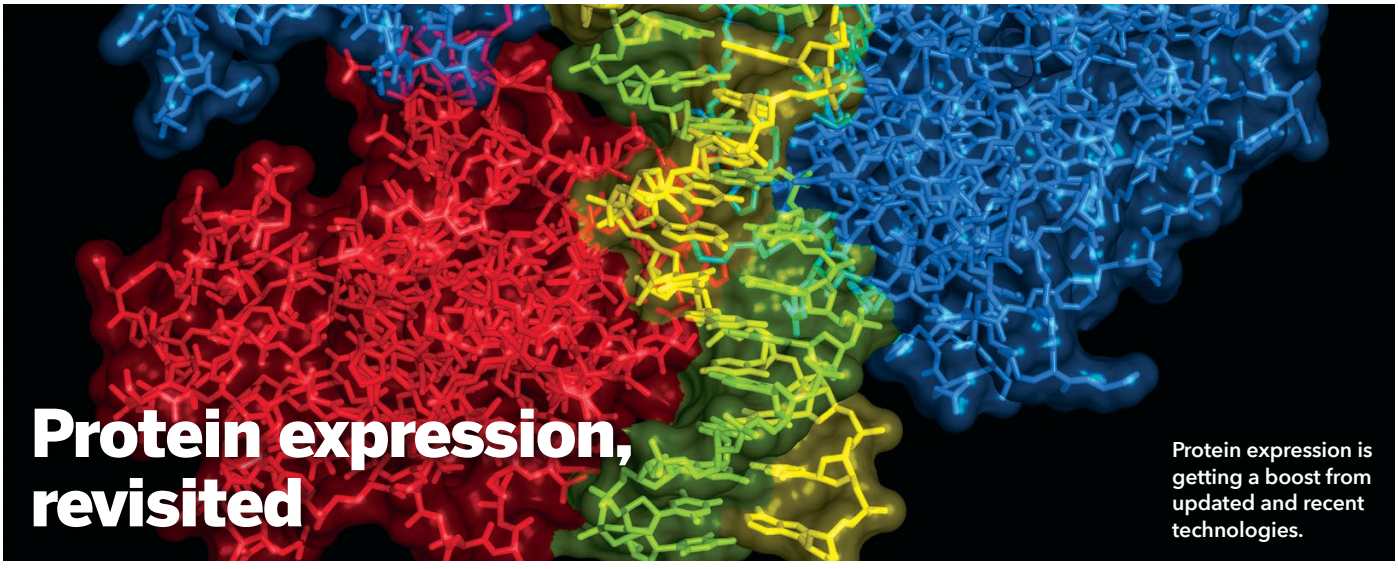


Protein expression, revisited

Protein expression is getting a boost from updated and recent technologies.

An integral tool in biological research, protein expression is increasingly important as a pharmaceutical production mechanism for biological therapeutics. The ability to direct the ultimate nature of proteins depends largely on the degree of control researchers have over the final product. The combination of old technologies—such as artificial enzymes—and new next-generation expression systems is revitalizing protein expression protocols. **By Caitlin Smith**

The advent of genome editing tools such as the CRISPR/Cas9 system has given scientists an affordable, accessible, and relatively simple method to alter genes. Yet precisely controlling protein expression through structural and posttranslational modifications takes on immense importance in the context of expressed therapeutic proteins, such as prescription medications. Controlling protein expression in other species, such as mosquitoes, may even aid researchers in eradicating malaria, the Zika and West Nile viruses, dengue fever, and other mosquito-borne illnesses.

Our quickly evolving molecular tools have far-reaching effects, as evidenced by the wide range of CRISPR applications already being used. "CRISPR systems are amazing resources for finding new tools to manipulate RNA and DNA," says Jamie Cate, professor in the departments of molecular and cell biology and biochemistry at the **University of California, Berkeley.** "I see the present as the biological equivalent of the transistor revolution for computers."

Other molecular tools are evolving alongside—and sometimes interacting with—CRISPR systems to broaden the scope and increase the potency and versatility of controls on protein expression. Artificial enzymes, for example, are shedding their limitations and gaining programmability. Molecular switches are becoming multifaceted, enabling greater fine-tuning. Next-generation expression systems can churn out greater yields of complex biological therapeutics than ever before. One thing is certain—as molecular controls become increasingly sophisticated, their myriad uses continue to expand.

