frontiers
in Microbiology

Positioning Diverse Type IV Structures and Functions Within Class 1 CRISPR-Cas Systems

Hannah N. Taylor[1], Eric Laderman[2], Matt Armbrust[1], Thomson Hallmark[1], Dylan Keiser[1], Joseph Bondy-Denomy[2] and Ryan N. Jackson[1]*

[1] Department of Chemistry and Biochemistry, Utah State University, Logan, UT, United States, [2] Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA, United States

Type IV CRISPR systems encode CRISPR associated (Cas)-like proteins that combine with small RNAs to form multi-subunit ribonucleoprotein complexes. However, the lack of Cas nucleases, integrases, and other genetic features commonly observed in most CRISPR systems has made it difficult to predict type IV mechanisms of action and biological function. Here we summarize recent bioinformatic and experimental advancements that collectively provide the first glimpses into the function of specific type IV subtypes. We also provide a bioinformatic and structural analysis of type IV-specific proteins within the context of multi-subunit (class 1) CRISPR systems, informing future studies aimed at elucidating the function of these cryptic systems.

Keywords: CRISPR, Cas, type IV, Cas7, Cas6, DinG helicase, CysH

INTRODUCTION

Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated (CRISPR-Cas) prokaryotic defense systems utilize Cas1 and Cas2 proteins, along with system-specific proteins such as Cas4, IHF, Csn2, and Cas9, to integrate foreign genetic material into the CRISPR locus, immunizing the cell against viruses and plasmids (Datsenko et al., 2012; Yosef et al., 2012; Nuñez et al., 2014, 2016; Heler et al., 2015; Rollie et al., 2015; Wang et al., 2015; Wei et al., 2015; Sternberg et al., 2016; Jackson S. A. et al., 2017; Kieper et al., 2018; Lee et al., 2018). To provide immunity, the CRISPR locus is transcribed and processed by RNA nucleases into CRISPR derived RNAs (crRNAs) (Brouns et al., 2008; Marraffini and Sontheimer, 2008; Haurwitz et al., 2010; Deltcheva et al., 2011). The crRNAs combine with Cas proteins to form ribonucleoprotein (RNP) complexes, which recognize and bind complementary nucleic acids. Binding induces cleavage of the foreign nucleic acid, protecting the cell (Brouns et al., 2008; Carte et al., 2008; Hale et al., 2008, 2009; Marraffini and Sontheimer, 2008; Garneau et al., 2010; Jackson R. N. et al., 2017; Hille et al., 2018).

Although all CRISPR systems use these general mechanisms to achieve immunity, the systems themselves are remarkably diverse, comprising two classes (1–2), six types (I–VI), and at least 33 subtypes (Yan et al., 2019; Makarova et al., 2020). In class 2 systems (types II, V, VI) a single Cas protein binds the crRNA to form the RNP complex, while class 1 RNP complexes (types I, III, IV) bind the crRNA with several proteins. Of the six CRISPR-Cas types, the least understood is type IV. Recent bioinformatic, biochemical, and structural studies of type IV CRISPR-Cas

systems have provided valuable insights into type IV system function. Here we compile known data on type IV systems, highlight recent advances in type IV system biology and biochemistry, and indicate questions concerning type IV systems that need to be addressed. Additionally, we provide phylogenetic analyses that suggest ancillary proteins associated with type IV systems have evolved Cas-specific functions.

TYPE IV SYSTEMS ARE MINIMAL, MOBILE CRISPR-CAS SYSTEMS

Distinguishable from other CRISPR-Cas systems, Type IV systems encode a distinct *cas7*-like gene (*csf2*), lack adaptation genes, rarely encode an obvious nuclease, and are primarily found on plasmids (Koonin and Makarova, 2017, 2019; Pinilla-Redondo et al., 2019). These unique features of type IV systems have made it difficult to predict the function of type IV systems.

All type IV systems encode homologs of proteins known to form multi-subunit RNP complexes, explaining their class 1 designation. However, the presence of specific genes, gene arrangements, and differences in gene sequences have been used to further classify type IV systems into three distinct subtypes (IV-A, IV-B, and IV-C) (Makarova et al., 2011, 2015, 2020). Types IV-A, IV-B, and IV-C each contain a subtype-specific gene (*dinG, cysH-like,* and *cas10-like*, respectively) and subtype-specific features (**Figure 1A**). Type IV-A operons encode the three core type IV genes (*csf1, csf2,* and *csf3*), an endoribonuclease (*cas6/csf5*), a CRISPR array, and a putative helicase (*dinG*). Type IV-B operons encode the three core type IV genes and a *cas11-like* gene but lack a CRISPR locus. Additionally, type IV-B operons contain an ancillary gene, labeled *cysH-like* because the predicted secondary structure of its protein product resembles the CysH enzyme (Shmakov et al., 2018; Faure et al., 2019). Type IV-C systems encode *csf2* and *csf3*, but in place of the *csf1* gene they encode a *cas10-like* gene with a putative HD-nuclease domain. They also encode the *cas11-like* gene observed in IV-B systems, and sometimes a *cas6* RNA endonuclease and CRISPR array (Pinilla-Redondo et al., 2019; Makarova et al., 2020).

It is curious that the type IV systems that encode CRISPR loci do so in the absence of adaptation genes. It has been hypothesized that these type IV systems commandeer adaptation machinery from other CRISPR-Cas types to maintain their CRISPRs, similar to some type III systems (Staals et al., 2013, 2014; Elmore et al., 2015; Bernheim et al., 2020). Supporting this hypothesis, recent bioinformatic work showed that some type IV-A subtypes co-localize with certain type I systems (e.g., I-F, I-E), and that spacers found within co-localized type IV CRISPR loci appeared to be selected with the same criteria utilized by the type I system adaptation machinery [e.g., both I-E and IV-A protospacers are flanked with an 5′-AAG-3′ protospacer adjacent motif (PAM)] suggesting there may be functional cross-talk between these systems (Pinilla-Redondo et al., 2019). To confirm this proposed cooperation, *in vivo* and *in vitro* experimental work that examines adaptation in type IV systems with adaptation proteins from co-localized systems is needed.

