Utilizing Type II Restriction Endonucleases (EcoR1 and BamH1) in Complex with gRNA for Precise Gene Editing

By Matthew Platzman

AUTHOR BIO

Matthew Platzman is a driven high school student from New Jersey. Despite his passion for competitive soccer, his true passion lies in research. Matthew has researched at Brown University, where he integrated quantitative proteomics for protein profiling. Currently, he is working at the NEPS Lab at Tufts University's HNRCA, where he's researching heterogeneity in muscle mass 's with resistance training for sarcopenia. Matthew also worked at the Rutgers Waksman Institute, where he researched the potential of *landoltia punctata* for bioremediation and published 6 DNA sequences via NCBI. He gained valuable experience working in a human cadaver laboratory at the University of Connecticut. Matthew has proven to be a diligent and passionate researcher with an eye for detail. He's dedicated to making meaningful contributions to the scientific community and is driven to pursue a career in scientific research. He enjoys sports and spending time with friends and family in his free time.

ABSTRACT

CRISPR-Cas9 is one of the leading and most commonly employed gene editing processes. This technique uses guide RNA (gRNA) to recognize and excise specific DNA sequences. However, this process frequently elicits off-target effects within the genome. In consequence, the future of CRISPR-Cas9 as a therapeutic remains uncertain. This review aims to provide a new perspective on an alternative to the specificity, efficiency, and safety of genome editing by using type II restriction endonucleases, EcoR1 and BamH1, to recognize a specific DNA sequence (the recognition site) and cleave the targeted DNA adjacent to or within the recognition site. This article discusses the potential of utilizing intrinsic properties of these restriction endonucleases and repurposing them to prevent off-target mutations that may result in cancers. Preliminary research exhibits the specificity of using restriction endonucleases, however, it remains understudied in the field of cancer therapeutics. This review focuses on the potential to employ EcoR1 and BamH1 to prevent DNA sequence mutations during gene editing.

Keywords: EcoR1, BamH1, restriction endonuclease, enzymes, CRISPR-Cas9, gene editing, DNA, cancer, genetics, molecular biology.

INTRODUCTION

In 2020, the Nobel Prize for Chemistry was awarded to Emmanuelle Charpentier and Jennifer Doudna for their resounding discovery of CRISPR-Cas9 genome editing (Westermann et al., 2021). Cas9, an RNA-guided enzyme employed in the CRISPR-Cas9 gene editing process, is adapted from an innate genome editing system found in bacteria (Guo et al., 2022). The process, although complex and precise, can be simplified and visualized as 4 steps: (1) the Cas9 enzyme configures a complex with the gRNA in a given cell, (2) the Cas9-gRNA complex attaches itself to the corresponding DNA sequence of the gRNA adjoined to a spacer, (3) the complex then 'snips' the double-helix, (4) DNA sequences are inserted into the genome, replacing the excised segment. There are 3 different ways CRISPR-Cas9 can be used: disruption, deletion, or insertion (Xu et al., 2020). Disruption is used to induce base pair changes that prevent the expression of certain proteins and biomarkers, while deletions remove portions of the genome, and insertion replaces DNA segments. This process is operated by utilizing Cas9, an endonuclease. A restriction endonuclease is derived mainly from certain strains of bacteria, such as *Escherichia coli*, and has the capability to cleave DNA segments with higher specificity and sensitivity.

However, Cas9 is not a *restriction* endonuclease. Thus, Cas9 is prone to cause mutagenesis within off-target sites within the genome (Guo et al., 2023). This review article pinpoints why CRISPR-Cas9 elicits off-target effects and introduces alternative techniques utilizing type II restriction endonucleases EcoR1 and BamH1 to prevent unintended mutagenesis. Specific genes have been targets of gene editing techniques due to their tendency to cause carcinogenesis. The BRCA1 gene was originally discovered in 1994 and the BRCA2 gene in 1995 (Makhnoon et al., 2022). Both BRCA1 and BRCA2 are tumor suppressors, encoding proteins that regulate cell growth and proliferation. BRCA1 has been found to interact with multiple DNA repair and recombination proteins such as Rad51, the Rad50-MRE11-Nibrin complex, Bloom's helicase, and the Fanconi D2 protein (Murthy et al., 2019). The specific role for BRCA1 within transcriptional regulation and proliferation is exemplified through interactions with CtIP, ER (estrogen receptor), HDAC, Rb, p53, RNA polymerase II holoenzyme, cycle D1, and c-myc (Murthy et al., 2019). Estrogen-responsive organs, such as the ovaries and the breasts, are at major risk when it comes to BRCA1 mutations. 55%-72% of women who inherit a *BRCA1* variant and 45%-69% of women who inherit a *BRCA2* variant will develop breast cancer between the ages of 70-80 (National Cancer Institute, 2020). Carcinogenesis will only occur when both alleles of BRCA1 and BRCA2 are mutated (Godet et al., 2017). Since BRCA1 is located on chromosome 17 and BRCA2 on chromosome 13, it is possible for mutations to be inherited by both parents as both genes are autosomal (Petrucelli et al., 2016).

