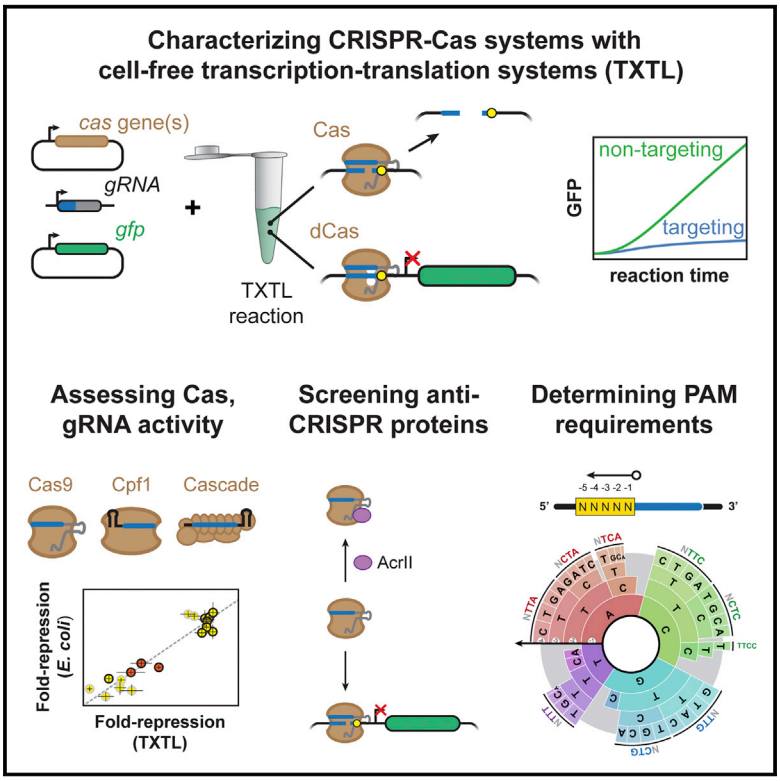
Technology

Molecular Cell

Rapid and Scalable Characterization of CRISPR Technologies Using an *E. coli* Cell-Free Transcription-Translation System

Graphical Abstract



Highlights

- Active expression of multiple CRISPR nucleases and gRNAs in an *E. coli* TXTL system
- Repression activity of dCas9 strongly correlated between TXTL and *E. coli*
- A TXTL-based PAM assay allowed characterization of five Cpf1 nucleases
- Determined specificities of 24 anti-CRISPR proteins against five Cas9 nucleases

Authors

Ryan Marshall, Colin S. Maxwell, Scott P. Collins, ..., Yonghua He, Chase L. Beisel, Vincent Noireaux

Correspondence

cbeisel@ncsu.edu (C.L.B.), noireaux@umn.edu (V.N.)

In Brief

Marshall et al. demonstrate that an *E. coli* cell-free transcription-translation (TXTL) system can be used to improve the speed and scalability of characterizing CRISPR nucleases and their accessory factors. The method will facilitate the discovery of uncharacterized CRISPR nucleases and anti-CRISPR proteins and aid the validation of designed gRNAs.

Marshall et al., 2018, Molecular Cell 69, 146–157
January 4, 2018
https://doi.org/10.1016/j.molcel.2017.12.007

CellPress

CellPress

Molecular Cell
Technology

Rapid and Scalable Characterization of CRISPR Technologies Using an *E. coli* Cell-Free Transcription-Translation System

Ryan Marshall,[1,4] Colin S. Maxwell,[2,4] Scott P. Collins,[2] Thomas Jacobsen,[2] Michelle L. Luo,[2] Matthew B. Begemann,[3] Benjamin N. Gray,[3] Emma January,[3] Anna Singer,[3] Yonghua He,[3] Chase L. Beisel,[2,*] and Vincent Noireaux[1,5,*]
[1]School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
[2]Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27695, USA
[3]Benson Hill Biosystems, St. Louis, MO 63132, USA
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: cbeisel@ncsu.edu (C.L.B.), noireaux@umn.edu (V.N.)
https://doi.org/10.1016/j.molcel.2017.12.007

SUMMARY

CRISPR-Cas systems offer versatile technologies for genome engineering, yet their implementation has been outpaced by ongoing discoveries of new Cas nucleases and anti-CRISPR proteins. Here, we present the use of *E. coli* cell-free transcription-translation (TXTL) systems to vastly improve the speed and scalability of CRISPR characterization and validation. TXTL can express active CRISPR machinery from added plasmids and linear DNA, and TXTL can output quantitative dynamics of DNA cleavage and gene repression—all without protein purification or live cells. We used TXTL to measure the dynamics of DNA cleavage and gene repression for single- and multi-effector CRISPR nucleases, predict gene repression strength in *E. coli*, determine the specificities of 24 diverse anti-CRISPR proteins, and develop a fast and scalable screen for protospacer-adjacent motifs that was successfully applied to five uncharacterized Cpf1 nucleases. These examples underscore how TXTL can facilitate the characterization and application of CRISPR technologies across their many uses.

INTRODUCTION

CRISPR technologies have proven to be broadly useful genome-editing tools for biomolecular research, biotechnology, human health, and agriculture (Barrangou and Doudna, 2016). These technologies rely on RNA-guided nucleases derived from prokaryotic CRISPR-Cas immune systems (Mohanraju et al., 2016). The nuclease specifically cleaves DNA or RNA sequences complementary to the guide portion of the RNA and flanked by a protospacer-adjacent motif (PAM). For DNA targets, this ability has allowed programmable DNA damage or repair-mediated genome editing. Furthermore, by disrupting endonuclease activity, these nucleases can be readily converted into programmable nucleic acid-binding proteins and used for applications in gene activation (CRISPRa) or repression (CRISPRi), base editing, and real-time imaging (Chen et al., 2016; Dominguez et al., 2016; Komor et al., 2017; Nelles et al., 2016; Qi et al., 2013).

While the vast majority of CRISPR-based technologies have relied on the DNA-targeting Cas9 nuclease from *Streptococcus pyogenes* (SpyCas9) (Barrangou and Doudna, 2016; Jinek et al., 2012), nature boasts a diverse collection of CRISPR-Cas systems that are currently sub-divided into two classes, six types, and 33 subtypes (Koonin et al., 2017; Makarova et al., 2015; Shmakov et al., 2015). Interrogation of the emerging subtypes has revealed nucleases with widely varying properties that can be smaller, recognize different PAMs, degrade DNA, target RNA, exhibit reduced propensity for off-target effects, or be more amenable to multiplexing in comparison to SpyCas9 (Abudayyeh et al., 2016; Kim et al., 2017; Kleinstiver et al., 2016; Mulepati and Bailey, 2013; Zetsche et al., 2017). Separately, the discovery of anti-CRISPR proteins that inhibit Type I-E, I-F, II-A, and II-C CRISPR-Cas systems offers means to tightly control nuclease activity (Bondy-Denomy et al., 2015; Pawluk et al., 2016a, 2016b; Rauch et al., 2017), with the likely existence of similar inhibitors for the other subtypes. However, despite this diversity, the vast majority of these proteins have been slow to be adopted as CRISPR technologies.

One major bottleneck is the long and tedious process of characterizing these proteins' basic properties and functions. To date, characterization has been performed with multiple methodologies based on *in vitro* biochemical assays or live cells. More recently, these assays have been modified for high-throughput analysis of PAM-binding requirements, off-target propensities, or large libraries of guide RNAs (gRNAs) through next-generation sequencing or imaging of arrayed nucleotides (Abudayyeh et al., 2016; Boyle et al., 2017; Jiang et al., 2013; Karvelis et al., 2015; Kleinstiver et al., 2016; Leenay et al., 2016; Pattanayak et al., 2013; Tsai et al., 2017). However, these assays consistently require days to weeks to perform due to the requirement for protein purification or for culturing and transforming live cells. Furthermore, these assays scale poorly when testing large sets of proteins or gRNAs. Given the growing

abundance of known CRISPR nucleases, the ease in gRNA design, and the growing prevalence of factors that interface with CRISPR-Cas systems, there remains a pressing need to develop rapid and scalable characterization methodologies.

DESIGN

Here, we address this need using an *Escherichia coli* cell-free transcription-translation (TXTL) system (Garamella et al., 2016; Shin and Noireaux, 2012). The system is based on *E. coli* lysates prepared from exponentially growing cells to preserve the natural transcriptional, translational, and metabolic machinery. This machinery contrasts with traditional TXTL systems, which rely on RNA polymerases from bacteriophages such as T7 as well as purified ribosomes. DNA encoding genes of interest can be added to the lysates in different amounts, resulting in gene expression within minutes to hours. While plasmid DNA is commonly used, TXTL can also accommodate linear DNA as long as RecBCD is inhibited using GamS or linear DNA containing Chi sites (Marshall et al., 2017). The TXTL reactions can also be conducted in volumes as small as a few microliters, allowing TXTL reactions to be scalably used in microtiter plates. Note that although we generally prepared our own *E. coli* lysates, a system based on our lysate-production protocols is available commercially (STAR Methods).

To express CRISPR-Cas systems, DNA encoding the cas genes and gRNAs are added to the TXTL mix, resulting in the generation of active nucleoprotein complexes. To measure activity of the complexes in TXTL, we use a reporter construct expressing the fluorescent protein deGFP, a slightly modified version of eGFP with identical fluorescent properties (Shin and Noireaux, 2012). Targeting the construct results in changes in deGFP fluorescence levels that can be measured quantitatively in real time. While deGFP or other fluorescent proteins lend to rapid measurement using a fluorescence microplate reader, other standard colorimetric or luminescence readouts like violacein or firefly luciferase can be used. Here, we demonstrate the utility of TXTL across a diverse set of Cas nucleases and show how it can be used to predict gRNA activity *in vivo*, characterize anti-CRISPR proteins, and elucidate recognized PAMs. Based on these findings, we expect TXTL to provide a powerful characterization tool for the expanding universe of CRISPR-Cas systems and proteins, and we show the applicability of TXTL beyond biomanufacturing, diagnostics, and genetic circuit prototyping (Carlson et al., 2012; Dudley et al., 2015; Garamella et al., 2016; Gootenberg et al., 2017; Hockenberry and Jewett, 2012; Jewett et al., 2008; Kanter et al., 2007; Karzbrun et al., 2014; Pardee et al., 2016; Sun et al., 2014; Swartz, 2006; Takahashi et al., 2015a, 2015b; Tayar et al., 2015). See Methods S1, Protocol 1 for more information about how to conduct TXTL with CRISPR.

RESULTS

SpyCas9 and dSpyCas9 Exhibit Robust Activity in TXTL

We initially examined the activity of the SpyCas9 nuclease. To monitor the dynamics of DNA cleavage by SpyCas9, we designed single-guide RNAs (sgRNAs) that target within the promoter, 5′ UTR, and coding sequence of a deGFP reporter construct. We measured the dynamics of deGFP production following the addition of the SpyCas9 plasmid, linear DNA encoding the sgRNA, and the deGFP reporter plasmid (Figure 1A). We found that the three tested sgRNAs resulted in greatly reduced deGFP concentrations in comparison to a non-targeting sgRNA (Figure 1B). Measurable repression was observed beginning after less than 1 hr into the TXTL reaction, indicating that this time span was required to express and assemble the Cas9-sgRNA ribonucleoprotein (RNP) complex and for the complex to bind and cleave the target DNA (Figures S1A and S1B). The rate of deGFP production then dropped quickly to approximately zero, consistent with irreversible DNA cleavage as confirmed by PCR amplification of the target site (Figure S1C). Interestingly, the onset of deGFP repression and the rate at which deGFP production dropped varied depending on the sgRNA used (Figures S1A and S1B), suggesting that the dynamics by which different guides target Cas9 to DNA can vary. These results indicate that an active SpyCas9-sgRNA complex can be expressed directly in TXTL, providing a dynamic and quantitative readout of nuclease activity in a few hours.

We also interrogated the activity of the catalytically dead version of SpyCas9 (dSpyCas9) commonly used for programmable DNA binding and gene regulation (Gilbert et al., 2014; Qi et al., 2013). DNA binding by dSpyCas9 can block transcriptional initiation or elongation of RNA polymerase in bacteria, offering a means to link DNA binding with reporter expression in TXTL (Bikard et al., 2013; Qi et al., 2013). To assess the expression and regulatory activity of dSpyCas9, we measured the concentration of deGFP over time in TXTL reactions for the same three targeting sgRNAs and the non-targeting sgRNA. Similar to reactions expressing SpyCas9, reactions expressing dSpyCas9 exhibited consistent deGFP repression, with deGFP production dropping after less than an hour (Figures S1A and S1B). dSpyCas9 repressed expression less strongly and exhibited some deGFP production even at the end of the TXTL reaction (Figures 1B and S1A) due to lack of irreversible target cleavage (Figure S1B). We also observed different extents of deGFP fold repression across the three targeting sgRNAs, similar to repression strengths reported in bacteria (Bikard et al., 2013; Qi et al., 2013).

