Consistent interpretation of coarse-grained peptide kinetics using Markov state models biased with external information

Simple, physics-based coarse-grained (CG) models have provided tremendous insight into the essential features of the protein folding process. Recent advancements in CG methodologies allow increased chemical detail and accuracy, while retaining the sampling efficiency to address problems intractable for atomically-detailed models. This beneficial speed-up, attained through a combination of reduced molecular friction and softer interaction potentials, comes at the cost of obscuring the connection to the true dynamical properties of the underlying system. Although it is possible to rescue the dynamics via a generalized Langevin formalism, this approach offers a daunting computational and conceptual challenge for complex biological molecules that give rise to hierarchical dynamics, i.e., kinetic processes coupled over various timescales. As an alternative, this work considers a Markov state modeling framework for characterizing and correcting the hierarchy of slow kinetic processes generated from CG simulations. In particular, the proposed Bayesian scheme identifies essential adjustments to a Markov state model, generated from CG simulations, in order to achieve consistent kinetics, with respect to given reference data for the system. We test the method on two CG peptide models and demonstrate that the resulting information may be directly and effectively employed for model reparametrization. In both cases, the reparametrization results in an improved hierarchy of slow kinetic processes, while retaining the fundamental properties of the original model. Finally, to better characterize the utility of the proposed methodology in the context of combining CG protein folding simulations with experimental reference data, we perform a detailed investigation of structure-kinetic relationships for CG helix-coil transitions.