ARTICLE

doi:10.1038/nature13011

DNA interrogation by the CRISPR RNA-guided endonuclease Cas9

Samuel H. Sternberg[1]*, Sy Redding[2]*, Martin Jinek[3]†, Eric C. Greene[4] & Jennifer A. Doudna[1,3,5,6]

The clustered regularly interspaced short palindromic repeats (CRISPR)-associated enzyme Cas9 is an RNA-guided endonuclease that uses RNA–DNA base-pairing to target foreign DNA in bacteria. Cas9–guide RNA complexes are also effective genome engineering agents in animals and plants. Here we use single-molecule and bulk biochemical experiments to determine how Cas9–RNA interrogates DNA to find specific cleavage sites. We show that both binding and cleavage of DNA by Cas9–RNA require recognition of a short trinucleotide protospacer adjacent motif (PAM). Non-target DNA binding affinity scales with PAM density, and sequences fully complementary to the guide RNA but lacking a nearby PAM are ignored by Cas9–RNA. Competition assays provide evidence that DNA strand separation and RNA–DNA heteroduplex formation initiate at the PAM and proceed directionally towards the distal end of the target sequence. Furthermore, PAM interactions trigger Cas9 catalytic activity. These results reveal how Cas9 uses PAM recognition to quickly identify potential target sites while scanning large DNA molecules, and to regulate scission of double-stranded DNA.

RNA-mediated adaptive immune systems in bacteria and archaea rely on CRISPRs and CRISPR-associated (Cas) proteins to provide protection from invading viruses and plasmids[1]. Bacteria harbouring CRISPR-Cas loci respond to viral and plasmid challenge by integrating short fragments of the foreign nucleic acid (protospacers) into the host chromosome at one end of the CRISPR array[2]. Transcription of the CRISPR array followed by enzymatic processing yields short CRISPR RNAs (crRNAs) that direct Cas protein-mediated cleavage of complementary target sequences within invading viral or plasmid DNA[3–5]. In type II CRISPR-Cas systems, Cas9 functions as an RNA-guided endonuclease that uses a dual-guide RNA consisting of crRNA and *trans*-activating crRNA (tracrRNA) for target recognition and cleavage by a mechanism involving two nuclease active sites that together generate double-stranded DNA (dsDNA) breaks[6,7].

RNA-programmed Cas9 has proven to be a versatile tool for genome engineering in multiple cell types and organisms[8–19]. Guided by either a natural dual-RNA complex or a chimaeric single-guide RNA[6], Cas9 generates site-specific dsDNA breaks that are repaired either by non-homologous end joining or homologous recombination. In addition, catalytically inactive Cas9 alone or fused to transcriptional activator or repressor domains can be used to alter transcription levels at sites targeted by guide RNAs[20–25]. Despite the ease with which this technology can be applied, the fundamental mechanism that enables Cas9–RNA to locate specific 20-base-pair (bp) DNA targets within the vast sequence space of entire genomes remains unknown.

Single-molecule visualization of Cas9

We used a single-tethered DNA curtains assay and total internal reflection fluorescence microscopy (TIRFM) to visualize the binding site distribution of single Cas9–RNA molecules on λ-DNA substrates (48,502 bp) (Fig. 1a)[26,27]. We purified *Streptococcus pyogenes* Cas9 containing a carboxy-terminal 3×-Flag tag that enabled fluorescent labelling using anti-Flag-antibody-coated quantum dots (QDs)[27,28], and generated guide RNAs (dual crRNA–tracrRNA) bearing complementarity to six different 20-bp sites within the λ-DNA (Fig. 1b and Extended Data Table 1). Neither the 3×-Flag tag nor QDs inhibited DNA cleavage by Cas9–RNA, and all guide RNAs were functional (Extended Data Fig. 1). Initial experiments used a nuclease-inactive Cas9 containing D10A and H840A point mutations (dCas9) that binds but does not cleave DNA[6]. QD-tagged dCas9–RNA localized almost exclusively to the expected target site (Fig. 1c and Supplementary Video 1). Furthermore, Cas9 could be directed to any region of the DNA by redesigning the RNA guide sequence (Fig. 1d and Extended Data Fig. 2), as anticipated[6,8,9]. Thus, DNA targeting by Cas9–RNA is faithfully recapitulated in the DNA curtains assay.

We next used *apo*-Cas9 protein to confirm that the binding observed in DNA curtains assays was due to Cas9–RNA and not *apo*-Cas9 lacking guide RNA. Notably, *apo*-Cas9 also bound DNA but exhibited no apparent sequence specificity (Fig. 1e, f). Attempts to measure the dissociation rate of DNA-bound *apo*-Cas9 were hampered by their exceedingly long lifetimes (a lower limit of at least 45 min was calculated; Fig. 1g). Biochemical experiments revealed an upper limit of ~25 nM for the equilibrium dissociation constant (K_d) of the *apo*-Cas9–DNA complex, compared to ~0.5 nM for the Cas9–RNA complex bound to a bona fide target (Extended Data Fig. 3).

To test whether DNA-bound *apo*-Cas9 could be distinguished from Cas9–RNA, we measured the lifetime of *apo*-Cas9 on DNA curtains before and after injection of crRNA–tracrRNA or heparin. *Apo*-Cas9 rapidly dissociated from nonspecific sites in the presence of either competitor (Fig. 1g), and this result was verified with bulk biochemical assays (Extended Data Fig. 3). In contrast, target-bound Cas9–RNA was unaffected by either heparin or excess crRNA–tracrRNA (Fig. 1g and Extended Data Fig. 3). These findings show that nonspecifically bound *apo*-Cas9 has properties distinct from those of Cas9–RNA complexes bound to cognate targets.

Initially, we used catalytically inactive dCas9 to avoid DNA cleavage. Notably, experiments with wild-type Cas9 also failed to reveal DNA cleavage; rather, Cas9–RNA molecules remained bound to their target sites, yielding identical results to those obtained using dCas9–RNA (Fig. 2a).

[1]Department of Chemistry, University of California, Berkeley, California 94720, USA. [2]Department of Chemistry, Columbia University, New York, New York 10032, USA. [3]Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA. [4]Department of Biochemistry and Molecular Biophysics and Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA. [5]Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA. [6]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. †Present address: Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland.
*These authors contributed equally to this work.

©2014 Macmillan Publishers Limited. All rights reserved

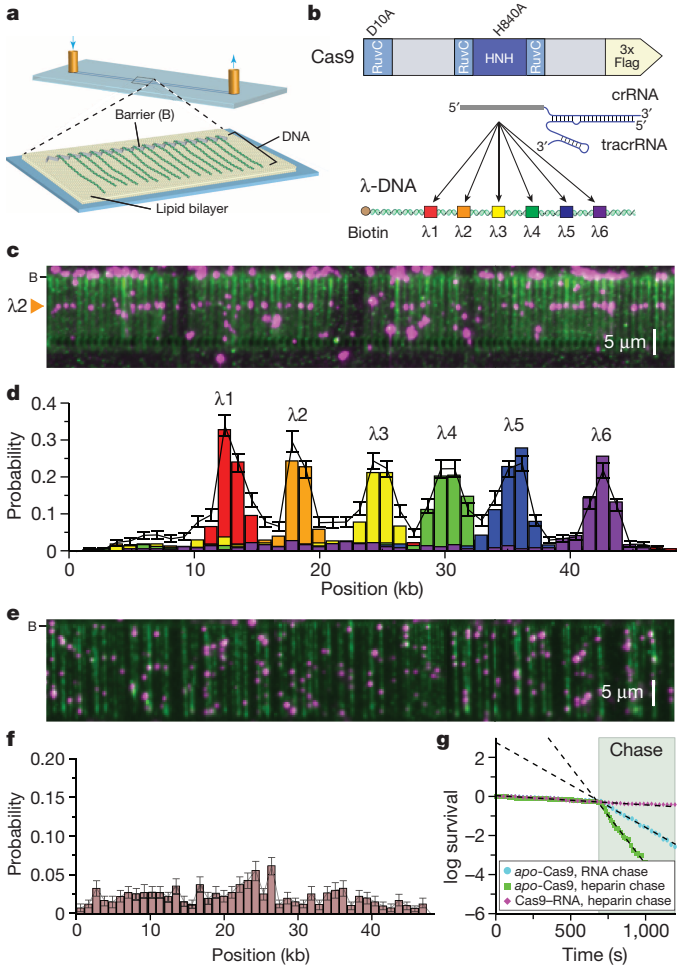


Figure 1 | DNA curtains assay for target binding by Cas9–RNA. **a**, Schematic of a single-tethered DNA curtain[26,27]. **b**, Wild-type Cas9 or dCas9 was programmed with crRNA–tracrRNA targeting one of six sites. **c**, YOYO1-stained DNA (green) bound by QD-tagged dCas9 (magenta) programmed with λ2 guide RNA. **d**, dCas9–RNA binding distributions; error bars represent 95% confidence intervals obtained through bootstrap analysis[28]. **e**, Image of *apo*-Cas9 bound to DNA curtains. **f**, Binding distribution of *apo*-Cas9 ($n = 467$); error bars represent 95% confidence intervals. **g**, Lifetimes of DNA-bound *apo*-Cas9 ($n = 205$) and Cas9–RNA ($n = 104$) after injection of 10 μg ml^{-1} heparin, and of *apo*-Cas9 ($n = 233$) after injection of 100 nM λ2 crRNA–tracrRNA.

We confirmed that the imaging conditions did not inhibit Cas9–RNA cleavage activity (Extended Data Fig. 1). These results suggested that Cas9–RNA might cleave DNA but remain tightly bound to both cleavage products, which was confirmed with biochemical gel-shift assays using 72-bp duplex DNA substrates (Extended Data Fig. 4). To determine whether stable product binding would prevent Cas9–RNA from performing multiple turnover cleavage, we conducted plasmid DNA cleavage assays at varying molar ratios of Cas9–RNA and target DNA and measured the rate and yield of product formation. Surprisingly, the amount of product rapidly plateaued at a level proportional to the molar ratio of Cas9–RNA to DNA, indicating that Cas9–RNA does not follow Michaelis–Menten kinetics (Fig. 2b). Turnover also did not occur with short duplex DNA substrates and is not stimulated by either elevated temperature or an excess of free crRNA–tracrRNA (Extended Data Fig. 5).

We next used a double-tethered DNA curtain (Fig. 2c)[29–31] to confirm that Cas9–RNA catalysed DNA cleavage in single-molecule assays. Although Cas9–RNA failed to dissociate from target sites even in the presence of 10 μg ml^{-1} heparin (Fig. 1g) or up to 0.5 M NaCl (not shown), injection of 7 M urea caused Cas9–RNA to release the downstream end containing the PAM, confirming that the DNA was cleaved at the expected target site (Fig. 2d). These findings show that Cas9–RNA remains tightly bound to both ends of the cleaved DNA, thus acting as a single-turnover enzyme.

Cas9–RNA finds targets by three-dimensional diffusion

To determine how Cas9–RNA locates DNA targets, we visualized the target search process using double-tethered DNA curtains. Site-specific DNA-binding proteins can locate target sites by three-dimensional collisions or through facilitated diffusion processes including one-dimensional sliding, hopping, and/or intersegmental transfer[32]; these mechanisms can be distinguished by single-molecule imaging[30,31,33]. For these assays, Cas9 programmed with λ2 guide RNA was injected into the sample chamber, buffer flow was terminated, and reactions were visualized in real time. These experiments revealed long-lived binding events at the target site and transient binding events at other sites on the DNA (Fig. 2e, f). We saw no evidence of Cas9–RNA associating with target sites by facilitated diffusion (either one-dimensional sliding and/or hopping); instead, all target association appeared to occur directly through three-dimensional collisions (Fig. 2e and Supplementary Video 2).

The shorter-lived, nonspecific binding events exhibited complex dissociation kinetics; the simplest model that describes the data is a double-exponential decay with lifetimes of ~3.3 and ~58 s (at 25 mM KCl) (Fig. 2f). These lifetimes were readily distinguished from the long lifetime of *apo*-Cas9 (Fig. 2f, inset). Furthermore, the experiments were conducted in the presence of a tenfold molar excess of crRNA–tracrRNA to exclude contamination from *apo*-Cas9. This result indicates that at least two and possibly more binding intermediates exist on the pathway towards cognate target recognition. Although nonspecific DNA binding typically involves electrostatic interactions with the phosphate backbone such that nonspecific lifetimes tend to decrease with increasing ionic strength[34], the lifetimes of Cas9–RNA bound at nonspecific DNA sites were not appreciably affected by salt concentration (Fig. 2f). One remarkable implication of this finding is that these Cas9–RNA non-target binding events have characteristics more commonly attributed to site-specific association[34,35].

