In This Issue:

ET&C FOCUS

Focus articles are part of a regular series intended to sharpen understanding of current and emerging topics of interest to the scientific community.

Alternative Approaches to Vertebrate Ecotoxicity Tests in the 21st Century: A Review of Developments Over the Last 2 Decades and Current Status

Adam Lillicrap,*† Scott Belanger,‡ Natalie Burden,§ David Du Pasquier,‖ Michelle R. Embry,# Marlies Halder,†† Mark A. Lampi,‡‡ Lucy Lee,§§ Teresa Norberg-King,‖‖ Barnett A. Rattner,## Kristin Schirmer,†††‡‡‡§§§ and Paul Thomas###

†Norwegian Institute for Water Research (NIVA), Oslo, Norway
‡Environmental Safety and Sustainability, Global Product Stewardship, Procter & Gamble, Mason, Ohio, USA
§National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), London, United Kingdom
‖WatchFrog, Evry, France
#ILSI Health and Environmental Sciences Institute, Washington, DC, USA
††European Commission Joint Research Centre, Ispra, Italy
‡‡ExxonMobil Biomedical Sciences, Annandale, New Jersey, USA
§§Faculty of Science, University of the Fraser Valley, Abbotsford, British Columbia, Canada
‖‖National Health and Environmental Effects Laboratory, Office of Research and Development, Mid-Continent Ecology Division-Duluth, US Environmental Protection Agency, Duluth, Minnesota, USA
##Patuxent Wildlife Research Center, US Geological Survey, Beltsville, Maryland, USA
†††Eawag–Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
‡‡‡Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Swiss Federal Institute of Technology, Zürich, Switzerland
§§§School of Architecture, Civil and Environmental Engineering, EPF Lausanne, Lausanne, Switzerland
###Consultancy for Environmental & Human Toxicology & Risk Assessment (Lyon Agency), L'Isle d'Abeau, France

Abstract—*The need for alternative approaches to the use of vertebrate animals for hazard assessment of chemicals and pollutants has become of increasing importance. It is now the first consideration when initiating a vertebrate ecotoxicity test, to ensure that unnecessary use of vertebrate organisms is minimized wherever possible. For some regulatory purposes, the use of vertebrate organisms for environmental risk assessments has been banned; in other situations, the number of organisms tested has been dramatically reduced or the severity of the procedure refined. However, there is still a long way to go to achieve a complete replacement of vertebrate organisms to generate environmental hazard data. The development of animal alternatives is based not just on ethical considerations but also on reducing the cost of performing vertebrate ecotoxicity tests and in some cases on providing better information aimed at improving environmental risk assessments. The present Focus article provides an overview of the considerable advances that have been made toward alternative approaches for ecotoxicity assessments over the last few decades. Environ Toxicol Chem 2016;35:2637–2646.* © *2016 SETAC*

Keywords—Ecotoxicity; Vertebrate; In vitro; In silico; 3Rs

* Address correspondence to ali@niva.no

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/etc.3603

Introduction

The book on the principles of humane experimental techniques by Russell and Burch [1] is now more than half of a century old, and still it is considered the seminal writing for alternative approaches to animal testing. This is where the idea of the 3Rs (reduction, refinement, and replacement) was conceived: that any experimental technique should consider a reduction in numbers of animals used; refinement of any procedures to minimize pain, suffering, and distress; and replacement of the use of animals wherever possible. Yet the

hazard assessment of chemicals and effluents that have limited toxicological information still relies heavily on animal tests, and hundreds of thousands of animals are used annually [2].

Although the concept of the 3Rs was initially focused on animal tests with mammalian species for ethical reasons, there has been significant progress of the last 2 decades toward developing alternative approaches for nonmammalian classes of vertebrates used in ecotoxicity testing: fish, amphibians, and birds. When fish were first used for ecotoxicology assessments, the most commonly used life stages were juveniles or adults of species with relatively long life cycles. Subsequently, increasing recognition of the importance of assessing the acute lethality of chemicals as well as the chronic, sublethal effects (survival, growth, fecundity) led to a transition from using larger and older life stages of fish to using earlier developmental stages and/or smaller species with comparatively shorter life cycles [3].

To obtain international acceptance within the scientific and regulatory communities, any new strategies and methods to reduce, replace, and refine animal tests should include consideration of an additional "3Rs"—namely, their reproducibility/reliability, ecological relevance, and regulatory acceptance. The variety of potential alternative approaches for ecotoxicity assessments is very broad but, simply put, includes new approaches of in silico, in vitro, and even in vivo test methods [4]. Bringing all of these tools together allows integrated test strategies to be developed with the ultimate goal of a more thorough understanding of the possible effects of chemicals on the environment. The present Focus article provides a collective overview of the different approaches that are now available for ecotoxicologists and risk assessors to evaluate potentially hazardous chemicals and minimize the use of vertebrates in ecotoxicity tests (summarized in Figure 1).

Regulatory Aspects of Vertebrate Ecotoxicity and the Need for Alternative Approaches

Hazard assessments of chemicals require ecotoxicity studies with vertebrate organisms such as fish. Fish tests are also required in some countries to assess the toxicity of effluents. The use of fish for effluent toxicity assessments likely exceeds their use for chemical hazard assessments. Because industrial effluents can be inherently variable and episodic in terms of flow and chemical constituents, tests are often required with a prescribed frequency to monitor the potential for adverse responses and risk. Thus, the use of fish in many countries is, for now at least, a permanent fixture and may even require constitutional laws to be changed before fish are not required to test the toxicity of effluents. Addressing the need for alternatives to vertebrates ranges from banning the use of fish in effluent assessments, to using fish embryo assays in place of later life stages for acute testing and experimentally assessing whether embryo assays can provide sufficient sensitivity for chronic/subchronic testing needs [5]. As treatment methods for

