**Graphical Abstract**

Reversibility during R-loop propagation ensures high target specificity.

Upon completion of R-loop, DNA target cleavage is inevitable.

**Highlights**

- DNA target binding by CRISPR-Cas12a is rate limiting for DNA cleavage.
- Multiple cleavage sites are present on each DNA strand.
- Specificity against mismatches suggests a late transition state for R-loop formation.
- Data explain *in vivo* cleavage patterns by Cas12a and differences from Cas9.

**Authors**

Isabel Strohkendl, Fatema A. Saifuddin, James R. Rybarski, Ilya J. Finkelstein, Rick Russell

**Correspondence**

rick_russell@cm.utexas.edu

**In Brief**

Strohkendl et al. dissect DNA binding and cleavage by CRISPR-Cas12a. They show that binding is functionally irreversible, yet Cas12a discriminates against mismatches with the target DNA extending beyond a seed region. These results suggest that R-loop propagation is readily reversible, enabling Cas12a to select DNA sequences more precisely than Cas9.
Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a

Isabel Strohkendl,1 Fatema A. Saifuddin,1,2 James R. Rybarski,1,2 Ilya J. Finkelstein,1,2 and Rick Russell1,3,*

1Department of Molecular Biosciences and the Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA
2Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, TX 78712, USA
3Lead Contact
*Correspondence: rick_russell@cm.utexas.edu
https://doi.org/10.1016/j.molcel.2018.06.043

SUMMARY

Class 2 CRISPR-Cas nucleases are programmable genome editing tools with promising applications in human health and disease. However, DNA cleavage at off-target sites that resemble the target sequence is a pervasive problem that remains poorly understood mechanistically. Here, we use quantitative kinetics to dissect the reaction steps of DNA targeting by Acidaminococcus sp Cas12a (also known as Cpf1). We show that Cas12a binds DNA tightly in two kinetically separable steps. Protospacer-adjacent motif (PAM) recognition is followed by rate-limiting R-loop propagation, leading to inevitable DNA cleavage of both strands. Despite functionally irreversible binding, Cas12a discriminates strongly against mismatches along most of the DNA target sequence. This result implies substantial reversibility during R-loop formation—a late transition state—and defies common descriptions of a “seed” region. Our results provide a quantitative basis for the DNA cleavage patterns measured in vivo and observations of greater reported target specificity for Cas12a than for the Cas9 nuclease.

INTRODUCTION

CRISPR-Cas systems are revolutionary new tools for gene editing applications (Hsu et al., 2014). The nuclease enzymes in these systems can be readily programmed by the rational design of a CRISPR RNA (crRNA) sequence. Targets are identified via protospacer-adjacent motif (PAM) recognition and formation of a helix between the guide RNA and the DNA or RNA target, which licenses cleavage of the target. The information inherent in the ~20 base pairs formed by the crRNA and the target is sufficient, in principle, to uniquely target a single sequence, even in large eukaryotic genomes.

Class 2 CRISPR-Cas systems are particularly useful because they use a single polypeptide for both recognition and cleavage of the target DNA or RNA. However, Cas9, the prototypical class 2 enzyme used in eukaryotic cells, invariably falls short of targeting a single unique site, as related “off-target” DNA sequences are also recognized and cleaved at significant levels (Stemberg and Doudna, 2015). The existence of off-target cleavage events is not surprising, as molecular interactions are never perfectly specific. For example, mismatches in DNA-RNA duplexes typically incur penalties of only 2–5 kcal/mol in solution (Sugimoto et al., 1995; Watkins et al., 2011). Even with one or two mismatches, a 20-bp helix is stable (Herschlag, 1991), and complexes formed with imperfect targets may be sufficiently long-lived to favor cleavage over enzyme dissociation (Bisaria et al., 2017). Despite the importance of these off-target cleavage events, we still lack a full understanding of the biophysical basis for the specificity of CRISPR-Cas enzymes.

Cas12a (also known as Cpf1) is a class 2 enzyme (Makarova et al., 2015; Shmakov et al., 2015; Zetsche et al., 2015) that has recently emerged as a more specific alternative to Cas9 (Figure 1A) (Kim et al., 2016; Swarts et al., 2017). Indeed, both Acidaminococcus sp. BV3L6 Cas12a (AsCas12a) and Lachnospiraceae bacterium Cas12a (LbCas12a) showed little or no tolerance for mismatches in mammalian cells (Kim et al., 2016, 2017; Kleinstiver et al., 2016; Töth et al., 2016; Tu et al., 2017; Zhong et al., 2017). In contrast, Cas9 discriminates strongly against mismatches only within the first ~10 bp of the RNA-DNA helix (R-loop) proximal to the PAM (Fu et al., 2013; Hsu et al., 2013; Slaymaker et al., 2016). Cas12a is also capable of processing its own precursor crRNA, unlike Cas9, and can, therefore, be used for multiplexed applications (Fonfara et al., 2016; Zetsche et al., 2017).

