Published in final edited form as:

Bioessays. 2014 January ; 36(1): 34–38. doi:10.1002/bies.201300135.

Exploiting CRISPR/Cas systems for biotechnology

Timothy R. Sampson^{1),2),3),*} and David S. Weiss^{2),3),4)}

¹⁾Department of Microbiology and Immunology, Microbiology and Molecular Genetics Program, Emory University School of Medicine, Atlanta, GA, USA

²⁾Emory Vaccine Center, Emory University, Atlanta, GA, USA

³⁾Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA

⁴⁾Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA

Abstract

The Cas9 endonuclease is the central component of the Type II CRISPR/Cas system, a prokaryotic adaptive restriction system against invading nucleic acids, such as those originating from bacteriophages and plasmids. Recently, this RNA-directed DNA endonuclease has been harnessed to target DNA sequences of interest. Here, we review the development of Cas9 as an important tool to not only edit the genomes of a number of different prokaryotic and eukaryotic species, but also as an efficient system for site-specific transcriptional repression or activation. Additionally, a specific Cas9 protein has been observed to target an RNA substrate, suggesting that Cas9 may have the ability to be programmed to target RNA as well. Cas proteins from other CRISPR/Cas subtypes may also be exploited in this regard. Thus, CRISPR/Cas systems represent an effective and versatile biotechnological tool, which will have significant impact on future advancements in genome engineering.

Keywords

biotechnology; Cas9; CRISPR; genome editing; RNAi

Introduction

CRISPR (clustered, regularly interspaced, short, palindromic repeats)/Cas (CRISPR-associated) systems are a unique prokaryotic defense against foreign genetic elements, such as those from bacteriophages or plasmids [1]. Functionally, CRISPR/Cas systems act as RNA-directed endonuclease complexes that are capable of targeting foreign nucleic acids in a sequence-specific fashion. Each individual system consists of a CRISPR array (crRNA array) composed of unique spacer sequences (21–72 base pairs [bp]) flanked by short (23–47 bp), repetitive and sometimes palindromic repeat sequences, as well as groups of

conserved Cas proteins encoded adjacent to the crRNA array [2]. Following transcription, the crRNA array is processed into individual CRISPR RNAs (crRNA) containing a spacer and a partial repeat. The spacers hybridize to complementary nucleic acid targets, triggering their degradation by Cas proteins [1]. A distinguishing feature of CRISPR/Cas systems compared to other sequence specific bacterial defenses, such as restriction modification systems, is that CRISPR/Cas systems are adaptive. Specific Cas proteins recognize foreign genetic elements and integrate their DNA as new spacer sequences into the crRNA array, ultimately allowing the bacteria to adapt and subsequently target these foreign elements [3].

Type II CRISPR/Cas systems are generally regarded as the simplest of the three known types, requiring only a single Cas protein, Cas9, for crRNA maturation and nucleic acid targeting and cleavage [4]. As the crRNA array is transcribed, a unique small RNA, termed trans-activating CRISPR RNA (tracrRNA), hybridizes to the repeat sequences, creating a double stranded RNA structure [4]. This complex is recognized and cleaved by RNase III, maturing the crRNA array into numerous small targeting crRNAs [4]. tracrRNA and the individual crRNAs are associated with Cas9, and upon hybridization of the crRNA with its complementary target (~30 bp length and 100% identity), two endonuclease domains of Cas9 each mediate cleavage of a strand of the targeted DNA [5]. This provides the prokaryotic cell with a sequence specific and adaptable mechanism for the cleavage of foreign DNA.

Cas9 as a programmable DNA targeting platform

Due to its specificity, coupled with the short sequence requirements for Cas9: crRNA targeting, it was hypothesized that Cas9 could be programmed to target and cleave any gene of interest within the bacterial cell. In fact, work by Jinek et al. [5] clearly demonstrated that introduction of a synthetic crRNA sequence capable of hybridizing to the DNA of a gene of interest, allows Cas9 to cleave that region. Targeting and subsequent cleavage requires a short (3–9 bp) sequence motif directly adjacent to the hybridized region, termed the Protospacer Adjacent Motif, or PAM, which is a necessary prerequisite for the Cas9:dual RNA complex to recognize the target sequence and prevent self-targeting.

