Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations | Graduate Studies

12-2022

Functional Characterization of the Newly Discovered Type V CRISPR-Cas Protein Cas12a2

Dylan J. Keiser
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

 Part of the Biochemistry Commons, and the Chemistry Commons

Recommended Citation

Keiser, Dylan J., "Functional Characterization of the Newly Discovered Type V CRISPR-Cas Protein Cas12a2" (2022). *All Graduate Theses and Dissertations*. 8664.
https://digitalcommons.usu.edu/etd/8664

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.

UtahStateUniversity
MERRILL-CAZIER LIBRARY

FUNCTIONAL CHARACTERIZATION OF THE NEWLY DISCOVERED TYPE V CRISPR-CAS PROTEIN CAS12A2

by

Dylan J. Keiser

A Thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Biochemistry

Approved:

_____	_____
Ryan N Jackson, Ph.D. Major Professor	Sean J. Johnson, Ph.D. Committee Member
_____	_____
Nick E. Dickenson, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Vice Provost of Graduate Studies

UTAH STATE UNIVERSTIY
Logan, Utah

2022

Copyright © Dylan J. Keiser, 2022

All Rights Reserved

ABSTRACT

Functional Characterization of the Newly Discovered Type V CRISPR-Cas Protein Cas12a2

by

Dylan J. Keiser, Master of Science

Utah State University, 2022

Major Professor: Dr. Ryan Jackson
Department: Chemistry and Biochemistry

The evolution of microbial immune systems as a result of selective pressure from frequent infections by bacteriophage has led to a vast diversity of microbial immune system function. Research into the mechanisms of these immune systems has led to advancements in Biotechnology. Our understanding of the restriction modification immune system led to advancement in molecular cloning, which helped facilitate the practice of recombinant protein expression. Our understanding of the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) immune system led to advancements in gene editing, viral diagnostics and many others. As we discover additional microbial immune systems, our studies of them could lead to similar advancements. Here work is presented on two uncharacterized CRISPR systems. First, the type IV-B CRISPR system, which lacks a canonical CRISPR array and contains an ancillary gene *cysH*. The hypothesis for the function of the type IV-B CRISPR system is as an anti-CRISPR system that inactivates other CRISPR systems. By designing a plasmid clearance assay, the groundwork for others in the Jackson Lab to test this hypothesis was set. Through sequence analysis, potential functions of ancillary protein CysH were identified.

Specifically, evidence of a potential pyrophosphatase active site and the absence of chromosomal CysH active site residues was found. These discoveries could mean that CysH functions through adenylation of nucleic acids. Second, presented is the characterization of the type V-A2 CRISPR System. By optimizing the expression and purification of effector protein Cas12a2, adequate protein was provided for biochemical assays that led to its characterization. The biochemical assays presented here show Cas12a2 functions through targeting of ssRNA. The binding of the ssRNA target is accompanied by indiscriminate nuclease activity *in trans* that leads to degradation of nonspecific ssDNA, ssRNA and dsDNA. The activity of Cas12a2 culminates in an abortive infection type mechanism of immunity. The findings here showcase a novel mechanism of microbial immunity that could be harnessed in Biotechnology to improve upon existing viral diagnostic tools, as well as offer addition use cases in areas of antivirals and cancer therapeutics.

(118 pages)

PUBLIC ABSTRACT

Functional Characterization of the Newly Discovered

Type V CRISPR-Cas Protein Cas12a2

Dylan J. Keiser

Similarly to people, bacteria are under the treat of infection by viruses. To circumvent these threats, bacteria evolve complex immune systems. Our understanding of some of these immune systems has led to many advancements in the field of Biotechnology including tools that made expressing proteins for study in a lab easier, tools that revolutionized the feasibility of gene editing, and tools that could change the way we think about viral diagnostics and cancer therapeutics. A certain type of immune system that bacteria use to fight virus is called a CRISPR system. Presented here is work to understand the function of two CRISPR immune systems with unknown functions. Through sequence analysis and the design of a plasmid clearance assay, a potential function of the type IV-B CRISPR system is proposed. In addition to the type IV-B CRISPR system, biochemical assays are used to determine the function of the type V-A2 protein Cas12a2. It is revealed that Cas12a2 functions using a unique mechanism of immunity that could be harnessed to improve tools in viral diagnostics, viral therapeutics, and cancer therapeutics.

ACKNOWLEDGMENTS

There are so many people that have been supportive and encouraging. My Girlfriend Sarah has been especially patient, encouraging and supportive as I finish up my schooling, so that we can move onto the next chapter of our lives. My parents and siblings have always had faith in me and were very encouraging through my entire education. Chris Monson, my undergraduate research advisor from SUU, taught me about research and got me excited about a career in it. Bruce Howard, from SUU, introduced me to the field of biochemistry and got me excited about a possible career in biochemistry. I'd also like to acknowledge my committee members Nick Dickenson and Sean Johnson for their support and advice. I'd especially like to acknowledge my major professor Ryan Jackson for helping me see my potential and being so encouraging and supportive. Funding for this research was provided by the NIH.

Dylan Keiser

CONTENTS

Page

Abstract iii

Public abstract v

Acknowledgements vi

Contents vii

List of tables ix

List of figures x

Introduction and Literature Review 1

- Bacteria and Phage are Locked in an Evolutionary Arms Race 1
- The Study of Microbial Immune Systems has led to Advancements in Biotechnology 1
- Diversity of Bacterial Immunity 2
- Abortive Infection Mechanisms 3
- CRISPR Systems are Adaptive Bacterial Immune Systems 5
- The Diversity of CRISPR-Cas Systems 7
- The Type IV-B CRISPR System is a CRISPR-Less System Found on Plasmid 9
- Cas12a2 is a Type V CRISPR Cas Protein Most Closely Related to Cas12a 10
- Future Directions 12

Functional Analysis of Novel Crispr Systems 14

- Abstract 14
- Introduction 15
- Results and Discussion 15
- Conclusion 33
- Methods 35

Functional Characterization of The Newly Discovered Type V Effector Protein Cas12a2 43

- Abstract 43
- Introduction 44
- Results and Discussion 46
- Conclusion 64

Methods....66

Summary and Future Directions....72

Abstract....72
Summary....73
Future Directions....74
Conclusion....80

References....81

Appendices....90

Curriculum Vitae....106

LIST OF TABLES

Page

1-1 Applications of Bacterial Immune Systems 2

1-2 Diversity of Bacterial Immune Systems 5

1-3 Diversity of CRISPR Systems 8

3-1 Comparison of Collateral Nuclease Activity Among CRISPR Systems 56

B-1 Comprehensive List of Plasmids Used 100

B-2 Comprehensive List of Oligonucleotides used in Cleavage Assays 103

LIST OF FIGURES

Page

1-1 The Mechanism of CRISPR Immunity 6

1-2 Type IV-B CRISPR Systems May Function as Potential Anti-CRISPR Systems 10

1-3 Cas12a2 is Most Closely Related to Cas12a 12

2-1 Analysis of The Function of Type IV-B CRISPR Systems.....17

2-2 CysH Sequence Alignments 22

2-3 Expression, Purification and in vitro Cleavage Assay of SmCas12a 24

2-4 Baby Hamster Kidney Cell Culture and Genome Editing.....26

2-5 BHK Transfection by with SuCas12a2 29

2-6 Improvement of Initial Expression and Purification Protocol of SuCas12a2 31

3-1 Expression and Purification of SuCas12a2 47

3-2 RNA Targeting 49

3-3 Accessibility of Cas12a2 Bound RNA Target for The Activation of an Additional Molecule of Cas12a2 51

3-4 Collateral Degradation.....54

3-5 A Comparison of Collateral Activity Between CRISPR Systems 57

3-6 Cas12a2 Requires PFS for Activation 60

3-7 Determination of Residues Important for Cas12a2 Function.....62

3-8 Dependence of Divalent Cation on SuCas12a2 Activity.....64

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Bacteria and Phage are Locked in an Evolutionary Arms Race

Viral infections are often thought of solely in regard to their impact on humans and other complex organisms. Less known is the fact that certain viruses infect microbes, including bacteria [1]–[3]. The survival of bacteria amongst these viruses, which out-number bacteria 10 to 1 [4]–[7], depends on the ability of bacterial immune systems to defend against viral infection. The evolution of these immune systems is the result of generations of selective pressure in which only the bacteria with immune systems capable of defending the viral threat survive. However, such selective pressure is not exclusive to the bacteria in this equation. Viruses must in turn evolve mechanisms to evade or disable bacterial immune systems to propagate their genes through the population [8], [9]. Selective pressure on both sides of the bacteria vs viruses equation has developed into an evolutionary arms race that has ultimately led to a vast diversity of bacterial immune systems and viral anti-bacterial immune system counterparts.

The Study of Microbial Immune Systems has led to Advancements in Biotechnology

Understanding how bacterial immune systems function has led to major discoveries in science and technology (**Table 1-1).** For instance, an understanding of Restriction Modification systems has led to advancements in molecular cloning [10]. Restriction enzymes coupled with a repurposing of phage enzymes such as T4 ligase allows for the creation of custom vectors that can be used for recombinant gene expression [10], [11]. Toxin-antitoxin systems improve molecular cloning by allowing

for an alternative way to select for the plasmid of interest without an over reliance on antibiotic selection [12]. The discovery and characterization of CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR Associated) systems has led to advances in genome editing, nucleic acid detection, cancer therapeutics, and regulation of gene expression [13]–[19]. Despite all these discoveries and advancements in biotechnology, there is a vast diversity of uncharacterized immune systems suggesting other mechanisms for disabling viruses exist that could potentially be repurposed for use in science and medicine.

Table 1-1: Applications of Bacterial Immune Systems

Immune System	Application(s)	Reference(s)
Restriction Modification Systems	Molecular cloning	[10]
Viral Anti-Defense Enzymes	Molecular cloning, Recombinant gene expression	[11]
Toxin-Antitoxin	Molecular cloning	[12]
CRISPR Cas9	Gene editing	[14]–[16], [20]
CRISPR Cas12a	Gene editing, Nucleic acid detection	[17]
CRISPR Cas13a	Nucleic acid detection, Cancer therapeutics, Regulation of gene expression	[18], [19]

Diversity of Bacterial Immunity

The evolutionary arms race between bacteria and viruses has led to a vast diversity of bacterial immune systems that have many functional differences. However, each immune system shares the goal of preserving enough of the bacterial population for genetic propagation. The mechanism of bacterial immunity can be classified under three categories: Degradation of phage nucleic acids, abortive infection, and inhibition of DNA and RNA synthesis [21] (**Table 1-2)**. The degradation of phage nucleic acids is done by

immune systems such as Restriction modification systems and the CRISPR-Cas systems. These systems both use sequence specific nuclease activity to cleave and eliminate foreign nucleic acid. At the core however, CRISPR-Cas systems and Restriction modification systems are very different. The Restriction modification system is classified as an innate immune system, while the CRISPR-Cas system is classified as an adaptive immune system. Innate immune systems target a wide variety of viruses using a shotgun approach, and they don't rely on memories of previous infection for immunity. The restriction modification system methylates host DNA to deter restriction nucleases from targeting host DNA. Invading nucleic acids are left unmethylated and therefore vulnerable to restriction nuclease activity [5]. The Restriction modification systems use a variety of restriction enzymes, that each target a unique sequence, to increase the probability of targeting a sequence found in the phage genome. CRISPR-Cas systems utilize a record keeping approach to recall sequence information from previous infections to be programmed into the immune system and used against subsequent infections [7]. As stated earlier, the goal of each immune systems is the preservation of bacteria for the propagation of genes. Genetic propagation can be maintained by either eliminating the infection before it harms its host or by sacrificing the host before the virus replicates and infects another host. Immune systems that result in the sacrifice of the host cell for the benefit of the population are called abortive infection systems.

Abortive Infection Mechanisms

Cell sacrifice can be achieved using a variety of mechanisms such as a compromise of membrane integrity, inhibition of protein synthesis, protein phosphorylation, and degradation of host nucleic acids [6]. The RexAB and AbiZ system

works by depolarizing the membrane as well as prematurely lysing the membrane upon recognition of the infection [6], [22]. The PrrC system is activated upon the inhibition of restriction enzyme EcoprrI by T4 phage and halts protein synthesis within the cell via the cleavage of Lys tRNAs [6], [23]. The Stk2 system recognizes the PacK phage protein and Stk2 is auto phosphorylated. Phosphorylation of Stk2 activates its kinase activity that leads to the phosphorylation of host cell proteins and eventually cell death [6], [24]. Toxin-antitoxin systems are systems in which both a protein toxic to the bacteria and the toxic protein's inhibitor are both expressed in the cell. An example of this is the RnlAB system. Under infection, expression of the RnlB antitoxin is inhibited and the unstable RnlB protein is quickly degraded [6], [25]. Degradation of RnlB results in the activation of RnlA, an endoribonuclease that degrades both phage and host cell RNA [6], [25]. CBASS systems produce cyclic oligonucleotides upon recognition of infection that are used to activate downstream proteins [6], [26]. The downstream effector proteins range from membrane ion channel proteins, to DNases and phospholipases [6]. Activation of the effector protein results in the demise of the host cell [6]. An example of an immune system that functions by inhibiting DNA and RNA synthesis of the virus are prokaryotic viperins. Viperins are proteins that synthesize modified nucleotides that act as chain terminators that inhibit DNA and RNA synthesis when incorporated into the growing nucleic acid strand [27]. Each of these immune systems is the result of generations of selective pressure to continue to adapt and survive in an environment of ongoing infection by phage. The vast diversity of immune systems this has led to is the subject of a lot of ongoing research. However, the focus of this thesis will be on CRISPR-Cas systems.

Table 1-2: Diversity of Bacterial Immune Systems		
Mechanism of Action	**Immune System(s)**	**Reference(s)**
Degradation of phage nucleic acid	CRISPR-Cas, Restriction modification	[5], [7]
Abortive Infection	CRISPR-Cas, RexAB/ABiZ, PrrC/Ecoprrl, RnlAB, PacK/Stk2, CBASS	[6], [22]–[26]
Inhibition of DNA and RNA Synthesis	Prokaryotic Viperins	[27]

CRISPR Systems are Adaptive Bacterial Immune Systems

CRISPR systems are adaptive bacterial immune systems [7]. The term CRISPR is an acronym for Clustered Regularly Interspaced Short Palindromic Repeats, and it refers to sequences of DNA, that contain short repeating segments that are interspaced by short, spacer sequences of DNA. The spacer sequences are identical to sequences from invading phage or other mobile genetic elements [28], [29]. The CRISPR is found adjacent to CRISPR associated (Cas) genes that code for the proteins essential for carrying out the function of the immune system [30]. There are several different types of CRISPR systems that have variances in their mechanisms of action [31]. However, each type executes immunity three generalized stages: adaptation, biogenesis, and interference. In adaptation foreign nucleic acid is recognized and a short segment from it is integrated into the CRISPR in the form of a spacer [32] (**Figure 1-1**). Each spacer serves as a memory of a previous infection and is the basis for the adaptive nature of the immune system. In biogenesis, the CRISPR sequence is transcribed into a long RNA molecule called a pre-CRISPR RNA (pre-crRNA) that is cleaved by ribonuclease activity into mature CRISPR

RNAs (crRNA) [33], [34] (**Figure 1-1**). Cas enzyme(s) bind to the crRNAs to form ribonucleoprotein complexes. In interference, the ribonucleoprotein complex recognizes invading nucleic acid complementary to the spacer within the crRNA. Self vs foreign nucleic acid is distinguished through the recognition of short 3-5 nucleotide sequences found adjacent to the target sequence called Protospacer Adjacent Motifs (PAMs). PAM sequences are not found adjacent to spacer sequences in the genome and thus are not targeted by the immune system [35]. Upon successful recognition of the foreign nucleic acid target, the interference complex cleaves/degrades the invading nucleic acid, eliminating it from the cell [36] (**Figure 1-1**).

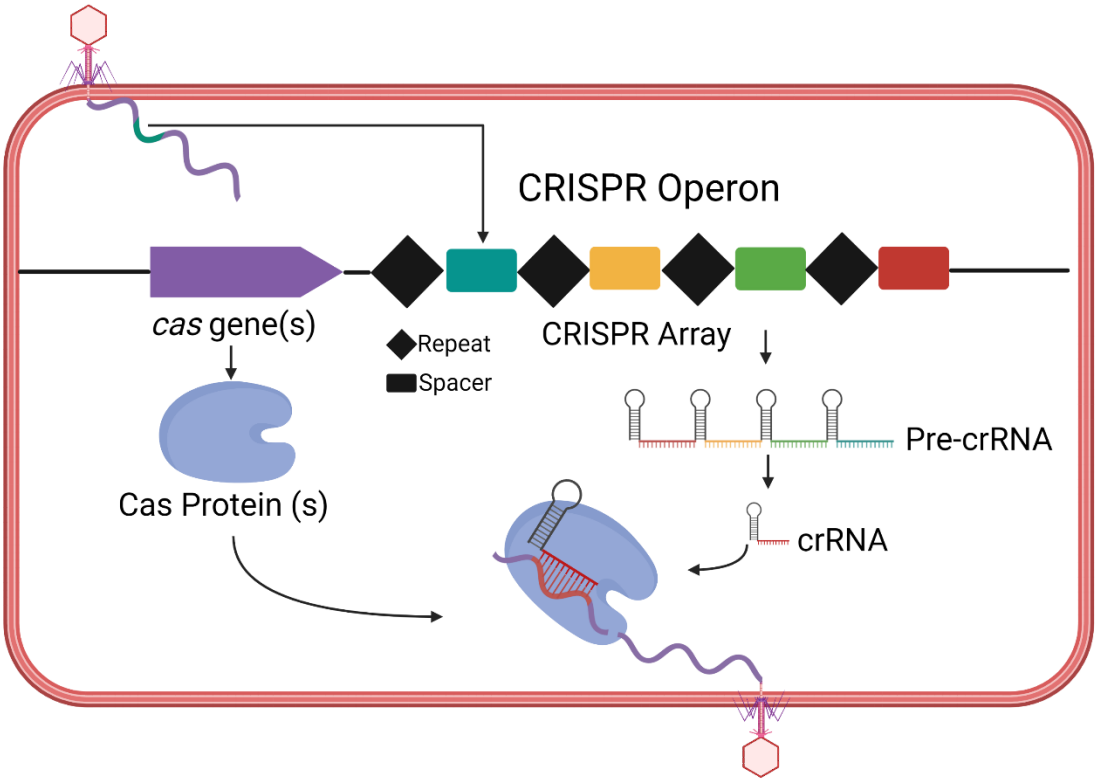


Figure 1-1 The Mechanism of CRISPR Immunity. CRISPR systems work first by integrating foreign nucleic acid into the host cell genome as a spacer within the CRISPR array. Transcription and RNA processing then creates small RNAs called CRISPR RNAs (crRNA) that are bound by CRISPR Associated (Cas) protein(s). Recognition of foreign nucleic acid complementary to the crRNA activates cleavage of the foreign nucleic acid by the Cas protein(s).

The Diversity of CRISPR-Cas Systems

The evolution of CRISPR systems has led to a highly diverse set of systems with different functionalities and specificities. There are six major types of CRISPR systems and many subtypes that fall under either class 1 or class 2 systems [31]. Class 1 systems function through the use of multiple proteins coming together to form a single effector complex [31]. Class 2 systems function through the use of a single protein acting as the effector complex [31]. The mechanism of immunity of CRISPR-Cas systems can be divided into systems that either target DNA or target RNA (**Table 1-3**). Type I systems involve the formation of multi-protein complex, Cascade, and following recruitment of Cas3 foreign DNA is targeted [37]. In the type II system, Cas9 forms a complex with crRNA and trans-activating RNA (tracrRNA) and also targets foreign DNA [38]. In the majority of type V systems, the effector protein, Cas12 forms a complex that targets DNA. Upon recognition of its target sequence, Cas12 often initiates indiscriminate nuclease activity that degrades non-specific ssDNA [17], [39]. CRISPR systems are not limited to targeting DNA, as some systems also target RNA.

The type V-G system utilizes a complex formed between Cas12g, a crRNA and a tracrRNA to form a complex that targets RNA. After recognition of its RNA target, Cas12g initiates indiscriminate nuclease activity to degrade non-specific ssDNA and

RNA [40]. Type III systems work through the formation of multi-protein Csm/Cmr complexes along with Cas10. A DNA target is transcribed by RNA polymerase to form an RNA molecule that is recognized by the Csm/Cmr and cleaved. Cas10 then cleaves the non-template strand of DNA [41]. An additional activity of Cas10 is the production of cyclic adenylates. These cyclic adenylates activate accessory nucleases that indiscriminately degrade non-specific RNA [42]. In the type VI system Cas13a recognizes and cleaves target RNA molecules. Then, recognition of RNA target sequences results in indiscriminate degradation of non-specific RNAs [43]. Recently it has been discovered that certain CRISPR-Cas systems such as the type III and type VI systems exhibit an abortive infection phenotype that comes as a result of indiscriminate RNAse activty [6], [44]–[47]. There are still many CRISPR systems to be discovered and understood. For instance, the function of the type IV-B system is unknown, largely due to the absence of a CRISPR within the operon. There also exists a newly discovered type V system, with effector protein Cas12a2 that has an unknown function. This thesis explores the function of type IV-B and type V-Cas12a2 CRISPR systems. It is expected that my efforts to characterize the type IV-B and the type V-Cas12a2 system will lead to knowledge that will be repurposed in biotechnology.

Table 1-3: Diversity of CRISPR Systems

Mechanism	CRISPR-Cas System(s)	Reference(s)
DNA targeting	Type I, Type II, Type III, Type V-A,B,C,E,F,H,I,K,J	[17], [37], [38], [41], [48]
RNA targeting	Type III, Type V-G, Type VI	[40], [41], [43], [49]

The Type IV-B CRISPR System is a CRISPR-Less System Found on Plasmids

The type IV-B system is found primarily on plasmids and lacks a CRISPR array required for immune system function [50], [51]. There also exists an ancillary gene found adjacent to the type IV-B operon; a variation of the *cysH* gene found in the sulfate assimilation pathway [51], [52]. The overall function of the type IV-B system is unknown. Given that the system lacks a CRISPR and is found on plasmids, we hypothesize that a possible function could be as an anti-CRISPR system. With all of the protein components required to form a complex around the crRNAs [53], the type IV-B system could function by soaking up crRNA from other systems, essentially rendering those systems inactive (**Figure 1-2)**. In chapter 2 I discuss the experimentation and preliminary data obtained for the characterization of the type IV-B system. I utilize plasmid clearance assays and bioinformatic analysis of CysH in order to ascertain a possible function of the type IV-B system. By using the type IV-A CRISPR system a host immune system, we discover a flaw in the reproducibility of the plasmid clearance assay that was later determined to be due to an inferior PAM sequence in the type IV-A system. In order to determine the anti-CRISPR capabilities of the type IV-B system, the plasmid clearance assay will need to be repeated with a more optimized PAM sequence. As for Cas-CysH, we show evidence for a possible lack of sulfate reductase functionality with instead a possibility for pyrophosphatase activity. This could be a key finding in developing possible activity assays for Cas-CysH.

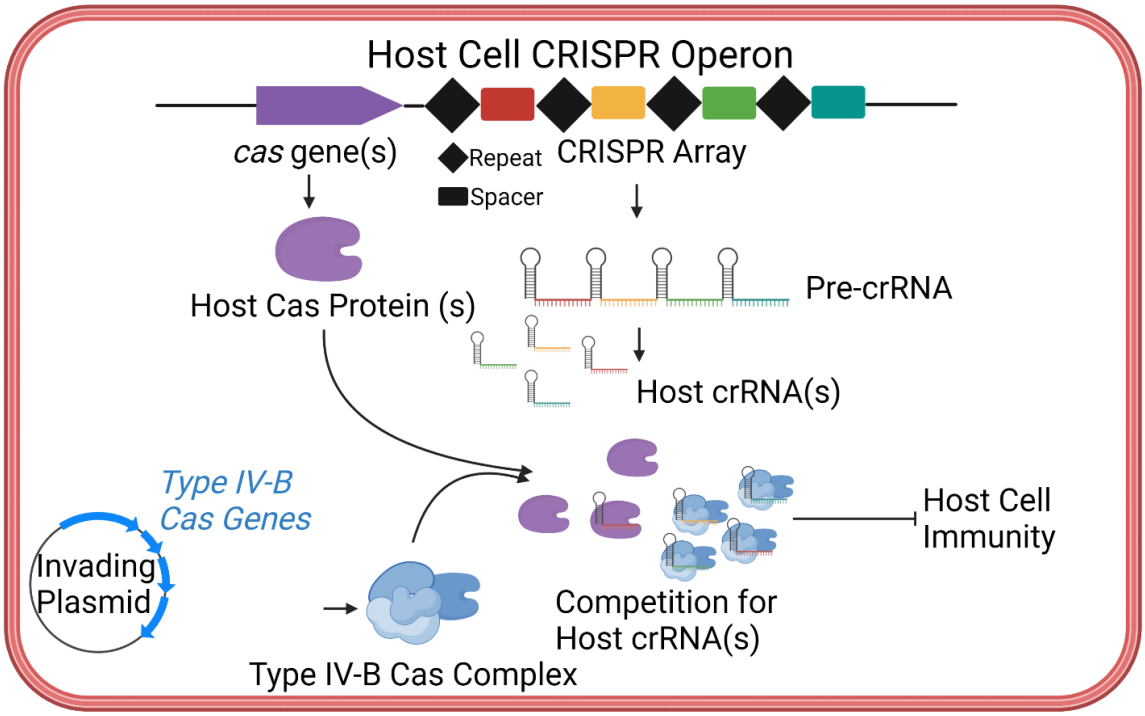


Figure 1-2 Type IV-B CRISPR System May Function as Potential Anti-CRISPR Systems. A possible function of the type IV-B CRISPR system may be to soak up host cell crRNAs to render the host cell CRISPR system unable to perform its functions. Therefore, allowing the foreign plasmid to propagate.

