The Molecular Initiating Event (MIE) is the initial interaction between a chemical and the biological system resulting in a cascade of biological events to an adverse outcome [1]. Mechanistic knowledge relates specific structural features of a chemical to its ability to bind covalently to proteins/DNA, key MIEs for endpoints such as skin/respiratory sensitisation, genotoxicity and organ toxicity, e.g., hepatotoxicity. Based on organic reaction mechanisms, previous work had defined structural fragments predicting covalent binding to proteins [2] and DNA [3,4]. They have been shown to be useful for chemical grouping.

Applications and Category Formation

- **Chemotypes** can be used to profile data sets to identify similar substances.
- The advantage of the possible coding of different properties is extension of that similarity beyond structural information.
- Groupings based on chemotypes can also be sub-categorised on the basis of ADME properties, e.g., oral bioavailability or skin permeation.
- Thus, more robust and meaningful categories, adapted to the particular assessment purpose, can be formed and used for read-across of toxicological information.

Mechanistic domains for potential protein and DNA binding identified in the COSMOS Cosmetics Inventory and by REFINE, forming a common definition of mechanistic categories which can be further refined. An example of further sub-categorizations, percentages of orally available substances per mechanistic domain and structural alert type are shown [5], along with guidance on exposure and bioavailability. The functional consideration of ADME properties allows for the prioritisation of substances for further safety assessment.

Structural Alerts

- Structural alert = fragment of a molecule identified as reacting with a biological system, i.e. the considered MIE, e.g. electrophile reacting with nucleophile (as S in cysteine, N in lysine).
- 108 alerts for protein binding and 111 alerts for DNA binding [1,4]. These alerts are compiled into profilers and have been included in the OECD QSAR Toolbox (www.qarbox.org), a software for grouping chemicals and supporting data filling for hazard assessment, and also implemented as KNIME workflows (www.knime.org).
- Generally, structural fragments are defined using Daylight SMARTS patterns.

Chemotypes

- The structural fragments coded in SMARTS have been refined to chemotypes, a more flexible format (CSRML – Chemical Structure Related Compound Lists) and refined by ADME considerations and physico-chemical parameters and constraints.
- Chemotypes can be defined in CSRML including a descriptor computing steric accessibility, so that sterically hindered ketones are excluded.

Conclusions

- Thus, refined screening of sets of chemicals is possible, as well as improved formulation of robust categories suitable for the application of read-across for in silico toxicity prediction and support in chemical safety assessment.
- The chemotypes are being made available through ToxPrint.org for use with the freely available ChimeraTyper software.

Cheminformatics Approaches to Tailor *In Silico* Profilers for Refined Category Formation to Support Chemical Safety Assessment

Andrea-Nicole Richarz\(^1\), Bruno Bienfait\(^2\), Steven J Ennoch\(^3\), Christof H Schwab\(^2\), Mark TD Cronin\(^1\), Chihae Yang\(^1,3\)

\(^1\) School of Pharmacy and Chemistry, Liverpool John Moores University, England; \(^2\) Molecular Networks GmbH, Erlangen, Germany; \(^3\) Altamira LLC, Columbus, OH, United States

Abstract # 1318

Safety and risk assessment of chemicals is increasingly supported by cheminformatics methods. In particular, the grouping of chemicals into categories allows for data gaps to be filled via read-across. In order to form meaningful categories, it is crucial to define the chemical and biological similarity robustly. The compounds are grouped together, e.g., by structural/mechanistic similarity, taking into account the "context" of the grouping and biological relevance. Consideration of mechanistic information is important, where grouping is done by Adverse Outcome Pathways. The aim of this work was to develop profilers to support chemical grouping by means of adapted "chemotypes", which can be tailored flexibly to the problem investigated and to apply them to inventories of chemicals.

**Chemotypes** can be used to profile data sets to identify similar substances. The advantage of the possible coding of different properties is extension of that similarity beyond structural information. Groupings based on chemotypes can also be sub-categorised on the basis of ADME properties, e.g., oral bioavailability or skin permeation. Thus, more robust and meaningful categories, adapted to the particular assessment purpose, can be formed and used for read-across of toxicological information.

Chemists can use the freely available ChemType software (https://chemotyper.org) to screen and profile chemicals in datasets and thus support in bespoke category formation. Chemotypes are defined by mechanistic considerations and refined by physico-chemical parameters and constraints.

**Chemotypes** allow for the "tailoring" of profilers to particular applications, including not only sub-structure information, but also physico-chemical properties and other constraints.

The chemotypes are being made available through ToxPrint.org for use with the freely available ChimeraTyper software.