To gain further insight into the nature of the target search mechanism, we measured the locations and corresponding lifetimes of all binding events (Fig. 2g). The off-target binding lifetime distributions did not vary substantially at different regions of the DNA. However, the number of observed binding events was not uniformly distributed along the substrate, indicating that some underlying feature of the λ-DNA might influence the target search. The λ-phage genome contains 5,677 PAM sites (~1 PAM per 8.5 bp), but it also has an unusual polar distribution of (A+T)- and (G+C)-rich sequences[27], which leads to an asymmetric distribution of PAMs (5′-NGG-3′ for *S. pyogenes* Cas9) (Fig. 2g). Pearson correlation analysis revealed that the Cas9–RNA binding site distribution was positively correlated with the PAM distribution ($r = 0.59$, $P < 0.05$) (Fig. 2h). When we repeated this experiment using a guide RNA having no complementary target sites within λ-DNA (spacer 2 crRNA)[6], we saw no change in the binding lifetimes and an even stronger correlation with the PAM distribution (Fig. 2h and Extended Data Fig. 2b, c). These results, together with the insensitivity of short-lived binding events to ionic strength, suggested that Cas9–RNA might bind specifically to PAMs and minimize interactions with non-PAM DNA while searching for potential targets.

A PAM is required for DNA interrogation

To test whether Cas9–RNA uses PAM recognition as an obligate precursor to interrogation of flanking DNA for potential guide-RNA complementarity, we used competition assays to monitor the rate of Cas9–RNA-mediated DNA cleavage (Fig. 3a, b). From these data we extracted the average amount of time that Cas9–RNA spends sampling each competitor DNA before locating and cleaving a radiolabelled substrate (Extended Data Fig. 6). In control experiments, reaction kinetics were not perturbed by the presence of an unlabelled competitor DNA lacking PAMs and bearing no sequence relationship to the crRNA,

©2014 Macmillan Publishers Limited. All rights reserved

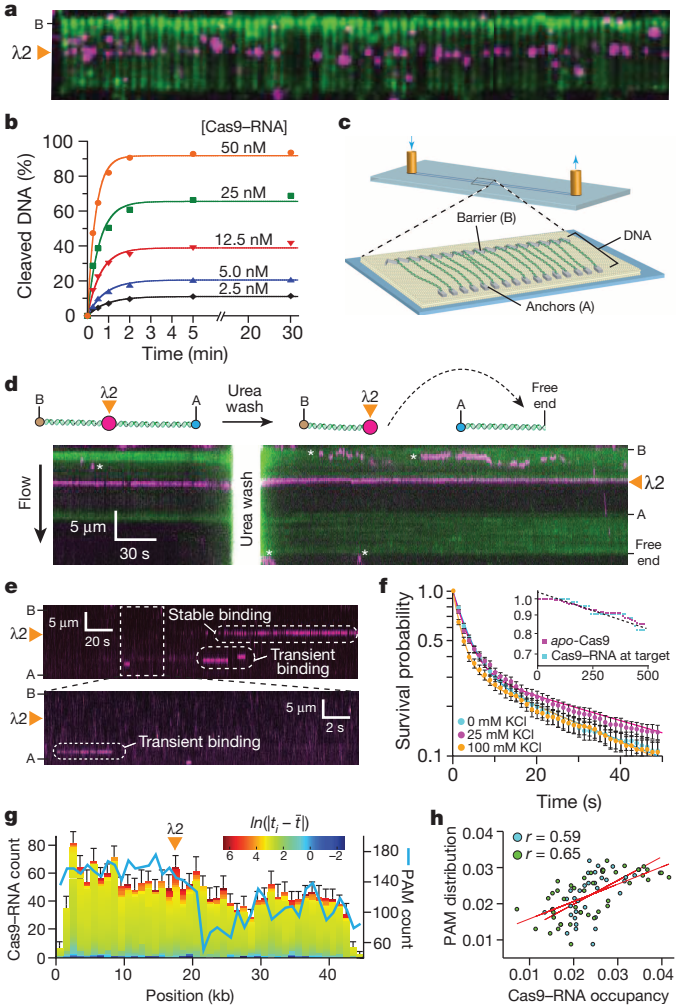


Figure 2 | Cas9–RNA remains bound to cleaved products and localizes to PAM-rich regions during the target search. **a**, Wild-type Cas9–RNA bound to DNA curtains. **b**, Cleavage yield of 25 nM plasmid DNA is proportional to [Cas9–RNA]. **c**, Schematic of a double-tethered DNA curtain[29–31]. **d**, Liberation of the cleaved DNA with 7 M urea; asterisks denote QDs that are attached to the lipid bilayer but not bound to the DNA. **e**, Kymographs illustrating distinct binding events. **f**, Survival probabilities (mean ± 70% confidence intervals) for non-target binding events. Solid lines represent double-exponential fits; $n = 502$ (0 mM KCl), $n = 632$ (25 mM KCl), and $n = 504$ (100 mM KCl). Inset: survival probabilities of DNA-bound *apo*-Cas9 ($n = 233$) and target DNA-bound Cas9–RNA ($n = 104$). **g**, Distribution of Cas9–RNA binding events ($n = 2{,}330$) and PAM sites; error bars represent 95% confidence intervals. Colour-coding reflects the binding dwell time (t_i) relative to the mean dwell time ($\bar{t}$). **h**, Correlation of PAM distribution and non-target Cas9–RNA binding distribution for λ2 (blue) and spacer 2 (green) guide RNAs.

whereas a competitor containing a PAM and fully complementary target sequence substantially reduced the cleavage rate of the radiolabelled substrate (Fig. 3b).

A series of competitors were tested that bore no crRNA guide sequence complementarity (Extended Data Table 1) but contained increasing numbers of PAMs (Fig. 3c). There was a direct correspondence between the number of PAMs and the ability of a DNA competitor to interfere with target cleavage, indicating that the lifetime of Cas9–RNA on competitor DNA increased with PAM density (Fig. 3c). This result persisted over a range of competitor DNA concentrations (Extended Data Fig. 6), and the same pattern of competition was observed for DNA binding by dCas9–RNA (Extended Data Fig. 7). These results demonstrate that the residence time of Cas9–RNA on non-target DNA lacking PAMs is negligible, and support the hypothesis that transient, non-target DNA binding events observed on the DNA curtains probably occurred at PAM sequences. Although Cas9–RNA complexes undoubtedly sample DNA lacking PAMs, these rapid binding events are neither detectable in single-molecule and bulk binding experiments (Extended Data Fig. 7), nor do they appreciably influence overall reaction kinetics in bulk biochemical assays.

We repeated the competition assay with a competitor bearing perfect complementarity to the crRNA, but with a single point mutation in the adjacent PAM (5′-TCG-3′)—like similarly mutated substrates[6], this competitor cannot be cleaved by Cas9–RNA (Extended Data Fig. 7). This competitor failed to inhibit DNA cleavage by Cas9–RNA and behaved comparably to the non-target competitor DNA lacking PAMs, despite the fact that it contained perfect complementarity to the crRNA (Fig. 3c). Together, these results demonstrate that PAM recognition is an obligate first step during target recognition by Cas9–RNA, as previously proposed[6].

Mechanism of RNA–DNA heteroduplex formation

After PAM recognition, Cas9–RNA must destabilize the adjacent duplex and initiate strand separation to enable base pairing between the target DNA strand and the crRNA guide sequence. Because Cas9 has no energy-dependent helicase activity, the mechanism of local DNA unwinding has been enigmatic, but must rely on thermally available energy. One possibility is that PAM binding could induce a general destabilization of the duplex along the length of the entire target sequence, leading to random nucleation of the RNA–DNA heteroduplex (Fig. 3d, top). Alternatively, PAM binding may cause only local melting of the duplex, with the RNA–DNA heteroduplex nucleating at the 3′ end of the target sequence next to the PAM and proceeding sequentially towards the distal 5′ end of the target sequence (Fig. 3d, bottom).

To distinguish between these models, we conducted cleavage assays with a panel of DNA competitors in which the length and position of complementarity to the guide RNA was systematically varied (Extended Data Table 1). These competitors were designed to distinguish between the random nucleation and sequential unwinding models for heteroduplex formation based on the predicted patterns of cleavage inhibition for each model (Fig. 3d). The ability of a competitor DNA to inhibit substrate cleavage by Cas9–RNA increased as the extent of complementarity originating at the 3′ end of the target sequence adjacent to the PAM increased (Fig. 3e). Inhibition increased markedly when 12 or more base pairs were complementary to the crRNA guide sequence, which agrees with the requirement for an 8–12-nucleotide seed sequence for the Cas9–RNA DNA cleavage reaction[6,36]. Notably, although competitors containing mismatches to the crRNA at the 5′ end of the target sequence competed effectively for Cas9–RNA binding, competitors containing mismatches to the crRNA at the extreme 3′ end immediately adjacent to the PAM were completely inert to binding (Fig. 3e). This was true even with a 2-bp mismatch followed by 18 bp of contiguous sequence complementarity to the crRNA. Therefore, when mismatches to the crRNA are encountered within the first two nucleotides of the target sequence, Cas9–RNA loses the ability to interrogate and recognize the remainder of the DNA. The pattern of inhibition observed with the different competitor DNAs indicates that sequence homology adjacent to the PAM is necessary to initiate target duplex unwinding until the reaction has proceeded sufficiently far (~12 bp, approximately one turn of an A-form RNA–DNA helix), such that the energy necessary for further propagation of the RNA–DNA heteroduplex falls below the energy needed for the reverse reaction. These findings indicate that formation of the RNA–DNA heteroduplex initiates at the PAM and proceeds through the target sequence by a sequential, stepwise unwinding mechanism consistent with a Brownian ratchet[37].

As a further test of this model, we used a DNA competitor that contained mismatches to the crRNA at positions 1–2 but was itself mismatched at the same two positions, forming a small bubble in the duplex. Despite the absence of sequence complementarity to the crRNA within

©2014 Macmillan Publishers Limited. All rights reserved

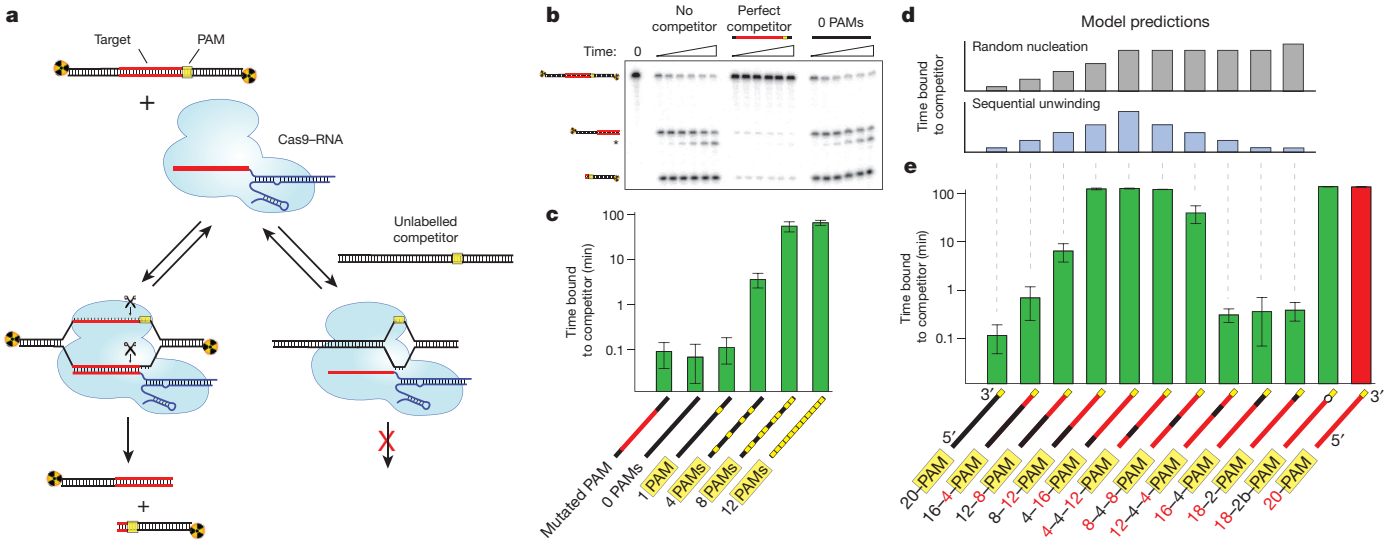


Figure 3 | Cas9–RNA searches for PAMs and unwinds dsDNA in a directional manner. **a**, Schematic of the competition cleavage assay. **b**, Cleavage assay with and without competitor DNAs. **c**, Quantification of competition data (mean ± s.d.). Competitor cartoon representations show PAMs (yellow) and regions complementary to the crRNA (red). **d**, Predicted data trends for the random nucleation or sequential unwinding models aligned with the corresponding data in **e**. **e**, Competition assays using substrates with variable degrees of crRNA complementarity, shown as in **c**. Numeric descriptions of the competitor DNAs indicate the regions of complementarity (red) or mismatches (black) to the crRNA sequence.

the DNA bubble, this substrate was a robust competitor and bound Cas9–RNA with an affinity nearly indistinguishable from that of a perfect substrate (Fig. 3e and Extended Data Fig. 8). Notably, this DNA could also be cleaved with near-wild-type rates (Extended Data Fig. 8). We speculate that the presence of the DNA bubble allowed Cas9–RNA to bypass the mismatches and reinitiate nucleation of the RNA–DNA heteroduplex downstream of the bubble, thereby propagating strand separation through the remainder of the target.