One striking question that arises is how different CRISPR-Cas enzymes can have different levels of specificity, despite using the chemically identical R-loop as the source of sequence-specific DNA binding. A common feature of CRISPR-Cas systems and other RNA-directed enzymes is that the targeting RNA includes a “seed” region. The seed region, which, for CRISPR-Cas enzymes, is a subset of nucleotides within the crRNA that base pairs with PAM-proximal nucleotides, is highly sensitive against mismatches and is, therefore, critical for target affinity and specificity (Gorski et al., 2017; Jinek et al., 2012; Wiedenheft et al., 2011). A crystal structure of Cas12a from Francisella novicida (FnCas12a) showed that the bound crRNA includes a 5-nt region that is pre-ordered for base-pair formation (Swarts et al., 2017) with PAM-proximal nucleotides. This pre-ordering of the crRNA may contribute to a PAM-proximal seed region, but it...
cannot account for the high cleavage specificity seen throughout the R-loop beyond these 5 nt. Indeed, little is known about how Cas12a achieves its high specificity despite the many published crystal structures (Dong et al., 2016; Gao et al., 2016; Stella et al., 2017; Swarts et al., 2017; Yamano et al., 2016, 2017) and in vivo targeting results. Here, we report a quantitative dissection of the binding and cleavage reactions of Cas12a, pre-loaded with a defined crRNA, toward a perfectly matched DNA target and a series of targets with single mismatches at each position within the R-loop. We find that DNA binding is rate limiting for cleavage, both for matched and mismatched targets. While this behavior generally abrogates specificity of enzymes, we find that Cas12a discriminates strongly against mismatches across most of the R-loop, suggesting reversibility within the process of R-loop formation and a seed region that defies the strict boundary observed for Cas9. Our results provide fundamental insights into DNA targeting by Cas12a, and the origins of greater DNA target specificity of Cas12a than Cas9. In addition, our work suggests strategies for further improvement of specificity through enzyme engineering.

RESULTS

A Kinetic Framework for Recognition and Cleavage of a Matched DNA Target

We first generated a complete kinetic framework for Cas12a on a matched DNA target. The rates of DNA binding and dissociation, as well as cleavage of each DNA strand, were measured using a crRNA-loaded AsCas12a (henceforth referred to as Cas12a) and an oligonucleotide DNA target (Figure 1B). Our target DNA included a 5′-TTTA-3′ PAM sequence, which conforms to the consensus 5′-TTTV-3′ PAM (Kim et al., 2017), adjacent to 24 bp that matched the crRNA targeting sequence. The target DNA was 32P labeled at the 5′ end of either the target strand (TS) or the non-target strand (NTS). The second-order rate constant for Cas12a binding ($k_{on}$) was determined by incubating various concentrations of Cas12a target relative to Cas12a, the fraction of processed DNA target is plotted over time. Specifically, the plot indicates the amount of DNA target that has been bound by Cas12a or released as PAM-proximal product relative to the amount of Cas12a. Over the time shown, released product was not observed, as expected from previous measurements and the very slow release of uncleaved DNA target measured herein; see (E).

See also Figures S1 and S2.
with a trace amount of NTS-labeled duplex. Reaction aliquots were quenched at various times by the addition of the same unlabeled duplex in excess and analyzed by native PAGE (Figure 1C). The dependence of the observed rate constant on Cas12a concentration gave a $k_{on}$ value of 1.1 ($\pm 0.1$) $\times 10^8$ M$^{-1}$ s$^{-1}$, within 1–2 orders of magnitude of the macromolecular diffusion limit. The $k_{on}$ value was the same within error for the nuclease-inactive Cas12a(D908A) (Yamano et al., 2016; Zetsche et al., 2013), indicating that cleavage of the DNA does not impact the observed binding rate (Figure S1). A hyperbolic dependence on protein concentration with an apparent plateau was observed for both Cas12a and Cas12a(D908A), indicating that binding includes two detectable steps. We infer that the Cas12a concentration that gives the half-maximal rate constant ($K_{1/2}$) of 1.1 ± 0.2 nM reflects the equilibrium constant for an initial complex, most likely resulting from PAM recognition (Gong et al., 2018). The maximal rate constant ($k_{max}$) of 0.13 ± 0.01 s$^{-1}$ reflects that a first-order step following initial binding, which we show later to be R-loop formation, limits the overall binding rate at high Cas12a concentrations. Cas12a(D908A) behaved similarly but with a somewhat larger $k_{max}$ value of 0.24 ± 0.06 s$^{-1}$ (Figure S1A), suggesting that the mutated residue within the RuvC domain may be involved in R-loop formation.

