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], Michelle L. Luo[2], Thomas Jacobsen[2], 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

*These authors contributed equally to this work

†Correspondence directed to cbeisel@ncsu.edu (C.L.B.) and noireaux@umn.edu (V.N.)

Key words: Cas9, Cascade, Cpf1, PAM, prototyping, synthetic biology, TXTL

Running title: Characterizing CRISPR with TXTL

ABSTRACT

CRISPR-Cas systems have offered versatile technologies for genome engineering, yet their implementation has been outpaced by the ongoing discovery of new Cas nucleases and anti-CRISPR proteins. Here, we present the use of *E. coli* cell-free transcription-translation systems (TXTL) to vastly improve the speed and scalability of CRISPR characterization and validation. Unlike prior approaches that require protein purification or live cells, TXTL can express active CRISPR machinery from added plasmids and linear DNA, and TXTL can output quantitative dynamics of DNA cleavage and gene repression. To demonstrate the applicability of TXTL, we rapidly measure guide RNA-dependent DNA cleavage and gene repression for single- and multi-effector CRISPR-Cas systems, accurately predict the strength of gene repression in *E. coli*, quantify the inhibitory activity of anti-CRISPR proteins, and develop a fast and scalable high-throughput screen for protospacer-adjacent motifs. These examples underscore the potential of TXTL to 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), allowing 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 regulation, base editing, and real-time imaging (Chen et al., 2016; Dominguez et al., 2016; Komor et al., 2017; Nelles et al., 2016).

While the vast majority of CRISPR-based technologies have relied on the DNA-targeting Cas9 nuclease from *Streptococcus pyogenes* (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 have 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 are 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 offer means to tightly control nuclease activity (Bondy-Denomy et al., 2015; Pawluk et al., 2016a; Pawluk et al., 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 characterization of 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 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 guide RNAs. Given the growing abundance of known CRISPR nucleases, the ease in guide rna design, and the growing prevalence of factors factors that interface with CRISPR-Cas systems, there remains a pressing need to develop rapid and scalable characterization methodologies.

Here, we address this need using an *Escherichia coli* cell-free transcription-translation system (TXTL) (Garamella et al., 2016; Shin and Noireaux, 2012). TXTL recapitulates gene expression and enzymatic activity following the addition of template DNA, allowing the quantitative and dynamic measurement of nuclease binding and cleavage at microliter scale in a few hours -- all without protein purification or live cells. We demonstrate the utility of TXTL across a diverse set of Cas nucleases and show how it can be used to predict guide RNA 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., 2011; Dudley et al., 2014; 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; Takahashi et al., 2015b; Tayar A., 2015).

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' untranslated region (UTR), and coding sequence of a deGFP reporter construct, a slightly modified version of eGFP with identical fluorescent properties (Shin and Noireaux, 2012). We then 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). The resulting fluorescence values were converted to a true reporter concentration based on a calibration curve made with purified eGFP. We found that the three tested sgRNAs resulted in greatly reduced deGFP concentrations in comparison to a non-targeting sgRNA after 16 hours in the TXTL reaction (Figure 1B). Measurable repression was observed beginning after less than one hour into the TXTL reaction, indicating that this time span was required to express and assemble the Cas9-sgRNA ribonucleoprotein complex (RNP) and for the complex to bind and cleave the target DNA (Figure S1A,B). 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 and of deGFP repression and the rate at which deGFP production dropped varied depending on the sgRNA used (Figure S1A,B), suggesting that varying efficiency by which different guides targeted Cas9 to DNA. 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 -- also called CRISPRi and CRISPRa (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, and the rate of deGFP production dropped after less an hour (Figure S1A,B). The catalytically-dead version repressed expression less strongly and exhibited some deGFP production even at the end of the TXTL reaction (Figures 1B, S1A). We confirmed that the dSpyCas9 did not cleave DNA by PCR amplification across the target location (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). These results illustrate that dSpyCas9 functions efficiently to repress gene expression in TXTL. Given the larger span of time in which regulatory activity can be observed for dSpyCas9 than for SpyCas9 (Figure S1A), we relied on CRISPR-based gene repression for all subsequent analyses.

To assess 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}, σ^{54}) or the 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, with each combination tested in triplicate. 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 the σ^{54} promoter (Figure S2), demonstrating that these sites were important for reporter expression. In total, we showed that TXTL can be used to rapidly and scalably assess sgRNA activity.

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 varying the amount of the dSpyCas9 plasmid reduced the total amount of GFP produced in a TXTL reaction (Figure S2A), suggesting that dSpyCas9 levels were limiting under these reaction conditions. Surprisingly, adding more dSpyCas9 plasmid elevated deGFP production for the non-targeting sgRNA, underscoring the need to add the same total amount of DNA when performing comparative studies 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 2B) (Schmidt et al., 2009). Appending the ssrA tag to deGFP resulted in an apparent enhancement of dSpyCas9-based repression (Figure 2B) due to the inability of deGFP to accumulate in the TXTL reaction (Figure S2B). This apparent enhancement offers a means to measure CRISPR-based repression under conditions that mimic dilution *in vivo*.

We next evaluated the impact of targeting 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 relaxed DNA may not efficiently recognize targets encoded on linear DNA. We measured the extent of dSpyCas9-based repression when targeting a plasmid or linear version of the deGFP reporter construct. The linear DNA was generated by PCR including flanks over one kilobase away from the closest target site, mitigating potential effects from the end of the dsDNA molecule. We measured the fold-repression of the deGFP produced by the reaction after five hours, which is shorter than our previous experiments because linear templates degrade after that time in our TXTL system

(Marshall et al., 2017). We found that dSpyCas9 was capable of repressing gene expression when targeting the transcribed sequence of the reporter gene on the linear template, but not when targeting the promoter (Figure 2C). These results suggest that the dSpyCas9 can block elongation but not initiation of the *E. coli* RNA polymerase on linear DNA. They also indicate that Cas nucleases interact differently with linear targets and plasmid targets, underscoring the need for further interrogation.

