Available online at www.sciencedirect.com

SciVerse ScienceDirect

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

Review

QSAR-like models: A potential tool for the selection of PhACs and EDCs for monitoring purposes in drinking water treatment systems – A review

Luis F. Delgado [a], *Philippe Charles* [b], *Karl Glucina* [b], *Catherine Morlay* [a,*]

[a] *Université Lyon 1, INSA-Lyon, MATEIS CNRS UMR 5510, Campus LyonTech La Doua, 7 av. Jean Capelle, F-69621 Villeurbanne, France*
[b] *CIRSEE, Suez-Environnement, 38 rue du Président Wilson, 78230 Le Pecq, France*

ARTICLE INFO

Article history:
Received 17 April 2012
Received in revised form
10 August 2012
Accepted 12 August 2012
Available online 28 August 2012

Keywords:
Emerging contaminants
Occurrence
Drinking water
Treatment process
QSAR-like model
Monitoring purposes

ABSTRACT

Recent studies have demonstrated the presence of trace-level pharmaceutically active compounds (PhACs) and endocrine disrupting compounds (EDCs) in a number of finished drinking waters (DWs). Since there is sparse knowledge currently available on the potential effects on human health associated with the chronic exposure to trace levels of these Emerging Contaminants (ECs) through routes such as DW, it is suggested that the most appropriate criterion is a treatment criterion in order to prioritize ECs to be monitored during DW preparation. Hence, only the few ECs showing the lowest removals towards a given DW Treatment (DWT) process would serve as indicators of the overall efficiency of this process and would be relevant for DW quality monitoring. In addition, models should be developed for estimating the removal of ECs in DWT processes, thereby overcoming the practical difficulties of experimentally assessing each compound. Therefore, the present review has two objectives: (1) to provide an overview of the recent scientific surveys on the occurrence of PhACs and EDCs in finished DWs; and (2) to propose the potential of Quantitative-Structure-Activity-Relationship-(QSAR)-like models to rank ECs found in environmental waters, including parent compounds, metabolites and transformation products, in order to select the most relevant compounds to be considered as indicators for monitoring purposes in DWT systems.

Contents

* *Corresponding author*. Tel.: +33 4 72 43 72 87.
E-mail address: catherine.morlay@univ-lyon1.fr (C. Morlay).
0043-1354/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.watres.2012.08.016

1. Introduction

Pharmaceuticals and Personal Care Products (PPCPs) as well as Endocrine Disrupting Compounds (EDCs) have been identified as Emerging Contaminants (ECs) in environmental waters. All these compounds are present at low ng/L levels in environmental samples and are currently unregulated chemicals. As highly sophisticated analytical methods are required to detect and quantify ECs in environmental waters, only a very small subset of these compounds has been investigated to date (Richardson, 2009).

Municipal and hospital wastewaters are the most important sources of human Pharmaceutically Active Compounds (PhACs), with contributions also from wastewater manufactures and landfill leachates, and from the disposal of unused medicines into the environment (Nikolaou et al., 2007). PhACs are introduced to the environment not only for human health care purposes but also through veterinary use for livestock, poultry, and fish farming. Various drugs are commonly given to farm animals to prevent illness and to increase the size of the animals (Richardson, 2009). It is estimated that approximately 4,000 active molecules are allowed for use in Europe as pharmaceutical ingredients covering an extreme variety of chemical structures, polarity and water solubility range including painkillers, antibiotics, antidiabetics, β-blockers, contraceptives, lipid regulators, antidepressants, and impotence drugs (Bruchet et al., 2005). Most EDCs are synthetic organic chemicals (e.g. PPCPs, surfactants, pesticides, brominated flame retardants, phthalates, etc.) introduced to the environment anthropogenically, although some (e.g. estrone, 17β-estradiol) occur naturally (Nikolaou et al., 2007). While not all PhACs are EDCs and vice-versa, it is clear that certain compounds fit both categories (i.e., certain oral contraceptive medications, thyroid hormones administered as medications, and estrogen replacement pharmaceuticals) (Snyder et al., 2003).

The possible effects on aquatic organisms and human health, associated with the consumption of water containing low concentrations of single compounds, have been presented in toxicology studies (Escher et al., 2005; Vosges et al., 2008). The metabolites and degradation by-products are also of concern, because they may have a similar or higher toxicity than the parent compounds. Information about the toxic effects of most of these compounds on living organisms is very limited today (Nikolaou et al., 2007). Currently, two FP7 European projects, CYTHOTHREAT and PHARMAS are focusing on genotoxic and cytotoxic anti-cancer drugs and on the induction of genetic resistance by antibiotics (Tourad et al., 2011). Some PhACs, although present at the ng/L level, have been shown to possess the potential to induce adverse effects in human embryonic cells HEK293 (Pomati et al., 2006). Nevertheless, researchers do not yet understand the exact risks from decades of continuous exposure to random combinations of low levels of PhACs, EDCs, and other organic compounds (i.e. synergistic effects) and research has to continue in order to clarify the toxicological significance of these trace EDCs/PPCPs in DW (Tourad et al., 2011). Hence, the long-term effects of the consumption of water containing low concentrations of micropollutants (μPs) will remain as an unanswered question for the near future (Nikolaou et al., 2007; Richardson, 2009). Therefore, the question of the removal of these compounds through conventional Drinking Water Treatment (DWT) processes is an emerging concern and is a real issue considering DW preparation (Weyer and Riley, 2001; Snyder et al., 2003; Reungoat et al., 2010).

Recent studies have demonstrated the presence of trace-level EDCs and PPCPs in finished DWs (Stackelberg et al., 2007; Yu et al., 2007; Kuster et al., 2008; Benotti et al., 2009; Vulliet et al., 2011; Kleywegt et al., 2011; Huerta-Fontela et al., 2011; Wang et al., 2012). Currently, there is no statutory maximum contaminant level for PhACs in DW. In the European Water Framework Directive on Priority Substances (Directive, 2008/105/EC) not any PhAC has been highlighted as a potential key hazardous pollutant. The proposed revision of the European Directive includes ibuprofen, diclofenac, α-ethinyloestradiol, β-oestradiol and perfluorooctane sulfonate (EC, 2011). In its third Contaminant Candidate List (CCL3), the US EPA has identified several pharmaceuticals ingredients, including estrogenic hormones and erythromycin (EPA, 2009).

ECs removal creates a unique challenge to DW preparation since: 1) the number of compounds detected is large and keeps increasing (an updated list of ECs can be found in the last reviews of Richardson (Richardson, 2009, 2010); 2) their physicochemical properties are highly diversified (Lei and Snyder, 2007) and 3) they occur in raw waters at concentrations as low as the part per trillion (Schwarzenbach et al., 2006).

Experimental data is most of the time unavailable to allow determining DWT processes efficiency for ECs removal (Ridder et al., 2010). Moreover, the removal of only a limited number of solutes can be experimentally investigated because of strong technical and financial constraints. Hence, a prioritisation scheme should be developed in order to select a small set of ECs that warrant further study in relation to the quality of DW. Furthermore, as the number of ECs is currently important and as it keeps growing because new chemicals are continuously synthesized, models should be developed that would provide an estimation of ECs removal in DWT processes.

The present review has two objectives: (1) to provide an overview of the recent scientific surveys on the occurrence of PhACs and EDCs in finished DWs; and (2) to propose the potential of QSAR-like models to rank ECs found in environmental waters, including parent compounds, metabolites and transformation products, in order to select the most relevant compounds to be considered as indicators for monitoring purposes in DWT systems.

2. Occurrence of PhACs and EDCs in drinking water

There is relatively sparse information regarding PhACs and EDCs occurrence in DW. Examples of pharmaceutical groups identified in DW in different countries include lipid regulators, antiepileptic drug, analgesics, antibiotic, β-blockers and iodinated X-ray contrast media. Table 1 shows examples of EDCs and PPCPs identified in DW.

The occurrence of 106 ECs, including some PhACs and potential EDCs, at different stages of a U.S. DWT plant (conventional treatment: raw-water screening, addition of Powered Activated Carbon (PAC), coagulation, primary disinfection, flocculation, sedimentation, sand filtration, Granular Activated Carbon (GAC) and secondary disinfection) was documented by Stackelberg et al. (2004). At least 11 and as many as 17 of the target compounds were detected in samples of finished DW. Carbamazepine was measured in finished DW at concentrations up to 258 ng/L. Carbamazepine was also detected in every sample of finished DW of the US DWT plant considered by Stackelberg et al. (2007) (conventional treatment: raw-water screening, clarification, disinfection, sand filtration/GAC adsorption, secondary disinfection) at an average concentration of 29 ng/L. In the same study, dehydronifedipine and three widely used non-prescription drugs like cotinine, caffeine and acetaminophen were also quantified but they were present less frequently and at lower concentrations (from 0.6 to 15 ng/L). Concentrations of individual compounds in finished DW were low and most of the time lower than 0.5 μg/L. Only the detergent-metabolite compound 4-nonylphenol was detected at concentrations exceeding 1 μg/L.