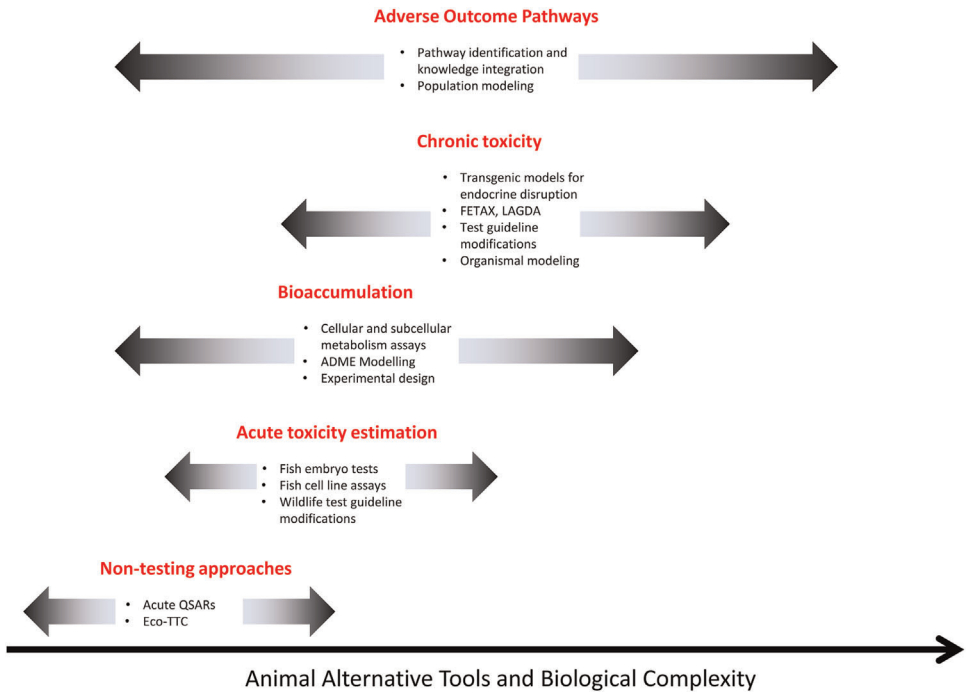

FIGURE 1: Overview of different areas of animal alternatives discussed in the present Focus article. ADME = absorption, distribution, metabolism, and excretion; FETAX = frog embryo teratogenesis assay *Xenopus*; LAGDA = larval amphibian growth and development assay; QSAR = quantitative structure–activity relationship; TTC = threshold for toxicological concern.

industrial and municipal wastewaters improve, acute toxicity tests of effluents may become redundant, and more focus will be placed on chronic toxicity testing of effluents. Currently, it is recognized that the lack of rapid tests that are capable of identifying possible chronic effects of effluents to fish is a shortcoming.

Legislation and data requirements for assessing the hazards of chemicals to the environment vary depending on the intended application of the substance (e.g., different regulations for pharmaceuticals, industrial chemicals, agrochemicals, biocides, and feed additives). Also, testing methods and data requirements differ among countries. Because companies will often seek global marketing of their products, this may mean that vertebrate tests are carried out to meet local requirements specified in a minority of countries or that additional species need to be tested for regional requirements.

Some legislation (e.g., the European Union's regulation on the Registration, Evaluation, Authorisation and Restriction of Chemicals [REACH]) stipulates that vertebrate tests should be carried out only as a last resort [6]. The REACH legislation also specifies that registrants must prove that they have considered alternative methods in their testing proposals before initiating any vertebrate testing. A recent report from the European Chemicals Agency showed that testing proposals now include the use of quantitative structure–activity relationship (QSAR) models to predict both bioaccumulation and short- and long-term toxicity in fish [7]. Furthermore, the US Environmental Protection Agency (USEPA) has routinely utilized QSAR approaches for new chemical registrations and other purposes since the 1980s [8]. A large application factor is assigned when QSARs for hazard assessment are used, necessitating the generation of in vivo data when volumes are large or expected toxicity is high. Read-across approaches, such as the use of data from analogous substances, are also being utilized increasingly to predict potential toxicity and bioaccumulation to omit additional testing proposals. The reliability of these predictions or the extent to which read-across approaches reduce the number of in vivo studies is unknown. Furthermore, a recent European Food Safety Authority aquatic guidance document [9] outlines opportunities where nontesting methods may be considered to address the information requirements for agrochemicals, including those related to pesticide metabolites. An example of a nontesting approach recognized in the European Food Safety Authority guidance document is the use of QSAR models for assessing the toxicity of cosmetic products, a consequence of the banning of animal testing for the safety testing of cosmetics in Europe (regulation 1223/2009 [10]). These examples illustrate that alternative approaches to animal testing are slowly being integrated into regulatory test programs; but because of the complex nature of these regulatory arenas, the extent to and speed with which different regulatory bodies implement alternative approaches still vary.

QSAR Models for Hazard Evaluation of Chemicals

Quantitative structure–activity relationship models are nontesting methods based on the similarity principle (i.e., analogous substances have similar biological activities and modes of action). These models assume that the physical, chemical, and biological properties of a molecule are related to its geometric or electronic properties (i.e., its structure). Parameters such as the log octanol–water partition coefficient (K_{OW}) or knowledge of the 2-dimensional or 3-dimensional structure can be used to predict (eco)toxicological endpoints. These models are constructed from databases of known properties using biomathematical approaches and require 3 components: a high-quality data set of chemicals with experimentally measured biological activities, a structure-related property data set for the chemical structures, and characterization of a relationship [11]. The models can be either mechanistic (using a priori knowledge of the mechanism of the studied activity) or descriptive (utilizing mathematical algorithms to select the most relevant descriptors to optimize predictions).

The quality of data sets used to train QSAR models is very important; therefore, only the most reliable experimental studies should be used for development of the model. The QSAR models that have been developed without prior study validation often suffer from poor predictive capacity. In addition, systematic faults in data and random input errors can lead to an apparently valid QSAR, but the resulting predictions may deviate significantly from experimental data. Thus, poor-quality data points from the training set used to develop the QSAR must be identified and omitted. Tropsha [12] underlined the importance of data quality when developing QSAR models, including chemical structure curation, outlier detection, and data set balancing.