When DNA expressing dSpyCas9, sgRNA, and reporter plasmid DNA 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 assembly speed of the dSpyCas9-sgRNA RNP complex. To test this hypothesis, we varied our initial protocol to pre-express dSpyCas9 and the sgRNA before adding the reporter plasmid for three hours. We reasoned that this period of pre-expression would allow the RNP to 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 2C), with measurable repression occurring as fast as six minutes, depending on the sgRNA used (Figure S2C,D). We also evaluated the effect of pre-expressing dSpyCas9 in cells prior to generating the TXTL lysate (Figure S2E). Interestingly, pre-expression of dSpyCas9 generally did not enhance repression (Figure S2F), suggesting that expression of the sgRNA and assembly of the ribonucleoprotein complex strongly contribute 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 minutes), but that DNA binding is faster (on the order of 5 minutes).

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

guide RNAs, 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 opted to compare dSpyCas9-based repression in TXTL and in *E. coli* by targeting 19 different locations within the deGFP reporter plasmid (Figure 3). 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 of poor recognition of NAG PAMs than NGG PAMs by dCas9 (Boyle et al., 2017; Jiang et al., 2013) and reduced 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*, and fold-repression was calculated based on endpoint measurements of deGFP in comparison to a non-targeting sgRNA control.

These experiments revealed a strong correlation (R^2 = 0.90) between the measurements in TXTL and in *E. coli* (Figure 3). Consistent with this correlation, we observed greatly reduced repression for targets flanked by NAG PAMs 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 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 other CRISPR nucleases can be measured with TXTL. Our results showing that active SpyCas9 RNPs can be expressed in TXTL suggested that other CRISPR-Cas systems would also function in TXTL. To test this, we evaluated effector proteins from two other CRISPR-Cas systems: the single-effector nuclease Cpf1 from the Type V-A system in *Francisella novicida*, and the multi-subunit complex Cascade from the Type I-E system in *E. coli*

(Figure 4). These were selected because Cpf1 nucleases exhibits many desirable properties over Cas9 nucleases 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 both capable of programmable gene repression in bacteria when using a catalytically-dead version of the nuclease (dFnCpf1) or by 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). To test CRISPR-based repression for each system, we designed three guide RNAs (a mature guide RNA for Cpf1, a repeat-spacer-repeat 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 guide RNAs that targeted the reporter gene promoter (Figure 4A). The measured repression by dFnCpf1 after 16 hours of reaction was weaker than that achieved by dSpyCas9, which we attribute to the longer delay in the onset of repression (Figure S3A,C). Interestingly, pre-expressing dFnCpf1 and its guide RNA for three hours reduced the delay in measurable repression by approximately two hours. This finding suggests that complex assembly and DNA binding is slower for the FnCpf1-gRNA complex than for the dSpyCas9-sgRNA complex.

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 and S3B,D). 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 and S3B,D), indicating that the complex -- once expressed and assembled -- rapidly binds DNA and efficiently blocks RNA polymerase recruitment. Unexpectedly, deGFP production was consistently lower when expressing EcCascade versus any of the other effector proteins that was exacerbated when EcCascade was pre-expressed (Figure 4B), suggesting

222 that EcCascade is interfering with the expression or activity in deGFP. In total, we show that
223 TXTL can be extended to the characterization of CRISPR-Cas systems requiring both single-
224 effector and multi-protein effector complexes.
225
226 **TXTL can quantify the inhibitory activity of anti-CRISPR proteins.** Recently, the discovery
227 of anti-CRISPR proteins that bind and inhibit Cascade and Cas3 from Type I-E and Type I-F
228 CRISPR-Cas systems and Cas9 from Type II-A and II-C systems raised the potential of using
229 these proteins to tightly regulate genome editing and gene regulation (Bondy-Denomy et al.,
230 2015; Pawluk et al., 2016a; Pawluk et al., 2016b; Rauch et al., 2017). We asked if TXTL could
231 be used to rapidly assess the inhibitory activity of potential anti-CRISPR proteins. We focused
232 on AcrIIA2 and AcrIIA4, two anti-CRISPR proteins that were recently reported to inhibit the
233 activity of SpyCas9 *in vitro* and in human cells (Rauch et al., 2017). To measure the inhibitory
234 activity of AcrIIA2 and AcrIIA4 against SpyCas9 in TXTL, we encoded each protein on a linear
235 expression construct that was expressed for two hours in the lysate prepared from cells
236 expressing dSpyCas9 (Figures 5A, S2E). We then added DNA encoding the GFP reporter
237 plasmid and linear DNA encoding a targeting or non-targeting sgRNA and measured deGFP
238 fluorescence over time.
239 We found that both AcrIIA2 and AcrIIA4 counteracted gene repression by dSpyCas9.
240 However, AcrIIA4 fully restored the amount of deGFP produced by the reaction compared to the
241 non-targeting control, whereas AcrIIA2 only restored deGFP production by 34%. Therefore,
242 TXTL can be used to quantify the inhibitory activity of anti-CRISPR proteins, where the resulting
243 values can guide the rapid identification of potent anti-CRISPR proteins that can more
244 effectively inhibit Cas nuclease activity.
245
246 **TXTL offers a rapid means of elucidating CRISPR PAMs.** One of the major barriers to the
247 functional characterization of new Cas nucleases is determining the recognized PAM

