Designer Proteins

Building machines of life from scratch

By RICHARD LIPKIN

Imagine that engines, like mushrooms, grew out of the ground.

The ancients wandering amid their fields might have stumbled upon strange iron hulks sprouting from the soil. Examining the metal contraptions, they might eventually have figured out that the intricate machines can do work, that their power can be harnessed, that they can generate electricity and move vehicles.

These early investigators probably would have taken the engines apart, determined how they were made, and learned how to operate them. Ultimately, they would have put their knowledge to the test by building one from scratch.

For today's molecular biologists, proteins offer this same allure and challenge. Often called the basic machines of life's biochemical factory, proteins carry out an extraordinary array of cellular functions and provide much of life's physical structure. Therefore, to master the intricacies of protein chemistry one must delve into life's fundamental mechanisms.

"Given that proteins are the machines of life, then to design proteins from scratch is like doing nanotechnology on the biological front," says Michael H. Hecht, a chemist at Princeton University. "To design and build a protein de novo is the ultimate test of our understanding of how proteins work. If we can learn to build these molecular machines, make them to order, and tailor them to our specifications, then we don't have to rely only on the [proteins] that nature provides."

The ability to tailor proteins to carry out very specific chemical tasks, thus harnessing their capacity to do molecular work, offers tremendous potential. Be it in medicine, industry, or environmental remediation, the customized protein beckons to biochemists as the steam engine once did to the earliest industrialists.

Recently, chemists have been reporting incremental successes at protein design, which are being hailed as milestones in the long march toward making useful biomolecules.

Among the latest achievements is the synthesis of beta-sheet proteins. Much harder to make than their alpha-helical cousins — which are long, winding strands — the beta-sheet proteins have been likened to "a greasy sandwich," with one side of each molecular sheet consisting of a slippery hydrophobic, or water-repelling, surface. Like pieces of bread spread with peanut butter, the two sheets come together — and stick.

Whereas a single alpha-helix consists of an elegant spiral of 15 to 20 amino acids, the beta sheet requires 60 to 70 amino acids. An alpha-helix protein will take shape in solution, yielding a broth of single, floating strands. Beta sheets, in contrast, are much less manageable. They form only as evenly matched pairs pressed together as a sandwich. And they must match face-to-face in just the right position.

"An alpha helix is like one hand clapping. It doesn't make much sound, but it can exist," says Bruce W. Erickson, a chemist at the University of North Carolina (UNC) at Chapel Hill. "But a beta sheet cannot exist alone. It must rest against another hydrophobic structure, most notably another beta sheet. And it's quite difficult to get them to fit together in just the right orientation."

"Imagine trying to make a sandwich by slapping two pieces of bread together from 2 feet away," he observes, "then trying to slide them into place so that the sandwich looks well made." Rather, he wants to design a beta-sheet protein whose two sandwich halves can easily find each other and fall readily into place.

For the engineered molecule to survive as a working protein and not just collapse into a garbled mess, it must fold into a specific shape. Tweaking the peptide's three-dimensional geometry into place presents one of the toughest tasks of protein design and synthesis. A successful structure depends entirely on a successful fold.

In a sense, designing an amino acid sequence that will fold into a properly shaped protein is like creasing a piece of origami paper so that, when immersed in a solution, the water's energy will fold it into a swan.

"Proteins don't really fold themselves," says Erickson. "Water folds proteins."

While some pieces of a protein molecule attract water, others are repelled by it. The art of protein design comes in putting just the right pieces in just the right places so that, merely by reacting to its watery environment, the complex chain of amino acids will collapse into exactly the right three-dimensional sculpture.

Much of a protein's crumpled form comes from the efforts of its water-repelling side chains to escape its aqueous surroundings and squeeze into the protein's protected core. To design a protein that will put itself through such contortions and then build it amino acid by amino acid requires extraordinary precision, physical intuition, and years of trial and error.

Hundreds of proteins may fail before one gives even a hint of success.

