Conformational regulation of CRISPR-associated nucleases

Article *in* Current Opinion in Microbiology · June 2017

DOI: 10.1016/j.mib.2017.05.010

4 authors, including:

Ryan Neal Jackson
Utah State University
26 PUBLICATIONS **1,540** CITATIONS

SEE PROFILE

Paul B.G. Van Erp
Montana State University
8 PUBLICATIONS **910** CITATIONS

SEE PROFILE

Published in final edited form as:

Curr Opin Microbiol. 2017 June ; 37: 110–119. doi:10.1016/j.mib.2017.05.010.

Conformational regulation of CRISPR-associated nucleases

Ryan N. Jackson[a], **Paul B.G. van Erp**[b], **Samuel H. Sternberg**[c], and **Blake Wiedenheft**[b]

[a]Department of Chemistry and Biochemistry, Utah State University, Logan UT 84322

[b]Department of Microbiology and Immunology, Montana State University, Bozeman MT 59717

[c]Caribou Biosciences, Inc. Berkeley, CA, 94710

Abstract

Adaptive immune systems in bacteria and archaea rely on small CRISPR-derived RNAs (crRNAs) to guide specialized nucleases to foreign nucleic acids. The activation of these nucleases is controlled by a series of molecular checkpoints that ensure precise cleavage of nucleic acid targets, while minimizing toxic off-target cleavage events. In this review, we highlight recent advances in understanding regulatory mechanisms responsible for controlling the activation of these nucleases and identify emerging regulatory themes conserved across diverse CRISPR systems.

Introduction

Nucleases that degrade DNA and RNA are indispensable for diverse biological functions [1]. To avoid toxicity associated with aberrant activity, nucleases are often controlled by substrate binding at (orthosteric), or away from (allosteric) the active site. Recent evidence suggests that CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat)-associated (Cas) nucleases, which are essential components of adaptive immune systems that protect bacteria and archaea from infection by viruses and plasmids, rely on orthosteric and allosteric control [2,3]. Presumably these regulatory mechanisms evolved to efficiently eliminate foreign DNA or RNA, while avoiding autoimmune reactions associated with destruction of the bacterial or archaeal genome. The programmable nature of CRISPR-associated nucleases (e.g. Cas9) has given rise to a powerful new method for genome engineering, and a comprehensive understanding of how these nucleases function is necessary for safe implementation [4–7].

CRISPR RNA-guided adaptive immune systems are structurally and functionally diverse, consisting of two Classes (1 and 2), six Types (I–VI), and more than nineteen subtypes distinguished by CRISPR repeat sequence and *cas* genes [2,8–10] (Figure 1). Despite this diversity, all CRISPR systems rely on specialized nucleases to execute three stages of adaptive immunity; acquisition, CRISPR RNA biogenesis, and interference [2,3]. During acquisition, a nuclease active integrase complex comprised of Cas1 and Cas2 proteins inserts foreign nucleic acid targets (about 20 – 40 nt in length) called 'protospacers' into the spacer-repeat array at the leader end of the CRISPR locus. CRISPR loci are transcribed, and Cas or

---

Correspondence: R.N.J. (ryan.jackson@usu.edu), P.B.G.v. E. (pbgvanerp@gmail.com), or B.W. (bwiedenheft@gmail.com).

RNAse III enzymes process these transcripts into libraries of small CRISPR-derived RNAs (crRNA). Each crRNA assembles with Cas proteins into surveillance complexes that use the crRNA to bind complementary DNA (Type I, II, V) or RNA (Type III, VI) (Figure 1). Target binding induces conformational rearrangements within the surveillance complex that activate either *cis*-acting nuclease domains located within the complex, or *a trans*-acting nuclease that is recruited for destruction of bound targets.

Structures of Cas nucleases involved in spacer acquisition (e.g. Cas1) and crRNA biogenesis (e.g. Cas6), suggest substrate binding at the enzymatic active site (i.e. orthosteric) induces conformational rearrangements that regulate the nuclease [11–13]. In contrast, Cas nucleases involved in interference rely on a combination of orthosteric and allosteric activation mechanisms. Here we review the regulatory mechanisms that control Cas nucleases involved in interference.

Regulation of Class 2 Cas nucleases

Class 2 systems consist of three different Types (i.e. Type II, V, and VI) that encode single-subunit Cas nucleases [2,8,9] (Figure 2). The Type II (Cas9) systems gained considerable notoriety in 2012 and early 2013 when the programmable nature of these RNA-guided nucleases was exploited for precise cleavage of DNA in human cells (for a review of the primary literature, see [4–6]). In the last five years, these nucleases have ushered in a new era of genome editing technologies that have considerable potential for applications in molecular medicine, industrial biotechnology, and agriculture.

Type II systems

The implementation of Type II systems for genome editing applications requires a fundamental understanding of how these RNA-guided nucleases assemble, bind to double-stranded DNA (dsDNA), and then cleave both strands of the dsDNA target. Initial biochemical work on the Cas9 ribonucleoprotein from *Streptococcus pyogenes* showed that the crRNA base-paired to a *trans*-activating tracrRNA, and that these two RNAs could be fused into a single-guide RNA (sgRNA) capable of directing the Cas9 nuclease to a complementary DNA [14,15]. Subsequent structural and biochemical studies demonstrated that Cas9 is a dynamic enzyme, and that target cleavage requires sequential completion of multiple regulatory checkpoints.

Cas9 adopts a bi-lobed architecture in which two adjacent nuclease domains (RuvC and HNH) reside within the nuclease lobe (NUC lobe), and an alpha-helical lobe (also known as the recognition or REC lobe) interacts with the guide RNA–target DNA heteroduplex [16–21]. In the absence of RNA, Cas9 adopts a nuclease-inactive conformation, which undergoes a dramatic rearrangement upon binding the guide RNA [16] (Figure 2A). Cas9 relies on base-pairing between the crRNA guide and DNA targets, but single-molecule experiments performed using Cas9 from *Streptococcus pyogenes* showed that Cas9 must first recognize a tri-nucleotide motif called a PAM (Protospacer Adjacent Motif) located next to the DNA target sequence [22]. In addition to PAM recognition, high-affinity binding requires precise complementarity between the DNA target and a PAM-proximal “seed sequence” followed by directional unwinding of the DNA from the seed to the PAM-distal end of the target [23]

(Figure 2A). Chromatin immunoprecipitation followed by sequencing (ChIP-seq), revealed that as many as 15 mismatches are allowed outside of the seed sequence for stable binding, but incomplete RNA-DNA hybrids are often not cleaved by Cas9 [24–26].

