Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Molecular architectures and mechanisms of Class 2 CRISPR-associated nucleases

Carmela Garcia-Doval and Martin Jinek

Prokaryotic Class 2 CRISPR-Cas systems mediate adaptive immunity against invasive genetic elements by means of standalone effector proteins that function as RNA-guided nucleases. The effectors Cas9 and Cas12 generate double-strand breaks in DNA substrates, which has been exploited for genome editing applications. In turn, Cas13 enzymes function as RNA-guided ribonucleases whose non-specific activity is triggered by target RNA binding. In this review, we highlight recent structural investigations of Cas9, Cas12 and Cas13 nucleases that have illuminated many aspects of their molecular mechanisms. In particular, these studies have highlighted the role of guide RNA seed sequences in facilitating target recognition and the importance of conformational transitions in controlling target binding and cleavage.

Address
Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Corresponding author: Jinek, Martin (jinek@bioc.uzh.ch)

Current Opinion in Structural Biology 2017, **47**:157–166

This review comes from a themed issue on **Protein–nucleic acid interactions**

Edited by **Stephen Cusack** and **Christoph W Müller**

For a complete overview see the Issue and the Editorial

Available online 3rd November 2017

http://dx.doi.org/10.1016/j.sbi.2017.10.015

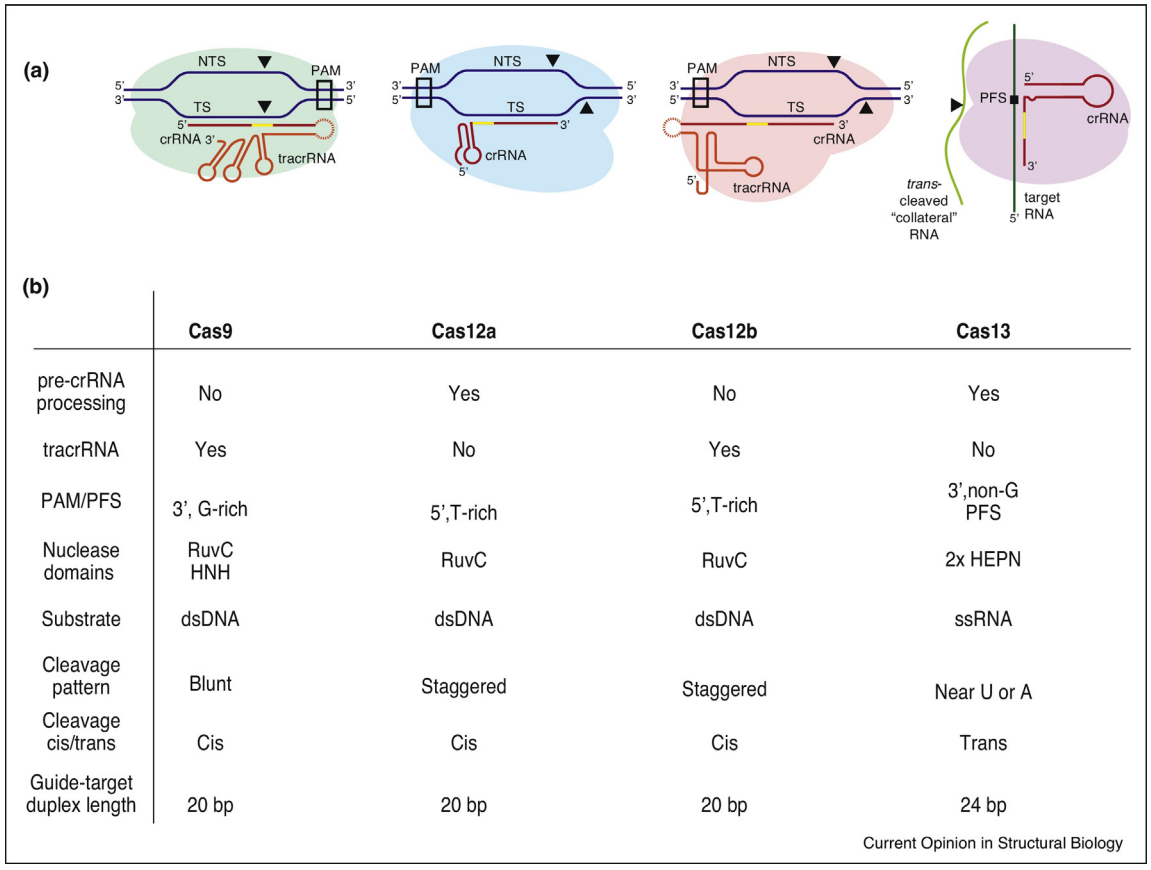
0959-440/© 2017 Elsevier Ltd. All rights reserved.

Introduction

CRISPR-Cas (for Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated) are RNA-guided genome defence systems found in many bacteria and most *Archaea* that provide adaptive immunity against foreign genetic elements such as bacteriophages and plasmids [1–4]. CRISPR-Cas systems target and ultimately degrade foreign nucleic acids by means of CRISPR RNA (crRNA) guides that associate with Cas proteins to form the interference machinery [5]. CRISPR-Cas systems have been classified into two classes (Classes 1 and 2) and six types (types I–VI), with distinct mechanisms of guide RNA biogenesis and target interference [6,7••]. In Class 1 systems, comprising types I, III and IV, interference is mediated by means of a multi-Cas protein complex bound to a mature crRNA guide. In turn, Class 2 systems, comprising type II and the more recently described types V and VI, target nucleic acids by crRNA-guided single-protein nuclease effectors. Until the recent discovery of type VI systems, only type III systems were known to target RNA [8–10], while the remaining CRISPR-Cas types all target double-stranded DNA (dsDNA) [3]. To ensure correct target recognition and avoid self-targeting, the DNA-targeting Class 1 and 2 systems rely on the recognition of a short sequence, termed protospacer adjacent motif (PAM), in the vicinity of the target site complementary to the crRNA.