Bruchet et al. (2005) investigated the occurrence of 21 ECs, antibiotics and X-ray contrast agents, in the Seine River (France) and in finished DW. The DWT chain included flocculation (with a low dosage of PAC only in case of pesticide occurrence), sand filtration followed by ozonation and final chlorine disinfection. Only four X-ray contrast agents persisted into finished DW at concentrations up to 60 ng/L: iopamidol (60 ng/L), ioxitalamic acid (12 ng/L), diatrizoate (32 ng/L) and iomeprol (11 ng/L).

Source water, finished DW, and distribution system water from 19 U.S. water utilities were analyzed for 51 ECs between 2006 and 2007 by Benotti et al. (2009). The 11 most frequently detected compounds were atenolol, atrazine, carbamazepine, estrone, gemfibrozil, meprobamate, naproxen, phenytoin, sulfamethoxazole, tris (2-chloroethyl) phosphate (TCEP), and trimethoprim. Of the 11 most frequently detected compounds in source water, naproxen, trimethoprim, and estrone were not detected in any finished waters. Median concentrations of detected PhACs and EDCs in finished DW were less than 10 ng/L, except for atrazine (49 ng/L), bisphenol-A (25 ng/L), galaxolide (31 ng/L), nonylphenol (93 ng/L), metolachlor (16 ng/L). Some of these median concentrations, however, were biased by low frequencies of detection. Target compounds were detected less frequently in finished waters as compared to source waters, and less frequently in distribution system waters as compared to source or finished waters.

Vulliet et al. (2011) assessed the levels of 51 compounds including PhACs and hormones in DWs from 8 French DWT plants. The DWT plants utilize various processes including coagulation, flocculation, primary disinfection, sand filtration or GAC adsorption and final disinfection. At least one target compound was quantified in all finished DW samples. Among the 51compounds, 25 were present at least in one of the finished DW samples. Salicylic acid was the most frequently detected compound, showing concentrations up to 19 ng/L. Carbamazepine (maximal concentration 10.7 ng/L) and atenolol (maximal concentration 2 ng/L) were also present in more than 30% of the finished DWs.

Huerta-Fontela et al. (2011) studied the occurrence of 55 compounds, including PhACs, hormones and some of their metabolites in Llobregat River (Spain) used for DW production and their removal through a DWT plant which includes prechlorination, coagulation, sand filtration, ozonation, GAC adsorption and post-chlorination. 35 out of 55 target compounds were detected in the raw water with concentrations up to 1.2 μg/L. The results showed that the DWT plant could completely remove all the target compounds detected in raw waters (i.e. 35 compounds) except 5 of them. Despite their persistence, the removal yields of these 5 PhACs were higher than 95%. Phenytoin, atenolol and hydrochlorotiazide were the 3 PhACs which were the most frequently found in finished DWs at concentrations about 10 ng/L. Sotalol and carbamazepine epoxide were found in less than a half of the samples at lower concentrations, about 2 ng/L.

Kleywegt et al. (2011) reported the results of the Ontario Ministry of the Environment (Canada) survey (carried out in 2006) on ECs which included PhACs, hormones and bisphenol A in both source waters and finished DWs. The survey collected 258 samples over a 16 month period from selected source waters and 17 finished DWs (source waters were: 8 from rivers; 7 from lakes; and 2 from groundwaters) and analyzed them for 48 ECs. The range of treatment processes used in these DWT plants included media and membrane filtration, GAC adsorption, and disinfection (chlorine, ozone, and ultraviolet irradiation). Among the 46 reportable ECs, 23, 22 and 18 were detected in source waters, finished DWs, or both, respectively, on at least one occasion. DWs produced from treatment plants using river and lake source waters accounted for more than 90% of the occurrences. The most frequently detected compounds in finished DWs (in more than 10% of all the analyzed samples) were: carbamazepine (25%), ibuprofen (15%), gemfibrozil (15%) and bisphenol A (12%). Lincomycin, sulfamethoxazole, acetaminophen, benzafibrate and trimethoprim showed lower frequencies of detection (2% or less) in the finished DW samples.

3. Quantitative-Structure-Activity-Relationship (QSAR) models – Validation principles

QSAR is a broadly used tool for developing relationships between the effects of a series of molecules (e.g. activities and properties of interest) with their structural properties. It is used in many areas of science (Cronin, 2010). For example, in order to minimize experimental testing work in drug design, the pharmaceutical industry applies QSAR models that can predict drug metabolic activity and toxicity a priori using

Table 1 – Examples of PPCPs/EDCs identified in drinking water before chlorination process (DWbC), in finished drinking water (DW), in distribution system (DS) and in tap water (TW).

Compound	Compound class	Sample type	Maximum concentration ng/L	Location	Reference
Acetaminophen (Paracetamol)	Analgesic and antipyretic	DW	17	Canada	Kleywegt et al., 2011
		DW	45	France	Vulliet et al., 2011
		DWbC	210	France	Togola and Budzinski, 2008
		DW	15.6	Spain	Boleda et al., 2011
AMDOPH	Phenazone metabolite	DW	900	Germany	Reddersen et al., 2002
		DW	550	Germany	Zühlke et al., 2007
Amitryptiline	Antidepressant, anti-anxiety	DWbC	1.4	France	Togola and Budzinski, 2008
Androstenedione	Steroid hormone	DW	2.8	France	Vulliet et al., 2011
Androsterone	Steroid hormone	DW	1	France	Vulliet et al., 2011
Atenolol	β-blocker	DW	2	France	Vulliet et al., 2011
		DW	18	USA	Benotti et al., 2009
		DS	0.84	USA	Benotti et al., 2009
		DW	23	Spain	Huerta-Fontela et al., 2011
Atrazine	Herbicide, Potential EDC	DW	870	USA	Benotti et al., 2009
		DS	930	USA	Benotti et al., 2009
		TW	28	USA	Garcia-Ac et al., 2009
Bezafibrate	Lipid regulator	DW	27	Germany	Ref in Jones et al. (2005)
		DW	2.2	France	Vulliet et al., 2011
		DW	1	Canada	Kleywegt et al., 2011
Bisphenol A	Component of plastics, potential EDC	DW	25	USA	Benotti et al., 2009
		DW	99	Canada	Kleywegt et al., 2011
		TW	423	China	Wang et al., 2012
		DW	220	USA	Stackelberg et al., 2007
		DW	0.76	USA	Chen et al., 2006
		DW	420	USA	Stackelberg et al., 2004
		TW	25	Spain	Casajuana and Lacorte, 2003
Bleomycin	Anti-neoplastic	DW	13	UK	Ref in Jones et al. (2005)
Carbamazepine	Anti-epileptic	DW	258	USA	Stackelberg et al., 2004
		DW	140	USA	Stackelberg et al., 2007
		TW	5.6	USA	Garcia-Ac et al., 2009
		DW	24	Canada	Ref in Jones et al. (2005)
		DW	32	France	Vulliet et al., 2011
		DWbC	43.2	France	Togola and Budzinski, 2008
		DW	100	Netherlands	Stolker et al., 2004
		DW	135	USA	Chen et al., 2006
		DW	18	USA	Benotti et al., 2009
		DS	10	USA	Benotti et al., 2009
		DW	601	Canada	Kleywegt et al., 2011
Carbamazepine epoxide	Metabolite of Carbamazepine	DW	2	Spain	Huerta-Fontela et al., 2011
Clofibric acid	Metabolite of lipid regulators clofibrate, etofibrate, thefibrate	DW	270	Germany	Ref in Jones et al. (2005)
		DW	5.3	Italy	Ref in Jones et al. (2005)
		DW	1.1	USA	Chen et al., 2006
		DW	100	Netherlands	Stolker et al., 2004
Codein	Opioidanalgesic	DW	30	USA	Stackelberg et al., 2007
Dehydronifedipine	Metabolite of the PhAC Nifedipine	DW	4	USA	Stackelberg et al., 2004
Diatrizoate	X-ray contrast agents	DW	32	France	Bruchet et al., 2005
		DW	1200	Germany	Ref in Mompelat et al. (2009)
		TW	100	France	Wenzel et al., 2003
Diazepam	Psychiatric drug, antianxiety	DW	10	UK	Ref in Jones et al. (2005)
		DW	23.5	Italy	Ref in Jones et al. (2005)
		DW	0.33	USA	Benotti et al., 2009
Diclofenac	Nonsteroidal anti-inflammatory	DW	6	Germany	Ref in Jones et al. (2005)
		DWbC	2.5	France	Togola and Budzinski, 2008
		DW	1	France	Vulliet et al., 2011
Enrofloxacin	Antibiotic	DW	13	Canada	Kleywegt et al., 2011
Erythromycin	Antibiotic	DW	155	Canada	Kleywegt et al., 2011
		DW	2	Spain	Boleda et al., 2011
		DW	4.9	USA	Ye et al., 2007
Fenofibric acid	Lipid regulator	DW	1	France	Vulliet et al., 2011
Flumequine	Antibiotic	DW	2.5	USA	Ye et al., 2007

