Big moves in protein structure prediction and design

December 16 2015

A modeling the molecular structure of proteins in the Institute of Protein Design at the University of Washington, where advanced computer algorithms for protein structure prediction are continuously being refined. Credit: Institute for Protein Design/University of Washington

The potential of modular design for brand new proteins that do not exist

in the natural world is explored Dec. 16 in the journal *Nature*. The reports are the latest in a recent series of developments toward custom-designing proteins.

Naturally occurring proteins are the nanoscale machines that carry out nearly all the essential functions in living things.

While it has been known for more than 40 years that a protein's sequence of amino acids determines its shape, it has been challenging for scientists to predict a protein's three-dimensional structure from its amino acid sequence.

Conversely, it has been difficult for scientists to devise brand new amino acid sequences that fold up into hitherto unseen structures. A protein's structure dictates the types of biochemical and biological tasks it can perform.

The *Nature* papers look at one type of natural construction: proteins formed of repeat copies of a structural component. The researchers examined the potential for creating new types of these proteins. Just as the manufacturing industry was revolutionized by interchangeable parts, originating protein molecules with the right twists, turns and connections for their modular assembly would be a bold direction for biotechnology.

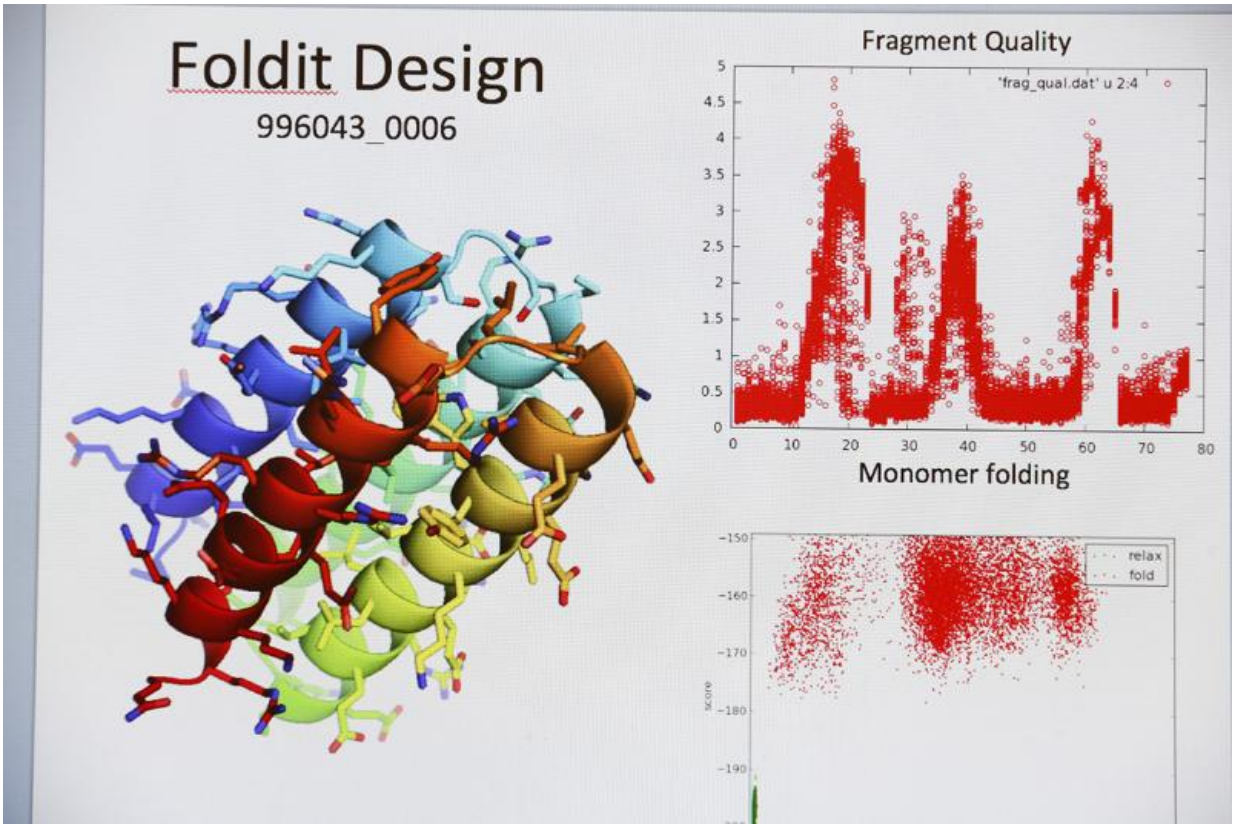
The papers are 'Exploring the repeat protein universe through computational design' and 'Rational design of alpha-helical tandem repeat proteins with closed architecture.' The findings suggest the possibilities for producing useful protein geometries that exceed what nature has achieved.

The work was led by postdoctoral fellows TJ Brunette, Fabio Parmeggiani and Po-Ssu Huang in the lab of David Baker at the University of Washington Institute for Protein Design and Lindsey

Doyle and Phil Bradley at the Fred Hutchinson Cancer Research Institute in Seattle.

