Multiple molecular superpositioning with a common core structure

Malte Schokolowski, Christoph Noack, Niklas Paulicks, Marius Rüve, Mike Trzaska, Jule Würfel, Christian Meyenburg, Thorben Schulze, Matthias Rarey

Universität Hamburg, ZBH – Center for Bioinformatics, Hamburg, Germany

Molecular alignments play a pivotal role in drug design, especially when information about the structure of a target protein is lacking.

With a set of known binders, ligand-based approaches offer the chance to derive a pharmacophore model, facilitating the search for further potential ligands for the same binding site. Crucial for these ligand-based methods is that the ligands bind comparatively and have a similar binding geometry. Since compound series in drug discovery projects often share a common core structure, scaffold superposition can be utilized for an initial alignment of the molecules. While many alignment methods exist, approaches utilizing common core structures are still a niche. [1]

Based on these foundations, we developed an aligner for multiple small molecules called CoAler (https://github.com/ciw-project-2023/coaler) which utilizes the common core structure and superimposes dozens of small molecules in seconds to minutes on a standard desktop computer. CoAler is based on an algorithm called MolAlign, [2] which takes pre-generated conformers, performs pair-wise alignments for each conformer of each ligand and generates multiple ligand assemblies based on the pair-poses generated. Our implementation is developed in C++ using the open-source toolkit for cheminformatics RDKit [3] and is free to use for your own applications and workflows.

[1] S. Hönig, C.Lemmen, M. Rarey, *WIREs Comput Mol Sci*, **2022**, 13(2), e1640

[2] Chan, S.L., *J Comput Aided Mol Des,* **2017**, *31*, 523–546.

[3] RDKit: Open-source cheminformatics. https://www.rdkit.org.