Computational Toxicology 5 (2018) 8–15

Contents lists available at ScienceDirect

Computational Toxicology

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

A new classification algorithm based on mechanisms of action

Franklin J. Bauer[a,b,*], Paul C. Thomas[a], Samuel Y. Fouchard[b], Serge J.M. Neunlist[b]

[a] KREATiS SAS, 23 rue du Creuzat, 38080 L'ISLE D'ABEAU, France
[b] Université de Haute-Alsace, Laboratoire de Chimie Organique et Bioorganique, EA4566, 2 rue des Frères Lumière, F-68093 MULHOUSE CEDEX, France

ARTICLE INFO

Keywords:
Mechanism of action
Classification
QSAR
Ecotoxicity
Toxicity
Quantum calculations

ABSTRACT

A good understanding of Mechanisms of Action (MechoAs) and appropriate methods to determine them is crucial for the accurate prediction of toxicity using *in silico* techniques. Different MechoAs can be related to different QSAR models to predict toxicity values. Therefore, we defined a set of MechoAs, based on Molecular Initiating Events, the first step of Adverse Outcome Pathways (AOPs). In the most common classification algorithms used to predict Modes of toxic Action, the prediction domain is often limited by the need to identify known structural alerts. To circumvent this limitation, we developed a new algorithm to predict MechoAs principally based on mammals and fish, using molecular modelling to obtain calculated molecular parameters. Comparing the Verhaar scheme (as modified by Enoch et al. (2008)) with the MechoA method for the same validation set, MechoA achieved 69% correct classifications as opposed to 45% for the Verhaar scheme, 17% misclassifications for both, 13% classifications slightly different from the literature for our algorithm. No substances fell into zones where two possible MechoAs couldn't be differentiated from each other, while 1% of the molecules were out of the prediction domain of the algorithm as opposed to 38% using the Verhaar scheme. Thus, this model enhances precision of correct AOP identification for *in silico* toxicity predictions.

Introduction

Under current chemical regulations, the use of *in silico* techniques is encouraged and used to save time, cut costs and reduce animal testing [1]. A good understanding of Mechanisms of Action (MechoAs) and appropriate methods to determine them is crucial for the efficient prediction of toxicity using *in silico* techniques, notably QSARs (Quantitative Structure-Activity Relationships) which are increasingly used to predict ecotoxicity, as a regulatory endpoint to replace studies, as can be seen on the ECHA disseminated database [2]. In this paper, the terms Mode of Action (MoA) and MechoA have different meanings. MechoA refers to the molecular interaction that a molecule will undergo, leading to a biological outcome, which can be the key starting point of the Adverse Outcome Pathway (AOP) for this substance, i.e. the Molecular Initiating Event (MIE). Allen and co-workers [3] defined an MIE as "the initial interaction between a molecule and a biomolecule or biosystem that can be causally linked to an outcome via a pathway". Thus, MIE and MechoA are equivalent. However, in this publication, we prefer to use the abbreviation MechoA instead of MIE as we want to underline that these events can serve to classify the toxicity class of molecules, in the same way as MoAs, because they are the basis of how a substance impacts an organism's integrity, at a molecular level. On the other hand, the MoA is not so clearly defined, often referring to the pathological effects that can be seen at the whole organism level in terms of behaviour or death i.e. at the other end of the AOP [4]. In this context, MechoAs and MoAs are both related to AOPs. But an MoA can be the result of very different MechoAs. Conversely, the same MechoA but different MoAs can occur mainly when comparing the AOPs across different species.

