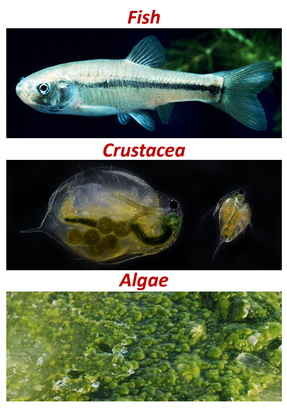
Development of an *in Silico* Structural Profiler facilitating Mechanistically-grounded Classification of Aquatic Toxicants

James W. Firman[1], Franklin J. Bauer[2], Geoff Hodges[3], Bruno Campos[3], Jayne Roberts[3], Steve Gutsell[3], Paul C. Thomas[2], Mark Bonnell[4] and Mark T.D. Cronin[1]

1. School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, England
2. KREATiS SAS, L'Isle d'Abeau, France
3. Safety and Environmental Assurance Centre, Unilever, Sharnbrook, England
4. Science and Risk Assessment Directorate, Environment & Climate Change Canada, Gatineau, Canada

Introduction



- **The environmental risk assessment of chemicals (in an aquatic context) is a process often impeded by a general shortage of suitable experimental data.**
- **With ~350,000 substances registered for use globally, alternative methods are required enabling ready, rational identification of intrinsically hazardous compounds.**
- ***In silico* (computational toxicology) approaches, such as the chemical classification and mode of toxic action assignment schemes devised by Verhaar *et al.* [1] and Russom *et al.* [2], have gained prominence as practical, cost-effective means through which these aims might be achieved.**
- **Such "first generation" profilers are, by now, decades old – holding restricted coverage of chemical space and taxa, and offering limited mechanistic resolution.**
- **We have, accordingly, developed an updated, "second generation" system: expanding breadth both of chemistry and of the species considered (spanning fish, crustacea and algae), whilst grounding conclusions in the context of the molecular initiating event (MIE) framework [3,4].**
- **The form and structure of this novel scheme, alongside its implementation as a practical screening tool (within KNIME software), are each discussed beneath.**
- **Further consideration is given to its future enhancement, including proposed integration alongside the related KREATiS MechoA rule-set [5,6].**

Construction and form of classification scheme

Origins and definitive organisational features

The collection of information relating to mechanisms of aquatic toxicity is described by Sapounidou *et al*. [3].

Scheme organisation draws inspiration from "first generation" profilers, with focus extended to provide further considered of mechanism at the level of the MIE (taxonomical coverage additionally expanded).

Five tiers of characterisation are incorporated, as described below and summarised within Fig. 1:

Tier 1: Activity domain

Overarching categorisation of principal chemical interaction

- **NARCOSIS:** *Non-specific, solubility-associated disruption to structural integrity of cellular membranes*
- **REACTIVE:** *Intrinsic, general chemical reactivity towards biomolecules (either pre- or post-metabolism)*
- **SPECIFIC:** *Selective in scope, arising through defined enzyme- or biological receptor-mediated effect*

Tiers 2 and 3: Mechanistic grouping

Offering, as appropriate, progressively enhanced degrees of resolution at the level of the biological "key event"

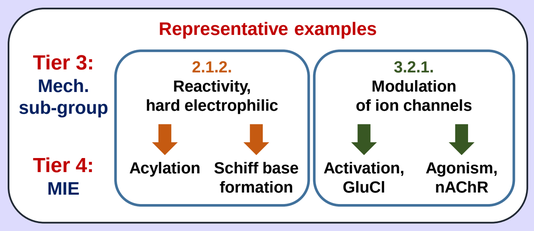
Tier 2: Mechanistic group *(10 classes)*

Divided further…

Tier 3: Mechanistic sub-group *(25 classes)*

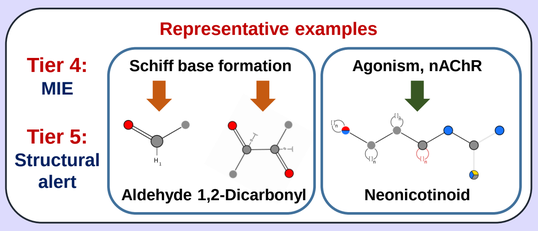
Tier 4: Molecular initiating event

Fundamental interaction between chemical and biological system (60 classes), in turn serving to trigger a key event



Tier 5: Structural alerts

Expert-defined structural features identifying molecules liable to induce a given MIE (183 distinct rules)