Over the last 20 yr, QSAR models have been applied in many domains, such as rapid screening approaches, hazard prediction, risk assessment, regulatory applications, and drug discovery. Using QSAR models, new substances can be screened initially as persistent, bioaccumulative, and toxic (PBT), and/or as carcinogenic, mutagenic, or reproductively toxic (CMR). Also, QSAR models may help a risk assessor identify flawed empirical study results or even aid in providing more accurate results than routine laboratory assays (e.g., for substances that are difficult to test because of low water solubility or other issues). In 2007, the Organisation for Economic Co-operation and Development (OECD) published 5 governing rules to ensure that QSAR models are of acceptable quality and thus a valid alternative to experimental data. These principles have also been integrated into REACH, which requires that a QSAR model must have a defined endpoint, unambiguous algorithm, defined domain of applicability, appropriate measures of goodness of fit, robustness and predictivity, and mechanistic interpretation wherever possible. In the future, high-accuracy QSAR models capable of providing results sufficient to replace empirical

studies will be emphasized; but for this to occur, the modes of action for each chemical structure tested must be well understood.

Thresholds of Toxicological Concern

The threshold for toxicological concern (TTC) concept is well established for assessing the safety of indirect food-contact substances to humans. This approach uses existing data to determine threshold exposure levels below which there would be no appreciable risk based on a de minimis value for toxicity identified for many chemicals. The TTC relies on existing information rather than requiring additional testing for screening-level assessments. An extension of the human safety TTC concept for application in environmental situations, termed the "ecological TTC" or "eco-TTC," has been explored by several groups and further developed as a multisector project by a Health and Environmental Sciences Institute committee [13]. The eco-TTCs summarize the wealth of ecotoxicological information as predicted no-observed effect concentrations on diverse chemical substances in the form of statistical (probability-based) distributions. In addition, eco-TTCs can be developed that allow prediction of untested chemicals based on structural attribute (category), mode of action, or functional use. The approach may be useful for assessing chemicals at early tiers of the risk-assessment process, providing hazard perspective on chemicals that lack QSAR models, guiding product development discussions, and assisting read-across or category justifications. This approach has the potential to reduce the need for vertebrate testing in many situations.

Progress and Development of In Vitro Assays for Predicting the Acute Toxicity of Chemicals in Fish

The primary interaction between chemicals and life-forms begins at the cell surface or inside the cell. Thus, studies at the cellular level are of key importance in ecotoxicology. The availability of cell lines capable of being continuously cultured began approximately 50 yr ago. More recently, fish cell lines have been used to study the modes of action of environmental contaminants (e.g., Bols et al. [14]). The effects of chemicals on processes such as xenobiotic metabolism, DNA damage, membrane transport, and oxidative stress, and on specialized functions such as phagocytosis and the synthesis of specific proteins have all been evaluated using fish cell lines. Fish cell lines have also been used for toxicity testing, and cytotoxicity assays with fish cell lines have provided reliable information in surveillance programs that target environmental chemicals [15]. For example, a fish gill cell line was able to distinguish effluents from a paper mill that were nontoxic to fish from those that were acutely toxic [16]. Tanneberger et al. [17] demonstrated that a fish gill cytotoxicity assay showed a nearly 1:1 correlation to acute fish toxicity data for a wide range of organic chemicals. The use of the fish gill cell line assay has recently been subjected to an international round-robin test, demonstrating its robustness and interlaboratory reproducibility, and is currently being evaluated as a potential new standard method within the International Organization for Standardization.

Refinements to the traditional monolayer growth of cells in culture are proving useful in ecotoxicological studies. For example, Baron et al. [18] developed a primary rainbow trout hepatocyte spheroidal culture system that can be maintained for over 1 mo and used for in vitro studies of chemical metabolism by fish. Another example is a primary fish gill cell culture model which allows active transport processes through the gill epithelium to be examined [19]. Sophisticated technologies, incorporating enhanced methodologies such as fluidic biochips with fish cell lines, are also being developed for the rapid evaluation of water safety [20].

The Use of Fish Embryos as a Surrogate Life Stage

Fish embryos increasingly have been used as a surrogate for juvenile fish because they are considered in many jurisdictions to be a nonprotected life stage (i.e., not subject to the same animal welfare considerations as juvenile and adult fish). In Germany, for example, a test using fish embryos (zebrafish, exposure 48 h postfertilization [21]) has replaced the fish test (golden ide, 48-h exposure [22]) for assessing the acute toxicity of effluents. The fish embryo acute toxicity test for testing the acute toxicity of chemicals was adopted as a new OECD test guideline (OECD 236 [23]) only after substantial validation was performed [24]. It is a promising alternative test to predict acute lethality with similar sensitivity to the acute fish lethality test [25]. Nevertheless, its broad regulatory acceptance to replace the fish acute toxicity test is still not clear.

Most fish embryo toxicity tests have been conducted with zebrafish; however, Braunbeck et al. [26] demonstrated their applicability to both fathead minnow and medaka embryos. Using zebrafish embryos offers some benefits over the use of fathead minnows, because the spawning behavior of zebrafish enables the collection of embryos to be synchronized, and tests can commence with embryos at an earlier stage of development (e.g., 4-cell stage zebrafish embryos can be used in contrast to fathead minnow embryos, which can be difficult to obtain before the 32-cell stage). Species differences in stages of development and chorion permeability could play a role in affecting the sensitivities of the different species of embryos to contaminants and in the development of broad guidelines covering a wide range of fish species. Böhler [27] found that the chorion of fathead minnow embryos is thicker and has narrower pores than that of the zebrafish, suggesting that the chorion of fathead minnow embryos is potentially less permeable than the chorion of zebrafish embryos. Scanning

electron microscopy clearly shows visual differences in chorion structure between fish species (Figure 2). It is hypothesized that different species could exhibit different toxicant sensitivities at different stages of embryogenesis, as has been observed for other life stages of fish, although this will be difficult to distinguish from inherent differences in species sensitivities to toxicants.

Concerns regarding the chorion as a barrier and the metabolic capacity of embryos versus larval and juvenile fishes have been raised. Indeed, it is recognized that some very large molecules, such as quaternary ammoniums, do not pass through the chorion. For this reason, the fish embryo toxicity test was extended in duration from 48 h to 96 h to encompass the hatching process and to capture any effects of substances which may not pass through the chorion. Differences in the metabolic capacity of fish embryos could underestimate the toxicity of certain substances, such as allyl alcohol. Conversely, fish embryos may also be more sensitive to certain substances than juvenile fish, so a balanced opinion needs to be considered on the suitability of embryos as alternatives to the fish acute toxicity test. Notably, Belanger et al. [25] reviewed fish embryo toxicity and fish acute toxicity data and found that the relationship between the tests had a slope value near 1.0 with an intercept near the origin, indicating that the results of the tests are closely related. This correlation was similar to those between acute toxicity within different species of fish.