The PAM triggers Cas9 nuclease activity

The results presented above indicate that PAM recognition has a central role in target recognition, and that introduction of a small bubble in

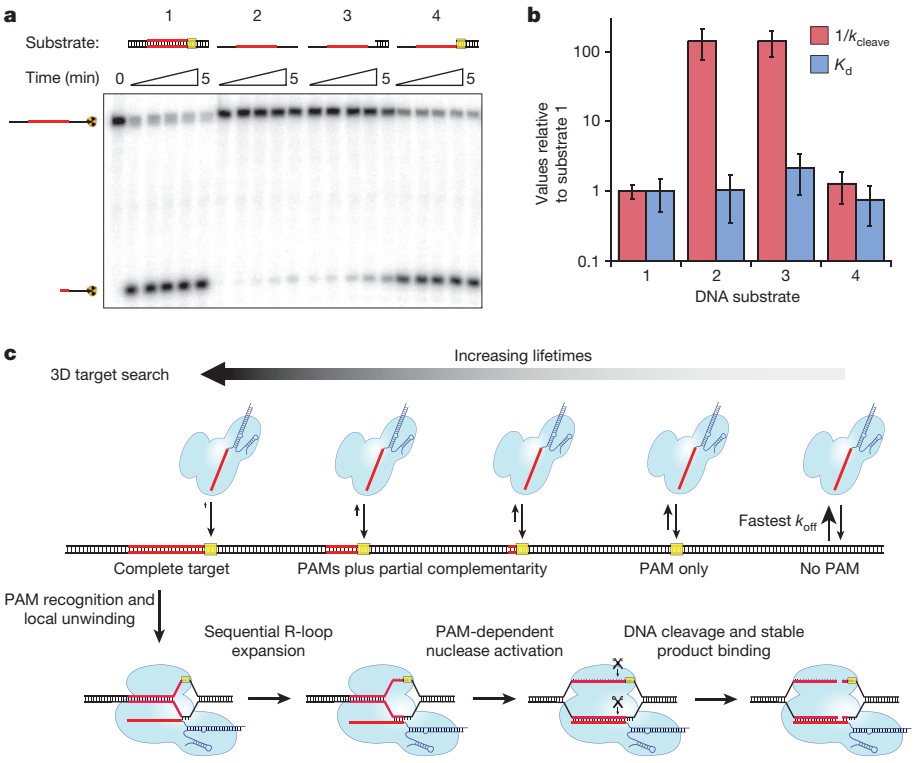


Figure 4 | PAM recognition regulates Cas9 nuclease activity. **a**, Cleavage assay with single-stranded, double-stranded and partially double-stranded substrates. **b**, Relative affinities and cleavage rates (mean ± s.d.). **c**, Model for target search, recognition and cleavage by Cas9–RNA. The search initiates through random three-dimensional collisions. Cas9–RNA rapidly dissociates from non-PAM DNA, but binds PAMs for longer times and samples adjacent DNA for guide RNA complementarity, giving rise to a heterogeneous population of intermediates. At correct targets, Cas9–RNA initiates formation of an RNA–DNA heteroduplex, and R-loop expansion propagates via sequential unwinding. The DNA is cleaved, and Cas9–RNA remains bound to the cleaved products.

©2014 Macmillan Publishers Limited. All rights reserved

the DNA target eliminates the need for RNA–DNA heteroduplex formation immediately adjacent to the PAM. One might expect PAM recognition to be dispensable for Cas9–RNA-mediated recognition and cleavage of a single-stranded DNA (ssDNA) target. Surprisingly, however, a ssDNA substrate was cleaved more than two orders of magnitude slower than a dsDNA substrate (Fig. 4a, b), despite the fact that dCas9–RNA bound both dsDNA and ssDNA substrates with similar affinities (Fig. 4b and Extended Data Fig. 9).

Importantly, Cas9–RNA recognizes the 5′-NGG-3′ PAM on the non-target DNA strand[6], so the ssDNA substrates did not contain a PAM but rather the complement to the PAM sequence. We proposed that the absence of the PAM on the ssDNA might explain why an otherwise fully complementary target is resistant to cleavage. To test this possibility, we prepared hybrid substrates with varying lengths of dsDNA at the 3′ flanking sequence (Fig. 4a). Cleavage assays revealed that the ssDNA target strand could be activated for cleavage in the presence of flanking dsDNA that extended across the PAM sequence, but that this activating effect was lost when the dsDNA was truncated immediately before the PAM (Fig. 4a, b and Extended Data Fig. 9). Gel-shift experiments confirmed that these results were not a consequence of discrimination at the level of binding (Fig. 4b). Rather, the presence of the 5′-NGG-3′ PAM on the non-target strand was critical for some step of the reaction that occurred after binding. These data suggest that the PAM acts as an allosteric regulator of Cas9–RNA nuclease activity.

Discussion

Our results suggest a model for target binding and cleavage by Cas9–RNA involving an unanticipated level of importance for PAM sequences at each stage of the reaction (Fig. 4c). Although details may differ, we propose that PAM interactions may function similarly for other CRISPR RNA-guided surveillance complexes[38–44]. The Cas9–RNA target search begins with random collisions with DNA. However, rather than sampling all DNA equivalently, Cas9–RNA accelerates the search by rapidly dissociating from non-PAM sites, thereby reducing the amount of time spent at off-targets. Only upon binding to a PAM site does Cas9–RNA interrogate the flanking DNA for guide RNA complementarity, as was previously proposed for Cas9 (ref. 6) and a distinct CRISPR RNA-guided complex (Cascade)[42]. A requirement for initial PAM recognition also eliminates the potential for self-targeting, as perfectly matching targets within the bacterial CRISPR locus are not flanked by PAMs. Our results indicate that PAM recognition coincides with initial destabilization of the adjacent sequence, as demonstrated by experiments using a bubble-containing DNA substrate, followed by sequential extension of the RNA–DNA heteroduplex. This mechanism explains the emergence of seed sequences, because mismatches encountered early in a directional melting-in process would prematurely abort target interrogation. Moreover, the complex dissociation kinetics observed on non-target λ-DNA would arise from heterogeneity in the potential target sites as Cas9–RNA probes sequences adjacent to PAMs for guide RNA complementarity. Binding to a correct target then leads to activation of both nuclease domains. This step also requires PAM recognition, providing an unanticipated level of PAM-dependent regulation that may ensure further protection against self-cleavage of the CRISPR locus. Notably, Cas9–RNA does not dissociate from the cleaved DNA except under extremely harsh conditions, indicating that Cas9–RNA may remain bound to the cleaved site *in vivo*[4] and require other cellular factors to promote recycling. Finally, our data indicate that efforts to minimize off-target effects during genome engineering using Cas9–RNA complexes need only consider off-targets adjacent to a PAM, because potential targets lacking a PAM are unlikely to be interrogated[36,45–47].

METHODS SUMMARY

Cas9 from *S. pyogenes* containing a C-terminal 3×-Flag tag was purified as described[6]. crRNAs and tracrRNAs were either transcribed *in vitro* with T7 polymerase or made synthetically. Cas9–RNA complexes for single-molecule experiments were reconstituted by incubating Cas9 and a 10× molar excess of crRNA–tracrRNA in reaction buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 5% glycerol, 1 mM dithiothreitol (DTT)) for ~10 min at 37 °C, before injecting 1–2 nM into the flow cell. Single-tethered DNA curtains were prepared and TIRFM conducted as described[26,27]. dCas9–RNA positions were determined by fitting a two-dimensional Gaussian to individual molecules. Double-tethered DNA curtains were prepared as described[31,33], and position and lifetime measurements were determined from kymographs generated for each DNA molecule. Duplex DNA substrates used in biochemical experiments were prepared from synthetic oligonucleotides (Integrated DNA Technologies), and their sequences can be found in Extended Data Table 1. Bulk competition cleavage assays were conducted at room temperature in reaction buffer and contained ~1 nM radiolabelled λ1 target DNA, 10 nM Cas9–RNA complex and 500 nM competitor DNA. Aliquots were removed at each time point and quenched by the addition of gel loading buffer supplemented with 50 mM EDTA. Reaction products were resolved by 10% denaturing polyacrylamide gel electrophoresis, visualized by phosphorimaging, and quantified using ImageQuant (GE Healthcare). Competition experiments were analysed to determine the survival probability of the target DNA. After determining the conditional survival probability for each time course, the change in survival probability relative to a reaction in the absence of competitor was calculated. Integration of the resulting curve over the total reaction time yielded the mean time spent on competitor DNA for each experiment.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 September; accepted 23 December 2013.
Published online 29 January 2014.

1. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482,** 331–338 (2012).
2. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
3. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).
4. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468,** 67–71 (2010).
5. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607 (2011).
6. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–821 (2012).
7. 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. USA* **109,** E2579–E2586 (2012).
8. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339,** 823–826 (2013).
9. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–823 (2013).
10. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2,** e00471 (2013).
11. Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnol.* **31,** 227–229 (2013).
12. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153,** 910–918 (2013).
13. Bassett, A. R., Tibbit, C., Ponting, C. P. & Liu, J.-L. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 System. *Cell Rep.* **4,** 220–228 (2013).
14. Gratz, S. J. *et al.* Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194,** 1029–1035 (2013).
15. Friedland, A. E. *et al.* Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nature Methods* **10,** 741–743 (2013).
16. Shan, Q. *et al.* Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature Biotechnol.* **31,** 686–688 (2013).
17. Li, J.-F. *et al.* Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nature Biotechnol.* **31,** 688–691 (2013).
18. Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D. G. & Kamoun, S. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nature Biotechnol.* **31,** 691–693 (2013).
19. Xie, K. & Yang, Y. RNA-guided genome editing in plants using A CRISPR-Cas system. *Mol. Plant* **6,** 1975–1983 (2013).
20. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152,** 1173–1183 (2013).
21. Bikard, D. *et al.* Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41,** 7429–7437 (2013).
22. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154,** 442–451 (2013).
23. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nature Methods* **10,** 977–979 (2013).
24. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods* **10,** 973–976 (2013).

