Mapping Conformational Pathways between known Function Protein States

Kevin Molloy[1], **Amarda Shehu**[1,2,3]

http://www.cs.gmu.edu/~ashehu
amarda@gmu.edu | kmolloy1@gmu.edu

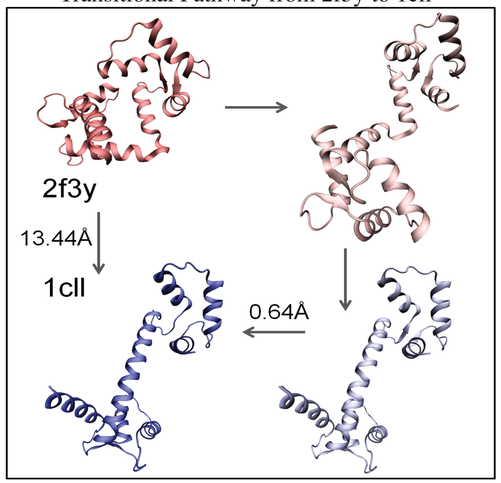
[1]Department of Computer Science [2]Department of Bioinformatics and Computational Biology [3]Department of Bioengineering
George Mason University, Fairfax VA, 22030

Abstract

Understanding and elucidating the conformational rearrangements that protein systems undergo to transition between different functional states is an important but challenging problem. Molecular dynamics approaches can obtain transitional trajectories, however the computational complexity of the problem results in impractical running times.

We introduce a robotics-inspired method to connect two functional states of a protein by computing conformational pathways with credible energy profiles. Unlike work that focuses on either small systems with a few amino acids [2], or very large systems modeled at a coarse level of detail [3], this work focuses on small to medium length proteins. Foregoing dynamics allows computing conformational pathways in an efficient manner. Molecular dynamics techniques can be employed to map conformational pathways to actual transition trajectories [1].

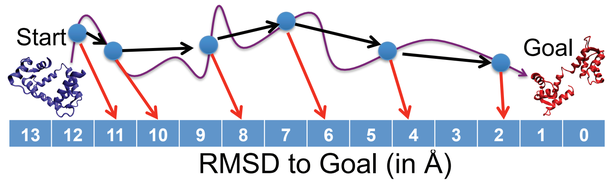
Transitional Pathway from 2f3y to 1cll



The diagram above illustrates two states of calmodulin, which are 13.44 Å apart in terms of lRMSD. Our method elucidates pathways that are less than 1 Å from the goal state.

Methods

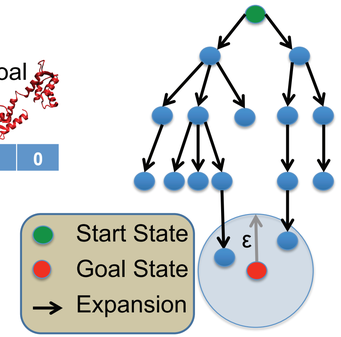
This method roots a tree at one of the functional states and biases the growth of the tree towards the other given functional state (the goal). The algorithm is comprised of two main steps: node selection and node expansion. The selection method employs a discretization layer of a progress coordinate, lRMSD to the goal (shown below). Three methods are evaluated: QUAD, LINEAR, and UNIFORM .

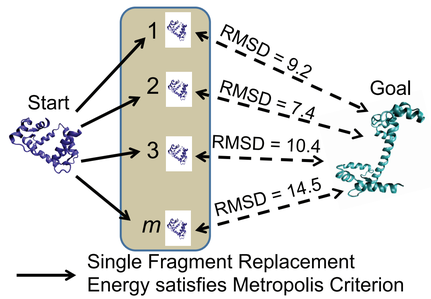


$$w(l) = \frac{1}{1+l^2} + \varepsilon \quad \text{QUAD}$$

$$w(l) = \frac{1}{1+l} + \varepsilon \quad \text{LINEAR}$$

$$w(l) = \frac{1}{\#\ of\ \_levels} \quad \text{UNIFORM}$$



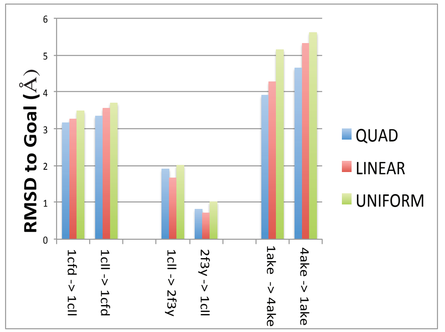


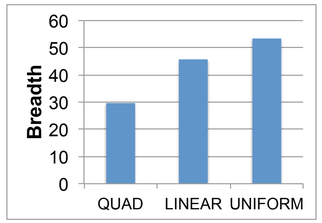
The expansion step creates *m* candidate child conformations. Each candidate is the result of performing a single molecular fragment replacement on the parent, where the resulting energy is accepted by the Metropolis criterion. From the *m* candidates, the one closest to the goal state is identified. A local bias is employed that conditionally adds this closest candidate to the tree only if it improves in proximity to the goal over its parent.

Results

The method is tested on all 6 directed pairs that can be defined over the three given states of calmodulin and between two states of AdK.

First, we analyzed the results in terms of the depth of the tree. The bar graph to the right shows that the goal state is approached within the 3 - 4 Å tolerance region using all three global bias schemes. The weighting schemes perform similarly on calmodulin, but show significant differences on AdK. The results on calmodulin are in qualitative agreement with those observed in experimentation and simulation [4][5].





$$b = \sum_{i=0}^{h} \frac{(i+1) * d_i}{h}$$

An important aspect of the search is the diversity of pathways that are found between the start and goal states. We use a heuristic (formula shown on left) to measure the diversity of conformations at each level of the tree, which is measurement of breadth. The heuristic measures the maximum pairwise distance in terms of RMSD at a given level of the tree. This metric reduces the weight of differences in levels closer to the goal. The graph to the left illustrates that our quadratic scheme has the lowest amount of diversity, while our linear and uniform selection schemes increase the diversity.

The linear biasing offers a good compromise between low lRMSD to the goal and diversity.

Conclusions

We present a robotics inspired exploration framework for elucidating energetic credible conformational pathways between two functional states of a protein. Molecular fragment replacement allows the tree to grow with physically-realistic conformations while reducing the size of the search space. Future work will consider the role of temperature in allowing the sampling to cross energy barriers in complex systems (like AdK). The employment of alternative progress coordinates will also be investigated.

Acknowledgements

Grant No. 1016995

References

1 P. Majek, H. Weinstein, and R. RIber, "Pathways of conformational transitions in proteins", Taylor and Francis group, 2008, ch 13, pp. 185-203.

2 L. Jaillet, F.J. Corcho, J.-J. Perez, and J. Cortes. "Randomized tree construction algorithm to explore energy landscapes", *J Comput. Chem*, vol. 32, no 16 pp 3464-3474, 2011.

3 N. Haspel, M. Moll, M. L. Baker, W. Chiu, and K.L. E., "Tracing conformational changes in proteins", *BMC Struct. Biol*, vol 10, No Suppl 1, p S1, 2010.

4 J. Evenas, S. Forsen, A. Malmendal, and M. Akke, "Backbone dynamics and energetics of a calmodulin domain mutant exchanging between closed and open conformations,", *J. Mol. Biol.*, vol. 289, no. 3, p. 603-617, 1999.

5 B.W. Zhang, D. Jasnow, and D. M. Zuckermann, "Efficient and verified simulation of a path ensemble for conformational change in a united-residue model of calmodulin", *Proc. Natl. Acad. Sci USA*, vol 104, no. 46, pp 18043-18048, 2007.