Perhaps most importantly, the requirement for tracrRNA-mediated maturation could be abrogated when a synthetic double-stranded targeting RNA (a guide RNA or gRNA) was engineered (Fig. 1A) [5]. This specific RNA combines the targeting features of the crRNA with the dsRNA structure formed by the tracrRNA:crRNA complex, allowing this single RNA to fulfill the requirements of both small RNAs (sRNA) [5]. This development increases the portability of the Cas9 system between both prokaryotic and eukaryotic models, by simplifying the components needed to initiate cleavage of targets by Cas9 (e.g. there is no longer a need for an accessory RNA [tracrRNA], nor the need for RNase III to mature the crRNA:tracrRNA complex) [5].

These advances paved the way for multiple groups to port the Cas9 system into a number of different experimental models. To date, Cas9 genome editing has been successfully performed in bacterial cells, yeast, plants, nematodes, fruitflies, zebrafish, rodents, and human cells (both transformed cell lines and embryonic stem cells), solidifying its use as a

convenient tool to site-specifically edit the genomes of multiple species [5–13]. Editing by Cas9 can occur in one of two ways (Fig. 1A). When both strands of the DNA target are cleaved by Cas9, the cell can undergo non-homologous end joining (NHEJ) to repair this double strand break. NHEJ often results in the loss or addition of nucleotides, causing frame-shift mutations or early stop codons and ultimately, the loss of function of the targeted gene (Fig. 1A) [12, 14]. Alternatively, following the cleavage of DNA, a donor construct containing a selectable (or non-selectable) marker flanked by sequences adjacent to the cleavage site can be introduced. This donor construct acts as a template for homology-directed repair (HDR) and results in the insertion of the marker into the targeted site (Fig. 1A) [12, 14].

In order to decrease the frequency of NHEJ and subsequently increase the relative frequency of HDR, a specific Cas9 point mutant was engineered [12, 14]. Generation of a single point mutant in the active site of one of the two critical Cas9 endonuclease domains (RuvC-I) renders Cas9 unable to create double strand breaks [5]. Rather, the remaining intact endonuclease domain cleaves only one targeted strand, resulting in nicked DNA [5]. In the presence of a donor construct, these nicks are preferentially repaired by HDR. Use of this form of Cas9 significantly increases the frequency of marked mutations [12, 14]. Together, these recent developments have increased the ease of generating site-specific mutations in many organisms, and enhanced the ability to do so in systems that have been recalcitrant to genetic manipulation, hence opening numerous avenues for genetic research.

Engineering Cas9 to control transcription

While the aforementioned applications of Cas9 allow manipulation of genomic content, there are cases in which gene disruption is not necessarily possible or desirable. Instead, altering the expression of the gene (or multiple genes) of interest could provide more useful information, without permanently altering the cell's genome. Because the ability of Cas9 to associate with targeted DNA does not depend on its endonuclease activities, an engineered Cas9 mutant that cannot cleave DNA can be employed [15, 16]. Termed dCas9, this protein completely lacks catalytic activity against DNA, due to a point mutation within each of its critical endonuclease domains (both RuvC-I and HNH) [5, 15, 16]. When guided to the target DNA by a gRNA, dCas9 binds, but does not cleave the target and instead prevents transcription by blocking RNA polymerase binding or elongation (Fig. 1B) [15, 16]. By guiding dCas9 to different locations along the gene sequence, or by targeting multiple sites within the same gene simultaneously, dCas9 can be used to modulate the level of repression [15, 16]. Thus, Cas9 has the capability to not only edit genomes, but also to act as a transcriptional repressor.

The ability of dCas9 to site specifically bind DNA without causing cleavage can be exploited even further. dCas9 can be tethered to a transcriptional activator, such as the omega subunit of RNA polymerase within a bacterial system or VP64 in a mammalian system [16–19]. Utilizing a gRNA to program the dCas9-activator fusions to a region upstream of the transcriptional promoter, these systems can successfully recruit RNA polymerase resulting in an increase in transcription of the gene of interest (Fig. 1C). Similar to the ability of dCas9 to repress expression, this transcriptional activation can be modulated

and tuned to differentially express a specific gene [16–19]. The ability to control both activation and repression of genes of interest with Cas9 allows a number of experimental systems to be probed. For instance, bacterial systems currently rely on inducible and repressible promoters, requiring both the introduction of a new promoter to the gene of interest, and subsequently the addition of a small molecule. The dCas9-activation or repression systems would successfully bypass many of these steps and allow experimental questions to be addressed more quickly and in both prokaryotic and eukaryotic systems that currently lack controllable expression systems.