Cas9, the archetypal Class 2 effector nuclease found in type II CRISPR-Cas systems, associates with a dual-guide RNA structure consisting of a mature crRNA duplexed with a trans-activating CRISPR RNA (tracrRNA) [11,12] (Figure 1). The guide RNA-bound Cas9 complex is capable of binding and cleaving dsDNA substrates complementary to the crRNA in a PAM-dependent manner, generating a double-strand DNA break (DSB) [13,14]. The crRNA and tracrRNA can be artificially fused to generate a chimeric single-chain guide RNA (sgRNA) [13]. Thanks to its programmable sequence specificity, Cas9 has emerged as a powerful and extremely versatile molecular tool for genome editing in numerous organisms and cell types. The utility of Cas9 extends beyond genome editing to RNA-programmable transcriptional and epigenetic control of gene expression relying on catalytically inactive Cas9 (dCas9). The success of CRISPR-Cas9 in these applications, combined with its limitations due to, for example, off-target activity and PAM requirement, has prompted extensive search for programmable RNA-guided nucleases in other CRISPR-Cas system types [15]. This resulted in the recent discoveries of type V and VI CRISPR-Cas systems, both of which rely on single-subunit effectors for RNA-guided interference. Type V systems are defined by the presence of Cas12 proteins (Cas12a/Cpf1 and Cas12b/C2c1 in type V-A and V-B systems, respectively), which function as RNA-guided, DSB-generating deoxyribonucleases [7••,16••,17] (Figure 1). By contrast, Cas13 (also known as C2c2), the effector protein of type VI systems, functions as RNA-guided ribonuclease whose non-specific, trans-acting RNase activity is triggered by base pairing of the crRNA guide to a single-stranded target RNA [18•] (Figure 1). Both Cas12 and Cas13 nucleases offer considerable practical utility. Cas12a enzymes from type V-A systems have emerged as promising alternative genome editing nucleases that complement the capabilities of Cas9. In turn, the non-specific, 'collateral effect'

Figure 1



	Cas9	Cas12a	Cas12b	Cas13
pre-crRNA processing	No	Yes	No	Yes
tracrRNA	Yes	No	Yes	No
PAM/PFS	3', G-rich	5',T-rich	5',T-rich	3',non-G PFS
Nuclease domains	RuvC HNH	RuvC	RuvC	2x HEPN
Substrate	dsDNA	dsDNA	dsDNA	ssRNA
Cleavage pattern	Blunt	Staggered	Staggered	Near U or A
Cleavage cis/trans	Cis	Cis	Cis	Trans
Guide-target duplex length	20 bp	20 bp	20 bp	24 bp

Class 2 CRISPR-associated effector nucleases. **(a)** Schematic diagrams of Cas9, Cas12a, Cas12b and Cas13 proteins. The crRNA is represented in red, the tracrRNA in orange and dashed lines represent the linkers used to create chimeric sgRNAs. Seed sequences are highlighted in yellow. The cleavage sites are indicated with triangles and the position of the PAM or PFS is indicated with a box. **(b)** Functional and mechanistic characteristics of Cas9, Cas12a, Cas12b and Cas13.

RNase activity of Cas13 has been exploited in applications allowing specific nucleic acid detection with attomolar sensitivity [19]. The exploitation and ongoing development of CRISPR-associated nucleases for genome editing and gene/transcript-targeting applications rely on detailed knowledge of their molecular mechanisms. In this review, we aim to highlight the contribution of recent structural studies towards our understanding of the molecular architectures and mechanistic aspects of Class 2 CRISPR effector nucleases from type II, V and VI CRISPR-Cas systems.

Type II: Cas9

The Cas9 nuclease specifically recognizes target dsDNA substrates complementary to a 5′-terminal 20-nucleotide sequence in the guide crRNA [11,12] (Figure 1). This occurs by the formation of an R-loop structure in which one of the DNA strands (the target strand) base-pairs to the guide RNA, while the other (the non-target strand) is displaced. Substrate DNA recognition is strictly dependent on the presence of a cognate PAM downstream of the target site in the DNA and requires complementarity of the target to a 'seed' sequence in the guide RNA [13,20]. Biochemical and single-molecule biophysical studies have established that PAM recognition licenses R-loop formation by facilitating strand separation in the dsDNA and RNA-target DNA hybridization [21,22]. Subsequently, the HNH and RuvC nuclease domains of Cas9 catalyse the cleavage of the target and non-target DNA strands, respectively, in a divalent cation-dependent manner [13]. Over the last four years, extensive structural studies of *Streptococcus pyogenes* Cas9 (SpCas9) have defined its molecular architecture and provided atomic-detail insights into its mechanism of RNA-guided DNA targeting and cleavage. SpCas9 is a bilobed molecule consisting of nuclease (NUC) and alpha-helical recognition (REC) lobes. The NUC lobe contains the HNH and RuvC nuclease domains as well as the C-terminal Wedge (WED) and PAM-interacting (PI) domains [23,24] (Figure 2a).

