Comparison of the Performance and Domains of *In Silico* Schemes to Classify Environmental Chemicals

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

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 and Aims

- ***In silico*** **classification schemes are commonly applied to assist in grouping of chemicals ahead of read-across and the application of quantitative structure-activity relationships (QSARs).**
- **These tools, which can be considered to be *in silico* New Approach Methodologies (NAMs), assist greatly in data gap filling for regulatory and other purposes.**
- **A new scheme has been developed from the findings of Sapounidou et al. [1] and its implementation is described in detail by Firman et al. [2] and poster 1.02.P-Mo014. Fig 1 shows the levels of detail captured in the new scheme.**

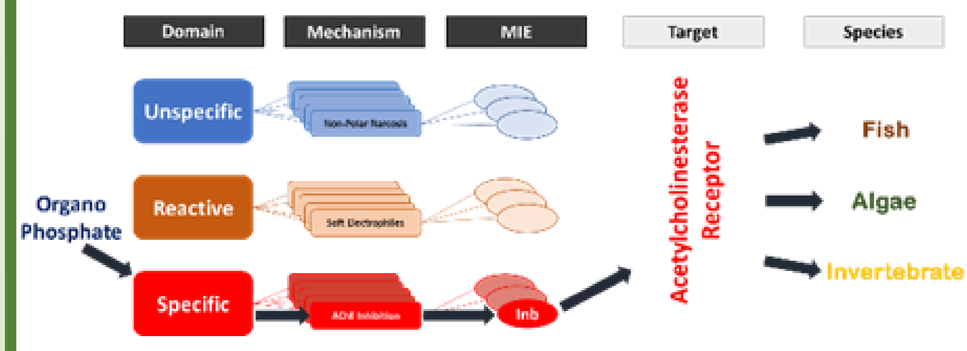

Figure 1. General structure of the Sapounidou *et al.* [1] scheme, showing domains and related information. Illustrated with reference to a compound with an organophosphate alert

- **The aim of this investigation was to compare the performance and applicability domains of four classification schemes.**

Methods

Classification Schemes Evaluated

Four classification schemes were evaluated:

- **Verhaar *et al.* [3] implemented in the OECD QSAR Toolbox ver 4.4.1**
- **Russom *et al.* [4] implemented in Chemprop ver 7.1.0**
- **MechoA ver 2.2 [5,6] implemented in iSafeRat© Desktop ver 2.1.0**
- **Sapounidou *et al.* [1] implemented as described below**

Implementation of the Sapounidou *et al.* [1] Scheme

- **The novel classification scheme reported by Sapounidou *et al.* [1] was described by 183 structural alerts, as reported in Firman *et al.* [2].**
- **Structural alerts are related to a molecular initiating event (MIE) and defined against existing mechanistic knowledge (into a narcotic, reactive or specific mode of action).**
- **Structural alerts were captured as SMARTS strings to create a profiler.**
- **The profiler is publicly available in a KNIME Workflow [7], see Fig 2.**

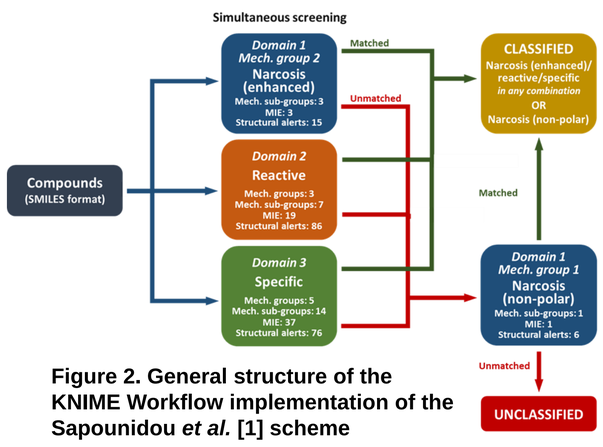

Figure 2. General structure of the KNIME Workflow implementation of the Sapounidou *et al.* [1] scheme

