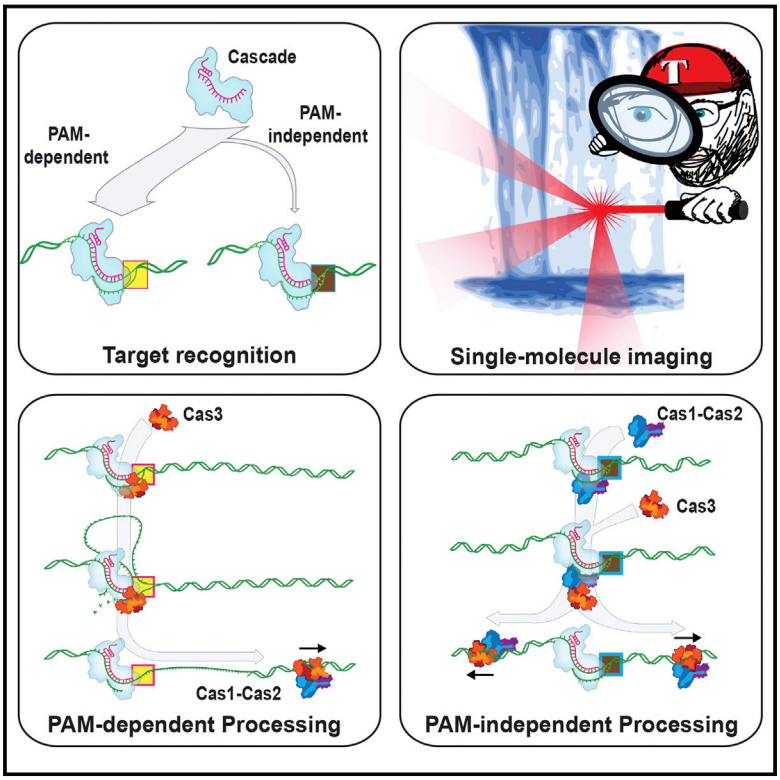
Article

Cell

Surveillance and Processing of Foreign DNA by the *Escherichia coli* CRISPR-Cas System

Graphical Abstract



Highlights

- Cascade can locate targets through PAM-dependent and PAM-independent pathways
- PAM-dependent recognition enables direct recruitment of the Cas3 translocase/nuclease
- PAM-independent recognition requires Cas1-Cas2 for Cas3 recruitment
- Cas1-Cas2 serve as *trans*-acting factors that regulate Cas3 activities

Authors

Sy Redding, Samuel H. Sternberg, Myles Marshall, ..., Blake Wiedenheft, Jennifer A. Doudna, Eric C. Greene

Correspondence

doudna@berkeley.edu (J.A.D.), ecg2108@cumc.columbia.edu (E.C.G.)

In Brief

Single-molecule analysis of the bacterial Cascade complex reveals two distinct pathways leading to differential regulation of Cas3 DNA translocase and nuclease activities.

Redding et al., 2015, Cell *163*, 854–865
November 5, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2015.10.003

Cell

Article

Surveillance and Processing of Foreign DNA by the *Escherichia coli* CRISPR-Cas System

Sy Redding,[1,9] Samuel H. Sternberg,[2] Myles Marshall,[8] Bryan Gibb,[8,10] Prashant Bhat,[3,11] Chantal K. Guegler,[2,12] Blake Wiedenheft,[4] Jennifer A. Doudna,[2,3,5,6,7,*] and Eric C. Greene[8,*]

[1]Department of Chemistry, Columbia University, New York, NY 10027, USA
[2]Department of Chemistry, University of California, Berkeley, CA 94720, USA
[3]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
[4]Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717, USA
[5]Innovative Genomics Initiative, University of California, Berkeley, CA 94720, USA
[6]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA
[7]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[8]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA
[9]Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA
[10]Present address: Department of Life Sciences, New York Institute of Technology, Old Westbury, NY 11568, USA
[11]Present address: David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA
[12]Present address: Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA
*Correspondence: doudna@berkeley.edu (J.A.D.), ecg2108@cumc.columbia.edu (E.C.G.)
http://dx.doi.org/10.1016/j.cell.2015.10.003

SUMMARY

CRISPR-Cas adaptive immune systems protect bacteria and archaea against foreign genetic elements. In *Escherichia coli,* Cascade (CRISPR-associated complex for antiviral defense) is an RNA-guided surveillance complex that binds foreign DNA and recruits Cas3, a *trans*-acting nuclease helicase for target degradation. Here, we use single-molecule imaging to visualize Cascade and Cas3 binding to foreign DNA targets. Our analysis reveals two distinct pathways dictated by the presence or absence of a protospacer-adjacent motif (PAM). Binding to a protospacer flanked by a PAM recruits a nuclease-active Cas3 for degradation of short single-stranded regions of target DNA, whereas PAM mutations elicit an alternative pathway that recruits a nuclease-inactive Cas3 through a mechanism that is dependent on the Cas1 and Cas2 proteins. These findings explain how target recognition by Cascade can elicit distinct outcomes and support a model for acquisition of new spacer sequences through a mechanism involving processive, ATP-dependent Cas3 translocation along foreign DNA.

INTRODUCTION

Many prokaryotes harbor an RNA-guided adaptive immune system comprised of a genetic locus called CRISPR (clustered regularly interspaced short palindromic repeats) and the CRISPR-associated (Cas) genes (Barrangou and Marraffini, 2014; van der Oost et al., 2014; Westra et al., 2012a). The CRISPR locus was first identified in *Escherichia coli* as an unusual series of 29-bp repeats separated by 32-bp spacer sequences (Ishino et al., 1987). It was later recognized that these spacers were derived from foreign genetic elements, suggesting the CRISPR locus might serve as an RNA-guided immune system (Bolotin et al., 2005; Makarova et al., 2006; Mojica et al., 2005). It is now known that CRISPR-Cas immunity is conferred through integration of short DNA fragments into the CRISPR locus, and these spacer sequences record the history of past infections (Barrangou and Marraffini, 2014; van der Oost et al., 2014; Westra et al., 2012a). The CRISPR locus is transcribed, and the resultant transcript is processed into shorter CRISPR-RNAs (crRNAs), each containing a sequence complementary to a previously encountered foreign DNA element.

CRISPR-Cas systems are classified as types I, II or III, which can be distinguished based on the presence of the signature *Cas3*, *Cas9*, or *Cas10* genes, respectively (Barrangou and Marraffini, 2014; Westra et al., 2012a). Type I are the most common, and much of our understanding of type I CRISPR-Cas systems comes from studies of *E. coli* Cascade (CRISPR-associated complex for antiviral defense), which is comprised of the five proteins Cse1, Cse2, Cas7, Cas5e, and Cas6e. These proteins assemble on a 61-nt crRNA, yielding a 405- kDa complex. The crRNA contains the 32-nt spacer sequence, which directs Cascade to sequences (protospacers) in foreign DNA, leading to formation of an R-loop intermediate. Cascade then recruits Cas3, which has an N-terminal histidine-aspartate (HD) nuclease domain and C-terminal superfamily 2 (SF2) helicase domain, to degrade the DNA (Mulepati and Bailey, 2013; Sinkunas et al., 2013).

Cascade must discriminate between spacer sequences found in the bacterial chromosome and those found in foreign DNA. This discrimination is thought to be accomplished through recognition of a trinucleotide sequence motif called the protospacer-adjacent motif (PAM; 5′-A[A/T]G-3′ for *E. coli* Cascade), which is adjacent to the protospacer in foreign DNA, but absent in the CRISPR locus. Strict sequence

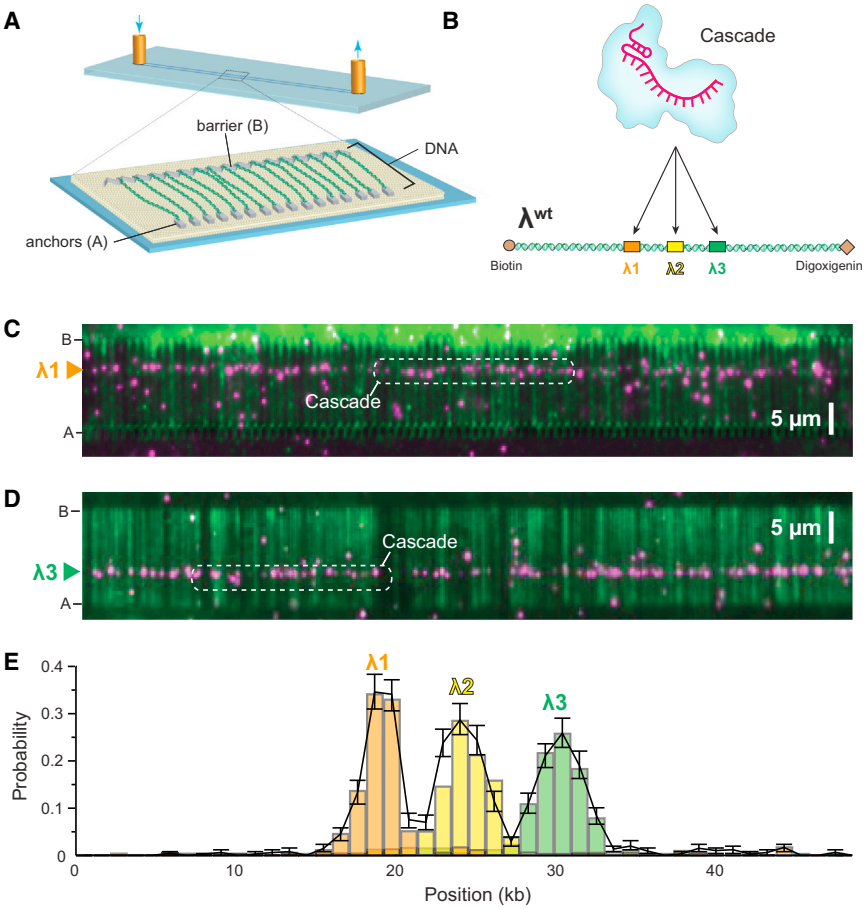


Figure 1. Programmed Target Binding by *E. coli* Cascade

(A) Overview of DNA curtains.