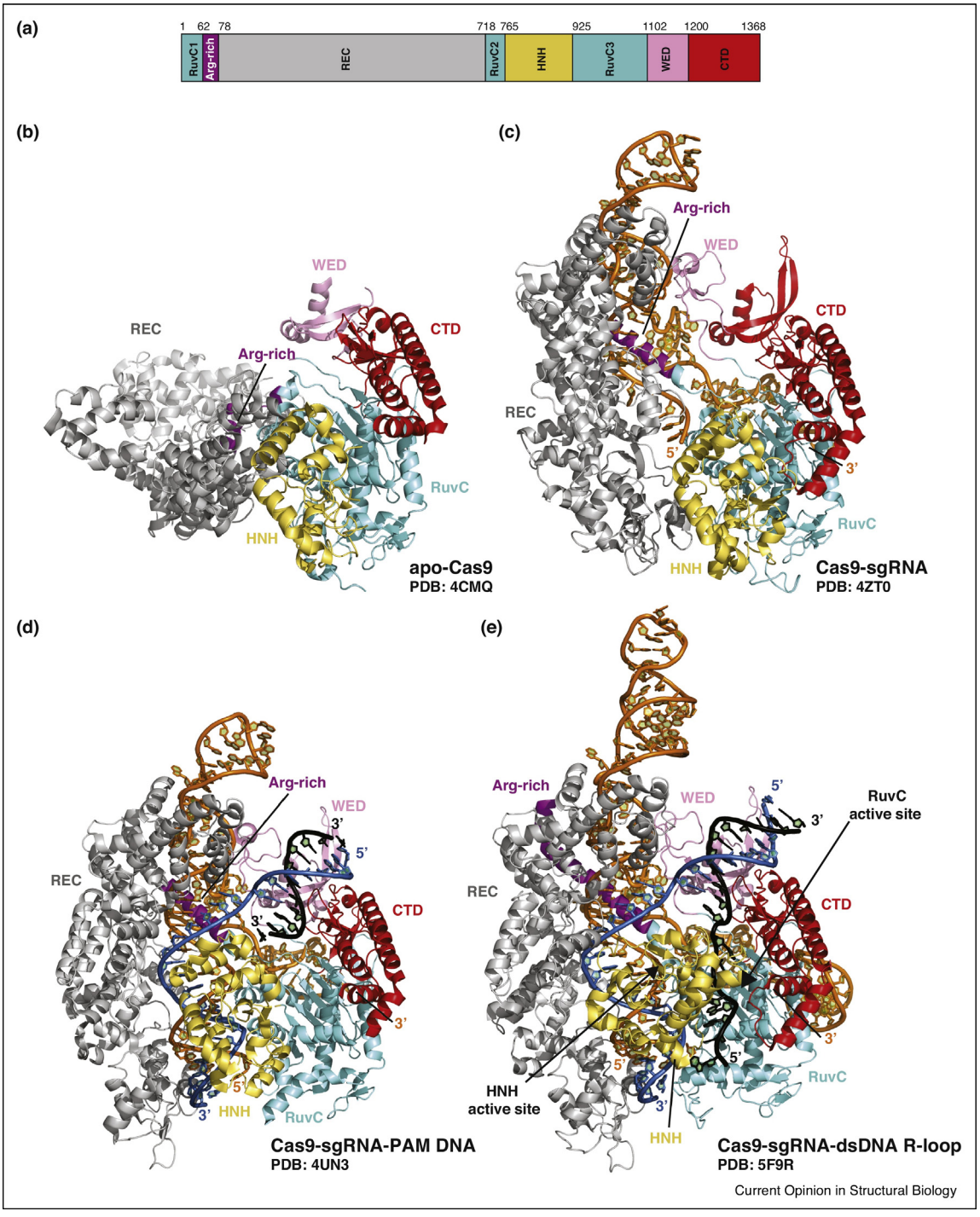
Comparisons of the apo-SpCas9 and DNA-bound complex structures, together with electron microscopic studies, indicate that Cas9 undergoes a major conformational rearrangement upon guide RNA binding, converting from an open, autoinhibited conformation to a closed form in which the NUC and REC lobes surround a central channel that accommodates the guide RNA-target DNA heteroduplex [23] (Figure 2b,c). The rearrangement is driven by extensive interactions with three stem-loop structures in the 3′-terminal part of the sgRNA, mediated in part by the arginine-rich Bridge Helix in Cas9 [25,26]. The 5′-NGG-3′ PAM is specifically recognized by SpCas9 in a cleft formed by the WED and PI domains, the latter of which provides two arginine residues that contact the major groove edges of the two guanine bases in the PAM [26]. PAM recognition positions the target DNA strand in register with the seed sequence of the guide RNA, such that interactions between the phosphate group immediately upstream of the PAM in the target DNA strand and the so-called 'phosphate-lock' loop in the Cas9 WED domain induce a kink in the target DNA strand, thereby facilitating strand separation in the substrate dsDNA [26]. In the binary SpCas9–sgRNA complex, interactions with the Bridge Helix induce the seed sequence of the sgRNA guide to adopt a pre-ordered, A-form conformation, thereby promoting target DNA binding by reducing the entropic penalty of RNA–DNA heteroduplex formation [27••]. The structure of SpCas9 in complex with a near-complete R-loop in turn reveals that interactions of the non-target DNA strand with the RuvC domain depend on a conformational rearrangement that positions the HNH domain for cleavage of the target DNA strand. This conformational coupling between the two nuclease domains is mediated by two hinge segments (L1 and L2) connecting the HNH domain to the RuvC [28••] (Figure 2d,e). These structural observations are in good agreement with FRET studies showing that the HNH domain is dynamic and samples a conformational equilibrium between inactive and active conformations [29]. Finally, cryoelectron microscopic reconstructions of SpCas9 in complex with a full R-loop reveal that the DNA is bent by ~30°, which is thought to facilitate strand separation in the target DNA [28••].

Structural studies of orthologous Cas9 proteins from *Staphylococcus aureus*, *Francisella novicida* and *Campylobacter jejuni* have revealed considerable structural diversity within the Cas9 enzyme family attuned to the specific recognition of distinct PAM sequences and crRNA-tracrRNA guide scaffolds [24,30,31]. Nevertheless, the basic domain architecture and features such as the phosphate-lock loop have been observed in all Cas9 structures determined to date, pointing to a conserved mechanism underpinning DNA binding and cleavage. Altogether, structural studies of Cas9 nucleases have highlighted the importance of conformational rearrangements along the reaction pathway. These studies have thus not only provided a mechanistic framework for understanding the functions of Cas9 but also for further development of Cas9-based genome editing. One of the constraints of Cas9 function is the requirement for a PAM in the DNA target. To address this limitation, protein engineering has been used to develop artificial SpCas9 variants (VQR, EQR and VRER) that recognize non-canonical PAM sequences, thereby expanding the spectrum of targetable DNA sequences [32]. Recent structures of these SpCas9 variants bound to their preferred PAMs revealed that PAM recognition occurs by an induced fit mechanism, in which the engineered amino acid substitutions not only mediate sequence-specific interactions with the PAM but also facilitate conformational rearrangements in the bound DNA to ensure optimal fit [33,34]. Finally, bacteriophages have recently been shown to encode inhibitory anti-CRISPR proteins that counteract Cas9-mediated immunity [35•,36–38]. Structures of the SpCas9–sgRNA complex bound to the AcrIIA4 protein show that AcrAII4 inhibits SpCas9 activity by acting as a structural mimic of the PAM DNA, thereby blocking the PAM binding site in SpCas9 [39•,40•].

Type V: Cas12a and Cas12b

CRISPR-associated nucleases of the Cas12 family, which includes Cas12a and Cas12b proteins, are characterized by the presence of a conserved RuvC domain but lack an HNH nuclease domain. Cas12a, previously known as Cpf1 (CRISPR from *Prevotella* and *Francisella* 1), is the defining component of type V-A CRISPR-Cas systems. Unlike Cas9, Cas12a is guided solely by a processed crRNA and does not require a tracrRNA for DNA targeting [16••] (Figure 1). crRNA processing is mediated by an endoribonuclease activity contained within Cas12a itself and occurs by a divalent cation-independent mechanism [17,41••]. Whereas Cas9 proteins often recognize G-rich PAMs and generate DSBs with blunt ends or 1-nucleotide overhangs, Cas12a enzymes recognize a T-rich PAM upstream of the target region and cleave DNA at a PAM-distal site to generate a staggered double stranded break with 5-nucleotide 5′-overhangs [16••]. Owing to these complementary features, several Cas12a enzymes have been repurposed for genome editing in mammalian cells and plants, thus expanding the molecular genome editing toolkit [42]. Cas12b, previously known as C2c1 (Class 2 candidate 1), is found in type V-B systems. Similar to Cas12a, Cas12b recognizes a T-rich PAM upstream of the target region in the dsDNA substrate. Unlike Cas12a, however, the enzyme is instead guided by a dual RNA structure consisting of a crRNA and a tracrRNA and generates DSBs with 7-nucleotide 5′ overhangs [43••,44]. As for Cas9, the Cas12b can be programmed using chimeric sgRNAs to cleave DNA *in vitro* [43••,44] but Cas12b enzymes have not been utilized for genome editing applications to date.