Erickson's journey into protein design began 12 years ago, when, at a conference in France on gene sequences, he and Jane S. Richardson, a biochemist at Duke University in Durham, N. C., conceived of a way to build a novel bell-shaped protein from scratch. Back in the United States, synthesis of the first such betabellin proteins began.

"The first eight molecules, which took 5 years, were failures," says Erickson. "Then betabellin 9 began to work." Refining, testing, and redesigning have taken his group up to betabellin 17.

Now Erickson, working with Yibing Yan, also a chemist at UNC, is reporting details of the design, synthesis, and characterization of several variations on the betabellin theme. One of them, betabellin 15, appears to demonstrate activity similar to that of a naturally occurring enzyme.

"We're attempting to redesign betabellin 15, using computer-assisted models, to enhance the enzymatic activity by repacking its interior so the hydrophobic residues fit together better. These are very exciting results to get after 12 years of work," says Erickson, whose most recently published findings, on betabellin 14D, appeared in the July PROTEIN SCIENCE.

"A long line of dedicated people has struggled to solve the protein's inherent structural problems," he adds, highlighting the painstaking nature of adding a bond here, a crosslink there, then purifying and analyzing the compound to see if the design worked.

"Ten years ago, we'd have given our eyeteeth to be half as far as we are today," Erickson observes.

Taking a slightly different approach to the beta-sheet conundrum, biochemist Thomas P. Quinn of the University of Missouri in Columbia has reported successfully designing and synthesizing from scratch a protein he calls a "betadoublet."

Itself a sandwich-style protein composed of two sheets pressed together, the betadoublet uses specially placed disulfide bonds as glue. Along with Duke's Jane Richardson and other colleagues, Quinn described the synthesis in the Sept. 13 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES.

"This motif is found in many naturally occurring proteins, such as immunoglobulin," says Quinn. The overall goal, he adds, is to use "design principles to make and evaluate a stable, three-dimensional beta-sandwich protein, using only naturally occurring amino acids." The end product should ideally resemble a native, or naturally occurring, protein.

For inspiration, Quinn looked to the betabellin series of proteins, but he then decided to branch off in a different direction. Instead of building the molecule piecemeal, he created a synthetic gene to manufacture beta sheets and spliced it into *Escherichia coli* bacteria. Once the bacteria expressed the gene, they cranked out beta sheets. The group then purified the proteins and searched for ones with the desired three-dimensional structures.

"There's a big hurdle we have to get over," says Quinn, "and that is to get the designed protein to have a stable, unique core." In naturally occurring proteins, dangling "side chains" normally fold neatly into place within the molecule's interior. Such a careful collapse, he says, is essential to its structure.

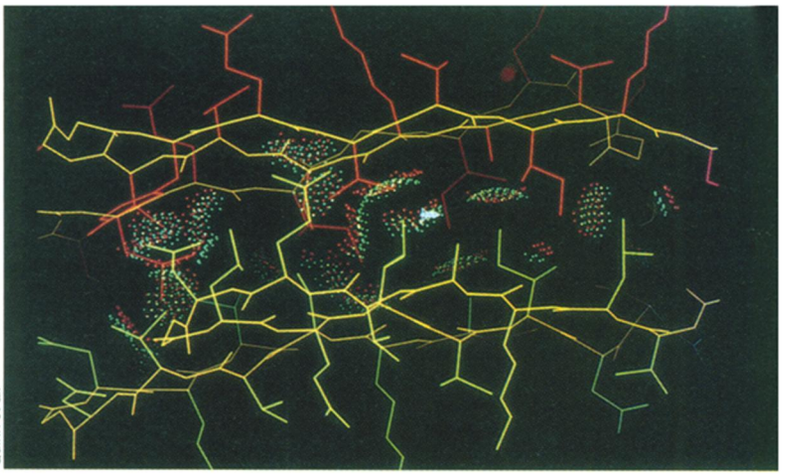
The trouble with designed proteins, Quinn points out, is that they "usually adopt the correct secondary structures and condense into compact molecules." Unfortunately, those critical side chains "often flop around, which means the core isn't quite right."

