Biochemical and structural characterization of a Type IV restriction endonuclease SauUSI

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Vinayak Sadasivam Tumuluri

20152019

Indian Institute of Science Education and Research, Pune.

2022

Dedicated to Mrs. Savita Sadasivam, Mrs. Saraswati Gudibande and

Mr. Dinesh Gudibande

CERTIFICATE

I certify that the work presented in the thesis titled “Biochemical and structural characterization of a Type IV restriction endonuclease SauUSI” submitted by Vinayak Sadasivam Tumuluri represents original research carried out by him at IISER Pune under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Prof. Saikrishnan Kayarat

Date: 4th August 2022

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources which have not been properly cited, or from whom proper permission has not been taken when needed.

Vinayak Sadasivam Tumuluri

(Registration number: 20152019)

Date: 3rd August 2022

TABLE OF CONTENTS

ABSTRACT

Restriction modification enzymes are major barriers to horizontal gene transfer (HGT) in bacteria. These enzymes are diverse in the nature of targets that they recognize, their co-factor requirements and mechanisms of action. They are broadly classified as modification independent and modification dependent restriction enzymes. SauUSI is a modification dependent restriction endonuclease and is known to thwart the acquisition of Vancomycin resistance genes in *Staphylococcus aureus*. Here, I report a detailed biochemical and structural characterization of SauUSI where in the presence of ATP the enzyme can cleave DNA having a single or multiple target site/s. Remarkably in substrates with two or more target sites, the DNA region flanked by two target sites is shred into smaller fragments. I also report the first crystal structure of SauUSI which reveals a stable dimer with the interface at the nuclease domains of the two protomers. The architecture of SauUSI facilitates cleavage of both DNA with a single or multiple target site/s and provides a molecular basis for target recognition. ATP hydrolysis powers DNA translocation and we propose that in a DNA flanked by two target sites, translocating SauUSI molecules pile up against a target site bound SauUSI molecule eventually leading to multiple double stranded DNA breaks causing irreparable DNA shredding. The lesser efficient single-site cleavage is attributed to a switch-based mechanism. I also report for the first time that an ATP dependent restriction endonuclease; SauUSI, functions as a *bona fide* single-stranded DNA nuclease. Interestingly, this nucleolytic activity is target-site and nucleotide independent. Further, SauUSI functions as an endonuclease on ssDNA because of its ability to cleave substrates with gaps and mismatch bubbles. However, this nuclease activity is restricted to ssDNA as ssRNA is refractory to cleavage by SauUSI. This ssDNA endonucleolytic activity of SauUSI is downregulated by the presence of single-stranded DNA binding protein and Adenine nucleotides such as ATP, dATP and ADP, thus providing a paradigm for protecting bacterial self-DNA. I hypothesize that this ssDNA cleavage activity could act as a barrier to modes of HGT such as transduction involving ssDNA phages, conjugation and natural transformation.

SYNOPSIS

This thesis submitted to IISER Pune for the partial fulfillment of the degree of Doctor of Philosophy (Ph.D.) focusses on elucidating the mechanism of action of a Type IV bacterial restriction enzyme SauUSI, which cleaves methylated DNA. The introduction to the thesis is a literature survey on molecular motors that traverse nucleic acid contours, Nucleotide triphosphate (NTP) dependent Restriction modification (RM) systems, their mechanisms of action, their role as innate defense systems in bacteria and a brief introduction of SauUSI which is a major barrier to horizontal gene transfer (HGT) in *Staphylococcus aureus*. Chapter 2 focusses on the detailed purification strategy of SauUSI to attain high degree of purity for biochemical and structural studies. In brief, I adopted a three column (Ni-NTA affinity chromatography, anion exchange chromatography and size exclusion chromatography) strategy to purify SauUSI with a C-terminal 6X Histidine tag. In chapter 3, I have elucidated the biochemical and structural characterization of SauUSI. In brief, SauUSI has the ability to cleave double-stranded DNA (dsDNA) which has either a single or multiple target site/s. The structure of SauUSI deciphered by X-ray crystallography reveals that it is a stable dimer with the nuclease domains of the two protomers forming the interface. The molecular architecture provides a basis for target site recognition. The target-site dependent ATPase activity of SauUSI powers translocation along dsDNA and we propose that it leads to a pile up of SauUSI molecules eventually causing multiple dsDNA breaks culminating in shredding of DNA flanked by two target sites. I attribute the single-site cleavage by SauUSI to a switch-based mechanism. While investigating the nucleolytic ability of SauUSI, I discovered that it has the ability to cleave single-stranded DNA, which is the focus of chapter 4. This cleavage activity of SauUSI is the first evidence that a restriction endonuclease can function as a *bona fide* single-stranded endonuclease as well. Further, this activity was negatively regulated by increasing concentrations of adenine nucleoside di/triphosphates and also in the presence of single-stranded DNA binding protein. We hypothesize that this novel ssDNA cleavage can act as a barrier to HGT involving ssDNA such as conjugation, single-stranded phage transduction and natural transformation. In toto, this thesis attempts to uncover the true potential of SauUSI as a potent barrier to HGT through biochemical and structural insights.

ACKNOWLEDGEMENTS

My journey as a doctoral candidate has been nothing short of extraordinary and I am deeply indebted to the IISER Pune community for providing a platform to excel on diverse fronts. I sincerely thank Prof. Saikrishnan Kayarat for giving me an opportunity of a lifetime to work in his laboratory. His passion, attitude, enthusiasm and intensity toward science is admirable and infectious. I joined the lab in 2016 and was really fortunate to have an overlap with the 'forebears', Dr. Manasi Kulkarni Khasnis, Dr. Neha Nirwan, Dr. Ishtiyaq Ahmed and Dr. Mahesh Kumar Chand who helped me immensely in my initial days to get accustomed to the laboratory environment, work culture and routine laboratory practices. Dr. Mahesh Kumar Chand was kind enough to allow me to shadow him and help him out with a few basic biochemical experiments during my laboratory rotation. Discussions with my immediate seniors Ms. Sujata Sharma and Dr. Vishal Aadhav on the basics of molecular biology and biochemistry have held me in good stead throughout my journey. I have had a fruitful collaboration with Dr. Om Prakash Chouhan who helped out with the initial structure building and biochemistry on the SauUSI system. I have also been really fortunate to have had three dynamic MS thesis students work with me; Ms. Vrunda Rajgor (MS University Baroda) who helped me with the biochemical characterization of SauUSI; Ms. Lakshmi Sriram who helped me establish the hypothesis that DNA shredders can have biotechnological applications in the processing of field samples for NGS and metagenomic studies; Mr. Abhinav Masih who helped me immensely by doing what nobody else dared to- establishing a phage system to study the efficacy of RM systems as barriers to HGT.

Credit for the fervent laboratory environment must be given to Mr. Akash Singh Kailashprasad, Ms. Ufaq Bhat, Mr. Rushik Bhatti, Mr. Akhilesh Pratap Singh, Dr. Om Prakash Chouhan, Ms. Sujata Sharma, Mr. Vishal Aadhav, Mr. Ashwin Uday, Ms. Geetanjali Ghosh, Ms. Nikki Dutt, Mr. Aman Kumar Yadav, Mr. Abhinav Masih, Mr. Rajdip Sarkar, Mr. Manas Mahaveer, Mr. Muhammed Navas and Mr. Animesh Anand. Their zeal, curiosity and congenial attitude made it an absolute pleasure to work in the laboratory even during unearthly hours. A special shout out to the 'Niggas' Mr. Akash Singh Kailashprasad, Ms. Ufaq Bhat and Mr. Rushik Bhatti who apart from

being brilliant colleagues are dear friends of mine and have encouraged me during times of frustration and exasperation.

I have met several amazing people at IISER Pune who have inspired me with their work and intellect and many of them have ended up being friends for life: Mr. Susovan Sarkar, Ms. Shikha Dagar, Ms. Sandhya Bhatia, Ms. Mrinmayee Bapat, Ms. Rupali Sathe, Ms. Sushmitha Hegde, Dr. Sukrut Kamerkar, Dr. Devika Andhare, Mr. Krishnendu Roy and Ms. Joyeeta Chakraborty.

Dr. Radha Chouhan and Dr. Jeetender Chugh were kind enough to be part of my research advisory committee and assess my progress. I am indebted to Dr. Gayathri Pananghat and her lab members for their inputs which have helped me during the course of my PhD. I am grateful to Dr. Sudha Rajamani; my faculty mentor who always lent an ear at any time of need.

A part of the smooth sailing of my laboratory experience has to be attributed to the biology office and bio-stores at IISER Pune for facilitating all my experiments through timely provision of reagents and other consumables.

I was fortunate to have had the opportunity to perform X-ray diffraction experiments and these were performed at ESRF (European Synchrotron Radiation Facility, Grenoble) and DLS (Diamond light source, Oxfordshire).

I must mention Prof. Shekhar Mande and his former PhD student Dr. Swastik Phulera who were instrumental in inspiring me to take up science as a career when I was an intern under the Indian academy of science (IAS) summer research fellowship scheme at NCCS, Pune in 2014.

Cricket is a game really close to me and I thank the IISER Pune sports club for choosing me to lead IISER Pune and the Biology department in the Inter IISER Sports meet (2019) and in IISER Premier league (2019, 2020 and 2022) respectively. Winning IPL 2019 and being Runners up in IISM 2019 was a memorable experience.

Finally, I must thank my family for showing immense faith and trust in me throughout this journey.

FIGURES AND TABLES

Chapter 1: Introduction to molecular motors, restriction-modification systems and their role in preventing HGT

Chapter 2: Heterologous Expression and High Degree Purification of the Restriction Endonuclease SauUSI

Chapter 3: Mechanism of DNA cleavage by the endonuclease SauUSI – a major barrier to horizontal gene transfer and antibiotic resistance in *Staphylococcus aureus*

Chapter 4: Characterizing a novel ssDNA nuclease activity of a type IV restriction endonuclease SauUSI

CHAPTER 1: Introduction to molecular motors, restriction-modification systems and their role in preventing HGT

Abstract

Molecular motors play seminal roles in a variety of cellular processes and are hence crucial for maintaining homeostasis. Molecular motors that traverse nucleic acid contours are classified as helicases. Here I discuss the classification of helicases and highlight some of features with respect to their structure and functioning. I also discuss the role of enzymes possessing modular helicase domains in the process of bacterial defense. Interestingly the characterized helicase domains present in Restriction modification (RM) enzyme function as translocases or molecular switches rather than *bona fide* helicases. RM and CRISPR-Cas systems being strong barriers to horizontal gene transfer (HGT) aid in protecting the host bacterium against pathogenic foreign DNA. However, these enzymes can also have a negative impact when the host is challenged with antibiotics where the only means of survival is to take up genes horizontally that confer resistance. Understanding the nucleolytic potency of these barriers to HGT can further our understanding on the phenomenon of antibiotic resistance and novel ways of countering its deleterious effects. This chapter acts as prelude to my work on SauUSI, a Type IV restriction endonuclease which is a major barrier to HGT in *Staphylococcus aureus*.

Introduction

Proteins are known to be the structural and functional units of life. Proteins known as 'molecular motors' are involved in an array of fundamental processes such intracellular cargo transport, DNA replication, transcription, translation and membrane transport through channels to name a few (1, 2). Briefly, a molecular motor functions through cyclic conformational changes in its structure which is usually brought about by a cofactor such as nucleotide triphosphates (NTPs) (3). Molecular motors such as Dynein, Kinesin and myosin use the network of microtubules and microfilaments within the cell for the transport of cargo (vesicles) and seminal processes such as muscle contraction (4). On the other hand, nucleic acids molecular motors use nucleic acid tracks or polynucleotide chains (either DNA or RNA) to perform crucial roles in a myriad of genetic processes such as DNA replication, transcription, translation, resolving Holliday junctions, DNA repair and recombination etc (5, 6).

Nucleic acid molecular motors are classified as helicases or translocases based on their ability to couple NTP hydrolysis to unwind stretches of double stranded polynucleotide chain or remodeling DNA/RNA by migrating on double stranded polynucleotide chains without unwinding them, respectively (7). Based on the preferred polarity of DNA/RNA for the process of unwinding/translocation, nucleic acid molecular motors are classified as: Type A translocation in the 3'→5' orientation and Type B translocation in the 5'→3' orientation (Figure 1A) (8, 9). Further, helicases are classified as α or β depending on whether their substrates are single-stranded or double-stranded polynucleotide chains respectively (Figure 1A) (8, 9). However, a more detailed classification of nucleic acid molecular motors is based on their molecular architecture and categorized into various superfamilies (SFs). Superfamilies 1 and 2 (SF 1 and SF 2) are non-ring forming helicases whereas Superfamilies 3 – 6 (SF 3 – SF 6) are annular (Figure 1B and 1C) (2, 8, 9).

Non-ring forming helicases

SF 1 and SF 2 comprise the largest group of characterized helicases (2). These systems have been extensively studied using tools of biochemistry, biophysics and structural biology and hence we have an understanding of how these enzymes function in certain contexts. They are characterized by the presence of two RecA-like protein folds called the Rec 1A and Rec 2A sub-domains which come together and form the complete helicase fold (Figure 1B) (10). There architecture consists of two distinct regions viz., the NTP binding pocket and the nucleic acid interaction surface. The NTP binding pocket is present in a cleft between the Rec 1A and Rec 2A sub-domains whereas the surface of these sub-domains possesses residues that interact with the nucleic acid backbone or form stacking interactions with the nucleobases. In addition to the aforementioned sub-domains, SF 1 helicases also possess Rec 1B and Rec 2B accessory subdomains which are considered single insertions within the polypeptide chain.

The two RecA sub-domains are composed primarily of five parallel beta strands (β 1 – β 5) connected to each other by four alpha helices (α 1 – α 4) (11). Interestingly, the two RecA domains share a high structural homology with each other, however, biochemically they have distinct functions. There are seven conserved (I, Ia – VI) motifs interspersed between these two sub-domains which contribute to the distinct roles played by Rec 1A and Rec 2A. Motif I, Ia, II and III are present in Rec 1A whereas Motifs IV, V and VI are present in Rec 2A (Figure 1D). The most conserved of these motifs are Motif I (Walker A; consensus sequence: GKT/S), Motif II (Walker B; consensus sequence: For SF 1 – DEY/F/AQD and SF 2 – DExD/H) and Motif VI (Arginine finger; consensus sequence: QxxGRxxR) which form the ATP binding and hydrolysis active site. Motifs Ia, IV and V play an important role in interacting with the polynucleotide chain and Motif III, also known as the coupler motif, unifies ATP hydrolysis to nucleic acid translocation or other downstream processes (12–17).

The SF 1 helicases characterized so far are Type α because they function primarily when a single stranded region or overhang is present, whereas SF 2 helicases can be either Type α or β (9). In-spite of this difference the shared feature of the two RecA like subdomains indicates a conservation in their mechanism of converting chemical energy from NTP hydrolysis to

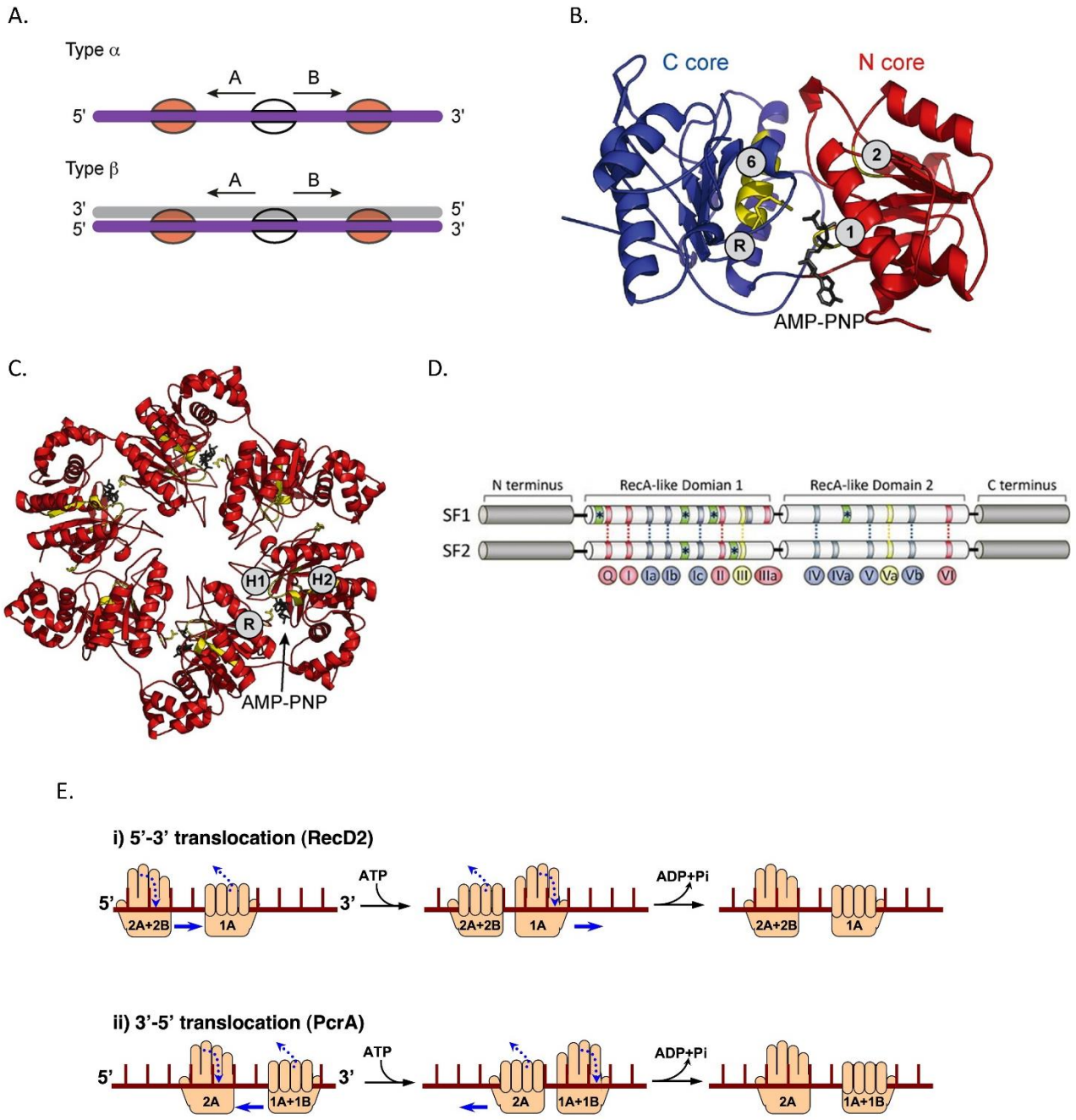


Figure 1: Types of helicases and mechanism of action of SF 1 helicases: A) Helicases are classified based on the nature of their substrates (single stranded nucleic acids: Type α and double stranded nucleic acids: Type β). Further, they are also classified based on their preferred polarity of translocation (3' → 5': Type A and 5' →3': Type B). B) Structure of the helicase domain of PcrA (SF 1 helicase) consisting of the N-core (Rec 1A) and C-core (Rec 2A). This structure also shares similarities with SF 2 helicases as well. AMPPnP (ATP analogue) is bound to the cleft region consisting of the Walker A (motif I), Walker B (motif II) and Arginine finger (motif VI). C) Annular helicase T7 gene 4 bound to ATP analogue (AMPPnP in black) at four sites. SF 3 – SF 6 are annular helicases share similarities in consisting of either a AAA+ or RecA core with the NTP binding site at the interface of protomers. D) The RecA core in SF 1 and SF 2 helicases primarily consists of Rec 1A (N- core with conserved motifs I, Ia, II and III) and Rec 2A (C-core with conserved motifs IV, V and VI). In addition to these seven conserved motifs there are other accessory motifs (Q, Ib, Ic, IIIa, Iva, Va and Vb). E) Mechanism of translocation of SF 1B (RecD2) and SF 1A (PcrA) helicase which involves movement of the two RecA domains with respect to each other upon ATP binding and hydrolysis eventually leading to the "inchworm" mechanism of translocation on nucleic acids. SF 2 helicases also perform a similar mode of action. (A – D) Adapted from *Singleton et al., 2007, Annual Reviews in Biochemistry*. (D) Adapted from *Saikrishnan et al., 2009, Cell.*

mechanical energy. Due to the availability of high-resolution structures of the SF 1A helicse PcrA in different conformations along with elaborate biochemical studies a general scheme has been proposed for the mechanism of how SF 1 and SF 2 helicases utilize the energy from ATP hydrolysis to bring about unwinding/translocation along polynucleotide chains (11, 18). The process is cyclic hence making these enzymes processive for translocation. The cycle begins by the binding of the enzyme to single stranded DNA. The initial binding of PcrA has the Rec 1A bound in a tight conformation and Rec 2A preceding it is in a weakly bound conformation. Then ATP binding leads to closure of the ATP binding cleft bringing Rec 1A in close proximity to Rec 2A. In such a state Rec 1A is in a weakly bound conformation whereas Rec 2A adopts a tightly bound conformation. Upon ATP hydrolysis, the protein returns to its initial bound form with Rec 1A in a tightly bound conformation and Rec 2A in a weak bound conformation (18, 19). However, the conformational changes resulting from the nucleotide hydrolysis move the domains to +1 position at the start of the next cycle thereby causing translocation of the enzyme in the forward direction (3' → 5') along ssDNA (Figure 1E) (18). Interestingly, RecD2 a SF 1B helicase, undergoes a similar cyclic process where the Rec 1A sub-domain precedes the Rec 2A sub-domain when interacting with ssDNA(20). Upon ATP binding Rec 2A moves towards Rec 1A and the hydrolysis of ATP followed by the release of ADP and Pi causes the Rec 1A sub-domain to move into a +1 position and poising the enzyme for a new cycle of ATP binding and hydrolysis. This slight deviation from the general scheme presented in case of PcrA contributes to a change in translocation polarity for RecD2 that is 5' → 3' (Figure 1E). The aforementioned mechanisms are called "inchworm mechanisms" of translocation and it accounts for the processive and concerted movement of the two RecA sub-domains on nucleic acid tracks (18, 20, 21). Structural studies on PcrA also revealed an interesting perspective on integrating the process of unwinding to the inchworm mechanism of translocation. It was found that PcrA bound the ssDNA component of a 3' overhang substrate with high affinity in the absence of nucleotide whereas, in the presence of an ATP analogue the 1B and 2B sub-domain adopted a conformation ideal for interacting with the dsDNA component of a 3' overhang substrate (18). These two distinct conformations of the protein along with the difference in position of interaction with the ssDNA component suggested that the free energy of binding of PcrA to ssDNA in the absence of nucleotide stabilized the transient fraying of the

DNA at the ssDNA – dsDNA junction while the ATP hydrolysis causes a conformational change to propel the enzyme in the forward direction thus reconciling ATP dependent translocation to unwinding activity (18, 22).

Interestingly, apart from acting as *bona fide* helicases that is, proteins that unwind polynucleotide chains, some proteins containing the helicase domain can also function as translocases (23–25), that is, enzymes that migrate on intact dsDNA without unwinding but bring about the remodeling of these polynucleotide chains; or as molecular switches (26–29) that is, enzymes that use the energy from nucleotide hydrolysis for certain distinct downstream processes. Helicases and translocases share a lot of similarity in their mechanism of action as both these types of enzymes utilize energy from NTP hydrolysis for directional movement on polynucleotide chains. Most helicases (PcrA, RecD2, HCV NS3) require a single-stranded nucleotide overhang for binding post which, ATP hydrolysis causes their translocation on that strand (18, 20, 22). As a consequence of their translocation along with destabilization of the dsDNA region at the ssDNA – dsDNA these helicases bring about the unwinding of double-stranded nucleic acids. Translocases (LlaBIII, HsdR) on the other hand do not require single-stranded nucleotide overhangs for binding and can efficiently bind to one of the strands in a polynucleotide duplex upon which they translocate by utilizing the energy of NTP hydrolysis without unwinding the duplex (30, 31). Both helicases and translocases undergo multiple rounds of NTP binding and hydrolysis without dissociating from their nucleic acid substrates thereby causing them to be highly processive in their function of unwinding or translocation (9). On the other hand, molecular switches (DEAD – Box RNA Helicases) utilize the energy of NTP hydrolysis in a non-processive manner i.e., a single catalytic cycle of ATP binding and hydrolysis is adequate for their biological function eventually leading to a conformation where the enzyme dissociates from its substrate (26, 27).

Annular helicases

SF 3 – 6 are called annular helicases because their enzymatically active form is a toroidal structure formed by six protomers (Figure 1C) (9). These hexamers across the different superfamilies share a high structural homology as they possess either a AAA+ or RecA like ATPase core at the interface of the monomers which consists an Arginine finger from the adjacent protomer (Figure 1C) (9).

The pore formed by the six protomers is the channel that a polynucleotide chain would take to interact with the hexamer (32). Interestingly, these helicases do not require the presence of overhangs to load onto a polynucleotide chain. Rather, they either adopt an open conformation which allows them to load onto one of the strands of a nucleic acid duplex or they polymerize around one of the strands of the duplex thus forming a stable hexamer only after binding to their nucleic acid substrate (33). This class of helicases function mainly as conventional helicases and can traverse polynucleotide chains in the kilobase – megabase range making them extremely processive (9). A hexamer forms six NTP binding pockets and this provides an interesting paradigm for NTP binding and mechanism of hydrolysis (34). Depending on the hexameric helicase, the three states viz., NTP bound (T), NDP bound (D) and empty (E) are known to interconvert in either a: 1) Sequential manner (T7 gp 4) where, three contiguous protomers are in T-state, two in the D-state and one in the E-state. Hence all the protomers are in a catalytic state and this sequence of conformations move in a cyclic fashion across the six protomers (35). 2) Concerted manner (SV 40 Large T antigen) where, all the six protomers are either in the T, D or E-states and the interconversion of these states occurs simultaneously (36). 3) Stochastic manner (ClpX) where, all the six protomers though catalytically active do not have a fixed sequence of interconversion of states (37).

Similar to other motor enzymes, these class of helicases couple NTP hydrolysis to nucleic acid translocation followed by unwinding. The central pore formed by the hexamer consists of conserved loops that have the ability to trace the spiral of the DNA and interact with its backbone (32). Upon NTP hydrolysis, the loops in the central pore undergo conformational changes which leads to a rearrangement of interactions with the polynucleotide chain thus pulling it through the pore in a “rotary inchworm mechanism” (9). A consequence of translocation of these helicases around only one of the strands of double stranded nucleic acids causes the other strand to be mechanically stripped from the duplex (32).

The prokaryotic form of life consists of many representatives of these helicase superfamilies to perform various physiological processes and efforts are directed to study these systems to unravel their mechanisms of action. The fundamental process of warding off the threat from mobile genetic elements (MGEs) in bacteria involves the use of complex multi-domain multi-

functional proteins many of which consist of the modular helicase domain functioning as translocases or molecular switches.

Bacterial defense systems

Bacteria and Archaea are under constant threat from their natural predators, bacteriophages (38, 39). Bacteriophages are viral entities that have the ability to take over the physiology of the host prokaryote for self-amplification eventually leading to the lysis of the host and spread of infection amongst the prokaryotic population (39, 40). To counter the deleterious effects of bacteriophages and other MGEs bacteria and archaea have evolved a whole battery of mechanisms that protect them (Figure 2A) (41). These mechanisms are broadly classified on the basis of mode of action upon foreign genetic material infection. Chemical defenses have been studied in certain *Streptomyces sp.* in which the bacterium produces an anti-phage small molecule which is a nucleic acid intercalator thus preventing the amplification of the phage genome (42). Abortive infection (Abi) mechanisms involve the activation of suicide proteins which eventually cause cell death of the host bacterium thus limiting the spread of the bacteriophage infection amongst the population (43). Phage infection in *Vibrio cholerae* stimulates the synthesis of 3'-5' cyclic guanosine monophosphate-adenosine monophosphate (cGAMP) which eventually activates CapV causing damage to the inner membrane thus destabilizing the cell membrane and causing lysis of the host cell (44). Analogously, the Gabija defense system of *Bacillus cereus* consists of an operon encoding a nucleotide-sensing nicking endonuclease GajA and a putative helicase GajB which provide robust defense against phage infection. *In-vitro* studies on GajA reveal that it has the potential to cleave bacterial self-genome in a nucleotide depleted state in an attempt to limit the spread of phage infection (45). However, the anti-phage defense systems that are predominantly present in prokaryotes are the ones that target foreign/phage genetic material such as the Restriction-modification (RM) systems and clustered regularly interspaced short palindromic repeats associated Cas (CRISPR-

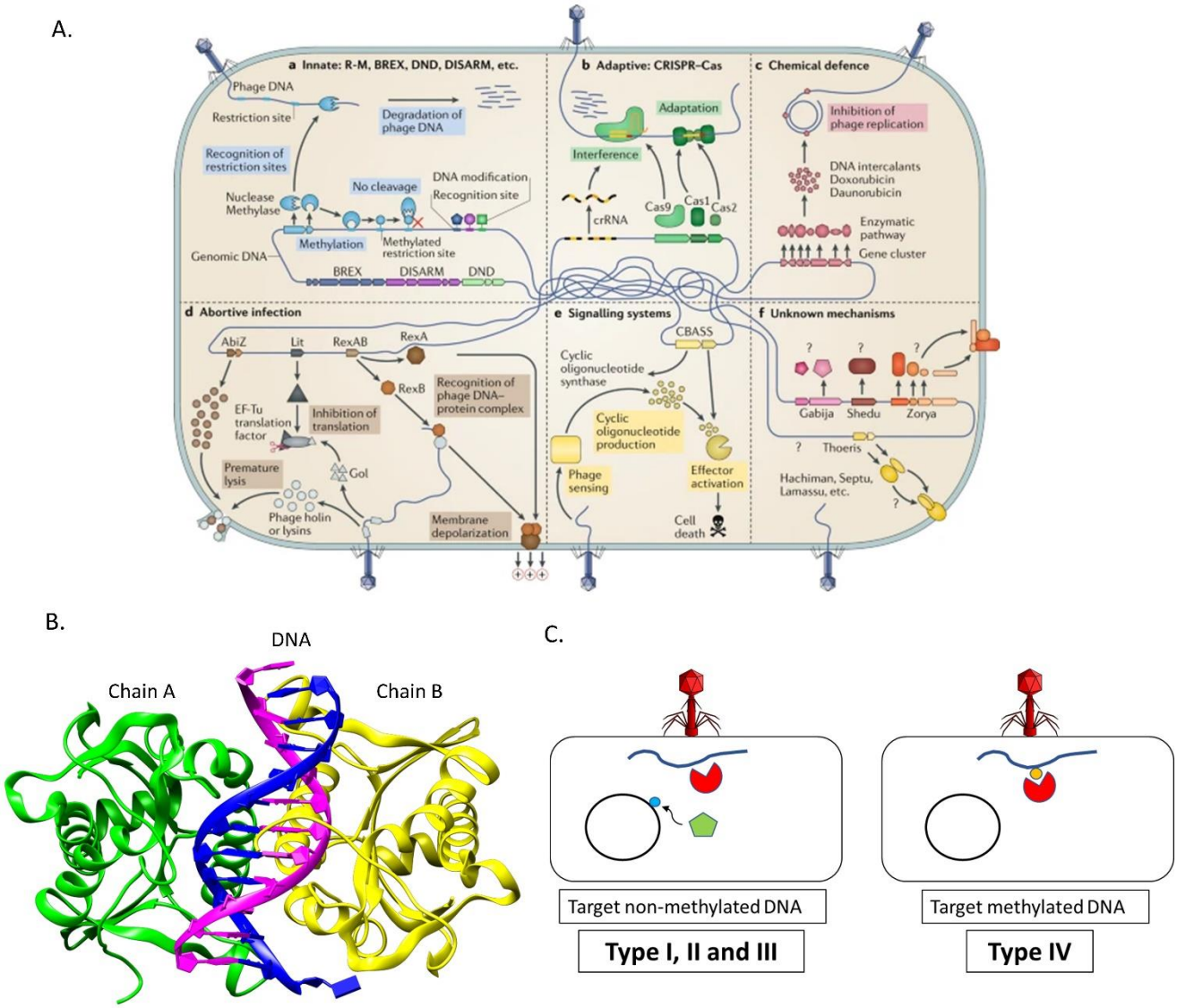


Figure 2: Diversity of bacterial defense systems: A) Cartoon representation of the bacterial defense systems that exist. The most abundant systems are the RM and CRISPR-Cas systems that target foreign nucleic acids. B) Structure of a dimer (Chain A: green and Chain B: yellow) of BamHI bound to dsDNA (blue and magenta) (PDB: 1BHM). Type II restriction endonucleases have a relatively simple architecture with a target-recognition domain and nuclease domain fused together. C) Cartoon representation of modification-independent (Type I, II and III) and modification-dependent (Type IV) restriction systems. Type I, II and III restriction endonucleases (red pacman) target unmodified foreign DNA (blue chain) and partner with methyltransferases (green) which methylate at cognate sites (blue) on the bacterial genome (black circle). However, Type IV restriction endonucleases target modified foreign DNA (blue chain with yellow circle) and do not partner with cognate methyltrasnferases. (A) Adapted from *Bernheim and Sorek, 2020, Nature Reviews Microbiology*.

Cas) systems (46, 47). RM systems are called the innate immune response of bacteria and constitute two components; a restriction component, which is a target-site specific endonuclease that degrades foreign DNA and a modification component, which methylates the cognate target sites on the host genome thus preventing self-DNA cleavage (48, 49). Certain restriction endonucleases such as BamHI, EcoRI, EcoRV etc., are known to cause precise dsDNA breaks hence have been tapped for routine laboratory biotechnological applications (50). On the other hand, the CRISPR-Cas system functions by incorporating short stretches of viral genetic code as "spacers" between designated palindromic repeat regions in bacteria (47, 51). These spacers are then transcribed and undergo processing to form CRISPR RNA (crRNA) which guide the potent Cas family of nucleases to incoming pathogenic foreign DNA (52, 53). Due to the specific nature of double stranded DNA breaks that certain CRISPR systems can bring about, the CRISPR-Cas9 system has been extensively used as a genome editing tool (54).

Further, with the advent of metagenomic studies and high-throughput sequencing technology many novel defense systems are coming to the fore and we are now beginning to understand the real potential and complete repertoire of bacterial defense systems (38, 41).

Diversity of bacterial RM systems

Restriction-modification systems are the most abundant foreign nucleic acid targeting mechanisms present in ~75 % of sequenced prokaryotic genomes (41). RM systems were first described by Werner Arber and Daisy Dussoix in 1962 when they observed bacterial host specific modification of λ phage genetic material (55). A particular strain of bacteria was infected only by those λ phages that had the same modification which was present in the bacterial strain. Attempts to infect different bacterial strains with λ phages that had "different" modifications were futile. This distinct nature of phages having host specificity was also shown by experiments conducted by Salvador Luria and Mary Human (56). Based on all these observations, Werner Arber and Stuart Linn proposed the presence of bacterial enzymes that specifically "restrict" and cleave foreign phage DNA which were unmodified (57). It was further proposed that bacteria harbor enzymes that "modify" self-DNA and when phages possessed these favorable modifications they flourished amongst compatible bacterial strains. Gunther Stent envisaged

that the modification was a methyl (-CH_3) group covalently bound to one of the nucleobases which eventually prompted Werner Arber to show that methionine was necessary to be present in the growth media for efficient modification to occur(58). This hypothesis was based on the discovery of modified cytosine (5-methyl cytosine) in the genome of *Mycobacterium tuberculosis* by Johnson and Coghill in 1925 but was further substantiated with the discovery of DNA and RNA methyltransferases in 1963 which catalyze the transfer of the -CH3 group from S-adenosyl Methionine (SAM) onto specific nucleobases (59). We now know that the modification could be 5-methylcytosine (m5C), *N*6-methyladenine (m6A), or *N*4-methylcytosine (m4C) depending on the methyltransferase performing the activity (60).

The restriction enzyme that Werner Arber and Stuart Linn discovered was named EcoB because it was derived from an *Escherichia coli* B – strain. Soon enough other restriction enzymes such as EcoK were discovered. However, it was in the year 1970 that Hamilton Smith and Kent Wilcox biochemically characterized the first restriction enzyme called HindII from *Haemophilus influenzae* (61, 62). Hamilton Smith and colleagues were able to confirm Werner Arber and Stuart Linn's hypothesis of restriction enzymes specifically cleaving foreign phage DNA but being inactive against bacterial self-DNA. They further characterized it to discover that it cleaves DNA with a specific target-site making it extremely sensitive to nucleobase sequences. The use of restriction enzymes as a way to cut DNA was demonstrated in 1971 by Daniel Nathans and Kathleen Danna when they showed a distinct DNA band pattern when eukaryotic virus SV 40 genomic DNA was treated with HindII and visualized through gel electrophoresis (63). The use of restriction enzymes as a means to perform molecular cloning was first demonstrated by Janet Mertz and Ronald Davis using the restriction enzyme EcoRI which produces staggered ends rather than blunt ends which HindII produces (64). Werner Arber, Hamilton Smith and Daniel Nathans were awarded the Nobel Prize in Physiology or Medicine in 1978 for their "discovery of Restriction endonucleases" which would eventually lead to the development of the field of recombinant DNA technology (65).

However, enzymes like EcoRI, HindII, HindIII, BamHI etc are categorized as Type II RM systems because they are site-specific cutters, have a relatively simple 3D architecture with a fused target recognition and nuclease domain and their function is not dependent on a nucleotide

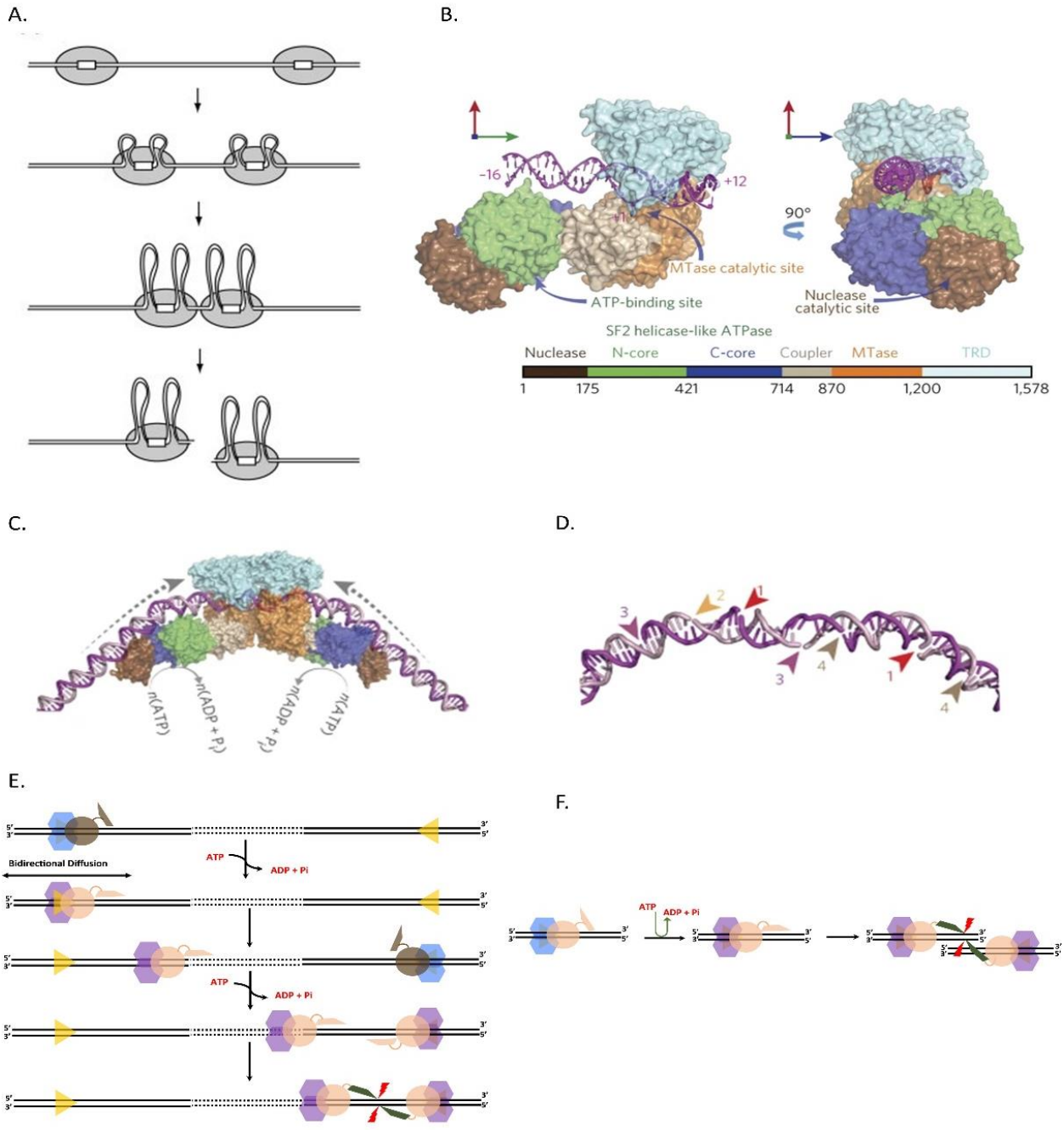


Figure 3: Mechanism of action of Type I, ISP and III RM systems: A) Type I RM systems bind DNA containing two bipartite target sites. Upon ATP hydrolysis, the two opposing HsdR subunit begin to loop DNA while the whole Enzyme complex remains bound to the target site eventually the looping leads to the collision of two enzyme entities where dsDNA cleavage is catalyzed. B) Surface view of the structure of LlaBIII (Type ISP RM system) bound to DNA. Type ISP RM systems consist of all the modular domains present in a single polypeptide. C) Model showing two Type ISP RM entities colliding with each other after translocating on DNA powered by ATP hydrolysis. As seen, Type ISP translocate of DNA without causing its looping. D) Experimentally determined positions of cleavage after collision of Type ISP RM systems. The colored arrow heads represent nicks on either strand of DNA. Due to the presence of numerous such nicks, DNA eventually undergoes irreparable cleavage. E) Type III RM systems function as molecular switches and undergo a 1D diffusion mode of sliding on DNA. Upon encountering another Type III enzyme at the opposing target-site hydrolyzing ATP, cleavage is catalyzed. F) 3D diffusion model of cleavage of DNA with single target site. Both the bound enzymes need to be in an ATP hydrolyzing state for cleavage to occur. (A) Adapted from *Ishikawa et al., 2010, DNA Research.* (B – D) Adapted from *Chand et al., 2015, Nature Chemical Biology.* (E) Adapted from *Ahmed et al., 2018, Nucleic Acids Research.*

cofactor (Figure 2B) (50). But majority of the RM systems that bacteria possess are not site-specific cutters and are composed of multi sub-unit/domains and are categorized as Type I, ISP, III and IV RM systems (48)(66). For the sake of simplicity, we call Type I, ISP, II and III modification-independent restriction systems and Type IV modification-dependent restriction systems (Figure 2C) (48).

