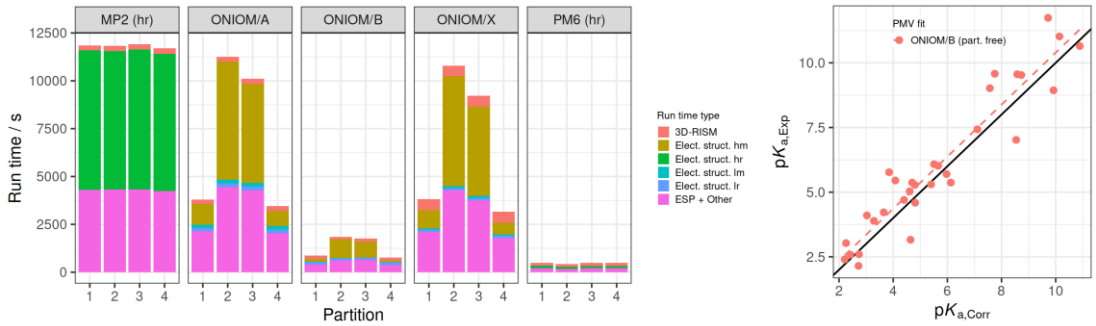
Accurate prediction of acidity constants with an ONIOM embedded cluster RISM approach

Lennart Eisel,[1] Nicolas Tielker,[1] Matthias Hennemann,[2] Timothy Clark,[2] Stefan M. Kast[1]

[1]*Fakultät für Chemie und Chemische Biologie, Technische Universität Dortmund, 44227 Dortmund, Germany*

[2]*Department Chemie und Pharmazie, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91052 Erlangen, Germany*

The accurate prediction of physicochemical properties such as acidity constants (pK_a) plays a decisive role for, e.g., modelling protein function or in the development of drug-like molecules. As cellular environments show a variety of different pH values, the pK_a of a titratable residue or drug therefore heavily influences the mode and strength with which a compound may bind to its target. Our previous strategy for predicting pK_a values for small molecules focused on modelling the compounds' thermodynamics in an aqueous environment by means of our "embedded cluster RISM" (EC-RISM) solvation model and first-principles quantum mechanical (QM) calculations. [1] However, the size of most biomolecular systems prevents the application of these QM methods, thus prohibiting the accurate and granular modelling of the solvent environment with EC-RISM. Since the first suggestion by Warshel and Levitt [2] multiscale methods have emerged as an effective tool to model large-scale chemical processes in various environments, thus offering a route to expand the range of system sizes that can be modelled via EC-RISM theory.



Here we present a novel multiscale solvation model which integrates a subtractive ONIOM(QM:SQM) [3] description of the solute into the EC-RISM formalism, combining established high-level QM methodology with EMPIRE [4] for the low-level semiempirical (SQM) component. We show how the set of equations used within this model can be derived, by similar approximations to the ones used for the ONIOM-PCM solvation model. [5] Extending previous schemes, [1] we develop an empirical correction to the solute's Gibbs energy, which is free of any ONIOM partitioning error. The resulting model is then validated for the SAMPL6 pK_a challenge dataset. [1,6] Our promising results show that the novel ONIOM-EC-RISM approach ranks equal in prediction quality with our previous full-QM EC-RISM methodology, while simultaneously drastically reducing the required computational cost. Thus, the model offers a step forward for modelling systems in solution up to biomolecular targets.

[1] N. Tielker, L. Eberlein, S. Güssregen, S. M. Kast, *J. Comput. Aided Mol. Des.* **2018**, *32*, 1151–1163.
[2] A. Warshel, M. Levitt, *J. Mol. Biol.* **1979**, *103*, 227–249.
[3] S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma, M. J. Frisch, *Theo. Chem.* **1999**, *461*, 1–21.
[4] M. Hennemann, T. Clark, *J. Mol. Model.* **2014**, *20*, 2331.
[5] T. Vreven, B. Mennucci, C. O. da Silva, K. Morokuma, J. Tomasi, *J. Chem. Phys.* **2001**, *115*, 62–72.
[6] M. Işık, ..., J. D. Chodera, *J. Comput. Aided Mol. Des.*, **2021**, *35*, 131–166.