Synthetic Epigenetics To Engineer Regulation

Theresa B. Loveless* and Chang C. Liu*

Departments of Biomedical Engineering, Chemistry, and Molecular Biology & Biochemistry, University of California, 3120 Natural Sciences II, Irvine, California 92697, United States

The often-quoted Richard Feynman dictum "what I cannot create, I do not understand" inspires many fields at the intersection of science and engineering. Synthetic biologists, for example, frequently test models of how biology works by building simple versions of biological phenomena in question from scratch, with the view that if those versions then have the intended behavior, the models and logic guiding their construction were largely correct.[1] In the best cases, this synthetic approach provides both strong confirmatory evidence for specific biological models and an engineerable platform that can support further modification for applications. This is especially true if the components of the synthetic system are modular and orthogonal, allowing them to integrate into natural biological systems without perturbing them in unintended ways.

In a recent issue of *Cell*, Park et al. masterfully take the synthetic approach to mammalian epigenetics.[2] Epigenetics concerns the study of heritable traits that do not occur at the level of DNA sequence, and an important mechanism for mammalian epigenetic regulation is the establishment of chemical modifications on specific regions of DNA.[3] Park et al. have recapitulated this mechanism through a synthetic system capable of site-specific DNA methylation, propagation of those methyl marks to effect memory, and actuation of gene expression changes through those modifications in human cells. This work therefore not only shows that the simple model of chemical modification, propagation, and actuation is enough to capture the key elements defining epigenetics but also provides a new layer of synthetic regulation that promises to be highly useful in mammalian cell engineering.

To achieve their platform for synthetic epigenetics, Park et al. introduced a chemical modification naturally found in bacteria [6-methyl-adenosine (m6A)] into mammalian cells (Figure 1). They first created a "writer" protein that adds m6A at specific DNA sequences. The writer module is based on *dam* methylase, which adds m6A to GATC sequences in *Escherichia coli* and is active when expressed in human cells. By fusing *dam* to a zinc finger (ZF), the authors reasoned that they could establish m6A marks at specific loci. However, in its native context, *dam* methylates the commonly occurring GATC sequence wherever it occurs. To restrict the activity of *dam* to only ZF-targeted sequences, Park et al. made a series of single mutations predicted to reduce the catalytic activity of *dam* and chose one that had >1000-fold increased activity toward target loci compared to six untargeted loci. Armed with a *dam* variant (DAM*) that can establish m6A marks locally, Park et al. then showed that the binding domain of the restriction enzyme DpnI can act as a reader of m6A in mammalian cells. This reader domain (RD) can be fused to DAM* to propagate m6A marks at nearby and complementary GATC sites in DNA, thereby spreading methylation across a locus and cell divisions.

The RD can also be fused to effectors that activate or repress transcription (EDs), thus creating m6A-dependent transcriptional regulators. Taken together, the initial ZF−DAM* fusion establishes an epigenetic mark that spreads through the action of the RD−DAM* fusion to achieve memory and regulates gene expression through RD−ED fusions to drive a persistent transcriptional state.

One particularly exciting feature of this work is the modularity of the parts involved, which can be exploited to control both the duration and strength of regulatory programs. For example, if one transiently expresses the ZF−DAM* fusion and constitutively expresses an RD−ED fusion, the regulatory response is predictably short-lived. This is because while the RD−ED fusion turns on transcription of the associated gene, as new DNA is synthesized and cells divide, the m6A marks established by the ZF−DAM* fusion are diluted out. However, when an RD−DAM* fusion is added to the equation, transcriptional activation can be maintained even 20 cell divisions after the original ZF−DAM* fusion is turned off, because the RD−DAM* fusion propagates m6A. In fact, the RD−DAM* fusion achieves not only the maintenance of the m6A mark but also the epigenetic hallmark of spreading, wherein the m6A mark is propagated as far as 1500 bp from the original targeted site. Spreading depends on the density of target GATC sites in a locus, which can be designed by the user to great effect. For example, with a change in the number of GATC sites, the strength of activation or repression by an RD−effector fusion can be tuned. Likewise, the closer GATC sites are to the optimal distance for RD−effector action at a gene, the greater the effect, as well.

