Latest News

September 19, 2005
Volume 83, Number 38
p. 7

COMPUTATIONAL CHEMISTRY

Protein Structure From Scratch

Method to generate high-resolution protein structures is best so far

Celia Henry

COURTESY OF DAVID BAKER

PROTEIN FOLDERS Misura (from left), Bradley, and Baker developed a new computational method to predict the structure of small proteins.

Scientists have known for more than 40 years that the amino acid sequence of a protein dictates its structure. Therefore, it should be possible to forecast how a protein will fold just by knowing the sequence. Yet scientists are still searching for methods that will accurately predict a protein structure from scratch. Now, biochemistry professor David Baker and postdocs Philip Bradley and Kira M. S. Misura of the University of Washington, Seattle, have developed what they believe is the best method so far for generating high-resolution protein structures from the amino acid sequence (*Science* **2005,** *309,* 1868).

Predicting protein structure involves two key challenges: calculating the protein's energy accurately and identifying which among the myriad of conformations has the minimum energy. The assumption--usually a good one--is that the native structure is the lowest-energy conformation. Because trying all possible conformations is computationally impossible, "you have to have a search that is focused on the regions where the global energy minimum might lie," Baker says.

Baker and his coworkers have devised a two-step method to find the lowest-energy conformation of an amino acid sequence. First, they do a low-resolution search of the energy landscape using only the backbones of the target protein, as well as those of 15-50 homologous proteins with similar sequences. The other protein sequences are key in the folding simulation, because they allow the method to explore other parts of the energy landscape and locate low-energy conformations that are different but similar to those that the target protein sequence reveals, Baker says. "It's a way of making this initial low-resolution sampling broader." Then, the low-energy regions found are used in a high-

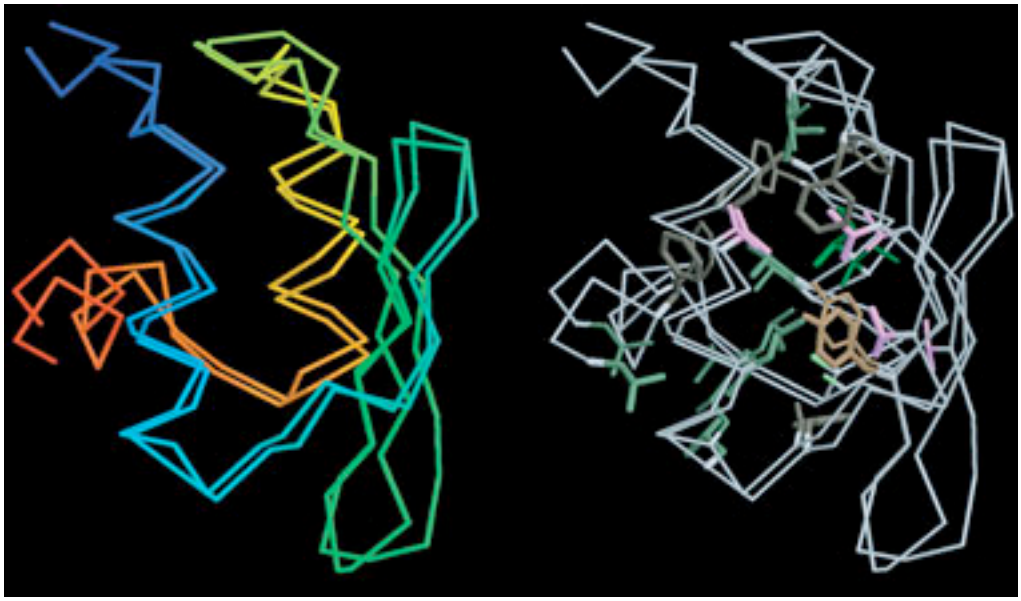
resolution energy search using only the target protein but with all of its atoms.

The team used the method to predict the structures of 16 small proteins. Among them is the structure of a protein--from *Thermus thermophilus*--used in a competition called the Sixth Critical Assessment of Techniques for Protein Structure Prediction. The team predicted its folded structure with a resolution of 1.6 Å without prior knowledge. The result "showed the best accuracy by far for an ab initio prediction in the competition's 10-year history," says competition organizer Tim Hubbard of the Wellcome Trust Sanger Institute in Hinxton, England.

The method is far from perfect: It could predict the structure accurately for only six of 16 proteins the team tried to solve. The problem, Baker says, is one of sampling rather than energy calculation. "Sampling is a major bottleneck to being able to predict structure more generally."

Baker is continuing to improve the accuracy of the method. "Before the method is really ready for prime-time general use, it's got to be more accurate and more consistent," he says. "It's not a replacement for real structural biology like NMR spectroscopy or X-ray crystallography until essentially every time you build a model and take your lowest-energy structure, it's correct."

The method so far is limited to proteins with fewer than 85 amino acids, but Baker is optimistic. "It's not a conceptual problem with larger proteins," he says. "With more computing power and maybe better algorithms, it should be possible."



COURTESY OF DAVID BAKER

PREDICTION The computational method devised by Baker and coworkers predicted the structure of a *Thermus thermophilus* protein with 1.6-Å resolution. The backbone trace is shown on the left, and the core side chains are shown on the right.

Chemical & Engineering News
ISSN 0009-2347
Copyright © 2005