To begin assessing the scalability of TXTL reactions, we expanded from one promoter to four promoters, each targeted by two sgRNAs and driving expression of deGFP. Each promoter is dependent on a unique alternative sigma factor (σ^{28}, σ^{38}, and σ^{54}) or T7 RNA polymerase, where each transcription factor is supplied on an added expression plasmid under the control of a σ^{70} promoter (Garamella et al., 2016; Shin and Noireaux, 2012) (Figure 1C). By including a *degfp*-targeting sgRNA, a non-targeting sgRNA, and a σ^{70} promoter, we tested a total of 50 promoter-sgRNA combinations. Each sgRNA was designed to target either strand across the recognition elements of its cognate promoter; the only exception was the P_{28a} promoter, where both sgRNAs were designed to target the non-template strand due to the lack of an NGG PAM on the template strand. The TXTL reactions confirmed that measurable repression was only observed when an sgRNA was matched with its target (Figure 1C), with the strength of repression ranging between 7-fold and 105-fold. We also targeted binding sites for the NtrC operator sites within

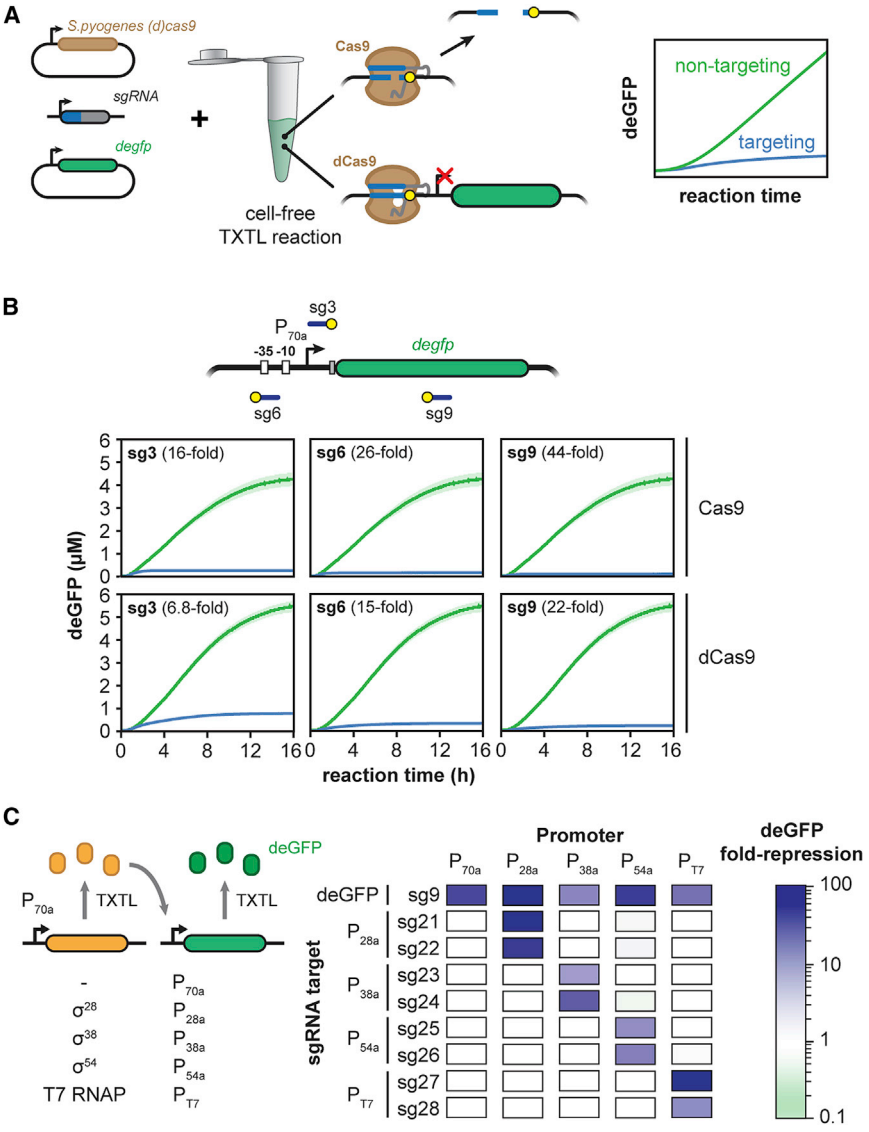


Figure 1. *S. pyogenes* Cas9 Functions Efficiently in TXTL

(A) Schematic of using TXTL to dynamically and quantitatively measure the activity of Cas9 and dCas9.

(B) Time series showing deGFP concentration for cell-free reactions expressing (d)Cas9 and a non-targeting sgRNA (green) or targeting sgRNAs (blue). Target locations include the sequence matching the guide (blue line) and the PAM (yellow circle). Error bars represent the SEM from at least six repeats.

(C) Alternative sigma factors σ^{28}, σ^{38}, and σ^{54} and the T7 polymerase can be expressed in TXTL from the P_{70a} promoter and activate their cognate promoters P_{28a}, P_{38a}, P_{54a}, and P_{T7}, respectively. A matrix showing dSpyCas9-based repression of promoters dependent on σ^{28}, σ^{38}, σ^{54}, and the T7 polymerase is shown. An sgRNA targeting each promoter or the *gfp* gene body was expressed along with each sigma factor or polymerase and a reporter gene driven by the sigma factor of its cognate promoter. Values represent the mean of at least three repeats.

the σ^{54} promoter (Figure S1D), demonstrating that these sites were important for reporter expression.

Multiple Factors Impact the Measured Activity of dSpyCas9 in TXTL

After characterizing SpyCas9 and dSpyCas9 in TXTL, we sought to determine parameters that affect the measured activity. First, we found that dSpyCas9 was limiting based on more dSpyCas9 plasmid reducing the total amount of GFP produced in the TXTL reaction (Figure S2A). Surprisingly, adding more dSpyCas9 plasmid elevated deGFP production for the non-targeting sgRNA, underscoring the need to add the same total amount of DNA such as when comparing the activity of different targeting sgRNAs.

Second, we evaluated the impact of destabilizing deGFP to emulate reporter turnover or dilution *in vivo*. We appended an ssrA degron tag recognized by the ClpXP protease to the C terminus of deGFP (Figure S2B) (Schmidt et al., 2009). Appending the tag resulted in greater total dSpyCas9-based repression (Figure S2C) due to the inability of deGFP to accumulate in the TXTL reaction.

We next evaluated how CRISPR-based repression is influenced by linear versus plasmid DNA. Linear DNA can be generated without cloning and thus can be readily used in TXTL (Marshall et al., 2017; Sun et al., 2014). However, some Cas nucleases have been shown to require supercoiled DNA for binding (Westra et al., 2012), suggesting that targets encoded on relaxed, linear DNA may not allow for efficient repression. We measured the extent of dSpyCas9-based repression when targeting a plasmid or linear version of the deGFP reporter construct. We found that dSpyCas9 was capable of repressing gene expression when targeting the reporter gene on the linear template; however, negligible repression was observed when targeting the promoter region (Figure 2A). These results suggest that dSpyCas9 can block elongation, but not initiation, of the *E. coli* RNA polymerase on linear DNA. Regardless of the underlying mechanism, linear DNA should not be used when assessing the ability of dCas9 to block transcription initiation.

When DNA expressing dSpyCas9, an sgRNA, and the fluorescent reporter were added together at the beginning of a TXTL experiment, we observed a transient period in which the reporter gene was expressed before the onset of dSpyCas9-based repression (Figure 1B). We hypothesized that this period of transient expression was due to the slow expression and assembly of the dSpyCas9-sgRNA RNP complex. To test this hypothesis, we varied our initial protocol to pre-express

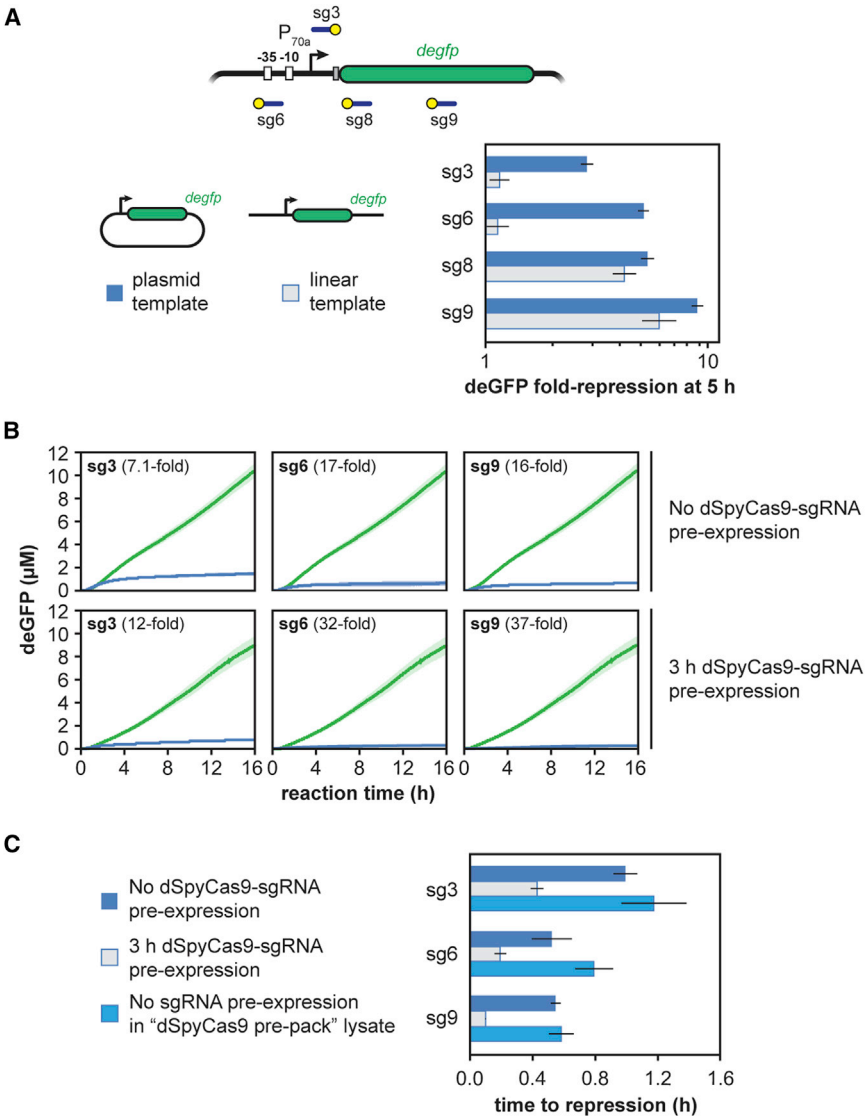


Figure 2. Multiple Factors Affect dSpyCas9-Based Repression of Reporter Gene Expression in TXTL

(A) Fold repression produced by a TXTL reaction when deGFP is expressed from either a targeted plasmid (dark) or linear (light) construct. Error bars represent the SEM from at least three repeats.

(B) Time series showing deGFP concentration in TXTL for cell-free reactions expressing dSpyCas9 and a targeting sgRNA. The reporter plasmid is added to the reaction either at the same time as dSpyCas9 and the sgRNA (top row) or after 3 hr (bottom row). Error bars represent the SEM from at least five repeats.

(C) Time to repression for the curves from (B), as well as for "dSpyCas9 pre-pack." Error bars represent the SEM from at least five repeats.

dSpyCas9 and the sgRNA before adding the reporter plasmid for 3 hr. We reasoned that this period of pre-expression would allow the RNP to be expressed and assemble and that this would shorten the time until we observed repression of the reporter gene. Consistent with our hypothesis, pre-expressing dSpyCas9 and the sgRNA reduced the time before the reporter gene was repressed (Figure 2B), with measurable repression occurring as fast as 6 min (Figures 2C and S2D). We also evaluated the effect of pre-expressing dSpyCas9 in cells prior to generating the lysate in a "dSpyCas9 pre-pack" mix (Figure S2E). Interestingly, reactions in this lysate exhibited similar times to repression as reactions with no dSpyCas9 or sgRNA pre-expression (Figure 2C), indicating that assembly of the RNP complex primarily contributes to the delayed onset of dSpyCas9-based repression. Taken together, these results suggest that dCas9 RNP assembly is relatively slow (on the order of 30 min), but DNA binding is fast (on the order of 5 min), in line with the timescales of Cas9:gRNA complex formation and DNA binding *in vitro* in the presence of non-specific RNA competitors (Mekler et al., 2016).

The Strength of dSpyCas9-Based Repression Strongly Correlates between TXTL and *E. coli*

Given the speed and scalability of employing TXTL to characterize CRISPR nucleases and gRNAs, an ensuing question is how well the quantified activity correlates to *in vivo* settings and recapitulates known phenomena (Chappell et al., 2013). To begin addressing this question, we compared dSpyCas9-based repression in TXTL and in *E. coli* by targeting 19 different locations within the deGFP reporter plasmid (Figure 3A). The particular locations were chosen to also evaluate the impact of strong PAMs (NGG) and weak PAMs (NAG) as well as targeting the template or non-template strand within the promoter or transcribed region. We chose these comparisons because Cas9 recognizes NAG PAMs more weakly than NGG PAMs (Boyle et al., 2017; Jiang et al., 2013) and exhibited impaired gene repression in bacteria when targeting the template strand within the transcribed region versus any other location (Bikard et al., 2013; Qi et al., 2013). The experiments were conducted by encoding the dSpyCas9, sgRNA, and deGFP reporter on separate, compatible plasmids for parallel testing in TXTL and in *E. coli*.