Artificial enzymes for targeting DNA and RNA

Although it's a powerful tool, even the CRISPR/Cas9 system has occasional drawbacks. One is the requirement for a protospacer adjacent motif (PAM) sequence just upstream of the intended editing site. The PAM sequence is a binding signal for the Cas9 enzyme, so its presence is necessary, but also limiting (if a PAM sequence doesn't exist upstream of a desired editing site, one needs to be created). Huimin Zhao, theme leader for biosystems design at the **Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign,** attempted to develop a CRISPR-like genome editing tool that would remove this constraint, and offer more versatility by permitting editing without a PAM sequence.

Zhao's lab developed an artificial restriction enzyme (ARE) system based on PfAgo, an Argonaute protein from the archaeon *Pyrococcus furiosus*, a prokaryote (single-celled) microbe. This platform surpasses traditional type II restriction enzymes—molecular scissors that cleave DNA at predefined sites (*1*).

"We can readily create an unlimited number of artificial restriction enzymes, with desired sequence specificity and defined sticky ends of varying lengths, using a simple system consisting of a single protein, PfAgo, and two DNA guides for targeting a specific dsDNA [double-stranded DNA] sequence," says Zhao. "This system can be multiplexed, in that the same protein can be loaded with multiple DNA guides for targeting several sites simultaneously."

Because the DNA guides show PfAgo where to cut, the system can be programmed to cut virtually anywhere. Another advantage is that PfAgo has longer recognition sequences (typically 16 base pairs) than do traditional restriction enzymes (which typically recognize 4 to 8 base pairs); a longer recognition sequence makes it more likely to find a unique cleavage site. Unlike previous artificial restriction enzymes, PfAgo can generate defined and longer sticky ends when cutting DNA, which aid in subsequently attaching DNA fragments together.

Zhao's lab is eager to apply its new technology. "We also developed a PfAgo/ARE-based direct cloning method to clone large, natural-product biosynthetic gene clusters for discovery of novel natural products that can potentially be used cont.>

as antibiotics and anticancer drugs," says Zhao. His newly formed company, **Modular Bioscience,** will explore PfAgo's use in medical diagnostics, such as liquid biopsies, and single-nucleotide polymorphism and pathogen detection.

Another CRISPR-associated Argonaute protein from the bacterium *Marinitoga piezophila* (MpAgo) is in the crosshairs of Cate's lab, which focuses on protein-translation mechanisms and regulation (*2*). "Our goal was to make an RNA-targeting technology that we could use to explore RNA biology in human cells," says Cate. They wanted to target RNA using an RNA-guided protein, as an alternative to labeling RNAs with covalent tags.

Postdoctoral researcher Audrey Lapinaite, a member of Cate's team, found the main challenge was "how to load the guide RNA into MpAgo inside cells," says Cate. She solved this problem by assembling guide RNA–protein complexes (RNPs) in vitro, then using the RNPs in cell-based experiments. Lapinaite found that a modification of the 5'-nucleotide of the guide RNA resulted in easily programmable RNPs with high affinity to their fully complementary RNA targets. Furthermore, the modified RNPs show high specificity, and can discriminate between RNA substrates differing by only one nucleotide.

Cate is excited to explore the prospects for manipulating RNAs using MpAgo RNPs. "We are most eager to use MpAgo for RNA targeting in human cells to explore human RNA biology, such as in imaging and proteomics experiments," he says. Because MpAgo is derived from bacteria, its use for therapeutic purposes is unlikely—but Cate hopes to apply this system to gain insight into how endogenous human Argonautes work.

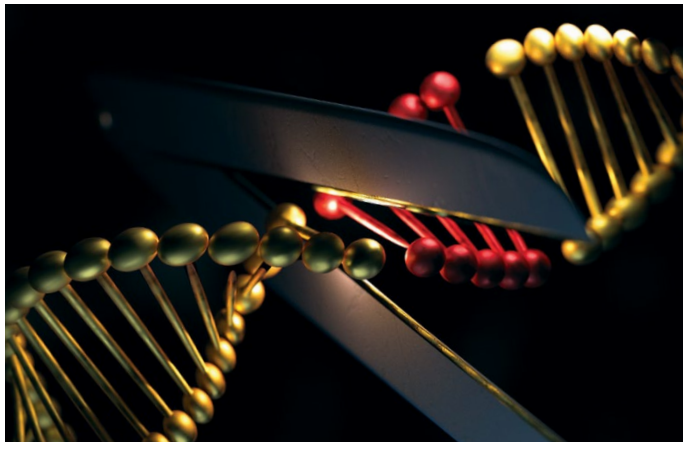
Molecular switches

Given the enormous capabilities of CRISPR/Cas9 genome editing, the search is on for ways to make the system inducible, so that editing can be turned on or off at specific times. A United Kingdom–based collaboration between Yu-Hsuan Tsai from **Cardiff University** and Anthony Perry from the **University of Bath** reveals a new type of inducible CRISPR switch (*3*). Previous switches had drawbacks, such as leakiness of editing activity in the absence of signal, or a reliance on antibiotic use, which can increase the risk of antibiotic resistance. "We envisaged the use of an artificial, nonphysiological amino acid [that] would address these problems," says Tsai.

