



Minimizing off-target effects in CRISPR-Cas9 genome editing

Shi-Jie Chen 

Received: 27 June 2019 / Accepted: 10 July 2019 / Published online: 17 July 2019
© Springer Nature B.V. 2019

CRISPR (clustered regularly interspaced short palindromic repeats), originally an antiviral immune system adopted by bacteria and archaea, is repurposed and developed into a highly efficient tool for genome editing. Central to the CRISPR system is a complex machinery formed by a Cas protein, guide RNA (gRNA or sgRNA), and the target DNA. Two key factors that determine the specificity of CRISPR gene editing are (a) the hybridization between Cas/sgRNA and the target, as directed by the sequence recognition at the protospacer adjacent motif (PAM) site and the DNA target site, and (b) the subsequent specific conformational changes in the Cas/sgRNA/DNA complex for the cleavage reaction.

CRISPR has proven to be a highly versatile tool for gene editing with tremendous potential in a wide range of problems such as gene therapy, drug discovery, and genetic modification in plant technology. However, the accuracy and reliability of the CRISPR technology are severely hampered by the off-target effects, namely, the unintended cleavage of DNA at sites whose sequences show mismatches with the guide RNA (gRNA or sgRNA) (Wang and Wang, 2019). Therefore, reducing the off-target effect becomes a timely critical issue in CRISPR genome editing technology.

Class II CRISPR with Cas9 system, which involves a single Cas protein, is the most well studied and explored among the different types of CRISPR systems. Significant efforts have been made to develop effective methods to reduce the off-target effects. The general strategy of these methods is to stabilize on-target binding stability and destabilize off-target stability. The purpose of this short article is to present a very brief overview and assessments of the different methods and potential future developments.

The first type of method employs strategies to strengthen the on-target stability (Ran et al. 2013; Tsai et al. 2014; Shen et al. 2014; Guilinger et al. 2014), for example, through the use of a Cas9 nickase mutant or dimeric Cas9 proteins complexed with pairs of sgRNAs (Ran et al. 2013). These approaches effectively introduce double checkpoints for target recognition by increasing the number of matched base pairs in the target site. These methods are highly effective and can lead to a significantly reduction in off-target frequency (e.g., by 50–1000 folds in cell lines; Ran et al. 2013). These methods, however, are not without limitations. For example, these methods involve multiple components of the CRISPR-Cas9 system, which could pose challenges in gene delivery that requires the concurrent delivery of two guide RNAs.

The second type of method focuses on monomeric Cas9/sgRNA system. The primary strategy of the method is to destabilize off-target binding without sacrificing on-target cleavage efficiencies. Along this line, multiple approaches have been developed. For example, truncated (less than 20 nucleotides) gRNA sequences have

S.-J. Chen (✉)
Department of Physics, Department of Biochemistry, and UM
Informatics Institute, University of Missouri, Columbia, MO
65211, USA
e-mail: chenshi@missouri.edu

been used to weaken the gRNA-DNA duplex stabilities (at the off-target sites). The approach has been shown to cause a reduction in off-target cleavage by 5000-fold (Fu et al. 2014). Another notable approach is to destabilize the function structure of the CRISPR complex through Cas9 protein modification. For this purpose, researchers have developed several highly effective Cas9 mutants such as e high fidelity (SpCas9-HF1; Kleinstiver et al. 2016), enhanced specificity (eSpCas9(1.1); Slaymaker et al. 2016), and hyper-accurate (HypaCas9; Chen et al. 2017) Cas9 variants.

These highly successful designs for the Cas9 variants are based on the physical interactions in the Cas9/sgRNA/DNA complex. For example, the DNA cleavage requires opening up of the DNA duplex to form the R-loop, therefore, destabilizing the R-loop (at the off-target sites) can effectively reduce the cleavage efficiency. Physically, the R-loop structure is stabilized by DNA-Cas9/RNA interactions (mainly through the non-target DNA strand in the R-loop). Thus, Cas9 mutations can disrupt the Cas9/sgRNA-DNA R-loop interactions can significantly reduce the cleavage efficiency. For this purpose, Slaymaker et al. designed and engineered the Cas9 protein where the charge distribution in Cas9 is altered (Slaymaker et al. 2016). Because DNA is a highly charged polymer, changing the Cas9 charge distribution can immediately impact the electrostatic interaction and destabilize the stability of the DNA-Cas9/sgRNA complex.

