Innovations

Bind every sequence

Sangamo BioSciences, Inc.

Chemistry & Biology October 1999, 6:R281–R282

1074-5521/99/$ – see front matter

Biology is all about turning things on and off. Pharmaceuticals are part of that same picture: most protein drugs turn on or block a receptor, and most small molecule drugs turn off a protein target. Each of those recognition events is unique, because the binding surface on each protein is unique.

Sangamo BioSciences, Inc. (Richmond, California) thinks it has found a more streamlined, rational approach to this problem. Sangamo's target is DNA, which shows enough variation to allow selective targeting, but enough consistency to allow the use of a single framework for thousands of different activators and inhibitors. Within that framework — the family of zinc finger transcription factors — Sangamo is inventing a new genetic code. They hope that the code will allow them to rationally target any site in the genome and thus turn on or off any gene.

A transcription factor from Legoland

Edward Lanphier became the founder and CEO of Sangamo in 1995. By then he had seen the inside of several pharmaceutical and biotechnology companies, and had experienced the difficulties of developing gene therapy firsthand.

Lanphier was discussing gene therapy strategy at Johns Hopkins University (Baltimore, Maryland) when he came across a whole new world of opportunity: zinc finger proteins (ZFPs). "There were a lot of people [in gene therapy] working on vectors and not many people thinking about what they would deliver once it did work," says Jeremy Berg of Johns Hopkins, who is now on the Sangamo scientific advisory board. ZFPs looked like a good bet. This was based on the structure of the protein bound to DNA, which was solved first by Carl Pabo (Massachusetts Institute of Technology, Cambridge, Massachusetts) for Zif268, a mouse zinc finger transcription factor.

Each of the three zinc fingers of Zif268 recognizes a three-base subsite in the nine-base binding site. The 30-amino-acid zinc-finger is too small to fold by itself, but is glued together by a coordinating zinc atom. An α-helix from each finger lies in the DNA major groove and contacts three contiguous bases, with each finger interacting with the DNA independently. This modularity contrasts with the unpredictability of the many transcription factors that bind as dimers. Dimeric proteins require palindromic recognition sites and are sensitive to poorly understood quaternary interactions.

Sangamo develops custom-designed switches for any gene.

"What became clear is that zinc fingers are the only class of transcription factor that binds in this repeated and rational manner to DNA, and therefore the only transcription factor that is amenable to rational design," says Lanphier. "Over the years that has proven to be true." According to Berg, "they're a natural for mixing and matching."

Deciphering the code

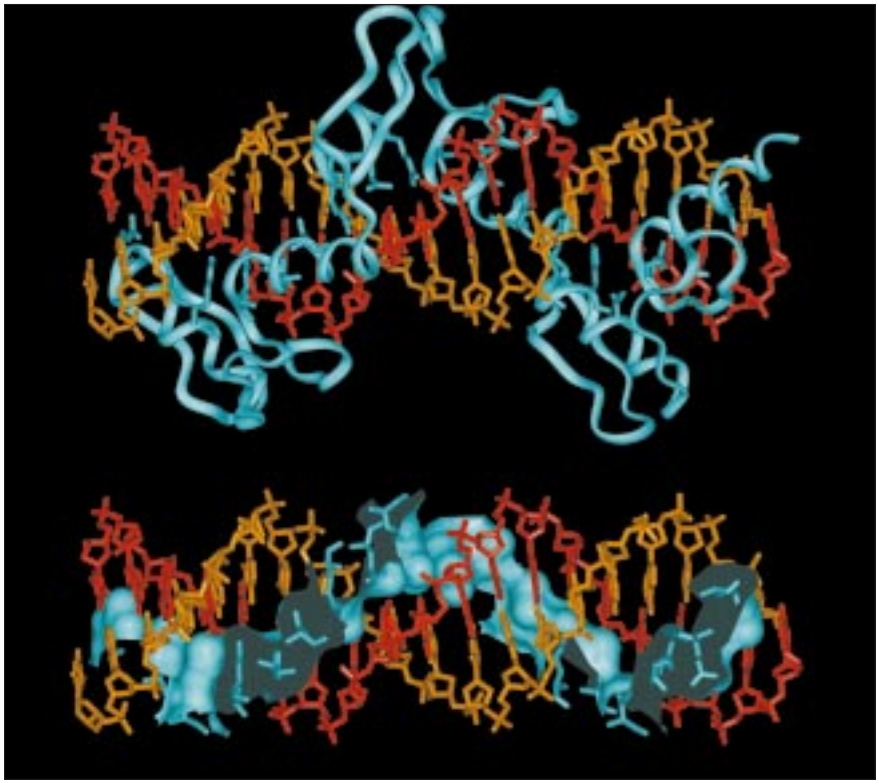
Once the structure was in, researchers like Berg, Pabo, Aaron Klug (Medical Research Council, Cambridge, UK) and Carlos Barbas (Scripps Research Institute, La Jolla, California) focused on the relationship between protein and DNA sequences. Could they come up with a one-to-one relationship, so that 64 interchangeable fingers could be designed for the 64 possible 3-base subsites?