248 sequences. While numerous experimental methods have been developed for PAM
249 determination, they consistently rely on *in vitro* assays that require protein purification or on cell-
250 based assays that require culturing and transforming of live cells (Karvelis et al., 2017; Leenay
251 and Beisel, 2017). We therefore asked if TXTL could also be used as the basis of PAM
252 determination assays to avoid protein purification and exploit the large library sizes that can be
253 screened with *in vitro* assays. Paralleling prior *in vitro* and *in vivo* DNA cleavage assays (Jiang
254 et al., 2013; Zetsche et al., 2015), our devised assay relies on introducing a library of potential
255 PAM sequences flanking a site targeted by an expressed guide RNA and Cas nuclease (Figure
256 5B). After incubating the pooled PAM library and DNA encoding the Cas nuclease and guide
257 RNA, the pool of uncleaved target sequences is PCR-amplified and subjected to next-
258 generation sequencing. By comparing the relative frequency of individual sequences within the
259 library before and after sequencing, we can quantify the depletion of each library member and
260 therefore how well the nuclease recognizes each sequence as a PAM. Critically, the assay can
261 be completed in ~10 hours from when the DNA constructs are in hand to when PCR products
262 are submitted for sequencing.

263 As a proof-of-principle demonstration, we assessed the PAM requirements of the well-
264 characterized FnCpf1 nuclease. Previous assays have revealed that FnCpf1 requires a TTN
265 motif on the 5' end of the target for efficient cleavage using the guide RNA-centric orientation,
266 although CTN can also be recognized (Fonfara et al., 2016; Leenay et al., 2016; Zetsche et al.,
267 2015). We created a five-nucleotide library of potential PAM sequences 5' to the guide RNA
268 target. *In vitro* PAM assays are known to yield less specific PAM sequences for higher nuclease
269 concentrations (Karvelis et al., 2015), so we assessed the determined PAMs in our assay for
270 reaction times ranging from one hour to six hours following the addition of the FnCpf1 and guide
271 RNA expression constructs (Figure S4). Figure 5 depicts the identified PAM sequences based
272 on the depletion of individual nucleotides at each position (Figure 5C) or specific motifs (Figure
273 5D) as well as a PAM wheel capturing the relative depletion of each sequence across the library

(Figure 5E) (Leenay et al., 2016). We did not include the -5 position in the PAM wheel because this position showed no appreciable bias across the four possible nucleotides.

Our TXTL-based PAM assay recapitulated the canonical 5' TTN PAM for FnCpf1 while revealing other distinct features. At a reaction time of six hours, NTTTN was the most active motif. Within this motif, ATTTA was the most depleted by ~3-fold more than the next most depleted sequences (Data file S1). We further found that NNCTN also supported efficient cleavage in line with previous reports (Fonfara et al., 2016; Zetsche et al., 2015), while a T at the -1 position was detrimental to cleavage. Taken together, these results indicate that a TXTL-based assay can elucidate PAM requirements for CRISPR-Cas systems, where the assay opens the door to the rapid and scalable characterization of PAM requirements for novel Cas nucleases.

DISCUSSION

We have demonstrated that *E. coli* cell-free TXTL can be used for the rapid characterization of CRISPR nucleases, guide RNAs, 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. We exploit these capabilities by (1) demonstrating DNA binding and cleavage by multiple types of Cas nucleases, (2) extracting the dynamics of component expression, complex formation, and DNA targeting, (3) show that the strength of gene repression strongly correlates between TXTL and *E. coli*, (4) characterize the inhibitory strength of anti-CRISPR proteins, and (5) devise a rapid PAM determination assay. With recent advances in DNA synthesis as well as liquid handling systems, these approaches 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. The issue is that 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 these multi-subunit 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 (see below) 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 could be used to investigate spacer acquisition across the diversity of Cas acquisition proteins found in nature. Our demonstration of the strong correlation between CRISPR-based repression *in vivo* and *in vitro* also suggests that TXTL could provide a rapid means to functionally validate guide RNAs before they are deployed for genome editing. However, more research is need to determine how binding or cleavage activity in TXTL correlates to genome editing and gene regulation in prokaryotic or eukaryotic cells or whether cellular factors such as DNA repair pathways or heterochromatin formation principally determine targeting efficiency. Next, we envision the development of further high-throughput screening assays using TXTL, such as assessing the sequence-dependence of guide RNA activity. 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 ribonucleoprotein 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.

ACKNOWLEDGEMENTS

We thank Ryan Leenay for help generating the FnCpf1 PAM wheel, and Jennie Fagan for cloning the plasmid expressing SpyCas9. 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 National Institutes of Health (grant 1R35GM119561-01, C.L.B.), and the National Science Foundation (grant CBET-1403135 to C.L.B.).

AUTHOR CONTRIBUTIONS

C.L.B. designed the experiments and wrote the paper. V.N. designed the experiments and wrote the paper. C.S.M. designed and performed the experiments and wrote the paper. R.M. designed and performed the experiments and wrote the paper. M.L.L. designed and performed the experiments. T.J. performed the experiments. S.P.C. performed the experiments.

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.

Bikard, D., Jiang, W.Y., 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 U S A *114*, 5461-5466.

Carlson, E.D., Gan, R., Hodgman, C.E., and Jewett, M.C. (2011). Cell-free protein synthesis: Applications come of age. Biotechnol Adv.

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. Annual review of biophysics *45*, 1-23.