Cas12a2 is a Type V CRISPR Cas Protein Most Closely Related to Cas12a

Type V CRISPR systems are highly diverse [48]. The most well understood system is the type V-A system that utilizes the effector protein Cas12a. Cas12a utilizes a RuvC nuclease active site to cleave dsDNA [17]. In some type V-A systems, after Cas12a binds its DNA target, it exhibits indiscriminate nuclease activity that degrades non-specific ssDNA [39]. A newly discovered type V CRISPR-Cas protein, Cas12a2 is most closely related to Cas12a, although, Cas12a2 only shares 10-20 percent sequence identity with Cas12a [54]. Phylogenetic analysis of Cas12a2 reveals that Cas12a2 falls on

its own separate clade indicating an evolutionary distinct protein **(Figure 1-3)**. Other known type V effector proteins also fall on individually distinct clades. Phylogenetic analysis also revealed frequent occurrences of Cas12a and Cas12a2 in the same CRISPR operon suggesting a potentially synergistic functionality **(Figure 1-3)** [54]. Sequence comparison between Cas12a and Cas12a2 reveals both proteins contain a RuvC domain (**Figure 1-3)**. However, Cas12a2 contains a Domain of Unknown Function (DUF) as well as a zinc finger motif that are not present in Cas12a **(Figure 1-3)**. The Nuc domain found in Cas12a is also not present in Cas12a2 **(Figure 1-3)**. These sequence differences suggested to us that Cas12a2 would have a biochemical function distinct from Cas12a. In chapter 2, optimization of protein purification, and the genome editing potential of Cas12a2 are discussed. We discover a possibility of targeted cell death in mammalian cells that could be repurposed for therapeutics. In chapter three, we show that Cas12a2 functions as an RNA targeting enzyme with a robust collateral nuclease activity that includes non-specific substrates such as ssDNA, dsDNA, and ssRNA. The collateral nuclease activity of Cas12a2 ultimately results in the death of the cell, indicating the discovery of a new CRISPR system with an abortive infection mechanism.

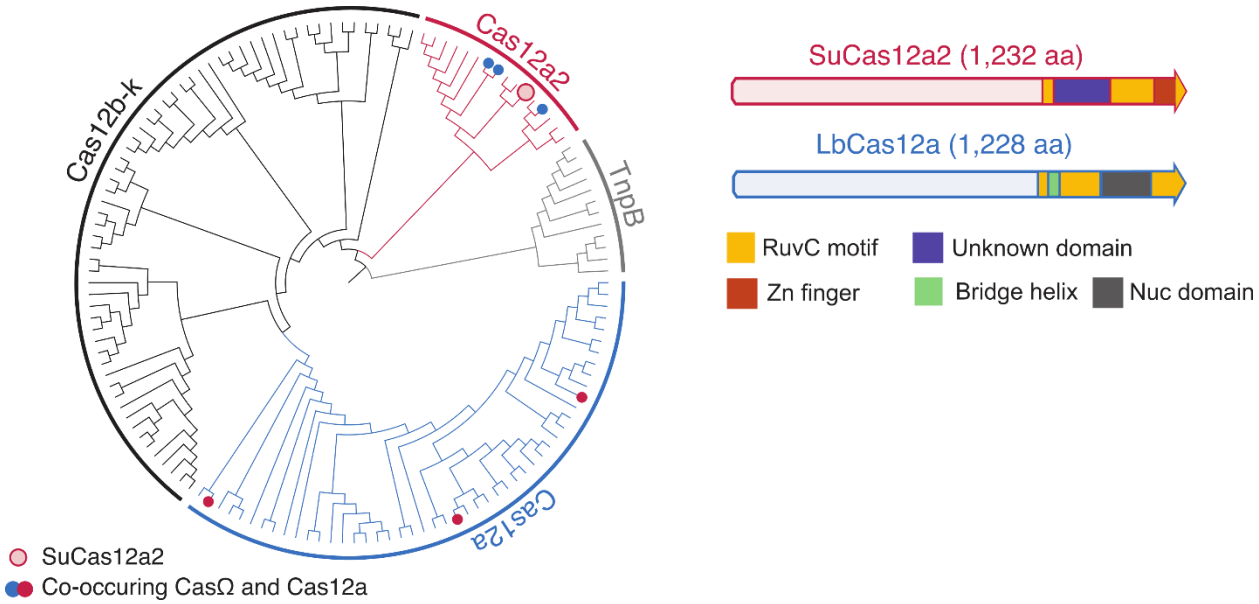


Figure 1-3 Cas12a2 is Most Closely Related to Cas12a. Phylogenetic analysis of Cas12a2 and comparison with other type V Cas sequences results in Cas12a2 falling on its own separate clade (left). Occasional co-occurrences of Cas12a and Cas12a2 in the same organism were also found (left, blue dots). A comparison of the amino acid sequences between Cas12a and Cas12a2 reveals both contain RuvC domains. However, Cas12a2 contains a domain of unknown function (DUF) and a zinc finger motif that is not present in Cas12a, and Cas12a contains a NUC domain that is not present in Cas12a2 (right). (Data and figure provided by Chase Beisel and Oleg Dmytrenko)

Future Directions

Of everything learned about the type IV-B and type V-A variant systems there is still much to discover. Gaining an understanding of the uncharacterized type IV-B and type V-A2 immune systems would not only lead to a better understanding of microbial immune systems, but it may also lead to advancements in biotechnology; much like the discovery of the mechanism of Cas9 led to a repurposing of it as a genome editor [14]–[16], [20]. Since, then research into other CRISPR systems has led to further advancements and different use cases for CRISPR based biotechnological tools[55], [56].

By learning more about different CRISPR systems, more tools may be added to the biotechnology toolbox, and there remains a lot to be uncovered. In chapter 4, we discuss the next steps and possible future directions for these projects. For the type IV-B system, this includes a discussion on potential improvements to the plasmid clearance assay as well as potential assays for the analysis of Cas-CysH function. For Cas12a2, we discuss any unanswered questions in terms of its mechanism as well as the development of any potential applications.

Acknowledgments

Figures 1-1 and 1-2 were created using Biorender.com

CHAPTER 2

FUNCTIONAL ANALYSIS OF NOVEL CRISPR SYSTEMS

Abstract

This chapter reports work I performed on multiple projects in the Jackson Lab before I narrowed down my focus to determining the function of Cas12a2. First, I evaluated the function of the type IV-B CRISPR system as an anti-CRISPR and its ancillary protein utilizing plasmid curing assays and bioinformatics respectively. Second, a bioinformatic analysis of the type IV-B ancillary gene *cysH* led to the discovery of similarities and differences between *cascysH* and non-cas*cysH* genes. Third, I investigated the activity of the Cas12a protein that is encoded with Cas12a2 in the same CRISPR system operon. This investigation revealed that SmCas12a behaves similar to Cas12a from other organisms, specifically as an RNA-guided dsDNA targeting protein. Fourth, I evaluated the genome editing potential of Cas12a2 in mammalian cells. It was determined that Cas12a2 did not induce insertion/deletions (indels) in mammalian cells. However, I did find evidence of Cas12a2 inducing cell death, which supports our current understanding of Cas12a2 activity. Fifth, I describe how I optimized the expression and purification protocol to greatly increase the yield of Cas12a2 after purification. Sixth, I performed initial biochemical evaluation of Cas12a2 involved cleavage assays targeting Cas12a2 to dsDNA. These assays mostly showed inconsistent results that was later solved by switching to assays in which Cas12a2 targeted RNA. The work presented here set the groundwork for various ongoing projects in the Jackson Lab including the project that is the focus of this thesis, the type V CRISPR protein, Cas12a2.

Introduction

In this chapter, I highlight the work I did in various projects. I first describe my efforts to evaluate the function of the type IV-B CRISPR system and its ancillary protein CysH. To do this, I used plasmid curing assays to determine if the IV-B system functions as an anti-CRISPR system and used amino acid sequence analysis to highlight a potential function of CysH. Next, I discuss my initial works aimed at understanding the function of the type V protein Cas12a2, including the optimization of Cas12a2 expression and purification, initial attempts at repurposing Cas12a2 as a gene editor in mammalian cells, and my initial work for determining Cas12a2 activity with *in vitro* assays. I also describe how I successfully determined the activity of the Cas12a protein from *Smithella sp.* (SmCas12a), a protein that occurs adjacent to Cas12a2 in the *Smithella* type V CRISPR system, in an effort to understand why these two nucleases co-occur in the same operon. Finally, I highlight the unfinished areas of my research and give what I believe are the next steps for pursuing these projects.

Results and Discussion

Determining The Function of The Type IV-B CRISPR System

The type IV-B proteins Csf1-Csf4 have previously been shown to form a complex around an RNA molecule [53]. However, the function of the system is unknown. The type IV-B CRISPR system is unique in that it does not contain a CRISPR and occurs frequently on plasmids [50]. A possible function of the IV-B system could be that it works as an anti-CRISPR system. Anti-CRISPR systems combat CRISPR immunity. Invasion of a cell with a plasmid containing a type IV-B system could possibly inactivate the CRISPR system by binding crRNAs from the host cell's immune system with the

type IV-B proteins resulting in a limitation of available crRNAs and an inhibition of CRISPR activity [50]. The type IV-B system is also often found adjacent to a variation of the *cysH* gene, that lacks the typical CysH protein active site. Gaining insight into the function of this CysH variant could be the key to understanding the function of the type IV-B system, assuming their functions are tied together.

To examine if the type IV-B system functions as an anti-CRISPR we inserted the genes of the type IV-B system from *Mycobaterium sp. JS623* into a plasmid containing a target sequence. A plasmid curing assay was conducted that compared Colony Forming Units (CFUs) between cells containing a type IV-A system transformed with target plasmid, and target plasmid with a type IV-B system (**Figure 2-1A**). An increase in the CFUs with the target strand, when in the presence of the type IV-B system, would indicate anti-CRISPR activity.

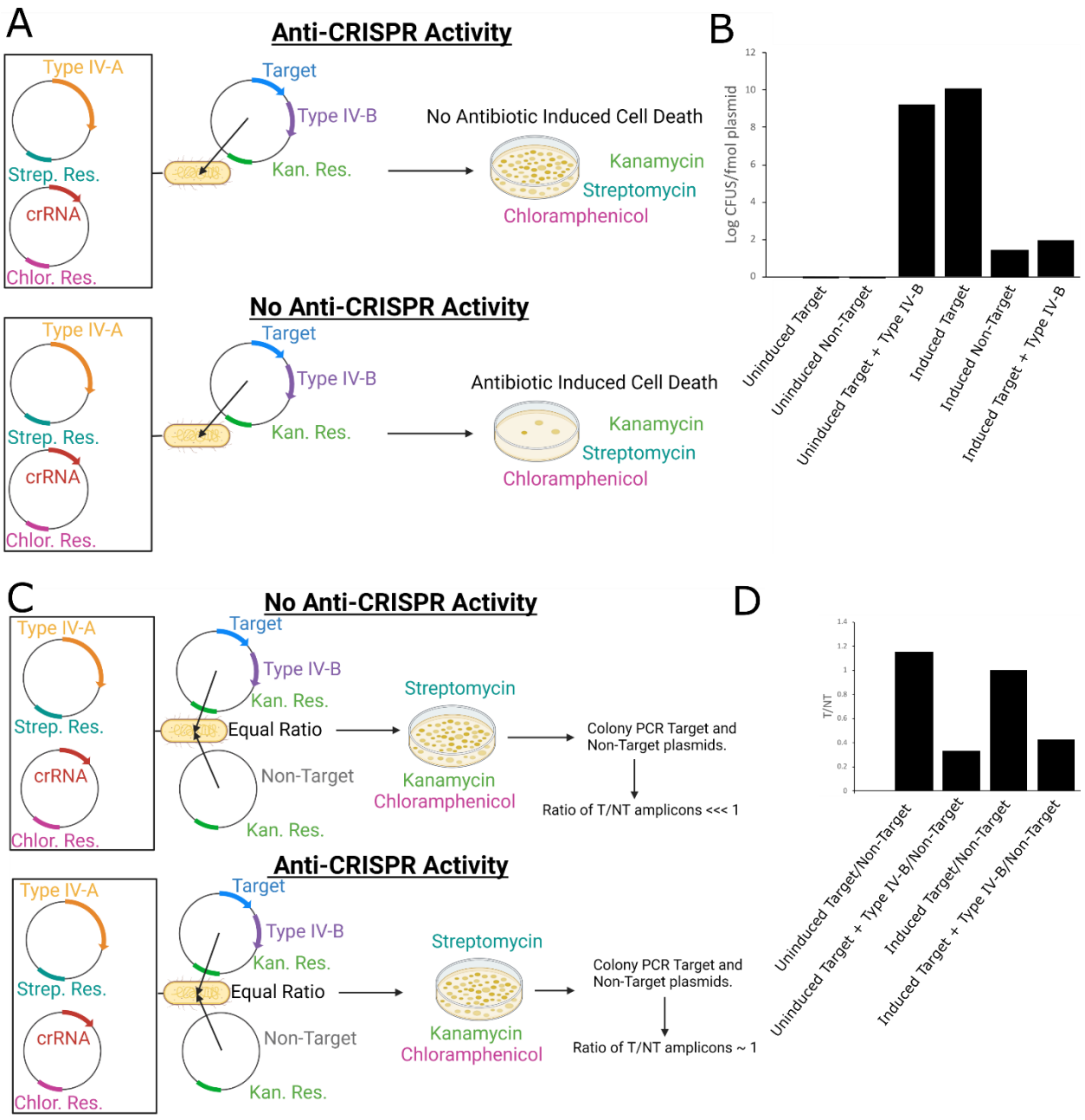

A
Anti-CRISPR Activity
Type IV-A
Strep. Res.
crRNA
Chlor. Res.
Target
Type IV-B
Kan. Res.
No Antibiotic Induced Cell Death
Kanamycin
Streptomycin
Chloramphenicol
No Anti-CRISPR Activity
Antibiotic Induced Cell Death
B
Log CFUS/fmol plasmid
Uninduced Target
Uninduced Non-Target
Uninduced Target + Type IV-B
Induced Target
Induced Non-Target
Induced Target + Type IV-B
C
No Anti-CRISPR Activity
Equal Ratio
Non-Target
Streptomycin
Kanamycin
Chloramphenicol
Colony PCR Target and Non-Target plasmids.
Ratio of T/NT amplicons <<< 1
Anti-CRISPR Activity
Ratio of T/NT amplicons ~ 1
D
T/NT
Uninduced Target/Non-Target
Uninduced Target + Type IV-B/Non-Target
Induced Target/Non-Target
Induced Target + Type IV-B/Non-Target

Figure 2-1. Analysis of The Function of Type IV-B CRISPR Systems. Initial Plasmid curing assay shows inconclusive results. (A) Workflow for CFU based plasmid curing assay. Chemically competent cells containing the type IV-A CRISPR system are transfected with target plasmid containing the type IV-B CRISPR system or with target alone. Cells are then plated with anti-biotics and counted. The presence of anti-CRISPR activity would result in the retention of the target plasmid and survival in kanamycin (top). A lack of anti-CRISPR activity for the type IV-B system would result in the destruction of the target plasmid and loss of kanamycin resistance, which would result in low CFUs (bottom). (B) initial data, represented as a bar graph for CFU based plasmid curing assay. The Y-axis indicates LogCFUs/fmol plasmid transformed. The different transformations are noted below the X-axis. (C) Workflow for PCR based plasmid curing assay. Chemically competent cells containing the type IV-A CRISPR system are transformed with equal ratios of target + type IV-B/non-target plasmid. Cells were plated then a series of colony PCR reactions were used to determine the ratio of transformed plasmids. The absence of anti-CRISPR activity would result in a target + type IV-B plasmid/non-target plasmid ratio of much less than 1 (top). The presence of anti-CRISPR activity would be indicated by a ratio of target + type IV-B plasmid/non-target plasmid of 1 (bottom). (D) Initial data for PCR based plasmid curing assay. The Y-axis shows the ratio of target/non-target plasmid found with colony PCR. The transformations are noted below the X-axis. (A,C)

A variation of the plasmid curing assay was conducted in which cells containing the type IV-A system were transfected with both a non-target plasmid and a target plasmid or a non-target plasmid and target plasmid with the type IV-B system (**Figure 2-1C**). By transforming equal ratios of target and non-target plasmid, colony PCR should show an equal number of cells containing the target as the non-target under non-targeting conditions. If targeting is occurring, the proportion of cells containing the non-target plasmid should be significantly greater than the proportion of cells containing the target plasmid as a result of cleavage of the target plasmid by the immune system. The addition of the type IV-B system would reduce the amount of target plasmids cleaved if it was able to function as an anti-CRISPR system (**Figure 2-1C**).

The preliminary data for the plasmid curing assay highlight a pitfall of the type IV-A CRISPR plasmid curing assay in terms of its reproducibility. Unfortunately, outside

of Valerie Crowley's hands the plasmid curing assay was inconsistent. The controls with target and non-target plasmid showed a significant reduction of CFUs for the non-target plasmid, indicating that the immune system was not functioning as shown before [57], as a functioning immune system would show less CFUs for the target relative to the non-target (**Figure 2-1B**). The results from these controls render the results from the experiment with the type IV-B system inconclusive. However, the results did highlight the need to develop a more reproducible type IV-A CRISPR plasmid curing assay, which has now been done in the Jackson Lab by Olive Redman, who determined that the PAM sequence we were using was non-optimal. Thus, this assay should be repeated using the target sequence adjacent to an optimal PAM.

The results from the PCR based plasmid curing assay support what was observed with the CFU based plasmid curing assay, which suggest the type IV-A CRISPR immune system is not functioning properly, as there was little difference between the uninduced and induced cells for the target/non-target controls. A functioning immune system would show a target/non-target ratio of much less than one (**Figure 2-1D**). A ratio of approximately one, does indicate that equal amounts of target and non-target plasmids were transformed (**Figure 2-1D**). However, expression of the host type IV-A system did not show targeting as it has before [57]. Transformation with the non-target plasmid and the target plasmid with the type IV-B system showed a much a lower ratio of target plasmid/non-target plasmid in both uninduced and induced conditions (**Figure 2-1D**). A possible explanation for the higher number of cells containing the non-target plasmid could be due to how easily the plasmids are taken up by the cells. The addition of the type IV-B system to the target plasmid increased its size by about 3kb, which puts the

size of the target plasmid with the type IV-B system about 3 kb higher than the non-target plasmid. It is possible that the increase in size reduced its ability to be taken up by cells. A possible way around this would be to insert the type IV-B system into the non-target plasmid as well. Overall, these data show that in order to determine the function of the type IV-B system, a more robust plasmid curing assay using the type IV-A CRISPR system needed to be developed.

With the discovery of a more optimal PAM sequence for the type IV-A system by Olive Redman, evaluating the anti-CRISPR functionality of the type IV-B system using the PCR based plasmid curing assay should work much better now. However, before running the assay, I would suggest that it may be necessary to clone the type IV-B system on the non-target plasmid as well for consistency.

Sequence Analysis of The Type IV-B Ancillary Protein CysH

To better understand the function of the ancillary type IV-B protein CysH, I analyzed the amino acid sequences of type IV-B proteins from several organisms. The results of my work were described in a perspective on the type IV CRISPR system published in Frontiers in Microbiology, in which I am a co-author [51]. We hypothesized that the CysH protein encoded by the type IV-B systems are distinct from the CysH proteins encoded elsewhere in the bacterial genome used for sulfur acquisition. To test this hypothesis another graduate student in the lab, Thom Hallmark and I, independently compared sequences of Cas-CysH (found within type IV-B systems) to other non-Cas-CysH sequences within the same organism.

Notably, the non-Cas-CysH has been previously described as a PAPS/APS reductase integral to the sulfate assimilation pathway [58]–[61], and it functions using a

catalytic cysteine located at its c-terminus [52]. The residues ECGLH in the active site are highly conserved among active CysH PAPS/APS reductases [61]. However, the active site residues ECGLH are not found in Cas-CysH sequences, suggesting a different function (**Figure 2-2A**). Cas-CysH sequences also showed the presence of a possible pyrophosphatase domain with strong conservation of the residues SxGxDS shown to coordinate ATP in pyrophosphatases [62], [63] (**Figure 2-2B**). The presence of these residues was not conserved in non-Cas-CysH sequences (**Figure 2-2B**). A lack of a catalytic cysteine means CysH probably does not function as a PAPS/APS reductase. However, the presence of a pyrophosphatase domain could indicate it functions as a pyrophosphatase, possibly to adenylate nucleic acids.

To evaluate the role CysH plays in the function of the type IV-B system, I propose a two-pronged approach. First, a variation of a plasmid curing assay testing for anti-CRISPR functionality of the type IV-B system could be repeated with the addition of the *cysH* gene into the target/non-target plasmids. If CysH plays a role in the potential anti-CRISPR functionality of the type IV-B system, its ability to enhance the inhibitory effect of the type IV-B system over the immunity of the type IV-A system would be observed. Second, determining the function of CysH individually could be done using *in vitro* biochemical assays. In brief, if CysH functions by adenylating nucleic acid, an assay that measures the migration alpha p32 labeled ATP after incubation with CysH and nucleic acid could be used to determine this.

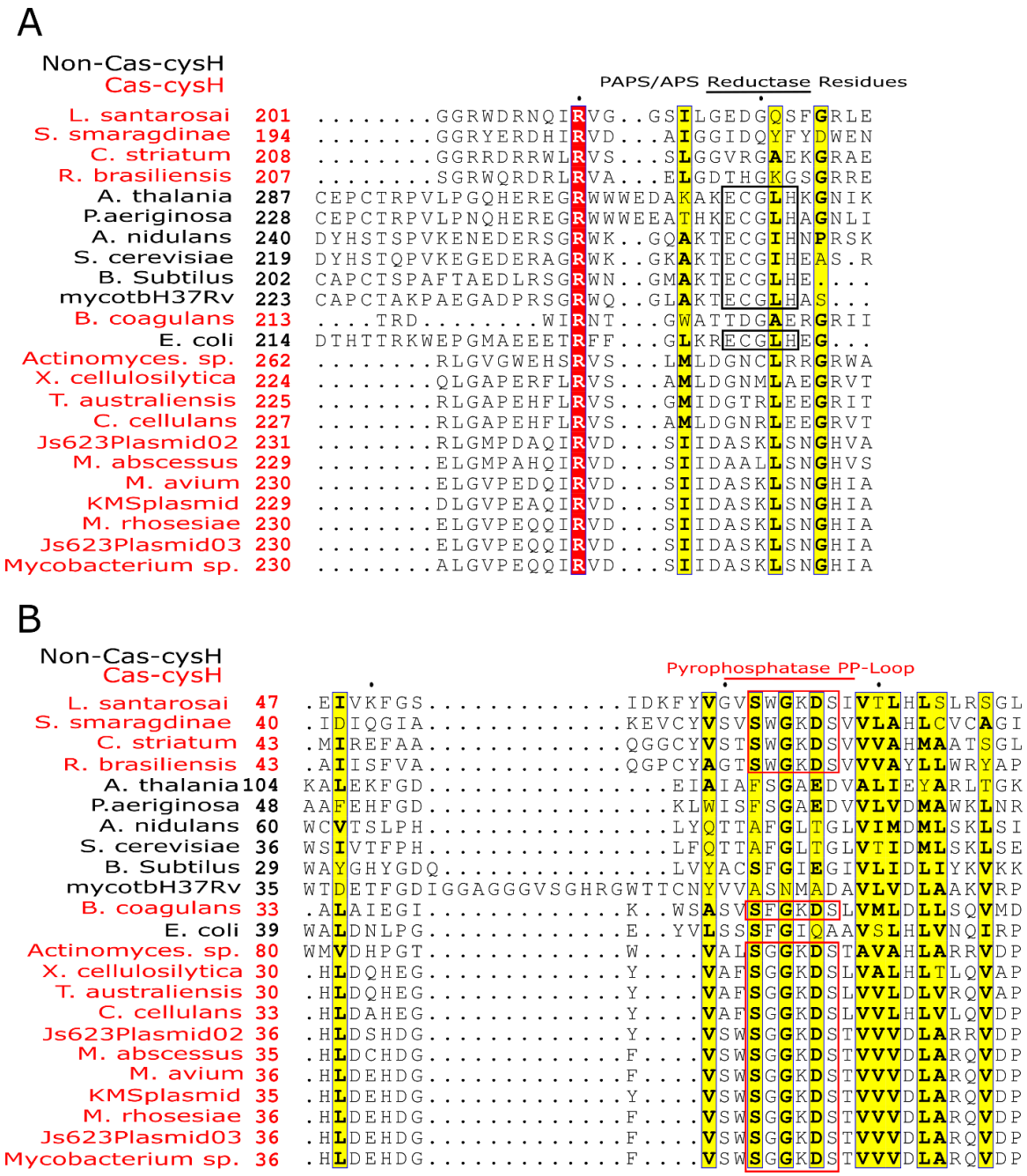


Figure 2-2. CysH Sequence Alignments. Alignments of Cas-CysH (red) and Non-Cas-CysH (black) sequences show unique differences between sequences. (A) Alignment of PAPS/APS reductase active site residues. The boxed region shows active site residues ECGLH. (B) Alignment of pyrophosphatase loop. The boxed regions show residues conserved in a pp-motif found in ATP pyrophosphatases, SxGxDS.