These experiments revealed a strong correlation ($R^2 = 0.90$) between the measurements in TXTL and in *E. coli* (Figure 3A). The correlation was based on a log-log plot rather than a linear-scale plot due to differences between continuous cellular processes and batch TXTL processes. Consistent with the strong correlation, we observed greatly reduced repression for targets flanked by NAG versus NGG PAMs. Furthermore, repression was consistently weaker when targeting the template strand of the transcribed region in comparison to targeting the non-template strand of the transcribed region or either strand within the vicinity

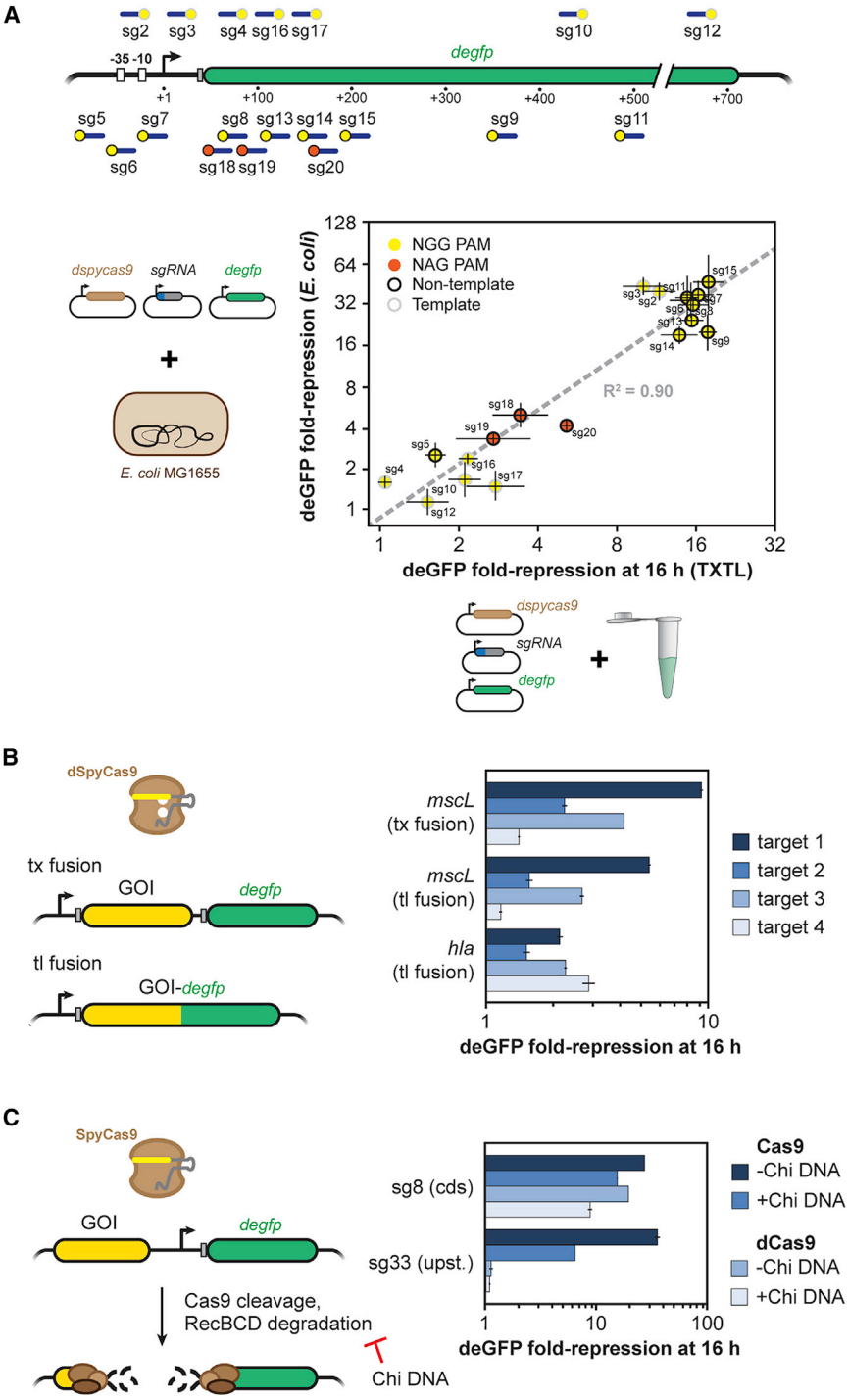


Figure 3. TXTL Can Be Used to Assess the Activity of sgRNAs

(A) A schematic of where each guide binds in the *gfp* promoter and gene body (top). The location of the target and PAM is indicated by a blue line and a yellow or orange dot, respectively. The fold repression of GFP production by dCas9-based repression for each sgRNA *in vivo* and *in vitro* (bottom). Points are colored by whether the guide is adjacent to an NGG (yellow) or NAG (orange) PAM, and whether the sgRNA targets the non-template strand (black ring) or template strand (gray ring). Error bars represent the SEM from at least three repeats.

(B) Assessing non-*gfp* targeting sgRNAs used by dCas9. The sequence or gene of interest is transcriptionally or translationally fused upstream of *degfp*. Fold repression was measured in TXTL for four targeting sgRNAs when *degfp* is fused to *mscL* or *hla*. Error bars represent the SEM from at least three repeats.

(C) Assessing non-*gfp* targeting sgRNAs used by Cas9. The sequence or gene of interest was inserted upstream of the promoter driving expression of deGFP. In the absence of a RecBCD inhibitor, cleavage by Cas9 leads to rapid degradation of the plasmid and loss of GFP expression. Fold repression was measured in TXTL when targeting in the *gfp* coding sequence (sg8) or upstream of the promoter (sg8) with Cas9 or dCas9, and in the presence or absence of the RecBCD inhibitor Chi site containing DNA. Error bars represent the SEM from at least three repeats.

of the promoter. This direct comparison therefore showed that TXTL can reasonably predict *in vivo* activity of CRISPR-Cas systems, at least for gene repression by dSpyCas9 in *E. coli*.

The Activity of Targeting Other Genes Can Be Measured with TXTL

To demonstrate that we can target virtually any sequence or gene, we devised schemes to indirectly link the activity of dCas9 or Cas9 to deGFP expression. In the case of dCas9, our scheme involved introducing the target sequence or gene as a transcriptional or translational fusion to the 5′ end of the *degfp* gene. Based on this setup, any sgRNA-mediated transcriptional silencing would downregulate the downstream expression of deGFP. As a proof-of-principle demonstration, we fused the *Staphylococcus aureus* alpha-hemolysin gene *hla* or the *E. coli* mechanosensory channel gene *mscL* to the N terminus of eGFP, or we introduced *mscL* upstream of deGFP as an operon-like DNA assembly with separate ribosome-binding sites. Two sets of four different sgRNAs were designed to target the non-template strand of *hla* or *mscL*. For each of the fusion constructs, we measured GFP fluorescence in TXTL reactions expressing dCas9 and a targeting or non-targeting sgRNA (Figure 3B). The TXTL reactions showed variable repression of deGFP in comparison to the non-targeting control. Our results also show the consistency of transcriptional repression between the MscL-eGFP transcriptional and translational fusions.

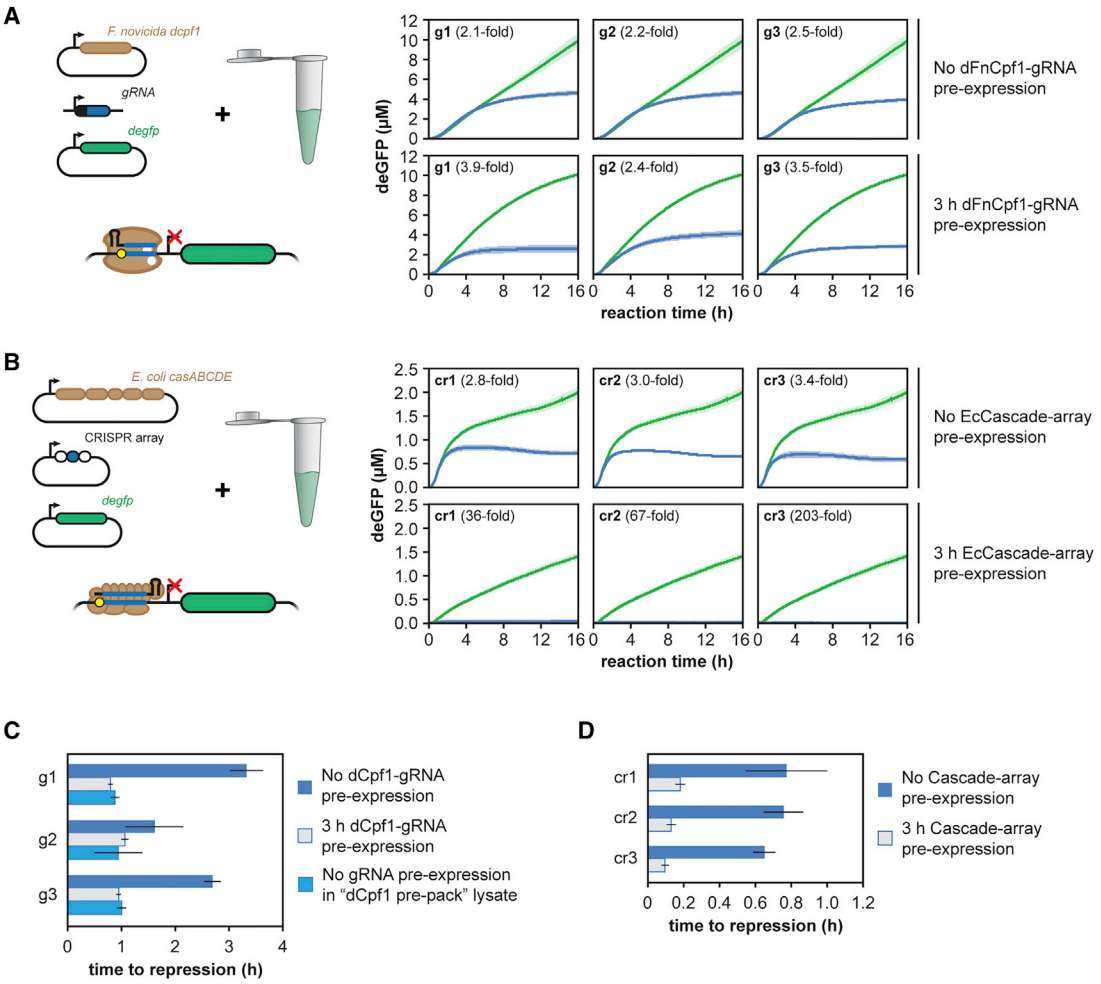


Figure 4. Single Effector and Multi-protein Effector Cas Proteins Function Efficiently in TXTL

(A and B) Time series of reporter gene expression in TXTL for cell-free reactions expressing (A) a catalytically inactive version of the Type V-A Cpf1 nuclease from *Francisella novicida* or (B) the Type I-E Cascade complex from *E. coli*. The protein or set of proteins was expressed along with a non-targeting gRNA (green) or one of three gRNAs (blue) designed to target the promoter of the deGFP reporter construct. The reporter plasmid is added to the reaction either at the same time as the constructs expressing the Cas protein(s) and the gRNA (top row) or after 3 hr (bottom row). Error bars represent the SEM from six repeats.

(C) Time to repression for the curves from (A), as well as for "dFnCpf1 pre-pack." Error bars represent the SEM from at least five repeats.

(D) Time to repression for the curves from (B). Error bars represent the SEM from at least four repeats.

In the case of Cas9, our scheme involved targeting a sequence upstream of the *degfp* promoter on the reporter plasmid. Without an inhibitor of RecBCD, we hypothesized that if Cas9 cleaved the sequence, the cleaved plasmid would be rapidly degraded, leading to a loss of deGFP expression. To test this hypothesis, we designed two sgRNAs: one targeting upstream of the promoter and another targeting the coding sequence of *degfp*. Targeting upstream of the promoter greatly reduced GFP expression when using Cas9, but not dCas9, while targeting within the coding sequence of *gfp* reduced expression for both Cas9 and dCas9, thus demonstrating that cleavage was required for quenching (Figure 3C). Note that adding DNA containing Chi sites, a RecBCD inhibitor, reduced repression by Cas9 when targeting upstream of the promoter, which shows the mechanism of quenching was RecBCD degradation of the cleaved plasmid. These results illustrate how to employ TXTL with dCas9 or Cas9 to measure the activity of sgRNAs.

The Activity of Other CRISPR Nucleases Can Be Measured with TXTL

Our results showed that active SpyCas9 RNPs can be expressed in TXTL and suggested that other CRISPR-Cas systems would also function in TXTL. To test this possibility, we evaluated effector proteins from two other CRISPR-Cas systems: the single-effector nuclease Cpf1 (Cas12a) from the Type V-A system in *F. tularensis subsp. Novicida* U112 and the multi-subunit complex Cascade from the Type I-E system in *E. coli* (Figures 4A and 4B). These effector proteins were selected because Cpf1 exhibits desirable properties over Cas9, while Type I CRISPR-Cas systems are the most prevalent system type in nature but are more challenging to characterize because multiple proteins

form the effector complex (Makarova et al., 2015; Zetsche et al., 2015). Previous studies have shown that both systems are capable of programmable gene repression in bacteria when using a catalytically dead version of the nuclease (dFnCpf1) or expressing Cascade in the absence of the Cas3 endonuclease, respectively (Leenay et al., 2016; Luo et al., 2015; Rath et al., 2015; Zetsche et al., 2015). We designed three gRNAs (a mature gRNA for Cpf1 and a repeat-spacer-repeat array for Cascade) targeting distinct sites within the P70a promoter, with each site flanked by an appropriate PAM.

We first found that dFnCpf1 was capable of repressing gene expression with each of the three gRNAs that targeted the reporter gene promoter (Figures 4A and S3A). Although dFnCpf1 eventually blocked deGFP expression to a similar extent as dCas9, the measured repression by dFnCpf1 was less than dSpyCas9 due to the longer delay in the onset of repression (Figures 4C and S3A). Consistent with this, pre-expressing dFnCpf1 and its gRNA for 3 hr reduced the delay in measurable repression by approximately 2 hr. Interestingly, a ''dFnCpf1 pre-pack'' TXTL lysate yielded a similar reduction in the time to repression. These results indicate that, in contrast to dCas9, the expression of dFnCpf1 is a major contributor to the delay in repression.

We also found that EcCascade could elicit gene repression in TXTL despite the need to coordinately express five proteins. The total reduction of GFP produced by the reaction was modest because of the delay in the onset of strong repression (Figures 4B, 4D, and S3B). However, pre-expressing EcCascade and the CRISPR RNA strongly reduced the total amount of GFP produced due to strong repression shortly after the addition of the reporter plasmid (Figures 4B, 4D, and S3B), indicating that the complex rapidly binds DNA and efficiently blocks RNA polymerase recruitment. Unexpectedly, deGFP production was lower when expressing EcCascade versus any of the other effector proteins (Figure 4B). This decrease in deGFP production was even more prominent when EcCascade was pre-expressed. Adding pure recombinant eGFP to a TXTL reaction expressing EcCascade resulted in only marginally lower fluorescence than a TXTL reaction expressing SpyCas9 or a no-nuclease control (Figure S3C). Therefore, the non-specific, inhibitory effect of EcCascade can be traced to deGFP expression rather than its stability or fluorescence properties. In sum, we showed that TXTL can be extended to the characterization of CRISPR-Cas systems requiring both single-effector and multi-protein effector complexes.

