

Navigating protein landscapes with a machine-learned transferable coarse-grained model

Cecilia Clementi

¹ Department of Physics, Freie Universität, Berlin, Germany
E-mail: cecilia.clementi@fu-berlin.de

Abstract – The most popular and universally predictive protein simulation models employ all-atom molecular dynamics (MD), but they come at extreme computational cost. The development of a universal, computationally efficient coarse-grained (CG) model with similar prediction performance has been a long-standing challenge. By combining recent deep learning methods with a large and diverse training set of all-atom protein simulations, we have developed a bottom-up CG force field with chemical transferability, which can be used for extrapolative molecular dynamics on new sequences not used during model parametrization. We have demonstrated that the model successfully predicts folded structures, intermediates, metastable folded and unfolded basins, and the fluctuations of intrinsically disordered proteins while it is several orders of magnitude faster than an all-atom model. This showcases the feasibility of a universal and computationally efficient machine-learned CG model for proteins.

Keywords – *Coarse-graining, machine-learning, molecular dynamics*