Figure 2



Domain architecture and structures of Cas9 and its complexes. **(a)** Domain structure of *Streptococcus pyogenes* Cas9 (SpCas9). **(b)** Crystal structure of apo-SpCas9 (PDB: 4CMQ). In panels (b–e), SpCas9 domains are colored according to (a); the sgRNA is represented in orange; the DNA target strand in blue and the DNA non-target strand in black. **(c)** Crystal structure of the binary SpCas9–sgRNA complex (PDB: 4ZT0). **(d)** Crystal structure of SpCas9 in complex with an sgRNA and a partially duplexed, PAM-containing DNA target (PDB: 4UN3). **(e)** Crystal structure of SpCas9 bound to an sgRNA and a duplex DNA substrate that forms an R-loop upon base pairing with the sgRNA (PDB: 5F9R). The HNH and RuvC nuclease domain active sites are marked with arrows.

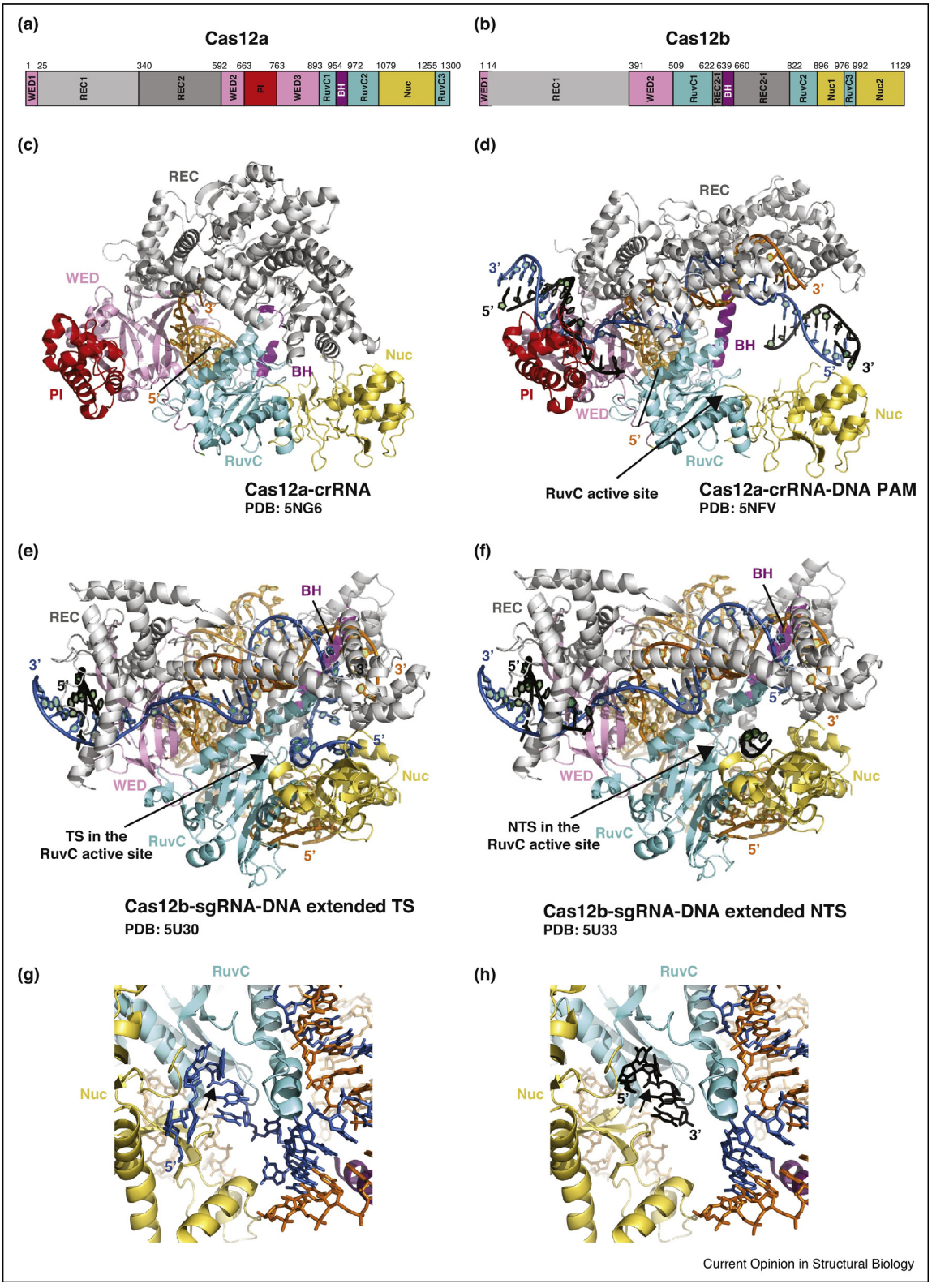
Initial crystallographic studies of Cas12a from *Acidaminococcus sp.* (AsCas12a) in a complex with a crRNA and a partially-duplexed target DNA revealed an overall bilobed architecture consisting of REC and NUC lobes that surround a central positively charged channel that accommodates the crRNA-target DNA heteroduplex [45] (Figure 3c,d). The N-terminal REC lobe is divided in two alpha-helical recognition (Rec1 and Rec2) domains, whereas the C-terminal NUC lobe consists of the WED, PI, RuvC and Nuc domains, with both lobes connected *via* the WED domain [45,46••,47] (Figure 3a). The Nuc domain lacks structural homology to other protein folds. Despite analogous nomenclature, there is limited structural similarity between the Rec, WED and PI domains of Cas12a and Cas9 [46••,47]. The 5′ handle of the crRNA adopts a pseudoknot structure formed by non-canonical base pairing interactions with the loop of a five base-pair stem-loop. crRNA binding to Cas12a is further stabilized by two hydrated Mg^{2+} ions, which substantially contribute to the overall affinity for the crRNA [41••,45]. A recent structure of *Francisella novicida* Cas12a (FnCas12a) in a binary complex with a crRNA shows that five nucleotides immediately downstream of the pseudoknot in the crRNA adopt a preorganized, A-form helical conformation, with the bases exposed to the solvent [41••] (Figure 3c). This suggests that DNA recognition by Cas12a nucleases relies on a seed sequence in the crRNA. Although seed sequence preordering was not observed in the *Lachnospiraceae bacterium* Cas12a (LbCas12a) complex [45], this interpretation is in agreement with biochemical studies showing that mismatches between the crRNA seed sequence and the target DNA inhibit DNA cleavage by Cas12a nucleases [16••,17,41••]. Overall, these insights thus indicate that crRNA binding primes Cas12a for target DNA binding by a mechanism similar to that of Cas9 and other nucleic acid-guided nucleases such as Argonaute proteins [45].

