

Electronic Thesis and Dissertation Repository

10-11-2023 2:00 PM

Kinetic Analyses of SaCas9[D10E] in vitro

Claire Zhang, *Western University*

Supervisor: Edgell, David, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biochemistry

© Claire Zhang 2023

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Biochemistry Commons](#)

Recommended Citation

Zhang, Claire, "Kinetic Analyses of SaCas9[D10E] in vitro" (2023). *Electronic Thesis and Dissertation Repository*. 9816.

<https://ir.lib.uwo.ca/etd/9816>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

CRISPR is a well-known adaptive defense mechanism that gained attention through its ability to be easily reprogrammable. Over time, the native CRISPR-Cas system evolved to tolerate mismatches, broadening its cleavage preferences; however, this development poses concerns with off-target cleavage in gene editing. Here, we introduced a D10E mutation into SaCas9 to potentially reduce off-target cleavage. To characterize SaCas9[D10E] and its tolerance for mutations, we designed 21 different substrates that each contained a single transversion in a therapeutically relevant EMX1-1 gene. Through a series of *in vitro* cleavage assays, SaCas9[WT] and SaCas9[D10E] were compared across these substrates. A kinetic analysis of SaCas9[D10E] revealed trends in initial cleavage rates as PAM proximal mutations exhibited reduced cleavage activity and PAM distal mutations displayed enhanced activity. Furthermore, the ability of SaCas9[D10E] was highlighted through a competition assay that displayed discrimination between single nucleotide differences.

Keywords

CRISPR, gene editing, enzyme kinetics, *in vitro* cleavage assay

Lay Summary

CRISPR is an adaptive defense mechanism that originated in bacteria and is currently utilized as a gene editing tool. In bacteria, when a virus invades the cell, CRISPR functions via guided targeted protein-RNA complex that generates a double stranded break in inserted viral DNA to destroy it. Since then, CRISPR has evolved to tolerate mutations to broaden its cleavage specificity. While broadening its cleavage specificity is good for bacteria, it results in off-target cleavage, which reduces its gene editing capabilities. The focus of my thesis is to find a solution to its tolerance and improve its accuracy in gene editing.

There are many pathways that we can take to reduce off-target cleavage, but, here, we chose to generate a conservative mutation of D10E in Cas9 to alter cleavage specificity and hypothesized that it can potentially reduce off-target cleavage.

This thesis established that the SaCas9[D10E] variant is unable to reduce off-target cleavage. Rather, D10E can discriminate between single mutations of on and off targets while also highlighting trends in cleavage rates of SaCas9[D10E].

Co-Authorship Statement

Dr. David Edgell and Specific Biologics Incorporated (SBI) conceived the initial *in vitro* experiments as well as generation of SaCas9[D10E].

All experiments were done by Claire Zhang with the exception of the competition assay in which Olha Haydaychuk worked in parallel, each generating data for D10E and WT, respectively.

Acknowledgments

There are so many people that I would like to thank for my graduate experience. I would like to, first, thank Dr. David Edgell for giving me the opportunity to be in his lab and take part in his research. I'm very thankful to have been a part of it all and for all your mentorship. I would also like to thank all the members of the Edgell Lab as well as my advisory committee, Dr. Murray Junop and Dr. Brian Shilton, for their support and guidance throughout my masters.

Mahsa Farmanbar and Olha Haydaychuk deserve a big thank you for being a huge part of my life over the past two years. We started together and now we're finishing our degree together and I'm going to miss you guys. I would not have been able to survive grad school without the both of you.

To Omar Hassan, I am so lucky that you were so supportive of me throughout this entire process whether it was late night phone calls when I was stuck in the lab or bringing me food when I forgot to eat that day. Through all the tears and ice cream, you've kept me sane and pushed me to be motivated every single day and I really can't thank you enough. I love you.

Thank you to my parents who have supported me throughout my entire life and education. I know that you've put so much of your effort to get me to this point, and I appreciate all your sacrifice.

Finally, Tuna and Miso, I dedicate this thesis to you guys for all the days you've waited for me to get home.



List of Tables

Table 1. DNA sequence of EMX1-1 off-targets	23
Table 2. Kinetic analyses of SaCas9[WT] and SaCas9[D10E] against on and off targets	25
Table 3. Golden Gate primers.....	48
Table 4. Primers used for Golden Gate mutagenesis for RuvC active site variants	48

List of Figures

Figure 1. Schematic of two common double-stranded break repair mechanisms: non-homologous end joining (NHEJ) and homologous recombination (HR)	2
Figure 2. Illustration of different gene editing enzymes. Image generated in BioRender.....	4
Figure 3. The three steps that CRISPR-Cas9 uses for its adaptive defense mechanisms: adaptation, biogenesis, and interference.....	6
Figure 4. SaCas9 crystal structure. Image retrieved from PDB: 5CZZ.	8
Figure 5. Close-up of the SaCas9 RuvC active site with catalytic residues of D10, E477, H701, and D704.	10
Figure 6. Conformational structure of SpCas9 as apo (PDB: 4CMQ), RNA bound (PDB: 4ZT0), RNA-DNA bound (PDB: 4UN3), and primed active SpCas9/precatalytic-SpCas9 (PDB: 5F9R).	11
Figure 7. Schematic of divalent metal ion cleavage in the RuvC domain active site.	13
Figure 8. Diagram of an in vitro cleavage time course assay and the results that come about using linear or supercoiled DNA.	17
Figure 9. Example of an in vitro time course cleavage assay with EMX1-1 supercoiled substrate performed with SaCas9[WT] and SaCas9[D10E].....	18
Figure 10. Schematic of generated substrates with single transversions.	20
Figure 11. Initial in vitro cleavage assay from Specific Biologics Incorporated (SBI) with SaCas9[WT] and SaCas9[D10E] against supercoiled DNA (target unknown)	21
Figure 12. In vitro cleavage assay of RuvC active site variants.....	22
Figure 13. Example of an in vitro cleavage assay of 5 nM on- and off-target linear substrates (Tv 3, 4, 10, 15, and 21) against SaCas9[WT] and SaCas9[D10E]..	24
Figure 14. Michaelis-Menten curves of SaCas9[WT] and SaCas9[D10E].....	26
Figure 15. Schematic of EMX1 on-target.	28
Figure 16. Analysis of on- and off-target EMX1-1 supercoiled DNA against WT and D10E.....	32
Figure 17. Visual representation of k_1 and k_2 rates.	33
Figure 18. In vitro cleavage assay of SaCas9[WT] and SaCas9[D10E] against substrates at all positions.	33
Figure 19. In vitro cleavage assay of SaCas9[WT] and SaCas9[D10E] against substrates at positions all positions.	34
Figure 20. In vitro cleavage assay of SaCas9[D10E]/[H557A].....	35
Figure 21. Comparison of on and off targets via a competition assay with WT and D10E variants.	37
Figure 22. Analysis of competition assay.	38
Figure 23. Protein Purification on a 8% SDS-PAGE gel.	49

List of Abbreviations

BH: bridge helix

Bps: base pairs

EtBr: ethidium bromide

D10E: SaCas9[D10E]

DEHA: SaCas9[D10E]/[H557A]

DSB: double-stranded break

Nts: nucleotides

PI: PAM-interacting

WT: wild-type or SaCas9[WT]

ZFN: zinc finger nucleases

ZFP: zinc finger protein

Tv/Transversions: reference to generated off-target substrates. Number pairing (Tv 4 represents the position of the single transversion)

Table of Contents

Abstract.....	i
Lay Summary.....	ii
Co-Authorship Statement.....	iii
Acknowledgments.....	iv
List of Tables	v
List of Figures.....	vi
List of Abbreviations	vii
1 « Introduction ».....	1
2 « Materials and Methods ».....	15
2.1 Purification of SaCas9 variants.....	15
2.2 Active Site Mutant Cloning: Golden Mutagenesis	15
2.3 Production of sgRNA.....	16
2.4 In Vitro Time Course Cleavage Assays with Linear Substrate	17
2.5 In Vitro Time Course Cleavage Assays with Supercoiled Substrate.....	18
2.6 Competition Assay.....	19
2.7 Statistical Analysis.....	19
3 « Results».....	20
3.1 Target Substrates.....	20
3.2 SaCas9[D10E] has the highest <i>in vitro</i> activity among RuvC active site variants	21
3.3 Rate of SaCas9[D10E] discrimination between targets with single nucleotide substitutions	23
3.4 The initial nicking rate is dependent on the position of single transversions	27
3.5 The first nicking step is performed by the HNH domain.....	35
3.6 SaCas9[D10E] can discriminate between on and off-targets	36

4 « Discussion» 39

4.1 Kinetic Comparisons Between Single Mutations 39

4.2 Independent nicking rates demonstrate discrimination with SaCas9[D10E] 40

4.3 The initial nicking step is coordinated through the conformational change of HNH
..... 41

4.4 Future Work 43

4.4.1 Transition vs Transversion Comparisons43

4.4.2 Mammalian Work43

4.5 Conclusions 44

References 45

Appendix 48

Curriculum Vitae 50

1 « Introduction »

Chapter 1 – Introduction

1.1 – Gene Editing

1.1.1 – Brief History

In 1927, while investigating the characteristics of fruit flies, Herman Müller discovered that the frequency of random mutations increases when exposed to x-rays¹. This discovery kickstarted the field of gene editing, documenting for the first time how genes can undergo change.

The first directed genetic modification occurred in yeast and mammalian cells in the 1970s and 80s^{2,3}, performed by inducing DNA damage and relying on the cell's endogenous DNA repair mechanisms to repair said damage. The two most common repair mechanisms are non-homologous end joining (NHEJ) and homologous recombination (HR), with the latter being considered more favourable for producing accurate and predictable repair outcomes²⁻⁴.

Current methods for inducing DNA damage include zinc finger nucleases (ZFNs)^{5,6}, transcription activator-like effector nucleases (TALENs)⁷, and the clustered regularly interspaced palindromic repeats (CRISPR) system⁸⁻¹⁰. One feature shared amongst these methods is the ability to bind DNA and induce a double-stranded break through a nuclease domain.

1.1.2 – Repair Mechanisms

The first precise nucleases functioned by generating a double-stranded break (DSB) and taking advantage of the cell's natural repair mechanisms. Non-homologous end-joining (NHEJ) and homologous recombination (HR) are considered the two most well-understood repair pathways (Fig. 1).

NHEJ is considered the dominant DSB repair pathway ¹¹, relying on ligation between DNA ends that lack homology. This tends to result in low fidelity repairs, frequently resulting in deletions or unpredictable products ¹².

In comparison, HR requires an additional copy of the genome harbouring a DSB ¹³. This results in much more processing compared to the NHEJ repair pathway, resulting in higher fidelity repairs. As a result, most uses of gene editing depend on HR repair, due to the accurate incorporation of exogenous DNA.

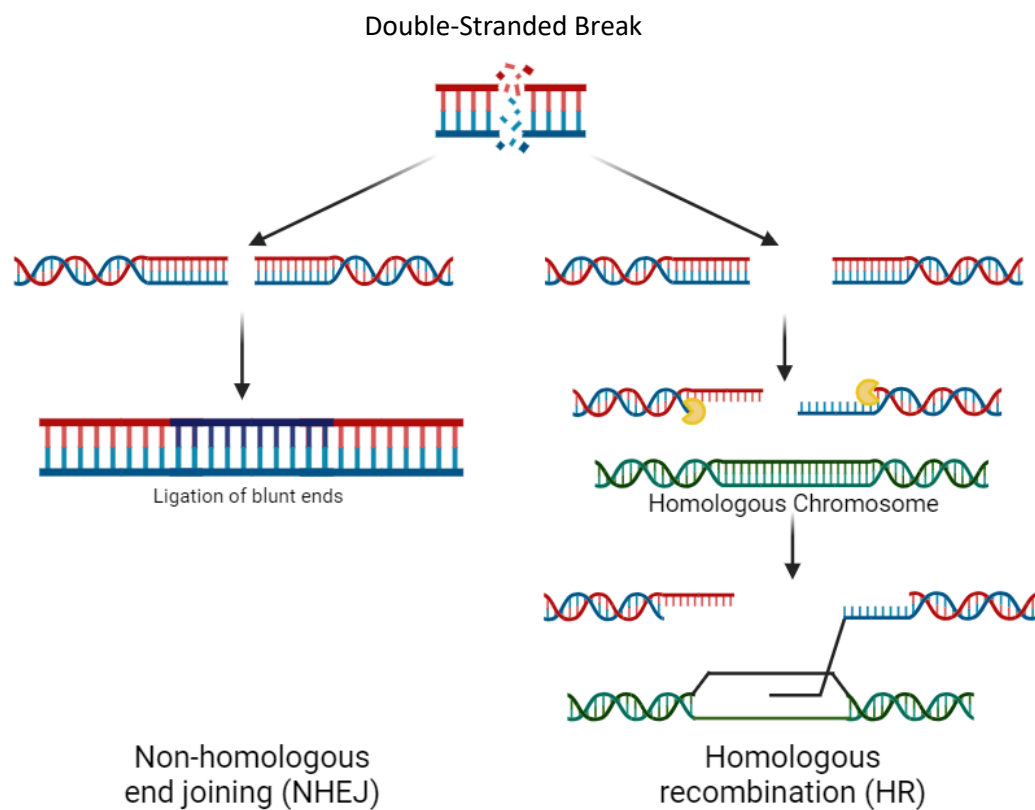


Figure 1. Schematic of two common double-stranded break repair mechanisms: non-homologous end joining (NHEJ) and homologous recombination (HR). Image created through Biorender.

1.1.3 – Zinc Finger Nucleases (ZFNs)

Zinc finger nucleases (ZFNs) were one of the first precision nucleases. ZFNs function as a chimera of zinc finger proteins (ZFP) and a nonspecific nuclease domain that is responsible for cutting targeted DNA sites (FokI restriction endonuclease) (Fig. 2.)^{5,14}. The ZFPs consists of three Cys₂His₂zinc fingers that each bind a zinc(II) ion to form the structural basis of the protein. The primary advantage of using ZFPs was in its ability to recognize nucleotide triplets, and specific arrangements of multiple ZFPs allowed scientists to target DNA sequences of interest. Gene editing occurred when exogenous DNA was introduced to the cell after inducing a DSB. The exogenous DNA was subsequently inserted into the cut site through homology directed repair (HDR).

1.1.4 – Transcription activator-like effector nucleases (TALENs)

Transcription activator-like effector nucleases (TALENs) are another engineered nuclease that consists of discrete DNA binding and nuclease domains. The DNA binding domain involves transcription activator-like effector (TALE) proteins that are found in the plant bacterial genus *Xanthomonas*⁷. These proteins are transcription factors that bind to promotor sequences consisting of tandem repeats (~34 amino acid sequence) and a repeat variable diresidue (RVD) at residues 12 and 13, which aids in DNA specificity^{15,16}. Like the ZFNs, TALEs were fused to the FokI endonuclease, producing a chimera that induces DSBs at specific DNA sequences. Similarly to ZFNs, TALE proteins can be combined to produce TALENs that recognize different promoter sequences. The resulting effects and steps to gene editing are the same, the nuclease domain generates a DSB, and exogenous DNA is integrated into a cell's genome at the cut site via HR (Fig. 2).

While ZFNs and TALENs are valuable tools, neither are without their limitations. ZFNs struggle to recognize a larger subset of DNA substrates due to the difficulty of modifying protein-DNA contacts. In contrast, TALENs are known to be large compared to other precision nucleases, which can lead to complications regarding delivery of the nuclease and genetic payload¹⁷.

1.1.5 – Meganucleases

Meganucleases are a more recent introduction to the gene editing field. These are large nucleases (~44 – 56 kDa) that are subsequently highly specific, as they can recognize large DNA sequences (~12-40 base pairs (bps))¹⁸. Meganucleases are a highly diverse family of proteins, and can be found in bacteria, archaea, and in plants.