371 Dominguez, A.A., Lim, W.A., and Qi, L.S. (2016). Beyond editing: repurposing CRISPR-Cas9
372 for precision genome regulation and interrogation. Nature reviews. Molecular cell biology *17*,
373 5-15.
374 Dudley, Q.M., Karim, A.S., and Jewett, M.C. (2014). Cell-free metabolic engineering:
375 Biomanufacturing beyond the cell. Biotechnol J.
376 Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lecrivain, A.L., Bzdrenga, J., Koonin,
377 E.V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of
378 dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic Acids Res *42*,
379 2577-2590.
380 Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-
381 associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature *532*,
382 517-521.
383 Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The All E. coli TX-TL Toolbox
384 2.0: A Platform for Cell-Free Synthetic Biology. ACS Synth Biol.
385 Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes,
386 C., Panning, B., Ploegh, H.L., Bassik, M.C.*, et al.* (2014). Genome-Scale CRISPR-Mediated
387 Control of Gene Repression and Activation. Cell *159*, 647-661.
388 Gootenberg, J.S., Abudayyeh, O.O., Lee, J.W., Essletzbichler, P., Dy, A.J., Joung, J., Verdine,
389 V., Donghia, N., Daringer, N.M., Freije, C.A.*, et al.* (2017). Nucleic acid detection with
390 CRISPR-Cas13a/C2c2. Science *356*, 438-442.
391 Hockenberry, A.J., and Jewett, M.C. (2012). Synthetic in vitro circuits. Curr Opin Chem Biol *16*,
392 253-259.
393 Jewett, M.C., Calhoun, K.A., Voloshin, A., Wuu, J.J., and Swartz, J.R. (2008). An integrated
394 cell-free metabolic platform for protein production and synthetic biology. Mol Syst Biol *4*, 220.
395 Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of
396 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., and Siksnys, V. (2017). Methods for decoding Cas9 protospacer adjacent motif (PAM) sequences: A brief overview. Methods *121-122*, 3-8.

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 biology *16*, 253.

Karzbrun, E., Tayar, A.M., Noireaux, V., and Bar-Ziv, R.H. (2014). 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. Nature communications *8*, 14500.

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.

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 chi sites enhances DNA stability and gene expression in E. coli cell-free transcription-translation systems. Biotechnol Bioeng *114*, 2137-2141.

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 e212.

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.

448 Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M.,
449 Sontheimer, E.J., Maxwell, K.L., *et al.* (2016a). Naturally Occurring Off-Switches for CRISPR-
450 Cas9. Cell *167*, 1829-1838 e1829.
451 Pawluk, A., Staals, R.H., Taylor, C., Watson, B.N., Saha, S., Fineran, P.C., Maxwell, K.L., and
452 Davidson, A.R. (2016b). Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in
453 diverse bacterial species. Nature microbiology *1*, 16085.
454 Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A.
455 (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of
456 Gene Expression. Cell *152*, 1173-1183.
457 Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P.R., and Lundgren, M. (2015). Efficient
458 programmable gene silencing by Cascade. Nucleic Acids Res *43*, 237-246.
459 Rauch, B.J., Silvis, M.R., Hultquist, J.F., Waters, C.S., McGregor, M.J., Krogan, N.J., and
460 Bondy-Denomy, J. (2017). Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. Cell *168*,
461 150-158 e110.
462 Schmidt, R., Bukau, B., and Mogk, A. (2009). Principles of general and regulatory proteolysis by
463 AAA+ proteases in Escherichia coli. Res Microbiol *160*, 629-636.
464 Shin, J., and Noireaux, V. (2012). An E. coli cell-free expression toolbox: application to synthetic
465 gene circuits and artificial cells. ACS Synth Biol *1*, 29-41.
466 Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E.,
467 Minakhin, L., Joung, J., Konermann, S., Severinov, K., *et al.* (2015). Discovery and Functional
468 Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol Cell *60*, 385-397.
469 Sun, Z.Z., Yeung, E., Hayes, C.A., Noireaux, V., and Murray, R.M. (2014). Linear DNA for Rapid
470 Prototyping of Synthetic Biological Circuits in an Escherichia coli Based TX-TL Cell-Free
471 System. Acs Synthetic Biology *3*, 387-397.
472 Swartz, J. (2006). Developing cell-free biology for industrial applications. J Ind Microbiol
473 Biotechnol *33*, 476-485.

474 Takahashi, M.K., Chappell, J., Hayes, C.A., Sun, Z.Z., Kim, J., Singhal, V., Spring, K.J., Al-
475 Khabouri, S., Fall, C.P., Noireaux, V.*, et al.* (2015a). Rapidly Characterizing the Fast
476 Dynamics of RNA Genetic Circuitry with Cell-Free Transcription-Translation (TX-TL) Systems.
477 ACS Synth Biol *4*, 503-515.
478 Takahashi, M.K., Hayes, C.A., Chappell, J., Sun, Z.Z., Murray, R.M., Noireaux, V., and Lucks,
479 J.B. (2015b). Characterizing and prototyping genetic networks with cell-free transcription-
480 translation reactions. Methods.
481 Tayar A., K.E., Noireaux V., Bar-Ziv R. (2015). Propagating gene expression fronts in a one-
482 dimensional coupled system of artificial cells. Nature Physics.
483 Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J., and Joung, J.K. (2017).
484 CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-
485 targets. Nature methods *14*, 607-614.
486 Westra, E.R., van Erp, P.B., Kunne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S.,
487 Jore, M.M., Semenova, E., Severinov, K.*, et al.* (2012). CRISPR immunity relies on the
488 consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and
489 Cas3. Mol Cell *46*, 595-605.
490 Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S.,
491 Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A.*, et al.* (2015). Cpf1 is a
492 single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759-771.
493 Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M.,
494 Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S.*, et al.* (2017). Multiplex
495 gene editing by CRISPR-Cpf1 using a single crRNA array. Nat Biotechnol *35*, 31-34.
496

497 **MATERIALS AND METHODS**

498 **Preparation of TXTL lysates and reactions.** The *E. coli* cell-free TXTL lysate was prepared
499 from BL21 Rosetta 2 from Novagen as described previously (Caschera and Noireaux, 2014).