(B) Schematic of *E. coli* Cascade programmed with a crRNA targeting one of three different binding sites (designated λ1, λ2, and λ3) on λ^{WT}.

(C) Wide-field TIRF microscopy image showing QD-tagged Cascade (magenta) bound to DNA (green) at λ1.

(D) Wide-field image showing Cascade bound at λ3.

(E) Binding distribution for Cascade targeted to each of the three protospacers; error bars here and all subsequent binding distributions represent 95% confidence intervals obtained through bootstrap analysis.

requirements present a potential weakness because mutations in either the PAM or protospacer can allow foreign DNA to escape CRISPR-Cas immunity (Semenova et al., 2011). However, bacteria can rapidly restore immunity using a positive-feedback loop to update the CRISPR locus (Datsenko et al., 2012; Fineran et al., 2014). The mechanism of primed spacer acquisition (priming) remains perhaps one of the most poorly understood aspects of CRISPR-Cas immunity (Datsenko et al., 2012; Fineran et al., 2014; Heler et al., 2014). Priming requires Cascade with a crRNA bearing at least partial complementarity to the escape target, suggesting Cascade must be able to locate targets even when they bear mutations sufficient to escape immunity (Datsenko et al., 2012). Priming also requires Cas3 (Datsenko et al., 2012) and the Cas1-Cas2 complex (Nuñez et al., 2014), which integrate new sequences into the CRISPR locus (Nuñez et al., 2015). It is not known how these complexes elicit the priming response to foreign elements bearing escape mutations.

Here, we use single-molecule imaging to visualize individual Cascade complexes as they search for protospacers within the bacteriophage λ genome. Our work reveals PAM-dependent and PAM-independent search pathways. The PAM-dependent pathway is highly efficient and allows Cascade to recruit Cas3 for strand-specific degradation of the target genome. The PAM-independent pathway is less efficient, but Cascade can still bind tightly to the DNA, ensuring that it can initiate the sequence of molecular events that precede primed spacer acquisition. Through this pathway, Cas3 recruitment becomes strictly dependent on Cas1-Cas2, and Cas1-Cas2 also attenuate Cas3 nuclease activity and enable Cas3 to rapidly translocate in either direction along the foreign DNA. These results establish Cas1-Cas2 as a *trans*-acting factor necessary for the recruitment and regulation of Cas3 at escape targets. Based on our findings, we propose a mechanistic framework describing how Cascade, Cas1, Cas2, and Cas3 work together to process and disable foreign genetic elements.

RESULTS

DNA Curtain Assay for Target Binding by Cascade

We sought to establish a DNA curtain assay using total internal reflection fluorescence (TIRF) microscopy for visualizing the behavior of Cascade on individual molecules of wild-type phage λ DNA (λ^{WT}) (Figure 1A; Supplemental Experimental Procedures) (Greene et al., 2010). In brief, the surface of a microfluidic sample chamber was coated with a lipid bilayer, and DNA molecules were anchored to the bilayer through a biotin-streptavidin interaction. The DNA was then pushed to the leading edges of nanofabricated barriers to lipid diffusion, and the downstream ends were anchored to pedestals through antibody-hapten linkage (Gorman et al., 2010, 2012). Cascade was prepared with one of three crRNAs targeted to different regions of λ^{WT} and then labeled with antiFLAG-quantum dots (QDs) attached to the 3xFLAG-tagged Cas6e subunit (Figure 1B). When visualized on DNA curtains, Cascade bound to target sites corresponding to DNA sequences complementary to the three different crRNAs (Figures 1C–1E). Cascade remained bound for at least 57 min; this lifetime represents a lower limit for the Cascade-protospacer interaction because these measurements are limited by the

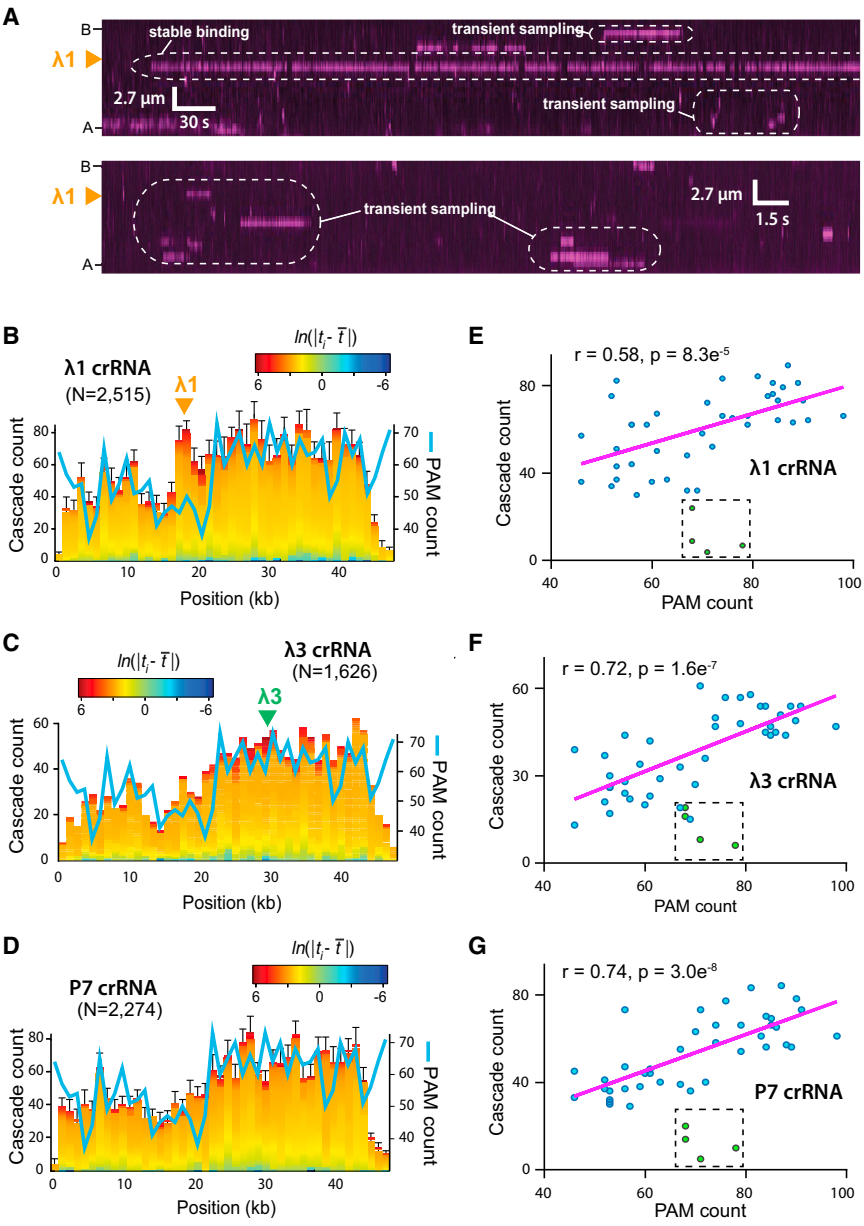


Figure 2. Cascade Searches for PAMs while Interrogating Foreign DNA

(A) Kymographs highlighting examples of Cascade binding events over two different time regimes (see scale bars). Examples of transient sampling and stable recognition are highlighted.

(B–D) Distribution of PAMs (blue line) and transient binding events for Cascade programmed with (B) the λ1-crRNA, (C) the λ3-crRNA, or (D) a P7-crRNA. Count refers to number of occurrences within 1 kbp of DNA. The locations of the λ1 and λ3 target sites are indicated, and the heat map color-coding reflects the binding dwell time (t_i) relative to the mean dwell time ($\bar{t}$).

(E–G) Correlation of PAMs with the transient binding events for Cascade programmed with (E) the λ1-crRNA, (F) the λ3-crRNA, or (G) P7-crRNA, as indicated. Outlying data points (colored green and boxed) reflect underrepresented binding events at PAM sites near the ends of the DNA; detection of binding at these sites is hindered by the chromium barriers.

See also Figure S1.

stability of the 3xFLAG-antiFLAG interaction (Sternberg et al., 2014). Stable binding was not observed for Cascade bearing a control crRNA (P7-crRNA) that was not complementary to λ^{WT}.

PAM-Dependent Target Recognition

Next, we sought to determine how Cascade locates protospacers by visualizing reactions in real-time. Most Cascade (> 75%) appeared immediately at the protospacer without exhibiting any evidence of microscopically detectable motion along the DNA (Figure 2A). This finding leads us to conclude that Cascade located the protospacer through a pathway that was dominated by 3D diffusion at the microscopic scale. Based on our optical resolution limits, these experiments provide an upper limit on any potential 1D diffusion by Cascade of no more than ~250 bp, although we do not exclude that possibility that Cascade may diffuse shorter distances along the DNA (Gorman et al., 2010, 2012). The remaining fraction of Cascade molecules (<25%) underwent optically detectable 1D diffusion; we did not pursue a detailed analysis of this behavior because it coincided with a loss of binding specificity and appeared to arise from Cascade aggregates (data not shown).