Methodology and Results of SpCas9-Mediated Off-Target Mutations in the vas-7280^{CRISPRh} Strain

In 2021, the Simoni Lab published their findings in the *Proceedings of the National Academy of Sciences* (*PNAS*), where they determined the rationale behind off-target mutations correlating to CRISPR-Cas9 editing (Garrood et al., 2021). The Simoni Lab used mosquitoes (*Anopheles gambiae*), which are vectors for human malaria (Garrood et al., 2021). They compared the propensity of off-target mutations in four different gene-drive

strains. A gene-drive causes certain strains to inherit genes more frequently than Mendelian genetics would allow for. In the first strain, they used a non-germline restricted promoter from Streptococcus pyogenes (SpCas9), a gRNA, and two recognition sites in the mosquito genome. Under these conditions, the frequency of off-target mutations reached 1.42%, which is considerably high at the microcellular level (Garrood et al., 2021). To identify off-target mutations of gRNA that targeted the gene *AGAP007280*, the Simoni Lab used CIRCLE-seq (Garrood et al., 2021), which is characterized by its circular structure for research of cleavage effects. Using the CIRCLE-seq process, the researchers discovered that there were 98 off-target effects apart from gene *AGAP007280* (Garrood et al., 2021).

They selected the top 15 off-target mutations with the highest CIRCLE-seq read counts and 5 additional sites to gauge whether the off-target mutations in the assay displayed evidence of insertion and/or deletion of nucleotides (Garrood et al., 2021). They then used targeted amplicon sequencing, which allows for the analysis of genetic variation, on pools of mosquitoes to search for the intended on-target sites to be cleaved (Garrood et al., 2021). 19 of the 20 total sites returned sequencing reads— the remaining site was excluded from the experiment due to the inability to be sequenced (Garrood et al., 2021). An additional site was removed from the experiment due to highly repetitive sequencing and polymorphisms (Garrood et al., 2021).

Within the on-target sites, there were many mutations in 3 strains targeting *AGAP007280* (Garrood et al., 2021). Using the strain that was predicted to cause the most sensitivity, vas-7280^{CRISPRh}, the Simoni Lab uncovered insertions and deletions (indels) above the threshold frequency at five distinct sites (off-2, -4, -6, -11, and -19) in all 5 sample generations (Garrood et al., 2021). Under these conditions, indel frequencies ranged from 0.03% to 1.42% of the total sequencing reads per off-target mutation site (Garrood et al., 2021). Four of the five sites were found in introns or exons, (*AGAP000774*, *AGAP011092*, *AGAP000061*, and *AGAP000042*), all of which had four or fewer discrepancies with the on-target sites (Garrood et al., 2021). Studies indicate that wild *Anopheles* mosquito populations show high degrees of polymorphisms within their genomes (Garrood et al., 2021). Several of the studied sites exhibited off-target mutations within the vas-7280^{CRISPRh} gene-drive population and had several allele variants; this impacted off-target cleavage through the formation of sites that resembled on-target sequences (Garrood et al., 2021).

The Simoni Lab tested the hypothesis that different innate polymorphisms can affect mutation frequencies at off-target sites (Garrood et al., 2021). They analyzed the relationship between indel frequencies and the presence of single nucleotide polymorphisms (SNPs) in sites made distinct by CIRCLE-seq (Garrood et al., 2021). The amplicon sequencing at the off-4 site displayed five total base pair mismatches with gRNA-7280 (Garrood et al., 2021). They analyzed the frequency of mutations for each individual allele and found that only the reference allele showed proof of cleavage, whereas no indels were found in the variant allele that had three mismatches (Garrood et al., 2021). The CRISPR-Cas9 process, although reportedly specific and sensitive, is actually faulty.

How EcoR1 can be Used in Complex with Programmed gRNA to Excise Mutated Sequences

EcoR1 is a type II restriction endonuclease isolated from *E. coli*. It cleaves double-stranded DNA into fragments with specificity and is a

part of the restriction modification system in prokaryotes, which defends against foreign DNA entry. EcoR1 binds and recognizes the sequence 5' GAATTC 3', cutting both DNA strands between the G and A nucleotides. EcoR1 is a homodimer of a 31 kilodalton subunit, having one globular domain being the *α/β* structure. By using time-resolved second harmonic (SH) generation spectroscopy, scientists were able to observe EcoR1 in real-time (Doughty et al., 2011). They discovered that the binding of EcoR1 to the recognition sequence of a 90 base pair (bp) DNA duplex attached to colloidal microparticles occurred rapidly (Doughty et al., 2011).

