Abstract Submitted
for the MAR15 Meeting of
The American Physical Society

Building toy models of proteins using coevolutionary information[1] RYAN CHENG, MOHIT RAGHUNATHAN, JOSE ONUCHIC, Rice University — Recent developments in global statistical methodologies have advanced the analysis of large collections of protein sequences for coevolutionary information. Coevolution between amino acids in a protein arises from compensatory mutations that are needed to maintain the stability or function of a protein over the course of evolution. This gives rise to quantifiable correlations between amino acid positions within the multiple sequence alignment of a protein family. Here, we use Direct Coupling Analysis (DCA) to infer a Potts model Hamiltonian governing the correlated mutations in a protein family to obtain the sequence-dependent interaction energies of a toy protein model. We demonstrate that this methodology predicts residue-residue interaction energies that are consistent with experimental mutational changes in protein stabilities as well as other computational methodologies. Furthermore, we demonstrate with several examples that DCA could be used to construct a structure-based model that quantitatively agrees with experimental data on folding mechanisms. This work serves as a potential framework for generating models of proteins that are enriched by evolutionary data that can potentially be used to engineer key functional motions and interactions in protein systems.

[1]This research has been supported by the NSF INSPIRE award MCB-1241332 and by the CTBP sponsored by the NSF (Grant PHY-1427654).

Ryan Cheng
Rice University

Date submitted: 13 Nov 2014 Electronic form version 1.4