

BACTERIAL GENETICS

What CRISPR memories are made of

The CRISPR–Cas system is a prokaryotic defence mechanism against invading mobile elements that depends on the acquisition of foreign DNA to establish immunological memory. CRISPR arrays comprise

repeat sequences that are separated by spacer sequences. Following infection, an invader-derived short DNA sequence is integrated as a new spacer at the leader end of the CRISPR array. Spacers are transcribed, and the pre-CRISPR RNA (pre-crRNA) is processed into small crRNAs that form

a CRISPR ribonucleoprotein (crRNP) complex with Cas proteins. This complex scans foreign DNA, and complementary sequences (known as protospacers) are cleaved by Cas nucleases, dependent on the presence of a protospacer-adjacent motif (PAM). Two studies now report mechanistic details of how bacteria select and integrate viral DNA fragments into their genome.

Heler *et al.* cloned the *Streptococcus pyogenes* type II-A CRISPR–Cas locus into *Staphylococcus aureus*, which lack CRISPR–Cas loci, and infected the bacteria with staphylococcal phages. PCR analysis showed that most bacteria acquired spacers that matched protospacers with a downstream NGG PAM sequence (with N being any nucleotide), which is specifically recognized by *S. pyogenes* Cas9 during target cleavage. Spacer integration was abolished in mutants lacking Cas1, Cas2 or Csn2 (which have all been implicated in spacer acquisition); however, these three

proteins were not sufficient for this process, as expansion of the CRISPR array was undetectable following overexpression of Cas1, Cas2 and Csn2 in mutant strains lacking Cas9 or trans-acting RNA (tracRNA), which is required to target the Cas9 nuclease to specific DNA sites. Moreover, biochemical experiments revealed that Cas1, Cas2, Csn2 and Cas9–tracRNA interact and form a complex. These findings suggested that Cas9 has a role in PAM-dependent spacer acquisition.

To investigate this hypothesis, the authors generated *S. aureus* mutants carrying the CRISPR systems of *S. pyogenes* or *Streptococcus thermophilus* with each other's *cas9* genes. Mutants carrying the *S. pyogenes* CRISPR locus with *S. thermophilus cas9* acquired spacers that matched protospacers flanked by NGGNG PAM sequences (which are specific for targets of *S. thermophilus* CRISPR), whereas expression of *S. pyogenes* Cas9 shifted the PAM specificity to NGG sequences. Finally, the authors demonstrated that the nuclease activity of Cas9, but not of Cas1, was dispensable for spacer acquisition, whereas the Cas9 PAM-binding domain was required for the selection of protospacers with the NGG PAM sequences. In summary, these results suggest that formation of the Cas1–Cas2–Csn2–Cas9 complex enables Cas9 to guide Cas1 to distinct PAM sequences during CRISPR adaptation to ensure the selection of functional protospacers with the correct PAM.

Núñez *et al.* designed an *in vitro* system using a purified Cas1–Cas2 complex from *Escherichia coli* (which uses AAG PAM sequences), double-stranded protospacer DNA and acceptor plasmid DNA with an

inserted CRISPR locus, and they showed that the Cas1–Cas2 complex was sufficient to catalyse spacer acquisition. Furthermore, the authors reported that site-specific integration at the CRISPR locus required 3'-OH ends in the protospacer DNA and target DNA supercoiling.

In *E. coli*, newly acquired spacers start with a 5' G nucleotide that originates from the last nucleotide of the AAG PAM. Notably, using high-throughput sequencing the authors showed that the 3'-OH end of the C nucleotide that pairs with the 5' G nucleotide in the protospacer attacks the target DNA minus strand at the repeat border distal to the leader sequence. Following half-site integration, the 3'-OH end of the opposite protospacer strand targets the leader–repeat border on the opposite side, resulting in full integration of the spacer. This mechanism resembles retroviral integration and DNA transposition, although the involvement and identity of a putative DNA polymerase and DNA ligase to complete integration remain to be determined. Together, these results suggest that recruitment of the Cas1–Cas2 complex specifically to CRISPR repeat sequences enables sequence- and structure-specific spacer integration.

In summary, both studies further our understanding of the mechanisms underlying selection and integration of protospacers into CRISPR loci to create immunological memory of past invaders.

Andrea Du Toit

ORIGINAL RESEARCH PAPERS Heler, R. *et al.* Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature* <http://dx.doi.org/10.1038/nature14245> (2015) | Núñez, J. K. *et al.* Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature* <http://dx.doi.org/10.1038/nature14237> (2015)

Don't forget

mechanistic details of how bacteria select and integrate viral DNA fragments into their genome

