CRISPR control of virulence in *Pseudomonas aeruginosa*

Clustered regularly interspaced short palindromic repeats (CRISPR) and their associated genes (*cas*) are essential components of an adaptive immune system that protects bacteria and archaea from viral infection. Now a recent paper published in *Cell Research* suggests that the Type I-F immune system in *Pseudomonas aeruginosa* may also be involved in post-transcriptional regulation of virulence.

Bacteria and archaea integrate short fragments of foreign nucleic acids into CRISPR loci. CRISPR loci maintain a molecular record of previously encountered infectious agents and short CRISPR-derived RNAs (crRNAs) guide Cas proteins to complementary targets for sequence-specific degradation. Although the phage defense capabilities of CRISPR-Cas systems are well established, these systems are phylogenetically and mechanistically diverse, and a number of alternative roles have been identified [1, 2].

Phylogenetic studies have identified six types (i.e., Types I-VI) and 19 subtypes (e.g., IA-F) of CRISPR-Cas systems. In *Francisella novicida* (Type II-B), the RNA-guided DNA targeting immune system appears to have evolved an ancillary role for regulating endogenous gene expression. In 2013, Sampson *et al*. [3] identified a small CRISPR-associated RNA that guides the Cas9 protein to a complementary mRNA encoding a lipoprotein. RNA-guided delivery of Cas9 to the mRNA decreased the stability of this transcript, which was shown to be important for immunoevasion by the pathogen. These results demonstrated a remarkable and unexpected versatility of these systems

by showing that RNA-guided DNA targeting systems can also participate in post-transcriptional regulation of virulence. Now Li *et al*. [4] present data suggesting that the Type I-F immune system in *Pseudomonas aeruginosa* (*P. aeruginosa*), formerly considered to exclusively target DNA, may also mediate sequence-specific cleavage of RNA.

In *P. aeruginosa*, numerous spacers in the CRISPR loci are complementary to sequences in bacteriophage genomes [5, 6]. While some of the spacers are perfectly complementary, others contain several mismatches. In 2011, Cady *et al*. [7] showed that one of the spacers with five mismatches to an integrated viral genome (i.e., a prophage called DMS3) was involved in prophage-dependent inhibition of biofilm formation. At the time it was thought that phage-dependent inhibition of biofilm formation might be due to crRNA-guided targeting of the RNA, but it was later shown that imperfect base pairing between the crRNA and the integrated prophage DNA caused DNA damage (i.e., low level autoimmunity) and cell death during biofilm formation [8]. Interestingly, this same crRNA can provide partial resistance to phage when the number of mismatches are reduced from 5 to 4 and complete phage resistance when the mismatches are removed entirely [9].

In *P. aeruginosa*, each crRNA assembles with four different Cas proteins (i.e., Csy 1-4) to form a surveillance complex, called the Csy complex or Cascade I-F (Figure 1) [6]. The Csy complex binds to foreign DNA in sequential steps, initiated by detection of two consecutive guanine-cytosine base pairs (G-C/G-C) located adjacent to

the complementary DNA target [10]. The complementary target sequence is referred to as a protospacer (i.e., origin of the spacer) and the consecutive GC base pairs are referred to as the Protospacer Adjacent Motif (PAM). Previous DNA binding studies have shown that the PAM and DNA target base-pairing with the first 8 bases of the crRNA, the so-called "seed" region, are critical for high-affinity binding [6]. Csy complex binding to a target with the correct PAM, base pairing within the seed, and sufficient base pairing throughout the remainder of the protospacer, trigger recruitment of a trans acting nuclease-helicase protein called Cas3 that degrades the target DNA. In the absence of a functional Cas3 protein (i.e., *cas3* knockout or in the presence of a Cas3 inhibitor protein), the Csy complex can act as a programmable transcriptional repressor through stable binding to the bacterial genome [11].

While the general rules for foreign DNA recognition and Cas3 recruitment are established, RNA targeting by this system has not been observed. Now in a recent paper published by *Cell Research*, Li *et al*. [4] present evidence suggesting that this system is also capable of crRNA-guided recognition of mRNA targets. Perhaps motivated by the work in *Francisella novicida*, Li *et al*. initially set out to determine whether crRNAs in *P. aeruginosa* might regulate expression of virulence genes. Indeed, they show that deletions of specific components of the CRISPR-Cas systems correlate with upregulation of certain virulence factors and increased pathogenicity in mice. Specifically, they identify a spacer (i.e., CRISPR 1 spacer 12) with partial complementarity