TXTL Can Quantify the Inhibitory Activity of Anti-CRISPR Proteins

Recently, the discovery of anti-CRISPR proteins that bind and inhibit Cascade and Cas3 from Type I-E and Type I-F CRISPR-Cas systems and Cas9 from Type II-A and II-C systems raised the potential of using these proteins to tightly regulate genome editing and gene regulation (Bondy-Denomy et al., 2015; Pawluk et al., 2016a, 2016b; Rauch et al., 2017). We tested if TXTL could be used to rapidly assess the inhibitory activity of potential anti-CRISPR proteins. We initially focused on AcrIIA2 and AcrIIA4, two anti-CRISPR proteins recently reported to inhibit the activity of SpyCas9 *in vitro* and in human cells (Rauch et al., 2017). Both proteins counteracted gene repression by dSpyCas9 (Figures 5A and S2E), although AcrIIA4 was a more potent inhibitor than AcrIIA2 in line with prior measurements in *E. coli* (Rauch et al., 2017).

We next tested a panel of 24 anti-CRISPR proteins, including seven previously reported proteins from defined classes (e.g., AcrIIA2 and AcrIIC1) (Hynes et al., 2017; Pawluk et al., 2016a), as well as up to four homologs of each protein. The panel of anti-CRISPR proteins was tested against five different Cas9 nucleases: the Type II-A SpyCas9, the VQR variant of SpyCas9 that recognizes an altered NGA PAM (SpyCas9^{VQR}) (Kleinstiver et al., 2015), the Type II-A CRISPR1 Cas9 from *Streptococcus thermophilus* (Sth1Cas9), the Type II-C Cas9 from *Neisseria meningitidis* (NmeCas9), and the Type II-C Cas9 from *Campylobacter jejuni* (CjeCas9). In total, we tested 120 different combinations (Figure 5B).

Across the set of tested anti-CRISPR proteins and Cas9 nucleases, we identified widely ranging specificities and activities. Many of the anti-CRISPR proteins inhibited a single nuclease or subtype of nucleases, such as AcrIIA2 and two of its homologs inhibiting only SpyCas9 and SpyCas9^{VQR} or AcrIIC1 and one of its homologs inhibiting both of the Type II-C Cas9 nucleases, NmeCas9 and CjeCas9, in line with a recent study of AcrIIC1 (Harrington et al., 2017). Finally, some anti-CRISPR proteins appeared to differentiate between SpyCas9 and SpyCas9^{VQR}: two (AcrIIA2-3 and AcrIIA4-3) preferred SpyCas9 while one (AcrIIA5-2) preferred SpyCas9^{VQR}.

Surprisingly, the TXTL-based measurements suggested that other anti-CRISPR proteins can inhibit beyond these boundaries. For instance, AcrIIA3 and AcrIIC2 inhibited the activity of the II-A Sth1Cas9 and both II-C Cas9 nucleases. AcrIIC2 and AcrIIC3 showed much stronger inhibition of NmeCas9 than CjeCas9, paralleling recent *in vitro* cleavage experiments with NmeCas9 and CjeCas9 (Harrington et al., 2017). Most strikingly, three of the tested homologs (AcrIIA3-1, AcrIIA3-3, and AcrIIA5-2) appeared to inhibit all tested Cas9 nucleases. While highly intriguing, these anti-CRISPR proteins strongly reduced deGFP expression even with the non-targeting sgRNA (Figure S4). As a result, these anti-CRISPR proteins may have limited expression of Cas9 and the targeting sgRNA, thereby confounding any definitive conclusions about inhibitory activity. Similarly, two of the tested anti-CRISPR proteins (AcrIIA3-2 and AcrIIA3-4) completely inhibited GFP expression even with the non-targeting sgRNA, preventing us from discerning any inhibitory activity. Overall, these experiments demonstrate that TXTL can be used to rapidly and scalably assess the inhibitory activity of diverse anti-CRISPR proteins, and they revealed anti-CRISPR proteins with widely ranging specificities.

TXTL Offers a Rapid Means of Elucidating CRISPR PAMs

One of the major barriers to the functional characterization of new Cas nucleases is determining the recognized PAM sequences. While numerous experimental methods have been developed for PAM determination, they consistently rely on *in vitro* assays that require protein purification or on cell-based assays that require culturing and transforming of live cells (Karvelis et al., 2017; Leenay and Beisel, 2017). We therefore asked if TXTL could also be used as the basis of PAM determination assays. Paralleling prior *in vitro* and *in vivo* DNA cleavage assays

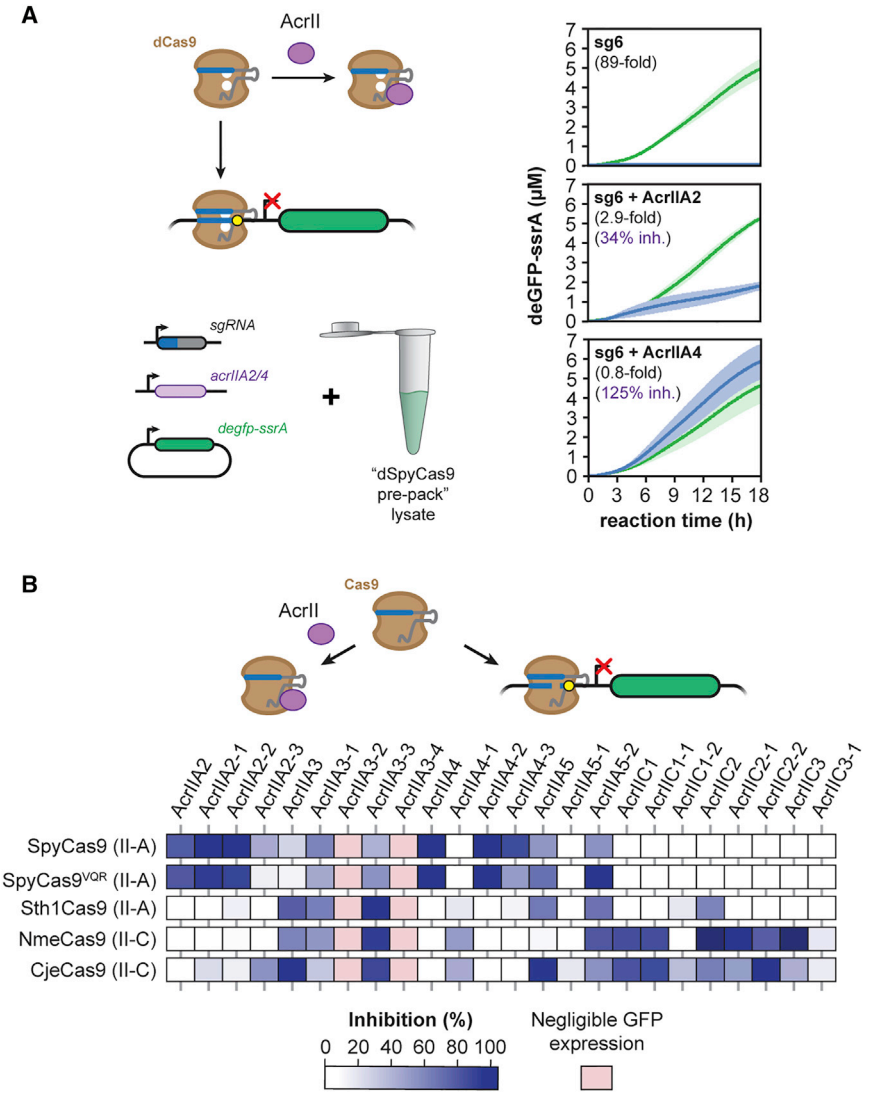


Figure 5. TXTL Can Be Used to Rapidly Characterize Anti-CRISPR Proteins

(A) Time series of deGFP-ssrA expression in TXTL for cell-free reactions also expressing dSpyCas9, an sgRNA, and one of two anti-CRISPR proteins, AcrIIA2 and AcrIIA4, shown to inhibit SpyCas9 activity. Each reaction was performed with a targeting sgRNA (blue) or a non-targeting sgRNA (green). Error bars represent the SEM from at least three repeats.

(B) A matrix showing the percentage inhibition for 24 different anti-CRISPR proteins on five different Cas9. Samples with no appreciable GFP expression in the presence of the anti-CRISPR protein are designated with light red. Values represent the mean of at least three technical replicates.

(Jiang et al., 2013; Zetsche et al., 2015), our devised assay relies on introducing a library of potential PAM sequences flanking a site targeted by an expressed gRNA and Cas nuclease (Figure 6A). After incubating the pooled PAM library and DNA encoding the Cas nuclease and gRNA, the pool of uncleaved target sequences was PCR-amplified and subjected to next-generation sequencing. By comparing the relative frequency of individual sequences within the library before and after cleavage, we could quantify the depletion of each library member and therefore how well the nuclease recognized each sequence as a PAM. Critically, the assay was completed in 10–20 hr from when the DNA constructs were in hand to when PCR products were submitted for sequencing. See Methods S1, Protocol 2 for more information about how to conduct the TXTL-based PAM determination assay.

As a proof-of-principle demonstration, we assessed the PAM requirements of the well-characterized FnCpf1 nuclease. Previous assays revealed that FnCpf1 prefers a TTN motif on the 5′ end of the target for efficient cleavage using the gRNA-centric orientation, although CTN can also be recognized (Fonfara et al., 2016; Leenay et al., 2016; Zetsche et al., 2015). *In vitro* PAM assays are known to yield less specific PAM sequences for higher nuclease concentrations (Karvelis et al., 2015), so we assessed the determined PAMs in our assay for reaction times ranging from 1 to 6 hr following the addition of the FnCpf1 and gRNA expression constructs (Figure S5A). Figure 6 depicts the identified PAM sequences based on the depletion of individual nucleotides at each position in a 5-nt library (Figure 6B) or specific motifs (Figure 6C), as well as a PAM wheel capturing the relative depletion of each sequence across the library (Figure 6D) (Leenay et al., 2016). The assay recapitulated the canonical 5′ TTN PAM while revealing NTTTN as the most active motif, with ATTTA as the most depleted by ~3-fold more than the next most depleted sequences (Data S1). We further found that NNCTN supported efficient cleavage in line with previous reports (Fonfara et al., 2016; Zetsche et al., 2015) and a T at the −1 position was detrimental to cleavage, indicating a consensus YTV PAM. NTCN also supported cleavage if there was a T at the −4 position and a C or an A at the −1 position. Importantly, the PAM distribution did not become less specific at later cleavage times unlike prior *in vitro* PAM assays (Karvelis et al., 2015).

To further demonstrate the applicability of the TXTL-based PAM assay, we performed the assay on five phylogenetically diverse and previously uncharacterized Cpf1 nucleases as well as one Cpf1 that had undergone limited characterization (Figure S5B). Initial analyses revealed that the nucleases exhibited varying levels of activity on the TTTC PAM when targeting the GFP plasmid (Figure S5C). Interestingly, the Cpf1 from *Moraxella bovoculi* 237 (MbCpf1) exhibited greatly enhanced activity at a reaction temperature of 37°C than the standard 29°C (Figure S5C). The resulting PAM assays revealed two general

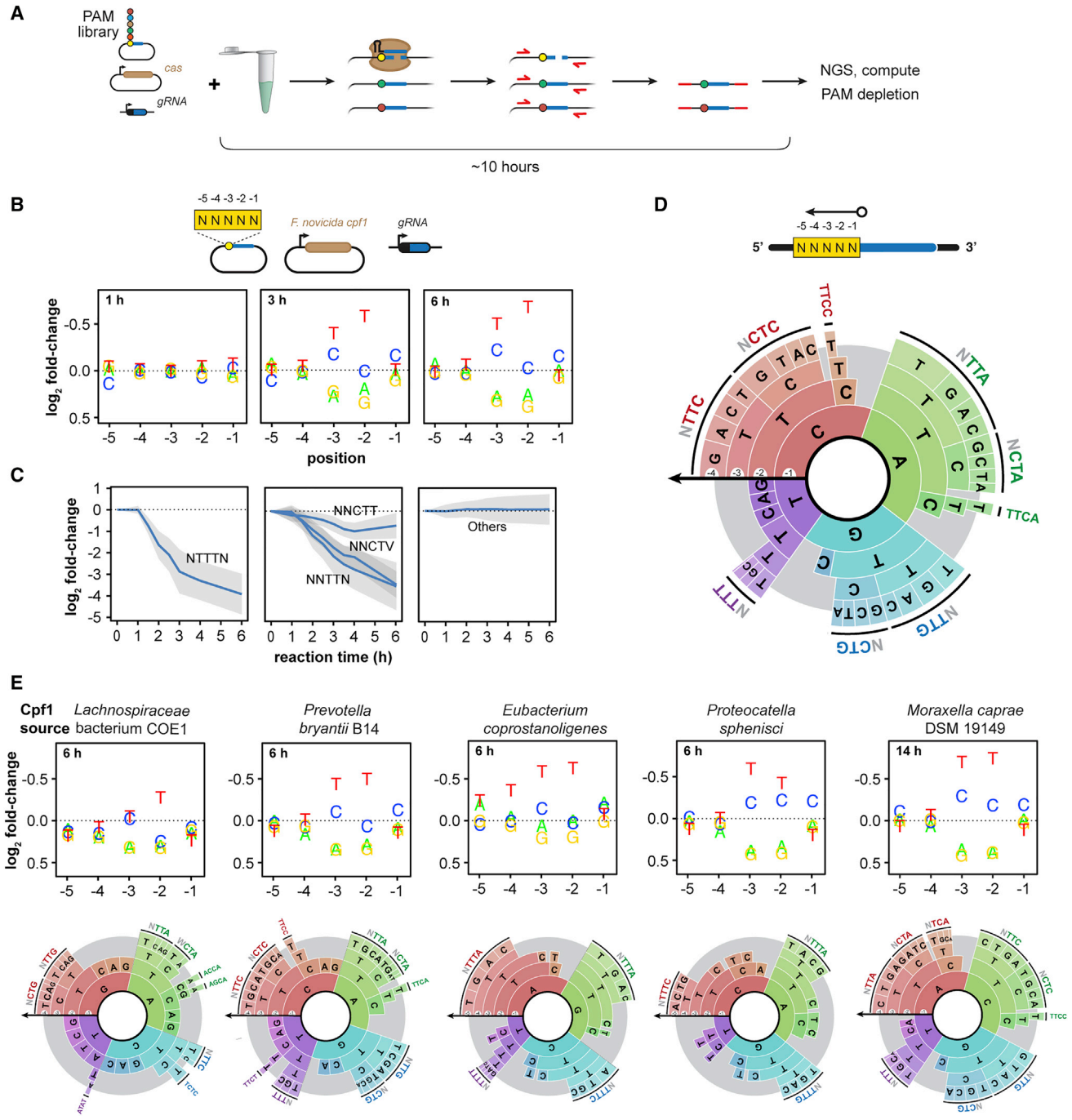


Figure 6. TXTL Can Be Used to Determine CRISPR PAMs

(A) Schematic of a TXTL-based cleavage assay to determine the PAM sequences recognized by Cas nucleases.