Can Cas9 interact with RNA targets?

Cas9 targeting of DNA has been exploited to allow genome editing and transcriptional repression and activation. While extremely useful, targeting of DNA alone has some potential pitfalls. For instance, in a eukaryotic system, repression of one transcript may result in the loss of multiple splice variants, when targeting of only one is warranted. Alternatively, binding of Cas9 to DNA may be inhibited by DNA structure or chromosomally bound proteins, preventing successful access to the targeted gene. It would therefore be a significant enhancement of Cas9 technology if differential targeting of either RNA or DNA targets could occur.

Recently, we demonstrated that the Cas9 endonuclease encoded by the bacterial pathogen *Francisella novicida* is involved in the repression of a specific transcript [20]. Our data suggest that *F. novicida* Cas9 (FnCas9) is guided to this mRNA by the tracrRNA and a novel small RNA termed scaRNA (small, CRISPR/Cas-associated RNA). Subsequently, the Cas9:scaRNA:tracrRNA: mRNA interaction results in decreased mRNA stability and low levels of expression of the target (Fig. 2A) [20]. It is therefore tempting to speculate that FnCas9, and possibly other Cas9s, may facilitate programmable RNA targeting (Fig. 2B). Furthermore, the FnCas9 system likely represents a tool that can be harnessed to understand both the structural and sequence requirements that determine how Cas9 can preferentially target RNA or DNA.

Many questions remain regarding how FnCas9 is capable of targeting RNA. This property may be dictated by the FnCas9 protein itself. While it contains all the predicted endonuclease motifs present in other Cas9 proteins [20, 21], none of the most conserved residues in these regions are required for its ability to target mRNA [20]. Instead, only a single domain, an arginine-rich motif (ARM) has been identified as important for mRNA targeting by Cas9 [20]. While the endonuclease motifs and ARM are highly conserved between all Cas9 species, regions of significant dissimilarity do exist [21]. This may suggest that FnCas9 has unique domains, not identified in other Cas9 proteins, which allow interaction with RNA targets. Alternatively, FnCas9 may simply act as a scaffold, allowing a dsRNA structure to form between the tracrRNA and the targeted mRNA, resulting in degradation of the target by endogenous RNases, which detect and cleave dsRNA (Fig. 2A). Another intriguing possibility is that rather than being dictated by FnCas9 per se, targeting of RNA is mediated by nonidentity interactions between the guiding RNA and the target. This process could require short regions of identity interspersed by regions of dissimilarity, as we predict to occur between tracrRNA and its targeted mRNA in *F. novicida* [20].

Alternatively, RNA targeting may simply require a longer sequence interaction, rather than the short, ~20 base interactions between a gRNA and its DNA target.

By understanding these interactions, there exists the potential to program FnCas9 to target RNAs of interest in the context of various biological systems. This would allow FnCas9 to function as a new form of RNA interference. Currently, it is unknown if FnCas9 can be reprogrammed to target diverse RNAs, or if it is capable of functioning in the eukaryotic cytosol. Similarly, a distinct CRISPR/Cas subtype (Type III) present in *Pyrococcus furiosus* has been demonstrated to target RNA substrates [22]. While this system requires a complex of six proteins (Cmr1, 3–6, and Cas10), it has been shown to utilize longer targeting RNAs, suggesting that it may have increased specificity compared to a Cas9-based system [22]. Development of these systems could provide benefits over RNAi systems currently used in eukaryotic models. Namely, it would be hypothesized that FnCas9 or the *P. furiosus* Cmr complex would not rely on any canonical RNAi host factors, such as Dicer or the components of the RISC complex, allowing their efficient use in model systems which may have intrinsic inhibition of RNAi. Therefore, the exploitation and engineering of FnCas9, or other RNA-targeting CRISPR/Cas systems, as programmable RNA-directed RNA targeting systems would likely be beneficial experimental tools in the study of numerous biological systems.

Caveats of Cas9 technologies

While the use of Cas9 as an efficient tool in the study of a wide array of organisms has been clearly demonstrated, there are caveats that must be further clarified in order to maximize its functionality. Firstly, while Cas9 is directed to its binding site by a gRNA with near 100% identity to the target, off-target effects do occur, specifically at sequences lacking identity in the 5' region of the gRNA [17, 23–25]. This means that during targeting there may be unforeseen pleiotropic effects that confound the conclusions of experiments.

