Valerie Daggett, University of Washington

Unraveling Protein Folding:

High-Performance Computing Helps Shed Light on Diseases

Proteins are the workhorses of the body, and the enigmatic process by which a protein folds into a specific three-dimensional structure has vast implications for human health. Researchers at the University of Washington are using innovative computer simulation techniques to reveal secrets of protein folding that could lead to improved diagnosis and treatment of diseases such as Alzheimer's, Parkinson's and many forms of cancer.

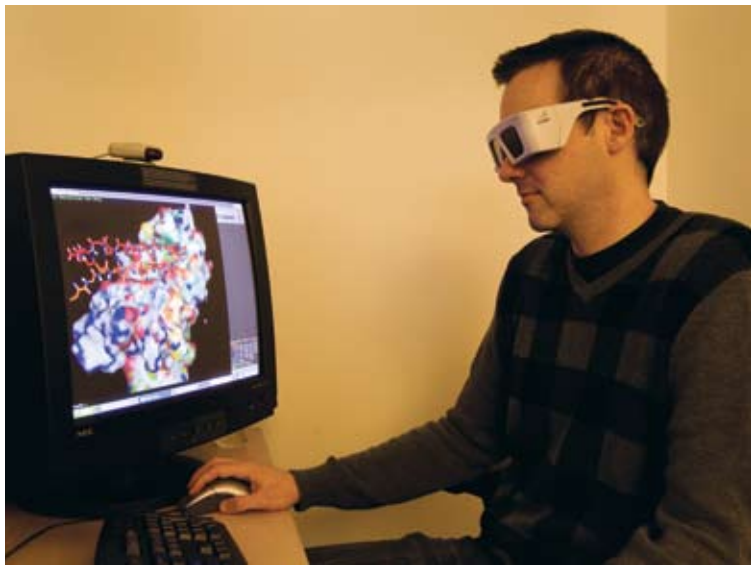
More than a century after German physician Alois Alzheimer first published his observations about a progressive brain disorder that impaired patients' memory, thinking and behavior, Alzheimer's disease remains an incurable and largely baffling condition. Roughly 26 million people worldwide are afflicted with Alzheimer's, and the number of cases is projected to quadruple by 2050, according to a 2007 study by the Johns Hopkins Bloomberg School of Public Health.[1]

A research team at the University of Washington in Seattle is working to shed more light on the causes and potential treatments of neurological diseases such as Alzheimer's by unraveling—literally and figuratively—the mysterious process by which proteins fold themselves into unique, three-dimensional structures that dictate their specific function. While scientists have determined that incorrectly folded proteins are responsible for Alzheimer's and a host of other illnesses such as Parkinson's disease, cystic fibrosis, Creutzfeldt-Jakob disease and many forms of cancer, little is known about how a protein undergoes this transformation.

"We are using computer simulations and other experimental methods to precisely characterize the structural changes that occur in protein folding, particularly changes that can trigger the onset of a disease," says Valerie Daggett, a professor of bioengineering who leads the Daggett Research Group at the University of Washington. Daggett and her colleagues believe that the insights gained from these simulations will help uncover general rules of protein folding that medical researchers can use to create more successful diagnostic techniques, drug therapies and disease prevention strategies.

Software, financial and technical support from Microsoft External Research is helping the Daggett Research Group carry out its computer simulations, expand its database of protein folding images and make the results available broadly for other scientists to study. For example, the group uses high-performance computing systems built on Microsoft Windows® Compute Cluster 2003, Windows Server® 2008, SQL Server® 2008 and SQL Server Analysis Services to handle the massive computational demands involved.

"Our protein simulations and structures involve hundreds of terabytes of data," says Daggett. "Before Microsoft became involved, there really wasn't a way for us to handle it all—we had data separated over hundreds of computer disks." The team could analyze a



Gene Hopping, a senior research fellow at the University of Washington, uses 3-D computer imaging technology to study a simulation of the protein folding process.

small number of proteins in detail but struggled to compare details about the unfolding process across all simulations.

"With this high-performance computing framework and the collaborative support that Microsoft Research provides, we can now ask bigger-picture scientific questions that enable us to really go after these diseases," says Daggett. "We have also been able to streamline our work because we're much better organized."