To directly measure R-loop formation, we used stopped-flow fluorescence with a DNA substrate that included a 2-aminopurine (2-AP) at either position 4 or position 18 of the NTS to report on the early and late stages of R-loop propagation (Figure S2A). The fluorescence of 2-AP is strongly quenched by stacking when it is base paired, making 2-AP a sensitive probe for monitoring that a first strand is base paired, making 2-AP a sensitive probe for monitoring the presence of Mg$^{2+}$, with a $k_{off}$ value of 5.9 ($\pm 0.4$) $\times 10^{-6}$ s$^{-1}$ ($t_{1/2} = 32 \pm 2$ hr). As a control, we compared the wild-type Cas12a and Cas12a(D908A) in the presence of EDTA to block DNA cleavage. The $k_{off}$ values were 2.4 ($\pm 0.3$) and 2.3 ($\pm 0.3$) $\times 10^{-6}$ s$^{-1}$ for Cas12a and Cas12a(D908A), respectively (Figure S1B), indicating that the D908A substitution does not affect the lifetime of the complex. In addition, Mg$^{2+}$ increases the lifetime of the ternary complex by 40-fold, perhaps by favoring the conformational transition that follows R-loop formation or by contributing more directly to R-loop stability in the complex.

The very slow dissociation from the matched target suggested that release of the PAM-proximal DNA product after cleavage would also be very slow and would result in a stoichiometric burst of processed target DNA. Indeed, we observed such a stoichiometric burst (Figure 1F). The lack of detectable enzyme turnover is consistent with previous observations of Cas12a as a “single turnover” enzyme (Chen et al., 2018; Singh et al., 2018), and the stoichiometry of the product indicates that Cas12a is fully active and that each Cas12a molecule binds and cleaves one DNA target.

Our measured rate constants for binding and dissociation of Cas12a from the matched DNA target give a calculated $K_D$ value of 54 ± 4 fM ($k_{on}/k_{off}$). This value is $10^3$-fold lower than the recently reported value of 0.1 nM from single-molecule fluorescence experiments (Singh et al., 2018) and $10^4$-fold lower than that determined via binding affinity assays with FnCas12a in the absence of Mg$^{2+}$ ion (Fontana et al., 2016). The previous studies used direct measurements of the fraction of DNA target bound rather than kinetics measurements, and they did not provide the >32-hr incubation required to reach equilibrium (i.e., at least one half-life of the bound complex), which would lead to apparently weaker binding of Cas12a. Although our measured equilibrium constant is extraordinarily low for an enzyme-substrate interaction, reflecting a very stable complex, the value is three orders of magnitude higher than the predicted $K_D$ value for a 20-bp RNA:DNA duplex of the same sequence (~0.02 fM) (Sugimoto et al., 1995).

We next measured cleavage of the NTS and the TS by individually radiolabeling each strand in separate reactions (Figure 2A). For each strand of the target DNA, we used hyperbolic fits to the observed rate constants from multiple Cas12a concentrations to determine the observed rate constants for cleavage. We obtained a maximal rate constant of 5.2 ($\pm 0.6$) $\times 10^{-2}$ s$^{-1}$ for NTS cleavage and a 10-fold lower value for TS cleavage, 5.1 ($\pm 0.2$) $\times 10^{-3}$ s$^{-1}$. The slower TS cleavage is consistent with a recent report that the NTS is expected to line the nuclease lobe of Cas12a, which contains the RuvC nuclease domain (Stella et al., 2017; Swarts et al., 2017) (Figure 1A), and is thus poised for cleavage. In contrast, the TS:crRNA heteroduplex (R-loop) lines the recognition lobe and may require substantial movement of the TS and/or the single Ruvc nuclease domain to position the TS for cleavage (Stella et al., 2017; Swarts et al., 2017).

Close inspection of the DNA cleavage reactions revealed multiple products for each strand (Figure 2A). For the NTS, the initial cleavage event occurred at position 16 and was followed by apparenttrimming of up to 4 nt, with a rate constant of 1.2 ($\pm 0.1$) $\times 10^{-2}$ s$^{-1}$, resulting in the accumulation of faster
Together, our results show that DNA cleavage by Cas12a occurs orders of magnitude faster than dissociation from a matched DNA target, indicating that essentially every complete binding event of Cas12a results in DNA cleavage (Gisaria et al., 2017). A key consequence of this irreversible binding is that the specificity of Cas12a against mismatches must be determined by the relative binding kinetics, not by differences in equilibrium binding.

**Single Mismatches throughout the R-Loop Slow Cas12a Binding**

To measure the specificity of Cas12a against mismatches between the crRNA and target DNA, we introduced single base-pair changes in the DNA at each position throughout the R-loop (Table S1). The $K_{on}$ values were reduced substantially for mismatched DNA targets, with pronounced effects along nearly the entire length of the R-loop (Figures 3 and S3; Table S2). PAM-proximal mismatches gave ~150- to 600-fold decreases, and PAM-distal mismatches gave up to 65-fold decreases. Additionally, we found that the identity of the nucleotide mismatch can impact the binding rate, as C16G slowed binding by nearly 70-fold, while C16T slowed binding by only 7-fold. A mismatch at position 21 did not decrease the rate of binding, supporting previous findings that a conserved aromatic residue (W382) stacks on position 20 of the RNA-DNA duplex and prevents further R-loop propagation (Swarts et al., 2017; Yamano et al., 2016).