THE *cas7*-LIKE GENE, *csf2*, DISTINGUISHES TYPE IV FROM OTHER CLASS 1 SYSTEMS

Initial bioinformatic analyses proposed *csf1* as the type IV *cas* signature gene (Makarova et al., 2015). However, some type IV systems lacking *csf1* have been identified, necessitating that the type IV *cas7* homolog, *csf2,* be used to classify type IV systems (Pinilla-Redondo et al., 2019). In type I and type III systems, Cas7-like proteins bind the crRNA guide within a helical backbone of a multi-subunit RNP complex and make direct interactions with other protein subunits (Jore et al., 2011; Lintner et al., 2011; Staals et al., 2013; Jackson et al., 2014; Mulepati et al., 2014; Osawa et al., 2015). Similarly, a recent cryo-EM structure of the type IV-B RNP complex revealed that Csf2 proteins bind RNA within a helical backbone, indicating a conserved function for Cas7-like proteins in all class 1 systems (Zhou et al., 2021). Despite this conservation, the sequence and structure of Csf2 is distinguishable from other Cas7 proteins (Makarova et al., 2011; **Supplementary Figure 1**). For example, when representative Cas7 sequences from all class 1 subtypes were aligned and a phylogenetic tree created, Csf2 sequences clustered on a separate branch from type I and type III Cas7 sequences (**Supplementary Figure 2A** and **Supplementary Methods**). Csf2 is distinct from other Cas7 homologs but appears to be most closely related to type III, supporting evolutionary hypotheses that type IV systems diverged from type III systems (Koonin and Makarova, 2019; Pinilla-Redondo et al., 2019; Makarova et al., 2020). Interestingly, an alignment of only Csf2 sequences shows clustering of Csf2 from each type IV subtype on its own branch, illustrating the intrinsic diversity of type IV subtypes and suggesting subtype-specific functional distinctions (**Supplementary Figure 2B**). It is worth noting that the type IV-B Csf2 subunit structure is most similar to the structure of the Cas7 homolog in type III-A systems, Csm3 (Zhou et al., 2021). Csm3 contains a catalytic aspartate that cleaves RNA targets (Tamulaitis et al., 2014). Alignment of target-bound Csm3 with Csf2 indicates that, although Csf2 also contains a conserved aspartate residue in a similar location, it is not in a position amenable for target cleavage (Zhou et al., 2021). Additional structural studies of type IV complexes bound to nucleic acid targets and complementary biochemical assays are needed to determine whether Csf2 is capable of RNA nuclease activity.

TYPE IV-A SYSTEMS ARE DEFENSE SYSTEMS WITH AN UNKNOWN MECHANISM OF ACTION INVOLVING A DinG HELICASE

Recently, a type IV-A system from *Pseudomonas aeruginosa* was shown to exhibit crRNA-guided defense against plasmids (Crowley et al., 2019), consistent with an analysis of type IV CRISPR spacers that suggested type IV-A systems disproportionately target plasmids (Pinilla-Redondo et al., 2019). Notably, earlier bioinformatic work indicated that
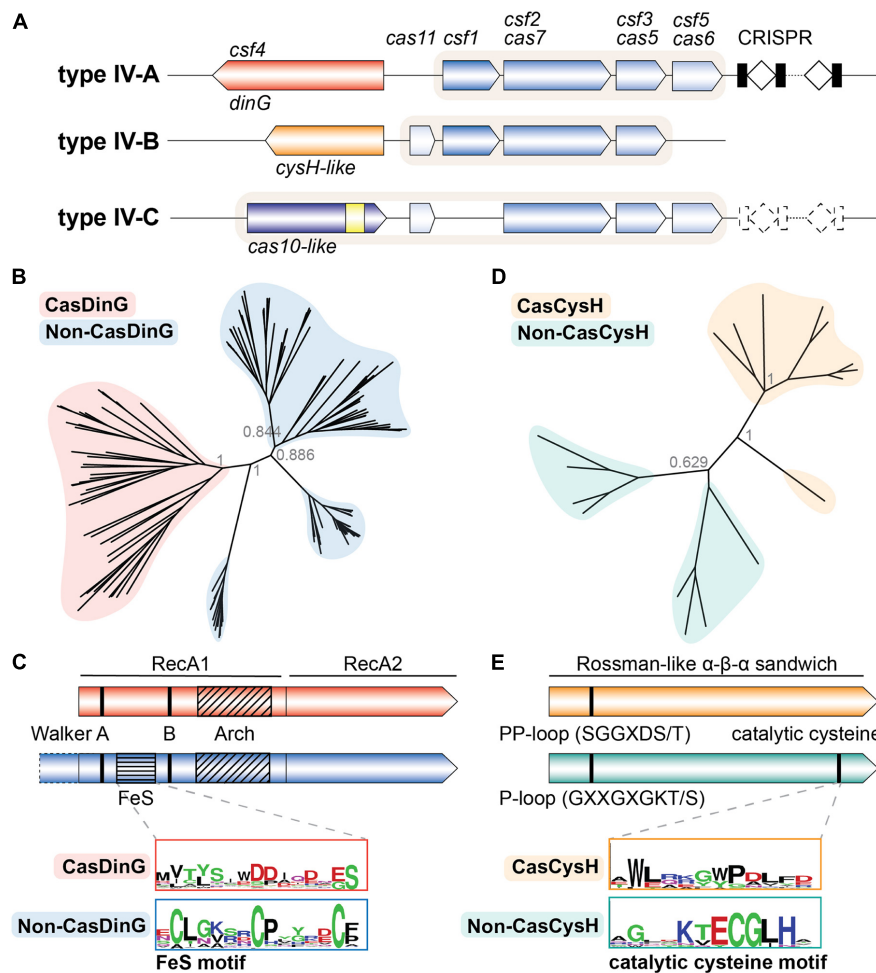
FIGURE 1 | The type IV Cas accessory proteins have evolved a Cas specific function. **(A)** Classification schematic of type IV CRISPR-Cas systems. A typical locus is represented for each type IV subtype. Dashed lines indicate components that are sometimes not encoded by the subtype. Shaded backgrounds highlight which gene products form the ribonucleoprotein (RNP) complex. The yellow square in the IV-C *cas10-like* large subunit represents an HD nuclease domain. **(B)** Phylogenetic tree of Cas- and non-CasDinG sequences. Posterior probabilities are shown. **(C)** Cartoons of Cas- and non-CasDinG sequences indicating positions of certain helicase motifs and domain architecture. Weblogos (Crooks, 2004) of the FeS cluster region in non-CasDinG (below, blue outline) and CasDinG (top, red outline) are shown. **(D)** Phylogenetic tree of Cas- and non-CasCysH sequences. Posterior probabilities are shown. **(E)** Cartoons of Cas- and non-CasCysH sequences. CasCysH is predicted to adopt the Rossman-like α - β - α fold observed in non-CasCysH structures. Positions and sequences of P- and PP-loops are indicated. Weblogos of the catalytic cysteine in non-CasCysH (bottom, teal outline) and CasCysH (top, orange outline) are shown.

many type IV-A spacers target viruses and prophage sequences encoding putative anti-CRISPRs, suggesting type IV-A systems also actively target viruses (Shmakov et al., 2017; Yin et al., 2019; Nobrega et al., 2020). However, direct data, such as viral plaque assays, are needed to confirm that type IV-A systems protect against viral attack.

Structural and biochemical work on a type IV-A complex from *Aromatoleum aromaticum* and IV-A Cas6 from *Mahella australiensis* demonstrated that the RNA endonuclease Csf5/Cas6 processes a crRNA upon which Csf1, Csf2, Csf3, and Csf5 form an RNP complex (Özcan et al., 2018; Taylor et al., 2019). At least three distinct crRNA processing endoribonucleases are encoded by Type IV-A systems (Cas6, Csf5, and Cas6e) (Makarova et al., 2020; **Supplementary Figure 3A**). Sequence alignments between biochemically characterized and putative

type IV Csf5/Cas6 enzymes revealed Csf5 enzymes cleave RNA with arginine active site residues, while type IV Cas6 and Cas6e enzymes utilize histidine/tyrosine active site residues (**Supplementary Figures 3B, 4**). Despite these obvious differences in endoribonucleases, we hypothesize that in all type IV-A systems the Csf1, Csf2, Csf3, and Csf5/Cas6 proteins bind to the processed crRNA to form a multi-subunit complex that binds complementary nucleic acid.