In addition to its use in acute toxicity, the fish embryo toxicity test may be a powerful tool in the development of adverse outcome pathways (AOPs), because approximately 90% of the genome is active during embryogenesis. Furthermore, its use in mode-of-action assessments could also enhance predictions of chronic ecotoxicological effects in fish.

Alternative Approaches to Chronic Ecotoxicity Testing

Chronic toxicity tests, ranging from subchronic effluent tests to full–life cycle assays, use more fish than any other aquatic test, including acute toxicity testing. Therefore, efforts to reduce test frequency or replace chronic tests altogether are welcomed. To date, alternative methods have focused on replacing the use of whole fish in acute toxicity testing, which is an appropriate priority; but now focus should be placed on developing alternatives to chronic ecotoxicity tests. Both embryos and newly hatched fish have been used for subchronic and chronic testing because these life stages have been shown to be quite sensitive. The fish early life stage test (OECD 210 [28]) uses embryos and newly hatched fish for assessing chronic toxicity effects. The early life stages of fish have long been recognized as a robust model for toxicological studies and have been shown to be predictive of chemical effects in full–life cycle tests [3,29]. However, methods for more predictive, shorter, and more reliable testing have been pursued. Environment Canada [30] has a test protocol using salmonids with 3 test options: embryo tests for frequent or routine monitoring, an embryo/alevin test for measuring effects on multiple phases of development, and an embryo/alevin/fry test for definitive investigations.

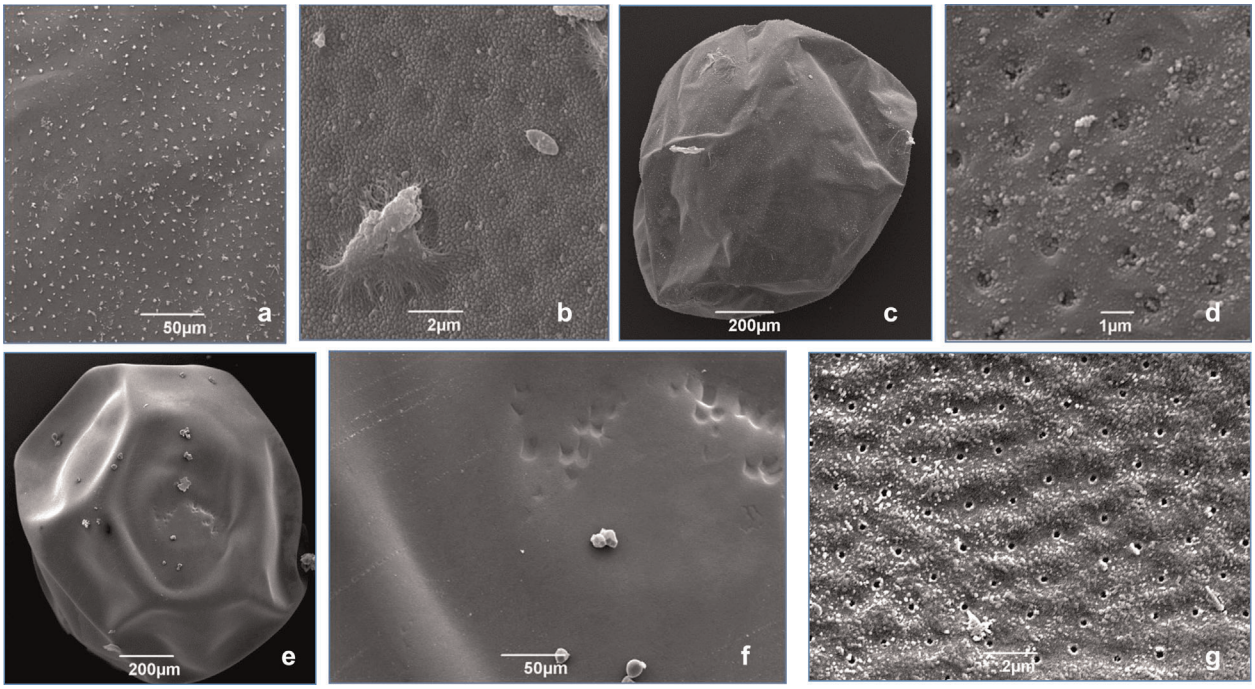

FIGURE 2: Scanning electron micrographs of 4-h-postfertilization zebrafish (a–d) and fathead minnow embryos (e–g). Micrographs (b) and (c) show the external surface of the zebrafish chorion, and (d) is the internal surface where pores are visible. Micrographs (f) and (g) are the external surface of the fathead minnow chorion where pores are visible. Images by A. Lillicrap.

Stadnicka-Michalak et al. [31] have recently shown that tests using a fish cell line, combined with mechanism-based computational models, could replace tests using juvenile fish. This model predicts reduced fish growth based on inhibition of fish cell growth and shows good agreement with measured in vivo data for pesticides in 2 fish species. This promising step toward fish cell lines and models as an alternative to using whole fish for chronic toxicity testing is simple, inexpensive, and rapid, requiring only in vitro data to calibrate the model. Intelligent utilization of existing assays, prioritized and rational testing, and the use of integrated decision schemes may aid in reducing unnecessary tests.

Early attempts to develop alternatives to chronic toxicity testing with fish were aimed at understanding the potential of using existing assays to provide a perspective on chronic toxicity. Extended fish embryo tests have been used in this regard, using acute no-observed-effect concentration statistics to extrapolate to fish early–life stage tests with some success, although the precision of the relationship is not satisfactory. A review of the fish early–life stage test guideline [28] identified a fundamental problem in that the statistical power of the assay was often insufficient to provide robust conclusions [32]. This led to a revision of OECD test guideline 210, requiring more replication per concentration, which unfortunately increases the number of fish used in the procedure. If fish are to be used, judgments about hazard should at least justify the use of animals; and in this case, the power of the test was improved by using additional fish.