The important contacts to DNA bases are made by three amino acids in each α-helix. Inspection of known ZFPs and structures suggested that some elements of a code existed, but far more information came from phage display experiments. Pabo, Klug, Barbas and a group from Genentech (South San Francisco, California) expressed libraries of mutated zinc fingers on the surface of phage, and tested them for their binding to DNA. Sadly, the modularity of each zinc finger was incomplete.

Pabo's response was to evolve each finger in sequence, so that it was selected in the context of its newly modified neighbor. This gave a jump in affinity, and provided invaluable information about the types of influences of one finger on the binding of a neighboring finger, but it was not practical for a company hoping for an efficient production line. Sangamo did not have the time or resources to create and screen three sequential phage libraries for each new DNA sequence. The problem only increased when Barbas introduced six-finger proteins to increase affinity (Figure 1).

Sangamo could, of course, simply ignore the inter-finger contacts. "To a first approximation I think pasting things together works," says Berg, "but then there are also more sophisticated additions that you can make." Zif268, for example, has just one key inter-finger contact — from an aspartate to the complementary cytosine in each preceding GNN subsite. Barbas has included this contact in his set of 16 fingers that he has selected for binding to GNN subsites. The result is a set of fingers that are fully modular in the context of the nearly 17 million possible $(GNN)_6$ sequences.

Sangamo has been expanding on such approaches and learning from each of its contract jobs. Lanphier says that constructing a new protein



A protein with six zinc fingers. The structures are based on the coordinates of the three-finger protein Zif268, which contacts primarily one DNA strand (shown here in red). The top view shows the backbone structure of the protein with DNA-contacting residues as stick representations. Each of the six zinc-finger modules features an α helix, which lies in the DNA major groove and functions as the reading head. The helical path of the protein is emphasized by the lower picture, which shows only the three amino acid residues from each finger that contact the DNA. Figure courtesy of David Segal and Carlos Barbas.

"is not automatic, but Sangamo has built an empirically derived database that we draw on for every target, so it's far from starting at ground zero. We're not at the point where we sit down at the computer, put in the sequence and, bang, we have a protein, but we're moving in that direction."

Hold me tight (but not too tight)

One of the unresolved questions in zinc finger design is just how much affinity is desirable. Less can be more. Barbas used site-directed mutagenesis to improve the specificity (or discrimination) of his GNN set, and found that it was the proteins with the lowest affinities that had the best specificities. Natural transcription factors bind a family of closely related sites, and it may be unreasonable to expect complete specificity from a designed protein. "If you listen to the PR and say you can regulate one site in the genome I think that's incredibly naive and probably irrelevant," says Berg. "If you bind with reasonable affinity you will almost certainly bind other sites with some affinity and you won't be approaching equilibrium."

The study of six-finger proteins has yielded clues. They show modest increases in *in vitro* affinity compared to three-finger proteins, but their biological effectiveness is far greater. Perhaps the increased affinity, by lengthening the time spent binding to DNA, makes the difference between a protein that works and a protein that binds but falls off before doing its job. Proteins that are close to this threshold will work the best in cells.

Go ahead – try it out

Theorizing is all very well, but affinity and specificity problems are sufficiently nebulous that the only true test may be experience. Plus Sangamo needs cash flow. So far the company has made deals with 11 partners to supply ZFPs that regulate a specific number of targets. These proteins are for use in functional genomics — determining the function of a gene by turning it on or off, either in cells or transgenic animals. The transcriptional activation or repression domains linked to the ZFP can function at a distance from the transcriptional start site, so the binding site for the ZFP can be within the coding sequence. Thus there is no need to isolate the promoters corresponding to expressed sequence tags (ESTs).

ZFPs are preferable to antisense because they can turn genes either on or off, and their target is a linear molecule, rather than the extensively folded mRNA. Lanphier hopes to add small-molecule control of the ZFPs in the near future, and he is investigating large-scale mutation detection with an unnamed chip company.

Sangamo hopes to hit the clinic in 2001 with ZFPs that turn on the production of well characterized proteins such as vascular endothelial growth factor (VEGF). Patent infringement should not be a problem as Sangamo is targeting the endogenous gene. Gene delivery is another matter. "N years down the road, where N is between 2 and 50, gene delivery will be standard practice," says Berg. Lanphier feels the necessary technology is already winding its way through clinical trials, but, at least in the short term, Berg remains cautious. "In the long run I think this technology could be very important for therapeutic applications," he says. "It's the pot of gold on the horizon."

William A. Wells
1095 Market Street #516, San Francisco, CA 94103-1628, USA; wells@biotext.com.