Tsai and Perry's groups used genetic code expansion to make a genome sensitive to an artificial signal. In cell lines and mouse embryos, researchers "expanded" the genetic code by inserting a toolkit that makes the expression of Cas9 (the enzyme required for genome editing) dependent on the presence of lysine derivative Lys (Boc). This non-natural amino acid is ideal for this purpose because it is cheap, safe, readily obtainable, and easy to administer to cells or whole animals.

The researchers showed that the presence of Lys (Boc) resulted in genomic editing, while in its absence no genomic editing occurred. The success of Tsai and Perry's switch may be due in part to the difference in strategies: "Our approach controls the translation of the functional Cas9 protein, whereas previous methods [use] posttranslational control, such as modulating the activity of translated protein by different stimuli," says Tsai.

Future applications of their work include gene drives, which virtually ensure that all progeny will inherit genetic changes that rapidly spread throughout a population of animals (in herds of livestock, for instance). Because the effect is inducible, its power can be more safely controlled. This new, inducible CRISPR switch may be more suitable "in clinical therapeutic genome editing in situ, or gene drives in which environmentally compatible control is paramount," says Tsai.



Next-gen expression systems may see further record expression levels with the incorporation of genome editing tools like CRISPR/Cas9.

Another type of molecular switch features an artificial enzyme, and was developed by a Swiss collaboration headed by Tom Ward, director of **NCCR (National Centre of Competence in Research) Molecular Systems Engineering,** a multidisciplinary initiative comprising researchers at the **University of Basel** and **ETH Zurich** (Swiss Federal Institute of Technology). Ward's lab took a modular design approach, combining their knowledge of artificial metalloenzymes with expertise in cell-penetrating modules from Stefan Matile's lab in the Department of Organic Chemistry at the University of Geneva, and a synthetic gene-switch module produced by Martin Fussenegger's lab in the Department of Biosystems Science and Engineering at ETH Zurich.

This new cell-penetrating disulfide module lets an artificial metalloenzyme enter a cell without harming it, "reminiscent of a Trojan horse," says Ward. Upon entry, the enzyme catalyzes a reaction that uncages a hormone. The synthetic gene-switch module detects the newly uncaged hormone and responds by turning on expression of the fluorescent indicator luciferase. "The gene switch is thus turned on by an abiotic reaction operated by an artificial enzyme," says Ward.

Now that the collaboration has established proof-of-concept for their system (*4*), they are actively pursuing biomedical applications that include interacting with a host protein—for example, a protein expressed by cancer cells that enables precise targeting. "Carbonic anhydrase is overexpressed on the surface of many cancer cell lines," says Ward. "We could use this protein to accumulate the artificial metalloenzyme either on the surface or inside of cancer cells; the system is turned on only when the artificial metalloenzyme binds to carbonic anhydrase."

Next-generation expression systems

Advances in molecular control of gene expression are also expanding the opportunities for protein expression systems. As drug manufacturers look for cheaper ways to express new biologic drugs, next-gen expression systems are gaining traction. Currently, the standard expression system is mammalian Chinese

Featured participants

AbSci
www.absci.com

Cardiff University
www.cardiff.ac.uk

Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign
www.igb.illinois.edu

Dyadic
www.dyadic.com

ETH Zurich
www.ethz.ch/en.html

Modular Bioscience
modbiosci.com

NCCR Molecular Systems Engineering
www.nccr-mse.ch/en/home

University of Basel
www.unibas.ch/en.html

University of Bath
www.bath.ac.uk

University of California, Berkeley
www.berkeley.edu

hamster ovary (CHO) cell lines. However, the limitations of CHO cells—such as high cost and comparatively slow doubling times—make the next-gen systems worth considering for expressing new biopharmaceuticals, especially to help reduce the costs of important medicines.

For example, the SoluPro protein expression platform from **AbSci** can generate soluble, properly folded proteins at extremely high titers, currently 4g/L of full-length antibody and >20g/L of other complex products, using *Escherichia coli* expression. Typically, some human proteins can't be expressed in *E. coli*, and require mammalian cell lines for proper folding. AbSci's technology includes two innovations that make it possible to produce such proteins in *E. coli*, while taking advantage of the simplicity and lower costs of this expression system.