Analysis of *Smithella* Cas12a Function

One of the first questions about Cas12a2 was why it often co-occurs adjacent to Cas12a in the operon of the type V systems. Before designing experiments to probe how their activities work together, we wanted to first understand how they function separately. Other orthologs of Cas12a have previously been characterized [17]. However, given how *cas12a2* is adjacent to *cas12a* we wondered about the potential of an alternative function for the encoded Cas12a protein. In the organism *Smithella*, Cas12a was previously found to have no activity [17]. However, we found reason to believe that the wrong direct repeat sequence was used because there are two direct repeat sequences found in *Smithella.* The direct repeat that bears the closest resemblance to other Cas12a direct repeats was not used in this study. We hypothesized that using the alternative direct repeat found in the *Smithella* genome, we would observe activity.

Recombinant expression and purification of SmCas12a resulted in approximately 1 mg of soluble protein per gram of cells (**Figure 2-3A**). A double transformation with a crRNA containing plasmid allowed SmCas12a to be co-expressed with its crRNA. Additionally, expression of apo SmCas12a was successful. Cleavage assays with SmCpf1:crRNA complex and a PCR product complimentary to the crRNA resulted in observable cleavage within an hour, consistent with the activities observed with other Cas12a orthologs (**Figure 2-3B**). Thus, this work laid a foundation for future biochemical studies aimed at understanding the interplay between Cas12a and Cas12a2 proteins when they are co-expressed in a cell.

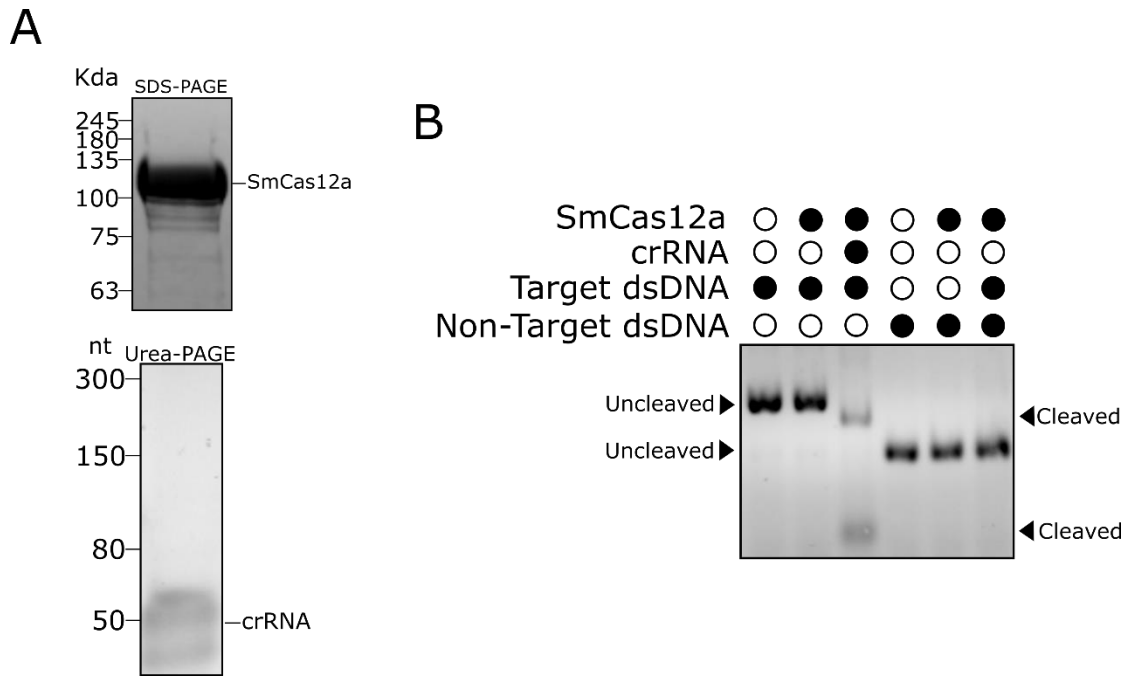


Figure 2-3 Expression, Purification and *in vitro* Cleavage Assay of SmCas12a. SmCas12a can be recombinantly expressed in *E. coli* and possesses RNA guided dsDNA cleavage activity. (A) Expression and purification results for SmCas12a. SDS-PAGE, stained with Coomassie blue, of concentrated SmCas12a SEC fractions (Top) show protein corresponding to the approximate size of SmCas12a. Urea-PAGE, stained with SYBR Gold, of phenol:chloroform extracted RNA from concentrated SmCas12a SEC fractions (Bottom) show RNA corresponding to the approximate size of processed crRNA. (B) Agarose gel of DNA cleavage assay of PCR product by SmCas12a. DNA stained with ethidium bromide.

Cas12a2 Mammalian Cell Editing

The discovery of Cas12a2 by Benson Hill was accompanied by data that Cas12a2 induced insertion/deletions (indels) in the rice plant *Oryza sativa* [54]. However, there was no evidence that Cas12a2 could edit mammalian cells. To determine the feasibility of using Cas12a2 as a novel mammalian cell editor, I collaborated with Zhongde Wang from the department of Animal, Dairy, and Veterinary Sciences of Utah State University,

who has experience with genome editing in the rodent *Mesocricetus auratus* using CRISPR Cas9 [64]. Through assistance with Dr. Wang's lab, I designed and carried out the experimentation used to evaluate the genome editing feasibility of Cas12a2 (**Figure 2-4**).

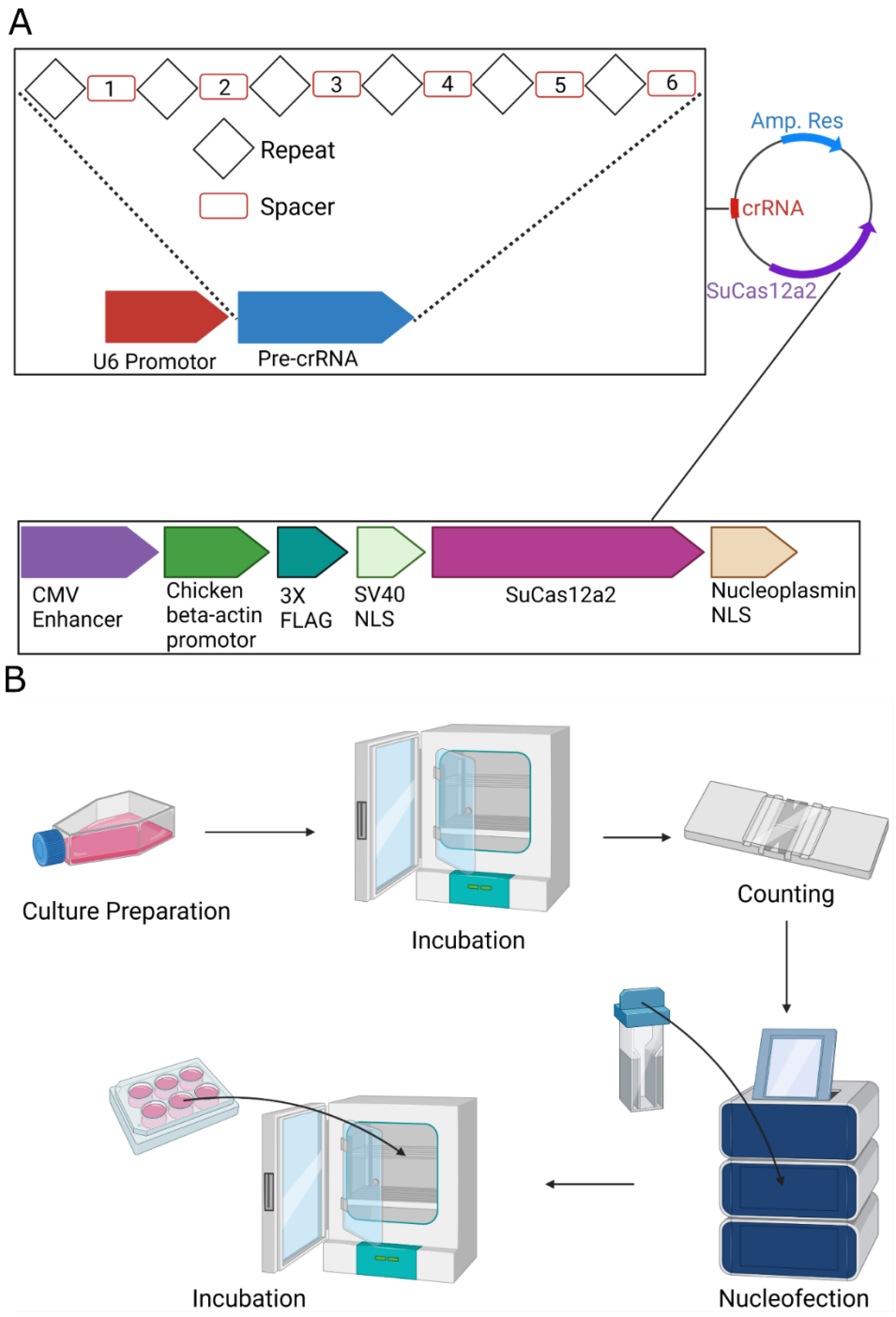


Figure 2-4. Baby Hamster Kidney Cell Culture and Genome Editing. (A) Vector design for SuCas12a2 genome editing experiments. Each spacer targets a different gene. 1) Mx2 2) Mavs 3) Ifnar2 4) Tlr2 5) Tlr4 6) Tlr10 (B) Workflow for cell culture and transfection.

Cloning of Plasmids for Mammalian Cell Expression

A *Mesocricetus auratus* codon optimized gene of *cas12a2* was cloned in the plasmid pX330-U6-Chimeric_BB-CBh-hSpCas9, a gift from Feng Zhang (Addgene plasmid # 42230) in the place of *cas9.* Already included in the plasmid is the chicken-beta-actin promotor and the CMV enhancer that drives expression of genes in mammalian cells [65]. Attached on N-terminus and C-terminus are nuclear localization signals that promote the importation of the protein into the nucleus [66]. A pre-crRNA targeting six genes was cloned into the plasmid in place of the gRNA scaffold under the U6 promotor that drives expression of small RNAs [67]

Evaluation of Indels

Evaluation of whether or not SuCas12a2 could edit mammalian cells was done with the T7 Endonuclease I assay. T7 Endonuclease I cleaves ssDNA located at bubbles resulting from mismatched bases of two strands of DNA [68]. If DNA from some cells of a culture of mammalian cells has been edited, then not all of the PCR products of the edited region will have the same sequence with a PCR reaction with primers flanking the target gene (**Figure 2-5A**). The random nature of bases inserted or deleted during non-homologous end joining [69] means that the sequences of PCR products will vary at the site of indels between edited and non-edited cells. Melting and subsequent annealing of PCR products allows for mismatches to occur between edited and non-edited strands at the site of mutation (**Figure 2-5A**). Addition of T7 Endonuclease I produces a cleavage band that can be visualized on a gel (**Figure 2-5A**). The T7 endonuclease I assay was used to evaluate indels produced by Cas12a2 at 6 target locations. There appeared to be

no difference in band migration of transfected and not transfected DNA, suggesting Cas12a2 was incapable of mammalian cell editing (**Figure 2-5B**).

Evaluation of Cell Death

Our collaborators discovered that targeting by Cas12a2 resulted in the death of *E. coli* cells containing the Cas12a2 immune systems instead of the survival of cells observed with most CRISPR systems. The discovery of Cas12a2 induced cell death in *E. coli* led to a hypothesis that a possible reason for lack of observed indels caused by Cas12a2 in mammalian cells was due to the death of cells targeted by Cas12a2. To evaluate this possibility, cells were transfected with either a plasmid promoting the expression of GFP or Cas12a2:crRNA. Live cells were counted 24 hours and 48 hours after transfection, and untransfected cells were used as a control (**Figure 2-5C**). Preliminary data show a decrease in cell counts upon transfection by both GFP and Cas12a2 (**Figure 2-5D**). Cells transfected with Cas12a2 show an even further decrease in cell counts compared to GFP (**Figure 2-5D**). Each culture was started with the same number of cells, suggesting that the reduction in cell counts is due to the transfection and/or the action of the transfected protein. Although, additional controls are needed to reach a definitive conclusion, these data provided adequate incentive for further investigation.

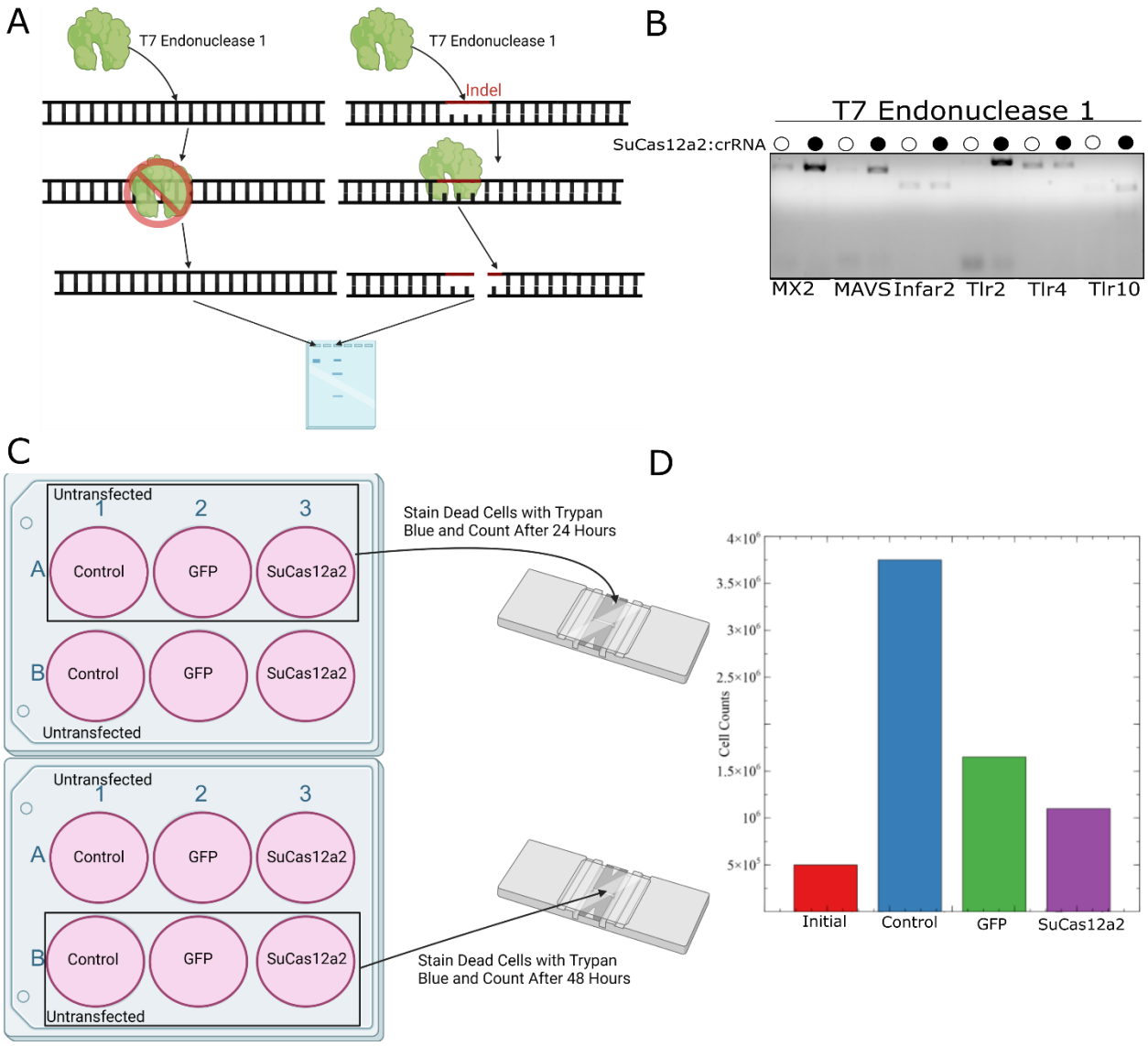


Figure 2-5 BHK Transfection by with SuCas12a2. SuCas12a2 does not cause indels when transfected into BHK cells. (A) Workflow for T7 Endonuclease I assay. (B) T7 Endonuclease I data. (C) Workflow for cell death assay. (D) Results from cell death assay. (A,C)

Optimization of SuCas12a2 Expression and Purification Protocol

Cas12a2 is a newly discovered type V CRISPR protein most closely related to Cas12a, although, Cas12a2 only shares 10-20 percent sequence identity with Cas12a [54]. Cas12a2 also occurs alongside Cas12a in certain genomes, suggesting a potentially synergistic functionality. One way to better understand the role of Cas12a2 is to characterize its activity using *in vitro* biochemical assays. The first step in running *in vitro* biochemical assays is to develop an efficient expression and purification protocol that produces enough clean protein for assays and structural studies. In this section I discuss my efforts to optimize the established expression and purification protocol of SuCas12a2, in order to obtain higher quantities of SuCas12a2.

The expression protocol for SuCas12a2 when I joined the lab was an autoinduction based protocol [70]. The soluble protein obtained was quickly depleted by setting up crystal trays and biochemical assays. I made a few modifications to the expression protocol that resulted in much larger quantities of soluble protein. I switched to an IPTG induction method (0.1 mM), expressed in Nico21 (DE3) cells, and grew them in a much more nutrient rich media, Terrific Broth.

The modified expression protocol for SuCas12a2 resulted in much larger quantities of soluble protein. To visualize the increase in expression, **Figure 2-6** compares the chromatograms resulting from two nickel affinity chromatography steps, one with old autoinduction protocol and one with the modified IPTG induction protocol. The column used, the FPLC system and UV detector were identical in the comparison. The only differences are the expression protocol and the date of the protein prep. The UV absorbance at 280 nm was much higher for the protein prep with the modified expression

protocol than for the autoinduction expression protocol (**Figure 2-6**). Overall, by modifying the expression protocol of SuCas12a2 I was able to obtain much more protein that allowed more biochemical assays to be ran, and more crystal trays to be setup at a higher protein yield per prep of protein. The work I did to optimize the expression and purification of SuCas12a2 has led to two manuscripts from the Jackson Lab that are currently in press.

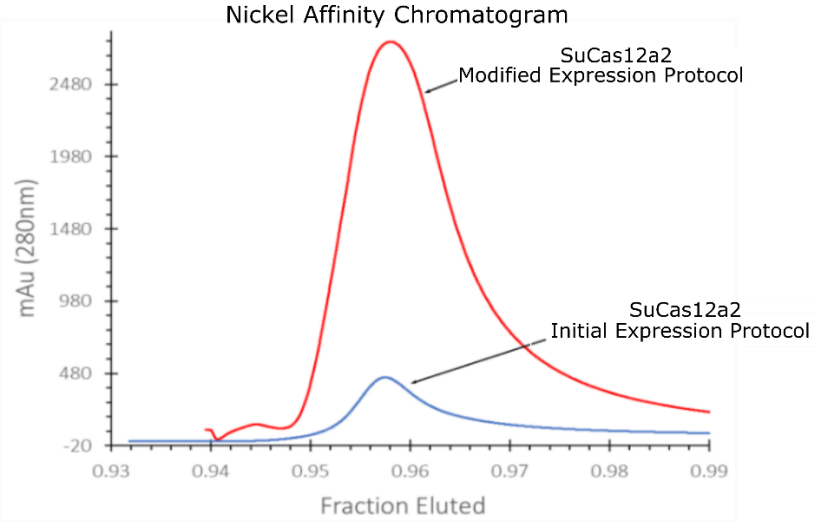


Figure 2-6: Improvement of Initial Expression and Purification Protocol of SuCas12a2. An IPTG induction method in Terrific Broth media enhances expression of SuCas12a2. A comparison of chromatograms obtained from nickel affinity chromatography between SuCas12a2 expressed in cells using an auto induction method [70] (blue line) and an IPTG induction method in Terrific Broth (red line).

Analysis of SuCas12a2 Cleavage Activity

The initial hypothesis for the activity of SuCas12a2 was dsDNA cleavage through a RuvC mediated, RNA guided mechanism. This hypothesis was based on the observation of plasmid targeting using a plasmid curing assay that is discussed in chapter 3. We also found reason to believe that the similarity between Cas12a2 and Cas12a

would mean similar function, i.e DNA targeting. The results were inconsistent. Occasionally, no cleavage would occur, other times complete degradation of both complementary and non-complementary DNA would occur. We hypothesized that certain fractions of purified enzyme were actively cleaving DNA and others were not. To identify which fractions, I ran cleavage assays on different fractions that eluted, from different columns. In this way I was able to narrow down the complete degradation activity to a single fraction from an anion exchange chromatography elution. I found that this degradation activity was intensified by a pre-incubation with RNAse A, and simply letting the protein sit in the fridge for a few days. We later discovered SuCas12a2 to be an RNA targeting protein that collaterally degrades ssDNA, dsDNA and ssRNA, which will be further elaborated on in chapter 3. The discovery of the true function of SuCas12a2 shines light on a possible explanation for our initial observations in regard to DNA targeting. We believe that because the crRNA is coexpressed with SuCas12a2, there may also be small amounts of the reverse complement of the crRNA being transcribed. The reverse complement of the crRNA could be recognized as an RNA target provided that the reverse complement of the direct repeat is not present, as RNA flanked by repeat sequence was found to be inhibitory and will be discussed in chapter 3. We hypothesized that increase in activity over time and by RNAse A treatment could be explained by degradation of the inhibitory direct repeat reverse complement to reveal compatible ssRNA for activation. In addition to this characterization of SuCas12a2 degradation activity, Valerie Crowley observed DNA degradation with SmCas12a2. These observations can now be explained by an RNA-triggered activation mechanism described in chapter 3.

Conclusion

In this chapter, I described how I designed and cloned the plasmids necessary for evaluating the anti-CRISPR functionality of the type IV-B system, and I identified pitfalls in the plasmid curing assay with the type IV-A systems. However, since my experiments, the Jackson Lab has discovered the PAM sequence used by the type IV-A system that has led to better reproducibility with the assay. The next steps to evaluate type IV-B system activity as an anti-CRISPR would be to repeat the assay I developed but using the correct PAM sequence.

Along with an evaluation of the anti-CRISPR potential of the type IV-B system, I have also set forth the groundwork for characterizing the ancillary protein CysH. Through bioinformatic analysis, I have found that Cas-Cys*H* and non-Cas-CysH protein sequences differ. The active site residues, necessary for sulfate assimilation activity, found in non-Cas-CysH protein are not present in Cas-CysH. While a conserved region necessary for pyrophosphatase activity is found solely in Cas-CysH proteins. To evaluate the activity of Cas-CysH, I believe a starting place would be to look into pyrophosphatase activity in the form of nucleic acid adenylation. Pyrophosphatase activity is found primarily in enzymes that adenylate certain substrates. I hypothesize that a possible substate for Cas-CysH could be nucleic acids because it is found adjacent to a CRISPR system that presumably works on nucleic acids.

I successfully expressed and purified SmCas12a and showed evidence of RNA guided DNA cleavage activity. This work could be useful for direct comparison of Cas12a and Cas12a2 activities from an organism that encodes both proteins.

I designed and cloned plasmids necessary for genome editing evaluation of SuCas12a2. Although I found no evidence of genome editing in the form of indels, I did find data suggesting that SuCas12a2 induces cell death in mammalian cells. However, additional experimentation with necessary controls will be important for a more definitive answer.

The optimization of the expression and purification of SuCas12a2 led to higher protein yields. By utilizing an IPTG induction method and Nico21-DE3 cells, I was able to obtain much larger amount of protein in comparison to the initial autoinduction methodology in Bl21-DE3 cells. With a new expression and purification protocol, the Jackson lab was able to successfully evaluate the biochemical function of SuCas12a2 as well as the Cyro-EM structure with the help of UT Austin. The work of which has led to two manuscripts now under review.

The work done in characterizing SuCas12a2 activity was rocky at the start as we were looking for evidence of DNA targeting when in reality SuCas12a2 was found to target RNA. In chapter 3, I discuss in more detail the work I did in characterizing the function of SuCas12a2.

As stated before, a lot of my work was spent chasing multiple questions at once until one showed promise. A lot of my initial work has set the groundwork for multiple projects in the Jackson Lab. Eventually, we identified the function of SuCas12a2, and I narrowed my focus to characterizing SuCas12a2 using *in vitro* nucleic acid cleavage/degradation assays. My work in this area will be discussed at length in chapter 3.

Methods

Plasmids used in this chapter can be found in **Table B-1**.