(B) Plots showing the fold change in the representation of a nucleotide at each variable position in the PAM library as a result of FnCpf1 activity in comparison to the original PAM library. Note that the y axis is inverted to highlight nucleotides that are depleted.

(C) Time series showing the depletion of selected motifs by FnCpf1 matching the consensus sequence in the sequencing libraries is shown. Error bars show the SD of the fold change.

(D) A PAM wheel showing the determined PAM sequences recognized by FnCpf1. PAM sequences are read proceeding from the outside to the inside of the circle, and the arc length directly correlates with the extent of PAM depletion. The −5 position was not shown for clarity.

(E) Plots showing the fold change in the representation of a nucleotide at each variable position in the PAM library in comparison to the original PAM library (top) and PAM wheels showing the determined PAM sequences (bottom) for five uncharacterized Cpf1 nucleases. A sixth Cpf1 nuclease (MbCpf1) previously characterized in Zetsche et al. (2015) is also reported in Figure S5.

consensus PAMs following that of FnCpf1 (YTV) or AsCpf1 (TTTV) (Figures 6E and S4D) (Zetsche et al., 2015). Interestingly, the PAM preferences did not correlate with the extent of homology (Figures 6E and S4B). Furthermore, some Cpf1 nucleases exhibited PAM preferences within these consensus sequences, such as the Cpf1 from *Lachnospiraceae* bacterium COE1 (Lb6Cpf1) accommodating ACCA and AGCA PAMs (Figure 6E). We proceeded to test four of these Cpf1 nucleases (McCpf1, EcCpf1, Pb2Cpf1, and Lb6Cpf1) for genome-editing activity using an established recombination system in rice callus (Begemann et al., 2017). All four exhibited detectable indel formation at a targeted site downstream of a TTTC PAM site based on T7EI assays and Sanger sequencing of the edited loci from antibiotic-resistant rice callus (Figures S5E and S5F). Taken together, these results indicate that a TXTL-based assay can elucidate PAM requirements for CRISPR nucleases, opening the door to the characterization of PAM requirements across both natural and engineered CRISPR nucleases.

DISCUSSION

We have demonstrated that *E. coli* cell-free TXTL can be used for the rapid and scalable characterization of CRISPR nucleases, gRNAs, and anti-CRISPR proteins. Unlike *in vitro* biochemical assays or cell-based assays, TXTL does not require any protein purification or cell culturing and transformation. TXTL also allows exquisite control over the reaction conditions and the amount of the DNA templates, and it can provide a dynamic and quantitative readout of nuclease activity within a few hours. With recent advances in DNA synthesis as well as liquid handling systems, the approaches here could be readily scaled to hundreds of reaction conditions or constructs.

CRISPR-Cas systems are remarkably diverse, with significant sequence, structural, and functional diversity (Koonin et al., 2017). This diversity exists even within a single subtype; for instance, Cas9 proteins from the well-characterized Type II-A subtype can exhibit less than 10% sequence homology at the amino acid level and show a range of recognized PAM lengths and sequences (Fonfara et al., 2014). However, only a few representative nucleases have been characterized for the other subtypes. This is particularly striking for Type I and III CRISPR-Cas systems, the most prevalent types in prokaryotes. These systems rely on multiple proteins to form the effector complex, requiring the purification or expression of multiple proteins in defined stoichiometries that has complicated their widespread characterization. TXTL is ideally suited to characterize the multisubunit effector complexes from these systems because linear, chemically synthesized DNA encoding each subunit can be combined in a single TXTL reaction. The expressed complex can then be characterized in a variety of ways, such as determining its assembly kinetics and PAM requirements. Evaluating numerous systems from one subtype could help reveal how CRISPR-Cas systems evolved and the structural basis of PAM recognition through mapping sequence-function relationships.

While our results demonstrate the promise of TXTL for characterizing CRISPR-Cas systems, there are multiple opportunities to further expand the utility of TXTL. For example, TXTL may be explored to investigate spacer acquisition across the diversity of Cas acquisition proteins found in nature. While it remains to be seen if acquisition can be recapitulated in TXTL, PCR-based or next-generation sequencing-based approaches could be readily used to assess the frequency and composition of integrated spacers (Fineran et al., 2014; Levy et al., 2015). We also found that dCas9-based repression correlated between TXTL and *E. coli* and devised different means of assessing the activity of any sgRNA, opening the possibility of using TXTL to validate gRNA designs. While more work is needed to determine which factors primarily determine gRNA activity in different cellular settings (e.g., chromatin state or sequence biases in DNA repair), factors that impact gRNA stability, the assembly of RNP complexes, or efficiency of DNA binding and cleavage would be identified in TXTL. Finally, measuring the time to repression with or without pre-expression of the CRISPR machinery opens the potential of using TXTL to rapidly measure the kinetics of RNP complex assembly and function under *in vivo*-like conditions. Through the demonstrations reported here and through further extensions, TXTL has the potential to make a widespread impact on the characterization of CRISPR-Cas systems and their transition into a new generation of CRISPR technologies.

Limitations

While TXTL can greatly accelerate the characterization of CRISPR-Cas systems, TXTL does possess limitations that restrict its use. For instance, gene expression in TXTL is limited to 25°C–42°C, with an optimal temperature of 29°C (Shin and Noireaux, 2010). While this range allowed us to enhance the activity of the MbCpf1 (Figure S5C), it could limit the characterization of CRISPR-Cas systems native to thermophilic prokaryotes. TXTL reactions also lose activity after ~16 hr, potentially limiting the use of GFP as a readout for poorly expressed nucleases. Additionally, there may be some variability in reaction rates between different batches of TXTL. However, this limitation is overcome by including internal standards for each batch (e.g., non-targeting sgRNAs) when measuring CRISPR-based activity. Next, TXTL is missing factors common to eukaryotic cells, such as long and diverse DNA sequences that impact the dynamics of target search, chromatin and other nucleoid proteins that impact the structure and availability of DNA, and proteins responsible for repair involved in eukaryotic genome editing. Finally, factors for CRISPR RNA processing besides RNase III would need to be identified (e.g., tracrRNAs), as *E. coli* would be unlikely to naturally express these factors.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Preparation of TXTL lysates and reactions
  - Assembly of expression constructs
  - Fluorescence time-course measurements in TXTL
  - Amplification of DNA targets in TXTL
  - Fluorescence measurements in *E. coli*

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, one data file, and one methods file and can be found with this article online at https://doi.org/10.1016/j.molcel.2017.12.007.

ACKNOWLEDGMENTS

We thank Ryan Leenay for help generating the FnCpf1 PAM wheel, Jennie Fagan for cloning the plasmid expressing SpyCas9, and Luciano Maraffini for providing the pCas9 plasmid (Addgene plasmid #42876). Preliminary experiments were performed during the 2016 CSHL Synthetic Biology summer course. This material is based upon work supported by the Defense Advanced Research Projects Agency (contract HR0011-16-C-01-34, V.N. and C.L.B.), the Office of Naval Research (award N00014-13-1-0074, V.N.), the NIH (grant 1R35GM119561-01 to C.L.B.), and the National Science Foundation (grant CBET-1403135 to C.L.B.).

AUTHOR CONTRIBUTIONS

V.N., C.L.B., R.M., and M.L.L. conceived of the original project idea. V.N., C.L.B., R.M., C.S.M., and M.L.L. designed the TXTL characterization experiments. R.M. and C.S.M. performed the TXTL characterization experiments. S.P.C. performed the comparisons between TXTL and *E. coli*. S.P.C. and R.M. designed and performed the anti-CRISPR experiments. C.S.M. and T.J. designed and performed the PAM determination assays. M.B.B., B.N.G., E.J., A.S., and Y.H. identified the novel Cpf1 nucleases and designed and performed the genome-editing experiments in rice callus. V.N., C.L.B., R.M., and C.S.M. wrote the paper.

DECLARATION OF INTERESTS

C.L.B. is a co-founder and scientific advisory board member of Locus Biosciences and submitted provisional patent applications on CRISPR technologies. M.B.B., B.N.G., E.J., A.S., and Y.H. are current employees of Benson Hill Biosystems, which has proprietary CRISPR technologies on which patent applications have been filed. Noireaux laboratory receives research funds from Arbor Biosciences, a distributor of myTXTL cell-free protein expression kit.

Received: August 4, 2017
Revised: November 14, 2017
Accepted: December 6, 2017
Published: January 4, 2018

REFERENCES

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*, aaf5573.

Barrangou, R., and Doudna, J.A. (2016). Applications of CRISPR technologies in research and beyond. Nat. Biotechnol. *34*, 933–941.

Begemann, M.B., Gray, B.N., January, E., Gordon, G.C., He, Y., Liu, H., Wu, X., Brutnell, T.P., Mockler, T.C., and Oufattole, M. (2017). Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. Sci. Rep. *7*, 11606.

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. *41*, 7429–7437.

Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M.F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K.L., and Davidson, A.R. (2015). Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. Nature *526*, 136–139.

Boyle, E.A., Andreasson, J.O.L., Chircus, L.M., Sternberg, S.H., Wu, M.J., Guegler, C.K., Doudna, J.A., and Greenleaf, W.J. (2017). High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. Proc. Natl. Acad. Sci. USA *114*, 5461–5466.

Carlson, E.D., Gan, R., Hodgman, C.E., and Jewett, M.C. (2012). Cell-free protein synthesis: applications come of age. Biotechnol. Adv. *30*, 1185–1194.

Caschera, F., and Noireaux, V. (2014). Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription-translation system. Biochimie *99*, 162–168.

Chappell, J., Jensen, K., and Freemont, P.S. (2013). Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. Nucleic Acids Res. *41*, 3471–3481.

Chen, B., Guan, J., and Huang, B. (2016). Imaging specific genomic DNA in living cells. Annu. Rev. Biophys. *45*, 1–23.

Dominguez, A.A., Lim, W.A., and Qi, L.S. (2016). Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat. Rev. Mol. Cell Biol. *17*, 5–15.

Dudley, Q.M., Karim, A.S., and Jewett, M.C. (2015). Cell-free metabolic engineering: biomanufacturing beyond the cell. Biotechnol. J. *10*, 69–82.

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

Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lécrivain, A.L., Bzdrenga, J., Koonin, E.V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic Acids Res. *42*, 2577–2590.

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

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The all E. coli TX-TL toolbox 2.0: a platform for cell-free synthetic biology. ACS Synth. Biol. *5*, 344–355.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. Cell *159*, 647–661.

Gootenberg, J.S., Abudayyeh, O.O., Lee, J.W., Essletzbichler, P., Dy, A.J., Joung, J., Verdine, V., Donghia, N., Daringer, N.M., Freije, C.A., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. Science *356*, 438–442.

Harrington, L.B., Doxzen, K.W., Ma, E., Liu, J.J., Knott, G.J., Edraki, A., Garcia, B., Amrani, N., Chen, J.S., Cofsky, J.C., et al. (2017). A broad-spectrum inhibitor of CRISPR-Cas9. Cell *170*, 1224–1233.e15.

Hockenberry, A.J., and Jewett, M.C. (2012). Synthetic in vitro circuits. Curr. Opin. Chem. Biol. *16*, 253–259.

Hynes, A.P., Rousseau, G.M., Lemay, M.L., Horvath, P., Romero, D.A., Fremaux, C., and Moineau, S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. Nat. Microbiol. *2*, 1374–1380.

Jewett, M.C., Calhoun, K.A., Voloshin, A., Wuu, J.J., and Swartz, J.R. (2008). An integrated cell-free metabolic platform for protein production and synthetic biology. Mol. Syst. Biol. *4*, 220.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. *31*, 233–239.

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

Kanter, G., Yang, J., Voloshin, A., Levy, S., Swartz, J.R., and Levy, R. (2007). Cell-free production of scFv fusion proteins: an efficient approach for personalized lymphoma vaccines. Blood *109*, 3393–3399.

Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. Genome Biol. *16*, 253.

Karvelis, T., Gasiunas, G., and Siksnys, V. (2017). Methods for decoding Cas9 protospacer adjacent motif (PAM) sequences: a brief overview. Methods *121-122*, 3–8.

