Development of the Mirabilis Knowledge Base for PMI Purge Prediction: Knowledge, Experience & Reaction Mining

James A. McManus, Samuel J. Webb, Michael J. Burns

James.McManus@LhasaLimited.org
Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS

Introduction

Virtually all API syntheses will involve mutagenic or potentially mutagenic impurities (PMIs), which may end up in the final API and require a risk assessment. However, if the processes involved can be expected to purge the PMI through chemical and/or physical means, there may not be a need for analytical testing. ICH M7 Control Option 4 outlines that "a risk assessment can be based on physicochemical properties and process factors that influence the fate and purge of an impurity".[1]

Physicochemical parameters	Purge factor		
Reactivity	High reactivity = **100**	Moderate reactivity = **10**	Low reactivity = **1**
Solubility	Freely soluble = **10**	Moderately soluble = **3**	Sparingly soluble = **1**
Volatility	Boiling point >20 °C *below* that of the reaction/ process solvent = **10**	Boiling point within ±20 °C of that of the reaction/ process solvent = **3**	Boiling point >20 °C *above* that of the reaction/ process solvent = **1**
Ionisability	Acidity/basicity of PMI *significantly different* from that of the desired product.		
Physical process: Chromatography	**10** – **100** based on *extent of separation*		
Physical process: Scavenger resin	Evaluated on an individual basis		

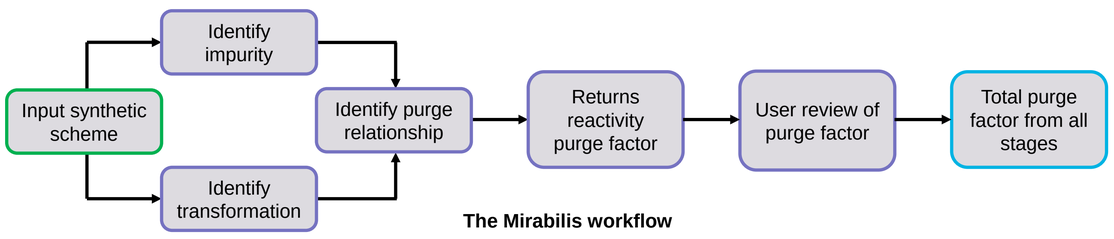
Table of physiochemical parameters and their related purge factor scores.

Purge factors are based on key physicochemical properties including reactivity, solubility, volatility and purification processes (e.g. chromatography). The product of all purge factors for each step gives the total purge factor for the PMI across the synthesis. Reactivity can be rationalised by chemical knowledge and has the highest scoring of 100.[2,3]

Mirabilis

Mirabilis is scientifically robust software providing an industry-standardised approach for calculating reactivity purge factors of PMIs.[4]

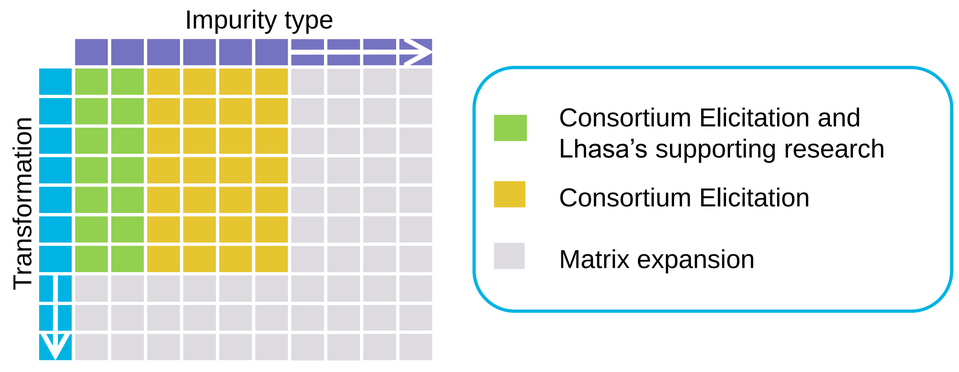
Mirabilis



The Mirabilis workflow

- Mirabilis is developed by Lhasa Limited in conjunction with an industrial consortium to provide a consensus of reactivity purge factors, based on the expert knowledge and experience of multiple chemists.
- This *in silico* method allows for greater consistency and reproducibility.
- Mirabilis uses a conservative approach and will deliberately under-predict the true purge to act as a fail-safe.
- Mirabilis provides a living knowledge base to assist scientists and regulators in making and reporting purge predictions.
- Mirabilis generates a report which is suitable for inclusion in a regulatory submission with robust scientific evidence to support the predicted reactivity purge factors.