(continued on next page)

Table 1 – (*continued*)

Compound	Compound class	Sample type	Maximum concentration ng/L	Location	Reference
Fluoxetine	Antidepressant	DW	0.82	USA	Benotti et al., 2009
		DS	0.64	USA	Benotti et al., 2009
Galaxolide	Fragance	DW	33	USA	Benotti et al., 2009
Gemfibrozil	Lipid regulator	DW	70	Canada	Ref in Jones et al. (2005)
		DW	4	Canada	Kleywegt et al., 2011
		DW	2.1	USA	Benotti et al., 2009
		DS	1.2	USA	Benotti et al., 2009
Hydrochlorothiazide	Diuretic drug	DW	7	Spain	Huerta-Fontela et al., 2011
Ibuprofen	Nonsteroidal	DW	3	Germany	Ref in Jones et al. (2005)
	anti-inflammatory	DWbC	0.6	France	Togola and Budzinski, 2008
		DW	1350	USA	Ref in Mompelat et al. (2009)
		DW	1.3	France	Vulliet et al., 2011
		DW	25	Canada	Kleywegt et al., 2011
Iomeprol	X-ray contrast agents	DW	11	France	Bruchet et al., 2005
		TW	12	France	Wenzel et al., 2003
Iopamidol	X-ray contrast agents	DW	60	France	Bruchet et al., 2005
		TW	180	France	Wenzel et al., 2003
Iopromide	X-ray contrast agents	DW	17.2	Spain	Boleda et al., 2011
		TW	29	France	Wenzel et al., 2003
Ioxitalamic acid	X-ray contrast agents	DW	12	France	Bruchet et al., 2005
Ketoprofen	Nonsteroidal anti-inflammatory	DW	7	France	Vulliet et al., 2011
		DWbC	3	France	Togola and Budzinski, 2008
Levonorgestrel	Hormonal contraceptives	DW	10	France	Vulliet et al., 2011
Lincomycin	Antibiotic	DW	1413	Canada	Kleywegt et al., 2011
Meprobamate	Antianxiety	DW	42	USA	Benotti et al., 2009
		DS	40	USA	Benotti et al., 2009
		DW	5.9	USA	Ref in Mompelat et al. (2009)
Metoprolol	β-blocker	DW	1	France	Vulliet et al., 2011
Monensin Na	Antibiotic (widely used in animal feeds)	DW	76	Canada	Kleywegt et al., 2011
Naproxen	Nonsteroidal anti-inflammatory	DW	0.5	France	Vulliet et al., 2011
		DWbC	0.2	France	Togola and Budzinski, 2008
Nonylphenol	Nonionic surfactant degradation	DW	100	USA	Benotti et al., 2009
	product, potential EDCs	DS	110	USA	Benotti et al., 2009
		TW	24	Spain	Casajuana and Lacorte, 2003
		DW	1100	USA	Stackelberg et al., 2007
		DW	72	USA	Chen et al., 2006
		DW	15	Germany	Kuch and Ballschmiter, 2001
Norethindrone	Hormonal contraceptives	DW	6.8	France	Vulliet et al., 2011
Oxazepam	Antianxiety	DW	2.5	France	Vulliet et al., 2011
Oxolinic acid	Antibiotic	DW	4	USA	Ye et al., 2007
Ofloxacin	Antibiotic	DW	1.6	USA	Chen et al., 2006
Phenazone	Analgesic and anti-pyretic	DW	400	Germany	Reddersen et al., 2002
		DW	250	Germany	Ref in Mompelat et al. (2009)
		DW	300	Germany	Zühlke et al., 2007
Phenytoin (Dilantin)	Anticonvulsant	DW	19	USA	Benotti et al., 2009
		DS	16	USA	Benotti et al., 2009
		DW	1.3	USA	Ref in Mompelat et al. (2009)
		DW	10	Spain	Huerta-Fontela et al., 2011
Progesterone	Steroid hormone	DW	0.93	Spain	Kuster et al., 2008
		DW	10.7	France	Vulliet et al., 2011
		DW	0.57	USA	Benotti et al., 2009
Propyphenazone	Analgesic and anti-pyretic	DW	120	Germany	Reddersen et al., 2002
		DW	80	Germany	Ref in Mompelat et al. (2009)
		DW	90	Germany	Zühlke et al., 2007
Roxithromycin	Antibiotic	DW	41	Canada	Kleywegt et al., 2011
		DW	18.1	France	Vulliet et al., 2011
		DW	1.4	USA	Ye et al., 2007
Salicylic acid	Anti-acne treatments	DW	19	France	Vulliet et al., 2011
Sotalol	β-blocker	DW	3	Spain	Huerta-Fontela et al., 2011
Stigmasterol	EDC	TW	63	France	Wenzel et al., 2003
β-Sitosterol	EDC	TW	179	France	Wenzel et al., 2003

Table 1 – (continued)

Compound	Compound class	Sample type	Maximum concentration ng/L	Location	Reference
Sulphamethoxazole	Antibiotic	DW	100	Netherlands	Stolker et al., 2004
		DW	2	Canada	Kleywegt et al., 2011
		DW	0.5	USA	Chen et al., 2006
		DW	3.4	USA	Ye et al., 2007
		DW	0.8	France	Vulliet et al., 2011
		DW	3.0	USA	Benotti et al., 2009
		DS	0.32	USA	Benotti et al., 2009
Tetracycline	Antibiotic	DW	15	Canada	Kleywegt et al., 2011
Triclosan	Antibacterial/Antimicrobial	DW	1.2	USA	Benotti et al., 2009
		DW	734	USA	Ref in Mompelat et al. (2009)
Trimethoprim	Antibiotic	DW	15	Canada	Kleywegt et al., 2011
		DW	1	France	Vulliet et al., 2011
Tylosin	Antibiotic (used in veterinary medicine)	DW	31	Canada	Kleywegt et al., 2011
		DW	1.7	Italy	Zuccato et al., 2006
		DW	4.2	USA	Ye et al., 2007

molecular descriptors (Kruhlak et al., 2007). Molecular descriptors are divided into two main classes: experimental measurements, such as log K_{ow}, pK_a, molar refractivity, dipole moment, polarizability, and, in general, physico-chemical properties; and theoretical molecular descriptors, which are derived from a symbolic representation of the molecule such as molecular surface areas, molecular interaction fields (Cronin, 2010; Sippl, 2010).

The Organization for Economic Cooperation and Development (OECD) has proposed principles for the validation of QSARs, which have been promoted widely (OECD, 2004). As suggested by Cronin (2010) the OECD principles might be of more importance for evaluating and characterizing a QSAR and hence to determine whether an individual prediction is valid. Thus, reliable QSAR model must include the following information (OECD, 2004): (1) a defined endpoint; (2) an unambiguous algorithm; (3) a defined domain of applicability; (4) appropriate measures of goodness-of-fit, robustness and predictive ability; (5) a mechanistic interpretation, if possible.

The final stage of a QSAR analysis consists of statistical validation in order to assess the significance of the model and hence its ability to predict biological activities of other (novel) compounds. In most QSAR case studies published in the literature, the leave-one-out (LOO) cross-validation procedure has been used for this purpose. The output of this procedure is the cross-validated (Q^2) which is commonly regarded as an ultimate criterion of both robustness and predictive ability of a model (Sippl, 2010).

The simplest cross-validation method is LOO, where one compound at a time is removed from the dataset, and the $N-1$ remaining compounds are used to predict its value. This is done systematically N times. The resulting N predictions are then compared with measured values for their respective compounds to enable computation of Q^2 (Ridder et al., 2010). A more robust and reliable method is the leave-several-out cross-validation. For example, in the leave-20%-out cross-validation, 80% of the compounds are randomly selected for the generation of a model, which is then used to predict the remaining compounds. This operation must be repeated numerous times in order to obtain reliable statistical results. The leave-20%-out, or also the more demanding leave-50%-out, cross-validation results are much better indicators for the robustness and the predictive ability of a QSAR model than the usually used LOO procedure (Golbraikh and Trophsa, 2002).