Analysis of the 3D events revealed long-lived binding to the protospacers, as well as transient binding events all along the λ^{WT} DNA (Figure 2A). The λ^{WT} genome contains a total of 3,151 PAM sites (5′-A[A/T]G-3′), corresponding to ~1 PAM per 15.4 bp, which are asymmetrically distributed across the phage genome. Cascade did not randomly sample the DNA, instead the transient binding events were correlated with the PAM distribution (Figures 2E–2G), as we have reported for *Streptococcus pyogenes* Cas9 (Sternberg et al., 2014). Control reactions using Cascade programmed with P7-crRNA revealed a similar pattern of transient binding (Figures 2B–2G), and we could detect no binding activity for Cascade lacking Cse1 (data not shown), which is the subunit responsible for PAM recognition (Sashital et al., 2012). Cascade programmed with either λ1-crRNA or λ3-crRNA displayed many reversible binding events at their targets, which are revealed by the ~50% increased prevalence of longer-lived intermediates at both of these target sites relative to non-target sites (Figures 2B, 2C, and S1A), and also by the peak in binding at λ1 for the λ1-crRNA, which is observable due to the overall lower density of PAM sites in this region of DNA (Figures 2B and 2C). This

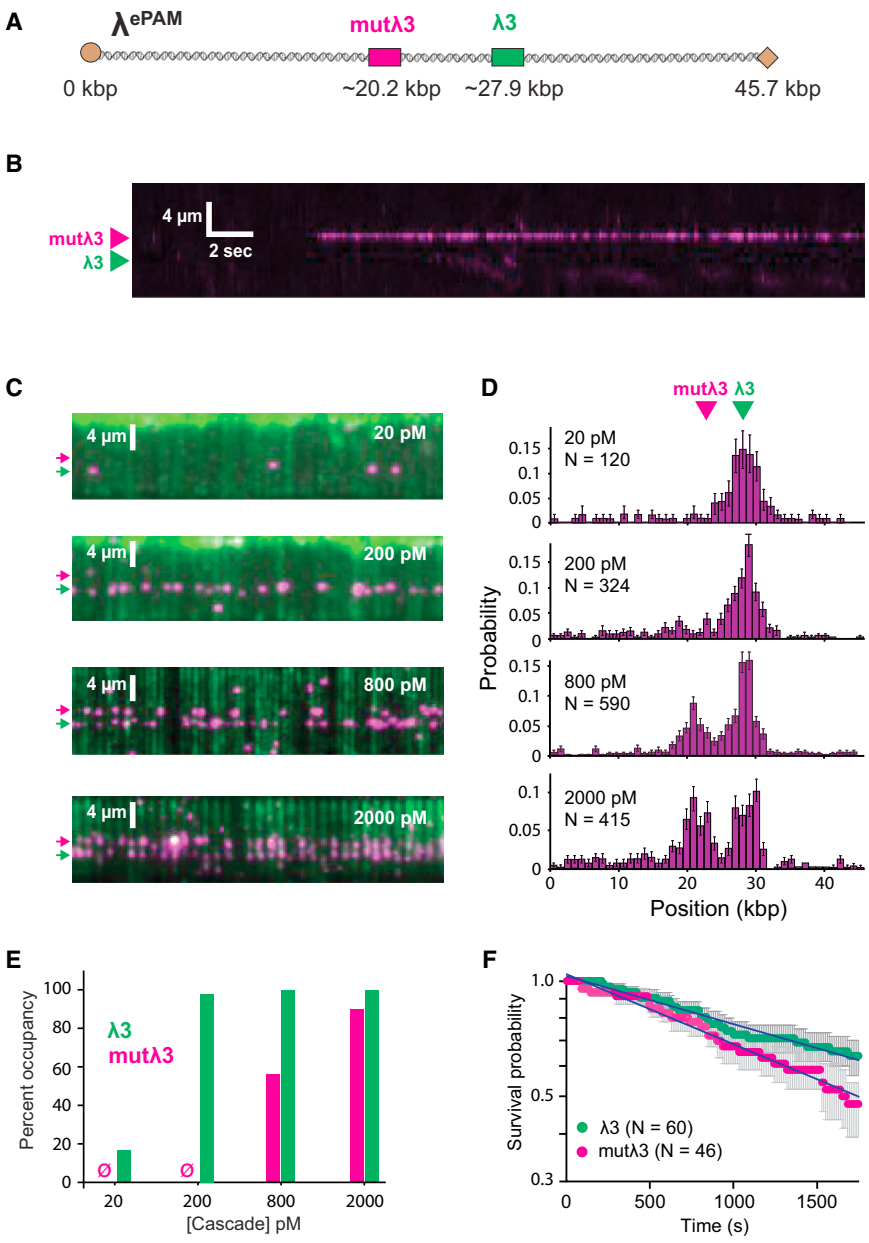


Figure 3. Recognition of Escape PAM Mutants

(A) Schematic of λ^{ePAM} bearing two identical protospacers, one with a cognate PAM (λ3) and the other with an escape PAM (mutλ3).

(B) Kymograph highlighting example of Cascade binding to the mutλ3 through 3D diffusion.

(C) Wide-field images showing binding to each of the two targets at different Cascade concentrations following a 10-min incubation. Arrowheads indicate the locations of the λ3 (green) and mutλ3 (magenta) targets.

(D) Binding distributions showing relative occupancy at each Cascade concentration.

(E) Quantification of percent occupancy; Ø indicates no detectable binding.

(F) Survival probability plots for Cascade bound to the two targets; error bars here and all subsequent survival probability plots represent 70% confidence intervals obtained through bootstrap analysis.

See also Table S1.

category of long-lived, but reversible, binding events at the protospacers likely represents abortive engagement, suggesting Cascade must often make multiple attempts before stably engaging the protospacer, similar to what we have observed for Cas9 (Sternberg et al., 2014).

The transient binding events exhibited double-exponential decays similar to *S. pyogenes* Cas9 (Sternberg et al., 2014), with lifetimes of ~3 and ~25 s (Figures S1A–S1D), indicating that at least two intermediates exist on the pathway toward target recognition. The lifetimes of these intermediates were not appreciably affected by either salt concentration or temperature (Figure S1D), similar to findings for Cas9 (Sternberg et al., 2014). These characteristics, more commonly attributed to site-specific association, provide further evidence that the initial observed interactions are based on a sequence-dependent association with PAM sites rather than on nonspecific interactions with the DNA phosphate backbone.

PAM-Independent Target Recognition

Next, we sought to determine whether and how Cascade locates targets that lack a canonical PAM. For this, we generated a new phage construct (λ^{ePAM}) bearing two duplicate targets (Figure 3A). One of the protospacers (λ3) was adjacent to a cognate PAM [5′-ATG-3′], whereas second protospacer (mutλ3) was adjacent to a mutated PAM [5-ATT-3′]. This escape PAM (ePAM) was chosen because it enables an invading DNA to escape the CRISPR-Cas machinery, but still elicits a rapid priming response (Datsenko et al., 2012; Fineran et al., 2014). Surprisingly, Cascade could still bind both protospacers, and binding of mutλ3 still occurred through 3D diffusion (Figure 3B), but recognition of mutλ3 was much less efficient than recognition of λ3 (Figures 3C–3E). This difference was evidenced by the ~10-fold higher Cascade concentration necessary to achieve similar levels of occupancy at both protospacers. Despite the large difference in initial recognition, the lifetimes of Cascade at λ3 and mutλ3 were comparable (57 and 40 min, respectively; Figure 3F). We conclude that PAMs increase the efficiency of target recognition, but that Cascade is still capable of protospacer recognition and high-affinity binding in the absence of a cognate PAM, and this conclusion is consistent with previous studies (Szczelkun et al., 2014).