©2014 Macmillan Publishers Limited. All rights reserved

25. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnol.* **31,** 833–838 (2013).
26. Fazio, T., Visnapuu, M.-L., Wind, S. & Greene, E. C. DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging. *Langmuir* **24,** 10524–10531 (2008).
27. Visnapuu, M.-L. & Greene, E. C. Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition. *Nature Struct. Mol. Biol.* **16,** 1056–1062 (2009).
28. Finkelstein, I. J., Visnapuu, M.-L. & Greene, E. C. Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature* **468,** 983–987 (2010).
29. Gorman, J., Fazio, T., Wang, F., Wind, S. & Greene, E. C. Nanofabricated racks of aligned and anchored DNA substrates for single-molecule imaging. *Langmuir* **26,** 1372–1379 (2010).
30. Gorman, J., Plys, A. J., Visnapuu, M.-L., Alani, E. & Greene, E. C. Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nature Struct. Mol. Biol.* **17,** 932–938 (2010).
31. Wang, F. *et al.* The promoter-search mechanism of *Escherichia coli* RNA polymerase is dominated by three-dimensional diffusion. *Nature Struct. Mol. Biol.* **20,** 174–181 (2013).
32. von Hippel, P. H. & Berg, O. G. Facilitated target location in biological systems. *J. Biol. Chem.* **264,** 675–678 (1989).
33. Gorman, J. *et al.* Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair. *Proc. Natl Acad. Sci. USA* **109,** E3074–E3083 (2012).
34. von Hippel, P. H. & Berg, O. G. On the specificity of DNA-protein interactions. *Proc. Natl Acad. Sci. USA* **83,** 1608–1612 (1986).
35. Rohs, R. *et al.* Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* **79,** 233–269 (2010).
36. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnol.* **31,** 233–239 (2013).
37. Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **438,** 460–465 (2005).
38. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733–740 (2009).
39. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463,** 568–571 (2010).
40. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl Acad. Sci. USA* **108,** 10098–10103 (2011).
41. Wiedenheft, B. *et al.* RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl Acad. Sci. USA* **108,** 10092–10097 (2011).
42. Sashital, D. G., Wiedenheft, B. & Doudna, J. A. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* **46,** 606–615 (2012).
43. Hou, Z. *et al.* Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl Acad. Sci. USA* **110,** 15644–15649 (2013).
44. Esvelt, K. M. *et al.* Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nature Methods* **10,** 1116–1121 (2013).
45. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnol.* **31,** 839–843 (2013).
46. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnol.* **31,** 827–832 (2013).
47. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnol.* **31,** 822–826 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank P. Bhat, A. Smith and K. Zhou for technical assistance, and members of the Doudna and Greene laboratories and J. Cate for discussions and critical reading of the manuscript. S.H.S. acknowledges support from the National Science Foundation and National Defense Science & Engineering Graduate Research Fellowship programs. Funding was provided by the National Institutes of Health (GM074739 to E.C.G.) and the National Science Foundation (MCB-1154511 to E.C.G. and MCB-1244557 to J.A.D.). M.J. was a Research Specialist, E.C.G. is an Early Career Scientist, and J.A.D. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions S.H.S. generated RNAs, conducted biochemical and single-molecule experiments, and assisted with single-molecule data analysis. S.R. conducted single-molecule experiments and data analysis, and assisted with the design and analysis of biochemical assays. M.J. cloned and purified Cas9, and assisted with the design and interpretation of initial single-molecule experiments. S.H.S., S.R., M.J., E.C.G. and J.A.D. discussed the data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of the paper. The authors declare competing financial interests: details are available in the online version of the paper. Correspondence and requests for materials should be addressed to J.A.D. (doudna@berkeley.edu) or E.C.G. (ecg2108@columbia.edu).

©2014 Macmillan Publishers Limited. All rights reserved

METHODS

Cas9 and RNA preparation. Wild-type Cas9 and D10A/H840A dCas9 from *S. pyogenes* were purified as described[6], and a 3×-Flag tag was cloned onto the C terminus of Cas9 for single-molecule experiments. crRNAs (42 nucleotides in length) were either ordered synthetically (Integrated DNA Technologies) or transcribed *in vitro* with T7 polymerase using single-stranded DNA templates (Extended Data Table 1), as described[48]. tracrRNA was also transcribed *in vitro* and contained nucleotides 15–87 following the numbering scheme used previously[6]. crRNA–tracrRNA duplexes were prepared by mixing equimolar concentrations of each RNA in hybridization buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$), heating to 95 °C for 30 s, and slow cooling.

DNA curtains post-steady-state binding measurements. Post-steady-state binding assays in Fig. 1 were performed with single-tethered DNA curtains[26,27]. First, 100 nM 3×-Flag-tagged dCas9 was reconstituted with 1 μM crRNA–tracrRNA targeting the desired region of λ-DNA by incubating for ~10 min at 37 °C in reaction buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 5% glycerol, 1 mM dithiothreitol (DTT)). 10 nM dCas9–RNA was then incubated with λ-DNA (100 pM) for ~15 min at 37 °C in 40 mM Tris-HCl pH 7.5, 25 mM KCl, 1 mg ml^{-1} BSA, 1 mM $MgCl_2$ and 1 mM DTT, before being diluted to 1 nM and injected into the flow cell. The flow cell was then washed with 3–5 ml of imaging buffer containing 40 mM Tris-HCl, 25 mM KCl, 1 mg ml^{-1} BSA, 1 mM $MgCl_2$, 1 mM DTT, 0.75 nM YOYO1 (Life Technologies), 0.8% glucose, and 0.2× glucose oxidase/catalase. Finally, 0.5 nM anti-Flag antibody-coated QDs were incubated in the flow cell for 5 min, followed by a wash of 1–2 ml of imaging buffer. Curtains were imaged and dCas9–RNA positions determined by fitting a two-dimensional Gaussian to individual molecules[31,33]. The data from all six dCas9–RNA complexes (λ1–λ6) were combined, and error bars for the combined set were generated by bootstrap methods[31,33].

DNA curtains equilibrium binding measurements. Binding position and lifetime measurements in Fig. 2 were performed using the λ2 crRNA–tracrRNA and double-tethered DNA curtains[31,33]. Cas9 was reconstituted with a 10× excess of crRNA–tracrRNA and incubated with anti-Flag antibody-coated QDs for ~10 min. Cas9–RNA was then diluted to 2 nM in imaging buffer containing 0–100 mM KCl as indicated, and injected into the flow cell. The approximate ionic strength for imaging buffer containing 0, 25 and 100 mM KCl is 32, 57 and 132 mM, respectively, given the expected ionization of Tris-HCl at pH 7.5 and presence of 1 mM $MgCl_2$. Videos were recorded at 50, 25, or 10 Hz, and the position of each binding event was determined from the *y*-coordinate within kymographs generated for each DNA molecule. The lifetime of each binding event was defined as the difference between the first frame and last frame in which the QD-tagged Cas9–RNA was observed. To analyse lifetimes, all binding events were synchronized, and the probability that a binding event survived up to a particular time was determined as the number of Cas9–RNA complexes bound at time *t* divided by the number initially bound. Position data were pooled to generate the binding distribution histogram, which was binned at 1,078 bp per bin[27]. Error bars for the binding distributions and survival probabilities were determined by bootstrap methods[31,33].

To test for cleavage in our single-molecule assays, Cas9–RNA bound to double-tethered DNA was exposed to a 7 M urea wash in the presence of flow (Fig. 2d). Similar experiments were conducted on single-tethered curtains in the absence of YOYO1, and 42% ($n = 150$) of target-bound molecules remained bound to the upstream DNA product after the urea wash. The remaining 58% ($n = 206$) either remained bound to the downstream product containing the PAM, dissociated from the DNA altogether, or remained bound to the upstream fragment but lost their florescent tag.

Bulk binding and cleavage experiments. The plasmid DNA substrate used in Fig. 2b contained a λ2 target sequence cloned into the EcoRI and BamHI sites on pUC19. Oligoduplex DNA substrates (Extended Data Table 1) were 55 bp in length and were prepared by mixing together complementary synthetic oligonucleotides (Integrated DNA Technologies) in hybridization buffer, heating to 95 °C for 1–2 min followed by slow-cooling, and purifying on a 5% native polyacrylamide gel (0.5× TBE buffer with 5 mM $MgCl_2$) run at 4 °C. When assayed directly, DNA substrates were 5′-radiolabelled using [γ-^{32}P]-ATP (Promega) and T4 polynucleotide kinase (New England Biolabs) or 3′-radiolabelled using [α-^{32}P]-dATP (Promega) and terminal transferase (New England Biolabs) (Extended Data Fig. 4). The substrates presented in Fig. 4b were prepared by 5′-radiolabelling only the target strand, hybridizing it to a 10× excess of the indicated unlabelled complementary strand, and gel purifying the partial/full duplex by 10% native gel electrophoresis.

Cas9–RNA complexes were reconstituted before cleavage and binding experiments by incubating Cas9 and the crRNA–tracrRNA duplex for 10 min at 37 °C in reaction buffer. Binding experiments used dCas9 (except as indicated in Extended Data Fig. 4) and either equimolar crRNA–tracrRNA or a 10× molar excess of crRNA–tracrRNA over dCas9 (Extended Data Figs 3 and 8). Binding reactions contained 0.1–1 nM DNA and increasing *apo*-dCas9 or dCas9–RNA concentrations, and were incubated at 37 °C for 1 h before being resolved by 5% native polyacrylamide gel electrophoresis (0.5× TBE buffer with 5 mM $MgCl_2$) run at 4 °C. DNA was visualized by phosphorimaging, quantified with ImageQuant (GE Healthcare), and analysed with Kaleidagraph (Synergy Software).

Cleavage assays were conducted in reaction buffer at room temperature and analysed by 1% agarose gel electrophoresis and ethidium bromide staining (Fig. 2b) or 10% denaturing polyacrylamide gel electrophoresis and phosphorimaging. Aliquots were removed at each time point and quenched by the addition of gel loading buffer containing 50 mM EDTA (25 mM at 1×). Reactions contained ~1 nM radiolabelled DNA substrate and 10 nM Cas9–RNA (competition experiments, Fig. 3b, c, e and Extended Data Fig. 6) or 100 nM Cas9–RNA (Fig. 4a, b and Extended Data Figs 1, 7, 8 and 9). Competition experiments used λ1 target DNA and were supplemented with 500 nM unlabelled competitor DNAs (Fig. 3c, e) or an extended concentration range of competitor DNAs (Extended Data Fig. 6). All oligoduplex DNA cleavage experiments were visualized by phosphorimaging and quantified with ImageQuant (GE Healthcare).

Analysis of cleavage competition assays. The competition experiments were analysed to determine the survival probability of the radiolabelled target DNA, $S(t)$. In principle, the survival probability should begin at 1 and go to 0, but in practice the reaction rarely proceeds to completion. Therefore, we conditioned the survival against the probability of a particular amount of target DNA being cleaved. This conditional survival probability, $S^*(t)$, relates to the survival probability as follows:

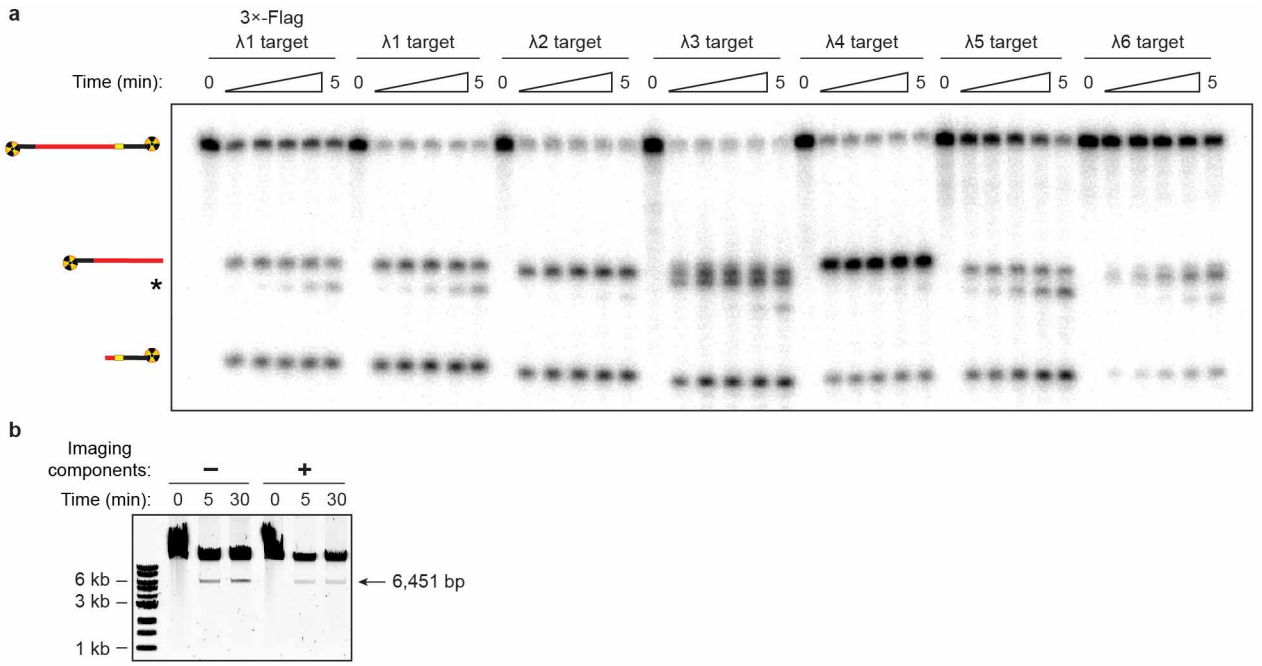
$$S^*(t) = \frac{S(t) - S(\infty)}{1 - S(\infty)}$$

All reactions in the presence of competitor DNA that reached ~90% completion were conditioned against their final values, whereas reactions uncompleted after 2 h were conditioned to the reaction in the absence of competitor DNA. For each reaction, we then obtained the change in the survival probability of the target DNA, $\Delta P_s(t)$, in the presence of competitor DNA. Finally, $\Delta P_s(t)$ was integrated over the 2-h reaction time. For reactions that reached completion in the presence of competitor within 2 h, this analysis yields the change in the mean relaxation time of the reaction (or the inverse of the average reaction rate). In cases where the reaction did not reach completion by the 2-h time point, this analysis instead yields a mean time spent on competitor DNA during the 2 h reaction. Notably, this analysis makes no assumptions about the nature of the reaction or the dynamic changes in the reactive species.

