Topological constraints and challenges imposed by knots in proteins
Cesar Ramirez1, Jeffrey K. Noel2, Mauricio Baez1
1Department of Biology, University of Chile, Santiago, Chile
2Center for Theoretical Biological Physics, University of California San Diego, USA

Protein knots are intriguing topologies that have been gradually populating the Protein Data Bank, and are now being recognized as significant structural motifs. The existence of knots in proteins is challenging in both experimental and theoretical settings. One of the most interesting questions is how these proteins reach the native state, when starting from a fully extended polypeptide chain. Experimentally, this problem is complicated to address because unknotted configurations are difficult to distinguish from knotted configurations, as they differ by only subtle chain crossings. This difficulty is overcome in theoretical investigations since the protein coordinates are explicitly available.

Therefore, to address the question of how proteins knots are able to tie, we investigated thermodynamic and kinetic folding of the smallest knotted proteins known, VirC2 and MJ0366, from the ribbon-helix-helix (RHH) family of proteins by simulation. Structure-based models, folding models that are based in energy landscape theory, were investigated through molecular dynamics simulations. To better understand the role of side-chain packing in the specificity to the knotted topology, we used two levels of graining, a coarse-grained and an all-atom geometry, in the simulations.

Kinetic folding simulations for both proteins show a preordered and looped, but unknotted, intermediate state. From this ensemble of conformations, either backtracking events or threading the loop by plugging or slipknotting to reach the native state can be observed. In coarse-grained simulations, the occurrence of this intermediate state is more likely as the temperature is lowered from folding temperature, but it is not observed at folding temperature, in contrast to the behavior in all-atom simulations. Moreover, thermodynamic data in coarse-grained simulations indicates that protein folding corresponds to a two-state mechanism, but all-atom simulations for MJ0366 show that the intermediate state is meta-stable, so is better characterized by a three-state mechanism.

To have a deeper understanding of the constraints imposed by knotted topologies, we compared these results with structure-based simulations of Arc repressor, an unknotted dimer that is also a member of the RHH family of proteins and shares a high structural similarity with VirC2 and MJ0366. While coarse-grained folding simulations proceed through a two-state mechanism, all-atom simulations show that Arc repressor folds through a three-state mechanism, in a similar fashion to the behavior observed for MJ0366 and VirC2. The intermediate state for Arc repressor in all-atom simulations displays some of the features observed for the intermediate state in MJ0366 (beta-sheet formation, low number of native contacts).

There are two main lessons from this work. First, that there are many different possible folding routes for knots, but all involve threading a native-like loop. Variations include threading either the N- or C-terminus and threading via either slipknotting or plugging. The second lesson is gleaned from comparing the knotted versus unknotted topologies between similar native folds. The larger free energy barrier in VirC2 compared to Arc shows that folding knots imparts a topological constraint that defines the rate-limiting step to folding in our simulations.