In addition, over the past several months, researchers at the Institute for Protein Design at the University of Washington, the Fred Hutch, and their colleagues at other institutions have described several other advances in two long-standing problem areas in building new proteins from scratch.

"It has been a watershed year for protein structure predictions and design," said UW Medicine researcher David A. Baker, UW professor of biochemistry, Howard Hughes Medical Institute investigator, and head of the UW Institute for protein design.



Volunteer citizen scientists around the world participate in research on protein structure prediction through the computer program Foldit. Credit: Institute for protein Design/University of Washington

The protein structure problem is figuring out how a protein's chemical makeup predetermines its molecular structure, and in turn, its biological role. UW researchers have developed powerful algorithms to make unprecedented, accurate, blind predictions about the structure of large proteins of more than 200 amino acids in length. This has opened the door to predicting the structures for hundreds of thousands of recently discovered proteins in the ocean, soil, and gut microbiome.

Equally difficult is designing amino acid sequences that will fold into brand new protein structures. Researchers have now shown the possibility of doing this with precision for protein folds inspired by naturally occurring proteins.

More importantly, researchers can now devise amino acid sequences to fashion novel, previously unknown folds, far surpassing what is predicted to occur in the natural world.

The new proteins are designed with help from volunteers around the globe participating in the Rosetta@home distributed computing project. The custom-designed amino acid sequences are encoded in synthetic genes, the proteins are produced in the laboratory, and their structures are revealed through X-ray crystallography. The computer models in almost all cases match the experimentally determined crystal structures with near atomic level accuracy.

Researches have also reported new protein designs, all with near atomic level accuracy, for such shapes as barrels, sheets, rings and screws. This

adds to previous achievements in designing protein cubes and spheres, and suggests the possibility of making a totally new class of protein materials.

By furthering advances such as these, researchers hope to build proteins for critical tasks in medical, environmental and industrial arenas. Examples of their goals are nanoscale tools that: boost the immune response against HIV and other recalcitrant viruses, block the flu virus so that it cannot infect cells, target drugs to cancer cells while reducing side effects, stop allergens from causing symptoms, neutralize deposits, called amyloids, thought to damage vital tissues in Alzheimer's disease, mop up medications in the body as an antidote, and fulfill other diagnostic and therapeutic needs. Scientists are also interested in new proteins for biofuels and clean energy.

In addition to this week's report on modular construction of proteins with repeating motifs, here are some other recent developments:

- Evolution offers clues to shaping proteins: The function of many proteins tends to stay the same across species, even after their amino acid sequences have changed over billions of years of evolution. Locating co-evolved pairs of amino acids helps calculate their proximity when the molecule folds. UW graduate student Sergey Ovchinnikov applied this co-evolution DNA sequence analysis in an *E-Life* paper published Sept. 3, 2015, 'Large-scale determination of previously unsolved protein structures using evolutionary information.' The effort illuminated for the first time the structures of 58 families of proteins that have hundreds of thousands of additional, structurally related family members.

  "This achievement was a grand slam home run in the history of protein structure prediction," said Baker.

- Barrel-fold design: Baker lab postdoctoral fellow Po-Ssu Huang, together with Birte Höcker at the Max Planck Institute for Developmental Biology in Germany discovered the elusive design principles for a barrel-shaped fold underpinning many natural enzyme molecules. The custom designed barrels folds, built at the Institute for Protein Design, were presented Nov. 23, 2015 in the *Nature Chemical Biology* paper, 'De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy.'

  This achievement has opened the door for bioengineers to generate totally new enzymes that speed up chemical reactions by positioning smaller molecules in custom barrel compartments.

- Self-assembling apparatus: Ordered protein arrays along a flat plane are found in bacteria, the heart, and other muscles. Overcoming protein interaction complexities, researchers at UW Institute for Protein Design and the Janelia Research Campus of the Howard Hughes Medical Institute programmed proteins to self-assemble into novel symmetric, 2-dimensional sheets of protein lattice patterns. UW graduate student Shane Gonen in the Baker lab, with his brother Tamir Gonen at Janelia, described their work in the June 19, 2015 issue of *Science*, 'Design of ordered two-dimensional arrays mediated by non-covalent protein-protein interfaces.'

  This research has applications for self-assembling protein nanomaterials, especially efficient sensors or light harvesters.

- Precision sculpting: Protein designers are continuously refining principles for fashioning ideal protein structures. A paper in the Oct. 6, 2015, *Proceedings of the National Academy of Sciences*, 'Control over overall shape and size in de novo designed proteins'

, further explains methods for systematically varying protein architecture inspired by nature. Such finesse could optimize designed proteins to take on the proper shapes for desired functions. This work was led by Baker lab graduate student Yu-Ru Lin in collaboration with Nobuyasu Koga at the Institute for Molecular Science in Japan.

More information: TJ Brunette et al. Exploring the repeat protein universe through computational protein design, *Nature* (2015). DOI: 10.1038/nature16162

Lindsey Doyle et al, Rational design of α-helical tandem repeat proteins with closed architectures, Nature (2015). DOI: 10.1038/nature16191

Provided by University of Washington