It remains unclear whether type IV RNP complexes bind single stranded RNA [like the type III Csm and Cmr complexes (Hale et al., 2009; Samai et al., 2015)] or double stranded DNA [like the type I Cascade complexes (Brouns et al., 2008)] and how type IV complexes distinguish self from non-self targets. RNPs that target dsDNA usually rely on a protein-mediated binding event with a specific non-self sequence adjacent to

the complementary target, called a PAM (protospacer adjacent motif) (Mojica et al., 2009; Westra et al., 2012, 2013). PAM binding provides the energy for target duplex unwinding and interrogation of the DNA by the crRNA-guide. Work by Pinilla-Redondo et al. (2019) identified a consensus PAM (5′-AAG-3′) flanking protospacers targeted by a subset of type IV-A systems, suggesting type IV-A systems rely on PAM recognition to license binding. However, the consensus PAM may only reflect a preference of the acquisition machinery, which may explain why consensus PAM sequences have not been identified in all IV-A systems. Reliance on a specific PAM sequence for type IV-A RNP interference remains to be confirmed experimentally, but it should be noted that a promiscuous PAM recognition mechanism may indicate that the type IV complexes have evolved to accommodate the preferences of diverse Cas1 and Cas2 proteins that use different PAM sequences in spacer acquisition.

Interestingly, the structural similarities of the type IV-B complex to the type III Csm complex suggest that type IV complexes may target RNA (Zhou et al., 2021). Instead of recognizing a "non-self" PAM to license base pairing with a double-stranded DNA target, RNPs that bind RNA generally use a "self recognition" mechanism to distinguish self from non-self sequences (Marraffini and Sontheimer, 2010). Self-sequence located in the flanking regions of a bound RNA can base pair with the direct repeat of the crRNA disrupting downstream activation of effector nucleases (You et al., 2019). Self-sequences are inhibitory to overall immune function (Marraffini and Sontheimer, 2010; Elmore et al., 2016; Estrella et al., 2016; Kazlauskiene et al., 2016; Zhang et al., 2016; Han et al., 2017; Liu et al., 2017), but in some systems only a subset of non-self protospacer flanking sequences [called RNA-PAMs (rPAM) in type III systems or protospacer flanking sites (PFS) in type VI systems] are activating (Marraffini and Sontheimer, 2010; Abudayyeh et al., 2016; Elmore et al., 2016). We suspect that one or more Csf subunits may be responsible for PAM recognition to license DNA binding or rPAM recognition to activate immunity. We anticipate that in vivo PAM screens and biochemical binding assays with purified type IV-A RNPs will reveal the type IV-A self vs. non-self recognition mechanism.

Type IV-mediated plasmid clearance required all type IV-A system genes (csf1, csf2, csf3, csf5, and dinG/csf4) and a CRISPR containing a spacer complementary to a target plasmid sequence adjacent to a 5′-TTC-3′ PAM (Crowley et al., 2019). Because deleting the dinG gene or mutating the ATPase active site residues (DEAH-box) fully disrupted plasmid clearance, we hypothesize that RNP complex binding recruits the type IV-associated DinG (CasDinG) helicase to the bound target and CasDinG-mediated ATP binding and hydrolysis performs work, such as duplex unwinding, that is essential for plasmid clearance. Such a mechanism is similar to the more extensively studied type I Cas3 helicase-nuclease that unwinds and degrades dsDNA targets bound by the type I Cascade RNP complex (Beloglazova et al., 2011; Mulepati and Bailey, 2011; Sinkunas et al., 2011).

Both DinG and Cas3 classify as superfamily 2 helicases but, unlike Cas3, CasDinG proteins have no identifiable nuclease domain and have yet to be biochemically or structurally characterized (Fairman-Williams et al., 2010;

Makarova et al., 2020). DinG helicases are generally involved in DNA recombination and repair, and are classified by amino acid sequence motifs involved in ATP binding and hydrolysis and nucleic acid binding and translocation (Lewis et al., 1992; Voloshin et al., 2003; Voloshin and Camerini-Otero, 2007; McRobbie et al., 2012; Wu and Brosh, 2012; Thakur et al., 2014; Cheng and Wigley, 2018). The motifs are located across two RecA helicase domains (**Supplementary Figure 5**). The first helicase domain also harbors two insertions, an iron sulfur cluster domain, and an arch domain, which are both important for duplex strand splitting (Ren et al., 2009; Peissert et al., 2020).

Since non-CasDinG helicases and their homologs have been extensively studied biochemically and structurally, we hypothesized that an in-depth comparison of CasDinG with non-CasDinG sequences would provide insight to CasDinG function. To investigate the relationship of CasDinG to other DinG helicases, we compiled CasDinG and non-CasDinG sequences from organisms containing a type IV-A system and generated a phylogenetic tree (**Supplementary Methods**; **Figure 1B**). Interestingly, CasDinG and non-CasDinG sequences clustered separately even when the sequences were retrieved from the same organism, suggesting CasDinG is functionally distinct from non-CasDinG. Notably, CasDinG helicases contain insertions within the first RecA domain of the same lengths as the iron-sulfur and arch insertions, but they lack homology with non-CasDinG sequences, including the residues important for coordinating the iron-sulfur cluster (**Figure 1C**). Sequence differences in these regions suggest these inserts may be a source of functional distinctions important for defense activities. Many functions for CasDinG have been hypothesized, including a role in displacing bound RNP complexes, cleaving bound targets with an unidentified nuclease activity (perhaps housed within an insert), or recruitment of endogenous nucleases to bound targets (Grodick et al., 2014). Notably, DinG helicases have been observed in a few type I and III systems (Dwarakanath et al., 2015; Makarova et al., 2020), indicating an evolutionary link and suggesting that some CasDinG activities essential for type IV immunity may have been co-opted by other class 1 systems.

In summary, recent bioinformatic and in vivo studies have indicated type IV-A systems protect prokaryotes from plasmids and viruses, but the mechanisms that underpin how the Csf RNP complex and CasDinG work together to provide immunity remain to be determined.

TYPE IV-B SYSTEMS ENCODE AN RNP COMPLEX OF UNKNOWN FUNCTION AND A SPECIALIZED CysH-LIKE PROTEIN WITH PUTATIVE ATP α-HYDROLASE ACTIVITY

Unlike type IV-A and IV-C subtypes, type IV-B systems lack a CRISPR locus and a crRNA processing enzyme, and are associated with an ancillary gene identified as cysH-like by the HHpred secondary structure prediction and alignment tool (Zimmermann et al., 2018; Makarova et al., 2020; **Figure 1A**).