The dSpyCas9-loaded TXTL lysate followed the same procedure, only the *E. coli* cells harbored a plasmid constitutively expressing dSpyCas9 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.

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. Fluorescence intensity measurements were quantified using a linear standard curve, spanning over three orders of magnitude, produced with pure recombinant eGFP (Cell Biolabs Inc., San Diego, CA). Error bars are the standard error of measurement based on at least three replicates run on the same day. GFP production rates were calculated by two-point numerical differentiation and smoothed with a five-point quadratic polynomial. 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.

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

552 multichannel pipette and loaded back into the BioTek Synergy H1. Single-point fluorescence at
553 485/528 nm excitation/emission as well as OD_{600} were measured from each well. Endpoint
554 fluorescence values were background subtracted using the fluorescence from cells lacking
555 deGFP and then normalized by the endpoint, background-subtracted OD_{600} value. Fold
556 repression was then calculated as the normalized fluorescence of the non-targeting sgRNA
557 strain divided by the normalized fluorescence of the targeting sgRNA strain.
558
559 **TXTL-based PAM assay.** The pET vector expressing FnCpf1 was a kind gift of Benson Hill
560 Biosystems. A PAM library with five randomized nucleotides flanking the spacer sequence was
561 prepared as described previously (Leenay et al., 2016). The crRNA expressing the FnCpf1
562 crRNA targeting the PAM library was expressed from the gBlock CSM-GB099. A TXTL reaction
563 was assembled as described above except the reaction contained 0.5 mM IPTG and the
564 following DNA components: 0.2 nM P70a-T7RNAP, 2 nM pET-FnCpf1, 2 μM of Chi6 annealed
565 oligos, 0.5 nM of the 5N PAM library, 2 nM of linear DNA expressing the crRNA. The reaction
566 was split into 5 μL reactions as above and incubated at 29˚C. Samples were then collected at
567 the specified time by cutting away the cap mat sealing individual reactions and immediately
568 freezing the reaction at -20˚C until subsequent use. The adapters were attached by PCR
569 amplification of the TXTL reaction using NEB's Q5 High-FIdelity DNA Polymerase with RL133
570 and RL134. The PCR reaction was then purified using Ampure beads before attaching unique
571 Nextera indices to each sample by PCR amplification. After a final PCR purification using
572 Ampure beads, each sample was normalized to 10 nM in a total volume of 20 uL and submitted
573 for next-generation sequencing with 150 single-end reads on an Illumina MiSeq machine. The
574 PAM wheel was generated as described previously (Leenay et al., 2016).
575