Type I RM systems are multi subunit proteins encoded by the *hsd* (Host specificity determinant) operon composed of three components HsdS (Specificity), HsdM (Methylation) and HsdR (Restriction). Incidentally the restriction endonuclease EcoB discovered by Werner Arber and Stuart Linn belonged to the Type I restriction modification systems. Type I RM systems usually occur in two stoichiometric configurations viz., a restriction and modification active R_2M_2S configuration and a restriction deficient modification active M_2S configuration (67, 68). The restriction activity of Type I RM systems requires SAM, ATP and Mg^{2+} whereas the modification activity requires only SAM and Mg^{2+} as cofactors (67, 68). The target-site for Type I RM systems is a bipartite sequence separated by 5-8 non-specific nucleotides (68, 69). For cleavage to occur, the DNA must contain two or more target sites. Once the enzyme binds to its target-site it begins to hydrolyze ATP and translocates on dsDNA acting like a molecular motor eventually encountering another enzyme complex at which point cleavage occurs (30). Hence all Type I RM systems cut approximately half-way between the two bipartite target-sites (Figure 3A). The M_2S configuration and the R_2M_2S configuration (in the absence of ATP) methylate DNA in the presence of SAM and Mg^{2+} thereby protecting self-DNA from the deleterious effects of restriction (70, 71). The HsdR subunit of Type I RM systems possess a SF2 helicase-like ATPase hence explaining its ability to integrate ATP hydrolysis to translocation causing long range communication between two enzyme complexes and eventually catalyzing cleavage of dsDNA. The SF2 helicase-like ATPase functions like a translocase rather than a canonical helicase (difference mentioned in previous sections) and contains the seven (I, Ia – VI) conserved motifs including the Walker A and Walker B motifs (72, 73). Each active Type I RM complex contains two HsdR subunit which means that it has two active ATPase active sites. Based on electron micrographs and the molecular architecture of Type I RM systems it was determined that enzyme remains stationarily bound to its target site even after ATP hydrolysis due to concerted action of two SF2 helicase-like ATPase

in inverted orientation (74, 75). The two SF2 helicase-like ATPases acting in tandem results in the DNA being pulled by looping towards the enzyme complex (76). Type ISP a subtype of Type I RM systems, is so named because the restriction component, the SF2 helicase-like ATPase, target recognition component and the methyltransferase component are present as multiple domains on a single-polypeptide (SP) (Figure 3B) (31, 77, 78). Each SP consists of one domain for each of the aforementioned components and like Type I RM systems, two protein entities upon ATP hydrolysis translocate and collide in an opposing direction for successful cleavage to occur and the position of dsDNA breaks is approximately halfway between the two target sites (Figure 3C and 3D) (31). However, unlike Type I RM systems Type ISP RM systems actively translocate on DNA without causing looping (31).

Type III RM systems are also multi subunit protein complexes made up of modification (Mod) and restriction (Res) components. The Mod component consists of a target recognition domain and a methyltransferase domain whereas the Res component is made up of a nuclease domain and a SF2 helicase-like ATPase. *E. coli* consists of two Type III RM systems EcoP1I and EcoP15I, housed on autonomous plasmids, which are well studied (79). Due to the lack of complete structural information on Type III RM systems the functioning of this class of enzymes has been deduced primarily through detailed biophysical and biochemical experiments. The enzyme complex has two stochiometric configurations, a restriction and methylation active RM_2 state and a restriction deficient and methylation active M_2 state. For efficient restriction activity, Type III RM systems require ATP, SAM and Mg^{2+} cofactors however, methylation can occur only in the presence of SAM and Mg^{2+} (80). Type III RM systems require two target sites in inverted orientation for high efficiency cleavage to occur indicating long range communication between two protein complexes (80, 81). However, a series of biochemical, biophysical and single-molecule studies have indicated that the long-range communication between the two protein complexes in Type III RM systems is distinct from Type I and Isp RM systems (82, 83). The lack of high ATPase activity deduced by single-turnover reactions has led to the hypothesis that the SF2 helicase-like ATPase domain (though sharing structural homology with other canonical SF2 helicases) in Type III RM systems functions as a molecular switch (29, 83, 84). An initial burst of ATP hydrolysis leads to a conformational change in the protein complex making the enzyme diffusion competent along the

1D contour of dsDNA (84). This causes a passive mode of translocation leading to convergence with a stationary protein complex bound to the target site in an ATP hydrolyzing state. The two enzymes catalyze cleavage 25 – 28 bp away from the target-site of the stationary enzyme (Figure 3E) (85). Recent studies have also shown cleavage of dsDNA having single-target site which occurs at high enzyme concentration (86)(86, 87). This unique property of Type III RM systems is attributed to a 3D diffusion mode of cleavage where, two target site bound ATP hydrolyzing protein complexes can interact with each other eventually catalyzing cleavage of both the enzyme bound DNA molecules (Figure 3F) (87).

Type IV restriction systems unlike other restriction endonucleases recognize modified DNA and hence lack a cognate methyltransferase (88). The presence of modification-dependent restriction enzymes is hypothesized to be an effect of the evolutionary arms race between host (Bacterium) and pathogen (Phage/MGEs) (89). The commonality between all the enzymes belonging to this class of restriction systems is their ability to recognize and cleave modified DNA. The modifications recognized could range from base specific methylation (-CH_3; m5C, m4C or m6A), cytosine hydroxymethylation (-CH_2OH; hm5C) to glucosyl-5-hydroxymethylcytosine (g5hmC) depending on the target recognition protein fold of these enzymes (90–92). Apart from this, Type IV restriction systems are a very diverse class containing both proteins with a relatively simple architecture like DpnI and multi-subunit protein complexes like McrBC. Due to their varied 3D architecture and protein folds they lack conservation in cofactor requirement and the nature of target-sites recognized on dsDNA. McrBC is a Type IV restriction systems made up of two components; McrB a polypeptide containing a target recognition domain and GTPase domain and McrC a polypeptide containing the endonuclease domain. These two components form a tetradecameric architecture with two McrB hexamers brought together by a dimer of McrC (93)(94). McrB contains the AAA+ protein fold and hence functions as an annular molecular motor forming GTP binding sites at the interface of protomers. McrBC requires two target-sites with cleavage occurring close to one of the target sites upon GTP hydrolysis (95–97). Recent structural advances in understanding the nature of nucleotide hydrolysis have revealed a sequential mode of GTP hydrolysis by McrB which is stimulated by McrC (94).

The presence of these strong barriers to HGT determines the gene pool amongst bacterial populations and hence can even modulate the ability of bacteria to gain antibiotic resistance genes (98, 99). Pathogenic bacteria can cause a variety of life-threatening maladies to human beings and livestock. Humans were exposed to these pathogenic bacteria mainly during surgical procedures and hence post-operative deaths were a common sight in early 1800s (100, 101). However, the mortality rates reduced when Joseph Lister introduced the process of sterilization of surgical instruments with phenol (carbolic acid) in late 1800s (100, 101). However, the morality rates of injured soldiers plummeted during world wars only after the discovery of Penicillin by Sir Alexander Fleming (In 1928) and Prontosil Rubrum by Gerhard Domagk (In 1932) (102, 103). These wonder compounds were named antibiotics and research and development has led to the discovery of many different natural and chemically synthesized antibiotics which are presently to treat bacterial infections.

Introduction to Antibiotic resistance and the role of REases in modulating HGT

Antibiotics are compounds that are used to specifically neutralize bacteria. Based on the nature of the antibiotic and the concentrations used, they can be either bacteriostatic and bactericidal (104, 105). Antibiotics usually function by interrupting certain physiologically important processes in bacteria such as, inhibition of cell wall synthesis, disruption of fundamental processes such as transcription or translation etc. (106). However, the evolutionary arms race between bacteria and antibiotics has led to novel ways by which bacteria circumvent the effects of these compounds and as a result, antibiotic resistance is known to naturally occur in niches of co-existing microflora. These novel mechanisms could vary from the presence of enzymes for degrading antibiotics to the presence of efflux pumps to flush out antibiotics (107). Antibiotic resistance is a naturally occurring phenomenon in certain bacteria to counter the harmful effects of metabolites released by other bacteria or fungi in a common niche (108, 109). But, the indiscriminate use and abuse of these wonder drugs by us in controlling bacterial infection has led to either specifically selecting for or evolving novel antibiotic resistant variants of certain clinically relevant bacterial species (110).

We are fast exhausting our repertoire of antibiotics that are effective against bacteria and hence there is a need for novel ways of combating this predicament. To understand how to combat antibiotic resistant organisms it is very important to elucidate:

1. Types of antibiotic resistance
2. The mechanisms of antibiotic resistance
3. The dynamics of the natural barriers to horizontal gene transfer

Types of antibiotic resistance:

Types of antibiotic resistance are varied and may differ based on the Gram status of the bacteria. But they can broadly be categorized as: Non-responsiveness and physical barriers to antibiotics; intrinsic resistance to a subset of susceptible bacteria due to chromosomal mutations that are inheritable; and acquired resistance due to mobile genetic elements (MGEs).

Non-responsiveness and physical barriers to antibiotics: For antibiotics to work effectively, they need access to their intracellular targets. Bacterial cells in the stationary phase of their life cycle tend to be non-responsive to antibiotic exposure due to this metabolically and reproductively dormant state (111, 112). Certain Gram-negative bacteria have a thick lipopolysaccharide layer thereby preventing the antibiotic from penetrating the cell. Also, certain bacteria grow in biofilms thereby increasing the inaccessibility of the cells towards the center of the biofilm to antibiotics (113, 114). However, these mechanisms are not resistance in its true sense because they are non-specific and as a result a special term called 'persistence' is used to explain these phenomena (115).

Intrinsic resistance to a subset of susceptible bacteria due to chromosomal mutations: Certain bacterial populations might have subsets of cells that harbor chromosomal mutations which render them resistant to antibiotics (107). Upon antibiotic exposure, these resistant cells dominate and increase in number while the wild type cells (susceptible) die out. However, the maintenance of these resistant mutants is not an evolutionarily cost-effective preposition and as a result in the absence of the antibiotic these mutants die out while the wild type cells take precedence (116). These chromosomal mutations could lead to: modifications of targets for antibiotics thereby rendering the antibiotic useless; efflux pumps binding to antibiotics to eject

them out of the cell; increase in enzyme concentrations and or enzyme efficiency to specifically degrade antibiotics etc. (107)

Acquired resistance due to mobile genetic elements (MGEs): horizontal gene transfer is the process of foreign mobile genetic elements by bacteria through the modes of transformation, conjugation and transduction (117). Horizontal gene transfer is known to occur across bacterial species thereby making it a potent evolutionary force. Through the process of HGT gene cassettes potentially containing antibiotic resistance genes can be transferred from one bacterium to another (118)(119). This process of acquired resistance possess one of the greatest challenges to the present healthcare system.

The mechanisms of antibiotic resistance:

Enzymatic degradation of antibiotics: Many Gram-negative and certain Gram-positive bacteria possess enzymes that naturally degrade β-lactam antibiotics such as penicillin (120–122). These enzymes called β-lactamases are known to target the amide bond of the antibiotic ring and breaking it leads to the inactivation of the antibiotic. In case of some Gram-positive bacteria like staphylococci the β-lactamase are transferred horizontally on a plasmid (123).

Modifications on the targets of the antibiotics: β-lactam antibiotics such as penicillin usually function by obstructing the synthesis of the cell wall by binding to PBPs (penicillin binding proteins) (107). These proteins are necessary for the cross linking of the peptidoglycan layer to maintain the coherence of the cell wall in bacteria. However, certain strains of bacteria have alternative PBPs that can perform the task of crosslinking in spite of the presence of the antibiotics and are hence resistant to it (124, 125).

Efflux pumps: Enzyme machinery present that can extrude harmful chemicals is the basis for many bacteria possessing the ability to resist antibiotics. One of the first discovered efflux pumps was that for the antibiotic Tetracycline in *E. coli* which functioned by using the energy from a proton gradient (126, 127). This efflux pump is encoded by the *tet* genes and is mainly found in gram-negative bacteria. Another example is the efflux pump encoded by the *mef* genes which confer resistance to macrolides such as erythromycin (128).

Modifications on the antibiotics: One of the most efficient ways that bacteria have evolved to circumvent the effects of antibiotics is by modifying antibiotic molecules thereby inactivating them. The modifications that bacteria bring about on these molecules are; phosphorylation of aminoglycosides and chloramphenicol, acetylation of chloramphenicol and streptogramins and adenylation of aminoglycosides and lincosamides (129–131). One of the most studied mechanisms is the acetylation of chloramphenicol catalyzed by chloramphenicol acetyltransferases (CATs) (132). Different types of CATs have been characterized from different gram-positive and gram-negative bacteria conferring either low-level or high-level resistance towards chloramphenicol. These enzymes are primarily present on MGEs making it an easily transferrable form of antibiotic resistance between bacterial populations.

The dynamics of the natural barriers to horizontal gene transfer:

As mentioned earlier, bacteria have an entire battery of mechanisms to prevent HGT as mobile genetic elements might not always be beneficial for a bacterium. However, MGEs might also harbor genes that confer antibiotic resistance hence, understanding the interaction of barriers to HGT might be seminal in understanding the spread of antibiotic resistance. In *Pseudomonas aeuroginosa* and *Acetinobacter baumannii* it was revealed that the presence of functional CRISPR-Cas systems led to the reduction in presence of intracellular integrative conjugative elements (ICEs), prophages and plasmids (133, 134). In fact, in *Klebsiella pneumoniae* and *Enterococcus faecium,* presence of CRISPR-Cas systems has led to very few AMR genes being present in their genomes (135, 136). These results highlight the negative effect that barriers to HGT like the CRISPR-Cas system have on AMR gene acquisition.

Another studied example of the aforementioned phenomenon is a gene encoding for a restriction endonuclease in *S. aureus* identified as a major barrier to horizontal gene transfer. Interestingly, this gene is present in methicillin-resistant *S. aureus* (MRSA) but is either absent or truncated in certain strains of vancomycin-resistant *S. aureus* suggesting that this restriction endonuclease is a potent barrier to HGT and might play a role in preventing the bacterium from transforming into a vancomycin-resistant form (137, 138). It is hypothesized that to procure the vancomycin resistance genes through HGT certain strains of MRSA evolved to either express a truncated

version of the protein or to completely lose the gene. Based on a preliminary bioinformatic analysis, this restriction endonuclease was categorized as a Type III R-M system (137). However, upon initial biochemical characterization, it was revealed that it cleaves methylated cytosine in the context of the target site 5'-S^{5m}CNGS-3' (S = C/G and N = A/T/G/C) (139). This restriction endonuclease was named SauUSI and categorized as a Type IV Restriction system (139).

About SauUSI

SauUSI is a single polypeptide Type IV Restriction system consisting a N-terminal Phospholipase-D (PLD) nuclease, a SF2 helicase-like ATPase and a C-terminal SET and RING associated (SRA) target recognition domain (139, 140). Biochemical assays on SauUSI revealed that it uses energy from ATP hydrolysis to catalyze double stranded DNA breaks on methylated or hydroxymethylated DNA (139, 140). The presence methylation or hydroxymethylation on cytosine (m5C or hm5C) is an absolute requirement for the enzymatic activity as methylation on adenine or non-methylated DNA substrates are refractive to cleavage. A preliminary model proposed based on the initial biochemical characterization suggested that SauUSI required two or more target sites for its cleavage activity (139). However, this thesis focuses on the novel cleavage mechanisms that we have unraveled while performing a detailed biochemical and structural characterization of SauUSI (140). Further, we propose a mechanism of cleavage which reconciles all our findings.

References

1. Howard,J. (2002) Mechanics of motor proteins. In Flyvbjerg,F., Jülicher,F., Ormos,P., David,F. (eds), *Physics of bio-molecules and cells. Physique des biomolécules et des cellules*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 69–94.

2. Gorbalenya,A.E. and Koonin,E. V (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.*, **3**, 419–429.

3. Schliwa,M. and Woehlke,G. (2003) Molecular motors. *Nature*, **422**, 759–765.

4. Hackney,D.D. (1996) The kinetic cycles of myosin, kinesin, and dynein. *Annu. Rev. Physiol.*, **58**, 731–750.

5. Patel,S.S. and Donmez,I. (2006) Mechanisms of Helicases*. *J. Biol. Chem.*, **281**, 18265–18268.

6. Matson,S.W. and Kaiser-Rogers,K.A. (1990) DNA helicases. *Annu. Rev. Biochem.*, **59**, 289–329.

7. Ye,J., Osborne,A.R., Groll,M. and Rapoport,T.A. (2004) RecA-like motor ATPases--lessons from structures. *Biochim. Biophys. Acta*, **1659**, 1–18.

8. Caruthers,J.M. and McKay,D.B. (2002) Helicase structure and mechanism. *Curr. Opin. Struct. Biol.*, **12**, 123–133.

9. Singleton,M.R., Dillingham,M.S. and Wigley,D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.

10. Story,R.M. and Steitz,T.A. (1992) Structure of the recA protein-ADP complex. *Nature*, **355**, 374–376.

11. Subramanya,H.S., Bird,L.E., Brannigan,J.A. and Wigley,D.B. (1996) Crystal structure of a DExx box DNA helicase. *Nature*, **384**, 379–383.

12. Walker,J.E., Saraste,M., Runswick,M.J. and Gay,N.J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.*, **1**, 945–951.

13. Scheffzek,K., Ahmadian,M.R., Kabsch,W., Wiesmüller,L., Lautwein,A., Schmitz,F. and Wittinghofer,A. (1997) The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, **277**, 333–338.

14. Dillingham,M.S., Soultanas,P. and Wigley,D.B. (1999) Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res.*, **27**, 3310–3317.

15. Banroques,J., Cordin,O., Doère,M., Linder,P. and Tanner,N.K. (2011) Analyses of the functional regions of DEAD-box RNA 'helicases' with deletion and chimera constructs tested in vivo and in vitro. *J. Mol. Biol.*, **413**, 451–472.

16. Thomä,N.H., Czyzewski,B.K., Alexeev,A.A., Mazin,A. V, Kowalczykowski,S.C. and Pavletich,N.P. (2005) Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nat.*

Struct. Mol. Biol., **12**, 350–356.

17. Rocak,S. and Linder,P. (2004) DEAD-box proteins: the driving forces behind RNA metabolism. *Nat. Rev. Mol. Cell Biol.*, **5**, 232–241.

18. Velankar,S.S., Soultanas,P., Dillingham,M.S., Subramanya,H.S. and Wigley,D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84.

19. Dillingham,M.S., Wigley,D.B. and Webb,M.R. (2000) Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: measurement of step size and translocation speed. *Biochemistry*, **39**, 205–212.

20. Saikrishnan,K., Powell,B., Cook,N.J., Webb,M.R. and Wigley,D.B. (2009) Mechanistic basis of 5'-3' translocation in SF1B helicases. *Cell*, **137**, 849—859.

21. Yarranton,G.T. and Gefter,M.L. (1979) Enzyme-catalyzed DNA unwinding: Studies on *Escherichia coli rep* protein. *Proc. Natl. Acad. Sci.*, **76**, 1658–1662.

22. Kim,J.L., Morgenstern,K.A., Griffith,J.P., Dwyer,M.D., Thomson,J.A., Murcko,M.A., Lin,C. and Caron,P.R. (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure*, **6 1**, 89–100.

23. Dürr,H., Flaus,A., Owen-Hughes,T. and Hopfner,K.-P. (2006) Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.*, **34**, 4160—4167.

24. Ryan,D.P. and Owen-Hughes,T. (2011) Snf2-family proteins: chromatin remodellers for any occasion. *Curr. Opin. Chem. Biol.*, **15**, 649–656.

25. Stanley,L.K., Seidel,R., van der Scheer,C., Dekker,N.H., Szczelkun,M.D. and Dekker,C. (2006) When a helicase is not a helicase: dsDNA tracking by the motor protein EcoR124I. *EMBO J.*, **25**, 2230—2239.

26. Linder,P. and Jankowsky,E. (2011) From unwinding to clamping - the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.*, **12**, 505–516.

27. Jankowsky,E. (2011) RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.*, **36**, 19–29.

28. Szczelkun,M.D., Friedhoff,P. and Seidel,R. (2010) Maintaining a sense of direction during long-range communication on DNA. *Biochem. Soc. Trans.*, **38**, 404–409.

29. Szczelkun,M.D. (2011) Translocation, switching and gating: potential roles for ATP in long-range communication on DNA by Type III restriction endonucleases. *Biochem. Soc. Trans.*, **39**, 589–594.

30. Studier,F.W. and Bandyopadhyay,P.K. (1988) Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. Natl. Acad. Sci.*, **85**, 4677–4681.

31. Chand,M.K., Nirwan,N., Diffin,F.M., van Aelst,K., Kulkarni,M., Pernstich,C., Szczelkun,M.D. and Saikrishnan,K. (2015) Translocation-coupled DNA cleavage by the Type ISP restriction-

modification enzymes. *Nat. Chem. Biol.*, **11**, 870–877.

32. Enemark,E.J. and Joshua-Tor,L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, **442**, 270–275.

33. Davey,M.J. and O'Donnell,M. (2003) Replicative helicase loaders: ring breakers and ring makers. *Curr. Biol.*, **13**, R594-6.

34. Bujalowski,W. and Klonowska,M.M. (1993) Negative cooperativity in the binding of nucleotides to Escherichia coli replicative helicase DnaB protein. Interactions with fluorescent nucleotide analogs. *Biochemistry*, **32**, 5888–5900.

35. Crampton,D.J., Mukherjee,S. and Richardson,C.C. (2006) DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase. *Mol. Cell*, **21**, 165—174.

36. Gai,D., Zhao,R., Li,D., Finkielstein,C. V and Chen,X.S. (2004) Mechanisms of Conformational Change for a Replicative Hexameric Helicase of SV40 Large Tumor Antigen. *Cell*, **119**, 47–60.

37. Martin,A., Baker,T.A. and Sauer,R.T. (2005) Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature*, **437**, 1115–1120.

38. Hampton,H.G., Watson,B.N.J. and Fineran,P.C. (2020) The arms race between bacteria and their phage foes. *Nature*, **577**, 327–336.

39. Fortier,L.-C. and Sekulovic,O. (2013) Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence*, **4**, 354–365.

40. Ackermann,H.W. (1998) Tailed bacteriophages: the order caudovirales. *Adv. Virus Res.*, **51**, 135–201.

41. Bernheim,A. and Sorek,R. (2020) The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.*, **18**, 113–119.

42. Kronheim,S., Daniel-Ivad,M., Duan,Z., Hwang,S., Wong,A.I., Mantel,I., Nodwell,J.R. and Maxwell,K.L. (2018) A chemical defence against phage infection. *Nature*, **564**, 283–286.

43. Durmaz,E. and Klaenhammer,T.R. (2007) Abortive phage resistance mechanism AbiZ speeds the lysis clock to cause premature lysis of phage-infected Lactococcus lactis. *J. Bacteriol.*, **189**, 1417–1425.

44. Cohen,D., Melamed,S., Millman,A., Shulman,G., Oppenheimer-Shaanan,Y., Kacen,A., Doron,S., Amitai,G. and Sorek,R. (2019) Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature*, **574**, 691–695.

45. Cheng,R., Huang,F., Wu,H., Lu,X., Yan,Y., Yu,B., Wang,X. and Zhu,B. (2021) A nucleotide-sensing endonuclease from the Gabija bacterial defense system. *Nucleic Acids Res.*, **49**, 5216–5229.

46. Oliveira,P.H., Touchon,M. and Rocha,E.P.C. (2014) The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Res.*, **42**, 10618–10631.

47. Hille,F., Richter,H., Wong,S.P., Bratovič,M., Ressel,S. and Charpentier,E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell*, **172**, 1239–1259.

48. Tock,M.R. and Dryden,D.T.F. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.*, **8**, 466–472.

49. Pingoud,A. and Jeltsch,A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.*, **29**, 3705–3727.

50. Pingoud,A., Wilson,G.G. and Wende,W. (2014) Type II restriction endonucleases--a historical perspective and more. *Nucleic Acids Res.*, **42**, 7489–7527.

51. Amitai,G. and Sorek,R. (2016) CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.*, **14**, 67–76.

52. Jansen,R., Embden,J.D.A. van, Gaastra,W. and Schouls,L.M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.*, **43**, 1565–1575.

53. Nussenzweig,P.M. and Marraffini,L.A. (2020) Molecular Mechanisms of CRISPR-Cas Immunity in Bacteria. *Annu. Rev. Genet.*, **54**, 93–120.

54. Doudna,J.A. and Charpentier,E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**, 1258096.

55. DUSSOIX,D. and ARBER,W. (1962) Host specificity of DNA produced by Escherichia coli. II. Control over acceptance of DNA from infecting phage lambda. *J. Mol. Biol.*, **5**, 37–49.

56. LURIA,S.E. and HUMAN,M.L. (1952) A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.*, **64**, 557–569.

57. Arber,W. and Linn,S. (1969) DNA modification and restriction. *Annu. Rev. Biochem.*, **38**, 467–500.

58. Arber,W. and Gitschier,J. (2014) The inventiveness of nature: an interview with Werner Arber. *PLoS Genet.*, **10**, e1004879.

59. Johnson,T.B. and Coghill,R.D. (1925) Researches on pyrimidines. C111. The discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus1. *J. Am. Chem. Soc.*, **47**, 2838–2844.

60. Tourancheau,A., Mead,E.A., Zhang,X.-S. and Fang,G. (2021) Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nat. Methods*, **18**, 491–498.

61. Smith,H.O. and Wilcox,K.W. (1970) A restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *J. Mol. Biol.*, **51**, 379–391.

62. Kelly,T.J.J. and Smith,H.O. (1970) A restriction enzyme from Hemophilus influenzae. II. *J. Mol. Biol.*, **51**, 393–409.

63. Danna,K. and Nathans,D. (1971) Specific cleavage of simian virus 40 DNA by restriction endonuclease of Hemophilus influenzae. *Proc. Natl. Acad. Sci. U. S. A.*, **68**, 2913–2917.

64. Mertz,J.E. and Davis,R.W. (1972) Cleavage of DNA by R 1 restriction endonuclease generates cohesive ends. *Proc. Natl. Acad. Sci. U. S. A.*, **69**, 3370–3374.

65. Szczelkun,M.D. and Halford,S.E. (2013) Restriction Endonuclease. In Maloy,S., Hughes,K. (eds), *Brenner's Encyclopedia of Genetics (Second Edition)*. Academic Press, San Diego, pp. 184–189.

66. Dryden,D.T., Murray,N.E. and Rao,D.N. (2001) Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Res.*, **29**, 3728–3741.

67. Bourniquel,A.A. and Bickle,T.A. (2002) Complex restriction enzymes: NTP-driven molecular motors. *Biochimie*, **84**, 1047–1059.

68. Murray,N.E. (2000) Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol. Mol. Biol. Rev.*, **64**, 412–434.

69. Rao,D.N., Saha,S. and Krishnamurthy,V. (2000) ATP-dependent restriction enzymes. *Prog. Nucleic Acid Res. Mol. Biol.*, **64**, 1–63.

70. Kennaway,C.K., Obarska-Kosinska,A., White,J.H., Tuszynska,I., Cooper,L.P., Bujnicki,J.M., Trinick,J. and Dryden,D.T.F. (2009) The structure of M.EcoKI Type I DNA methyltransferase with a DNA mimic antirestriction protein. *Nucleic Acids Res.*, **37**, 762–770.

71. Kennaway,C.K., Taylor,J.E., Song,C.F., Potrzebowski,W., Nicholson,W., White,J.H., Swiderska,A., Obarska-Kosinska,A., Callow,P., Cooper,L.P., *et al.* (2012) Structure and operation of the DNA-translocating type I DNA restriction enzymes. *Genes Dev.*, **26**, 92–104.

72. Loenen,W.A.M., Dryden,D.T.F., Raleigh,E.A. and Wilson,G.G. (2013) Type I restriction enzymes and their relatives. *Nucleic Acids Res.*, **42**, 20–44.

73. Fairman-Williams,M.E., Guenther,U.-P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.

74. Rosamond,J., Endlich,B. and Linn,S. (1979) Electron microscopic studies of the mechanism of action of the restriction endonuclease of Escherichia coli B. *J. Mol. Biol.*, **129**, 619–635.

75. Yuan,R., Hamilton,D.L. and Burckhardt,J. (1980) DNA translocation by the restriction enzyme from E. coli K. *Cell*, **20**, 237–244.

76. Jindrova,E., Schmid-Nuoffer,S., Hamburger,F., Janscak,P. and Bickle,T.A. (2005) On the DNA cleavage mechanism of Type I restriction enzymes. *Nucleic Acids Res.*, **33**, 1760–1766.

77. Kong,J. and Josephsen,J. (2002) The ability of the plasmid-encoded restriction and modification system LlaBIII to protect Lactococcus lactis against bacteriophages. *Lett. Appl. Microbiol.*, **34**, 249–253.

78. Madsen,A. and Josephsen,J. (2001) The LlaGI restriction and modification system of Lactococcus lactis W10 consists of only one single polypeptide. *FEMS Microbiol. Lett.*, **200**, 91–96.

79. Arber,W. and Wauters-Willems,D. (1970) Host specificity of DNA produced by Escherichia coli. XII. The two restriction and modification systems of strain 15T-. *Mol. & Gen. Genet. MGG*, **108**, 203—217.

80. Meisel,A., Bickle,T.A., Krüger,D.H. and Schroeder,C. (1992) Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. *Nature*, **355**, 467–469.

81. van Aelst,K., Tóth,J., Ramanathan,S.P., Schwarz,F.W., Seidel,R. and Szczelkun,M.D. (2010) Type III restriction enzymes cleave DNA by long-range interaction between sites in both head-to-head and tail-to-tail inverted repeat. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 9123–9128.

82. Ramanathan,S.P., van Aelst,K., Sears,A., Peakman,L.J., Diffin,F.M., Szczelkun,M.D. and Seidel,R. (2009) Type III restriction enzymes communicate in 1D without looping between their target sites. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 1748—1753.

83. Schwarz,F.W., Tóth,J., van Aelst,K., Cui,G., Clausing,S., Szczelkun,M.D. and Seidel,R. (2013) The Helicase-Like Domains of Type III Restriction Enzymes Trigger Long-Range Diffusion Along DNA. *Science (80-.).*, **340**, 353–356.

84. Tóth,J., Bollins,J. and Szczelkun,M.D. (2015) Re-evaluating the kinetics of ATP hydrolysis during initiation of DNA sliding by Type III restriction enzymes. *Nucleic Acids Res.*, **43**, 10870–10881.

85. Hadi,S.M., Bächi,B., Shepherd,J.C., Yuan,R., Ineichen,K. and Bickle,T.A. (1979) DNA recognition and cleavage by the EcoP15 restriction endonuclease. *J. Mol. Biol.*, **134**, 655–666.

86. Möncke-Buchner,E., Rothenberg,M., Reich,S., Wagenführ,K., Matsumura,H., Terauchi,R., Krüger,D.H. and Reuter,M. (2009) Functional Characterization and Modulation of the DNA Cleavage Efficiency of Type III Restriction Endonuclease EcoP15I in Its Interaction with Two Sites in the DNA Target. *J. Mol. Biol.*, **387**, 1309–1319.

87. Ahmad,I., Kulkarni,M., Gopinath,A. and Saikrishnan,K. (2018) Single-site DNA cleavage by Type III restriction endonuclease requires a site-bound enzyme and a trans-acting enzyme that are ATPase-activated. *Nucleic Acids Res.*, **46**, 6229–6237.

88. Bickle,T.A. and Krüger,D.H. (1993) Biology of DNA restriction. *Microbiol. Rev.*, **57**, 434–450.

89. Tesfazgi Mebrhatu,M., Wywial,E., Ghosh,A., Michiels,C.W., Lindner,A.B., Taddei,F., Bujnicki,J.M., Van Melderen,L. and Aertsen,A. (2011) Evidence for an evolutionary antagonism between Mrr and Type III modification systems. *Nucleic Acids Res.*, **39**, 5991–6001.

90. Wang,H., Guan,S., Quimby,A., Cohen-Karni,D., Pradhan,S., Wilson,G., Roberts,R.J., Zhu,Z. and Zheng,Y. (2011) Comparative characterization of the PvuRts1I family of restriction enzymes and their application in mapping genomic 5-hydroxymethylcytosine. *Nucleic Acids Res.*, **39**, 9294–9305.

91. Szwagierczak,A., Brachmann,A., Schmidt,C.S., Bultmann,S., Leonhardt,H. and Spada,F. (2011) Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine. *Nucleic Acids Res.*, **39**, 5149–5156.

92. Kisiala,M., Copelas,A., Czapinska,H., Xu,S.-Y. and Bochtler,M. (2018) Crystal structure of the modification-dependent SRA-HNH endonuclease TagI. *Nucleic Acids Res.*, **46**, 10489–10503.

93. Panne,D., Müller,S.A., Wirtz,S., Engel,A. and Bickle,T.A. (2001) The McrBC restriction endonuclease assembles into a ring structure in the presence of G nucleotides. *EMBO J.*, **20**, 3210–3217.

94. Nirwan,N., Singh,P., Mishra,G.G., Johnson,C.M., Szczelkun,M.D., Inoue,K., Vinothkumar,K.R. and Saikrishnan,K. (2018) Hexameric assembly of the AAA+ protein McrB is necessary for GTPase activity. *Nucleic Acids Res.*, **47**, 868–882.

95. Pieper,U. and Pingoud,A. (2002) A mutational analysis of the PD...D/EXK motif suggests that McrC harbors the catalytic center for DNA cleavage by the GTP-dependent restriction enzyme McrBC from Escherichia coli. *Biochemistry*, **41**, 5236–5244.

96. Panne,D., Raleigh,E.A. and Bickle,T.A. (1999) The McrBC endonuclease translocates DNA in a reaction dependent on GTP hydrolysis. *J. Mol. Biol.*, **290**, 49–60.

97. Stewart,F.J. and Raleigh,E.A. (1998) Dependence of McrBC cleavage on distance between recognition elements. *Biol. Chem.*, **379**, 611–616.

98. Lerminiaux,N.A. and Cameron,A.D.S. (2019) Horizontal transfer of antibiotic resistance genes in clinical environments. *Can. J. Microbiol.*, **65**, 34–44.

99. Thomas,C.M. and Nielsen,K.M. (2005) Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nat. Rev. Microbiol.*, **3**, 711–721.

100. Gilmore,O.J. (1977) 150 years after. A tribute to Joseph Lister. *Ann. R. Coll. Surg. Engl.*, **59**, 199–204.

101. Jessney,B. (2012) Joseph Lister (1827-1912): a pioneer of antiseptic surgery remembered a century after his death. *J. Med. Biogr.*, **20**, 107–110.

102. Sulek,K. (1968) [Nobel prize for Gerhard Domagk in 1939 for discovery of the antibacterial activity of prontosil]. *Wiad. Lek.*, **21**, 1089.

103. Tan,S.Y. and Tatsumura,Y. (2015) Alexander Fleming (1881-1955): Discoverer of penicillin. *Singapore Med. J.*, **56**, 366–367.

104. JAWETZ,E. and GUNNISON,J.B. (1953) Antibiotic synergism and antagonism; an assessment of the problem. *Pharmacol. Rev.*, **5**, 175–192.

105. Ocampo,P.S., Lázár,V., Papp,B., Arnoldini,M., Abel zur Wiesch,P., Busa-Fekete,R., Fekete,G., Pál,C., Ackermann,M. and Bonhoeffer,S. (2014) Antagonism between bacteriostatic and bactericidal antibiotics is prevalent. *Antimicrob. Agents Chemother.*, **58**, 4573–4582.

106. Reygaert,W.C. (2018) An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiol.*, **4**, 482—501.

107. Munita,J.M. and Arias,C.A. (2016) Mechanisms of Antibiotic Resistance. *Microbiol. Spectr.*, **4**.

108. Nguyen,B.-A.T., Chen,Q.-L., He,J.-Z. and Hu,H.-W. (2020) Microbial regulation of natural antibiotic resistance: Understanding the protist-bacteria interactions for evolution of soil resistome. *Sci. Total Environ.*, **705**, 135882.

109. Allen,H.K., Moe,L.A., Rodbumrer,J., Gaarder,A. and Handelsman,J. (2009) Functional metagenomics reveals diverse $β$-lactamases in a remote Alaskan soil. *ISME J.*, **3**, 243–251.

110. Hawkey,P.M. (1998) The origins and molecular basis of antibiotic resistance. *BMJ*, **317**, 657–660.

111. Keren,I., Kaldalu,N., Spoering,A., Wang,Y. and Lewis,K. (2004) Persister cells and tolerance to antimicrobials. *FEMS Microbiol. Lett.*, **230**, 13–18.

112. Wood,T.K., Knabel,S.J. and Kwan,B.W. (2013) Bacterial Persister Cell Formation and Dormancy. *Appl. Environ. Microbiol.*, **79**, 7116–7121.

113. Mah,T.-F. (2012) Biofilm-specific antibiotic resistance. *Future Microbiol.*, **7**, 1061–1072.

114. Van Acker,H., Van Dijck,P. and Coenye,T. (2014) Molecular mechanisms of antimicrobial tolerance and resistance in bacterial and fungal biofilms. *Trends Microbiol.*, **22**, 326–333.

115. Harms,A., Maisonneuve,E. and Gerdes,K. (2016) Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science*, **354**.

116. Martinez,J.L. (2014) General principles of antibiotic resistance in bacteria. *Drug Discov. Today. Technol.*, **11**, 33–39.

117. Heuer,H. and Smalla,K. (2007) Horizontal gene transfer between bacteria. *Environ. Biosafety Res.*, **6**, 3–13.

118. Manson,J.M., Hancock,L.E. and Gilmore,M.S. (2010) Mechanism of chromosomal transfer of Enterococcus faecalis pathogenicity island, capsule, antimicrobial resistance, and other traits. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 12269–12274.

119. Courvalin,P. (1994) Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. *Antimicrob. Agents Chemother.*, **38**, 1447–1451.

120. Bush,K. (2013) Proliferation and significance of clinically relevant β-lactamases. *Ann. N. Y. Acad. Sci.*, **1277**, 84–90.

121. Bush,K. and Jacoby,G.A. (2010) Updated functional classification of beta-lactamases. *Antimicrob. Agents Chemother.*, **54**, 969–976.

122. Bush,K. and Bradford,P.A. (2016) β-Lactams and β-Lactamase Inhibitors: An Overview. *Cold Spring Harb. Perspect. Med.*, **6**.

123. Fuda,C.C.S., Fisher,J.F. and Mobashery,S. (2005) Beta-lactam resistance in Staphylococcus aureus: the adaptive resistance of a plastic genome. *Cell. Mol. Life Sci.*, **62**, 2617–2633.

124. Reygaert,W. (2009) Methicillin-resistant Staphylococcus aureus (MRSA): molecular aspects of antimicrobial resistance and virulence. *Clin. Lab. Sci. J. Am. Soc. Med. Technol.*, **22**, 115–119.

125. Beceiro,A., Tomás,M. and Bou,G. (2013) Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world? *Clin. Microbiol. Rev.*, **26**, 185–230.

126. McMurry,L., Petrucci,R.E.J. and Levy,S.B. (1980) Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, **77**, 3974–3977.

127. Blanco,P., Hernando-Amado,S., Reales-Calderon,J.A., Corona,F., Lira,F., Alcalde-Rico,M., Bernardini,A., Sanchez,M.B. and Martinez,J.L. (2016) Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants. *Microorganisms*, **4**.

128. Mingoia,M., Morici,E., Brenciani,A., Giovanetti,E. and Varaldo,P.E. (2014) Genetic basis of the association of resistance genes mef(I) (macrolides) and catQ (chloramphenicol) in streptococci. *Front. Microbiol.*, **5**, 747.

129. Schwarz,S., Kehrenberg,C., Doublet,B. and Cloeckaert,A. (2004) Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS Microbiol. Rev.*, **28**, 519–542.

130. Robicsek,A., Strahilevitz,J., Jacoby,G.A., Macielag,M., Abbanat,D., Hye Park,C., Bush,K. and Hooper,D.C. (2006) Fluoroquinolone-modifying enzyme: a new adaptation of a common aminoglycoside acetyltransferase. *Nat. Med.*, **12**, 83–88.

131. Ramirez,M.S. and Tolmasky,M.E. (2010) Aminoglycoside modifying enzymes. *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.*, **13**, 151–171.

132. Shaw,W. V (1983) Chloramphenicol acetyltransferase: enzymology and molecular biology. *CRC Crit. Rev. Biochem.*, **14**, 1–46.

133. Wheatley,R.M. and MacLean,R.C. (2021) CRISPR-Cas systems restrict horizontal gene transfer in Pseudomonas aeruginosa. *ISME J.*, **15**, 1420–1433.

134. Tyumentseva,M., Mikhaylova,Y., Prelovskaya,A., Tyumentsev,A., Petrova,L., Fomina,V., Zamyatin,M., Shelenkov,A. and Akimkin,V. (2021) Genomic and Phenotypic Analysis of Multidrug-Resistant Acinetobacter baumannii Clinical Isolates Carrying Different Types of CRISPR/Cas Systems. *Pathog. (Basel, Switzerland)*, **10**.

135. Wang,G., Song,G. and Xu,Y. (2020) Association of CRISPR/Cas System with the Drug Resistance in Klebsiella pneumoniae. *Infect. Drug Resist.*, **13**, 1929–1935.

136. Price,V.J., McBride,S.W., Hullahalli,K., Chatterjee,A., Duerkop,B.A. and Palmer,K.L. (2019) Enterococcus faecalis CRISPR-Cas Is a Robust Barrier to Conjugative Antibiotic Resistance Dissemination in the Murine Intestine. *mSphere*, **4**.

137. Corvaglia,A.R., François,P., Hernandez,D., Perron,K., Linder,P. and Schrenzel,J. (2010) A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical Staphylococcus aureus strains. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 11954—11958.

138. Monk,I.R., Shah,I.M., Xu,M., Tan,M.-W. and Foster,T.J. (2012) Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *MBio*, **3**.

139. Xu,S.-Y., Corvaglia,A.R., Chan,S.-H., Zheng,Y. and Linder,P. (2011) A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300. *Nucleic Acids Res.*, **39**, 5597–5610.

140. Tumuluri,V.S., Rajgor,V., Xu,S.-Y., Chouhan,O.P. and Saikrishnan,K. (2021) Mechanism of DNA cleavage by the endonuclease SauUSI: a major barrier to horizontal gene transfer and antibiotic

resistance in Staphylococcus aureus. *Nucleic Acids Res.*, **49**, 2161–2178.