Initial structures of DNA-bound Cas9 could not explain why stable binding and incomplete R-loop formation fail to activate Cas9 cleavage. In these structures, the active site of the HNH nuclease domain is ~30 Å away from the scissile phosphate of the target DNA strand [17,18,27] (Figure 2A), suggesting that a major conformational rearrangement was necessary to activate Cas9-mediated DNA cleavage. To test this hypothesis and clarify the substrate features required for Cas9 activation, Förster Resonance Energy Transfer (FRET)-based assays were used to measure conformational changes in the HNH domain upon binding to perfectly complementary and mismatched DNA targets [28]. This work revealed that the HNH nuclease domain adopts an inactive conformation on mismatched DNA substrates that are stably bound by Cas9, explaining why off-target DNA binding events do not always lead to off-target DNA cleavage. Full complementarity between target DNA and the crRNA guide drives conformational changes in the HNH domain that trigger target-strand cleavage, and allosterically activates cleavage of the non-complementary strand by the RuvC domain, ensuring concerted production of double-strand breaks [28].

A recent structure of Cas9 bound to a long dsDNA substrate provides a structural basis for the concerted activation of the Cas9 nuclease domains [29] (Figure 2). As observed in all other Cas9 structures, the HNH domain is inserted between motifs II and III in the RuvC domain, and is connected to these motifs by two peptide linkers, L1 and L2. Comparison with previous Cas9 structures reveals that L1 and L2 undergo extensive rearrangement after binding a dsDNA-target. L1 interacts with the PAM-distal side of the RNA-DNA hybrid, while a conserved phenylalanine on L2 stacks with the fourth nucleobase of the protospacer on the non-complementary strand. These interactions stabilize a ~180° rotation of the HNH domain and position the active site in proximity to the scissile phosphate of the complementary strand. Additionally, these rearrangements position the scissile phosphate of the non-complementary strand into the active site of the RuvC nuclease. To avoid steric clashes, these large conformational rearrangements can only be achieved if both L1 and L2 move in concert, which explains why HNH and RuvC activation is coupled. Cas9 nucleases from organisms other than *S. pyogenes* share a similar HNH and RuvC domain linkage, suggesting this concerted activation mechanism may be conserved across all Cas9 enzymes [16,19–21].

Type V systems

Like Cas9, the Type V enzymes – Cas12a, Cas12b, and Cas12c; also known as Cpf1, C2c1, and C2c3 – contain a putative nuclease domain (Nuc) inserted between motifs II and III of a RuvC domain [9,30–35]. Point mutations in the Nuc domain of Cas12a (Cpf1) inhibited cleavage of the complementary strand, while mutations in the RuvC active site disrupted cleavage of both strands [31]. These data suggested RuvC-mediated cleavage of the non-complementary strand is a pre-requisite for cleavage of the complementary strand by the Nuc domain. However, a structure of Cas12b (C2c1) bound to a long single-stranded DNA suggests the RuvC nuclease may be responsible for cleaving both strands of the DNA in sequential order [34]. Indeed, a recent study of Cas12a provides biochemical evidence that a single active site within the RuvC domain cleaves both DNA strands using the same catalytic mechanism, and that activation of this nuclease function requires substantial conformational rearrangements to unblock the previously occluded active site [36]. While it appears that Cas12 enzymes recognize DNA targets in a distinct way from Cas9, both enzyme families have evolved mechanisms to sequester the active sites in their nuclease domains until RNA-DNA heteroduplex formation triggers activation.

Type VI systems

Unlike most Class 2 systems, which target dsDNA, the Type VI Cas nucleases (Cas13, also known as C2c2) bind and cleave ssRNA using two HEPN (Higher Eukaryote and Prokaryote Nucleotide-binding) domains [9,37–39] (Figure 2B). The HEPN RNase domains are activated by ssRNA base-pairing with the crRNA guide, but nuclease activity is suppressed if complementarity extends beyond the guide and into the repeat-derived portion of the crRNA [37]. Complementarity to the repeat-derived sequence of the crRNA may explain the inhibition of Cas13 (C2c2) nuclease activity, similar to regulation of DNase activity in Type III systems (see below). Notably, unlike other Class 2 systems, RNA binding by Cas13 activates a non-sequence specific and trans-acting RNase activity [37,38]; a property which has recently been harnessed for sensitive RNA and DNA diagnostics [40]. Thus, target-induced conformational rearrangements are required to activate the promiscuous HEPN nuclease domains. While recent structures have shed light on Cas13 conformational changes that occur upon crRNA binding [39], further structural and biophysical studies are necessary to fully understand the conformational rearrangements that occur upon target RNA binding and the mechanism of RNA cleavage *in trans*.

Regulation of Class 1 nucleases that are essential for interference

Class 1 CRISPR systems are phylogenetically and functionally diverse, but a unifying feature of these systems is that they all rely on multi-subunit crRNA-guided surveillance complexes for detection of invading nucleic acids [8,41]. Class 1 systems consist of two well-studied Types (Type I and III), and a recently identified Type IV system that has not been experimentally tested and is beyond the scope of this review.

Type I systems

Phylogenetic studies have divided Type I systems into seven sub-types, I-A to I-F and I-U [8]. All Type I CRISPR systems rely on multi-subunit surveillance complexes for crRNA-guided detection of dsDNA targets, and a Cas3 nuclease/helicase that is responsible for target degradation.