Cas12a undergoes a conformational transition upon substrate DNA binding that involves a rearrangement of the REC lobe that results in the opening of a central channel that accommodates the crRNA–target DNA heteroduplex [46••,47,41••,49•]. Unlike in Cas9, the cognate 5′-(T)TTV-3′ PAM is recognized by both sequence-specific contacts and shape-specific readout of the narrow minor groove of the DNA duplex, which together ensure steric exclusion of non-cognate base-pairs [46••,47]. LbCas12a and AsCas12a can also recognize suboptimal C-containing PAMs due to conformational plasticity in the PAM binding site [16••,50]. Artificial variants of Cas12a have additionally been engineered to cleave DNA substrates containing non-canonical PAMs and their structural studies have revealed specific interactions underpinning PAM recognition [50,51].

Although the guide segment of Cas12a-associated crRNAs is 24 nucleotides long, the crRNA–DNA hybrid comprises only 20 base-pairs and is capped at the PAM-distal end by stacking interaction with a conserved aromatic residue [46••]. Crystal structures of FnCas12a in complex with a complete R-loop, which likely represents a pre-cleavage structure, and in a cleavage product-bound complex show that Cas12a makes extensive interactions with the displaced non-target DNA strand downstream of the PAM, thereby facilitating strand separation and crRNA-target DNA hybridization [41••,49•] (Figure 3d). Upon R-loop formation, the conformational rearrangement of the REC lobe also exposes the DNA nuclease active site located in a cleft spanning the interface of the RuvC and Nuc domains, thereby activating the enzyme for DNA cleavage [49•]. In AsCas12a and LbCas12a, mutations of conserved residues in the RuvC domain interfere with cleavage of both substrate DNA strands, while mutation of a conserved arginine residue in the Nuc domain impacts mostly target DNA strand cleavage. This was originally interpreted to suggest that the RuvC domain initially cleaves the non-target DNA strand, which is a prerequisite for subsequent cleavage of the target DNA strand by the Nuc domain. However, a more extensive mutational analysis of FnCas12a suggests that both DNA strands are cleaved by the RuvC active site [41••]. These conclusions are further supported by recent structural studies of Cas12b, as detailed below.

Similar to Cas12a, Cas12b exhibits a bilobed molecular architecture but it lacks a PI domain and its Nuc domain has a distinct fold [43••,44] (Figure 3b). The sgRNA scaffold forms an intricate pseudoknot structure composed of five duplex segments. Structures of the binary complex of *Alicyclobacillus acidoterrestris* Cas12b (AacCas12b) with an sgRNA guide reveal preordering of a putative seed sequence in the guide RNA, although AacCas12b appears to be sensitive to mismatches along most of the guide RNA–target DNA heteroduplex. Insights into the mechanism of DNA cleavage have come from crystal structures of AacCas12b in ternary complexes with an sgRNA guide and a set of DNA substrates [43••] (Figure 3e–h). In these complexes, the guide RNA–target DNA heteroduplex comprises 20 base-pairs, while either the target or the displaced non-target DNA strand contain single-stranded 5′ or 3′ extensions, respectively. The structures of the complexes reveal that both the target and non-target strand extensions can independently occupy a single catalytic pocket located at the interface of the RuvC and Nuc domains. In the complex containing the extended target DNA strand, the 5′-terminal extension forms a sharp kink as it inserts into the catalytic pocket. A conserved arginine residue in the Nuc domain contributes to the coordination of the scissile phosphate in the RuvC active site, in agreement with mutational studies implicating the Nuc domain in the cleavage of the target DNA strand in Cas12a [43••] (Figure 3e,g). The positioning of the target and non-target DNA strands in the RuvC active site is consistent with the locations of the

Figure 3



Molecular architectures and nucleic acid complexes of Cas12a and Cas12b. **(a)** Domain structure of *Francisella novicida* Cas12a (FnCas12a). **(b)** Domain structure of *Alicyclobacillus acidoterrestris* Cas12b (AacCas12b). **(c)** Crystal structure of FnCas12a in binary complex with a crRNA guide (PDB: 5NG6). In panels (c, d), Cas12a domains are colored according to (a); the crRNA is depicted in orange; the DNA target strand in blue and

cleavage sites in the substrate DNA, which together yield a staggered double-strand DNA break. Thus, the DNA-bound structures of Cas12b provide compelling evidence that Cas12 nucleases cleave both strands in dsDNA substrates using a single active site. As the two DNA strands can only bind in the RuvC domain active site in a mutually exclusive manner, this implies that the two strands are cleaved sequentially. Whether these cleavage events occur in a defined order remains to be established by future biochemical studies.