Finally, Gramatica (2010) suggested that for the most stringent evaluation of model applicability for prediction of new chemicals, external validation verified by Q^2_{EXT} or R^2_{EXT} (R^2 is the linear regression coefficient when calculated removals are correlated with experimental data) is recommended as the last step after model development. The preferred model will be that with the highest prediction parameter values and the most balanced results between the cross-validation parameters on the training chemicals and the predictive power (Q^2_{EXT} or R^2_{EXT}) verified later on the external prediction chemicals.

QSAR models are also used by environmental protection agencies (US EPA, Danish EPA) and in the European Union in order to elaborate the regulation relative to water quality (Cronin et al., 2003). Implementing on the same idea and considering molecular descriptors, models have been developed in order to estimate the removal of several ECs under typical and controlled treatment conditions (e.g. laboratory and/or pilot experiments). These QSAR derived models are named QSAR-like models and are lightly different from QSARs as they can consider some treatment operating conditions into their equation.

For a correct validation of removal estimation, QSAR models validation principles need to be also applied to QSAR-like models.

4. QSAR-like models

Only a few QSAR-like models used to predict μPs removal by some DWT processes were presented in literature. Some models were proposed for membrane filtration (Yangali-Quintanilla et al., 2010), chlorination, ozonation (Lei and Snyder, 2007; Sudhakaran et al., 2012) and adsorption (Brasquet and Le Cloirec, 1999; Redding et al., 2009; Ridder et al., 2009, 2010). Recent QSAR-like models developed for DWT processes are shown in Table 2. The estimation of the μPs removal by a DWT process calculated using these QSAR-

like models was validated by evaluating the model prediction accuracy (R^2) and, in some cases, its predictive power (Q^2).

Lei and Snyder (2007) utilized 3D molecular structural information to develop QSAR-like models to better understand the removal of several EDCs and PPCPs during ozonation and chlorination processes (Table 2). It was suggested that the removal of the selected compounds was a function of descriptors including geometrical properties (volume and surface areas such as molecular, hydrophobic, hydrophilic, π and weakly polar surface area), physicochemical properties (octanol-water partition coefficient K_{ow}; Henry's Law constant, water solubility, polarizability) and structural characteristics (functional groups). The authors observed that the removal obtained by ozonation was largely determined by weakly polar surface area, followed by π surface area and the number of functional groups that can be oxidized. The removal obtained by chlorination was mainly dependent upon the solute ionization potential in addition to the extent of the removal obtained by ozonation (see Table 2, chlorine removal equation).

Ridder et al. (2009) developed different QSAR-like models to determine the removal of 21 PhACs by adsorption onto GAC in either ultrapure or treated surface water, and on either used or freshly regenerated AC. The treated surface water originated from Weesperkarspel water treatment plant, after coagulation, filtration, ozonation and softening pretreatment, and had a pH value of 8. Ultrapure water (pH 4) was produced from tap water, using GAC adsorption, ion exchange and reverse osmosis. The used AC was collected from the full scale GAC adsorbers at Weesperkarspel after a runtime of more than 6 months. Before use, all GAC was sieved and the fraction 0.63–0.71 mm was collected. Fine particles were separated and removed from the AC sample using sedimentation in ultrapure water. Finally, the AC was dried at 105 °C for 24 h. Removal data used to develop the QSAR-like model was determined from adsorption isotherms (8 week equilibrium time). The model includes the following parameters: solute hydrophobicity (expressed as log *D*, the pH-corrected value of the octanol–water partition coefficient log K_{ow}), solute charge and carbon dose. Solute molecular weight was also considered as a model input parameter, but this solute property appeared to poorly relate to solute removal. In the model, solute charge was represented by a simplified parameter which was −1 for negatively charged solutes, 0 for neutral solutes, and +1 for positively charged solutes. This simplified parameter was then multiplied with the percentage of dissociated or protonated solute at the solution pH, to obtain the variable named *charge* used in the model. The resulting proposed equations are shown in Table 2. In all the validation sets, cyclophosphamide removal was over-predicted at low carbon dose. The authors suggested that the over-prediction could be related to the absence of aromatic ring in cyclophosphamide molecular structure, and subsequently, to the inability of cyclophosphamide to form π–π bonds with the AC surface.

Redding et al. (2009) developed a QSAR-like model for the adsorption of 29 PPCPs and EDCs on 3 ACs (Table 2). Rapid small-scale column tests were used to evaluate the removal of the target compounds. Most bed life time data for removing the target EDCs/PPCPs were describable by a normalized QSAR-like model including both solute and AC properties. With regard to the AC, the pertinent characteristics were its pore volume corresponding to pore width in the range 0.4–100 nm, its slurry pH, as an indicator of the carbon surface chemistry, and its apparent density in the column (kept constant in this study). With regard to the EDC/PPCP compounds, the pertinent characteristics were: (1) the molecule compactness, as depicted by the compound 8th-order simple-path Chi index ($^8\chi_p$ – unit less) (Kier and Hall, 1986); (2) the molecule hydrophobic surface area (FOSA; $Å^2$/molecule), as an indicator of the fraction of the molecular surface area that can hydrophobically interact with the AC graphene layers (which are hydrophobic by nature); and (3) the volume of Avogadro's number of molecules (CV, mL/mol). Through the process of developing this QSAR-like model, the authors identified more than 40 molecular descriptors that showed a poor statistical significance. For example, log K_{ow} was not significant, even though it is commonly considered to be an important parameter when considering sorption at the ppm-level (Westerhoff et al., 2005).

Considering a single AC (Filtrasorb 400), Ridder et al. (2010) proposed QSAR-like models based on multivariate linear regression to predict equilibrium carbon loading for solutes reflecting a wide range of solute properties (excluding PhACs, Table 2). In order to improve prediction accuracy, groups of solutes showing similar properties were defined and removals were predicted for each group separately. According to the authors, the main limitations of their models are: i) they describe solute removal for only one specific AC, and ii) that removal was measured in ultrapure water. The authors concluded that the carbon loading could be accurately predicted only after the set of solutes with broad variation in solute properties was subdivided into four groups: aliphatic without H-bond donor/acceptor (d/a) groups, aliphatic with H-bond d/a groups, aromatic without H-bond d/a groups, aromatic with H-bond d/a groups. Hydrophobic partitioning (i.e. log *D* value) was relevant for all solutes. Finally, the authors observed that the influence of the presence of H-bond d/a groups was independent of the amount of H-bond donor or H-bond acceptor groups.

Yangali-Quintanilla et al. (2010) proposed a QSAR-like model to predict the rejection of several PhACs and EDCs by nanofiltration membranes (Table 2). Compound properties describing hydrophobicity (log K_{ow}, log *D*), polarity (dipole moment), and size (molecular parameters such as length, depth, equivalent width, weight and molar volume) were used as descriptors in the model. The molecular length is defined as the distance between the two most distant atoms. The molecular width and molecular depth (width > depth) are measured by projecting the molecule on the plane perpendicular to the length axis and the equivalent molecular width is defined as the geometric mean of width and depth. The acid dissociation constant (pK_a) was used to determine the speciation of the organic compound in ionic species at pH 7. Based on pK_a and log K_{ow} values, the compounds were classified as hydrophilic neutral, hydrophilic ionic, hydrophobic ionic and hydrophobic neutral. Compounds with log $K_{ow} > 2$ were referred to as hydrophobic, therefore those with log $K_{ow} < 2$ were hydrophilic. Nevertheless, this classification was not used in the building up of the model. The QSAR model identified that the most important variables which influenced the

Table 2 – Examples of QSAR-like models used to predict the removal of organic micropollutants in drinking water treatment processes.

Selected compounds	Classification criteria	Experimental conditions	QSAR-like model equation/Molecular descriptors	Ref.

