Computational Toxicology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computational Toxicology

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

A workflow for identifying metabolically active chemicals to complement *in vitro* toxicity screening

Jeremy A. Leonard[a], Caroline Stevens[b], Kamel Mansouri[a,c,d], Daniel Chang[e], Harish Pudukodu[f], Sherrie Smith[f], Yu-Mei Tan[f,*]

[a] *Oak Ridge Institute for Science and Education, Oak Ridge, TN, USA*
[b] *National Exposure Research Laboratory, United States Environmental Protection Agency, Athens, GA, USA*
[c] *National Center for Computational Toxicology, United States Environmental Protection Agency, Research Triangle Park, NC, USA*
[d] *ScitoVation LLC, Research Triangle Park, NC, USA*
[e] *Office of Pollution and Prevention of Toxics, United States Environmental Protection Agency, Washington, D.C., USA*
[f] *National Exposure Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, NC, USA*

ARTICLE INFO

Keywords:
Metabolite prediction
QSAR
ER
in vitro screening

ABSTRACT

The new paradigm of toxicity testing approaches involves rapid screening of thousands of chemicals across hundreds of biological targets through use of *in vitro* assays. Such assays may lead to false negatives when the complex metabolic processes that render a chemical bioactive in a living system are unable to be replicated in an *in vitro* environment. In the current study, a workflow is presented for complementing *in vitro* testing results with *in silico* and *in vitro* techniques to identify inactive parents that may produce active metabolites. A case study applying this workflow involved investigating the influence of metabolism for over 1,400 chemicals considered inactive across 18 *in vitro* assays related to the estrogen receptor (ER) pathway. Over 7,500 first-generation and second-generation metabolites were generated for these *in vitro* inactive chemicals using an *in silico* software program. Next, a consensus model comprised of four individual quantitative structure activity relationship (QSAR) models was used to predict ER-binding activity for each of the metabolites. Binding activity was predicted for ~8–10% of metabolites in each generation, with these metabolites linked to 259 *in vitro* inactive parent chemicals. Metabolites were enriched in substructures consisting of alcohol, aromatic, and phenol bonds relative to their inactive parent chemicals, suggesting these features are potentially favorable for ER-binding. The workflow presented here can be used to identify parent chemicals that can be potentially bioactive, to aid confidence in high throughput risk screening.

Introduction

The traditional toxicity testing paradigm involving use of animal studies requires a large investment in time and resources [1,2]. Therefore, the thousands of environmental chemicals currently in commerce [3] and the hundreds of chemicals that are registered on an annual basis [4] render such a low-throughput chemical-by-chemical approach unsustainable. Moreover, extrapolating toxicity results from animals dosed with a specific chemical at high concentrations to relevant human health outcomes arising from much lower concentrations often involves many uncertainties [1,2].

Recognizing these issues, the National Toxicology Program proposed a new "roadmap" for toxicity testing in the 21st century that involved development of rapid profiling strategies to reduce or refine the use of animal studies while remaining scientifically sound and promoting human and animal welfare [5]. High-throughput (HT) *in vitro* screening assays arose from this initiative to act as a means for rapidly investigating the effects of thousands of chemicals across hundreds of biological endpoints linked to disease outcomes relevant to both human and ecosystem health [6,7]. These screening results from HT assays can be used to prioritize chemicals for more extensive *in vivo* testing [8,9].

Abbreviations: AOP, adverse outcome pathway; CERAPP, Collaborative Estrogen Receptor Activity Prediction Project; EDC, endocrine disrupting chemical; ER, estrogen receptor; HT, high-throughput; IUPAC, International Union of Pure and Applied Chemistry; InChI, International Chemical Identifier; MIE, molecular initiating event; NCCT, National Center for Computational Toxicology; OCHEM, Online Chemical Database with Modeling Environment; OECD, Organization for Economic Cooperation and Development; QSAR, quantitative structure activity relationship; SMILES, simplified molecular input line-entry system; UNISTRA, University of Strasbourg, France

* Corresponding author at: 109 T.W. Alexander Dr. Mail Code E205-01, Research Triangle Park, NC 27709, USA.
E-mail addresses: leonard.jeremy@epa.gov (J.A. Leonard), stevens.caroline@epa.gov (C. Stevens), kmansouri@scitovation.com (K. Mansouri), chang.daniel@epa.gov (D. Chang), tan.cecilia@epa.gov (Y.-M. Tan).

http://dx.doi.org/10.1016/j.comtox.2017.10.003
Received 10 July 2017; Received in revised form 14 October 2017; Accepted 18 October 2017
2468-1113/ Published by Elsevier B.V.