The sensitivity of Cas12a to mismatches indicates that almost all positions of the R-loop contribute significantly to the binding rate constant. This result was surprising, because when an RNA-DNA duplex forms in the absence of a protein co-factor, the transition state is reached after the formation of only a few base pairs (Pörschke, 1977; Woodside et al., 2006b). Thus, only these base pairs contribute to the rate and specificity of duplex formation (Woodside et al., 2006a). In contrast, our results indicate that, when Cas12a-crRNA binds to its target DNA, the formation of the R-loop has a later transition state, with more base pairs formed. Indeed, assuming that the R-loop begins to form adjacent to the PAM sequence (Sterberg et al., 2014; Szczelkun et al., 2014), the sensitivity of the binding rate to PAM-distal mismatches implies that propagation of the R-loop remains readily reversible until nearly the entire loop is formed. Biologically, this late transition state is expected to permit Cas12a to discriminate against mismatches in spite of the irreversible nature of the overall binding process.

To further explore the effects of sequence mismatches on Cas12a binding, we measured Cas12a dissociation from a series of DNA targets. All but one mismatch within the R-loop increased the dissociation rate (Figures 3 and S3C; Table S2), as is expected for destabilizing changes to the bound complex. Intriguingly, the largest increases in the dissociation rate were caused by PAM-distal mismatches. This effect has also been observed for Cas9 (Boyle et al., 2017; Szczelkun et al., 2014). PAM-distal mismatches increased the rate of dissociation 6- to 80-fold, except one substitution that gave a smaller increase of 3-fold (T14C), perhaps because this mismatched target can form a dG-rU wobble base pair. In contrast, PAM-proximal mismatches increased the rate of Cas12a dissociation only up to...
5-fold. The larger effects of PAM-distal mismatches on dissociation support the late transition state model proposed from our $k_{on}$ results.

The combined effects of mismatches on $k_{on}$ and $k_{off}$ reveal a high level of thermodynamic discrimination by Cas12a for its perfect target DNA (Figure 3). Mismatches throughout the R-loop weaken binding substantially, with typical effects ranging from 50-fold to approximately 1,000-fold. Although the effects of single mismatches are large, the resulting complexes are still quite stable, with sub-nanomolar affinity. The thermodynamic effects of mismatches do not display a detectable trend across the R-loop, in contrast to some expectations for seed regions (see Discussion).

**Mismatches Do Not Impact DNA Cleavage Rates after R-Loop Formation**

We next measured cleavage of each strand of a subset of the mismatched targets by using labeled targets as described earlier and saturating Cas12a concentrations (Figures 4A and S3D; Table S2). Mismatches decreased the observed rate constant for NTS cleavage in a manner that paralleled the slower R-loop formation determined from our binding data (Figure 4B). The correspondence of these rate constants across a range of $>100$-fold suggests that the rate of NTS cleavage is limited by propagation of the R-loop.

Because TS cleavage is approximately 10-fold slower than NTS cleavage for the matched target, we expected that mismatches that slow R-loop propagation, and thus NTS cleavage, by less than 10-fold would not decrease the observed rate of TS cleavage substantially. For mismatches that give larger effects, cleavage of both strands would be rate limited by R-loop propagation, so that the observed rate of TS cleavage would be decreased to that of R-loop propagation and NTS cleavage (modeled by the solid curve in Figure 4C). Indeed, the effects of both sets of mismatches conformed to this simple model.

**Cas12a Binding Remains Rate Limiting at a Physiological Mg$^{2+}$ Concentration**

Biological Mg$^{2+}$ concentrations are between 0.2 and 1 mM in mammalian nuclei (Gunther, 2006) and only slightly higher in bacterial cells (Lusk et al., 1968), much lower than the 5–10 mM used in most in vitro studies. To extend our kinetic framework to a cellular milieu, we also conducted reactions at the lower Mg$^{2+}$ concentration of 1 mM (Figure S4). Cas12a binding to its perfect target DNA was slowed by 30%, to 8.1 (±1.2) $\times$ 10$^{-7}$ M$^{-1}$ s$^{-1}$, and Cas12a (D908A) dissociation was accelerated 6-fold, to a value of 3.3 (±0.6) $\times$ 10$^{-5}$ s$^{-1}$. Cleavage of the NTS, with an observed rate constant of 0.020 ± 0.001 s$^{-1}$, remained orders of magnitude faster than Cas12a dissociation. Thus, at a near-physiological Mg$^{2+}$ concentration, Cas12a remains an efficient DNA nuclease. The increased dissociation rate demonstrates the importance of Mg$^{2+}$ for stability of the ternary complex, but the affinity remains very high ($K_D = 410 \pm 100$ fM), and DNA binding is still rate limiting for cleavage. Thus, it is likely that DNA binding remains rate limiting for cleavage in vivo and that target specificity is determined by differences in the rate constant for Cas12a binding, not differences in affinity.