Unlike ZFNs and TALENs, meganucleases are capable of both recognizing and cleaving DNA within the same domain. As previously mentioned, exogenous DNA is incorporated into a cell's genome by stimulating the cell's repair machinery (specifically HR repair¹⁹) (Fig. 2). However, meganucleases are limited to a single cognate target sequence, greatly limiting its ability to target other DNA sequences. Like ZFNs and TALENs, meganucleases recognize DNA through protein-DNA mediated contacts which results in difficulties with engineering meganucleases to target other DNA substrates. Furthermore, the large size of meganucleases makes delivery more difficult than other alternatives.

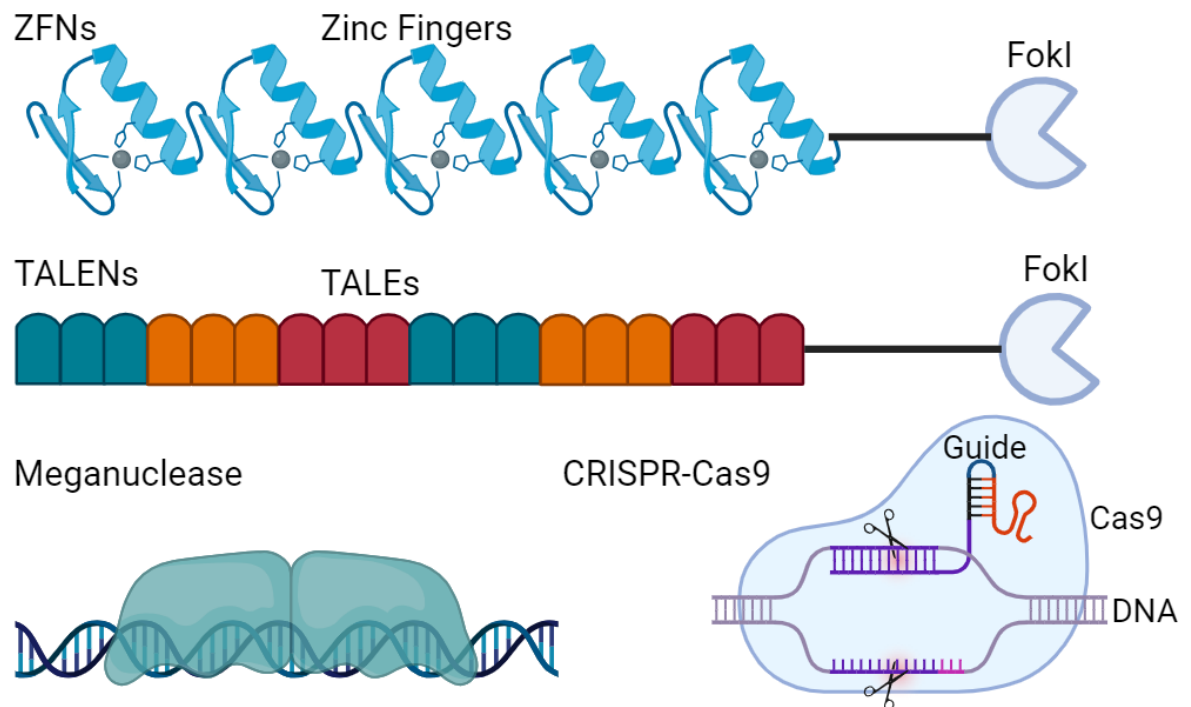


Figure 2. Illustration of different gene editing enzymes. Image generated in BioRender.

1.2 – CRISPR Overview

CRISPR is an adaptive defense mechanism discovered in prokaryotes that function by integrating foreign DNA ²⁰. Type II CRISPR-CRISPR associated protein (Cas) is a well-characterized and common type of CRISPR system used in gene editing. In prokaryotes, this adaptive mechanism revolves around three major steps: adaptation, biogenesis, and interference (Fig. 3)²¹.

In the adaptation phase, a protospacer adjacent motif (PAM) sequence is required for Cas9 to recognize a target protospacer for integration. Once a protospacer is targeted, Cas9 recruits Cas1 and Cas2 to integrate the target sequence into its CRISPR array ²². During biogenesis, a long precursor CRISPR-RNA (pre-crRNA) is processed with a trans-activating CRISPR RNA (tracrRNA), which forms an RNA duplex with the pre-crRNA. RNase III recognizes the RNA duplex and further processes it into a mature crRNA ²³. Finally, during interference when foreign DNA invades the cell, the crRNA and tracrRNA duplex guides Cas9 to introduce a double-strand break (DSB) and degradation the exogenous DNA molecule ²⁴.

To properly function, the CRISPR system must form a ribonucleoprotein (RNP), consisting of Cas9, a crRNA, and a tracrRNA ²³. The crRNA and tracrRNA first hybridizes, followed by binding to Cas9, which then, undergoes a conformational change through 3'-stem-loops ²⁵. The RNP complex recognizes a PAM sequence and binds to a target site, it displaces a DNA strand and results in the formation of an R-loop structure ²⁶. The HNH and RuvC nuclease domains in Cas9 nicks the top and bottom strand in sequence, resulting in DSB at the target sequence ²⁶. Following DNA cleavage, the cut site is usually repaired using non-homologous end joining (NHEJ) and, insertions or deletions can be incorporated at the cut site resulting in DNA manipulation.

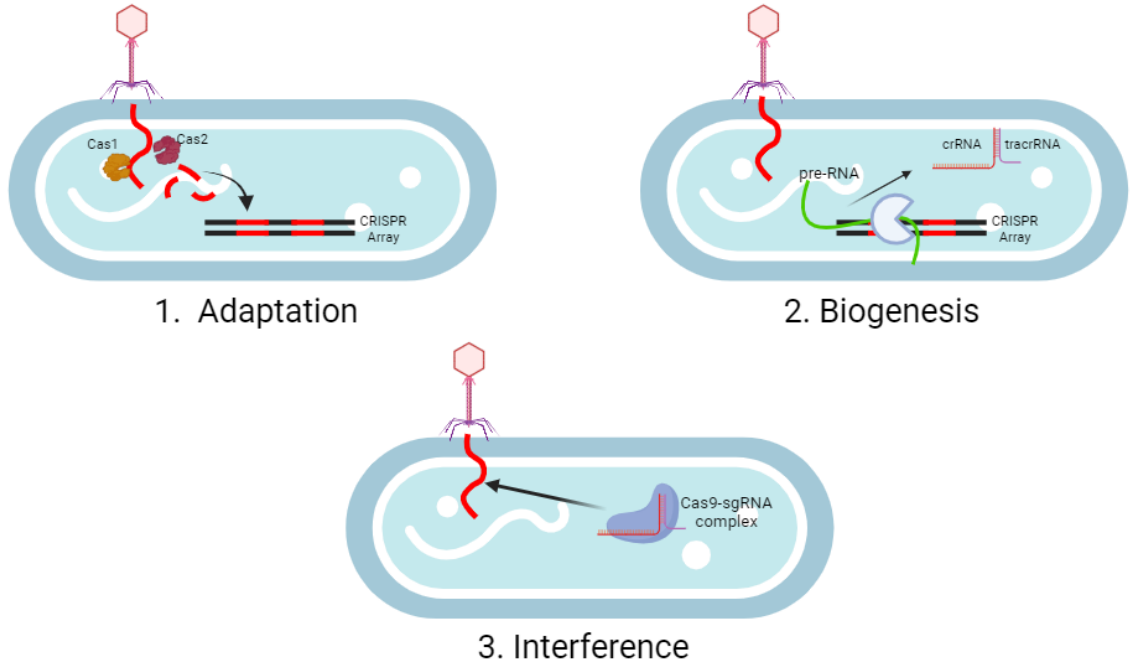


Figure 3. The three steps that CRISPR-Cas9 uses for its adaptive defense mechanisms: adaptation, biogenesis, and interference.

1.2.1 – *Streptococcus pyogenes* Cas9

Since the discovery of CRISPR, scientists primarily focused on the well-characterized *Streptococcus pyogenes* Cas9 (SpCas9). Unlike other endonucleases, SpCas9 can generate DSBs through RNA:DNA mediated contacts. However, it requires a PAM sequence (5' – NGG – 3') adjacent to the ~20-bp target sequence. SpCas9 uses two nuclease domains to generate DSBs: the HNH and RuvC domains, which cut the target and non-target strand, respectively. SpCas9 is frequently used in research since it is small (at 1368 amino acids long)²⁷, which more easily facilitates delivery. This protein follows single turnover enzyme kinetics (in which a single enzyme molecule converts a single molecule of substrate into a product), as it remains bound to DNA after cleavage²⁸.

1.2.2 – *Staphylococcus aureus* Cas9

Staphylococcus aureus Cas9 (SaCas9) is a structural and functional homolog of SpCas9. It is slightly smaller in size (1053 amino acids) and requires a lengthier PAM sequence (5' – NNGRRT – 3') (Fig.4)²⁹. A primary advantage of using SaCas9 over SpCas9 is that it follows multi-turnover enzyme kinetics³⁰. Due to its size and efficacy advantages, SaCas9 is considered more promising in potential gene editing over the more commonly used SpCas9.

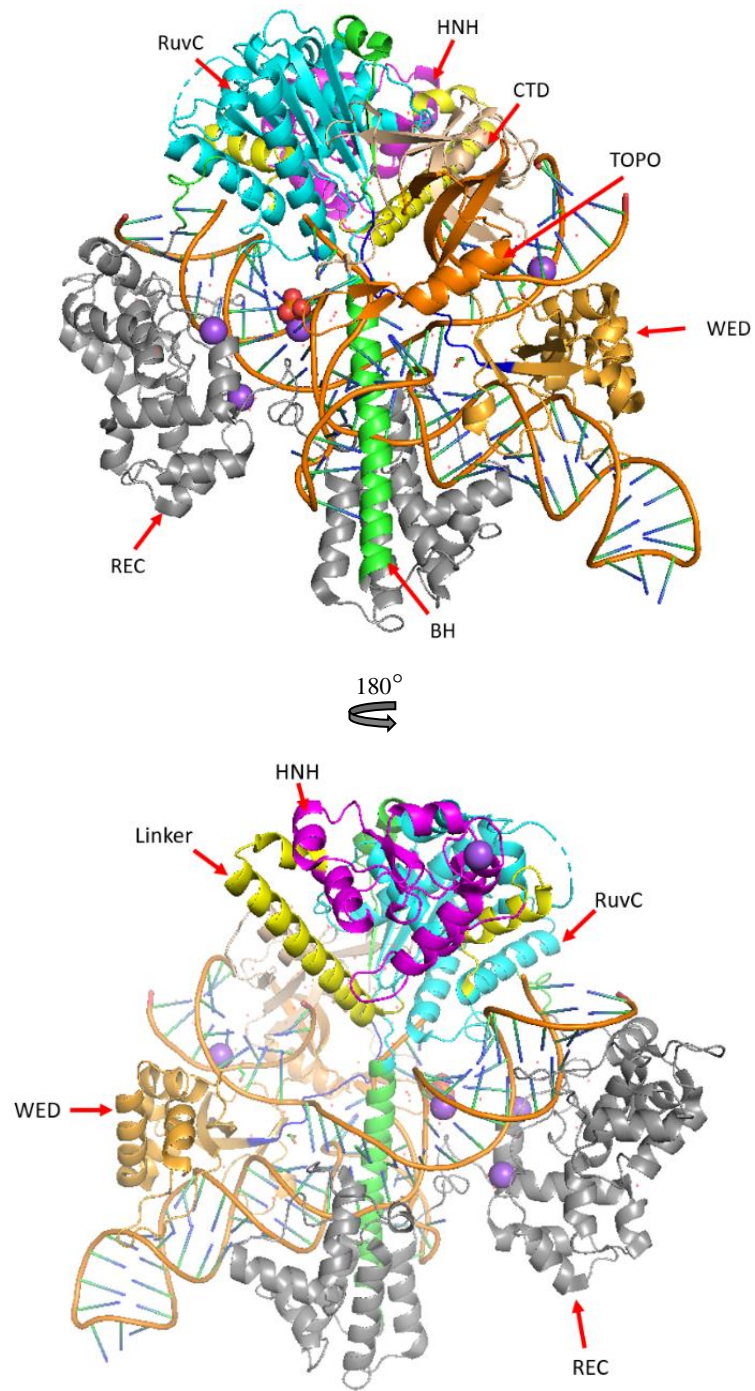


Figure 4. SaCas9 crystal structure with labelled domains. Image retrieved from PDB: 5CZZ.

1.2.2.1 – SaCas9 Domains

1.2.2.1.1 – Nuclease Lobe (NUC)

SaCas9 consists of two lobes, the recognition lobe (REC) and the nuclease lobe (NUC) that are connected through a bridge helix (BH; residues 41-73) and a linker loop (residues 426-434) (Fig. 4)²⁹.

The nuclease lobe (NUC) is the second lobe of SaCas9, consisting of the HNH domain, the RuvC domain, wedge domain (WED), and the PAM-interacting domain (PI). The NUC lobes function to cleave by sequentially nicking the target then the non-target strand³¹.

The HNH domain, one of the two nuclease domains, contains a $\beta\beta\alpha$ -metal fold that cleave the target strand following a single metal ion cleavage mechanism^{27,29}. For complete Cas9 cleavage, studies have shown that the HNH domain must undergo a conformational change to its active formation³².

The RuvC domain is the second nuclease domain, sharing similar structure to an RNase H fold²⁷. RuvC cleaves the non-target strand (including the PAM sequence) following a divalent metal ion cleaving mechanism. It contains 4 active site residues that coordinate the movement of the divalent metal ion: Asp10, Glu477, His701, and Asp704 (Fig. 5). Asp10 (D10) is critical in stabilizing the transition state of the cleavage reaction mutations to this residue renders RuvC non-functional^{33,34}.

The wedge domain is relatively uncharacterized²⁹. The PI domain facilitates the unwinding of DNA and the formation of the heteroduplex.

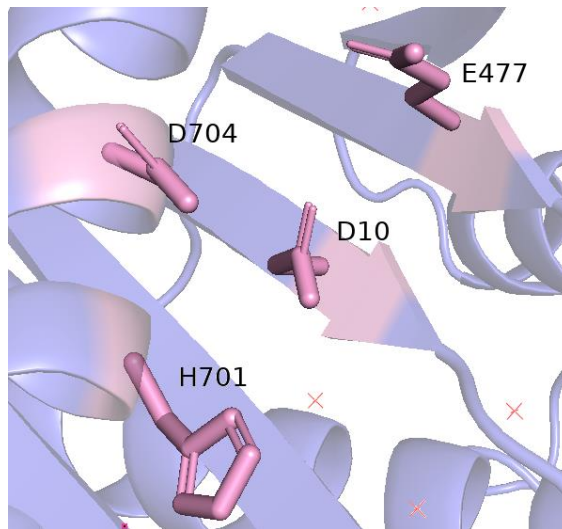


Figure 5. Close-up of the SaCas9 RuvC active site with catalytic residues of D10, E477, H701, and D704 without Mg^{2+} . Image retrieved from PDB: 5CZZ file.

1.2.2.1.2 – Recognition Lobe (REC)

The REC lobe (spans residues 41-425) has three main functions: the recognition of nucleic acids, the coordination of HNH conformational changes, and the locking of the HNH domain in the active state ³⁵.

There are currently believed to be three different portions of the REC lobe: REC1, REC2, and REC3. Prior to DNA cleavage, REC3 binds to the RNA-DNA duplex and activates the HNH nuclease domain by re-orientating REC2, which initiates conformational changes in Cas9 ^{35,36}. After this recognition state, REC3 separates from heteroduplex and REC1 and REC2 develop salt bridge interactions with the HNH domain, beginning the transition stage. Finally, REC2 moves parallel to REC3 (opposite of HNH) and REC1 anchors the HNH domain in its active state, stabilized through ionic interactions ³⁵.

REC3 is also associated with the proofreading. There are 5 conserved residue clusters associated with proofreading – 4 in REC3 and 1 in the HNH-RuvC linker 2 (L2) ³⁶. While details are still unclear, there is thought to be a conformational checkpoint during an intermediate cleavage state ³⁴. At this point, Cas9 will dissociate from its target if it recognizes a mismatch.