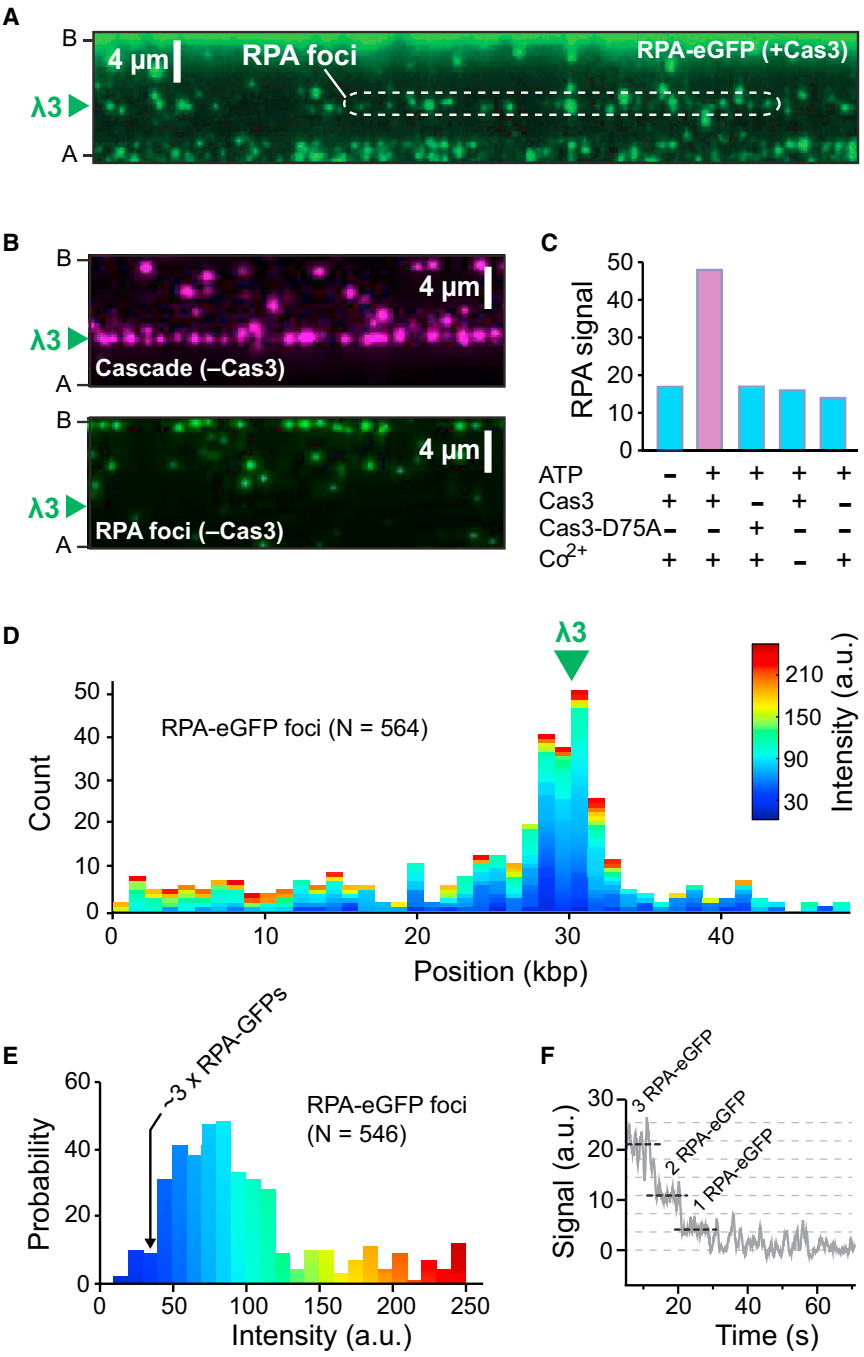


Figure 4. Cas3 Generates an ssDNA Gap at the λ3 Protospacer

(A) Image showing RPA-eGFP foci at λ3 for reactions with unlabeled Cascade and unlabeled Cas3.

(B) Control images showing that RPA-eGFP foci are not present when Cas3 is omitted from the reactions; the upper and lower panels show the same field of view.

(C) Requirements for RPA-eGFP foci formation at λ3.

(D) Distribution of RPA-eGFP foci in reactions containing both Cascade and Cas3; Count refers to the number of occurrences within 1 kbp of DNA.

(E) Signal intensities for RPA-eGFP foci. The intensity of a focus comprised of three molecules of RPA-eGFP is indicated, and each successive bin corresponds to ~1 additional molecule of RPA-eGFP. The heat map color-coding in (D) and (E) are the same.

(F) Representative stepwise photobleaching curve used to estimate the number of RPA-eGFP molecules in each focus.

See also Figure S2.

Cas3 Recruitment Leads to Disruption of the Target DNA Duplex

Next, we sought to visualize Cascade-dependent recruitment of Cas3. Cas3 interacts with Cse1, and the displaced single-stranded DNA (ssDNA) strand that is generated by R-loop formation (Hochstrasser et al., 2014; Mulepati and Bailey, 2013; Sinkunas et al., 2013). Upon recruitment, Cas3 first nicks the DNA and is thought to then translocate in the 3′→5′ direction along the non-target strand, while unwinding and degrading duplex DNA through an ATP-, Mg^{2+}-, and Co^{2+}-dependent mechanism (Mulepati and Bailey, 2011, 2013; Sinkunas et al., 2013). Cas3 degrades both target DNA strands in bulk biochemical assays (Mulepati and Bailey, 2013; Sinkunas et al., 2013). However, these measurements use relatively high concentrations of Cas3 (50 nM–1 μM) (Hochstrasser et al., 2014; Mulepati and Bailey, 2011, 2013; Sinkunas et al., 2011, 2013), suggesting that DNA degradation may be due to the action of multiple Cas3 molecules, only the first of which is directly recruited by Cascade. Given these considerations, it is plausible that the initial Cascade-recruited molecule of Cas3 only introduces a small nick or ssDNA gap in the target DNA (Mulepati and Bailey, 2013).

We reasoned that if Cas3 was initially generating ssDNA after loading at Cascade, then this might be revealed in reactions with low concentrations of Cas3 (4 nM), followed by the addition of eGFP-tagged replication protein A (RPA), which binds ssDNA. When RPA-eGFP was added after Cascade and Cas3, bright eGFP foci were detected at the λ3 protospacer (Figure 4A). Formation of RPA-eGFP foci was dependent on Cascade, Cas3, ATP, and Co^{2+} and the conditions under which we detected RPA-eGFP foci paralleled the conditions necessary for plasmid degradation in bulk biochemical assays (Figures 4B and S2C). Furthermore, RPA-eGFP foci were not observed for a Cas3 nuclease mutant (D75A) (Figures 4C and S2A–S2C). Notably, the DNA in the single-molecule assays

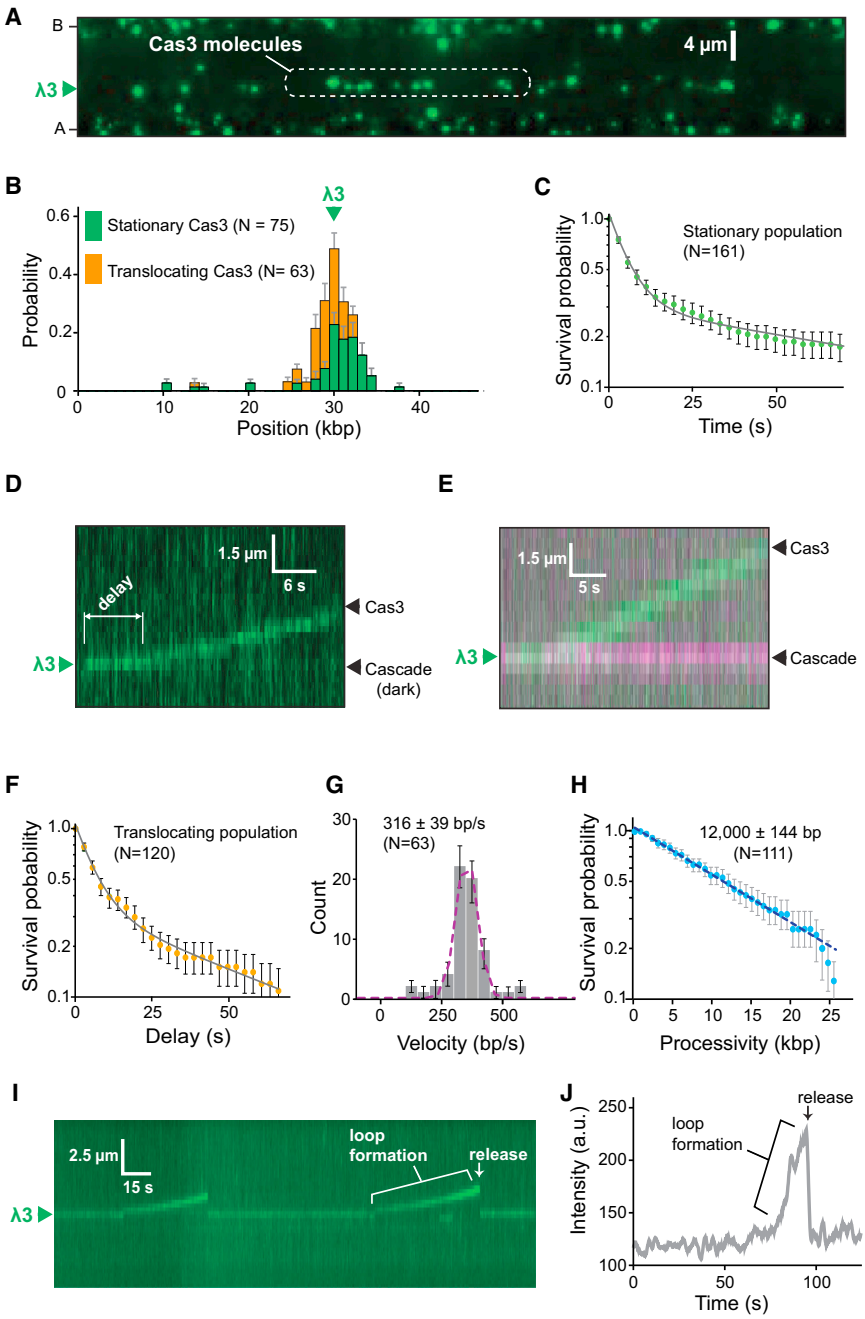


Figure 5. Cascade-Mediated Recruitment of Cas3

(A) Image showing that QD-tagged Cas3 is recruited to unlabeled Cascade at λ3.

(B) Binding of Cas3 to λ3. The distribution is segregated into the translocation (orange) and stationary (green) Cas3 populations.

(C) Survival probabilities of the stationary Cas3 population.

(D) Kymograph illustrating the translocation of Cas3 away from λ3 in a reaction with unlabeled Cascade. The delay period prior to the initiation of Cas3 translocation is indicated.

(E) Two-color experiment showing that Cas3 (green) translocates away from Cascade (magenta).

(F) Survival probability (delay time) of the translocating population of Cas3 prior to moving away from λ3.

(G) Cas3 velocity distribution.

(H) Cas3 processivity distribution.