A recent structural study recombinantly expressed and purified a *Mycobacterium sp. JS623* IV-B Csf RNP complex containing four type IV-B proteins (Csf1, Csf2, Csf3, and Cas11) (Zhou et al., 2021). Interestingly, RNA sequencing revealed the type IV-B Csf complex bound small heterogeneous RNAs, instead of co-expressed type I-E crRNAs from the *Mycobacterium sp. JS623* plasmid, suggesting a possible function other than CRISPR-mediated defense. A high resolution cryo-EM structure of the complex revealed several Csf2 subunits bind an RNA within a helical filament, while Cas11 subunits form a minor filament that contacts the larger filament at Csf2 dimer interfaces (Zhou et al., 2021). This structure of intertwined large and small protein filaments is similar to other class 1 RNP complexes, suggesting similar function as an RNA-guided complex that binds complementary targets (Jore et al., 2011; Lintner et al., 2011; Staals et al., 2013; Jackson et al., 2014; Mulepati et al., 2014; Osawa et al., 2015; **Supplementary Figure 6**).

Several observations are currently confounding an understanding of the type IV-B complex function. First, electron density for Csf1 and Csf3 subunits was not clearly observed within the structure, although SDS-PAGE indicated their presence in the purified complex. Thus, the structure and function of these important proteins remains unknown. Second, because the IV-B Csf complex bound heterogenous RNA, it remains unknown whether the Csf complex lacks sequence-specific preference for small RNAs or if the RNA(s) that the complex would normally bind were not available in the recombinant expression conditions. Finally, the role of the strictly conserved ancillary CysH-like protein and how it may interact with the complex is unknown.

The key to understanding the function of type IV-B systems likely lies with the uncharacterized, but ubiquitous, type IV-B accessory *cysH*-like gene (Shmakov et al., 2018; Faure et al., 2019). Typical CysH proteins are phosphoadenosine phosphosulfate (PAPS) reductases involved in sulfate assimilation. Structures reveal CysH proteins fit within a family of enzymes that adopt a Rossman-like α–β–α sandwich fold that binds nucleotides (InterPro IPR014729) (Blum et al., 2020). CysH proteins also contain a P-loop motif (GXXGXGKT/S consensus sequence) that binds nucleotide phosphates, and a conserved C-terminal cysteine that performs nucleophilic attack on the PAPS β-sulfate, hydrolyzing PAPS at the α-phosphate and forming a covalent thiosulfanate intermediate during sulfur reduction (Savage et al., 1997; Carroll et al., 2005; Chartron et al., 2007). Interestingly, the DndC protein from the recently discovered DND bacterial immune system also belongs to the PAPS reductase family, and uses a similar mechanism to incorporate sulfur into the backbone of chromosomal DNA through a disulfide cysteine (You et al., 2007; Wang et al., 2018; Faure et al., 2019). These phosphorothioate modifications serve as an epigenetic signature that allows the DND system to distinguish self from non-self DNA (Wang et al., 2018). The predicted structural homology between the type IV-B CysH (CasCysH) and DndC proteins justifies speculation that CasCysH proteins perform a similar function. However, a closer analysis of type IV-B CasCysH sequences suggests that if CasCysH does epigenetically modify

DNA, it will not be through the formation of phosphorothioates. Although HHPred predicts CasCysH adopts a Rossman-like α–β–α sandwich fold, the catalytic cysteine important for sulfonate reduction in non-CasCysH and phosphothiolation of DNA by DndC is absent. Additionally, the P-loop sequence of CasCysH is more similar to the PP-motif (pyrophosphatase motif) (SGGXDS/T consensus sequence) observed in ATP PPases (Bork and Koonin, 1994; **Figure 1E**).

To better understand CasCysH activity and to explore the relationship between non-CasCysH and CasCysH proteins, sequences from organisms encoding both Cas- and non-CasCysH were aligned and phylogenetic trees determined. As was seen with Cas- and non-CasDinG, CasCysH sequences cluster separately from non-CasCysH sequences even when the sequences were retrieved from the same organisms (**Figure 1D**). Together, with our more in depth sequence analysis these differences suggest CasCysH evolved to preserve nucleotide binding without sulfonucleotide reduction.

Non-CasCysH enzymes fall within the larger classification of ATP α-hydrolases, which include N-type ATP PPases (Savage et al., 1997). Unlike non-CasCysH and DndC, N-type ATP PPases catalyze sequential reactions involving substrate AMPylation, instead of the formation of covalent enzyme substrate intermediates requiring nucleophilic attack from a catalytic cysteine (Chartron et al., 2007; Wang et al., 2018). The absence of a catalytic cysteine suggests that the role of CasCysH is to stabilize the AMPylation of specific substrates through catalysis of ATP α-hydrolase activity. We hypothesize that such an activity could be used to modify nucleic acids bound by the type IV-B RNP for immune system purposes, gene regulation, or the formation of secondary messengers. Future biochemical studies aimed at defining the function of CasCysH and its interactions with the IV-B Csf RNP complex will be critical for understanding type IV-B systems.

