position of the target site in the 5′UTR with respect to the translation start site does not have a strong effect. Could it be that the distance from the poly-A tail, or maybe from the stop codon, to the LIN41 site determines whether a target site causes translational repression or mRNA decay? All of these questions need to be addressed in the future to begin understanding the rules governing each of the two regulatory mechanisms.

In the last decade, the field of post-transcriptional regulation has been continually expanding its knowledge base; day-to-day, novel factors are described, cis-elements are recognized, and new mechanisms are discovered. As Aeschimann et al. (2017) demonstrated, even when the factors and elements are known, alternative pathways can be used by a single factor. New discoveries bring light to the complex nature of mRNA regulation and protein production. This complexity increases the resolution to fine-tune gene expression and makes post-transcriptional gene regulation so fascinating.

REFERENCES

Aeschimann, F., Kumari, P., Bartake, H., Gaidatzis, D., Xu, L., Ciosk, R., and Großhans, H. (2017). Mol Cell *65*, this issue, 476–489.

Ambros, V. (2011). Curr. Opin. Genet. Dev. *21*, 511–517.

Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Science *336*, 233–237.

Bazzini, A.A., Del Viso, F., Moreno-Mateos, M.A., Johnstone, T.G., Vejnar, C.E., Qin, Y., Yao, J., Khokha, M.K., and Giraldez, A.J. (2016). EMBO J. *35*, 2087–2103.

Djuranovic, S., Nahvi, A., and Green, R. (2012). Science *336*, 237–240.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). FEBS Lett. *582*, 1977–1986.

Hüttelmaier, S., Zenklusen, D., Lederer, M., Dictenberg, J., Lorenz, M., Meng, X., Bassell, G.J., Condeelis, J., and Singer, R.H. (2005). Nature *438*, 512–515.

Mishima, Y., and Tomari, Y. (2016). Mol. Cell *61*, 874–885.

Park, E., and Maquat, L.E. (2013). Wiley Interdiscip. Rev. RNA *4*, 423–435.

Building the Class 2 CRISPR-Cas Arsenal

Kevin M. Lewis[1] and Ailong Ke[1,*]
[1]Department of Molecular Biology and Genetics, Cornell University, 251 Biotechnology Building, Ithaca, NY 14853, USA
*Correspondence: ailong.ke@cornell.edu
http://dx.doi.org/10.1016/j.molcel.2017.01.024

Adaptation of CRISPR-Cas9 for genome-editing applications has revolutionized biomedical research. New single-component effector CRISPR systems are emerging from the bioinformatics pipeline. How can we best harness their power? Three new studies will no doubt facilitate this transition by generating the C2c1 and C2c2 structure snapshots in different functional states.

CRISPRs, along with the adjacent *cas* (CRISPR-associated) operon, form an RNA-based adaptive immunity against foreign genetic elements in prokaryotes. CRISPR-Cas has become the center of attention since the invention of CRISPR-Cas9-based genome-editing technology (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013). A dazzling line of CRISPR-Cas9 applications quickly emerged: high-throughput knockout/activation/repression screening, cell lineage tracing, and gene drive applications, just to name a few. The demand for more robust and precise CRISPR-based genome-editing tools is ever growing. While some efforts are focusing on improving the existing CRISPR-Cas9 tools, others are tapping into the power of alternative CRISPR-Cas systems.

CRISPR-Cas systems can be classified as either class I, which uses multi-component effector complexes to achieve RNA-guided nucleic acid targeting and degradation, or class II, which uses a single-component effector (i.e., Cas9). Each class further contains at least three types and multiple subtypes therein (Makarova et al., 2015). In 2015, three new class 2 CRISPR-Cas systems were identified bioinformatically and given the names of C2c1, C2c2, and C2c3 (for class 2, candidate 1, 2, or 3, respectively) (Shmakov et al., 2015). Based on sequence homology, C2c1 and C2c3 were classified together with Cpf1 as type V-B, V-C, and V-A, respectively, whereas C2c2 was designated as type VI. The same study concluded that C2c1 is a dual-RNA-guided DNase and that lysate of human cells expressing C2c1 and crRNA:tracrRNA could cleave its target DNA (Shmakov et al., 2015). These findings suggested that a C2c1-based genome-editing method is conceptually feasible, although it was not directly demonstrated, possibly due to the lack of sufficient mechanistic understanding.