**DISCUSSION**

Cas12a is currently a subject of intense interest for its simplicity and potential for high target specificity in vivo. Here, we provide a detailed kinetic mechanism that outlines a mechanistic basis for its extraordinary specificity. A quantitative kinetic analysis of the physical and chemical steps of DNA targeting shows that binding of Cas12a to target DNA involves two kinetically separable steps (Figure 5). An initial step likely represents PAM recognition, which is rapid and reversible. Next, the R-loop forms with a rate constant of $\sim$0.1 s$^{-1}$, as revealed by a plateau in the binding rate constant and by 2-AP fluorescence experiments. The same rate constant is obtained whether the 2-AP probe is near the PAM-proximal end or the PAM-distal end, indicating that the R-loop propagates quickly after initiating adjacent to the PAM. After R-loop formation, the NTS is cleaved rapidly, so that R-loop formation limits the cleavage rate under saturating conditions.
Following this cleavage event, there is additional trimming of 4 nt of the NTS and cleavage of the TS 10-fold more slowly. The DNA cleavage steps are not impacted detectably by single mismatches between the crRNA and the target DNA. Instead, the decreased observed cleavage rates arise from decreases in the rate of R-loop formation.

DNA cleavage is less precise than previously thought, as the NTS is trimmed toward the PAM and the TS is cleaved at multiple positions. These cleavage events imply that there is considerable mobility of the NTS and the RuvC domain and that there is flexibility in positioning of the TS in the nuclease active site. Additionally, the multiple cleavage events are consistent with recent data indicating that the RuvC domain becomes highly active upon R-loop formation, rapidly cleaving exogenous ssDNA (Chen et al., 2018; Li et al., 2018). Biologically, the trimming of the target DNA ends is significant, because it leads to non-cohesive DNA ends, which require filling in by a polymerase but could also be subjected to resection during DNA repair (Chang et al., 2017). Additionally, the rapid release of the PAM-distal cleavage product following TS cleavage (Singh et al., 2018; Figure 1E) would produce an unprotected 5' overhang available for binding by repair factors or annealing to a new DNA with microhomology.

Our kinetic analysis showed that substrate dissociation is orders of magnitude slower than DNA cleavage. This result indicates that DNA binding is rate limiting for cleavage under sub-saturating conditions, and, therefore, DNA binding controls the level of target specificity. Importantly, this behavior is robust, with DNA binding remaining rate limiting for cleavage at low, physiological Mg\(^{2+}\) concentration and for DNA targets that include single mismatches with the crRNA. Despite DNA binding being rate limiting for DNA cleavage, substantial penalties were observed for mismatches throughout the R-loop. This surprising result implies a late transition state for R-loop formation, with R-loop propagation being readily reversible until most of the 20 base pairs have formed (Figure 5). The late transition state may arise as a consequence of the concurrent displacement of the non-target DNA strand, allowing a reversible "exchange" reaction between the DNA duplex and the R-loop that may be augmented by features of Cas12a. Alternatively or in addition, features of the contacts of the R-loop and/or the NTS with Cas12a may contribute to the reversibility of R-loop formation.

The pattern of mismatch effects on Cas12a specificity is consistent with some descriptions of the roles and origins of a seed region but defies others. The decrease in binding kinetics and, consequently, in cleavage specificity is generally larger in PAM-proximal than PAM-distal positions, supporting descriptions of a seed region that invoke increased specificity close to the PAM (Gorski et al., 2017; Swarts et al., 2017). On the other hand, substantial specificity remains present for more PAM-distal positions, which differs from Cas9 and other studied enzymes that use seed regions for target recognition (Jinek et al., 2012; Wang et al., 2009). Additionally, the effects of mismatches on DNA target affinity are uniform across the R-loop, in contrast to descriptions of seed regions that invoke higher thermodynamic specificity originating from pre-ordering of the crRNA in the PAM-proximal region (Gorski et al., 2017). The kinetic origin of the seed region for Cas12a probably reflects that the R-loop...
forms through a defined pathway, starting from the PAM, and that the PAM-proximal base pairs are formed in the transition state for R-loop formation to a greater extent than the PAM-distal base pairs.

Our results have important implications for applications of Cas12a in vivo. The finding that Cas12a binding to target DNA is rate limiting for cleavage suggests that the levels of discrimination against off-target cleavage in cells are determined by the extent to which the rate of target binding, not the affinity, is decreased by a given mismatch (Bisaria et al., 2017). DNA supercoiling in vivo may facilitate rate-limiting R-loop formation, as observed for Cas9 and the multisubunit Cascade complex (Szczechun et al., 2014; Westra et al., 2012), but there is no expectation of a change in the magnitudes of the mismatch penalties or their distribution within the R-loop. Indeed, the reported patterns of off-target DNA cleavage in vivo show decreased discrimination at the PAM-distal end and in the center of the R-loop (Kim et al., 2016; Kim et al., 2017; Kleinstiver et al., 2016), closely resembling the effects of mismatches on the binding rate constant rather than effects on equilibrium binding (Figure 3).

