4CF13-2013-000090

Abstract for an Invited Paper
for the 4CF13 Meeting of
the American Physical Society

Designing Foldable Protein Sequence Through Zipping Contacts[1]

SEFIKA OZKAN, Arizona State University

Earlier experiments suggest that the evolutionary information (conservation of amino acids and coevolution between amino acids) encoded in protein sequences is necessary and sufficient to specify the fold of a protein family. However, there is no computational work to quantify the effect of such evolutionary information on the folding process. Here we simulate a repertoire of native and artificial WW domain sequences using a physics-based protein structure search method called ZAM (Zipping and Assembly method), which samples conformational space effectively towards native-like conformations through zipping and assembly search mechanism. We explore the sequence-structure relationship for WW domains and find that the coevolution information has a remarkable influence on local contacts of N-terminal β-turn of WW domains . This turn would not form correctly if lack of such information. Moreover, through maximum likelihood approach, we identify five local contacts that play a critical role in folding. Using the contact probability of those five local contacts at the early stage of folding, a classification model is built. This enables us to predict the foldability of a WW sequence with 81% accuracy. Based on this classification model, we re-design the unfoldable WW domain sequences and make them foldable by introducing a few mutations that leads to stabilization of these critical contacts.

[1]NIH 1U54GM094599