SCOPE OF THE THESIS

This thesis submitted to IISER Pune for the partial fulfillment of the degree of Doctor of Philosophy (Ph.D.) focusses on elucidating the mechanism of action of a Type IV bacterial restriction enzyme SauUSI, which cleaves methylated DNA. Biochemical assays on SauUSI have revealed a unique mode of DNA cleavage. In addition to cleaving DNA having a single target sequence, SauUSI nucleolytically degrades DNA having at least two of its target sequences by catalyzing multiple double-stranded DNA breaks between the target sites. This results in the DNA being shredded into multiple fragments. We have deciphered the structure of the apo-form of the protein which reveals that it is a homodimer with the nuclease domain at the interface, which provides a molecular basis for single-site and two-site DNA cleavage. While investigating the nucleolytic ability of SauUSI, we uncover that it has the ability to cleave single-stranded DNA. This cleavage activity of SauUSI is the first evidence that a restriction endonuclease can function as a *bona fide* single-stranded endonucleases as well. Further, this activity was negatively regulated by increasing concentrations of adenine nucleoside di/triphosphates. We hypothesize that this novel ssDNA cleavage can act as a barrier to HGT involving ssDNA such as conjugation, single-stranded phage-based transduction and natural transformation. Interestingly, the strong negative regulation conferred by Adenine nucleoside di/triphosphates suggests an abortive infection (Abi) mode of bacterial defense. Hence the biochemical and structural characterization of SauUSI reveals its true potential as a potent barrier to HGT in bacteria, giving a new perspective on antibiotic resistance and acquisition of antibiotic resistance factors.

CHAPTER 2: Heterologous Expression and High Degree Purification of the Restriction Endonuclease SauUSI

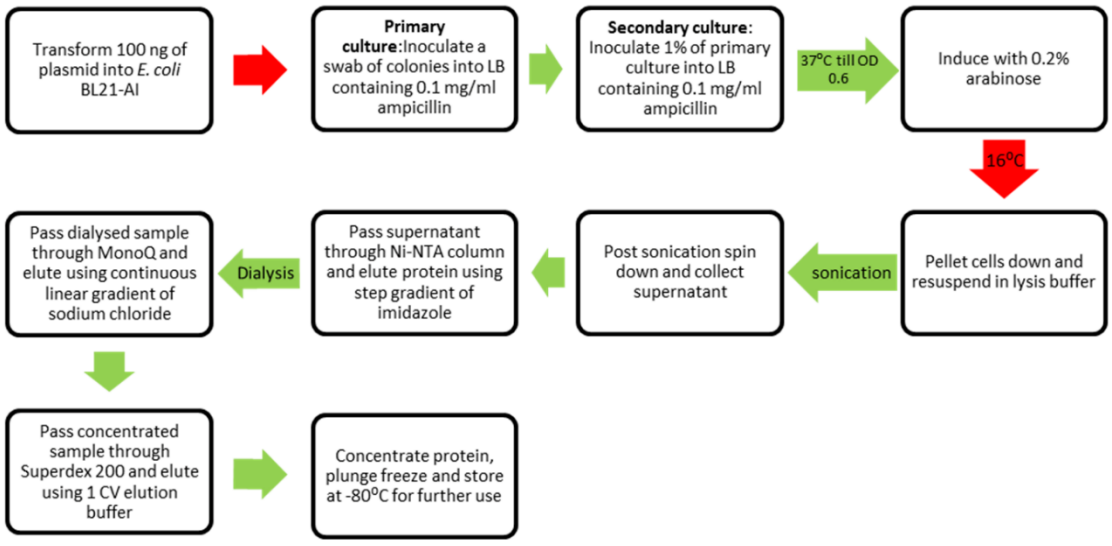
Reprinted from: Tumuluri et al., *Bio-protocol*, 2022.

DOI: 10.21769/BioProtoc.4275

Abstract

Mechanisms that target and destroy foreign nucleic acids are major barriers to horizontal gene transfer (HGT) in prokaryotes. Amongst them, restriction-modification (R-M) systems are found in ≥75% of the sequenced genomes in Bacteria and Archaea. Due to their high target sequence specificity and potent nucleolytic activity, R-M systems are used as a paradigm to elucidate the mechanisms of DNA binding and cleavage. Since these enzymes modulate HGT, they are one of the machineries implicated in the ability of a bacterium to gain antibiotic resistance. This protocol provides a detailed purification strategy for the Type IV restriction endonuclease SauUSI from *Staphylococcus aureus*. This protocol eventually leads to ≥95% purity of protein which can then be used for crystallographic and biochemical purposes.

Graphic abstract:



Workflow for purification of SauUSI

Introduction

Restriction-modification (R-M) systems function as the innate immune system of bacteria thus, preventing not only infection by bacteriophages but also other potentially harmful mobile genetic elements. There are a variety of bacterial R-M systems that differ in their target site epigenetic status, oligomeric structure, cofactor requirements and mode of DNA cleavage (1). SauUSI is a Type IV restriction endonuclease present in *Staphylococcus aureus* that binds specifically to the methylated DNA target sequence 5'-S^{5m}CNGS-3' (S=C/G and N=A/T/G/C) (2). SauUSI has the ability to cleave DNA by two distinct mechanisms they are a) a highly efficient convergence-based cleavage mechanism when there are two or more target sequences and b) a lesser efficient switch-based mechanism on DNA with only one target sequence (3). Studies on the cleavage activity of SauUSI highlight its role in preventing the acquisition of exogenous genetic elements that have the ability to transform methicillin-resistant *Staphylococcus aureus* (MRSA) to vancomycin-resistant *Staphylococcus aureus* (VRSA) (3–5).

For the *in-vitro* biochemical characterization and crystallographic studies, SauUSI was purified post expression in the heterologous *Escherichia coli* BL21-AI cells*. For this purpose the sauUSIR gene was* cloned into a pHis17 (ampicillin resistant) vector (6) between the NdeI (upstream) and BamHI (downstream) restriction sites (Figure 1) using a restriction-free (RF) cloning strategy (3). The RF cloning primarily consisted of two sequential polymerase chain reactions (PCRs); the first reaction was to amplify the *sauUSIR* gene with flanking regions complementary to the pHis17 vector and the second reaction was an extension PCR to incorporate the amplified *sauUSIR* gene into the vector. The reaction was treated with DpnI to cleave the parent template and subsequentially transformed into *Escherichia coli* NEB Turbo® cells*.* The presence of the sauUSIR gene was confirmed by DNA sequencing.

Materials and Reagents

1. 100 mm Petri dishes (HiMedia, catalog number: PW1305-1x100NO)
2. 50 ml beaker (Borosil)
3. 200 ml conical flask (Borosil)
4. 2000 ml culture flask (Borosil)
5. 50 ml centrifuge tubes (Eppendorf)
6. 15 ml centrifuge tubes (Eppendorf)
7. 1.5 ml fraction tubes (Eppendorf)
8. 0.5 ml fraction tubes (Eppendorf)
9. SnakeSkin™ dialysis tubing 10000 MWCO (ThermoFisher, catalog number: 68100)
10. 0.2 µm syringe filter (Whatman™, catalog number: 9914-2502)
11. 0.2 µm filter membranes (Millipore, catalog number: GVWP04700)
12. Whatman filter paper (No. 1) (Filtration laboratory, catalog number: T-06648-13)
13. Lysogeny broth (LB) (HiMedia, catalog number: M1245-1KG)
14. Luria-Bertani (LB) Agar (HiMedia, catalog number: M1151-1KG)
15. Ampicillin (Sigma-Aldrich, catalog number: A9518-25G)
16. Arabinose (SRL, catalog number: 52392)
17. Sodium Chloride (NaCl) (Sigma-Aldrich, catalog number: S9888-1KG)
18. Tris(hydroxymethyl)aminomethane (Tris-Base) (Sigma-Aldrich, catalog number: 252859-500G)
19. Hydrochloric acid (HCl) for molecular biology (Qualigens, catalog number: Q29505)
20. Imidazole (Sigma-Aldrich, catalog number: I2399-500G)
21. Magnesium Chloride ($MgCl_2$) (Sigma-Aldrich, catalog number: M8266-100G)

22. (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate) (CHAPS) (Sigma-Aldrich, catalog number: 226947-5G)
23. Glycerol for molecular biology (Sigma-Aldrich, catalog number: 356352-1L-M)
24. Ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich, catalog number: ED-500G)
25. 1,4-Dithiothreitol (DTT) (Sigma-Aldrich, catalog number: 10708984001)
26. Glycine (Sigma-Aldrich, catalog number: G8898-1KG)
27. Sodium dodecyl sulfate (SDS) (Sigma-Aldrich, catalog number: L3771-100G)
28. Ammonium persulfate (APS) (Sigma-Aldrich, catalog number: A3678-25G)
29. Tetramethylethylenediamine (TEMED) (Sigma-Aldrich, catalog number: T9281-25ML)
30. Acrylamide (Sigma-Aldrich, catalog number: A8887-500G)
31. Bis-acrylamide (Sigma-Aldrich, catalog number: 146072-500G)
32. Bio-Rad Precision Plus Protein™ Standards (Bio-Rad, catalog number: 1610374)
33. 2× Laemmli sample buffer (Bio-Rad, catalog number: 1610737EDU)
34. Coomassie Brilliant Blue (R-250) (Sigma-Aldrich, catalog number: 1125530025)
35. Ethanol for molecular biology
36. Glacial Acetic acid for molecular biology
37. Liquid Nitrogen

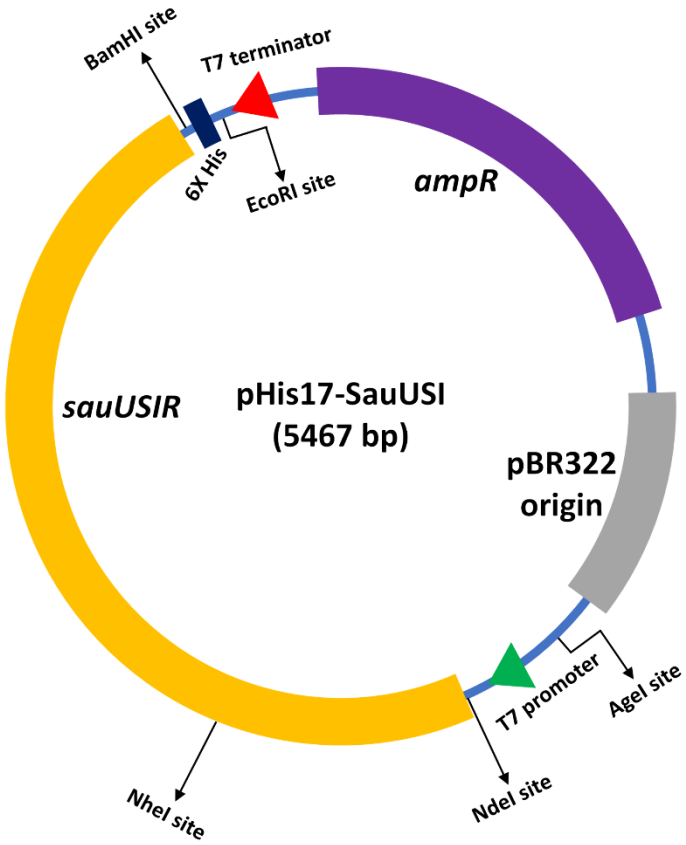


Figure 1: Cartoon representation of the expression vector integrated with *sauUSIR* gene: The *sauUSIR* gene is 2859 bp and codes of a 953 amino acid single polypeptide protein. This gene is integrated between NdeI and BamHI sites in a pHis 17 vector thus providing a C-terminal 6× Histidine tag for ease of purification.

Stocks solutions (see Recipes)

1. 20% (w/v) Arabinose
2. 5 M NaCl
3. 2 M Tris-HCl (pH = 8)
4. 2 M Imidazole
5. 4% (w/v) CHAPS
6. 2 M $MgCl_2$
7. 0.5 M EDTA (pH = 8)
8. 1 M DTT

Buffers for purification (see Recipes)

1. Lysis buffer
2. Buffers for affinity (Ni-NTA) chromatography (Buffer A and Buffer B)

3. Buffer for Dialysis (Buffer B_0)
4. Buffers for anion exchange chromatography (Buffer B_{50} and Buffer B_{1000})
5. Buffer for size exclusion chromatography (Buffer B_{100})
6. SDS-PAGE (see Recipes)
7. 10× stock of TGS (Tris base, glycine and SDS) running buffer
8. Coomassie staining solution
9. Coomassie destaining solution
10. Solutions required for Gel mix
11. 10% resolving gel
12. 5% stacking gel

Equipment

1. Thermoblock (Eppendorf ThermoMixer C, catalog number: 5382000015)
2. Shaker incubator (ThermoScientific, model: MXQ 6000)
3. Avanti J26S-XP centrifuge (Beckman Coulter, model: Avanti J26S-XP, Rotor: JLA-9.1000)
4. Optima-L-100K ultracentrifuge (Beckman Coulter, model: Optima L-100K, Rotor: Type 45 Ti)
5. GE AKTAprime plus FPLC system (GMI, catalog number: 8149-30-0004)
6. Bio-Rad NGC™ chromatography system (Bio-Rad, catalog number: 788001)
7. Vaccum filter assembly (Millipore, model: XX1014720)
8. Sonicator (Sonics-Vibra cell) (Sonics, model: VCX 130)
9. HisTrap™ HP 5 ml Ni-NTA column (GE, catalog number: 17524802)
10. MonoQ™ 10/100 GL column (GE, catalog number: 17516701)
11. Superloop™ (50 ml) (GE, catalog number: 18111382)
12. Superdex 200™ 16/600 pg column (GE, catalog number: GE28-9893-35)
13. SDS-PAGE apparatus (Bio-Rad, catalog number: 1658001FC)

14. Vivaspin^{R} 10000 MWCO centrifugal concentrator (Sartorius, catalog number: Z614602-12EA)

Software

1. Microsoft Excel
2. GraphPad Prism
3. ImageJ

Procedure

A. Bacterial cell transformation

1. Thaw BL21 (AI) competent cells on ice.
2. Under sterile conditions add 100 ng of plasmid to 50 µl of competent cells and place on ice for 12-15 min.
3. Transfer the cells to a thermoblock (heat shock) at 37°C for 3 min or at 42°C for 1.5 min, respectively.
4. Transfer the cells back to ice for 2 min.
5. Add 100 µl of 1× LB to the cells and place at 37°C for 10 min and plate all the cells on a LB Agar plate containing 0.1 mg/ml Ampicillin.
6. Incubate the plate at 37°C overnight.

B. Culture growth

1. Primary culture

   Inoculate a swab (~15-20 colonies) of the transformed cells into 40 ml of 1× LB containing 0.1 mg/ml Ampicillin and grow at 37°C in a shaker incubator (180 RPM) for 5-6 h.

2. Secondary culture

a. Inoculate 1% of primary culture in 1 litre of 1× LB containing 0.1 mg/ml Ampicillin and grow at 37°C in a shaker incubator (180 RPM) and monitor the optical density (OD) of the secondary culture.

b. Once the OD reaches 0.4-0.5 change the temperature of the shaker incubator (180 RPM) to 16°C.

c. Once the OD at 600 nm reaches 0.6 induce the culture with 0.2% (w/v) Arabinose and grow overnight (12-14 h) at 16°C.

C. Cell pelleting and lysate clarification

1. Pelleting down of cells

a. Pellet the cells down using a Avanti J26X-SP centrifuge at an RCF of 5180 (~4,000-5,000 RPM) for 20 min maintaining the temperature at 4°C and then discard the supernatant.

b. Either store the pellet at -80°C after plunge freezing with liquid nitrogen or process it immediately for lysis.

2. Bacterial cell lysis

a. Weigh the cell pellet and then resuspend it in lysis buffer, (volume of lysis buffer to be added is 5 times the weight of the cell pellet *i.e*., 5 ml of lysis buffer for every 1 g of cell pellet).

b. Lyse the cells on ice using a sonicator set at 60% amplitude with a pulse of 1 s and a recovery of 3 s for 3 min. Post this, allow the lysate to sit on ice for 5 min. Continue the sonication protocol for another 2 min.

3. Lysate clarification

a. Centrifuge the lysate using an Optima-L-100K ultracentrifuge for 50 min at 100,000 *× g.* maintaining the temperature at 4°C.

b. Transfer the lysate to a pre-chilled tube or conical flask for loading onto the Ni-NTA column.

D. Ni-NTA column chromatography (perform at 4°C using a GE AKTAprime plus FPLC system)

1. Equilibrate a GE HisTrap™ HP 5 ml Ni-NTA column with 3 column volumes (CV) of Buffer A.

2. Load the supernatant onto the columns at a slow flow rate ~1 ml/min.

3. After the supernatant is loaded, wash the column with 20 CV Buffer A at a flow rate of 2 ml/min.

4. Elute the sample using a step gradient (5%, 10%, 20%, 30%, 50% and 100%) of Buffer B. Gradients of Buffer B were made by mixing of Buffer A and Buffer B. Elute two fractions of 5 ml each at a particular step.

5. Load alternate fractions on a 10% SDS-PAGE gel with an appropriate protein ladder (Bio-Rad Precision Plus Protein™ Standards) (Figure 2). The protein usually elutes in the step gradients of 5% to 30% of Buffer B. Subsequently, based on the SDS-PAGE gel, pool the fractions with most protein, which is usually fractions corresponding to 10%, 20% and 30% of Buffer B (Figure 2).

E. Dialysis (perform at 4°C using SnakeSkin™ dialysis tubing 10000 MWCO)

Dialyse the pooled fractions (to remove excess salt and imidazole) using a 10,000 MWCO dialysis tubing against 2 litres of Buffer B_0 for 3 h.

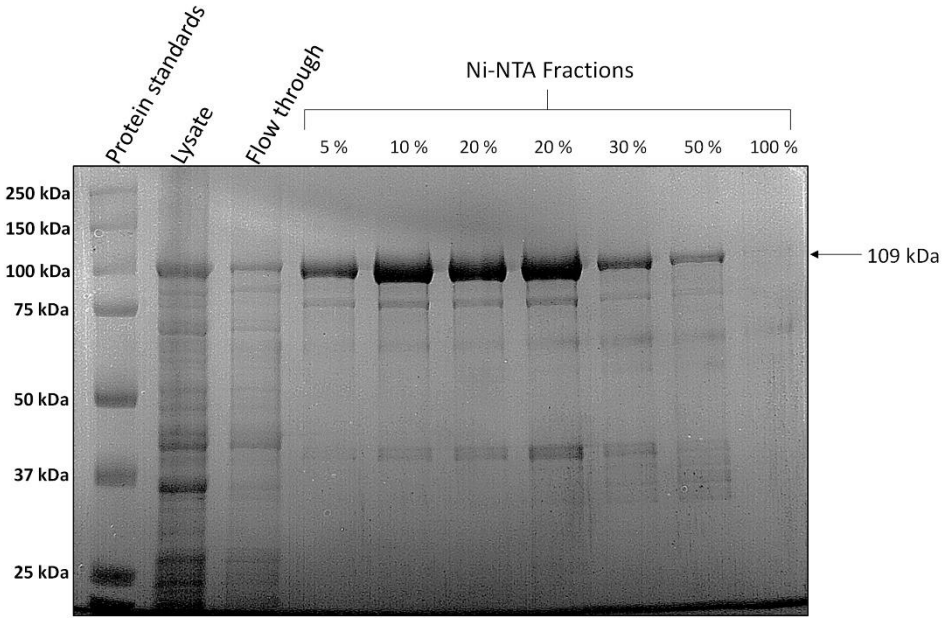


Figure 2: Representative 10% SDS-PAGE after Ni-NTA chromatography: 10 µl of sample (5 µl of the protein fraction mixed with 5 µl of 2× Laemmli sample buffer) is loaded on a 10% SDS-PAGE gel. The fractions with the most protein are pooled and dialysed. The SDS-PAGE gel is stained with Coomassie R-250.

F. Anion exchange chromatography (perform at 4°C using a Bio-Rad NGC™ chromatography system)

1. Equilibrate a GE MonoQ™ 10/100 GL column with 3 CV of Buffer B_{50}.
2. Using a 50 ml GE Superloop™ load the dialysed sample onto the column and wash the column with 10 CV of Buffer B_{50}.
3. Elute the sample in 1 ml fractions using a linear gradient (0%-50%) of Buffer B_{1000} across 20 CV. Monitor the Absorbance at 280 nM (UV_{280}) and conductivity measurements. The protein begins to elute when the conductivity reaches ~38 mS/cm (Figure 3A).
4. Based on the chromatograph, load alternate fractions of the protein peak on a 10% SDS-PAGE gel with an appropriate protein ladder (Bio-Rad Precision Plus Protein™ Standards) (Figure 3B). Once the presence of the protein is confirmed, pool the fractions.

G. Concentration of pooled fractions (perform at 4°C using a Vivaspin Turbo 10,000 MWCO centrifugal concentrator)

1. Equilibrate the centrifugal concentrator using Buffer B_{100}.
2. Add the pooled fractions into the concentrator and spin at a RCF of 3180 (~4,000-5,000 RPM) till the sample is concentrated to ~600 µl, which takes approximately 1 h.

H. Size exclusion chromatography (perform at 4°C using a Bio-Rad NGC™ chromatography system)

1. Equilibrate a GE Superdex 200™ 16/600 pg column with 1.5 CV of Buffer B_{100}.
2. Load the concentrated protein using a 1 ml loop onto the column.
3. Elute the sample in 1 ml fractions across 1 CV of Buffer B_{100}. This elution position represents a molecular weight of ~200 kDa which is in accordance with the dimeric state of SauUSI (~219 kDa) (Figure 4A).

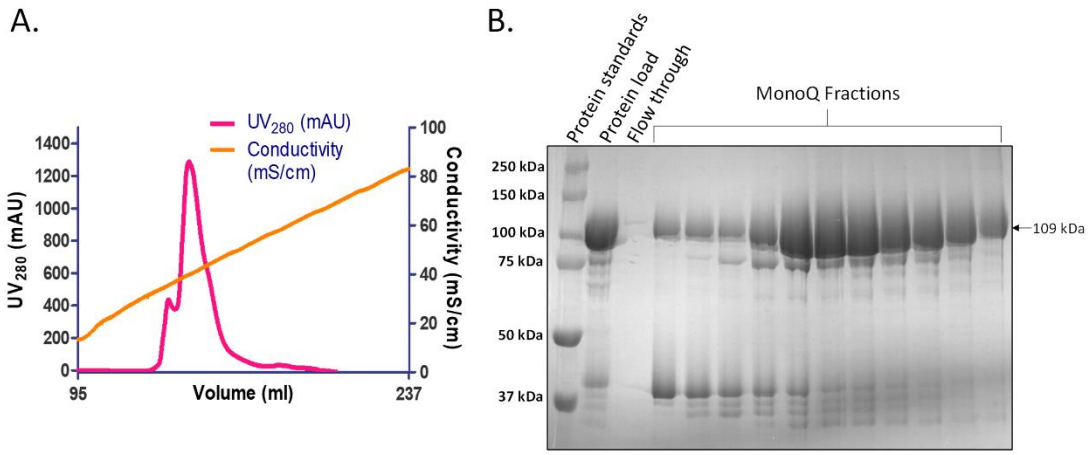


Figure 3: Representative chromatogram and 10% SDS-PAGE after anion exchange chromatography: A. Elution profile of SauUSI after anion exchange chromatography using MonoQ™ 10/100 GL column. The protein begins to elute as the conductivity approaches ~38 mS/cm. B. Representative 10% SDS-PAGE (stained with Coomassie R-250) image after anion exchange chromatography. 10 µl of sample (5 µl of the protein fraction mixed with 5 µl of 2× Laemmli sample buffer) is loaded on the gel. The fractions with the most protein are pooled and concentrated using Vivaspin Turbo 10000 MWCO centrifugal concentrator.

Load the relevant fractions on a 10% SDS-PAGE gel with an appropriate protein ladder (Bio-Rad Precision Plus Protein™ Standards) (Figure 4B). Once the presence of the protein is confirmed, pool the fractions (since the sample is denatured prior to loading onto a SDS-PAGE, the protein would run at the monomeric molecular weight of ~109 kDa).

I. Concentration of pooled fractions (perform at 4°C using a Vivaspin Turbo 10,000 MWCO centrifugal concentrator)

1. Equilibrate the centrifugal concentrator using Buffer B_{100} and add the pooled fractions into the concentrator and spin at a RCF of 3,180 3,180 RCF (~4,000-5,000 RPM).
2. Divide the concentrated protein (~10 mg/ml) into suitable aliquots (~5 µl), plunge freeze in liquid nitrogen and store at -80°C.

J. Running of 10% SDS-PAGE (Using Bio-Rad SDS-PAGE apparatus)

1. Add 5 µl of sample to 5 µl of 2× Laemmli sample buffer and heat at 99°C for 10 min.

2. Load the samples onto a 10% SDS-PAGE immersed in a running tank containing 1× TGS running buffer. Allow the gel to run at a constant voltage of 200 V for 45 min.
3. After the run, allow the gel to the stain in Coomassie (R-250) staining solution with gentle heating using a commercial microwave until the solution boils.
4. Destain the gel in Coomassie destaining solution with gentle heating using a commercial microwave until the solution boils.

Data analysis

The original research paper for the above purification protocol was published in Tumuluri *et al*. (2021), https://doi.org/10.1093/nar/gkab042.

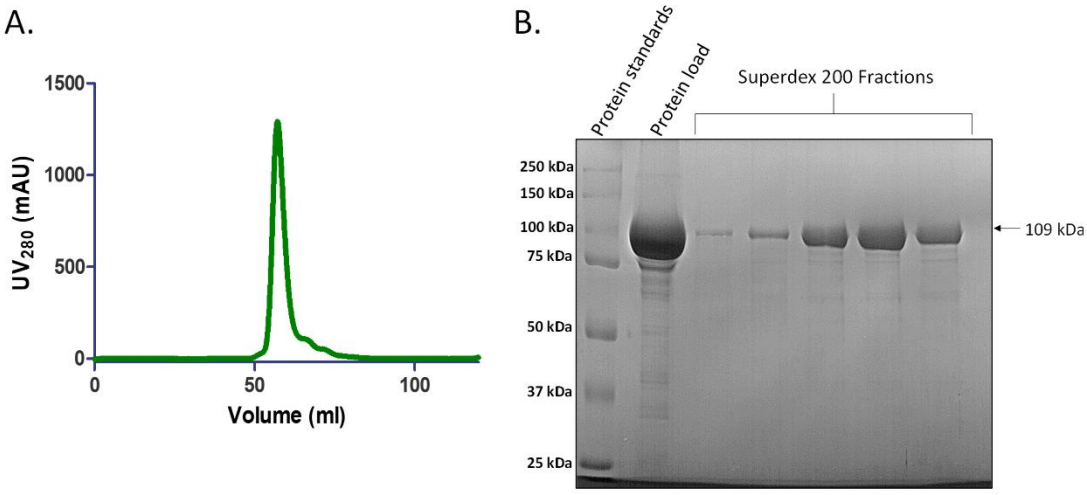


Figure 4: Representative chromatogram and 10% SDS-PAGE after size exclusion chromatography: A. Elution profile of SauUSI after anion exchange chromatography using Superdex 200™ 16/600 GL column. The protein begins to elute at ~0.45 CV. B. Representative 10% SDS-PAGE (stained with Coomassie R-250) image after size exclusion chromatography. 10 µl of sample (5 µl of the protein fraction mixed with 5 µl of 2× Laemmli sample buffer) is loaded on the gel. Size exclusion chromatography helps in separating the protein of interest from other contaminating proteins based on molecular weight eventually leading to ≥95% purity of protein. It helps in eluting the protein in an appropriate buffer for further experiments. The fractions with protein are pooled and concentrated using Vivaspin Turbo 10,000 MWCO centrifugal concentrator.

Recipes

A. Materials that need to be Autoclaved

1. 50 ml 1× Lysogeny broth

   Dissolve 1.25 g of LB broth using reverse osmosis (RO) water made up to 50 ml in a 200 ml conical flask. Close the mouth with a cotton plug.

2. 1,000 ml 1× Lysogeny broth

   Dissolve 25 g of LB broth using RO water made up to 1,000 ml in a 2,000 ml conical flask. Close the mouth with a cotton plug. The composition of HiMedia LB (per litre) is, 10 g NaCl, 10 g Peptone and 5 g Yeast extract.

B. Stocks solutions

1. 20% (w/v) Arabinose

   Dissolve 5 g of Arabinose in RO water made up to 25 ml. After dissolution, filter the solution into a sterile 50 ml centrifuge tube using a sterile 0.2 µm syringe filter.

2. 5 M NaCl

   Dissolve 146.1 g of NaCl in RO water made up to 500 ml.

3. 2 M Tris-HCl (pH = 8)

   Dissolve 121.14 g of Tris base in 300 ml of RO water. Adjust the pH using concentrated HCl. Make the volume up to 500 ml after pH adjustment.

4. 2 M Imidazole

   Dissolve 68.07 g of Imidazole in RO water and made up to 500 ml.

5. 4% (w/v) CHAPS

   Dissolve 1 g CHAPS in RO water made up to 25 ml.

6. 2 M $MgCl_2$

Dissolve 1.90 g of $MgCl_2$ in RO water made up to 10 ml.

7. 0.5 M EDTA (pH = 8)

   Dissolve 3.362 g of disodium EDTA salt in RO water (adjust the pH using NaOH pellets) made up to 20 ml and sterilize by autoclaving.

8. 1 M DTT

   Dissolve 3.09 g of DTT in 0.01 M sodium acetate (pH=5.2) made up to 20 ml. Filter the solution using a 0.2 µm syringe filter.

C. Buffers for purification

Note: Prepare the buffers as per the compositions given below and prechill at 4°C before use.

1. Lysis buffer (500 ml)

50 mM Tris-HCl (pH = 8) (12.5 ml, 2 M)

500 mM NaCl (50 ml, 5 M)

25 mM Imidazole (6.25 ml, 2 M)

5 mM $MgCl_2$ (1.25 ml, 2 M)

10% glycerol (50 ml, 100%)

0.2% CHAPS (25 ml, 4%)

Make the volume up to 500 ml in RO water and filter using a vacuum filter assembly attached to a 0.2 µm filter.

2. Buffers for affinity (Ni-NTA) chromatography (Buffer A and Buffer B) (500 ml each)

50 mM Tris-HCl (pH = 8) (Buffer A 12.5 ml, Buffer B 12.5 ml, 2 M)

500 mM NaCl (Buffer A 50 ml, Buffer B 50 ml, 5 M)

25 mM Imidazole (Buffer A 6.25 ml, 2 M)

500 mM Imidazole (Buffer B 125 ml, 2M)

Make the volume up to 500 ml in RO water and filter using a vacuum filter assembly attached to a 0.2 µm filter. Allow the buffers to degas for 20 min using the same assembly.

3. Buffer for Dialysis (Buffer B_0) (2 L)

50 mM Tris-HCl (pH = 8) (50 ml, 2 M)

1 mM EDTA (4 ml, 0.5 M)

1 mM DTT (2 ml, 1 M)

4. Buffers for anion exchange chromatography (Buffer B_{50} and Buffer B_{1000}) (500 ml each)

50 mM Tris-HCl (pH = 8) (Buffer B_{50} 12.5 ml, Buffer B_{1000} 12.5 ml, 2 M)

50 mM NaCl (Buffer B_{50} 5 ml, 5 M)

1000 mM NaCl (Buffer B_{1000} 100 ml, 5 M)

1 mM EDTA (Buffer B_{50} 1 ml, Buffer B_{1000} 1 ml, 0.5 M)

1 mM DTT (Buffer B_{50} 0.5 ml, Buffer B_{1000} 0.5 ml, 1 M)

Make the volume up to 500 ml in RO water and filter using a vacuum filter assembly attached to a 0.2 µm filter. Allow the buffers to degas for 20 min using the same assembly.

5. Buffer for size exclusion chromatography (Buffer B_{100}) (500 ml)

50 mM Tris-HCl (pH = 8) (12.5 ml, 2 M)

100 mM NaCl (10 ml, 5 M)

1 mM DTT (0.5 ml, 1 M)

Make the volume up to 500 ml in RO water and filter using a vacuum filter assembly attached to a 0.2 µm filter. Allow the buffers to degas for 20 min using the same assembly.

D. SDS-PAGE

1. 10X stock of TGS (Tris base, glycine and SDS) running buffer

   Dissolve 30 g of Tris base, 144 g of glycine and 10 g of SDS in RO water made up to 1,000 ml. Working concentration: 1X.

2. Coomassie staining solution

   Dissolve 2.5 g of Coomassie Brilliant Blue (R-250) in solution containing 500 ml Ethanol, 100 ml Glacial Acetic acid and 100 ml of RO water. After complete solvation, filter the solution through a Whatman filter paper (No. 1).

3. Coomassie destaining solution

   Add 200 ml of Ethanol, 100 ml of Glacial Acetic acid to 700 ml of RO water.

4. Solutions required for Gel mix

a. 30% Acrylamide/Bis-Acrylamide (AB) mix

   Dissolve 58 g of Acrylamide and 2 g of Bis-Acrylamide in RO water made up to 200 ml.

b. 1.5 M Tris-HCl (pH = 8.8)

   Dissolve 90.85 g of Tris base in 300 ml of RO water. Adjust the pH using concentrated HCl. Make the volume up to 500 ml after pH adjustment.

c. 1.5 M Tris-HCl (pH = 6.8)

   Dissolve 90.85 g of Tris base in 300 ml of RO water. Adjust the pH using concentrated HCl. Make the volume up to 500 ml after pH adjustment.

d. 10% SDS

   Dissolve 1 g of SDS in RO water made up to 10 ml.

e. 10% APS

   Dissolve 1 g of APS in RO water made up to 10 ml.

5. 10% resolving gel

   Water (RO) (4 ml)

   AB-mix (3.3 ml, 30%)

   Tris-HCl (pH = 8.8) (2.5 ml, 1.5 M)

   SDS (0.1 ml, 10%)

   APS (0.1 ml, 10%)

   TEMED (0.004 ml)

6. 5% stacking gel

   Water (RO) (1.4 ml)

   AB-mix (0.33 ml, 30%)

   Tris-HCl (pH = 6.8) (0.25 ml, 1.5 M)

   SDS (0.02 ml, 10%)

   APS (0.02 ml, 10%)

   TEMED (0.002 ml)

References

1. Loenen,W.A.M., Dryden,D.T.F., Raleigh,E.A., Wilson,G.G. and Murrayy,N.E. (2014) Highlights of the DNA cutters: A short history of the restriction enzymes. *Nucleic Acids Res.*, **42**, 3–19.

2. Xu,S.Y., Corvaglia,A.R., Chan,S.H., Zheng,Y. and Linder,P. (2011) A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300. *Nucleic Acids Res.*, **39**, 5597–5610.

3. Tumuluri,V.S., Rajgor,V., Xu,S.Y., Chouhan,O.P. and Saikrishnan,K. (2021) Mechanism of DNA cleavage by the endonuclease SauUSI: A major barrier to horizontal gene transfer and antibiotic resistance in Staphylococcus aureus. *Nucleic Acids Res.*, **49**, 2161–2178.

4. Monk,I.R., Shah,I.M. and Xu,M. (2012) Transforming the Untransformable : Application of Direct Transformation To manipulate Genetically Staphylococcus and Staphylococcus epidermidis. *MBio*, **3**, e00277-11.

5. Corvaglia,A.R., François,P., Hernandez,D., Perron,K., Linder,P. and Schrenzel,J. (2010) A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical Staphylococcus aureus strains. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 11954–8.

6. Kunzelmann,S. and Webb,M.R. (2009) A biosensor for fluorescent determination of ADP with high time resolution. *J. Biol. Chem.*, **284**, 33130–33138.

CHAPTER 3: Mechanism of DNA cleavage by the endonuclease SauUSI – a major barrier to horizontal gene transfer and antibiotic resistance in *Staphylococcus aureus*

Adapted from: Tumuluri et al., *Nucleic Acids Research*, 2022.

DOI: 10.1093/nar/gkab042

Abstract

Acquisition of foreign DNA by *Staphylococcus aureus,* including vancomycin resistance gene, is thwarted by the ATP-dependent endonuclease SauUSI. Deciphering the mechanism of action of SauUSI could unravel the reason how it singularly plays a major role in preventing horizontal gene transfer (HGT) in *S. aureus*. Here, we report a detailed biochemical and structural characterization of SauUSI, which reveals that in the presence of ATP, the enzyme can cleave DNA having a single or multiple target site/s. Remarkably, in the case of multiple target sites, the entire region of DNA flanked by two target sites is shred into smaller fragments by SauUSI. Crystal structure of SauUSI reveals a stable dimer held together by the nuclease domains, which are spatially arranged to hydrolyze the phosphodiester bonds of both strands of the duplex. Thus, the architecture of the dimeric SauUSI facilitates cleavage of either single-site or multi-site DNA. The structure also provides insights into the molecular basis of target recognition by SauUSI. We show that target recognition activates ATP hydrolysis by the helicase-like ATPase domain, which powers active directional movement (translocation) of SauUSI along the DNA. We propose that a pile-up of multiple translocating SauUSI molecules against a stationary SauUSI bound to a target site catalyzes random double-stranded breaks causing shredding of the DNA between two target sites. The extensive and irreparable damage of the foreign DNA by shredding makes SauUSI a potent barrier against HGT.

Introduction

Antimicrobial resistance (AMR) is an outstanding threat to the health care system worldwide (1). One serious concern is the resistance in *Staphylococcus aureus* – an opportunistic pathogen responsible for the majority of the nosocomial and community-acquired infections in humans. Presently, the first-line of treatment involves antibiotics penicillin and methicillin. However, over the years, many strains of *S. aureus* have grown resistant to them (2, 3, 4). Vancomycin is a glycopeptide antibiotic that is known to treat methicillin-resistant *S. aureus* (MRSA). The ability of MRSA to acquire vancomycin-resistance through horizontal gene transfer (HGT) is a looming health crisis (5, 6). HGT is the process of acquisition of foreign DNA by bacteria through transduction, transformation or conjugation, and is a biological driver of bacterial evolution, ecology and pathogenicity (7, 8). Restriction endonucleases (REases) that nucleolytically degrade foreign DNA are one of the main barriers to HGT in bacteria (9, 10). A striking example is an almost impregnable barrier created by the REase SauUSI to the entry of foreign DNA into *S. aureus* (11, 12, 13).

SauUSI was first discovered in the MRSA strains UAMS-1 and SA564, and was found to be one of the two main barriers to HGT in many strains of *S. aureus*; the REase Sau1 being the other (11, 12, 14, 15). Based on the epigenetic status of their substrate DNA, REases are broadly classified as modification-independent Type I, ISP, II or III REases or modification-dependent Type IV REases (16). Modification-independent REases cut DNA having non-modified target sites, while modification-dependent REases cleave DNA only if the target sites have a specific modified base. SauUSI is a modification-dependent Type IV REase, and Sau1 is a modification-independent Type I REase. The dynamics of the REases, i.e., the expression levels of REases, the types of the REases acting in tandem and their nucleolytic potential in the bacterial cell, dictate the rate of HGT, and thus potentially the ability of a bacterium to gain resistance to an antibiotic. Inactivation of the *sauUSI* gene leads to a manifold increase in the transformation efficiency of *S. aureus* (11, 17, 18). Interestingly, deletion of the Sau1 in these strains does not have such a profound effect (11, 13, 19). Furthermore, the absence of SauUSI results in hyper-susceptibility to the gain of the

vancomycin-resistance gene from *Enterococcus faecalis* (11, 13). What makes SauUSI such a potent barrier to HGT is not clear.

SauUSI is made of a single polypeptide chain containing three domains - a nuclease domain belonging to the phospholipase D (PLD) family, an ATPase domain belonging to the Superfamily 2 (SF2) of helicase, and a target recognition domain (TRD) (11, 12). The Type IV REase SauUSI cleaves DNA only in the presence of the target sequence 5'-S^{5m}CNGS-3', where S=G or C and ^{5m}C can be 5-methylcytosine or 5-hydroxymethylcytosine (12). It was previously reported that SauUSI cleaves only those DNA having two or more target sites. Based on the site of their cleavage, REases can be categorized as site-specific or random cutters. Site-specific REases, such as Type II and Type III REases, cleave inside or at a fixed distance outside the target sequence. In contrast, the random cutters, such as Type I and ISP REases, cut randomly as far away as thousands of base pairs (bp) from their target sequence. Preliminary analysis showed that SauUSI cleaves DNA 3 to 18 bp outside of the target sequence (12).

SauUSI requires ATP as a cofactor for its nucleolytic activity (12). REases that require ATP as a cofactor for DNA cleavage usually belong to Type I, ISP or III (20, 21). In the case of Type I and ISP REases, ATP hydrolysis powers translocation of the enzyme along the DNA that allows it to cleave DNA away from the target sequence, while the hydrolysis of ATP by Type III REases primarily serves as a switch to turn on its nucleolytic activity (20, 22, 23, 24). The role of ATP hydrolysis in DNA cleavage by SauUSI is unknown. Here, through complementary biochemical and structural studies, we unravel the mechanism of DNA cleavage by SauUSI and the role of ATP hydrolysis for its nucleolytic activity and gain insights into what makes this enzyme a potent barrier to HGT.