Most of the time, separate models are needed for each MechoA or chemical class to obtain acceptable levels of accuracy. However, in the Verhaar classification [5], except for MoAs 1 & 2, MoA categories are more broadly defined so a specific QSAR cannot be allocated to a single MoA. Moreover, the most widely used method to determine MoAs, the Verhaar scheme, does not always correctly allocate chemicals to the appropriate categories despite several updates [6–7] and many common chemical classes still remain "unclassifiable" (MoA5). Another classification scheme was published by Russom et al. [4]. This system is more detailed and could more easily be used to split chemicals into categories that can be predicted with the same QSAR model if it was made publicly available, but the number of categories in this scheme is also far from comprehensive. More recently, another general classification was published by Barron and coworkers [8], with 6 broad MoAs including 31 specific MoAs. A statistical method to predict these MoAs and to generate QSAR models for each MoA was also published by this group [9] but with less specific MoAs than in the publication of 2015.

* Corresponding author at: KREATiS SAS, 23 rue du Creuzat, 38080 L'ISLE D'ABEAU, France.
E-mail address: franklin.bauer@kreatis.eu (F.J. Bauer).

https://doi.org/10.1016/j.comtox.2017.11.001
Received 14 September 2017; Received in revised form 31 October 2017; Accepted 7 November 2017
Available online 08 November 2017

The method is based on machine learning algorithms: linear discriminant analysis (LDA) and random forest (RF). The algorithms select randomly up to 8 from a pool of 970 descriptors to generate several prediction models, each model relates to one MoA. For the LDA algorithm, an MoA is selected if the dedicated model predicts a probability > 50% that the substance acts according to that MoA. For the RF algorithm, an MoA is selected if > 50% of the decision trees predict this MoA. Although, these criteria seem quite unrestrictive, the authors report 84.5–87.7% correct MoA classifications for the validation set. Nevertheless, despite these impressive results, a comparison of Verhaar, Russom and Barron methods by Kienzler and co-workers [10] was less optimistic. This publication revealed that a high proportion of chemicals are unclassified by both Russom and Barron methods. For those compounds which are classified by the 3 schemes, the predicted MoAs were in concordance for only 42% of the molecules. Finally, several methods to predict MoAs or MechoAs are available but have narrow applicability domains [11–15]. For instance, a classification method using MechoAs, proposed by Schultz [15], effectively differentiates MechoAs, using a simple and justifiable descriptor (pKa), but with an applicability domain limited to phenols, with only 2 MechoAs being discriminated (polar narcosis and uncoupling of oxidative phosphorylation). Consequently, there is a need for a classification scheme or algorithm with a wider applicability domain based on precise and relevant category definitions corresponding to well-understood mechanisms of action. A model matching these requirements appears useful, if not essential, to the appropriate use of QSARs [4] leading to a more reliable toxicity prediction. Nevertheless, certain classes of compounds are especially challenging to classify. For instance, non-polar and polar narcosis, two MoAs as defined by Verhaar and co-workers, are still not yet completely understood, even if several hypotheses and likely explanations have been proposed [16]. As another example, the different toxicity mechanisms of phenols are not completely described [17] and the criteria to separate out the categories remain somewhat unclear [18]. Moreover, the classification methods already available apply to one species (or group of species), e.g. to fish [4,9,13], or to aquatic organisms [5,19], or to rodents [20]. In our work, we compiled data from various species, mostly rodents and fish, and thus, we derive prediction MechoAs for these species, which differ significantly in lifestyle and environment (and additional information can be given for other species, but with lesser weight of evidence).

The aim of this work is to predict the MechoA only from the chemical structure, just as previous methods have attempted to do with MoAs [4–5,13]. Firstly, it is reasonable to define chemical categories based on the first molecular event, which is the simplest thing we can predict in the whole AOP if we base our prediction only on the chemical structure. Secondly, it may be easier to construct single QSAR models covering complete categories using MechoAs rather than MoAs. Our new MechoAs definitions were chiefly inspired by the work of Russom and co-workers [4]. In parallel, we developed a new *in silico* method to predict the MechoA of a molecule from its structure using a tiered selection approach. This new method is applicable to a wide range of organic chemicals (hydrocarbons, halides, ethers, esters, ketones, aldehydes, alcohols, phenols, anilines, nitrobenzenes, carbamates, sulphur-containing compounds, organophosphorus compounds, etc.). The performance of the algorithm developed in this work was compared with that of Verhaar's scheme using a dataset taken from the literature.