1.2.2.2 – Structure Conformations of SaCas9

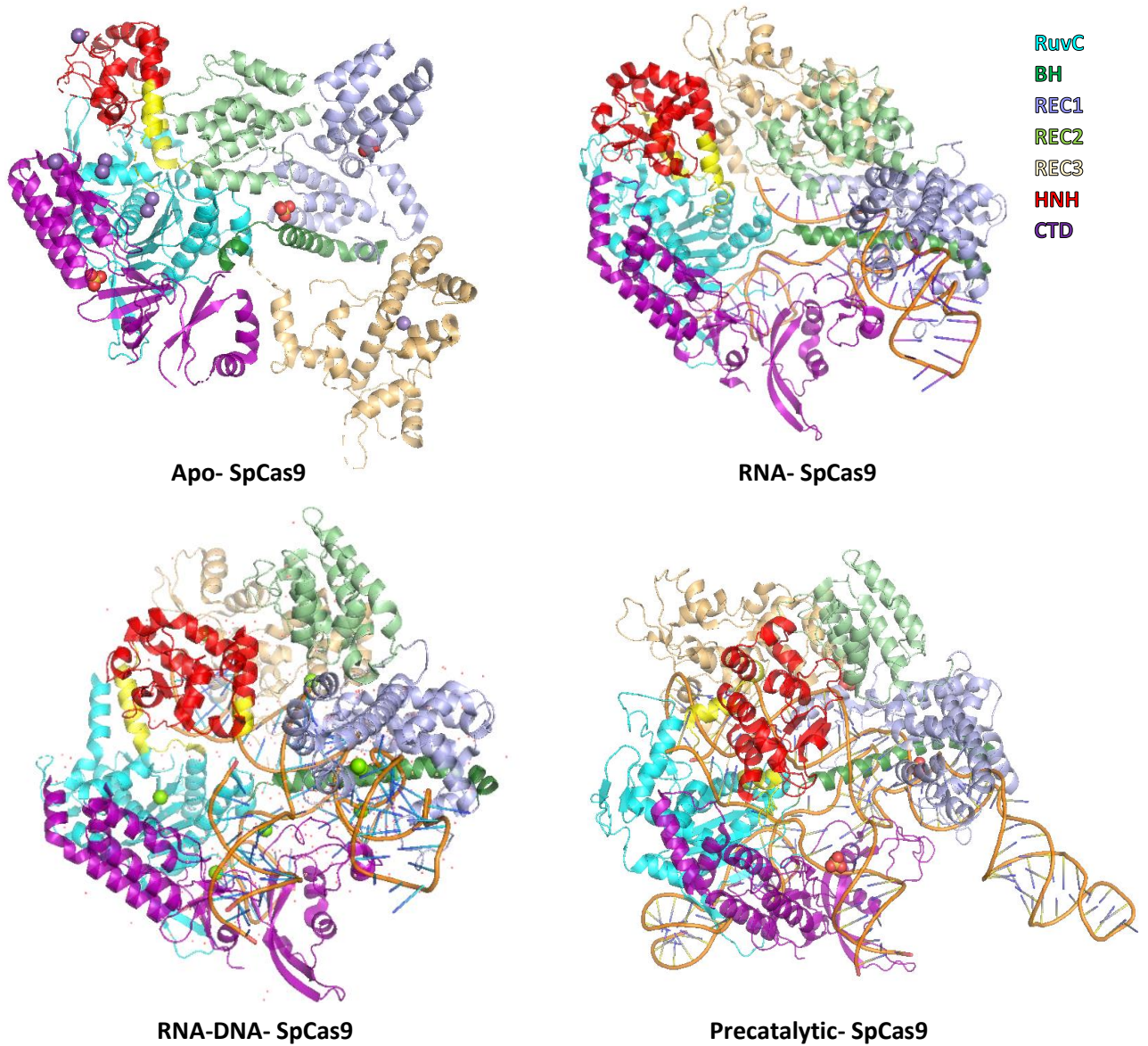


Figure 6. Conformational structure of SpCas9 as apo (PDB: 4CMQ), RNA bound (PDB: 4ZT0), RNA-DNA bound (PDB: 4UN3), and primed active SpCas9/precatalytic-SpCas9 (PDB: 5F9R). Both SpCas9 and SaCas9 have similar conformational changes. These four crystal structures represent open (free channel for DNA to pass through) and closed (does not allow DNA to pass) states for SaCas9 as well.

There are two major Cas9 conformational states: open or closed. In the apo state, Cas9 maintains a closed conformation in the absence of nucleic acids. In this state, the HNH domain is unordered and distant from the binding cleft/catalytic site ²⁵(Fig. 6). In the presence of RNA or the RNA:DNA heteroduplex, Cas9 begins to move by rotating the HNH domain towards the catalytic center and creates a central channel for the heteroduplex. The conformation is in a similar open state in the pre-catalytic state (primed for cleavage). The conformations of both SpCas9 and SaCas9 are structurally similar, with slight differences in the WED, PI, and the REC domains.

1.2.2.3 – DNA Binding and Targeting

Initial binding of Cas9 requires the presence of a protospacer adjacent motif (PAM), a recognition sequence recognized by an RNA guide. In SpCas9, Arg1333 and 1335 in the PAM interacting domain (PI) bind to the major groove of the guanine bases in the PAM (5' – NGG – 3' in this case) ³⁷. Meanwhile, Lys1107 and Ser1109 (PI domain) bind to the minor groove and create a phosphate lock loop. This movement stabilizes the first few bases of the target DNA ^{24,27,37}.

Complementarity of the guide RNA strand to the target DNA strand (leading to an R-loop formation²⁵) will allow the DNA to unzip. Consequently, the DNA can bind to RNA, leaving Cas9 to cleave between the third and fourth nucleotides from the PAM ^{27,37,38}.

1.2.2.4 – Catalytic Mechanism of SaCas9

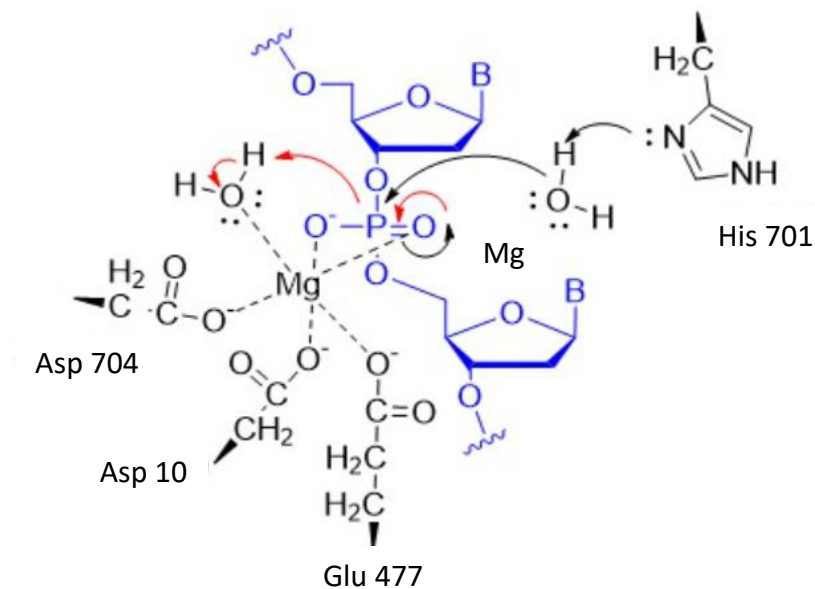


Figure 7. Schematic of divalent metal ion cleavage in the RuvC domain active site. Two divalent metal ion cleavage mechanism is facilitated by D10, E477, H701, and D704 using a Mg²⁺ ion. Image edited from Tuft.

After recognition and binding, DSBs are introduced by cleavage by both the HNH and RuvC nuclease domains, relying on either a one or two divalent metal ion mechanism, respectively.

The HNH domain utilizes a histidine to deprotonate a water molecule resulting in a hydroxide ion which attacks the scissile phosphate^{27,39,40}. It simultaneously uses an aspartic acid and two asparagine residues that mediate the coordination of a magnesium ion^{27,39,40}.

The RuvC domain requires a divalent metal ion that is used to coordinate cleavage in the non-target strand. It has 4 active site residues, one of which facilitates the movement of the divalent metal ion^{27,29}. Like the HNH domain, the active site histidine residue deprotonates the water molecule to attack the scissile phosphate, while the two aspartic acid and glutamic acid residues are used to stabilize the transition state (Fig. 7)^{27,29}.

1.3 – Reasoning/Hypothesis/Objectives

While there are many advantages to using SaCas9 for gene editing, it has evolved to tolerate mismatches in the DNA target sequence. This inaccuracy is considered undesirable for gene editing. Here, we focused on characterizing conservative mutations of active site residues.

A study by McMurrough et al., (2014) and McMurrough et al., (2018) assessed mutations in the active sites of meganucleases that used a two-metal ion cleavage mechanism and found that mutations from aspartic acid to glutamic acid (or vice versa) resulted in a change of cleavage preferences. Based on this study, we focused on the RuvC domain as it follows a similar divalent metal ion mechanism.

As mentioned in 1.2.2.1, RuvC is dependent on 4 catalytic residues for its cleavage reactions. While all 4 residues are potential options, Asp10 is known to be very important for stabilizing the transition state of MgB and was a promising residue to mutate. Similarly, to McMurrough et al., (2018), we mutated Asp10 to Glu10 based on several assumptions:

1. D10E would result in a change in substrate preferences which can alter its discrimination between mismatches ⁴¹(McMurrough et al., 2018)
2. A larger residue will cause some steric hindrance that would slow down the cleavage process that may allow for the proofreading function of SaCas9 to be enhanced.
3. D10E would destabilize the transition state, which could help with dissociation of SaCas9 for mismatched targets.

Based on these assumptions, we hypothesized that a conservative mutation of Asp10 to Glu10 would result in a reduction of off-target cleavage.

2 « Materials and Methods »

2.1 Purification of SaCas9 variants

Plasmid pET11a containing *Staphylococcus aureus* Cas9 WT and D10E were obtained from Specific Biologics Incorporated. Bacterial expression of WT and D10E plasmids were transformed into T7 Express (New England Biolabs - NEB). Cells were cultured at 37°C in LB media until it reached an OD₆₀₀ of 0.6-0.8, after which expression was induced with 1 mM IPTG at 16°C for 20 hours. After induction, cell culture was spun for 10 minutes at 4°C with a speed of 6000g. Pellet was resuspended in binding buffer (10 ml per gram of pellet) with the addition of a protease inhibitor. Sonification was used to lyse the cells at an amplitude of 100% for 10 rounds with on and off times of 10 seconds and 20 seconds, respectively. Lysed cells were spun for an hour at 4°C with a speed of 4000g.

Protein was filtered with a 0.22µm filter and was run on a HisTrap HP His tag protein purification column. 10 column volumes of binding buffer (0.5 M NaCl, 20 mM Tris – pH 8, 10 mM imidazole, 5% glycerol) were used for the first washing step. 10 column volumes of a 50 mM elution buffer (0.5 M NaCl, 20 mM Tris – pH 8, 50 mM imidazole, 5% glycerol) was used for the second washing steps. Increasing levels of imidazole (100, 200, 300, and 500 mM) were used to elute protein and were concentrated using MilliporeSigma™ Amicon™ Ultra-0.5 Centrifugal Filter Units (Fisher Scientific) (Appendix – Figure 20).

2.2 Active Site Mutant Cloning: Golden Mutagenesis

To generate the active site mutants, golden gate methods from Püllmann et al., 2019 were used⁴². Golden gate primers used are on Appendix – Table 4.

2.3 Production of sgRNA

Production of sgRNA was done through a one-pot *in vitro* transcription. A HiScribe T7 High Yield RNA Synthesis Kit from NEB was used. In a 20 μL total volume, Klenow, 6.67 μM of sgRNA oligo, 6.67 μM of universal tracr oligo, 0.125 mM of dNTPs, 10 μL of 1x RiboMAX Express T7 Buffer (half of total volume), 2 μL of T7 Express Enzyme Mix, and 2 μL of nuclease free water was put into a single PCR tube and incubated at 37°C for 4 hours⁴³. No clean-up is required. RNA concentration was determined by Nanodrop.

2.4 In Vitro Time Course Cleavage Assays with Linear Substrate

SaCas9 and its variants were diluted using storage buffer (50 mM Tris-HCl (pH 8), 250 mM NaCl, 1 mM DTT, 5% glycerol) to a working concentration of 140 nM of the enzyme with sgRNA of 280 nM. Proteins were assayed with a reaction mixture of 100 mM NaCl, 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 1 mM DTT, and either 2.5, 5, 7.5, or 10 nM substrate. Protein and the reaction mixture were incubated separately for 10 minutes at 37°C. The assay contained 9 individual time-points with the addition of protein (Fig. 8). Reactions were stopped using 0.5 mM EDTA, 20 mg/ml RNase A, 25 mg/ml Proteinase K, and 1x D-PBS (0.9 mM CaCl₂, 2.7 mM KCl, 0.5 mM MgCl₂-6H₂O, 0.1368 M NaCl, 15.2 mM Na₂HPO₄) and were incubated for 30 minutes at 37°C. Initial rates were calculated by the decay of reactant using an exponential decay function. Initial rates were plotted against substrate concentration to fit to a Michaelis-Menten model to determine K_m and V_{max} on Graphpad Prism 9. Experiments were done in triplicate.

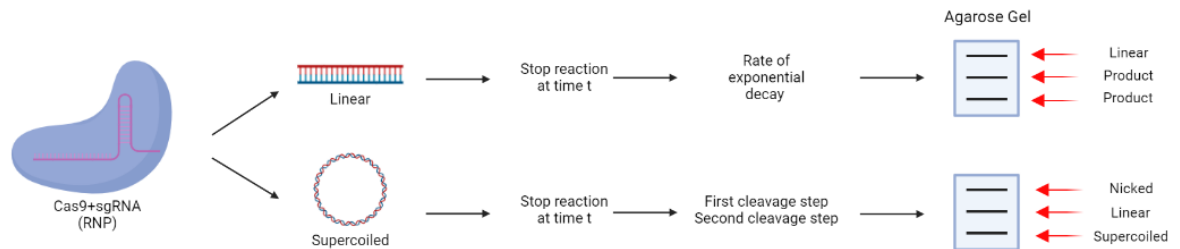


Figure 8. Diagram of an *in vitro* cleavage time course assay. With linear DNA, we can observe the overall rate and with supercoiled DNA, we can observe individual nicking steps (k_1 and k_2).

2.5 In Vitro Time Course Cleavage Assays with Supercoiled Substrate

SaCas9 and its variants were diluted using storage buffer (50 mM Tris-HCl (pH 8), 250 mM NaCl, 1 mM DTT, 5% glycerol) to a working concentration of 140 nM of the enzyme with sgRNA of 280 nM. Proteins were assayed with a reaction mixture of 100 mM NaCl, 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 1 mM DTT, and either 2.5, 5, 7.5, or 10 nM substrate. Protein and the reaction mixture were incubated separately for 10 minutes at 37°C. The assay contained 9 individual time-points with the addition of protein. Reactions were stopped using 0.5 mM EDTA, 20 mg/ml RNase A, 25 mg/ml Proteinase K, and 1x D-PBS (0.9 mM CaCl₂, 2.7 mM KCl, 0.5 mM MgCl₂-6H₂O, 0.1368 M NaCl, 15.2 mM Na₂HPO₄) and were incubated for 30 minutes at 37°C. Reactions were run on 1% agarose gels (Fig. 9). Initial nicking rates (k_1) were calculated by the decay of reactant using an exponential decay function (1). Second cleavage rates (k_2) were calculated using (2) and (3).

$$(1) S = S_0 e^{-k_1 t}$$

$$(2) N = S_0 \left(\frac{k_1}{k_1 - k_2} \right) (-e^{-k_1 t} + e^{-k_2 t})$$

$$(3) L = S_0 \left(1 + \frac{1}{k_1 - k_2} \right) (k_2 e^{-k_1 t} - k_1 e^{-k_2 t})$$

Equation (1) uses S to represent supercoiled DNA, (2) uses N to represent nicked DNA, and (3) uses L to represent linear DNA.