The Daggett Research Group has compiled nearly 5,000 independent simulations of more than 650 protein folds—roughly half of all known protein structures—as part of its work in the field of molecular dynameomics. This research seeks to bridge a gap in scientists' understanding of the transitions and structures that occur in a protein from the time it is created to when it achieves its folded state, a process that takes place in microseconds to minutes.

"The static structure of a protein doesn't show you anything about how it moves or behaves, which is critical in order to really understand the protein's function and its pathology if it is associated with a disease like Alzheimer's or cancer," explains Daggett. "I believe that this lack of dynamics is one of the biggest obstacles to more effective drug design."

To simulate pathways of the protein folding process, Daggett's team actually works in reverse: starting with a protein in its native, folded structure and applying heat to make it unfold. The group's research has shown that protein unfolding follows the same pattern as folding, and creating a computer simulation from an existing structure requires far less sampling of structures.

Multiple different simulations are needed in order to capture an adequate record of the unfolding process. To perform these simulations, the Daggett Research Group developed a software code called "in lucem Molecular Mechanics" (ilmm) that runs on Windows Compute Cluster Server 2003 and employs Isaac Newton's equations of motion to chart the path of every atom in the protein molecule over time.

As they home in on the folding pathways of proteins and begin identifying significant structural changes along the way, such as the instant when a protein misfolds and potentially becomes toxic, biomedical researchers will be able to investigate the underlying causes of those changes. "Then we can create diagnostic agents and drug therapies that target specific protein structures," says Daggett.

"With diseases such as Alzheimer's, the symptoms usually occur only after there has already been a great deal of damage," she adds. "If physicians can detect the onset of the disease sooner and introduce therapy to inhibit protein misfolding or toxicity, a cure should someday be possible."

Daggett's team is in the second phase of testing diagnostic and therapeutic compounds to target the protein structures that

contribute to Alzheimer's, Creutzfeldt-Jacob and bovine spongiform encephalopathy—commonly known as mad cow disease. In addition to their potential for inhibiting the structural changes that make a protein toxic, the compounds could eventually yield more effective methods of screening food supplies and blood supplies for infectious forms of these diseases.

Byron Caughey, a senior investigator in the Rocky Mountain Laboratories at the U.S. National Institutes of Health, says Daggett's computer modeling of protein structures is contributing to significant gains in biomedical research.

"Some of the most persistent problems in protein science involve understanding how proteins misbehave by misfolding and assembling into toxic aggregates," says Caughey. "Computational simulations are especially valuable in studying the process of protein misfolding and its connection to diseases. Valerie has made groundbreaking contributions to these efforts."

Many of the computer simulations and related data are available on the Daggett Research Group's Web site for general use, and team members collaborate regularly with other researchers who are studying specific proteins. The research group has also applied for patents related to its work.

Although Daggett cautions that breakthroughs in disease diagnosis and treatment stemming from her team's protein research are still years away from reaching healthcare professionals and patients, she is encouraged by the initial results.

"A growing number of diseases are associated with protein unfolding, so this is a place where we can really intervene," she says. "Understanding the changes involved could have tremendous implications for millions of people who live with these conditions."

[1] Ron Brookmeyer, Elizabeth Johnson, Kathryn Ziegler-Graham and H. Michael Arrighi. "Forecasting the Global Burden of Alzheimer's Disease," *Alzheimer's and Dementia* 3.3 (2007): 186-191.

"A growing number of diseases are associated with protein unfolding, so this is a place where we can really intervene. Understanding the changes involved could have tremendous implications for millions of people who live with these conditions."

—Valerie Daggett, professor of bioengineering, University of Washington

At a Glance

Project: Molecular Dynameomics

Location: University of Washington, Seattle

Project Principal:

Valerie Daggett, professor of bioengineering, University of Washington

Web Sites:

www.dynameomics.org
depts.washington.edu/daglab

Microsoft Technologies:

Windows Compute Cluster 2003, Windows Server 2008, SQL Server 2008, Visual Studio® 2008, Internet Information Services 7