In order to combat non-target interactions, numerous strategies could be used. Creating gRNA secondary structures that inhibit non-target interactions, altering the length of the gRNA, or utilizing paired nickases to create off-set dsDNA breaks, all decrease the amount of off-target effects [17, 23–26]. Alternatively, understanding the protein structure requirements of Cas9 recognition of the PAM to initiate cleavage may allow Cas9 proteins to be engineered with more stringent specificity. In fact, the PAM for the Cas9 protein from *Neisseria meningitidis* (NmCas9) was recently described. Its 5'-NNNNGATT-3' sequence as well as the PAM required for targeting by Cas9 from *Streptococcus thermophilus* (5'-NNAGAAW-3') are both more stringent than that required by *Streptococcus pyogenes* Cas9 (5'-NGG-3'), which is currently the most commonly used for biotechnological applications [5, 13, 27]. Cas9 proteins with more stringent PAMs may result in fewer off-target effects, since cleavage requires the presence of a less frequent sequence. Comparisons between various Cas9 systems may permit the elucidation of the requirements for PAM recognition, allowing even more specific Cas9 proteins to be engineered.

Conclusions and outlook

Despite the caveats discussed above, Cas9 represents one of the most significant advances in genome engineering technologies. By exploiting this intricate genetic targeting system from prokaryotes, we are now able to quickly edit or modulate genomes, including those that have until now remained refractory to such interventions. In fact, such engineering technologies have recently been multiplexed to allow genome editing, and transcriptional activation and repression within the same cell, utilizing Cas9 orthologs from different species [28]. Furthermore, the observation that FnCas9 may act to directly target RNA opens new avenues of RNA interference that have not yet been fully appreciated.

Perhaps most excitingly, the continued development of the Cas9 targeting systems may provide the foundation for their future use in gene therapy. By delivering Cas9 and specific gRNAs and donor constructs, this system may be used to introduce or remove genetic information in order to treat various genetic disorders. Future comparisons of the activity of Cas9 proteins from diverse species will certainly reveal how these proteins can be engineered to interact with either DNA or RNA, and how to increase target specificity; a necessary requirement if therapeutic capabilities are to be used in the future.

Additionally, other types of CRISPR/Cas systems may also prove useful as tools to target and alter specific sequences. Other than the RNA targeting *P. furiosus* Type III system, DNA targeting by Type I or Type III CRISPR/Cas systems utilizes a complex of nucleases distinct from Cas9 of Type II systems. While engineering these systems may be slightly more complicated since they require multiple protein components, these CRISPR/Cas variants may have benefits compared to Cas9 systems due to the potential for different tolerances for mis-matches in the target sequence, length requirements of the targeting RNA, and the specificity of cleavage. Ultimately, such engineering of not only Cas9 systems, but other CRISPR/Cas types as well, will maximize the utility and versatility of these prokaryotic-derived restriction systems as experimental tools and therapeutics.

Acknowledgments

Due to space constraints, we were unable to cite all relevant studies, and sincerely apologize to those authors whose work was not mentioned. This work was supported by National Institutes of Health (NIH) grants U54-AI057157 from the Southeastern Regional Center of Excellence for Emerging Infections and Biodefense and R56-AI87673 to D.S.W., who is also supported by a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award. T.R. S. was supported by the NSF Graduate Research Fellowship, as well as the ARCS Foundation.

References

1. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. *Nature*. 2012; 482:331–8. [PubMed: 22337052]
2. Makarova KS, Haft DH, Barrangou R, Brouns SJ, et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*. 2011; 9:467–77. [PubMed: 21552286]
3. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res*. 2012; 40:5569–76. [PubMed: 22402487]
4. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. 2011; 471:602–7. [PubMed: 21455174]
5. Jinek M, Chylinski K, Fonfara I, Hauer M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012; 337:816–21. [PubMed: 22745249]

