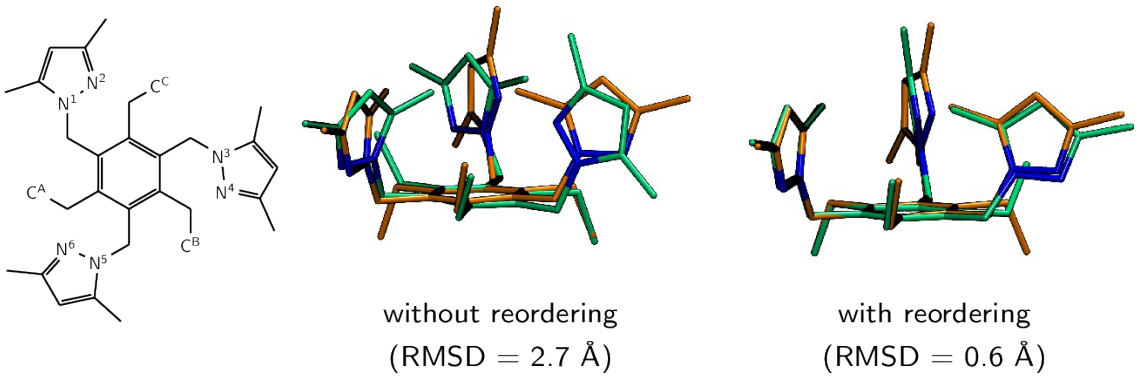
The Hungarian method revisited – Usage of combinatorial optimisation for structural comparison of conformers

Conrad Hübler

Institut für Physikalische Chemie, Technische Universität Freiberg, Leipziger Straße 29, 09599 Freiberg

Conformational search problems are mostly associated with finding the global minimum geometry of a structure of interest. The analyses of the potential surface using various molecular modelling methods reveal, that the most stable structures at a given lower level of theory (for example semiempirical methods) do not necessarily coincide with the most stable structures at a higher level of theory (for example density functional theory).[1] This may be overcome by re-evaluating and/or reoptimising all structures at a higher level of theory, resulting in an increase of the overall computational cost. Furthermore, molecules with some kind of symmetry may give rise to another challenge: Although the presence of symmetrically equivalent atoms can be helpful in quantum chemical calculations, “topological symmetry” in (sub)structures may prohibit the successful identification of identical or similar conformations in a set of conformational search results. As geometries are in general stored as a list of atoms with fixed order, the root mean square deviation (RMSD) between two structures may not represent the true similarity unless the atomic indices are reassigned.



A general approach to the combinatorial optimisation problem was first proposed by Kuhn[2] and improved by Munkres.[3] The so-called Hungarian method was implemented for the use of structure comparison in a *python* tool available at https://github.com/charnley/rmsd.[4] However, there is a good chance that the plain Kuhn-Munkres approach does not result in the best reassignment of the atoms and hence the identification of duplicate or similar structures may fail. During the analysis of the supramolecular binding behavior of tripodalic receptors, both the reassignment problem of the atoms in case of two structures and a general conformational filtering with respect to the “topologically symmetric” structures were analysed in detail and a combined approach is now implemented in the program *curcuma*.[5] The conformational filter approach includes the Hungarian method with a well defined pre-aligning of the geometry for each pair of two structures and to pre-judge whether two structures are potentially similar and reassigning atomic indices may be fruitful. The pre-judging is based on the comparison of the total energies, rotational constants and Vietoris–Rips barcodes[6] which all are invariant with respect to the ordering of the atomic indices. However, they are in most cases not sufficient to correctly identify similar structures on their own.

[1] R. Sure, S. Grimme, *Chem. Commun.*, **2016**, *52*, 9893-9896. [2] H. W. Kuhn, *Nav. Res. Logist.*, **1955,** *2*, 83-97. [3] J. Munkres, *SIAM J. Appl. Math.*, **1957**, *5*, 32–38. [4] Calculate Root-mean-square deviation (RMSD) of Two Molecules Using Rotation, GitHub, http://github.com/charnley/rmsd, <1.5.1> [5] C. Hübler. (2020). conradhuebler/curcuma: Curcuma . Zenodo. https://doi.org/10.5281/zenodo.4302723. [6] U. Bauer, *J Appl. and Comput. Topology* **2021**, *5*, 391–423.