Scientists unlock secrets of protein folding

September 17 2007

A team led by biophysicist Jeremy Smith of the University of Tennessee and Oak Ridge National Laboratory has taken a significant step toward unraveling the mystery of how proteins fold into unique, three-dimensional shapes.

Using ORNL's Cray XT4 Jaguar supercomputer as well as computer systems in Italy and Germany, the team revealed a driving force behind protein folding involving the way its constituents interact with water. The team's results are being published in this week's edition of the *Proceedings of the National Academy of Sciences*.

Proteins are the workhorses of the body, taking on a wide variety of tasks. They fight infections, turn food into energy, copy DNA and catalyze chemical reactions. Insulin is a protein, as are antibodies and many hormones.

Scientists are still very interested in deciphering how proteins work.

A protein is a string of amino acids, and what it does is determined by the shape it takes. That shape is determined by the sequence of the amino acids. Like a piece of biological origami, the protein folds itself into the form necessary to carry out its job. Without the shape the protein would be worthless.

"Understanding the mechanism by which proteins fold up into unique three-dimensional architectures is a holy grail in molecular biology," explained Smith, who holds the first UT-ORNL Governor's Chair and is

a member of the Biochemistry and Molecular Biology Department at UT.

"Unfortunately, if you give me the sequence of amino acid building blocks in the protein, I cannot tell you what the structure would be," he said. "If I had been able to do that with a computer a while ago, the work behind about a dozen Nobel prizes -- those awarded for experimental work on protein structure determination -- would not have been necessary."

Working on a smaller chain of amino acids known as a peptide, the group showed that the folding is determined largely by how parts of the peptide interact with water. Areas that shun water are said to be hydrophobic, and the team's results show that the way water wets these hydrophobic areas determines the ultimate shape and behavior of the peptide.

In particular, the team determined that small hydrophobic areas of the peptide, up to the size of a water molecule, induce different behavior in water than larger hydrophobic areas, and that this difference is crucial for the folding. This insight builds on the work of another team, based at the University of California–Berkeley.

"David Chandler and his colleagues at Berkeley have a theory stating that hydrophobicity is qualitatively different on different length scales," Smith said. "If you have small hydrophobic molecules or groups that are themselves roughly the size of a water molecule, the water doesn't seem to be too bothered by these groups. But when you get hydrophobic entities as long as several water molecules, the water molecules have a problem with that. They can't cloak themselves around the hydrophobic surface anymore, and there is a dewetting or drying effect as they are repelled from the surface.

"Our simulations have shown that Chandler's theory works for peptides, and, moreover, that the drying effect determines which structure our peptide adopts. It's kind of 'dry it off then fold it up.'"

Smith said his team's achievement was made possible by high-performance computing, noting that Jaguar is currently rated the second most powerful computing system in the world. Smith also said that his team will need increasingly powerful supercomputers for additional simulation. While the team so far has been able to simulate about a microsecond in the life of a peptide, they must eventually be able to increase that time a thousand-fold, to milliseconds, and simulate proteins that are 10 to 100 times as large as the peptides.

"The runs were a couple of microseconds, which was adequate for the peptide that was simulated," Smith explained. "But the team is looking forward to increased computing capacity as it moves forward. The technique used is molecular dynamics simulation, and it needs high-performance leadership supercomputing to reach the length and timescales needed to fold a complete functional protein in the computer. With the projected capability improvements in Jaguar over the next couple of years, we will soon be approaching that goal."

Smith made it clear that the achievement would represent a watershed in the field.

"When we do eventually find out how to calculate protein structure from sequence," he said, "then a major revolution will come upon us, as we will have the basis to move forward with understanding much of biology and medicine, and the applications will range from rationally designing drugs to fit clefts in protein structures to engineering protein shapes for useful functions in nanotechnology and bioenergy."

Source: University of Tennessee