Chemography-Guided Analysis of Reaction Path Networks by GTM

Philippe Gantzer[1], Ruben Staub[1], Alexandre Varnek[1, 2]

[1] *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Kita 21, Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0021, Japan.*

[2]*Laboratory of Chemoinformatics, UMR 7140, CNRS, University of Strasbourg, 67081 Strasbourg, France.*

Ab initio kinetic studies provide microscopic insights into reaction mechanisms by exploring reaction paths. Resulting reaction networks are represented as connected graphs in which nodes represent stationary structures whereas edges correspond to chemical transformations. The visualization and analysis of large chemical reaction networks becomes rather challenging when conventional graph-based approaches are used. As an alternative, we proposed to use the Generative Topographic Mapping (GTM) algorithm[2], a probabilistic dimensionality reduction method describing data distributions on a 2-dimensional map.[1]

Reaction path networks for ethylene hydrogenation with a model Wilkinson's catalyst were obtained with the help of the Artificial Force Induced Reaction (AFIR)[3] method, coupled with either Density Functional Theory or Neural Network Potential (NNP) for potential energy surface calculations. A series of GTMs, each accommodating some 10^6 geometries, identified the areas of high and low energy, some zones populated by physically irrelevant 3D structures and those corresponding to different steps of the reaction path. The expansion of the reaction chemical space explored during an AFIR-based reaction path search was monitored by successive projections on GTM manifold. The design of new D^3 descriptors allowed to distinguish different molecular structures while regrouping structures possessing similar geometries on GTM.[4]

Bibliography :

[1] R. Staub, P. Gantzer, Y. Harabuchi, S. Maeda, A. Varnek, Molecules 2023, 28, 4477.
[2] H. A. Gaspar, I. I. Baskin, G. Marcou, D. Horvath, A. Varnek, Mol. Inf. 2015, 34, 348–356.
[3] Maeda, S.; Taketsugu, T.; Morokuma, K., J. Comput. Chem. 2014, 35, 166–173.
[4] Gantzer P, Staub R, Harabuchi Y, Maeda S, Varnek A., ChemRxiv. 2024