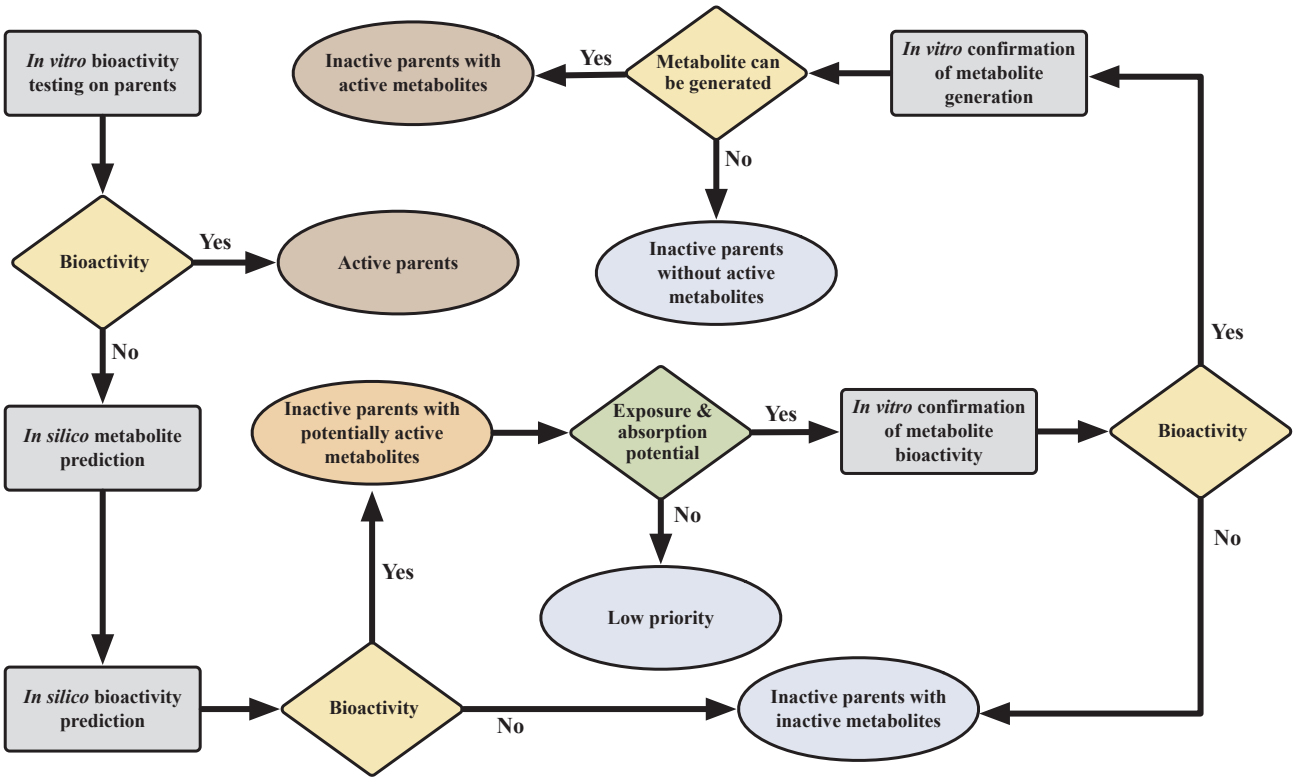

Fig. 1. Workflow for identifying parent chemicals than can become bioactive in a living system.

While *in vitro* assays have significantly expanded the field of toxicity testing, interpreting the results of such studies is not without challenges. False positives may arise for chemicals that are active in *in vitro* assays through means other than direct binding to an intended technological target, or by altering pathways that are not specifically related to that target (e.g., dyes that interfere with an assay relying upon fluorescent signal for detection) [10–12]. A false positive signal may also stem from non-selective cell cytotoxicity, rather than from chemical interaction with the intended enzyme or receptor targets [13]. A high incidence in the number of false positives resulting from *in vitro* testing can lead to an unnecessary waste of time and cost, even though some risk managers may be less concerned with false positives in order to avoid incorrectly labeling some chemicals as safe [14].

Much of the activity caused by endocrine disrupting chemicals (EDCs) can be tested within numerous *in vitro* assays that are part of the US Environmental Protection Agency's Toxicity Forecaster (ToxCast) HT screening program [15]. EDCs, such as estrogenic chemicals, are of concern due to permanent adverse health outcomes that may arise from perturbations in signaling processes during early-life exposures, at a period when proper communication among cells is critical for normal development [16,17]. Adverse outcomes resulting from perturbations on pathways related to estrogen receptor (ER) binding can be investigated for each of the relevant individual ToxCast assays, but interpreting results from multiple assays as a whole may help screen out false positives arising from assay interference or certain characteristics in each assay's technological design. A recent study attempted to reduce such false positives by developing an orthogonal model to integrate results across a total of 18 assays that are related to perturbations on the ER pathway [18].

Of greater concern, however, is the chance that some chemicals may register as negatives in *in vitro* assays while possessing the ability to induce adverse health outcomes under *in vivo* conditions [19]. False negatives, like false positives, can be the result of assay design or interference, as chemicals that are active *in vivo* may have difficulty interacting with their intended *in vitro* technological target [20]. False negatives can also result from pharmacokinetic behaviors that are unable to be replicated in most *in vitro* assays. Specifically, a chemical found inactive in an *in vitro* assay may undergo bioactivation in an *in vivo* environment to a metabolite capable of perturbing an *in vivo* target [21].

Thousands of putative metabolites, as well as their bioactivities and chemical properties or features mediating those bioactivities, can be predicted through use of computational methods in order to avoid the cost, time, and difficulty associated with deriving such information through *in vitro* or *in vivo* means [22,23]. Though numerous existing studies employ cheminformatics and docking approaches to create models for predicting chemical binding activity, only few consider the binding affinities of those chemicals' metabolites as well [24–27]. Here, we propose an approach that integrates *in silico* model predictions of metabolites with quantitative structure activity relationship (QSAR) models predicting activity for a specific biological endpoint, in order to identify *in vitro* inactive parents that may have *in vivo* active metabolites. The utility of this approach is demonstrated in a case study that first predicts metabolites for all chemicals considered inactive in the aforementioned orthogonal model for 18 ER-related assays [18], followed by use of a consensus QSAR model that identifies chemical structural features influencing interactions with the ER and predicts ER binding potential. Metabolites that are predicted to be active are then linked to their respective inactive parents, and these parents may in turn undergo further *in vitro* analyses to confirm generation of their respective active metabolites, followed by evaluation of those metabolites' bioactivities. This workflow is proposed to complement HT *in vitro* toxicity testing methods through its ability to identify chemicals with the potential for bioactivation in a living system, for more accurate prioritization of such chemicals during risk screening and chemical prioritization.