Type VI systems: Cas13

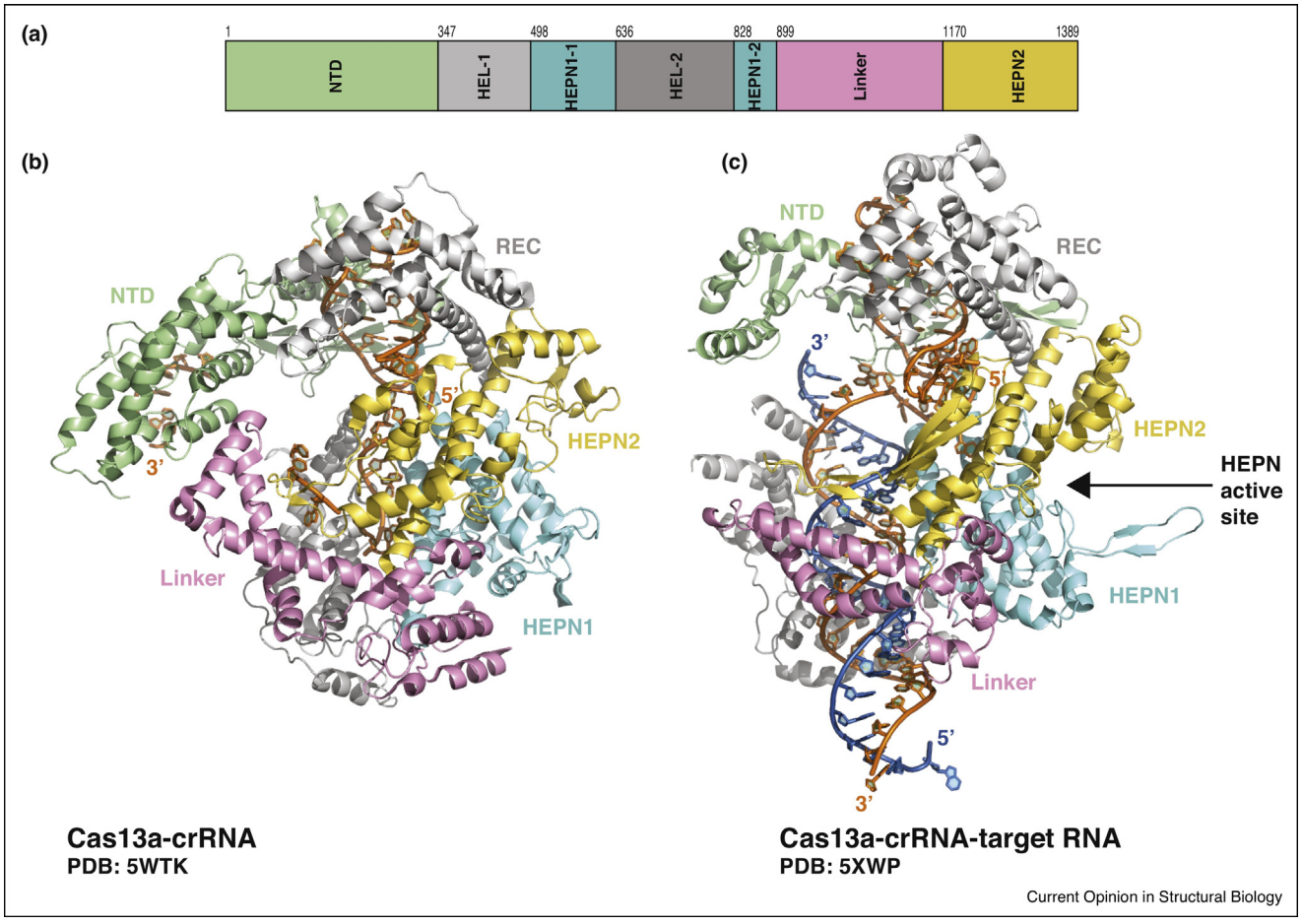
By contrast to type II and type V CRISPR-associated nucleases, which target and cleave dsDNA, Cas13 effectors (Cas13a–c from type VI-A, VI-B and VI-C systems, respectively) function as RNA-guided ribonucleases [52] (Figure 1). Through base pairing to a crRNA guide, Cas13 recognizes single-stranded RNA target sequences, which activates an RNase activity catalyzed by a pair of HEPN (Higher Eukaryote and Prokaryote Nucleotide-binding) ribonuclease domains within the enzyme (Figures 1 and 4a). Target RNA recognition is dependent on the presence of a non-G nucleotide at the 3′ protospacer flanking site (PFS). Upon activation, Cas13 is capable of degrading not only the target RNA but also 'collateral' (*i.e.* non-target) RNAs, suggesting that unlike other Class 2 CRISPR-associated nucleases, which cleave bound target nucleic acids in *cis*, Cas13 degrades its RNA substrates in *trans* [18•]. Cas13a enzymes comprise two clades that initiate degradation by preferential endonucleolytic cleavage at adenosine or uridine residues, respectively [18•,53••,54]. In addition to its RNA-guided effector RNase activity, Cas13 enzymes also possess a distinct RNase activity responsible for endonucleolytic processing of crRNA precursors in type VI systems [53••].

Crystal structures of *Leptotrichia shahii* Cas13a in the apo-form and in complex with a crRNA reveal a bilobed structure composed of a REC lobe and a NUC lobe (Figure 4b). Despite the analogous nomenclature, there is no structural homology with Cas9 or Cas12 proteins [55••]. The recognition (REC) lobe contains an N-terminal domain (NTD) and a Helical-1 domain. The surface of the Helical-1 domain is positively charged and together with the NTD domain forms the crRNA binding channel. The NUC lobe consists of two HEPN domains connected by a Linker domain, and a Helical-2 domain that contacts the 3′ end of the crRNA [55••]. Upon crRNA binding, the Helical-2 domain rotates towards the Linker and HEPN2 domains, closing the crRNA channel (Figure 4a) [55••]. Although the 5′-terminal stem-loop and the 3′ end of the crRNA are specifically recognized by LshCas13a and buried in the complex, the central part of the guide segment of the crRNA is unstructured and exposed to bulk solvent to prime hybridization to the target RNA. This is consistent with biochemical experiments demonstrating that the central segment of the guide RNA acts as a seed sequence to nucleate base pairing with target RNAs. Importantly, the LshCas13a structures reveal that Cas13 enzymes harbor two independent RNase active sites. The two HEPN domains contain conserved Arg-Asn-X_3-His motifs essential for enzymatic activity and form an intramolecular dimer in which the catalytic motifs form a single composite RNase active site responsible for crRNA-dependent RNA cleavage [54]. In turn, a conserved pocket in the Helical-1 domain catalyzes crRNA processing.