(I) Kymograph showing an example of Cas3 repeatedly looping the DNA.

(J) Intensity profile showing the increase in Cas3 fluorescence signal coinciding with DNA loop formation.

See also Figure S3 and Movie S1.

was not liberated from the flow cell surface and there was no evidence for long tracts of RPA-eGFP, indicating that Cas3 only generated a small ssDNA gap. To estimate the size of the ssDNA gaps, we measured the intensity of the RPA-eGFP foci (Figures 4D and 4E) and then used photobleaching steps to roughly estimate the number of RPA-eGFP molecules present (Figure 4F; Supplemental Experimental Procedures). We estimate that the average focus contained ~8–10 molecules of RPA-eGFP, corresponding to ~240–300 nt of ssDNA. These results suggest that the first Cas3 molecule recruited by Cascade makes a short ssDNA gap adjacent to the protospacer.

Cas3 Recruitment to Target-Bound Cascade

Next, we sought to visualize the behavior of fluorescently tagged Cas3 (Figure S3A; Supplemental Experimental Procedures). We were unable to detect stable binding of Cas3 to Cascade when ATP or Mg^{2+} were omitted or when ATP was replaced with ADP or AMP-PNP (data not shown). However, Cas3 bound stably to Cascade when ATP and Mg^{2+} were included in the reactions (Figures 5A and 5B). Cas3 located Cascade through 3D diffusion during initial recruitment (see Figure 5D). Once bound, ~55% of the Cas3 molecules remained stationary within optical resolution limits (Figure 5B). These seemingly stationary molecules exhibited two distinct lifetimes: one population with a lifetime (τ_1) of ~6 s and a second population (τ_2) with a lifetime of ≥1 min (Figures 5C, S3B, and S3C). These findings suggest that Cas3 transiently samples target-bound Cascade before transitioning into a more stably bound state and that entry into this longer-lived state requires ATP hydrolysis. Interestingly, once a longer-lived Cas3 binding event was observed at a given molecule of Cascade, then that particular Cascade complex appeared incapable of recruiting any additional Cas3 at the protein concentrations used in these assays.

Cas3 Is a Highly Processive Molecular Motor

Many of the Cas3 molecules (~45%) translocated along the DNA (Figures 5B, 5D, and 5E). In these instances, Cas3 was recruited

to Cascade at the λ3 protospacer and then moved rapidly away from the protospacer in a direction consistent with 3′→5′ translocation on the non-target strand, as expected from bulk biochemical experiments (Mulepati and Bailey, 2013). There was no evidence that Cas3 translocation could initiate from any other location on the DNA other than the λ3 protospacer, and Cas3 translocation was entirely dependent on the presence of Cascade. Remarkably, Cascade remained tightly bound to the protospacer even after Cas3 had begun translocating along the DNA (Figure 5E). Moreover, once Cas3 had translocated away from Cascade, then no additional molecules of Cas3 could bind to or translocate away from that particular Cascade complex.

Cas3 exhibited a short delay prior to moving away from Cascade (Figure 5D); analysis of these delay times revealed two lifetimes that were similar to the τ_1 and τ_2 lifetimes for the stationary Cas3 population, suggesting that the observed intermediates reflected the same underlying molecular processes (Figures 5C, 5F, S3B, and S3C). Cas3 traveled at a mean velocity of ~316 bp/s for 12,000 bp before stalling or dissociating from the DNA (Figures 5G and 5H), and >99% of molecules exhibited unidirectional movement (Figures 5D and 5E; Movie S1; see below). Three key observations suggested that Cas3 was not extensively degrading the DNA during translocation. First, there was no evidence that the translocating population of Cas3 caused double-strand breaks. Second, we saw no evidence for long ssDNA tracts when reactions were chased with RPA-eGFP. Finally, if Cas3 had generated tracts of ssDNA long enough to be optically detected, then Cascade would also appear to move in the same direction because of the change in persistence length that accompanies the conversion of dsDNA to ssDNA, but Cascade always remained stationary at the protospacer. We conclude that Cas3 is a highly processive molecular motor that first generates a small ssDNA gap and then translocates in 3′→5′ direction along the non-target DNA strand away from Cascade.

Evidence for Looped DNA Intermediates

Surprisingly, in addition to our observation that Cas3 recruitment and translocation did not coincide with the ejection of Cascade from the DNA, inspection of the Cas3 translocation trajectories revealed evidence that the contacts between Cas3 and Cascade were not immediately broken. In many instances (14%), Cas3 began to translocate along the DNA, but then returned almost instantaneously to the original binding site (Figure 5I). This behavior coincided with an increase in Cas3 fluorescence, suggesting that the molecules were pulled closer to the surface of the flow cell because of increased tension on the DNA. These observations are most consistent with looped DNA intermediates, where Cas3 maintains contact with Cascade, while simultaneously translocating for a short distance along the flanking duplex DNA (Figure S3D). We conclude that Cas3 can initially remain bound to Cascade as it begins translocating along the DNA and that a subset of these molecules generates optically detectable DNA loops.

PAM Is Essential for Cascade-Mediated Recruitment of Cas3

Next, we sought to determine whether Cascade could recruit Cas3 to mutλ3, and if so, whether the properties of Cas3 differ in the absence of a cognate PAM. Interestingly, Cas3 did not colocalize with Cascade at mutλ3 (Figure S4A), and we were unable to detect even transient binding of Cas3 to Cascade at the mutλ3 protospacer. We were also unable to detect RPA-eGFP foci at mutλ3 (Figures S4B–S4E), and Cas3 did not cleave plasmid substrates bearing the mutλ3 protospacer (see below). We conclude that Cascade cannot recruit Cas3 to DNA in the absence of a cognate PAM, in agreement with previous bulk biochemical experiments (Hochstrasser et al., 2014; Mulepati and Bailey, 2013).

PAM-Independent Recruitment of Cas3 by Cas1-Cas2

Cas3 is required for primed sequence acquisition (Datsenko et al., 2012), suggesting that alternative pathways must exist to recruit Cas3 to escape targets. Cas1 and Cas2 are universally conserved across CRISPR types and are also necessary for primed sequence acquisition, suggesting the possibility that these proteins may work in concert with Cascade to promote the recruitment of Cas3 to escape targets. Therefore, we next asked whether the Cas1-Cas2 complex might affect target recognition, target processing, or both, in reactions with Cas3. Attempts to generate fluorescently tagged Cas1 or Cas2 yielded inactive proteins, and therefore these experiments utilized wild-type (unlabeled) Cas1-Cas2.

Remarkably, the addition of Cas1-Cas2 enabled the recruitment of Cas3 to mutλ3 and also ~3-fold enhanced recruitment of Cas3 to λ3 (Figures 6A and 6B; Movie S1 and S2). The velocity and processivity of Cas3 were not altered by Cas1-Cas2 (Figures S5A and S5B). However, Cas3 recruited to the escape target behaved markedly different from Cas3 that was recruited to cognate protospacer. Most strikingly, Cas3 targeted to mutλ3 could rapidly translocate in either direction away from Cascade (Figure 6C; Movie S3). Moreover, Cas3 exhibited only a ~6 s delay prior to moving away from mutλ3, but there was no evidence for the second longer-lived intermediate (τ_2) that was always observed at λ3 (Figures S3B, S3C, and S5C). There was also no evidence for ssDNA gaps at mutλ3 in the presence of Cas1-Cas2 (Figure S5D), and bulk biochemical assays with Cascade, Cas1-Cas2, and Cas3 revealed no nicking or cleavage of plasmids with the mutλ3 protospacer (Figures S6A and S6B), even though Cascade was capable of binding the mutλ3 protospacer in bulk assays (Figure S6C). Finally, there was no evidence for Cas3-mediated DNA looping at mutλ3 in reactions with Cas1-Cas2 (Figure 6C). Together, these results show that Cas1-Cas2 are necessary to recruit Cas3 to mutλ3 and attenuate the nuclease activity of Cas3 at these escape targets, enabling Cas3 to translocate away from Cascade in either direction along the foreign DNA.

Cas1-Cas2 also appeared to affect the behavior of Cas3 at the λ3 protospacer. Specifically, Cas1-Cas2 partially attenuated Cas3 nuclease activity in bulk biochemical assays (Figures S6A and S6B), and the presence of Cas1-Cas2 also enabled iterative Cas3 binding and translocation events from the same Cascade complex bound to the λ3 protospacer (Movie S4). This observation was in stark contrast to reactions done in the absence of Cas1-Cas2, where we never detected evidence of multiple Cas3 recruitment events to the same Cascade complex. These findings suggest that Cas1-Cas2 not only enhances the recruitment of Cas3 to Cascade bound at λ3 but may also enable iterative Cas3 loading events.