Several hypotheses exist concerning the function of type IV-B CRISPR-Cas systems. As they lack both a CRISPR array and an obvious nuclease it seems unlikely that type IV-B systems function as independent CRISPR-Cas defense systems (Makarova et al., 2011; Faure et al., 2019; Zhou et al., 2021). It has been suggested that type IV-B systems could bind crRNAs derived from other CRISPR systems, forming IV-B RNP complexes that perform RNA-guided defense (Makarova et al., 2011, 2015; Koonin and Makarova, 2019). As type IV systems are generally encoded on plasmids, such a crRNA scavenging system could be passed between organisms, acting as a mobile defense system. Interestingly, it was recently shown that sometimes type IV-B systems colocalize with specific class 1 systems, suggesting a cooperative function (Pinilla-Redondo et al., 2019). However, the same study showed that type IV-B systems are most often observed without other CRISPR systems, supporting an alternative hypothesis that proposes IV-B systems may protect plasmids from RNA-guided defense mechanisms by sponging up and inactivating small guide RNAs, including crRNAs (Koonin and Makarova, 2017; Faure et al., 2019; Pinilla-Redondo et al., 2019). Such an anti-guide-RNA activity could give plasmids containing
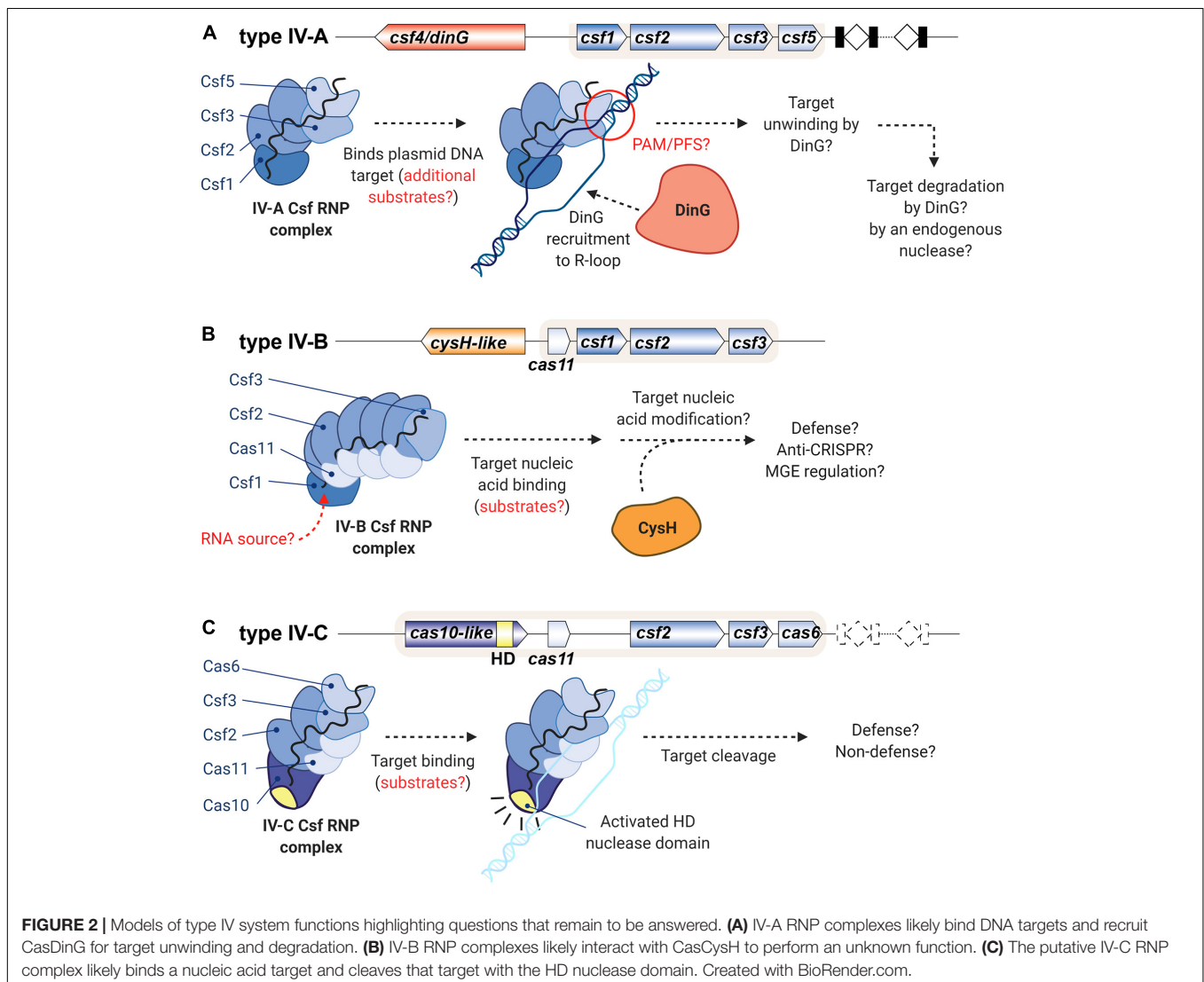
a type IV-B system a selective advantage (Shmakov et al., 2018; Koonin and Makarova, 2019). Although intriguing, neither of these hypotheses explain the role of the highly conserved ancillary protein CasCysH, suggesting the true function of IV-B systems may be more intricate than has so far been proposed.

THE NEWLY CLASSIFIED TYPE IV-C SYSTEM HIGHLIGHTS THE DIVERSE NATURE OF TYPE IV CRISPR-Cas SYSTEMS

Only recently did bioinformatics studies classify the subtype IV-C CRISPR-Cas system (Pinilla-Redondo et al., 2019; Makarova et al., 2020). Type IV-C systems lack a Csf1 subunit, and instead encode a Cas10-like subunit with an HD nuclease domain (**Figure 1A**). Type III CRISPR-Cas systems also encode

Cas10, which is the large subunit of the RNP complex. In type III systems Cas10 has nuclease activity and synthesizes cyclic oligoadenylate second messengers (Jung et al., 2015; Kazlauskiene et al., 2017; Niewoehner et al., 2017). The type IV Cas10 contains an HD nuclease domain but not a nucleotide cyclase motif "GGDD," suggesting it has nuclease but not cyclic adenylate synthetase activity (Pinilla-Redondo et al., 2019). Interestingly, the HD domain motifs of type IV Cas10 are more similar to the HD domain of Cas3 than the type III Cas10 (Aravind and Koonin, 1998; Makarova et al., 2020). The presence of a *cas10-like* gene in type IV-C systems and the similarity between the type III-A and type IV-B RNP complexes support proposals that type IV and type III CRISPR-Cas systems share a common ancestor (Pinilla-Redondo et al., 2019; Makarova et al., 2020; Zhou et al., 2021). Experimental work is needed to better understand the function of these fascinating systems.

Several variants of type IV systems have been identified in bioinformatics studies and clinical samples which include

FIGURE 2 | Models of type IV system functions highlighting questions that remain to be answered. **(A)** IV-A RNP complexes likely bind DNA targets and recruit CasDinG for target unwinding and degradation. **(B)** IV-B RNP complexes likely interact with CasCysH to perform an unknown function. **(C)** The putative IV-C RNP complex likely binds a nucleic acid target and cleaves that target with the HD nuclease domain. Created with BioRender.com.

type IV systems; without a *csf1*, with a *csf1-csf3* fusion, with a *recD* helicase instead of *dinG*, and in association with IncH1b plasmids (Crowley et al., 2019; Pinilla-Redondo et al., 2019; Kamruzzaman and Iredell, 2020; Newire et al., 2020). We expect further study and analysis of these diverse systems will reveal unique mechanisms and functions.

UNANSWERED QUESTIONS CONCERNING TYPE IV BIOCHEMISTRY AND BIOLOGICAL FUNCTION

Throughout this perspective we have highlighted pressing questions concerning type IV CRISPR-Cas system structures and functions. Here we suggest models for the function of each type IV subtype and indicate areas which require further understanding. Type IV-A systems have been shown to form RNP complexes and prevent targeted plasmid transformation, but they have not been shown to target viruses nor is the mechanism of crRNA-guided defense clear (**Figure 2A**). Understanding the targets of the type IV-A system is critical to understanding the full scope of its defense activity. The presence of a helicase within the system suggests the need to unwind a duplex substrate. We suspect that the type IV-A system targets dsDNA, as it can defend against invasive plasmids (Crowley et al., 2019). However, CasDinG could also be important for unwinding duplex secondary structure within ssRNA targets or for targeting dsRNA phages (Poranen and Tuma, 2004). Remaining questions include the targeting parameters of the complex (DNA vs RNA, seed sequence, mismatch tolerance), how the complex distinguishes self from non-self, the role of CasDinG in immunity, and how targets are cleared without an identifiable nuclease domain within the system. We speculate that the IV-A RNP complex will bind to a dsDNA target and recruit CasDinG to the resulting R-loop, allowing CasDinG to unwind the target. To clear the target from the cell, we hypothesize that either an endogenous nuclease will degrade the unwound nucleic acid, or CasDinG harbors an intrinsic nuclease activity not predicted by the sequence.