Acenaphthene, Acenaphthylene, Acetaminophen, Aldrin, Androstenedione, Anthracene, Atrazine, Benzophenone, Bisphenol A, Benzo[a]anthracene, Benzo[a]pyrene, Benzo[b]fluoranthene, Benzo[k] fluoranthene, Caffeine, Carbamazepine, Chrysene, Diazepam, Diclofenac, Dieldrin, Dilantin, Endrin, Erythromycin, Estradiol, Estriol, Estrone, Ethinyl estradiol, Fluoranthene, Fluorene, Fluoxetine, Galaxolide, Gemfibrozil, Heptachlor, Heptachlor epoxide, Hydrocodone, Hydrocortisone, Indolebutyric acid, Isobutylparaben, Ibuprofen, Iopromide, Meprobamate, Methoxychlor, Metolachlor, Mirex, Musk ketone, Naphthalene, Naproxen, Octylphenol-4t, Oxybenzone, Pentoxifylline, Phenanthrene, Progesterone, Propylparaben, Pyrene, Sulfamethoxazole, Testosterone, Triclocarban, Triclosan, Trimethoprim.	N A	Three different surface waters:
		Colorado River water, Ohio River water and Passaic River water.
		The water was filtered to remove particulate matter and then spiked with target compounds at a concentration range of 10–250 ng/L.
		Process:
		Ozonation and Chlorination
			For all target compounds:
			% ozone removal = 67.3 + 0.0506 * PISA + 5.20 * (#metab) − 4.34 * (#rtvFG) − 0.114 * (WPSA).
			$R^2 = 0.84$
			For all target compounds:
			% chlorine removal = 106.8 + 0.791 * % ozone removal + 7.89 * (#rtvFG) + 4.80 (QPlog Pow) + 0.175 * (FISA) − 15.0 * (IP)
			$R^2 = 0.71$
			PISA = Molecular π components of surface area
			#metab = number of metabolites
			#rtvFG = Number of reactive functional groups,
			WPSA = Molecular weakly polar components of surface area
			QPlog Pow = Octanol/water partition coefficient
			FISA = Molecular hydrophilic surface area
			IP = Ionization potential
				Lei and Snyder, 2007.

Acetaminophen, Androstenedione, Atrazine, Caffeine, Carbamazepine, N,N-diethyl-meta-toluamide (DEET), Diazepam, Diclofenac, Dilantin, Erythromycin, Estradiol, Estriol, Estrone, Ethynylestradiol, Fluoxetine, Gemfibrozil, Hydrocodone, Ibuprofen, Lopromide, Meprobamate, Naproxen, Oxybenzone, Pentoxifylline, Progesterone, Sulfamethoxazole, tris(2-chloroethyl) phosphate (TCEP), Testosterone, Triclosan, Trimethoprim	N A	Rapid small-scale column tests. Three lignite variants: HYDRODARCO 4000 (HD4000), steam-modified HD4000, and methane/steam-modified HD4000. Target compounds at a concentration ranged of 100–200 ng/L. Native Lake Mead Nevada water **Process: Activated Carbon adsorption**	For all target compounds: $BV_P = \left(\frac{(TPV \times \rho_{mc})(e^{0.2812 \times pH_Z})}{CV \times C_0}\right)\left(0.2758 \times {}^{8}\chi_P + 0.0011 \times FOSA\right)$ BV_p = bed volumes to characteristic initial breakthrough for each of the compounds, TPV = pore volume, ρ_{mc} = apparent density, CV = molecular volume, C_o = influent concentration, ${}^{8}\chi_p$ = 8th order simple-path Chi index (depicts the molecule's compactness), FOSA = molecule's hydrophobic surface area. $R^2 = 0.861$	Redding et al., 2009.

(*continued on next page*)

Table 2 (*continued*)

Selected compounds	Classification criteria	Experimental conditions	QSAR-like model equation/Molecular descriptors	Ref.

Acetaminophen, Androstenedione, Atrazine, Benzopyrene, Caffeine, Carbamazepine, N,N-diethyl-meta-toluamide (DEET), Diazepam, Diclofenac, Dilantin, Erythromycin, Estradiol, Estriol, Estrone, Ethynylestradiol,Fluorene, Fluoxetine, Galaxolide, Gemfibrozil, Hydrocodone, Ibuprofen, Iopromide, Lindane, Meprobamate, Metolachlor, Musk ketone, Naproxen,Oxybenzone, Pentoxifylline, Progesterone, Sulfamethoxazole, tris(2-chloroethyl) phosphate (TCEP), Testosterone, Triclosan, Trimethoprim.	N A	Four different surface waters:
		Colorado River water (CRW), Ohio River water (ORW), Passaic River water (PRW) and synthetic water (SRW) prepared by adding Suwannee River isolate natural organic matter (purchased from International Humic Substances Society) to de-ionized water with sodium bicarbonate added as a pH buffer.
		Two studies were performed:
		one based on an ozone (O_3) dosage and the other on ozone and hydrogen peroxide dosage, representing an advanced oxidation process (AOP)
		Process:
		Ozonation and advanced oxidation
			For all target compounds:
			CRW• Ln rml (AOP) = 9.77 − 0.63(E_{LUMO} − E_{HOMO}) − 0.194 EA + 0.02 #ring
			R^2 = 0.902. Q^2_{Loo} = 0.868
			• Ln rml (O_3) = 12.45 − 0.95(E_{LUMO} − E_{HOMO}) − 0.32 MON
			R^2 = 0.866. Q^2_{Loo} = 0.792
			ORW
			• % rml (AOP) = 318.23 − 26.52(E_{LUMO} − E_{HOMO}) − 6.11 EA + 0.41#ring
			R^2 = 0.922. Q^2_{Loo} = 0.897
			• % rml (O_3) = 310.31 − 25.91(E_{LUMO} − E_{HOMO}) − 8.64 EA − 2.11 #X + 0.32#ring
			R^2 = 0.915. Q^2_{Loo} = 0.853
			PRW
			• Ln rml (AOP) = 15.17 − 1.33(E_{LUMO} − E_{HOMO}) − 0.56 EA + 0.06 #in56
			R^2 = 0.862. Q^2_{Loo} = 0.790
			• Ln rml (O_3) = 18.15 − 1.63(E_{LUMO} − E_{HOMO}) − 0.33 MON
			R^2 = 0.887. Q^2_{Loo} = 0.851
			SRW
			• % rml (AOP) = 286.67 − 22.41(E_{LUMO} − E_{HOMO}) −52.68 O/C
			R^2 = 0.854. Q^2_{Loo} = 0.702
			• % rml (O_3) = 345.54 − 29.18(E_{LUMO} − E_{HOMO}) − 5.59 EA
			R^2 = 0.862. Q^2_{Loo} = 0.78
			
			% rml = percent removal of organic micropollutant
			Ln rml = natural logarithm of percent-removal
			E_{LUMO} − E_{HOMO} = Energy difference between lowest unoccupied and highest occupied molecular orbital.
			EA = Electron-Affinity.
			#ring = Number of ring atoms.
			#in56 = Number of ring atoms in 5 or 6 membered ring.
			MON = mean oxidation number indicates the oxidation state of the compounds.
			#X = number of halogens in a molecule.
			O/C = Oxygen to carbon ratio.
				Sudhakaran et al., 2012.

Caffeine, acetaminophen, phenacetin, phenazone, metronidazole	Hydrophilic neutral	Ultrapure water, 20 °C Ionic strength of 10 mM as KCl	For all target compounds: **Rejection = 265.150 eqwidth − 117.356 depth + 81.662 length − 5.229log D + 1358.090 SR − 1447.817** (N = dataset = 106, R^2 = 0.75 Q^2_{Loo} = 0.72) Eqwidth = molecular equivalent width, Depth = molecular depth, Length = molecular length, SR = magnesium sulphate salt rejection, log D = distribution coefficient (pH-corrected log K_{ow}), log K_{ow} = octanol–water partitioning coefficient	Yangali-Quintanilla et al., 2010.
Sulfamethoxazole	Hydrophilic ionic	Polyamide NF membranes. (NF-200 and NF-90, Dow Chemical Co.)		
Carbamazepine, 17b-estradiol, estrone, bisphenol A, nonylphenol, atrazine	Hydrophobic neutral	**Process: Membrane Nanofiltration (NF)**		