Materials and Methods

Cloning and expression of SauUSI

The *sauUSIR* gene followed by a BamHI site and C-terminal 6xHis-Tag downstream and flanked by a NdeI site upstream was inserted into a pHis17 (Ampicillin resistant) vector by a PCR based restriction-free cloning strategy. NEB turbo electrocompetent cells were transformed with the plasmid DNA. The positive clone was confirmed by sequencing for the presence of the wild type

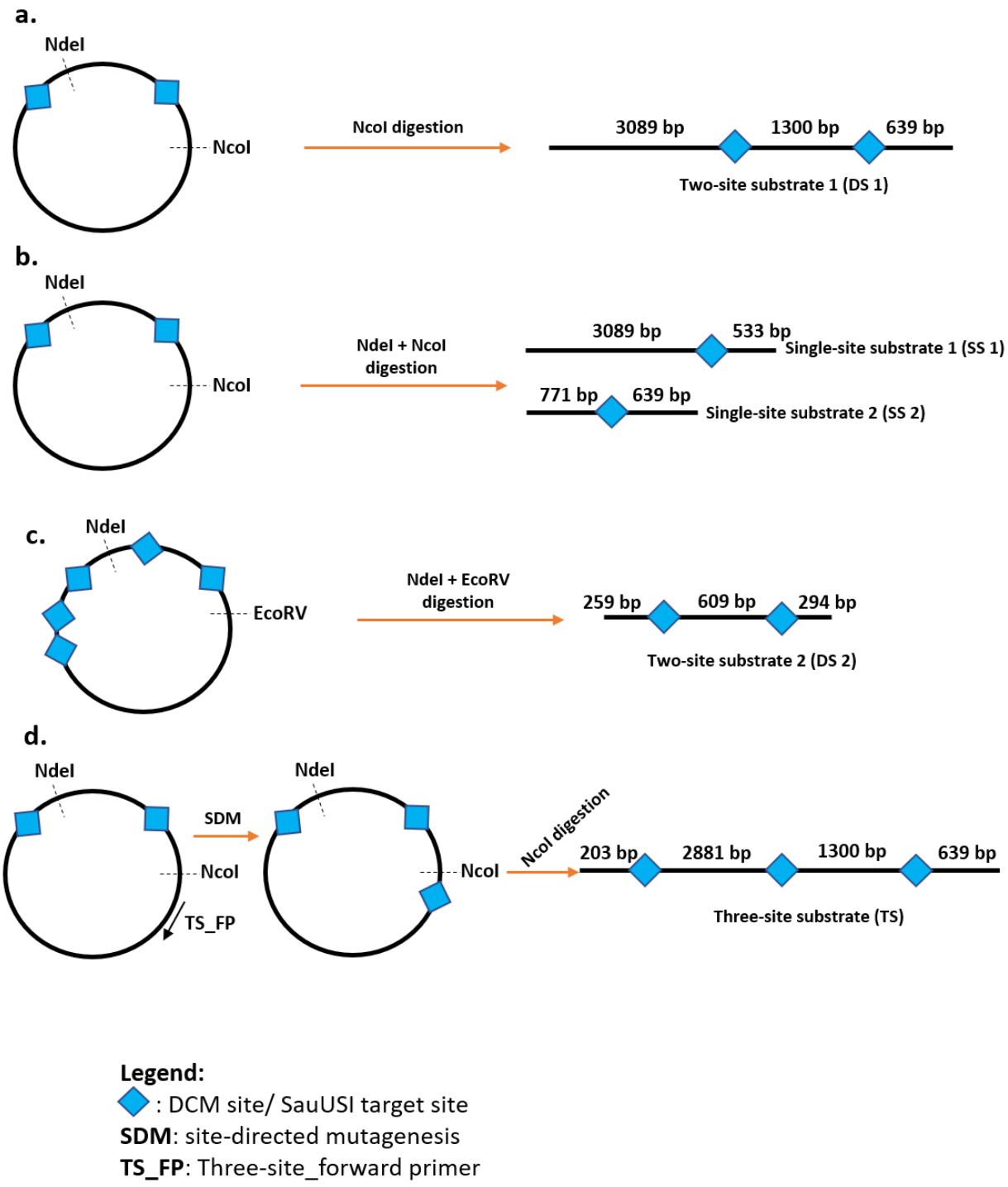


Figure 1: Schematic of the substrates used: (A) Schematic of the two-site plasmid (DS 1). The plasmid was grown in a *dcm*$^+$ strain of *E. coli* and was purified and linearized with NcoI-HF. (B) Schematic of the single-site substrates SS 1 and SS 2. The two-site plasmid was digested with NcoI-HF and NdeI, and the two fragments were extracted from an agarose gel and purified. (C) DS2 was generated by digesting a plasmid with five sites (grown in a *dcm*$^+$ strain of *E. coli*) with NdeI and EcoRV followed by an agarose gel purification (D) The three-site plasmid (TS) was generated using site-directed mutagenesis of the two-site plasmid. The three-site plasmid was processed similar to the two-site plasmid.

sauUSIR gene. Small scale expression of SauUSI WT suggested that it was best expressed in BL21(AI) cells grown in LB (Lysogeny Broth) supplemented with ampicillin (0.1 mg/ml) up to 0.6 OD at 37^{o}C, induced with 0.2% arabinose and left to grow overnight at 16^{o}C.

Purification of SauUSI for biochemical assays

The purification of SauUSI was performed using a three-column strategy at 4^{o}C. 2 L culture was pelleted down using an Avanti J26X-SP at an RCF of 5180 for 20 mins. The pellet was resuspended in a lysis buffer (50 mM Tris-HCl, pH 8, 500 mM NaCl, 5 mM $MgCl_2$, 25 mM Imidazole, 10% Glycerol and 0.2% CHAPS) and lysed in a sonicator for 5 mins using a pulse of 1 sec and recovery of 3 secs at a 60% amplitude. Clarification of the cell lysate was achieved by ultra-centrifuging the sample at 100,000 g using an Optima-L-100K ultracentrifuge. The supernatant was loaded onto a GE HisTrap^{TM} HP 5 ml column and washed thoroughly with Buffer A (50 mM Tris-HCl, pH 8, 500 mM NaCl and 25 mM Imidazole). Elution of the protein was achieved by using a step gradient of Buffer B (50 mM Tris-HCl, pH 8, 500 mM NaCl and 500 mM Imidazole). Presence of protein was confirmed using a 10% SDS-PAGE gel, and the appropriate fractions were pooled for further processing. The pooled fractions were dialyzed against Buffer B0 (50 mM Tris-HCl, pH 8, 1 mM EDTA and 1 mM DTT) using Snakeskin dialysis tubing 10,000 MWCO for 3 hours. The dialyzed sample was then loaded onto a GE MonoQ^{TM} 10/100 GL column using Buffer B50 (50 mM Tris-HCl, pH 8, 50 mM NaCl, 1 mM EDTA and 1 mM DTT). Elution of the protein was achieved using a linear gradient (0%-50%) of Buffer B1000 (50 mM Tris-HCl, pH 8, 1000 mM NaCl, 1 mM EDTA and 1 mM DTT). The fractions containing protein (confirmed by a 10% SDS-PAGE gel) were pooled and concentrated using a Vivaspin Turbo 10K MWCO centricon. The concentrate was loaded onto a GE Superdex 200^{TM} increase 10/300 GL column and protein was eluted using Buffer B100 (50 mM Tris-HCl, pH 8, 100 mM NaCl and 1 mM DTT). Pure protein was analyzed on a 10% SDS-PAGE gel, concentrated appropriately and plunge frozen using liquid nitrogen and stored at -80^{o}C for biochemical studies.

DNA substrates used for cleavage assay

The substrates used in the cleavage assays were plasmids and methylation was achieved *in vivo* using *dcm*$^{+}$ DH5α *E. coli* and *in vitro* using commercial M.MspI (NEB). Dcm recognizes the DNA

sequence 5'-CCAGG-3' and converts it to 5'-C^{5m}CAGG-3' whereas M.MspI recognizes the DNA sequence 5'-CCGG-3' and converts it to 5'-^{5m}CCGG-3'. The sequences methylated by Dcm or M.MspI are subsets of the SauUSI target site 5'-S^{5m}CNGS-3'. Methylation and non-methylation were achieved by transforming plasmids into DH5α *E. coli* and NEB Turbo respectively. Plates were kept overnight at 37^{o}C, and single colonies were picked and inoculated in LB (0.1 mg/ml of ampicillin) and grown until turbidity. The cells were pelleted down, and plasmid preparation was performed using QIAprep® miniprep spin protocol. The plasmid represented in figure 2a is a derivative of *pUC18*. The two-site substrate 1 (DS 1) and the three-site substrate (TS) are derivatives of the *pHis17* vector which were linearized with NcoI-HF at 37^{o}C for 1 hour according to the NEB protocol. Then NcoI-HF was heat-inactivated by incubating the reaction at 65^{o}C for 20 mins (Figure 1). The single-site substrates 1 and 2 (SS 1 and SS 2) were obtained by double digesting the two-site substrate with NdeI and NcoI-HF followed by a 1% agarose gel extraction using the QIAquick® extraction protocol (Figure 1). The two-site substrate 2 (DS 2) was obtained by digesting a plasmid, also a derivative of the *pHis17* vector, having five SauUSI sites with NdeI and EcoRV, followed by a 1% agarose gel extraction (Figure 1). The 120 bp methylated, hemimethylated and non-methylated substrates were generated by an extension of two partially complementary 70 bp DNA oligos using a PCR machine.

DNA cleavage assays

DNA cleavage assays were performed to test the nuclease activity of SauUSI. Initial cleavage assays were done using 150 ng of a circular plasmid. The substrate (plasmid) was incubated with 150 nM SauUSI WT and 2 mM ATP in the NEBuffer™ 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, 1 mM DTT pH 7.9) for 1 hour at 37^{o}C. Post incubation, the reaction was stopped by adding 1X Gel Loading Dye Purple and heating the reaction to 65^{o}C for 20 mins. The cleavage fragments were analyzed on a 1% Agarose gel. Cleavage assays involving DS 1, SS 1 and TS were carried out with ~3 nM DNA, 2 mM ATP and variable concentrations of the protein as mentioned in the respective figures. All other parameters were as mentioned above. Time-dependent assays of DS 1 and SS 1 were carried out across a time range.

Cleavage assays involving TS 2 and SS 2 were carried out with ~10 nM bDNA and 2 mM ATP across a concentration range of protein, which is mentioned in the respective figures. Cleavage products were resolved using a 2% agarose gel. 38 nM of the 120 bp methylated, hemimethylated and non-methylated substrates were used for cleavage by 100 nM SauUSI and 2 mM ATP and resolved on an 8% native PAGE. The position of the circular plasmid and the cleaved fragments were determined using LabImage® (Kapelan Bio-imaging). Quantification of the intensities of the fragments from the DS 1, SS 1 and TS upon digestion was performed relative to the intensity of the corresponding fragments from a BstNI reaction. The BstNI reaction was performed to correct for the staining effect due to the variable length of the DNA fragments. The absolute intensities of the fragments obtained upon cleavage of the respective plasmids by BstNI were compared with the absolute intensity of the corresponding substrate (linearized plasmid) band. The ratio of the individual intensities so obtained was treated as the maximum intensity possible for a particular fragment. Next, a similar analysis was performed for DNA digested by SauUSI. The ratio of the relative intensity of a fragment obtained by SauUSI digestion with that of the relative intensity of the corresponding fragment obtained by BstNI digestion was calculated and plotted. The error bars represent the standard error of the mean across three independent assays. The error bars for the cleavage assay of SS 1 with 400 nM DNA represent the standard error of the mean across two independent assays. Cleaved products were analyzed using ImageJ, and their relative intensities were plotted using GraphPad Prism.

Run-off sequencing

A 200 µl mixture containing 4 µg of a single-site substrate, 400 nM SauUSI and 2 mM ATP was incubated at 37^{o}C till the reaction reached completion (checked on a 1% agarose gel). On completion of the reaction, the DNA fragments were purified using the QIAquick® protocol. The purified fragments were sequenced (Sanger sequencing) using forward and reverse primers.

SEC-MALS

1 mg/ml of the SauUSI was injected into a GE Superdex 200^{TM} column equilibrated with 50 mM Tris-HCl, pH 8, 50 mM NaCl and 1 mM DTT. The apparatus was connected to a light scattering diode array and a differential refractive index detector (Wyatt Technology). The data was

analyzed using the ASTRA software 6.1.7.17. The molar mass was calculated from the light scatter and differential refractive index. Before use, the column was calibrated using bovine serum albumin.

Crystallization

To obtain phase information, we purified the Se-Met derivative of SauUSI (SauUSI^{SelMet}). SauUSI^{SelMet} was crystallized by the hanging-drop vapor diffusion method using a 1:1 protein: reservoir buffer at 291K. The crystals were set in a 24-well plate with the reservoir buffer containing 0.1 M sodium citrate, 11%-14% PEG 4000 and 0.25 M-0.4 M ammonium sulphate. Ethylene glycol was used as a cryoprotectant.

X-ray data collection and processing

Crystals were first screened for diffraction quality using an in-house X-ray diffraction system of Rigaku MicroMax 007 X-ray generator with a Mar Research 345D detector. The crystals that diffracted best at home source were retrieved and taken for diffraction studies at synchrotron facilities at Diamond Light Source (DLS) Oxfordshire, UK, and European Synchrotron Research Facility (ESRF), Grenoble. The diffracted data was indexed and processed using XDS (25). The data sets were scaled and merged using AIMLESS (26).

SAD phasing, structure solution and refinement

The structure of SauUSI^{SelMet} was solved using experimentally determined phases to 3.1 Å by Single-wavelength Anomalous Dispersion (SAD) method (Table 2). The selenium positions were identified using the program SOLVE-RESOLVE (27, 28) in the Phenix suite of crystallography programs (29). In all, 37 seleniums were located in the asymmetric unit. A simple Fourier map calculated using the SAD phases was used to build the structure of SauUSI manually. Initial electron density maps were visualized, and further model building was carried out using COOT (30). Structure refinement was carried out using *phenix.refine* (31).

Table 1: Primers used in the study.

Primer name	Sequence of primers (5'->3')
SauUSI_forward primer	GAAGGAGATATACATATGAGTAGATTACTAAATGATTTCAATC
SauUSI_reverse primer	GATGATGATGATGATGGGATCCATTTGTTAGATAACGATATATATC
SauUSI^{H119A} reverse primer	CTCAAAAATATATCCTTTGGCGGCGAATCCAGCAATAT
Triplex-site reverse primer	CACCTTCATCTCTGACATTTCTTTCTTCTTTCTTTGCTCGAGACCAGCTAAAATCCTGAAAGA
120 bp methylated forward primer	ATGAGACATGGTTTCTTGAGTGTATATGAACTATTCAAATCAGTGTCTTCGGACTACGC^{5m}CAGGAGAATTT
120 bp methylated reverse primer	ACGTGGTCAATAGCACAACAATGTCGTCACAAGCTCTTCACATCTATGCTAAATTCTC^{5m}CTGGCGTAGTCC
120 bp non-methylated forward primer	ATGAGACATGGTTTCTTGAGTGTATATGAACTATTCAAATCAGTGTCTTCGGACTACGCCAGGAGAATTT
120 bp non-methylated reverse primer	ACGTGGTCAATAGCACAACAATGTCGTCACAAGCTCTTCACATCTATGCTAAATTCTCCTGGCGTAGTCC
Three-site forward primer	GATGATACTTCCACTGGATCCAGGTTTGAAGTAACCAGAAAT
Sequencing forward primer	CACCACTTCAAGAACTCTGTAGCA
Sequencing reverse primer	CAAAATTATTTCTAGAGGGAAACCGT
EMSA methylated forward primer	GCCGCGGATCCTTCACAAGACGATTAC/5MedC/GGCGCTGAAAGACCAGCATATGCAATTTGACT
EMSA methylated reverse primer	AGTCAAATTGCATATGCTGGTCTTTCAGCG/5MedC/CGGTAATCGTCTTGTGAAGGATCCGCGGC
EMSA non-methylated forward primer	GCCGCGGATCCTTCACAAGACGATTACCGGCGCTGAAAGACCAGCATATGCAATTTGACT
EMSA non-methylated reverse primer	AGTCAAATTGCATATGCTGGTCTTTCAGCGCCGGTAATCGTCTTGTGAAGGATCCGCGGC

DNA binding studies

The substrates used for the DNA binding study were generated using an annealing PCR (Table 1). Different concentrations of SauUSI were allowed to incubate with the DNA on ice for 20 min. EMSA was carried out on a 5% Native PAGE gel and was stained with ethidium bromide.

Site-directed mutagenesis (Nuclease inactive mutant)

The nuclease inactive mutant, i.e., SauUSI^{H119A}, was developed using the quick-change site-directed mutagenesis strategy. The mutation was confirmed by sequencing the plasmid. Purification of SauUSI^{H119A} was performed using the protocol described above using *E. coli* BL21 (DE3) cells.

ATPase assay

The malachite green method was used to determine the ATPase stimulation in a DNA dependent and independent manner. 5 nM SauUSI^{H119A} was allowed to incubate with 2 mM ATP, 2 mM ATP & 10 nM non-methylated DNA and 2 mM ATP & 10 nM methylated DNA separately in a time-dependent manner. The inorganic phosphate so released upon ATP hydrolysis formed a complex with molybdate and could be estimated colorimetrically at 630 nm. The results were compared to an appropriate phosphate standard curve to obtain inorganic phosphate release v/s time in minutes was fit using the one-phase association function in GraphPad Prism. The equation used was $Y=Y_{max1}(1-e^{-K.X})$. The error bars represent the standard error of the mean across three independent experiments.

Triplex displacement assay

The substrate of the triplex displacement assay had two components: A Triplex Forming Oligo (TFO) 100 nM of which was labelled at the 5' end using T4 polynucleotide kinase in the presence of ^{32}P-$^{\gamma}$ATP at 37^{o}C for 30 mins. Post the incubation, the reaction was stopped by heat inactivating T4 PNK at 65^{o}C for 20 mins. MicroSpin columns (GE) was used for purifying the TFO and stored at -30^{o}C for future use. The second component was a plasmid DNA consisting of an enzyme binding cassette and a triplex binding site (TBS) 604 bp downstream of the enzyme binding cassette. Methylation and linearization of the plasmid were done with a similar protocol to the

plasmids used for cleavage assays. To form the triplex, 50 nM of the linearized plasmid was incubated along with 25 nM TFO, 10 mM MES (pH 5.5) and 200 mM $MgCl_2$ at 57^{o}C for 15 minutes. Subsequently, the reaction was allowed to cool to 20^{o}C and kept overnight. The triplex displacement reaction was carried out at 20^{o}C at different time points by incubating the SauUSI^{H119A} with the triplex DNA and protein for 2 mins following which cold TFO and ATP were added. The curve (% of triplex displaced v/s time in minutes) was fit using the one-phase association function in GraphPad Prism. The equation used was $Y=Y_{max1}(1-e^{-K.X})$. The error bars represent the standard error of the mean across three independent experiments.

Phenogram generation

The SauUSI primary sequence was used as a query for a pBLAST against REBASE and a tnBLAST in NCBI. Sequences above the 30% sequence identity were collated in JalView and aligned using Clustal Omega. Redundant sequences were removed by keeping the redundancy threshold at 90%. A primary tree was generated in JalView using the neighbor joining method. The tree was viewed using the iTOL software and rooted to SauUSI (32).

Results

Cleavage products of SauUSI differ from those of site-specific cutters

On analyzing plasmid DNA treated with SauUSI and ATP, we observed that it appeared as a smear on an agarose gel (Figure 1a). Furthermore, the smearing was more pronounced as the number of target sites increased. For example, a plasmid DNA purified from *E. coli* (*dcm*$^{+}$), which has five 5'-C^{5m}CWGG-3' sites, on treatment with SauUSI and ATP resulted in a smear with the peak intensity at ~2300 bp (Figure 2a, Figure 3a). The theoretical length of the largest fragment expected for SauUSI as a site-specific cutter is 2068 bp (Figure 3b). The site-specific ATP-independent Type II REase BstNI, which recognized the same but non-methylated target site, i.e. 5'-CCWGG-3', cut the unmodified plasmid into discreet fragments with the longest being ~2200 bp (Figure 2a, Figure 3c, lane 5).

This plasmid when purified from *E. coli* (*dcm*$^{-}$) and treated with MspI methyltransferase (M.MspI) resulted in thirteen modified 5'-^{5m}CCGG-3' sites, ten of which were target sites of SauUSI (Figure

3d). Treatment of the modified plasmid with SauUSI and ATP resulted in a smear with the peak at ~530 bp (Figure 3d). On the other hand, the expected length of the largest fragment was 657 bp if SauUSI was a single-site cutter (Figure 3b). The unmodified plasmid, on incubation with the site-specific Type II REase MspI (R.MspI), which cuts at the target site 5'-CCGG-3', resulted in discreet fragments with the longest being ~510 bp (Figure 3c, lane 3, 3d). Next, incubating the plasmid purified from *E. coli* (*dcm*$^+$) with M.MspI resulted in a 15-site SauUSI substrate, which upon cleavage resulted in a smear with the peak at ~350 bp (Figure 3d). In contrast, double digestion of the unmodified plasmid with the REases R.MspI (cut at thirteen 5'-^{5m}CCGG-3' sites) and BstNI (cut at five 5'-CCWGG-3' sites) resulted in discreet fragments with the longest being ~470 bp (Figure 3c, lane 7, 3d). The expected size of the largest fragment if SauUSI was a site-specific cutter was 495 bp (Figure 3b). Thus, the above analysis demonstrated that the cleavage product of SauUSI was different from a site-specific cutter, such as the Type II REases, which suggested that the enzyme had a distinct mechanism of DNA cleavage.

SauUSI can nucleolytically cleave single-site DNA

To gain a better understanding of the nucleolytic activity of SauUSI, we focused our attention on a linear DNA substrate (DS 1) containing two target sites (5'-C^{5m}CTGG-3') separated by 1300 bp, which was generated by linearizing a 5034 bp plasmid DNA purified from *E. coli* (*dcm*$^+$) (Figure 2b). It was previously proposed that SauUSI cleaves only that DNA that have at least two target sites and that the cleavage is close to one of the two target sites (12). Hence, we expected the SauUSI treated two-site substrate to yield fragments of length around 4394 bp and 1944 bp. However, a nuclease assay of the two-site substrate with enzyme concentration varying from 10 nM to 400 nM revealed three distinct bands corresponding to ~3089 bp, ~ 1300 bp and ~ 639 bp resulting from double-strand (ds) DNA breaks at or near both the target sites (Figure 2b). Additionally, less intense bands viz., 4394 bp and 1944 bp corresponding to DNA cleavage only at one of the two target sites were visible at lower concentrations of the enzyme (Figure 2b).

The disappearance of the 4394 bp and 1944 bp DNA with increasing concentration of SauUSI suggested that the cleavage of the two-site substrate happened in a step-wise manner, with the

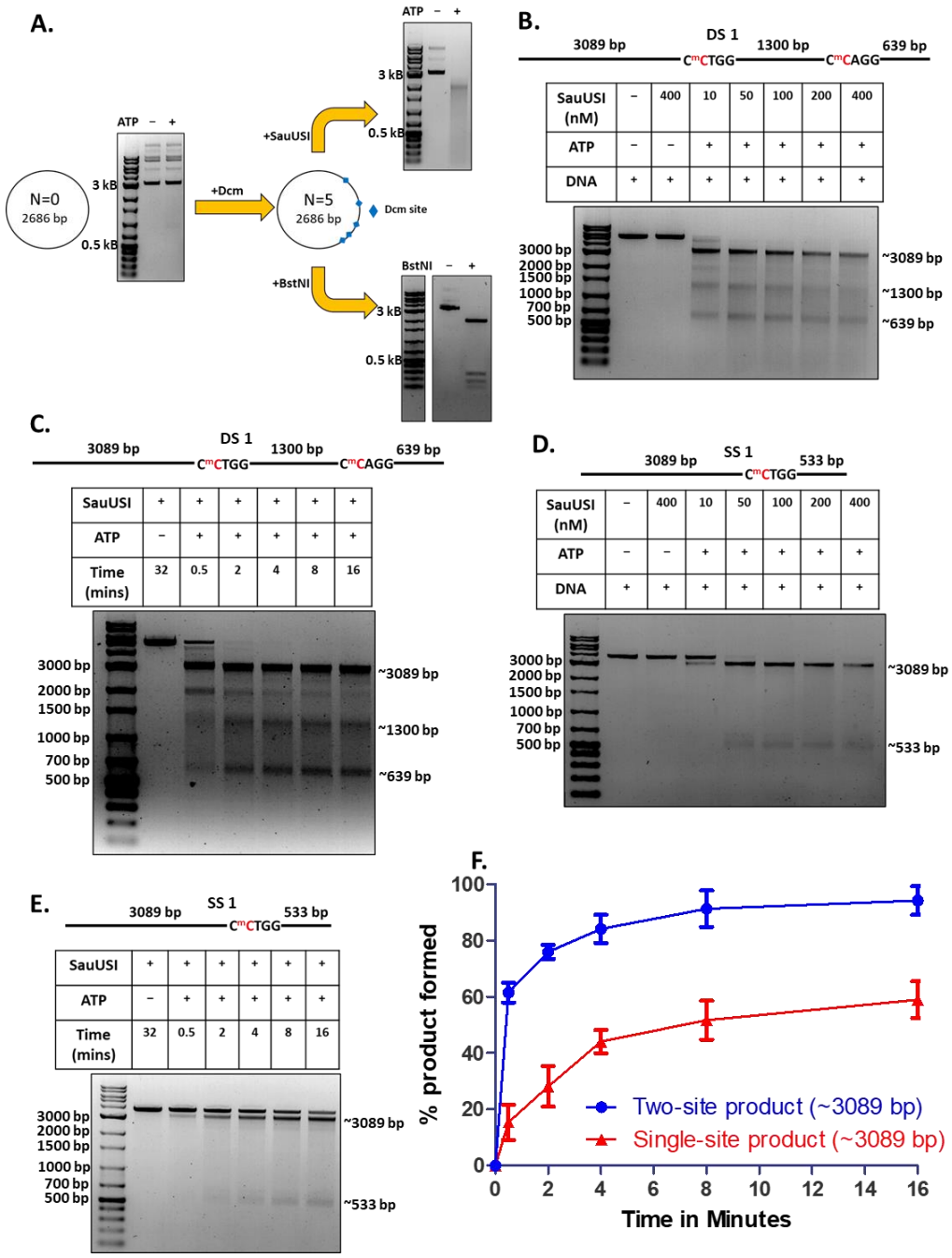


Figure 2. Nuclease activity of SauUSI: (A) Schematic of the plasmid used for cleavage. N represents the number of target sites of SauUSI in a particular plasmid. Methylation was brought about *in vivo* to generate the target sites of SauUSI. The corresponding cleavage pattern (on a 1% agarose gel) of SauUSI on the respective plasmids. The reactions with and without nucleotide (ATP) are compared against a DNA marker. (B) The two-site substrate (DS 1) used for the cleavage assay; the separation between the two sites is 1300 bp. The cytosine highlighted in red represents methylation. The target site is methylated on both the strands. Representative 1% agarose gel for two-site cleavage using varying concentrations (10 nM - 400 nM) of SauUSI. (C) Single-site substrate (SS 1) used for the cleavage assay. The cytosine highlighted in red represents methylation. Methylation is present on both the strands. The target site is methylated on both the strands. Representative 1% agarose gel for single-site cleavage using varying concentrations (10 nM - 400 nM) of SauUSI. (D) Representative 1% agarose gel for two-site cleavage (DS 1) carried over a time period of 30 seconds to 16 minutes. (E) Representative 1% agarose gel for single-site (SS 1) cleavage carried over a time period of 30 seconds to 16 minutes. (F) Percentage of product (3089 bp) formed as a function of time (DNA concentration 3 nM and protein concentration 100 nM) (n=3). In case of DS 1 the 3089 bp fragment is formed by either a single cut at the nearest target site or from sequential two-site cleavage.

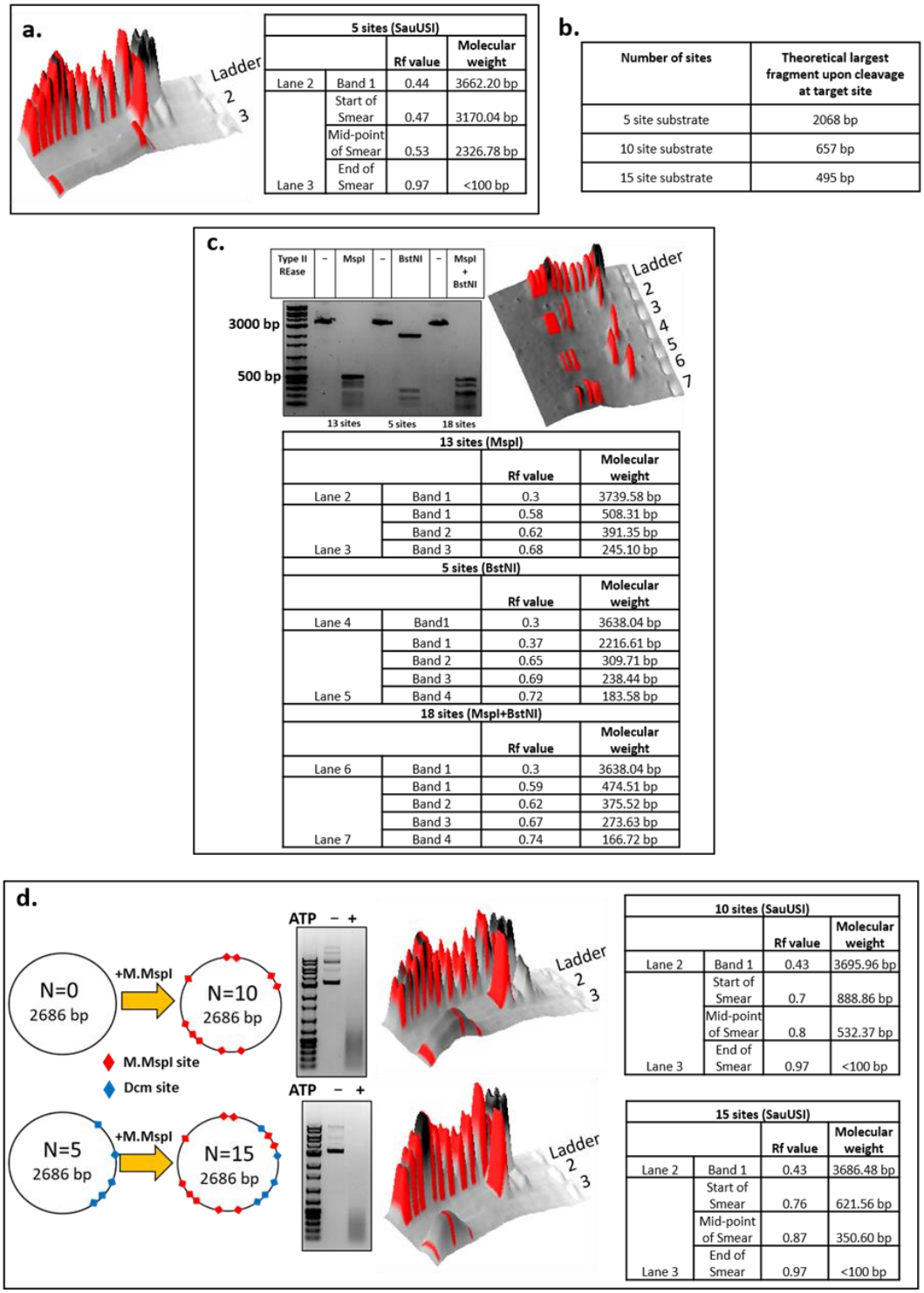


Figure 3: Analysis of DNA cleavage: (A) 3-D representation of cleavage of SauUSI on a plasmid with five sites (derived from figure 2a). The red color represents the region of the peak quantified. The adjoining table provides details of the Rf value and the calculated molecular weight. (B) A table containing theoretically largest fragments formed if plasmids were cleaved at the SauUSI target sites. (C) 1% Agarose gel of the cleavage of plasmid having multiple sites (part of the gel also shown in figure 1a) by MspI, BstNI and a combination of the two. 3-D representation of the cleavage pattern. The table of Rf values and calculated molecular weights is also given. (D) Schematic of the plasmid used for cleavage. N represents the number of target sites of SauUSI in a particular plasmid. The names on the arrows are the methyltransferases used (either *in vitro* or *in vivo*) to generate the methylated target sites of SauUSI. The corresponding cleavage pattern (on a 1% agarose gel) of SauUSI on the respective plasmids. The reactions with and without nucleotide (ATP) are compared against a DNA marker. 3-D representation of cleavage of SauUSI on a plasmid with 10 sites and 15 sites respectively. The red color represents the region of the peak quantified. Since SauUSI cleavage produces smears, therefore the start and mid-points of the peaks are quantified. The adjoining tables provide details of the Rf value and the respective calculated molecular weights.

first cleavage event happening at one of the two target sites followed by cleavage at the second site. This was corroborated by a time-dependent cleavage assay carried out at an enzyme concentration of 100 nM, which showed diminished intensities of 4394 bp and 1944 bp DNA with increasing incubation time (Figure 2c). This result led us to investigate if SauUSI could cut a single-site substrate. For this, a 3627 bp linear DNA (SS 1) with a single SauUSI target site was generated (see Methods). The target site was flanked by 3089 bp and 533 bp DNA (Figure 2d). On performing a SauUSI concentration-dependent DNA cleavage assay, we found that the DNA was cut close to the target site (Figure 2d), albeit with an efficiency lower than that noted for the two-site substrate (Figure 2e, 2f).

SauUSI cuts DNA at multiple locations away from the target site

Previously, it was reported that SauUSI causes multiple nicks close to the target site (12), which led us to find the locations of cleavage. To identify the location of cleavage, we carried out the following experiments. We performed a cleavage assay using SauUSI and BstNI on a two-site substrate (DS 2) which had the sites spaced 609 bp apart and visualized them on an agarose gel (Figure 4a). It was evident that the fragments generated upon treatment with SauUSI migrated slightly lower than that of the fragments generated by BstNI (Figure 4a). This indicated that SauUSI cleaved the DNA away from the target site. Furthermore, when we treated a 1415 bp single-site substrate (SS 2) with SauUSI, a pair of doublets were visible on the gel about the size of the fragments generated by the single-site cutter BstNI (Figure 4b). This suggested that SauUSI made multiple dsDNA breaks beyond the target site.

To better resolve the fragments generated by SauUSI cleavage at multiple points, we used a 120 bp substrate with a target site flanked by 57 bp and 58 bp DNA on either side (Figure 4c). Multiple fragments were observed on the cleavage of the 120 bp methylated DNA by SauUSI (Figure 4c). An eyeball estimation of the sizes of the fragments indicated three distinct types: fragments that were ~100 bp, which would form if cleavage occurred ~35 bp away from the target site; fragments that were ~75 bp and >35 bp, which would form if cleavage occurred ~15 bp away from the target site (Figure 4d). The fragments that are ~75 bp and >35 bp seem to be of greater

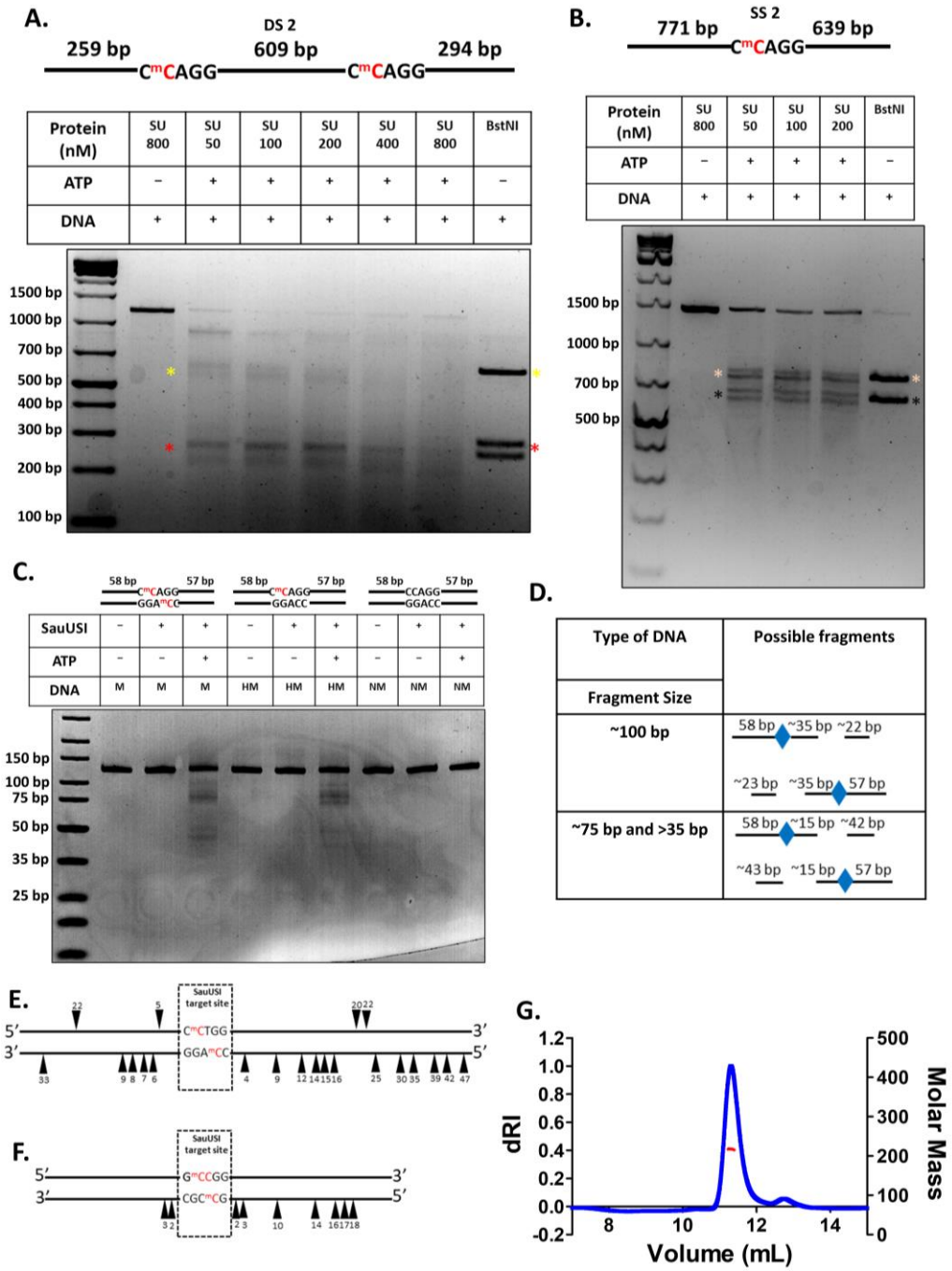


Figure 4: Position of cleavage by SauUSI and oligomeric status of SauUSI: (A) A two-site substrate (DS 2) used for the cleavage assay; the separation between the two sites is 609 bp. The cytosine highlighted in red represents methylation. The target site is methylated on both the strands. Representative 2% agarose gel for the two-site cleavage using varying concentrations (50 nM - 800 nM) of SauUSI. The yellow and red asterisks represent the central (~609 bp upon BstNI treatment) and flanking regions (~294 bp and ~259 bp upon BstNI treatment) of the DNA post cleavage, respectively. (B) A single-site substrate (SS 2) used for the cleavage assay. The cytosine highlighted in red represents methylation. The target site is methylated on both the strands. Representative 2% agarose gel for the single-site cleavage using varying concentrations (50 nM - 200 nM) of SauUSI. The beige and black asterisks represent the two fragments (~771 bp and ~639 bp upon BstNI treatment) of the DNA post cleavage. (C) 120 bp methylated, hemimethylated and non-methylated substrates used for the cleavage assay. The cytosine highlighted in red represents methylation. Representative 8% Native PAGE gel for the 120 bp substrates. (D) Interpretation of the possible cleavage products from panel c. (E) Cartoon representation of the sequencing of the single-site DNA after cleavage by SauUSI. The arrowheads represent nicks on the DNA substrate, the position of the arrowheads illustrates on which strand of the DNA the nicks were observed and how far away they are, in base pairs, from the target site. (F) Cartoon representation of the nicks on a two-/multi-site substrate after cleavage by SauUSI identified using run-off sequencing (data from Xu. et al., 2011). (G) SEC-MALS chromatogram of SauUSI shows the refractive index signal with the derived molar mass indicated by a red horizontal line.

intensity than the ~100 bp fragment which indicates that dsDNA breaks occur predominantly ~15 – 20 bp away from the target site.

Additionally, we studied the cleavage of a 120 bp hemi-methylated single-site substrate (Figure 4c). The cleavage pattern was similar to that of the methylated DNA but not the same, possibly because the enzyme could bind to the target sites in only a single orientation, unlike the two orientations possible in case of a palindromic fully methylated site. However, the cleavage efficiency of both the substrates was low, indicating the requirement of a minimum substrate length for efficient cleavage. Single-site cleavage was not observed previously, possibly because of the comparatively short length of the DNA used for the assay (12). As expected, a 120 bp DNA with non-methylated target sequence was not cleaved by SauUSI (Figure 4c).

Cleavage of a single-site substrate (SS 1) was used to map the location of the nicks causing the dsDNA breaks using run-off sequencing. The sequencing was performed both in the forward and reverse directions to locate nicks on the lower strand and upper strand of the DNA, respectively (Figure 5 and 6). Sequencing in the forward direction located nicks 4 bp, 9 bp, 12 bp, 14 bp, 15 bp, 16 bp, 25 bp, 30 bp, 35 bp, 39 bp, 42 bp & 47 bp to the right-hand side (RHS) and 6 bp, 7 bp, 8 bp, 9 bp & 33 bp to the left-hand side (LHS) of the target site (Figure 4e and 5). Sequencing in the reverse direction also showed multiple nicks on the RHS (20 bp and 22 bp) and LHS (5 bp and 22 bp) of the target site (Figure 4e and 6). The presence of nicks ~15 – 20 bp away from the target site corroborated with the cleavage data obtained from the 120 bp methylated substrate. However, since the efficiency of cleavage for the 120 bp substrate is low, the optimal range of cleavage may differ for longer substrates, which are cut more efficiently by SauUSI.

The noise in the sequencing data increased at >45 bp away from the target site, thereby making it difficult to interpret if there were nicks beyond this distance. The position of the nicks that we observed were consistent with the location of the cleavage mapped by Xu et al. (12) from a two-site substrate and a multi-site substrate (Figure 4f). The small variation in the location of cleavage reported by the two studies could be a result of the difference in the sequence of the DNA used, which can affect the exact location of cleavage. Effect of sequence on the location of cleavage has been reported previously in the case of Type ISP REases (22).