Screening of Compounds

- ***All Schemes*** **- Over 5,500 single chemical substances, derived from priority pollutants and other substances, were screened using the four classification schemes.**
- ***Sapounidou Scheme*** **– A more detailed analysis of the Sapounidou *et al.* [1] scheme was undertaken with over 70,000 structures from REACH, pharmaceuticals, cosmetics, pesticide and biocides inventories.**
- **In order to perform a meaningful comparison of classifications, the output from the schemes was assigned to one of the following classifications (domains in Fig 1):**
  - **Narcosis e.g. non-specific, reversible, Verhaar Classes 1 & 2**
  - **Reactive e.g. non-specific (electrophilic) reactivity, Verhaar Class 3**
  - **Specific e.g. receptor-binding, Verhaar 4**
  - **Non classified – compounds that could not be classified into one of the three classes above**
- **In all screening activities, the following information was captured:**
  - **Whether the compound was classification to a mechanism / mode or not (the latter analogous to Verhaar Class 5)**
  - **The general mechanistic domain e.g. Classes 1-4 in Verhaar / domain in Fig 1.**

References

1. Sapounidou M *et al.* 2021. *Environ. Sci. Technol.* 55, p. 1897-1907. DOI: 10.1021/acs.est.0c06551
2. Firman JW *et al.* 2022. *Environ. Sci. Technol.* 56, p. 17805–17814. DOI: 10.1021/acs.est.2c03736
3. Verhaar HJM *et al.* 1992. *Chemosphere.* 25, p. 471-491. DOI: 10.1016/0045-6535(92)90280-5
4. Russom CL *et al.* 1997. *Environ. Toxicol. Chem.* 16, p. 948-967. DOI: 10.1002/etc.5620160514
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
7. James Firman. (2022). Sapounidou mechanistic profiler. Zenodo. https://doi.org/10.5281/zenodo.7100972

Results

- **All schemes are publicly available and provide a transparent means of classifying single chemicals up to large chemical inventories.**
- **The use of the classification schemes, most notably Verhaar and Russom, is well established for grouping (and read-across) and for allocation to QSARs.**
- **Fig 3 shows the classifications obtained by the four schemes for over 5,500 compounds:**
  - **Classification rates varied: MechoA classified over 90% of compounds, Verhaar fewer than 45%.**
  - **Only the Chemprop implementation of the Russom considers applicability domain, with only 13% of compounds in domain.**
  - **As well as extending chemical coverage, the MechoA and Sapounidou schemes have greater taxonomic coverage.**

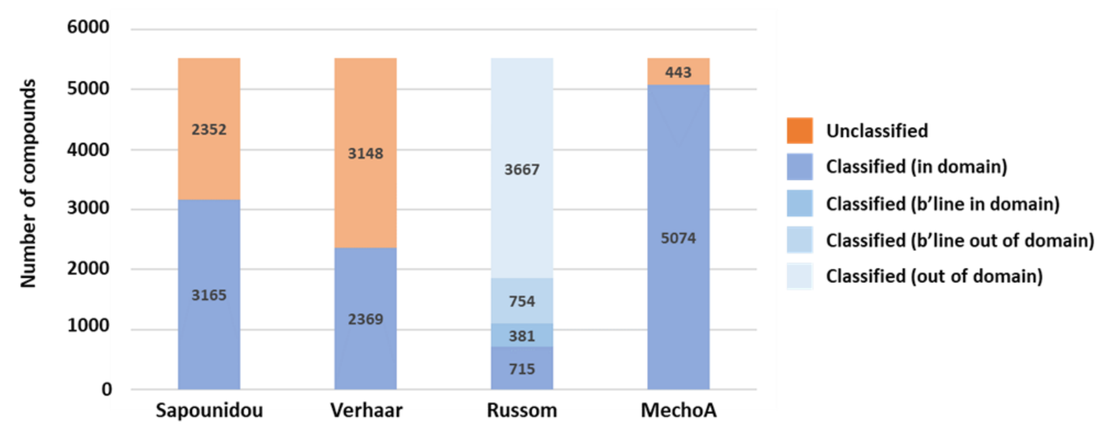

Figure 3. Number of compounds within screening inventory either classified or unclassified, according to Sapounidou, Verhaar, Russom and MechoA schemes.