Expression and Purification of Smithella Cas12a (Cpf1)

N-terminal 6x His-tagged SmCas12a was expressed in *E. coli* Bl21 (DE3) cells from a pET plasmid (lab archive #1424) along with a 3X CRISPR from a pACYC plasmid (lab archive #1425). 1 L of LB media and was inoculated with 15 mL of overnight growth, started from a fresh transformation, and was grown at 37°C until an $O.D_{600}$ of 0.6. The cells were cold shocked on ice for 15 minutes, to induce the expression of cold-shock proteins that assist with protein expression at lower temperatures [71], [72]. The cells were induced with 0.133 mM IPTG followed by a 16-18 hour incubation at 18°C. Cells were harvested by centrifugation at 8k RPM for 25 minutes. The cell pellets were lysed by sonication in Lysis buffer (20mM Tris-pH 8.0, 500 mM NaCl, 10 mM imidazole) in the presence of leupeptin, aprotinin, pepstatin, and lysozyme. The lysate was clarified by centrifugation at 36,400 x g for 35 minutes. The clarified lysate was vacuum filtered using 0.45μm filter. The lysate was applied to a Histrap crude FF 5 mL column equilibrated in lysis buffer. The column was washed with 10 CV of 98% lysis buffer + 2 % Ni-elution buffer (20mM Tris-pH 8.0, 500 mM NaCl, 250 mM imidazole), and eluted with 8 CV of ni-elution buffer. Fractions containing SmCas12a were desalted into low salt buffer (20mM Tris-pH 8.0, 50 mM KCl) using a Hiprep 26/10 desalting column. The fractions containing SmCas12a were applied over a Hitrap Heparin HP 5 mL column equilibrated in low salt buffer. The column was washed with 5 CV of 98% low salt buffer, 2% high salt buffer (20mM Tris-pH 8.0, 1 M KCl), and eluted using a

linear gradient elution from 98% low salt buffer + 2% high salt buffer to 100% high salt buffer over 12 CV. The fractions containing SmCas12a were concentrated to 1 mL and then applied to a Hiload 16/600 superdex 75 pg SEC column equilibrated in SEC buffer (20 mM HEPES-pH 8.0, 500 mM KCl, 1mM DTT). The fractions containing SmCas12a were concentrated and stored at -80°C. The presence and purity of SmCas12a was analyzed using 12% SDS-PAGE.

PCR Amplification of pET27b Target dsDNA

In a 25 μL reaction, the following components were mixed: 0.5 μL of forward primer (lab archive #904, 10 μM stock), 0.5 μL of reverse primer (lab archive #905, 10 μM stock), 0.25 μL of template DNA (lab archive #1096, at least 100 ng/μL stock is ideal), 11.25 μL nuclease free water, 12.5 uL of Taq 2x master mix. For good yields, a master mix worth 16 reactions was made by multiplying everything by 17 to make up for pipetting error. The following temperature program was run on a thermocycler: Initial denaturation: 30 seconds at 95°C, (recuring denaturation: 15 seconds at 95°C, annealing: 30 seconds at 50°C, extension: 1 min at 68°C)x35, final extension: 5 minutes at 68°C. PCR reactions were cleaned with a PCR clean up kit. For good concentrations, one column was used per 8 PCR reactions with a 20 μL elution volume in nuclease free water.

Smithella Cas12a in Vitro Cleavage Assay

In a 20 μL reaction, the following containing ~300 ng target PCR product and 370 nM SmCas12a:crRNA was prepared in 1X cleavage buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT). The reaction was incubated at 25°C for 90

min, cleaned with a PCR clean up kit, and analyzed on 2% agarose with ethidium bromide.

Initial Sulfuricurvum Cas12a2 Expression and Purification Protocol

N-terminal 6x His-tagged SuCas12a2 WT and mutant constructs were expressed in *E. coli* Bl21(DE3) cells from a plasmid either lacking (apo) (lab archive #1416) or containing a three-spacer CRISPR array (lab archive #1408). The cells were grown in 1L auto induction media [70] at 37°C for 5 hours followed by 24°C for 24 hours. Cells were harvested by centrifugation at 8k RPM for 25 minutes. The cell pellets were lysed by sonication in Lysis buffer (25mM Tris-pH 7.2, 500 mM NaCl, 10 mM imidazole, 2 mM $MgCl_2$, 5 mM TCEP, 10% glycerol) in the presence of leupeptin, aprotinin, pepstatin, PMSF and lysozyme. The lysate was clarified by centrifugation at 36,400 x g for 35 minutes. The lysate was applied to a Histrap crude FF 5 mL column equilibrated in lysis buffer. The column was washed with 10 CV of 90% lysis buffer + 10% Ni-elution buffer (25mM Tris-pH 7.2, 500 mM NaCl, 500 mM imidazole, 2 mM $MgCl_2$, 5 mM TCEP, 10% glycerol), and eluted with 8 CV of ni-elution buffer. Fractions containing SuCas12a2 were desalted into low salt buffer (25 mM Tris-pH 7.2, 50 mM NaCl, 2 mM $MgCl_2$, 5 mM TCEP, 10% glycerol) using a Hiprep 26/10 desalting column. The fractions containing SuCas12a2 + CR were applied to a Hitrap Q HP anion exchange column, and SuCas12a2 to a Hitrap SP HP cation exchange column equilibrated in low salt buffer. The column was washed with 10 CV of 90% low salt buffer, 10% high salt buffer (25 mM Tris-pH 7.2, 1 M NaCl, 2 mM $MgCl_2$, 5 mM TCEP, 10% glycerol), and eluted 8 CV high salt buffer. The fractions containing SuCas12a2 were concentrated to 1

mL and then applied to a Hiload 26/600 superdex 200 pg SEC column equilibrated in SEC buffer (100 mM HEPES-pH 7.2, 150 mM KCl, 5 mM TCEP). The fractions containing SuCas12a2 were concentrated and stored at -80°C.

Modified Sulfuricurvum Cas12a2 Expression and Purification Protocol

N-terminal 6x His-tagged SuCas12a2 WT and mutant constructs were expressed in *E. coli* Nico21(DE3) cells from a plasmid either lacking (apo) (lab archive #1416) or containing a three-spacer CRISPR array (lab archive #1408). 1 L of TB media was inoculated with 20 ml of overnight growth, started from a fresh transformation, and was grown at 37°C until an OD_{600} of 0.6. The cells were then cold shocked on ice for 15 minutes and induced with 0.1 mM IPTG, followed by a 16-18 h incubation at 18°C. Cells were harvested by centrifugation. Cells were lysed by sonication in Lysis buffer (25 mM Tris-pH 7.2, 500 mM NaCl, 10 mM imidazole, 2 mM $MgCl_2$, 10% glycerol) in the presence of leupeptin, aprotinin, pepstatin, aebsf, and lysozyme. The lysate was clarified by centrifugation at 36,400 x g for 35 minutes. The clarified lysate was added to 5 mL of Ni-NTA resin and incubated at 4°C for 30 minutes. The resin was washed with 10 fractions of lysis buffer (10 mL per fraction). The protein was eluted with 10 fractions of Ni-elution buffer (25mM Tris-pH 7.2, 500 mM NaCl, 250 mM imidazole, 2 mM $MgCl_2$, 10% glycerol) (4 mL per fraction). The fractions containing SuCas12a2 were desalted using a Hiprep 26/10 desalting column into low salt buffer (25mM Tris-pH 7.2, 50 mM NaCl, 2 mM $MgCl_2$, 10% glycerol). The fractions containing SuCas12a2 + CR were applied to a Hitrap Q HP anion exchange column, and SuCas12a2 to a Hitrap SP HP cation exchange column. The column was washed with 90 percent low salt buffer + 10 percent high salt buffer (25mM Tris-pH 7.2, 1 M NaCl, 2 mM $MgCl_2$, 10% glycerol). A

linear gradient elution was used from 10 percent high salt buffer + 90 percent low salt buffer to 100 percent high salt buffer over 10 CV. The fractions containing SuCas12a2 were concentrated using a 100 MWKO concentrator to about 1 mL. The concentrated sample was applied to a Hiload 26/600 superdex 200 pg SEC column equilibrated in SEC buffer (100 mM HEPES-pH 7.2), 150 mM KCl, 2mM $MgCl_2$, 10% glycerol). The fractions containing SuCas12a2 were concentrated and stored at -80°C. The presence and purity of SuCas12a2 was analyzed using 12% SDS-PAGE.

Mammalian Cell Experiments: Working Organism and Vector Design.

The ability to induce indels in mammalian cells was evaluated using baby hamster kidney cells (BHK). A total of 6 genes were selected as possible targets with 3 possible sites on each target. Each possible target utilized a different PAM, which were shown to be recognized by Cas12a2. The experiments were done using transfection of a vector containing a pre-crRNA with 6 spacer sequences and a Cas12a2 gene codon optimized for *Mesocricetus auratus*. Transcription of the pre-crRNA was initiated using a U6 promotor recognized by RNA polymerase III, shown to be important is transcription of small-RNAs [67]. Transcription of the Cas12a2 gene was initiated using a chicken-beta actin promotor with a CMV enhancer shown to function well in mammalian cell expression [65]. The N-terminus of Cas12a2 is linked to a 3X FLAG tag and a SV40 nuclear localization signal (NLS). The C-terminus of Cas12a2 is linked to a nucleoplasmin NLS.

Mammalian Cell Cloning

SuCas12a2 codon optimized for *Mesocricetus auratus* and a 6-spacer containing crRNA were cloned into a pX330 backbone from addgene #42230 using Gibson assembly (SGI-DNA Gibson Assembly HiFi 1-step). Three construct variations were made containing differing spacer sequences in the crRNA.

Mammalian Cell Culture

Baby hamster kidney cells (BHK) were cultured in DMEM media with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin, 0.4% essential amino acids, and 0.2% non-essential amino acids. Cells were grown at 37°C and 5% CO_2 in a 25 cm^2 cell culture flask and passaged using trypsin once they reached ~100% confluency. Cells were counted using a hemocytometer and 5 x 10^5 were transfected with 3 μg of plasmid using a P3 Primary Cell 4D-Nucleofector X Kit L. Cells were further cultured in a 6 well culture plate. For cell death experiments, cells were stained with trypan blue and counted using a hemocytometer after 24 and 48 hours. For genome editing, genomic DNA was harvested using a DNeasy Blood and Tissue Kit from Qiagen after 48 hours.

PCR Amplification of Target Sites

BHK DNA was isolated using a DNeasy Blood & Tissue Kit from Qiagen. Primers adjacent to target sites were used in PCR amplification by Taq polymerase or Q5 polymerase. PCR reactions were cleaned using a PCR clean up kit.

T7 Endonuclease I cleavage

In a 10 μL reaction, ~300-500 ng of PCR product was prepared in 1X NEB 2 and denatured and reannealed using the following program on a thermocycler: 95°C for 5 min

followed by a temperature change to 85°C at -2°C/s, then a change to 4°C at a rate of -0.1°C/s. 2000 units of T7 Endonuclease I (NEB #E3321) was added and the reaction was incubated at 37°C for 1 hour. The reaction was analyzed using 1 % agarose stained with ethidium bromide.

CFU based Plasmid Curing Assay

Bl21-AI cells were transformed with the type iv-a system and made chemically competent. Half the cells were grown in inducing conditions (1 mM IPTG for CRISPR-Cas genes, 0.2% L-arabinose for T7 RNA polymerase), the other half were not. 100 μL from both groups of cells, once chemically competent, were transformed with 1 μL of plasmid (target, non-target, or target + type iv-b) at 55 ng/ μL using heat shock at 42°C. 90 μL of the cells were transferred to 5 mL of LB with appropriate antibiotics to select for each plasmid. The induced cells were supplemented with 1 mM IPTG and 0.2% L-arabinose. Both were grown overnight at 37°C. Growths were plated on LB agar with antibiotics and the induced cells with 1 mM IPTG and 0.2% L-arabinose. To optimize for number of colonies resulting for each plate, a dilution series was made. The uninduced cells were diluted by factors of 10^5, 10^6, 10^7, and 10^8. The induced cells were diluted by factors of 10^1-10^8 in ten-fold increments. The plates were incubated overnight at 37°C and the number of colonies counted and multiplied by the dilution factor.

PCR based Plasmid Curing Assay

100 μL of induced and uninduced chemically competent Bl21-AI cells were transformed with ~13 fmol each of non-target and target plasmid (non-target/target, or non-target/target + type iv-b) by heat shock at 42°C. 90 μL of each transformation was

transferred to 410 μL of LB with the induced cells being supplemented with 1 mM IPTG and 0.2% L-Arabinose. The cells were incubated at 37°C for 1 hour then split and diluted by factors of 2, 5, and 10 then plated on LB agar with appropriate antibiotics and the induced cells with 1 mM IPTG and 0.2% L-arabinose. A total of 94 colonies, split evenly between plates (uninduced target/non-target, uninduced target + type IV-B/non-target, induced target/non-target, and induced target + type IV-B/non-target), were used for colony PCR using primers flanking both the target site and non-target site. The number of colonies containing the target site and those containing the non-target site were counted and a ratio between them was calculated.

Acknowledgments

Figures 2-4, 2-5A, and 2-5C were created using biorender.com

CHAPTER 3

FUNCTIONAL CHARACTERIZATION OF THE CRISPR TYPE V EFFECTOR PROTEIN CAS12A2

Abstract

Previous research of CRISPR systems has led to advancements in biotechnology. A newly characterized type V CRISPR protein Cas12a2 shows potential as a novel tool. Initial research conducted by collaborators and others in the Jackson Lab using *in vivo* assays showed that Cas12a2 was a functioning immune system that induced an abortive infection phenotype. However, much about the mechanism of Cas12a2 remained unknown. In this study, the mechanism of Cas12a2 based immunity is determined using *in vitro* biochemical assays. We show that Cas12a2 functions through an RNA targeting mechanism. Upon RNA targeting, Cas12a2 is activated to collaterally degrade ssDNA, ssRNA and dsDNA. We show that Cas12a2 utilizes a RuvC domain and zinc-finger motif in a Mg^{2+} dependent manner. To highlight the uniqueness of the collateral nuclease activity of Cas12a2, a comparison against other CRISPR-Cas proteins Cas12a, Cas12g and Cas13a is conducted. The results highlight that Cas12a2 is unique in that it can collaterally degrade ssDNA, ssRNA and dsDNA while Cas12a, Cas12g and Cas13a are limited to just ssDNA and/or ssRNA. The collateral nuclease activity discovered of Cas12a2 provides a mechanism for Cas12a2 abortive infection immunity found using *in vivo* assays. The novel mechanism of Cas12a2 immunity has the potential to be utilized as new tools or as an improvement of existing tools in biotechnology.

This chapter is based on my contribution to a manuscript in press:
Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Introduction

Many different tools in biotechnology have risen from the research of microbial immune systems. Restriction modification systems have been repurposed as a revolutionary tool of molecular cloning. Genome editing has been possible for many years with tools such as Transcription Activator Like Endonucleases (TALENs) and Zinc Finger Nucleases (ZFN) [73]–[75]. Unfortunately, each TALEN or ZFN had to be specifically designed to target only one sequence. Changing the target sequence required modification of the DNA binding proteins involved that made the time and cost of use often impractical. The discovery of the RNA guided mechanism of Cas9 led to large advancements in genome editing by overcoming many of the drawbacks of TALENs or ZFNs [76]. The ability to target different sequences in the genome simply by programming Cas9 with an RNA guide made using a Cas9 based genome editor a much more practical option over a TALEN or zinc finger nuclease-based editor, which required several months to design and purify a tool to make a single edit [76]. As prior research of microbial immune systems has led to advancements in biotechnology, it stands to reason that as more bacterial immune systems are understood, further advancements may be found.

Recently, the biotech company Benson Hill discovered a potential CRISPR nuclease similar to type V CRISPR endonuclease Cas12a, but with enough of a

difference to warrant interest, called Cas12a2 [54]. Sequence analysis of Cas12a2 showed the presence of a RuvC nuclease domain suggesting similar function, although with an overall sequence identity of 10-15% compared to Cas12a [54]. Interestingly, it was also found that Cas12a2 often co-exists in genomes alongside Cas12a suggesting that Cas12a2 shares a distinct mechanism that complements the dsDNA targeting of Cas12a [54].

Through the work of collaborators and other members of the Jackson Lab, we learned a few things about the function of Cas12a2. First, using *in vitro* biochemical assays, Valerie Crowley and Hannah Domgaard from the Jackson Lab showed that Cas12a2 processes its own RNA guide through RNA cleavage activity. Second, Valerie and Gina Neumann from Benson Hill showed that Cas12a2 functions as an immune system by clearing transformed target plasmid using cell-based assays. Third, in addition to clearing transformed target plasmid, collaborators Gina Neumann and Oleg Dmytrenko showed that Cas12a2 triggered an abortive infection phenotype using cell-based assays. In this study, we use biochemical assays to determine how Cas12a2 functions as an immune system to clear transformed target plasmid and cause abortive infection. We find that Cas12a2 targets RNA complementary to its crRNA. The binding of target RNA activates nuclease activity that indiscriminately degrades ssDNA, dsDNA and ssRNA. Our findings highlight a novel CRISPR mechanism with potential for applications in biotechnology.

Results and Discussion

Expression and Purification of SuCas12a2

Expression and purification of SuCas12a2 involved an expression plasmid containing SuCas12a2 with an N-terminal Histag and a 3X crRNA under separate T7 promotors. SuCas12a2 and crRNA were expressed in DE3 cells and purified over Ni affinity, anion exchange and size exclusion chromatography columns (**Figure 3-1A**). The resulting SEC chromatogram with absorbance measurements at 280nm and 254nm hows nucleic acid tracking along with the protein presumably due to bound crRNA (**Figure 3-1B**). SDS-PAGE analysis shows the presence of protein at 135 kDa, the expected molecular weight for SuCas12a2. Phenol:chloroform extraction of RNA and analysis using Urea-PAGE shows RNA just below 50 nt corresponding to a crRNA (**Figure 3-1B**).

In chapter 2, I discussed the work I did in enhancing the expression of SuCas12a2. To recap, the initial expression protocol of SuCas12a2 was autoinduction. The yield of Cas12a2 obtained was low and was quickly depleted using biochemical assays and ideally not enough for structural studies. By switching to an IPTG induction in nutrient rich TB media, the expression of SuCas12a2 was greatly increased (**Figure 3-1C**).

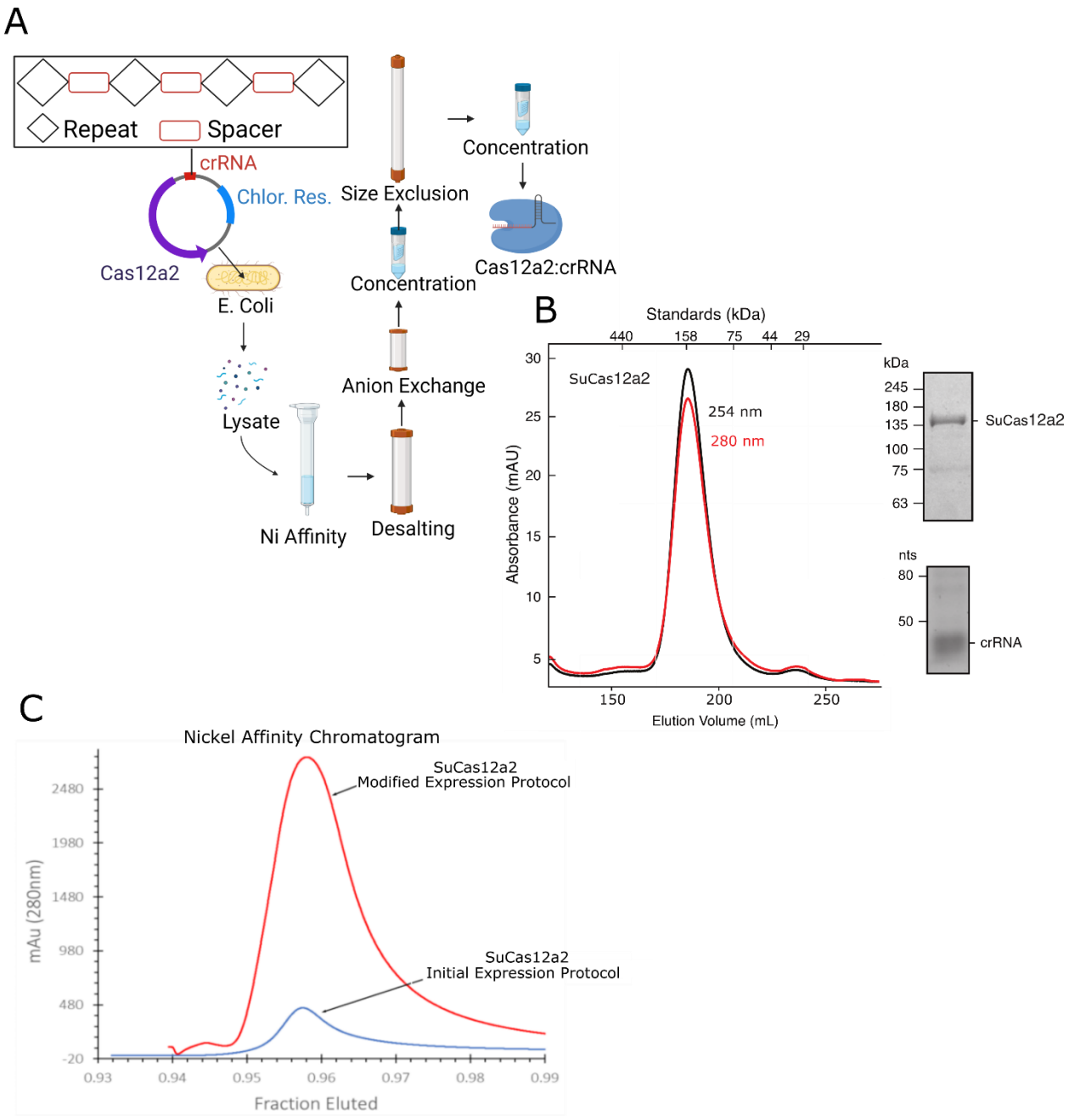

A
Repeat
Spacer
crRNA
Chlor. Res.
Cas12a2
E. Coli
Lysate
Ni Affinity
Desalting
Anion Exchange
Concentration
Size Exclusion
Concentration
Cas12a2:crRNA
B
Standards (kDa)
440
158
75
44
29
SuCas12a2
254 nm
280 nm
Absorbance (mAU)
Elution Volume (mL)
kDa
245
180
135
100
75
63
SuCas12a2
nts
80
50
crRNA
C
Nickel Affinity Chromatogram
SuCas12a2
Modified Expression Protocol
SuCas12a2
Initial Expression Protocol
mAu (280nm)
Fraction Eluted

Figure: 3-1. Expression and Purification of SuCas12a2. Expression of SuCas12a2 is enhanced over autoinduction through the use of IPTG induction in nutrient rich TB media. (A) Recombinant expression and purification of SuCas12a2 workflow. (B) Expression and purification results for SuCas12a2. Size exclusion chromatogram (Left). SDS-PAGE of concentrated SuCas12a2 SEC fractions (Right, Top). Urea-PAGE of phenol:chloroform extracted RNA from concentrated SuCas12a2 SEC fractions (Right, Bottom) show processed CRISPR RNA. (C) Nickel affinity chromatogram shows higher A_{280} for SuCas12a2 expressed with IPTG in TB media (red) over Cas12a2 expression using autoinduction (blue).

Cas12a2 Targets ssRNA

Based on the similarities with Cas12a with regards to a RuvC domain and its target preference of dsDNA [17], and previous *in vivo* data suggesting Cas12a2 could clear a dsDNA plasmid, we hypothesized that Cas12a2 could have similar preference for dsDNA, but because we had other data suggesting Cas12a2 caused an abortive infection phenotype, we wanted to examine the possibility that Cas12a2 works through another mechanism. To determine Cas12a2 target preference, an assay was performed in which various 5'FAM labeled nucleic acids (ssDNA, dsDNA, and ssRNA) containing target sequences complementary to the Cas12a2 crRNA, were individually incubated with recombinantly expressed SuCas12a2:crRNA (**Figure 3-2A**). Surprisingly, it was found that instead of targeting DNA as expected, SuCas12a2 targets RNA (**Figure 3-2B**).

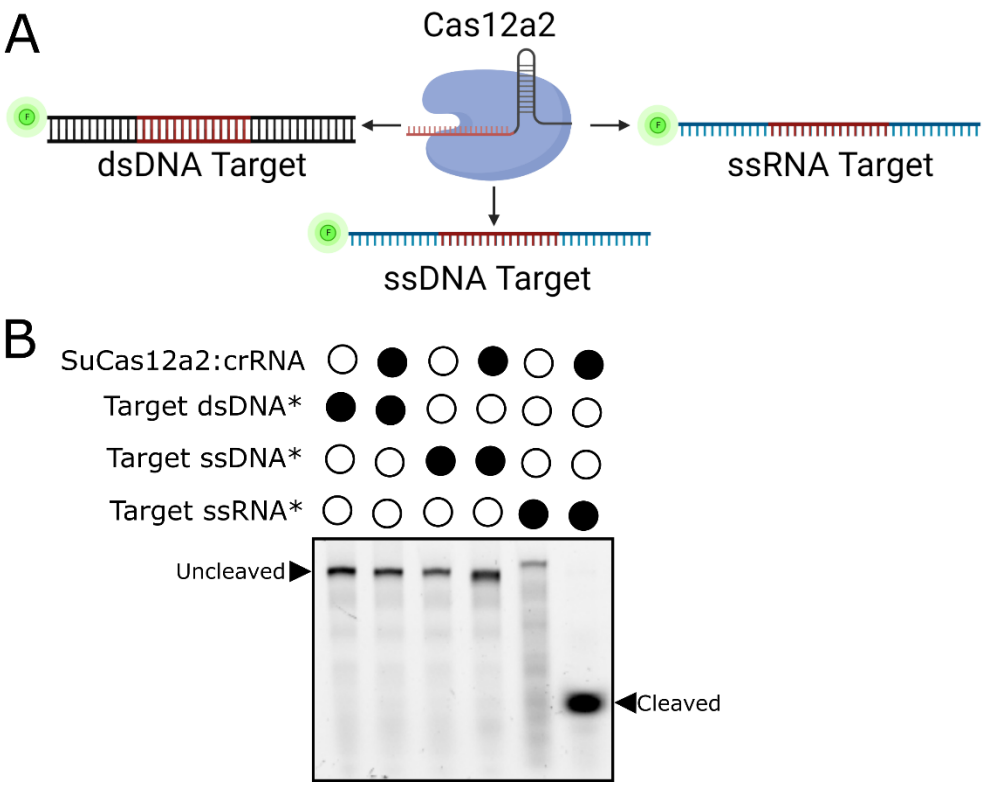


Figure: 3-2. RNA Targeting. SuCas12a2 is an RNA targeting protein. (A) Workflow for assay to determine nucleic acid target. (B) Nucleic acid digestion assay of different FAM-labeled nucleic acid targets by SuCas12a2 * denotes FAM labeled substrate.