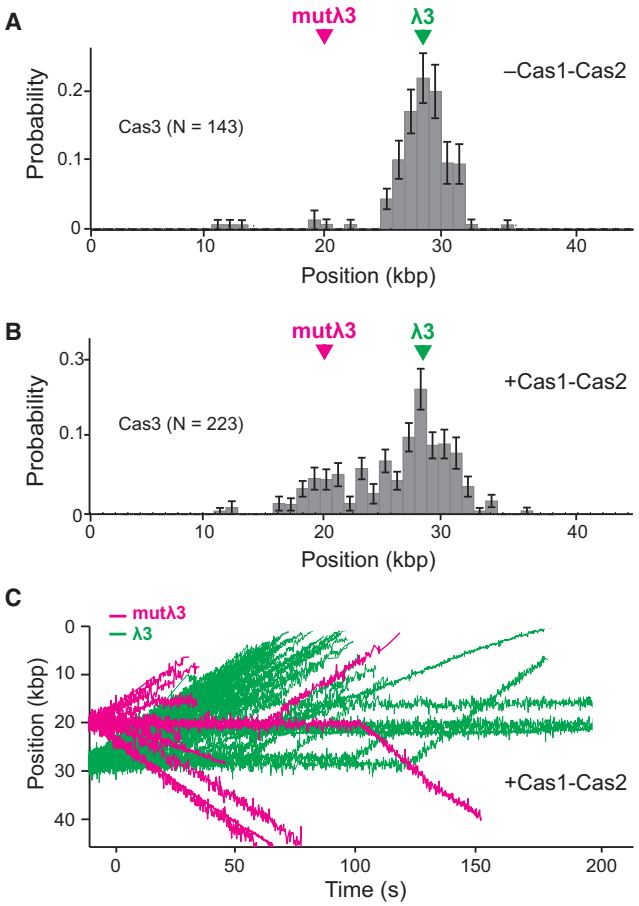


Figure 6. Cas1-Cas2-Mediated Recruitment of Cas3 to Escape Targets

(A) Binding distribution of Cas3 on λ^{ePAM} in the absence of Cas1-Cas2.

(B) Cas3 binding distribution histogram on λ^{ePAM} in the presence of Cas1-Cas2.

(C) Overlaid trajectories showing examples of Cas3 translocation events originating from either the λ3 protospacer (green) or the mutλ3 protospacer (magenta). Of the trajectories originating from mutλ3, 59% of the Cas3 molecules move toward the downstream anchor points, and the remaining 41% travel in the opposite direction.

See also Figures S4 and S5 and Movies S2, S3, and S4.

DISCUSSION

CRISPR-Cas immunity involves complex interplay among multiple macromolecular components, with the potential for overlapping or convergent pathways. Our work reveals two distinct pathways for target recognition and processing and shows that the choice of pathway is dictated by the presence or absence of a PAM sequence adjacent to the targeted protospacer (Figure 7).

A Conserved Mechanism for PAM-Dependent Target Recognition

Our results support a model in which an initial search for PAM sequences is the predominant mode of DNA surveillance by *E. coli* Cascade (Figure 7A). Once a PAM is identified, Cascade interrogates the flanking DNA for sequence complementarity to the crRNA via directional unwinding of the DNA beginning at the PAM, and identification of a matching protospacer leads to stable capture and R-loop formation (Rutkauskas et al., 2015). This PAM-dependent search process is strikingly similar to that of *S. pyogenes* Cas9, the crRNA-guided surveillance complex in type II CRISPR-Cas systems, which also initiates the search by looking for PAMs (Sternberg et al., 2014). In addition, the type IF CRISPR-Cas system of *Pseudomonas aeruginosa* also searches for PAM sequences before probing the flanking DNA for sequence complementarity to the crRNA (Rollins et al., 2015). The type II CRISPR-Cas systems require only a single polypeptide for target recognition and cleavage, whereas type I CRISPR-Cas systems require large multimeric complexes for target recognition and a separate *trans*-acting protein (Cas3) for DNA cleavage. Cas9 and Cse1 share no amino acid sequence homology, and the Cas9 PAM (5′-NGG-3′) and the Cascade PAM (5′-A[A/T]G-3′) are located on opposite ends of the protospacer and on different DNA strands (Jinek et al., 2012; Sashital et al., 2012). Given these differences, there was no reason to assume that *S. pyogenes* Cas9 and *E. coli* Cascade would search for target sites using the same general mechanism. The similarities between Cascade and Cas9 suggest that an initial search for PAMs may be a broadly conserved mechanism for DNA surveillance among the type I and type II CRISPR-Cas systems.

Facilitated Diffusion versus Reduced Complexity

It is often assumed that site-specific DNA binding proteins accelerate target searches relative to 3D diffusion by facilitated diffusion, which reduces the dimensionality of the search process through 1D sliding, hopping, and/or intersegmental transfer (von Hippel and Berg, 1989). However, there is little evidence supporting this general assumption (Halford, 2009). The Cascade target search is remarkably similar to that of Cas9's, which also exhibits no evidence of 1D sliding (Sternberg et al., 2014). Instead, we find that Cascade and Cas9 both appear to optimize their target searches by reducing the complexity of the sequence space that is sampled while surveying DNA. They accomplish this task by first looking for a small portion of the overall binding site, the PAM, before probing the flanking DNA for sequences complementary to the crRNA, which provides an additional layer of discrimination enabling Cascade to sample and reject incorrect targets (Rutkauskas et al., 2015; Sternberg et al., 2014). The effectiveness of this strategy can be illustrated by considering that based on sequence composition alone Cascade can avoid ~90% of the λ genome just by utilizing the PAM as an initial recognition signal, while kinetically ignoring other sequences. The finding that much higher Cascade concentrations are necessary to achieve similar occupancy at protospacers with an escape PAM compared to those with a cognate PAM also reflects the effectiveness of reducing search complexity.

PAM-Dependent Target Processing

The PAM-dependent pathway requires only Cascade to recruit Cas3 to protospacers (Figure 7B). Cas3 first transiently samples Cascade before transitioning into a stably bound complex.

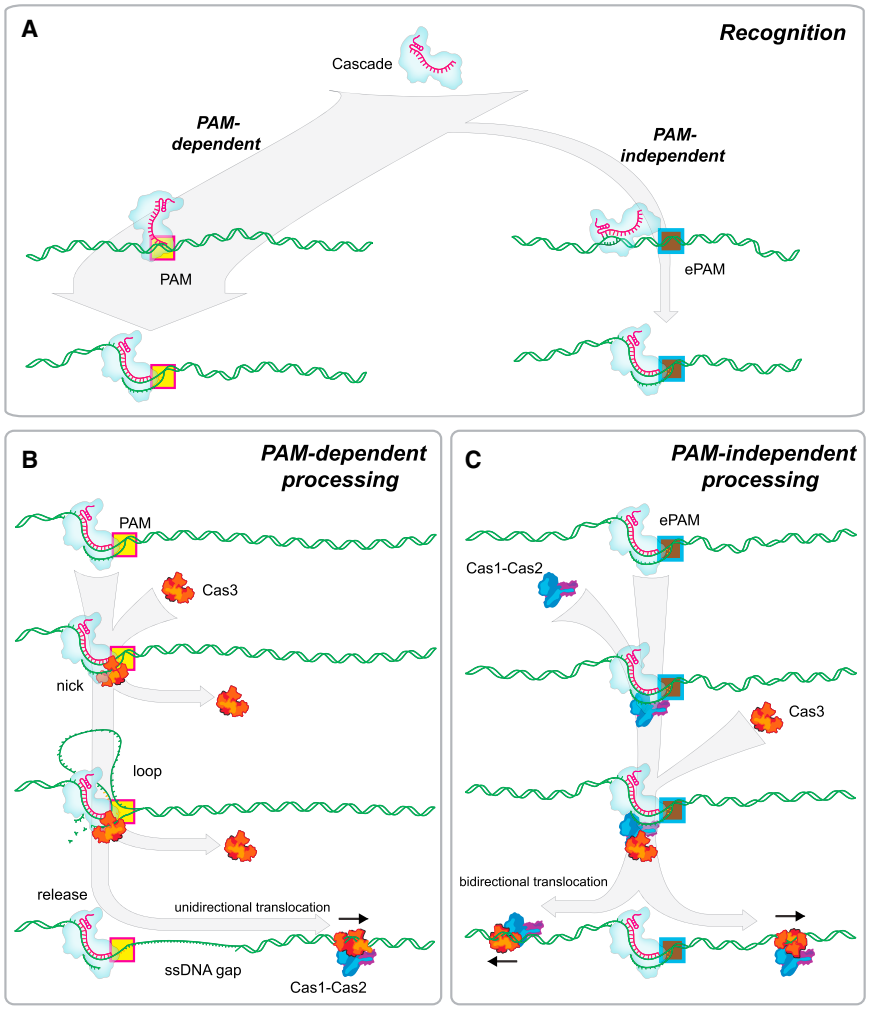


Figure 7. Model for Foreign DNA Recognition and Processing by Cascade, Cas1, Cas2, and Cas3

(A) The predominant mechanism for protospacer recognition is through the PAM-dependent pathway.

(B) PAM-dependent processing involves the recruitment of Cas3 to the protospacer by Cascade. Cas3 nicks the R-loop and generates an ssDNA gap; Cas3 can dissociate at either of these two steps. Cas3 then breaks free from Cascade and travels unidirectionally along the DNA.

(C) PAM-independent processing requires Cas1-Cas2 to recruit Cas3. Cas3 is loaded onto the DNA in one of two possible orientations through a mechanism that attenuates Cas3 nuclease activity. Cas3 then travels in either direction along the DNA as part of a spacer acquisition complex.

See also Figure S6.