Naproxen, ibuprofen	Hydrophobic ionic			
71 Organic micropollutants (Excluding pharmaceutical compounds)	Aromatic, no H-bond d/a	Ultrapure water/fresh carbon F400 activated carbon Organics-free water, pH (5.3–8.0), 24 °C **Process: Activated Carbon adsorption**	For **Aromatic, no H-bond d/a compounds**: **log qe (μmol/g) = 0.525 * log D + 0.454 * log Ce (μg/L) + 1.22** (N = dataset = 30. $R^2 = 0.84$ $Q^2_{Loo} = 0.80$)	Ridder et al., 2010.
	Aromatic, H-bond d/a		For **Aromatic, H-bond d/a compounds**: **log qe (μmol/g) = 0.140 * log D + 0.237 * log Ce(μg/L) + 2.53** ($N = 22$. $R^2 = 0.81$ $Q^2_{Loo} = 0.76$)	
	Aliphatic, no H-bond d/a		For **Aliphatic, no H-bond d/a compounds**: **log qe (μmol/g) = 0.574 * log D + 0.181 * polarizability + 0.574 * log Ce(μg/L) − 0.68** ($N = 48$. $R^2 = 0.84$ $Q^2_{Loo} = 0.81$)	
	Aliphatic, H-bond d/a		For **Aliphatic, H-bond d/a compounds**: log qe (μmol/g) = 0.196 * log D + 0.319 * log Ce(μg/L) + 2.28 ($N = 34$. $R^2 = 0.61$ $Q^2_{Loo} = 0.55$) qe = Carbon loading. Ce = equilibrium concentration. log D = distribution coefficient.	
Fenoprofen, Clofibric acid, Ibuprofen, Ketoprofen, Diclofenac, Gemfibrozil, Bezafibrate, Naproxen	Negatively Charged	carbon NORIT GAC830 P (0.63–0.71 mm) Initial concentration of 2 μg/L for all pharmaceuticals 12 °C	For all target compounds in **Ultrapure water and fresh regenerated carbon**: **$Ce/C_0 = -0.019 \cdot \log D - 0.029 \cdot charge - 0.284 \cdot \log CC + 0.46$** (Max CC: 44.4 mg/l. N = dataset = 34. $R^2 = 0.58$ $Q^2_{Loo} = 0.47$)	Ridder et al., 2009.
Phenazone, Cyclophosphamide, Aminopyrine, Carbamazepine, Pentoxifylline	Neutral	**Process: Activated Carbon adsorption**	For all target compounds in **Ultrapure water and preloaded carbon**: **$Ce/C_0 = -0.042 \cdot \log D - 0.227 \cdot charge - 0.284 \cdot \log CC + 0.57$** (Max CC: 44.4 mg/l. $N = 40$. $R^2 = 0.63$ $Q^2_{Loo} = 0.50$)	
Terbutaline, Propanolol, Sotalol, Salbutamol, Pindolol, Atenolol, Metoprolol, Clenbuterol, Aminopyrine	Positively Charged		For all target compounds in **Surface water and preloaded carbon**: **$Ce/C_0 = -0.042 \cdot \log D - 0.143 \cdot charge - 0.545 \cdot \log CC + 1.09$** (Max CC: 88.9 mg/l. $N = 62$. $R^2 = 0.78$ $Q^2_{Loo} = 0.75$)	
			CC: Carbon concentration. Ce = equilibrium concentration, C_0 = initial concentration. log D = distribution coefficient. Charge = charged solute fraction * (−1 for negatively charged solutes, 0 for neutral solutes, and +1 for positively charged solutes).	

rejection of organic solutes were log *D*, equivalent width, depth and length of the solute and the magnesium sulphate salt rejection of the membrane.

Sudhakaran et al. (2012) developed QSAR-like models for both ozonation and the combination of ozone and hydrogen peroxide oxidation process (the latter process being called an Advanced Oxidation Process, AOP). Their work was focused on PPCPs and pesticides present at ng/L levels in 4 natural waters (Table 2). Bench-scale experiments were carried out in order to determine the μPs removal percentages. The energy difference between the lowest unoccupied and the highest occupied molecular orbitals ($E_{LUMO} - E_{HOMO}$) appeared in all the QSAR-like models developed which indicated that it was an important parameter in understanding ozonation mechanisms. According to the authors, the models developed were boundary conditioned to the water quality parameters such as pH, dissolved organic carbon and alkalinity.

5. Discussion

Some previous studies have suggested several selection criteria in order to prioritize ECs to be monitored during water treatment. Bruchet et al. (2005) suggested selecting a priority list of 10, up to 25 ECs representative of various therapeutic classes, known for their resistance towards both wastewater and DWT processes and their persistence in the environment. In particular, the authors suggested that sulfamethoxazole (for antibiotics) and iopamidol (for X-ray contrast agents) could be relevant indicators. A compound such as iopamidol, which combines a non-toxic character and a higher persistence than other ECs towards DWT processes, would hence represent a crucial indicator to guarantee that DW is free from significant levels of drugs. Carbamazepine has been found in many countries and was proposed as a good representative of the antiepileptics. Finally, galaxolide, which is consumed in large quantities in Europe, could be used as an indicator for cosmetics.

Benotti et al. (2009) evaluated the efficiency of ozone and chlorine treatment for the elimination of 20 PhACs and 25 EDCs. The authors suggested a list of 6 potential indicator compounds divided into 3 groups according to their likelihood of being removed by: (1) both chorine and ozone (estrone and trimethoprim); (2) ozone only (N,N-diethyl-meta-toluamide, atenolol) or (3) neither chlorine nor ozone (atrazine, meprobamate).

De Voogt et al. (2009) proposed 3 priority lists of pharmaceuticals relevant for the water cycle according to 7 sets of criteria extracted from a literature review of prioritisation studies (a total of 25 publications). The criteria were: regulation, consumption, physicochemical properties, toxicity, occurrence, persistence in the environment and resistance to water treatment. Among the 153 compounds identified in the review they prioritised 44 PhACs, which were then classified into 3 lists: 1. High priority pharmaceuticals (10 compounds) including PhACs that were mentioned in 5 or more of the reference documents cited, and that fulfilled more than 4 of the 7 criteria. 2. Priority pharmaceuticals (18 compounds) corresponding to PhACs that were mentioned in more than 2 of the reference documents cited, and that fulfilled more than 2 criteria. 3. Lower priority pharmaceuticals (16 compounds) consisting of PhACs that were mentioned in 2 of the reference documents cited, and that fulfilled 2 or more criteria. As pointed out by the authors, these lists include compounds that were identified in the current literature and therefore the lists need to be regularly updated as new compounds will be identified.

Recently, a comprehensive summary of ranking approaches of PPCPs and EDCs for monitoring purposes was reported by Kumar and Xagoraraki (2010). The authors also developed a ranking system for prioritizing PPCPs and EDCs to be monitored in several US source waters and finished DWs considering the following 4 criteria: (1) occurrence (prevalence, frequency of detection), (2) removal in DWT plants, (3) ecological effects (bioaccumulation, ecotoxicity) and (4) health effects. The health effects were characterized using 7 attributes: carcinogenicity, mutagenicity, impairment of fertility, central nervous system acting, endocrine effects, immunotoxicity and developmental effects. According to the authors, the ranking system they propose uses both scientific and subjective data to determine the relative importance of the various criteria. The authors highlighted the facts that 1) health effects and treatment criteria influenced considerably on the ranking system and 2) there is a need for further data collection on these two criteria.

Currently, very little is known about potential effects on human health associated with chronic exposure to trace levels of multiple μPs (i.e. synergistic effects) through routes such as DW (Jones et al., 2005; Tourad et al., 2011). While some frequently detected μPs, such as PhACs, were designed to be ingested, others (e.g. flame retardants, solvents, and other personal care and industrial chemicals) were not designed for human consumption. Some of the latter compounds (e.g. bisphenol A) are known or suspected endocrine disruptors and may be potent reproductive toxics even at low concentrations (Stackelberg et al., 2004). Consequently, it would be prudent to reduce the levels of μPs in DW and hence

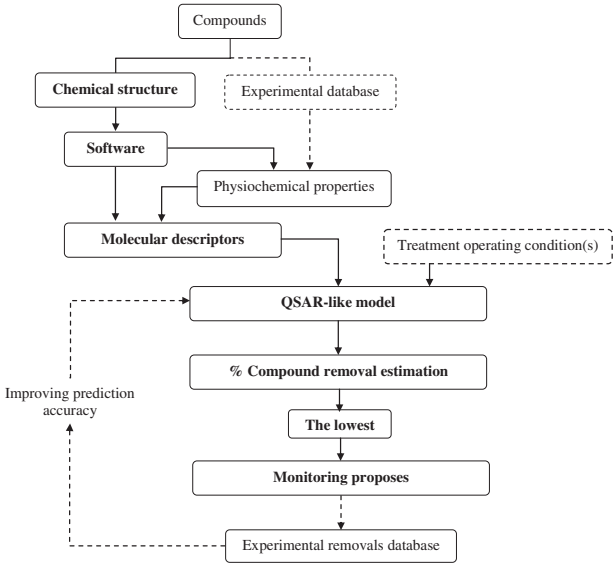

Fig. 1 – Overview of the proposed QSAR-like model analysis system for the selection of PhACs and EDCs for monitoring purposes.

a removal criterion seems to be the most important criterion to be considered in order to prioritize the μPs, and more specifically ECs, to be monitored during DWT. Therefore, the short list of priority compounds to be taken as indicators of DW quality should be set selecting the ECs showing the lowest removals in current DWT chains (assuming that a higher removal would thus be obtained for all the other ECs).

