OLIMPIC - Overcoming current Limitations In Metabolism Prediction of Industrial Chemicals

Molecular Networks
Inspiring Chemical Discovery

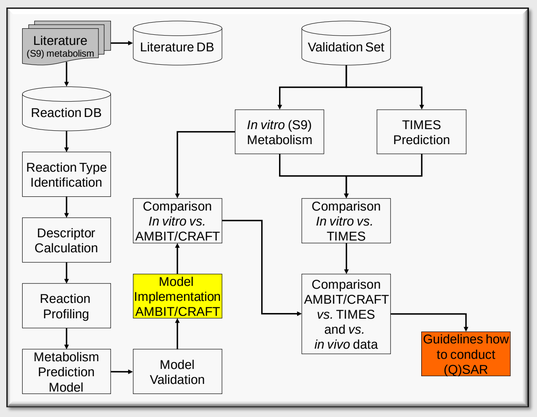
L. Terfloth,[*] A. Tarkhov,[*] J. Gasteiger,[*] M. Bausen,[§] E. Fabian,[§] R. Finking[§], T. Bernshausen[#]
[*]Molecular Networks GmbH, Henkestr. 91, 91052 Erlangen, Germany. http://www.molecular-networks.com
[§]BASF SE, 67056 Ludwigshafen, Germany [#]Archeiron, 65620 Waldbrunn, Germany

Introduction

Strategic combinations and tiered application of alternative testing methods to replace or minimize the use of animal models in risk assessment are attracting much attention. The aim of the CEFIC-LRI funded OLIMPIC project is to provide improvements in the prediction of metabolism of industrial chemicals in mammals. The holistic investigation will consider *in vivo*, *in vitro* and *in silico* data in order to identify discrepancies between the different methods and to enhance the predictive power of *in silico* models for metabolism. In this poster we will report on the compilation of a data set on the metabolism of chemicals in rats, the measurements of S9 i*n vitro* metabolism data for additional 30 compounds, the evaluation of the software package TIMES, and the chemoinformatics reactivity modeling. The metabolic degradation of the food dye curcumin (E100) will be shown exemplarily.

Workflow

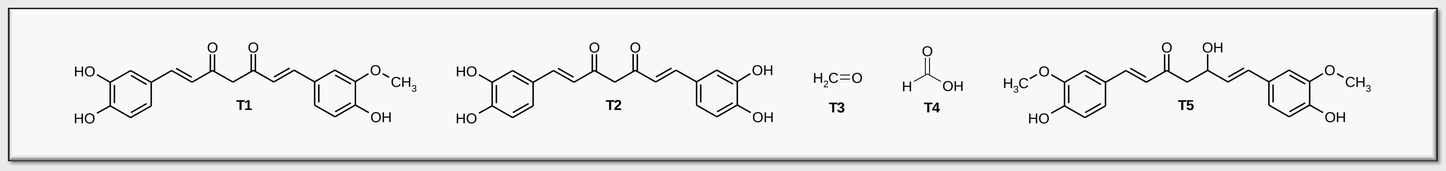
In the figure below the sequence of the individual workpackages is assembled to a workflow.



Evaluation of TIMES

OASIS TIMES was used to generate *in silico* the metabolites of curcumin. Curcumin was in the applicability domain of TIMES. TIMES predicts five metabolites of curcumin.

Scheme of curcumin metabolites predicted by TIMES



Although only the metabolite M2 (S9-fraction incubation), which is equal to T1 (TIMES evaluation) was detected *in vitro* and *in silico*, the main functionalization reactions were predicted correctly. The metabolite M2 resp. T1 was also detected *in vitro* by Tamvakopoulos et al.[1]

Data Set

The training data set based on literature data on rat S9 metabolism (data set 1) comprises 358 reactions. A second data set covering metabolism by rat microsomes was compiled. This data set on metabolism by rat microsomes (data set 2) comprises 3,390 reactions.

The data on rat S9 metabolism is quite limited. Therefore, the chemoinformatics investigation is focused on data set 2 extending the scope of the chemoinformatics modeling from rat S9 metabolism to rat microsomes.