Animal Alternatives to Fish Bioaccumulation Testing

Bioaccumulation of chemicals in aquatic organisms is governed by the process of absorption, distribution, metabolism, and excretion. For fish, this can be conceptually represented as a function of rate constants including dietary uptake, fecal egestion, uptake and elimination through the gills, metabolic transformation, and growth dilution (see Figure 3). Each of these different processes can be quantified to some degree, and each depends significantly on the physicochemical properties of a chemical. One such physicochemical characteristic is the log K_{OW}, which indicates the potential of a chemical to partition into lipids. This is a classic criterion for screening chemicals for their potential to bioaccumulate; but because log K_{OW} is not predictive of any of the absorption, distribution, metabolism, and excretion processes involved in bioaccumulation, reliance on log K_{OW} alone can result in overestimations of the bioaccumulation potential of a substance.

There are also QSAR models available for predicting bioaccumulation potential, which include simple models based on the log K_{OW} approach. In addition, Arnot and Gobas [33] developed a bioaccumulation model that includes biotransformation rate estimates for different trophic levels of fish which may be considered to be more realistic than QSAR

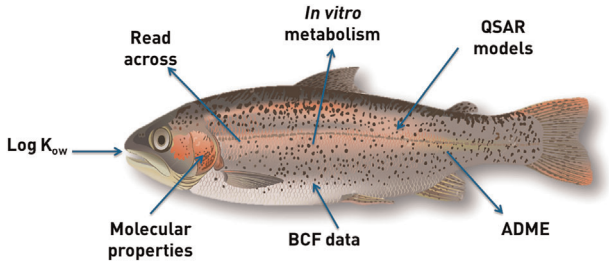

FIGURE 3: Alternative approaches for predicting bioaccumulation in fish adapted from Lillicrap et al. [34] (based on the conceptual model proposed by Arnot and Gobas [33]). The major routes of chemical uptake and elimination in an aquatic organism are a function of the following rate constants: dietary uptake (kD), gill uptake ($k1$), gill elimination ($k2$), metabolic transformation (kM), fecal egestion (kE), and growth dilution (kG). ADME = absorption, distribution, metabolism, and excretion; BCF = bioconcentration factor; K_{OW} = octanol–water partitioning coefficient; QSAR = quantitative structure–activity relationship.

models based on log K_{OW} alone. Similarly, molecular size (and some molecular properties) can also be used as an indicator for reduced bioaccumulation potential. In this case, bioaccumulation potential is based on the assumption that uptake of a chemical through biological membranes is dependent on its molecular size. This is only relevant for some chemicals as it does not take into account active uptake and elimination processes or biotransformation of the chemical to forms that can be eliminated more readily from the body. Metabolism may be predicted using in vitro test systems such as S9 fractions, fish hepatocyte cell cultures, or 3-dimensional spheroidal cultures.

Other more simplistic approaches have recently been accepted at the OECD, including a minimized in vivo approach for assessing the bioaccumulation of chemicals. Ultimately, using all available data together in a weight-of-evidence approach is likely to aid bioaccumulation assessments without the need for full fish bioaccumulation tests. Using tiered assessment strategies, such as the approach suggested by Lillicrap et al. [34], may also further strengthen bioaccumulation assessments in the future.

Alternative Methods and Study Design Refinements for Wildlife Toxicity Tests

Development of new and refinement of existing procedures for assessing effects of potentially toxic chemicals in wildlife (i.e., through testing on amphibians, reptiles, birds, and mammals, which is termed "wildlife toxicology") have received limited attention in comparison to procedures for fish. During the past 20 yr, however, considerable effort has been expended to protect diverse groups of wildlife by developing scaling and extrapolation factors derived from traditional avian and mammalian test species (e.g., scaling factors and indices [35], species sensitivity distributions [36]). While such methods are of value for extrapolating median lethal dose estimates for avian and mammalian test species to

predict potential toxicity of a chemical in other wildlife species, there is considerable uncertainty for predictions that entail large taxonomic differences, particularly for compounds with limited data and for endpoints other than mortality. In some circumstances, it may be appropriate to test more closely related species, particularly when assessing risk for threatened or endangered wildlife.

From a regulatory perspective, some new procedures have greatly reduced the number of individuals traditionally used ($n \approx 40$) in acute toxicity tests. For example, OECD test guideline 223 avian acute oral test [37] can generate a limit dose, a 50% lethal dose, and a slope (for a dose–response curve) with as few as 14 animals. Application of the up-and-down procedure has been used to estimate the 50% lethal dose with 95% confidence intervals in reptiles using fewer than 15 test animals [38]. In addition to such new test procedures, reanalysis of existing data suggests that avian acute testing of formulated plant protection products is largely unnecessary if data are already available for the active substance [39]. In a research setting, computer simulations and creative designs (e.g., use of a washout period or sequential dosing schemes) have also been devised to generate data on desired endpoints using fewer animals than in traditional toxicity studies.

The use of in vitro test systems to predict toxicity and risk of contaminants to wildlife lags well behind methods currently employed for aquatic and mammalian species. Cell culture and molecular assay systems have been used to predict interspecies differences in sensitivity to arylhydrocarbon receptor–active compounds (e.g., Farmahin et al. [40]), but the use of such methods for other groups of compounds has yet to be conducted in wildlife. To the best of our knowledge, in vitro methods to select starting doses for laboratory mammalian toxicology studies have yet to be applied to wildlife toxicity tests. While high-throughput mammalian screening assay systems are being used to identify the activity of thousands of chemicals, such efforts have yet to be undertaken with proteins, cells, or embryos derived from wildlife.

Amphibian Ecotoxicology and Recent Advances

Ecotoxicity tests with amphibians generally use the species *Xenopus laevis*. The first amphibian test to be standardized was a developmental toxicity test: the frog embryo teratogenesis assay *Xenopus* test [41]. Because of the small amount of compound required and the rapidity of the test (96 h), the frog embryo teratogenesis assay *Xenopus* is well suited as a screening assay. The test is conducted on fertilized mid-blastula eggs whose development is followed through organogenesis. The teratogenic potential of a compound is determined after analysis of mortality and malformations. The whole test is performed at embryonic, nonprotected stages of development and is an alternative to the use of mammalian species for the assessment of developmental toxicity.