Based on this single removal criterion, QSAR-like models show a high potential to allow selecting indicator compounds to be considered for monitoring proposes, besides predicting ECs removal in DWT processes, taking into account the solute properties and, in some cases, a few treatment operating conditions: Once a QSAR-like model has been developed and validated, one could estimate the removal of ECs in DWT processes by using their molecular descriptors. Furthermore, molecular descriptors (including physiochemical properties) can be estimated using commercial softwares (e.g. ADME, Dragon, Chem3D Ultra, QikProp, Marvin) considering the EC chemical structure, thereby overcoming the need to conduct experimental measurements. In this way, the removal of ECs could be estimated starting with their chemical structure, hence enabling estimating the removal of newly detected environmental ECs, as well as ECs metabolites and ECs transformation products, for which the chemical structure is known. However, QSAR-like models are only able to estimate ECs removal under the same operating conditions in which they were developed. Since the extent of ECs removal is expected to vary from one process to another, a QSAR-like model should be developed for each DWT process (i.e. ozonation, chlorination, nanofiltration, AC adsorption…). Moreover, by classifying the solutes of a dataset into specific groups of compounds that show similar properties, the model prediction accuracy should be improved. Thereafter, considering the removal estimations of all the ECs predicted by a QSAR-like model, the compounds showing the lowest predicted removals would need to be monitored, as indicator compounds, in DWT plants and would warrant further studies (e.g. development of analytical methods, occurrence surveys, evaluation of treatability and toxicity, etc).

Under this proposed scheme, QSAR-like models would estimate the removal of ECs in DWT processes (including parent compounds, metabolites and transformation products) and would help select compounds, or a specific group containing similar solute properties, for monitoring purposes. A scheme of such an analysis system is presented in Fig. 1.

Finally, some already existing data can be used either to develop or improve the QSAR-like models. For instance, a comprehensive dataset of removal values obtained for 62 EDCs and PPCPs considering several DWT processes was published by Westerhoff et al. (2005) and Snyder et al. (2007). The removal data was generated from surface waters having complex matrixes, which included background contaminations and natural organic matter, in typical operating conditions. This database should be used for developing new QSAR-like models. Furthermore, it is suggested that the removal of compounds, expressed in percentage, should be used as the defined endpoint of these new QSAR-like models in order to facilitate further selection of indicator compounds based on the removal criterion.

6. Conclusion

This review shows that a large and still growing variety of ECs, such as PhACs and EDCs, has been detected in finished DWs of a number of industrialized countries. The chronic potential risks to human health of these trace compounds and, for some of them, their resistance towards DWT processes, highlight the need to remove ECs during DW preparation. These factors, in addition to the highly diversified physicochemical properties of these compounds and the subsequent practical difficulties to fully assess their removal during DW preparation, make it necessary to develop more QSAR-like models. According to the QSAR-like model analysis system proposed in this paper, these models would serve i) to predict the removal of ECs towards DWT processes and ii) to identify those ECs that warrant further study because of their predicted low removal for a given DWT process. The latter compounds could usefully serve as relevant indicators for monitoring purposes of DWT processes.

Acknowledgments

This work was carried out in the frame of the Sharing Knowledge Program in the Field of Environment, set up between Rhone Alps Region (France) and the Municipality of Shanghai (China), within the sub-program Securing Drinking Water Supply. The authors wish to thank Rhone-Alps Region and Suez-Environnement research centre (CIRSEE) for their financial support.

REFERENCES

Benotti, M., Trenholm, R., Vanderford, B., Holady, J., Stanford, B., Snyder, S., 2009. Pharmaceuticals and endocrine disrupting compounds in U.S. drinking water. Environmental Science and Technology 43, 597–603.

Boleda, M.R., Galceran, M.T., Ventura, F., 2011. Behavior of pharmaceuticals and drugs of abuse in a drinking water treatment plant using combined conventional and ultrafiltration and reverse osmosis treatments. Environmental Pollution 159, 1584–1591.

Brasquet, C., Le Cloirec, P., 1999. QSAR for organics adsorption onto activated carbon in water: what about the use of neural networks. Water Research 33 (17), 3603–3608.

Bruchet, A., Hochereau, C., Picard, C., Decottignies, V., Rodrigues, J.M., Janex-Habibi, M.L., 2005. Analysis of drugs and personal care products in French source and drinking waters: the analytical challenge and examples of application. Water Science and Technology 52 (8), 53–61.

Casajuana, N., Lacorte, S., 2003. Presence and release of phthalic esters and other endocrine disrupting compounds in drinking water. Chromatographia 57, 649–655.

Chen, M., Ohman, K., Metcalfe, C., Ikonomou, M.G., Amatya, P.L., Wilson, J.J., 2006. Pharmaceuticals and endocrine disruptors in wastewater treatment effluents and in the water supply system of Calgary, Alberta, Canada. Water Quality Research Journal of Canada 41, 351–364.

Cronin, M.T.D., 2010. Quantitative structure-activity relationships (QSARs) – applications and methodology. In: Puzyn, T.,