- **Fig 4 shows the classification to mechanistic domains of the 5,500 compounds:**
  - **The Sapounidou scheme classifies more compounds as reactive, MechoA more as specific.**
  - **The Sapounidou scheme provides direct relation to the MIE.**

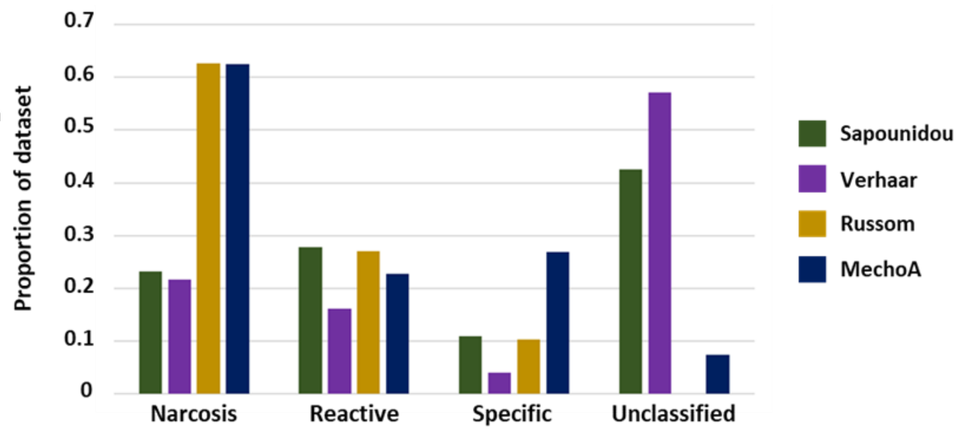

Figure 4. Comparison of domain-level mechanistic classifications from each of the four schemes.

- **Fig 5 shows the classification of over 75,000 compounds by the Sapounidou scheme:**
  - **The pesticides have the most specifically-acting compounds, expected as pesticide MIEs were the basis of the specific domain.**
  - **The pharmaceuticals have the greatest number of unclassified compounds.**

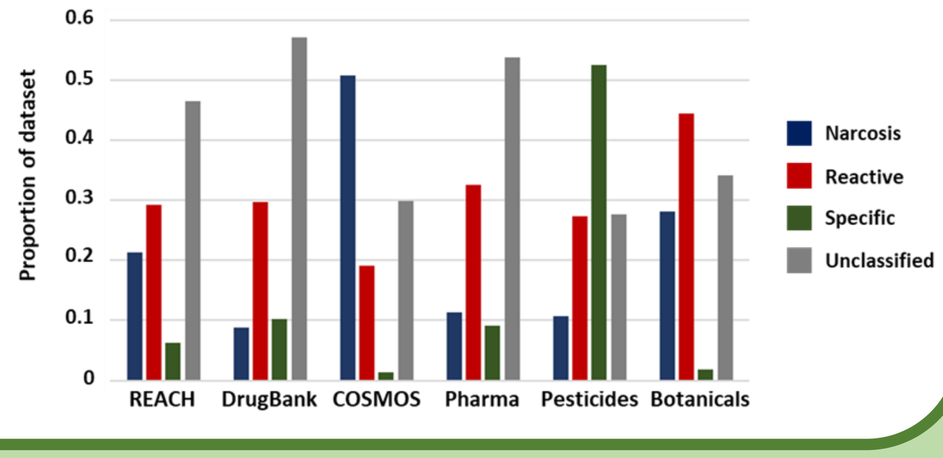

Figure 5. Analysis of the extent of coverage offered by Sapounidou scheme across representative compound inventories, presented at the domain level.

Conclusions

- **The Sapounidou *et al*. [1] classification scheme has been successfully implemented as SMARTS strings into a KNIME workflow.**
- **MechoA has the greatest coverage and classifies the greatest proportion of compounds to a definitive mechanism of action.**
- **The Sapounidou *et al*. scheme has more detail than other schemes for reactive mechanisms; MechoA for specific mechanisms.**
- **There is potential to combine and optimise the MechoA and Sapounidou *et al.* schemes.**
- **Areas where more information is required in the classification schemes include:**
  - **Specific modes of action of pharmaceuticals.**
  - **Greater granularity and distinction of narcotic mechanisms of action.**
  - **Endocrine disruption.**

Further information → Mark Cronin: m.t.cronin@ljmu.ac.uk