Modeling of Complex Chemical Reaction Networks

Nico Domschke[1], Nora Beier[1], Thomas Gatter[1], Bernd Abel[2], Peter F. Stadler[1]

[1] Bioinformatics, Leipzig University, Haertelstraße 16-18, 04107, Leipzig, Germany
[2] Wilhelm-Ostwald-Institut für Physikalische und Theoretische Chemie, Leipzig University, Linnéstraße 2, 04103, Germany

In the field of chemical analysis in order to research complex reactions mixtures advanced computational modeling is necessitated by a series of inherent challenges, including complex compositions, unknown intermediates and combinatorial limits to achieving exhaustive analysis.
This study introduces a methodological pipeline tailored for the analysis of MS/MS spectra associated with complex reaction mixtures.

First MØD, a rule-based network generation framework rooted in category theory, is utilized.[1] It is designed as a tool for very fast chemical space exploration, employing double-pushout diagrams to specify reaction rules. This assures mathematical consistency, formal reaction reversibility and associative composability of rules and reactions. The application of the "AlChemy" rule set collection ensures adherence to chemistry-like behavior.[2] The framework iteratively computes all reachable products, mitigating a combinatorial explosion by different pruning strategies, by which chemically unlikely products can be recognized and rejected.

Following the generation of potential products, a MS/MS spectra predictor such as CFM-ID is used to predict possible fragments.[3] These fragments, along with their corresponding base molecule and mass peaks, are stored in a database. Subsequently, when presented with the MS/MS spectra of a complex reaction mixture, the database can be queried on a peak-by-peak basis. This information facilitates the indentification of the most probable candidate molecules and most likley formation pathway (figure 1).

To demonstrate this approach, we turn to the extensively studied Miller-Urey Experiment. Through this pipeline, we offer a systematic and computationally efficient method for unraveling complexities inherent in reaction mixtures, paving the way for a better understanding of reaction mixtures.

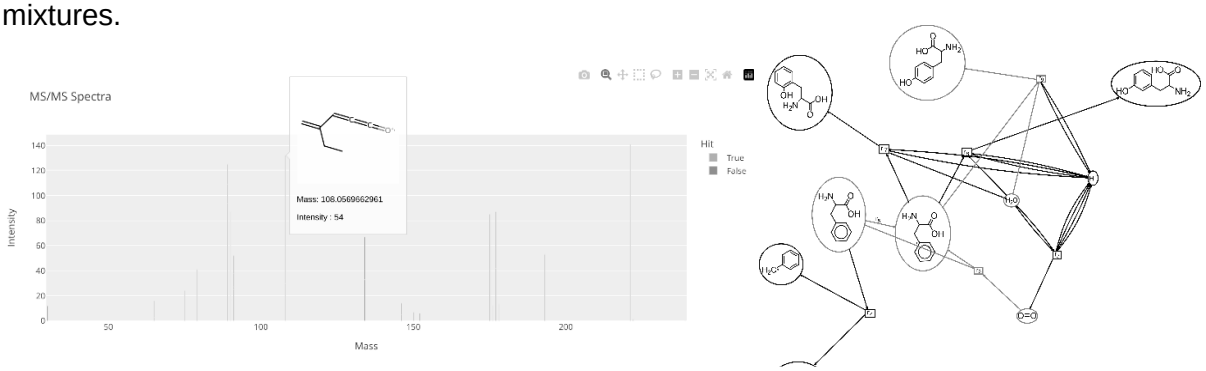

Figure 1 : (left) Analysis of candidates in MS/MS Spectra of complex reaction mixture. Overlaps in Predicted fragments and predicted fragments are highlighted (right) Derived likeliest reaction path to found candidates.

Bibliography :

[1] J. L. Andersen; C. Flamm; D. Merkle; P. F. Stadler. J. Syst. Chem., 4 (2013) 4.
[2] A. Wollos; R. Roszak; A. Żądlo-Dobrowolska; W. Beker; B. Mikulak-Klucznik; G. Spólni; M. Dygas; S. Szymkuć; B. A. Grzybowski Science, 369 (2020) 1-12.
[3] F. Wang, D. Allen, S. Tian , E. Oler , V. Gautam, R. Greiner, TO Metz, DS Wishart. Nucleic Acids Research 50 (2022) W165-W174.