Our results also highlight a key difference between Cas12a and Cas9. As noted earlier, the positions of strong discrimination against mismatches are confined to a PAM-proximal “seed” region for Cas9, as measured both in cells by DNA cleavage and in vitro by binding experiments (Boyle et al., 2017; Hsu et al., 2013; Singh et al., 2016; Sternberg et al., 2014). Analogous to our results for Cas12a and supported by recent kinetics experiments on Cas9 (Gong et al., 2018), it appears likely that DNA binding also limits the rate of cleavage by Cas9 in vivo. Thus, the limited seed region for Cas9 specificity implies that the transition state for R-loop formation is reached upon formation of fewer base pairs than for Cas12a. We conclude, from this difference, that the position of the transition state for binding can be influenced by properties of the Cas endonuclease, not just by the intrinsic reversibility of the R-loop, which would be equivalent for Cas12a and Cas9. Further, the variation in transition state position suggests the possibility of engineering nucleases to shift the transition state even further into the PAM-distal region of the R-loop, extending the region and the overall levels of specificity for gene editing applications.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCE TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **METHOD DETAILS**
  - Cas12a Cloning and Purification
  - DNA Substrate Preparation
  - Cas12a-crRNA Assembly
  - Measurement of Cas12a binding kinetics
  - Measurement of Cas12a dissociation kinetics
  - Measurement of DNA strand cleavage rates
  - Measurement of R-loop propagation rate
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND SOFTWARE AVAILABILITY**

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes four figures and two tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.06.043.

**ACKNOWLEDGMENTS**

We thank Prof. Ken Johnson and Dr. Shanzhong Gong for use of the stopped-flow fluorimeter and for guidance with the 2-AP fluorescence experiments, Prof. Along Ke for providing a Twin-Strep-Sumo expression vector, and Prof. Dan Herschlag and members of the Russell and Finkelstein labs for comments on the manuscript. This work was supported by NIGMS grants P01GM066275 (to R.R., co-investigator) and R01GM124141 (to I.J.F.) and by Welch Foundation grants F-1563 (to R.R.) and F-1808 (to I.J.F.).

**AUTHOR CONTRIBUTIONS**

Conceptualization, all authors; Methodology, I.S., F.A.S., I.J.F., and R.R.; Investigation, I.S. and F.A.S.; Writing – Original Draft, I.S. and R.R.; Writing – Review & Editing, all authors; Funding Acquisition, I.J.F. and R.R.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: March 7, 2018
Revised: June 13, 2018
Accepted: June 26, 2018
Published: August 2, 2018

**REFERENCES**


STAR METHODS

KEY RESOURCE TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial and Virus Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli BL21 star (DE3)</td>
<td>Thermo Fisher</td>
<td>Cat. No.: C601003</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AsCas12a</td>
<td>Finkelstein Lab</td>
<td>N/A</td>
</tr>
<tr>
<td>AsCas12a(D908A)</td>
<td>Finkelstein Lab</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Deposited Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyzed data</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>Raw data</td>
<td>Mendeley dataset</td>
<td><a href="https://doi.org/10.17632/w8d5fn5vxn.1">https://doi.org/10.17632/w8d5fn5vxn.1</a></td>
</tr>
<tr>
<td><strong>Oligonucleotides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min. pre-crRNA: 5'-UAAUUUCUCUCUUGUAGA UUGUGAUAAGUGGAAUGCCAUGUGGA</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>DNA Oligos listed in Supplemental Info</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This paper, Table S2</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td><strong>Recombinant DNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pIF254- Cas12a</td>
<td>Finkelstein Lab</td>
<td>N/A</td>
</tr>
<tr>
<td>pIF255- Cas12a(D908A)</td>
<td>Finkelstein Lab</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Software and Algorithms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinetic Explorer</td>
<td>KinTek Corporation</td>
<td>Version 6.1</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strep-Tactin Superflow</td>
<td>IBA Life Sciences</td>
<td>Cat. No.: 2-4033-001</td>
</tr>
<tr>
<td>HiLoad 16/600 Superdex pg</td>
<td>GE Healthcare</td>
<td>Cat. No.: 28989335</td>
</tr>
</tbody>
</table>

CONTACT FOR REAGENT AND RESOURCE SHARING

Requests for further information and reagents should be directed to the Lead Contact, Rick Russell (rick_russell@cm.utexas.edu).