MechoA definitions

For the purpose of this work we defined six general MechoA classes, which are listed below together with a generic description:

1. Membrane destabilization: accumulation of molecules in cell membranes without specific reaction. Membrane destabilization corresponds to a direct narcotic effect, where the main hypothesis is that cell membrane integrity is impaired by the accumulation of a xenobiotic in the lipid bilayer [21].
2. Hydrolytic prodestabilization: a mixture of both direct accumulation and enzymatic hydrolysis generating membrane destabilizers. Hydrolytic prodestabilization relates to compounds that are not only membrane destabilizers, but further to enzymatic hydrolysis, generate also new destabilizer products. This occurs mainly with esters [22].
3. Reactive toxicity: non-enzymatic reactions with endogenous compounds proteins, DNA). Reactive toxicity occurs when a reactive compound creates adducts with proteins and/or DNA by covalent binding, either directly or through generation of reactive oxygen species [23].
4. Pro-active toxicity: metabolic transformation of the molecule into more toxic compounds. Similarly to hydrolytic prodestabilization, pro-active toxicity is the MechoA of compounds that are first metabolised to become highly toxic compounds, i.e. reactants (MechoA 3), indirect enzyme disruptors (MechoA 5) and direct dockers (MechoA 6).
5. Indirect enzyme disruption: modification of the environment of an enzyme, preventing its normal activity. Enzymes function under specific conditions. Indirect enzyme disruption occurs when these conditions, such as pH, or the relative abundancy of the oxidized and reduced forms of a cofactor, are changed.
6. Direct docking interaction: binding to a docking site of a key protein (enzyme, receptor, ion channel). Direct docking interaction includes all mechanisms that involve a specific interaction with enzymes, receptors or ion channels by binding to a docking site (which can be the active site or an allosteric site). This can result in the inhibition or blocking of the target protein or on the contrary the induction of it.

MechoA 2, which includes an enzymatic hydrolysis, is actually an example of "Metabolic transformation of the molecule" for MechoA 4, but this was separated out from MechoA 4 for two reasons:

First, hydrolysis splits the molecule into two compounds which are more hydrophilic than the parent, decreasing the narcotic effect, which is mostly a detoxification process. However, hydrolysis can generate highly toxic compounds in some cases. This is the case for dinoseb acetate which is rapidly and quantitatively hydrolysed when administered to rats, releasing dinoseb, which is an oxidative phosphorylation uncoupler [24].

Second, MechoA 4 concerns the molecules which are activated by metabolism to give more active compounds than the parent (i.e. with an adverse biological activity). However, some exceptions may occur: metabolic oxidation, not hydrolysis, can also generate a less reactive or less toxic compound as is the case for benzaldehyde which is rapidly oxidized to benzoic acid [25].

These two exceptions, leading to different outcomes, explain the need for subcategorization of the generic MechoA framework described above.

However, even if they share a similar or have the same MIE, the different categories defined in the list above needed to be split further into subgroups as presented in Table 1 such that one single MechoA subgroup relates to one QSAR model per endpoint (e.g. daphnid 48 h EC50 value). The numbering of the subgroups in MechoA 4 are consistent with the MechoA of the product generated by the metabolisation of the parent compound. Thus, MechoA 4.1 refers to compounds that will be transformed into MechoA 1 products. For MechoA 4.3, the end products are reactants (MechoA 3) and so on. The MechoA of RedOx cycling is assigned the number 4.4 as it is metabolised repeatedly in a cyclic way. This mechanism is harmful, because the xenobiotics will metabolise between a reduced and an oxidized form, generating reactive oxygen species (ROS) each time, and depleting NAD(P)H stocks [26].

The designation of these detailed MechoAs is important, because they separate out compounds that need specific QSAR models to predict