Karzbrun, E., Tayar, A.M., Noireaux, V., and Bar-Ziv, R.H. (2014). Synthetic biology. Programmable on-chip DNA compartments as artificial cells. Science *345*, 829–832.

Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K., Cho, H.Y., Song, D.W., Lee, K.J., Jung, M.H., Kim, S., et al. (2017). In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. Nat. Commun. *8*, 14500.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P., Li, Z., Peterson, R.T., Yeh, J.R., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature *523*, 481–485.

Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nguyen, N.T., Welch, M.M., Lopez, J.M., McCaw, Z.R., Aryee, M.J., and Joung, J.K. (2016). Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. Nat. Biotechnol. *34*, 869–874.

Komor, A.C., Badran, A.H., and Liu, D.R. (2017). CRISPR-based technologies for the manipulation of eukaryotic genomes. Cell *168*, 20–36.

Koonin, E.V., Makarova, K.S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. *37*, 67–78.

Leenay, R.T., and Beisel, C.L. (2017). Deciphering, communicating, and engineering the CRISPR PAM. J. Mol. Biol. *429*, 177–191.

Leenay, R.T., Maksimchuk, K.R., Slotkowski, R.A., Agrawal, R.N., Gomaa, A.A., Briner, A.E., Barrangou, R., and Beisel, C.L. (2016). Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. Mol. Cell *62*, 137–147.

Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., and Sorek, R. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature *520*, 505–510.

Luo, M.L., Mullis, A.S., Leenay, R.T., and Beisel, C.L. (2015). Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. Nucleic Acids Res. *43*, 674–681.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. *13*, 722–736.

Marshall, R., Maxwell, C.S., Collins, S.P., Beisel, C.L., and Noireaux, V. (2017). Short DNA containing χ sites enhances DNA stability and gene expression in E. coli cell-free transcription-translation systems. Biotechnol. Bioeng. *114*, 2137–2141.

Mekler, V., Minakhin, L., Semenova, E., Kuznedelov, K., and Severinov, K. (2016). Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3′-terminal segment of guide RNA. Nucleic Acids Res. *44*, 2837–2845.

Mohanraju, P., Makarova, K.S., Zetsche, B., Zhang, F., Koonin, E.V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. Science *353*, aad5147.

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

Nelles, D.A., Fang, M.Y., O'Connell, M.R., Xu, J.L., Markmiller, S.J., Doudna, J.A., and Yeo, G.W. (2016). Programmable RNA tracking in live cells with CRISPR/Cas9. Cell *165*, 488–496.

Pardee, K., Slomovic, S., Nguyen, P.Q., Lee, J.W., Donghia, N., Burrill, D., Ferrante, T., McSorley, F.R., Furuta, Y., Vernet, A., et al. (2016). Portable, on-demand biomolecular manufacturing. Cell *167*, 248–259.e12.

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. *31*, 839–843.

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., and Davidson, A.R. (2016a). Naturally occurring off-switches for CRISPR-Cas9. Cell *167*, 1829–1838.e9.

Pawluk, A., Staals, R.H., Taylor, C., Watson, B.N., Saha, S., Fineran, P.C., Maxwell, K.L., and Davidson, A.R. (2016b). Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. Nat. Microbiol. *1*, 16085.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.

Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P.R., and Lundgren, M. (2015). Efficient programmable gene silencing by Cascade. Nucleic Acids Res. *43*, 237–246.

Rauch, B.J., Silvis, M.R., Hultquist, J.F., Waters, C.S., McGregor, M.J., Krogan, N.J., and Bondy-Denomy, J. (2017). Inhibition of CRISPR-Cas9 with bacteriophage proteins. Cell *168*, 150–158.e10.

Schmidt, R., Bukau, B., and Mogk, A. (2009). Principles of general and regulatory proteolysis by AAA+ proteases in Escherichia coli. Res. Microbiol. *160*, 629–636.

Shin, J., and Noireaux, V. (2010). Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. J. Biol. Eng. *4*, 8.

Shin, J., and Noireaux, V. (2012). An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. ACS Synth. Biol. *1*, 29–41.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol. Cell *60*, 385–397.

Sun, Z.Z., Yeung, E., Hayes, C.A., Noireaux, V., and Murray, R.M. (2014). Linear DNA for rapid prototyping of synthetic biological circuits in an Escherichia coli based TX-TL cell-free system. ACS Synth. Biol. *3*, 387–397.

Swartz, J. (2006). Developing cell-free biology for industrial applications. J. Ind. Microbiol. Biotechnol. *33*, 476–485.

Takahashi, M.K., Chappell, J., Hayes, C.A., Sun, Z.Z., Kim, J., Singhal, V., Spring, K.J., Al-Khabouri, S., Fall, C.P., Noireaux, V., et al. (2015a). Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS Synth. Biol. *4*, 503–515.

Takahashi, M.K., Hayes, C.A., Chappell, J., Sun, Z.Z., Murray, R.M., Noireaux, V., and Lucks, J.B. (2015b). Characterizing and prototyping genetic networks with cell-free transcription-translation reactions. Methods *86*, 60–72.

Tayar, A.M., Karzbrun, E., Noireaux, V., and Bar-Ziv, R.H. (2015). Propagating gene expression fronts in a one-dimensional coupled system of artificial cells. Nat. Phys. *11*, 1037–1041.

Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J., and Joung, J.K. (2017). CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. Nat. Methods *14*, 607–614.

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

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759–771.

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol. *35*, 31–34.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
BL21 Rosetta 2	Novagen	71405
MG1655	Table S1	Table S1
5-alpha Electrocompetent Cells (High Efficiency)	NEB	C2987I
Chemicals, Peptides, and Recombinant Proteins		
reGFP	Cell Biolabs	STA-201
Taq DNA Polymerase	NEB	M0273
Q5 DNA Hot Start High-Fidelity DNA Polymerase	NEB	M0493
Potassium Glutamate	Sigma	G1149
Magnesium Glutamate	Sigma	49605
Maltodextrin	Sigma	419672
PEG8000	Sigma	P2139
Critical Commercial Assays		
myTXTL	Arbor Biosciences	507024
ZymoPURE Plasmid Midi Prep Kit	Zymo Research	D4200
NEBuilder HiFi DNA Assembly Cloning Kit	NEB	E5520S
Q5 Site-Directed Mutagenesis Kit	NEB	E0554S
Deposited Data		
Cpf1 NGS data	Table S1	Table S1
Oligonucleotides		
Primers	Table S1	Table S1
Nextera Indexing primers	Illumina	https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/experiment-design/illumina-adapter-sequences-1000000002694-03.pdf
Recombinant DNA		
Plasmids and gBlocks	Table S1	Table S1
Software and Algorithms		
PAM counting script (Python)	This paper	https://bitbucket.org/csmaxwell/crispr-txtl-pam-counting-script
Other		
Ampure XP beads	Beckman Coulter	A63880
Detailed protocols (TXTL reaction and PAM assay)	Methods S1	Methods S1

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests should be directed to the Lead Contact, Vincent Noireaux (noireaux@umn.edu).

METHOD DETAILS

Preparation of TXTL lysates and reactions

The *E. coli* cell-free TXTL lysate was prepared from BL21 Rosetta 2 from Novagen as described previously (Caschera and Noireaux, 2014). Currently, the TXTL lysate is also commercially available through the company Arbor Biosciences under the product name myTXTL. The dSpyCas9-loaded and dFnCpf1 TXTL lysates followed the same procedure, only the *E. coli* cells harbored a plasmid constitutively expressing dSpyCas9 or dFnCpf1 from a J23108 promoter. TXTL reactions were composed of 33% volume crude extract and the remaining 67% volume with the following components: energy mix, amino acid mix, cofactors, ions, and DNA. A typical TXTL reaction contains 50 mM HEPES pH 8, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM

coenzyme A, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-PGA, 1.5% PEG8000, 30 mM maltodextrin, 3 mM of each of the 20 amino acids, 90 mM K-glutamate, and 4 mM Mg-glutamate. TXTL reactions were conducted in volumes of 5 μl at 29-30°C. When expressing from linear DNA template, 2 μM of annealed oligos containing six copies of the χ-site sequence (Chi6; for details see Marshall et al., 2017) was added to the reaction. The linear DNA construct for expressing deGFP used to produce the data for Figure 2A was generated by PCR and included flanks over one kilobase away from the closest target site, mitigating potential effects from the end of the dsDNA molecule. To measure the inhibitory activity of AcrIIA2 and AcrIIA4 against SpyCas9 in TXTL, each protein was encoded on a linear expression construct that was expressed for two hours in the "dSpyCas9 pre-pack" lysate (Figures 5A and S2E). Then, DNA encoding the deGFP reporter plasmid and linear DNA encoding a targeting or non-targeting sgRNA was added and deGFP fluorescence was measured over time. To test the panel of anti-CRISPR proteins against the five Cas9 nucleases, we pre-expressed for three hours the catalytically active Cas9 from a plasmid, a targeting or non-targeting sgRNA and the anti-CRISPR protein from linear DNA. Then we added plasmid DNA encoding the GFP reporter. We measured the endpoint GFP expression after an additional 18 hours for each combination of Cas9 and anti-CRISPR protein (Figure 5B)

Assembly of expression constructs

Plasmids were constructed using standard techniques. The sequence of each gBlock and plasmid used in this experiment is available in the supporting information (Table S1). The plasmid expressing catalytically active SpyCas9 was generated by amplifying the transcriptional unit expressing SpyCas9 from pCas9 but excluding the CRISPR array from using primers CSMpr1132/1157 and cloning it into the backbone pCSM117. pCas9 was a gift from Luciano Maraffini (Addgene plasmid #42876). gBlocks were ordered from IDT and amplified with CSMpr1105/1106 before being PCR purified. All constructed plasmids were verified by Sanger sequencing of the inserted sequences.

Fluorescence time-course measurements in TXTL

Fluorescence kinetics measurements were performed principally using the reporter plasmid P70a-deGFP expressing a truncated version of eGFP (deGFP, 25.4 kDa, 1 mg/mL = 39.38 μM) (Shin and Noireaux, 2012). Fluorescence was measured on a Biotek H1m plate reader using a 96-well V-bottom plate (Corning Costar 3357) and Ex 485 nm, Em 528 nm. Time-course measurements were run for at least 16 hours at 29-30°C, with an interval of 3 minutes between reads.

Amplification of DNA targets in TXTL

TXTL reactions with the SpyCas9 or dSpyCas9 plasmid, sgRNA DNA template, and the P70a-deGFP expression plasmid were incubated at 29°C for three hours, then diluted 200X in water. 1 μl of this dilution was then used as the DNA template in a 50 μl PCR reaction with 55°C annealing temperature, 45 s extension time, and 25 cycles. The PCR reaction was conducted using Taq polymerase and primers RM01s and RM05as to produce a 1074-bp amplicon. PCR products were visualized on a 0.8% agarose gel.

Fluorescence measurements in *E. coli*

E. coli K-12 MG1655 cells with plasmids expressing dSpyCas9 and deGFP was transformed by electroporation with plasmids encoding the sgRNAs targeting sites (g1-19 targeting) on *degfp*, as well as a non-targeting control (g-nt). To measure each strain, three colonies from a freshly streaked plate was inoculated into 2 mL of LB with 34 μg/ml chloramphenicol (Cm), 50 μg/ml ampicillin (Amp), and 50 μg/ml kanamycin (Kan). The strains were cultured for 16-hours at 37°C shaking at 250 rpm. The cultures were then diluted 1:10,000 in LB with Cm, Amp, and Kan to a final volume of 100 μL within a black 96-well assay plate. Each culture was diluted into two wells as technical replicates. Cultures were incubated at 37°C with single orbital shaking at 425 rpm on a 3 mm diameter circle within a BioTek Synergy H1 for 20 hours at which cultures were in stationary phase. Cultures were resuspended with a multichannel pipette and loaded back into the BioTek Synergy H1. Single-point fluorescence at 485/528 nm excitation/emission as well as OD_{600} were measured from each well.

TXTL-based PAM assay

The pET vector expressing FnCpf1 was a kind gift of Benson Hill Biosystems. A PAM library with five randomized nucleotides flanking the spacer sequence was prepared as described previously (Leenay et al., 2016). The crRNA expressing the FnCpf1 crRNA targeting the PAM library was expressed from the gBlock CSM-GB099. A TXTL reaction was assembled as described above except the reaction contained 0.5 mM IPTG and the following DNA components: 0.2 nM P70a-T7RNAP, 2 nM pET-FnCpf1, 2 μM of Chi6 annealed oligos, 0.5 nM of the 5N PAM library, 2 nM of linear DNA expressing the crRNA. The reaction was split into 5 μL reactions as above and incubated at 29°C. Samples were then collected at the specified time by cutting away the cap mat sealing individual reactions and immediately freezing the reaction at −20°C until subsequent use. The adapters were attached to the PAM library by PCR amplification of the TXTL reaction using NEB's Q5 Hot Start High-Fidelity DNA Polymerase (NEB, M0493) with RL133 and RL134. The PCR reaction was then purified using Ampure XP (Beckman Coulter, A63880) beads before attaching unique Nextera indices to each sample by PCR amplification. After a final PCR purification using Ampure XP beads, each sample was normalized to 10 nM in a total volume of 20 uL and submitted for next-generation sequencing with 150 single-end reads on an Illumina MiSeq machine. For the other six Cpf1s, the assay was conducted in the same manner, where the *cpf1* genes were expressed either in a T7 expression plasmid (Lb6Cpf1, Pb2Cpf1, EcCpf1, and PsCpf1) or from linear DNA (McCpf1 and MbCpf1), while a processed crRNA derived from the

organism's native CRISPR array was expressed from linear DNA. See Methods S1, Protocol 2 for more information about conducting the TXTL-based PAM determination assay.