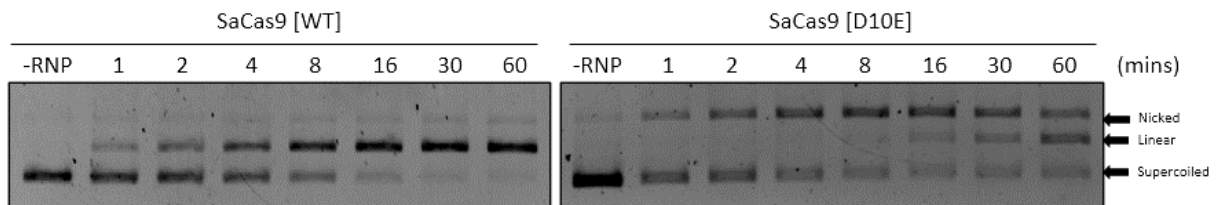


Figure 9. Example of an in vitro time course cleavage assay with EMX1-1 supercoiled substrate performed with SaCas9[WT] and SaCas9[D10E]. Two types of products are formed: nicked and linear.

2.6 Competition Assay

SaCas9[WT] and SaCas9[D10E] were diluted using storage buffer (50 mM Tris-HCl (pH 8), 250 mM NaCl, 1 mM DTT, 5% glycerol) to a working concentration of 250 nM of the enzyme with sgRNA of 500 nM. Proteins were assayed with a reaction mixture of 100 mM NaCl, 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 1 mM DTT, and equimolar of on-target:transversion substrate (10 nM total). Protein and the reaction mixture were incubated separately for 10 minutes at 37°C. The assay contained 9 individual time-points with the addition of protein. Reactions were stopped using 0.5 mM EDTA, 20 mg/ml RNase A, 25 mg/ml Proteinase K, and 1x D-PBS (0.9 mM CaCl₂, 2.7 mM KCl, 0.5 mM MgCl₂-6H₂O, 0.1368 M NaCl, 15.2 mM Na₂HPO₄) and were incubated for 30 minutes at 37°C. Rates were calculated by the decay of reactant through an exponential decay function (1).

2.7 Statistical Analysis

Statistical analysis of *in vitro* cleavage assays was analyzed through ImageLab. Band intensities were based on band percentage per lane for standardization. Bands were analyzed in GraphPad Prism using:

$$(1) S = S_0 e^{-k_1 t}$$

$$(2) N = S_0 \left(\frac{k_1}{k_1 - k_2} \right) (-e^{-k_1 t} + e^{-k_2 t})$$

$$(3) L = S_0 \left(1 + \frac{1}{k_1 - k_2} \right) (k_2 e^{-k_1 t} - k_1 e^{-k_2 t})$$

Error bars used standard error of means (\pm SEM).

3 « Results »

3.1 Target Substrates

Empty spiracles homeobox 1 (EMX1) is a gene that encodes transcription factors for telencephalic developments. Specifically, EMX1 proteins play a role in positional identity, the production of neural stem cells, and differentiation of neuronal phenotypes⁴⁴. Mutations in EMX1 is associated with human brain diseases and is a main therapeutically relevant target used as a candidate for CRISPR-Cas treatment. Target site EMX1 for SaCas9 (5'-GGCCTCCCCAAAGCCTGGCCAGGGAGT-3') (Appendix – Table 2) was inserted into pBbA8k-RFP used as a substrate for *in vitro* cleavage assays. Primers were phosphorylated and annealed then were inserted into the backbone using Golden Gate cloning⁴⁵. Substrates with single mutations will be referred to “Tv #”, in which the number refers to the position of the transversion (Fig. 10).

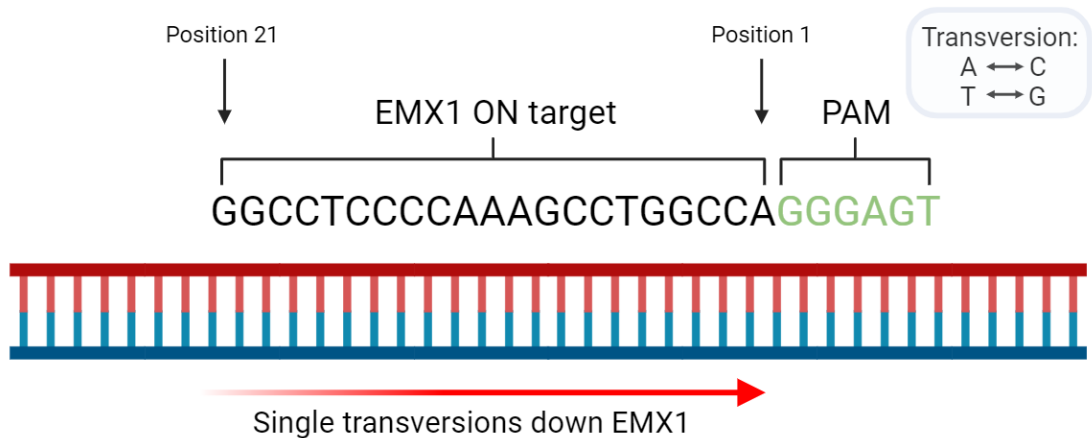


Figure 10. Schematic of generated substrates with single transversions. Positions number is the position adjacent from the PAM. i.e. Position 2 is considered PAM proximal, but position 15 is considered PAM distal.

3.2 SaCas9[D10E] has the highest *in vitro* activity among RuvC active site variants

Previous work at SBI was performed with SaCas9[D10E] (D10E) which displayed a reduced rate in the second nicking step against supercoiled EMX1-1 without eliminating the ability to fully cleave (Fig. 11.). These observations suggest that D10E can potentially change its substrate preferences. However, while D10E is promising, I hypothesized that making other substitutions in the active site would generate Cas9 variants with different cleavage profiles than wild-type while maintaining similar properties to SaCas9[D10E]. To evaluate these substitutions, the 3 catalytic residues of RuvC (E477, H701, and D704) were conservatively mutated to test this hypothesis, resulting in the E477D, H701K, and D704E mutants. To characterize their cleavage abilities, I determined the rate of product formation by the three conservative mutants against supercoiled pBbA8k-RFP carrying the EMX1-1 on-target *in vitro* over 1 hour. E477D and D704E had similar cleavage rates as compared to the other variants (Fig. 12); however, these preliminary cleavage assays were not performed in equal protein concentrations. To accurately characterize RuvC active site variants, *in vitro* cleavage assays with linear, PCR-amplified-substrate (1kb) containing the EMX1-1 on target site were conducted with a 28:1 protein:DNA ratio. Cleavage activity was reduced with E477D and D704E compared to WT and D10E (Fig. 12). As reduced cleavage activity is considered undesirable for *in vivo* gene editing, all further profiling was conducted with D10E.

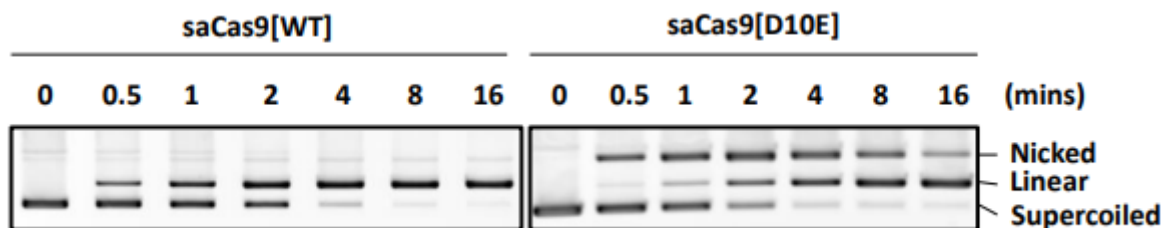


Figure 11. Initial *in vitro* cleavage assay from Specific Biologics Incorporated (SBI) with SaCas9[WT] and SaCas9[D10E] against supercoiled DNA (target unknown). This figure is used to show that the second nicking step slows for SaCas9[D10E] in comparison to WT.

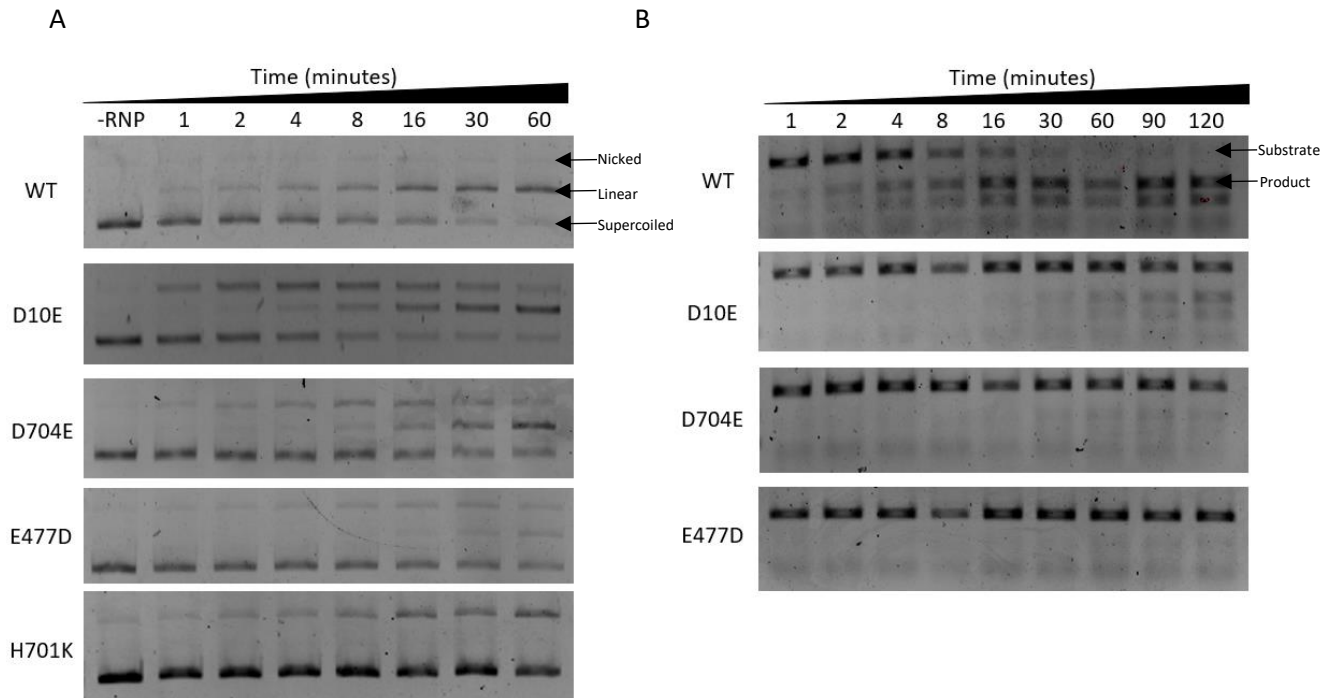


Figure 12. In vitro cleavage assay of RuvC active site variants. (A) Product formation over time for supercoiled EMX1-1 (3.97 kbp) on-target against SaCas9 mutants. Protein:DNA ratio of WT and D10E is 28:1 and other variants had a ratio of 14:1. (B) Linear EMX1-1 on-target assayed against SaCas9 mutants with protein concentrations of 70 nM and 2.5 nM of DNA.

3.3 Rate of SaCas9[D10E] discrimination between targets with single nucleotide substitutions

To evaluate disparities in cleavage activity between WT and D10E SaCas9, five off-target EMX1-1 linear substrates with transversions at positions 3, 4, 10, 15, and 21 (Table 1) were tested with *in vitro* time course cleavage assays (Fig. 13). These transversions were used to sample different positions either inside or outside of the seed region (first 10-12 bps). The resulting reactions were run on an agarose gel and band intensities were determined as described in the Methods sections. A kinetic analysis of WT and D10E was performed to determine the V_{\max} (maximum velocity), K_m (the amount of substrate to reach half of V_{\max}), and K_{cat} (turnover rate). These kinetic analyses can be useful to understand the catalytic mechanism of the newly generated SaCas9[D10E].

Table 1. DNA sequence of EMX1-1 off-targets for linear *in-vitro* cleavage assays. Red lettering corresponds with the position that has the single transversion.

Name	Sequence (5' → 3')
Tv 3	GGCCTCCCCAAAGCCTGGCCA
Tv 4	GGCCTCCCCAAAGCCTGTACA
Tv 10	GGCCTCCCCAACGCCTGGACA
Tv 15	GGCCTCACCAAAGCCTGGACA
Tv 21	TGCCTCCCCAAAGCCTGGACA

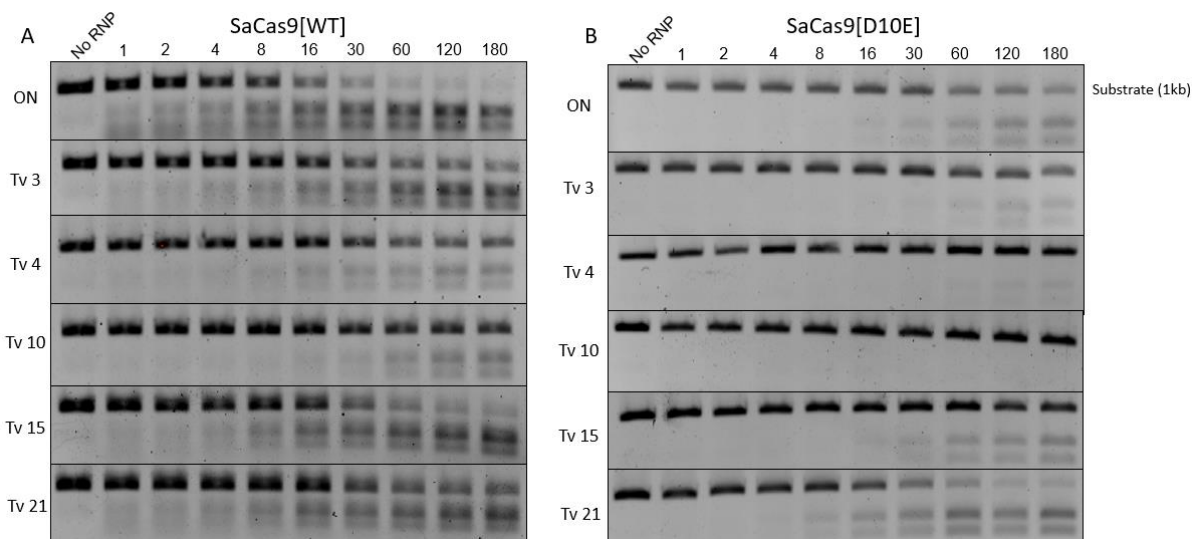


Figure 13. Example of an *in vitro* cleavage assay of 5 nM on- and off-target linear substrates (Tv 3, 4, 10, 15, and 21) against SaCas9[WT] and SaCas9[D10E]. Reactions were run on a 1% agarose gel and stained with EtBr. (A) Cleavage of SaCas9[WT] (B) Cleavage of SaCas9[D10E].

With WT SaCas9, the V_{max} decreased for all substrates except for Tv 4 and Tv 15 (Table 2). However, we chose to exclude Tv 4 from further analysis as we observed unstable enzyme kinetics and poor correlation for this transversion. The K_m of SaCas9[WT] against substrates with transversions outside of the seed region (Tv 15 and Tv 21) were higher compared to SaCas9[WT] against the EMX1-1 on-target. In contrast, the K_{cat} values of most off-target substrates were reduced. Tv 15 was the sole exception, which had similar k_{cat} values to the on-target substrate (Table 2).