Another exciting feature of this synthetic epigenetic system is its orthogonality to existing host epigenetic regulation. Though there is debate, m6A DNA methylation is thought to be rarely used in mammalian cells,[4] so synthetic m6As should not perturb natural epigenetic mechanisms. Therefore, the epigenetic programs installed through the use of m6A-based writers and readers are encoded in a new regulatory language, which makes their behavior highly predictable. While it remains to be seen how m6A modifications interact with host DNA-binding proteins and possibly host m6A modifications, it is likely that these new programs will not become tangled with the natural epigenetic networks of the host. In short, Park et al. have achieved an effective regulatory engineering platform for mammalian cells in which complex properties of epigenetic circuits like memory can be studied and programmed in a simple and reliable manner.
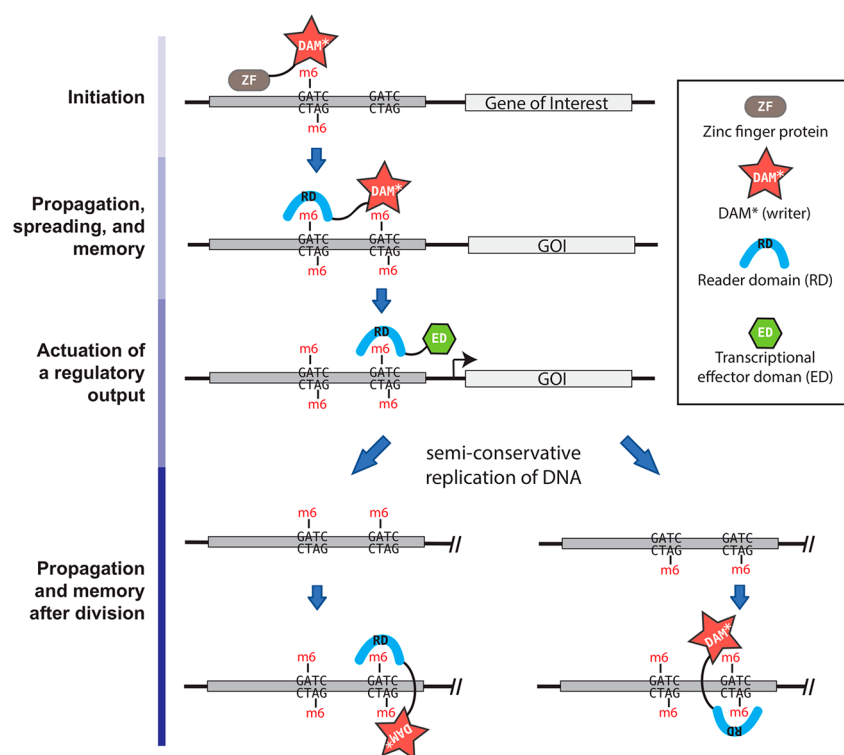
Figure 1. A synthetic epigenetic system allows for the establishment and propagation of m6A as a new layer of regulatory control in mammalian cells. The ZF−DAM* protein initiates methylation at a desired locus; the RD−DAM* protein propagates the m6A mark to nearby GATC sequences, and the RD−ED protein recognizes m6As to recruit transcriptional machinery to regulate a gene of interest. After cell division, m6A marks are lost on the newly synthesized strand but reestablished by the RD−DAM* protein.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: theresa.loveless@uci.edu.

*E-mail: ccl@uci.edu.

ORCID ⓘ

Chang C. Liu: 0000-0002-3290-2880

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Yeh, B. J., and Lim, W. A. (2007) Synthetic biology: lessons from the history of synthetic organic chemistry. *Nat. Chem. Biol. 3*, 521−525.

(2) Park, M., Patel, N., Keung, A. J., and Khalil, A. S. (2019) Engineering epigenetic regulation using synthetic read-write modules. *Cell 176*, 227−238.

(3) Bonasio, R., Tu, S., and Reinberg, D. (2010) Molecular signals of epigenetic states. *Science 330*, 612−616.

(4) Schiffers, S., Ebert, C., Rahimoff, R., Kosmatchev, O., Steinbacher, J., Bohne, A.-V., Spada, F., Michalakis, S., Nickelsen, J., Muller, M., and Carell, T. (2017) Quantitative LC-MS provides no evidence for m⁶dA or m⁴dC in the genome of mouse embryonic stem cells and tissues. *Angew. Chem., Int. Ed. 56*, 11268−11271.