Although we determined that Cas12a2:crRNA is activated by binding to target RNA, it was unclear whether an activating RNA is able to activate many Cas12a2 enzymes, or if a single target RNA is limited to the activation of one Cas12a2 enzyme.

To better understand of the nature of Cas12a2 activation, an assay was developed in which WT SuCas12a2 would compete with RuvC nuclease dead SuCas12a2 (D1213A). The D1213A mutant reduces the ability of the active site to coordinate a magnesium ion that is essential for nuclease activity [77]. Assuming the only difference in activity between WT SuCas12a2 and SuCas12a2 (D1213A) is the nuclease activity, SuCas12a2 (D1213A) should be able to still bind target RNA. In this assay, SuCas12a2

(D1213A) (200 nM) was incubated with FAM labeled target RNA (150 nM) (**Figure 3-3AI**). After 30 min of incubation given for SuCas12a2 (D1213A) to bind target RNA, WT SuCas12a2 (150 nM) was added (**Figure 3-3AII**). The hypothesis is that target RNA bound by SuCas12a2 (D1213A) will be inaccessible to activate WT SuCas12a2. To simulate what would happen if target RNA was able to activate WT SuCas12a2, a control containing additional unlabeled target RNA (200 nM) was also added alongside WT SuCas12a2 (**Figure 3-3AIII**). The results of this assay show the target RNA band shifting to a higher molecular weight species upon addition of SuCas12a2 (D1213A) (**Figure 3-3BI**). These observations are most likely not due to binding interactions with the protein because the reactions were phenol:chloroform extracted before running on the gel. However, this shift could be due to a binding interaction between the target RNA and the crRNA. More importantly, the results show a lack of cleavage of the target RNA. The addition of WT SuCas12a2 did not result in any observable cleavage of the target RNA (**Figure 3-3BII**). The lack of target RNA cleavage by WT SuCas12a2 could be explained by a lack of accessible target RNA available to activate WT SuCas12a2. This would signify that once an RNA target is bound by Cas12a2, it is inaccessible to another molecule of Cas12a2. The addition of extra target RNA resulted in the cleavage of the FAM labeled target RNA bound by SuCas12a2 (D1213A) (**Figure 3-3BIII**). Presumably, WT SuCas12a2 was activated by the addition of excess target RNA, and this led to cleavage of the ends of target RNA bound to SuCas12a2 (D1213A) *in trans* as it does not seem likely that the binding of newly added target RNA by WT SuCas12a2 resulted in dissociation of SuCas12a2 (D1213A) and its bound target RNA.

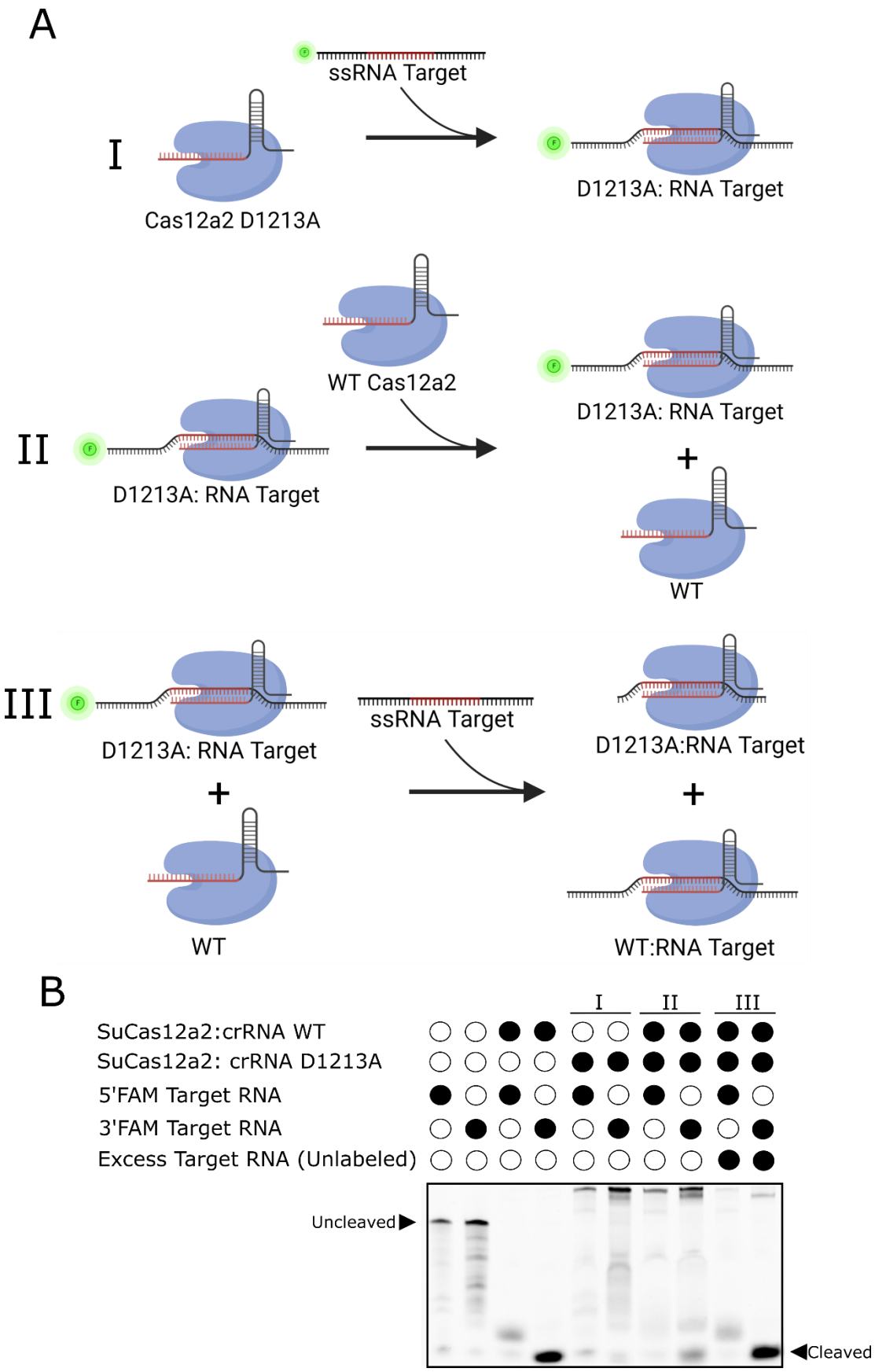

A
I
ssRNA Target
Cas12a2 D1213A
D1213A: RNA Target
II
WT Cas12a2
D1213A: RNA Target
D1213A: RNA Target
+
WT
III
D1213A: RNA Target
ssRNA Target
D1213A:RNA Target
+
+
WT
WT:RNA Target
B
I
II
III
SuCas12a2:crRNA WT
SuCas12a2: crRNA D1213A
5'FAM Target RNA
3'FAM Target RNA
Excess Target RNA (Unlabeled)
Uncleaved
Cleaved

Figure: 3-3. Accessibility of Cas12a2 Bound RNA Target for the Activation of an Additional Molecule of Cas12a2. Target RNA bound by Cas12a2 restricts access for the binding and activation of another molecule of Cas12a2. (A) Workflow for assay to determine if Cas12a2 can be activated by a previously bound molecule of target RNA. I) RuvC nuclease mutant D1213A is incubated with FAM labeled target RNA. II) After 30 min, WT Cas12a2 is added to the solution containing Cas12a2 D1213A bound to FAM labeled RNA target. III) After 30 min, WT Cas12a2 and excess unlabeled target RNA is added to the solution containing Cas12a2 D1213A bound to FAM labeled RNA Target (B) Urea-PAGE gel imaging for fluorescence (FAM) showing the results of the assay described above. I, II, and III correspond to I, II and III described above.

Cas12a2 Binding to Target ssRNA Activates Indiscriminate Nuclease Activity

Collateral nuclease activity by Cas proteins is a common strategy used by CRISPR immune systems [39], [40], [78]. For instance, it has been shown that LbCas12a collaterally degrades ssDNA once activated by dsDNA or ssDNA [39]. LwCas13a has been shown to collaterally degrade ssRNA after activation by a ssRNA target [78]. AbCas12g has been shown to collaterally degrade both ssDNA and ssRNA after activation by a ssRNA target [40]. Collateral nuclease activity has not been shown to be important for immune system functions for Cas12a and Cas12g [79]. However, in regards to Cas13a, collateral RNAse activity leads to cell dormancy [47]. To determine the existence and extent of collateral activity by SuCas12a2, recombinantly expressed SuCas12a2:crRNA was incubated with a target RNA and different non-target nucleic acids (FAM-labeled ssDNA, dsDNA, and ssRNA, or supercoiled plasmid) (**Figure 3-4A**). The reactions were quenched at different time points to show degradation over time. The results show complete collateral degradation of ssDNA, dsDNA, ssRNA, and supercoiled plasmid (**Figure 3-4B,C**). The ability to collaterally degrade plasmid after activation by target RNA supports the initial *in vivo* data that indicated clearance of a

target plasmid and could also explain why targeting caused cells to die or go dormant. Collaborators found that the presence or absence of a promotor region adjacent to the target sequence on the plasmid did not change the outcome. Cryptic transcription of the target plasmid is likely the cause for the presence of an RNA target in the cell when a promotor is not available.

The resulting rates of degradation suggest a faster degradation for single stranded species because at identical starting concentrations the single stranded nucleic acids are fully degraded much earlier in the time course (**Figure 3-4B**). However, the difference in rates of degradation of single stranded species vs double stranded species could be attributed to a 2x increase in the number of phosphodiester bonds for dsDNA at the same length as ssDNA. The slower degradation of dsDNA could also signify a mechanism in which the strands must first be separated before one strand is cleaved at a time in the RuvC active site similarly to Cas12a [80]. Collateral degradation of plasmid by SuCas12a2 shows a transition from supercoiled plasmid to a mixture of linear and nicked early in the reaction (**Figure 3-4C**). The collateral degradation of supercoiled plasmid eventually forms primarily linear plasmid, followed by complete degradation resulting in a smear (**Figure 3-4C**). The ability to cleave the supercoiled plasmid without access to free ends, show that SuCas12a2 acts as an endonuclease, and the transition from supercoiled to nicked to linear suggests nickase activity being the primary nuclease activity. Non-specific collateral degradation of plasmid has been observed in Cas12a [81]. At a molar ratio of protein-complex:supercoiled plasmid of ~2:1, SuCas12a2 was able to fully degrade plasmid in 1 hour. At a similar molar ratio, it took Cas12a 24 hours

to fully degrade supercoiled plasmid [81]. This suggests a faster rate of collateral nuclease activity of SuCas12a2 over Cas12a.

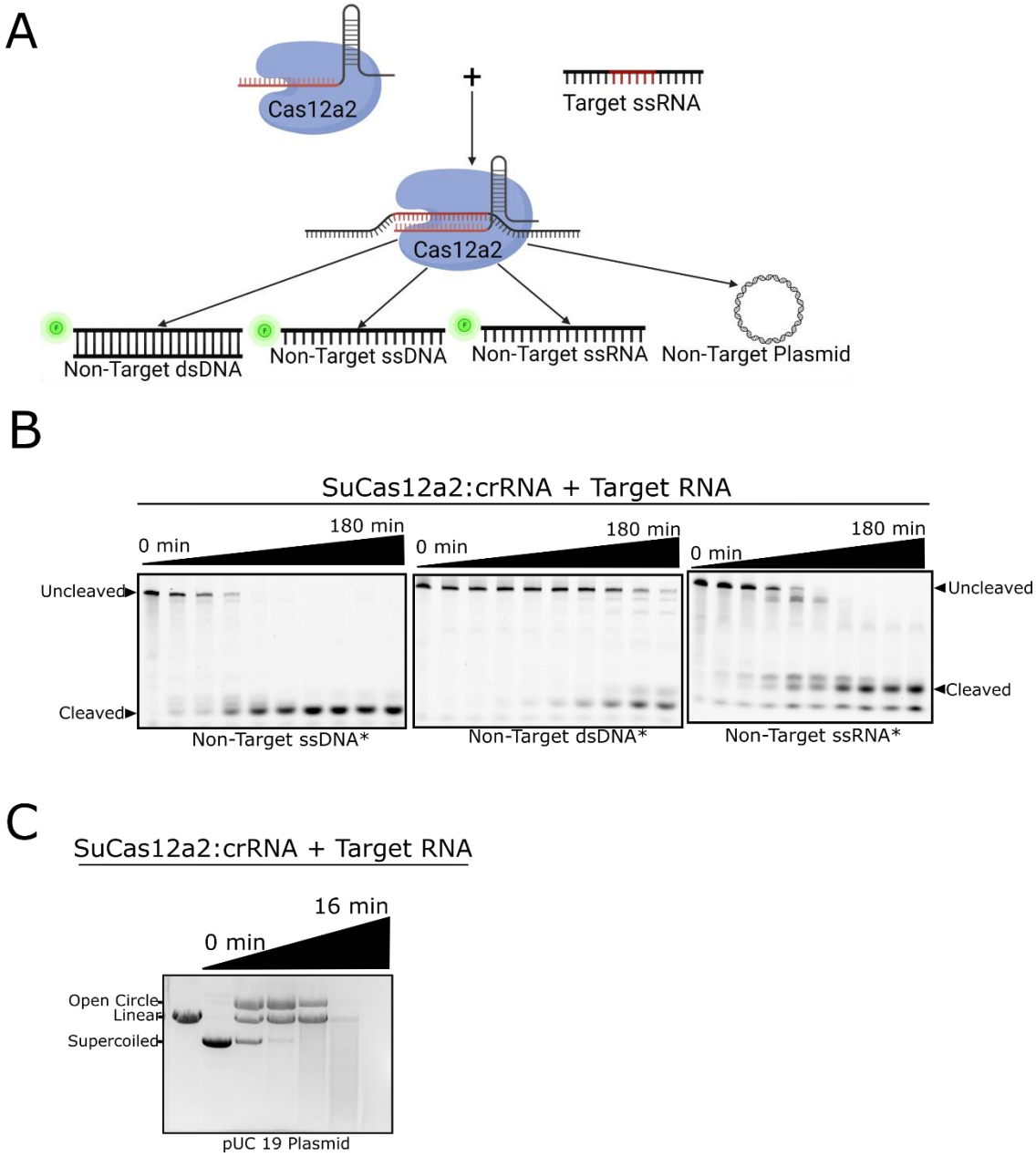


Figure: 3-4. Collateral Degradation. Once activated by target RNA, SuCas12a2 collaterally degrades ssDNA, ssRNA and dsDNA. (A) Workflow for SuCas12a2 collateral degradation assays. (B) Time course assays of collateral degradation of ssRNA, ssDNA, and dsDNA by SuCas12a2. Urea-PAGE of FAM-labeled collateral substrates (C) Time course of collateral cleavage of pUC 19 plasmid by SuCas12a2. Plasmid DNA visualized using ethidium bromide.

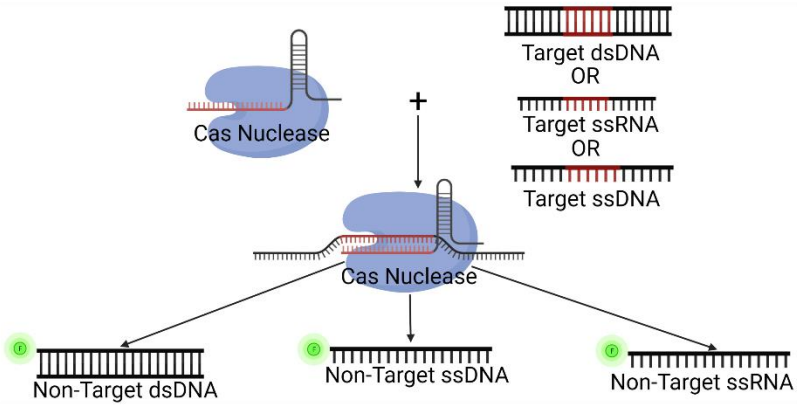
To obtain a direct comparison between SuCas12a2 and other Cas nucleases possessing collateral activity, an assay was developed in which SuCas12a2, LbCas12a, LwCas13a, and AbCas12g would be tested for collateral degradation of different 5'FAM labeled non-target nucleic acids (ssDNA, dsDNA, and ssRNA) after the addition of different unlabeled nucleic acid target activators (ssDNA, dsDNA, and ssRNA) (**Figure 3-5A**). The results are summarized in **Table 3-1** and mostly confirm what has been shown in the literature for Cas12a, Cas13a, and Cas12g [39], [40], [78]. Cas12a collaterally degraded ssDNA upon activation by ssDNA and dsDNA targets (**Figure 3-5B**). Interestingly, Cas12g at 50°C collaterally degraded ssDNA, and ssRNA upon activation by ssRNA target. However, there was also partial degradation of collateral ssDNA, and ssRNA after activation by ssDNA target. (**Figure 3-5B**). This is in contradiction with what was previously observed with Cas12g in that collateral activity was only observed after activation with ssRNA and not ssDNA [40]. Cas13a collaterally degrades ssRNA upon activation by a ssRNA target (**Figure 3-5B**). Activation of SuCas12a2 was limited to a ssRNA target. However, SuCas12a2 showed complete collateral degradation of ssRNA, ssDNA and dsDNA (**Figure 3-5B**). Ultimately, the observed collateral nuclease activity by SuCas12a2 is significant as no other Cas nuclease characterized to date can degrade ssDNA, ssRNA and dsDNA. While collateral

degradation of dsDNA is especially significant as it could lead to degradation of genomic DNA that could explain the abortive infection phenotype.

Table 3-1: Comparison of Collateral Nuclease Activity Among CRISPR Systems

Cas Enzyme	Target Substrate(s)	Collateral Substrate(s)
LbCas12a	ssDNA, dsDNA	ssDNA
LwCas13a	ssRNA	ssRNA
AbCas12g	ssDNA, ssRNA	ssDNA, ssRNA
SuCas12a2	ssRNA	ssDNA, ssRNA, dsDNA

A


Cas Nuclease
+
Target dsDNA
OR
Target ssRNA
OR
Target ssDNA
Cas Nuclease
Non-Target dsDNA
Non-Target ssDNA
Non-Target ssRNA

B

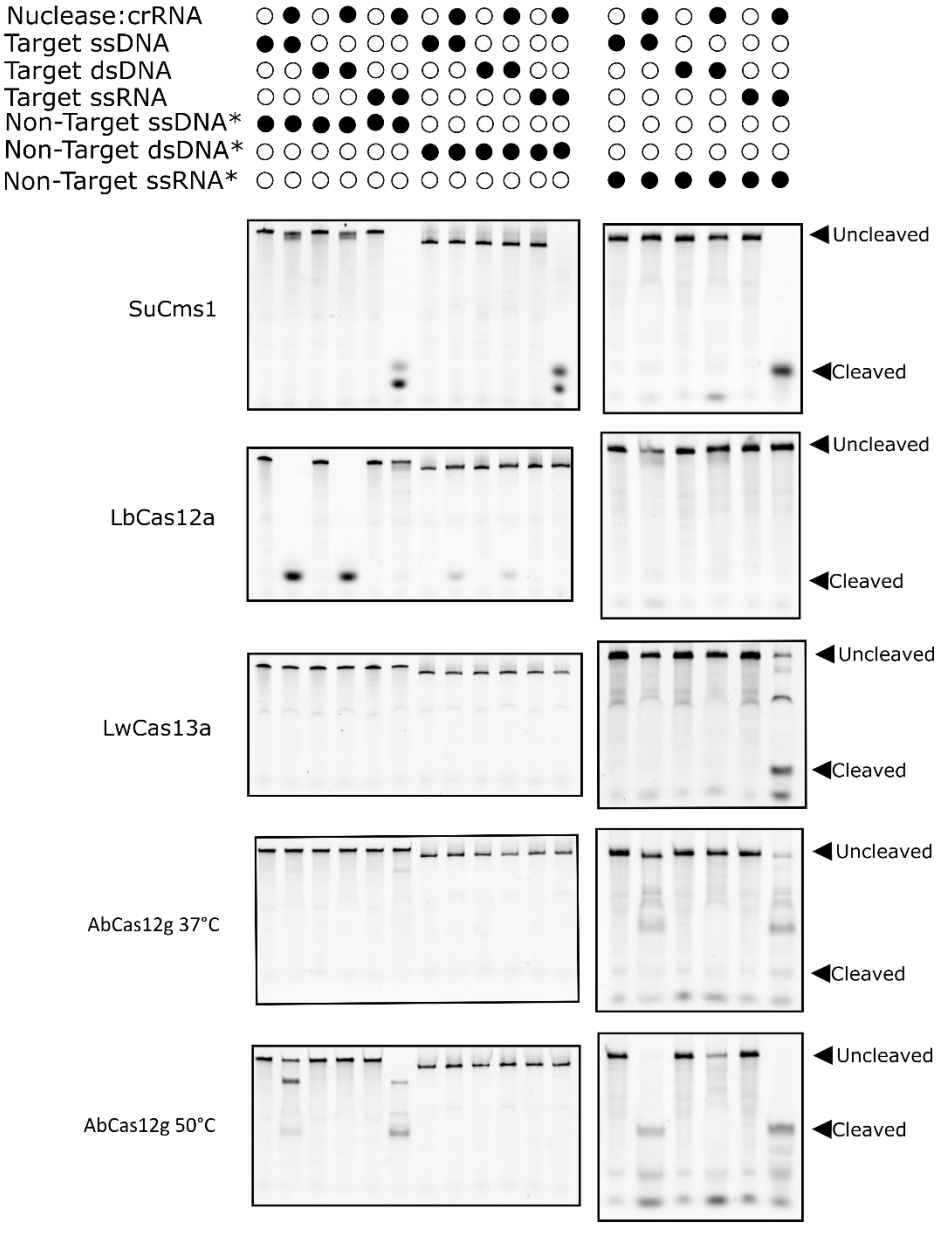

Nuclease:crRNA
Target ssDNA
Target dsDNA
Target ssRNA
Non-Target ssDNA*
Non-Target dsDNA*
Non-Target ssRNA*
SuCms1
LbCas12a
LwCas13a
AbCas12g 37°C
AbCas12g 50°C
Uncleaved
Cleaved

Figure: 3-5. A Comparison of Collateral Activity Between CRISPR Systems. The collateral activity of Cas12a2 is compared against Cas12a, Cas12g and Cas13 (A) Workflow for assay comparing activating and collateral substrates between SuCas12a2, LbCas12a, LwCas13a, and AbCas12g. (B) Collateral activity comparison assay between SuCas12a2, LbCas12a, LwCas13a, and AbCas12g testing both different activator substrates and different collateral substrates. Urea-PAGE gels show digestion of FAM-labeled collateral substrates * denotes FAM-labeled substrate. Reaction was run for 1 hour.

Cas12a2 Requires a Protospacer Flanking Sequence for Activity

The ability of Cas12a2 to collaterally degrade ssDNA, ssRNA and dsDNA substrates was not observed in other compared CRISPR systems. However, much about the mechanism remains unknown. Before the discovery of RNA targeting by Cas12a2, a collaborator conducted a plasmid curing assay using a PAM library (data not shown). The data suggest that Cas12a2 required a PAM sequence, although it tolerated a wide variety of different PAM sequences. With regards to RNA targeting, this suggests that a Protospacer Flanking Sequence (PFS) is important for Cas12a2 activation. To test this hypothesis, we monitored collateral cleavage of different non-target substrates (ssRNA, ssDNA, and dsDNA) by SuCas12a2:crRNA with a non-self-target with a GAAAG 3' PFS, derived from the PAM CTTTC previously shown to be a functioning PAM (data not shown), a self-target with a flanking sequence complementary to the direct repeat of the crRNA, and a target with no flanking sequence (**Figure 3-6A**). The results showed that collateral degradation of ssRNA, ssDNA, and dsDNA was only observed with the RNA target containing a GAAAG 3' PFS (**Figure 3-6B**). In comparison with other RNA targeting CRISPR systems the need for a PFS sequence to activate is uncommon. As PAM/PFS sequence preferences tend not to be found for non-self-sequences [40], [82]–[84]. The Type III-A,B systems have been shown to function regardless of the sequence

of the 3’ non-self-flanking sequence of the RNA target [11–13]. However, for the Cmr type III-B complex, the presence of 3’ flanking nucleotides is essential [82]. However, the non-self-sequence is irrelevant as a 6A sequence was effective at activating cleavage [82]. Data on the importance of 3’ flanking nucleotides was not found for the type III-A Csm complex. Cas12g, a type V-G protein also shows no evidence of a 3’ flanking sequence preference [40]. However, Cas13 requires a C, A, or U in the first position of the 3’ flanking sequence [78].