Foreign DNA detection by the Type I-E system in *Escherichia coli* relies on a large crRNA-guided complex called Cascade (CRISPR-associated complex for antiviral defense) [42,43]. Cascade searches for targets by first localizing to PAM sequences, which are recognized by the Cas8e subunit (also called Cse1) [23,44–46]. PAM detection destabilizes the target duplex and facilitates crRNA-guided sampling of the PAM-proximal DNA seed sequence [47]. Complementarity between the crRNA-guide and the seed leads to directional unwinding of the DNA duplex that proceeds away from the PAM and triggers a conformational change in Cascade that is necessary for recruitment of the *trans*-acting Cas3 nuclease/helicase, which degrades the DNA target [23,48–51]. However, there is an accumulating body of work showing that Cascade may adopt distinct conformational states based on interactions with specific PAM and protospacer combinations, and that these distinct conformations may regulate the activity of Cas3 [44,49,51–56].

Cascade elicits two distinct immune responses upon target binding: interference or priming (Figure 3). During interference, the Cas8e subunit adopts a “closed” conformation capable of recruiting a nuclease active Cas3, which processively degrades the target DNA [44,50,51,57–59]. However, strict sequence requirements present a potential weakness in the immune system, because mutations in the PAM, seed, or specific positions of the protospacer, allow viruses to escape CRISPR-Cas immunity [47,60,61]. Bacteria with Type I immune systems can restore immunity against PAM or protospacer “escape” mutants using a positive feedback loop that rapidly updates the CRISPR locus with new spacers derived from the phage or plasmid genome [61–64]. This process of rapid acquisition, called “priming”, requires a target that is partially complementary to the crRNA-guide, Cas3, and two proteins that are essential for adaptation (i.e. Cas1 and Cas2) [62,65].

Mutated targets that elicit a priming response correspond to an “open” conformation of the Cas8e subunit, which does not effectively recruit nuclease active Cas3 [44,51,52] (Figure 3). However, in the presence of Cas1 and Cas2, Cascade bound to a priming target efficiently recruits nuclease-repressed Cas3 [44]. Recent structural and biochemical studies performed using the Type I-F immune system from *Pseudomonas aeruginosa*, in which Cas2 and Cas3 are fused into a single polypeptide (Cas2/3), indicated that Cas1 and Cas2/3 form a complex, and that Cas1 represses Cas2/3 nuclease activity until activated by the target bound surveillance complex [66].

Collectively, these results suggest that the conformational state of Cascade recruits a nuclease active Cas3 for direct degradation (interference) or a nuclease repressed Cas1-2/3 complex for rapid adaptation (priming) [44]. However, in both instances Cas3 cleavage products are precursors for CRISPR adaptation [67], suggesting that priming targets do elicit some level of Cas3 nuclease activity and that primed adaptation also takes place during the interference response. It is clear Cas1, 2, and 3 work together to generate and integrate new sequences into CRISPR loci, but the mechanism of allosteric regulation of Cas3 activity by the conformational state of Cascade remains poorly understood.

Type III systems

Like Type I systems, the Type III systems also rely on large multi-subunit crRNA-guided surveillance complexes that are divided into sub-types (i.e., Type III-A to III-D) [8]. Initially these subtypes were functionally divided into either DNA or RNA targeting immune systems. However, a series of recent experiments revealed a sophisticated mechanism that explains how crRNA-guided binding to complementary RNA allosterically activates a non-sequence specific DNase activity (Figure 3) [68]. This model reconciles earlier distinctions between RNA or DNA targeting and reveals that Type III systems cleave both RNA and DNA targets.

The RNase and DNase active sites in Type III complexes are spatially separated, functionally integrated, and temporally controlled. RNA binding drives a conformational rearrangement that allosterically activates non-sequence-specific DNase activity in the Cas10 subunit [69–75]. The RNA target is cleaved in regular 6-nt intervals and structures of Type III complexes have identified residues in the backbone (Cas7-family) and belly subunits that are necessary for RNA cleavage [41,68,76]. According to the current model, periodic cleavage of the RNA target, leads to dissociation of the RNA fragments, which restores the complex to a DNase inactive state [68,70–72].

In addition to complementarity between the crRNA-guide and the RNA protospacer, sequence 3′ of the protospacer regulates DNase activity. Complementarity between the crRNA and the RNA target that extends beyond the guide and into the 5′-repeat of the crRNA inhibits DNase activation in all Type III systems [70–74,75,77], but requirements for DNase activation seem to vary between systems. In *Thermotoga maritima*, the DNase function is activated by protospacers lacking flanking sequence, suggesting complementation between the crRNA-guide and RNA target is the only signal necessary for DNase activation and base pairing with the 5′-repeat is inhibitory [71]. However, in several other systems, hybridization with the protospacer sequence alone is not sufficient for DNase activation. In these systems, the DNase remains inactive unless a non-complementary sequence extends from the protospacer into the 5′-repeat [72,73,75], and in some systems (e.g. Type III-B Cmr system in *P. furiosus*) this 3′ non-complementary RNA sequence must meet specific sequence requirements [70]. The mechanistic distinctions between these systems and the precise role of the 3′ RNA in controlling DNase activation awaits structural studies of Type III complexes bound to RNAs with 3′-flanking sequence.

Perspective

CRISPR-associated nucleases are phylogenetically and functionally diverse, yet common regulatory themes are beginning to emerge. All CRISPR systems that target DNA rely on protein mediated PAM recognition and crRNA-guided recognition of the protospacer. However, binding does not necessarily result in target cleavage. Recent data indicates that different PAM and protospacer combinations result in distinct conformational states that regulate nuclease activity. These new insights are revealing additional layers of sophistication in prokaryotic adaptive immune responses, and understanding how target selection impacts the activity of these nucleases will have important implications for applications in genome editing.

Acknowledgments

Research in the Wiedenheft lab is supported by the National Institutes of Health (P20GM103500, P30GM110732-03, R01GM110270, and R01GM108888), the National Science Foundation EPSCoR (EPS-110134), the M. J. Murdock Charitable Trust, a young investigator award from Amgen, Gordon and Betty Moore Foundation, and the Montana State University Agricultural Experimental Station. Research in the Jackson Lab is supported by Utah State University New Faculty Start-up funding from the Department of Chemistry and Biochemistry, the Research and Graduate Studies Office, and the College of Science.

