Integration of spacers into CRISPR loci requires the Cas1/Cas2 integrase complex, frequently in combination with Cas4 exonuclease. However, several CRISPR-Cas systems lack Cas4. Whether Cas4-like activity is dispensable in these systems or provided by an unidentified actor was not known. In this issue of the *Journal of Biological Chemistry*, Ramachandran et al. show that in subtype I-E systems, Cas4-like activity is supplied by DnaQ-superfamily exonucleases, providing a beautiful example of cellular machinery moonlighting in support of CRISPR-Cas adaptive immunity.

Despite extreme interest in using CRISPR-Cas for gene editing, nature did not invent CRISPR-Cas for this purpose. Rather, CRISPR-Cas is an adaptive immune system found in many bacteria and most archaea, where the immune response is manifested in three stages: (i) adaptation, (ii) CRISPR RNA (crRNA) maturation, and (iii) interference. During adaptation, CRISPR-associated (Cas) proteins integrate segments of invading viral DNA (“protospacers”) into CRISPR loci. These “spacers” are clustered within a single locus, where they are separated by DNA repeats, giving rise to the namesake clustered regularly interspaced short palindromic repeats, or CRISPRs (Fig. 1). In stage ii, the CRISPR locus is transcribed, and the pre-CRISPR RNA is cleaved within the repeats by Cas6 or cellular nucleases to produce crRNA. In stage iii, the crRNA is loaded into nucleoprotein complexes that survey the cell for invading nucleic acid complementary to the crRNA. Upon recognition of invading protospacers, Cas nuclease activity destroys the invading genetic material. CRISPR loci thus hold a record of past infections that immunize the cell against recurring infections (1–3). CRISPR-Cas systems are extremely diverse and are grouped into two classes, six types, and 33 subtypes (4). Whereas this largely reflects the composition and mechanisms of the interference complexes, it also indicates variation in other facets of CRISPR-Cas, including spacer integration.

CRISPR loci are preceded by AT-rich leader sequences, and new spacers are inserted at the leader end of the array. This results in spacers archived from newest to oldest along the array (Fig. 1). In addition, protospacers selected for integration generally contain a 2–6-base “protospacer adjacent motif” (PAM), which in type I and type II systems is also needed for interference. Critically, although the PAM is removed from the spacer during integration, spacers must be inserted in the correct orientation relative to the outgoing PAM. In type I systems, PAMs lie upstream of the protospacer. Thus, the PAM end of the spacer should be inserted in a leader-proximal orientation. In contrast, type II systems (Cas9) with downstream PAMs should preferentially insert spacers in a leader-distal orientation. The machinery required for spacer integration must therefore not only insert the DNA in the correct place, but in the correct orientation.

This machinery includes the ubiquitous Cas1 and Cas2 proteins. In *vitro*, Cas1-Cas2 will insert artificially processed protospacers containing 5-base, 3’ overhangs into synthetic CRISPR arrays. The process begins with nucleophilic attack by the 3’ ends of the spacer at alternate ends of the first repeat. This staggered insertion is critical, as it generates the additional repeat needed to accommodate a new spacer. Cas1-Cas2 cannot, however, insert unprocessed sequences, meaning additional factors are required *in vivo*.

These include RecBCD in Gram-negative bacteria and AddAB in Gram-positives that install extended 3’ ends prior to recognition by Cas1-Cas2. And in several type I systems, including subtype I-E, integration host factor (IHF) mediates Cas1-Cas2 recognition of the CRISPR leader. In type II systems, Cas9, Csn2, and Cas4 have been implicated in a variety of roles. Cas4 also functions in type I systems, playing a role in final spacer trimming and PAM cleavage. However, Cas4 is not found in type I-E and I-F systems. How do these types lacking Cas4 complete the insertion?

In this issue of the *Journal of Biological Chemistry*, Ramachandran et al. (5) resolve this conundrum. Working with the type I-E system from *Escherichia coli*, they first confirmed that Cas1-Cas2-IHF will integrate preprocessed spacers with 5-nt 3’ overhangs but is incapable of integrating unprocessed spacers with 15-nt overhangs. Noting that type I-E Cas2 is occasionally fused to DnaQ-like domains, where its 3’-5’ exonuclease activity trims 3’ spacer overhangs (6), they then considered host DnaQ superfamily members for this role. Indeed, inclusion of DnaQ (and ExoT, but not a third family member) supported integration of unprocessed spacers. Further, in contrast to integration of processed spacers in the absence of DnaQ (7), integration of unprocessed spacers showed a strong bias for PAM-compatible orientations (60:1 versus 1:1), and mutation of the PAM diminished this preference.