Taxonomical applicability

Assigned to accompany each MIE: 53 categories, from species-level upwards (across fish, crustacea, algae)

e.g., **1.1.1. Non-polar narcosis : Accumulation in membrane-based phospholipids : All taxa and species**
3.3.7. Protein biosynthesis disruption : Inhibition of chitin synthase : *Daphnia magna*

Scheme composition

Tier 1 Domain	Tier 2 Mechanistic group	Tier 3 Mechanistic sub-group	Tier 4 MIE	Tier 5 S. alert
1. Narcosis	1.1. Non-polar narcosis	1.1.1. Non-polar	1	6
	1.2. Enhanced narcosis	1.2.1. Polar	1	13
		1.2.2. Alkyl amine	1	1
		1.2.3. Carboxylic acid ester	1	1
2. Reactive	2.1. Electrophilic	2.1.1. Soft	3	32
		2.1.2. Hard	7	16
		2.1.3. Pre-reactive (electrophilic)	5	26
	2.2. Nucleophilic	2.2.1. Nucleophilic	1	0
	2.3. Free radical generation	2.3.1. Radical damage of tissues	1	1
		2.3.2. Redox cycling	1	9
		2.3.3. Pre-reactive (free radical generation)	1	2
3. Specific	3.1. Enzyme inhibition	3.1.1. Acetylcholinesterase inhibition	1	2
		3.1.2. Photosynthesis inhibition	3	8
	3.2. Ion channel modulation	3.2.1. Modulation of ion channels	8	13
	3.3. Cellular function disruption	3.3.1. Amino acid biosynthesis disruption	3	6
		3.3.2. Cell structure disruption	1	1
		3.3.3. Fatty acid biosynthesis disruption	3	8
		3.3.4. Nucleic acid biosynthesis disruption	2	2
		3.3.5. Steroid biosynthesis disruption	2	5
		3.3.6. Carotenoid biosynthesis disruption	3	5
		3.3.7. Protein biosynthesis disruption	1	2
		3.3.8. Developmental disruption	4	9
	3.4. Mitochondrial disruption	3.4.1. Electron transport inhib. (specific)	3	6
		3.4.2. Electron transport inhib. (non-specific)	1	2
	3.5. Nuclear receptor modulation	3.5.1. Modulation of nuclear receptors	2	7

Figure 1

Implementation of scheme as structural profiling tool

Coding and operation (KNIME workflow)

Structural alerts were encoded as SMARTS strings and integrated into a workflow within KNIME software (www.knime.com) [4].

This tools allows for ready batch processing of substances (presented within the SMILES format).

Form of workflow is illustrated within Fig. 2. Compounds may identify with multiple putative toxic mechanisms.

This tool is freely available for download (link in Fig. 2), and may be used without restriction for screening purposes.

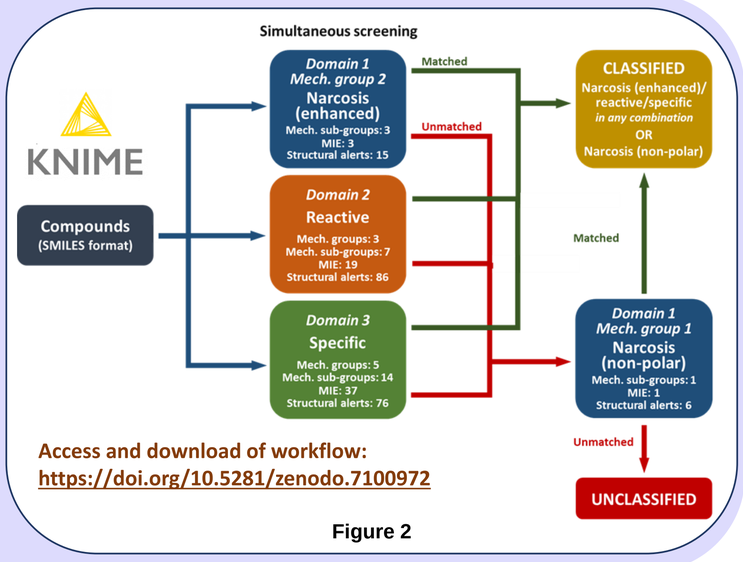


Figure 2

Illustrative screening application

Fig. 3 provides an illustration of the profiling tool in practical use.

Depicted is passage of five representative substances (A-E), each triggering a distinct structural alert aligned to MIE underlying the disruption of amino acid biosynthesis (3.3.1).