Plant Transformation Construct Cloning

Cpf1 coding sequences were optimized for monocot codon usage preferences and synthesized by GenScript. Cloning of Cpf1 constructs and guide RNA constructs was performed as described previously (Begemann et al., 2017). Briefly, each Cpf1 open reading frame was expressed from an enhanced Cauliflower Mosaic Virus 35S promoter (2X35S promoter); guide RNAs were expressed from the rice U6 promoter (OsU6 promoter). Rice (*Oryza sativa* cv. Kitaake) callus was transformed biolistically as described previously (Begemann et al., 2017). Briefly, gold microcarriers (0.6 μm in diameter) were coated with plasmids encoding the Cpf1 gene of interest and a guide RNA targeting the rice CAO1 gene, along with a hygromycin resistance cassette containing a hygromycin resistance gene expressed from the maize ubiquitin promoter (ZmUbi promoter). Bombarded rice callus was placed on selection medium containing hygromycin (50mg/L) for 3 weeks, and resistant callus pieces were subcultured.

QUANTIFICATION AND STATISTICAL ANALYSIS

Fluorescence time-course measurements in TXTL

Fluorescence intensity measurements were quantified using a linear standard calibration curve, spanning over three orders of magnitude, produced with pure recombinant eGFP (Cell Biolabs, STA-201, San Diego, CA). GFP production rates were calculated by two-point numerical differentiation and smoothed with a five-point quadratic polynomial. All fold-repression values for plasmid reporter constructs represent the ratio of deGFP concentrations after 16 hours of reaction for the non-targeting over the targeting sgRNA. For the linear reporter construct (Figure 2A), we measured the fold-repression of the deGFP produced by the reaction after five hours because linear templates incubated with Chi site-containing DNA degrade after that time in our TXTL system (Marshall et al., 2017). The time to repression was calculated based on the earliest time in which the rates of deGFP production diverge for the targeting sgRNA and the non-targeting sgRNA. For the experiments measuring the inhibitory activity of the anti-CRISPR proteins, inhibition was calculated from endpoint expression values after 18 hours of expression, using the following formula:

$$Inhibition(\%) = 100\% \left(\frac{\frac{GFP_{t,acr}}{GFP_{nt,acr}} - \frac{GFP_t}{GFP_{nt}}}{1 - \frac{GFP_t}{GFP_{nt}}} \right)$$

See Methods S1, Protocol 1 for more information about assessing the activity of CRISPR nucleases using TXTL.

Fluorescence measurements in *E. coli*

Endpoint fluorescence values were background subtracted using the fluorescence from cells lacking deGFP and then normalized by the endpoint, background-subtracted OD_{600} value. Fold-repression was then calculated as the normalized fluorescence of the non-targeting sgRNA strain divided by the normalized fluorescence of the targeting sgRNA strain.

TXTL-based PAM assay

The PAM wheel was generated as described previously (Leenay et al., 2016). Next-generation sequencing data is accessible through NCBI Biosample under BioProject: PRJNA415429. Links to the accession #'s are located in Table S1. See Methods S1, Protocol 2 for more information about conducting the TXTL-based PAM determination assay.

Sequence Alignments and Phylogenetic Trees

A multiple-sequence alignment of Cpf1 nuclease amino acid sequences was performed using MUSCLE with default parameters. A phylogenetic tree of Cpf1 protein sequences was built from this alignment using FigTree version 1.4.3. Midpoint tree rooting and cladogram transformation of branches were used for final tree visualization.

DATA AND SOFTWARE AVAILABILITY

Next-generation sequencing data for the PAM assay is accessible through NCBI Biosample under BioProject: PRJNA415429. Links to the accession #'s are located in Table S1. Krona plots used to generate the PAM wheels are found in Data S1.

Molecular Cell, Volume *69*

Supplemental Information

Rapid and Scalable Characterization of CRISPR Technologies Using an *E. coli* Cell-Free Transcription-Translation System

Ryan Marshall, Colin S. Maxwell, Scott P. Collins, Thomas Jacobsen, Michelle L. Luo, Matthew B. Begemann, Benjamin N. Gray, Emma January, Anna Singer, Yonghua He, Chase L. Beisel, and Vincent Noireaux

SUPPLEMENTAL INFORMATION

Rapid and scalable characterization of CRISPR technologies using an *E. coli* cell-free transcription-translation system

Ryan Marshall[1]*, Colin S. Maxwell[2]*, Scott P. Collins[2], Thomas Jacobsen[2], Michelle L. Luo[2], Matthew B. Begemann[3], Benjamin N. Gray[3], Emma January[3], Anna Singer[3], Yonghua He[3], Chase L. Beisel[2]*, Vincent Noireaux[1]*

[1]School of Physics and Astronomy
University of Minnesota Minneapolis, MN, 55455

[2]Department of Chemical and Biomolecular Engineering
North Carolina State University, Raleigh, NC 27695

[3]Benson Hill Biosystems
St. Louis, MO 63132, USA

Table of Contents

Supplemental Figure Legends

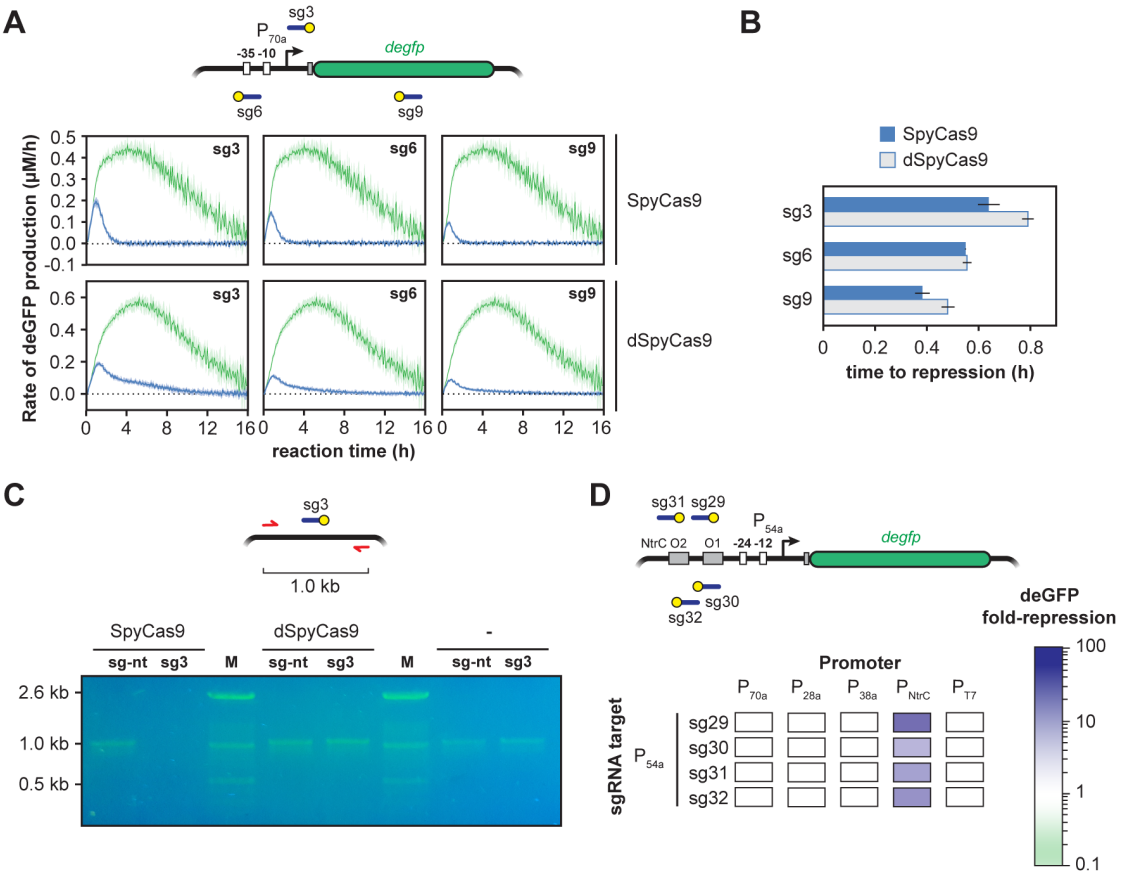


Figure S1. Related to Figure 1. The *S. pyogenes* Cas9 functions in TXTL. **A.** deGFP production rates from the time series curves from Figure 1B. Rates are shown for a targeting sgRNA (blue) and a non-targeting sgRNA (green). Rates are calculated by two-point numerical differentiation and smoothed with a five-point quadratic polynomial. The dark lines and light regions represent the average and S.E.M. of at least three runs. **B.** The number of hours before repression was observed in the time series curves from Figure 1B is shown. Values and error bars represent the average and S.E.M. of at least three replicates. **C.** 0.8% agarose gel showing cleavage of P70a-deGFP by SpyCas9, but not dCas9. A PCR of P70a-deGFP was performed after three hours of expression in the TXTL reaction, with primers flanking the g3 site in P70a-deGFP upstream by 197 bp and downstream by 877 bp. The full PCR product is 1074 bp. **D.** dSpyCas9-based repression by targeting NtrC binding sites. Top: Schematic of P54a-deGFP plasmid showing -24 and -12 consensus regions, NtrC binding sites and sgRNA target locations. Bottom: A matrix showing dSpyCas9-based repression using sgRNAs that target the NtrC binding sites. Values represent the mean of at least three repeated TXTL reactions.

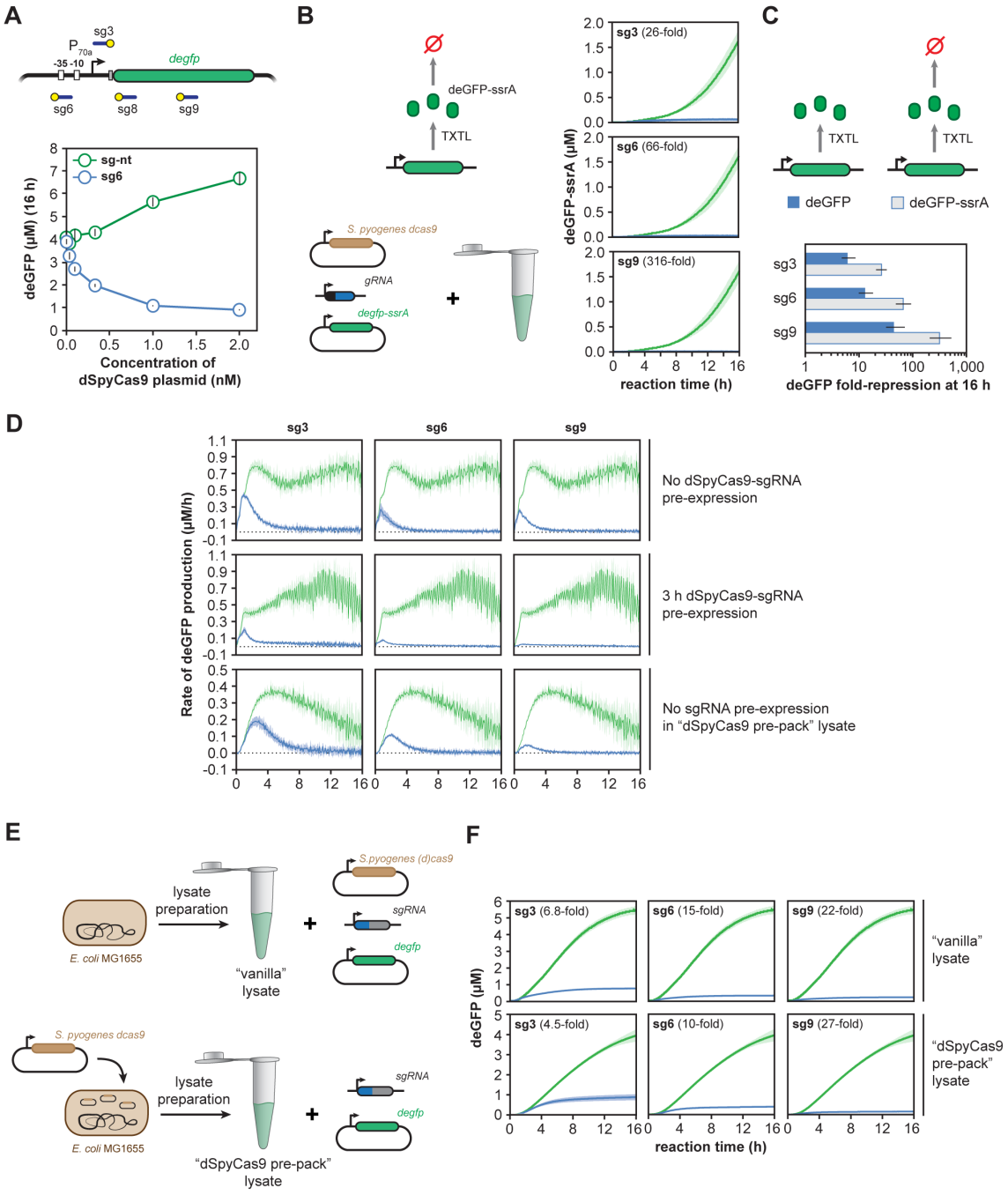


Figure S2. Related to Figure 2. Multiple factors affect dSpyCas9-based repression of reporter production in TXTL. **A.** Endpoint deGFP concentrations for TXTL reactions expressing the reporter plasmid, either a targeting (g6) sgRNA or a non-targeting (g-nt) sgRNA, and varying concentrations of the dCas9 plasmid. Values and error bars represent the average and S.E.M. of at least three replicates. **B.** Time course of deGFP-ssrA expression, where the reporter plasmid was targeted by a targeting sgRNA (blue) or a non-targeting sgRNA (green). The ssrA degron tag is recognized by the ClpXP protease that results in rapid turnover of the fusion protein. The dark lines and light

regions represent the average and S.E.M. of at least three runs. **C.** Fold-repression for reporter constructs encoding deGFP or deGFP-ssrA. Fold-repression is the ratio of deGFP concentrations after 16 hours of reaction for the non-targeting (green) over the targeting (blue) sgRNA. Error bars represent the S.E.M. from at least three repeated TXTL reactions. **D.** deGFP production rates from the time series curves from Figure 2B, as well as for TXTL reactions where dSpyCas9 was expressed in cells prior to generating the lysate. Rates are calculated by two-point numerical differentiation and smoothed with five-point quadratic polynomial. The dark lines and light regions represent the average and S.E.M. of at least five runs. **E.** Pre-expressing dSpyCas9 protein in *E. coli* prior to generating the TXTL lysate. A schematic for preparing either “vanilla” or “dSpyCas9 pre-pack” lysate is shown. **F.** Time series of deGFP concentration for vanilla or dSpyCas9 pre-pack cell-free reactions expressing either non-targeting sgRNA (green) or targeting sgRNAs (blue) is shown. The dark lines and light regions represent the average and S.E.M. of at least five runs.

