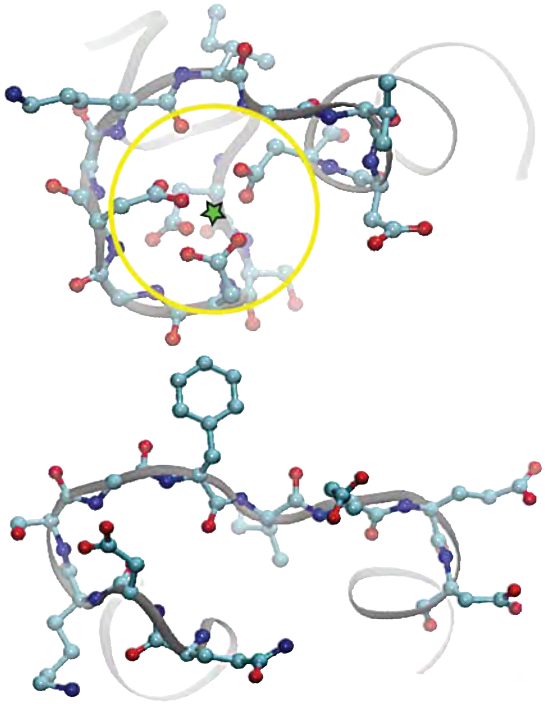
SimbiosNews

BY JOY KU, PhD

Enhanced Function Recognition in Protein Trajectories over Space and Time



Glazer's study showed that molecular dynamics simulations can increase the accuracy of predictions about a molecule's functional sites, as compared with a static image. Shown above are two frames from a simulation. (top) A molecular configuration with a likely calcium-binding site, shown by the green star. (bottom) The same molecule from a different time point in the simulation that is unlikely to have any calcium-binding sites. A static image would show the molecule in only one of the configurations.

If a picture's worth a thousand words, then a motion picture, such as that provided by molecular dynamics (MD) simulations, must contain a wealth of information. It's this potential payoff that has Simbios investing in the infrastructure to speed up these simulations (see last issue's article on OpenMM) and to store and share the trajectories that are generated.

Research, such as that of **Dariya Glazer,** a graduate student in genetics at Stanford University whose work is partially supported by Simbios, illustrates the insights that MD simulations can provide. In recent work, she demonstrated that simulating how a molecule moves over time can lead to more accurate predictions of molecular functional sites, such as active enzymatic or drug-binding sites. Her poster based on this work received an Outstanding Poster Award at the Intelligent Systems for Molecular Biology (ISMB) conference in Toronto in July 2008.

"Dariya has shown that simulating the motion of these proteins using molecular dynamics can markedly improve the ability to detect function and should probably be routinely employed by these algorithms," says **Russ Altman, MD, PhD**, Glazer's advisor who is also a professor of bioengineering at Stanford University and a principal investigator for Simbios.

Most function prediction algorithms use data from experimental techniques, such as X-ray crystallography, which show a molecule's structure at one particular moment in time. But the molecule may not be in a functional configuration at that instant, resulting in incorrect predictions.

"Unfortunately, most prediction methods forget that the beautiful crystal structure of a protein is nothing but a snapshot of what the actual protein looks like *in vivo*," says **Marco Punta, PhD**, a research scientist in biochemistry and molecular biophysics at Columbia University and the posters committee chair at ISMB this year.

To increase her odds of making accurate predictions, Glazer used MD simulations to model molecular motion. From the resulting molecular trajectories, she extracted hundreds of frames and then applied traditional function prediction methods to each of those frames to identify potential binding sites for calcium.

For one protein she examined, the static structure offered no hints of a binding site. "It was only during the simulations that the binding site adopted an appropriate conformation so that the algorithm could identify it," she says.

After running the MD simulations, Glazer faced a new problem: how to combine the information from the tens to hundreds of structures that presented with potential binding sites. "It wasn't always obvious to us whether the data represented a single binding site or several." So she came up with a multi-tiered clustering scheme to identify the independent sites.

With the MD simulations and her new clustering scheme, Glazer was able to identify upwards of 60 percent more true calcium binding sites. In fact, one of the prediction methods could not identify any binding sites at all without the MD.

Glazer is looking forward to experimenting with MD simulations on graphics processing units (GPUs), something that Simbios is making possible. "I'll be able to investigate more complicated functions in larger systems," she says. "The possibilities are exciting." □

DETAILS

Glazer tested her approach using two different prediction algorithms: FEATURE (http://simtk.org/ home/feature), which looks at about 80 different properties to make its prediction; and a valence method, which uses a molecule's local charge to determine where it might bind.

Her test cases consisted of both a calcium-bound and a non-calcium-bound version of five different proteins. The simulation trajectories generated for these molecules will be made available at http://simtk.org/home/mdfxnpredict.

Simbios (http://simbios.stanford.edu) is a National Center for Biomedical Computing located at Stanford University.