One creation is a semioxidized cytoplasm that produces soluble, disulfide-linked proteins. "Cytoplasmic production, which traditionally is limited by inclusion body formation, is desirable because it has significantly higher capacity than the periplasm, places no restrictions on protein size, and achieves dramatically shorter production cycles of one to two days, compared to secretion-based expression systems," says Sean McClain, AbSci's founder and CEO. The second innovation is SoluPro's dual inducible promoters, which can be independently controlled. This enables "tuning" of protein production rates by optimizing for the best protein-folding and titer.

"The SoluPro system's ability to properly fold proteins overcomes a large majority of the limitations found with traditional *E. coli* expression," says McClain. "We have successfully produced novel antibody scaffolds as well as IgG1 and IgG4 molecules where effector function is not desirable." Many of these novel antibody scaffolds, which are gaining increasing traction in development pipelines, are challenging to produce in CHO cells. "AbSci is keenly focused on these hard-to-produce next-generation antibody scaffolds, including bispecifics, Fc [fragment crystallizable]-fusion proteins, and other multispecific products, which SoluPro is ideally suited to manufacture efficiently," he says.

Dyadic offers another next-gen expression platform, the fungal-based C1 gene expression system. Dyadic's scientists took advantage of a serendipitous mutation that resulted in a several 100-fold increase in protein productivity of the filamentous *Myceliophthora thermophila* fungus (which the company nicknamed "C1," then created a molecular toolset to turn the fungus into a recombinant expression host). They are currently focused on biomedical efforts, applying the C1 expression platform to making biologic vaccines and drugs more affordable and accessible. Though filamentous fungus may sound exotic, in nature they are natural secreters (C1 has a 2-hour doubling time, compared to about 20 hours for CHO cells). They also use defined synthetic media, which is cheaper and avoids the need for viral inactivation required for CHO cells.

Today, biosimilars are being produced ineffectively by cell lines such as CHO cells, according to Dyadic's CEO Mark Emalfarb. However, says Emalfarb, "We can produce up to 2 to 5 times more, in a third of the time, and at a fraction of CHO media costs [as compared to] the reported average industry CHO productivity." And because C1 secretes its product into the media, the downstream protein harvesting steps are also simpler than those in other systems—for example, *E. coli* requires lysing of cells and purifying product from cell fractions, adding more complexity and cost.

Dyadic's C1 platform efficiently produces full-length monoclonal antibodies, antibody heavy and light chains, Fc-fusion proteins, Fabs (antigen-binding fragments), bispecific antibodies, and vaccines. "We can also make VLPs [virus-like particles], which are theoretically more potent types of vaccines, and are more difficult to express in general," says Emalfarb. "We even have a secreted VLP now, so less is lost in downstream processing."

Dyadic is also engineering C1 to produce human-like glycosylation in different forms, to allow pharmaceutical companies to evaluate them. "Unlike CHO cells, C1 cells are monoclonal cells," explains Matthew Jones, Dyadic's chief commercial officer. "C1 has the potential to produce more consistent, more homogeneous glycostructures for companies to evaluate, to test which ones may work better." A recently announced collaboration with the biopharmaceutical company Sanofi-Aventis Deutschland GmbH will study the use of the C1 technology to express different types of therapeutic compounds, such as vaccines and protein-based biologics.

Both *E. coli*- and yeast-based expression systems share other advantages, such as not requiring the expensive viral clearance steps needed in CHO cells. In addition, the lower costs of next-gen expression systems can in turn lower drug prices, while allowing drug manufacturers to maintain a profit margin—which incentivizes manufacturers to produce drugs with lower prices and make them available to the public. Both McClain and Emalfarb hope their companies' technologies will encourage manufacturers to market drugs that are good for society but not otherwise profitable, such as new influenza vaccines that cost less but work better.

None of this would be possible without the genetic tools that harness expression systems. As scientists achieve finer controls over genomes, the number of expression products will soar. For example, next-gen expression systems may see further record expression levels with the incorporation of genome editing tools like CRISPR/Cas9. As these tools continue to grow in sophistication, the benefits to patients will continue to increase—and may transform biomedicine in the process.

References

1. B. Enghiad, H. Zhao, *ACS Synth. Biol.* **6**, 752–757 (2017).
2. A. Lapinaite, J. A. Doudna, J. H. D. Cate, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3368–3373 (2018).
3. T. Suzuki *et al.*, *Sci. Rep.* **8**, 10051 (2018).
4. Y. Okamoto *et al.*, *Nature Comm.* **9**, 1943 (2018).

Caitlin Smith is a freelance science writer in Portland, Oregon.