Further provided is an example of corresponding workflow output.
For each substance is listed:

- **The title(s) of the specific structural alert(s) against which it is matched.**
- **An overview of the MIE to which each alert corresponds.**
- **The biological site at which the MIE occurs.**
- **Identities of the aligning Tier 1, 2 and 3 classifications.**
- **A generalised descriptive summary of the postulated mechanism.**
- **The accompanying range of taxonomical applicability.**

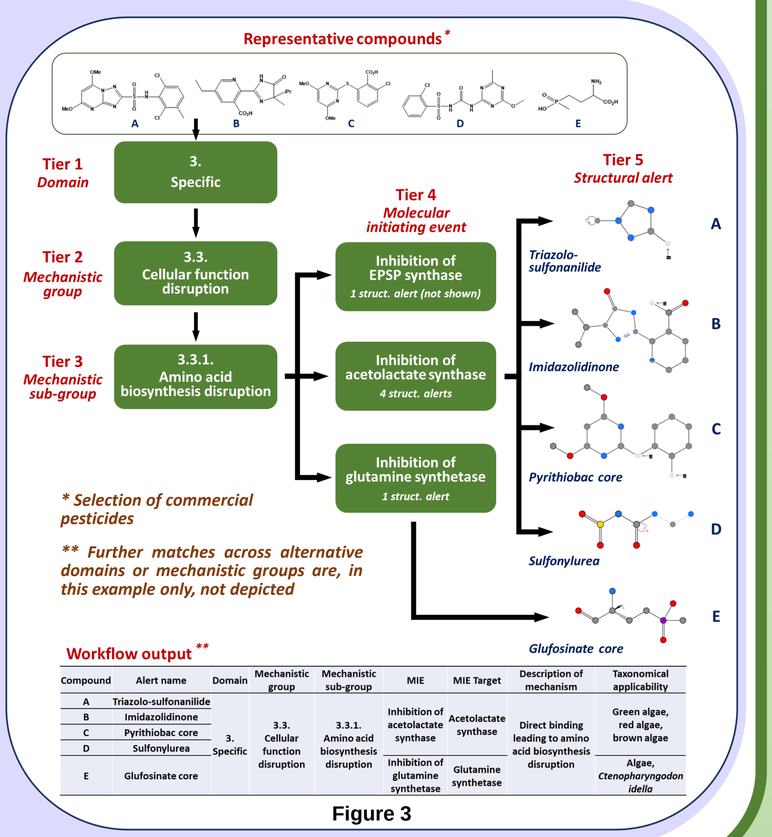


Figure 3

Future development: Expansion and integration

- **A full analysis of the chemical-space coverage offered by the profiler, in its current iteration, is presented within poster 1.02.P-Mo015.**
- **Whilst superior to the "first generation" schemes (those of Verhaar and Russom) when considering the mechanistic characterisation of both reactive and specific domains, shortcomings are noted in relation to identification of narcotics.**
- **Future updates will refine narcosis rules, and see further expansion across the "specific" domain (i.e., extended incorporation of pharmaceuticals, endocrine-disrupting chemicals etc).**
- **Collaborative integration of this rule-set, alongside that underlying the KREATiS MechoA scheme [5,6], remains ongoing (please refer to oral presentation 1.02.T-01 for further detail).**
- **The resultant profiling tool, titled "MechoA+", shall be included within future releases of the OECD QSAR Toolbox software.**

References

1. Verhaar HJM *et al.* 1992. *Chemosphere*. 25, p. 471-491. DOI: 10.1016/0045-6535(92)90280-5
2. Russom CL *et al.* 1997. *Environ. Toxicol. Chem.* 16, p. 948-967. DOI: 10.1002/etc.5620160514
3. Sapounidou M *et al.* 2021. *Environ. Sci. Technol.* 55, p. 1897-1907. DOI: 10.1021/acs.est.0c06551
4. Firman JW *et al.* 2022. *Environ. Sci. Technol.* 56, p. 17805–17814. DOI: 10.1021/acs.est.2c03736
5. Bauer FJ *et al*. 2018. *Comput. Toxicol.* 5, p. 8-15. DOI: 10.1016/j.comtox.2017.11.001
6. Bauer FJ *et al*. 2018. *Comput. Toxicol.* 7, p. 36-45. DOI: 10.1016/j.comtox.2018.06.004

Corresponding author → James Firman: J.W.Firman@ljmu.ac.uk