References and recommended reading

Paper of particular interest, published within the period of review, have been highlighted as:

• of special interest

•• of outstanding interest

1. Yang W. Nucleases: diversity of structure, function and mechanism. Q Rev Biophys. 2010; 44:1–93. [PubMed: 20854710]
2. Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, van der Oost J. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. Science. 2016; 353:aad5147. [PubMed: 27493190]
3. Wright AV, Nuñez JK, Doudna JA. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell. 2016; 164:29–44. [PubMed: 26771484]
4. Sternberg SH, Doudna JA. Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol Cell. 2015; 58:568–574. [PubMed: 26000842]
5. Jiang W, Marraffini LA. CRISPR-Cas: New Tools for Genetic Manipulations from Bacterial Immunity Systems. Annu Rev Microbiol. 2015; 69:209–228. [PubMed: 26209264]
6. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. Cell. 2017; 168:20–36. [PubMed: 27866654]
7. van Erp PB, Bloomer G, Wilkinson R, Wiedenheft B. The history and market impact of CRISPR RNA-guided nucleases. Curr Opin Virol. 2015; 12:85–90. [PubMed: 25914022]
8. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJJ, Charpentier E, Haft DH, et al. An updated evolutionary classification of CRISPR–Cas systems. Nat Rev Microbiol. 2015; 13:722–736. [PubMed: 26411297]
9. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Minakhin L, Joung J, Konermann S, Severinov K, et al. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol Cell. 2015; 60:385–397. [PubMed: 26593719]
10. Burstein D, Harrington LB, Strutt SC, Probst AJ, Anantharaman K, Thomas BC, Doudna JA, Banfield JF. New CRISPR–Cas systems from uncultivated microbes. Nature. 2017; 542:237–241. [PubMed: 28005056]
11. Nuñez JK, Harrington LB, Kranzusch PJ, Engelman AN, Doudna JA. Foreign DNA capture during CRISPR–Cas adaptive immunity. Nature. 2015; 527:535–538. [PubMed: 26503043]
12. Wang J, Li J, Zhao H, Sheng G, Wang M, Yin M, Wang Y. Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. Cell. 2015; 163:840–53. [PubMed: 26478180]
13. Sashital DG, Jinek M, Doudna JA. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. Nat Struct Mol Biol. 2011; 18:680–687. [PubMed: 21572442]
14. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011; 471:602–607. [PubMed: 21455174]
15. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012; 337:816–821. [PubMed: 22745249]

16•. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, et al. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. Science. 2014; 343:1247997. This manuscript describes the apo Cas9 structures of *Streptococcus pyogenes* and *Actinomyces naeslundii*, which revealed apo-Cas9 adopts a catalytically inactive conformation. Electron Microscopic reconstructions of guide-RNA bound Cas9 revealed conformational rearrangements enable DNA binding. [PubMed: 24505130]

17•. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. Cell. 2014;

156:935–949. This paper presents the crystal structure of *Streptococcus pyogenes* Cas9 loaded with sgRNA and bound to target DNA. [PubMed: 24529477]

18•. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 2014; 513:569–573. This manuscript describes the crystal structure of *Streptococcus pyogenes* Cas9- sgRNA complex bound to a DNA target with 5′-NGG-3′ PAM. This structure explains how the PAM sequence is recognized and how it contributes to duplex strand separation. [PubMed: 25079318]

19•. Nishimasu H, Cong L, Yan WX, Ran FA, Zetsche B, Li Y, Kurabayashi A, Ishitani R, Zhang F, Nureki O. Crystal Structure of Staphylococcus aureus Cas9. Cell. 2015; 162:1113–1126. This paper reports the crystal structure of *Staphylococcus aureus* Cas9 in complex with sgRNA and target DNA. It reveals how the canonical 5′-NNGRRT-3′ PAM is recognized and highlights structural conservation and divergence compared to spyCas9. [PubMed: 26317473]

20•. Yamada M, Watanabe Y, Gootenberg JS, Hirano H, Ran FA, Nakane T, Ishitani R, Zhang F, Nishimasu H, Nureki O. Crystal Structure of the Minimal Cas9 from Campylobacter jejuni Reveals the Molecular Diversity in the CRISPR-Cas9 Systems. Mol Cell. 2017; 65:1109–1121. This paper reports the crystal structure of *Campylobacter jejuni* Cas9 bound to sgRNA and target DNA. The structure reveals that the guide RNA contains a triple helix and explains how the canonical 5′-NNNVRYM-3′ PAM is recognized on both DNA strands. [PubMed: 28306506]

21•. Hirano H, Gootenberg JS, Horii T, Abudayyeh OO, Kimura M, Hsu PD, Nakane T, Ishitani R, Hatada I, Zhang F, et al. Structure and Engineering of Francisella novicida Cas9. Cell. 2016; 164:950–961. This paper reports the crystal structure of *Francisella novicida* Cas9 bound to sgRNA and target DNA. The structure shows how the canonical 5′-NGG-3′ PAM is recognized. Structure guided mutations are used to create a variant that recognizes a relaxed 5′-YG-3′ PAM. [PubMed: 26875867]

22•. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014; 507:62–67. The authors use single molecule techniques to show that Cas9 first interrogates the PAM before hybridization with the protospacer. [PubMed: 24476820]

23••. Szczelkun MD, Tikhomirova MS, Sinkunas T, Gasiunas G, Karvelis T, Pschera P, Siksnys V, Seidel R. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc Natl Acad Sci U S A. 2014; 111:9798–9803. This paper reports single-molecule DNA supercoiling experiments that show in Cascade and Cas9 R-loop formation proceeds directionally starting at the PAM. PAM mutations affect R-loop formation rates while protospacer distal mutations affect R-loop stability. [PubMed: 24912165]

24. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol. 2014; 32:670–676. [PubMed: 24752079]

25. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, Iafrate AJ, Le LP, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol. 2014; 33:187–197. [PubMed: 25513782]

26. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 2015; 520:186–191. [PubMed: 25830891]