The reduction in cleavage rate in the presence of competitor DNA is directly proportional to the time that Cas9–RNA spends bound to each competitor. In each reaction, Cas9–RNA encounters competitor DNA on average more frequently than the target DNA, and the time Cas9–RNA spends interrogating a competitor has the cumulative effect of slowing the overall reaction. The presented models merely state that the amount of time spent on competitor DNA will be proportional to the 'observed' complementarity between crRNA and bound DNA, that is, the number of canonical Watson–Crick base pairs that can be formed. It then directly follows that in the case where the R-loop is randomly nucleated (regardless of nucleation size), the time bound to competitor DNA will simply scale with the total amount of complementarity between competitor DNA and crRNA (Fig. 3d, top). However, in the case where the R-loop is nucleated from a particular site, that is, the 3′ end of the target sequence directly adjacent to the PAM, the time bound to competitor DNA will scale proportionally to the length of contiguous complementarity between the crRNA and DNA beginning from the nucleation site (Fig. 3d, bottom).

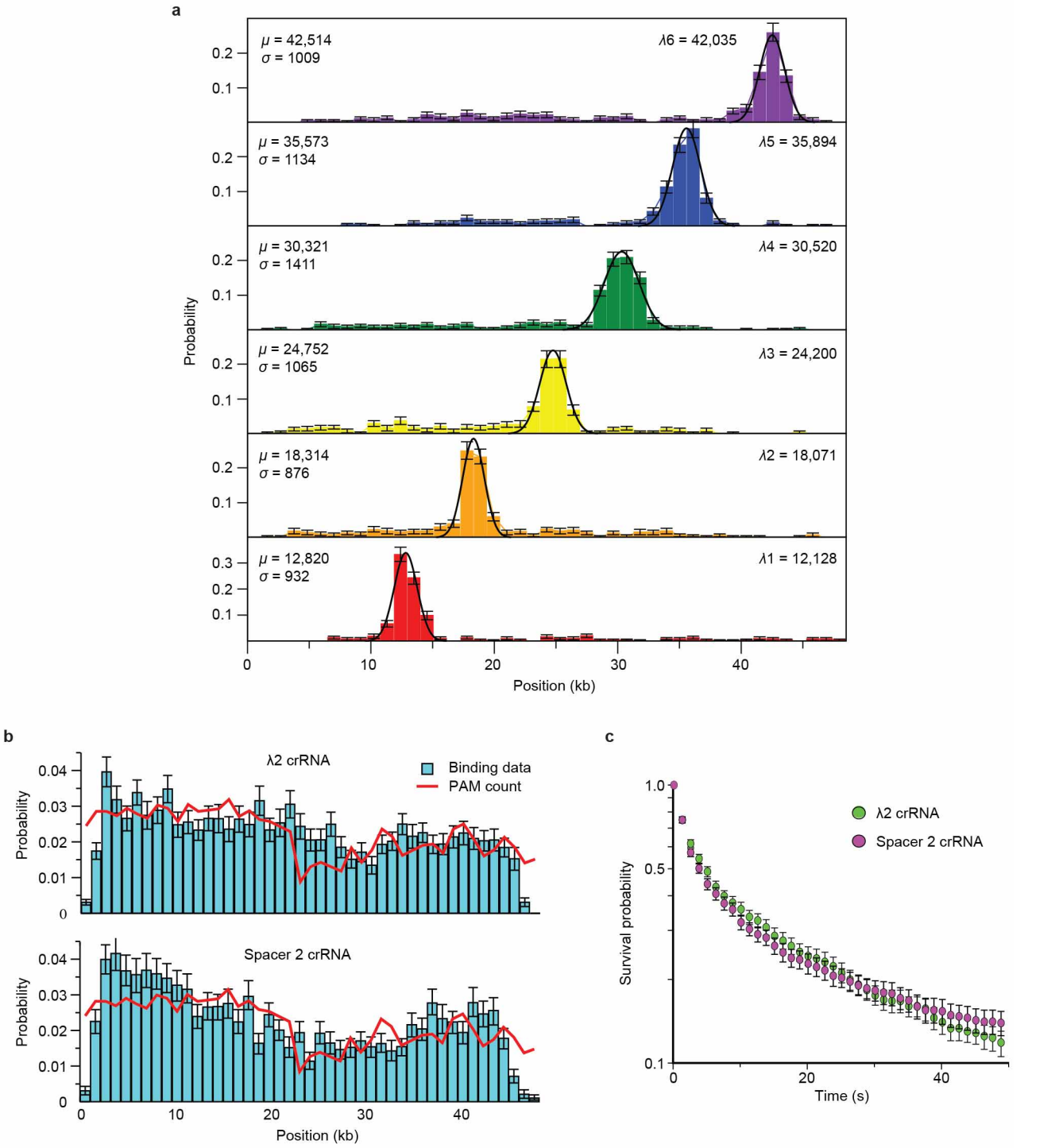
48. Sternberg, S. H., Haurwitz, R. E. & Doudna, J. A. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* **18,** 661–672 (2012).

©2014 Macmillan Publishers Limited. All rights reserved



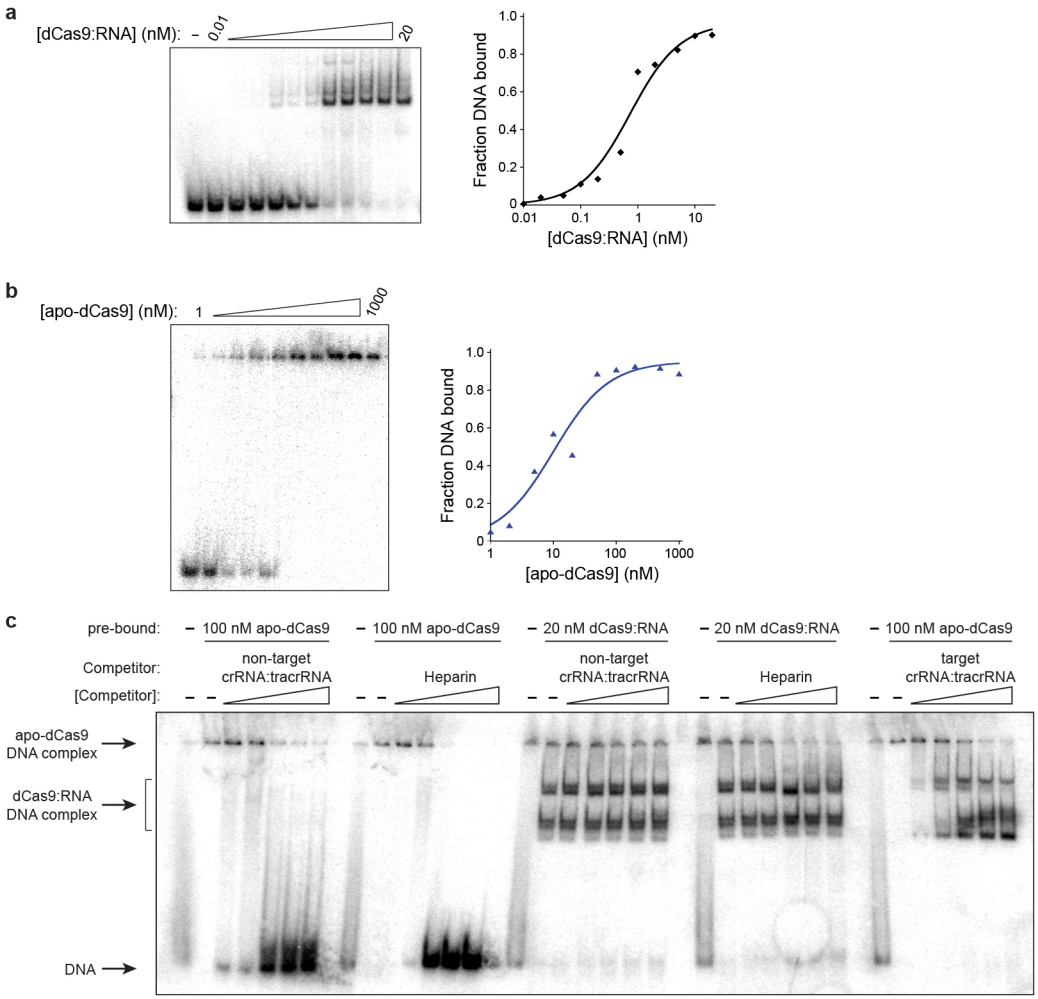
Extended Data Figure 1 | Activity assays of reagents used in single-molecule experiments. **a**, Cleavage assays were conducted using radiolabelled 55-bp DNA substrates that contained the six λ-DNA sequences targeted in Fig. 1d. Each DNA substrate (~1 nM) was incubated with 100 nM Cas9–RNA complex reconstituted using the corresponding guide RNA, and reaction products were resolved by 10% denaturing polyacrylamide gel electrophoresis (PAGE). Reactions contained 3×-Flag-tagged Cas9 (where indicated) or untagged, wild-type Cas9. The asterisk denotes further trimming of the non-target strand. **b**, Cleavage assay of λ-DNA under conditions identical to those used in single-molecule experiments. Full-length λ-DNA (25 ng μl^{-1}) was incubated with 10 nM Cas9–RNA reconstituted using the λ6 guide RNA, and reaction products were resolved by agarose gel electrophoresis. Successful cleavage is expected to generate DNA products that are 42,051 and 6,451 bp in length. When present, imaging components included anti-Flag antibody-coated quantum dots, YOYO1, BSA, glucose, and glucose oxidase/catalase.