Reactivity Knowledge Matrix



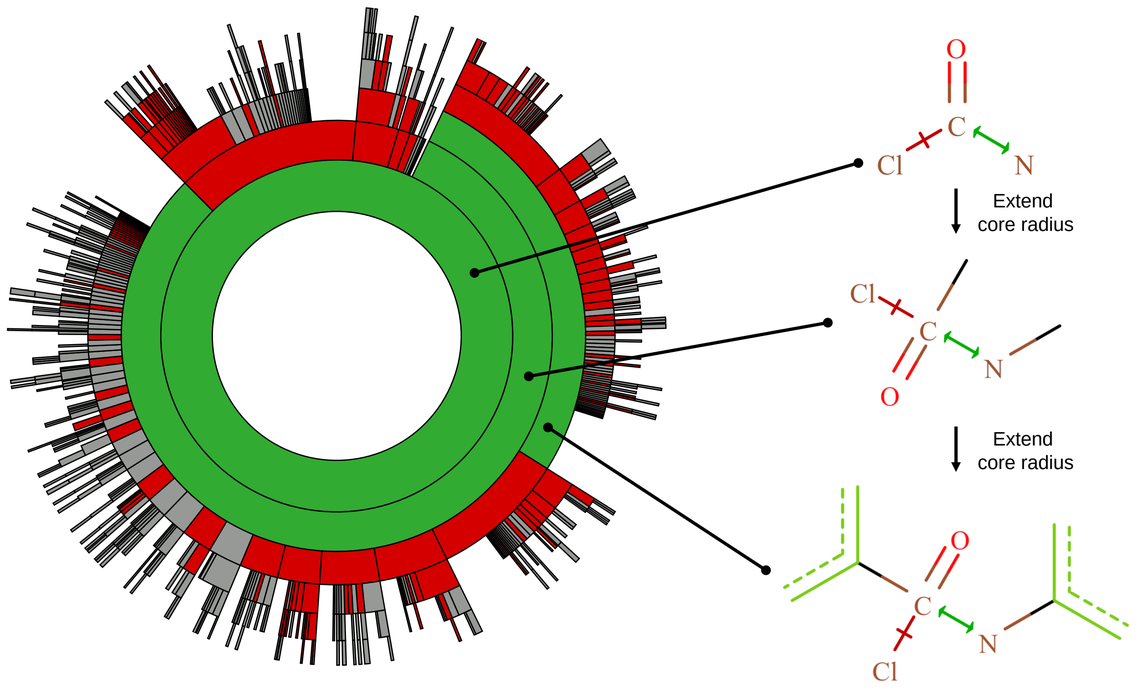
Representation of the Mirabilis 3.1 knowledge matrix

- Reactivity purge factors are contained within a knowledge matrix of common transformation and impurity types.
- Purge values were originally obtained through an elicitation process to achieve a consensus for each cell in the matrix.
- Lhasa scientists augment the elicitation to include research on the reaction using published literature (mechanisms of purge, scope and limitations, examples and references to support the purge).
- The knowledge matrix is being rapidly expanded with more transformations and impurities.

Lhasa's Prototype Reaction Mining Tool

To provide users with supporting information for the expanding knowledge matrix ahead of in-depth research, reaction mining can be utilised on the United States Patent Office (USPTO) patent grants.

- Over 1.4 million text-mined reactions provided by NextMove software.[5]
- The tool extracts property descriptions (including chemical structures, time, temperature and yield) and then uses reaction mapping to generate reaction cores.
- Facilitates rapid analysis of reaction data by searching the database for matching cores.
- Assists chemists in their assessment of a purge argument.

