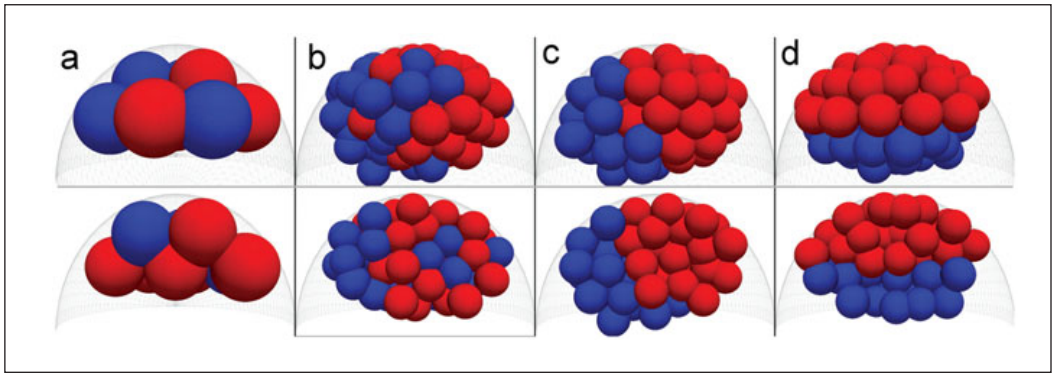
author of the paper. On top of that, they included a handful of genes, concentrations of expressed proteins and the interactions between those genetic products. They found that geometric organization, a mechanical factor, can influence the direction of division or the polarization of the cell. At the same time selected genes can affect the elasticity or adherence properties of the cell. "This connection works both ways: the genes can influence the mechanics and the mechanics can influence the genes," he says. This model, which was published in *PLoS Computational Biology* in May, provides the most inclusive simulation of these factors developed to date.

Even with this progress, researchers are just at the initial stages of building models of how cells respond to their environments. Further work will rely on close collaborations between computational and experimental researchers. "We're just scratching the surface," Zaman says. "I hope more people start to look at modeling complex behaviors of *in vivo* environments and go beyond what we already know." □



Carsten Peterson and his colleagues at Lund University in Sweden have combined genetic and mechanical signals to accurately simulate the layering of cells in the early embryo. Here, two different views (external above and cross-sectional below) show various simulations including (a) the pre-set "salt & pepper" pattern of the two cell types that determine cell orientation (GATA6 in red; NANOG in blue) next to the blastocoelic surface (gray); (b) the effect of random movements alone; (c) the effect of setting different adhesion properties for each cell type (a layer forms but positioning isn't stable or in the right place); and (d) the addition of stronger adhesion between the NANOG cells and the surrounding trophectoderm, which stabilizes the endoderm in the correct position next to the blastocoel. Reprinted from Krupinski, P, et al., Simulating the Mammalian Blastocyst—Molecular and Mechanical Interactions Pattern the Embryo, PLoS Comput Biol 7(5)1-11(2011).

DE NOVO PROTEIN DESIGN: Designing Novel Proteins that Interact

By Sarah Webb, PhD

By stringing together amino acids in a prescribed sequence that then folds into a defined structure, nature designs proteins to perform specific functions. Nowadays, computational researchers are doing some protein designing of their own—and it's bearing some valuable fruit.

The goal is to come up with new proteins to perform specific functions and recognize or bind to specific substrates, says **Jeffery Saven, PhD**, professor of chemistry at the University of Pennsylvania. "What matters is what they can do and what they can recognize," Saven says.

In nature, proteins acquire changes to their sequence of amino acids that lead to new functional forms. In the lab, researchers make chemical changes to an amino acid sequence and test it to see whether it functions in ways they can understand. But by taking the initial design work *in silico*, researchers can simplify the experimental workload by honing in on candidates worthy of laboratory work. Essentially, researchers computationally create a multitude of novel amino acid sequences, predict and build models of the new proteins' likely structures, and model or simulate how they will interact with other molecules.

Although this process is no easy matter, progress is being made, as described in a recent review by Pantazes *et al.* in *Current Opinion in Structural Biology*. Most notable, perhaps, are the efforts aimed at modeling binding to other proteins and designing new enzymes.

Designs For Binding With Hot Spots

Protein design requires overcoming the difficult challenge of getting a novel protein to bind to another protein at the correct site, in the correct orientation, and with high affinity. To address that problem, **David Baker, PhD**, professor of biochemistry at the University of Washington, and his colleagues developed a new and generalizable approach that focuses on a specific patch on a defined target. They computationally place disembodied protein side-chains next to the patch to determine how they interact in the hotspot. Only when they are satisfied with those interactions do they attach it to a protein scaffold with a shape that is complementary for anchoring the hotspot proteins. They then use computational methods to recalculate the energies and make other adjustments aimed at ensuring the appropriate "hotspot" contacts. The researchers also employ a experimental strategy, "yeast display" which expresses designed

proteins on the surface of yeast cells, allowing the researchers to test a higher number of designed proteins through flow cytometry in the laboratory than they could by traditional expression and purification methods.

In their test case, Baker and his team designed two different novel proteins that bound in the correct orientation—and quite tightly—to the hemaglutinin protein from the especially virulent H1N1 influenza strain from 1918. Those structures could serve as the basis for a new type of flu drug. The work was published in *Science* in May 2011.