©2014 Macmillan Publishers Limited. All rights reserved



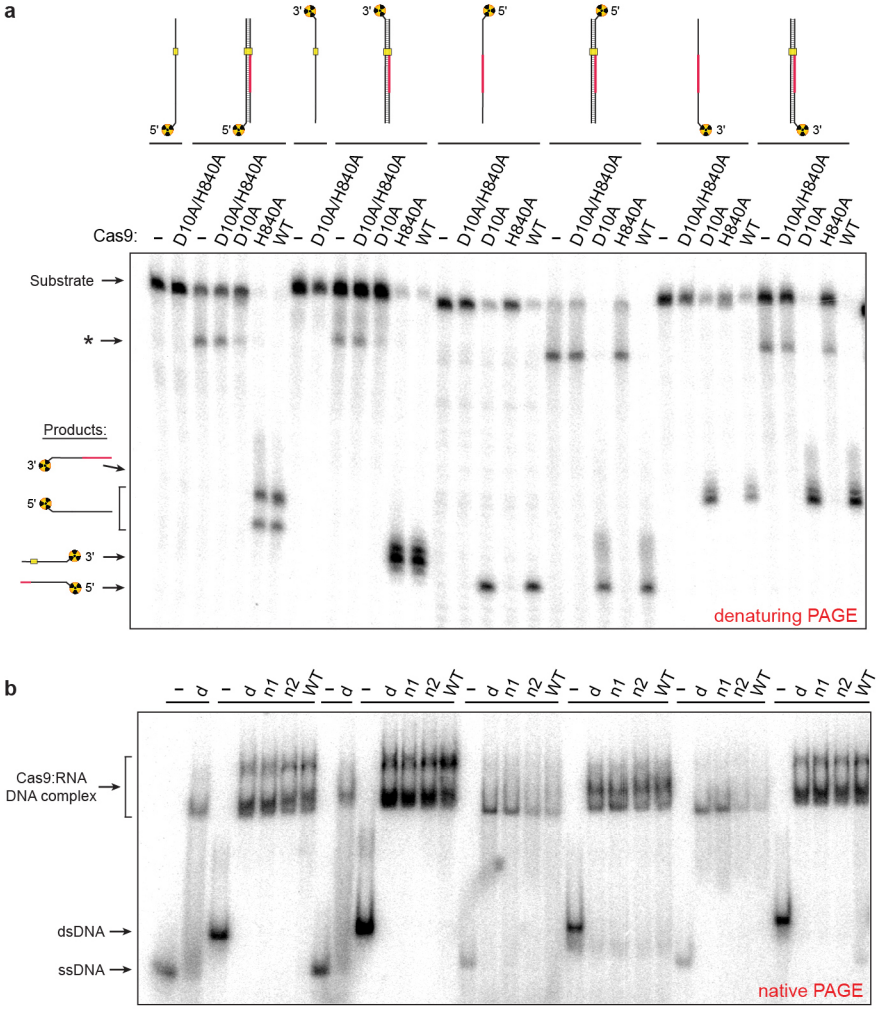
Extended Data Figure 2 | Binding histograms and Gaussian fits for λ-DNA target binding, and analysis of off-target binding. **a**, Binding distributions for dCas9 programmed with λ1–λ6 guide RNAs were measured as described in Methods, and the data from each individual experiment were then bootstrapped and fit with a Gaussian curve. Error bars represent 95% confidence intervals; n = 366, 378, 373, 420, 397 and 363 for experiments with λ1–λ6 guide RNAs, respectively. Shown in number of base pairs is the mean, μ, and standard deviation, σ, obtained from each Gaussian fit, as well as the expected location of each target site in λ-DNA. **b**, Distribution of Cas9–RNA binding events for λ2 crRNA (n = 2,330, top) and spacer 2 crRNA (n = 2,190, bottom); error bars represent 95% confidence intervals. The density of PAM sites throughout the λ-DNA substrate is shown in red. **c**, Survival probabilities for non-target binding events with λ2 (n = 632) and spacer 2 (n = 607) crRNA; error bars represent 70% confidence intervals. Data were collected at 25 mM KCl.

©2014 Macmillan Publishers Limited. All rights reserved



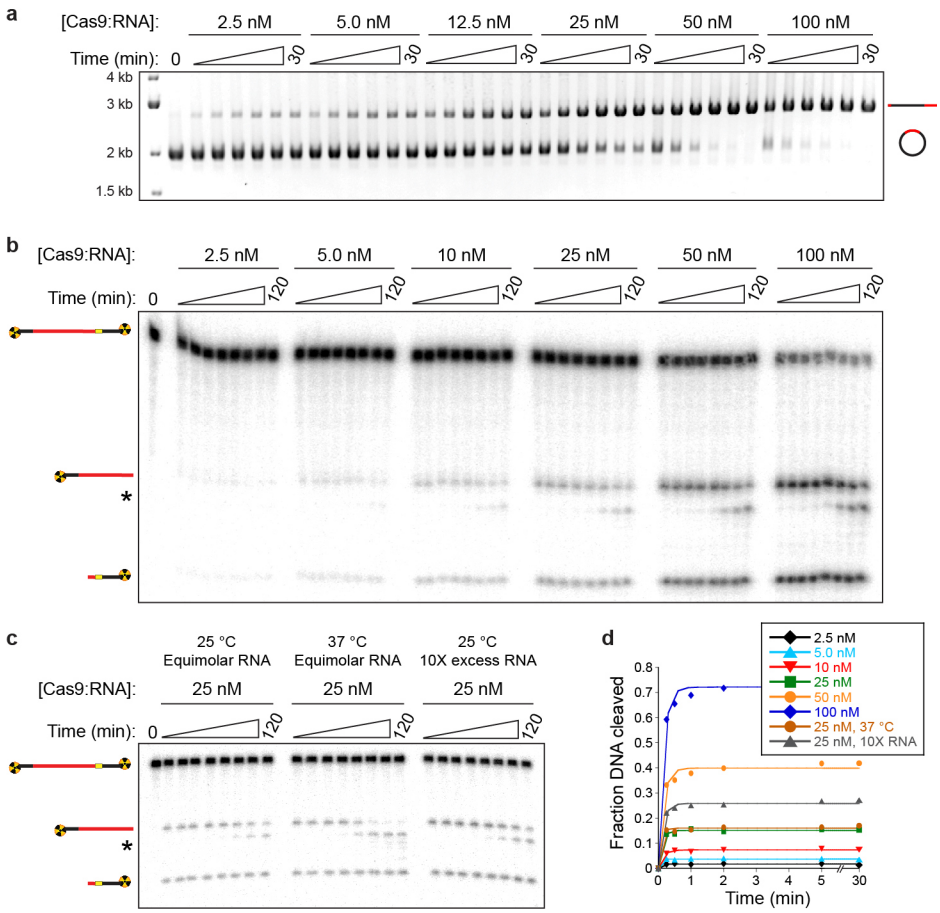
Extended Data Figure 3 | DNA binding by *apo*-dCas9 and dCas9–RNA. **a**, Electrophoretic mobility gel shift assay (left) with radiolabelled 55-bp target DNA and increasing concentrations of dCas9–RNA, using a 10× excess of crRNA–tracrRNA over dCas9. The quantified data (right) were fit with a standard binding isotherm (solid line), and data from three such experiments yielded an equilibrium dissociation constant (K_d) of 0.49 ± 0.21 nM. **b**, Results for *apo*-dCas9 shown as in panel **a**. Data from three independent experiments yielded a K_d of 26 ± 15 nM. **c**, crRNA–tracrRNA duplex and heparin dissociate *apo*-dCas9 bound to nonspecific DNA, but not dCas9–RNA complexes bound to target DNA. 55-bp DNA substrates were pre-incubated with the indicated reagent for 15 min at 37 °C, at which point non-targeting crRNA–tracrRNA duplex (10–1,000 nM) or heparin (0.01–100 μg ml^{-1}) was added. Reactions were incubated an additional 15 min at 37 °C and then resolved by 5% native PAGE. Reactions at the far right show that *apo*-dCas9 pre-bound to target DNA can be dissociated by complementary crRNA–tracrRNA and re-bind the same DNA in complex with RNA. Note the distinct mobilities of DNA in complex with *apo*-dCas9 versus DNA in complex with dCas9–RNA.

©2014 Macmillan Publishers Limited. All rights reserved



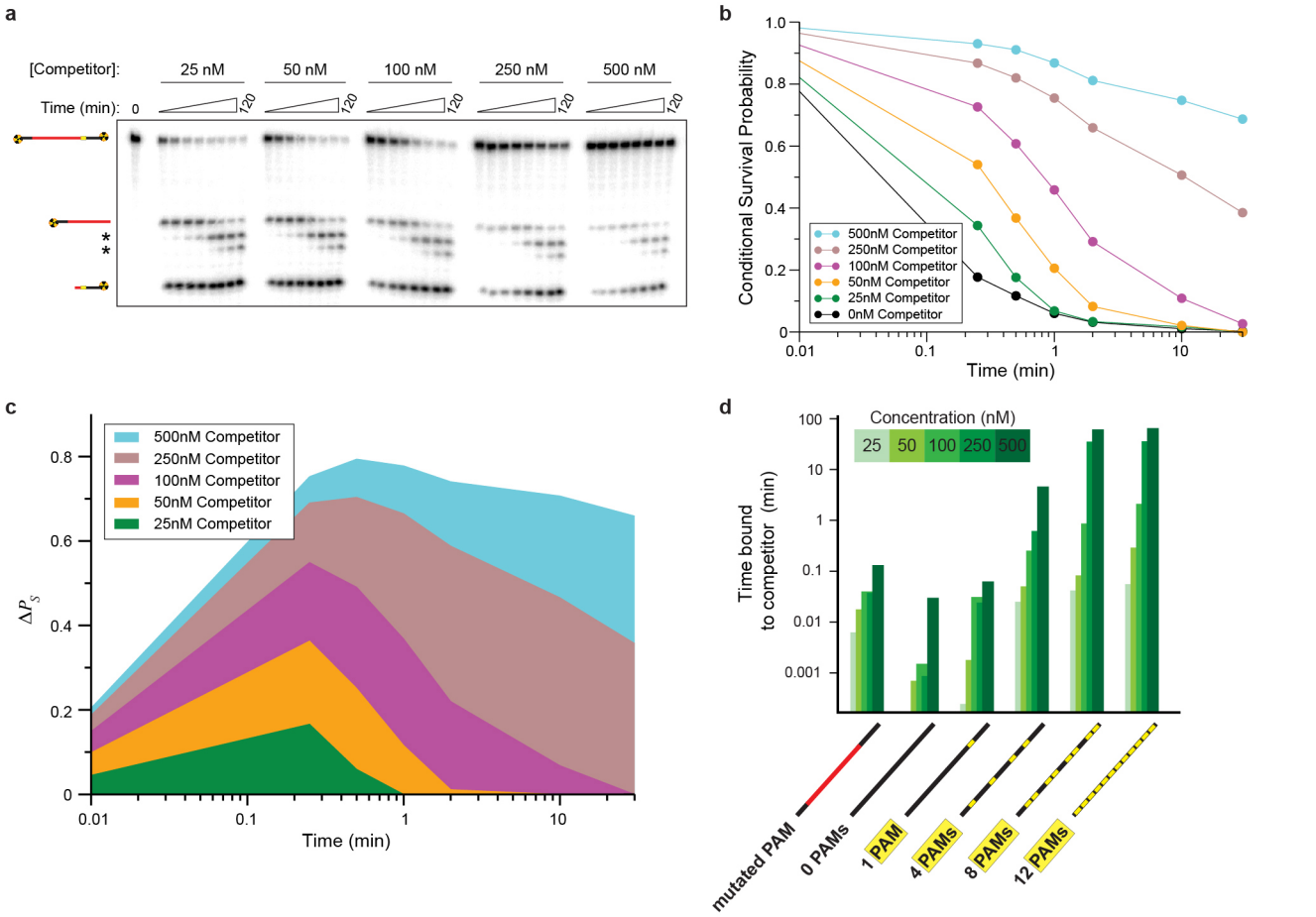
Extended Data Figure 4 | Target DNA cleavage products remain bound to Cas9–RNA. **a**, **b**, DNA substrates 72 nucleotides in length were radiolabelled at either their 5′ or 3′ ends and annealed to an unlabelled complementary strand, where indicated (top). The non-target strand contains the PAM (yellow box), whereas the target strand contains the sequence complementary to crRNA (red). Each DNA substrate (~1 nM) was incubated with 100 nM Cas9–RNA complex for 30 min at room temperature, using nuclease-inactive D10A/H840A Cas9 (d), both nickase mutants (D10A, n1; H840A, n2), and wild type (WT). Half the reaction volume was quenched with formamide gel loading buffer containing 50 mM EDTA and analysed by 10% denaturing PAGE to verify the expected cleavage pattern of each sample (**a**). The other half of each reaction was analysed by 5% native PAGE to determine whether the radiolabelled DNA fragment remained bound to Cas9–RNA (**b**). Aside from an apparent reduced affinity for the single-stranded target strand after cleavage, wild-type Cas9–RNA shows an affinity for all four possible DNA products that is indistinguishable from the affinity of dCas9–RNA for uncleaved DNA substrates. Note that the order of samples in **a** and **b** is identical. The additional band present for double-stranded DNA substrates in panel **a** results from incomplete denaturation and partial migration of intact duplex into the gel (marked with an asterisk).