In the existence of cofactor Mg^{2+}, the researchers observed DNA cleavage, the detachment of EcoR1 from the DNA substrate, and the diffusion of the 74-bp fragment into the bulk solution that they ran, leaving the 16-bp fragment perfectly intact (Doughty et al., 2011). Since type II restriction enzymes cut at extremely specific recognition sites, enzymes such as EcoR1 can be utilized for precise cleavage in gene editing with gRNA. This gRNA will contain the palindromic complementary recognition sequence of 5' GAATTC 3' and be programmed to find the recognition site using targeted amplicon sequencing. Thus, when the gRNA-EcoR1 complex is inserted into the genome, it can effectively locate the target segment and attach itself adjacent to the recognition sequence. EcoR1 will begin cleaving adjacent to the site, excising the mutated base pairs.

How BamH1 Can complex with gRNA

BamH1, isolated from *Bacillus amyloli*, is another type II restriction endonuclease. It was found that BamH1 is remarkably similar in structure to EcoR1 in regards to their active sites (Newman et al., 1994). BamH1 recognizes short sequences of DNA, specifically the 6 bp sequence 5' GGATCC 3' and cleaves adjacent to the 5' guanine (Adedeji Olulana et al., 2022). The cleavage results in sticky ends that are 4 bp long (Arzanova et al., 2022). BamH1 is composed of a central β-sheet, which resides between α-helices (Adedeji Olulana et al., 2022). BamH1 undergoes a series of conformational changes upon recognition sequence identification, allowing the DNA to maintain its innate B-DNA conformation without distorting itself to ease enzyme binding (Adedeji Olulana et al., 2022). Akin to EcoR1, the potential of utilizing BamH1 in complex with gRNA is compelling. This gRNA will be able to recognize the mutated sequence in the genome through targeted amplicon sequencing and will contain the palindromic recognition sequence 5' GGATCC 3'. The BamH1-gRNA complex will bind adjacent to a mutation in the genome and cleave the DNA segment. The BamH1 restriction endonuclease provides the field of gene editing with an alternative method of genetic manipulation to treat mutations in genes such as BRCA1 and BRCA2 that are associated with diseases and cancers.

Conclusion

The field of oncology is like a puzzle. Scientists and researchers are slowly completing the puzzle piece by piece. Gene editing is one of the leading techniques in the realm of genetics and biology, and its potential to alter protein expression can be leveraged in the field of cancer therapeutics. The current standard in gene editing, CRISPR-Cas9, commonly produces off-target effects that can be detrimental. As discussed, the Simoni Lab tested CRISPR-Cas9 gene-drives to determine whether this method induced off-target mutations in *Anopheles* mosquitoes. They compared the propensity of off-target mutations among the gene-drives, concluding that the CRISPR-Cas9 process frequently induces off-target mutations.

This review discussed the potential alternative of using EcoR1 and/or BamH1 in a complex with gRNA to mitigate the off-target effects caused by the Cas9 enzyme, which can be crucial in treating mutation-derived diseases such as cancer. Annually, 287,850 new cases of invasive breast cancer will be diagnosed in women and about 43,250 will die from breast cancer (Arzanova et al., 2022). In addition, breast cancer diagnoses make up 1/3 of all cancer cases in women each year in the United States (Breast Cancer Statistics, n.d.). The utilization potential of these two enzymes is limitless. Not only do the EcoR1 and BamH1 complexes provide higher efficiency and specificity in gene editing for cancer therapeutics, but this alternative method can certainly revolutionize the treatment of other mutation-derived diseases such as sickle cell anemia, Tay-Sachs, cystic fibrosis, and many more.

This efficient alternative method will shorten research and bench-to-bedside timelines since there will be much fewer unintended effects to correct. This alternative method is particularly exciting knowing that there are over 3000 type II restriction endonucleases. One can imagine using any recognition site of choice to direct cleavage for gene editing. Type II restriction endonucleases, such as EcoR1 and BamH1, can facilitate the advancement of research in genetics and all therapeutic fields.