FIGURE LEGENDS

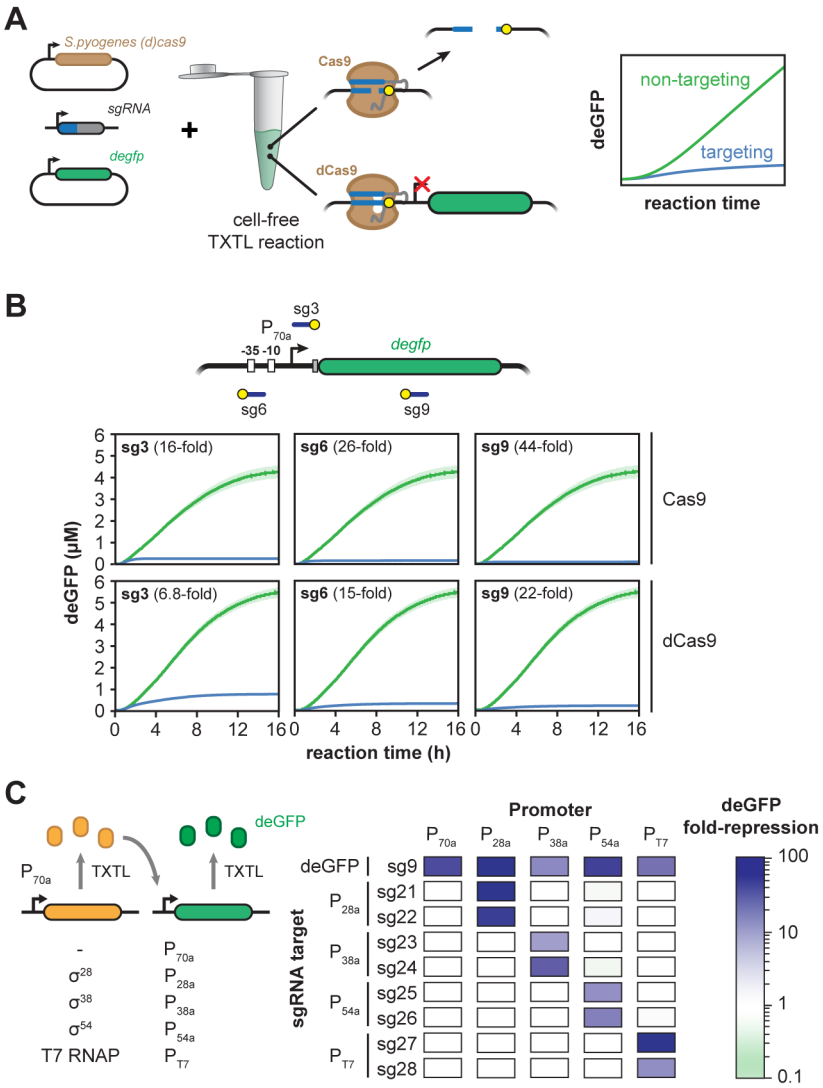


Figure 1. *S. pyogenes* Cas9 functions efficiently in TXTL. **A.** Schematic of using TXTL to dynamically and quantitatively measure the activity of Cas9 and the catalytically-dead Cas9 (dCas9) repression based on repressing expression of a deGFP reporter construct. Plasmids expressing a reporter gene (deGFP) and SpyCas9 or dSpyCas9 as well as linear DNA expressing an sgRNA are added to the cell-free reaction. sgRNAs direct the Cas9 or dCas9 ribonucleoprotein (RNP) complex to the reporter construct, leading to cleavage or binding by the RNP, respectively. Cleavage or binding can be monitored based on GFP fluorescence over time in a microplate reader. **B.** Time series showing deGFP concentration for cell-free reactions

586 expressing either Cas9 or dCas9 and a non-targeting sgRNA (green) or targeting sgRNAs
587 (blue). Target locations include the sequence matching the guide (blue line) and the PAM
588 (yellow circle). **C.** Alternative sigma factors σ28, σ38, and σ54 and the T7 polymerase can be
589 expressed in TXTL from the P_{70a} promoter and activate their cognate promoters P_{28a}, P_{38a}, P_{54a},
590 and P_{T7}, respectively. A matrix showing dSpyCas9-based repression of promoters dependent
591 on σ28, σ38, σ54 and the T7 polymerase is shown. An sgRNA targeting each promoter or the
592 GFP gene body was expressed along with each sigma factor or polymerase and a reporter
593 gene driven by the sigma factor of its cognate promoter. For each condition, repression was
594 calculated after a reaction time of 16 hours in comparison to a non-targeting sgRNA.
595

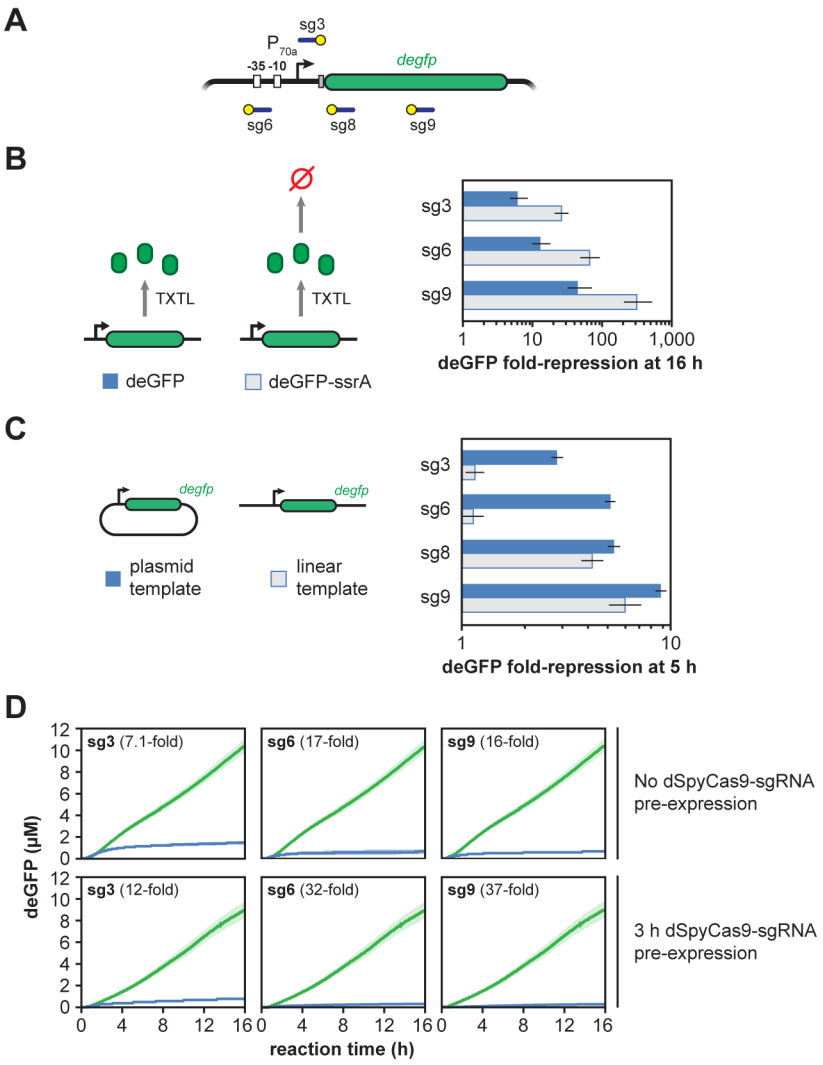
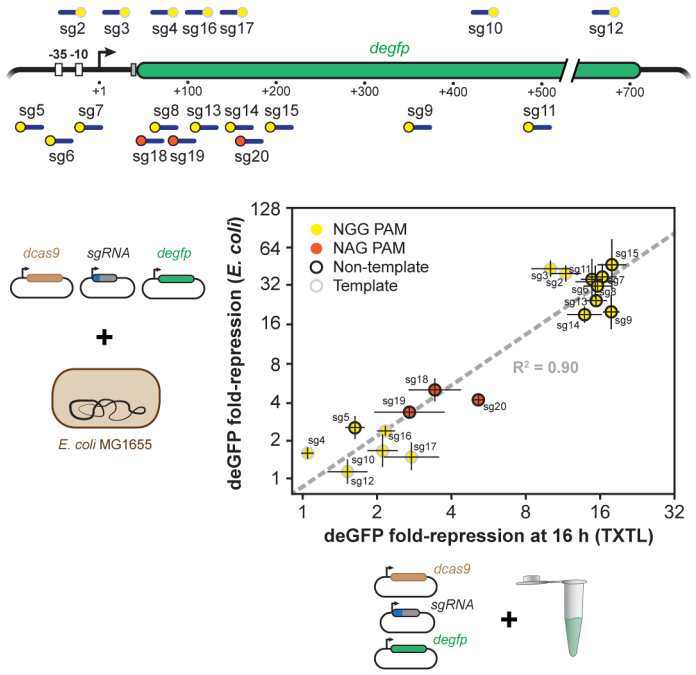