©2014 Macmillan Publishers Limited. All rights reserved



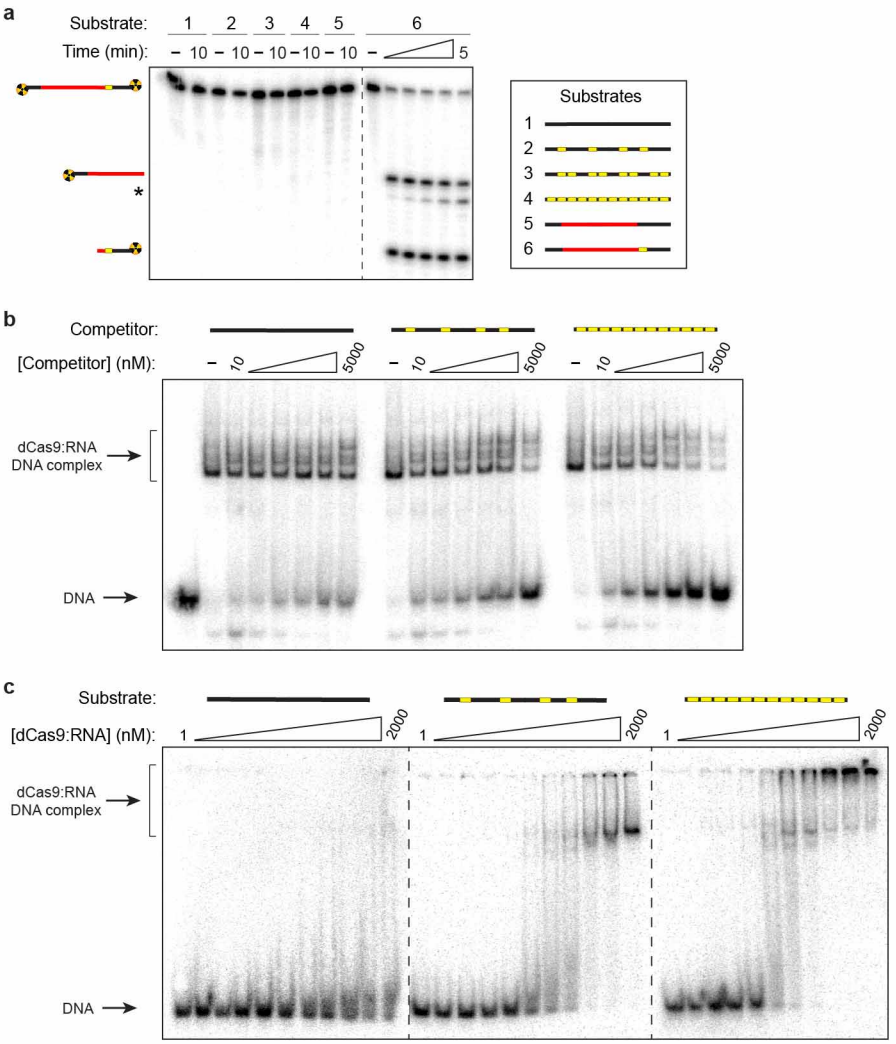
Extended Data Figure 5 | Cas9–RNA acts as a single-turnover enzyme. **a**, Agarose gel electrophoresis (1%, TAE buffer) was used to assess cleavage of plasmid DNA containing a λ2 target sequence as a function of Cas9–RNA concentration. DNA (25 nM) was incubated with the indicated concentration of Cas9–RNA, and aliquots were removed at each time point and quenched with gel loading buffer containing 50 mM EDTA. The gel was stained with ethidium bromide, and the quantified data are presented in Fig. 2b. **b**, Similar turnover experiments were conducted with 25 nM radiolabelled λ2 oligoduplex substrates and increasing concentrations of Cas9–RNA. Cleavage data were visualized by phosphorimaging; an asterisk denotes further trimming of the non-target strand. **c**, Turnover experiments with 25 nM Cas9–RNA were repeated at 37 °C and with a 10× excess of crRNA–tracrRNA over Cas9; neither condition significantly stimulates turnover. **d**, Quantified data from experiments in panels **b** and **c** show that each reaction reaches its maximum yield after ~1 min and does not increase with further incubation time, demonstrating that Cas9–RNA exhibits single-turnover activity. Note that the observed requirement for a slight stoichiometric excess of Cas9–RNA over DNA to reach reaction completion is probably a result of our enzyme preparations not being 100% active. Although modest turnover (2.5-fold) was observed at a single enzyme–substrate stoichiometry in ref. 6, our results clearly demonstrate that the reaction yield remains proportional to the molar ratio between Cas9–RNA and DNA across a range of concentrations.

©2014 Macmillan Publishers Limited. All rights reserved



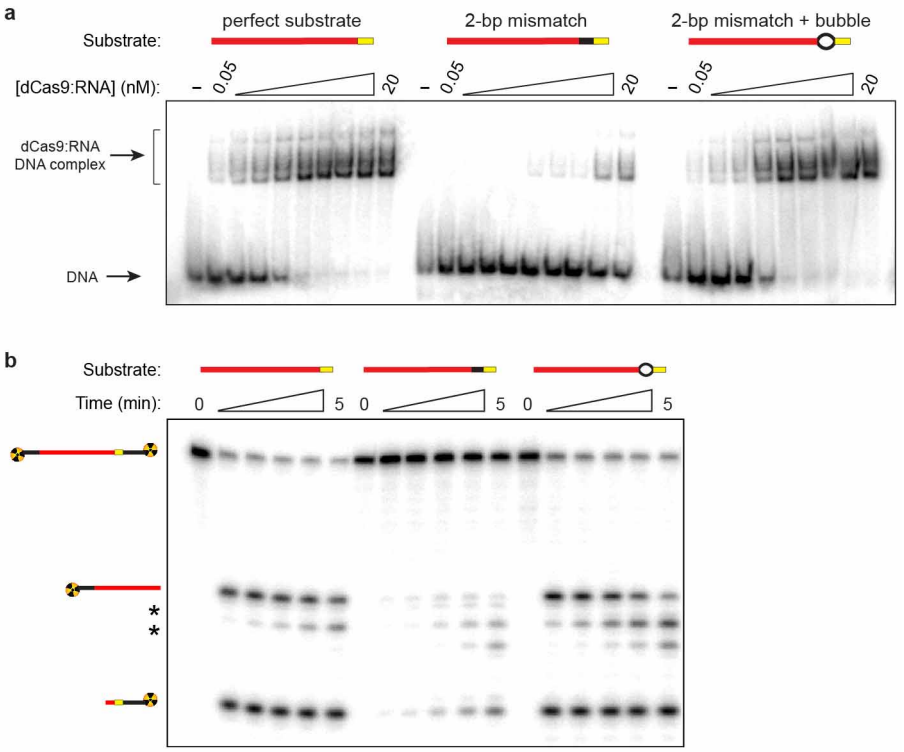
Extended Data Figure 6 | Analysis of competition cleavage assays. **a**, Representative cleavage assays as a function of competitor DNA concentration, using a competitor containing 12 PAM sites. Radiolabelled λ1 target DNA (1 nM) was incubated with 10 nM Cas9–RNA and increasing concentrations of the competitor, and reaction products at each time point were resolved by 10% denaturing PAGE. Cleavage data were visualized by phosphorimaging; an asterisk denotes further trimming of the non-target strand. **b**, Shown are the conditional survival probabilities for the radiolabelled target DNA at each concentration of 12-PAM competitor. **c**, Shown is the change in survival probability of the target DNA, $\Delta P_s(t)$, for each 12-PAM competitor concentration. The area under each curve represents the amount of time that Cas9–RNA spent on the competitor DNA during the reaction. **d**, Competition data with a panel of substrates that have no complementarity to the guide RNA and variable numbers of PAMs, and a perfect target sequence with single-base-pair mutation in the PAM. The data are presented similarly to Fig. 3c, but the time bound to competitor is shown for all five concentrations of competitor tested.

©2014 Macmillan Publishers Limited. All rights reserved



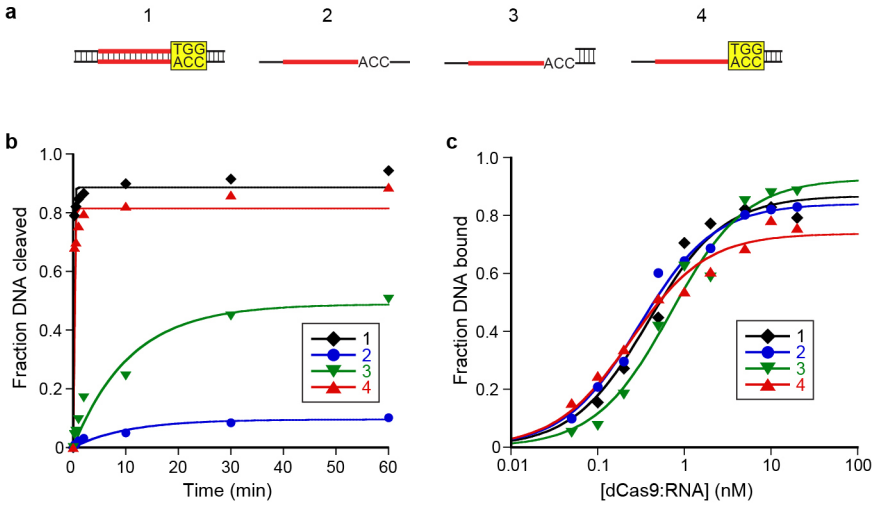
Extended Data Figure 7 | PAM sites in non-target DNA are bound specifically by dCas9–RNA. **a**, None of the competitors from Fig. 3c can be cleaved, including one that bears full complementarity to the crRNA but contains a single-base-pair mutation in the PAM. Radiolabelled competitor DNAs and target DNA (1 nM) were incubated with 100 nM wild-type Cas9–RNA for the indicated time, and reaction products were assessed by 10% denaturing PAGE. The asterisk denotes further trimming of the non-target strand. **b**, PAM-rich competitor DNAs interfere with target DNA binding by dCas9–RNA. The same radiolabelled 55-bp target DNA from Fig. 3b, c was pre-mixed with increasing concentrations of the indicated competitor DNA and then incubated with 10 nM dCas9–RNA for 60 min at 37 °C. Binding reactions were resolved by 5% native PAGE. **c**, dCas9–RNA has increased affinity for non-target DNA containing multiple PAM sequences. The indicated radiolabelled DNA substrates (~0.02 nM) were incubated with increasing concentrations of dCas9–RNA for 60 min at 37 °C, and reactions were resolved by 5% native PAGE. The observed well-shifting at high concentrations may result from multiple dCas9–RNA molecules binding the same DNA substrate.

©2014 Macmillan Publishers Limited. All rights reserved



Extended Data Figure 8 | Cas9–RNA binds and cleaves bubble-containing DNA substrates with mismatches to the crRNA that are otherwise discriminated against within the context of perfect duplexes. a, dCas9–RNA has weak affinity for a substrate containing a 2-bp mismatch to the crRNA (middle), whereas a substrate presenting the same mismatches within a small 2-nucleotide bubble (right) is bound with an affinity nearly indistinguishable from a perfect target substrate (left), in agreement with data presented in Fig. 3e. The indicated DNA substrates were incubated with increasing concentrations of dCas9–RNA for 60 min at 37 °C, and reactions were resolved by 5% native PAGE. **b**, The same bubble-containing substrate in panel **a** is cleaved with similar kinetics as a perfect substrate (compare right and left time courses), whereas a perfectly base-paired substrate with the same pattern of complementarity to the crRNA is cleaved with substantially reduced kinetics (middle). Radiolabelled DNA substrates (1 nM) were incubated with 100 nM wild-type Cas9–RNA for the indicated time, and reaction products were resolved by 10% denaturing PAGE. The asterisk denotes further trimming of the non-target strand.

©2014 Macmillan Publishers Limited. All rights reserved



Extended Data Figure 9 | PAM recognition activates the nuclease activity of Cas9. **a**, The indicated DNA substrates were prepared using the λ2 target sequence where the flanking region extending beyond the PAM was 16 bp (cleavage experiments) or 26 bp (binding experiments). **b**, For cleavage experiments, substrates were prepared by annealing the radiolabelled target strand (that is, substrate 2) to a 5× excess of cold complement, and 1 nM DNA was reacted with 50 nM Cas9–RNA at room temperature. Reaction products were resolved by 10% denaturing PAGE, and the quantified data were fit with single-exponential decays (solid lines). Results from three independent experiments yielded apparent pseudo-first-order cleavage rate constants of $9.0 \pm 2.0\,min^{-1}$ (substrate 1), $0.067 \pm 0.027\,min^{-1}$ (substrate 2), $0.066 \pm 0.024\,min^{-1}$ (substrate 3) and $7.3 \pm 3.2\,min^{-1}$ (substrate 4), and are presented as values relative to substrate 1 in Fig. 4b. Rate constants for substrates 2 and 3 are probably overestimates, as the reactions did not approach completion and the data were best fit with amplitudes well below 1. **c**, For binding experiments, substrates were gel purified after annealing the radiolabelled target strand to a 10× excess of cold complement. Binding reactions contained ~0.1 nM DNA and increasing concentrations of dCas9–RNA, and were incubated at 37 °C for 1 h before being resolved by 5% native PAGE. The quantified data were fit with standard binding isotherms (solid lines). Results from three independent experiments yielded apparent K_d values of 0.27 ± 0.14 nM (substrate 1), 0.28 ± 0.12 nM (substrate 2), 0.59 ± 0.18 nM (substrate 3) and 0.21 ± 0.06 nM (substrate 4), and are presented as values relative to substrate 1 in Fig. 4b.