Leszczynski, J., Cronin, M.T.D. (Eds.), Recent Advances in QSAR Studies. Springer, pp. 3–11.
Cronin, M.T.D., Walker, J.D., Jaworska, J.S., Comber, M.H.I., Watts, C.D., Worth, A.P., 2003. Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances. Environmental Health Perspectives. 111, 1376–1390.
De Voogt, P., Janex-Habibi, M.L., Sacher, F., Puijker, L., Mons, M., 2009. Development of a common priority list of pharmaceuticals relevant for the water cycle. Water Science & Technology 59.1, 39–46.
Directive 2008/105/EC of the European Parliament and of the Council of 16 December 2008 on environmental quality standards in the field of water policy, OJ L 348, 24.12.2008, pp. 84–97.
European Commission (EC), 2011. Review of the List of Priority Substances Under the Water Framework Directive.
Environmental Protection Agency (EPA), 2009. Contaminant Candidate List 3. http://www.epa.gov/ogwdw000/ccl/ccl3.html#ccl3.
Escher, B.I., Bramaz, N., Eggen, R.I.L., Richter, M., 2005. In vitro assessment of modes of toxic action of pharmaceuticals in aquatic life. Environmental Science & Technology 39 (9), 3090–3100.
Garcia-Ac, A., Segura, P.A., Viglino, L., Fürtos, A., Gagnon, C., Prévost, M., Sauvé, S., 2009. On-line solid-phase extraction-LC-MS/MS using large volume injections: validation for trace organic contaminants in surface and drinking water. Journal of Chromatography A 1216, 8518–8527.
Gramatica, P., 2010. Chemometric methods and theoretical molecular descriptors in predictive QSAR modeling of the environmental behavior of organic pollutants. In: Puzyn, T., Leszczynski, J., Cronin, M.T.D. (Eds.), Recent Advances in QSAR Studies. Springer, pp. 327–366.
Golbraikh, A., Trophsa, A., 2002. Beware of q2! Journal of Molecular Graphics and Modelling 20, 269–276.
Huerta-Fontela, M., Galceran, M.T., Ventura, F., 2011. Occurrence and removal of pharmaceuticals and hormones through drinking water treatment. Water Research 45, 1432–1442.
Jones, A.J., Lester, J.N., Voulvoulis, N., 2005. Pharmaceutical: a threat to drinking water? Trends in Biotechnology 23, 163–167.
Kier, L.B., Hall, L.H., 1986. Molecular Connectivity in Structure–Activity Analysis. Research Studies Press, Letchworth, Hertfordshire, England.
Kleywegt, S., Pileggi, V., Yang, P., Hao, C., Zhao, X., Rocks, C., Thach, S., Cheung, P., Whitehead, B., 2011. Pharmaceuticals, hormones and bisphenol A in untreated source and finished drinking water in Ontario, Canada – Occurrence and treatment efficiency. Science of the Total Environment 409, 1481–1488.
Kruhlak, N.L., Contrera, J.F., Benz, R.D., Matthews, E.J., 2007. Progress in QSAR toxicity screening of pharmaceutical impurities and other FDA regulated products. Advanced Drug Delivery Reviews 59, 43–55.
Kuch, M.H., Ballschmiter, K., 2001. Determination of endocrine-disrupting phenolic compounds and estrogens in surface and drinking water by HRGC-(NCI)-MS in the pictogram per litre range. Environmental Science and Technology 35, 3201–3206.
Kumar, A., Xagoraraki, I., 2010. Pharmaceuticals, personal care products and endocrine-disrupting chemicals in U.S. surface and finished drinking waters: a proposed ranking system. Science of the Total Environment 48, 4972–5989.
Kuster, M., Lopez de Alda, M.J., Hernando, M.D., Petrovic, M., Martin-Alonso, J., Barcelo, D., 2008. Analysis and occurrence of pharmaceuticals, estrogens, progestagens and polar pesticides in sewage treatment plant effluents, river water and drinking water in the Llobregat river basin (Barcelona, Spain). Journal of Hydrology 358, 112–123.
Lei, H., Snyder, S.A., 2007. 3D QSPR models for the removal of trace organic contaminants by ozone and free chlorine. Water Research 41, 4051–4060.
Mompelat, S., Le Bot, B., Thomas, O., 2009. Occurrence and fate of pharmaceutical products and by-products, from resource to drinking water. Environment International 35, 803–814.
Nikolaou, A., Meric, S., Fatta, D., 2007. Occurrence patterns of pharmaceuticals in water and wastewater environments. Analytical Bioanalytical Chemistry 387, 1225–1234.
Organisation for Economic Cooperation and Development (OECD), 2004. The Report from the Expert Group on (Quantitative) Structure Activity Relationship ([Q]SARs) on the Principles for the Validation of (Q)SARs, Series on Testing and Assessment No. 49 (ENV/JM/MONO(2004)24). France, Paris. Available at: http://www.oecd.org/.
Pomati, F., Castiglioni, S., Zuccato, E., Fanelli, R., Vigetti, D., Rossetti, C., Calamari, D., 2006. Effects of a complex mixture of therapeutic drugs at environmental levels on human embryonic cells. Environmental Science and Technology 40 (7), 2442–2447.
Reddersen, K., Heberer, T., Dünnbier, U., 2002. Identification and significance of phenazone drugs and their metabolites in ground- and drinking water. Chemosphere 49, 539–544.
Redding, A.M., Cannon, F.S., Snyder, S.A., Vanderford, B.J., 2009. A QSAR-like analysis of the adsorption of endocrine disrupting compounds, pharmaceuticals, and personal care products on modified activated carbons. Water Research 43, 3849–3861.
Reungoat, J., Macova, M., Escher, B.I., Carswell, S., Mueller, J.F., Keller, J., 2010. Removal of micropollutants and reduction of biological activity in a full scale reclamation plant using ozonation and activated carbon filtration. Water Research 44, 625–637.
Richardson, S.D., 2009. Water analysis: emerging contaminants and current Issues. Analytical Chemistry 81, 4645–4677.
Richardson, S.D., 2010. Environmental mass spectrometry: emerging contaminants and current Issues. Analytical Chemistry 82, 4742–4774.
Ridder, D.J., Villacorte, L., Verliefde, A.R.D., Verberk, J.Q.J.C., Heijman, S.G.J., Amy, G.L., Van Dijk, J.C., 2010. Modelling equilibrium adsorption of organic Micropollutants onto activated carbon. Water Research 44, 3077–3086.
Ridder, D.J., McConville, M., Verliefde, A.R.D., Van der Aa, L.T.J., Heijman, S.G.J., Verberk, J.Q.J.C., Rietveld, L.C., Van Dijk, J.C., 2009. Development of a predictive model to determine micropollutant removal using granular activated carbon. Drinking Water Engineering and Science 2, 57–62.
Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., Von Gunten, U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313, 1072–1077.
Sippl, W., 2010. 3D-QSAR – Applications, recent advances, and limitations. In: Puzyn, T., Leszczynski, J., Cronin, M.T.D. (Eds.), Recent Advances in QSAR Studies. Springer, pp. 103–125.
Snyder, S.A., Westerhoff, P., Yoon, Y., Sedlak, D.L., 2003. Pharmaceuticals, personal care products, and endocrine disruptors in water: implications for the water industry. Environmental Engineering Science 20 (5), 449–469.
Snyder, S.A., Wert, E.C., Lei, H.X., Westerhoff, P., Yoon, Y., 2007. Removal of EDCs and Pharmaceuticals in Drinking and Reuse Treatment Processes. Awwa Research Foundation, Denver, CO, p. 331.
Stackelberg, P.E., Furlong, E.T., Meyer, M.T., Zaugg, S.D., Henderson, A.K., Reissman, D.B., 2004. Persistence of pharmaceutical compounds and other organic wastewater contaminants in a conventional drinking-water treatment plant. Science of the Total Environment 329 (1–3), 99–113.
Stackelberg, P.E., Gibs, J., Furlong, E.T., Meyer, M.T., Zaugg, S.D., Lippincott, R.L., 2007. Efficiency of conventional drinking

water-treatment processes in removal of pharmaceuticals and other organic compounds. Science of the Total Environment 377, 255–272.

Stolker, A.M.M., Niesing, W., Hogendoorn, E.A., Versteegh, J.F.M., Fuchs, R., Brinkman, U.A.T., 2004. Liquid chromatography with triplequadrupole or quadrupole-time of flight mass spectrometry for screening and confirmation of residues of pharmaceuticals in water. Analytical and Bioanalytical Chemistry 378, 955–963.

Sudhakaran, S., Calvin, J., Amy, G.L., 2012. QSAR models for the removal of organic micropollutants in four different river water matrices. Chemosphere 87 (2), 144–150.

Togola, A., Budzinski, H., 2008. Multi-residue analysis of pharmaceutical compounds in aqueous samples. Journal of Chromatography A 1177, 150–158.

Tourad, E., Roig, B., Sumpter, J.P., Coetsier, C., 2011. Drug residues and endocrine disruptors in drinking water: risk for humans? International Journal of Hygiene and Environmental Health 214, 437–441.

Vosges, M., Braguer, J.C., Combarnous, Y., 2008. Long-term exposure of male rats to low-dose ethinylestradiol (EE2) in drinking water: effects on ponderal growth and on litter size of their progeny. Reproductive Toxicology 25 (2), 161–168.

Vulliet, E., Cren-Olivé, C., Grenier-Loustalot, M.F., 2011. Occurrence of pharmaceuticals and hormones in drinking water treated from surface waters. Environmental Chemistry Letters 9, 103–114.

Wang, H.X., Zhou, Y., Juang, Q.W., 2012. Simultaneous screening of estrogens, progestogens, and phenols and their metabolites in potable water and river water by ultra-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry. Microchemical Journal 100, 83–94.

Wenzel, A., Muller, J., Ternes, T., 2003. Study on Endocrine Disrupters in Drinking Water. Final report. ENV.D.1/ETU/2000/0083. ESWE Institute for Water research and Water Technology GmbH, Wiesbaden-Schierstein, Germany.

Westerhoff, P., Yoon, Y., Snyder, S., Wert, E., 2005. Fate of endocrine-disruptor, pharmaceutical, and personal care product chemicals during simulated drinking water treatment processes. Environmental Science and Technology 39, 6649–6663.

Weyer, P., Riley, D., 2001. Endocrine Disruptors and Pharmaceuticals in Drinking Water. AWWA Research Foundation and American Water Works Association, Denver, CO, USA.

Yangali-Quintanilla, V., Sadmani, A., McConville, M., Kennedy, M., Amy, G., 2010. A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes. Water Research 44, 373–384.

Ye, Z., Weinberg, H.S., Meyer, M.T., 2007. Trace analysis of trimethoprim and Sulfonamide, Macrolide, Quinolone, and Tetracycline antibiotics in Chlorinated Drinking Water Using Liquid Chromatography Electrospray Tandem Mass Spectrometry. Analytical Chemistry 79 (3), 1135–1144.

Yu, Z., Peldszus, S., Huck, P.M., 2007. Optimizing gas chromatographic–mass spectrometric analysis of selected pharmaceuticals and endocrine-disrupting substances in water using factorial experimental design. Journal of Chromatography A 1148, 65–77.

Zuccato, E., Castiglioni, S., Fanelli, R., Reitano, G., Bagnati, R., Chiabrando, C., Pomati, F., Rossetti, C., Calamari, D., 2006. Pharmaceuticals in the environment in Italy: causes, occurrence, effects and control. Environmental Science and Pollution Research 13, 15–21.

Zühlke, S., Duennbier, U., Heberer, T., 2007. Investigation of the behaviour and metabolism next term of pharmaceutical residues during purification of contaminated ground water used for drinking water supply. Chemosphere 69, 1673–1680.