Computational Toxicology 7 (2018) 36–45

Contents lists available at ScienceDirect

Computational Toxicology

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

High-accuracy prediction of mechanisms of action using structural alerts

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, Université de Strasbourg, CNRS, LIMA UMR 7042, F-68100 Mulhouse, France

ARTICLE INFO

Keywords:
Mechanisms of toxic action
Classification
Ecotoxicology
Mammalian toxicology
Structural alerts
Decision tree
Modes of Action

ABSTRACT

Knowing the mechanism of action (MechoA) of substances is a crucial first step in determining an Adverse Outcome Pathway and in risk assessment, especially when using *in silico* models to predict (eco)toxicity. We developed a set of structural alerts associated with specific MechoAs based on (eco)toxicity data on hundreds of chemicals and compiled them into a new method to predict MechoAs with high accuracy and with simple rules. The MechoA scheme classifies substances into 6 general MechoAs and 23 detailed MechoA sub-groups. The rules are mainly based on data on mammals and fish. We used a training set of 301 chemicals, and a validation set of 493 molecules. We achieved with this method 91.4% correct classifications for the training set and 92.2% for the validation set. This model is both simpler and performs better than the previous (quantic chemistry based) model we developed and we recommend its use for AOP compilation and for risk assessment. This model will be continuously enhanced with the addition of new rules and minor corrections as they are discovered.

1. Introduction

Mode of action (MoA) is a concept that was defined several decades ago [1] and its underlying mechanisms been discussed by several authors. In 1992, Verhaar and co-workers issued a publication [2] where they gathered existing information on MoAs to generate the first MoA classification scheme for a wide range of organic chemicals. Since then, the Verhaar scheme has been updated in order to better characterise the organic chemical family [3,4]. Despite these updates, this scheme still does not cover many chemical groups and MoAs [5]. Several other prediction methods to predict modes of action already exist and some are listed below. The MoA classification of Russom and co-workers [6] was associated with a method of prediction based on topological rules, followed by the use of MoA-specific QSARs. The final MoA selected is the one for which the QSAR model predicts the lowest LC50. More recently, MOATox, a method to predict MoAs for aquatic organisms has been published [7] classifying compounds into 6 MoA categories and 31 sub-categories, using machine-learning methods. In the human health field, prediction methods have been developed to predict the potential for specific Mechanisms of Action (MechoAs), such as carcinogenicity [8], mutagenicity [9] or acetylcholinesterase inhibition [10]. However, no classification methods for MechoAs, have yet unified these fields such that environmental and mammalian toxicologists can consider Adverse Outcome Pathways (AOP) using the same universal language. Some of these existing decision trees are limited to a specific endpoint, related to a specific MechoA, some others predict different MoAs to cover the main existing toxic pathways, but these methods do not include a number of chemical classes (e.g. esters in the case of the Verhaar scheme) and are therefore incomplete, and finally some methods require complicated machine learning methods, which cannot be easily interpreted.

The aim of the work presented here was to map all possible Mechanisms of Action (MechoAs) in a single scheme. The MechoA is the starting point of the AOP, which describes how the toxicity occurs in a living organism. It appears more logical to predict MechoAs from the structure, rather than MoAs, because MechoAs are related to the first key step of an AOP, while MoA definition may be confusing, rather related to the end of the AOP, based on effects at the whole organism level. The intention was to implement a single, uniform, prediction method and classification that can be used by the ecotoxicological and toxicological community universally, employing the same language and underlying methodology. The intention was also to make the method available as a free user-friendly program that can be run as a software tool on any computer.

Using the classification scheme for MechoAs recently published and the knowledge of MechoAs that was gathered and compiled for hundreds of substances [11], we built structural rules related to each MechoA and found strong relationships between the molecular structure and the MechoA of substances. As a first attempt, we tried to predict the MechoAs from structure using quantum parameters to separate out compounds in groups with distinct electronic properties (molecular orbitals energies, partial charges). We obtained insufficiently accurate

* Corresponding author.
E-mail address: paul.thomas@cehtra.com (P.C. Thomas).

https://doi.org/10.1016/j.comtox.2018.06.004
Received 23 February 2018; Received in revised form 14 June 2018; Accepted 19 June 2018
Available online 20 June 2018

results with this first method, obtaining 69% correct classifications for the validation set. Moreover, the method was quite time-consuming to run, requiring geometry optimization with the Restricted Hartree-Fock method and the basis set 6-31G(d,p). In order to achieve satisfying results with a much lower level of CPU utilization, we developed this new method with simple structural alerts, not needing any quantum parameters to be calculated.

The MechoA definitions that we use, presented in this previous publication [11] are based on 6 general MechoAs divided into 23 MechoA sub-categories. These classes are reproduced below now including a new subgroup that we recently uncovered (MechoA 1.3), and renaming class 2, formerly called “hydrolytic pro-destabilisation”, and now named “enzymatic hydrolysis”. The previous name was considered misleading, as the membrane destabilisation effect occurs mainly via the parent compound, prior to hydrolysis. Enzymatic hydrolysis will cause other toxic effects which may be more or less severe depending on the degradation product (narcotic or non-narcotic).

