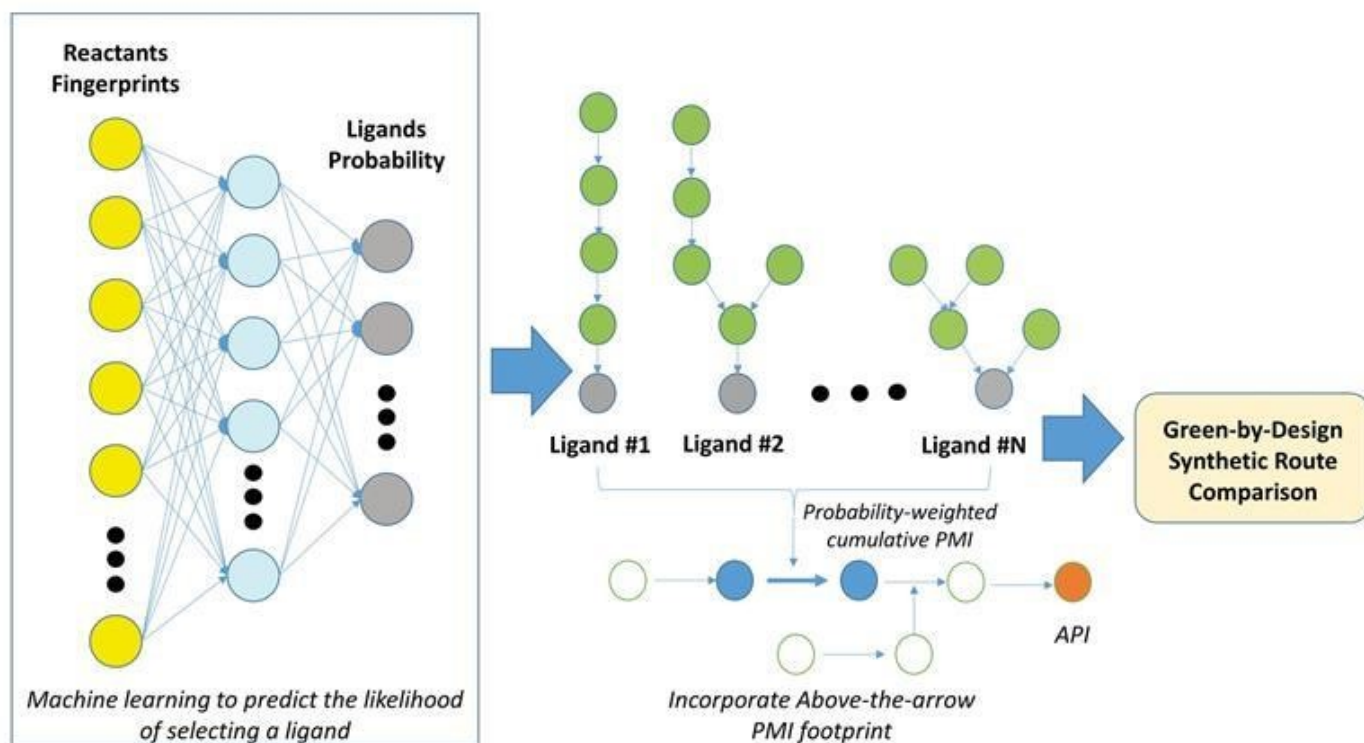# CHEMISTRY\\WORLD

Ligand selector steers C–N cross-couplings down most sustainable path

BY HARRIET BREWERTON | 2 APRIL 2019

Tool is step towards strategy that considers reagents and reactants above and below the arrow

Source: © Martin Eastgate/Bristol-Myers Squibb

Scientists have devised a framework for incorporating machine-learned ligand prediction into predictive route comparisons, to enable greener chemistry outcomes

Researchers have used machine learning to develop a tool that predicts which ligands for a metal-catalysed coupling reaction will result in a synthetic route with the lowest

environmental and financial cost. The idea could be expanded into a system to help pharmaceutical organisations select how to manufacture a drug.

Pharmaceuticals often have complex synthetic routes with several possible paths to the final product. Scientists designing these routes need to pick the optimal one and, historically, such decisions centre on safety, efficiency, cost and product quality.

Given the positive correlation between reaction cost and sustainability, Jun Li and Martin Eastgate at Bristol-Myers Squibb, US, have now designed a machine learning approach that can predict the synthetic route with the lowest environmental impact. Environmental impact is gauged using the cumulative mass intensity ratio – the mass of all the materials used in the synthesis divided by the mass of the final product. Higher values mean more wasted materials and a higher impact.

Li and Eastgate's tool works on transition metal-catalysed carbon–nitrogen coupling reactions involving phosphine ligands, which frequently feature in pharmaceutical syntheses. Literature reports of coupling reactions with phosphine ligands served as the dataset for the system; the molecular features of ligand electrophiles and nucleophiles provide the input variables, and the phosphine ligands that result in successful reactions are the output. They found that their tool predicts which ligands will provide a successful reaction, and which ones provide the lowest cumulative mass intensity.

Machine intelligence expert Ross King at the University of Manchester, UK, says the research 'tackles how best to design synthetic paths that not only have high-yields, but are also of low financial and environmental cost is an important subject area, and an area that will only grow in importance. This is yet another successful application of machine learning in chemistry'.

The work 'helps draw attention to two key challenges in synthesis design that could benefit from computational assistance: ligand selection for catalytic reactions, and evaluation of a route's greenness after taking that selection into account,' comments Connor Coley whose research at Massachusetts Institute of Technology, US, uses data and automation to streamline discovery in the chemical sciences. 'Hopefully, we will see many more studies

like this one that help bring quantitative metrics into route selection beyond cost and number of reaction steps.'

'We hope this work will help researchers make better decisions during route design,' comments Eastgate. 'Keeping sustainability and efficiency in mind, on a holistic level, during these decisions – through predictions and an easy to use app – will help provide greater context to these key decisions and help researchers choose route options with the highest chance of being the most sustainable'.

References

This article is free to access until 14 May 2019
J Li and M D Eastgate, *React. Chem. Eng.*, 2019, DOI: 10.1039/c9re00019d