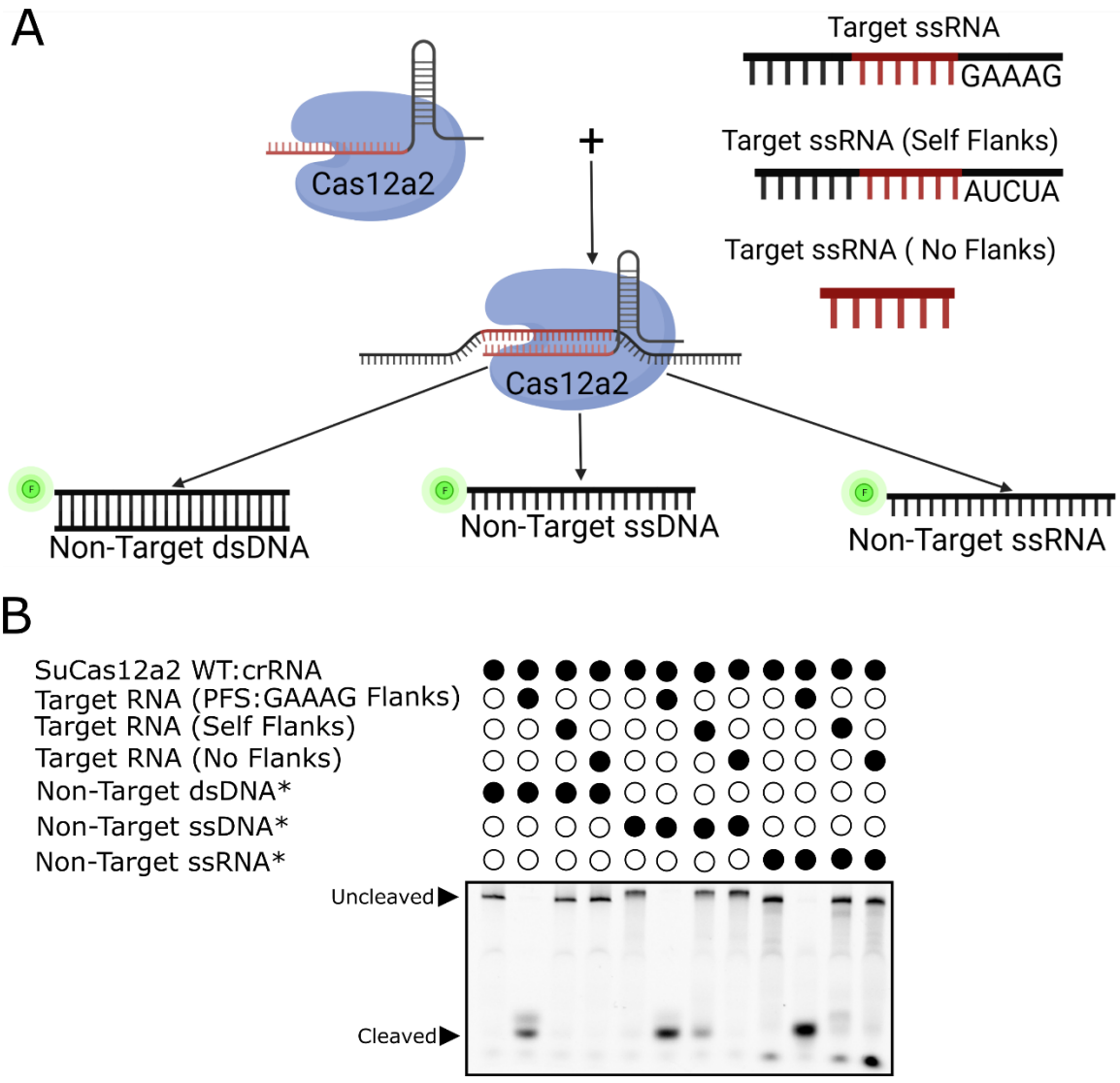


Figure: 3-6. Cas12a2 Requires PFS for Activation. SuCas12a2 activity relies on the presences of a protospacer flanking sequence. (A) Workflow for assay determining the effect of different target RNA flanks on activity. (B) Assay for determining the effect of different target RNA flanks on SuCas12a2 activity. Urea-PAGE gel shows digestion of FAM-labeled collateral substrates *denotes FAM-labeled substrate

Cas12a2 Cleavage/Degradation Activity is RuvC Mediated

The importance of the RuvC domain for nuclease activity in other CRISPR systems has been documented [16], [17], [40], [85]–[90]. The dependence on residues, within the RuvC domain, D848, E1063, and D1213 for activity of Cas12a2 has been observed *in vivo*. It has also been shown that the point mutation D1213A abolishes RNA degradation *in vitro* (**Figure 3-3BI**). However, the importance of these residues for collateral degradation of DNA *in vitro* is unknown. An assay was performed in which the collateral degradation of non-target dsDNA by SuCas12a2:crRNA (WT, D848A, E1063A, and D1213A) could be examined. We observed that each of the RuvC point mutations abolished activity *in vitro* (**Figure 3-7B**), suggesting that together they make up the nuclease active site.

A Zinc-Finger Motif Enhances Cas12a2 Activity

Additional sequence analysis and alignment revealed a putative zinc finger motif with four conserved cysteine residues in the form CXXC…..CXXC presumably making up the metal coordination site (**Figure 3-7C**). The presence of similar zinc-finger motifs has been found in other CRISPR systems and DNA binding proteins [89], [91]–[94]. In order to eliminate the ability of the zinc finger to coordinate a metal, the four cysteine residues, C1170, C1173, C1188, and C1191 were mutated to serines using site directed mutagenesis. The impact of these mutations was tested by monitoring the collateral degradation of non-target nucleic acid substrates (ssRNA, ssDNA, and dsDNA) by SuCas12a2:crRNA (WT, and C1170S/C1173S/C1188S/C1191S) after activation by a target RNA. The mutations reduced activity (**Figure 3-7D**). The highest impact of these residues does seem to be on the degradation of dsDNA. However, given that rates of

degradation are much faster for single-stranded substrates (**Figure 3-4B**), if mutation of the zinc-finger simply causes an overall reduction in rates across the board, then the amount of single stranded degradation would still be greater than the amount of double stranded degradation in the same amount of time.

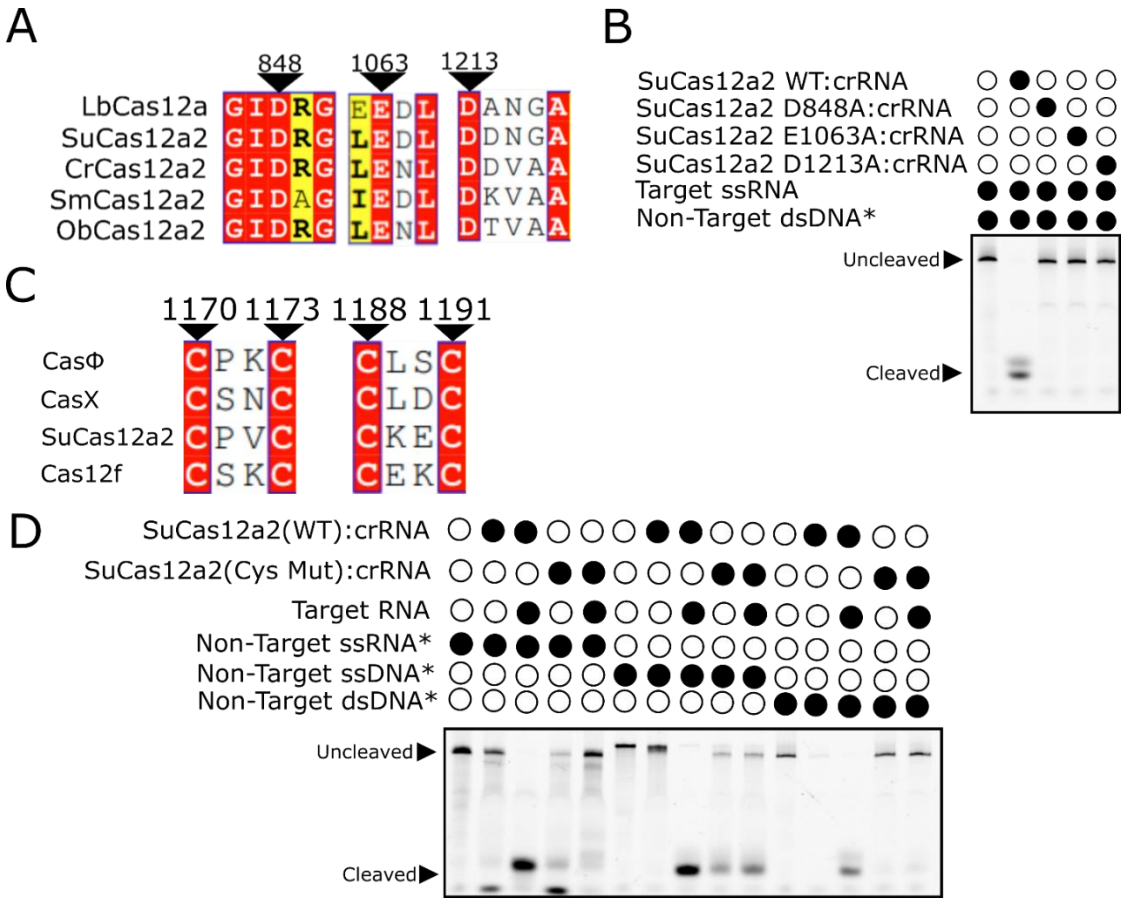


Figure: 3-7. Determination of Residues Important for Cas12a2 Function. SuCas12a2 function requires acidic residues in the RuvC domain and is enhanced by cysteine residues in the zinc-finger motif. (A) Alignment of acidic RuvC residues between CRISPR-Cas proteins. Acidic residues chosen for mutation are marked. (B) Collateral cleavage assay of dsDNA with Cas12a2 containing select point mutations in the RuvC nuclease domain. Urea-PAGE gel shows digestion of FAM-labeled collateral substrates *denotes FAM-labeled substrate. (C) Alignment of zinc-finger motif between different CRISPR-Cas proteins. Residues chosen for mutation are marked. (D) Collateral cleavage assay using Cas12a2 with cysteine residues in zinc-finger mutated. Urea-PAGE gel shows digestion of FAM-labeled collateral substrates *denotes FAM-labeled substrate

The Presence of Mg^{2+} is Important for Cas12a2 Activity

An additional common requirement for nuclease activity of other nucleases is a divalent cation present in the active site [16], [95]–[98]. The role of the divalent cation is to coordinate negatively charged amino acids and shield the negative charge of the phosphate backbone from the nucleophilic water molecule [77], [99]. Commonly, the divalent cation used is magnesium, although others have been shown to work [95]–[97]. To test the importance of magnesium on activity of SuCas12a2, recombinantly expressed and purified SuCas12a2:crRNA was first treated with EDTA to chelate the 2 mM Mg^{2+} found in the buffer. The protein sample was then dialyzed into a Mg^{2+} free buffer and used for the analysis of activity in the absence of magnesium. Collateral degradation of ssDNA and dsDNA by SuCas12a2:crRNA after activation by target RNA was monitored in the presence and absence of $MgCl_2$ as well as $ZnSO_4$, to rule out the possibility of chelated Zn^{2+} from the zinc finger disrupting activity. For collateral degradation of ssDNA, $MgCl_2$ was required for activity. However, there appeared to be some background collateral dsDNA activity in the absence of $MgCl_2$, but the addition of $MgCl_2$ enhanced activity (**Figure 3-8**). It seems unlikely that $MgCl_2$ would be more critical for degradation of ssDNA than dsDNA, so the background dsDNAse activity could most likely be attributed to magnesium contamination. Interestingly, regardless of the presence of $MgCl_2$, if $ZnSO_4$ was present in the reaction, activity was completely inhibited (**Figure 3-8**). To shift the equilibrium in favor of Mg^{2+} binding in the RuvC active site, a 10-fold molar excess of $MgCl_2$ was used over $ZnSO_4$. However, this excess may not have been enough given the stability of transition-metal complexes predicted by the Irving-William

series that predicts a higher stability for Zn (II) complexes over Mg (II) complexes [100]. This most likely translates into a higher affinity for Zn^{2+} for the active site. Regardless, these observations suggest that Cas12a2 functions using a similar magnesium dependent mechanism as other Cas nucleases.

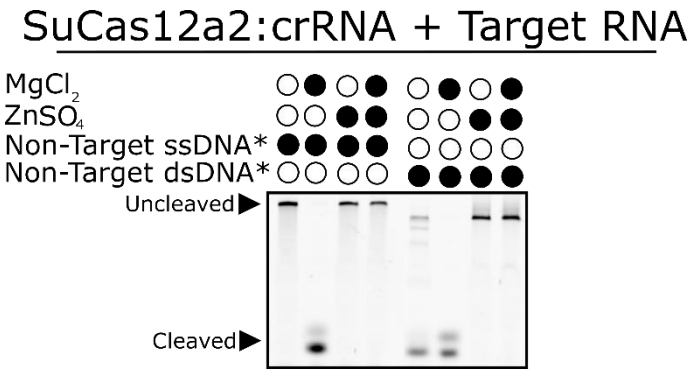


Figure: 3-8 Dependence of Divalent Cation on SuCas12a2 Activity. SuCas12a2 function is enhanced by the presence of Mg^{2+} and is inhibited by the presence of Zn^{2+}. Collateral cleavage assay in the presence or absence of Mg^{2+} and Zn^{2+}. Urea-PAGE gel shows digestion of FAM-labeled collateral substrates *denotes FAM-labeled substrate.

Conclusion

We believe the discovery and characterization of the Cas nuclease Cas12a2 will lead to many advancements in the field of biotechnology. Just as the discovery of RNA guided dsDNA cleavage activity of Cas9 led to advancements in genome editing [14]–[16], [20], and the discovery of collateral ssDNA degradation activity of Cas12a led to advancements in diagnostic tests [39], [101]–[104]. We believe that the novel activity of Cas12a2 could be repurposed. For instance, the detection of the presence of RNA viruses in a diagnostic kit could be made easier using Cas12a2 as it targets RNA directly and

would not require a reverse transcription step as it does with Cas12a [101], [102]. The much broader range of collateral nuclease activity by Cas12a2 over Cas12a also gives it the advantage of using a greater variety of reporter molecules. Additionally, the compatibility of Cas12a2 with Cas12a crRNAs would allow for the direct integration of Cas12a2 into already implemented Cas12a based diagnostic kits. In addition to its potential to greatly impact the diagnostic realm, Cas12a2 may also show promise in its mechanism of targeted cell death. The ability to cause cell death based on the presence of specific RNA could be beneficial in terms of cancer therapeutics, as well as the ability to eliminate a population of unedited cells transfected with Cas9 or Cas12a for the purpose of gene knockouts. However, given these potential use cases, there is still much to learn about Cas12a2. Although Cas12a2 mediated cell death has been observed in bacteria, it has yet to be tested in human cells. A larger genome and more complex DNA repair mechanisms could possibly overcome the collateral nuclease activity of Cas12a2 that leads to cell death. Similar to other CRISPR tools, there is also the issue of off-target effects [81], [105]–[107]. Given that Cas12a2 has much less stringent PAM/PFS requirements, and has been shown to have a higher tolerance for mismatches, off-target effects may be an even bigger issue with Cas12a2. By learning more about Cas12a2, there may be potential to enhance its activity and specificity, potentially mitigating its downsides just as a better understanding of Cas9 and Cas12a has led to enhancements in their activity [108]–[114]. The work presented here ultimately goes to show that there is still much to be uncovered in the CRISPR field, and the next revolutionary tool may be just around the corner.

Methods

Plasmids and nucleotide sequences used in this chapter can be found in Table B-1 and Table B-2 respectively.

Cas12a2 DNA and RNA Phenol:Chloroform Extraction Technique

A volume of phenol/chloroform/isoamyl alcohol (25:24:1, pH 6.8/8.0) equal to the sample volume was added to the sample. The sample was mixed by flicking the bottom of the tube. The layers were separated by spinning the tube in a mini-spin-6 centrifuge for about 30 seconds or more. About 50-70 percent of the top (aqueous) layer was removed and put in a new tube. The sample was then added to the appropriate gel loading dye for analysis.

Cas12a2 RNA Phenol:Chloroform Extraction Technique

A volume of acidic phenol equal to the sample volume was added to the sample. The sample was mixed by flicking the bottom of the tube. The layers were separated by spinning the tube in a mini-spin-6 centrifuge for about 30 seconds or more. About 50-70 percent of the top (aqueous) layer was removed and put in a new tube. An equal volume of chloroform was added to the sample and mixed by flicking the tube. The layers were separated by spinning the tube in a mini-spin-6 centrifuge for about 30 seconds or more. About 50-70 percent of the top (aqueous) layer was removed and put in a new tube. The sample was then added to the appropriate gel loading dye for analysis.

Urea-PAGE

The following is the recipe for two 12% UREA-PAGE gels. In a 50 mL conical tube the following containing 9.6 g urea, 4 mL of 5x TBE, 6 mL of 40% acrylamide, and

water to a volume of 20 mL was prepared. The solution was mixed by vortexing then allowed to warm up to room temperature or until the urea completely dissolved. 20 μL of TEMED and 160 μL of 10% Ammonium Persulfate (APS) were added to initiate polymerization. The solution was pipetted in between two glass plates to form the gel. Once polymerized, the gel plates were rinsed and a gel running apparatus was setup. The wells of the gel were rinsed with 1x TBE running buffer and pre ran at 200 V for at least 30 min, or until the samples were ready for analysis. Once the samples were added, the gel was ran at 200 V for about 30 min.

Analysis of Target RNA Protection by Cas12a2

In multiple 10 μL reactions, the following containing: 150 nM 5' and 3' FAM labeled target RNA, and 200 nM SuCas12a2 (D1213A):crRNA were prepared in 1x NEB 3.1 and incubated at 37°C for 30 min. To select reactions, SuCas12a2 (WT):crRNA was added to a concentration of 150 nM. To other reactions target ssRNA was added to a concentration of 200 nM. The reactions were incubated at 37°C for an additional hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Determination of Nucleic Acid Target Substrate

In a 10 μL reaction, the following containing: 1x NEB 3.1, 100 nM FAM labeled crRNA complementary target nucleic acid (ssRNA, ssDNA, or dsDNA), 250 nM SuCas12a2:crRNA was prepared in 1x NEB 3.1. The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using the FDF-PAGE method outlined in [115] and visualized for fluorescein fluorescence.

Cas12a2 Analysis of Collateral Cleavage of Labeled Nucleic Acids

In a 100 μL reaction, the following containing: 100 nM FAM labeled non-target nucleic acid substrates (ssRNA, ssDNA, or dsDNA), 100 nM target ssRNA, 100 nM SuCas12a2:crRNA was prepared in NEB 3.1. The reactions were incubated at 37°C for three hours. Timepoints were taken by extracting 10 μL from the master mix and phenol:chloroform extracting at 1, 2, 5, 10, 15, 30, 60, 120 and 180 minutes. The reactions were analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Cas12a2 Collateral Cleavage Assay of Supercoiled Plasmid

In a 100 μL reaction, the following containing 7 nM supercoiled pUC 19 plasmid, 25 nM target RNA, and 14 nM SuCas12a2:crRNA was prepared in 1x NEB 3.1. The reaction was incubated at 37°C and at 1, 2, 4, 6, 10, 15, 25, 40, and 60 minutes timepoints were taken by removing 10 μL from the master reaction and phenol:chloroform extracting. The reactions were analyzed using 1% agarose stained with ethidium bromide.

Cas12a2, Cas12a, Cas13a, and Cas12g Activity Comparison

In a 10 μL reaction, the following containing 200 nM target nucleic acid (ssDNA, ssRNA, or dsDNA), 100 nM FAM labeled collateral substrate (ssDNA, ssRNA, or dsDNA), 250 nM EnGen LbaCas12a (New England Biolabs (M0653S), and 500 nM of its cognate crRNA was prepared in 1x NEB 2.1. The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

In a 10 μL reaction, the following containing 200 nM target nucleic acid (ssDNA, ssRNA, or dsDNA), 100 nM FAM labeled collateral substrate (ssDNA, ssRNA, or dsDNA), 250 nM LwCas13a (Molecular Cloning Laboratories Cas13a-100), and 500 nM of its cognate crRNA was prepared in 1X Cas9 buffer (20 mM HEPES (pH 6.5), 5 mM $MgCl_2$, 100mM NaCl, 100 μM EDTA). The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

In a 10 μL reaction, the following containing 200 nM target nucleic acid (ssDNA, ssRNA, or dsDNA), 100 nM FAM labeled collateral substrate (ssDNA, ssRNA, or dsDNA), 250 nM AbCas12g (Addgene plasmid #120879, expressed and purified by Thom Hallmark), 500 nM of its cognate crRNA, and 1 μM of its cognate tracrRNA was prepared in 1x NEB 3.1. The reaction was incubated at 37°C or 50°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

In a 10 μL reaction, the following containing 200 nM target nucleic acid (ssDNA, ssRNA, or dsDNA), 100 nM FAM labeled collateral substrate (ssDNA, ssRNA, or dsDNA), and 250 nM SuCas12a2:crRNA was prepared in 1x NEB 3.1. The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Analysis of Target RNA Flanking Region Importance

In a 10 μL reaction, the following containing 300 nM target RNA (PFS GAAAG flanks, self-flanks, or no flanks), 100 nM FAM labeled collateral substrate (ssDNA, ssRNA, or dsDNA), and 250 nM SuCas12a2:crRNA was prepared in 1X NEB 3.1. The

reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Analysis of RuvC Point Mutations on dsDNA Collateral Cleavage

In a 10 μL reaction, the following containing 300 nM target RNA, 100 nM FAM labeled dsDNA collateral substrate and 250 nM SuCas12a2:crRNA (WT, D848A, E1063A, or D1213A) was prepared in 1X NEB 3.1. The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Analysis of Zinc Finger Cysteine Mutations on Collateral Cleavage

In a 10 μL reaction, the following containing 300 nM target RNA, 100 nM FAM labeled collateral substrate (ssDNA, ssRNA, or dsDNA) and 250 nM SuCas12a2:crRNA (WT, C1170S/C1173S/C1188S/C1191S) was prepared in 1X NEB 3.1. The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Analysis of Dependency of Metal Ions on Cleavage Activity

Metal ions were chelated from concentrated samples of SuCas12a2:crRNA by incubating on ice for 5-10 min with 5 mM EDTA. Sample was added to a Slide-A-Lyzer dialysis cup (Thermofisher Scientific) and placed in a 1.5 mL microcentrifuge tube filled with 750 μL buffer (25 mM HEPES, pH 7.2, 150 mM KCl). The dialysis apparatus was incubated on a rocker at 4°C and switched to a tube with fresh buffer after 1 hour then left to incubate overnight. The protein sample was extracted from the dialysis cup and its $A_{280\ nm}$ measured. In a 10 μL reaction, the following containing 300 nM target RNA, 100

nM FAM labeled collateral substrate (ssDNA, dsDNA), 200 nM SuCas12a2:crRNA (EDTA treated and dialyzed), and divalent cations (10 mM $MgCl_2$ and/or 1 mM $ZnSO_4$) was prepared in 1X cleavage buffer (50 mM Tris-HCl, pH 7.9, 100 mM NaCl, 100 μg/mL BSA). The reaction was incubated at 37°C for 1 hour and then phenol:chloroform extracted and analyzed using 12% Urea-PAGE and visualized for fluorescein fluorescence.

Acknowledgments

Figures 3-1A, 3-2A, 3-3A, 3-4A, 3-5A, and 3-6A were created using Biorender.com

CHAPTER 4

SUMMARY AND FUTURE DIRECTIONS

Abstract

The work I have done in characterizing the Type IV-B and Type V-A2 CRISPR systems has both set the groundwork for others in the Jackson Lab to build upon and led to a better understanding of these immune systems. The function of newly discovered Type IV-B and Type V-A2 was evaluated in chapters 2 and 3. In chapter 2, to determine if the Type IV-B system could act as an anti-CRISPR, a plasmid clearance assay showed inconsistent results that was due to the use of an inferior non-self PAM sequence with the type IV-A system. A potential function of type IV-B ancillary protein CysH was analyzed using sequence alignments and proposed to function through nucleic acid adenylation using pyrophosphatase activity. In chapter 3, the function and mechanism of type V-A2 protein Cas12a2 was determined using *in vitro* biochemical assays. Cas12a2 was determined to target RNA that subsequently activated collateral nuclease activity to degrade ssDNA, ssRNA and dsDNA that led to an abortive infection phenotype in cells. Although this thesis has revealed much about the type IV-B and type V-A2 CRISPR systems, many mechanistic questions remain to be answered. In this chapter I highlight the next steps for each research project and offer potential directions to learn more about the mechanism of Cas12a2 activation, possible considerations when developing tools that utilize Cas12a2, and ways to test the hypothesis of an anti-CRISPR mechanism of the type IV-B system and its ancillary protein CysH.

Summary

Mechanism of Cas12a2

The work described in this thesis revealed many significant components of the function and mechanism of Cas12a2. Its ability to facilitate immunity starts at its ability to process its own RNA guide. From there, Cas12a2 finds a target RNA sequence that is complementary to the guide RNA and binds. An activation event happens, presumably from the binding energy provided through interaction with the target RNA, and Cas12a2 becomes a rampant indiscriminate nuclease. The nuclease activity of Cas12a2 functions using magnesium dependent activity facilitated through the RuvC domain. It appears that a zinc-finger motif also assists in activity. *In vivo*, the end result for Cas12a2 nuclease activity is cellular dormancy. Even with all these discoveries, many mechanistic questions remain.

