Abstract ID: 92778

Student: Platero Rochart Daniel

Area of Research: Computational and structural science

PhD Programme: PhD Molecular Medicine (MolMed)

Semester: 3

Computer aided (re)-design of enzymes for therapeutics: First steps on energy barriers prediction

Daniel Platero-Rochart; Pedro A. Sánchez-Murcia

Enzyme manipulation has proven to be an effective way to modulate properties such as catalytic activity, substrate specificity, and (thermo)stability. Nowadays, advances in cloning and expression have allowed the use of experimental techniques such as directed evolution (DE) for protein engineering campaigns. Nevertheless, DE has unpredictable outcomes and can be an expensive process. Therefore, in-silico methods like steered QM/MM molecular dynamics (sMD) for calculating free energy barriers, constitute an alternative for a rational design. However, the use of sMD for comprehensive mutation libraries is still a time- and resource-demanding approach. In the present work, we explored the possibility of using Machine Learning approaches to overcome the sMD limitations for the exploration of the mutational landscape in enzyme engineering.

Employing two kernel regression methods (kernel ridge regression and ElasticNet) and different chemical representations, we were able to obtain good regression scores for the prediction of the pulling work for each sMD frame. We also demonstrate that using initial snapshots (directly from molecular dynamics (MD) simulations) the predicting power is remarkably low, being needed structures closer to the transition state.

On one hand, our results show that kernel regression methods along with a simple representation such as the Coulomb matrix, are capable of learning and predicting pulling work from an entire sMD trajectory. On the other hand, while using only a single structure, representation can play an important role. Moreover, results strongly depend on the selected snapshot for the training process. We also show that the inherent variability of the sMD makes the delivery of predictive models a challenging task.