Structural biologists have moved exceptionally fast to provide molecular details for C2c1-mediated DNA targeting process. In a recent issue of *Molecular Cell*, Liu et al. (2017a) reported that the crystal structure of *A. acidoterrestris* C2c1 bound to a 111-nt single guide RNA (sgRNA), fused from crRNA and tracrRNA. The C2c1 structure is indeed architecturally similar to that of Cpf1 (Yamano et al., 2016; Zetsche et al., 2015). Both consist of an α-helical recognition lobe (REC) and a nuclease lobe (NUC).

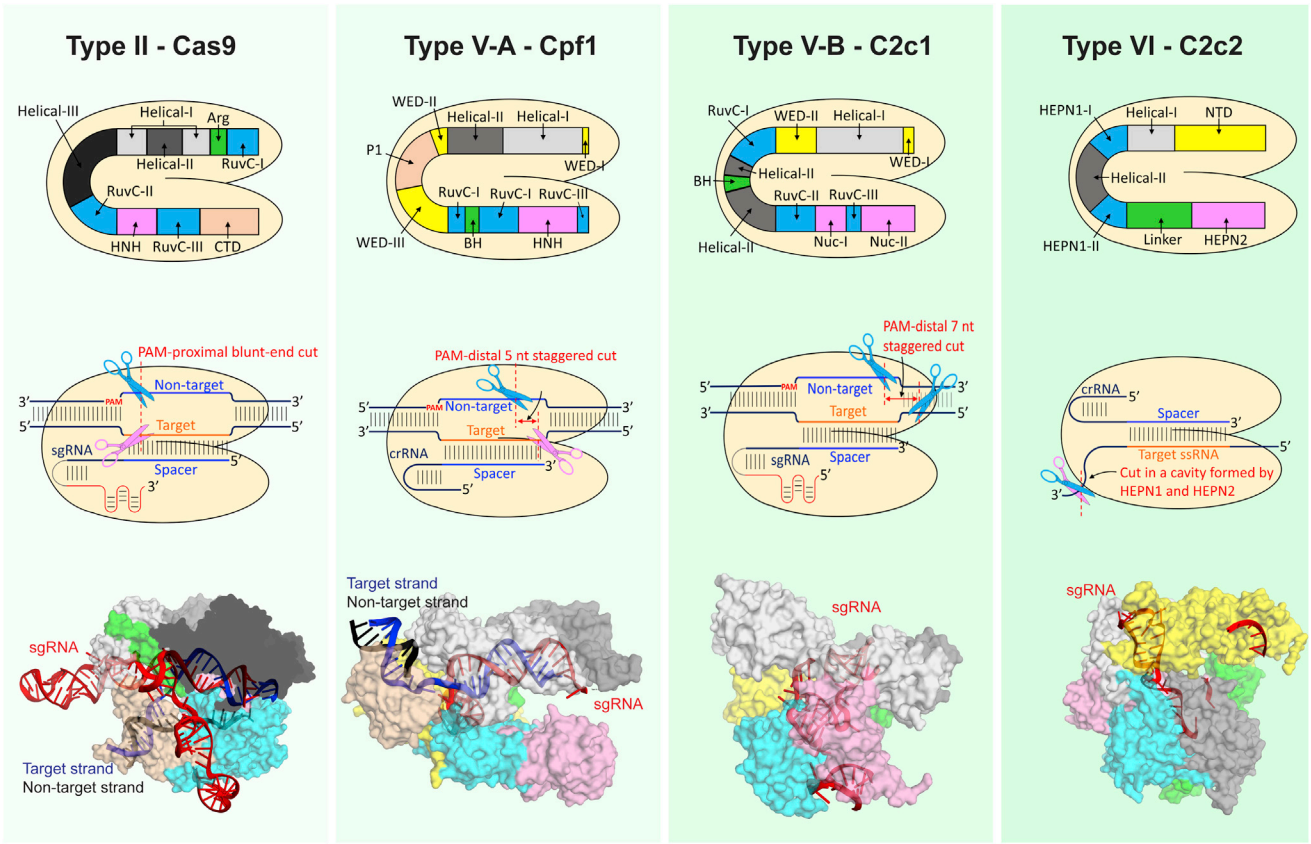


Figure 1. Comparison of Domain Architecture, Biochemical Activity, and 3D Structure among Class II CRISPR-Cas Systems
Top: an approximation of the domain arrangement in class II effector proteins relative to their bi-lobed architecture. Middle: schematics for guide RNA arrangement, PAM recognition, R-loop formation, and DNA/RNA cleavage inside class II effector complexes. Bottom: representative crystal structures of the class II effector complexes, showing the arrangement of protein domains, guide RNAs, and substrates in 3D. The color scheme is according to the top panel.

The NUC lobe further contains a WED/OBD (oligonucleotide-binding) domain, a RuvC domain, a Nuc domain, and a bridge helix (BH). Some of these domains are interrupted and shuffled in the primary sequence, only to fold into intact domains in 3D. Distinct from Cpf1, but similar to Cas9, C2c1 requires tracrRNA for the biogenesis and recruitment of the crRNA guide (Shmakov et al., 2015). Importantly, the C2c1 structure nullified the bioinformatically predicted crRNA:tracrRNA secondary structure (Shmakov et al., 2015) by defining a different crRNA:tracrRNA base-pairing scheme and the presence of a pseudoknot structure. Without question, such information will result in improved sgRNA