DNA Targeting

Viral Immunity

RNA Targeting

Endogenous Gene Regulation
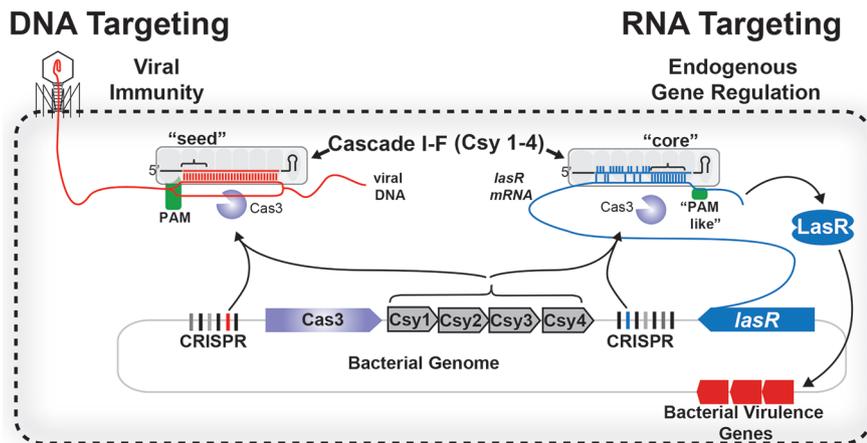
Figure 1 DNA targeting controls viral infection, while RNA targeting controls expression of virulence genes. DNA binding by the CRISPR RNA-guided surveillance complex (Cascade I-F or Csy) relies on recognition of a double-stranded PAM and base pairing at the seed. Binding to the *lasR* mRNA requires base pairing at the opposite end of the crRNA, a region called the "core", and recognition of a single-stranded PAM-like sequence. In both cases, target binding recruits Cas3 for degradation.

to the *lasR* gene, a master regulator that detects quorum-sensing (QS) signals. Deletions of *cas3*, any of the *csy* genes or the CRISPR, result in an upregulation of *lasR* and other downstream virulence factors. To interrogate the rules for crRNA-guided RNA cleavage, the authors developed an *in vitro* assay, and systematically introduced single-base changes at each of the 11 nucleotides with complementarity to the crRNA. By analogy to the "seed" region that is required for DNA binding at the 5′ end of the crRNA, a similar "core" region was identified at the 3′ end of the crRNA (Figure 1). Like the seed, the core also consists of 8 nucleotides that must be complementary to license Cas3 for RNA cleavage. In addition to the core region, the authors also identify a "PAM-like" sequence (i.e., 5′-GGN-protospacer-3′), albeit on the opposite side of the canonical double-stranded PAM (5′-protospacer-GG-3′) [10]. The mechanism of recognizing a single-stranded "PAM-like" sequence will require further investigation.

The work by Li *et al.* generates many questions for future study of Type I CRISPR-Cas function and mechanism. For example, are the rules for RNA binding (i.e., core and single-stranded PAM) broadly applicable to other RNA targets, and if so can these rules be used to predict other RNA targets? How many crRNAs in *P. aeruginosa* and other bacteria with Type I CRISPR-Cas regulate endogenous transcripts? Could pull-down experiments be used to show a direct interaction between the Csy complex and *lasR in vivo*, and does this interaction require any other cellular co-factors? How does CRISPR-mediated repression of *lasR* fit with recent data showing that CRISPRs are upregulated by auto-inducers of quorum sensing [12]? Do auto-inducers activate LasR, leading to CRISPR-Cas expression and subsequent downregulation of the *lasR* transcript in a negative feedback loop?

A final point to consider is the origin of the "*lasR* spacer". We assume that it was not acquired from the *lasR* gene due to the number of mismatches, though it is possible that the mismatches evolved over time. One possible alternative is that the spacer was originally acquired from foreign DNA, and that limited

complementarity to the *lasR* gene is a beneficial coincidence. In rare situations where acquired spacers provide a selective advantage, then we would expect conservation of these spacers over time. One prediction of this hypothesis might be that "older" spacers could serve a regulatory role, while "newer" spacers are more likely to target foreign DNA. While possible, we are currently unaware of data suggesting that older spacers are maintained for alternative functions in endogenous gene regulation, but this could be a well-kept secret hidden among divergent spacer sequences.

Blake Wiedenheft[1],
Joseph Bondy-Denomy[2]

[1]*Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717, USA;* [2]*Department of Microbiology and Immunology, University of California San Francisco, San Francisco, CA, USA*
Correspondence: Blake Wiedenheft[a],
Joseph Bondy-Denomy[b]
[a]E-mail: bwiedeneheft@gmail.com
[b]E-mail: joseph.bondy-denomy@ucsf.edu

References

1   Westra ER, Buckling A, Fineran PC. *Nat Rev Microbiol* 2014; **12**:317-326.
2   Makarova KS, Wolf YI, Alkhnbashi OS, *et al. Nat Rev Microbiol* 2015; **13**:722-736.
3   Sampson TR, Saroj SD, Llewellyn AC, *et al. Nature* 2013; **497**:254-257.
4   Li R, Fang L, Tan S, *et al. Cell Res* 2016; **26**:1273-1287.
5   Zegans ME, Wagner JC, *et al. J Bacteriol* 2009; **191**:210-219.
6   Wiedenheft B, van Duijn E, Bultema JB, *et al. Proc Natl Acad Sci USA* 2011; **108**:10092-10097.
7   Cady KC, O'Toole G. *J Bacteriol* 2011; **193**:3433-3445.
8   Heussler GE, Cady KC, Koeppen K, *et al. MBio* 2015; **6**:e00129-15.
9   Cady KC, Bondy-Denomy J, Heussler GE, *et al. J Bacteriol* 2012; **194**:5728-5738.
10 Rollins MF, Schuman JT, Paulus K, *et al. Nucleic Acids Res* 2015; **43**:2216-2222.
11 Bondy-Denomy J, Garcia B, Strum S, *et al. Nature* 2015; **526**:136-139.
12 Høyland-Kroghsbo NM, Paczkowski J, Mukherjee S, *et al. Proc Natl Acad Sci USA* 2017; **114**:131-135.