REFERENCES

Westermann, L., Neubauer, B., & Köttgen, M. (2021). Nobel Prize 2020 in Chemistry honors CRISPR: a tool for rewriting the code of life. *Pflugers Archiv : European journal of physiology*, *473*(1), 1–2. https://doi.org/10.1007/s00424-020-02497-9

Guo, N., Liu, J. B., Li, W., Ma, Y. S., & Fu, D. (2022). The power and the promise of CRISPR/Cas9 genome editing for clinical application with gene therapy. *Journal of advanced research*, *40*, 135–152. https://doi.org/10.1016/j.jare.2021.11.018

Xu, Y., & Li, Z. (2020). CRISPR-Cas systems: Overview, innovations and applications in human disease research and gene therapy. *Computational and structural biotechnology journal*, *18*, 2401–2415. https://doi.org/10.1016/j.csbj.2020.08.031

Guo, C., Ma, X., Gao, F., & Guo, Y. (2023). Off-target effects in CRISPR/Cas9 gene editing. *Frontiers in bioengineering and biotechnology*, *11*, 1143157. https://doi.org/10.3389/fbioe.2023.1143157

Makhnoon, S., Arun, B., & Bedrosian, I. (2022). Helping Patients Understand and Cope with BRCA Mutations. *Current oncology reports*, *24*(6), 733–740. https://doi.org/10.1007/s11912-022-01254-8

Murthy, P., & Muggia, F. (2019). Women's cancers: how the discovery of BRCA genes is driving current concepts of cancer biology and therapeutics. *Ecancermedicalscience*, *13*, 904. https://doi.org/10.3332/ecancer.2019.904

National Cancer Institute. (2020, November 19). *BRCA Mutations: Cancer Risk & Genetic Testing*. National Cancer Institute. https://www.cancer.gov/about-cancer/causes-prevention/genetics/brca-fact-sheet

Godet, I., & Gilkes, D. M. (2017). BRCA1 and BRCA2 mutations and treatment strategies for breast cancer. *Integrative cancer science and therapeutics*, *4*(1), 10.15761/ICST.1000228. https://doi.org/10.15761/ICST.1000228

Petrucelli, N., Daly, M. B., & Tuya Pal. (2016, December 15). *BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer*. Nih.gov; University of Washington, Seattle. https://www.ncbi.nlm.nih.gov/books/NBK1247/

Garrood, W. T., Kranjc, N., Petri, K., Kim, D. Y., Guo, J. A., Hammond, A. M., Morianou, I., Pattanayak, V., Joung, J. K., Crisanti, A., & Simoni, A. (2021). Analysis of off-target effects in CRISPR-based gene drives in the human malaria

mosquito. *Proceedings of the National Academy of Sciences*, *118*(22), e2004838117. https://doi.org/10.1073/pnas.2004838117

Doughty, B., Kazer, S. W., & Eisenthal, K. B. (2011). Binding and cleavage of DNA with the restriction enzyme EcoR1 using time-resolved second harmonic generation. *Proceedings of the National Academy of Sciences*, *108*(50), 19979–19984. https://doi.org/10.1073/pnas.1115498108

Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., & Aggarwal, A. K. (1994). Structure of restriction endonuclease BamHI and its relationship to EcoRI. *Nature*, *368*(6472), 660–664. https://doi.org/10.1038/368660a0

Adedeji Olulana, A. F., Choi, D., Inverso, V., Redhu, S. K., Vidonis, M., Crevatin, L., Nicholson, A. W., & Castronovo, M. (2022). Noncanonical DNA Cleavage by BamHI Endonuclease in Laterally Confined DNA Monolayers Is a Step Function of DNA Density and Sequence. *Molecules (Basel, Switzerland)*, *27*(16), 5262. https://doi.org/10.3390/molecules27165262
Arzanova, E., & Mayrovitz, H. N. (2022). *The Epidemiology of Breast Cancer* (H. N. Mayrovitz, Ed.). PubMed; Exon Publications. https://www.ncbi.nlm.nih.gov/books/NBK583819/

Breast Cancer Statistics | How Common Is Breast Cancer? (n.d.). Www.cancer.org. https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html#:~:text=Breast%20cancer%20is%20the%20most

Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., & Aggarwal, A. K. (1994). Structure of restriction endonuclease bamhi phased at 1.95 A resolution by MAD analysis. *Structure (London, England : 1993)*, *2*(5), 439–452. https://doi.org/10.1016/s0969-2126(00)00045-9

Restriction enzyme analysis O F eds and Celo virus ... - semantic scholar. Semantic Scholar. (2007). https://www.semanticscholar.org/paper/Restriction-enzyme-analysis-o-f-EDS-and-CELO-virus/fe956b299c0cbd76ca86dbb52434490666b14b0d

Bairwa, A., Venkatasalam, E. P., Subhash, S., & Dipta, B. (2022). *Management of potato cyst nematodes (Globodera spp.) using biotechnological approaches*. SpringerLink. https://link.springer.com/chapter/10.1007/978-981-19-0264-2_12