental chemicals. Toxicol. Sci. 95 (1), 5–12.

Escher, B.I., Allinson, M., Altenburger, R., Bain, P.A., Balaguer, P., Busch, W., Crago, J., Denslow, N.D., Dopp, E., Hilscherova, K., Humpage, A.R., Kumar, A., Grimaldi, M., Jayasinghe, B.S., Jarosova, B., Jia, A., Makarov, S., Maruya, K.A., Medvedev, A., Mehinto, A.C., Mendez, J.E., Poulsen, A., Prochazka, E., Richard, J., Schifferli, A., Schlenk, D., Scholz, S., Shiraishi, F., Snyder, S., Su, G., Tang, J.Y., Burg, B.V., van der Linden, S.C., Werner, I., Westerheide, S.D., Wong, C.K., Yang, M., Yeung, B.H., Zhang, X., Leusch, F.D., 2014. Benchmarking organic micropollutants in wastewater, recycled water and drinking water with *in vitro* bioassays. Environ. Sci. Technol. 48 (3), 1940–1956.

Jarosova, B., Ersekova, A., Hilscherova, K., Loos, R., Gawlik, B.M., Giesy, J.P., Blaha, L., 2014. Europe-wide survey of estrogenicity in wastewater treatment plant effluents: the need for effect-based monitoring. Environ. Sci. Pollut. Res. 21 (18), 10970–10982.

Leusch, F.D.L., de Jager, C., Levi, Y., Lim, R., Puijker, L., Sacher, F., Tremblay, L.A., Wilson, V.S., Chapman, H.F., 2010. Comparison of five in vitro bioassays to measure estrogenic activity in environmental waters. Environ. Sci. Technol. 44 (10), 3853–3860.

Leusch, F.D.L., Khan, S.J., Laingam, S., Prochazka, E., Froscio, S., Trinh, T., Chapman, H.F., Humpage, A., 2014. Assessment of the application of bioanalytical tools as surrogate measure of chemical contaminants in recycled water. Water Res. 49, 300–315.

Macova, M., Toze, S., Hodgers, L., Mueller, J.F., Bartkow, M., Escher, B.I., 2011. Bioanalytical tools for the evaluation of organic micropollutants during sewage treatment, water recycling and drinking water generation. Water Res. 45 (14), 4238–4247.

National Resource Management Ministerial Council, Environment Protection and Heritage Council, National Health and Medical Research Council (NRMMC/EPHC/NHMRC), 2009. Australian Guidelines for Water Recycling: Managed Aquifer Recharge. National Water Quality Management Strategy Document No. 24, p. 267.

Piersma, A.H., Bosgra, S., van Duursen, M.B., Hermsen, S.A., Jonker, L.R., Kroese, E.D., van der Linden, S.C., Man, H., Roelofs, M.J., Schulpen, S.H., Schwarz, M., Uibel, F., van Vugt-Lussenburg, B.M., Westerhout, J., Wolterbeek, A.P., van der Burg, B., 2013. Evaluation of an alternative in vitro test battery for detecting reproductive toxicants. Reprod. Toxicol. 38, 53–64.

Reif, D.M., Martin, M.T., Tan, S.W., Houck, K.A., Judson, R.S., Richard, A.M., Knudsen, T.B., Dix, D.J., Kavlock, R.J., 2010. Endocrine profiling and prioritization of environmental chemicals using ToxCast data. Environ. Health Perspect. 118 (12), 1714–1720.

Schriks, M., van Leerdam, J.A., van der Linden, S.C., van der Burg, B., van Wezel, A.P., de Voogt, P., 2010. High-resolution mass spectrometric identification and quantification of glucocorticoid compounds in various wastewaters in the Netherlands. Environ. Sci. Technol. 44 (12), 4766–4774.

Snyder, S.A., 2014. Emerging chemical contaminants: looking for greater harmony. J. Am. Water Works Assoc. 106 (8), 38–52.

Sonneveld, E., Riteco, J.A., Jansen, H.J., Pieterse, B., Brouwer, A., Schoonen, W.G., van der Burg, B., 2006. Comparison of in vitro and in vivo screening models for androgenic and estrogenic activities. Toxicol. Sci. 89 (1), 173–187.

State Water Resources Control Board (SWRCB), 2013. Recycled Water Policy, Resolution 2013-003. Sacramento, CA, p. 34.
Suzuki, G., Sata, K., Isobe, T., Brouwer, A., Nakayama, K., 2014. Glucocorticoid (GR) receptor agonist profiling in the effluents from sewage treatment plants in Japan. In: Poster Presented at the 2014 SETAC North America Meeting, Vancouver, BC, Canada, Nov 9-13, 2014.
Tang, J.Y.M., Busetti, F., Charrois, J.W.A., Escher, B.I., 2014. Which chemicals drive the biological effects in wastewater and recycled water? Water Res. 60, 289–299.
Ternes, T.A., Joss, A., Siegrist, H., 2004. Scrutinizing pharmaceuticals and personal care products in wastewater treatment. Environ. Sci. Technol. 38 (20), 392A–399A.
U.S. Environmental Protection Agency (USEPA) Endocrine Disruptor Screening Program, 2012. Prioritization of the Endocrine Disruptor Screening Program Universe of Chemicals for an Estrogen Receptor Adverse Outcome Pathway using Computational Toxicology Tools. OCSPP/OW/ORD/NCCT, p. 142.
USEPA Endocrine Disruptor Screening Program, 2014. Integrated Bioactivity and Exposure Ranking: a Computational Approach for the Prioritization and Screening of Chemicals in the Endocrine Disruptor Screening Program. FIFRA SAP Meeting. Docket ID EPA–HQ–OPP–2014–0614.
van der Linden, S.C., Heringa, M.B., Man, H.Y., Sonneveld, E., Puijker, L.M., Brouwer, A., van der Burg, B., 2008. Detection of multiple hormonal activities in wastewater effluents, surface water, using a panel of steroid receptor CALUX bioassays. Environ. Sci. Technol. 42 (15), 5814–5820.
van der Linden, S.C., von Bergh, A.R., van Vught-Lussenburg, B.M., Jonker, L.R., Teunis, M., Krul, C.A., van der Burg, B., 2014. Development of a panel of high-throughput reporter-gene assays to detect genotoxicity and oxidative stress. Mutat. Res. Genet. Toxicol. Environ. Mutagen. 760, 23–32.
Vanderford, B.J., Mawhinney, D.B., Trenholm, R.A., Zeigler-Holady, J.C., Snyder, S.A., 2011. Assessment of sample preservation techniques for pharmaceuticals, personal care products, and steroids in surface and drinking water. Anal. Bioanal. Chem. 399 (6), 2227–2234.
Vanderford, B.J., Drewes, J.E., Eaton, A., Guo, Y.C., Haghani, A., Hoppe-Jones, C., Schluesener, M.P., Snyder, S.A., Ternes, T., Wood, C.J., 2014. Results of an inter-laboratory comparison of analytical methods for contaminants of emerging concern in water. Anal. Chem. 86 (1), 774–782.
WateReuse Research Foundation (WRRF), 2011. Direct Potable Reuse: a Path Forward. WRRF, Alexandria, VA, p. 114.