General MechoAs together with a generic description:

1. Membrane destabilization: accumulation of molecules in cell membranes without specific reaction
2. Enzymatic hydrolysis: a mixture of both direct accumulation and enzymatic hydrolysis generating acidity
3. Reactive toxicity: spontaneous non-enzymatic reactions with endogenous compounds (proteins, DNA)
4. Pro-active toxicity: metabolic transformation of the molecule into biologically active compounds (e.g. into reactive compounds, or into inhibitors of enzymes)
5. Indirect enzyme disruption: modification of the environment of an enzyme, preventing its normal activity
6. Direct docking interaction: binding to a docking site of a key protein (enzyme, receptor, ion channel).

Detailed MechoAs:

2. Materials and methods

In order to prepare a scheme dependent on specific classification rules, the first step was to compile a database of MechoAs for a large number of substances, and to identify the relationship between the structures of chemicals and the MechoA for these substances. All this work has been based on an in-depth literature search for each group of compounds, without using automated tools to gather data or rank them or identify structural alerts. Depending on the chemicals and the MechoAs that were investigated, the published data available to identify the MechoAs in different species were not of the same nature. For some compounds, we could find publications where the mechanism by which toxicity occurs was already well identified and reported for different organisms, e.g. for quinones [12]. For other compounds, differences in toxicological effects between 2 similar chemicals were observed which could be related to a difference in their structure, related for instance to a different reactivity or docking ability for a given receptor. An extensive review of each MechoA and structural alert associated with the references used to build it is included in the thesis by Bauer [13]. The training set used to develop structural alerts and the subsequent decision tree is composed of 301 molecules, which come from the complete dataset (both the training and validation sets) of our first method and can be found in the supporting information file of the corresponding publication [11]. The external validation set for the method presented here contains 487 molecules and is composed of molecules not used in the training set, taken from the internal database of KREATiS, and for which acute fish LC50 data are available. This was supplemented with substances with specific MechoAs that were not present in the KREATiS database at the time. The data for the validation set are presented in the supporting information file associated with the present paper. In both training and validation sets, the MechoA information was taken from various sources in the published literature, among which [3,9,14,15]. The reference for each MechoA assignment is also given in the tables in supporting information files. Both the training and the validation sets include members of these chemical families (but this list is not exhaustive): alkanes, alkenes, aromatics,

Table 1
MechoA subgroups.

MechoA	Subgroup	Mechanism detail
1: membrane destabilization	1.1 Nonpolar narcotics	Physical perturbation of membrane integrity with different affinities for non-polar and polar chemicals, and with stronger interaction for positively charged molecules.
	1.2 polar narcotics	
	1.3 cationic narcotics	
2: enzymatic hydrolysis	2.1 hydrolysis to destabilizers	Direct membrane destabilization + Enzymatic hydrolysis, giving:
	2.2 hydrolysis to actives	2.1 only narcotic products
		2.2 at least one non-narcotic product
3: reactive toxicity	3.1 hard electrophiles	3.1 adduct formation with amino and thiol protein residues
	3.2 soft electrophiles	3.2 adduct formation through Michaël addition with thiol protein residues
	3.3 spontaneous radical-generating compounds	3.3 homolytic cleavage of a weak bond generating 2 radicals, generating oxidative stress
4: pro-active toxicity	4.1 readily detoxified compounds	4.1 the metabolised compound has a MechoA 1
	4.3 pro-reactants	4.3 the metabolised compound has a MechoA 3
	4.4 RedOx cyclers	4.4 compounds undergoing RedOx cycling, generating oxidative stress
	4.5 indirect pro-disruptors	4.5 the metabolised compound has a MechoA 5
	4.6 direct pro-dockers	4.6 the metabolised compound has a MechoA 6
5: indirect enzyme disruption	5.1 Oxidative Phosphorylation Uncouplers	5.1 cyclic transport of H + across the mitochondria inner membrane, preventing the generation of ATP
	5.2 acids and bases	5.2 diffusion across cell membranes in neutral form, and then release/retrieval of H + in cytosol
	5.3 other mechanisms	5.3 e.g. consumption of NADH pool, scavenging of Ca^{2+} ions, *etc.*
06:00	6.1 AChE inhibitors	6.1 inhibition of Acetylcholinesterase
Direct docking site interaction	6.2 AChR binders	6.2 binding to nicotinic or muscarinic Acetylcholine Receptors, as an agonist or antagonist
	6.3 dopamine transport disruptors	6.3 induction of the transport of dopamine into the synapse and inhibition of its recycling back into the nerve cell
	6.4 metal chelators	6.4 chelation of metal ions, such as Fe^{2+}, thus blocking or degrading the enzymes containing this metal ion in their active site
	6.5 Photosystem II ET inhibitors	6.5 inhibition of a protein involved in the electron transport chain in the photosystem II of plant cells
	6.6 ion channel modulators	6.6 docking to an ion channel’s binding site causing its activation or deactivation
	6.7 other mechanisms	6.7 e.g. binding to glycine receptors, inhibition of tubulin polymerisation, *etc.*