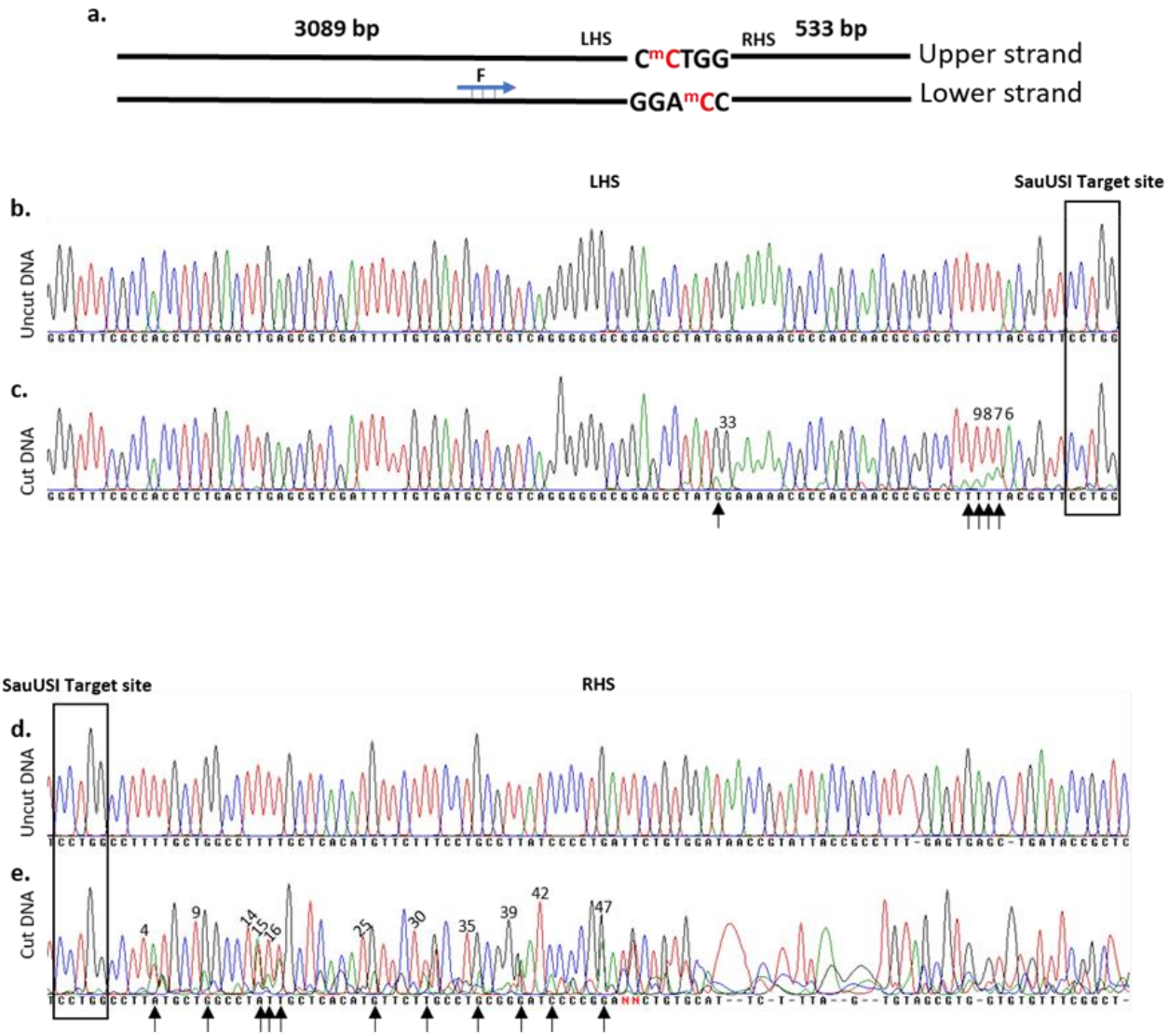


Figure 5: Sequencing assay (forward direction) to locate nicks in the lower strand: (a) Schematic of the DNA substrate used for sequencing. The cytosines highlighted in red represents methylation. F refers to the forward primer. The DNA preceding the target site is annotated LHS, and the DNA succeeding the target site is annotated RHS. (b) Sequencing of the uncut DNA substrate on the LHS used for the assay. (c) Sequencing of the cleaved fragment of DNA in the forward direction on the LHS. The arrows on the doublet peaks indicate nicks identified by sequencing. The numbers represent the position of cleavage in base-pairs away from the SauUSI target site. (d) Sequencing of the uncut DNA substrate on the RHS used for the assay. (e) Sequencing of the cleaved fragment in the forward direction on the RHS. The arrows on doublet peaks indicate nicks. The numbers represent the position of cleavage in base-pairs away from the SauUSI target site. It is to be noted that the sequencing result gets very noisy (multiple overlapping peaks) >45 bp from the SauUSI target site.

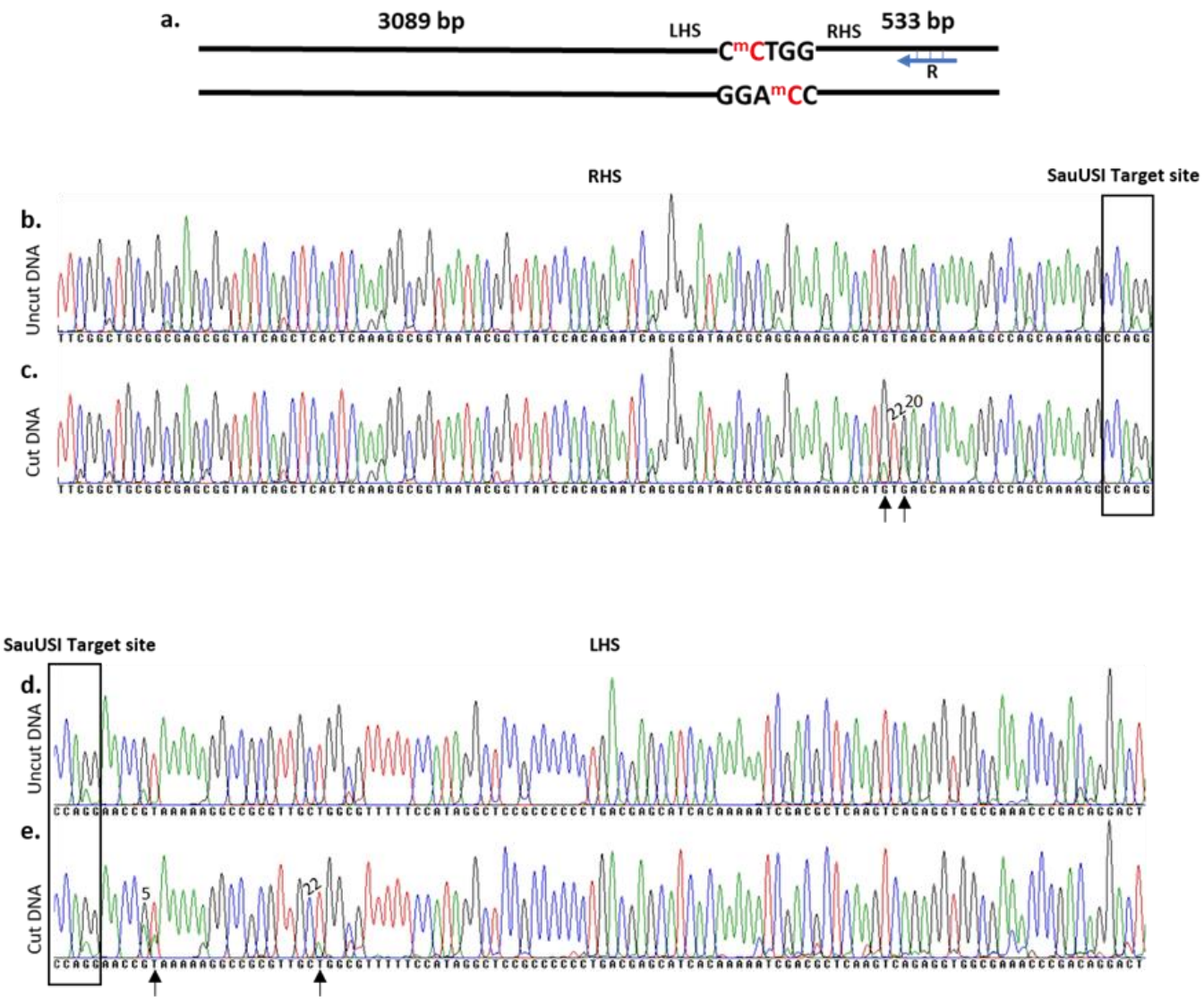


Figure 6: Sequencing assay (reverse direction) to locate nicks in the upper strand: (a) Schematic of the DNA substrate used for sequencing. The cytosines highlighted in red represents methylation. R refers to the reverse primer. The annotations of LHS and RHS are as in supplementary figure 4 for the sake of convenience. (b) Sequencing of the uncut DNA substrate on the RHS used for the assay. (c) Sequencing of the cleaved fragment of DNA in the reverse direction on the RHS. The arrows on the doublet peaks indicate nicks identified by sequencing. The numbers represent the position of cleavage in base-pairs away from the SauUSI target site. (d) Sequencing of the uncut DNA substrate on the LHS used for the assay. (e) Sequencing of the cleaved fragment in the reverse direction on the LHS. The arrows on the doublet peaks indicate nicks. The numbers represent the position of cleavage in base pairs away from the SauUSI target site.

SauUSI exists as a dimer in solution

dsDNA break requires cleavage of at least two spatially close phosphodiester bonds, one each from the two strands of the duplex. Usually, an endonuclease catalyzes DNA cleavage by one of the following means. A constitutive homodimeric enzyme binds to the target site, and each of the two nuclease active sites cleaves one of the two phosphodiester bonds, as in the case of many Type II REases (33); or the two active sites of a monomeric enzyme cleave the two phosphodiester bonds as in homing endonuclease PI-*Sce*I(34); or a single active site formed by a dimeric endonuclease cuts the strands sequentially as in the Type II REase BfiI (35); or two monomeric enzymes associate transiently to catalyze the cleavage of the two neighboring phosphodiester bonds as in the case of the REase FokI or the Type III RM enzymes (23, 36),; or the two nuclease domains come together in *cis* by translocation along the DNA as in the case of Type I and Type ISP RM enzymes (22, 37). Thus, an analysis of the oligomeric structure of the enzyme can provide insights into the mechanism of DNA cleavage. We characterized the oligomeric structure of SauUSI using size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS). A clear monodisperse peak was observed from the SEC-MALS run (Figure 4g). The scatter at 659 nm corresponded to a mean molecular weight (M_{avg}) of 217.7 (±0.022%) kDa, which is approximately twice the theoretical mass of the monomeric protein (109.9 kDa). Therefore, we concluded that SauUSI existed as a dimer in solution.

Molecular architecture of SauUSI

The homodimeric assembly of SauUSI suggested that the enzyme could perform dsDNA breaks by employing two nuclease active sites to catalyze the hydrolysis of the phosphodiester bonds on either strand. Usually, such dimeric endonucleases have the oligomeric interface formed by the nuclease domains, with their respective active sites being spatially close to each other (38). To ascertain the architecture of SauUSI, we crystallized the apo-enzyme and determined the structure using X-ray crystallography at a resolution of 3.1 Å (Table 2). The structure revealed that the single polypeptide protein SauUSI is a dimer and is made of four domains - the N-terminal nuclease domain, followed by a Superfamily 2 (SF2) helicase-like ATPase domain, and an α-helical coupler domain that connects the ATPase to the target recognition domain (TRD) (Figure 7a, 7b,

Table 2: Crystallographic data and refinement statistics.

Structure	SauUSI^{SELMET}
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	86.6 175.7 89.4
α, β, γ (deg)	90, 115.7, 90
Wavelength (Å)	0.9795
Resolution (Å) (Highest resolution shell)	87.8-3.1 (3.22-3.10)
R_{merge} (%) overall	0.08 (1.103)
I/σ	13.1 (1.6)
Completeness (%)	99.9 (100)
Redundancy	6.9 (7.1)
Refinement	
Resolution (Å)	56.8–3.1
No. of reflections	15469
R_{work}/R_{free} (%)	24.5/29.7
No. of atoms	12275
No. of ions	4
No. of water molecules	33
RMS deviations in	
Bond lengths (Å)	0.005
Bond angles (deg)	0.752
PDB ID	7CLG

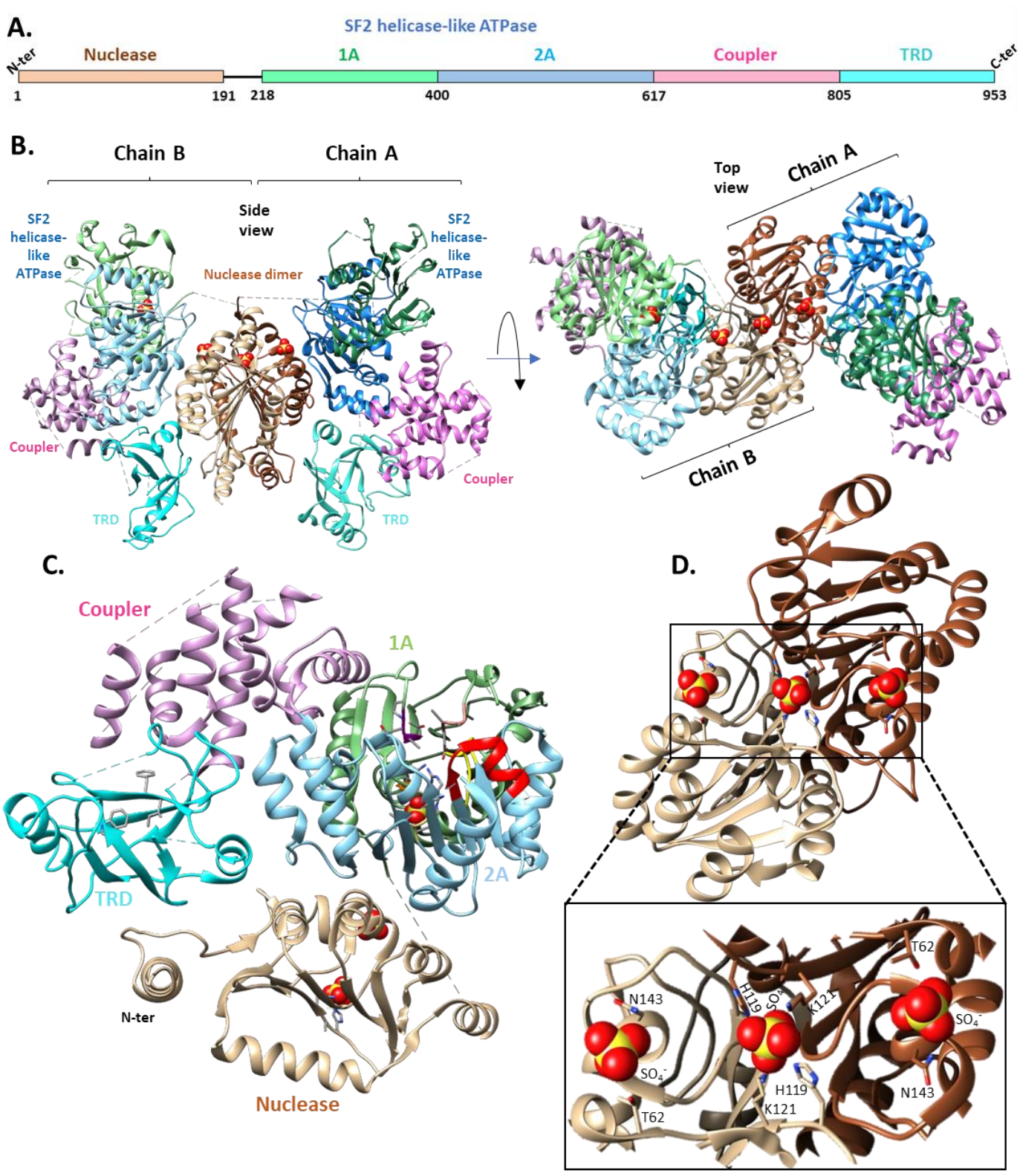


Figure 7: Molecular architecture of SauUSI: (A) The primary domain arrangement in a SauUSI protomer, the line connecting the nuclease domain to the SF2 helicase-like ATPase signifies an unstructured linker. (B) Ribbon diagram of two views of the crystal structure of the dimeric SauUSI. Each structural domain of the protomers is colored distinctly. The sulfates are represented as spheres. (C) Structure of a protomer of SauUSI with nuclease domain in tan; the 1A and 2A domains of the ATPase having the RecA fold in green and blue, respectively; the coupler domain in pink; the TRD (SRA) domain in cyan. Certain important motifs/residues of domains are highlighted in sticks. (D) Structure of the nuclease from the two protomers highlighting the dimeric interface. A zoomed view of the DNA binding region of the nuclease dimer identified using the bound sulfate ions shown as spheres.

and 7c). The structure also revealed that the interaction of the nuclease domains of the respective monomers stabilized the dimeric assembly (Figure 7b and 7d).

The SauUSI nuclease domain (residues 1-191) belongs to the Phospholipase D (PLD) family nucleases with the conserved 119-H(X)K(X)4E-126 motif. The dimeric interface of SauUSI formed by the two nuclease domains had a buried surface area of 1970 $Å^2$ and was reminiscent of the interface seen in the crystal structure of the PLD family nuclease BfiI (PDB ID: 2C1L) (Figure 7b and 7d) (39). In the dimeric assembly of SauUSI, the nuclease catalytic site appeared pre-formed by the 'HxK' motifs from the two protomers. Three sulfate ions, which possibly mimic the phosphate backbone of a substrate DNA, highlighted the path that the DNA would take on binding to the nuclease domain (Figure 7b and 7d). One of them was bound to the active site residues His-119 and Lys-121 of both the protomers, while the other two interacted with the main chain of Ser-142, and the side chains of Thr-62 and Asn-143 from the two protomers, respectively. Sulfate ions were part of the crystallization condition.

Experiments on BfiI, a Type II REase with a PLD nuclease domain, have shown that the single active site (formed by the two protomers) cuts both the strands of DNA by undergoing a 180^o conformational switch and that both the active site histidines (from the two protomers) are not functionally equivalent. Instead, the roles of the histidines are sequential. One of the histidines functions as a nucleophile that attacks the phosphate backbone and forms a histidine-phosphate intermediate and then this covalent intermediate is broken with the help of the second histidine (40). However, this needs to be verified in the context of SauUSI.

The nuclease domain of SauUSI is connected to the ATPase domain by a long linker (residues 192-220). The linker could not be built as the corresponding electron density was absent. In case of BfiI, it is hypothesized that the linker between the PLD nuclease and TRD plays a role in repressing nucleolytic activity of the PLD domain by mimicking the nucleic acid backbone (39). The linker in SauUSI is 26 amino acids long and possess negatively charged residues Glu-194, Glu-199, Glu-203, Glu-205 and Asp-213. Alternatively, the linker could structurally hinder the accessibility of the nuclease active site thereby potentially acting as a regulator of the nuclease activity. These alternatives need further examining in SauUSI.

The RM systems CglI and NgoAVII also possess PLD nucleases and SF2 helicase-like ATPases (41). However, unlike SauUSI, NgoAVII and CglI are multi-subunit enzymes having R_2H_2 stoichiometry, where R and H are the two subunits. The R-subunit consists of the PLD nuclease domain and B3 TRD whereas, the H-subunit consists of the ATPase domain along with the accessory domains Z1 and C. Though, NgoAVII/CglI contain a PLD nuclease and an SF2 ATPase, the differences in their domain arrangement from that of SauUSI hint at a distinct mode of DNA cleavage in the two types of enzymes.

Recognition of the methylated target sequence by the TRD activates the ATPase

At the C-terminus of the protein is the TRD (residues 806 to 953), which is linked to the ATPase domain by the coupler domain (residues 618 to 805). The structure of the TRD revealed that it has the fold of an SRA (SET and RING associated) domain, a fold that specializes in recognition of 5-methylcytosine in DNA (Figure 6a). A search using the DALI server(42) identified the SRA domain of SuvH6 from *Arabidopsis thaliana* (PDB ID: 6A5N) (43), a histone methyltransferase, which belongs to the SuvH family, to be the closest structural homologue of the SauUSI SRA domain. Based on the DNA-bound structure of SuvH6 and the corresponding structure-based sequence alignment, a model of SauUSI SRA domain bound to the target sequence was generated (Figure 8a).

The model predicted SauUSI SRA to approach a DNA duplex from its minor groove side and flip out the 5-methylcytosine of the target sequence into a hydrophobic pocket lined by the residues Tyr-831, Phe-841, Ile-842 and Trp-868, while Met-829 inserted into the cavity formed in the DNA by base flipping (Figure 8a). Asp-858 is positioned to make base-specific hydrogen bonds with the flipped cytosine, Tyr-831 is positioned to interact with the methyl group to establish specificity for the methylated cytosine, and Trp-868 could stack against and stabilize the flipped base. Additionally, the region between 848 and 854 is disordered in the apo-structure, which may interact with the flipped base upon DNA binding, as predicted based on the role of the structurally equivalent loop in SuvH6. SauUSI can also recognize target sequences having 5-hydroxymethylcytosine (5hmC) and cleave such DNA (12). Studies on SuvH5, another close homolog of the SRA domain of SauUSI, indicate that the mechanism of recognition of 5-

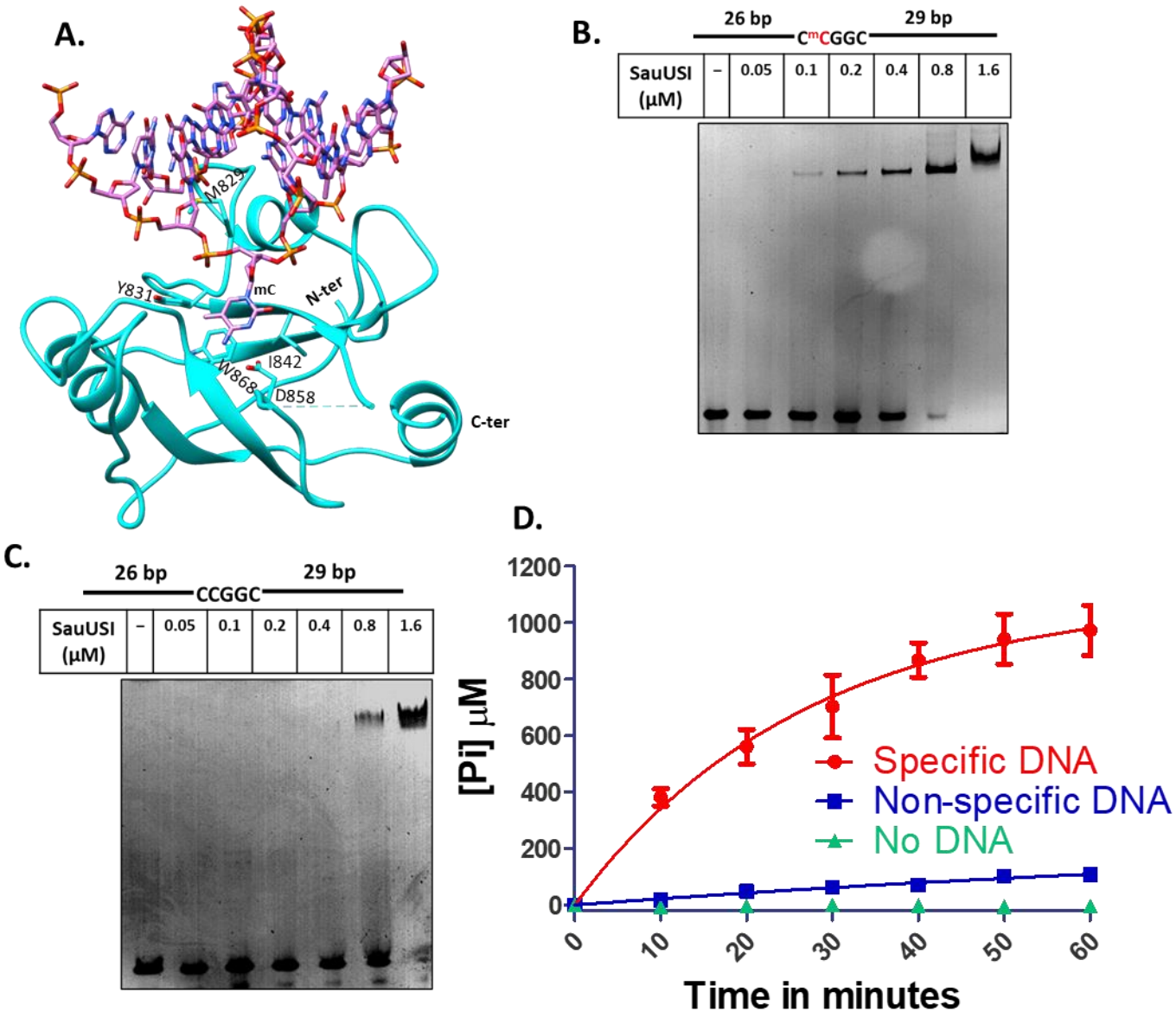


Figure 8. Target recognition by the SRA domain and DNA dependent ATP stimulation of SauUSI: (A) Structure of the SRA domain of SauUSI with DNA modeled. The DNA is from the structure of the structural homolog SUVH6, an H3K9 histone methyltransferase from *Arabidopsis thaliana* (PDB ID: 6A5N). (B) The specific DNA substrate used for EMSA. The cytosine highlighted in red represents methylation. The target site is methylated on both the strands. Representative 5% native PAGE EMSA gel using specific DNA substrate (250 nM) and assay performed over a protein concentration range of 0.05 μM – 1.6 μM. (C) The non-specific DNA substrate used for EMSA. Representative 5% Native PAGE EMSA gel using non-specific DNA substrate (250 nM) and assay performed over a protein concentration range of 0.05 μM-1.6 μM. (D) ATPase stimulation assay carried out using the malachite green method, represented by inorganic phosphate released (on the Y-axis) over time (on the X-axis) in reactions containing specific DNA, non-specific DNA and no DNA (n=3).

methylcytosine and 5hmC is similar, and the binding affinities are comparable (44). We hypothesize a similar situation for the SRA domain of SauUSI.

Base specific interaction with 5-methylcytosine in the target sequence was found to be essential for SauUSI to discriminate between a specific DNA substrate and a nonspecific DNA. Replacement of 5-methylcytosine by cytosine in the recognition sequence reduced the affinity of SauUSI for the DNA as observed by electrophoretic mobility shift assay (EMSA) using a 60 bp DNA (Figure 8b and 8c). A shift in the DNA was detected at 0.1 µM of SauUSI for specific DNA, whereas for non-specific DNA it was at 0.8 µM. Furthermore, the ATPase activity of SauUSI was stimulated on recognition of the target sequence (Figure 8d). A nuclease inactive point mutant, SauUSI^{H119A}, was used for studying the ATPase activity of the enzyme to prevent cleavage of the DNA in the assay. We found that the ATPase activity of 5 nM of SauUSI^{H119A} incubated with 2 mM ATP was negligible in the absence of DNA as measured by quantifying the phosphate released using the malachite green ATPase assay. Addition of 10 nM non-specific DNA (in which the target sites were not methylated) resulted in only a marginal increase in the activity. However, in the presence of 10 nM specific DNA, a significant amount of ATP was hydrolyzed. Since the enzyme used was a nuclease inactive point mutant of SauUSI, there is a possibility of multiple rebinding events of the enzyme to the substrate. Hence, the phosphate release recorded cannot be directly compared to the cleavage rates.

The manifold increase in the ATPase activity of the enzyme upon recognition of a specific DNA could be a result of two possible steps occurring sequentially. One, the SRA domain's intrinsic preference to bind methylated DNA over non-methylated DNA, thereby leading to the engagement of the ATPase with only specific DNA. Two, activation of the ATPase upon binding to a specific DNA, which results in conformational changes transduced from the target-sequence-bound SRA domain to the ATPase and due to the binding of the ATPase to DNA. The coupler domain that connects these two domains may play a role in this allostery. The coupler was made of two subdomains: a subdomain of four short helices (618-669) and another made of at least six helices (673-805). The electron density corresponding to residues 710 to 773 of the coupler domain was poor, and the region could only be built partially.

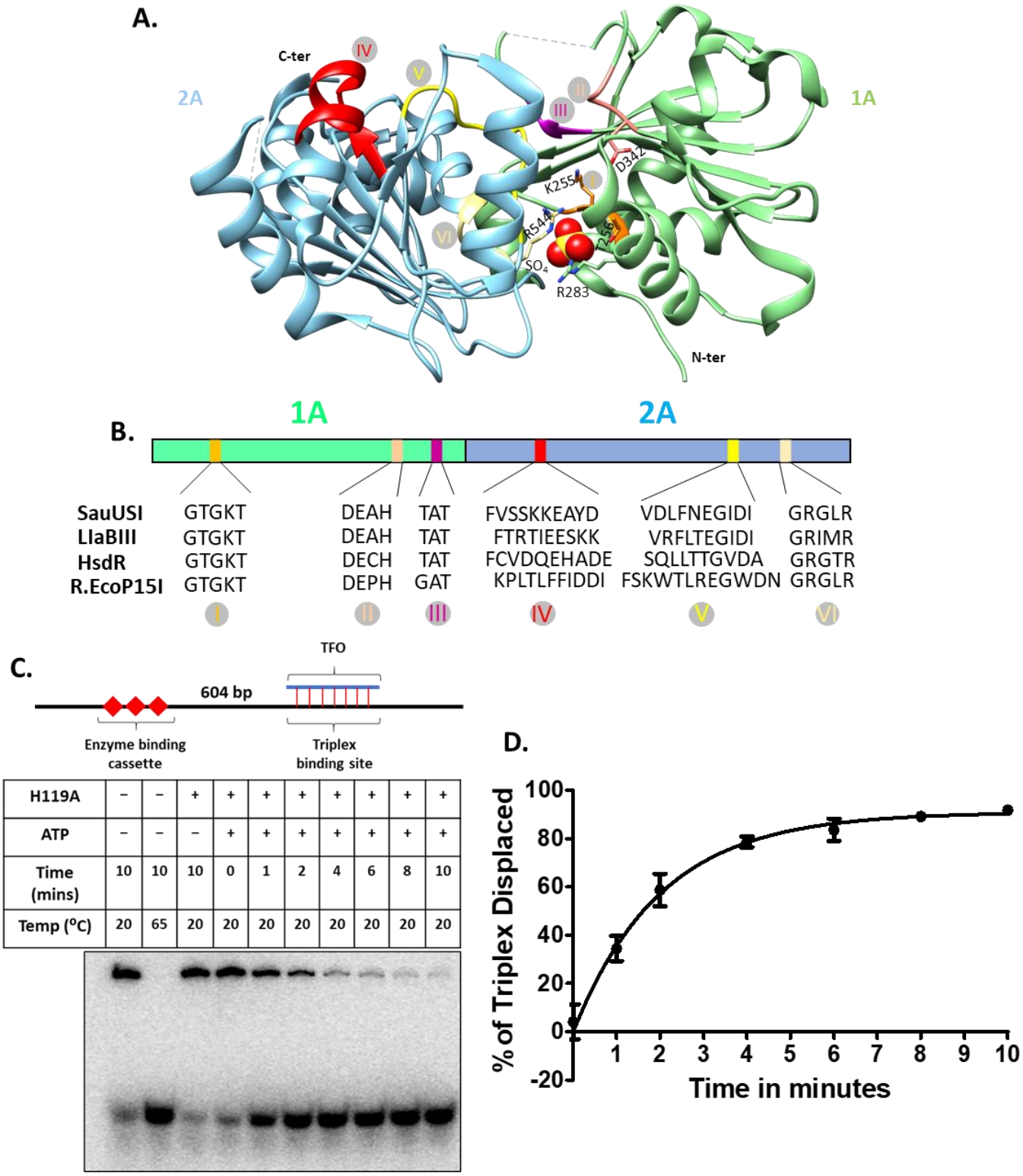


Figure 9: ATPase-driven DNA translocation by SauUSI: (A) The primary domain architecture of the SF2-helicase like ATPase along with the canonical motifs marked in specific colors. The canonical motifs of the SF2-helicase like ATPase of SauUSI are compared to that of LlaBIII (Type ISP REase), HsdR (Type I REase) and EcoP15I (Type III REase). (B) Structure of the ATPase domain with the two RecA folds colored distinctly. The catalytically important Walker A and Walker B residues, along with a sulfate ion (shown as a sphere) bound at the ATP binding site are shown. Also, illustrated are the other canonical motifs in the same color code as in panel a. (C) The DNA used for the triplex displacement assay has an enzyme binding cassette located 604 bp away from the triplex binding site. A representative PAGE gel of the triplex displacement assay carried out over a time period of 1 minute to 10 minutes. (D) Translocase activity of SauUSI^{H119A} monitored by the percentage of triplex displaced as a function of time (n=3).

The ATPase powers DNA translocation by SauUSI

The SauUSI ATPase belongs to the SF2 DExH helicase family. In the apo-structure of SauUSI, the two ATPases from each monomer are spatially separate and do not interact with each other. The SF2 helicase-like ATPase domain features two RecA-like sub-domains, i.e., 1A & 2A (residues 221 to 400 and 401 to 617) and the ATP binding pocket is located in a cleft between them, which in the crystal structure of the apo-enzyme was occupied by a sulfate ion (Figure 9a). A search for structural homologs of the ATPase of SauUSI using DALI server (42) revealed it to be similar to the ATPase domains of HsdR of the Type I REase from *Vibrio vulnificus* (PDB ID: 3H1T) and the Type ISP REase LlaBIII (PDB ID: 4XQK) (22, 45). The 1A and 2A sub-domains of the SauUSI ATPase structurally aligned well with the corresponding sub-domains of LlaBIII and HsdR (Figure 10). LlaBIII has a β-hairpin loop extending from the 1A sub-domain, which is essential for DNA translocation (46). SauUSI has a short-disordered loop at the structurally equivalent position, which may have a role in DNA interaction and translocation.

A sequence alignment of the ATPase of SauUSI with the corresponding domain in LlaBIII, HsdR and EcoP15I revealed the six conserved canonical motifs distributed amongst the two RecA-like sub-domains (figure 9b). Motif I (Walker A) and motif VI (arginine finger) are involved in ATP binding, whereas motif II (Walker B) is involved in ATP hydrolysis. The nucleic acid interacting module of the SF2 helicase-like ATPase involves motif IV and motif V. Motif III helps in coupling ATP hydrolysis to nucleic acid translocation. The sulfate ion interacts with Thr-256 of the Walker A motif, Arg-283, and Arg-544 of motif VI (Figure 9a).

Previously, it has been shown that ATP hydrolysis is essential for activation of the SauUSI nuclease (12). The fact that SauUSI cut DNA close to one of the two target sites and that it could cut DNA having a single target site was reminiscent of Type III REase. In a Type III REase, the ATPase functions as a switch to conformationally activate the enzyme upon ATP hydrolysis to perform long-range diffusion along the DNA and/or cleave it (23, 24). Interestingly, the same family of ATPase in Type I and Type ISP REases function as dsDNA translocating motor that actively translocates the enzyme along the DNA. The convergence of two translocating enzymes results

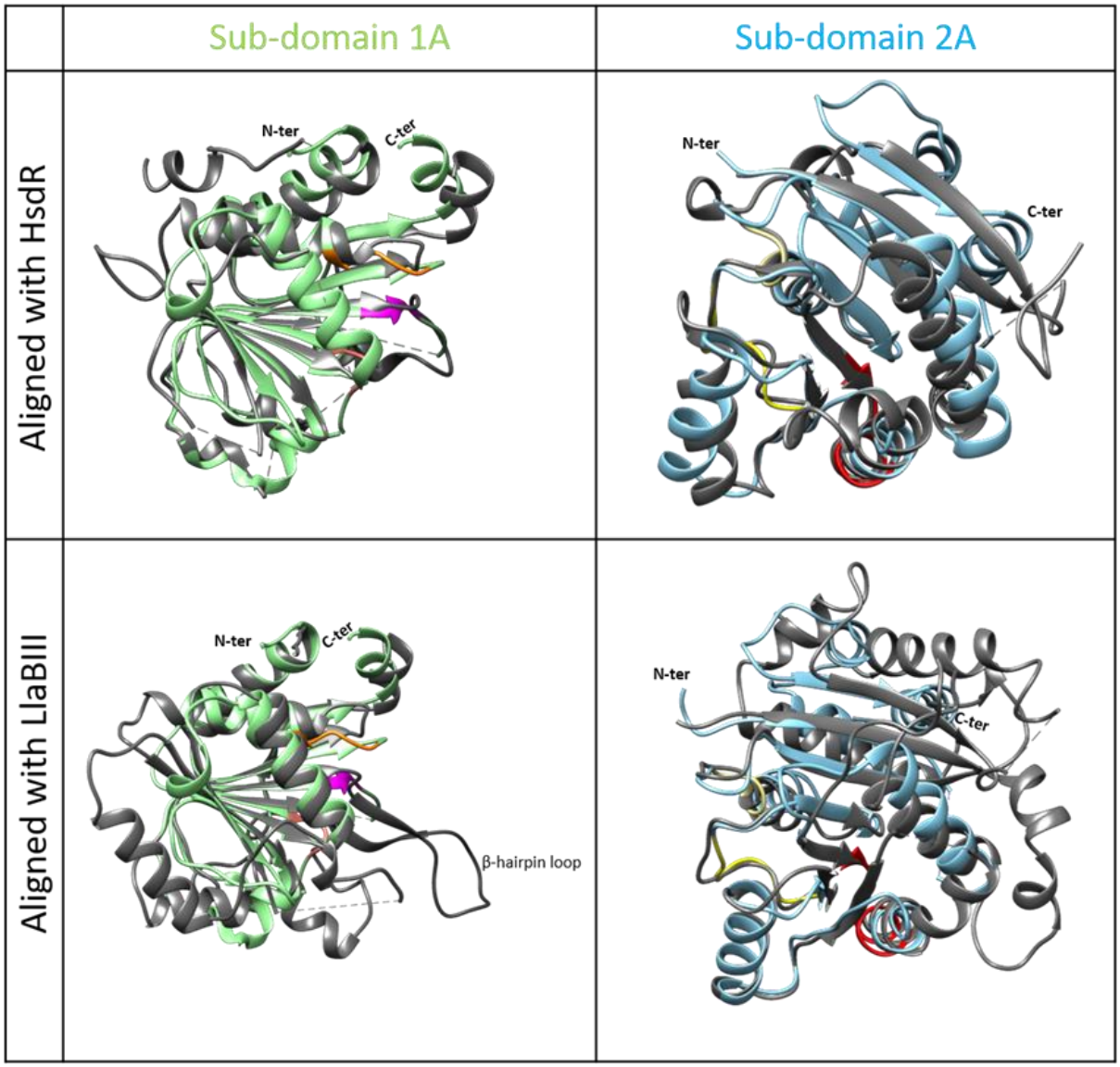


Figure 10: Alignment of the 1A and 2A sub-domains of SauUSI with that of HsdR and LlaBIII: The 1A (green) and 2A (blue) sub-domains of the SF2 helicase-like ATPase domain of SauUSI structurally aligned with the respective subdomains (grey) of HsdR (PDB ID: 3H1T) and LlaBIII (PDB ID: 4XQK).

in DNA cleavage at the point of the meeting, which is random and located somewhere between two target sites (22, 37).

To find out if the SauUSI ATPase, which is homologous to the ATPases of both Type I/ISP and Type III RM enzymes, functioned as a switch or as a motor, we performed a triplex displacement assay. The assay monitors the ability of an enzyme to displace a TFO (triplex forming oligo) from a DNA (47). A translocating ATPase motor can displace the TFO from the DNA, while an ATPase switch cannot displace the TFO (48). For the assay, we used a substrate DNA having a SauUSI binding cassette containing three closely spaced target sites and a TFO binding site 604 bp downstream of the cassette. The TFO was radiolabeled to visualize them on a non-denaturing polyacrylamide gel (PAGE). SauUSI^{H119A} was used for the assay to prevent DNA cleavage. Displacement of the TFO, if any, by 30 nM SauUSI^{H119A} was monitored as a function of time (Figure 9c and 9d). It was observed that in the presence of ATP, over 85% of the TFO was displaced within 10 minutes of the reaction (Figure 9d), indicative of a translocase activity by SauUSI. Displacement of the TFO was not observed in the absence of ATP. We modelled a DNA into the target recognition domain and the SF2 helicase-like ATPase by structurally aligning two close homologs (TRD of SuvH6 and the SF2 helicase-like ATPase of LlaBIII) and noticed that the orientation of the TRD and the SF2 helicase-like ATPase was such that its DNA binding surface faced away from the expected path of the nuclease-bound DNA (as predicted by the position of sulfates). It was apparent that only a large conformational change could engage both the ATPase and the nuclease with the DNA simultaneously unless the DNA takes a convoluted path that wraps around the enzyme. Hence, the apo-structure of SauUSI may be that of an inactive state. Target recognition and ATP hydrolysis are steps towards the activation of the SauUSI nuclease.

SauUSI nucleolytically shreds DNA having target sites at both its ends

Like Type I or ISP RM enzymes, SauUSI translocated DNA, however, unlike the former, SauUSI cut DNA close to one of the two target sites rather than somewhere in between the target sites. Furthermore, unlike Type I or ISP REases, SauUSI was able to cut linear DNA having a single target site (see above). These observations implied that though ATP hydrolysis was essential for the nucleolytic activity of SauUSI, the convergence of two such enzymes by DNA translocation was

not essential. As discussed above, the mode of DNA cleavage by SauUSI appeared distinct from canonical Type I/ISP, II or III REases. This deduction was further strengthened when a careful examination of the cleavage product of the two-site substrate (DS 1) by SauUSI on an agarose gel revealed that of the three primary fragments, the intensity of the central ~1300 bp fragment was disproportionately lower than the ~3089 kb and ~639 bp fragments that flank the two target sites, respectively. This was clear when the cleavage pattern of the two-site substrate by SauUSI was compared with the cleavage pattern of the same DNA, obtained from treatment with the Type II REase BstNI, which cuts at 5'-CC/^{5m}CTGG-3', the methylated version of which is the target site for SauUSI (Figure 11a).

The reduction in the intensity of the ~1300 bp fragment became more pronounced with increasing concentration of SauUSI and was close to zero at 400 nM of enzyme concentration (Figure 11b). In contrast, the reduction in the intensity of the ~3089 bp and ~639 bp fragments were much less. Similarly, neither of the two DNA fragments generated by the single-site cleavage by SauUSI showed a disproportionate decrease in intensity (inset Figure 11b). This led us to hypothesize that the DNA between two target sites underwent further nucleolytic processing. To test this hypothesis, we converted the two-site substrate to a three-site substrate (TS) by introducing a new site 203 bp away from one end of the DNA (Figure 11c). A comparative cleavage assay performed using the two-site substrate (3 nM) and the three-site substrate (3 nM) with 100 nM SauUSI and 2 mM ATP showed that there was a sharp decrease in the intensity of the ~2881 bp fragment formed by the processing of the three-site substrate by SauUSI (Figure 11c and 11d). This was very similar to the decrease observed in case of the ~1300 bp fragment of the two-site substrate. This assay confirmed that the DNA between two target sites underwent further nucleolytic processing. A similar trend was observed with the substrates DS 2 (Figure 4a). As the nucleolytic activity of SauUSI resulted in a smear when visualized on an agarose gel, which extended to almost the bottom of the gel, we concluded that the DNA was cut at multiple locations along the entire length of the DNA resulting in shredding. In comparison, DNA flanking the two target sites showed minimal smearing, suggesting that they were, if at all, minimally shred.