27•. Jiang F, Zhou K, Ma L, Gressel S, Doudna JA. A Cas9-guide RNA complex preorganized for target DNA recognition. Science. 2015; 348:1477–1481. This paper describes the crystal structure of *Streptococcus pyogenes* Cas9 bound to sgRNA and reveals the seed sequence of the single-guide-RNA is pre-ordered in a pseudo A-form configuration. [PubMed: 26113724]

28••. Sternberg SH, LaFrance B, Kaplan M, Doudna JA. Conformational control of DNA target cleavage by CRISPR–Cas9. Nature. 2015; 527:110–113. This paper reports single molecule work that shows Cas9 binding and cleavage of target DNA first requires recognition of a PAM. [PubMed: 26524520]

29••. Jiang F, Taylor DW, Chen JS, Kornfeld JE, Zhou K, Thompson AJ, Nogales E, Doudna JA. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. Science. 2016; 351:867–871. This paper reported the structure of Cas9 bound to a duplex DNA target with a non-complementary R-loop. The structure captures the HNH nuclease domain in an active conformation and explains the concerted cleavage mechanism of duplex DNA. [PubMed: 26841432]

30. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell. 2015; 163:759–771. [PubMed: 26422227]
31. Yamano T, Nishimasu H, Zetsche B, Hirano H, Slaymaker IM, Li Y, Fedorova I, Nakane T, Makarova KS, Koonin EV, et al. Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. Cell. 2016; 165:949–962. [PubMed: 27114038]
32. Dong D, Ren K, Qiu X, Zheng J, Guo M, Guan X, Liu H, Li N, Zhang B, Yang D, et al. The crystal structure of Cpf1 in complex with CRISPR RNA. Nature. 2016; 532:522–526. [PubMed: 27096363]
33. Gao P, Yang H, Rajashankar KR, Huang Z, Patel DJ. Type V CRISPR-Cas Cpf1 endonuclease employs a unique mechanism for crRNA-mediated target DNA recognition. Cell Res. 2016; 26:901–913. [PubMed: 27444870]
34. Liu L, Chen P, Wang M, Li X, Wang J, Yin M, Wang Y. C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism. Mol Cell. 2017; 65:310–322. [PubMed: 27989439]
35. Yang H, Gao P, Rajashankar KR, Patel DJ. PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease. Cell. 2016; 167:1814–1828. [PubMed: 27984729]
36. Swarts DC, van der Oost J, Jinek M. Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. Molecular Cell. 2017; 66:221–233.e4. [PubMed: 28431230]
37. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DBT, Shmakov S, Makarova KS, Semenova E, Minakhin L, et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science. 2016; 353:aaf5573. [PubMed: 27256883]
38. East-Seletsky A, O'Connell MR, Knight SC, Burstein D, Cate JHD, Tjian R, Doudna JA. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature. 2016; 538:270–273. [PubMed: 27669025]
39. Liu L, Li X, Wang J, Wang M, Chen P, Yin M, Li J, Sheng G, Wang Y. Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. Cell. 2017; 168:121–134. [PubMed: 28086085]
40. Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, Verdine V, Donghia N, Daringer NM, Freije CA, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. Science. 2017; 356:438–442. [PubMed: 28408723]
41. Jackson RN, Wiedenheft B. A Conserved Structural Chassis for Mounting Versatile CRISPR RNA-Guided Immune Responses. Mol Cell. 2015; 58:722–728. [PubMed: 26028539]
42. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, van der Oost J. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008; 321:960–964. [PubMed: 18703739]
43. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, Wiedenheft B, Pul U, Wurm R, Wagner R, et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat Struct Mol Biol. 2011; 18:529–536. [PubMed: 21460843]

44••. Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, Guegler CK, Wiedenheft B, Doudna JA, Greene EC. Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System. Cell. 2015; 163:1–12. The authors use single molecule techniques to show that nuclease active Cas3 is recruited to Cascade bound to interference targets. Cascade bound to priming targets only recruits nuclease inactive Cas3 in the presence of Cas1 and Cas2. Furthermore, they show that Cas3 nuclease activity produces short segments of ssDNA.

45•. Hayes RP, Xiao Y, Ding F, van Erp PBG, Rajashankar K, Bailey S, Wiedenheft B, Ke A. Structural basis for promiscuous PAM recognition in type I–E Cascade from E. coli. Nature. 2016; 530:499–503. Crystal structure of Cascade bound to DNA containing a double stranded PAM. This structure explains the PAM recognition mechanism and PAM preferences of Cascade. [PubMed: 26863189]

46. Sashital DG, Wiedenheft B, Doudna JA. Mechanism of foreign DNA selection in a bacterial adaptive immune system. Mol Cell. 2012; 46:606–615. [PubMed: 22521690]

47. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, van der Oost J, Brouns SJJ, Severinov K. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci U S A. 2011; 108:10098–10103. [PubMed: 21646539]

48. Rutkauskas M, Sinkunas T, Songailiene I, Tikhomirova MS, Siksnys V, Seidel R. Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. Cell Rep. 2015; 10:1534–1543.

49. van Erp PBG, Jackson RN, Carter J, Golden SM, Bailey S, Wiedenheft B. Mechanism of CRISPR-RNA guided recognition of DNA targets in Escherichia coli. Nucleic Acids Res. 2015; 43:8381–8391. [PubMed: 26243775]

50. Hochstrasser ML, Taylor DW, Bhat P, Guegler CK, Sternberg SH, Nogales E, Doudna JA. CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. Proc Natl Acad Sci U S A. 2014; 111:6618–6623. [PubMed: 24748111]

51••. Xue C, Whitis NR, Sashital DG. Conformational Control of Cascade Interference and Priming Activities in CRISPR. Immunity Mol Cell. 2016; 64:826–834. The authors show that the Cse1 subunit of Cascade adopts a closed conformation when bound to interference targets and an open conformation when bound to priming targets. They demonstrate that the closed conformation is associated with interference while the open conformation attenuates interference, but increases priming. [PubMed: 27871367]