Figure 2. Multiple factors affect dSpyCas9-based repression of reporter gene expression in TXTL. **A.** Locations within the deGFP reporter construct targeted with the sgRNAs. Target locations include the sequence matching the guide (blue line) and the PAM (yellow circle). **B.** Fold-repression for reporter constructs encoding deGFP or deGFP-ssrA. The ssrA degron tag is recognized by the ClpXP protease that results in rapid turnover of the fusion protein. Fold-repression is the ratio of deGFP concentrations after 16 hours of reaction for the non-targeting (green) over the targeting (blue) sgRNA. **C.** Fold-repression produced by a TXTL reaction when deGFP is expressed from either a targeted linear or plasmid construct. **D.** Time series showing deGFP concentration in TXTL for cell-free reactions expressing dSpyCas9 and a targeting

606 sgRNA. The reporter plasmid is added to the reaction either at the same time as dCas9 and the

607 sgRNA (top row) or after a 3 hour pre-expression of dSpyCas9 and the sgRNA (bottom row).



608

609 **Figure 3.** dSpyCas9-based repression *in vivo* and *in vitro* are well-correlated. A schematic of
610 where each guide binds in the GFP promoter and gene body is shown (top). The location of the
611 target and PAM are indicated by a blue line and a yellow or orange dot, respectively. The fold-
612 repression of GFP production by CRISPRi for a guide RNA *in vivo* and *in vitro* is shown
613 (bottom). Points are colored by whether the guide is adjacent to an NGG (yellow) or NAG
614 (orange) PAM, and whether the sgRNA targets the non-template (black ring) or non-template
615 (gray ring) strand.

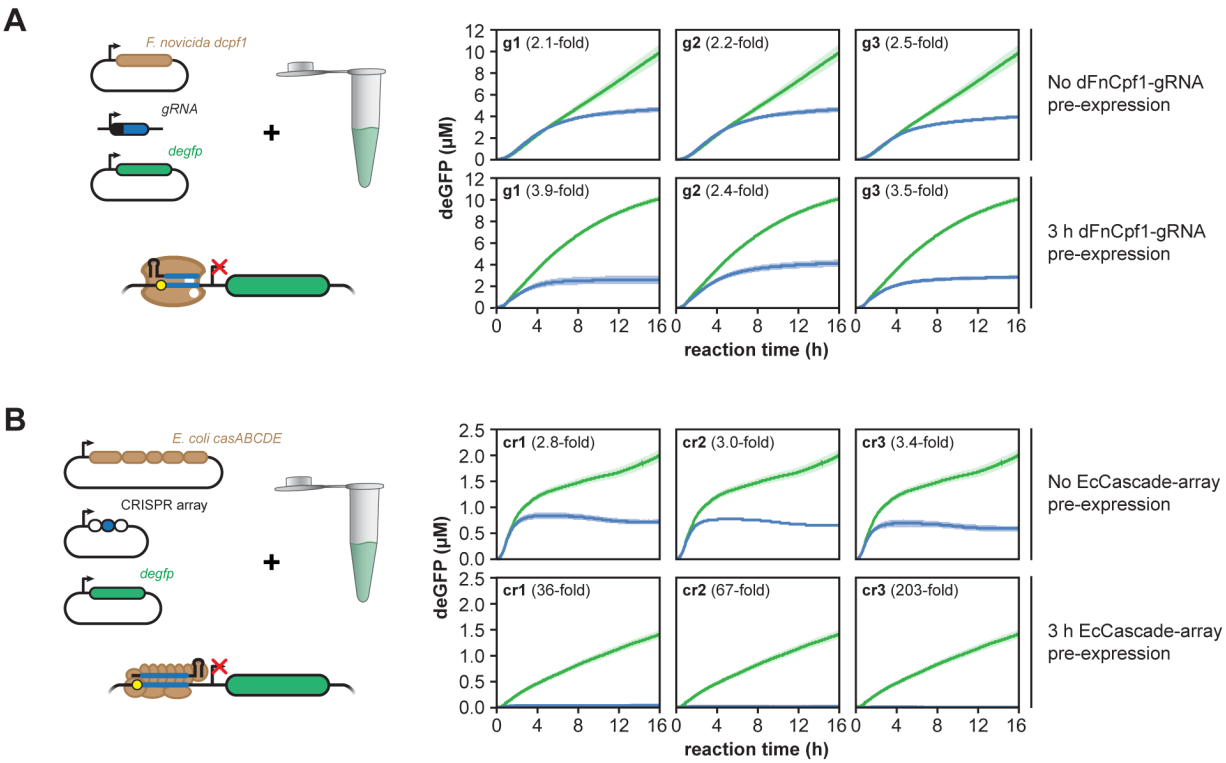


Figure 4. Single effector and multi-protein effector Cas proteins function efficiently in TXTL. Time series of reporter gene expression in TXTL for cell-free reactions expressing **A.** a catalytically inactive version of the Type V-A Cpf1 nuclease from *Francisella novicida* or **B.** the Type I-E Cascade complex from *E. coli*. The protein or set of proteins were expressed along with a non-targeting guide RNA (green) or one of three guide RNAs (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 guide RNA (top row) or after three hours of pre-expression (bottom row). The guide RNAs were expressed as a mature CRISPR RNA (FnCpf1) or as a repeat-spacer-repeat (EcCascade).