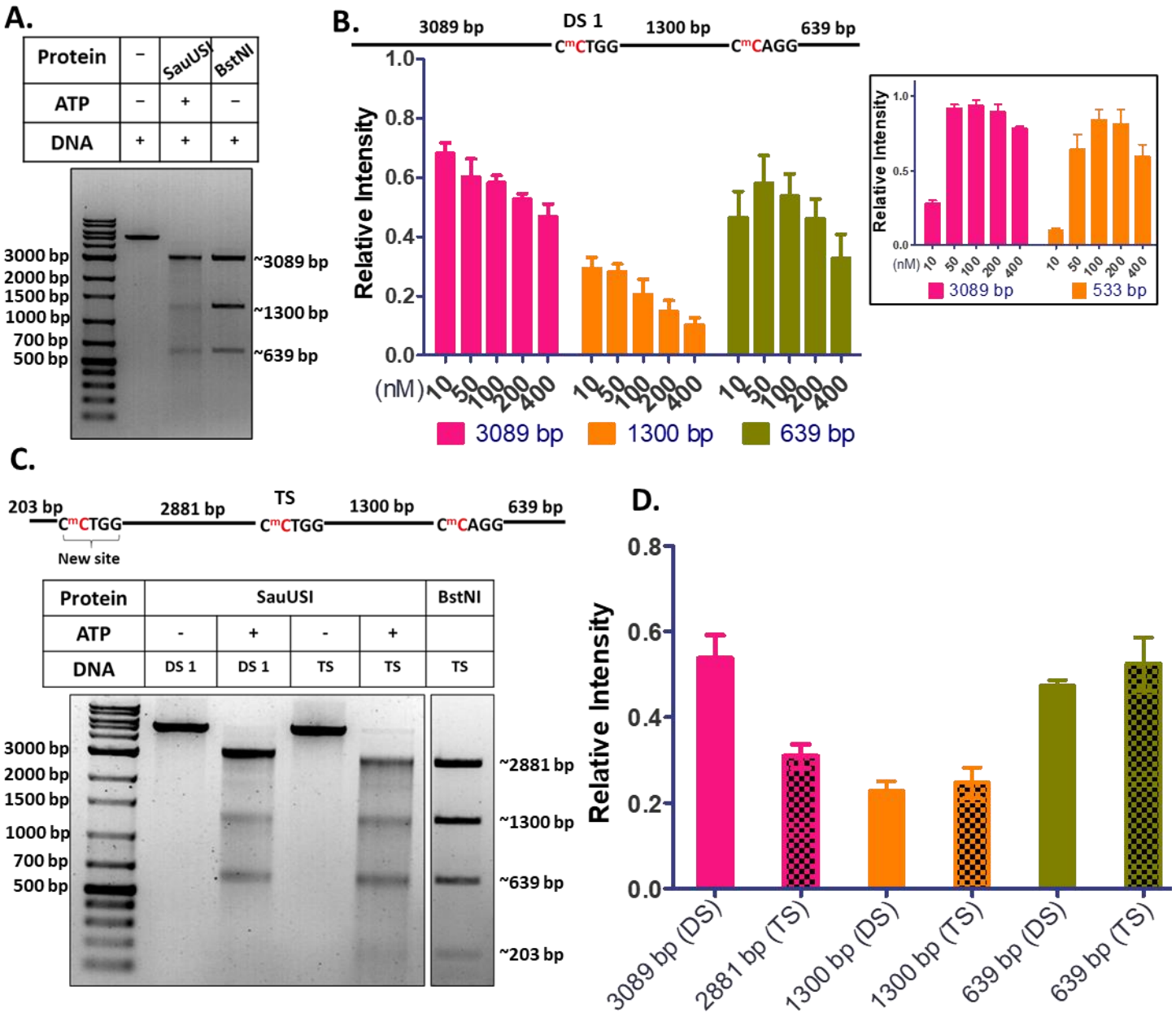


Figure 11: Reduced intensity of major fragments: (A) Two-site (DS 1) cleavage pattern comparison between SauUSI (100 nM) and BstNI (Type II REase) using 3 nM DNA at 37°C. (B) Schematic of the two-site substrate. The cytosine highlighted in red represents methylation. The target site is methylated on both the strands. Quantification (relative intensity) of the three major cleavage fragments of the two-site substrate across the concentration gradient of SauUSI, assay conditions as discussed in Figure 2b (n=3). (*Inset*) The quantification (relative intensity) of the two major cleavage fragments of the single-site substrate (SS 1) across varying concentrations of SauUSI, as shown in Figure 1c (n=2). (C) The three-site substrate used for the cleavage assay using 3 nM DNA and 100 nM SauUSI incubated at 37°C for 1 hour. The cytosine highlighted in red represents methylation. Representative gel of the three-site substrate (TS) cleavage pattern in comparison to the cleavage seen with a two-site substrate (DS 1). For comparison, the product of the three-site substrate on cleavage with the REase BstNI is also shown. (D) The quantification (relative intensity) of the major cleavage products between the two-site substrate (DS 1) cleavage pattern and the three-site substrate (TS) cleavage pattern (n=3).

Discussion

The biochemical and structural analysis of SauUSI, a major barrier to HGT in *S. aureus*, revealed that the overall architecture of the enzyme and its mode of DNA cleavage were different from those of a *bona fide* Type I, ISP, II or III REases. Structural and SEC-MALS studies revealed that the single polypeptide chain of SauUSI is made of the nuclease, the ATPase, the coupler and the TRD, and exists as a homodimer in solution. The linear domain arrangement in SauUSI is similar to that of the ATP-independent Type II REase MmeI (lacking the ATPase domain) and the ATP-dependent Type ISP REases, both of which have a methyltransferase domain as well (22, 49). However, the latter two, unlike SauUSI, are monomeric. We think that only one of the two TRDs of SauUSI will recognize and bind to the palindromic target site 5'-S^{5m}CNGS-3', enabling SauUSI to assemble on the target site in two different directions (50, 51). Binding of the dimeric enzyme to the target site allosterically activates the ATPase domain, which was experimentally observed as stimulation of ATP hydrolysis upon addition of substrate DNA. Hence, on a two-site substrate, there are four possible arrangements of SauUSI on the DNA, i.e. head-to-head, head-to-tail, tail-to-tail and tail-to-head (Figure 12a). For the sake of simplicity, we shall focus on one of these possibilities, i.e. the head-to-head orientation (Figure 12b). As demonstrated using the triplex displacement assay, upon ATP hydrolysis SauUSI would translocate along the DNA. The translocating enzyme eventually converges in *cis* on to a stationary SauUSI molecule bound to the second target site. We propose that the stationary enzyme acts as a roadblock stalling the translocating enzyme, and predict that the physical convergence of the two enzymes stimulates the nuclease activity of the stationary enzyme resulting in dsDNA break. Our prediction of the stimulation of the nuclease upon convergence derives from the observation that a two-/multi-site substrate is cleaved with higher efficiency than a single-site substrate. Note that the convergence of two translocating SauUSI molecules, i.e. enzymes that have left their respective target sites, does not result in DNA cleavage, because, if true, such an event would have had predominantly resulted in cleavage at a location somewhere in between the two target sites where the two enzymes converge, which we did not observe. Nicks on dsDNA upon stalling of enzyme due to structural constraints of a DNA -protein complex or due to the presence of a roadblock has been observed in case of Type I REases (52, 53).

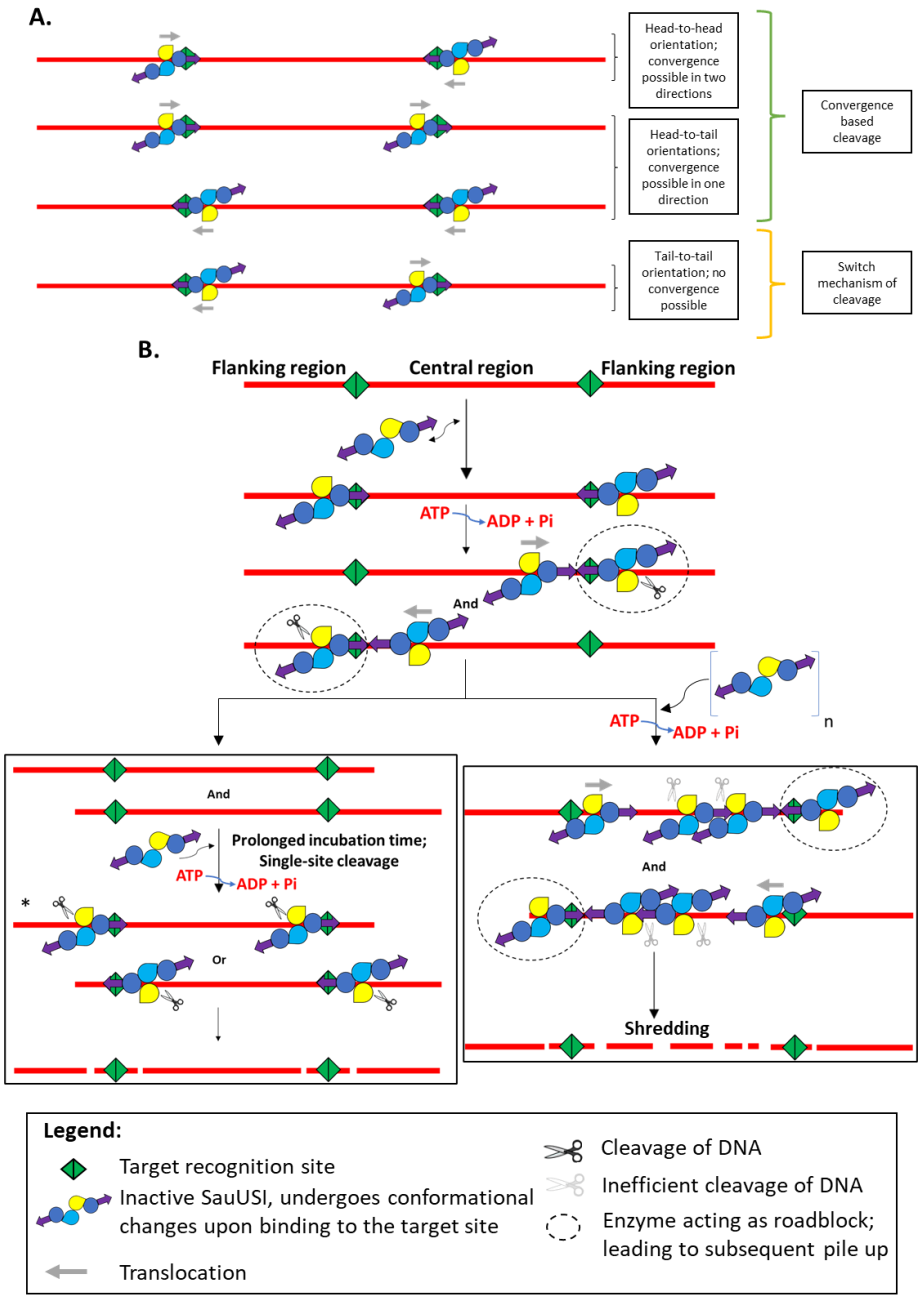


Figure 12: Model for the cleavage by SauUSI: (A) The different orientations that SauUSI can bind when there are two target sites. The head-to-head and head-to-tail orientations lead to possible convergence events leading to DNA cleavage. In the tail-to-tail orientation, however, there is no convergence event possible; therefore, it is cleaved by the switch mechanism (similar to single-site cleavage). Even though the probability of binding to each of these orientations is equal, cleavage seems to occur more rapidly when there is a possibility of enzymes converging. (B) A cartoon representing the model of cleavage of a two-site substrate when SauUSI molecules are bound in the head-to-head orientation. SauUSI is a homodimer with the interface at the nuclease domain (yellow and blue colors representing the nuclease domains). Each monomer has one helicase domain (dark blue circle) and one TRD (purple arrow). The enzyme binds to the DNA (red) possessing the target sites (green) undergoing a conformational change (represented by a slight change in the orientation in the cartoon). Upon ATP hydrolysis, a translocating SauUSI can converge with a stationary SauUSI (roadblock; black dotted circle around the roadblock) bound to another target site. The convergence stalls the translocating enzyme, and DNA is cleaved by the stationary enzyme (black pair of scissors). There are two routes that the cleavage can proceed depending on the concentration of the enzyme. When SauUSI is present in a limited concentration, the preceding convergence event leads to cleavage producing two two-site fragments, these fragments upon prolonged incubation with SauUSI and ATP hydrolysis undergo cleavage due to the switch activity of the ATPase eventually leading to discreet bands. However, when the SauUSI is present in excess, then, there is multiple binding and translocation of SauUSI in either direction leading to a pile-up of enzymes along the central region. This, in turn, leads to stalling of multiple enzymes in the central region. Inefficient nucleolytic activity (grey pair of scissors) of the stalled enzymes leads to random dsDNA breaks, thus shredding the central DNA fragment. Starred steps represent single-site cleavage events.

Based on the location of the nuclease domain with respect to the target site of a Type ISP REase or MmeI bound to DNA, we predict that in the head-to-head arrangement the nuclease domain of the stationary SauUSI will be engaged with the DNA on the flanking region (Figure 12b). As a result, DNA cleavage by the stationary SauUSI will occur close to the target site on the flanking side resulting in two DNA fragments. The resulting fragment from the flanking side will not have a target site. The other fragment will have two target sites, and will also have the stationary and the stalled SauUSI molecules bound to the cleaved end. Successively, another molecule of SauUSI can bind to the free target site of this fragment and can, upon ATP hydrolysis, catalyze DNA cleavage close to the site at the other flanking region.

A surprising discovery that we made was the shredding by SauUSI of a DNA fragment having target sites at both its ends: a feature never observed before in the case of other REases and endonucleases. While shredding of DNA due to multiple nicks by Type I and Type ISP REases has been reported previously, the shredding was limited to the region around the site of dsDNA break (22, 54). In contrast, the shredding by SauUSI was along the entire length of the DNA fragment caused by multiple cuts, which appeared as a smear on the agarose gel. We propose a model for DNA shredding in which a translocating SauUSI starting from a free target site converges and piles-up against the stationary and the stalled SauUSI molecules at another target site. At a higher concentration of the enzyme, more molecules of SauUSI would pile-up. The inefficient nucleolytic activity of the piled-up enzymes can result in random dsDNA break at multiple positions away from the target site, resulting in DNA fragments of varying lengths. The pile-up model, thus, not only explains DNA shredding but also accounts for the observed increase in shredding with increasing enzyme concentration.

We predict a similar mode of cleavage to occur in the case of a head-to-tail arrangement of SauUSI on a two-/multi-site substrate. However, in this arrangement, the convergence of the translocating SauUSI with a stationary SauUSI will result in cleavage close to the target site along the central region side. This will result in two DNA fragments each with one target site, with one site bound to the stalled SauUSI and the other to the stationary SauUSI. The DNA with the stalled SauUSI will have a free target site from where more SauUSI molecules can initiate translocation

and pile-up against the stalled enzyme, especially at high enzyme concentration. As proposed above, the pile-up can result in multiple dsDNA breaks to cause DNA shredding.

In the tail-to-tail arrangement of SauUSI, the target-site-bound enzymes will never converge (Figure 12a). DNA Cleavage in the absence of convergence may be observed because of the enzyme's ability to cleave a single-site substrate. We predict that upon allosteric activation of ATP hydrolysis, the target-site-bound stationary enzyme cuts close to the site resulting in dsDNA break as in the case of a single-site substrate (Figure 12b). DNA shredding will not occur in this case. It is possible that in the case of a single-site substrate or a substrate having target sites in a tail-to-tail orientation where the dsDNA breaks occur close to the target site, the ATPase of SauUSI functions as a switch to activate the nuclease, akin to that observed in the case of Type III REases (23, 24). The additional nicks that we observed away from the target site in the case of the single-site substrate may be a result of a transiently stalled translocating enzyme. In the case of palindromic target sites, the reduced efficiency of single-site cleavage in comparison to the two-site cleavage would mean that a convergence-based cleavage will be more rapid rather than a switch-based event (Figure 9b). Therefore, though the probabilities of SauUSI binding in different orientations are the same, the orientations that facilitate convergence would cleave DNA faster than the orientations that would rely on the switch-based mechanism of cleavage.

Shredding by SauUSI should result in extensive and irreparable damage to the integrity of any DNA recognized as foreign by the enzyme, thus making the enzyme a potent barrier against HGT. A chink in this armor is non-methylation or absence of the SauUSI target site, which will prevent the DNA from being recognized as foreign (13, 18). SauUSI is not unique to *S. aureus*; a search for homologs of SauUSI revealed its presence in many important lineages of bacteria. A phenogram so generated indicated the occurrence of SauUSI-like enzymes in both Gram-positive and Gram-negative bacteria with a major representation of the enzymes in Firmicutes (Figure 13, Table 3). Certain archaea belonging to the phylum Euryarchaeota were also found to have homologs of SauUSI. The presence of the homologs of SauUSI in minimalistic-genome bacteria belonging to the phylum Tenericutes was indicative of the importance of these enzymes as barriers to HGT. The phenogram also suggested to the acquisition of SauUSI-like enzyme by Gram-negative

Proteus cibarius from Gram-positive Firmicutes. The analysis revealed that SauUSI-like enzymes are common, and may function as important regulators of HGT in bacteria.

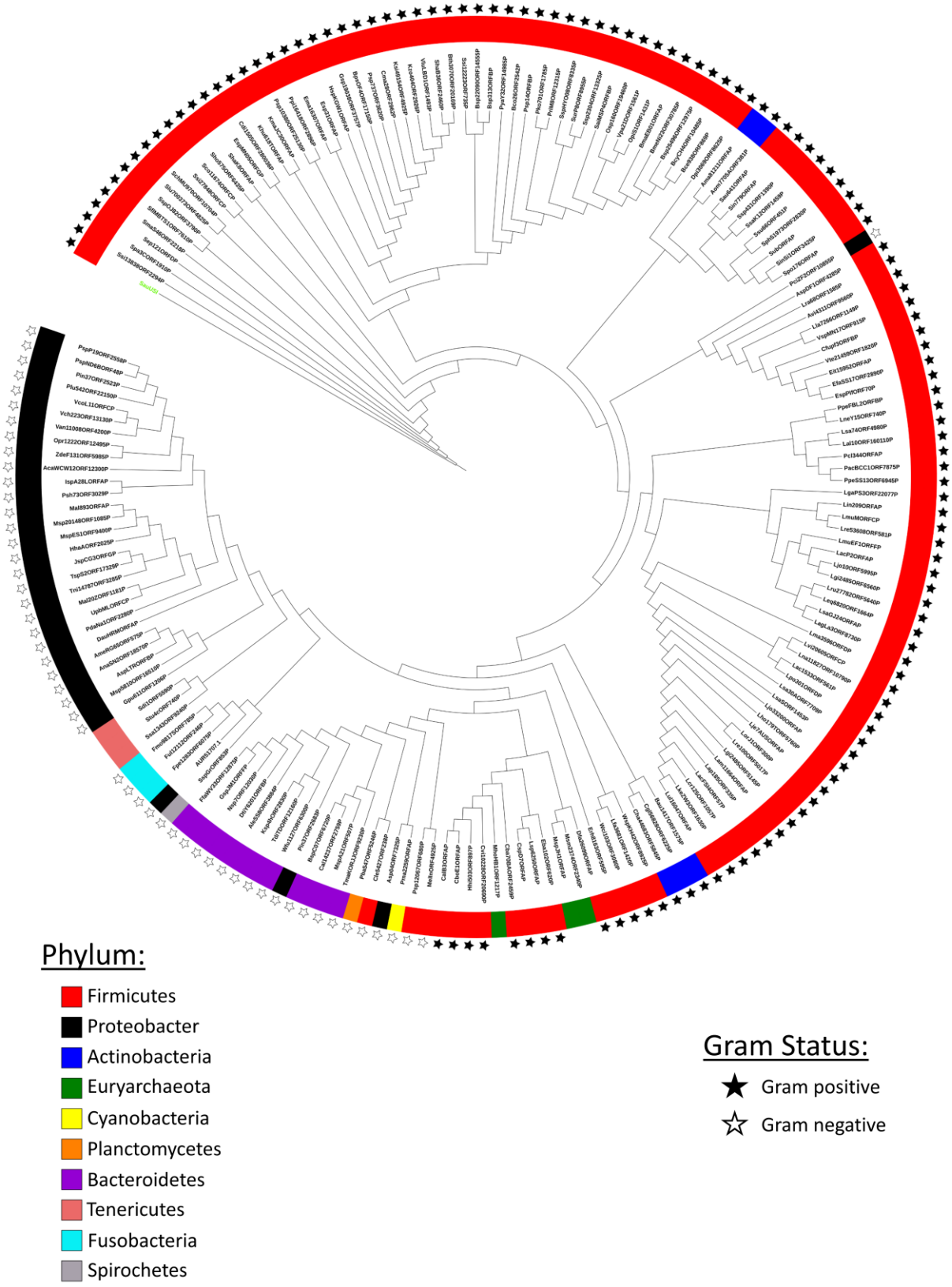


Figure 13: Distribution of SauUSI-like enzymes: A phenogram representing the distribution of SauUSI family of proteins across different bacterial phyla. The color on the strip represents the bacterial phyla, and the star represents if the organism is Gram-positive or negative. No star is an indication of an archaeal species or species with no cell wall. The node highlighted in fluorescent green is SauUSI. The gene and species name represented in the phenogram are given in Table 3.

Table 3: List of ORFs expressing SauUSI-like protein in different bacterial species. Sequences of these proteins were used to generate the phenogram shown in Figure 10.

REase	Organism	Gram Status	Phylum
SauUSI	*Staphylococcus aureus* subsp. *Aureus* USA300_FPR3757	Positive	Firmicutes
Ssi13838ORF2294P	*Staphylococcus simiae* NCTC13838	Positive	Firmicutes
Spa3CORF1910P	*Staphylococcus pasteuri* 3C	Positive	Firmicutes
Sep121ORFDP	*Staphylococcus epidermidis* CDC121	Positive	Firmicutes
SflMBTS1ORF7610P	*Staphylococcus fleurettii* MBTS-1	Positive	Firmicutes
SspOJ82ORF3790P	*Staphylococcus species* OJ82	Positive	Firmicutes
SmaS46ORF2218P	*Staphylococcus massiliensis* S46	Positive	Firmicutes
Slu700373ORF4825P	*Staphylococcus lutrae* ATCC 700373	Positive	Firmicutes
SchMU970ORF10704P	*Staphylococcus chromogenes* MU 970	Positive	Firmicutes
Sco11674ORFCP	*Staphylococcus condimenti*	Positive	Firmicutes
Ssi27848ORFCP	*Staphylococcus simulans*	Positive	Firmicutes
Sho575ORF6435P	*Staphylococcus hominis* FDAARGOS_575	Positive	Firmicutes
ShaK8ORFAP	*Staphylococcus haemolyticus* K8	Positive	Firmicutes
PciZF2ORF10855P	*Proteus cibarius* ZF2	Negative	Proteobacter
AspDF1ORF4285P	*Anthococcus species* DF1	Positive	Firmicutes
Avi4311ORF9560P	*Aerococcus viridans* CCUG4311	Positive	Firmicutes
Lra68ORF1585P	*Lactococcus raffinolactis* WiKim0068	Positive	Firmicutes
Cfupf3ORFBP	*Carnobacterium funditum* pf3	Positive	Firmicutes
Lla7266ORF1149P	*Lactococcus lactis* subsp. *Cremoris* P7266	Positive	Firmicutes
VspMN17ORF915P	*Vagococcus species* MN-17	Positive	Firmicutes
Vte21459ORF1820P	*Vagococcus teuberi*	Positive	Firmicutes
Eit15952ORFAP	*Enterococcus italicus* DSM 15952	Positive	Firmicutes
EfaSS17ORF2890P	*Enterococcus faecalis* S17	Positive	Firmicutes

EspPIfORF70P	*Enterococcus species* RIT-PI-f	Positive	Firmicutes
PpeFBL2ORFBP	*Pediococcus pentosaceus* FBL2	Positive	Firmicutes
LneY15ORF740P	*Lactobacillus nenjiangensis* SH-Y15	Positive	Firmicutes
Lsa74ORF4980P	*Lactobacillus sakei* WiKim0074	Positive	Firmicutes
Lal10ORF160110P	*Lactobacillus algidus* CMTALT10	Positive	Firmicutes
Pcl344ORFAP	*Pediococcus claussenii*	Positive	Firmicutes
PacBCC1ORF7875P	*Pediococcus acidilactici* BCC1	Positive	Firmicutes
PpeSS13ORF6945P	*Pediococcus pentosaceus*	Positive	Firmicutes
Khu618TORFAP	*Kurthia huakuii* LAM0618	Positive	Firmicutes
KmaJC30ORFAP	*Kurthia species* JC30 JC30T	Positive	Firmicutes
EspMN05ORFGP	*Enterococcus species* MN05	Positive	Firmicutes
Cdi1505ORF280036P	*Carnobacterium divergens* MFPA43A1505	Positive	Firmicutes
Ema16307ORFAP	*Exiguobacterium marinum*	Positive	Firmicutes
Esp31ORFAP	*Exiguobacterium* sp. GIC31	Positive	Firmicutes
HspKGW1ORFAP	*Halobacillus* species KGW1	Positive	Firmicutes
Gsp19038ORF3757P	*Geomicrobium* species	Positive	Firmicutes
PyaY32ORF14985P	*Pontibacillus yanchengensis* Y32	Positive	Firmicutes
Psp737ORF3620P	*Paenibacillus* species FSL H7-0737	Positive	Firmicutes
Cma28ORF2962P	*Carnobacterium maltaromaticum* LMA28	Positive	Firmicutes
ShaB36ORF2460P	*Salinicoccus halodurans* H3B36	Positive	Firmicutes
VluLBD1ORF1493P	*Vagococcus lutrae* LBD1	Positive	Firmicutes
Ksi49154ORF4925P	*Kurthia sibirica* ATCC 49154	Positive	Firmicutes
Kzo404ORF2926P	*Kurthia zopfii* NCTC404	Positive	Firmicutes
Bth3070ORF20169P	*Brochothrix thermosphacta* isolate EBP 3070	Positive	Firmicutes
Ssi12223ORF735P	*Solibacillus silvestris*	Positive	Firmicutes
Bsp313ORFBP	*Bacillus* species m3-13	Positive	Firmicutes
Bsp22090ORF14555P	*Bacillus* species FJAT-22090	Positive	Firmicutes

Psp10380ORF25130P	*Paenibacillus* species IHBB 10380	Positive	Firmicutes
Ppi16418ORF2896P	*Paenibacillus pini*	Positive	Firmicutes
BpsOF4ORF17150P	*Bacillus pseudofirmus* OF4	Positive	Firmicutes
Psp14ORFBP	*Paenisporosarcina* species TG-14	Positive	Firmicutes
Pko701ORF1785P	*Planococcus kocurii* HK 701	Positive	Firmicutes
PriM8ORF12315P	*Planococcus rifietoensis* M8	Positive	Firmicutes
Bco26ORF2542P	*Bacillus coagulans* 2-6	Positive	Firmicutes
SspHYO08ORF8335P	*Sporosarcina* species HYO08	Positive	Firmicutes
SurP8ORF8955P	*Sporosarcina ureae* P8	Positive	Firmicutes
Ssp2304ORF13325P	*Sporosarcina* species PTS2304	Positive	Firmicutes
SaiMSP4ORFBP	*Salinibacillus aidingensis* MSP4	Positive	Firmicutes
BmaEB01ORFAP	*Bacillus* species EB01	Positive	Firmicutes
Osp160ORF19460P	*Oceanobacillus* species 160	Positive	Firmicutes
Vpa21DORF1561P	*Virgibacillus pantothenticus* 21D	Positive	Firmicutes
OpiS1ORF1431P	*Oceanobacillus picturae* S1	Positive	Firmicutes
BmeNi23ORF30785P	*Bacillus megaterium* Ni2-3	Positive	Firmicutes
Bsp25496ORF12975P	*Bacillus* species FJAT-25496	Positive	Firmicutes
BcyCH4ORF10485P	*Bacillus cytotoxicus* CH_4	Positive	Firmicutes
Bce938ORF869P	*Bacillus cereus* RIVM_BC938	Positive	Firmicutes
Dpi3069ORF8625P	*Dolosigranulum pigrum* KPL3069	Positive	Firmicutes
SsaK12ORF1459P	*Streptococcus salivarius* K12	Positive	Firmicutes
Ssp431ORF1390P	*Streptococcus* species oral taxon 431	Positive	Firmicutes
Sau641ORFAP	*Streptococcus australis* ATCC 700641	Positive	Firmicutes
Sin779ORFAP	*Streptococcus infantis* ATCC 700779	Positive	Firmicutes
Ama81211ORFAP	*Alloscardovia macacae* UMA81211	Positive	Actinobacteria
Aom7705AORF381P	*Alloscardovia omnicolens* CMW7705A	Positive	Actinobacteria
Ssu66ORF451P	*Streptococcus suis* LSS66	Positive	Firmicutes

Sph51973ORF2830P	*Streptococcus phocae*	Positive	Firmicutes
SubORFAP	*Streptococcus uberis* 0140J	Positive	Firmicutes
SinSi1ORF3425P	*Streptococcus iniae* UEL-Si1	Positive	Firmicutes
Spo176ORFAP	*Streptococcus porcinus* Jelinkova 176	Positive	Firmicutes
Bau1417ORF1575P	*Brevibacterium aurantiacum* SMQ-1417	Positive	Actinobacteria
Cgl56828ORF6225P	*Corynebacterium glutamicum* USDA-ARS-USMARC-56828	Positive	Actinobacteria
Cha44683ORF5845P	*Corynebacterium halotolerans* YIM 70093	Positive	Actinobacteria
WspKH42ORF8925P	*Weissella* species 26KH-42	Positive	Firmicutes
Lfa3681ORF1420P	*Lactobacillus farciminis*	Positive	Firmicutes
Wci103ORF3080P	*Weissella cibaria* DmW_103	Positive	Firmicutes
LgaPS3ORF22077P	*Lactobacillus gastricus* PS3	Positive	Firmicutes
Lru27782ORF5640P	*Lactobacillus ruminis* ATCC 27782	Positive	Firmicutes
Leq6820ORF1664P	*Lactobacillus equi* DPC 6820	Positive	Firmicutes
LsaGJ24ORFAP	*Lactobacillus salivarius* GJ-24	Positive	Firmicutes
LagLa3ORF8730P	*Lactobacillus agilis* La3	Positive	Firmicutes
Lin209ORFAP	*Lactobacillus ingluviei* Autruche 4	Positive	Firmicutes
LmuMORFCP	*Lactobacillus mucosae* LM1	Positive	Firmicutes
Lre53608ORF581P	*Lactobacillus reuteri* ATCC 53608	Positive	Firmicutes
LmuEF1ORFFP	*Lactobacillus murinus* EF-1	Positive	Firmicutes
LacP2ORFAP	*Lactobacillus acidophilus* P2	Positive	Firmicutes
Ljo10ORF5995P	*Lactobacillus johnsonii* ZLJ010	Positive	Firmicutes
Lgi2485ORF6560P	*Lactobacillus gigeriorum* CRBIP 24.85	Positive	Firmicutes
Lac1533ORF561P	*Lactobacillus acidipiscis* ACA-DC 1533	Positive	Firmicutes
Lpo301ORFDP	*Lactobacillus pobuzihii* E100301	Positive	Firmicutes
Lma3596ORFDP	*Lactobacillus mali* KCTC 3596	Positive	Firmicutes
Lvi20605ORFCP	*Lactobacillus vini*	Positive	Firmicutes
Lna11827ORF10790P	*Lactobacillus nagelii* TMW 1.1827	Positive	Firmicutes

Lsa30AORF7709P	*Lactobacillus saerimneri* 30a	Positive	Firmicutes
LsaSORF1453P	*Lactobacillus salivarius* UCC118	Positive	Firmicutes
Lje7AUSORFAP	*Lactobacillus jensenii* SJ-7A-US	Positive	Firmicutes
Lho179TORF5760P	*Lactobacillus hominis* CRBIP 24.179	Positive	Firmicutes
Ljo33200ORFAP	*Lactobacillus johnsonii* ATCC 33200	Positive	Firmicutes
LorJ1ORF300P	*Lactobacillus oris* J-1	Positive	Firmicutes
Lre100ORF5017P	*Lactobacillus reuteri* 100-23	Positive	Firmicutes
Lgi2485ORF5145P	*Lactobacillus gigeriorum* CRBIP 24.85	Positive	Firmicutes
Lam11664ORFAP	*Lactobacillus amylolyticus* DSM 11664	Positive	Firmicutes
Lap185ORF335P	*Lactobacillus apis* ESL0185	Positive	Firmicutes
LacFSI4ORF57P	*Lactobacillus acidophilus* FSI4	Positive	Firmicutes
Lcr125ORF1057P	*Lactobacillus crispatus* 125-2-CHN	Positive	Firmicutes
Lul16047ORFAP	*Lactobacillus ultunensis* DSM 16047	Positive	Firmicutes
LkeZW3ORF1630P	*Lactobacillus kefiranofaciens* ZW3	Positive	Firmicutes
Erh8163ORF595P	*Erysipelothrix rhusiopathiae* NCTC8163	Positive	Firmicutes
Dfa26099ORFAP	*Dielma fastidiosa* strain JC13	Positive	Firmicutes
Msm2374ORF2340P	*Methanobrevibacter smithii* DSM 2374	Archaea	Euryachaeota
MspJH1ORFAP	*Methanobrevibacter* species JH1	Archaea	Euryachaeota
Eba102ORF620P	*Erysipelotrichaceae* bacterium SG0102	Positive	Firmicutes
Lsp6250ORFAP	*Longibaculum* species KGMB06250	Positive	Firmicutes
CspD7ORFAP	*Coprobacillus* species D7	Positive	Firmicutes
Cba70BAORF2459P	*Clostridiales bacterium* 70B-A	Positive	Firmicutes
MhoHB1ORF1217P	*Methanosarcina horonobensis* HB-1	Archaea	Euryachaeota
Cvi10228ORF20690P	*Clostridium vincentii* DSM 10228	Positive	Firmicutes
Hhi503ORF897P	*Hathewaya histolytica* NCTC503	Positive	Firmicutes
CboE1ORFAP	*Clostridium botulinum* E1 BoNT E Beluga	Positive	Firmicutes
CalB3ORFAP	*Clostridium algidicarnis* strain B3	Positive	Firmicutes

MelInORF4925P	*Megasphaera elsdenii indica*	Negative	Firmicutes
Psp12067ORF680P	*Phascolarctobacterium* species YIT 12067	Negative	Firmicutes
Pma2259ORFAP	*Prochlorococcus marinus* clone ASNC2259	Negative	Cyanobacteria
Asp04ORF7325P	*Acinetobacter* species TTH0-4	Negative	Proteobacter
Cle5427ORF238P	*Clostridium lentocellum*	Negative	Firmicutes
Pba547ORF5246P	*Planctomycetes bacterium* SRT547	Negative	Planctomycetes
Pin37ORF2683P	*Psychromonas ingrahamii* 37	Negative	Proteobacter
TmaKORJJORF9230P	*Tenacibaculum maritimum* TM-KORJJ	Negative	Bacteroidetes
MspA21ORF507P	*Myroides* species A21	Negative	Bacteroidetes
AleSS8ORF3884P	*Algibacter lectus*	Negative	Bacteroidetes
Cal14237ORF2759P	*Cellulophaga algicola* DSM 14237	Negative	Bacteroidetes
Ksp4hORF2830P	*Krokinobacter* species 4H-3-7-5	Negative	Bacteroidetes
TdiTDORF12160P	*Tenacibaculum dicentrarchi* AY7486TD	Negative	Bacteroidetes
Wfu1127ORF6300P	*Wenyingzhuangia fucanilytica* CZ1127	Negative	Bacteroidetes
BspC07ORF6720P	*Balneola* species EhC07	Negative	Bacteroidetes
DtiY6201ORFBP	*Dyadobacter tibetensis* Y620-1	Negative	Bacteroidetes
Nsp7ORF12030P	*Nonlabens* species MB-3u-79	Negative	Bacteroidetes
GspJM1ORFFP	*Gillisia* sp. JM1	Negative	Bacteroidetes
FfaWV33ORF12875P	*Flavobacterium faecale* WV33	Negative	Bacteroidetes
Ssa1343ORF9240P	*Spiroplasma sabaudiense* Ar-1343	Cell wall less	Tenericutes
Stu4cORF740P	*Spiroplasma turonicum* Tab4c	Cell wall less	Tenericutes
Sdi1ORF5590P	*Spiroplasma diminutum* CUAS-1	Cell wall less	Tenericutes
DauHRMORFAP	*Desulfobacterium autotrophicum* HRM2	Negative	Proteobacter
AcaWCW12ORF12300P	*Aeromonas caviae* WCW1-2	Negative	Proteobacter
ZdeF131ORF5985P	*Zobellella denitrificans* F13-1	Negative	Proteobacter
Opr1222ORF12495P	*Oceanisphaera profunda* SM1222	Negative	Proteobacter
Van11008ORF4200P	*Vibrio anguillarum* CNEVA NB11008	Negative	Proteobacter

Plu542ORF22150P	*Pseudoalteromonas luteoviolacea* S40542	Negative	Proteobacter
Vch223ORF13130P	*Vibrio cholerae* FDAARGOS_223	Negative	Proteobacter
VcoL11ORFCP	*Vibrio cholerae* L11	Negative	Proteobacter
Pin37ORF2523P	*Psychromonas ingrahamii* 37	Negative	Proteobacter
PspND6BORF48P	*Pseudoalteromonas* species ND6B	Negative	Proteobacter
PspP19ORF2558P	*Pseudoalteromonas* species P1-9	Negative	Proteobacter
Gpu611ORF1206P	*Glaciecola punicea* DSM 14233	Negative	Proteobacter
Msp5810ORF16510P	*Marinobacterium* species ST58-10	Negative	Proteobacter
AspLTRORFBP	*Alteromonas* species LTR	Negative	Proteobacter
AmeRG65ORF575P	*Alteromonas mediterranea* RG65	Negative	Proteobacter
AnaSN2ORF18570P	*Alteromonas naphthalenivorans* SN2	Negative	Proteobacter
Psh73ORF3029P	*Plesiomonas shigelloides* 302-73	Negative	Proteobacter
IspA28LORFAP	*Idiomarina* species A28L	Negative	Proteobacter
PdaNa1ORF2280P	*Photobacterium damselae* subsp. *Damselae* KC-Na-1	Negative	Proteobacter
UpbMLORFCP	Unknown delta proteobacterium MLMS-1	Negative	Proteobacter
Mal20ZORF1181P	*Methylomicrobium alcaliphilum* 20Z	Negative	Proteobacter
Tni14787ORF3285P	*Thioalkalivibrio nitratireducens*	Negative	Proteobacter
TspS2ORF17329P	*Thauera aminoaromatica* S2	Negative	Proteobacter
JspCG3ORFGP	*Janthinobacterium* species CG3	Negative	Proteobacter
HhaAORF2025P	*Halorhodospira halochloris* A	Negative	Proteobacter
MspES1ORF9400P	*Marinobacter* species ES-1	Negative	Proteobacter
Mal893ORFAP	*Marinobacter algicola* DG893	Negative	Proteobacter
Msp20148ORF1085P	*Marinobacter* species BSs20148	Negative	Proteobacter
SspGrORF853P	*Spirochaeta* species Grapes	Negative	Spirochetes
AUR51707.1 NgoFVII family restriction endonuclease [Neisseriaceae bacterium DSM 100970]	*Neisseria*	Negative	Proteobacter

Fmo9817SORF785P	*Fusobacterium mortiferum* ATCC 9817S	Negative	Fusobacteria
Ful12112ORF246P	*Fusobacterium ulcerans* NCTC12112	Negative	Fusobacteria
Fpe1283ORF6075P	*Fusobacterium periodonticum* KCOM 1283	Negative	Fusobacteria

References

1. Hernando-Amado,S., Coque,T.M., Baquero,F. and Martínez,J.L. (2019) Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nat. Microbiol.*, **4**, 1432–1442.

2. Lee,A.S., De Lencastre,H., Garau,J., Kluytmans,J., Malhotra-Kumar,S., Peschel,A. and Harbarth,S. (2018) Methicillin-resistant Staphylococcus aureus. *Nat. Rev. Dis. Prim.*, **4**, 1–23.

3. Hamilton,F. and MacGowan,A. (2019) A long history of β-lactams for MRSA. *Nat. Microbiol. news views*, **4**, 1604–1605.

4. Giesbrecht,P., Kersten,T., Maidhof,H. and Wecke,J. (1998) Staphylococcal Cell Wall: Morphogenesis and Fatal Variations in the Presence of Penicillin. *Microbiol. Mol. Biol. Rev.*, **62**, 1371–1414.

5. Haaber,J., Leisner,J.J., Cohn,M.T., Catalan-Moreno,A., Nielsen,J.B., Westh,H., Penadés,J.R. and Ingmer,H. (2016) Bacterial viruses enable their host to acquire antibiotic resistance genes from neighbouring cells. *Nat. Commun.*, **7**.

6. Arthur,M., Molinas,C., Depardieu,F. and Courvalin,P. (1993) Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in Enterococcus faecium BM4147. *J. Bacteriol.*, **175**, 117–127.

7. Thomas,C.M. and Nielsen,K.M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.*, **3**, 711–721.

8. Smets,B.F. and Barkay,T. (2005) Horizontal gene transfer: Perspectives at a crossroads of scientific disciplines. *Nat. Rev. Microbiol.*, **3**, 675–678.

9. Arber,W. and Linn,S. (1969) DNA Modification and Restriction. *Annu. Rev. Biochem.*, **38**, 467–500.

10. Bickle,T.A. and Kruger,D.H. (1993) Biology of DNA restriction. *Microbiol. Rev.*, **57**, 434–450.

11. Corvaglia,A.R., François,P., Hernandez,D., Perron,K., Linder,P. and Schrenzel,J. (2010) A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical Staphylococcus aureus strains. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 11954–8.

12. Xu,S.Y., Corvaglia,A.R., Chan,S.H., Zheng,Y. and Linder,P. (2011) A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300. *Nucleic Acids Res.*, **39**, 5597–5610.

13. Monk,I.R., Shah,I.M. and Xu,M. (2012) Transforming the Untransformable : Application of Direct Transformation To manipulate Genetically Staphylococcus and Staphylococcus epidermidis. *MBio*, **3**, e00277-11.

14. Cooper,L.P., Roberts,G.A., White,J.H., Luyten,Y.A., Bower,E.K.M., Morgan,R.D., Roberts,R.J., Lindsay,J.A. and Dryden,D.T.F. (2017) DNA target recognition domains in the Type i restriction and modification systems of Staphylococcus aureus. *Nucleic Acids Res.*, **45**, 3395–3406.