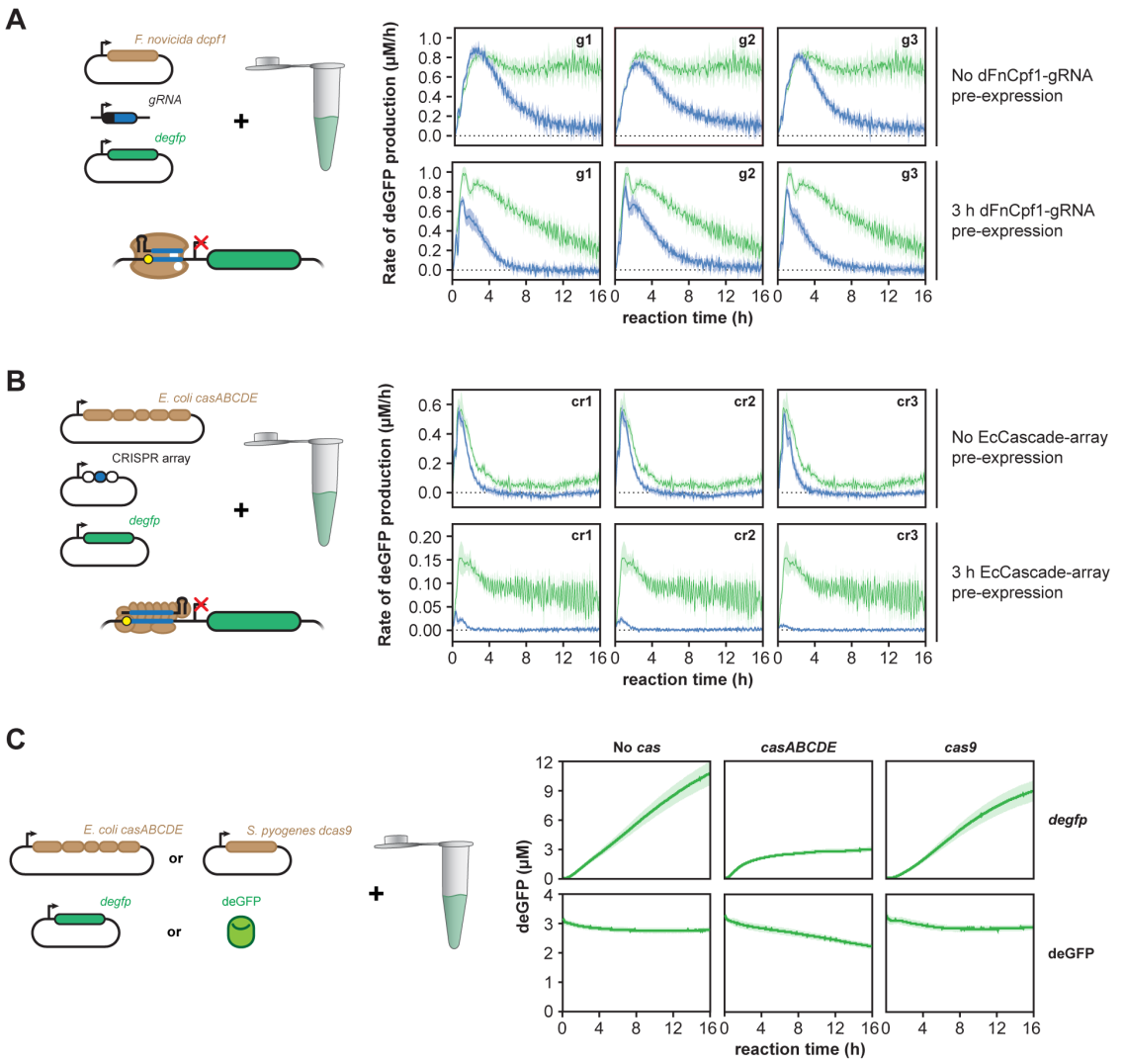


Figure S3. Related to Figure 4. Single effector and multi-protein effector Cas proteins function efficiently in TXTL. **A.** deGFP production rates from the time series curves from Figure 4A for a targeting guide RNA (blue) or a non-targeting guide RNA (green). Rates are calculated by two-point numerical differentiation and smoothed with five-point quadratic polynomial. The dark lines and light regions represent the average and S.E.M. of at least five runs. **B.** deGFP production rates from the time series curves from Figure 4B for a targeting guide RNA (blue) or a non-targeting guide RNA (green). Rates are calculated by two-point numerical differentiation and smoothed with five-point quadratic polynomial. The dark lines and light regions represent the average and S.E.M. of at least five runs. **C.** Time course of deGFP fluorescence in TXTL when deGFP is expressed from a plasmid (top) or added as purified recombinant protein (bottom). Reactions included the expression of EcCascade or the SpyCas9 or nothing (No *cas*). The dark lines and light regions represent the average and S.E.M. of at least six runs.

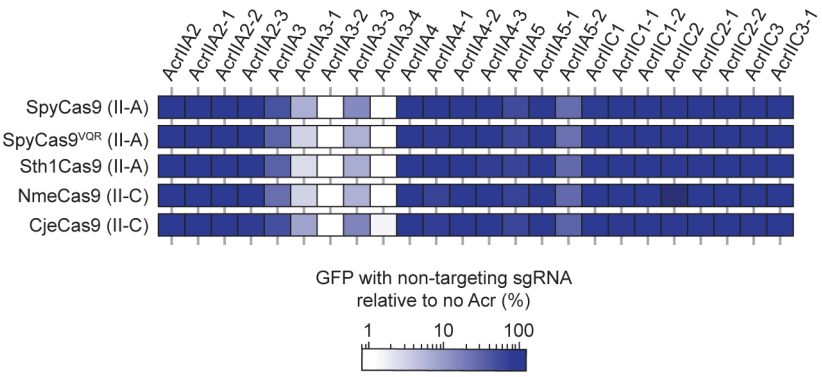


Figure S4. Related to Figure 5. Effect of anti-CRISPR proteins on expression of GFP in TXTL. The matrix shows the endpoint GFP expression for TXTL reactions with non-targeting sgRNA, the anti-CRISPR protein, and the Cas9. The values are reported in comparison to the same TXTL reaction without the anti-CRISPR protein. Endpoints were taken after 18 hours of incubation. Values represent the mean of at least three technical replicates.

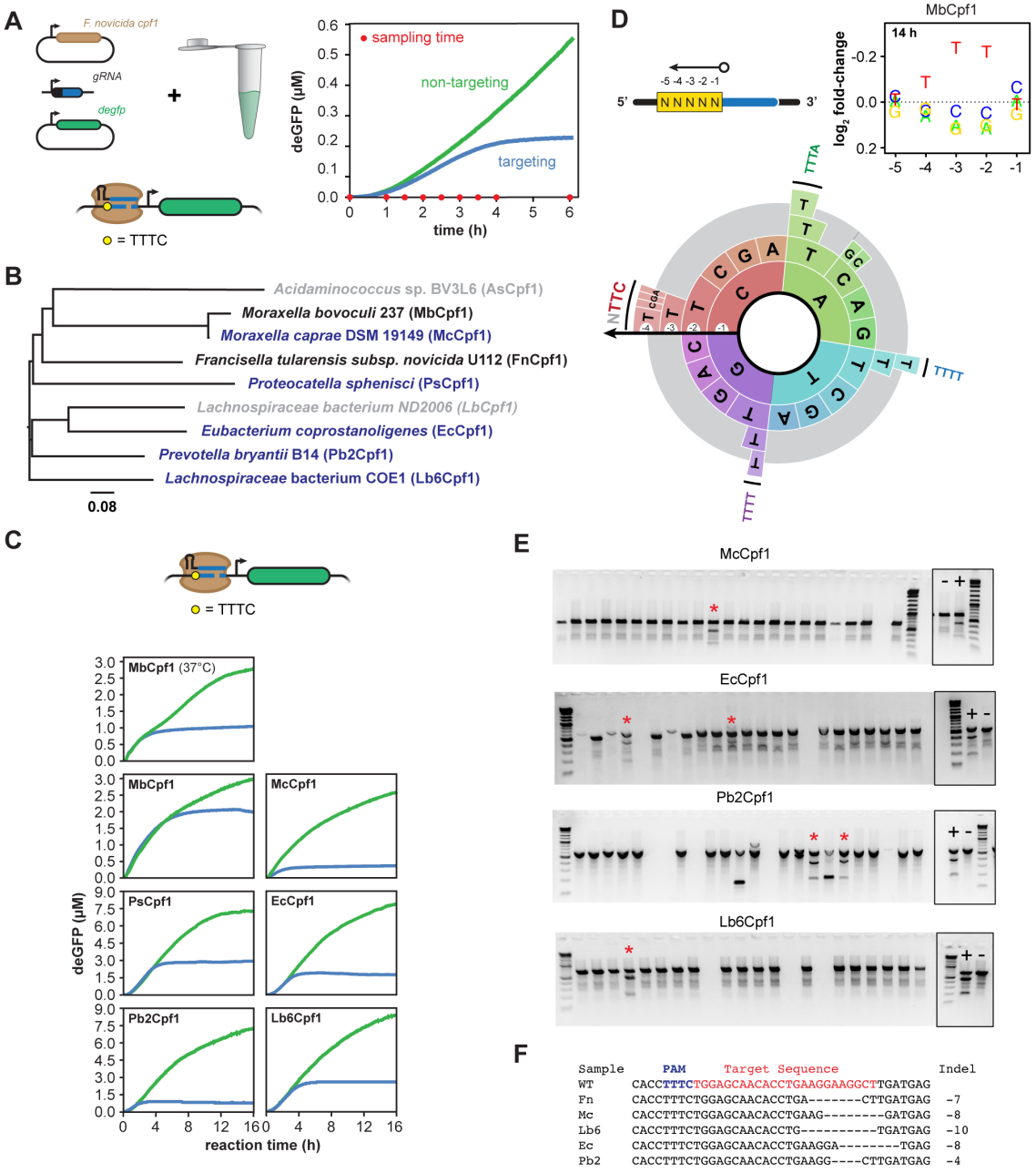


Figure S5. Related to Figure 6. Dynamics of DNA cleavage by FnCpf1. **A.** A time series is shown for deGFP concentration in TXTL reactions expressing FnCpf1, deGFP, and either a targeting guide RNA (g4) or non-targeting guide RNA (g-nt). Red dots show times where samples were collected in a parallel reaction expressing FnCpf1, a targeting guide RNA, and containing a 5N PAM library as part of the PAM determination assay (Figure 6A). **B.** Phylogenetic tree of Cpf1 nucleases that were previously uncharacterized (blue), previously characterized but also tested in this work (black), or previously characterized and not explicitly tested in this work (gray). Note that the PAM for MbCpf1 was previously reported in Zetsche et. al. *Cell* 2015, although the nuclease was not subsequently characterized. The tree was constructed using MUSCLE based on

the amino acid sequence of each nuclease. **C.** Time courses for TXTL reactions expressing a Cpf1 nuclease from a plasmid, a targeting or non-targeting sgRNA from linear DNA, and the deGFP reporter from a plasmid. In all cases, the target site was flanked by a 5' TTTC PAM. **D.** Results from the TXTL-based PAM determination conducted on MbCpf1. The assay was conducted at 37ºC with an incubation time of 14 hours. Note that depletion was limited compared to the other tested Cpf1 nucleases, resulting in a different scale for the vertical axis as compared to those in Figures 6B,E. Results of the assay are displayed as the fold-change of each nt at each position of the PAM library in comparison to the original PAM library (top) and the PAM wheel showing the determined PAM sequences (bottom). **E.** Representative T7EI assays screening rice callus for indels. Each lane corresponds to a separate callus piece, while + and – lanes correspond to positive and negative controls, respectively. Asterisks indicate samples that showed a T7EI signal indicative of an indel; these samples were further analyzed by Sanger sequencing to validate genome editing events. **F.** Genome editing events mediated by Cpf1 nucleases in rice. Representative deletions are shown that were produced by Lb, Mc, Lb6, and Ec Cpf1 nucleases and determined by Sanger sequencing. Dashes indicate deletions relative to the wild-type sequence.

Supplemental Table Legends

Supplemental Table S1. Related to Figures 1 - 6 and S1 - S5. Plasmid, gBlocks, bacterial cell stocks, and primers used in this study.

Supplemental Data

Supplemental Data S1. Related to Figure 6. Krona plots used to generate the PAM wheels. The plots capture the output of the TXTL-based PAM determination assays conducted on FnCpf1, Lb6Cpf1, Pb2Cpf1, EcCpf1, PsCpf1, McCpf1, and MbCpf1.

Supplemental Protocols

Supplemental Protocol S1. Related to Figure 1. Protocol for measuring the activity of CRISPR nucleases using TXTL.

Supplemental Protocol S2. Related to Figure 6. Protocol for performing the TXTL-based PAM determination assay.