Despite finding two successful proteins, it's worth noting that Baker's team lab-tested about 80 other possible designs that didn't bind. The low success rate raises questions. "We don't know what happened to the other 80 designs," Baker says. "Some might have folded into structures that didn't work." But he and others would like to be able to explain these failures, because it will help them build refine existing algorithms and build new ones that can better predict protein structures and energies. "People rarely do further analysis on failed designs: they don't have the resources, the money or the time. So we only tend to learn from success stories," says **Costas Maranas, PhD**, professor of chemical engineering at Pennsylvania State University, and a co-author of the *Current Opinion in Structural Biology* review.

By taking the initial design work *in silico*, researchers can simplify the experimental workload by honing in on candidates worthy of laboratory work.

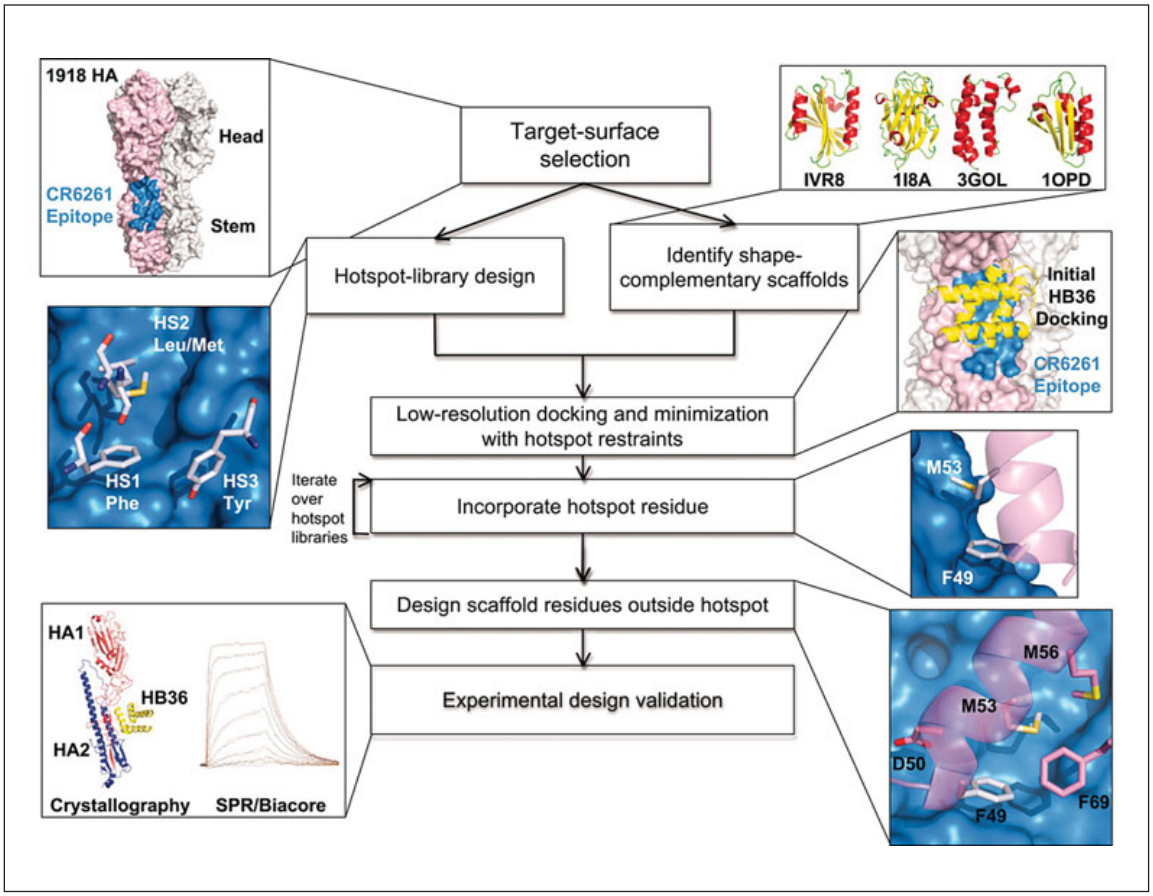
Designing Catalysts

Computational protein designers are also designing new catalysts, a problem that requires even more precision than designing binding partners. The most efficient natural enzymes speed chemical reactions up to 19 orders of magnitude (10^{19}) faster than the reaction would occur on its own. The fastest computationally designed enzymes, by contrast, enhance reaction rates by just 6 orders of magnitude (10^{6}). For example, in a 2008 *Nature* paper, Baker and his colleagues achieved this feat with a modest problem, a novel enzyme that catalyzed a Kemp elimination, a reaction not facilitated by existing biological enzymes.

Though a more difficult task than protein binding, building a new catalyst involves a similar strategy. Researchers first must design an appropriate enzyme active site with side chains that position a substrate in an appropriate position to facilitate the reaction. In this case, the researchers centered their active site around an activated serine. Then they had to graft that active site onto a protein backbone structure that would maintain those configurations. They used their RosettaMatch algorithm to sift through thousands of possible structures.

In further work published in 2010 in *Protein Science*, Baker and his team sought to explain what makes some designed catalysts more efficient than others. "There are many places where this process can go wrong," Baker says. A large part of the process is narrowing the experimental possibilities. By using molecular dynamics methods, they were able to rank their computational designs and thereby reduce the number that would need to be tested experimentally. Indeed, had they used this procedure in the 2008 *Nature* study, only 24 designs would have required experimental testing rather than 120.

Such research is a work in progress, Baker says. "What we proved is that you can make enzymes from scratch starting with a computer," he says. "Our challenge now is to make much more active catalysts." □



Flow chart of the key steps in the design of novel proteins. Reprinted with permission from Fleishman* et al.*, Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin,* Science *332:816-821 (2011).