SYNTHETIC BIOLOGY

De novo–designed riboregulators

RNA-based toggle switches designed from first principles show high dynamic range and orthogonality.

Sometimes the biggest challenge for a researcher is to ignore nature's tested and tried ways. Peng Yin and his collaborators at Harvard University recently shared this experience when looking to expand the number of RNA-based regulatory components for synthetic networks.

During the past decade, nature's wide variety of RNA-based transcriptional and post-transcriptional regulators has inspired the design of riboregulators that control transcription and translation in response to an input RNA signal. James Collins, also at Harvard and one of Yin's collaborators, pioneered the design of riboregulators in 2004. By inserting a complementary sequence upstream of the ribosome-binding site (RBS) that formed a stem-loop after transcription, he created a structure that could interfere with ribosome binding. Only the binding of a small noncoding RNA, expressed in *trans*, to this stem-loop opened the structure and allowed translation. Providing or withholding the RNA trigger could thus regulate the gene.

Although this RNA-based system had advantages over protein-based regulation in that it was easier to design and exerted less of a burden on cells, its dynamic range—the output over input signal—of up to 50-fold was much lower than that of protein regulators, which can achieve around 500-fold. Another limitation was the lower specificity because of sequence constraints between *cis* sequences that have to form a secondary structure and their complementary trigger RNAs. About a fifth of all known riboswitches showed cross-talk.

Thus, Yin, Collins and their joint postdoctoral fellow Alexander Green decided to design new riboregulators from scratch, basing them on what is known about RNA-RNA interactions rather than on existing examples (Green *et al.*, 2014). Their toehold design contains a switch RNA, the gene to be regulated and an upstream hairpin that includes the RBS; but instead of requiring the trigger RNA to bind to this stem-loop, the RNA binds to a linear toehold at the 5′

STRUCTURAL BIOLOGY

USING EVOLUTION TO PREDICT STRUCTURE

Researchers use sequence coevolution information to predict the structures of protein complexes.

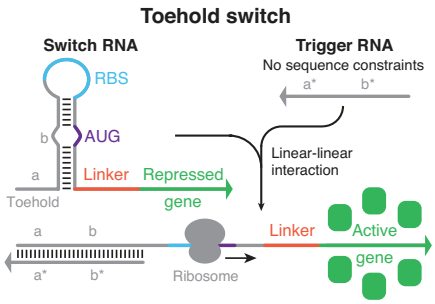
A lot can be learned about a protein's function by determining its three-dimensional (3D) structure. But structure is only part of the story: to truly understand what a protein does, one also needs to know what it interacts with, and how. Although protein interactome data continue to grow at a rapid clip, there remains little 3D structural information about the large majority of protein interactions.

Protein complexes are notoriously challenging to study using traditional structural methods such as X-ray crystallography. Hybrid approaches combining low-resolution or sparse experimental results with computational modeling are growing in popularity, but still, 3D structural information is limited for many biologically interesting complexes.

Debora Marks of Harvard University with her colleagues and collaborator Chris Sander of Memorial Sloan Kettering Cancer Center are tackling this important challenge from a different perspective. They are developing methods that use evolutionary sequence information to predict the 3D structures of proteins and protein complexes.

Marks and others have previously shown that by looking at the evolutionary record of a protein—that is, by aligning a large number of sequence homologs—one can identify coupled, coevolving residues. These coupled mutations are indicative of pairs of residues that are close in 3D space; such through-space interactions can be used as experimental constraints for structure modeling.

Marks's team reasoned that this concept could be extended to two-protein complexes by determining coevolving residues between, rather than within, proteins (Hopf *et al.*, 2014). The first step is to align sets of homologous sequences of two proteins presumed to interact. Because it is unknown whether a given protein interaction is conserved across evolutionary history, they restrict the analysis to protein pairs found in close proximity in the genome. They then link the paired sequences together and use their



Binding of the trigger RNA to the toehold opens the transcript for translation. Figure reproduced from Green *et al.*, Elsevier.

end of the hairpin, thus allowing far greater design flexibility. Upon binding of the trigger RNA to the toehold, the structure opens up and translation proceeds.

Their first generation of toehold switches included 168 varieties, of which 20 showed a dynamic range of greater than 100 when tested with a GFP reporter in *Escherichia coli*. "But," says Yin, "some did not work … so we empirically tested and distilled critical design parameters."

The second-generation riboregulators included four design changes: sequence alterations in the stem and loop as well as an increase in length of the toehold sequence and its shift further away from the RBS. The resulting toggle switches showed a 400-fold dynamic range, and only 1 of 13 switches did not work.

Although the group spent almost 2 years on the design, demonstrating different applications took only 4 months. Green used a toehold switch to sense or to regulate endogenous genes. He developed a multiplexed regulatory system in which the expression of four fluorescent proteins is regulated in parallel and designed a four-input AND gate.

The toehold switches also work *in vitro*. In a recently developed paper-based platform for diagnostics by the Collins group, a circuit containing a toehold switch is dried onto paper and upon rehydration can test for the presence of a trigger RNA (Pardee *et al.*, 2014).

Yin's group is now working on toehold switches for mammalian cells.

Nicole Rusk

RESEARCH PAPERS

Green, A.A. *et al.* Toehold switches: *de-novo*-designed regulators of gene expression. *Cell* **159**, 925–939 (2014).

Pardee, K. *et al.* Paper-based synthetic gene networks. *Cell* **159**, 940–954 (2014).

previously reported EVcouplings algorithm to statistically evaluate coevolving residues both within and between the proteins in the pairs. This new method and tool, which they call EVcomplex (http://www.evfold.org/), generates a score that indicates whether the residues predicted to interact in space are likely correct. Finally, they use the evolutionary coupling results to help determine the 3D arrangement of the two-protein complex using a protein-protein docking tool. Alternatively, another tool from the Marks lab called EVfold can be used for generating models of unknown structures.

To test their method, Marks's team made use of a recent data set consisting of 76 known 3D structures of binary protein-protein interactions in *Escherichia coli*. Using EVcouplings, they predicted contacts between the protein pairs and then generated 3D models for 15 of these complexes. Satisfyingly, 70% of the models were reasonably close to the experimentally determined structures, and known functional residue couplings were revealed. The team also predicted interprotein residue interactions for 32 two-protein complexes with unknown structures and showed that the method could predict which subunits within the large ATP synthase complex directly interact.

This work from the Marks lab follows on the heels of an approach similar in concept, though different in implementation, from David Baker's lab at the University of Washington. In May of this year, his team reported a method to predict which residues in a two-protein complex interact, also by using coevolution information (Ovchinnikov *et al.*, 2014).

A limitation of such approaches, of course, is the need for a large number of homologous sequences. But given the rapidly growing numbers of sequenced genomes, such evolutionary approaches are poised to become a powerful complement to traditional protein structure determination methods.

Allison Doerr

RESEARCH PAPERS

Hopf, T.A. *et al.* Sequence co-evolution gives 3D contacts and structures of protein complexes. *eLife* **3**, e03430 (2014).

Ovchinnikov, S. *et al.* Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* **3**, e02030 (2014).