15. Waldron,D.E. and Lindsay,J.A. (2006) Sau1: A novel lineage-specific type I restriction-modification

system that blocks horizontal gene transfer into Staphylococcus aureus and between S. aureus isolates of different lineages. *J. Bacteriol.*, **188**, 5578–5585.

16. Loenen,W.A.M., Dryden,D.T.F., Raleigh,E.A., Wilson,G.G. and Murrayy,N.E. (2014) Highlights of the DNA cutters: A short history of the restriction enzymes. *Nucleic Acids Res.*, **42**, 3–19.

17. Jones,M.J., Donegan,N.P., Mikheyeva,I. V. and Cheung,A.L. (2015) Improving transformation of Staphylococcus aureus belonging to the CC1, CC5 and CC8 clonal complexes. *PLoS One*, **10**, 1–14.

18. Johnston,C.D., Cotton,S.L., Rittling,S.R., Starr,J.R., Borisy,G.G., Dewhirst,F.E. and Lemon,K.P. (2019) Systematic evasion of the restriction-modification barrier in bacteria. *Proc. Natl. Acad. Sci. U. S. A.*, **166**, 11454–11459.

19. Veiga,H. and Pinho,M.G. (2009) Inactivation of the saul type i restriction-modification system is not sufficient to generate Staphylococcus aureus strains capable of efficiently accepting foreign DNA. *Appl. Environ. Microbiol.*, **75**, 3034–3038.

20. Dryden,D.T.F., Murray,N.E. and Rao,D.N. (2001) Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Res.*, **29**, 3728-3741.

21. Rao,D.N., Dryden,D.T.F. and Bheemanaik,S. (2014) Type III restriction-modification enzymes : a historical perspective. *Nucleic Acids Res.*, **42**, 45–55.

22. Chand,M.K., Nirwan,N., Diffin,F.M., Aelst,K. Van, Kulkarni,M., Pernstich,C., Szczelkun,M.D. and Saikrishnan,K. (2016) Translocation-coupled DNA cleavage by the Type ISP restriction-modification enzymes. *Nat. Chem. Biol.*, **11**, 870–877.

23. Ahmad,I., Kulkarni,M., Gopinath,A. and Saikrishnan,K. (2018) Single-site DNA cleavage by Type III restriction endonuclease requires a site-bound enzyme and a trans-acting enzyme that are ATPase-activated. *Nucleic Acids Res.*, **46**, 6229–6237.

24. Friedrich W. Schwarz, Júlia Tóth, Kara van Aelst, Guanshen Cui, Sylvia Clausing,M.D.S. and R.S. (2013) The Helicase-Like Domains of Type III Restriction Enzymes Trigger Long-Range Diffusion Along DNA. *Science.*, **340**, 353–357.

25. Kabsch,W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 125–132.

26. Evans,P.R. and Murshudov,G.N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **69**, 1204–1214.

27. Terwilliger,T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **56**, 965–972.

28. Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **55**, 849–861.

29. Adams,P.D., Afonine,P. V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W., *et al.* (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 213–221.

30. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta*

Crystallogr. Sect. D Biol. Crystallogr., **66**, 486–501.

31. Afonine,P. V., Mustyakimov,M., Grosse-Kunstleve,R.W., Moriarty,N.W., Langan,P. and Adams,P.D. (2010) Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 1153–1163.

32. Letunic,I. and Bork,P. (2019) Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.*, **47**, 256–259.

33. Pingoud,A. and Jeltsch,A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.*, **29**, 3705–3727.

34. Christ,F. (1999) The monomeric homing endonuclease PI-SceI has two catalytic centres for cleavage of the two strands of its DNA substrate. *EMBO J.*, **18**, 6908–6916.

35. Lagunavicius,A., Sasnauskas,G., Halford,S.E. and Siksnys,V. (2003) The metal-independent type IIs restriction enzyme BfiI is a dimer that binds two DNA sites but has only one catalytic centre. *J. Mol. Biol.*, **326**, 1051–1064.

36. Bitinaite,J., Wah,D.A., Aggarwal,A.K. and Schildkraut,I. (1998) Foki dimerization is required for dna cleavage. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 10570–10575.

37. Studier,W.F. and Bandyopadhyay,P.K. (2005) Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **85**, 1–5.

38. Pingoud,A., Wilson,G.G. and Wende,W. (2014) Type II restriction endonucleases - A historical perspective and more. *Nucleic Acids Res.*, **42**, 7489–7527.

39. Grazulis,S., Manakova,E., Roessle,M., Bochtler,M., Tamulaitiene,G., Huber,R. and Siksnys,V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 15797–15802.

40. Sasnauskas,G., Zakrys,L., Zaremba,M., Cosstick,R., Gaynor,J.W., Halford,S.E. and Siksnys,V. (2010) A novel mechanism for the scission of double-stranded DNA: BfiI cuts both 3′-5′ and 5′-3′ strands by rotating a single active site. *Nucleic Acids Res.*, **38**, 2399–2410.

41. Zaremba,M., Toliusis,P., Grigaitis,R., Manakova,E., Silanskas,A., Tamulaitiene,G., Szczelkun,M.D. and Siksnys,V. (2014) DNA cleavage by CgII and NgoAVII requires interaction between N- and R-proteins and extensive nucleotide hydrolysis. *Nucleic Acids Res.*, **42**, 13887–13896.

42. Holm,L. (2020) DALI and the persistence of protein shape. *Protein Sci.*, **29**, 128–140.

43. Li,X., Jake Harris,C., Zhong,Z., Chen,W., Liu,R., Jia,B., Wang,Z., Li,S., Jacobsen,S.E. and Du,J. (2018) Mechanistic insights into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA methylation. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, E8793–E8802.

44. Rajakumara,E., Nakarakanti,N.K., Nivya,M.A. and Satish,M. (2016) Mechanistic insights into the recognition of 5-methylcytosine oxidation derivatives by the SUVH5 SRA domain. *Sci. Rep.*, **6**, 1–11.

45. Uyen,N.T., Park,S.Y., Choi,J.W., Lee,H.J., Nishi,K. and Kim,J.S. (2009) The fragment structure of a

putative HsdR subunit of a type I restriction enzyme from Vibrio vulnificus YJ016: Implications for DNA restriction and translocation activity. *Nucleic Acids Res.*, **37**, 6960–6969.

46. Chand,M.K., Carle,V., Anuvind,K.G. and Saikrishnan,K. (2020) DNA-mediated coupling of ATPase , translocase and nuclease activities of a Type ISP restriction-modification enzyme. **48**, 2594–2603.

47. Firman,K. and Szczelkun,M.D. (2000) Measuring motion on DNA by the type I restriction endonuclease EcoR124I using triplex displacement. *EMBO J.*, **19**, 2094–2102.

48. Van Aelst,K., Tóth,J., Ramanathan,S.P., Schwarz,F.W., Seidel,R. and Szczelkun,M.D. (2010) Type III restriction enzymes cleave DNA by long-range interaction between sites in both head-to-head and tail-to-tail inverted repeat. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 9123–9128.

49. Callahan,S.J., Luyten,Y.A., Gupta,Y.K., Wilson,G.G., Roberts,R.J., Morgan,R.D. and Aggarwal,A.K. (2016) Structure of Type IIL Restriction-Modification Enzyme MmeI in Complex with DNA Has Implications for Engineering New Specificities. *PLoS Biol.*, **14**, 1–18.

50. Rajakumara,E., Law,J.A., Simanshu,D.K., Voigt,P., Johnson,L.M., Reinberg,D., Patel,D.J. and Jacobsen,S.E. (2011) A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo. *Genes Dev.*, **3**, 137–152.

51. Hashimoto,H., Horton,J.R., Zhang,X., Bostick,M., Jacobsen,S.E. and Cheng,X. (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature*, **455**, 826–829.

52. Jindrova,E., Schmid-Nuoffer,S., Hamburger,F., Janscak,P. and Bickle,T.A. (2005) On the DNA cleavage mechanism of Type I restriction enzymes. *Nucleic Acids Res.*, **33**, 1760–1766.

53. Janscak,P., MacWilliams,M.P., Sandmeier,U., Nagaraja,V. and Bickle,T.A. (1999) DNA translocation blockage, a general mechanism of cleavage site selection by type I restriction enzymes. *EMBO J.*, **18**, 2638–2647.

54. Endlichl,B. and Linn,S. (1985) The DNA Restriction Endonuclease of Escherichia coli B. *J. Biol. Chem.*, **260**, 5729–5738.

CHAPTER 4: Characterizing a novel ssDNA nuclease activity of a Type IV restriction endonuclease SauUSI

Abstract

Restriction endonucleases cleave exogenous DNA thus restricting horizontal gene transfer and phage infection of host bacterium. This nucleolytic activity occurs on double-stranded DNA (dsDNA) and is target site specific. Here we report that the Type IV ATP-dependent restriction endonuclease SauUSI from *Staphylococcus aureus* also possesses a hitherto unknown single-stranded DNA (ssDNA) endonuclease activity. We demonstrate that, unlike the dsDNA cleavage activity, ssDNA cleavage by SauUSI does not require divalent cation or ATP hydrolysis and is target-site and DNA methylation-status independent. Furthermore, we show that SauUSI can cut ssDNA gaps, overhangs, bubbles and loops but not ssRNA. The activity is inhibited at higher concentrations of magnesium ion, ATP, and the presence of single strand DNA binding protein. The ssDNA nuclease activity is thus tightly regulated and may protect the host DNA from damage by SauUSI.

Introduction

Nucleases are enzymes that hydrolyze the phosphodiester bond of nucleic acid backbones and are ubiquitous in all forms of life. They play a pivotal role in various key physiological processes such as replication, transcription and DNA repair apart from regulating transfer of genetic material in the bacterial world (1–6). Foreign genetic material such as plasmids, fragments of double-stranded DNA (dsDNA), double-stranded RNA (dsRNA), single-stranded DNA (ssDNA) or single-stranded RNA (ssRNA) are usually taken up by bacteria by either transformation, transduction or conjugation. More often than not, uptake of foreign DNA can prove detrimental to bacteria, and hence a variety of defense mechanisms to counter the potential catastrophic effects of horizontal gene transfer have evolved. Nucleases that regulate the transfer of genetic material are primarily part of the defense mechanisms of bacteria, such as restriction-modification (RM) systems, clustered regularly interspaced short palindromic repeats associated Cas (CRISPR-Cas) systems and non-specific endo/exonucleases.

RM systems are the most prominent bacterial defense system found in ˜75 % of the sequenced bacterial genomes (7). In comparison, CRISPR-Cas systems are found in ~40 % of the sequenced bacterial genomes. RM systems are usually composed of two components that act in tandem. The restriction component selectively cleaves foreign DNA and the modification component protects the host DNA from the restriction activity by methylating it at cognate sites. Restriction endonucleases discovered thus far are known to cleave only dsDNA and there has been no known example of a *bona fide* ssDNA endonuclease reported. Though, during the early days of research on restriction endonucleases, there were exciting reports of restriction endonucleases cleaving ssDNA (8–15), they could be explained as arising from cleavage of transient dsDNA formed by the ssDNA (8, 10–12, 16). Here, we demonstrate that the restriction endonuclease SauUSI from *Staphylococcus aureus*, which is a sequence-specific dsDNA shredder (17) , is also a *bona fide* non-specific ssDNA endonuclease.

Previous studies had characterized SauUSI as a modification-dependent Type IV REase that recognizes the target-site 5’-S^{5m}CNGS-3’ (S=C/G and N= A/T/G/C) and cleaves dsDNA multiple base pairs away from the target site (17, 18). SauUSI contains a N-terminal PLD nuclease, a SF2 helicase-like ATPase and a C-terminal SRA target recognition domain. Detailed biochemical and

structural studies revealed that it functions as a stable dimer in solution with the interface formed by the nuclease domains. Part of its impregnability as a barrier to horizontal gene transfer stems from its unique mechanism of cleavage wherein DNA present between two target sites is shred into multiple fragments. This shredding of DNA is brought about by the enzyme's ability to translocate on dsDNA fueled by the hydrolysis of ATP. In addition to this, SauUSI also cleaves dsDNA with a single target site in an ATP-dependent manner albeit at an efficiency lower than substrates having multiple target-sites.

One of the ways methicillin-resistant strains of *S. aureus* are converted to vancomycin-resistant ones is attributed to inactivation of the *sauUSI* gene which eventually leads to a manifold increase in horizontal gene transfer efficiency (2, 19). The inactivation of the *sauUSI* gene is hypothesized to occur when the bacterium is challenged by antibiotics where the only means of survival is to take up foreign DNA either from other bacteria (conjugation) or from the environment (natural transformation) (20–23). Presently, vancomycin and combinatorial antibiotics are the last line of defense against methicillin resistant *S. aureus* (MRSA). Hence, the emergence of vancomycin-resistant *S. aureus* (VRSA) is a burden on our present medical infrastructure and understanding the ways in which bacteria evolve antibiotic resistance is of paramount importance to devise novel ways to combat these pathogens.

In this study, we show for the first time that a restriction endonuclease, SauUSI, has the ability to cleave ssDNA, which is DNA sequence-independent and does not require ATP. We also show that ssDNA as short as 20 base long to very long ones, such as M13mp18 ssDNA, are efficiently cleaved my SauUSI in the absence of nucleotide and without any sequence specificity. We further demonstrate that SauUSI cleaves ssDNA endonucleolytically without a preference for either a 5' or 3' free end. We also find that magnesium ion (Mg^{2+}), ATP and single strand DNA binding protein (SSB) affect the ssDNA endonucleolytic activity of SauUSI, which may lend protection to the host DNA from the toxicity of the enzyme. This novel view on SauUSI, that is, its ability to cleave both dsDNA and ssDNA, could have major implications in determining the efficiency of HGT. Bacterial defense mechanisms like the CRISPR-Cas system have been shown to modulate the transfer of antibiotic resistance genes in bacterial populations (1, 3–5).

Materials and methods

Purifications of SauUSI

SauUSI and the nuclease-inactive mutant (SauUSI ^{H119A}) were purified using the protocol described earlier (17, 24). Both the *sauUSI* and *sauUSI* ^{H119A} were cloned into a *pHis17* expressed vector containing a C-terminal 6x His-Tag. A three-column purification strategy: Ni-NTA affinity chromatography, Anion exchange chromatography and size exclusion chromatography, at 4^{o} C was employed to obtain the protein to a high degree of homogeneity.

DNA Substrates

The sequence of substrates used in this study is mentioned in supplementary table. The dsDNA substrates used in this study were annealed using commercially synthesized ssDNA oligos. These oligos were annealed by sequentially decreasing the temperature from 95^{o}C to 25^{o}C in steps of 1^{o}C at an interval of 1 minute in each step. The substrates were run on a 12% Native PAGE to test the annealing efficiency against individual ssDNA oligos. The ssRNA used in the study was commercially synthesized. The canonical cleavage assays for long dsDNA substrates were generated either by PCR amplification (EcoP15I) or by linearizing plasmids (SauUSI, LlaBIII and BstNI).

Cleavage assays

Initial cleavage assays to test the activity of the enzyme were performed by incubating 250 nM DNA substrate and 200 nM SauUSI with 2 mM ATP in the NEBuffer™ 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate and 1 mM DTT pH 7.9) at 37^{o} C for 45 minutes. Temperature and incubation time were kept constant for all the assays (unless specified in the results section). Assays that were conducted in the presence of divalent cation chelator were performed in the presence of 16 mM EDTA. Assays that were conducted in the absence of magnesium acetate used a buffer containing 50 mM potassium acetate, 20 mM Tris-acetate pH 7.9 and 1 mM DTT. Subsequent characterization of cleavage assays were performed by keeping the concentration of SauUSI at 50 nM. The reactions were stopped by adding 1X Gel Loading Dye Purple followed by heating at 65 ^{o}C for 20 minutes. The cleavage fragments were resolved on a 12% native PAGE and visualized on a GE Amersham Typhoon gel scanner after staining with 1x SyBr gold staining solution. Some of the initial assays were visualized on a G-box UV imager. The

cleavage assays using different nucleotides (ATP, GTP, CTP, UTP, dATP, ADP and AMP) were performed using the NEBuffer™ 4. The combination cleavage assays were performed with 125 nM 60 bp dsDNA and 250 nM 36-base long ssDNA in NEBuffer™ 4. RNA cleavage assay was performed by incubating 250 nM ssRNA with a concentration gradient of SauUSI in NEBuffer™ 4 at 37^{o} C for 30 minutes. The reaction was stopped by adding 1X formamide loading buffer and heating at 95^{o} C for 10 minutes. The cleavage fragments were resolved on a 15% urea denaturing PAGE and visualized on a GE Amersham Typhoon gel scanner. Densitometric calculations for all the cleavage assays were performed using ImageJ software and plotted using GraphPad Prism.

ATP-dependent Cleavage assays

The cleavage assay using long dsDNA to test the activity of SauUSI, EcoP15I, LlaBIII and BstNI were performed as described earlier(17, 25, 26). 100 nM of SauUSI was incubated with DNA at 37^{o}C for 45 minutes. 100 nM of LlaBIII or EcoP15I were incubated with DNA and 2 mM ATP at 25^{o} C. BstNI was incubated with DNA at 60^{o}C. All the reactions were stopped using 1X Gel Loading Dye Purple followed by heating at 65 ^{o}C for 20 minutes.

DNA binding study

The DNA used for binding study is a 60 base ssDNA. 250 nM of ssDNA was incubated with SauUSI^{H119A} on ice for 20 minutes. EMSA was carried out on a 5% Native PAGE and the samples were loaded using a 1X STB (sucrose, Tris and bropmophenol blue) loading buffer. The gel was stained using a 1X SyBr gold staining and visualized on a GE Amersham Typhoon gel scanner.

Cleavage assay using M13mp18 ssDNA

100 nanograms of NEB M13mp18 ssDNA which corresponds to ~882 nM in a 10 microlitre reaction was incubated with 220 nM, 441 nM and 882 nM of *E. coli* SSB (single-stranded DNA binding protein) to form a SSB:ssDNA molar ratio (R) of 0.25, 0.5 and 1 respectively in a low-salt buffer (2 mM NaCl, 10 mM Tris pH 8 and 0.1 mM EDTA) at 25^{o} C for 2 hours (27, 28). The molarity of M13mp18 ssDNA was calculated assuming a $(SSB)_{35}$ mode of binding and that of *E. coli* SSB was calculated using its tetramer quaternary assembly.

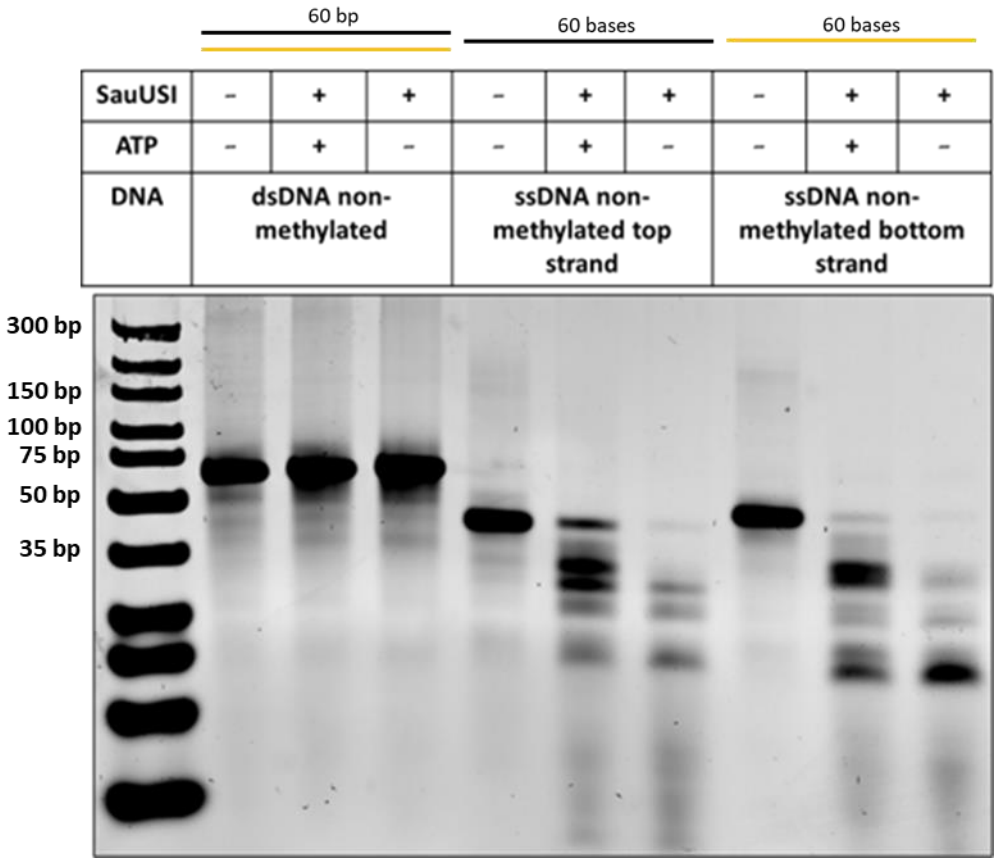


Figure 1: ssDNA cleavage is target site independent and can occur without nucleotide stimulation: Schematic of the substrates used for the cleavage assay. Representative 12% Native PAGE gel to study the cleavage assay for dsDNA (non-methylated) and both the ssDNAs (non-methylated top/bottom strands). The dsDNA substrate was generated by annealing the two ssDNAs using a temperature gradient protocol. dsDNA (non-methylated DNA: lane 2 – 4), ssDNA (non-methylated top strand: lane 5 – 7) and ssDNA (non-methylated bottom strand: lane 8 – 10) were treated with SauUSI either in the presence or absence of ATP and the cleavage pattern was compared with ds/ss DNA controls and run in the presence of ultra-low range DNA ladder.

The reactions were electrophoresed on a 0.5% agarose gel equilibrated with ethidium bromide in 0.5X TAE buffer and visualized on a E-gel Imager. Following ssDNA-SSB complex formation, 50 nM of SauUSI was added and incubated at 25^{o} C for 30 minutes and the reactions were visualized on a 0.5 % Agarose gel equilibrated with ethidium bromide which was electrophoresed as mentioned earlier.

Results

SauUSI cleaves ssDNA in an ATP- and target-site independent manner

Previous studies on SauUSI revealed that it has the ability to specifically cleave methylated DNA containing a target-site in an ATP-dependent manner and DNA which is non-methylated is not a substrate for cleavage. Non-methylated dsDNA is refractory to cleavage by SauUSI owing to its inability to stimulate the ATPase activity of the SF2 helicase-like ATPase domain which tightly regulates the nuclease activity. This was exemplified when a 60 bp non-methylated dsDNA was subjected to cleavage by 100 nM SauUSI in the presence of ATP (Figure 1, Lane 2, 3 and 4). As expected, the non-methylated DNA was not cleaved by SauUSI either in the presence or absence of ATP. The 60 bp dsDNA substrate was an annealed product of two 60 bases ssDNA. Interestingly, on treatment of two DNA single strands, which had complimentary sequences, with 100 nM SauUSI, 2 mM ATP and 10 mM magnesium acetate cleavage was observed on a 12 % native-PAGE gel (Figure 1, Lanes 5 - 10). Multiple fragments of ssDNA were produced, which pointed towards the cleavage being sequence-independent (Figure 1, Lane 6 and 9). The efficiency of cleavage seemed to improve in the absence of ATP (Figure 1, Lane 7 and 10).

ssDNA cleavage: a unique feature of SauUSI, inactivated by nuclease active site mutation

The N-terminal nuclease domain of SauUSI belongs to the Phospholipase-D (PLD) family of protein fold. A characteristic feature of such nucleases is the presence of a strong dimeric interface which brings two 'HxK' motifs from each protomer in close proximity to each other to form the enzymatic active site. The Histidine side chains from both the protomers act as nucleophiles (29, 30). The histidine in close proximity to the DNA phosphate backbone causes a nucleophilic attack forming a protein-DNA adduct which is resolved by a water molecule whose activation is facilitated by the histidine from the other protomer. This concerted action of two

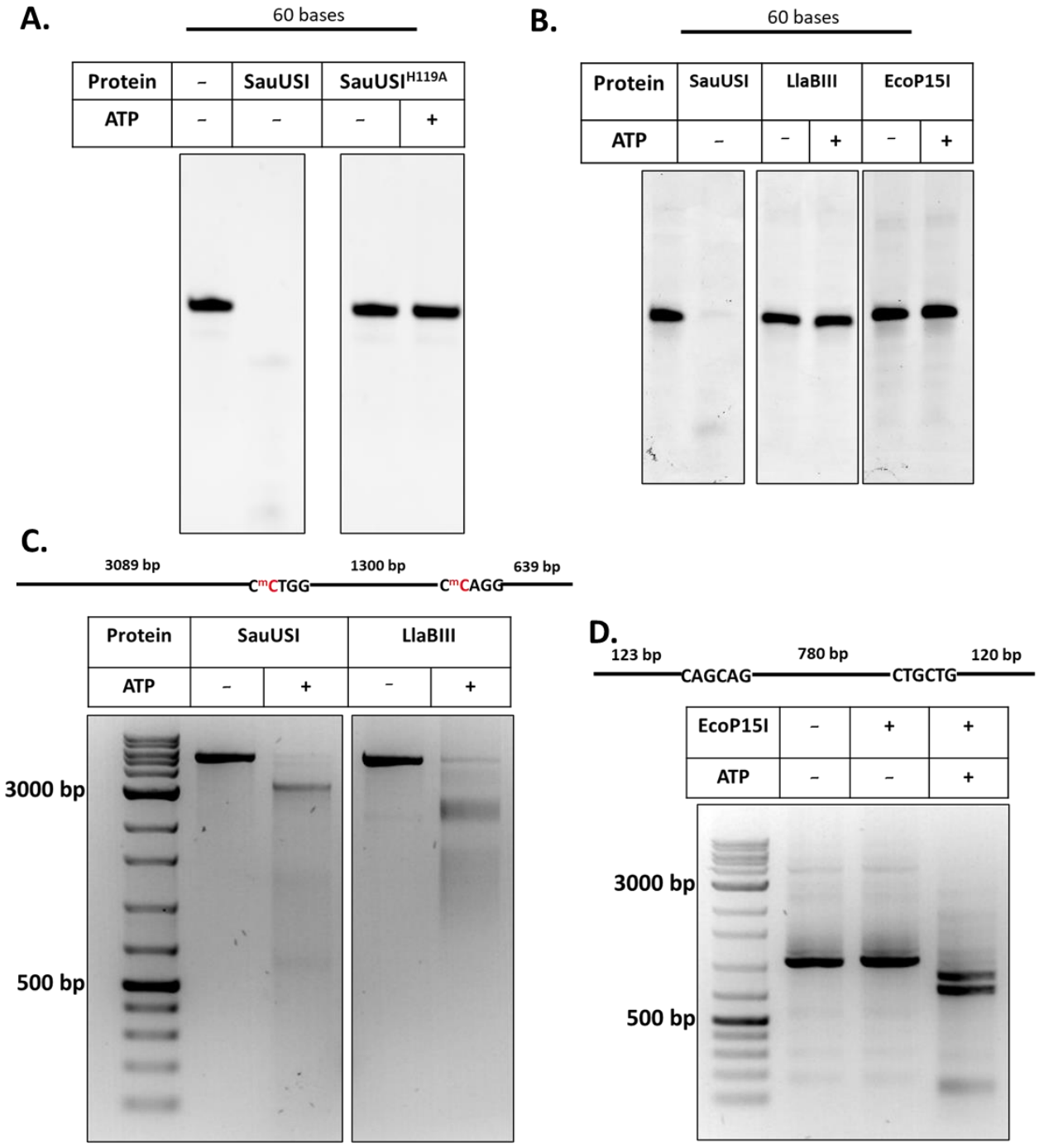


Figure 2: Cleavage of ssDNA does not occur in all types of R-M systems: (A). Schematic of the substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of ssDNA treated with SauUSI and SauUSI H119A (Nuclease inactive mutant) either in the presence or absence of ATP. (B) Schematic of the substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of ssDNA treated with SauUSI, LlaBIII and EcoP15I either in the presence or absence of ATP. (C) dsDNA substrate used for the cleavage assay. The substrate used is a linearized (with NcoI) plasmid and methylation was brought about *in vivo* to generate SauUSI target sites. Representative 1 % Agarose gel to study the canonical cleavage of SauUSI and LlaBIII in the presence or absence of ATP run in the presence of 1kB DNAmark ladder. (D) dsDNA substrate used for the cleavage assay. The substrate used was generated *in vitro* by PCR amplification . Representative 1 % Agarose gel to study the canonical cleavage of EcoP15I in the presence or absence of ATP run in the presence of 1kB DNAmark ladder.

histidines brings about the eventual scission of the phosphate backbone. Mutating the active-site histidine of the nuclease domain to alanine abolished the ssDNA cleavage activity of SauUSI (SauUSI^{H119A}) both in the presence and absence of 2 mM ATP (Figure 2a).

We tested if other ATP-dependent restriction enzymes LlaBIII (Type ISP RM enzyme) and EcoP15I (Type III RM enzyme) possess the ability to cut ssDNA. The ability of BstNI, a canonical ATP-independent Type II restriction endonuclease that recognizes the same target sequence as SauUSI for dsDNA cleavage, to cleave ssDNA was also tested. Like SauUSI, LlaBIII and EcoP15I require at least two sites in the head-to head orientation for dsDNA cleavage (25, 31–33). 250 nM of ssDNA was treated with 100 nM of LlaBIII and EcoP15I and incubated at 25^{o}C for 1 hour both in the presence and absence of 2 mM ATP (see Materials and Methods for details). Along with these reactions, a control reaction containing 250 nM ssDNA and 100 nM SauUSI either in the presence or absence of 2 mM ATP was set up and incubated at 37^{o}C for 1 hour. Upon performing the aforementioned assay, we concluded that while SauUSI cleaves ssDNA (Figure 2b, Lane 2), LlaBIII and EcoP15I did not possess this ability (Figure 2b, Lanes 3 - 6). BstNI was also not able to cleave ssDNA. In parallel, we also tested the DNA cleavage activity of these enzymes on their canonical two-site dsDNA under identical conditions. The cleavage of SauUSI, LlaBIII and EcoP15I on their respective substrates gave expected patterns of cleavage in the presence of 2 mM ATP (Figures 2c and 2d).

ssDNA cleavage does not require divalent cation and is EDTA resistant

Biochemical characterization of the PLD endonuclease Nuc and BfiI revealed that the histidine side chain at the active site functions as a nucleophile thus causing an enzymatic attack of the phosphate backbone of DNA. This is in stark contrast to canonical Type II REases which function by coordinating the active-site with two-divalent metal ions thus bringing about the activation of a water molecule which serves as a nucleophile for scission of the phosphate backbone of DNA (34). Therefore, BfiI and Nuc are categorized as metal-independent and EDTA resistant nucleases whereas majority of the Type II REases are metal-dependent and EDTA labile (30, 35). Previous experiments on SauUSI revealed that dsDNA cleavage is stimulated by ATP hydrolysis brought

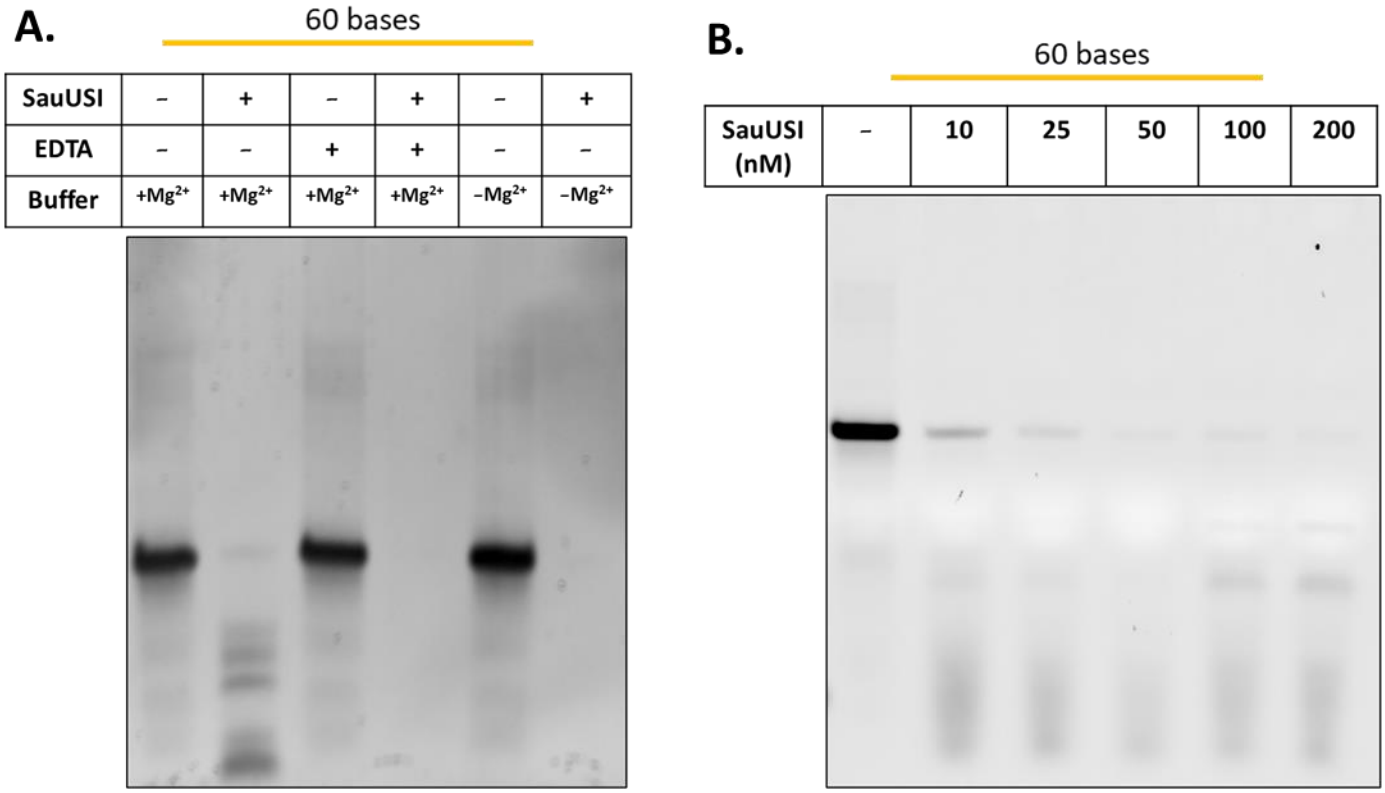


Figure 3: Cleavage of ssDNA can occur in the absence of divalent cations: (A). Schematic of the ssDNA substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of ssDNA treated with SauUSI either in the presence of EDTA, absence of EDTA or absence of Mg^{2+} from the reaction buffer. (B) Schematic of the ssDNA substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of ssDNA varying the concentration (10 nM - 200 nM) of SauUSI in a reaction buffer free of Mg^{2+}.

about by the SF2 helicase-like ATPase domain. SF2 helicase-like ATPases require Mg^{2+} ion to bring about stable binding of ATP to the nucleotide binding pocket (36). ATP hydrolysis in SauUSI leads to a confirmational change thereby resulting in translocation along dsDNA.

Since SauUSI possesses a N-terminal PLD nuclease domain whose ssDNA cleavage activity is ATP independent, we tested if it might function as a metal-independent and EDTA resistant nuclease on ssDNA. To this end, we performed a cleavage assay in the presence of the divalent cation chelator EDTA. Since the reaction buffer standardized so far contained 10 mM magnesium acetate, we used a comparable concentration of EDTA (16 mM) to ensure complete chelation of the divalent cations. Upon incubating 200 nM SauUSI with 250 nM ssDNA in 1X NEB4 buffer (cleavage buffer of SauUSI) supplemented with 16 mM EDTA at 37^{o}C for 45 minutes we observed cleavage (Figure 3a. Lane 3 and 4). SauUSI was also able to cleave ssDNA in the absence of magnesium acetate in the reaction mixture (Figure 3a. Lane 5 and 6). Hence, we concluded that ssDNA cleavage by SauUSI does not require divalent cation, which was consistent with the metal-independent nature of other characterized PLD nucleases. Interestingly, in the absence of magnesium acetate, SauUSI cleaved ssDNA better than in the presence of it (Figure 3a. Lane 2 and Lane 6). Performing a cleavage assay with dsDNA possessing two target sites of SauUSI in the absence of 10 mM magnesium acetate showed no detectable cleavage in the presence of ATP whereas, cleavage was achieved on supplementing the reaction buffer with 10 mM magnesium acetate in the presence of ATP. This was consistent with our assumption that Mg^{2+} was important for ATP hydrolysis by the ATPase domain of SauUSI that activated dsDNA cleavage in a target-site dependent manner. We think that Mg^{2+} is not required for the reaction chemistry of cleavage of the phosphodiester bonds of dsDNA *per se*.

Single-stranded cleavage is efficient and specific only for DNA and not RNA

Efficient cleavage of 250 nM of ssDNA was observed with as low as 10 nM SauUSI with optimal cleavage noted at 50 nM (Figure 3b). Time dependent cleavage assay at constant concentration of DNA (250 nM) and protein (50 nM) showed ~60 % cleavage efficiency within 2 minutes of reaction incubation at 37^{o} C and which increased to ~90 % by 6 minutes (Figure 4a, 4b). With increasing time, the ssDNA continued to be cut into shorter fragments. This highlights a potent

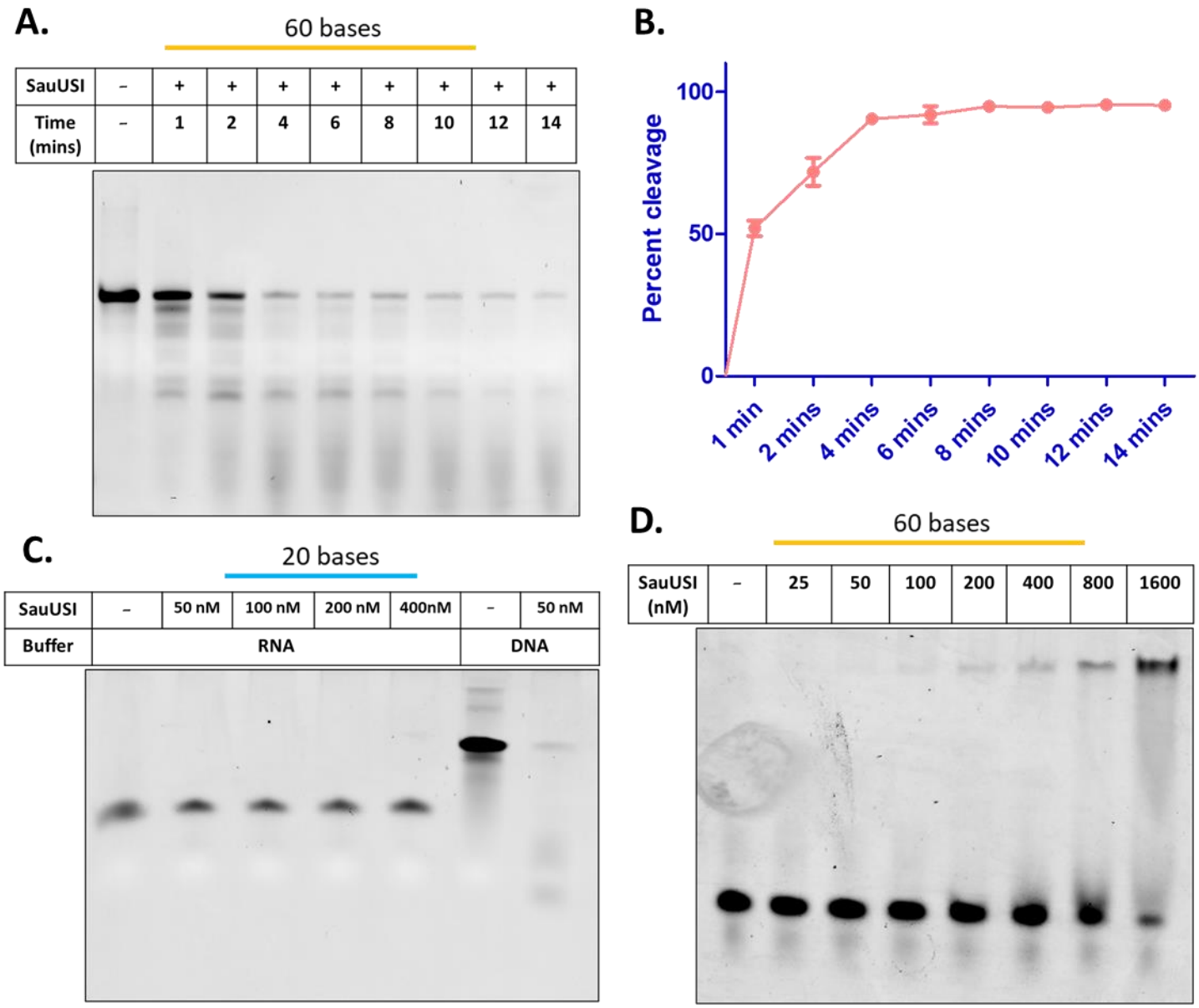


Figure 4: Cleavage of ssDNA is efficient and ssRNA is recalcitrant to cleavage: (A) Schematic of the ssRNA substrate used for the cleavage assay. Representative 20 % Native PAGE gel to study the time dependence (1 min – 14 min) of ssDNA cleavage using 50 nM of SauUSI in a reaction buffer free of Mg^{2+}. (B) Percent cleavage (y-axis) as a function of time (mins: x-axis) (n=3). (C) Schematic of the ssRNA substrate used for the cleavage assay. Representative 20 % Denaturing PAGE gel to study the cleavage pattern of ssRNA varying the concentration (50 nM - 400 nM) of SauUSI. The cleavage is in contrast to cleavage of ssDNA by 50 nM SauUSI. (D) ssDNA substrate used for EMSA study. Representative 6 % Native PAGE to study the binding of ssDNA over a SauUSI concentration range of 25 nM - 1600 nM.

single-stranded nuclease activity. We also found that the cleavage activity of SauUSI was specific to DNA as it did not cleave ssRNA (Figure 4c). We tested the affinity of SauUSI for ssDNA using electromobility shift assay (EMSA). Varying concentrations of nuclease inactive mutant SauUSI^{H119A} and 250 nM 60 base ssDNA was used for the experiment. More than 50% of the DNA shifted at a protein concentration of 1.6 μM, which showed that SauUSI had significant affinity for ssDNA (Figure 4d).