52. Xue C, Seetharam AS, Musharova O, Severinov K, Brouns SJJ, Severin AJ, Sashital DG. CRISPR interference and priming varies with individual spacer sequences. Nucleic Acids Res. 2015; 43:10831–10847. [PubMed: 26586800]

53•. Blosser TR, Loeff L, Westra ER, Vlot M, Kunne T, Sobota M, Dekker C, Brouns SJJ, Joo C. Two Distinct DNA Binding Modes Guide Dual Roles of a CRISPR-Cas Protein Complex. Mol Cell. 2015; 58:60–70. The authors show that Cascade adopts distinct conformations when bound to priming or interference targets. [PubMed: 25752578]

54. Jackson RN, Golden SM, van Erp PBG, Carter J, Westra ER, Brouns SJJ, van der Oost J, Terwilliger TC, Read RJ, Wiedenheft B. Crystal structure of the CRISPR RNA-guided surveillance complex from Escherichia coli. Science. 2014; 345:1473–1479. [PubMed: 25103409]

55. Mulepati S, Heroux A, Bailey S. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. Science. 2014; 345:1479–1484. [PubMed: 25123481]

56. Zhao H, Sheng G, Wang J, Wang M, Bunkoczi G, Gong W, Wei Z, Wang Y. Crystal structure of the RNA-guided immune surveillance Cascade complex in Escherichia coli. Nature. 2014; 515:147–150. [PubMed: 25118175]

57. Mulepati S, Bailey S. In Vitro Reconstitution of an Escherichia coli RNA-guided Immune System Reveals Unidirectional, ATP-dependent Degradation of DNA Target. J Biol Chem. 2013; 288:22184–22192. [PubMed: 23760266]

58. Sinkunas T, Gasiunas G, Waghmare SP, Dickman MJ, Barrangou R, Horvath P, Siksnys V. In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. EMBO J. 2013; 32:385–394. [PubMed: 23334296]

59. Westra ER, van Erp PBG, Kunne T, Wong SP, Staals RHJ, Seegers CLC, Bollen S, Jore MM, Semenova E, Severinov K, et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol Cell. 2012; 46:595–605. [PubMed: 22521689]

60. Westra ER, Semenova E, Datsenko KA, Jackson RN, Wiedenheft B, Severinov K, Brouns SJJ. Type I-E CRISPR-Cas Systems Discriminate Target from Non-Target DNA through Base Pairing-Independent PAM Recognition. PLoS Genet. 2013; 9:e1003742. [PubMed: 24039596]

61. Fineran PC, Gerritzen MJH, Suarez-Diez M, Kunne T, Boekhorst J, van Hijum SAFT, Staals RHJ, Brouns SJJ. Degenerate target sites mediate rapid primed CRISPR adaptation. Proc Natl Acad Sci U S A. 2014; 111:e1629–e1238. [PubMed: 24711427]

62. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat Comms. 2012; 3:945.

63. Savitskaya E, Semenova E, Dedkov V, Metlitskaya A, Severinov K. High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in E. coli. RNA Biol. 2013; 10:716–725. [PubMed: 23619643]

64. Richter C, Dy RL, McKenzie RE, Watson BNJ, Taylor C, Chang JT, McNeil MB, Staals RHJ, Fineran PC. Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. Nucleic Acids Res. 2014; 42:8516–8526. [PubMed: 24990370]

65. Vorontsova D, Datsenko KA, Medvedeva S, Bondy-Denomy J, Savitskaya EE, Pougach K, Logacheva M, Wiedenheft B, Davidson AR, Severinov K, et al. Foreign DNA acquisition by the I-F CRISPR–Cas system requires all components of the interference machinery. Nucleic Acids Res. 2015; 43:10848–10860. [PubMed: 26586803]

66. Rollins MF, Chowdhury S, Carter J, Golden SM, Wilkinson RA, Bondy-Denomy J, Lander GC, Wiedenheft B. Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. Proc Natl Acad Sci U S A. 2017; doi: 10.1073/pnas.1616395114

67••. Kunne T, Kieper SN, Bannenberg JW, Vogel AIM, Miellet WR, Klein MD, Suarez-Diez M, Brouns SJJ. Cas3-Derived Target DNA Degradation Fragments Fuel Primed CRISPR Adaptation. Mol Cell. 2016; 63:852–864. The authors show *in vitro* Cascade and Cas3 activity produces spacer precursors enriched in PAM sequences that are integrated into the CRISPR locus by Cas1–Cas2. These precursors are produced from both interference and priming targets. [PubMed: 27546790]

68. Tamulaitis G, Venclovas Č, Siksnys V. Type III CRISPR-Cas Immunity: Major Differences Brushed Aside. Trends Microbiol. 2017; 25:49–61. [PubMed: 27773522]

69. Taylor DW, Zhu Y, Staals RHJ, Kornfeld JE, Shinkai A, van der Oost J, Nogales E, Doudna JA. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. Science. 2015; 348:581–585. [PubMed: 25837515]

70••. Elmore JR, Sheppard NF, Ramia N, Deighan T, Li H, Terns RM, Terns MP. Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR–Cas system. Gene Dev. 2016; 30:447–459. The authors show that in the Type III-B Cmr complex from *Pyrococcus furiosus* binding of target RNA, and recognition of the 3′-side or "rPAM" of the target RNA, activates non-sequence specific DNase activity. [PubMed: 26848045]

71•. Estrella MA, Kuo FT, Bailey S. RNA-activated DNA cleavage by the Type III-B CRISPR–Cas effector complex. Gene Dev. 2016; 30:460–470. This paper reports that in the Type III-B Cmr complex from *Thermotoga maritima* binding of complementary ssRNA activates non-sequence specific ssDNA nuclease activity. [PubMed: 26848046]

72•. Kazlauskiene M, Tamulaitis G, Kostiuk G, Venclovas Č, Siksnys V. Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. Mol Cell. 2016; 62:295–306. This paper shows that in the Type III-A Csm complex from *Streptococcus thermophilus* binding of target RNA activates DNase activity, while RNA cleavage and release deactivates DNase activity. [PubMed: 27105119]

