THE AUTHOR FILE

Stan Fields and Doug Fowler

Pairing selection with sequencing, one can probe how proteins function.

For several years now, Stan Fields has wanted an easy, comprehensive tool to study a common question: how a protein's function is affected by genetic variance. He and members of his laboratory at the University of Washington thought they could design a combination DNA-protein microarray to test hundreds of thousands of slightly varying gene sequences of the same protein. Transcription and translation could be carried out on the chip, and then these arrays would show which amino acids contributed what to a protein's behavior.

Stan Fields (right) and lead author Doug Fowler.

Fields recruited postdoc Doug Fowler and graduate student Carlos Araya to turn the idea into reality. They found themselves struggling with technical detail after technical detail, trying to get the protein microarrays to work. Fowler decided to scuttle the array approach after a particularly discouraging presentation to his colleagues. "It was one of those lab meetings where you get up and say, 'I tried this, and it didn't work, and I tried this, and it didn't work'," he recalls. He was irritated by the number of hurdles necessary to accomplish something as conceptually simple as linking genotype to phenotype. "And then it hit me," he says. "The scientific community figured this out 20 years ago."

Fowler was thinking about a technique called phage display, widely used to evolve proteins with improved properties. In traditional forms of phage display, phage particles are engineered to produce a particular protein on their outer shells; thousands of variants are tested, and stringent selection rounds home in on the very few mutations that produce the best-functioning protein.

"Doug's insight was to say that you don't have to do that," says Fields. Instead of using phage display to produce a biochemical tool, this version would give a readout on a process: how various genetic sequences became more or less abundant through successive rounds of gentle selection. Sequences that became more common would represent beneficial mutations; sequences that became less common would represent deleterious ones. Whereas the Sanger sequencing typically used with phage display limits the number of sequences that can be examined practically to a few thousand, next-generation sequencing could allow hundreds of thousands of protein variants to be assessed, explains Fields, and so accomplish the goal initially envisioned for the microarrays. "That's been the driving motivation: using the ability to count DNAs to tell you whether proteins perform better or worse."

On page 741, Fields and colleagues report that the approach does indeed show how specific protein residues in a domain contribute to binding. The advance is not so much a new application of phage display, explains Fields, as a concept that can extend into many types of protein assays. He is eager to move this selection-plus-sequencing technique beyond *in vitro* phage display and into *in vivo* applications such as looking for enrichment of plasmids in cultured cells.

Pairing *in vivo* and *in vitro* assessments could provide particularly rich information. Mutations whose abundance shifts in *in vitro* screens would be likely to affect protein activities such as binding affinity or stability. However, mutations enriched or depleted only in cell-based assays might affect other properties such as localization, interaction or *in vivo* degradation. One valuable medical application could be to understand which mutations are likely to cause resistance to various drugs, says Fowler, who is now adapting the technique to study the Src kinase family of oncogenes.

Fields says he was not surprised that the approach using phage display worked on the first real try. "Biology always surprises you, but if technology works, it often works more or less the way you designed it." That is also what happened when Fields invented the yeast two-hybrid assay, the high-throughput technique for finding protein-protein interactions for which he is most famous.

The motivation, he says, is not overcoming a stumbling block to answering an intriguing biological question, but rather the pure joy of inventing a useful tool. It is something he learned working with two-time Nobel laureate Frederick Sanger, who worked out ways to sequence both proteins and DNA. "I don't think Fred ever had a biological problem that he wanted to solve," Fields says. "He just wanted to come up with technologies. And as a grad student I thought, 'if you come up with a cool method you can have a lot of impact'."

Monya Baker

Fowler, D.M. *et al.* High-resolution mapping of protein sequence-function relationships. *Nat. Methods* **7**, 741–746 (2010).