

MARCH 20, 2008

Building Enzymes from Scratch

In an undertaking so computationally demanding it required donated computer time from thousands of people around the world, Howard Hughes Medical Institute (HHMI) researchers have designed and built two functional enzymes never seen in nature.

Enzymes are nature's catalysts, and without them, vital biological tasks like converting sugar to energy or replicating DNA would take cells billions or even trillions of times longer than they do. Researchers have long sought to mimic nature's efficiency by creating custom enzymes that speed up sluggish industrial processes in the production of pharmaceuticals and fuels. The molecules' complexity, however, has hampered their efforts.

"Enzymes are some of nature's most miraculous creations, but we still don't really know how they work. Trying to create new ones from scratch will really tell us what is critical for something to be a good catalyst."

— David Baker

The HHMI-led team leveraged its understanding of the basic rules that govern protein function to overcome the problems and create two un-natural functional enzymes from scratch.

According to senior author David Baker, an HHMI investigator at the University of Washington, the methods can, in principle, be applied to any chemical reaction, and may lead to advances in pharmaceutical manufacturing, toxic waste cleanup, and many other fields.

The research team published its findings in two papers, one in the March 7, 2008, issue of *Science*, and the second March 19, 2008, in an advanced online publication of the journal *Nature*.

Like other proteins, enzymes are constructed from long chains of amino acids. Twenty amino acids are the basic structural building units of proteins and each has different properties. As a protein is synthesized, it folds spontaneously into a precise three-dimensional shape that represents a

balance between the repulsive and attractive interactions of the atoms in its amino acids and the water molecules that surround them.

An enzyme's shape is critical to its function because it creates a crevice called an active site specifically shaped to bind the enzyme's target molecule. Once the target molecule binds, atoms lining the active site interact with it - snipping starch into individual glucose molecules, for example. Without a precise fit, the enzyme doesn't work.

For their enzyme design project, the team chose two model reactions from the world of chemistry,. Chemists had studied these for a while, said Baker, so there was a pretty good idea of what would be needed [in an active site] to catalyze them. According to Baker, the key to creating their novel enzymes was designing an amino acid sequence that would fold up to create that active site.

Predicting the form a given amino acid chain will take is a specialty of Baker's. In October 2007, his team reported significant progress in the prediction of the structure of natural proteins based solely on their amino acid sequences. To make the predictions, Baker's team used Rosetta, a computer program they developed to model the atomic interactions that govern protein shape. However, said Baker, the calculations involved require very large amounts of computer time. So much time is needed that Baker needs access to thousands of computers.

When he first started predicting protein structures, Baker used computer clusters in his laboratory. Several years ago, he realized that the amount of computing required to make significant advances was far beyond what his laboratory could afford. But the computing capacity Baker needed was all around him, in the homes and businesses of ordinary people.

So Baker created Rosetta@home, an online community that pairs Rosetta to the Berkeley Open Infrastructure for Network Computing (BOINC). BOINC divvies up the calculations into manageable chunks and sends the chunks off to an army of volunteers around the world who donate their computer downtime to folding proteins. Today, Rosetta@home has nearly 190,000 members.

For the enzyme design project, Baker's team, led by senior fellows Daniela Rothlesberger and Eric Althoff and graduate students Lin Jiang and Alex Zanghellini, designed active sites they thought would speed up the chemical reactions. They then used the Rosetta@home network to find amino acid sequences that would fold to produce those active sites. After that step, they created actual genes encoding those amino acid sequences and inserted them into bacteria to see if the proteins they produced speeded up their reactions.

According to Baker, the enzymes worked, though not as well as those found in nature. Rather than speeding up the rates of reaction a trillion-fold, we're only getting on the order of 100,000-fold rate enhancements, he said. There is clearly something we're missing, and very important to [our research] is trying to figure out what that is.

To help, Baker's Israeli collaborators, Dan Tawfik and Olga Khersonsky (co-authors of the paper that appears in *Nature*), took one of the enzymes and forced it to evolve. Working in a test tube, the pair created thousands of versions of the enzyme with random mutations. By chance, some of these mutations sped their enzyme up. According to Baker, several rounds of directed evolution improved the enzyme's speed 200-fold, and analyzing the changes will help the team fine tune their computer models for future projects.

According to Baker, the studies' findings are important for two reasons. First, he says, enzymes are some of nature's most miraculous creations but we still don't really know how they work. This project of trying to create new ones from scratch will really tell us what is critical for something to be a good catalyst. The second, says Baker, is that there are huge numbers of important practical applications for enzymes. For example, he says, in speeding up the production of pharmaceuticals, creating new fuels, or cleaning up pollutants. One could imagine new enzymes that will help all of us, he said.