The function of type IV-B Csf RNP complexes is still unknown (**Figure 2B**). Many questions of type IV-B system function will be answered as the source of the RNA component of the IV-B RNP complex is discovered and the function of the accessory protein CasCysH is understood. We propose that the Csf RNP complex will bind a nucleic acid target and recruit CasCysH to modify the nucleic acid via an ATP α-hydrolase activity.

No biochemical studies have been performed with type IV-C systems, to date. We hypothesize that the IV-C Csf proteins will form an RNP complex with a crRNA and the Cas10-like subunit (**Figure 2C**). The IV-C Csf RNP complex will bind a nucleic acid target complementary to the crRNA and the HD nuclease domain of the Cas10-like subunit will cleave the target. Some IV-C systems have a CRISPR and a crRNA processing endonuclease and others do not, suggesting some IV-C systems may serve a crRNA-guided defense function while others may employ Cas proteins to perform an entirely different, non-defense function. Future studies should seek

to understand the role of Cas10 within the type IV Csf RNP complex and the overall function of type IV-C CRISPR-Cas systems.

To understand the function of type IV CRISPR-Cas systems, it is critical that we determine the structures and biochemical functions of the type IV subtype specific proteins: CasDinG, CasCysH, and Cas10-like. Phylogenetic trees suggest that the IV-A DinG and IV-B CysH have evolved to support a putative Cas specific function. The IV-C Cas10 also has a unique domain composition that likely supports a distinct function. We propose that, due to the different accessory proteins and subtype specific proteins encoded by the three subtypes, each type IV subtype will have a distinct mechanism of action and possibly distinct function. We highly anticipate future work detailing the mechanisms and functions of type IV RNP complexes and their accessory proteins.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021. 671522/full#supplementary-material

REFERENCES

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353:6299. doi: 10.1126/science.aaf5573

Aravind, L., and Koonin, E. V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23, 469–472. doi: 10.1016/S0968-0004(98)01293-6

Beloglazova, N., Petit, P., Flick, R., Brown, G., Savchenko, A., and Yakunin, A. F. (2011). Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* 30, 4616–4627. doi: 10.1038/emboj.2011.377

Bernheim, A., Bikard, D., Touchon, M., and Rocha, E. P. C. (2020). Atypical organizations and epistatic interactions of CRISPRs and cas clusters in genomes and their mobile genetic elements. *Nucl. Acids Res.* 48, 748–760. doi: 10.1093/nar/gkz1091

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., et al. (2020). The InterPro protein families and domains database: 20 years on. *Nucl. Acids Res.* 49, D344–D354. doi: 10.1093/nar/gkaa977

Bork, P., and Koonin, E. V. (1994). A P-loop-like motif in a widespread ATP pyrophosphatase domain: implications for the evolution of sequence motifs and enzyme activity. *Proteins* 20, 347–355. doi: 10.1002/prot.340200407

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., et al. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* 321, 960–964. doi: 10.1126/science.1159689

Carroll, K. S., Gao, H., Chen, H., Stout, C. D., Leary, J. A., and Bertozzi, C. R. (2005). A Conserved Mechanism for Sulfonucleotide Reduction. *PLoS Biol.* 3:e250. doi: 10.1371/journal.pbio.0030250

Carte, J., Wang, R., Li, H., Terns, R. M., and Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Develop.* 22, 3489–3496. doi: 10.1101/gad.1742908

Chartron, J., Shiau, C., Stout, C. D., and Carroll, K. S. (2007). 3′-Phosphoadenosine-5′-phosphosulfate reductase in complex with thioredoxin: a structural snapshot in the catalytic cycle. *Biochemistry* 46, 3942–3951. doi: 10.1021/bi700130e

Cheng, K., and Wigley, D. B. (2018). DNA translocation mechanism of an XPD family helicase. *ELife* 7:e42400. doi: 10.7554/elife.42400.001

Crooks, G. E. (2004). WebLogo: a Sequence Logo Generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Crowley, V. M., Catching, A., Taylor, H. N., Borges, A. L., Metcalf, J., Bondy-Denomy, J., et al. (2019). A Type IV-A CRISPR-Cas System in *Pseudomonas aeruginosa* Mediates RNA-Guided Plasmid Interference In Vivo. *CRISPR J.* 2, 434–440. doi: 10.1089/crispr.2019.0048

Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3:945. doi: 10.1038/ncomms1937

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607. doi: 10.1038/nature09886

Dwarakanath, S., Brenzinger, S., Gleditzsch, D., Plagens, A., Klingl, A., Thormann, K., et al. (2015). Interference activity of a minimal Type I CRISPR–Cas system from Shewanella putrefaciens. *Nucl. Acids Res.* 43, 8913–8923. doi: 10.1093/nar/gkv882

Elmore, J., Deighan, T., Westpheling, J., Terns, R. M., and Terns, M. P. (2015). DNA targeting by the type I-G and type I-A CRISPR-Cas systems of Pyrococcus furiosus. *Nucl. Acids Res.* 43, 10353–10363. doi: 10.1093/nar/gkv1140

Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., et al. (2016). Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR–Cas system. *Genes Develop.* 30, 447–459. doi: 10.1101/gad.272153.115

Estrella, M. A., Kuo, F.-T., and Bailey, S. (2016). RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes Develop.* 30, 460–470. doi: 10.1101/gad.273722.115

Fairman-Williams, M. E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr. Opinion Struct. Biol.* 20, 313–324. doi: 10.1016/j.sbi.2010.03.011

Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019). CRISPR–Cas in mobile genetic elements: counter-defence and beyond. *Nat. Rev. Microbiol.* 17, 513–525. doi: 10.1038/s41579-019-0204-7

Garneau, J. E., Dupuis, M. -È, Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523

Grodick, M. A., Segal, H. M., Zwang, T. J., and Barton, J. K. (2014). DNA-Mediated Signaling by Proteins with 4Fe–4S Clusters Is Necessary for Genomic Integrity. *J. Am. Chem. Soc* 136, 6470–6478. doi: 10.1021/ja501973c

Hale, C., Kleppe, K., Terns, R. M., and Terns, M. P. (2008). Prokaryotic silencing (psi)RNAs in Pyrococcus furiosus. *RNA* 14, 2572–2579. doi: 10.1261/rna.1246808

Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., et al. (2009). RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. *Cell* 139, 945–956. doi: 10.1016/j.cell.2009.07.040

Han, W., Li, Y., Deng, L., Feng, M., Peng, W., Hallström, S., et al. (2017). A type III-B CRISPR-Cas effector complex mediating massive target DNA destruction. *Nucl. Acids Res.* 45, 1983–1993. doi: 10.1093/nar/gkw1274

Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010). Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science* 329, 1355–1358. doi: 10.1126/science.1192272

Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., et al. (2015). Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature* 519, 199–202. doi: 10.1038/nature14245