designs, leading to more robust C2c1-sgRNA assembly and possibly more efficient RNA-guided DNA cleavage. A surprise from the crystal structure was that C2c1 lacks an identifiable PAM-interaction (PI) domain, common to both Cpf1 and Cas9. Liu et al. (2017a) speculated that C2c1 must have evolved surrogate structure features for PAM recognition and that based on comparison with the Cpf1:crRNA:DNA structure (Yamano et al., 2016), these features are likely located in the REC1 and WED domains. The companion biochemical data provided additional mechanistic insights. For example, C2c1 was shown to generate a 7-nt staggered cut at the PAM-distal side. The sticky ends could potentially be exploited in the knockin type of genome-editing applications through the non-homologous end-joining pathway.

Although very informative, the Liu et al. (2017a) study relied on structure comparisons to fill in the mechanistic details in the RNA-guided DNA cleavage process. In a separate study published in a recent issue of *Cell*, Yang et al. (2016) directly addressed these mechanistic questions by determining crystal structures of both the binary *A. acidoterrestris* C2c1:sgRNA complex and the ternary C2c1:sgRNA:DNA complex. Direct structural comparisons suggest that C2c1 undergoes a moderate rigid-body conformational adjustment to accommodate the binding of the PAM-proximal double-stranded DNA (dsDNA). The PAM recognition and R-loop formation mechanisms can be directly interpreted from the ternary complex. Lacking a PI domain, C2c1 engages its WED/OBD and Helix-I domains instead to recognize the PAM duplex from both the major groove and minor groove sides of DNA. Following PAM, a 20-nt region of the target DNA strand is specified by the crRNA guide by forming an A-form DNA/RNA heteroduplex. This approximates the R-loop