6. DiCarlo JE, Norville JE, Mali P, Rios X, et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 2013; 41:4336–43. [PubMed: 23460208]
7. Nekrasov V, Staskawicz B, Weigel D, Jones JD, et al. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat Biotechnol.* 2013; 31:691–3. [PubMed: 23929340]
8. Lo TW, Pickle CS, Lin S, Ralston EJ, et al. Precise and heritable genome editing in evolutionarily diverse nematodes using TAL-ENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics.* 2013; 195:331–48. [PubMed: 23934893]
9. Yu Z, Ren M, Wang Z, Zhang B, et al. Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics.* 2013; 195:289–91. [PubMed: 23833182]
10. Hwang WY, Fu Y, Reyon D, Maeder ML, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol.* 2013; 31:227–9. [PubMed: 23360964]
11. Wang H, Yang H, Shivalila CS, Dawlaty MM, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell.* 2013; 153:910–8. [PubMed: 23643243]
12. Mali P, Yang L, Esvelt KM, Aach J, et al. RNA-guided human genome engineering via Cas9. *Science.* 2013; 339:823–6. [PubMed: 23287722]
13. Hou Z, Zhang Y, Propson NE, Howden SE, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci USA.* 2013; 110:15644–9. [PubMed: 23940360]
14. Cong L, Ran FA, Cox D, Lin S, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013; 339:819–23. [PubMed: 23287718]
15. Qi LS, Larson MH, Gilbert LA, Doudna JA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell.* 2013; 152:1173–83. [PubMed: 23452860]
16. Bikard D, Jiang W, Samai P, Hochschild A, et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 2013; 41:7429–37. [PubMed: 23761437]
17. Mali P, Aach J, Stranges PB, Esvelt KM, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol.* 2013; 31:833–8. [PubMed: 23907171]
18. Cheng AW, Wang H, Yang H, Shi L, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 2013; 23:1163–71. [PubMed: 23979020]
19. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods.* 2013; 10:973–6. [PubMed: 23892895]
20. Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, et al. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature.* 2013; 497:254–7. [PubMed: 23584588]
21. Makarova KS, Aravind L, Wolf YI, Koonin EV. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol Direct.* 2011; 6:38. [PubMed: 21756346]
22. Hale CR, Zhao P, Olson S, Duff MO, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell.* 2009; 139:945–56. [PubMed: 19945378]
23. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* 2013; 41:9584–92. [PubMed: 23939622]
24. Pattanayak V, Lin S, Guilinger JP, Ma E, et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol.* 2013; 31:839–43. [PubMed: 23934178]
25. Hsu PD, Scott DA, Weinstein JA, Ran FA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol.* 2013; 31:827–32. [PubMed: 23873081]
26. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* 2013; 154:1380–9. [PubMed: 23992846]

27. Deveau H, Barrangou R, Garneau JE, Labonte J, et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol*. 2008; 190:1390–400. [PubMed: 18065545]
28. Esvelt KM, Mali P, Braff JL, Moosburner M, et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods*. 2013; 10:1116–21. [PubMed: 24076762]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Diagram of current Cas9 genome engineering technologies. **A:** Interaction of the Cas9:gRNA complex with its dsDNA target. Targeting gRNA sequence shown in dark blue, with the targeted sequence shown in light blue. Cas9 endonuclease motifs (shown as red circles) cleave either strand adjacent to the proto-spacer adjacent motif (PAM, stars). The subsequent repair by (I) non-homologous end joining (NHEJ) when both strands are cleaved can result in mutations (red). Alternatively, (II) homology-directed repair (HDR) can occur when a single target strand is cleaved and a repair donor is present (green). **B:** Cas9 with non-functional endonuclease motifs (gray circles, dCas9) can act as a transcriptional repressor by targeting a promoter region (light blue, dashed arrow) and preventing RNA polymerase association. **C:** dCas9 fused to a transcriptional activator (yellow star) can act as a transcriptional activator by associating upstream of a promoter (light blue, dashed arrow) and recruiting RNA polymerase.

Figure 2. Schematic of FnCas9 interaction with an RNA target. **A:** FnCas9 associates with a dsRNA complex formed by two small RNAs, tracrRNA (black) and the scaRNA (blue). Together, this allows tracrRNA to target an mRNA transcript (green). Subsequently, the mRNA target's stability is reduced and the transcript lost. This likely occurs by either currently unidentified FnCas9 activity or by the action of endogenous RNases. **B:** Schematic representation of a hypothetical tracrRNA:scaRNA hybrid which has been reprogrammed (gray) to target a new mRNA (orange).