Formation of this longer-lived species prevents any further Cascade-specific recruitment of Cas3, most likely because the first stably bound Cas3 cleaves the R-loop, which destroys the Cas3 binding site (Mulepati and Bailey, 2013). Consistent with this interpretation, formation of stable Cascade-Cas3 intermediates coincides with the appearance of a ~200- to 300-nt ssDNA gap adjacent to the protospacer. The first molecule of Cas3 does not appear to induce any damage other than creating this initial ssDNA gap. This finding is notably different from bulk biochemical assays, which reveal more extensive DNA degradation (Mulepati and Bailey, 2013; Sinkunas et al., 2013). This difference may be explained by the potential for recruitment of additional Cas3 molecules in the bulk biochemical assays through a Cascade-independent pathway, as previously suggested (Mulepati and Bailey, 2013). Consistent with this explanation, Cas3 is a potent ssDNA nuclease even in the absence of Cascade (Mulepati and Bailey, 2013; Sinkunas et al., 2011, 2013). Thus, the ssDNA gaps generated by the first molecule of Cas3 likely reflect an early intermediate in the degradation pathway and serve as an entryway for additional ssDNA-specific nucleases, including Cas3 or perhaps other host enzymes. Together, these findings suggest that the early stages of foreign DNA degradation involve the ATP-dependent recruitment of just one molecule of Cas3 through a mechanism that requires Cascade-specific contacts and an intact R-loop. This initial transient binding event exhibits a ~6-s lifetime (τ_1) before Cas3 transitions into a more stably bound intermediate. The first stably bound molecule of Cas3 then generates a short ssDNA gap, reflected in the delay time (τ_2) prior to moving away from Cascade, and after being released from Cascade, this Cas3 molecule can either dissociate into solution or continue traveling along the remaining duplex DNA. Any subsequent recruitment of Cas3 (or other nucleases) occurs through nonspecific interactions with the resultant ssDNA gap.

Cascade remains tightly bound to the DNA even after Cas3 generates an ssDNA gap and moves away from the protospacer. It is possible that continued presence of Cascade may distinguish these Cas3-generated gaps from other ssDNA gaps that can be produced during normal DNA metabolism, and Cascade may perhaps prevent host DNA repair proteins from filling in these gaps before the invading DNA is eventually destroyed.

PAM-Dependent Cas3 Motor Activities

Cas3 is a fast and highly processive molecular motor, which is recruited by Cascade through the PAM-dependent pathway and then translocates along the flanking DNA. This translocation does not coincide with any apparent DNA degradation or persistently unwound DNA. When Cas3 is recruited by Cascade through the PAM-dependent pathway, it always moves in the same direction along the DNA, consistent with expectations for 3′→5′ translocation along the non-target strand. A subset of Cas3 molecules also forms optically detectable looped intermediates, and Cas3 likely generates smaller DNA loops that cannot be observed in

our experiments, suggesting these looped intermediates may be a common feature of the PAM-dependent pathway (Figure 7B). Interestingly, similar looping behaviors have been reported for many different SF1 and SF2 helicases, including *Saccharomyces cerevisiae* Pif1 (Zhou et al., 2014), *Bacillus stearothermophilus* PcrA (Park et al., 2010), *E. coli* Rep (Myong et al., 2005), and *S. cerevisiae* Srs2 (Qiu et al., 2013). The looping behaviors exhibited by these proteins are thought to help establish and maintain a particular structural state of the DNA; for instance, PcrA and Srs2 repeatedly shuttle back and forth, while removing proteins from ssDNA proximal to an ssDNA/dsDNA junction to prevent aberrant recombination (Park et al., 2010; Qiu et al., 2013). Similarly, Pif1 repeatedly unwinds G-quadraplexes, ensuring that these structures do not inhibit DNA replication (Zhou et al., 2014). The looping activity observed for Cas3 may reflect attempts to dissociate from Cascade. Alternatively, looping may help keep the ssDNA gap clear of proteins, free of secondary structures or both, until the arrival of additional Cas3 molecules or other accessory nucleases.

PAM-Independent Target Recognition

Like the PAM-dependent search, the PAM-independent pathway also occurs by microscopic 3D diffusion, suggesting that Cascade must test for complementarity to the crRNA by either transiently melting the DNA or by taking advantage of the intrinsic breathing of the DNA duplex (Figure 7A). One primary difference between PAM-dependent and PAM-independent target recognition is that the efficiency of the PAM-independent pathway is comprised, such that a higher concentration of Cascade is required to achieve similar levels of occupancy at both targets. Despite this disparity in apparent association constants, Cascade can still bind tightly to DNA regardless of whether or not the protospacer has a canonical PAM. In both instances, the lifetime of the target-bound Cascade complexes is significantly longer that the typical doubling time of *E. coli*, a finding that is in good agreement with the results of magnetic tweezer experiments (Szczelkun et al., 2014). This tight binding would help ensure that even though escape target recognition is inefficient, in the rare instances in which an escape target is captured, Cascade would remain in place long enough to initiate downstream steps necessary for primed sequence acquisition (Figure 7C; see below). Interestingly, not all PAM mutations are equal with respect to Cascade, and the defect in binding with the ATT mutant PAM is more moderate that some other PAM mutations (Szczelkun et al., 2014). Future studies will be essential for testing the effects of other PAM mutations on target binding in these single-molecule assays.

Interestingly, recent single-molecule fluorescence resonance energy transfer (FRET) experiments have suggested that Cascade recognizes escape targets with substantially reduced fidelity, and interactions with these targets are characterized by a ~25-s lifetime (Blosser et al., 2015), which is identical to one of the nonspecific lifetimes observed in our experiments (Figure S1). We suggest that these shorter-lived complexes found by FRET reflect intermediates that have failed to transition into the more tightly bound complexes observed in our assays.

Importantly, PAM escape mutations reflect only a subset of mutations that can lead to a priming response, with the remainder occurring within the protospacer, but both types of escape mutants lead to similar priming responses (Datsenko et al., 2012; Fineran et al., 2014). We anticipate that Cascade will locate protospacer escape mutants through the normal PAM-dependent search pathway, but then may require Cas1-Cas2 to recruit Cas3 and initiate a priming response from this class of escape mutations.

Cas1-Cas2 Recruitment of Cas3 to Escape Targets

We demonstrate that the Cas1-Cas2 complex serves as a *trans*-acting factor necessary for the recruitment and regulation of Cas3 at protospacers bearing an escape PAM (Figure 7C). Recruitment may occur through one of two general mechanisms. Cas1-Cas2 may modify the structure of Cascade such that it can now directly recruit Cas3 by the same process as occurs during PAM-dependent recruitment. Alternatively, protein-protein contacts with Cas1-Cas2 may directly recruit Cas3 to the escape target through a mechanism that is distinct from the Cascade-dependent recruitment at cognate protospacers. Importantly, the behavior of Cas3 at the escape targets differs markedly from the behavior of Cas3 at cognate targets. First, Cas3 can translocate in either direction from the escape targets, implying that that Cas3 is loaded onto the flanking phage DNA through a different pathway than is observed at cognate protospacers. Second, there was no evidence that Cas3 generates ssDNA gaps at the escape targets, nor was there any evidence that Cas3 even nicked the DNA when loaded at escape targets, suggesting that the nuclease activity of Cas3 is fully attenuated at escape targets. The inability of Cas3 to cleave the escape target is also consistent with the fact that the vast majority of cells will die when infected with phage bearing an escape mutation, and immunity is only conferred for those rare survivors that successfully update the CRISPR locus (Datsenko et al., 2012). Third, Cas3 loaded at escape targets exhibited only a ~6-s lifetime prior to initiating translocation, but there was no evidence for the longer-lived intermediate (τ_2) that we have ascribed to ssDNA degradation. Fourth, there was no evidence for DNA looping when Cas3 initiated translocation from the escape target, suggesting that Cas3 is more readily released from Cascade at the escape target. Together, these observations suggest that Cas1-Cas2 recruits and loads Cas3 onto the DNA flanking the escape targets through a mechanism that is distinct from the Cascade-mediated mechanism that takes place at cognate protospacers.

Primed Acquisition of New Spacer Sequences

Together, our data provide direct support for a model of primed sequence acquisition involving Cas1-Cas2-mediated recruitment of Cas3 to Cascade at escape targets, followed by ATP-dependent translocation of Cas3 along the foreign DNA (Figure 7C). Cas3 can move in either direction away from the escape target, consistent with the expectation that new spacers can be acquired from either side of an escape target (Richter et al., 2014). Translocation of Cas3 away from the escape target does not induce DNA damage, and we speculate that Cas3 may be looking for an as-yet-unidentified signal (e.g., DNA sequence, partner protein, or both) necessary to activate its nuclease activity, or the nuclease activity of a partner protein, at some distal

location. Importantly, although the tagged Cas6e subunits remain bound to the protospacer after Cas3 translocation, we do not know whether the other Cas proteins are also left behind. It is possible that Cas3 takes a subset of Cascade components while translocating along the DNA. In fact, Cas3 is naturally linked with Cse1 in a single polypeptide chain in some systems, suggesting that Cse1 may have additional downstream functions during Cas3 translocation (Westra et al., 2012b). In addition, Cas1-Cas2 are essential to process and insert new spacer sequences into the CRISPR locus (Nuñez et al., 2014; Nuñez et al., 2015), and one attractive model is that Cas1-Cas2 travel with Cas3 as part of a larger spacer acquisition complex (Figure 7C), which would allow delivery of Cas1-Cas2 to sites distal to an escape target, where they would then be able to process the DNA to promote new spacer acquisition. In support of this model, studies in the closely related type 1F CRISPR-Cas system from *Pectobacterium atrosepticum* have shown that Cas3 interacts directly with Cas1 (Richter et al., 2012).