With respect to SaCas9[D10E], no cleavage activity was observed during initial kinetic evaluations of this mutant against Tv 10. For several transversions (Tv 3, Tv 4, and Tv 15), D10E displayed reduced V_{max} values to the on-target. In contrast, Tv 21 displayed a 3-fold increase in the V_{max} . The K_m of D10E against Tv 3, 4, 15, and 21 increased compared to the on-target. Furthermore, the k_{cat} of Tv 21 increased 3-fold, a finding mirrored in the K_m .

The catalytic efficiency (K_{cat}/K_m) of D10E was higher against the on-target substrate compared to WT. However, the cleavage efficiency for the WT variant was higher against the transversions, demonstrating its ability to tolerate single nucleotide substitutions.

V_{\max} and K_m values were relatively similar for WT against all substrates suggesting that WT could overcome transversions (Table 2). On the other hand, D10E displays different enzyme kinetics between the on-target substrates compared to the transversions. Therefore, D10E has the potential to discriminate between on-targets and be intolerant to individual mismatches (Fig. 14).

Table 2. Kinetic analyses of SaCas9[WT] and SaCas9[D10E] against on and off targets.

Table of V_{\max} , K_m , and k_{cat} of WT and D10E. Analyses were generated from GraphPad Prism9.

		ON	3	4	10	15	21
WT	V_{\max} (min^{-1})	1.82E-02	7.76E-03	N/A	1.11E-02	1.10E-01	1.20E-02
	K_m (nM)	5.86E+00	7.44E-01	Unstable	9.67E+00	5.21E+01	3.14E-01
	K_{cat} (min^{-1})	1.30E-04	5.55E-05	N/A	7.94E-05	7.84E-04	8.54E-05
	K_{cat}/K_m	2.22E-05	7.46E-05	N/A	8.22E-06	1.50E-05	2.72E-04
D10E	V_{\max} (min^{-1})	1.09E-02	4.79E-03	7.43E-03	N/A	6.57E-03	3.31E-02
	K_m (nM)	1.23E+00	2.40E+00	6.38E+01	N/A	5.05E+00	2.54E+00
	K_{cat} (min^{-1})	7.81E-05	3.42E-05	5.31E-05	N/A	4.69E-05	2.36E-04
	K_{cat}/K_m	6.35E-05	1.42E-05	8.32E-07	N/A	9.29E-06	9.32E-05

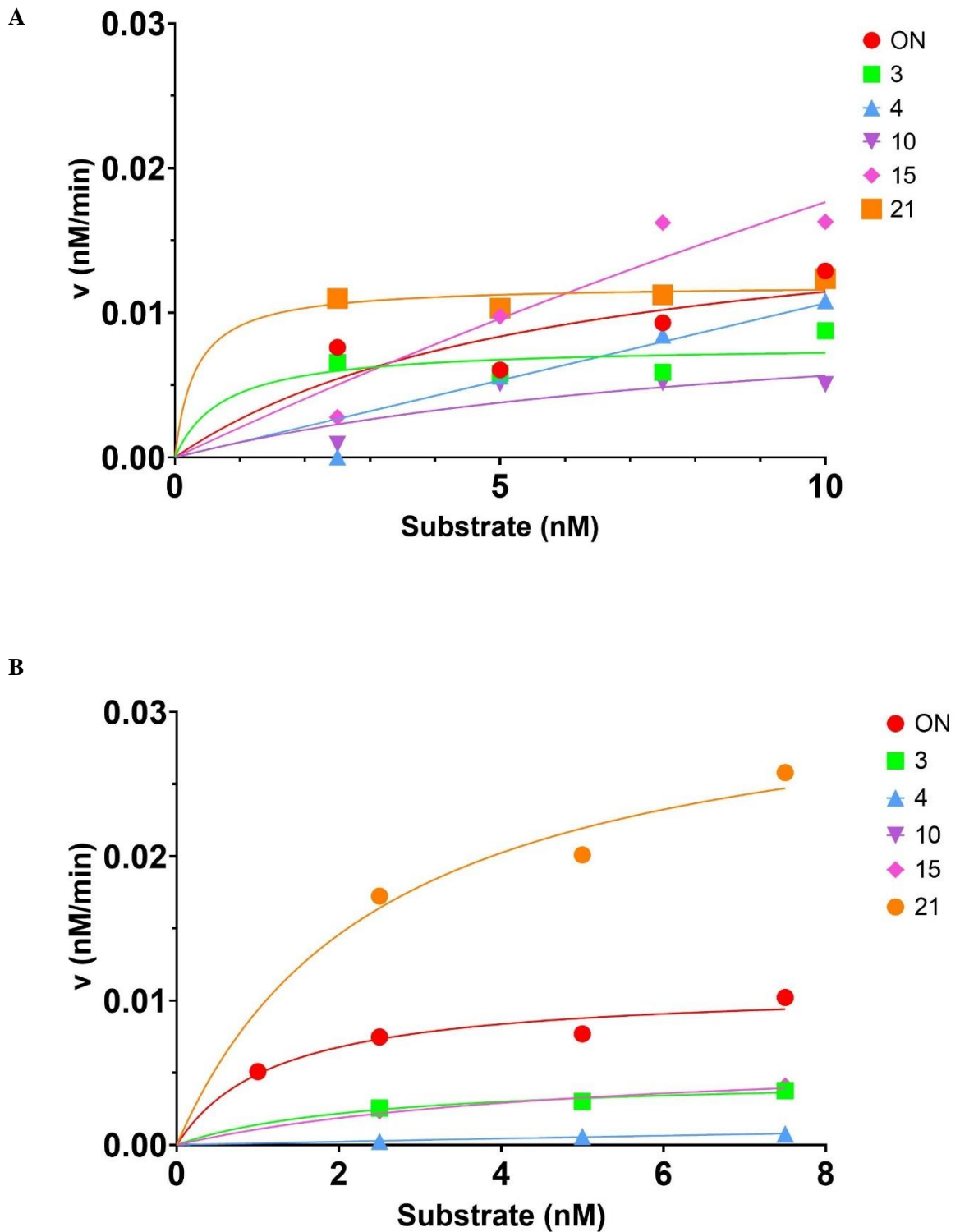


Figure 14. Michaelis-Menten curves of SaCas9[WT] and SaCas9[D10E]. In vitro cleavage assay of SaCas9[WT] and SaCas9[D10E] against substrates with transversions at positions 3, 4, 10, 15, and 21. Linear EMX1-1 targets assayed against (A) SaCas9[WT] and (B) SaCas9[D10E] against substrates with DNA concentrations between 2.5 and 10 nM.

3.4 The initial nicking rate is dependent on the position of single transversions

Studies involving *in vitro* cleavage assays typically only utilize linear substrates. However, it may be misleading to solely study SaCas9 in the context of linear substrates, considering that linear double stranded substrates cannot distinguish nicking activity and *in vivo* use of SaCas9 mutants will more commonly be used against topologically complex DNA. Therefore, it is important to understand the impact of substitutions in the active site of SaCas9 on the nicking reactions, which was performed by determining the independent rates of WT and D10E.

To find first and second strand nicking rates, *in vitro* cleavage assays were performed using WT and D10E against supercoiled substrates. The first nicking step (k_1 rates) was calculated from the exponential decay of supercoiled DNA and the second nicking step (k_2) was calculated from the decay of open circular nicked product and formation of linear product (Fig. 16; Fig. 17). It is important to note that k_1 and k_2 are usually associated with the HNH and RuvC cleavage domains, respectively³⁵ and looking at individual nicking steps could provide more insight into the cleavage mechanism of SaCas9[D10E].

I found that the k_1 rates of SaCas9[WT] against the on-target and each transversion displayed similar activity with notable local peaks at Tv3, Tv12, and Tv21 with rates of 0.1154, 0.1155, and 0.1843 min^{-1} , respectively (Fig. 18). As compared to the WT enzyme, single transversions were able to influence k_1 D10E rates, which were largely dependent on the position of the mutation. Transversions proximal to the PAM (in positions 1-12) (Fig. 15) displayed reduced k_1 rates whereas enhanced k_1 rates were observed with transversions in PAM distal regions (positions 13-21) nicking activity.

k_2 rates were assayed in a similar manner; however, because the WT enzyme rapidly cleaved supercoiled substrates, the first nicked product could not be observed. To calculate k_2 rates, nicked product needs to be observable for quantification. To slow down cleavage activity and visualize nicked product, the concentration of MgCl_2 in the reaction buffer was reduced 100-fold to 0.1 mM. Furthermore, the reaction time for the *in vitro* cleavage assays extended to 3 hours. The calculated k_2 with 0.1mM MgCl_2 for D10E displayed overall slower cleavage activity compared to WT (Fig.19).

Collectively, these data demonstrate that rate of the first nicking step (k_1) is dependent on the position of the transversion relative to the PAM site. The k_1 plot in Figure 16 displayed a similar pattern for WT and D10E with respect to the positional effect, except that the k_1 rates for WT overall are faster than for D10E.

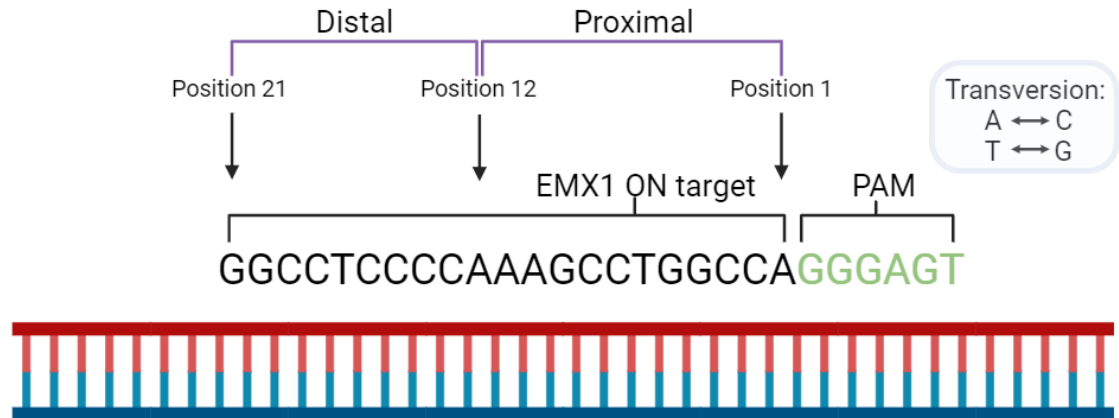
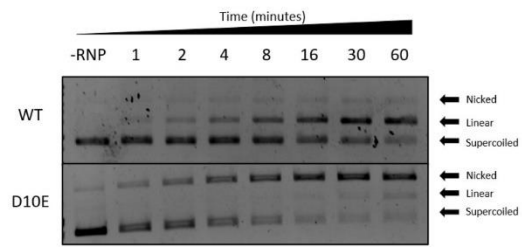
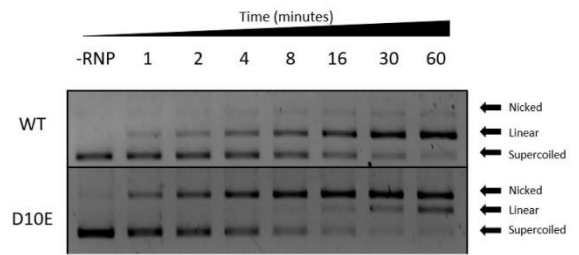


Figure 15. Schematic of EMX1 on-target. PAM (green) proximal positions are between position 1 and 12. PAM distal positions are between position 13 and 21.

