

BY KATHARINE MILLER

## Visualizing Markov State Models Using MSMExplorer

An unfolded protein can move through thousands of intermediate structures (conformations) before finding its properly folded state. One approach to understanding this process involves simulating a protein's behavior with molecular dynamics, and then analyzing the resulting trajectories statistically to build a Markov state model or MSM, which describes a series of conformational states and the probabilities of transitions between them. Interpreting these MSMs can lead to new insights. Now, a new application is available to address a key bottleneck that has slowed the process of interpretation: the lack of a simple, automated way to visualize MSMs.

“With a picture, you gain a rapid way to qualitatively analyze MSMs,” says **Bryce Cronkite-Ratcliff**, an undergraduate in computer science at Stanford University. He worked with **Vijay Pande, PhD**, professor of chemistry, structural biology and computer science at Stanford, to create MSMExplorer, an application for visualizing MSM data. “It can guide further quantitative analysis in new directions,” Cronkite-Ratcliff says. And as a bonus, the application can generate graphics for publications and posters, a task that most MSM researchers currently do by hand.

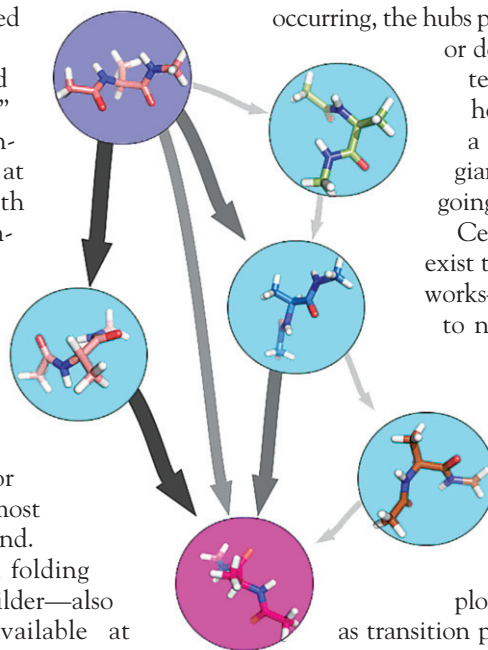
To create an MSM of a protein folding simulation, a program like MSMBuild—also out of Pande's lab and also available at [simtk.org](http://simtk.org)—runs simulations to watch how often a protein goes from one state to another, and clusters intermediate structures based on kinetic proximity—*i.e.*, how energetically easy it is to change from one structure to another. The resulting data matrices must then undergo extensive quantitative analysis to be understood. Visualizing the data as a network can guide researchers to ask relevant quantitative questions. “MSMExplorer is not a replacement for quantitative analysis, but it lets you see overall patterns that are otherwise hard to tease out of the numbers,” Cronkite-Ratcliff says.

A few years ago, for example, **Greg Bowman, PhD**,

then a graduate student in Pande's lab, crunched the numbers for a lot of MSMs and discovered the presence of kinetic hubs. These kinetic hubs represent a new and more physically accurate way to conceptualize how proteins fold, where the native (folded) state plays the role of a central hub (gathering point) kinetically, not just thermodynamically. The work was very labor intensive. But if MSMExplorer had existed at that time, Cronkite-Ratcliff says, the discovery would have been much more obvious. Because the visualization displays various protein-folding states and their connections in sizes that reflect their likelihood of occurring, the hubs pop out. “It's hard to look at a matrix or do calculations on a matrix and detect whether there's a kinetic hub,”

he says. “But it's pretty easy to open a picture and see whether there's a giant circle in the middle with spokes going out from it.”

Certainly other software programs exist that can build visualizations of networks—including Cytoscape and Gephi, to name just two. But MSMExplorer knows how to collect the MSM files it needs and automatically turn them into a useful visualization “right off the bat,” Cronkite-Ratcliff says. It also does a few things those other programs can't. For example, Cronkite-Ratcliff says, MSMExplorer includes some custom tools such as transition path theory (TPT) algorithms that



**MSMExplorer uses transition path theory (TPT) to depict the likelihood of a protein folding via various intermediates. Courtesy of: Bryce Cronkite-Ratcliff.**

extract and display the likely paths a protein will take to get from node A to node B, as well as how much activity occurs along each path. The nodes can also be displayed as images of the folded states they represent, a feature that can help users produce publication-worthy graphics. “You can make them look less like a hairball,” Cronkite-Ratcliff says.

According to Pande, the MSM technique is already used in 20 to 30 labs around the world and many others are interested in adopting it. “A way to visualize MSMs will empower many researchers now and many more in years to come,” he says. □



### DETAILS

**MSMExplorer is available for download at <https://simtk.org/home/msmexplorer> along with a walkthrough and video demonstration.**