

Knowledge-based De Novo Design using Reaction Vectors

H Patel¹

V Gillet¹, B Chen², M Bodkin³

¹ *University of Sheffield, Department of Information Studies, Sheffield, UK*

² *University of Sheffield, Department of Chemistry, Sheffield, UK*

³ *Eli Lilly UK, Windlesham, UK*

A number of *de novo* design tools have been described with the aim of generating novel molecules for drug design, however, they are limited in their ability to propose molecules which are synthetically feasible. Here we describe a novel method that utilises reaction vectors from databases of known reactions to generate structures of interest. The method has been implemented using the pipelining environment KNIME (1).

The reaction vector captures the changes that take place at the reaction centre, without the need for complex reaction mapping procedures (2). By first describing the individual components of a reaction using descriptors such as atom pairs, the overall reaction vector is generated using:

Reaction Vector = [Sum of product vectors] – [Sum of reactant vectors]

We show how reaction vectors can be used to generate novel molecules for synthesis based on simple transformations involving, for example, a simple functional group substitution, to more complex multi-component reactions of the form (R1 + R2 → P1 + P2). We demonstrate the application of the method to the design of known drugs from simple starting materials and a ‘cleaned’ reaction dataset, via mixing and matching of reaction transforms and reactants.

We also describe the how the method can be developed into an automated multi-objective application for *de novo* design.

References:

1. Konstanz Information Miner. www.knime.org
2. Broughton, H. B. et al. Methods for Classifying and Searching Chemical Reactions. United States Patent Application 367550, 25 Sept, 2000.