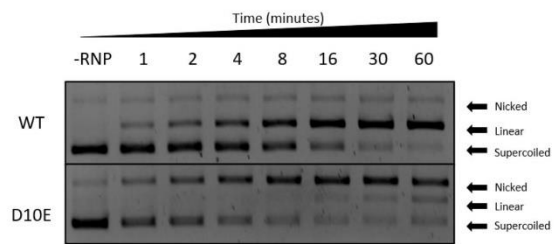
Transversion 1



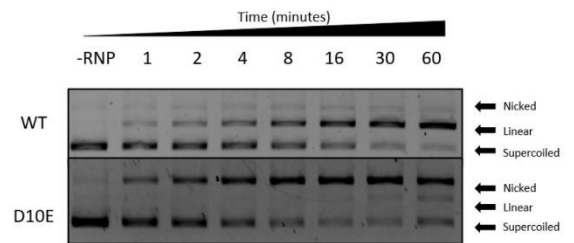
Transversion 2



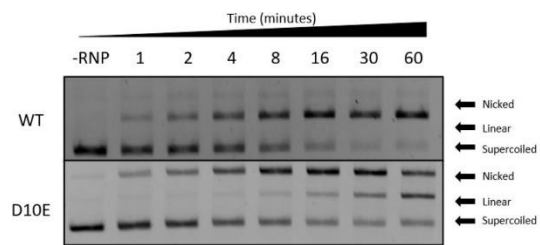
Transversion 3



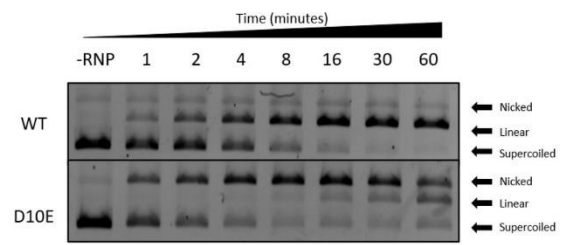
Transversion 4



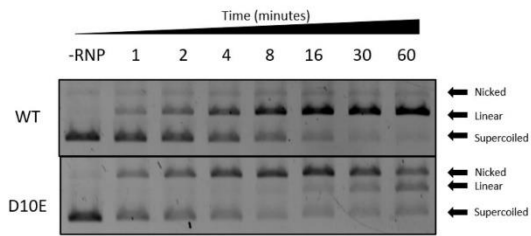
Transversion 5



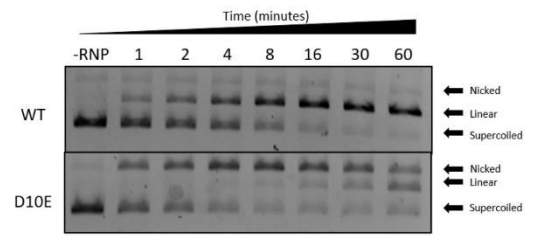
Transversion 6



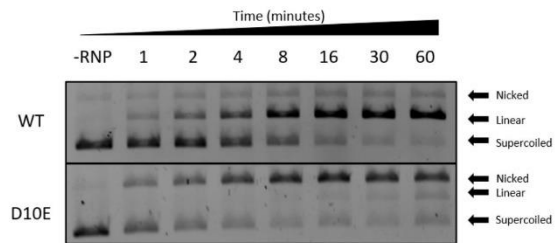
Transversion 7



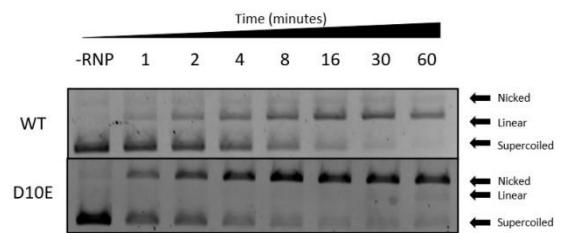
Transversion 8



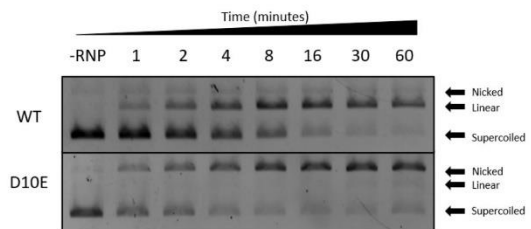
Transversion 9



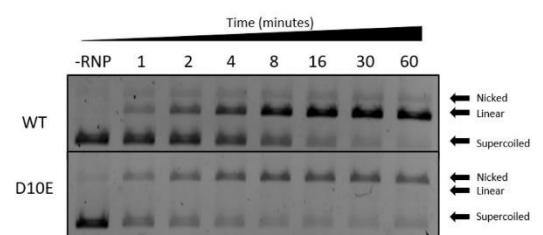
Transversion 10



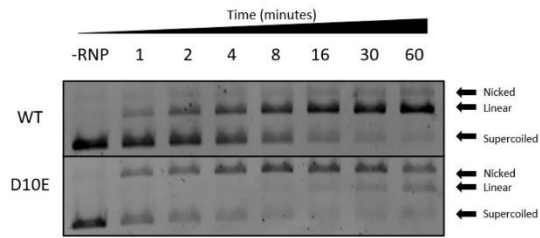
Transversion 11



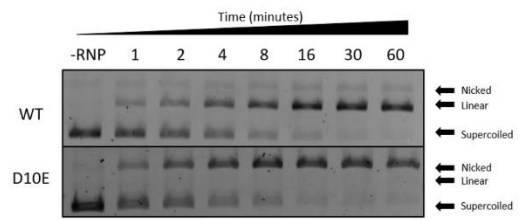
Transversion 12



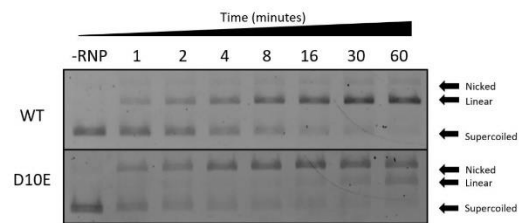
Transversion 13



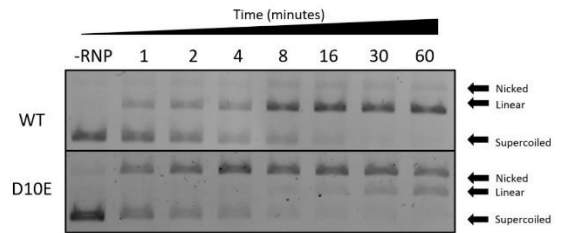
Transversion 14



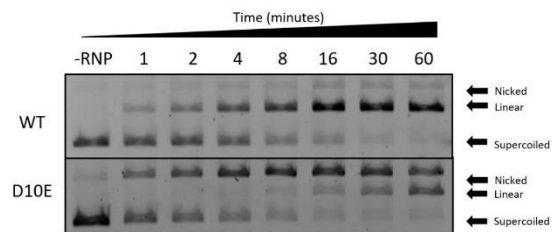
Transversion 15



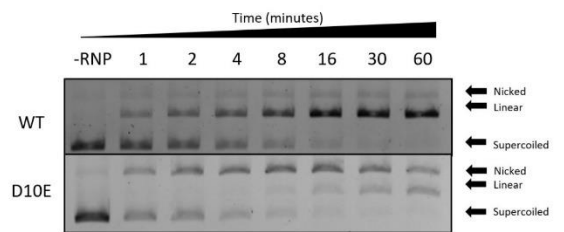
Transversion 16



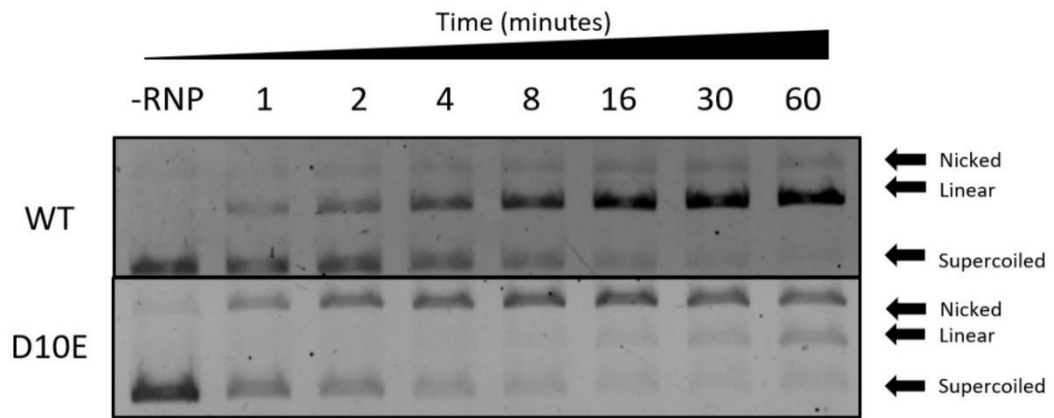
Transversion 17



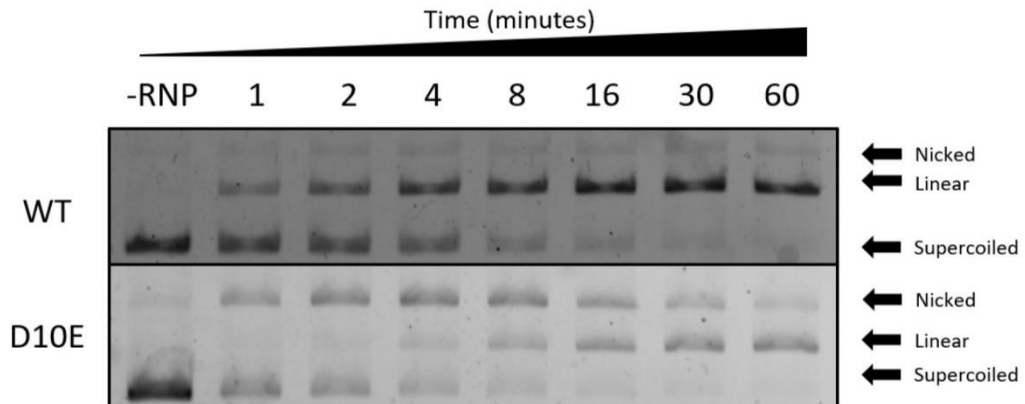
Transversion 18



Transversion 19



Transversion 20



Transversion 21

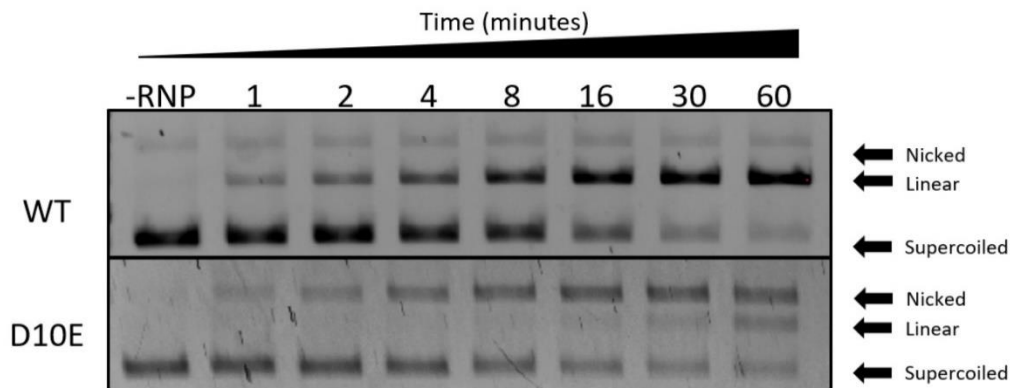


Figure 16. Analysis of on- and off-target EMX1-1 supercoiled DNA against WT and D10E. Agarose gels depict a representation for the corresponding graph shown in Figure 18 (N=3, SEM).

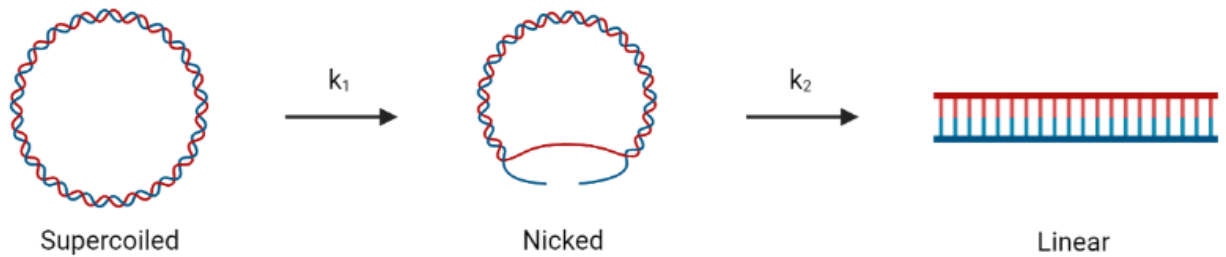


Figure 17. Visual representation of k_1 and k_2 rates. k_1 signifies the depletion of supercoiled and the formation of nicked DNA. k_2 represents the formation of linear and the depletion of nicked DNA.

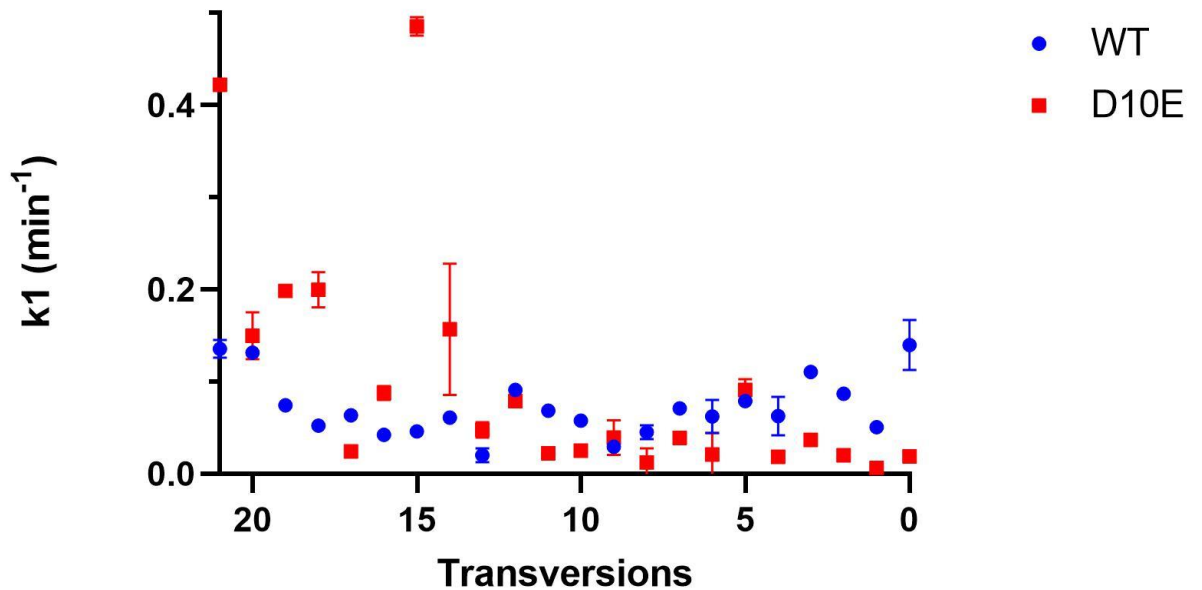


Figure 18. In vitro cleavage assay of SaCas9[WT] and SaCas9[D10E] against substrates at all positions. Supercoiled EMX1 targets assayed against SaCas9[WT] and SaCas9[D10E] against substrates with protein concentrations of 70 nM and DNA concentrations of 2.5 nM. k_1 rates calculated through exponential decay. Rates calculated by GraphPadPrism9 with $N=3$. Error bars = SEM.

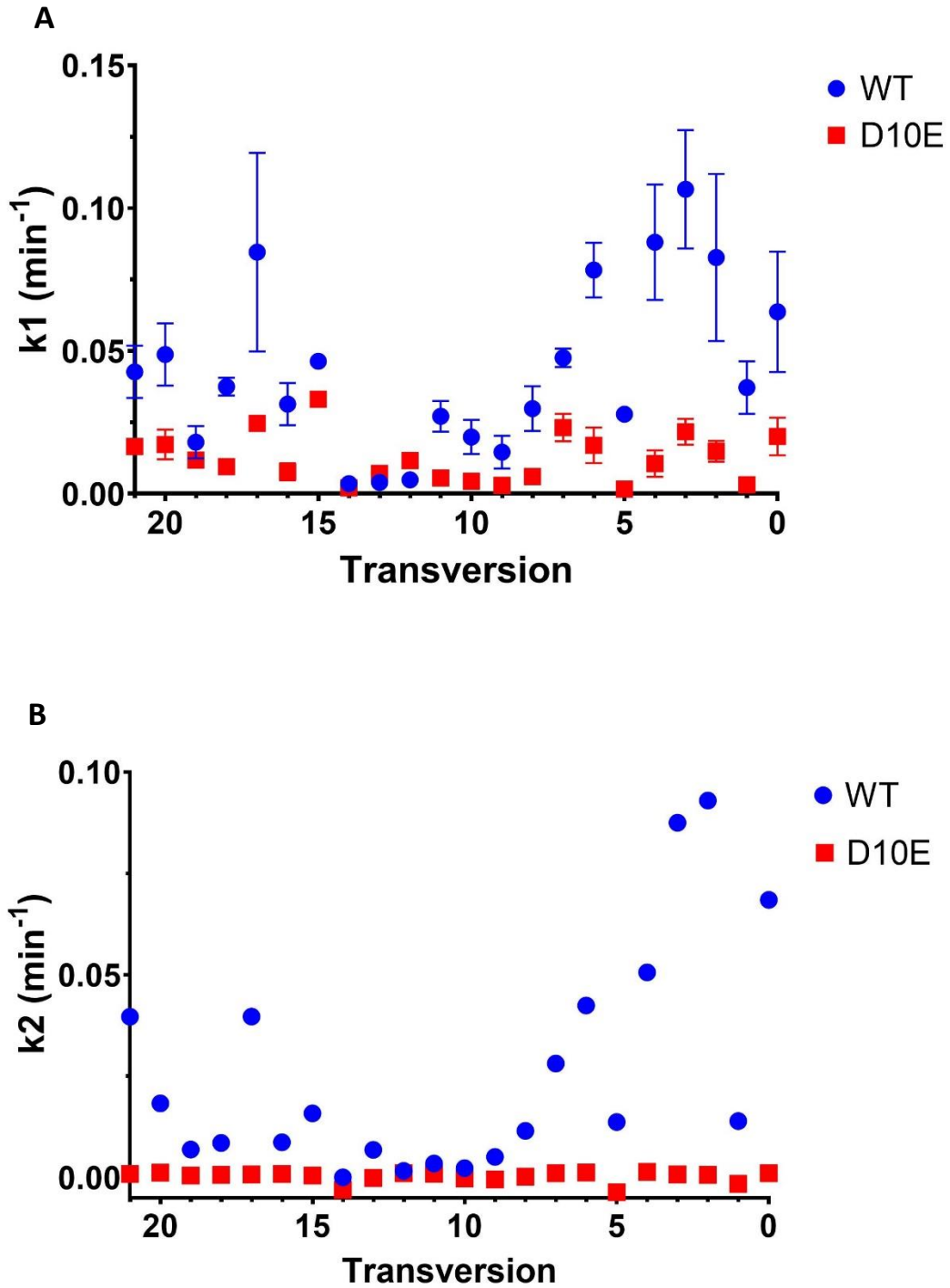


Figure 19. In vitro cleavage assay of SaCas9[WT] and SaCas9[D10E] against substrates at positions all positions. Supercoiled EMX1 targets assayed against SaCas9[WT] and SaCas9[D10E] against substrates with protein concentrations of 70 nM and DNA concentrations of 2.5 nM. (A) k1 rates calculated through exponential decay and (B) k2 rates calculated with decay of nicked product and formation of linear products. Rates calculated by GraphPadPrism9 with N=3. Error bars = SEM.

3.5 The first nicking step is performed by the HNH domain

While the HNH domain is known to be responsible for the first cleavage step³¹, independent rates from the previous *in vitro* cleavage assays suggests that the first cleavage step of SaCas9[D10E] may not be dependent on the initial nicking step by the HNH domain as mutation of the RuvC site affected nicking of the initial nicking step. To assess if the RuvC domain acts as the first nicking step, the catalytic activity of the HNH domain was knocked out through a SaCas9[D10E]/[H557A] (DEHA) double mutant, resulting in all subsequent cleavage activity performed by the RuvC domain. The DEHA mutant was incubated with the on-target substrate, Tv 3, Tv 10, and Tv 21 for 60 minutes.