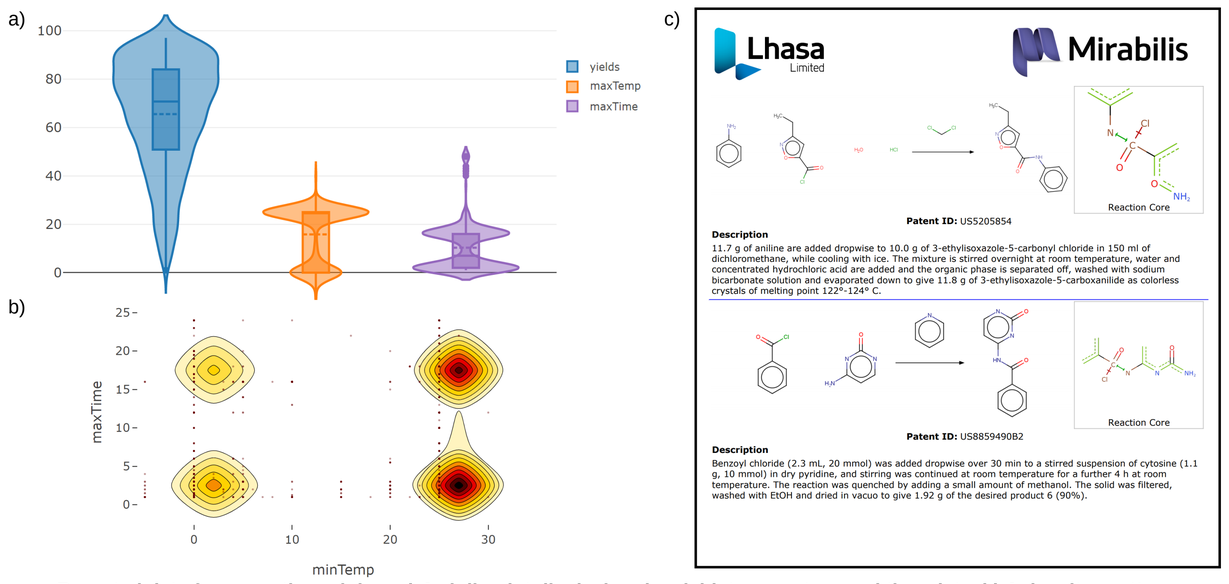


An interactive sunburst chart for all the reaction core types that match the query reaction scheme, in this case an amide formation. The central ring is the simplest core and each step away from the centre increases the core radius. Green = Selected cores, Red = Core matches query, Grey = Core no longer matches query.

The underlying database is labelled with impurity types, which can then be used to generate reaction networks describing the reactivity of functional groups of interest.

A chemist can then interact with these networks, selecting the relevant cores and setting limits for temperature and time ranges.

The prototype reaction mining tool can be used to identify similar reactions to the input reaction, which can be analysed to visualise the scope of yields and conditions. The user can then compare the mined data to their reaction conditions to assist with determining whether the expected reactivity of the impurity is within the scope of their process.



Reported data from reaction mining: a) A violin plot displaying the yield, temperature and time data. b) A density scatter plot to show correlation between time and temperature data. c) Supporting examples provided from patent literature.

Conclusion

Mirabilis is scientifically robust software that assists chemists with the reactivity purge calculation of PMIs and generates a report to be included in regulatory submissions.

Lhasa scientists can use the reaction mining tool to populate the new cells of the knowledge matrix with common conditions, solvents and reagents, as well as providing supporting examples from the patent literature. This allows for a fast approach to fill in the expanding knowledge matrix whilst the more in-depth research continues at Lhasa.

References and Acknowledgements

1) ICH M7 Guidelines, https://www.ich.org/products/guidelines/multidisciplinary/article/multidisciplinary-guidelines.html (accessed 20/08/2019).
2) Teasdale, A. et al., Org. Process Res. Dev., 2013, 17, 221-230.
3) Teasdale, A. et al., Org. Process Res. Dev., 2010, 14, 943-945.
4) Burns, M. J. et al., Org. Process Res. Dev., Manuscript submitted.
5) NextMove Software, https://www.nextmovesoftware.com/ (accessed 15/08/2019).
The authors would like to thank the Mirabilis consortium for their contributions.

shared **knowledge** • shared **progress**