73•. Liu TY, Iavarone AT, Doudna JA. RNA and DNA Targeting by a Reconstituted Thermus thermophiles Type III-A CRISPR-Cas System. PLoS One. 2017; 12:e0170552. This manuscript reports in the Type III-A Csm complex from *Thermus thermophilus* complementary RNA binding activates non-sequence specific ssDNase. DNase activation also requires a lack of complementarity with the 5′-repeat of the crRNA. [PubMed: 28114398]

74•. Zhang J, Graham S, Tello A, Liu H, White MF. Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. Nucleic Acids Res. 2016; 44:1789–1799. This paper shows that in the Type III-B and III-D complexes from *Sulfolobus solfataricus* complementary RNA binding activates cleavage of plasmid DNA *in vitro*. [PubMed: 26801642]

75•. Han W, Li Y, Deng L, Feng M, Peng W, Hallstrøm S, Zhang J, Peng N, Liang YX, White MF, et al. A type III-B CRISPR-Cas effector complex mediating massive target DNA destruction. Nucleic Acids Res. 2016; The authors show that in the Type III-B Cmr complex from *Sulfolobus islandicus* cleaves RNA binding activates ssDNase activity, but activation required non-complementarity between the 5′-handle of the guide and 3′-flank of the RNA target. doi: 10.1093/nar/gkw1274

76. Osawa T, Inanaga H, Sato C, Numata T. Crystal Structure of the CRISPR-Cas RNA Silencing Cmr Complex Bound to a Target Analog. Mol Cell. 2015; 58:418–430. [PubMed: 25921071]

77. Marraffini LA, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature. 2010; 463:568–571. [PubMed: 20072129]

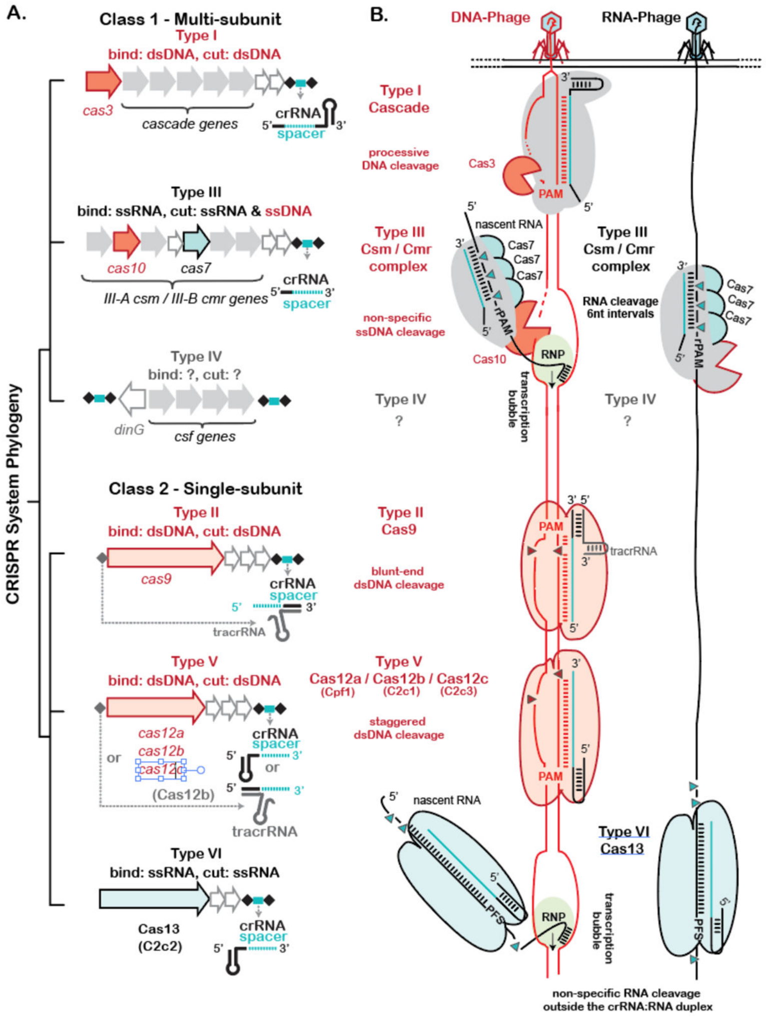

Figure 1. Diverse CRISPR systems defend against DNA- and RNA-based invaders

A. Phylogenetic tree of CRISPR system Types with nucleic acid substrates that are bound and cleaved. DNA targeting nucleases are labeled red and RNA targeting nucleases are labeled blue. Genes coding for multi-subunit complexes are indicated with the brackets. **B.** Cartoon representations of each CRISPR system targeting DNA phage, RNA phage, or both.

