Contribution ID : **199**   Type : **Poster Presentation**

Physics-informed ML hydration free energy models with EC-RISM augmentation: Physicality matters

The accurate prediction of hydration free energies (HFEs) is an important task in different areas, as the thermodynamic cost of solvation is one of the key factors in processes from protein-ligand binding to chemical reactions. Hence, much effort has been spent in predicting the HFE from electronic structure calculations and molecular dynamics simulations as well as, in recent years, machine learning (ML) methods,[1] for which large and reliable datasets are needed.

Training data dependence can be reduced using physics-informed ML, where measured or calculated properties representing the underlying physics of the molecules and their interactions are used as additional input features.[2] Here we present a physics-informed ML method that combines the Embedded Cluster Reference Interaction Site Model (EC-RISM)[3] with a Message Passing Neural Network (MPNN).[4] The solute is represented by a graph, where atoms are described by partial charges and Lennard-Jones parameters together with "local" atomic free energies (LFEs)[5,6] adding up to the total HFE. Augmenting with EC-RISM HFEs, we demonstrate state-of-the-art accuracy on independent HFE datasets, including SAMPL challenge data.[7,8] This model as well as a range of related approaches with varying numbers of physical features is then employed to predict the HFEs of tautomers. These HFEs are combined with high level gas phase data to calculate the aqueous tautomerization free energies for a curated Tautobase[9] subset. We compare the performance of our physically augmented HFE prediction approach to literature HFE models, both for independent HFE datasets and for the Tautobase subset. We demonstrate that a good HFE prediction performance does not automatically translate into performing well on the solvation contribution to the aqueous tautomerization free energy, illustrating the importance of physicality for predicting reaction thermodynamics in solution by accounting for relevant physical features.

Primary author(s) :   CHODUN, Christian

Co-author(s) :   STROBL, Michael (TU Dortmund university);   Dr. TIELKER, Nicolas (TU Dortmund University);   KAST, Stefan M. (TU Dortmund University)

Presenter(s) :   KAST, Stefan M. (TU Dortmund University)