	Data set 1 (rat S9)	Data set 2 (rat microsomes)
Unique comounds	414	3,924
Unique reactions	358	3,390
Unique substrates	77	1,043
Unique products	340	3,018

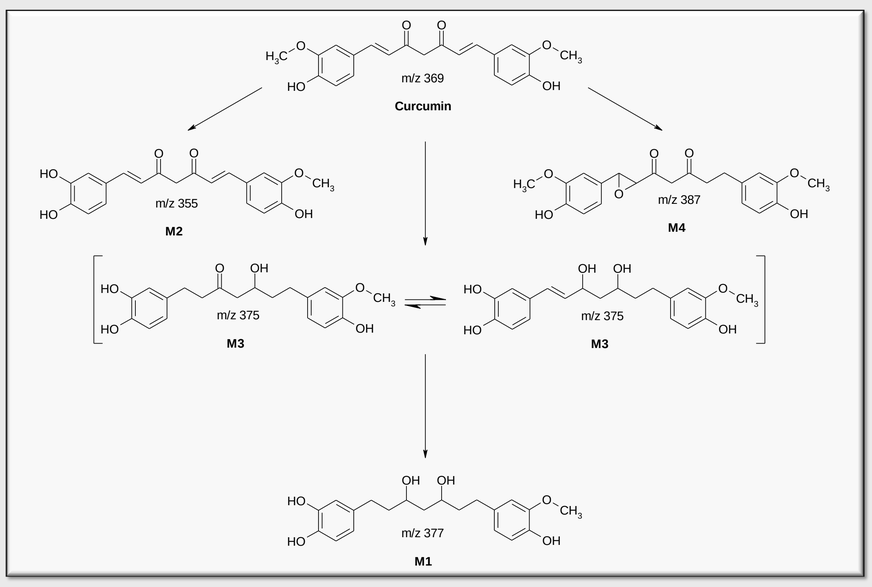
Statistics on data set 1 and data set 2

Experimental S9 Metabolism Measurements

Experimental setup

Metabolic clearance of curcumin (CAS: 000458-37-7) was performed by standard incubations in rat liver S9-fraction in a cofactor (NADPH) containing buffer system at 37 °C. After incubation proteins of incubates were precipitated by the addition of acetone and the supernatant was analyzed by LC/UV and LC/MS/MS.

Metabolic pathway of curcumin in rat liver S9-fraction



MOSES.Metabolism

Approach

- The reactions are clustered by their reaction centers in order to identify the most relevant reaction types.
- Probabilities for each reaction type are derived from the reaction database by statistical analysis.
- For the metabolite prediction the reaction rules are applied to the query compounds. The metabolites are ranked by the probability of the applied reaction rule.

Modeling of aromatic hydroxylation with physico-chemical descriptors

Starting from the 1,043 substrates in the data set 2 the reaction rule for the hydroxylation of aromatic phenyl rings yields a total number of potential 3,137 reactions. Among those 3,137 reactions are 284 observed and 2,853 non-observed reactions.

Partition & CV		number of runs	%correct predictions mean	stdev	min	max
training		1	99.2			
cross-validation	LOO	10	98.6	0.1	98.5	98.7
	10-fold	10	97.8	0.3	97.0	98.1
	5-fold	20	97.1	0.3	96.5	97.6
	2-fold	50	94.4	0.4	93.4	95.5

Predictability of random forest models for aromatic hydroxylation

Pred. / exp.	Observed	non-observed	sum	Sensitivity [%]
Observed	240	44	284	84.5
Non-observed	26	2827	2853	99.1
Sum	266	2871	3137	
Specificity [%]	90.2	98.5		

Confusion matrix for a 10-fold cross-validation run with 97.8% predictability and the classifier random forest

Depiction of the reaction centers of the top clusters

Cluster	Reaction center
1	
2	
3	cyclic
4	
5	acyclic
6	
7	N-demethylation of acyclic tertiary amines
8	
9	N-demethylation of cyclic teriary amines

Metabolite Prediction with the Current Model

- The metabolite M2 resulting from the O-demethylation of a methoxy group attached to an aromatic ring is predicted as main metabolite with the highest likelihood.
- Glucuronidation or sulphation of the aromatic hydroxy group are predicted at the second rank. The S9-fraction does not catalyze phase II reactions.

Acknowledgements

We gratefully acknowledge CEFIC-LRI for funding the project „Overcoming current Limitations in Metabolism Prediction of Industrial Chemicals - OLIMPIC” (LRI-Q2-MN081212).

References

[1] Tamvakopoulos C, Sofianos ZD, Garbis SD, Pantazis P. Analysis of the *in vitro* metabolites of diferuloylmethane (curcumin) by liquid chromatography--tandem mass spectrometry on a hybrid quadrupole linear ion trap system: newly identified metabolites. *Eur. J. Drug Metab. Pharmacokinet.* **2007**, *32*, 51-57.

Molecular Networks GmbH
Henkestr. 91
91052 Erlangen, Germany

Phone +49 9131 815668
Fax +49 9131 815669

www.molecular-networks.com
info@molecular-networks.com