Comparative evaluation of atom mapping algorithms for metabolic reactions

German Preciat Gonzalez (*gapreciat@gmail.com), Lemmer El Assal, Ines Thiele, Hulda Haraldsdóttir, **Ronan M.T. Fleming**(ronan.mt.fleming@gmail.com)

Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Project Goals: Comparison of the predictive accuracy of atom mapping algorithms.

The reaction mechanism of each chemical reaction in a metabolic network can be represented as a set of atom mappings, each of which relates an atom in a substrate metabolite to an atom of the same element in a product metabolite. Atom mapping data for metabolic reactions open the door to a growing list of applications [1, 2, 3, 4]. Complete manual acquisition of atom mapping data for a large set of chemical reactions is a laborious process. Many algorithms exist to predict atom mappings. How do their predictions compare to each other and to manually curated atom mappings? For more than five thousand metabolic reactions we compared the atom mappings predicted by six atom mapping algorithms [5, 6, 7, 8, 9, 10]. We also compared these predictions to those obtained by manual curation of atom mappings for over five hundred reactions distributed amongst all top level enzyme commission number classes. Five of the evaluated algorithms had similarly high prediction accuracy over 91% when compared to manually curated atom mapped reactions. On average, the accuracy of the prediction was highest for reactions catalysed by oxidoreductases and lowest for reactions catalysed by ligases. In addition to prediction accuracy, the algorithms were evaluated on their availability and advanced features such as the ability to identify equivalent atoms and reaction centres, and the option to map hydrogen atoms. In addition to prediction accuracy, we found that availability and advanced features were fundamental to the selection of an atom mapping algorithm.

References

[1] Wiechert W (2001) 13c Metabolic Flux Analysis. Metabolic Engineering 3: 195–206.

[2] Haraldsdóttir HS, Fleming RMT (2016) Identification of Conserved Moieties in Metabolic Networks by Graph Theoretical Analysis of Atom Transition Networks. PLOS Computational Biology 12: e1004999.

[3] Pey J, Planes FJ, Beasley JE (2014) Refining carbon flux paths using atomic trace data. Bioinformatics 30: 975–980.

[4] Kotera M, Okuno Y, Hattori M, Goto S, Kanehisa M (2004) Computational Assignment of the EC Numbers for Genomic-Scale Analysis of Enzymatic Reactions. Journal of the American Chemical Society 126: 16487–16498.

[5] First EL, Gounaris CE, Floudas CA (2012) Stereochemically Consistent Reaction Mapping and Identification of Multiple Reaction Mechanisms through Integer Linear Optimization. Journal of Chemical Information and Modeling 52: 84–92.

[6] ChemAxon (2015) Standardizer, was used for structure canonicalization and transformation, JChem 16.1.11.0, 2015, ChemAxon (http://www.chemaxon.com). URL `ChemAxon (http://www.chemaxon.com)`.

[7] Rahman SA, Torrance G, Baldacci L, Cuesta SM, Fenninger F, et al. (2016) Reaction Decoder Tool (RDT): extracting features from chemical reactions. Bioinformatics 32: 2065–2066.

[8] Kumar A, Maranas CD (2014) CLCA: Maximum Common Molecular Substructure Queries within the MetRxn Database. Journal of Chemical Information and Modeling 54: 3417–3438.

[9] Latendresse M, Malerich JP, Travers M, Karp PD (2012) Accurate Atom-Mapping Computation for Biochemical Reactions. Journal of Chemical Information and Modeling 52: 2970–2982.

[10] Kraut H, Eiblmaier J, Grethe G, Löw P, Matuszczyk H, et al. (2013) Algorithm for reaction classification. Journal of Chemical Information and Modeling 53: 2884–2895.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.