In addition to the frog embryo teratogenesis assay *Xenopus*, 2 OECD test guidelines using tadpoles have been published over the past 10 yr for the identification of endocrine disruptors. The amphibian metamorphosis assay [42] was developed to screen substances that may interfere with the normal functioning of the thyroid axis. The assay begins with premetamorphic tadpoles, which are exposed for 21 d, by which time metamorphosis should be complete. Measured endpoints include developmental stage, snout–vent length, hind limb length and weight, and histopathology of the thyroid gland. Substances with prothyroid activity will accelerate metamorphosis, and antithyroid compounds will slow metamorphosis and alter thyroid gland structure. The methods and the endpoints of the amphibian metamorphosis assay formed the basis for a more complex and more informative test: the larval amphibian growth and development assay [43]. This assay assesses early development, metamorphosis, survival, growth, and partial reproductive maturation. It also enables measurement of endpoints designed to characterize specific endocrine toxicity modes of actions targeting estrogen, androgen, or thyroid-mediated physiological processes. The larval amphibian growth and development assay serves as a higher-tier test with amphibians for collecting more comprehensive concentration–response information on adverse effects than amphibian metamorphosis assay and frog embryo teratogenesis assay *Xenopus* data for use in hazard identification and characterization as well as ecological risk assessment.

Transgenic Aquatic Models for Ecotoxicological Assessment of Endocrine Disruptors

Several transgenic models with fish and amphibians have recently been developed for use in ecotoxicological testing targeted at endocrine disruption. These tests use embryonic, nonprotected stages of development. The first example is a transgenic medaka strain (ChgH-GFP) harboring the green fluorescence protein (GFP) gene, which enables the activity of the estrogen axis to be visualized in vivo [44]. Recently, a Spiggin-GFP medaka fish line was created to identify disruption of the androgen axis [45]. The Spiggin-GFP medaka embryo model has been shown to be a sensitive, specific, and physiologically pertinent biosensor system for analyzing environmental androgen activity.

Two tests based on aquatic transgenic models are currently undergoing validation at the OECD. The EASZY assay (detection of endocrine active substances, acting though the estrogen receptor, using transgenic cyp19a1b-GFP zebrafish embryos) relies on transgenic embryos expressing GFP under the control of the brain aromatase promoter. This model presents a useful tool to identify compounds acting as estrogen mimics either directly or following metabolic activation within the zebrafish embryo. In addition to this is the *Xenopus* embryonic thyroid signaling assay, which uses transgenic ThβZIP-GFP *Xenopus* embryos. In this model,

GFP expression correlates with the level of expression of the ThβZIP gene, a direct thyroid hormone response gene and a well-characterized transcription factor involved in the control of amphibian metamorphosis [46]. This is a new tool that allows the identification of potential thyroid active chemicals, including those acting on nonreceptor targets or requiring metabolic activation.

AOPs as a Unifying Framework

Conventional fish studies in ecotoxicology have traditionally provided little mechanistic toxicity data as their main focus is on the apical endpoints of survival, growth, and reproduction. Because of these limitations, resource-efficient alternatives to conventional testing, including high-throughput in vitro and in silico screening assays, have been proposed as key components of a future testing paradigm for mechanism-based regulatory toxicology and ecotoxicology [47]. However, the predictive power of molecular or cellular perturbations (modeled or detected in vitro) for apical endpoints relevant to ecological risk assessment must be sufficiently high to minimize uncertainties and provide meaningful data for regulatory decision-making.

Adverse outcome pathways provide a conceptual framework for summarizing knowledge about linkages between key events at various levels of biological organization while providing an adverse outcome relevant to ecological risk assessment [48]. Predictive linkages are implicitly defined when developing AOPs, providing a scientific framework that represents key events at multiple levels of biological levels of organization. The AOP paradigm guides ecotoxicologists through the process of generating linkages that will support the development of new methods and models. Adverse outcome pathways are expected to lead to acceptance of alternative approaches as a credible basis for risk assessment and management. Ultimately, it is envisioned that such alternatives, supported by biologically based extrapolation tools (e.g., toxicokinetic and toxicodynamic models, concentration–duration–response models, species extrapolation models) could gradually replace whole-animal toxicity testing as the primary source of data generation to support ecological risk assessment and environmental decision-making. Funding for research on developing AOPs for environmental science is far less than that received for human safety systems biology (e.g., USEPA ToxCast program). Research is required to enhance the development of AOPs, and currently for regulatory purposes their acceptance is certainly even further off. However, it is clear that the AOP concept could provide significant benefits to environmental risk assessments once they become more commonplace and accepted.

The concept of the traditional 3Rs (reduction, refinement, and replacement) form the backbone of animal alternative approaches. However, the development and use of new alternative approaches for hazard assessments of chemicals and effluents will never be adopted without considering the entire body of the 6Rs as a whole (reduction, replacement, refinement, reproducibility, relevance and regulatory acceptance).

Moving Toward a Future Paradigm for Alternative Approaches in Ecotoxicology

As indicated previously, the fish embryo toxicity test (OECD 236) is currently being considered by various regulatory authorities, including the European Chemicals Agency, as an alternative to fish acute toxicity studies, whereas industry has broadly adopted its use with many contract laboratories now performing the assay. Other potential alternative methods, such as the use of fish gill cell lines for the prediction of fish acute toxicity (e.g., RTgill-W1 [17]), are currently undergoing formal validation processes within the International Organization for Standardization toward adoption as internationally recognized test standards. Before registrants and regulators can have confidence in any alternative approach, however, evidence must be established around the reliability and adequacy as well as the applicability domains and constraints of the approach. Although it is recognized that QSAR models need to be adequately validated, their applicability domain for the test chemical also needs to be verified. Furthermore, a profound understanding of the mode of action of the chemical is essential to be certain that the correct QSAR model is being used [49]. Along these lines, perhaps one of the greatest benefits of evaluating and developing alternative approaches is that it has encouraged ecotoxicologists to find better ways to identify characteristics such as modes of action, which can aid in deciding what types of tests and species are required.