Against the on-target EMX1 substrate, the DEHA mutant was unable to cleave over the course of 1 hour. Similar results were observed for substrates with transversions at position 3, 10, and 21 (Fig. 20).

Therefore, these findings would suggest that SaCas9[D10E] cleavage first occurs by binding and nicking by the HNH domain, prior to nicking by the RuvC domain. Results suggests that the independent cleavage trend we previously observed with k_1 results from the RuvC mutation can potentially affect the conformation of the HNH domain.

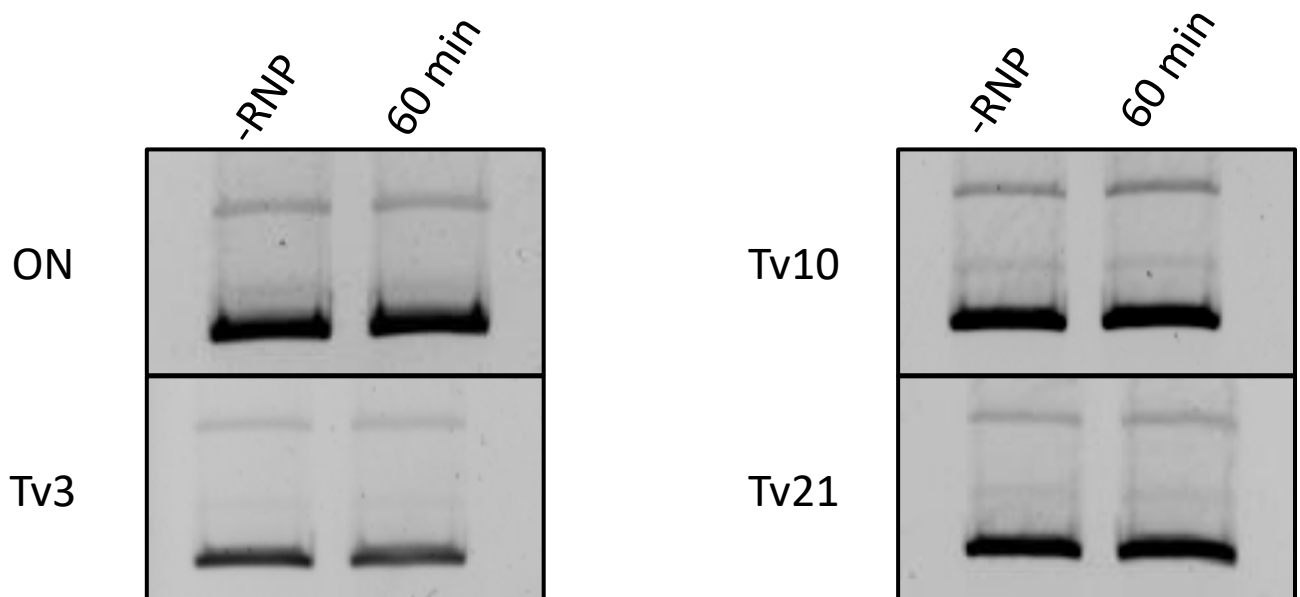


Figure 20. *In vitro* cleavage assay of SaCas9[D10E]/[H557A]. Assay was performed on EMX1-1 supercoiled DNA and ran on a 1% agarose gel and stained with EtBr.

3.6 SaCas9[D10E] can discriminate between on and off-targets

To determine if D10E can discriminate between on-targets and transversions, competition assays were performed by incubating SaCas9[D10E] with equimolar ratios of on-target and single transversion substrates. The following experiments were performed with Olha Haydaychuk. Linear substrates with single transversions at positions 1, 3, 10, and 21 were used in *in vitro* time course cleavage assays up to 3 hours with 5 nM of each substrate and a final protein concentration of 140 nM (N = 3).

No discrimination between on- and off-targets were observed with SaCas9[WT] when incubated with the on-target substrate and Tv 3, Tv 10, and Tv 21. However, the WT enzyme displayed preference for the on-target substrate when paired with Tv 1. The WT enzyme preferentially cleaves substrates with transversions that occur outside of the seed region. In contrast, D10E favoured the on-target over off-target transversions and showed preference for transversions in the PAM distal region (Fig. 21; Fig. 22). Similar to the k_1 rate trend that we previously observed with SaCas9[D10E], D10E k_1 values were reduced (transversion 1, 3, and 10) or enhanced (transversion 21) compared to activity against the on-target substrate. Altogether, these results demonstrate that D10E can differentiate between on-targets and single transversions on an EMX1 substrate.

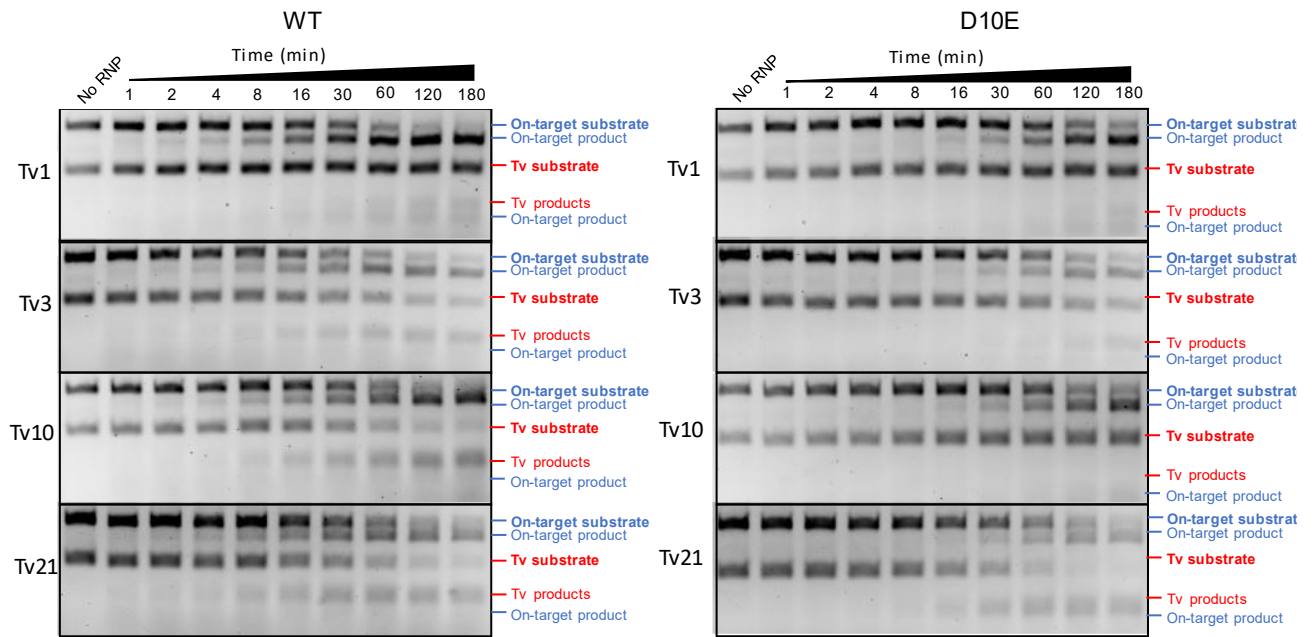


Figure 21. Comparison of on and off targets via a competition assay with WT and D10E variants. Cleavage assays were performed in triplicate (error bars indicate SEM) and resolved on 2.25% agarose gels and stained with EtBr.

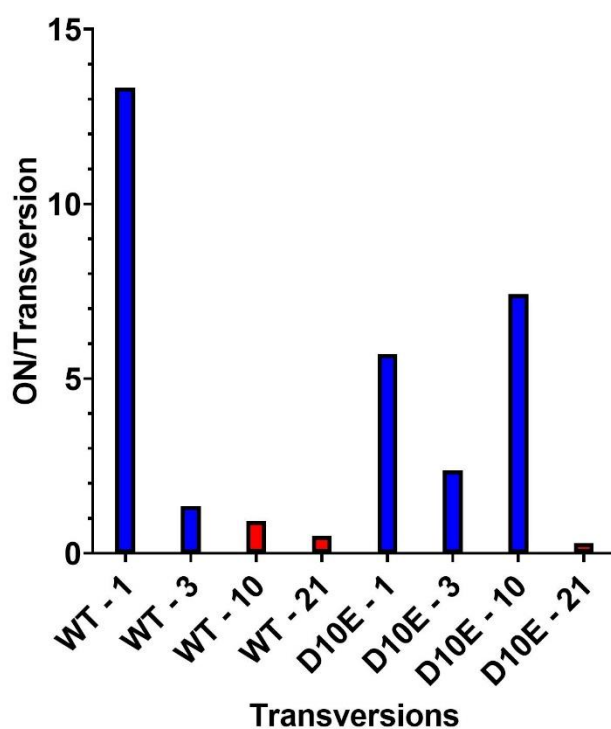
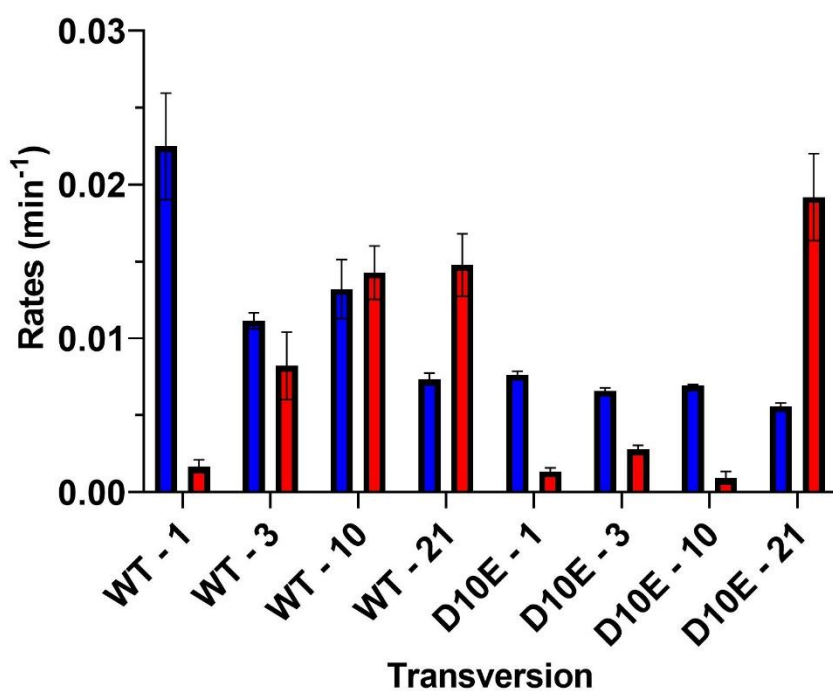


Figure 22. Analysis of competition assay. A) Rates of on target (blue) and transversions (red) (N = 3, \pm SEM) B) Rates of ON/Tv with blue representing preferences for on-target and red bars representing preference for transversions.

Chapter 4

4 « Discussion »

4.1 Kinetic Comparisons Between Single Mutations

The focus of this thesis was to determine if SaCas9[D10E] can reduce off-target cleavage. To see if that is true, SaCas9[D10E] was assayed along with its wild-type in an *in vitro* time course cleavage assay and their enzyme kinetics were compared. For single mutations, SaCas9[D10E] had decreased k_1 nicking rates for PAM proximal mutations and enhanced k_1 nicking rates for PAM distal mutations. This did not conclude that D10E can reduce off-target cleavage, however, it demonstrates that SaCas9[D10E] can differentiate between single mutations.

SaCas9[WT] is a multi-turnover enzyme³⁰; however, *in vitro* studies in this paper required an excess of protein to DNA ratios of 14:1 to 28:1 for observable cleavage since we did not see any cleavage with excess DNA to protein ratios. As our experiments were conducted with a molar excess of protein, we found SaCas9[WT] cleavage rates that were not consistent with other papers^{30,46}. Yourik et al., 2019 found a V_{max} between 1 to 3.3 min^{-1} for overall cleavage. They also determined a K_m between 8 to 10 nM. In contrast, the results from our study determined a V_{max} and K_m of SaCas9[WT] against on-target EMX1-1 of 0.018 min^{-1} and 5.86 nM, respectively, resulting in V_{max} findings that are 55-fold slower and K_m findings about 1.5-fold reduced. The rate of SaCas9[WT] could be attributed to protein purity (Appendix – Figure 23), the protein:sgRNA ratio PAM sequence, protein:DNA ratio, and different reaction buffers.

Overall, the WT enzyme displayed cleavage rates that were consistent across all transversions. However, we observed notable changes between on- and off-target substrates with the D10E variant. Due to the differences observed between transversions and D10E, the independent cleavage rates of each transversion substrate needed to be further characterized.

4.2 Independent nicking rates demonstrate discrimination with SaCas9[D10E]

The observed trends of cleavage rates for transversions in PAM proximal positions led to the initial hypothesis that D10E would have overall reduced rate of cleavage relative to wild-type SaCas9. However, this hypothesis seemed disproven as the presence of PAM distal transversions increased k_1 cleavage rates. The observed trend in k_1 was unexpected (under the assumption that nicking was domain dependent), as k_1 rates are often associated with HNH domain cleavage³⁶. Interestingly, it appeared that while the WT enzyme displayed similar cleavage rates across all transversions and on-target substrates, D10E displayed discrimination, irrespective of the mutation position. Further evidence for discrimination by D10E was demonstrated by the competition cleavage assay with linear DNA substrates, where D10E displayed clear evidence of discrimination, preferring towards one substrate over another. In contrast, the WT enzyme preferences displayed no discrimination.

Previous studies suggested that mismatches within the seed region impact heteroduplex formation due to reduced binding affinity³⁷. Conversely, the impact is reduced if the same mutations occur outside of the seed region. While this study found similar trends with PAM proximal mismatches, PAM distal mutations appear to deviate from current findings. Since we observed discrepancies with previous research, this study suggests that it is unlikely that binding affinity is the main factor influencing our observed cleavage rates.

Furthermore, Liu et al., 2020 suggested that slowing down the cleavage rate in general can increase discrimination as enzyme kinetics favours release of the substrate over cleavage. However, it is important to establish that discrimination is not just dependent on the relative rates of DNA cleavage. Rather, given the rapid formation of R-loops, discrimination hinges on the rates of both DNA release and cleavage⁴⁷.

Considering previous findings and observations in this study, we can conclude that SaCas9[D10E] can discriminate between substrates with no substitutions and substrates with single transversion substitutions.

4.3 The initial nicking step is coordinated through the conformational change of HNH

Previous studies have shown that the initial cleavage step of Cas9 first occurs through a conformational change that allows the HNH domain to bind the target strand^{27,29}. However, Babu et al., 2021 suggested this may not be the case for all Cas9 proteins. They raised the possibility that some Cas9 variants may possess two pathways to cleavage: a target strand pathway and a non-target strand pathway. They stated that with their L64P/K65P SpCas9 variant, any mutation that occurs in the target site would result in the new variant switching to the non-target pathway⁴⁸. Given the diversity of Cas9 homologs, there is a possibility that variations occur naturally. It is possible some of these variants could coordinate cleavage in a similar manner that induces the non-target pathway.

Babu et al., 2021 revealed a similar phenotype to what was observed for SaCas9[D10E] when they generated an SpCas9 variant (L64P/K65P) with mutations in the bridge helix (BH) (SpyCas9^{2Pro}). Babu et al., 2021 found that SpyCas9^{2Pro} could cleave a substrate with a mutation at position 18 but not a matched DNA target, suggesting that nuclease coordination of a mismatched target DNA is modulated by the BH. Their study proposed that weakened DNA base pairing resulted in enhanced RuvC activity⁴⁸. It may be possible to extend these conclusions to our own findings, as interactions with the BH due to D10E could have resulted in the increased cleavage activity we observed in Cas9.

Given that no cleavage activity was observed with the SaCas9[D10E]/[H557A] (DEHA) variant (for which HNH domain was made non-functional), it is clear that a mutation in the RuvC domain did not deviate from the conventional target strand pathway. As DEHA seemed to follow the target strand pathway, it is plausible that D10E could have affected the conformation of the REC2 and REC3 domain. This could have directly impacted interactions with PAM distal regions of our target strand³⁶.