Cas12a2 Mammalian Cell Experiments

As shown in chapter 2, We found no evidence of genome editing by SuCas12a2 in the form of insertions/deletions (indels) in *Mesocricetus auratus.* However, there may be promise in targeted cell death experiments. Preliminary data showed a decrease in live cell counts for cells transfected with SuCas12a2 in comparison to non-transfected and GFP transfected controls. Additional controls and replicates are required for a more definitive result.

Type IV-B Ancillary Protein CysH

Protein sequence between Cas-CysH and Non-Cas-CysH sequences revealed Cas-CysH sequences do not possess the active site residues ECGLH necessary for the canonical Paps/aps reductase activity of CysH. Instead, Cas-CysH sequences contain a pyrophosphatase loop (pp-loop) with the consensus sequence SxGxDS that is not found in Non-Cas-CysH.

Type IV-B CRISPR System

The ability of the type IV-B system to function as an anti-CRISPR was tested using a plasmid curing assay. The results were inconclusive because a working type IV-A immune system control was not available. This rendered the results obtained with the type IV-B system irrelevant. Since leaving the Jackson Lab, the inconsistent results were determined to be due to an inferior PAM sequence in the target plasmid. Since modifying the PAM sequence, consistent type IV-A plasmid clearance activity has been restored, suggesting that a return to this experiment will yield more concrete results.

Future Directions

Cas12a2 Mechanism of Activation

The activation mechanism of Cas12a2 is hypothesized to be target RNA binding. However, it is not known if binding is sufficient for activation, or if target RNA cleavage is required. It has been shown that the observed cleavage of target RNA by Cas12a2 is in trans. (**Figure 3-3**). A possible experiment to determine if cleavage in cis is important would be to utilize a target RNA with a phosphorothioate backbone. A cleavage resistant RNA target would be a possible choice for evaluating the importance of target RNA

cleavage on activation. If cleavage in cis is important for activation of collateral nuclease activity, then rates of collateral activity using a phosphorothioate target RNA would be reduced compared with rates of collateral activity using normal phosphate backboned substrates. An assay that compares rates of collateral activity after activation by either target RNA with a phosphate or a phosphorothioate backbone should be able to determine if binding or cleavage of the target RNA is required for activation of Cas12a2.

In chapter 3 we show that once RNA targets are bound by Cas12a2 they are inaccessible for activation of another molecule of Cas12a2 (**Figure 3-3**). A possible explanation is that the target sequence and/or PFS are protected within the binding site of Cas12a2. It would be interesting to know the degree of protection and specifically what residues are protected. An option I could think of to answer this question would be an RNA footprinting assay in which target RNA bound to Cas12a2 is subjected to hydroxy-radical cleavage. Protected areas of the RNA should be able resistant to cleavage. Gel electrophoresis could then be used to determine the size of the protected section of RNA.

According to *in vivo* data generated by collaborators, the protospacer flanking sequence is important for Cas12a2 activation, and Cas12a2 does show sequence preference in the flanking region [116] In chapter 3, we confirm that the presence of a PFS is important for activation of Cas12a2 with an *in vitro* assay (**Figure 3-6**). However, the various sequence preferences could be analyzed with an *in vitro* assay to determine why some PFS sequences appear to be more activating than others in the *in vivo* assay. We were able to show that a flanking sequence was required for activation and that the flanking sequence could not be the reverse complement of the direct repeat (self). One question that remains is what would happen if we included a target RNA with a flanking

sequence shown not to be activating based on *in vivo* data, or if we included a target RNA with a poly-A flanking sequence. Would the *in vitro* data agree with what has been shown with *in vivo* data?

The tolerance of mismatches for Cas12a2 was analyzed using cell based assays by collaborators [116]. They found that when using a 24 nt spacer in the crRNA and mutating select residues along the spacer and measuring immune system activity using plasmid clearance assays that mismatches are less tolerated on the 3' end of the spacer in the PFS distal region and that four mismatches throughout the sequence were required to abolish immune system function. It would be interesting to see how this translates to an *in vitro* experiment as well as what the length of the spacer says about the mismatch tolerance of Cas12a2. Perhaps a longer spacer would allow for more mismatches because the extra binding energy provided by the additional nucleotides is sufficient for activation. It would be interesting if it was possible to determine the number of crRNA:target RNA base pairs necessary for activation. A structural approach to determine the effect of a mismatched target RNA on the amino acid residues that interact with the target RNA:crRNA duplex could be insightful for determining how Cas12a2 is activated. Specifically, determining structures of Cas12a2:crRNA:target RNA with different degrees of complementarity between the target RNA and crRNA and evaluating the effect on amino acid residues that interact with crRNA:target RNA duplex. It could also be determined which specific interactions between the crRNA:target RNA duplex and the amino acid residues are required to produce conformational changes seen with target RNA binding in the recent Cryo-EM structure[117]. By understanding more about

how Cas12a2 tolerates mismatches to its target RNA, more can be understood about the potential off-target effects of Cas12a2 when utilizing in mammalian cells.

Cas12a2 Potential Applications

In terms of the targeted cell death experiment performed in chapter 2, more replicates need to be done to show precision in results. A few controls should also be done to show cell death is dependent on SuCas12a2 activity including: comparisons between SuCas12a2:crRNA, apo SuCas12a2, SuCas12a2 and a non-targeting crRNA, and SuCas12a2 (RuvC mut):crRNA. Additionally, it may be necessary to shift to human cells and target genes with a generally higher-level basal expression for a higher chance of success.

Thom Hallmark, in the Jackson Lab, has had success with using Cas12a2 as a tool for nucleic acid detection by using RNAse alert and DNAse alert reporter molecules. Upon successful recognition of a target RNA, Cas12a2 degrades the reporter molecules, which produces a detectable fluorescent signal [116]. The use of CRISPR systems as tools for nucleic acid detection has been shown before [39], [102]. Cas12a2 could be particularly useful because it can directly recognize RNA without the need for a reverse transcription step as is required when using Cas12a [102]. Additionally, the ability of Cas12a2 to function using Cas12a crRNAs would mean easy integration of Cas12a2 into already designed Cas12a detection kits. Additionally, the collateral activity of Cas12a2 allows a broader selection of reporter molecules including ssDNA, RNA and dsDNA. However, the potential applications of Cas12a2 should not be limited to a tool for nucleic acid detection because it may also possess potential as a disease therapeutic.

One of the most intriguing potential applications of Cas12a2 could be the ability to target and kill select cells. In the case of a viral infection, Cas12a2 could become active by targeting viral RNAs, resulting in the early death of the host cell before the virus replicates and infects neighboring cells. Cas12a2 may also have potential in cancer therapeutics, gene-editing enrichment, and fighting auto-immune disorders. By targeting a gene expressed primarily in cancer cells, Cas12a2 could kill off cancerous cells, while leaving healthy cells intact. Before such therapies become widespread, it would be necessary to determine the cytopathic effect Cas12a2 has in cells in non-targeting conditions. A Cas12a2 therapeutic that is toxic to all cells defeats the purpose of its targeted cell death application.

With each of these potential applications in mind, an important consideration should not be overlooked. SuCas12a2 required four mismatches in its target RNA before it wasn't activated [116]. This could potentially lead to off target effects. Therefore, a more stringent variant may need to be found or engineered.

Determination of Type IV-B Ancillary Protein CysH Function

After successful expression and purification of Cas-CysH, its role in the function of the Type IV-B system should be evaluated by determining its function *in vitro*. The discovery of a conserved pyrophosphatase loop in its sequence could signal a good starting point for the type of biochemical activity to evaluate. In enzymology, pyrophosphatase activity usually involves the transfer of AMP from ATP to another molecule. I hypothesize that Cas-CysH functions by adenylating nucleic acids. To test this hypothesis, an assay that analyzes the migration of alpha p32 labeled ATP on a gel after the incubation with CysH and nucleic acid could be utilized. The putative

pyrophosphatase residues in CysH have already been identified in the sequence. Site-directed mutagenesis of these residues and subsequent analysis of CysH activity is also important in determining the mechanism of CysH.

Evaluation of Type IV-B CRISPR System Anti-CRISPR Potential

Since finishing up my work in the Jackson Lab, the issues with the Type IV-A plasmid curing assay have been solved by discovering the PAM sequence used in the assay was incorrect. With a proper PAM sequence in place the assay is very reproducible and a plasmid curing assay with the Type IV-B system should be repeated. Before running the assay, I believe some modifications would be beneficial.

First, the type IV-B system should be cloned onto the non-target plasmid along with the target plasmid for consistency in transformation efficiency. Second, Type IV-B ancillary proteins such as CysH should be included in the assay to test their importance.

If the Type IV-B system is shown to function as an anti-CRISPR system, the mechanism for CRISPR inactivation should also be determined. The current hypothesis is that the type IV-B Cas proteins bind to cognate crRNAs of the host cell CRISPR system to prevent the immune system from functioning. A method for testing this hypothesis could be to mutate the RNA binding residues of the type IV-B complex and evaluate the effect on the anti-CRISPR activity. The role that ancillary protein CysH plays in anti-CRISPR activity could also be determined. A possible role of CysH could be to adenylate host cell crRNAs. Perhaps the true mechanism of type IV-B anti-CRISPR function is to bind to host cell crRNAs, adenylate, or even poly-adenylate the end(s) of the crRNA and prevent the host cell CRISPR system from utilizing the modified crRNAs.

Conclusion

The work highlighted in chapter 2 on the evaluation of the function of the type IV-B system set the groundwork for others in the Jackson lab to pursue. The work I did on designing an assay to evaluate type IV-B system function can now be used to test the hypothesis that the type IV-B system functions as an anti-CRISPR now that the use of the correct PAM sequence of the type IV-A system corrected issues with the reproducibility of its plasmid clearance assay. The sequence analysis of type IV-B ancillary protein CysH reveals a putative pyrophosphatase activity that could play a role in the function of the type IV-B immune system function as a whole.

The discovery of the function of type V-A2 CRISPR protein Cas12a2 highlighted in chapter 3 identifies a novel mechanism of CRISPR immunity that utilizes RNA targeting and subsequent destruction of dsDNA to elicit an abortive infection mechanism of immunity. The potential for applications of Cas12a2 including RNA detection, and targeted cell death is astounding. However, given that the activation of Cas12a2 collateral nuclease activity results in the demise of the cell, learning more about how Cas12a2 is activated and discovering ways to more precisely control its activation is important for safe and effective use in mammalian cells.

REFERENCES

[1] A. Stern and R. Sorek, “The phage-host arms race: Shaping the evolution of microbes,” *BioEssays*, vol. 33, no. 1, pp. 43–51, Jan. 2011.

[2] H. G. Hampton, B. N. J. Watson, and P. C. Fineran, “The arms race between bacteria and their phage foes,” *Nature*, vol. 577, no. 7790, pp. 327–336, 2020.

[3] S. Chibani-chennoufi, A. Bruttin, M. Dillmann, and H. Bru, “MINIREVIEW Phage-Host Interaction : an Ecological Perspective,” vol. 186, no. 12, pp. 3677–3686, 2004.

[4] Ø. Bergh, K. Y. Børsheim, G. Bratbak, and M. Heldal, “High abundance of viruses found in aquatic environments,” *Nature*, vol. 340, no. 6233, pp. 467–468, 1989.

[5] M. R. Tock and D. T. F. Dryden, “The biology of restriction and anti-restriction,” *Curr. Opin. Microbiol.*, vol. 8, no. 4, pp. 466–472, 2005.

[6] A. Lopatina, N. Tal, and R. Sorek, “Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy,” *Annu. Rev. Virol.*, vol. 7, pp. 371–384, 2020.

[7] R. Barrangou *et al.*, “CRISPR provides acquired resistance against viruses in prokaryotes.,” *Science*, vol. 315, no. 5819, pp. 1709–12, Mar. 2007.

[8] A. L. Borges, A. R. Davidson, and J. Bondy-Denomy, “The Discovery, Mechanisms, and Evolutionary Impact of Anti-CRISPRs,” *Annu. Rev. Virol.*, vol. 4, no. 1, pp. 37–59, 2017.

[9] D. H. Kruger and T. A. Bickle, “Bacteriophage survival: Multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts,” *Microbiol. Rev.*, vol. 47, no. 3, pp. 345–360, 1983.

[10] S. N. Cohen, A. C. Y. Chang, H. W. Boyer, and R. B. Helling, “Construction of biologically functional bacterial plasmids in vitro,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 70, no. 11, pp. 3240–3244, 1973.

[11] V. B. P. E.-R. Savva, “Biotechnology Tools Derived from the Bacteriophage/Bacteria Arms Race,” Rijeka: IntechOpen, 2020, p. Ch. 3.

[12] S. J. Unterholzner, B. Poppenberger, and W. Rozhon, “Toxin–antitoxin systems,” *Mob. Genet. Elements*, vol. 3, no. 5, p. e26219, 2013.

[13] J. Wyman *et al.*, “Multiplex Genome Engineering Using CRISPR/Cas Systems,” no. February, pp. 819–824, 2013.

[14] P. Mali *et al.*, “RNA-guided human genome engineering via Cas9,” *Science (80-.).*, vol. 339, no. 6121, pp. 823–826, 2013.

[15] M. Jinek, A. East, A. Cheng, S. Lin, E. Ma, and J. Doudna, “RNA-programmed genome editing in human cells,” *Elife*, vol. 2013, no. 2, pp. 1–9, 2013.

[16] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier, "A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity," vol. 337, no. August, pp. 816–822, 2012.

[17] B. Zetsche *et al.*, "Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System," *Cell*, vol. 163, no. 3, pp. 759–771, Oct. 2015.

[18] M. Kordyś, R. Sen, and Z. Warkocki, "Applications of the versatile CRISPR-Cas13 RNA targeting system," *Wiley Interdiscip. Rev. RNA*, no. April, pp. 1–30, 2021.

[19] Y. Zhang, Junxia; You, "CRISPR-Cas13a system: a novel approach to precision oncology," *Cancer Biol. Med.*, vol. 17, no. 1, pp. 6–8, 2020.

[20] L. Cong *et al.*, "Multiplex Genome Engineering Using CRISPR/Cas Systems," *Science (80-.).*, vol. 339, no. 6121, pp. 819 LP – 823, Feb. 2013.

[21] N. Tal and R. Sorek, "SnapShot: Bacterial immunity," *Cell*, vol. 185, no. 3, pp. 578-578.e1, 2022.

[22] D. H. Parma, M. Snyder, S. Sobolevski, M. Nawroz, E. Brody, and L. Gold, "The Rex system of bacteriophage λ: Tolerance and altruistic cell death," *Genes Dev.*, vol. 6, no. 3, pp. 497–510, 1992.

[23] M. Amitsur, I. Morad, D. Chapman-Shimshoni, and G. Kaufmann, "HSD restriction - Modification proteins partake in latent anticodon nuclease," *EMBO J.*, vol. 11, no. 8, pp. 3129–3134, 1992.

[24] F. Depardieu *et al.*, "A Eukaryotic-like Serine/Threonine Kinase Protects Staphylococci against Phages," *Cell Host Microbe*, vol. 20, no. 4, pp. 471–481, 2016.

[25] M. Koga, Y. Otsuka, S. Lemire, and T. Yonesaki, "Escherichia coli rnlA and rnlB compose a novel toxin-antitoxin system," *Genetics*, vol. 187, no. 1, pp. 123–130, 2011.

[26] D. Cohen *et al.*, "Cyclic GMP–AMP signalling protects bacteria against viral infection," *Nature*, vol. 574, no. 7780, pp. 691–695, 2019.

[27] A. Bernheim *et al.*, "Prokaryotic viperins produce diverse antiviral molecules," *Nature*, vol. 589, no. 7840, pp. 120–124, 2021.

[28] F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, and E. Soria, "Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements," *J. Mol. Evol.*, vol. 60, no. 2, pp. 174–182, 2005.

[29] A. Bolotin, B. Quinquis, A. Sorokin, and S. Dusko Ehrlich, "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin," *Microbiology*, vol. 151, no. 8, pp. 2551–2561, 2005.

[30] R. Jansen, J. D. A. Van Embden, W. Gaastra, and L. M. Schouls, "Identification of genes that are associated with DNA repeats in prokaryotes," *Mol. Microbiol.*, vol.

43, no. 6, pp. 1565–1575, 2002.

[31] E. V. Koonin, K. S. Makarova, and F. Zhang, “Diversity, classification and evolution of CRISPR-Cas systems,” *Curr. Opin. Microbiol.*, vol. 37, pp. 67–78, 2017.

[32] S. A. Jackson, R. E. McKenzie, R. D. Fagerlund, S. N. Kieper, P. C. Fineran, and S. J. J. Brouns, “CRISPR-Cas: Adapting to change,” *Science (80-.).*, vol. 356, no. 6333, 2017.

[33] E. Charpentier, H. Richter, J. van der Oost, and M. F. White, “Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity,” *FEMS Microbiol. Rev.*, vol. 39, no. 3, pp. 428–441, 2015.

[34] J. Behler and W. R. Hess, “Approaches to study CRISPR RNA biogenesis and the key players involved,” *Methods*, vol. 172, no. May 2019, pp. 12–26, 2020.

[35] S. A. Shah, S. Erdmann, F. J. M. Mojica, and R. A. Garrett, “Protospacer recognition motifs: Mixed identities and functional diversity,” *RNA Biol.*, vol. 10, no. 5, pp. 891–899, 2013.

[36] F. Hille, H. Richter, S. P. Wong, M. Bratovi, S. Ressel, and E. Charpentier, “Leading Edge The Biology of CRISPR-Cas: Backward and Forward,” 2018.

[37] E. R. Westra *et al.*, “CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3,” *Mol. Cell*, vol. 46, no. 5, pp. 595–605, 2012.

[38] R. Sapranauskas, G. Gasiunas, C. Fremaux, R. Barrangou, P. Horvath, and V. Siksnys, “The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli,” *Nucleic Acids Res.*, vol. 39, no. 21, pp. 9275–9282, 2011.

[39] J. S. Chen *et al.*, “CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity,” *Science (80-.).*, vol. 360, no. 6387, pp. 436–439, Apr. 2018.

[40] W. X. Yan *et al.*, “Functionally diverse type V CRISPR-Cas systems,” *Science (80-.).*, vol. 7271, no. December, p. eaav7271, 2018.

[41] N. C. Pyenson and L. A. Marraffini, “Type III CRISPR-Cas systems: when DNA cleavage just isn’t enough,” *Curr. Opin. Microbiol.*, vol. 37, pp. 150–154, 2017.

[42] M. Kazlauskiene, G. Kostiuk, Č. Venclovas, G. Tamulaitis, and V. Siksnys, “A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems,” *Science (80-.).*, vol. 357, no. 6351, pp. 605–609, 2017.

[43] O. O. Abudayyeh *et al.*, “RNA targeting with CRISPR-Cas13,” *Nature*, vol. 550, no. 7675, pp. 280–284, 2017.

[44] R. K. Lau *et al.*, “Structure and Mechanism of a Cyclic Trinucleotide-Activated Bacterial Endonuclease Mediating Bacteriophage Immunity,” *Mol. Cell*, vol. 77, no. 4, pp. 723-733.e6, 2020.

[45] S. Grüschow, C. S. Adamson, and M. F. White, “Specificity and sensitivity of an RNA targeting type III CRISPR complex coupled with a NucC endonuclease effector,” *Nucleic Acids Res.*, vol. 49, no. 22, pp. 13122–13134, 2021.

[46] B. Lowey *et al.*, “CBASS Immunity Uses CARF-Related Effectors to Sense 3′–5′- and 2′–5′-Linked Cyclic Oligonucleotide Signals and Protect Bacteria from Phage Infection,” *Cell*, pp. 1–12, 2020.

[47] A. J. Meeske, S. Nakandakari-Higa, and L. A. Marraffini, “Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage,” *Nature*, vol. 570, no. 7760, pp. 241–245, 2019.

[48] B. Tong, H. Dong, Y. Cui, P. Jiang, Z. Jin, and D. Zhang, “The Versatile Type V CRISPR Effectors and Their Application Prospects,” *Front. Cell Dev. Biol.*, vol. 8, no. February, pp. 1–11, 2021.

[49] A. M. Millen *et al.*, “Lactococcus lactis type III-A CRISPR-Cas system cleaves bacteriophage RNA,” *RNA Biol.*, vol. 16, no. 4, pp. 461–468, 2019.

[50] R. Pinilla-Redondo *et al.*, “Type IV CRISPR-Cas systems are highly diverse and involved in competition between plasmids,” *Nucleic Acids Res.*, vol. 48, no. 4, pp. 2000–2012, 2020.

[51] H. N. Taylor *et al.*, “Positioning Diverse Type IV Structures and Functions Within Class 1 CRISPR-Cas Systems,” *Front. Microbiol.*, vol. 12, no. May, pp. 1–10, 2021.

[52] U. Berendt, T. Haverkamp, A. Prior, and J. D. Schwenn, “Reaction mechanism of thioredoxin: 3’-phospho-adenylylsulfate reductase investigated by site-directed mutagenesis.,” *Eur. J. Biochem.*, vol. 233, no. 1, pp. 347–356, Oct. 1995.

[53] Y. Zhou *et al.*, “Structure of a type IV CRISPR-Cas ribonucleoprotein complex,” *iScience*, vol. 24, no. 3, p. 102201, 2021.

[54] M. B. Begemann *et al.*, “Characterization and Validation of a Novel Group of Type V, Class 2 Nucleases for in vivo Genome Editing.”

[55] Y. Tang and Y. Fu, “Class 2 CRISPR/Cas: An expanding biotechnology toolbox for and beyond genome editing 06 Biological Sciences 0604 Genetics,” *Cell Biosci.*, vol. 8, no. 1, pp. 1–13, 2018.

[56] M. Adli, “The CRISPR tool kit for genome editing and beyond,” *Nat. Commun.*, vol. 9, no. 1, 2018.

[57] V. M. Crowley *et al.*, “A Type IV-A CRISPR-Cas System in Pseudomonas aeruginosa Mediates RNA-Guided Plasmid Interference In Vivo,” *Cris. J.*, vol. 2, no. 6, pp. 434–440, Nov. 2019.

[58] H. Savage, G. Montoya, C. Svensson, J. D. Schwenn, and I. Sinning, “Crystal structure of phosphoadenylyl sulphate (PAPS) reductase: A new family of adenine nucleotide α hydrolase,” *Structure*, vol. 5, no. 7, pp. 895–906, 1997.

[59] K. S. Carroll, H. Gao, H. Chen, C. D. Stout, J. A. Leary, and C. R. Bertozzi, "A conserved mechanism for sulfonucleotide reduction," *PLoS Biol.*, vol. 3, no. 8, 2005.

[60] J. Chartron *et al.*, "Substrate Recognition, Protein Dynamics, and Iron-Sulfur Cluster in Pseudomonas aeruginosa Adenosine 5′-Phosphosulfate Reductase," *J. Mol. Biol.*, vol. 364, no. 2, pp. 152–169, 2006.

[61] J. A. Bick, J. J. Dennis, G. J. Zylstra, J. Nowack, and T. Leustek, "Identification of a new class of 5'-adenylylsulfate (APS) reductases from sulfate-assimilating bacteria," *J. Bacteriol.*, vol. 182, no. 1, pp. 135–142, 2000.

[62] P. Bork and E. V Koonin, "A P-loop-like motif in a widespread ATP pyrophosphatase domain: Implications for the evolution of sequence motifs and enzyme activity," *Proteins Struct. Funct. Bioinforma.*, vol. 20, no. 4, pp. 347–355, Dec. 1994.

[63] M. Fellner, R. P. Hausinger, and J. Hu, "A structural perspective on the PP-loop ATP pyrophosphatase family," *Crit. Rev. Biochem. Mol. Biol.*, vol. 53, no. 6, pp. 607–622, 2018.

[64] Z. Fan *et al.*, "Efficient gene targeting in golden Syrian hamsters by the CRISPR/Cas9 system," *PLoS One*, vol. 9, no. 10, pp. 1–9, 2014.

[65] S. J. Gray *et al.*, "Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors," *Hum. Gene Ther.*, vol. 22, no. 9, pp. 1143–1153, 2011.

[66] P. Liu *et al.*, "Enhanced Cas12a editing in mammalian cells and zebrafish," *Nucleic Acids Res.*, vol. 47, no. 8, pp. 4169–4180, 2019.

[67] M. R. Paule, "SURVEY AND SUMMARY Transcription by RNA polymerases I and III," *Nucleic Acids Res.*, vol. 28, no. 6, pp. 1283–1298, 2000.

[68] R. D. Mashal, J. Koontz, and J. Sklar, "Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases," *Nat. Genet.*, vol. 9, no. 2, pp. 177–183, 1995.

[69] C. W. N. S. S. J. R. B. A., Shafiq, "Non-homologous DNA end joining and alternative pathways to double-strand break repair," *Physiol. Behav.*, vol. 176, no. 3, pp. 139–148, 2017.

[70] F. W. Studier, "Protein production by auto-induction in high-density shaking cultures," *Protein Expr. Purif.*, vol. 41, no. 1, pp. 207–234, May 2005.

[71] K. Higuchi, T. Yabuki, M. Ito, and T. Kigawa, "Cold shock proteins improve E. coli cell-free synthesis in terms of soluble yields of aggregation-prone proteins," *Biotechnol. Bioeng.*, vol. 117, no. 6, pp. 1628–1639, 2020.