In the interim, better harmonization and mutual acceptance of data from the vertebrate-based studies that continue to be carried out across different global regions could lead to a substantial reduction in the number of animals used in regulatory ecotoxicity testing. These include greater cross-country acceptability of OECD test species and guidelines, particularly in countries with newly developing chemical safety regulations; a move toward extrapolating the toxicity of finished products from active substance data [39]; and reconsideration of key species choices [50]. The shortcomings as well as advantages of these and various other methodologies still need to be critically discussed as supplemental or alternative models to vertebrate testing. Nonetheless, important steps forward are being made in evaluating alternative approaches for basic and applied research for future ecotoxicological tests. However, obtaining regulatory acceptance of these new approaches remains a major challenge. We suggest that scientists continue to promote the advancement of the 6Rs (the traditional 3Rs in combination with the additional 3Rs) through appropriate stewardship and dialogue with regulators so that alternative approaches for generating high-quality ecotoxicity data for environmental risk assessments could become the norm rather than just a possibility.

Acknowledgment

This Focus article is a summary provided by the SETAC Global Animal Alternatives Advisory Group, which acknowledges the rich discussions in several recent SETAC regional annual meetings and World Congresses.

Disclaimer

The views, conclusions, and recommendations expressed in the present article are those of the authors and do not necessarily represent the views or policies of the European Commission or the USEPA. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.

REFERENCES

[1] Russell WMS, Burch RL. 1959. *The Principles of Humane Experimental Technique*. Methuen, London, UK.

[2] European Union. 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official J Eur Union* L276:34–79.

[3] McKim JM. 1977. Evaluation of tests with early life stages of fish for predicting long-term toxicity. *Journal of the Fisheries Research Board of Canada* 34:1148–1154.

[4] Scholz S, Ortmann J, Kluver N, Leonard M. 2014. Extensive review of fish embryo acute toxicities for the prediction of GHS acute systemic toxicity categories. *Regul Toxicol Pharmacol* 69:572–579.

[5] Jeffries MKS, Stultz AE, Smith AW, Stephens DA, Rawlings JM, Belanger SE, Oris JT. 2015. The fish embryo toxicity test as a replacement for the larval growth and survival test: A comparison of test sensitivity and identification of alternative endpoints in zebrafish and fathead minnows. *Environ Toxicol Chem* 34:1369–1381.

[6] European Commission. 2007. Regulation (EC) 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official J Eur Union* L136:3–280.

[7] European Chemicals Agency. 2014. The use of alternatives to testing on animals for the REACH regulation—Second report under Article 117(3) of the REACH regulation. Helsinki, Finland.

[8] Nabholz JV, Miller P, Zeeman M. 1993. Environmental risk assessment of new chemicals under the Toxic Substances Control Act (TSCA) section five. In Landis WG, Hughes JS, Lewis MA, eds, *Environmental Toxicology and Risk Assessment*. ASTM Special Technical Publication 1179. ASTM International, Philadelphia, PA, USA pp 40–55.

[9] European Food Safety Authority Panel on Plant Protection Products and Their Residues. 2013. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters (PPR). *EFSA Journal* 11:3290.

[10] European Union. 2009. Regulation (EC) 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products. *Official J Eur Union* 342:59–209.

[11] Organisation for Economic Co-operation and Development. 2007. OECD guidance document on the validation of (quantitative) structure–activity relationship [(Q)SAR] models. Series on Testing and Assessment, No. 69. ENV/JM/MONO(2007)2. Paris, France.

[12] Tropsha A. 2010. Best practices for QSAR model development, validation, and exploitation. *Mol Inform* 29:476–488.

[13] Belanger SE, Sanderson H, Embry MR, Coady K, Dezwart D, Farr BA, Gutsell S, Halder M, Sternberg R, Wilson P. 2015. It is time to develop ecological thresholds of toxicological concern to assist environmental hazard assessment. *Environ Toxicol Chem* 34:2864–2869.

[14] Bols NC, Dayeh VR, Lee LEJ, Schirmer K. 2005. Use of fish cell lines in the toxicology and ecotoxicology of fish: Piscine cell lines in environmental toxicology. In Mommsen TP, Moon TW, eds, *Biochemistry and Molecular Biology of Fishes*, Vol 6. Elsevier, Amsterdam, The Netherlands, pp 43–84.

[15] Schirmer K. 2006. Proposal to improve vertebrate cell cultures to establish them as substitutes for the regulatory testing of chemicals and effluents using fish. *Toxicology* 224:163–183.

[16] Dayeh VR, Schirmer K, Bols NC. 2002. Applying whole-water samples directly to fish cell cultures in order to evaluate the toxicity of industrial effluent. *Water Res* 36:3727–3738.

[17] Tanneberger K, Knöbel M, Busser FJM, Sinnige TL, Hermens JLM, Schirmer K. 2013. Predicting fish acute toxicity using a fish gill cell line–based toxicity assay. *Environ Sci Technol* 47:1110–1119.

[18] Baron MG, Purcell WM, Jackson SK, Owen SF, Jha AN. 2012. Towards a more representative in vitro method for fish ecotoxicology: Morphological and biochemical characterisation of three-dimensional spheroidal hepatocytes. *Ecotoxicology* 21:2419–2429.

[19] Stott LC, Schnell S, Hogstrand C, Owen SF, Bury NR. 2015. A primary fish gill cell culture model to assess pharmaceutical uptake and efflux: Evidence for passive and facilitated transport. *Aquat Toxicol* 159:127–137.

[20] Brennan LM, Widder MW, Lee LEJ, van der Schalie WH. 2012. Long-term storage and impedance-based water toxicity testing capabilities of fluidic biochips seeded with RTgill-W1 cells. *Toxicol In Vitro* 26:736–745.

[21] International Organization for Standardization. 2007. Water quality—Determination of the acute toxicity of wastewater to zebrafish eggs (*Danio rerio*). ISO 15088:2007(E). Geneva, Switzerland.

[22] Deutsches Institut für Normung. 2001. German standard methods for the examination of water, waste water and sludge—Subanimal testing (group T)—Part 6: Toxicity to fish. Determination of the non-acute-poisonous effect of waste water to fish eggs by dilution limits (T 6). DIN 38415-6. Berlin, Germany.

[23] Organisation for Economic Co-operation and Development. 2013. Test No. 236: Fish embryo toxicity (FET) test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.