"We can get 80 percent of the correct structure, but then the core gives us trouble," says Quinn. Currently, he and his research group are attempting to redesign the betadoublet's interior so that the core locks more firmly into place.

"The core is really critical," he stresses. "A protein's structure depends on the core. It's the last piece of the puzzle. If we can get over this hurdle, we're home free."

Meanwhile, Hecht and his colleagues at Princeton are pursuing what they call a general strategy for de novo protein design. For a protein to fold properly, they hypothesize, its amino acid sequence must contain hydrophilic, or water-loving, and hydrophobic sections in very specific locations. Less important, they stress, are the details of which amino acid goes where and in what order. Less worrisome too, they maintain, are certain details of the protein's side chains in determining how the chains collapse into the molecule's core.

Instead, Hecht's team highlights the necessity of having water-attracting and water-repelling cornerstones in just the right spots along the protein's amino acid sequence, leaving the rest of the structure to take care of itself.



Quinn *et al.*

***A** side view of a designed and synthesized betadoublet protein. The protein consists of two beta sheets (yellow) held together with disulfide bonds, forming a sandwich. At the protein's center lie hydrophobic amino acid side chains (red and green), which fold into the molecule's core.*

"The exact packing of the protein's three-dimensional jigsaw puzzle does not have to be set in advance," says Hecht. "You only have to specify the sequence location, not the identity, of the hydrophilic and hydrophobic residues."

The theory thus provides a recipe for assembling a protein that will fold in a specific way.

To test the hypothesis, Hecht and his coworkers engineered 48 genes specially designed to make proteins that would fold into a bundle of four helices. The proteins that these genes encoded all had one thing in common: the same pattern of water-loving and water-repelling units in the same key spots. In the end, 29 of those genes produced stable, compact proteins that folded more or less as predicted.

"The take-home message here is that the protein took care of itself," Hecht says. "We just had to provide the driving force of basic structure and hydrophobic collapse. Then everything else fell into place."

While Hecht's general strategy so far is geared to produce a four-helix bundle, this type of approach may prove useful later on for designing beta-sheet structures or other varieties of biomolecules.

Indeed, David A. Tirrell, a polymer scientist at the University of Massachusetts at Amherst, has such visions in mind. In the Sept. 2 SCIENCE, he and his colleagues describe techniques for exploring the control of "supramolecular organization" in genetically engineered polymers.

Such polymers, he says, could create an entirely new type of material.

"Appropriately designed artificial proteins," Tirrell says, "represent a new class of macromolecular materials, with properties potentially quite different from those of the synthetic polymers currently available and in widespread use." For instance, from a carefully designed artificial gene sequence, Tirrell has successfully synthesized a layered crystal made up of stacked beta sheets.

"Trying to build proteins is a very humbling experience," Quinn says. "The alternative approach is to take a native protein, tinker with its structure, and try to modify it. That's a valid approach, but tough, too. Over millions of years, evolution has honed each protein to achieve a stable shape to do a specific job. You can alter a few amino acids, but most of the time that has little effect because of the molecule's evolutionary stability.

"If you want to learn how proteins fold in general, just from the information encoded in their amino acid sequences, then you have to free yourself from some of the confines imposed by evolution. And the only way to do that, really, is to build one from scratch and see what happens."

"It's like learning a new code of nature with a new set of rules," Quinn adds. "The DNA code only has four bases in a linear sequence. Figuring that out took a long time, but look what it made possible. With proteins, there are 20 components in three dimensions" — that is, configurations of the 20 naturally occurring amino acids.

Indeed, deciphering this chemical architecture may come about as much from building up proverbial protein sand castles as from tearing them down — the traditional process of analysis and synthesis.

"Nature has been at this design game for a long time," Erickson notes. "Here we are, trying to do in a few years and a few labs what nature has been doing for 2 billion years — namely, running an open-ended experiment with billions of living laboratories, each one an individual whose life depends on its enzymes.

"This is nature's protein-engineering project," he adds. "We call it life." □