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The Biology of CRISPR-Cas: backward and Forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Jackson, R. N., Golden, S. M., Erp, P. B. G., van, Carter, J., Westra, E. R., et al. (2014). Crystal structure of the CRISPR RNA–guided surveillance complex from *Escherichia coli. Science* 345, 1473–1479. doi: 10.1126/science.1256328

Jackson, R. N., van Erp, P. B., Sternberg, S. H., and Wiedenheft, B. (2017). Conformational regulation of CRISPR-associated nucleases. *Curr. Opinion Microbiol.* 37, 110–119. doi: 10.1016/j.mib.2017.05.010

Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., and Brouns, S. J. J. (2017). CRISPR-Cas: adapting to change. *Science* 356:eaal5056. doi: 10.1126/science.aal5056

Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mole. Biol.* 18, 529–536. doi: 10.1038/nsmb.2019

Jung, T.-Y., An, Y., Park, K.-H., Lee, M.-H., Oh, B.-H., and Woo, E. (2015). Crystal Structure of the Csm1 Subunit of the Csm Complex and Its Single-Stranded DNA-Specific Nuclease Activity. *Structure* 23, 782–790. doi: 10.1016/j.str.2015.01.021

Kamruzzaman, M., and Iredell, J. R. (2020). CRISPR-Cas System in Antibiotic Resistance Plasmids in *Klebsiella pneumoniae. Front. Microbiol.* 10:2934. doi: 10.3389/fmicb.2019.02934

Kazlauskiene, M., Kostiuk, G., Venclovas, Č, Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* 357, 605–609. doi: 10.1126/science.aao0100

Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č, and Siksnys, V. (2016). Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: coupling DNA Degradation with the Target RNA Recognition. *Mol. Cell* 62, 295–306. doi: 10.1016/j.molcel.2016.03.024

Kieper, S. N., Almendros, C., Behler, J., McKenzie, R. E., Nobrega, F. L., Haagsma, A. C., et al. (2018). Cas4 Facilitates PAM-Compatible Spacer Selection during CRISPR Adaptation. *Cell Rep.* 22, 3377–3384. doi: 10.1016/j.celrep.2018.02.103

Koonin, E. V., and Makarova, K. S. (2017). Mobile Genetic Elements and Evolution of CRISPR-Cas Systems: all the Way There and Back. *Genome Biol. Evol.* 9, 2812–2825. doi: 10.1093/gbe/evx192

Koonin, E. V., and Makarova, K. S. (2019). Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc B. Lond. Biol. Sci.* 374:20180087. doi: 10.1098/rstb.2018.0087

Lee, H., Zhou, Y., Taylor, D. W., and Sashital, D. G. (2018). Cas4-Dependent Prespacer Processing Ensures High-Fidelity Programming of CRISPR Arrays. *Mol. Cell* 70, 48–59.e5. doi: 10.1016/j.molcel.2018.03.003

Lewis, L. K., Jenkins, M. E., and Mount, D. W. (1992). Isolation of DNA damage-inducible promoters in *Escherichia coli*: regulation of polB (dinA), dinG, and dinH by LexA repressor. *J. Bacteriol.* 174, 3377–3385.

Lintner, N. G., Kerou, M., Brumfield, S. K., Graham, S., Liu, H., Naismith, J. H., et al. (2011). Structural and Functional Characterization of an Archaeal Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated Complex for Antiviral Defense (CASCADE). *J. Biol. Chem.* 286, 21643–21656. doi: 10.1074/jbc.M111.238485

Liu, T. Y., Iavarone, A. T., and Doudna, J. A. (2017). RNA and DNA Targeting by a Reconstituted Thermus thermophilus Type III-A CRISPR-Cas System. *PLoS One* 12:e0170552. doi: 10.1371/journal.pone.0170552

Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V. (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* 6:38. doi: 10.1186/1745-6150-6-38

Makarova, K. S., Wolf, Y. I, Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. doi: 10.1038/s41579-019-0299-x

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569

Marraffini, L. A., and Sontheimer, E. J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science* 322, 1843–1845. doi: 10.1126/science.1165771

Marraffini, L. A., and Sontheimer, E. J. (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463, 568–571. doi: 10.1038/nature08703

McRobbie, A.-M., Meyer, B., Rouillon, C., Petrovic-Stojanovska, B., Liu, H., and White, M. F. (2012). *Staphylococcus aureus* DinG, a helicase that has evolved into a nuclease. *Biochem. J.* 442, 77–84. doi: 10.1042/BJ20111903

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740. doi: 10.1099/mic.0.023960-0

Mulepati, S., and Bailey, S. (2011). Structural and Biochemical Analysis of Nuclease Domain of Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated Protein 3 (Cas3). *J. Biol Chem.* 286, 31896–31903. doi: 10.1074/jbc.M111.270017

Mulepati, S., Héroux, A., and Bailey, S. (2014). Crystal structure of a CRISPR RNA–guided surveillance complex bound to a ssDNA target. *Science* 345, 1479–1484. doi: 10.1126/science.1256996

Newire, E., Aydin, A., Juma, S., Enne, V. I., and Roberts, A. P. (2020). Identification of a Type IV-A CRISPR-Cas System Located Exclusively on IncHI1B/IncFIB Plasmids in *Enterobacteriaceae*. *Front. Microbiol.* 11:1937. doi: 10.3389/fmicb.2020.01937

Niewoehner, O., Garcia-Doval, C., Rostøl, J. T., Berk, C., Schwede, F., Bigler, L., et al. (2017). Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. *Nature* 548, 543–548. doi: 10.1038/nature23467

Nobrega, F. L., Walinga, H., Dutilh, B. E., and Brouns, S. J. J. (2020). Prophages are associated with extensive CRISPR–Cas auto-immunity. *Nucl. Acids Res.* 48, 12074–12084. doi: 10.1093/nar/gkaa1071

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. doi: 10.1016/j.molcel.2016.04.027

Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528–534. doi: 10.1038/nsmb.2820

Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015). Crystal Structure of the CRISPR-Cas RNA Silencing Cmr Complex Bound to a Target Analog. *Mol. Cell* 58, 418–430. doi: 10.1016/j.molcel.2015.03.018

Özcan, A., Pausch, P., Linden, A., Wulf, A., Schühle, K., Heider, J., et al. (2018). Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. *Nat. Microbiol.* doi: 10.1038/s41564-018-0274-8

Peissert, S., Sauer, F., Grabarczyk, D. B., Braun, C., Sander, G., Poterszman, A., et al. (2020). In TFIIH the Arch domain of XPD is mechanistically essential for transcription and DNA repair. *Nat. Commun.* 11:1667. doi: 10.1038/s41467-020-15241-9

Pinilla-Redondo, R., Mayo-Muñoz, D., Russel, J., Garrett, R. A., Randau, L., Sørensen, S. J., et al. (2019). Type IV CRISPR–Cas systems are highly diverse and involved in competition between plasmids. *Nucl. Acids Res.* 48, 2000–2012. doi: 10.1093/nar/gkz1197

Poranen, M. M., and Tuma, R. (2004). Self-assembly of double-stranded RNA bacteriophages. *Virus Res.* 101, 93–100. doi: 10.1016/j.virusres.2003.12.009