METHOD DETAILS

Cas12a Cloning and Purification

An *E. coli* codon-optimized gene encoding AsCas12a was purchased from BioMatik. The gene was cloned into a custom pET-19 based expression vector pIF254 (WT) and pIF255 (D908A) and transformed into BL21 star (DE3) cells (Thermo Fisher). Cas12a contained a His6/Twin-Strep/SUMO N-terminal fusion. Single colonies were used to inoculate Terrific Broth supplemented with 50 μg/mL carbenicillin at 37°C. An aliquot of the starter culture (4 mL) was then used to inoculate 1 L of Terrific Broth. When the culture reached an OD600 of ~0.6, 1 mM IPTG was added and cultures were further incubated at 18°C for 24 hr. Cells were then lysed in 20 mM Na-HEPES, pH 8.0, 1 M NaCl, 1 mM EDTA, 5% glycerol, 0.1% Tween-20, 1 mM PMSF, 2000 U DNase (GoldBio), and 1X HALT Protease Inhibitor (Thermo Fisher) and homogenized. The lysate was clarified by centrifugation and was applied to a hand-packed Strep-Tactin Superflow gravity column (IBA Life Sciences). Cas12a was eluted with 20 mM Na-HEPES, 1 M NaCl, 5 mM MgCl2, and 5 mM dithiothreitol (DTT). Fractions containing Cas12a were determined by SDS-PAGE, pooled and concentrated to ~10 μM using a 30 kDa centrifuge concentrator (Millipore) (Figure S1C). Small aliquots were flash frozen in liquid nitrogen and stored at −80°C.

DNA Substrate Preparation

Oligonucleotides corresponding to the TS and NTS (Integrated DNA Technologies) (Table S1) were 5'-radiolabeled with [γ-32P]ATP (Perkin-Elmer) using T4 polynucleotide kinase (New England Biolabs). Radiolabeled oligonucleotides were purified by
polyacrylamide gel electrophoresis, eluted in TE buffer, and stored at -20°C. Target DNA duplexes were formed by incubating either the radiolabeled TS or NTS (25 nM) with four-fold excess of the unlabeled complementary strand in annealing buffer (50 mM Na-MOPS pH 7.0, 100 mM NaCl). The mixture was heated at 90°C for 5 min, slow-cooled to 25°C, and diluted to ~1 nM duplex (corresponding to the calculated concentration of the radiolabeled strand).

**Cas12a-crRNA Assembly**

The Cas12a-crRNA complex was assembled for experiments each day using purified Cas12a and a minimal precursor crRNA (pre-crRNA) that had a single uridine upstream of the Cas12a cleavage site (purchased from IDT). The crRNA precursor sequence is: 5'-UAUUUUCUCUUCUGUAGAUGUGAAAGCGAUGUGGA. Assembly reactions were performed by incubating Cas12a with a 2-fold excess of the precursor crRNA for 25 min at 25°C in 50 mM Na-MOPS, pH 7.0, 120 mM NaCl, 5 mM MgCl$_2$, 2 mM DTT, and 0.2 mg/ml bovine serum albumin. The assembled complex was then diluted, maintaining the same buffer conditions, and stored on ice until use.

**Measurement of Cas12a binding kinetics**

Reactions were performed in the same buffer conditions as the assembly reaction. DNA binding was initiated by the addition of a trace amount of target DNA duplex, radiolabeled on the NTS, to various concentrations of the assembled Cas12a-crRNA complex (1-10 nM). At various time points, 2 µl of the 20 µl binding reactions were withdrawn and added to 4 µl of an ice-cold chase solution consisting of 3X loading dye (15% glycerol, 5 mM Tris-Cl, pH 8.0, 0.015% xylene cyanol) supplemented with 50 mM Na-MOPS, pH 7.0, 120 mM NaCl, 20 mM EDTA, and 100 nM unlabeled perfect target duplex. Control reactions in which the unlabeled chase duplex were premixed before addition of Cas12a-crRNA resulted in minimal binding of the labeled DNA (< 10%), indicating that the unlabeled DNA target is an effective quench. Samples were electrophoresed on a 12% native gel at 4°C, exposed to a phosphorimager screen overnight, and then scanned using a Typhoon FLA 9500 (GE Healthcare). Bands of radioactivity corresponding to unbound and bound DNA complexes were quantified using ImageQuant 5.2 (GE Healthcare). For experiments performed at 1 mM Mg$^{2+}$, Cas12a was first assembled with the crRNA under standard conditions at 5 mM Mg$^{2+}$ and then diluted to 1 mM Mg$^{2+}$ prior to initiation of reactions. Values of $k_{on}$ and $k_{off}$ were determined from hyperbolic fits to the protein concentration dependence of the observed binding rates. In these plots, data points also included the $k_{off}$ value as a y-intercept point. For mismatched DNA targets, fits were constrained with $K_{d,2}$ = 1 nM as warranted, such that the decreases in $k_{max}$, which were highly constrained by the data, led to corresponding decreases in the $k_{on}$ values.