Early models suggested Cascade might diffuse away from the escape target (Datsenko et al., 2012). However, this model was later disfavored because the distribution of new spacers acquired from a circular plasmid was inconsistent with expectations for a diffusion-based mechanism, which would predict a strong bias toward acquisition of new spacer sequences near the original protospacer (Savitskaya et al., 2013). The high processivity of *E. coli* Cas3 (~12-kbp) explains why assays using relatively small plasmids (~5-kb) fail to yield a biased distribution of newly acquired spacer sequences as predicted by the original sliding hypothesis (Heler et al., 2014; Savitskaya et al., 2013). Interestingly, the type 1F CRISPR-Cas system from *P. atrosepticum* does exhibit a biased distribution of newly acquired spacers in response to an escape mutation (Richter et al., 2014). Assuming that priming occurs through a similar mechanism for the type 1F and type 1E CRISPR-Cas systems, our model predicts that *P. atrosepticum* Cas3 is less processive that *E. coli* Cas3, explaining why spacer acquisition bias can be observed in plasmid assays for *P. atrosepticum*.

Our data demonstrate that the first Cas3 molecule recruited to cognate protospacers through the PAM-dependent pathway can translocate rapidly away from Cascade before the DNA is destroyed. Moreover, the nuclease activity of Cas3 was partially attenuated by Cas1-Cas2 at cognate protospacers, allowing iterative Cas3 firing events presumably before the eventual destruction of the R-loop. Together, these observations suggest that priming might take place even when there is no escape mutation present in the invading DNA (Figure 7B). The ability to occasionally acquire new spacers in the absence of an escape mutation may allow microbes to routinely update the CRISPR locus even before foreign genetic elements have the opportunity to evade the CRIPSR/Cas immune response by acquiring new mutations.

EXPERIMENTAL PROCEDURES

Single-Molecule Assays

DNA curtains were fabricated by electron-beam lithography as previously described (Greene et al., 2010; Sternberg et al., 2014). A lipid bilayer was then deposited on the surface of the sample chamber; the anchor points were coated with anti-digoxigenin antibodies; and the DNA was anchored to the bilayer through a biotin-streptavidin linkage. The DNA was then aligned along the leading edges of the Cr diffusion barriers and coupled to the antibody-coated anchors through the application of hydrodynamic force. Cascade single-molecule binding assays were conducted in reaction buffer containing 40 mM Tris-HCl (pH 7.4), 1 mM $MgCl_2$, 25 mM KCl, 1 mg/ml BSA, 0.8% glucose, YOYO-1, and a glucose oxidase-catalase oxygen-scavenging system. The Cas6e subunit of Cascade was expressed with an N-terminal 3xFLAG tag, and the Cascade complex was labeled with antiFLAG-coated QDs (Invitrogen) for 10 min on ice prior to use. In experiments with Cas3, the YOYO-1 dye was omitted, and the reaction buffer was supplemented to contain 2 mM $MgCl_2$, 1 mM ATP, and 20 μM $CoCl_2$. Cas3 was labeled by incubation with streptavidin-coated QDs (Invitrogen) on ice for 20 min prior to injection onto the flow cell at 4 nM final concentration. RPA-eGFP labeling of ssDNA gaps was always performed at the end of the Cas3 experiments as a check for activity. In these assays, Cas3 was flushed from the sample chamber, followed by delivery of buffer containing 100 nM RPA-eGFP. The buffer flow was then terminated, and RPA-eGFP was incubated with the DNA for 10 min prior to imaging. Buffer conditions for experiments containing Cas1-Cas2 were identical to those above. Cas1 (8 nM) and Cas2 (16 nM) were pre-incubated on ice for 20 min and then mixed with Cas3 (4 nM) for an additional 5 min before being delivered to the flow cell. All single-molecule experiments were conducted at 25°C, unless otherwise indicated, and all data were collected and analyzed as previously described (Sternberg et al., 2014).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, one table, and four movies and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.10.003.

AUTHOR CONTRIBUTIONS

S.R. and S.H.S. conceived of the project with initial input from B.W. S.R. conducted single-molecule experiments and data analysis and designed all biochemical assays. S.H.S. purified proteins and assisted with experimental design and bulk biochemical experiments. M.M. and B.G. conducted bulk biochemical assays and purified proteins. C.K.G., P.B., and B.W. assisted with preparation and biochemical analysis of Cascade and Cas3. All authors discussed the data and co-wrote the manuscript.

ACKNOWLEDGMENTS

We thank Alison Marie Smith and Kaihong Zhou for technical assistance and members of the J.A.D. and E.C.G. laboratories for their helpful discussions and critical reading of the manuscript. S.H.S. acknowledges support from the NSF and National Defense Science & Engineering Graduate Research Fellowship programs. Funding was provided by the NIH (GM074739 to E.C.G.; GM108888 to B.W.), an HHMI Early Career Scientist Award (to E.C.G.), and the NSF (MCB-1154511 to E.C.G.; MCB-1244557 to J.A.D.). J.A.D. is an HHMI Investigator.

Received: June 1, 2015
Revised: August 3, 2015
Accepted: September 9, 2015
Published: October 29, 2015

REFERENCES

Barrangou, R., and Marraffini, L.A. (2014). CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Mol. Cell *54*, 234–244.

Blosser, T.R., Loeff, L., Westra, E.R., Vlot, M., Künne, T., Sobota, M., Dekker, C., Brouns, S.J., and Joo, C. (2015). Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. Mol. Cell *58*, 60–70.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology *151*, 2551–2561.

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.

Fineran, P.C., Gerritzen, M.J., Suárez-Diez, M., Künne, T., Boekhorst, J., van Hijum, S.A., Staals, R.H., and Brouns, S.J. (2014). Degenerate target sites mediate rapid primed CRISPR adaptation. Proc. Natl. Acad. Sci. USA *111*, E1629–E1638.

Gorman, J., Plys, A.J., Visnapuu, M.L., Alani, E., and Greene, E.C. (2010). Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. Nat. Struct. Mol. Biol. *17*, 932–938.

Gorman, J., Wang, F., Redding, S., Plys, A.J., Fazio, T., Wind, S., Alani, E.E., and Greene, E.C. (2012). Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair. Proc. Natl. Acad. Sci. USA *109*, E3074–E3083.

Greene, E.C., Wind, S., Fazio, T., Gorman, J., and Visnapuu, M.L. (2010). DNA curtains for high-throughput single-molecule optical imaging. Methods Enzymol. *472*, 293–315.

Halford, S.E. (2009). An end to 40 years of mistakes in DNA-protein association kinetics? Biochem. Soc. Trans. *37*, 343–348.

Heler, R., Marraffini, L.A., and Bikard, D. (2014). Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. Mol. Microbiol. *93*, 1–9.

Hochstrasser, M.L., Taylor, D.W., Bhat, P., Guegler, C.K., Sternberg, S.H., Nogales, E., and Doudna, J.A. (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. Proc. Natl. Acad. Sci. USA *111*, 6618–6623.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. *169*, 5429–5433.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., and Koonin, E.V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct *1*, 7.

Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. *60*, 174–182.

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.

Mulepati, S., and Bailey, S. (2013). In vitro reconstitution of an Escherichia coli RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. J. Biol. Chem. *288*, 22184–22192.

Myong, S., Rasnik, I., Joo, C., Lohman, T.M., and Ha, T. (2005). Repetitive shuttling of a motor protein on DNA. Nature *437*, 1321–1325.

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.

Nuñez, J.K., Lee, A.S., Engelman, A., and Doudna, J.A. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. Nature *519*, 193–198.

Park, J., Myong, S., Niedziela-Majka, A., Lee, K.S., Yu, J., Lohman, T.M., and Ha, T. (2010). PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. Cell *142*, 544–555.

Qiu, Y., Antony, E., Doganay, S., Koh, H.R., Lohman, T.M., and Myong, S. (2013). Srs2 prevents Rad51 filament formation by repetitive motion on DNA. Nat. Commun. *4*, 2281.

Richter, C., Gristwood, T., Clulow, J.S., and Fineran, P.C. (2012). In vivo protein interactions and complex formation in the Pectobacterium atrosepticum subtype I-F CRISPR/Cas System. PLoS ONE *7*, e49549.

Richter, C., Dy, R.L., McKenzie, R.E., Watson, B.N., Taylor, C., Chang, J.T., McNeil, M.B., Staals, R.H., and Fineran, P.C. (2014). Priming in the type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. Nucleic Acids Res. *42*, 8516–8526.

Rollins, M.F., Schuman, J.T., Paulus, K., Bukhari, H.S., and Wiedenheft, B. (2015). Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from Pseudomonas aeruginosa. Nucleic Acids Res. *43*, 2216–2222.

Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M.S., Siksnys, V., and Seidel, R. (2015). Directional R-loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. Cell Rep., S2211-1247(15)00135-7.

Sashital, D.G., Wiedenheft, B., and Doudna, J.A. (2012). Mechanism of foreign DNA selection in a bacterial adaptive immune system. Mol. Cell *46*, 606–615.

Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A., and Severinov, K. (2013). High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in E. coli. RNA Biol. *10*, 716–725.

Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. USA *108*, 10098–10103.

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. EMBO J. *30*, 1335–1342.

Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, M.J., Barrangou, R., Horvath, P., and Siksnys, V. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. EMBO J. *32*, 385–394.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.

Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc. Natl. Acad. Sci. USA *111*, 9798–9803.

van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. *12*, 479–492.

von Hippel, P.H., and Berg, O.G. (1989). Facilitated target location in biological systems. J. Biol. Chem. *264*, 675–678.

Westra, E.R., Swarts, D.C., Staals, R.H., Jore, M.M., Brouns, S.J., and van der Oost, J. (2012a). The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu. Rev. Genet. *46*, 311–339.

Westra, E.R., van Erp, P.B., Künne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., et al. (2012b). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol. Cell *46*, 595–605.

Zhou, R., Zhang, J., Bochman, M.L., Zakian, V.A., and Ha, T. (2014). Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA. eLife *3*, e02190.