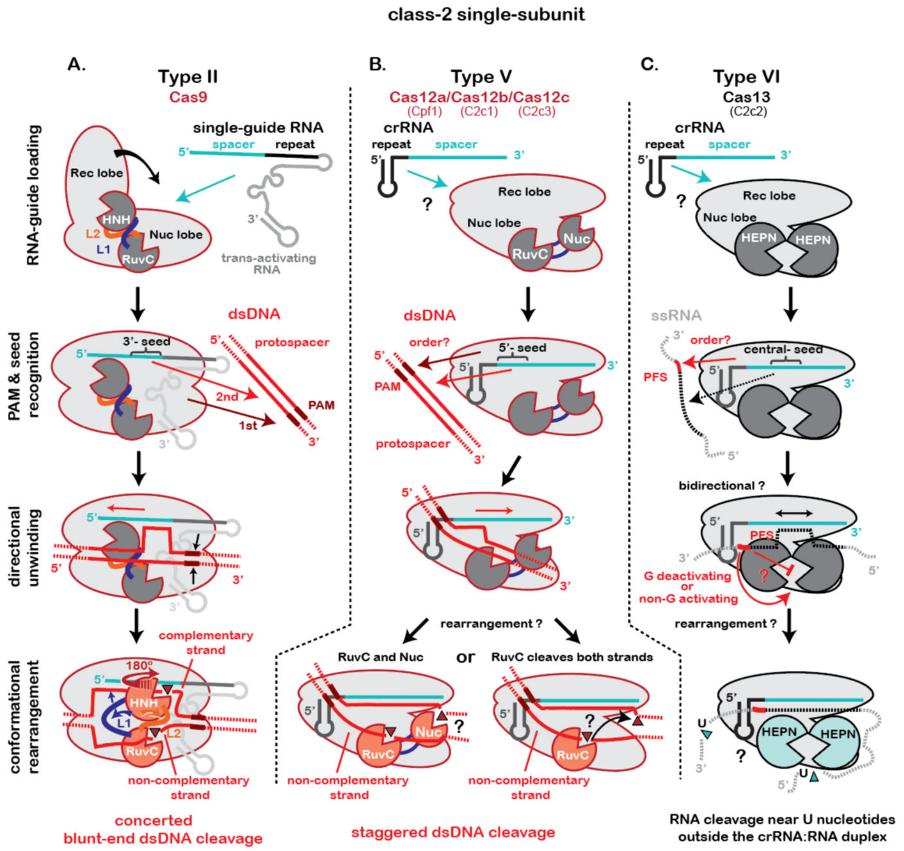

Figure 2. Class 2 Nuclease activation mechanisms

A. Activation of the Type II Cas9 nuclease (grey with red outline) relies on several regulatory checkpoints. Cas9 adopts a bi-lobed structure that consists of a nuclease-lobe (Nuc lobe) and an alpha-helical recognition lobe (Rec lobe). Single-guide-RNA loading causes a conformational rearrangement in the Rec lobe. To bind duplex DNA, a PAM motif (dark red) is recognized followed by 3′ seed complementation. Directional unwinding of the duplex to the PAM-distal side of the guide induces a conformational rearrangement of two peptide linkers (L1 and L2 colored blue and orange) that stabilize the active conformation of the HNH and RuvC nuclease domains (colored red) to make a blunt-end cut in duplex DNA. **B**. Type V nucleases (Cas12a, Cas12b, and Cas12c; also known as Cpf1, C2c1, and C2c3) also adopt a bi-lobed architecture, are loaded with an RNA-guide and use PAM and seed recognition to initiate target binding with DNA. However, the DNA cleavage mechanisms may involve the RuvC domain and Nuc domain, or just the RuvC domain. **C.** Like Type II and V, the Type VI (Cas13, also known as C2c2) proteins adopt a bilobed architecture and are programmed with an RNA-guide. Type VI nucleases bind RNA with a central seed within the RNA-guide, and an adjacent sequence called the PFS (Protospacer Flanking Site) plays a role in nuclease activation.

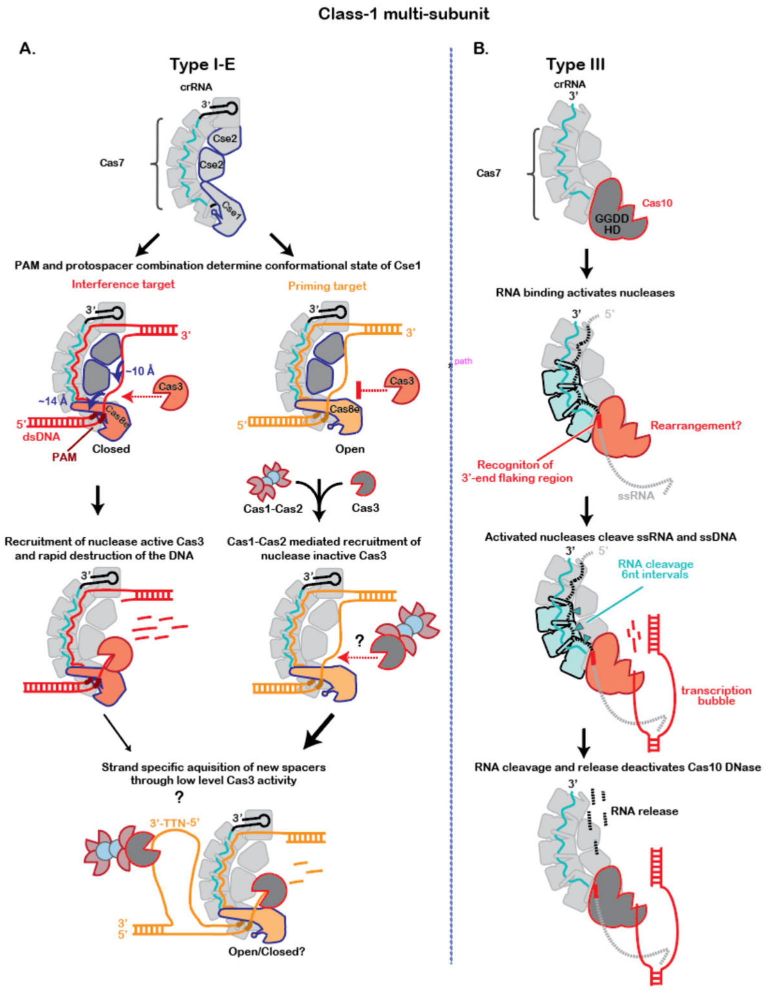

Figure 3. Regulation of Class 1 Cas nucleases

A. Type I-E Cascade binds to both interference (red) and priming (orange) DNA targets. Binding of dsDNA causes a conformational change in the Cas8e and Cse2 subunits (blue arrows). When bound to interference targets the Cas8e subunit adopts a closed conformation that recruits nuclease active Cas3 (red). When bound to priming targets the Cas8e subunit adopts an open conformation that does not recruit Cas3 directly, but relies on the presence of Cas1–Cas2 for the recruitment of nuclease repressed Cas3 (gray). Cas3 helicase activity produces looped sections of ssDNA which can act as sources of new spacers acquired during priming. **B.** In Type III systems the multi-subunit complexes bind ssRNA complementary to the crRNA-guide and RNase active sites in the Cas7 backbone subunits (blue) cleave the RNA at regular 6-nt intervals. In addition, RNA binding also activates non-specific DNase activity of the Cas10 subunit, which in some systems also requires recognition of the 3′-end flanking sequence (red). Cleavage and release of the RNA deactivates the Cas10 subunit (gray) and ensures temporal control.