**Measurement of Cas12a dissociation kinetics**

A DNase-inactive mutant (Cas12a(D908A)) was used for dissociation reactions in the presence of Mg$^{2+}$. Purified Cas12a(D908A) was loaded with the minimal pre-crRNA as described above. Cas12a(D908A)-crRNA was then mixed with a trace amount of target DNA duplex labeled on the NTS for at least one hour at 25°C to allow complete binding of the duplex. For mismatch targets that bound slowly, binding incubation periods were increased to 2 hr. Dissociation reactions were performed at 10 nM Cas12a(D908A) and initiated by adding the unlabeled perfect target DNA in 50-fold excess over the protein concentration to ensure no further binding or re-binding of labeled DNA during the time course. Aliquots (2 µl) were withdrawn over the course of ~24 hr, mixed with 2 µl of ice-cold 6X loading dye (30% glycerol, 10 mM Tris-Cl, pH 8.0, 0.03% xylene cyanol), and stored at 4°C. Samples were electrophoresed on a 12% native gel at 4°C, imaged with a phosphorimager, and analyzed as described above. For experiments performed at 1 mM Mg$^{2+}$ or in the absence of Mg$^{2+}$, Cas12a (wild-type or D908A) was assembled with the crRNA as described above and then diluted to 1 mM Mg$^{2+}$ during the DNA loading step or chelated with 10 mM EDTA.

**Measurement of DNA strand cleavage rates**

DNA duplexes were radiolabeled at the 5' end of either the TS or NTS. Cleavage reactions were performed using 2, 10, and 50 nM Cas12a (pre-assembled with crRNA) that were initiated by addition of a trace amount of labeled target DNA. At various times, 2 µl samples were quenched in 4 µl of denaturing quench (90% formamide, 20 mM EDTA, 0.01% bromophenol blue, and 0.005% xylene cyanol). Our denaturing quench also included 300 nM of an oligonucleotide with no complementarity to our target sequences to reduce retention on microfuge tube walls and in the gel wells. Samples were heated to 95°C for 3 min and immediately loaded onto a 20% denaturing polyacrylamide gel containing 7 M urea. Gels were exposed and imaged as described above. Bands corresponding to cleaved and uncleaved DNA substrate were quantified. For experiments performed at 1 mM Mg$^{2+}$, Cas12a was first assembled with the crRNA under standard conditions at 5 mM Mg$^{2+}$ and then diluted to 1 mM Mg$^{2+}$ when it was transferred into the cleavage reactions.

To model the expected dependence of the observed rate constant for TS cleavage on the R-loop formation rate constant, we used the inverse of the sum of the transit times (curve in Figure 4C), with the R-loop formation rate constant varied and the rate constant for TS cleavage held constant (1/$k_{obs} = 1/k_{R-loop} + 1/k_{TS}$). Simulations using Kinetic Explorer software (Kintek Corp.) for select combinations of rate constants showed good agreement between the values of $k_{obs}$ from this approximation and first-order fits to the simulated progress curves.
Measurement of R-loop propagation rate

NTSs were purchased as oligonucleotides with 2-aminopurine substituted at position 4 or 18 (IDT). The target duplex was formed by annealing the NTS and TS (10 μM each) in a thermocycler in 2-AP Annealing Buffer (10 mM Tris-Cl, pH 8.0, 50 mM NaCl, 1 mM EDTA). The duplex was then diluted to 600 nM with Running Buffer (100 mM NaCl, 50 mM Tris pH 8.0, 5 mM MgCl₂, 2 mM DTT) in an amber tube. Cas12a was assembled with a 2-fold excess of precursor crRNA for 25 min at room temperature in binding buffer (50 mM Tris-Cl, pH 8.0, 100 mM NaCl, 5 mM MgCl₂, 2 mM DTT). The SF 2004 series stopped-flow (KinTek Corp.) ports were washed with 600 μl (port volume) of 2 N NaOH, 2 N HCl and 5 X 600 μl of deionized water before experiments. For each experiment, 300 μl of the target DNA was loaded into port A and 300 μl of 1 μM Cas12a solution was loaded into port B. Triplicate datasets were combined and analyzed using the KinTek Global Kinetic Explorer to obtain rates.

QUANTIFICATION AND STATISTICAL ANALYSIS

All binding, dissociation, and cleavage reactions were performed at least three times at the indicated enzyme concentration. Rates of dissociation are shown as the mean ± SEM of the time courses fit by an exponential decrease. Rate constants for binding and cleavage are shown as the mean fit ± SE of the hyperbolic fit of the observed rate constants as a function of Cas12a concentration.

All gel-based assays were analyzed using ImageQuant 5.2 (GE Healthcare) and all data fitting was performed with Kaleidagraph. Weighted data fitting using nonlinear regression was used to determine rates, amplitudes, and standard errors of the fit.

Stopped-flow assays were performed in replicates and analyzed with KinTek Global Kinetic Explorer.

DATA AND SOFTWARE AVAILABILITY

The raw data reported in this paper are available from Mendeley Datasets at https://doi.org/10.17632/w8d5fn5vxn.1.