Ren, B., Duan, X., and Ding, H. (2009). Redox control of the DNA damage-inducible protein DinG helicase activity via its iron-sulfur cluster. *J. Biol. Chem.* 284, 4829–4835. doi: 10.1074/jbc.M807943200

Rollie, C., Schneider, S., Brinkmann, A. S., Bolt, E. L., and White, M. F. (2015). Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *ELife* 4:e08716. doi: 10.7554/eLife.08716

Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* 161, 1164–1174. doi: 10.1016/j.cell.2015.04.027

Savage, H., Montoya, G., Svensson, C., Schwenn, J. D., and Sinning, I. (1997). Crystal structure of phosphoadenylyl sulphate (PAPS) reductase: a new family of adenine nucleotide α hydrolases. *Structure* 5, 895–906. doi: 10.1016/S0969-2126(97)00244-X

Shmakov, S. A., Makarova, K. S., Wolf, Y. I., Severinov, K. V., and Koonin, E. V. (2018). Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc. Natl Acad. Sci.* 115, E5307–E5316. doi: 10.1073/pnas.1803440115

Shmakov, S. A., Sitnik, V., Makarova, K. S., Wolf, Y. I., Severinov, K. V., and Koonin, E. V. (2017). The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. *mBio* 8, 01397-17. doi: 10.1128/mBio.01397-17

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system: cas3 nuclease/helicase. *EMBO J.* 30, 1335–1342. doi: 10.1038/emboj.2011.41

Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., et al. (2013). Structure and Activity of the RNA-Targeting Type III-B CRISPR-Cas Complex of Thermus thermophilus. *Mol. Cell* 52, 135–145. doi: 10.1016/j.molcel.2013.09.013

Staals, R. H. J., Zhu, Y., Taylor, D. W., Kornfeld, J. E., Sharma, K., Barendregt, A., et al. (2014). RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. *Mol. Cell* 56, 518–530. doi: 10.1016/j.molcel.2014.10.005

Sternberg, S. H., Richter, H., Charpentier, E., and Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Mol. Cell* 61, 797–808. doi: 10.1016/j.molcel.2016.01.030

Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, È, Nwokeoji, A. O., Dickman, M. J., et al. (2014). Programmable RNA Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus. *Mol. Cell* 56, 506–517. doi: 10.1016/j.molcel.2014.09.027

Taylor, H. N., Warner, E. E., Armbrust, M. J., Crowley, V. M., Olsen, K. J., and Jackson, R. N. (2019). Structural basis of Type IV CRISPR RNA biogenesis by a Cas6 endoribonuclease. *RNA Biol.* 16, 1438-1447. doi: 10.1080/15476286.2019.1634965

Thakur, R. S., Desingu, A., Basavaraju, S., Subramanya, S., Rao, D. N., and Nagaraju, G. (2014). Mycobacterium tuberculosis DinG Is a Structure-specific Helicase That Unwinds G4 DNA Implications for Targeting G4 DNA as a Novel Therapeutic Approach. *J. Biol. Chem.* 289, 25112–25136. doi: 10.1074/jbc.M114.563569

Voloshin, O. N., and Camerini-Otero, R. D. (2007). The DinG Protein from *Escherichia coli* Is a Structure-specific Helicase. *J. Biol. Chem.* 282, 18437–18447. doi: 10.1074/jbc.M700376200

Voloshin, O. N., Vanevski, F., Khil, P. P., and Camerini-Otero, R. D. (2003). Characterization of the DNA Damage-inducible Helicase DinG from *Escherichia coli*. *J. Biol. Chem.* 278, 28284–28293. doi: 10.1074/jbc.M301188200

Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., et al. (2015). Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. *Cell* 163, 840–853. doi: 10.1016/j.cell.2015.10.008

Wang, L., Jiang, S., Deng, Z., Dedon, P. C., and Chen, S. (2018). DNA phosphorothioate modification—A new multi-functional epigenetic system in bacteria. *FEMS Microbiol. Rev.* 43, 109–122. doi: 10.1093/femsre/fuy036

Wei, Y., Terns, R. M., and Terns, M. P. (2015). Cas9 function and host genome sampling in Type II-A CRISPR–Cas adaptation. *Genes Develop.* 29, 356–361. doi: 10.1101/gad.257550.114

Westra, E. R., Semenova, E., Datsenko, K. A., Jackson, R. N., Wiedenheft, B., Severinov, K., et al. (2013). Type I-E CRISPR-Cas Systems Discriminate Target from Non-Target DNA through Base Pairing-Independent PAM Recognition. *PLoS Gene.* 9:e1003742. doi: 10.1371/journal.pgen. 1003742

Westra, E. R., van Erp, P. B. G., Künne, T., Wong, S. P., Staals, R. H. J., Seegers, C. L. C., et al. (2012). CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3. *Mol. Cell* 46, 595–605. doi: 10.1016/j.molcel.2012.03.018

Wu, Y., and Brosh, R. M. Jr. (2012). DNA helicase and helicase–nuclease enzymes with a conserved iron–sulfur cluster. *Nucl. Acids Res.* 40, 4247–4260. doi: 10. 1093/nar/gks039

Yan, W. X., Hunnewell, P., Alfonse, L. E., Carte, J. M., Keston-Smith, E., Sothiselvam, S., et al. (2019). Functionally diverse type V CRISPR-Cas systems. *Science* 363, 88–91. doi: 10.1126/science.aav7271

Yin, Y., Yang, B., and Entwistle, S. (2019). Bioinformatics Identification of Anti-CRISPR Loci by Using Homology, Guilt-by-Association, and CRISPR Self-Targeting Spacer Approaches. *mSystems* 4, e455–e419. doi: 10.1128/mSystems. 00455-19

Yosef, I., Goren, M. G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucl. Acids Res.* 40, 5569–5576. doi: 10.1093/nar/gks216

You, D., Wang, L., Yao, F., Zhou, X., and Deng, Z. (2007). A Novel DNA Modification by Sulfur: dndA Is a NifS-like Cysteine Desulfurase Capable of Assembling DndC as an Iron-Sulfur Cluster Protein in *Streptomyces lividans* [†]. *Biochemistry* 46, 6126–6133. doi: 10.1021/bi602615k

You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., et al. (2019). Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. *Cell* 176, 239–253. doi: 10.1016/j.cell.2018. 10.052

Zhang, J., Graham, S., Tello, A., Liu, H., and White, M. F. (2016). Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. *Nucl. Acids Res.* 44, 1789–1799. doi: 10.1093/nar/gkw020

Zhou, Y., Bravo, J. P. K., Taylor, H. N., Steens, J. A., Jackson, R. N., Staals, R. H. J., et al. (2021). Structure of a type IV CRISPR-Cas ribonucleoprotein complex. *iScience* 24:102201. doi: 10.1016/j.isci.2021.102201

Zimmermann, L., Stephens, S., Seung-Zin, N., Rau, D., Kubler, J., Lozajic, M., et al. (2018). A Completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243. doi10.1016/j.jmb.2017.12.007