[72] S. Phadtare, J. Alsina, and M. Inouye, "Cold-shock response and cold-shock proteins," *Curr. Opin. Microbiol.*, vol. 2, no. 2, pp. 175–180, 1999.

[73] M. Christian *et al.*, “Targeting DNA double-strand breaks with TAL effector nucleases,” *Genetics*, vol. 186, no. 2, pp. 756–761, 2010.

[74] M. Bibikova, M. Golic, K. G. Golic, and D. Carroll, “Targeted Chromosomal Cleavage and Mutagenesis in Drosophila Using Zinc-Finger Nucleases,” *Genetics*, vol. 1175, no. July, pp. 1169–1175, 2002.

[75] M. Bibikova, K. Beumer, J. K. Trautman, and D. Carroll, “Enhancing gene targeting with designed zinc finger nucleases,” *Science (80-.).*, vol. 300, no. 5620, p. 764, 2003.

[76] J. A. Doudna and E. Charpentier, “The new frontier of genome engineering with CRISPR-Cas9,” *Science (80-.).*, vol. 346, no. 6213, 2014.

[77] H. Iwasaki, “Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holiday junction resolvase,” vol. 92, no. August, pp. 7470–7474, 1995.

[78] O. O. Abudayyeh *et al.*, “C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector,” *Science (80-.).*, vol. 353, no. 6299, p. aaf5573, Aug. 2016.

[79] N. D. Marino, R. Pinilla-Redondo, and J. Bondy-Denomy, “CRISPR-Cas12a targeting of ssDNA plays no detectable role in immunity,” *Nucleic Acids Res.*, vol. 50, no. 11, pp. 6414–6422, 2022.

[80] D. C. Swarts and M. Jinek, “Mechanistic Insights into the cis- and trans-Acting DNase Activities of Cas12a,” *Mol. Cell*, vol. 73, no. 3, pp. 589-600.e4, Feb. 2019.

[81] K. Murugan, A. S. Seetharam, A. J. Severin, and D. G. Sashital, “CRISPR-Cas12a has widespread off-target and dsDNA-nicking effects,” *J. Biol. Chem.*, vol. 295, no. 17, pp. 5538–5553, 2020.

[82] J. Zhang *et al.*, “Structure and Mechanism of the CMR Complex for CRISPR-Mediated Antiviral Immunity,” *Mol. Cell*, vol. 45, no. 3, pp. 303–313, 2012.

[83] P. Samai, N. Pyenson, W. Jiang, G. W. Goldberg, A. Hatoum-Aslan, and L. A. Marraffini, “Co-transcriptional DNA and RNA cleavage during type III CRISPR-cas immunity,” *Cell*, vol. 161, no. 5, pp. 1164–1174, 2015.

[84] G. Tamulaitis *et al.*, “Programmable RNA Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus,” *Mol. Cell*, vol. 56, no. 4, pp. 506–517, 2014.

[85] 5 Patrick Pausch1, 2 †, Basem Al-Shayeb1, 3 †, Ezra Bisom-Rapp4, Connor A. Tsuchida1 and 8 Brady F. Cress1, 2, Gavin J. Knott1, 2, Jillian F. Banfield1, 6, Jennifer A. Doudna1, 2, 7, “CRISPR-CasΦ from huge phages is a hypercompact genome editor,” vol. 337, no. July, pp. 333–337, 2020.

[86] Z. Li, H. Zhang, R. Xiao, R. Han, and L. Chang, “Cryo-EM structure of the RNA-guided ribonuclease Cas12g,” *Nat. Chem. Biol.*, 2021.

[87] R. Xiao, Z. Li, S. Wang, R. Han, and L. Chang, “Structural basis for substrate recognition and cleavage by the dimerization-dependent CRISPR–Cas12f nuclease,” *Nucleic Acids Res.*, vol. 49, no. 7, pp. 4120–4128, 2021.

[88] M. Ming *et al.*, “CRISPR–Cas12b enables efficient plant genome engineering,” *Nat. Plants*, vol. 6, no. March, pp. 202–208, 2020.

[89] J.-J. Liu *et al.*, “CasX enzymes comprise a distinct family of RNA-guided genome editors,” *Nature*, vol. 566, no. 7743, pp. 218–223, Feb. 2019.

[90] L. B. Harrington *et al.*, “A scoutRNA Is Required for Some Type V CRISPR-Cas Systems,” *Mol. Cell*, vol. 79, no. 3, pp. 416-424.e5, 2020.

[91] P. Pausch, K. M. Soczek, D. A. Herbst, and B. Al-shayeb, “DNA interference states of the hypercompact CRISPR-CasΦ effector,” pp. 1–40.

[92] S. N. Takeda *et al.*, “Structure of the miniature type V-F CRISPR-Cas effector enzyme,” *Mol. Cell*, vol. 81, no. 3, pp. 558-570.e3, 2021.

[93] S. Hahn and S. Roberts, “The zinc ribbon domains of the general transcription factors TFIIB and Brf: Conserved functional surfaces but different roles in transcription initiation,” *Genes Dev.*, vol. 14, no. 6, pp. 719–730, 2000.

[94] M. Okuda *et al.*, “A novel zinc finger structure in the large subunit of human general transcription factor TFIIE,” *J. Biol. Chem.*, vol. 279, no. 49, pp. 51395–51403, 2004.

[95] I. Fonfara, H. Richter, M. Bratovič, A. Le Rhun, and E. Charpentier, “The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA,” *Nature*, vol. 532, no. 7600, pp. 517–521, Apr. 2016.

[96] X. Huang *et al.*, “Structural basis for two metal-ion catalysis of DNA cleavage by Cas12i2,” *Nat. Commun.*, vol. 11, no. 1, 2020.

[97] W.-C. Ahn *et al.*, “In vivo genome editing using the Cpf1 ortholog derived from Eubacterium eligens,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–7, 2019.

[98] T. Karvelis *et al.*, “PAM recognition by miniature CRISPR–Cas12f nucleases triggers programmable double-stranded DNA target cleavage,” *Nucleic Acids Res.*, vol. 48, no. 9, pp. 5016–5023, May 2020.

[99] W. Yang, *Nucleases: Diversity of Structure, Function and Mechanism*, vol. 176, no. 1. 2017.

[100] H. Irving and R. J. P. Williams, “637. The stability of transition-metal complexes,” *J. Chem. Soc.*, no. 0, pp. 3192–3210, 1953.

[101] J. P. Broughton *et al.*, “CRISPR – Cas12-based detection of SARS-CoV-2,” 2019.

[102] X. Ding *et al.*, “Ultrasensitive and visual detection of SARS-CoV-2 using all-in-one dual CRISPR-Cas12a assay,” *Nat. Commun.*, vol. 11, no. 1, pp. 1–10, 2020.

[103] J. S. Gootenberg *et al.*, “Nucleic acid detection with CRISPR-Cas13a/C2c2,”

Science (80-.)., vol. 356, no. 6336, pp. 438–442, 2017.

[104] L. Li *et al.*, "HOLMESv2: A CRISPR-Cas12b-Assisted Platform for Nucleic Acid Detection and DNA Methylation Quantitation," *ACS Synth. Biol.*, vol. 8, no. 10, pp. 2228–2237, 2019.

[105] P. D. Hsu *et al.*, "DNA targeting specificity of RNA-guided Cas9 nucleases," *Nat. Biotechnol.*, vol. 31, no. 9, pp. 827–832, 2013.

[106] Y. Fu *et al.*, "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells," *Nat. Biotechnol.*, vol. 31, no. 9, pp. 822–826, 2013.

[107] D. Kim, J. Kim, J. K. Hur, K. Wook Been, S. Yoon, and J.-S. Kim, "Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells," *Nat. Biotechnol.*, vol. 34, 2016.

[108] B. P. Kleinstiver *et al.*, "Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing," *Nat. Biotechnol.*, vol. 37, no. 3, pp. 276–282, Mar. 2019.

[109] L. Gao *et al.*, "Engineered Cpf1 variants with altered PAM specificities increase genome targeting range HHS Public Access Author manuscript," *Nat Biotechnol*, vol. 35, no. 8, pp. 789–792, 2017.

[110] A. A. Malzahn *et al.*, "Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis," *BMC Biol.*, vol. 17, no. 1, pp. 1–14, 2019.

[111] T. H. Hand *et al.*, "Catalytically Enhanced Cas9 through Directed Protein Evolution," *Cris. J.*, vol. 4, no. 2, pp. 223–232, 2021.

[112] T. Scott, R. Urak, C. Soemardy, and K. V. Morris, "Improved Cas9 activity by specific modifications of the tracrRNA," *Sci. Rep.*, vol. 9, no. 1, pp. 1–11, 2019.

[113] I. M. Slaymaker, L. Gao, B. Zetsche, D. A. Scott, W. X. Yan, and F. Zhang, "Rationally engineered Cas9 nucleases with improved specificity," *Science (80-.).*, vol. 351, no. 6268, pp. 84–88, 2016.

[114] H. Kim *et al.*, "Enhancement of target specificity of CRISPR-Cas12a by using a chimeric DNA-RNA guide," *Nucleic Acids Res.*, vol. 48, no. 15, pp. 8601–8616, 2020.

[115] C. J. Harris, A. Molnar, S. Y. Müller, and D. C. Baulcombe, "FDF-PAGE: A powerful technique revealing previously undetected small RNAs sequestered by complementary transcripts," *Nucleic Acids Res.*, vol. 43, no. 15, pp. 7590–7599, 2015.

[116] O. Dmytrenko *et al.*, "Cas12a2 elicits abortive infection via RNA-triggered destruction of double-strandd DNA," pp. 1–60, 2022.

[117] J. P. K. Bravo, T. Hallmark, B. Naegle, C. L. Beisel, N. Ryan, and D. W. Taylor, "Large-scale structural rearrangements unleash indiscriminate nuclease activity of

CRISPR-Cas12a2," pp. 1–39, 2022.

Appendices

APPENDIX A

PERMISSION OF COPYRIGHTED MATERIALS

Dylan Keiser <dylan.keiser@usu.edu>

Sat 10/29/2022 4:12 PM

All,

I am submitting my master's thesis in the coming weeks and in my thesis, I discuss my contributions to works listed below. The USU graduate school requires I obtain permission from all authors on the manuscripts before publishing my thesis. Could you please reply to this email with the following text copied and your name inserted? Keep in mind to do it for each manuscript for which you are listed as an author.

Thanks,

Dylan

I, (your name), hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Taylor, H. N., Laderman, E., Armbrust, M. J., Hallmark, T., Keiser, D., Bondy-Denomy, J., & Jackson, R. N. (2021). Positioning diverse type IV structures and functions within class 1 CRISPR-Cas systems. *Frontiers in Microbiology*.

I, (your name), hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN,

Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Gina Neumann <gcgneumann@gmail.com>

Sat 10/29/2022 5:18 PM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Gina Neumann, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Hannah Domgaard <hdomgaard@gmail.com>

Sat 10/29/2022 5:50 PM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Hannah Domgaard, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

---

Joe Bondy-Denomy <joebondy@gmail.com>

Sat 10/29/2022 6:10 PM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Joseph Bondy-Denomy (your name), hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Taylor, H. N., Laderman, E., Armbrust, M. J., Hallmark, T., Keiser, D., Bondy-Denomy, J., & Jackson, R. N. (2021). Positioning diverse type IV structures and functions within class 1 CRISPR-Cas systems. *Frontiers in Microbiology*.

Hannah Domgaard <hdomgaard@gmail.com>

Sat 10/29/2022 6:23 PM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Hannah Domgaard, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Beisel, Chase <chase.beisel@helmholtz-hiri.de>

Sun 10/30/2022 1:00 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Chase Beisel, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Taylor, H. N., Laderman, E., Armbrust, M. J., Hallmark, T., Keiser, D., Bondy-Denomy, J., & Jackson, R. N. (2021). Positioning diverse type IV structures and functions within class 1 CRISPR-Cas systems. *Frontiers in Microbiology*.

I, Chase Beisel, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Dmytrenko, Oleg <oleg.dmytrenko@helmholtz-hiri.de>

Sun 10/30/2022 8:05 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

Hi Dylan,

You have my permission!

Best of luck!

Oleg

Matthew Begemann <mbegemann@bensonhill.com>

Sun 10/30/2022 1:50 PM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Matthew Begemann, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Sent from my iPhone

Wandera, Katharina <katharina.wandera@helmholtz-hiri.de>

Mon 10/31/2022 2:03 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

Hi Dylan,

best of luck with your thesis!

I, Katharina Wandera, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN,

Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Vialetto, Elena <elena.vialetto@helmholtz-hiri.de>

Mon 10/31/2022 2:14 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Elena Vialetto, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Valerie Crowley <valerie.crowley@gmail.com>

Mon 10/31/2022 6:46 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Valerie Crowley, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Johannes Weber <joe.web.sw@gmail.com>

Tue 11/1/2022 3:21 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Johannes Weber, hereby give permission to Dylan Keiser to reprint the following publication in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Yannis .Mougiakos <ymouyiakos@googlemail.com>

Tue 11/1/2022 6:41 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Ioannis Mougiakos, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Josie Metcalf <josiemetcalf2010@gmail.com>

Wed 11/2/2022 4:47 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Josie Metcalf, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

Dylan Keiser <dylan.keiser@usu.edu>

Wed 11/16/2022 7:22 PM

BCC: Hannah Taylor <Hannah.Taylor@USU.edu>, <Eric.Laderman@UCSF.edu>, <Matt.armbrust5@gmail.com>, <Thom.hallmark@USU.edu>, <bgray@bensonhill.com>

All, I have not received responses from you. If I don't hear from you in 72 hours, then I will assume you approve.

Thom Hallmark <thom.hallmark@usu.edu>

Thu 11/17/2022 11:27 AM

To: Dylan Keiser <dylan.keiser@usu.edu>

I, Thomson Hallmark, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Taylor, H. N., Laderman, E., Armbrust, M. J., Hallmark, T., Keiser, D., Bondy-Denomy, J., & Jackson, R. N. (2021). Positioning diverse type IV structures and functions within class 1 CRISPR-Cas systems. *Frontiers in Microbiology*.

I, Thomson Hallmark, hereby give permission to Dylan Keiser to reprint the following publication(s) in part or in full in his thesis

Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits abortive infection through RNA-triggered destruction of double-stranded DNA

APPENDIX B

PLASMIDS AND NUCLEIC ACID SEQUENCES

Table B-1 Comprehensive List of Plasmids Used

Lab Archive #	Backbone	Tag(s)	Expression Gene(s)	Selection Marker	Use Case
1	pUC19	-	-	ampicillin	SuCas12a2 collateral plasmid degradation assay
1094	pET	-	Orza sativa caoI gene (with crRNA target)	kanamycin	Plasmid curing assay, template for PCR target
1095	pET	-	Orza sativa caoI gene (without crRNA target)	kanamycin	Plasmid curing assay, template for PCR non-target
1408	pACYC	N-Term 6x His	WT SuCas12a2, 3x crRNA	chlomphe-nicol	coexpression of WT SuCas12a2 and crRNA
1410	pACYC	N-Term 6x His	SuCas12a2 D848A, 3x crRNA	chlomphe-nicol	coexpression of SuCas12a2 D848A and crRNA
1411	pACYC	N-Term 6x His	SuCas12a2 D848A	chlomphe-nicol	expression of apo SuCas12a2 D848A
1412	pACYC	N-Term 6x His	SuCas12a2 E1063A, 3x crRNA	chlomphe-nicol	coexpression of SuCas12a2 E1063A and crRNA

1413	pACYC	N-Term 6x His	SuCas12a2 E1063A	chlomphe-nicol	expression of apo SuCas12a2 E1063A
1414	pACYC	N-Term 6x His	SuCas12a2 D1213A, 3X crRNA	chlomphe-nicol	coexpression of SuCas12a2 D1213A and crRNA
1415	pACYC	N-Term 6x His	SuCas12a2 D1213A	chlomphe-nicol	expression of apo SuCas12a2 D1213A
1416	pET	N-Term 6x His	WT SuCas12a2	ampicillin	Expression of apo SuCas12a2
1424	pET	N-Term 6x His	WT SmCas12a	ampicillian	Expression of apo SmCas12a
1425	pACYC	-	3X crRNA (Sm repeat)	chlomphe-nicol	Combined with protein expression vector in double transformation for coexpression of Cas protein and 3x crRNA
1453	pX330	N-Term 3X FLAG, N-Term SV40 NLS, C-Term Nucleoplasmin NLS	*Mesocricetus auratus* codon optimized SuCas12a2, 6x crRNA version 1	ampicillin	mammalian cell transfection and genome editing experiments in *Mesocricetus auratus*
1454	pX330	N-Term 3X FLAG, N-Term SV40	*Mesocricetus auratus* codon optimized	ampicillin	mammalian cell transfection and genome editing

		NLS, C-Term Nucleoplasmin NLS	SuCas12a2, 6x crRNA version 2		experiments in *Mesocricetus auratus*
1455	pX330	N-Term 3X FLAG, N-Term SV40 NLS, C-Term Nucleoplasmin NLS	*Mesocricetus auratus* codon optimized SuCas12a2, 6x crRNA version 3	ampicillin	mammalian cell transfection and genome editing experiments in *Mesocricetus auratus*
1565	pACYC	N-Term 6x His	WT SuCas12a2, 1x crRNA	chlomphe-nicol	coexpression of WT SuCas12a2 and crRNA
1586	pET	-	*Orza sativa* caoI gene (w/crRNA target), *Mycobacterium sp. JS623-2* Type IV-B CRISPR System Csf1-Csf4	kanamycin	Plasmid curing assay for anti-CRISPR analysis of the type IV-B CRISPR system
1738	pACYC	N-Term 6x His	SuCas12a2 C1170S/C1173S/C1188S/ C1191S, 3x crRNA	chlomphe-nicol	coexpression of SuCas12a2 C1170S/C1173S/C1188S/ C1191S, and crRNA
-	pACYC	N-Term 6x His	SuCas12a2 K784A	chlomphe-nicol	Expression of SuCas12a2 K784A

-	pACYC	N-Term 6x His	SuCas12a2 R785A	chlomphe-nicol	Expression of SuCas12a2 R785A
-	pACYC	N-Term 6x His	SuCas12a2 R788A	chlomphe-nicol	Expression of SuCas12a2 R788A

Table B-2. Comprehensive List of Oligonucleotides used in Cleavage Assays

Oligo Name	Sequence	Figure(s)
FAM dsDNA Target	5-FAM CTTTGCTGCCGCACTTGCTCATCAAGCCTTCCTTCAGGTGTTGCTC CAGAAAGGTGAGTTCTTCTTGTTGT	3-2B
FAM ssDNA Target	5'FAM- CTTTGCTGCCGCACTTGCTCATCAAGCCTTCCTTCAGGTGTTGCTC CAGAAAGGTGAGTTCTTCTTGTTGT	3-2B
FAM ssRNA Target	5'FAM- UCUUUGCUGCCGCACUUGCUCAUCAAGCCUUCCUUCAGGUGUU GCUCCAGAAAGGUGAGUUCUUCUUGUUGUU	3-2B, 3-3B
FAM dsDNA Non-Target	5'FAM- AACTGATATGACAATTGCGCGTAGCACGACGACGATATGACACT TGCGCATAACGACGACGATACAATGAT	3-4B, 3-5B, 3-6B, 3-7B, 3-7D, 3-8

FAM ssDNA Non-Target	5'FAM-AACTGATATGACAATTGCGCGTAGCACGACGACGATATGACACTTGCGCATAACGACGACGATACAATGAT	3-4B, 3-5B, 3-6B, 3-7D, 3-8
FAM ssRNA Non-Target	5’FAM-CAGAGAUAAGUGACGCGCGGCGAGUGGCGCGCCACGUCGGAAAUCUAGAGGCG	3-4B, 3-5B, 3-6B, 3-7D
ssRNA target	CGCACUUGCUCAUCAAGCCUUCCUUCAGGUGUUGCUCCAGAAAGGUGAGUUCU	3-3B, 3-4B, 3-4C, 3-5B, 3-6B, 3-7B, 3-7D, 3-8
ssRNA no flank	CAUCAAGCCUUCCUUCAGGUGUUGCUCCA	3-6B
ssRNA self flank	UCCUUUGGACCAUCAAGCCUUCCUUCAGGUGUUGCUCCAAUCUACAAUA	3-6B
HTssDNA target	TCTTTGCTGCCGCACTTGCTCATCAAGCCTTCCTTCAGGTGTTGCTCCAGAAAGGTGAGTTCTTCTTGTTGTT	3-5B

HT dsDNA target	TCTTTGCTGCCGCACTTGCTCATCAAGCCTTCCTTCAGGTGTTGCT CCAGAAAGGTGAGTTCTTCTTGTTGTT	3-5B
LbcrRNA	UAAUUUCUACUAAGUGUAGAUUGGAGCAACACCUGAAGGAAGG CUUGAUG	3-5B
Cas12g crRNA	UUUACCGGCUCUGACACCUGGAGCAACACCUGAAGGAAGGCU	3-5B
Cas13a crRNA	GACCACCCCAAUAUCGAAGGGGACUAAAACUUUGGAGCAACAC CUGAAGGAAGGCU	3-5B

CURRICULUM VITAE

Dylan Keiser

Department of Chemistry and Biochemistry

Utah State University, Logan UT 84322

801-674-2414

Dylan.Keiser@USU.edu

Education

- M.S Biochemistry: Utah State University 2018-Present
- B.S. Chemistry: Southern Utah University 2011-2018

Research Experience

Graduate

Utah State University- December 2018-Present

Advisor: Ryan Jackson

- Characterized the function of a newly discovered CRISPR Cas protein Cas12a2 using recombinant expression and purification, in vitro activity assays (nucleic acid digestion), and various gel electrophoresis techniques.

Undergraduate

Southern Utah University- November 2016-May 2018

Advisor: Chris Monson

- Worked on developing a copper-binding protein filter by working with microfluidic devices, supported lipid bilayers, and fluorescence microscopy.

Utah State University-Summer Internship June 2017-August 2017

Advisor: Nick Dickenson

- Used restriction enzyme DNA cloning to create different arrangements of two cloning site polycistronic DNA in e. coli and shigella bacteria. The goal was to develop a plasmid that coded for both wild type and mutated versions of the same protein.

Publications

Graduate

2022 Dmytrenko O, Neumann GC, Hallmark T, Keiser DJ, Crowley VM, Vialetto E, Mougiakos I, Wandera KG, Domgaard H, Weber J, Metcalf J, Begemann MB, Gray BN, Jackson RN, Beisel CL. CRISPR Cas12a2 elicits bacterial suicide through RNA-triggered destruction of double-stranded DNA (Manuscript in press)

2021 Taylor HN, Laderman E, Armbrust A, Hallmark T, Keiser D, Bondy-Denomy J, Jackson RN. Positioning Diverse Type IV Structures and Functions Within Class 1 CRISPR-Cas Systems Front. Microbiol vol. 12 DOI:10.3389/fmicb.2021.671522

Undergraduate

2021 Kock RS, Reynolds CM, Williams SJ, Keiser DJ, Monson CF Phosphatidylserine in Supported Lipid Bilayers Binds His-Tagged Proteins Journal of the Utah Academy 2021

Presentations

Oral

- Utah State University Hansen Retreat September 2020, *Biochemical Characterization of The Newly Discovered Type V CRISPR Associated Endonuclease Cas12a2*
- Utah State University Hansen Retreat September 2019, *Jackson Lab Overview*
- Southern Utah University Festival of Excellence, April 3, 2018, *The use of phosphatidylserine containing supported lipid bilayers as a method for separating copper binding proteins from non-copper binding proteins*, Dylan Keiser, Nate Mitchell, and Ryan Koch.
- Southern Utah University COSE Research Symposium, November 13, 2017, *Utilizing phosphatidylserine containing supported lipid bilayers as a method to separate copper binding proteins from non-copper binding proteins*, Dylan Keiser
- Southern Utah University Festival of Excellence, April 4, 2017, *Phosphatidylserine-containing Supported Lipid Bilayer as a Separation Medium and Novel Chromatography Column for Copper Binding Compounds*, Chris Reynolds and Dylan Keiser
- Southern Utah University COSE Research Symposium, November 7, 2016, Phosphatidylserine-*containing Supported Lipid Bilayer as a Separation Medium and Novel Chromatography Column for Copper Binding Compounds*, Chris Reynolds and Dylan Keiser

Poster

- Utah State University Hansen Retreat October 2021, *Characterization of a Newly Discovered Type V CRISPR Associated Protein Cas12a2*

Teaching Experience

Utah State University

Assisted students in understanding and applying chemistry laboratory practices and core concepts including proper lab technique, safety, and problem solving.

Courses

- General Chemistry (Fall 2020)
- General Chemistry Lab (Spring 2020)
- General Chemistry Recitation TA (Fall 2019)
- General Chemistry Lab (Spring 2019)
- General Chemistry Lab (Fall 2018)

Research Grants

- Southern Utah University Physical Science Department UGRASP grant October 2017

Scholarships

- Jim and Carolyn Stahl Scholarship 2019
  Utah State University
  Department of Chemistry and Biochemistry
- A.S. & Aline W. Skaggs Scholarship 2017-2018 academic year
  Southern Utah University
  Department of Physical Sciences

Skills

- Recombinant Protein Expression in E. Coli
- Protein Purification (FPLC AKTA system)
- Molecular Cloning
- Nucleic Acid Digestion Assays
- Gel Electrophoresis Techniques
- Radiation Trained (^{32}P labeled substrates)