©2014 Macmillan Publishers Limited. All rights reserved

Extended Data Table 1 | RNA and DNA substrates used in this study

Description	Sequence *	Used in
λ1 target sequence, 12,128 bp in λ-DNA	5′-GGCGCATAAAGATGAGACGCTGG-3′ 3′-CCGCGTATTTCTACTCTGCGACC-5′	Fig. 1b,d
λ2 target sequence, 18,071 bp in λ-DNA	5′-GTGATAAGTGGAATGCCATGTGG-3′ 3′-CACTATTCACCTTACGGTACACC-5′	Fig. 1b,d
λ3 target sequence, 24,200 bp in λ-DNA	5′-CTGGTGAACTTCCGATAGTGCGG-3′ 3′-GACCACTTGAAGGCTATCACGCC-5′	Fig. 1b,d
λ4 target sequence, 30,520 bp in λ-DNA	5′-CAGATATAGCCTGGTGGTTCAGG-3′ 3′-GTCTATATCGGACCACCAAGTCC-5′	Fig. 1b,d
λ5 target sequence, 35,894 bp in λ-DNA	5′-GGCAATGCCGATGGCGATAGTGG-3′ 3′-CCGTTACGGCTACCGCTATCACC-5′	Fig. 1b,d
λ6 target sequence, 42,035 bp in λ-DNA	5′-GGTGTGAAAGAACACCAACAGGG-3′ 3′-CCACACTTTCTTGTGGTTGTCCC-5′	Fig. 1b,d
Oligo for T7 promoter, in vitro transcription	5′-**TAATACGACTCACTATA**-3′	NA
ssDNA T7 template, λ1-targeting crRNA †	5′-CAAAACAGCATAGCTCTAAAACGCGTCTCATCTTTATGCGTC**TATAGTGAGTCGTATTA**-3′	NA
λ1-targeting crRNA ‡	5′-GACGCAUAAAGAUGAGACGCGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 1d, 3, ED1, 2a, 3, 6, 7, 8
λ2-targeting crRNA	5′-GUGAUAAGUGGAAUGCCAUGGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 1c,d,g, 2, 4a,b, ED2, 4, 5, 9
λ3-targeting crRNA	5′-CUGGUGAACUUCCGAUAGUGGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 1d, ED1, 2a
λ4-targeting crRNA	5′-CAGAUAUAGCCUGGUGGUUCGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 1d, ED1, 2a
ssDNA T7 template, λ5-targeting crRNA †	5′-CAAAACAGCATAGCTCTAAAACCTATCGCCATCGGCATTGCC**TATAGTGAGTCGTATTA**-3′	NA
λ5-targeting crRNA	5′-GGCAAUGCCGAUGGCGAUAGGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 1d, ED1, 2a
ssDNA T7 template, λ6-targeting crRNA †	5′-CAAAACAGCATAGCTCTAAAACTGTTGGTGTTCTTTCACACC**TATAGTGAGTCGTATTA**-3′	NA
λ6-targeting crRNA	5′-GGUGUGAAAGAACACCAACAGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 1d, ED1, 2a
Spacer 2 crRNA	5′-AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAUGCUGUUUUG-3′	Fig. 2h, ED2b,c
ssDNA T7 template, tracrRNA †	5′-AAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATGCTGTCC**TATAGTGAGTCGTATTA**-3′	NA
tracrRNA (nt 15-87)	5′-GGACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUU-3′	All expts with Cas9:RNA
λ1 target duplex §	5′-AGCAGAAATCTCTGCTGACGCATAAAGATGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATTTCTACTCTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3, ED1, 3, 6, 7a,b, 8
λ2 target duplex	5′-GAGTGGAAGGATGCCAGTGATAAGTGGAATGCCATGTGGCTGTCAAAATTGAGC-3′ 3′-CTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCGACAGTTTTAACTCG-5′	Fig. 4a,b, ED1, 4, 5b-d, 9b
λ3 target duplex	5′-AACGTGCTGCGGCTGGCTGGTGAACTTCCGATAGTGCGGGTGTTGAATGATTTCC-3′ 3′-TTGCACGACGCCGACCGACCACTTGAAGGCTATCACGCCCACAACTTACTAAAGG-5′	Fig. ED1
λ4 target duplex	5′-TCACAACAATGAGTGGCAGATATAGCCTGGTGGTTCAGGCGGCGCATTTTTATTG-3′ 3′-AGTGTTGTTACTCACCGTCTATATCGGACCACCAAGTCCGCCGCGCGTAAAAATAAC-5′	Fig. ED1
λ5 target duplex	5′-GAATGAACGATGCAGAGGCAATGCCGATGGCGATAGTGGGTATCATGTAGCCGCT-3′ 3′-CTTACTTGCTACGTCTCCGTTACGGCTACCGCTATCACCCATAGTACATCGGCGA-5′	Fig. ED1
λ6 target duplex	5′-AATCGATGGTGTCTCCGGTGTGAAAGAACACCAACAGGGGTGTTACCACTACCGC-3′ 3′-TTAGCTACCACAGAGGCCACACTTTCTTGTGGTTGTCCCCACAATGGTGATGGCG-5′	Fig. ED1
λ2 target duplex, cloned into pUC19 ‖	5′-AATTGAAAGTGATAAGTGGAATGCCATGTGGAAAC -3′ 3′- CTTTCACTATTCACCTTACGGTACACCTTTGCTAG-5′	Fig. 2b, ED5a
λ1 Competitor, Mutated PAM	5′-AGCAGAAATCTCTGCTGACGCATAAAGATGAGACGCTCGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATTTCTACTCTGCGAGCTCATGTTTGCAGTCGA-5′	Fig. 3c, ED7a
Competitor, 0 PAMs	5′-AGCTGCATAACGCGAAAAAATATATTTATCTGCTTGATCTTCAAATGTTGTATTG-3′ 3′-TCGACGTATTGCGCTTTTTTATATAAATAGACGAACTAGAAGTTTACAACATAAC-5′	Fig. 3b,c, ED7
Competitor, 1 PAM	5′-AGCAGAAATCTCTGCTCTGCGTATTTCTACTCTGCGTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGAGACGCATAAAGATGAGACGCACCTCATGTTTGCAGTCGA-5′	Fig. 3c,e
Competitor, 4 PAMs	5′-AGCTGCATAACGCGGGAAAATCCATTTATCTGCTTGATCTTCGGATGTTCCATTG-3′ 3′-TCGACGTATTGCGCCCTTTTAGGTAAATAGACGAACTAGAAGCCTACAAGGTAAC-5′	Fig. 3c, ED7
Competitor, 8 PAMs	5′-GGCTGCACCACGCGGGAAAATCCATTTAGGTGCTTCCTCTTCGGATGTTCCATTG-3′ 3′-CCGACGTGGTGCGCCCTTTTAGGTAAATCCACGAAGGAGAAGCCTACAAGGTAAC-5′	Fig. 3c, ED7a
Competitor, 12 PAMs	5′-GGCTGGACCACGCGGGAAAATCCACCTAGGTGGTTCCTCTTCGGATGTTCCATCC-3′ 3′-CCGACCTGGTGCGCCCTTTTAGGTGGATCCACCAAGGAGAAGCCTACAAGGTAGG-5′	Fig. 3c, ED7
λ1 Competitor, 16–4–PAM	5′-AGCAGAAATCTCTGCTCTGCGTATTTCTACTCACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGAGACGCATAAAGATGAGTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 12–8–PAM	5′-AGCAGAAATCTCTGCTCTGCGTATTTCTTGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGAGACGCATAAAGAACTCTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 8–12–PAM	5′-AGCAGAAATCTCTGCTCTGCGTATAAGATGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGAGACGCATATTCTACTCTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 4–16–PAM	5′-AGCAGAAATCTCTGCTCTGCCATAAAGATGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGAGACGGTATTTCTACTCTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 4–4–12–PAM	5′-AGCAGAAATCTCTGCTGACGGTATAAGATGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCCATATTCTACTCTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e, ED6
λ1 Competitor, 8–4–8–PAM	5′-AGCAGAAATCTCTGCTGACGCATATTCTTGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATAAGAACTCTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 12–4–4–PAM	5′-AGCAGAAATCTCTGCTGACGCATAAAGAACTCACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATTTCTTGAGTGCGACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 16–4–PAM	5′-AGCAGAAATCTCTGCTGACGCATAAAGATGAGTGCGTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATTTCTACTCACGCACCTCATGTTTGCAGTCGA-5′	Fig. 3e
λ1 Competitor, 18–2–PAM	5′-AGCAGAAATCTCTGCTGACGCATAAAGATGAGACCGTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATTTCTACTCTGGCACCTCATGTTTGCAGTCGA-5′	Fig. 3e, ED8
λ1 Competitor, 18–2b–PAM	5′-AGCAGAAATCTCTGCTGACGCATAAAGATGAGACGCTGGAGTACAAACGTCAGCT-3′ 3′-TCGTCTTTAGAGACGACTGCGTATTTCTACTCTGGCACCTCATGTTTGCAGTCGA-5′	Fig. 3e, ED8
λ2 target (55-nt), dsDNA	5′-GAGTGGAAGGATGCCAGTGATAAGTGGAATGCCATGTGGCTGTCAAAATTGAGC-3′ 3′-CTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCG-5′	Fig. 4a,b, ED1, 4, 5b-d, 9b
λ2 target (55-nt), ssDNA target strand	3′-CTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCG-5′	Fig. 4a,b, ED9b
λ2 target (55-nt), partial dsDNA without PAM	5′- GCTGTCAAAATTGAGC-3′ 3′-CTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCG-5′	Fig. 4a,b, ED9b
λ2 target (55-nt), partial dsDNA with PAM	5′- TGGGCTGTCAAAATTGAGC-3′ 3′-CTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCG-5′	Fig. 4a,b, ED9b
λ2 target (72-nt), dsDNA	5′-TTCGAAAGAGTGGAAGGATGCCAGTGATAAGTGGAATGCCATCTGGGCTGTCAAAATTGAGCAGACCAAAGA-3′ 3′-AAGCTTTCTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCGTCTGGTTTCT-5′	Fig. 4b, ED4, 9c
λ2 target (72-nt), ssDNA	3′-AAGCTTTCTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCGTCTGGTTTCT-5′	Fig. 4b, ED9c
λ2 target (72-nt), partial dsDNA without PAM	5′- GCTGTCAAAATTGAGCAGACCAAAGA-3′ 3′-AAGCTTTCTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCGTCTGGTTTCT-5′	Fig. 4b, ED9c
λ2 target (72-nt), partial dsDNA with PAM	5′- TGGGCTGTCAAAATTGAGCAGACCAAAGA-3′ 3′-AAGCTTTCTCACCTTCCTACGGTCACTATTCACCTTACGGTACACCCGACAGTTTTAACTCGTCTGGTTTCT-5′	Fig. 4b, ED9c

NA, not applicable.

* crRNA guide sequences and complementary DNA target strand sequences are shown in red. PAM sites (5′-NGG-3′) are highlighted in yellow on the non-target strand when adjacent to the target sequence, except for PAM competitors in which case all PAMs are highlighted.

† The reverse complement of the T7 promoter is indicated in bold.

‡ The second nucleotide of the λ1-targeting crRNA was mutated from G to A to match the λ1 target duplex sequence.

§ The underlined base pairs were mutated relative to the wild-type λ-DNA sequence to remove all PAM sites from the substrate other than the PAM immediately adjacent to the target sequence. The crRNA was mutated accordingly, as were the λ1 competitor DNAs.

‖ The duplex was cloned into EcoRI and BamHI sites on pUC19.

©2014 Macmillan Publishers Limited. All rights reserved