formation scenario, although the corresponding non-target strand is not included in structure. How does C2c1 generate the 7-nt staggered cuts? Does the cleavage involve a single nuclease center or two active sites? Similar questions were not fully resolved in the Cpf1 studies (Yamano et al., 2016; Zetsche et al., 2015). The C2c1 ternary complex structure was particularly insightful in this regard. Serendipitously, the excess single-stranded DNA (ssDNA) showed up as "extra densities" in the inactivated catalytic pocket of the RuvC domain, being accommodated in a sequence-nonspecific fashion. Based on the strand directionality, Yang et al. (2016) interpreted this observation as to reflect the binding and cleavage of the target DNA strand following the R-loop region. Impressively, Yang et al. (2016) carried out further detective work to probe for DNA cleavage mechanism by determining two additional C2c1 structures containing either a target or a non-target strand extension. Whereas the target-strand-extension structure verified their initial interpretation, the compensatory structure revealed that when extended by 20-nt at the 3′ end, the non-target DNA strand could also access the RuvC active site, presumably via an alternative (and shorter) route, rendering cleavage to this strand in the middle of the R-loop region. The thorough set of evidences converged in suggesting that the two strands of DNA are cut in a sequential fashion, by a single nuclease center in C2c1, to produce a staggered 5′-protruding double-strand break. This mechanism likely applies to all type V CRISPR-Cas systems.

While most CRISPR systems target dsDNA, C2c2 and a subset of others are specialized in RNA interference (Abudayyeh et al., 2016; Hale et al., 2009). There is considerable interest in utilizing C2c2 for RNA knockdown, editing, or in vivo tagging. C2c2 contains two HEPN RNase domains and causes crRNA-guided RNA cleavage as well as collateral RNA degradation (Abudayyeh et al., 2016). It is equipped with yet another RNase center responsible for its own crRNA (East-Seletsky et al., 2016). In a recent issue of *Cell*, Liu et al. (2017b) also reported the structures of *Leptotrichia shahii* C2c2, with and without the crRNA bound (Figure 1). The structure is particularly insightful in revealing the crRNA maturation mechanism. The crRNA was shown to undergo further processing inside the crystal, and the catalytic center responsible for this activity is located in the Helical-1 domain. A cluster of positively charged residues therein are functionally important for this activity. Without substrate in the structure, the mechanistic explanations for the crRNA-guided target searching and degradation process is less conclusive. The promiscuous cleavage activity in C2c2, which causes collateral RNA damage, could be rationalized to some extent based on the observation that the two HEPN active sites are solvent exposed rather than substrate oriented. Liu et al. (2017b) speculated that C2c2 may simply rely on crRNA to gain proximity to RNA targets, or alternatively, two C2c2s may work in a cooperative fashion to degrade each other's substrate. Substrate-bound C2c2 structures are needed to provide a crystal-clear picture, as none of the hypotheses could fully explain why disrupting one of the two physically separated HEPN active sites led to the complete inactivation of C2c2 (Abudayyeh et al., 2016; East-Seletsky et al., 2016).

In summary, the new wave of structural studies help synthesize a family portrait for class II CRISPR-Cas systems (Figure 1). Whereas all of these single-component CRISPR effector complexes adopt a similar bi-lobed architecture, their domain arrangement differs significantly, leading to idiosyncrasy in guide RNA accommodation and the DNA/RNA cleavage pattern (Figure 1). The high-resolution structural information will certainly inspire future creative efforts to harness their power. The dream to one day pick and choose desirable genome-editing tools by browsing a company catalog is one step closer.

REFERENCES

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). Science *353*, aaf5573.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Science *339*, 819–823.

East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H.D., Tjian, R., and Doudna, J.A. (2016). Nature *538*, 270–273.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). Cell *139*, 945–956.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). Science *337*, 816–821.

Liu, L., Chen, P., Wang, M., Li, X., Wang, J., Yin, M., and Wang, Y. (2017a). Mol. Cell *65*, 310–322.

Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., and Wang, Y. (2017b). Cell *168*, 121–134.e12.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H., et al. (2015). Nat. Rev. Microbiol. *13*, 722–736.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). Science *339*, 823–826.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Mol. Cell *60*, 385–397.

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I.M., Li, Y., Fedorova, I., Nakane, T., Makarova, K.S., Koonin, E.V., et al. (2016). Cell *165*, 949–962.

Yang, H., Gao, P., Rajashankar, K.R., and Patel, D.J. (2016). Cell *167*, 1814–1828.e12.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cell *163*, 759–771.