In fact, Chen et al., 2017 proposed that REC3 directly interacts with the RNA-DNA heteroduplex, particularly in the PAM distal region, which is crucial for re-orienting REC2. Mutations in REC3 seems to interrupt HNH progression past a conformational checkpoint. We propose that D10E could overcome this conformational checkpoint with a PAM distal mutation and potentially facilitate quicker movements of the HNH domain.

It is worth noting that the mutation impacted RuvC cleavage (second cleavage) as expected; however, the initial rates of cleavage present more insight on Cas9 cleavage mechanisms and should be further studied.

4.4 Future Work

4.4.1 Transition vs Transversion Comparisons

To better understand Cas9 substrate preferences, future studies can explore the effect of transition mutations on cleavage activity. This can be performed in addition to expanding target substrates to include other clinically relevant genes such as VEGFA and CFTR. This approach would shed light on whether substrate preferences of Cas9 are solely based on specific mutations these findings can be generalized for most genes. Continued work can aid in the gene editing capabilities of Cas9.

4.4.2 Mammalian Work

While evidence presented from this study demonstrated discrimination *in vitro*, it is important to account for the complexities of *in vivo* environments which introduce significant complexity compared to *in vitro* conditions. To bridge this gap, we can detect genome editing through a T7 assay using mammalian cells to provide insights into if the observed discrimination translates to a more biologically relevant context. Investigating D10E may allow for a better understanding of factors that influence activity, specificity, and potential off-target cleavage. With these findings, we will gain further insight into Cas9 cleavage and improve current gene editing practices.

4.5 Conclusions

The premise of this study was to evaluate SaCas9[D10E] to determine whether this variant of Cas9 can reduce off target cleavage. While SaCas9[D10E] seemed unable to reduce off target cleavage, evidence has displayed that this variant can discriminate between single nucleotide mismatches.

As well, certain mutations in the RuvC domain appear to influence HNH binding and coordination with target DNA substrates. Several studies suggest that this may be due to indirect interaction with REC2 and REC3. Further investigation is required to fully evaluate the cleavage mechanism behind SaCas9.

Although, the results obtained from this study provide insight into the function of Cas9 as well as advances in improvement in gene editing precision.

References

1. Muller, H. J. Artificial Transmutation of the Gene. *Science (1979)* **66**, 84–87 (1927).
2. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. Insertion of DNA sequences into the human chromosomal β -globin locus by homologous recombination. *Nature* **317**, 230–234 (1985).
3. Scherer, S. & Davis, R. W. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc Natl Acad Sci U S A* **76**, 4951 (1979).
4. Rudin, N., Sugarman, E. & Haber, J. E. Genetic and physical analysis of double-strand break repair and recombination in *Saccharomyces cerevisiae*. *Genetics* **122**, 519–534 (1989).
5. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science (1979)* **300**, 764 (2003).
6. Miller, J., McLachlan, A. D. & Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* **4**, 1609–1614 (1985).
7. Boch, J. *et al.* Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science (1979)* **326**, 1509–1512 (2009).
8. Carroll, D. Genome Editing: Past, Present, and Future. *Yale J Biol Med* **90**, 653–659 (2017).
9. Carroll, D. Genome Engineering with Targetable Nucleases. *Annu Rev Biochem* **83**, 409–439 (2014).
10. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* **169**, 5429–5433 (1987).
11. Guirouilh-Barbat, J. *et al.* Impact of the KU80 Pathway on NHEJ-Induced Genome Rearrangements in Mammalian Cells. *Mol Cell* **14**, 611–623 (2004).
12. Moore, J. K. & Haber, J. E. Cell Cycle and Genetic Requirements of Two Pathways of Nonhomologous End-Joining Repair of Double-Strand Breaks in *Saccharomyces cerevisiae*. *Mol Cell Biol* **16**, 2164–2173 (1996).
13. Hinnen, A., Hicks, J. B. & Fink, G. R. Transformation of yeast. *Proceedings of the National Academy of Sciences* **75**, 1929–1933 (1978).
14. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**, 1156 (1996).
15. Deng, D. *et al.* Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335**, 720–723 (2012).
16. Mak, A. N. S., Bradley, P., Cernadas, R. A., Bogdanove, A. J. & Stoddard, B. L. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335**, 716 (2012).

17. Gupta, R. M. & Musunuru, K. The Journal of Clinical Investigation The emergence of genome-editing technology. doi:10.1172/JCI72992.
18. Thierry, A. & Dujon, B. Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. *Nucleic Acids Res* **20**, 5625–5631 (1992).
19. Epinat, J.-C. *et al.* A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. doi:10.1093/nar/gkg375.
20. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science (1979)* **315**, 1709–1712 (2007).
21. Hille, F. & Charpentier, E. CRISPR-Cas: biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, (2016).
22. Wei, Y., Terns, R. M. & Terns, M. P. Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. doi:10.1101/gad.257550.114.
23. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. (2011) doi:10.1038/nature09886.
24. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
25. Jinek, M. *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science (1979)* **343**, 1247997–1247997 (2014).
26. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579–E2586 (2012).
27. Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).
28. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, (2014).
29. Nishimasu, H. *et al.* Crystal structure of *Staphylococcus aureus* Cas9. *Cell* **162**, 1113 (2015).
30. Yourik, P., Fuchs, R. T., Mabuchi, M., Curcuru, J. L. & Robb, G. B. *Staphylococcus aureus* Cas9 is a multiple-turnover enzyme. (2019) doi:10.1261/rna.
31. Sternberg, S. H., LaFrance, Benjamin, Kaplan, Matias & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR-Cas9. (2015) doi:10.1038/nature15544.
32. Zuo, Z. & Liu, J. Cas9-catalyzed DNA Cleavage Generates Staggered Ends: Evidence from Molecular Dynamics Simulations OPEN. *Nature Publishing Group* (2016) doi:10.1038/srep37584.
33. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems HHS Public Access. *Science (1979)* **339**, 819–823 (2013).

34. Dagdas, Y. S., Chen, J. S., Sternberg, S. H., Doudna, J. A. & Yildiz, A. A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Sci Adv* **3**, (2017).
35. Palermo, G. *et al.* Key role of the REC lobe during CRISPR-Cas9 activation by ‘sensing’, ‘regulating’, and ‘locking’ the catalytic HNH domain HHS Public Access. *Q Rev Biophys* **51**, (2018).
36. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407 (2017).
37. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. (2014) doi:10.1038/nature13579.
38. Puppulin, L. *et al.* Dynamics of Target DNA Binding and Cleavage by *Staphylococcus aureus* Cas9 as Revealed by High-Speed Atomic Force Microscopy. *ACS Nano* **17**, 4641 (2023).
39. Yoon, H., Zhao, L. N. & Warshel, A. Exploring the Catalytic Mechanism of Cas9 Using Information Inferred from Endonuclease VII. *ACS Catal* **9**, 1329–1336 (2019).
40. Li, C. L. *et al.* DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. *EMBO J* **22**, 4014 (2003).
41. Mcmurrough, T. A. *et al.* Active site residue identity regulates cleavage preference of LAGLIDADG homing endonucleases. *Nucleic Acids Res* **46**, 11990–12007 (2018).
42. Püllmann, P. *et al.* Golden Mutagenesis: An efficient multi-site-saturation mutagenesis approach by Golden Gate cloning with automated primer design. *Scientific Reports* | **9**, 10932 (2019).
43. Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science (1979)* **368**, 290–296 (2020).
44. Molecular mechanisms underlying cell fate specification in the developing telencephalon Schuurmans and Guillemot 27.
45. Engler, C., Kandzia, R. & Marillonnet, S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS One* **3**, 3647 (2008).
46. Shanzhong Gong, A. *et al.* DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity Article DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. (2018) doi:10.1016/j.celrep.2017.12.041.
47. Liu, M.-S. *et al.* Engineered CRISPR/Cas9 enzymes improve discrimination by slowing DNA cleavage to allow release of off-target DNA. doi:10.1038/s41467-020-17411-1.
48. Babu, K. *et al.* Coordinated Actions of Cas9 HNH and RuvC Nuclease Domains Are Regulated by the Bridge Helix and the Target DNA Sequence. *Biochemistry* **60**, 3783–3800 (2021).

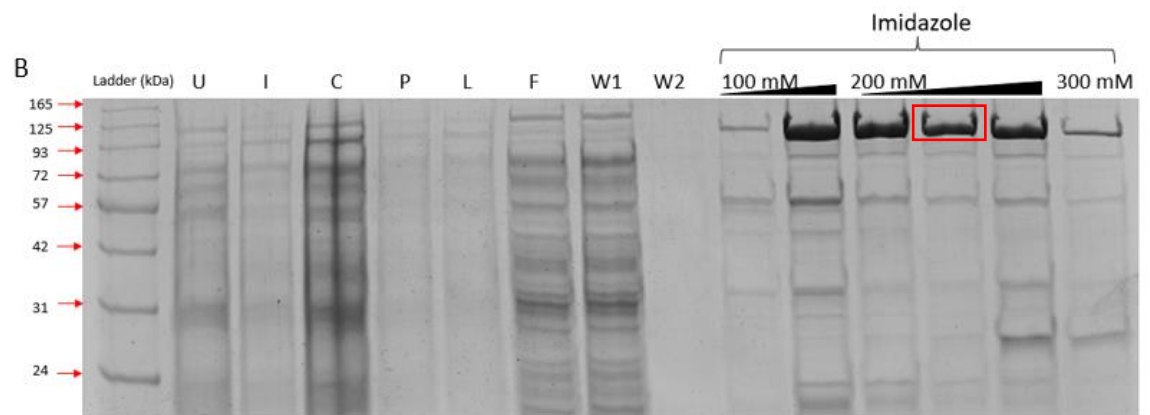
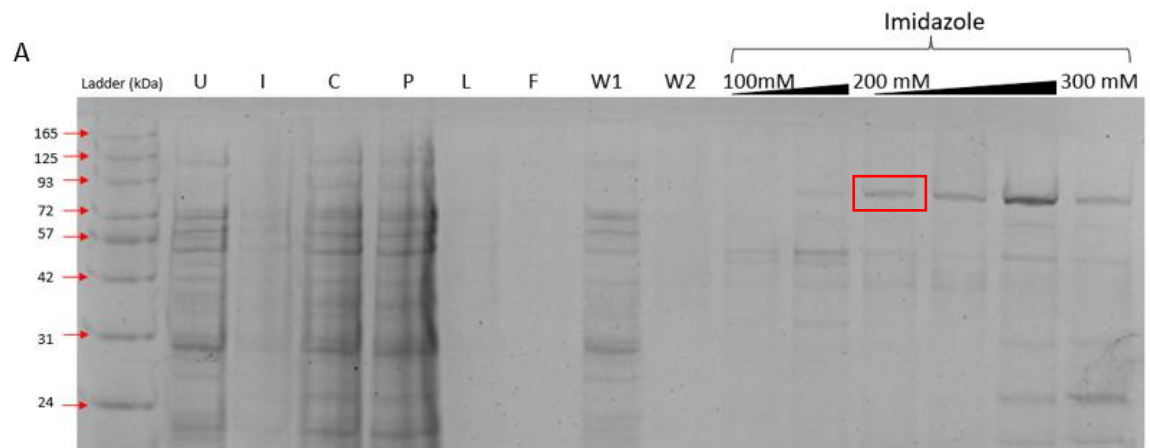
Appendix

Table 3. Golden Gate primers containing Bsal type IIS enzyme cassette with target (capitalized) from 5'-3'.

Transversion	Sense (5'-3')	Antisense (5'-3')
1	gatccGGCCTCCCCAAAGCCTGGCCCCGGGAGTc	tcgagACTCCCGGGCCAGGCTTTGGGGAGGCCg
2	gatccGGCCTCCCCAAAGCCTGGCAAGGGAGTc	tcgagACTCCCTTGCCAGGCTTTGGGGAGGCCg
3	gatccGGCCTCCCCAAAGCCTGGACAGGGAGTc	tcgagACTCCCTGTCCAGGCTTTGGGGAGGCCg
4	gatccGGCCTCCCCAAAGCCTGTCCAGGGAGTc	tcgagACTCCCTGGACAGGCTTTGGGGAGGCCg
5	gatccGGCCTCCCCAAAGCCTTGCCAGGGAGTc	tcgagACTCCCTGGCAAGGCTTTGGGGAGGCCg
6	gatccGGCCTCCCCAAAGCCGGGCCAGGGAGTc	tcgagACTCCCTGGCCCCGGCTTTGGGGAGGCCg
7	gatccGGCCTCCCCAAAGCATGGCCAGGGAGTc	tcgagACTCCCTGGCCATGCTTTGGGGAGGCCg
8	gatccGGCCTCCCCAAAGACTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGTCTTTGGGGAGGCCg
9	gatccGGCCTCCCCAAATCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGATTTGGGGAGGCCg
10	gatccGGCCTCCCCAACGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCGTTGGGGAGGCCg
11	gatccGGCCTCCCCACAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTGTGGGGAGGCCg
12	gatccGGCCTCCCCAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTGGGGAGGCCg
13	gatccGGCCTCCCCAAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTTGGGAGGCCg
14	gatccGGCCTCCACAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGTGGAGGCCg
15	gatccGGCCTACCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGTGAGGCCg
16	gatccGGCCTACCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGGTAGGCCg
17	gatccGGCCGCCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGGGCGGCCg
18	gatccGGCATCCCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGGGATGCCg
19	gatccGGACTCCCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGGGAGTCCg
20	gatccGTCCTCCCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGGGAGGACg
21	gatccTGCCTCCCCAAAGCCTGGCCAGGGAGTc	tcgagACTCCCTGGCCAGGCTTTGGGGAGGCAg

Table 4. Primers used for Golden Gate mutagenesis for RuvC active site variants. Red letters represent mutated base.

Description	Sequence (5'-3')
E477D Forward	CATTATCGA C CTGGCGC
E477D Reverse	G GTCGATAATGATATCGTTCCG
H701K Forward	CAAGCAC AA AGCGGAAGATG
H701K Reverse	CCG TTT GTGCTTGTAACC
D704E Forward	CGGAAGA A GCCTGATTATC
D704E Reverse	G CT TCTTCCGCGTGG



U – Uninduced	L – Loaded supernatant
I – Induced	F – Flowthrough
C – Cell lysate	W1 – Wash 1 (Binding Buffer – 10 mM)
P – Cell lysate pelleted	W2 – Wash 2 (Elution Buffer – 50 mM)

Figure 23. Protein Purification on a 8% SDS-PAGE gel. Red boxed band represents (A) SaCas9[WT] (B) SaCas9[D10E].

Curriculum Vitae

Name: Claire Zhang

Post-Secondary Education and Degrees: The University of Western Ontario, London, ON, Canada 2021-2023
Department of Biochemistry
Supervisor: Dr. David Edgell

The University of Western Ontario, London, ON, Canada 2017-2021
Honours Specialization in Biochemistry and Medical Biophysics

Graduate Course: BIOCHEM 9545Q: Bioinformatics I
BIOCHEM 9546R: Bioinformatics II
BIOCHEM 9501: Scientific Writing

Honours and Awards Mitacs Accelerate 2021

Dean's Honor List 2019-2021

Western Scholarship of Excellence 2017
\$2000

Extracurricular Activities: Social Media Team Member 2021-2023
RIOT, London, ON, Canada

Discovery Day Instructor May 2023
University of Western Ontario, London, ON, Canada

Publications (in preparation) Grainger, Ryan A., Buzon, Beverlee., Zhang, Claire., Medeiros, Braeden., Blue, Tyler., and Junop, Murray S. SNM1/Pso2 acts as a dual-function nuclease in interstrand crosslink repair.

Relevant Work Experience Undergraduate Researcher 2017-2021
University of Western Ontario, London, ON, Canada
Supervisor: Dr. Murray Junop