Further enquiry revealed that DnaQ digestion of the Cas1-Cas2-bound spacer is asymmetric, generating 4–5-nt overhangs on the PAM-distal strand, but 9–10-nt overhangs on the PAM-proximal strand. This prompted them to examine the effects of overhang length on integration while selectively blocking integration of one strand. In these “half-site” integration assays, single-strand integration occurs at either the lead-
er-repeat or repeat-spacer interfaces, and product size reveals the relative frequency of each event. Efficient leader-side integration was found for PAM-distal 5–7-nt overhangs, but only with 5-nt overhangs on the PAM-proximal strand. In contrast, spacer-side integration required exact 5-nt overhangs, independent of strand. DnaQ trimming thus gives an asymmetric spacer primed for integration of the PAM-distal strand, but not the PAM-proximal strand. Further, because leader side integration is faster than spacer-side integration, integration of the PAM-distal strand at the leader-repeat junction is favored, resulting in the “PAM-compatible” orientation. This asymmetry in both trimming and kinetics provides a satisfying explanation for PAM-dependent insertion of the captured sequence.

Whereas Cas4 binds the PAM in most type I systems (see Ref. 5 and references therein), Cas1 serves this role in type I-E and I-F systems lacking Cas4, initially protecting the PAM from cleavage by DnaQ. However, sequencing indicates eventual trimming and removal of the PAM. Whether this is mediated by Cas1 or DnaQ is not resolved, but Ramachandran et al. suggest that this follows formation of the half-site intermediate, which then allows spacer-side integration of the PAM-proximal strand, completing integration.

This work adds to the growing list of proteins with moonlighting roles in CRISPR-Cas. Other examples include the role for IHF in spacer integration discussed above (8, 9) and Rnase III, which assists in crRNA maturation in type II systems (10), thus explaining the absence of Cas6. DnaQ normally functions as the proofreading exonuclease subunit of DNA polymerase III. Whether it can serve its moonlighting role when incorporated in DNA polymerase III is not clear, but if so, this might also serve to recruit DNA polymerase III to fill in the complementary strands of the flanking single-stranded repeats. Regardless, Ramachandran et al. provide a satisfying explanation for the lack of Cas4 in type-1E systems, and their work also suggests that kinetically or conformationally controlled asymmetric states are a hallmark of PAM-compatible spacer acquisition. It will be interesting to see how this is accomplished in type II systems where the PAM is on the opposite end of the protospacer. Finally, it suggests that other non-Cas enzymes, especially those involved in cellular nucleic acid metabolism, may serve moonlighting roles in other CRISPR-Cas systems that lack expected subtype-specific components.

**References**

1. McGinn, J., and Marraffini, L. A. (2019) Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.* **17**, 7–12 CrossRef Medline
2. Hille, F., Richter, H., Wong, S. P., Bratović, M., Ressel, S., and Charpentier, E. (2018) The Biology of CRISPR-Cas: backward and forward. *Cell* **172**, 1239–1259 CrossRef Medline
3. Sorek, R., Lawrence, C. M., and Wiedenheft, B. (2013) CRISPR- mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* **82**, 237–266 CrossRef Medline
4. Koonin, E. V., Makarova, K. S., and Zhang, F. (2017) Diversity classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 CrossRef Medline
5. Ramachandran, A., Summerville, L., Learn, B. A., DeBell, L., and Bailey, S. (2019) Processing and integration of functionally oriented prespacers in the *E. coli* CRISPR system depends on bacterial host exonucleases. *J. Biol. Chem.* **295**, 3403–3414 CrossRef Medline
6. Drabavicius, G., Sinkunas, T., Silanskas, A., Gasiunas, G., Venclovas, Č., and Siksnys, V. (2018) DnaQ exonuclease-like domain of Cas2 promotes spacer integration in a type I-E CRISPR-Cas system. *EMBO Rep.* **19**, e45543 CrossRef Medline
7. Nuñez, J. K., Lee, A. S., Engelman, A., and Doudna, J. A. (2015) Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519**, 193–198 CrossRef Medline
8. Nuñez, J. K., Bai, L., Harrington, L. B., Hinder, T. L., and Doudna, J. A. (2016) CRISPR immunological memory requires a host factor for specificity. *Mol. Cell* **62**, 824–833 CrossRef Medline
9. Wright, A. V., Liu, J. J., Knott, G. J., Doxzen, K. W., Nogales, E., and Doudna, J. A. (2017) Structures of the CRISPR RNA maturation by trans-encoded small RNA and host factor Rnase III. *Nature* **471**, 602–607 CrossRef Medline