[24] Busquet F, Strecker R, Rawlings JM, Belanger SE, Braunbeck T, Carr GJ, Cenijn P, Fochtman P, Gourmelon A, Hubler N, Kleensang A, Knobel M, Kussatz C, Legler J, Lillicrap A, Martinez-Jeronimo F, Polleichtner C, Rzodeczko H, Salinas E, Schneider KE, Scholz S, van den Brandhof E-J, van der Ven LTM, Walter-Rohde S, Weigt S, Witters H, Halder M. 2014. OECD validation study to assess intra- and inter-laboratory reproducibility of the zebrafish embryo toxicity test for acute aquatic toxicity testing. *Regul Toxicol Pharmacol* 69:496–511.

[25] Belanger SE, Rawlings JM, Carr GJ. 2013. Use of fish embryo toxicity tests for the prediction of acute fish toxicity to chemicals. *Environ Toxicol Chem* 32:1768–1783.

[26] Braunbeck T, Boettcher M, Hollert H, Kosmehl T, Lammer E, Leist E, Rudolf M, Seitz N. 2005. Towards an alternative for the acute fish LC50 test in chemical assessment: The fish embryo toxicity test goes multi-species—An update. *ALTEX* 22:87–102.

[27] Böhler S. 2012. The fathead minnow embryo as a model for the development of alternative testing methods in ecotoxicology. PhD thesis. University of Heidelberg, Heidelberg, Germany.

[28] Organisation for Economic Co-operation and Development. 2013. Test No. 210: Fish, early-life stage toxicity test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.

[29] Norberg-King TJ. 1989. An evaluation of the fathead minnow seven-day subchronic test for estimating chronic toxicity. *Environ Toxicol Chem* 8:1075–1089.

[30] Environment Canada. 1998. Biological test method: Toxicity tests using early life stages of salmonid fish (rainbow trout). Ottawa, Ontario.

[31] Stadnicka-Michalak J, Schirmer K, Ashauer, R. 2015. Toxicology across scales: Cell population growth in vitro predicts reduced fish growth. *Sci Adv* 1:e1500302.

[32] Oris JT, Belanger SE, Bailer AJ. 2012. Baseline characteristics and statistical implications for the OECD 210 fish early-life stage chronic toxicity test. *Environ Toxicol Chem* 31:370–376.

[33] Arnot JA, Gobas FAPC. 2004. A food web bioaccumulation model for organic chemicals in aquatic ecosystems. *Environ Toxicol Chem* 23:2343–2355.

[34] Lillicrap A, Springer T, Tyler CR. 2016. A tiered assessment strategy for more effective evaluation of bioaccumulation of chemicals in fish. *Regul Toxicol Pharmacol* 75:20–26.

[35] Mineau P, Collins BT, Baril A. 1996. On the use of scaling factors to improve interspecies extrapolation of acute toxicity in birds. *Regul Toxicol Pharmacol* 24:24–29.

[36] Awkerman JA, Raimondo S, Barron MG. 2008. Development of species sensitivity distributions for wildlife using interspecies toxicity correlation models. *Environ Sci Technol* 42:3447–3452.

[37] Organisation for Economic Co-operation and Development. 2010. Test No. 223: Avian acute oral toxicity test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.

[38] Weir SM, Yu S, Talent LG, Maul JD, Anderson TA, Salice CJ. 2015. Improving reptile ecological risk assessment: Oral and dermal toxicity of pesticides to a common lizard species (*Sceloporus occidentalis*). *Environ Toxicol Chem* 34:1778–1786.

[39] Maynard SK, Edwards P, Wheeler JR. 2014. Saving two birds with one stone: Using active substance avian acute toxicity data to predict formulated plant protection product toxicity. *Environ Toxicol Chem* 33:1578–1583.

[40] Farmahin R, Manning GE, Crump D, Wu D, Mundy LJ, Jones SP, Hahn ME, Karchner SI, Giesy JP, Bursian SJ, Zwiernik MJ, Fredricks TB, Kennedy SW. 2013. Amino acid sequence of the ligand-binding domain of the aryl hydrocarbon receptor 1 predicts sensitivity of wild birds to effects of dioxin-like compounds. *Toxicol Sci* 131:139–152.

[41] ASTM International. 2016. Standard guide for conducting the frog embryo teratogenesis assay-*Xenopus*. E-1439. In *Annual Book of ASTM Standards*, Vol 11.06. Philadelphia, PA, USA.

[42] Organisation for Economic Co-operation and Development. 2009. Test No 231: Amphibian metamorphosis assay. *OECD Guidelines for the Testing of Chemicals*. Paris, France.

[43] Organisation for Economic Co-operation and Development. 2015. Test No 241: The larval amphibian growth and development assay (LAGDA). *OECD Guidelines for the Testing of Chemicals*. Paris, France.

[44] Spirhanzlova P, Leleu M, Sébillot A, Lemkine GF, Iguchi T, Demeneix BA, Tindall AJ. 2016. Oestrogen reporter transgenic medaka for non-invasive evaluation of aromatase activity. *Comp Biochem Physiol C Toxicol Pharmacol* 179:64–71.

[45] Sébillot A, Damdimopoulou P, Ogino Y, Spirhanzlova P, Miyagawa S, Du Pasquier D, Mouatassim N, Iguchi T, Lemkine GF, Demeneix BA, Tindall AJ. 2014. Rapid fluorescent detection of (anti)androgens with spiggin-gfp medaka. *Environ Sci Technol* 48:10919–10928.

[46] Fini JB, Le Mevel S, Turque N, Palmier K, Zalko D, Cravedi JP, Demeneix BA. 2007. An in vivo multiwell-based fluorescent screen for monitoring vertebrate thyroid hormone disruption. *Environ Sci Technol* 41:5908–5914.

[47] Villeneuve DL, Garcia-Reyero N. 2011. Vision & strategy: Predictive ecotoxicology in the 21st century. *Environ Toxicol Chem* 30:1–8.

[48] Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE, Villeneuve DL. 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem* 29:730–741.

[49] Thomas P, Dawick J, Lampi M, Lemaire P, Presow S, Van Egmond R, Arnot JA, Mackay D, Mayer P, Galay Burgos M. 2015. Application of the activity framework for assessing aquatic ecotoxicology data for organic chemicals. *Environ Sci Technol* 49:12289–12296.

[50] Wheeler JR, Maynard SK, Crane M. 2014. Are acute and chronic saltwater fish studies required for plant protection and biocidal product active substance risk assessment? *Environ Toxicol Chem* 33:703–707.