Most previous methods for the destabilization of (off-target) CRISPR complex have focused on the Cas9 protein; in a recent study, Kocak et al. developed a highly innovative approach by focusing on the sgRNA (Kocak et al. 2019). The basic strategy was to add a hairpin structure to the 5'-end of the sgRNA spacer region. The purpose of adding the additional hairpin is to impede the R-loop formation through steric repulsion between the hairpin and the R-loop or the Cas9/sgRNA complex. Experimental tests suggested that such a modified sgRNA design can lead to a remarkable increase in gene editing specificity by several orders of magnitude.

In parallel to the experimental efforts above, tremendous efforts have been made on computational models for the selection of the optimal DNA targets and the corresponding sgRNAs with minimum off-target effect. Most of the existing computational models are based on various data processing algorithms such as deep learning. These approaches have shown encouraging results for specific systems. However, the accuracy for the

purely database-derived parameters often suffers from the quality of the mixed data source. For example, the data collected from the different cell types, species, delivery modality, and dosage may affect the training quality. Because the experimental condition can significantly influence the on- and off-target gene editing efficiencies, the reliability for extracting important features over mixed experimental data using blind database-based algorithms can be limited.

Realizing the limitations above, researchers begin to tackle the CRISPR gene editing problem from a physical point of view. For example, the recently reported VfoldCAS model (Xu et al. 2017) predicts and analyses CRISPR gene editing efficiency by computing the free-energy landscape for the full conformational ensemble of the sgRNA-DNA system. The model successfully quantifies the intrinsic correlation between and thermodynamic stability of the CRISPR system and the gene editing efficiency. More recently, uCRISPR (Zhang et al. 2019), another free-energy-based model, leads to much improved accuracy in predicting on- and off-target efficiencies. The uCRISPR model synthesizes multiple key steps into a unified framework. Specifically, unlike most other approaches, which focus on the sgRNA spacer-DNA target binding site, uCRISPR also accounts for the stability (the probability for the formation) of the functional tracrRNA structure, the sequence-dependent interactions between Cas9 and the non-target DNA strand, and the PAM sequence effect on Cas9/sgRNA-PAM binding affinity. Furthermore, because the model breaks the gene editing result down to elementary interaction terms, the parameters in the model are less likely data-dependent and are more transferable between the different experimental systems. Future development of the computational efforts requires more extensive database for the different experimental conditions including the different cell types and species. With the computational framework such as uCRISPR, the larger database would enable the development of a new generation of computational tools for precision CRISPR for any given genome systems.

Acknowledgements The author would like to thank Drs. Dongsheng Duan and Dong Zhang for many useful discussions.

Funding information This research was supported by National Institute of Health grants R01-GM063732 (S-J. C) and R01-GM117059 (S-J. C).

References

- Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, et al. Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*. 2017;550:407–10.
- Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR–Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 2014;32:279–84.
- Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol*. 2014;32:577–82.
- Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529:490–5.
- Kocak DD, et al. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat Biotechnol*. 2019;37(6):657–66.
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154:1380–9.
- Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, et al. Efficient genome modification by CRISPR–Cas9 nickase with minimal off-target effects. *Nat Methods*. 2014;11:399–402.
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science*. 2016;351:84–8.
- Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32:569–76.
- Wang DC, Wang X. Off-target genome editing: a new discipline of gene science and a new class of medicine. *Cell Biol Toxicol*. 2019;35(3):179–83.
- Xu X, Duan DS, Chen SJ. CRISPR–Cas9 cleavage efficiency correlates strongly with target–sgRNA folding stability: from physical mechanism to off-target assessment. *Sci Rep*. 2017;7(1):143.
- Zhang D, Hurst T, Duan D, Chen SJ. Unified energetics analysis unravels SpCas9 cleavage activity for optimal gRNA design. *Proc Natl Acad Sci U S A*. 2019;116(18):8693–8.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.