CRISPR RNA-guided Surveillance in *Escherichia Coli*

Viruses that infect bacteria and archaea are the most abundant biological entities on the planet[1]. To defend themselves against these pervasive viral predators, bacteria have evolved sophisticated adaptive immune systems that rely on a repetitive chromosomal locus called CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat)[2,3]. Each CRISPR locus consists of short (30-40nt) repeats that are separated by viral or plasmid derived spacer sequences of about the same length. Upon viral challenge, the bacterial CRISPR evolves by integrating new fragments of foreign DNA. In this way, CRISPRs serve as molecular vaccination cards that maintain a genetic record of previously encountered genetic parasites.

CRISPR loci and CRISPR associated proteins (Cas) are extremely diverse, consisting of three Types (I, II and III) and ten subtypes (IA-F, IIA-B, IIIA-B)[4]. However, all CRISPR systems defend against viral attack by generating short CRISPR RNAs (crRNAs) that are incorporated into CRISPR-associated (Cas) complexes. Detection of foreign DNA by CRISPR surveillance complexes requires crRNA-guided hybridization to the DNA target, and recognition of a conserved di- or tri-nucleotide motif called a PAM (protospacer adjacent motif) [2,3]. Once bound by CRISPR RNA-guided machinery, DNA targets are degraded by dedicated nucleases.

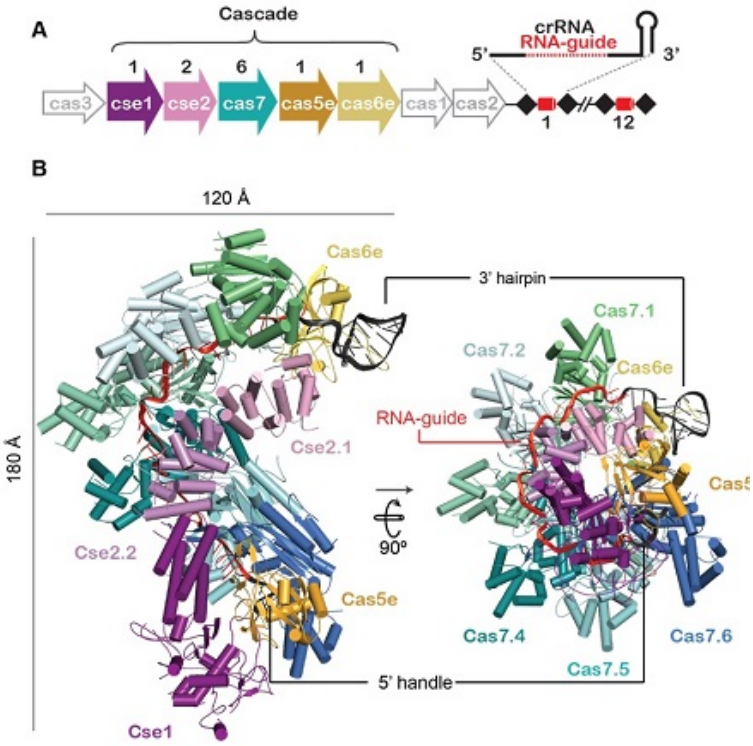

Figure 1. Overview of the Cascade crystal structure (PDB 4TVX). A**)** Shown is a schematic of the Type IE CRISPR operon in *E. coli*. The Cas genes are shown as arrows and the CRISPR repeat and guide sequences are shown as black squares and red barrels, respectively. The Cas genes *cse1*, *cse2*, *cas7*, *cas5e*, and *cas6e* (colored arrows) are expressed in an uneven stoichiometry (numbers above arrows) and combine with processed crRNA to produce Cascade. B) Orthogonal views of Cascade. Subunits are colored as in A.

Cascade (CRISPR-associated complex for antiviral defense) is a 405 kDa CRISPR RNA-guided surveillance complex that defends *Esherichia coli* against viral attack. Structures of Cascade determined using cryo-electron microscopy revealed a sea-horse-shaped architecture that undergoes a conformational rearrangement upon target binding, which may serve as a signal for recruiting the trans-acting nuclease called Cas3[5]. However, the 8-9 Å resolution EM structures left many unanswered questions about Cascade assembly and the mechanics of DNA targeting.

To better understand CRISPR-mediated immunity in *E. coli*, Cascade was crystallized and x-ray diffraction data were measured using Beam Line 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL). The Pilatus detector and Stanford Automated Mounting system were essential for successful large-unit-cell data collection to 3.24 Å. Molecular Replacement with phases from the 8 Å cryo-EM map were used with these data to solve the Cascade crystal structure.

The 3.24 Å Cascade structure explains in atomic resolution how 11 Cas proteins and a 61 nt crRNA assemble into a sea-horse-shaped ribonucleoprotein complex that identifies and binds DNA. The 3' and 5' ends of the crRNA are anchored at opposite ends of the complex, while the guide portion of the crRNA is displayed along the helical protein backbone. The backbone is formed by six interwoven Cas7 subunits that make critical contacts with the crRNA. Each Cas7 protein adopts a fold that resembles a right hand, with fingers, a palm and a thumb. An alpha-helix located in the palm of each Cas7 subunit induces two 90° turns in the crRNA, creating a "kink" at every sixth nucleotide that divides the crRNA-guide into six structurally similar segments. Each segment contains five solvent exposed nucleotides in pseudo-A-form conformation, but the thumb of an adjacent Cas7 subunit pinches the crRNA at the "kink" and buries the sixth nucleotide such that it cannot participate in DNA target hybridization.

The crystal structure of this 405 kDa ribonucleoprotein complex is an incredible technical accomplishment that provides a molecular blueprint for understanding the fundamental mechanisms of CRISPR-RNA guided detection of invading DNA. Single protein CRISPR-RNA guided surveillance complexes, such as the Type II Cas9 system, have been repurposed for programmable genome editing in eukaryotic cells and the Cascade structure paves the way for engineering of multi-subunit CRISPR systems to be used in similar biomedical applications.

Primary Citation

R. N. Jackson, S. M. Golden, P. B. G. van Erp, J. Carter, E. R. Westra, S. J. J. Brouns, J. van der Oost, T. C. Terwilliger, R. J. Read and B. Wiedenheft, "Crystal Structure of the CRISPR RNA-guided Surveillance Complex from *Escherichia coli*", *Science* **345**, 1473 (2014), DOE: 10.1126/science.1256328.

References

1. Z. Carl, "An Infinity of Viruses – Phenomena", *phenomena.nationalgeographic.com* (2013).
2. J. van der Oost, E. R. Westra, R. N. Jackson, and B. Wiedenheft, "Unravelling the Structural and Mechanistic Basis of CRISPR-Cas Systems", *Nat. Rev. Microbiol.* **12,** 479 (2014).
3. R. Sorek, C. M. Lawrence, and B. Wiedenheft, "CRISPR-mediated Adaptive Immune Systems in Bacteria and Archaea", *Annu. Rev. Biochem.* **82,** 237 (2013).
4. K. S. Makarova *et al.,* "Evolution and Classification of the CRISPR-Cas Systems", *Nat. Rev. Microbiol.* **9,** 467 (2011).
5. B. Wiedenheft *et al.,* "Structures of the RNA-guided Surveillance Complex from a Bacterial Immune System", *Nature* **477,** 486 (2011).

Contacts

Blake Wiedenheft, Montana State University
Ryan N. Jackson, Montana State University

SSRL is primarily supported by the DOE Offices of Basic Energy Sciences and Biological and Environmental Research, with additional support from the National Institutes of Health, National Institute of General Medical Sciences.