The structural basis of crRNA-guided targeting and interference by Cas13 has recently been revealed by a crystal structure of *Leptotrichia buccalis* Cas13a (LbuCas13a) bound to a crRNA and a target RNA, and a cryo-EM structure of LbuCas13a in complex with a crRNA [56••]. The binary complex structure reveals that the central segment of the crRNA (nucleotides 9–15) is pre-ordered in a helical-like conformation, further supporting the idea that it functions as a seed sequence to facilitate target RNA recognition [56••]. The structure of the target-bound complex shows that crRNA-bound Cas13a forms a 24 base-pair duplex with the target RNA and undergoes a substantial conformational rearrangement upon target RNA binding (Figure 4c). As the crRNA–target RNA duplex binds in a central channel within the NUC lobe, the Helical-2 domain rotates away from the HEPN2 domain, while the HEPN1 domain moves closer towards HEPN2. Consequently, the catalytic motifs of the two HEPN domains are brought into close proximity, resulting in catalytic activation [56••]. Since the composite HEPN domain active site is exposed on the external surface of the complex and located a considerable distance away from the crRNA–target RNA duplex, the structure provides compelling evidence that Cas13a catalyzes RNA degradation in *trans*, in agreement with biochemical studies demonstrating non-specific 'collateral' RNase activity of Cas13 enzymes. The target RNA thus acts as an allosteric effector rather than a

(Figure 3 Legend Continued) the DNA non-target strand in black. **(d)** Crystal structure of FnCas12a bound to a pre-crRNA mimic and a fully duplexed DNA, forming an R-loop structure upon base pairing with the crRNA (PDB: 5NFV). **(e)** Crystal structure of AacCas12b bound to an sgRNA and substrate DNA containing an extended target strand (PDB: 5U30). In panels (e–h) Cas12b domains are colored according to (b); the sgRNA is represented in orange; the DNA target strand in blue and the DNA non-target strand in black. **(f)** Crystal structure of the AacCas12b–sgRNA–DNA complex with an extended non-target DNA strand (PDB: 5U33). **(g)** Close-up view of the kinked target strand in the AacCas12b active site in the AacCas12b complex containing the extended target DNA strand. The scissile phosphate is marked with an arrow. **(h)** Close-up view of the non-target strand in the AacCas12b active site in the complex containing the extended non-target strand. The scissile phosphate is marked with an arrow.

Figure 4



Architecture and structures of Cas13a and its nucleic acid complexes. **(a)** Domain structure of *Leptotrichia shahii* Cas13a (LshCas13a). **(b)** Crystal structure of a binary complex of LshCas13a and a crRNA (PDB: 5WTK). In panels (b and c), Cas13a domains are colored according to (a); the crRNA is represented in orange and the target RNA strand in blue. **(c)** Crystal structure of *Leptotrichia buccalis* Cas13a (LbuCas13a) in complex with a crRNA and a target RNA (PDB: 5XWP).

substrate, triggering conformational activation of the HEPN domains to degrade other free RNAs in solution (including other molecules of the target RNA itself) in a non-specific manner [56••]. The structural insights thus lend further support to the hypothesis that type VI CRISPR-Cas systems function, at least in part, by inducing dormancy or programmed cell death [18•].

Conclusions

Class 2 CRISPR-associated nucleases have found widespread use as versatile tools for the detection and manipulation of nucleic acids in gene targeting and genome editing applications. Over recent years, extensive structural studies of these nucleases and their complexes have provided fundamental insights into their molecular mechanisms and advanced our understanding of their biological functions. In spite of their diverse evolutionary origins and distinct functionalities, Cas9, Cas12 and Cas13 enzymes share common mechanistic aspects. In all three nuclease families, target recognition is facilitated by structural pre-organization of a seed sequence in the bound crRNA guide, in common with other nucleic acid-guided effectors such as Argonaute proteins or the Cascade complex. In the absence of target binding, the crRNA-bound nucleases are autoinhibited by adopting catalytically inactive conformations (in Cas9 and Cas13) or steric occlusion of nuclease active sites (Cas12a/b). Their respective mechanisms of crRNA-guided target binding proceed through a coordinated series of conformational rearrangements and checkpoints, such as PAM recognition, that trigger catalytic activation only in response to a cognate target nucleic acid, thus minimizing off-targeting activity. Ongoing structural, biochemical and biophysical studies of Class 2 CRISPR-associated nucleases will continue to inform our understanding of their mechanisms and play a critical role in their further technological development.

Conflict of interest statement

MJ is a co-founder of Caribou Biosciences, Inc. and inventor on patents and patent applications related to CRISPR-Cas technologies.

Acknowledgements

We thank members of the Jinek group for discussions and critical reading of the manuscript. Research in the Jinek laboratory is supported a Swiss National Science Foundation project grant to MJ (SNSF 31003A_149393), by a European Research Council (ERC) Starting Grant ANTIVIRNA (ERC-StG-337284), and by funding from the Swiss National Competence Center for Research (NCCR) 'RNA & Disease'. CGD was supported by a Long-term Fellowship from the European Molecular Biology Organization (EMBO). MJ is an EMBO Young Investigator, Vallee Scholar of the Bert L & N Kuggie Vallee Foundation, and International Research Scholar of the Howard Hughes Medical Institute.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Marraffini LA: **CRISPR-Cas immunity in prokaryotes**. *Nature* 2015, **526**:55-61.

2. Sorek R, Lawrence CM, Wiedenheft B: **CRISPR-mediated adaptive immune systems in bacteria and archaea**. *Annu Rev Biochem* 2013, **82**:237-266.

3. van der Oost J, Westra ER, Jackson RN, Wiedenheft B: **Unravelling the structural and mechanistic basis of CRISPR-Cas systems**. *Nat Rev Microbiol* 2014, **12**:479-492.

4. Wiedenheft B, Sternberg SH, Doudna JA: **RNA-guided genetic silencing systems in bacteria and archaea**. *Nature* 2012, **482**:331-338.

5. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C: **Short motif sequences determine the targets of the prokaryotic CRISPR defence system**. *Microbiology* 2009, **155**:733-740.

6. 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.

7. •• 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.
This paper describes the discovery and characterization of Cas12b and Cas13a effectors.

8. Samai P, Pyenson N, Jiang W, Goldberg GW, Hatoum-Aslan A, Marraffini LA: **Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity**. *Cell* 2015, **161**:1164-1174.

9. 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**. *Genes Dev* 2016, **30**:447-459.

10. Estrella MA, Kuo FT, Bailey S: **RNA-activated DNA cleavage by the Type III-B CRISPR Cas effector complex**. *Genes Dev* 2016, **30**:460-470.

11. Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S: **The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA**. *Nature* 2010, **468**:67-71.

12. 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.

13. 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.

14. Gasiunas G, Barrangou R, Horvath P, Siksnys V: **Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria**. *Proc Natl Acad Sci* 2012, **109**:E2579-E2586.

15. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna JA: **RNA-programmed genome editing in human cells**. *Elife* 2013, **2013**:1-9.

16. •• 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.
This paper describes Cas12a as a single RNA-guided endonuclease that cleaves dsDNA and can be used for genome editing.

17. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E: **The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA**. *Nature* 2016, **532**:517-521.

18. • 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.
This paper reports the biochemical characterization of Cas13a.

19. 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.

20. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S: **Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus***. *J Bacteriol* 2008, **190**:1390-1400.

21. 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* 2014, **111**:9798-9803.

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.

23. 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** 1247997-1247997.

24. 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.

25. 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.

26. 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.

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 structure of Cas9 bound to a sgRNA and shows the structural preordering of the seed sequence.

28. •• 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 describes the structure of the ternary complex of Cas9–sgRNA–dsDNA and the conformational changes occurring upon R-loop formation.

29. Sternberg SH, LaFrance B, Kaplan M, Doudna JA: **Conformational control of DNA target cleavage by CRISPR-Cas9**. *Nature* 2015, **527**:110-113.

30. 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.

31. 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.e3.

32. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales APW, Li Z, Peterson RT, Yeh J-RJ *et al.*: **Engineered CRISPR-Cas9 nucleases with altered PAM specificities**. *Nature* 2015, **523**:481-485.

33. Anders C, Bargsten K, Jinek M: **Structural plasticity of PAM recognition by engineered variants of the RNA-guided endonuclease Cas9**. *Mol Cell* 2016, **61**:895-902.

34. Hirano S, Nishimasu H, Ishitani R, Nureki O: **Structural basis for the altered PAM specificities of engineered CRISPR-Cas9**. *Mol Cell* 2016, **61**:886-894.

35. • Pawluk A, Amrani N, Zhang Y, Garcia B, Hidalgo-Reyes Y, Lee J, Edraki A, Shah M, Sontheimer EJ, Maxwell KL *et al.*: **Naturally occurring off-switches for CRISPR-Cas9**. *Cell* 2016, **167** 1829–1838.e9.
This paper provides the first description of anti-CRISPR proteins inhibiting Cas9.

36. Rauch BJ, Silvis MR, Hultquist JF, Waters CS, McGregor MJ, Krogan NJ, Bondy-Denomy J, Abudayyeh OO, Gootenberg JS, Konermann S *et al.*: **Inhibition of CRISPR-Cas9 with bacteriophage proteins**. *Cell* 2016, **168** 150–158.e10.

37. Shin J, Jiang F, Liu J-J, Bray NL, Rauch BJ, Baik SH, Nogales E, Bondy-Denomy J, Corn JE, Doudna JA: **Disabling Cas9 by an anti-CRISPR DNA mimic**. *Sci Adv* 2017, **3**:e1701620.

38. Harrington LB, Doxzen KW, Ma E, Liu J-J, Knott GJ, Edraki A, Garcia B, Amrani N, Chen JS, Cofsky JC *et al.*: **A broad-spectrum inhibitor of CRISPR-Cas9**. *Cell* 2017, **170**:1224-1233 http://dx.doi.org/10.1016/j.cell.2017.07.037.

39. • Dong D, Guo M, Wang S, Zhu Y, Wang S, Xiong Z, Yang J, Xu Z, Huang Z: **Structural basis of CRISPR–SpyCas9 inhibition by an anti-CRISPR protein**. *Nature* 2017, **546**:436-439.
This paper describes the structure of a Cas9–sgRNA complex inhibited by the anti-CRISPR protein AcrII4.

40. • Yang H, Patel DJ: **Inhibition mechanism of an anti-CRISPR suppressor AcrIIA4 targeting SpyCas9**. *Mol Cell* 2017, **67** 117–127.e5.
See annotation to Ref. [39•].

41. •• Swarts DC, van der Oost J, Jinek M: **Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a**. *Mol Cell* 2017, **66** 221–233.e4.
This paper reports structures of binary and ternary Cas12a complexes, revealing the mechanism of crRNA processing by Cas12a, and providing insights into R-loop formation and DNA cleavage.

42. Fagerlund RD, Staals RHJ, Fineran PC: **The Cpf1 CRISPR-Cas protein expands genome-editing tools**. *Genome Biol* 2015, **16**:251.

43. •• 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.e12.
This paper reports structures of DNA-bound Cas12b complexes and elucidates the mechanism of DNA cleavage.

44. 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.

45. 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.

46. •• 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.
This paper reports the first Cas12a ternary complex structure and describes the mechanism of PAM recognition.

47. 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.

48. Künne T, Swarts DC, Brouns SJJ: **Planting the seed: target recognition of short guide RNAs**. *Trends Microbiol* 2014, **22**:74-83.

49. • Stella S, Alcón P, Montoya G: **Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage**. *Nature* 2017, **546**:559-563.
This paper reports the structure Cas12a bound to a cleaved DNA product, providing insights into R-loop formation.

50. Yamano T, Zetsche B, Ishitani R, Zhang F, Nishimasu H, Nureki O: **Structural basis for the canonical and non-canonical PAM recognition by CRISPR-Cpf1**. *Mol Cell* 2017, **67** 633–645.e3.

51. Nishimasu H, Yamano T, Gao L, Zhang F, Ishitani R, Nureki O, Zetsche B, Ishitani R, Zhang F, Nishimasu H *et al.*: **Structural basis for the altered PAM recognition by engineered CRISPR-Cpf1**. *Mol Cell* 2017, **67** 633–645.e3.

52. Koonin EV, Makarova KS, Zhang F: **Diversity, classification and evolution of CRISPR-Cas systems**. *Curr Opin Microbiol* 2017, **37**:67-78.

53. •• 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.
This paper presents the biochemical characterization of Cas13a, describing the pre-crRNA processing activity and target cleavage.

54. East-Seletsky A, O'Connell MR, Burstein D, Knott GJ, Doudna JA: **RNA targeting by functionally orthogonal type VI-A CRISPR-Cas enzymes**. *Mol Cell* 2017, **66** 373–383.e3.

55. •• 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.e12..
This paper reports the molecular architecture of Cas13a and its mechanism of crRNA binding.

56. •• Liu L, Li X, Ma J, Li Z, You L, Wang J, Wang M, Zhang X, Wang Y: **The molecular architecture for RNA-guided RNA cleavage by Cas13a**. *Cell* 2017, **170** 714–726.e10.
This paper reports the Cas13a structure bound to crRNA and a target RNA, providing mechanistics insights into its collateral RNA cleavage activity.