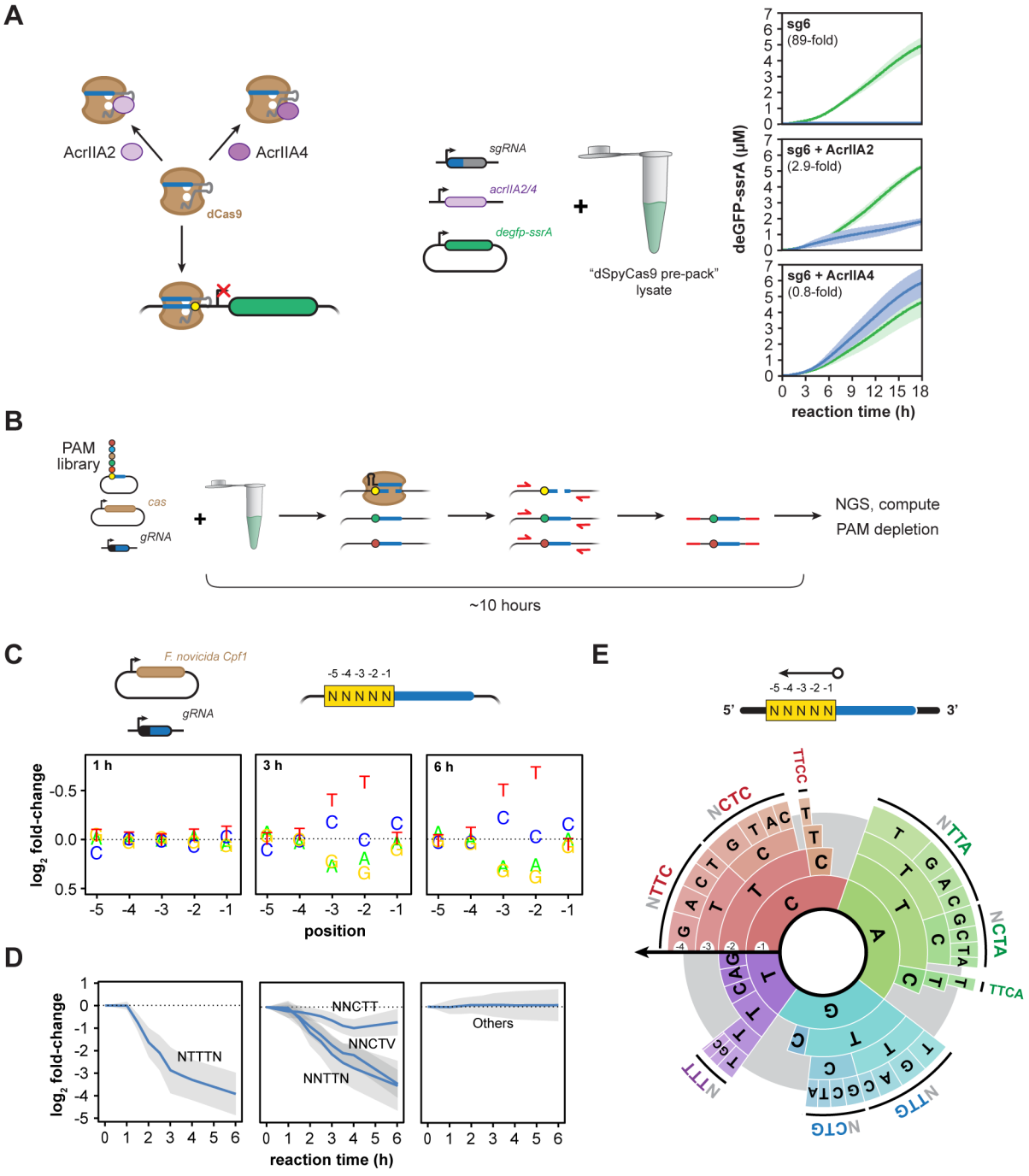


Figure 5. Using TXTL to characterize anti-CRISPR proteins and to determine PAMs. **A.** Time series of deGFP-ssrA expression in TXTL for cell-free reactions also expressing dSpyCas9, an sgRNA, and one of two anti-CRISPR proteins, AcrIIA2 and AcrIIA4, shown to inhibit SpyCas9 activity. Each reaction was performed with a targeting sgRNA (blue) or a non-targeting sgRNA (green). **B.** Schematic of a TXTL-based cleavage assay to determine the PAM sequences

recognized by Cas nucleases. Plasmids expressing the Cas nuclease and a guide RNA targeting a sequence flanked by randomized base pairs are combined and incubated in TXTL. Plasmids containing valid PAMs are cleaved by the Cas nuclease, while plasmids lacking a PAM are uncut. Sequencing libraries are then created by PCR amplifying across the target site, resulting in depletion of PAM sequences in the subsequent sequencing library. **C.** 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. Note that the y-axis is inverted to highlight nucleotides that are depleted. **D.** Time series showing the depletion of selected motifs matching the consensus sequence in the sequencing libraries are shown. Error bars show the standard deviation of the fold-change. **E.** 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.