SauUSI is a ssDNA endonuclease

Having established that SauUSI cuts ssDNA efficiently, we proceeded to find if the enzyme is an endonuclease or an exonuclease. Exonucleases are enzymes that trim/cleave DNA from either the 5' end or 3' end releasing approximately one nucleotide or a nucleoside at a time. On the other hand, endonucleases have the ability to cleave DNA within the DNA phosphate backbone without the strict requirement of end trimming. To distinguish between the two modes of cleavage we used four kinds of substrates: 1) substrates with a 5' and 3' overhang, 2) substrate with varying lengths of gaps 3) substrate with a mismatch bubble and 4) substrate with hairpin loop.

Substrates with overhangs were designed to have a 40 bp dsDNA region with either a 20-base long 5' or 3' overhang. Cleavage with 50 nM SauUSI did not show any preference for either the 5' or 3' termini (Figure 5a). This result indicated that SauUSI did not have a polarity for cleavage and that it might function as an endonuclease. If this were the case then, substrates with a ssDNA gap and mismatch bubble would be labile to cleavage as well. Cleavage assay with a DNA having a 20-base long ssDNA gap flanked by a 15 bp and 25 bp dsDNA region revealed that SauUSI specifically cleaved the ssDNA fragment leaving the dsDNA region unaffected establishing that the enzyme functioned as an endonuclease (Figure 5b. Lanes 2 and 3). Decreasing the ssDNA gap to 15 bases, 10 bases and 5 bases resulted in commensurate reduction in cleavage efficiency (Figure 5b. Lanes 4-9). Furthermore, nicked DNA was not cleaved by SauUSI. Consistent with these observations, we also found that SauUSI cleaved a DNA having a 15 base mismatch bubble flanked by 22 bp and 23 bp dsDNA. SauUSI selectively cleaved the unannealed single-stranded bubble and not the flanking dsDNA (Figure 5c). As expected, the enzyme also cleaved DNA

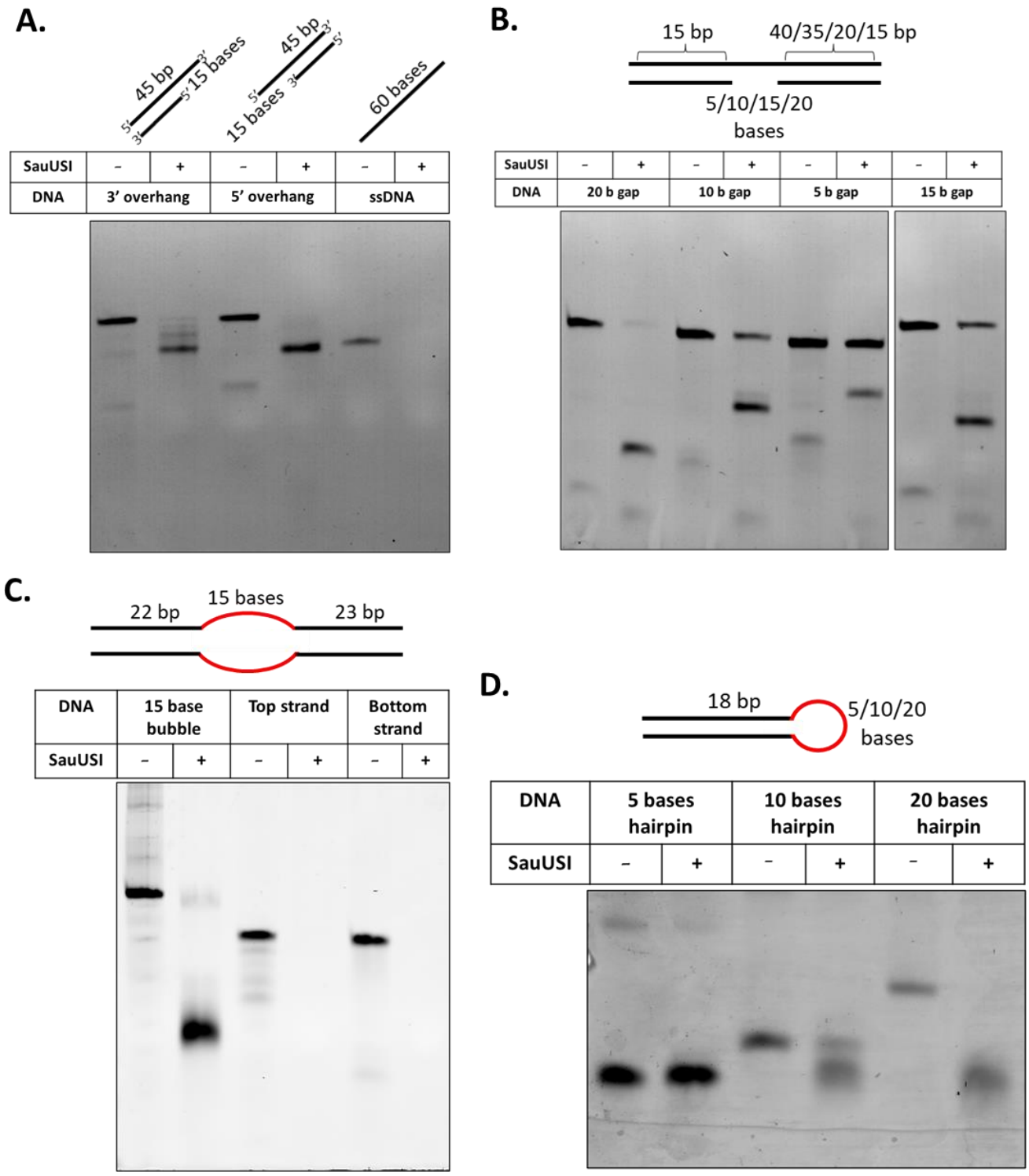


Figure 5: SauUSI functions as an endonuclease on ssDNA: (A). Schematic of the substrates used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of overhang substrates and ssDNA treated with SauUSI. (B) Schematic of the DNA substrates with gaps used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of DNA substrates with gaps (5/10/15/20 bases gaps) treated with SauUSI. (C) Schematic of the DNA substrate with a mismatch bubble (red) used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of DNA substrate with a mismatch bubble (red) and the individual ssDNA components treated with SauUSI. (D) Schematic of the DNA substrates with a hairpin loop (red) used for the cleavage assay. Representative 12 % Native PAGE gel to study the cleavage pattern of DNA substrate with a hairpin loop (red: 5/10/20 base hairpin) treated with SauUSI.

containing hairpin loop. While cleavage was observed for hairpin having loops of length 20 and 10 bases, a hairpin having a 5-base long loop remained uncleaved (Figure 5d).

Factors inhibiting ssDNA cleavage

As mentioned earlier, ssDNA is a transient intermediate of fundamental processes such as replication, transcription and DNA repair. However, ssDNA is sequestered and protected from damage in the cell by single strand DNA binding (SSB) protein. To check if SSB can protect ssDNA from the endonucleolytic activity of SauUSI, we incubated M13mp18 ssDNA with varying concentrations of *E. coli* SSB. We confirmed the binding of SSB to ssDNA based on the shift in the DNA position in an agarose gel on addition of SSB (Figure 6a) (27, 28). Cleavage assay with free M13mp18 ssDNA and SSB-bound M13mp18 ssDNA was carried out. While free M13mp18 ssDNA was nucleolytically degraded, SSB-bound M13mp18 ssDNA was protected from degradation by SauUSI (Figure 6b). This indicated a possible mode of ssDNA protection from SauUSI in the cell.

We had earlier found that the presence of Mg^{2+} lowered the ssDNA nuclease activity of SauUSI (Figure 3a). We had also noticed that ATP lowered the ssDNA cleavage activity (Figure 1). We studied this effect further by performing the cleavage assay by varying the concentrations of ATP from 50 μM to 4 mM. A reduction in efficiency of cleavage was observed as the concentration of ATP increased (Figure 7a). The inhibition was >85 % at 4 mM ATP and the IC_{50} calculated was ~ 760 μM (Figure 7b). We next tested if other nucleoside tri-phosphates possess the ability to inhibit cleavage. Interestingly, ATP and dATP conferred maximum inhibition to cleavage, whereas inhibition by GTP, CTP and UTP was comparatively lower (Figure 7c). ADP, on the other hand, appeared as inhibitory as ATP, while AMP had limited inhibition (Figure 7d). This result suggested a regulatory role for the ATPase domain in the ssDNA cleavage activity.

Discussion

We have demonstrated here that the restriction enzyme SauUSI has the ability to endonucleolytically cleave ssDNA. This is, to the best of our knowledge, the first example of a restriction enzyme being a *bona fide* ssDNA endonuclease. While there have been previous reports of restriction enzymes cleaving ssDNA, the cleavage was concluded to be the result of formation of transient dsDNA regions that were susceptible to cleavage. The observations

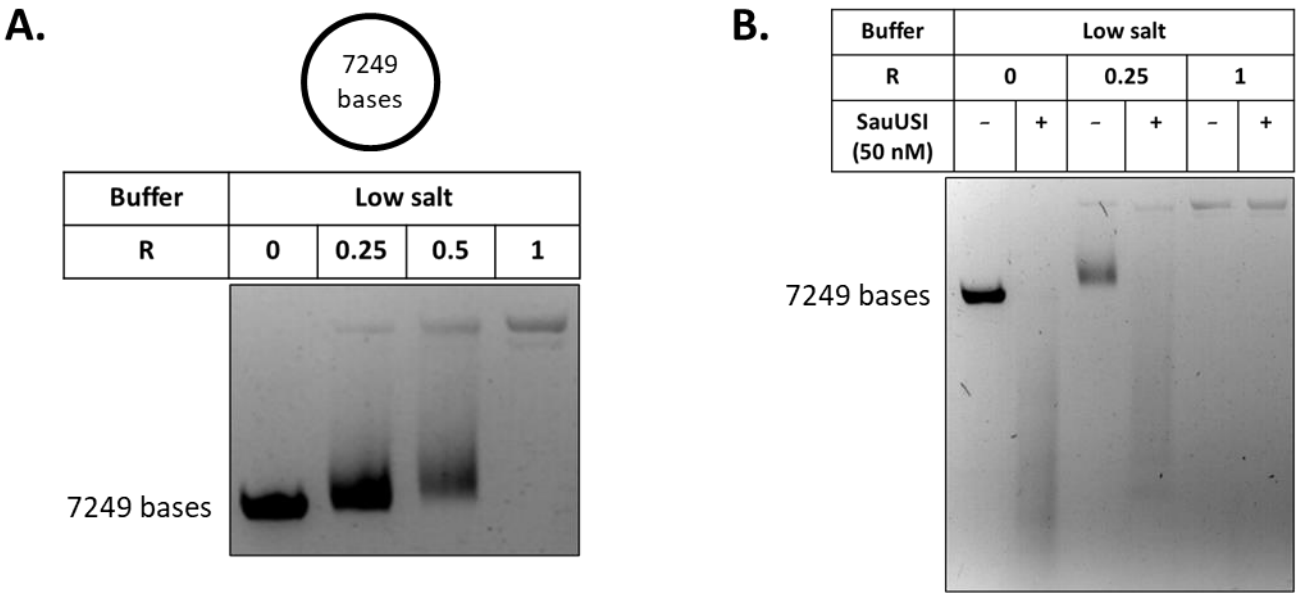


Figure 6: Regulation of ssDNA cleavage by single-stranded DNA binding protein: (A). M13mp18 circular ssDNA used for the cleavage assay. Representative Agarose gel to study the binding of *E. coli* single-stranded DNA binding (SSB) protein at different R values (ratio of SSB:ssDNA) to M13mp18 circular ssDNA. (B) Representative Agarose gel to study the cleavage pattern of unbound or bound M13mp18 (bound to *E. coli* SSB) treated with SauUSI performed over R values 0, 0.25 and 1.

supporting this conclusion were (i) the requirement of the same target site for ssDNA cleavage as for dsDNA cleavage (11, 13); (ii) correlation between the stability of the dsDNA structure and efficiency of cleavage (11, 16); the requirement of higher enzyme concentration for ssDNA cleavage than dsDNA cleavage (8–15) possibly due to the transient nature of the dsDNA substrate (16). In contrast, cleavage of ssDNA by SauUSI did not show any particular sequence requirement. Also, efficient ssDNA cleavage required only as much concentration of the enzyme as dsDNA cleavage; and did not require ATP – a requisite for dsDNA cleavage by SauUSI.

S. aureus is a notorious pathogen known to cause various human maladies ranging from minor skin infections to life threatening endocarditis (37)(38). The success of the organism as a pathogen is greatly determined by its ability to evade host immune response and acquire antibiotic resistance (39–43). Micrococcal Nuclease (MN), a secreted endonuclease encoded by *nuc1*, and the cell-wall bound nuclease encoded by *nuc2* are involved in immune evasion thus increasing pathogenesis (44–47). The diverse Cas nuclease repertoire of the CRISPR systems and the different types of R-M systems in *S. aureus* regulate the acquisition of foreign genetic material (2, 18, 48–51). The dsDNA cleavage activity of SauUSI has been identified as a potent barrier to the entry of methylated dsDNA either by transduction or transformation (2, 17–19). However, we also find that SauUSI can also function as a potent ssDNA endonuclease. Consequently, SauUSI can be a barrier to invasion of ssDNA phages, such as those belonging to the bacteriophage family Microviridae, or to the process of conjugation, which involves receipt of chromosomal or extrachromosomal material as ssDNA from donor bacteria (52). Conjugation between the donor *Enterococcus sp.* and the recipient *S. aureus* is a major mode of acquisition of antibiotic resistance (20).

The study also demonstrated that the enzyme endonucleolytically cleaved ssDNA in a variety of DNA structures, including DNA with single-stranded overhangs, gaps, bubbles and loops. Efficiency of cleavage of ssDNA gaps decreased with decreasing length of the gap and SauUSI did not cut nicked DNA. These ssDNA containing structures occur transiently but frequently in a cell at various stages of its physiology, including during DNA replication, repair and recombination.

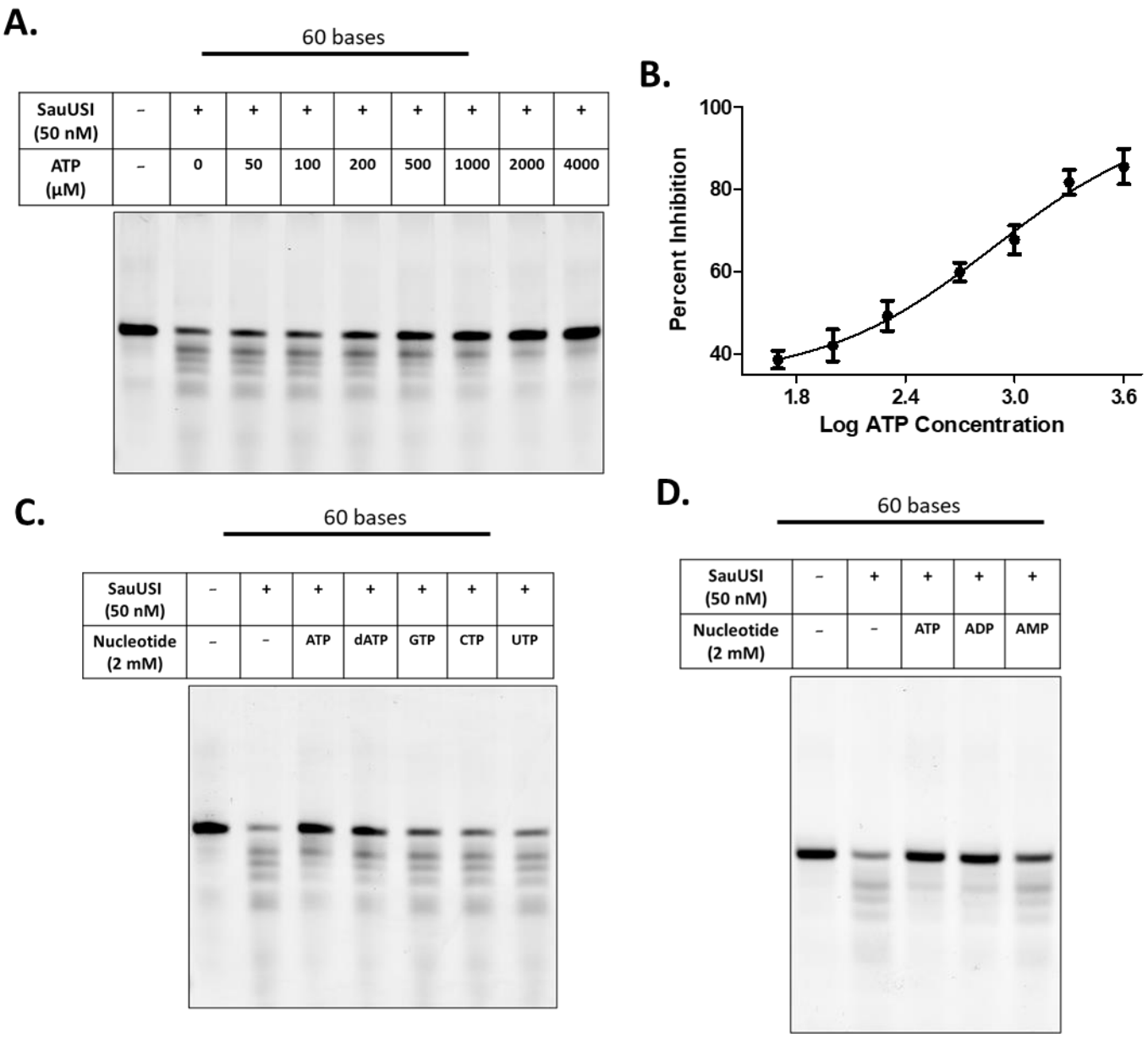


Figure 7: Regulation of ssDNA cleavage using nucleotides: (A) Schematic of the ssDNA substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the effect of an increasing concentration (0 – 4000 μ M) of ATP on the cleavage of ssDNA DNA substrate by SauUSI. (B) Percent inhibition plotted as a function of Log ATP concentration to calculate IC_{50} value for inhibition (C) Schematic of the ssDNA substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the effect of different nucleoside triphosphates on the cleavage of ssDNA DNA substrate by SauUSI. (D) Schematic of the ssDNA substrate used for the cleavage assay. Representative 12 % Native PAGE gel to study the effect of different Adenine nucleotides (ATP, ADP and AMP) on the cleavage of ssDNA DNA substrate by SauUSI.

Consequently, these important DNA structural intermediates will be susceptible to the ssDNA cleavage activity of SauUSI, if left unregulated. We found that the single-stranded endonucleolytic activity of SauUSI was inhibited by Mg^{2+}, ATP, dATP or ADP. However, inhibitory effects of AMP were much lower. The IC_{50} of ATP inhibition was measured to be 760 μM, a concentration which is lower than the concentrations of ATP in bacteria under normal physiological conditions.

Inhibition of the ssDNA nuclease activity of SauUSI by either ATP or ADP indicates that the inhibition is due to nucleotide binding rather than hydrolysis. Inhibition of the nuclease could be a result of conformational changes in SauUSI resulting from nucleotide binding. A similar proposal has been made for the regulation of ATPase-coupled nuclease *S. aureus* DinG (sarDinG) by ATP(53). sarDinG has a nuclease domain and an inactive helicase that can bind and hydrolyze ATP but cannot unwind DNA (53). The nuclease has a ssDNA 3'-exonuclease activity, which is negatively regulated by ATP concentration (53). It is proposed that ATP binding at the helicase domain causes a conformational change that inhibits the exonuclease. Apart from the inhibitory effect of ATP/ADP and Mg^{2+} on ssDNA cleavage activity of SauUSI, we found that SSB also prevented this cleavage *in vitro*. Thenceforth, we think that the ssDNA cleavage activity of SauUSI is tightly regulated under physiological conditions. This also leads us to question if the ssDNA cleavage activity of SauUSI would ever manifest and serve as a bacterial defense against foreign ssDNA.

We speculate that SauUSI could function as a bacterial defense against single-stranded DNA phages in a manner analogous to the nicking endonuclease GajA of the nucleotide sensing Gabija system in *Bacillus sp*(54). Like SauUSI, GajA is negatively regulated by nucleotide concentrations (dNTPs, NTPs, dNDPs and NDPs), while monophosphate nucleotides (ribo and deoxyribo derivatives) and nucleosides do not impede cleavage (54). It is proposed that during phage infection, active DNA transcription and replication deplete NTP and dNTP pool. Decrease in the trinucleotide pool would then activate the GajA nuclease causing degradation of phage DNA and host genome, resulting in abortive infection (54). A similar scenario can pan out when bacteria hosting SauUSI are infected by single-stranded DNA phages. Depletion of NTPs and dNTPs during phage DNA replication and transcription will activate cleavage of ssDNA, including the replicating

phage genome and that formed transiently in the host genome, causing abortive infection. If the foreign DNA entering the host bacteria is double-stranded and has the SauUSI target sites, shredding of the foreign dsDNA would be the prominent barrier to the entry (17).

S. aureus is a facultative intracellular pathogen and, hence, is subject to a host of DNA damaging agents such as reactive oxygen species (ROS) synthesized by the host defense mechanisms (55). As a result, they have evolved numerous DNA repair mechanisms that involves multiple nucleases acting in tandem (56). *sbcC* and *sbcD* encode the endonuclease SbcCD to resolve stalled replication forks (57); a homologue of YqfS is known to be involved in base-excision repair (56); UvrABC endonuclease complex plays a pivotal role in nucleotide-excision repair (56). The mismatch repair system, which helps correct aberrant incorporation of nucleotides during replication, is thought to be brought about by the nicking-endonuclease MutL (because of the absence of homologs of MutH) and excision of the erroneous region by the exonuclease RecJ (56, 58). dsDNA breaks are repaired by homologous recombination in *S. aureus*, which is initiated by trimming the ends by the RexAB helicase-nuclease to form 3'-overhangs crucial for the process of strand invasion (59). The holiday junction so formed is further resolved by another nuclease MutS2 (56, 60). SauUSI having the ability to cut a mismatch bubble, a ssDNA gap or an overhang, has the potential to moonlights as a DNA repair endonuclease *in-vivo*.

The ability of SauUSI to endonucleolytically cleave ssDNA leads to a variety of interesting possibilities and roles for the enzyme *in vivo* that remains to be unraveled. However, the ability of an enzyme to carry out a specific reaction does not necessarily imply an *in vivo* significance for that enzymatic reaction. An example of such a scenario is that noted for CRISPR-Cas12a recently. CRISPR-Cas12 has a strong ssDNA endonucleolytic activity *in vitro* (61), which does not have a physiological relevance (62). The present study provides a platform to perform detailed studies to understand the physiological role of the ssDNA cleavage activity of SauUSI, in particular during the process of transduction and conjugation.

References

1. Wheatley,R.M. and MacLean,R.C. (2021) CRISPR-Cas systems restrict horizontal gene transfer in Pseudomonas aeruginosa. *ISME J.*, **15**, 1420–1433.

2. Corvaglia,A.R., François,P., Hernandez,D., Perron,K., Linder,P. and Schrenzel,J. (2010) A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical Staphylococcus aureus strains. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 11954—11958.

3. Price,V.J., McBride,S.W., Hullahalli,K., Chatterjee,A., Duerkop,B.A. and Palmer,K.L. (2019) Enterococcus faecalis CRISPR-Cas Is a Robust Barrier to Conjugative Antibiotic Resistance Dissemination in the Murine Intestine. *mSphere*, **4**.

4. Wang,G., Song,G. and Xu,Y. (2020) Association of CRISPR/Cas System with the Drug Resistance in Klebsiella pneumoniae. *Infect. Drug Resist.*, **13**, 1929–1935.

5. Tyumentseva,M., Mikhaylova,Y., Prelovskaya,A., Tyumentsev,A., Petrova,L., Fomina,V., Zamyatin,M., Shelenkov,A. and Akimkin,V. (2021) Genomic and Phenotypic Analysis of Multidrug-Resistant Acinetobacter baumannii Clinical Isolates Carrying Different Types of CRISPR/Cas Systems. *Pathog. (Basel, Switzerland)*, **10**.

6. Nishino,T. and Morikawa,K. (2002) Structure and function of nucleases in DNA repair: shape, grip and blade of the DNA scissors. *Oncogene*, **21**, 9022–9032.

7. Bernheim,A. and Sorek,R. (2020) The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.*, **18**, 113–119.

8. Blakesley,R.W. and Wells,R.D. (1975) 'Single-stranded' DNA from phiX174 and M13 is cleaved by certain restriction endonucleases. *Nature*, **257**, 421–422.

9. Horiuchi,K. and Zinder,N.D. (1975) Site-Specific Cleavage of Single-Stranded DNA by a Hemophilus Restriction Endonuclease. *Proc. Natl. Acad. Sci. U. S. A.*, **72**, 2555–2558.

10. Reckmann,B. and Krauss,G. (1987) The cleavage of single-stranded DNA by the isoschizomeric restriction endonuclease HhaI and CfoI. *Biochim. Biophys. Acta*, **908**, 90–96.

11. Nishigaki,K., Kaneko,Y., Wakuda,H., Husimi,Y. and Tanaka,T. (1985) Type II restriction endonucleases cleave single-stranded DNAs in general. *Nucleic Acids Res.*, **13**, 5747–5760.

12. Blakesley,R.W., Dodgson,J.B., Nes,I.F. and Wells,R.D. (1977) Duplex regions in 'single-stranded' phiX174 DNA are cleaved by a restriction endonuclease from Haemophilus aegyptius. *J. Biol. Chem.*, **252**, 7300–7306.

13. Godson,G.N. and Roberts,R.J. (1976) dna, single stranded/*metab. *Virology*, **73**, 561–567.

14. Yoo,O.J. and Agarwal,K.L. (1980) Cleavage of single strand oligonucleotides and bacteriophage phi X174 DNA by Msp I endonuclease. *J. Biol. Chem.*, **255 22**, 10559–10562.

15. Hofer,B., Ruhe,G., Koch,A. and Köster,H. (1982) Primary and secondary structure specificity of the cleavage of 'single-stranded' DNA by endonuclease Hinf I. *Nucleic Acids Res.*, **10**, 2763–2773.

16. Nishigaki,K. (2018) Short lifetime structures appearing in RNA and DNA. *Brief. Funct. Genomics*, **18**, 174–181.

17. Tumuluri,V.S., Rajgor,V., Xu,S.-Y., Chouhan,O.P. and Saikrishnan,K. (2021) Mechanism of DNA cleavage by the endonuclease SauUSI: a major barrier to horizontal gene transfer and antibiotic resistance in Staphylococcus aureus. *Nucleic Acids Res.*, **49**, 2161–2178.

18. Xu,S.-Y., Corvaglia,A.R., Chan,S.-H., Zheng,Y. and Linder,P. (2011) A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300. *Nucleic Acids Res.*, **39**, 5597–5610.

19. Monk,I.R., Shah,I.M., Xu,M., Tan,M.-W. and Foster,T.J. (2012) Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *MBio*, **3**.

20. de Niederhäusern,S., Bondi,M., Messi,P., Iseppi,R., Sabia,C., Manicardi,G. and Anacarso,I. (2011) Vancomycin-resistance Transferability from VanA Enterococci to Staphylococcus aureus. *Curr. Microbiol.*, **62**, 1363–1367.

21. Noble,W.C., Virani,Z. and Cree,R.G. (1992) Co-transfer of vancomycin and other resistance genes from Enterococcus faecalis NCTC 12201 to Staphylococcus aureus. *FEMS Microbiol. Lett.*, **72**, 195–198.

22. Périchon,B. and Courvalin,P. (2009) VanA-type vancomycin-resistant Staphylococcus aureus. *Antimicrob. Agents Chemother.*, **53**, 4580–4587.

23. Maree,M., Thi Nguyen,L.T., Ohniwa,R.L., Higashide,M., Msadek,T. and Morikawa,K. (2022) Natural transformation allows transfer of SCCmec-mediated methicillin resistance in Staphylococcus aureus biofilms. *Nat. Commun.*, **13**, 2477.

24. Tumuluri,V.S. and Saikrishnan,K. (2022) Heterologous Expression and High Degree Purification of the Restriction Endonuclease SauUSI. *Bio-protocol*, **12**, e4275.

25. Chand,M.K., Nirwan,N., Diffin,F.M., van Aelst,K., Kulkarni,M., Pernstich,C., Szczelkun,M.D. and Saikrishnan,K. (2015) Translocation-coupled DNA cleavage by the Type ISP restriction-modification enzymes. *Nat. Chem. Biol.*, **11**, 870–877.

26. Ahmad,I., Kulkarni,M., Gopinath,A. and Saikrishnan,K. (2018) Single-site DNA cleavage by Type III restriction endonuclease requires a site-bound enzyme and a trans-acting enzyme that are ATPase-activated. *Nucleic Acids Res.*, **46**, 6229–6237.

27. Lohman,T.M., Overman,L.B. and Datta,S. (1986) Salt-dependent changes in the DNA binding co-operativity of Escherichia coli single strand binding protein. *J. Mol. Biol.*, **187**, 603–615.

28. Waldman,V.M., Weiland,E., Kozlov,A.G. and Lohman,T.M. (2016) Is a fully wrapped SSB-DNA complex essential for Escherichia coli survival? *Nucleic Acids Res.*, **44**, 4317–4329.

29. Sasnauskas,G., Halford,S.E. and Siksnys,V. (2003) How the BfiI restriction enzyme uses one active site to cut two DNA strands. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 6410–6415.

30. Gottlin,E.B., Rudolph,A.E., Zhao,Y., Matthews,H.R. and Dixon,J.E. (1998) Catalytic mechanism of the phospholipase D superfamily proceeds via a covalent phosphohistidine intermediate. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 9202–9207.

31. Ramanathan,S.P., van Aelst,K., Sears,A., Peakman,L.J., Diffin,F.M., Szczelkun,M.D. and Seidel,R. (2009) Type III restriction enzymes communicate in 1D without looping between their target sites. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 1748—1753.

32. van Aelst,K., Tóth,J., Ramanathan,S.P., Schwarz,F.W., Seidel,R. and Szczelkun,M.D. (2010) Type III restriction enzymes cleave DNA by long-range interaction between sites in both head-to-head and tail-to-tail inverted repeat. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 9123–9128.

33. Dryden,D.T., Murray,N.E. and Rao,D.N. (2001) Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Res.*, **29**, 3728–3741.

34. Pingoud,A., Fuxreiter,M., Pingoud,V. and Wende,W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell. Mol. Life Sci.*, **62**, 685–707.

35. Zaremba,M., Urbanke,C., Halford,S.E. and Siksnys,V. (2004) Generation of the BfiI Restriction Endonuclease from the Fusion of a DNA Recognition Domain to a Non-specific Nuclease from the Phospholipase D Superfamily. *J. Mol. Biol.*, **336**, 81–92.

36. Fairman-Williams,M.E., Guenther,U.-P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.

37. Tong,S.Y.C., Davis,J.S., Eichenberger,E., Holland,T.L. and Fowler,V.G.J. (2015) Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin. Microbiol. Rev.*, **28**, 603–661.

38. David,M.Z. and Daum,R.S. (2010) Community-Associated Methicillin-Resistant *Staphylococcus aureus*: Epidemiology and Clinical Consequences of an Emerging Epidemic. *Clin. Microbiol. Rev.*, **23**, 616–687.

39. Foster,T.J., Geoghegan,J.A., Ganesh,V.K. and Höök,M. (2014) Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. *Nat. Rev. Microbiol.*, **12**, 49–62.

40. Itoh,S., Hamada,E., Kamoshida,G., Yokoyama,R., Takii,T., Onozaki,K. and Tsuji,T. (2010) Staphylococcal superantigen-like protein 10 (SSL10) binds to human immunoglobulin G (IgG) and inhibits complement activation via the classical pathway. *Mol. Immunol.*, **47**, 932–938.

41. Pietrocola,G., Nobile,G., Rindi,S. and Speziale,P. (2017) Staphylococcus aureus Manipulates Innate Immunity through Own and Host-Expressed Proteases. *Front. Cell. Infect. Microbiol.*, **7**.

42. Pantosti,A., Sanchini,A. and Monaco,M. (2007) Mechanisms of antibiotic resistance in Staphylococcus aureus. *Future Microbiol.*, **2**, 323–334.

43. Ito,T., Okuma,K., Ma,X.X., Yuzawa,H. and Hiramatsu,K. (2003) Insights on antibiotic resistance of Staphylococcus aureus from its whole genome: genomic island SCC. *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.*, **6**, 41–52.

44. Tang,J., Zhou,R., Shi,X., Kang,M., Wang,H. and Chen,H. (2008) Two thermostable nucleases coexisted in Staphylococcus aureus: evidence from mutagenesis and in vitro expression. *FEMS Microbiol. Lett.*, **284**, 176–183.

45. Thammavongsa,V., Missiakas,D.M. and Schneewind,O. (2013) *Staphylococcus aureus* Degrades Neutrophil Extracellular Traps to Promote Immune Cell Death. *Science (80-.).*, **342**, 863–866.

46. Kiedrowski,M.R., Kavanaugh,J.S., Malone,C.L., Mootz,J.M., Voyich,J.M., Smeltzer,M.S., Bayles,K.W. and Horswill,A.R. (2011) Nuclease modulates biofilm formation in community-associated methicillin-resistant Staphylococcus aureus. *PLoS One*, **6**, e26714.

47. Yu,J., Jiang,F., Zhang,F., Hamushan,M., Du,J., Mao,Y., Wang,Q., Han,P., Tang,J. and Shen,H. (2021) Thermonucleases Contribute to Staphylococcus aureus Biofilm Formation in Implant-Associated Infections-A Redundant and Complementary Story. *Front. Microbiol.*, **12**, 687888.

48. Sussenbach,J.S., Monfoort,C.H., Schiphof,R. and Stobberingh,E.E. (1976) A restriction endonuclease from Staphylococcus aureus. *Nucleic Acids Res.*, **3**, 3193–3202.

49. Waldron,D.E. and Lindsay,J.A. (2006) Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into Staphylococcus aureus and between S. aureus isolates of different lineages. *J. Bacteriol.*, **188**, 5578–5585.

50. Cruz-López,E.A., Rivera,G., Cruz-Hernández,M.A., Martínez-Vázquez,A.V., Castro-Escarpulli,G., Flores-Magallón,R., Vázquez,K., Cruz-Pulido,W.L. and Bocanegra-García,V. (2021) Identification and Characterization of the CRISPR/Cas System in Staphylococcus aureus Strains From Diverse Sources. *Front. Microbiol.*, **12**, 656996.

51. Nishimasu,H., Cong,L., Yan,W.X., Ran,F.A., Zetsche,B., Li,Y., Kurabayashi,A., Ishitani,R., Zhang,F. and Nureki,O. (2015) Crystal Structure of Staphylococcus aureus Cas9. *Cell*, **162**, 1113–1126.

52. Olsen,N.S., Forero-Junco,L., Kot,W. and Hansen,L.H. (2020) Exploring the Remarkable Diversity of Culturable Escherichia coli Phages in the Danish Wastewater Environment. *Viruses*, **12**.

53. McRobbie,A.-M., Meyer,B., Rouillon,C., Petrovic-Stojanovska,B., Liu,H. and White,M.F. (2012) Staphylococcus aureus DinG, a helicase that has evolved into a nuclease. *Biochem. J.*, **442**, 77–84.

54. Cheng,R., Huang,F., Wu,H., Lu,X., Yan,Y., Yu,B., Wang,X. and Zhu,B. (2021) A nucleotide-sensing endonuclease from the Gabija bacterial defense system. *Nucleic Acids Res.*, **49**, 5216–5229.

55. Ha,K.P., Clarke,R.S., Kim,G.-L., Brittan,J.L., Rowley,J.E., Mavridou,D.A.I., Parker,D., Clarke,T.B., Nobbs,A.H. and Edwards,A.M. (2020) Staphylococcal DNA Repair Is Required for Infection. *MBio*, **11**.

56. Ha,K.P. and Edwards,A.M. (2021) DNA Repair in Staphylococcus aureus. *Microbiol. Mol. Biol. Rev.*, **85**, e0009121.

57. Cirz,R.T., Jones,M.B., Gingles,N.A., Minogue,T.D., Jarrahi,B., Peterson,S.N. and Romesberg,F.E. (2007) Complete and SOS-Mediated Response of *Staphylococcus aureus* to the Antibiotic Ciprofloxacin. *J. Bacteriol.*, **189**, 531–539.

58. Prunier,A.-L. and Leclercq,R. (2005) Role of mutS and mutL genes in hypermutability and recombination in Staphylococcus aureus. *J. Bacteriol.*, **187**, 3455–3464.

59. Clarke,R.S., Ha,K.P. and Edwards,A.M. (2021) RexAB Promotes the Survival of Staphylococcus aureus Exposed to Multiple Classes of Antibiotics. *Antimicrob. Agents Chemother.*, **65**, e0059421.

60. Burby,P.E. and Simmons,L.A. (2017) MutS2 Promotes Homologous Recombination in Bacillus subtilis. *J. Bacteriol.*, **199**.

61. Chen,J.S., Ma,E., Harrington,L.B., Costa,M. Da, Tian,X., Palefsky,J.M. and Doudna,J.A. (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science (80-.).*, **360**, 436–439.

62. Marino,N.D., Pinilla-Redondo,R. and Bondy-Denomy,J. (2022) CRISPR-Cas12a targeting of ssDNA plays no detectable role in immunity. *Nucleic Acids Res.*, **50**, 6414–6422.

CONCLUSION AND FUTURE PROSPECTS

This thesis submitted for evaluation encompasses a study on the mechanism of action of a Type IV restriction endonuclease SauUSI. Genetic studies on methicillin-resistant *Staphylococcus aureus* (MRSA) revealed that SauUSI was a potent barrier to horizontal gene transfer (1). This enzyme was categorized as a Type III RM system based on sequence alignment and bioinformatic analysis(1). However, experimental studies on purified SauUSI revealed that it specifically cleaved methylated DNA and was refractory to cleavage of non-methylated DNA hence, this enzyme was reclassified as a Type IV restriction endonuclease (2).

I began working on this system with the hope of understanding the nucleolytic potency of SauUSI. I was intrigued by the fact that a mutation rendering the *sauUSI* gene inactive in *S. aureus* could have a manifold increase in its HGT efficiency and potentially transform the organism into a vancomycin-resistant bacterium (1, 3). Interestingly, this phenomenon was not observed when other RM systems of *S. aureus* were mutated (1). The project initiated with me cloning the *sauUSI* gene into an *E. coli* expression vector *pHis17*. I optimized the bacterial culture conditions and purification protocol to attain a high degree of purity (4). High degree purity of protein was essential not only for biochemical studies but was of paramount importance for crystallographic purposes. Since SauUSI cleaves methylated DNA, I needed to develop a pipeline for obtaining methylated DNA in scalable amounts for biochemical assays. Towards this, I began by constructing synthetic plasmids using the restriction-free method of cloning to obtain substrates having two or three target sites for studying the nuclease activity and a substrate possessing a triplex binding site for studying translocation activity of SauUSI. Once these substrates were sequenced and confirmed for the presence of either the SauUSI target-site or triplex binding site, I transformed them into a *dcm*$^{+}$ strain of *E. coli* (DH5α) cells and extracted the plasmids. The plasmids thus extracted were methylated at SauUSI target sites and were used for characterizing the enzyme.

I used the aforementioned DNA substrates and performed a sequential biochemical characterization of the nuclease activity and ATP hydrolysis-based translocation activity of SauUSI(5). In parallel, I was able to obtain a crystal for SauUSI using the hanging-drop vapor

diffusion method which diffracted to a resolution of 3.1 Å. The diffraction experiment was carried out at the ESRF and DLS synchrotron facilities. To obtain phase information, I used the single anomalous dispersion method by obtaining a selenomethionine derivative protein crystal. The structure revealed that SauUSI was dimer with the interface at the nuclease domain of the two protomers. The structure was in agreement to the single-site cleavage activity that I discovered while using a one-site methylated substrate. Further, I observed that DNA flanked by SauUSI target sites were nucleolytically shredded to bits and was evident as a smear on an Agarose gel. I attributed the ability of SauUSI to shred DNA to its ATP driven translocase activity. Based on the structural information of the Apo-form of SauUSI taken together with the biochemical characterization, I was able to put forth a model for cleavage where, in the presence of two or more target sites a translocation-based pile of SauUSI molecules would lead to shredding of DNA whereas in the presence of a single-target site, a less efficient switch-based mechanism would cause cleavage of DNA(5). A limitation of this study is the lack of structural information of the binary (SauUSI + DNA) and ternary (SauUSI + DNA + AMPPNP) complex which might help in understanding the various conformational changes that SauUSI can adopt. The Apo-form of the structure presented here I believe is in an inactive conformation because the nuclease, the SF2 helicase-like ATPase and target recognition domains are spatially apart and only a major conformational change between the domains would allow for them interacting with DNA simultaneously. The nature and trigger for these conformational changes between domains remain elusive.

On performing a pilot experiment with single-stranded DNA I discovered that SauUSI has the ability to cleave ssDNA in a target-site and nucleotide independent manner. This revelation prompted me to further characterize this hitherto unknown activity of SauUSI. I was able to characterize this cleavage activity as an endonucleolytic mode rather than an exonucleolytic mode primarily because of its ability to cleave large circular ssDNA, gapped substrates and substrates with a mismatch bubble. My study, reveals that the nuclease activity of SauUSI on ssDNA is divalent cation independent which is in agreement to other members of the Phospholipase-D family of proteins. Further, I was able to decipher mechanisms that could regulate this ssDNA cleavage activity thus providing a premise for why bacterial self-DNA is

protected from this indiscriminate cleavage action of SauUSI *in-vivo*. My study revealed that Single-stranded DNA binding protein was a strong inhibitor to cleavage of ssDNA by SauUSI. Interestingly, my study also revealed that the presence of ATP/ADP is inhibitory to the ssDNA cleavage. Other nucleotides such as GTP, CTP, UTP and AMP do not have such an inhibitory effect. However, the conformational changes that are brought about by ATP/ADP upon binding to SauUSI remains to be explored. To understand if the ssDNA cleavage activity of SauUSI has any physiological bearing, it is of paramount importance to undertake a detailed *in-vivo* study on SauUSI to test if ssDNA modes of HGT such as transduction using ssDNA phages and conjugation can be restricted even in the absence of methylation. Further, to understand how SauUSI would interact with ssDNA, crystallographic studies must be initiated on a nuclease inactive mutant of SauUSI using appropriate substrates.

In toto, my thesis provides a platform for further *in-vitro*, *in-vivo* and structural studies to focus on exploring all the facets of SauUSI activity. Detailed studies on enzymes like SauUSI can help not only in understanding the coupling of nucleotide binding and hydrolysis to translocation/cleavage but can also help discern the dynamics of spread of antibiotic resistance amongst bacterial populations.

References:

1. Corvaglia,A.R., François,P., Hernandez,D., Perron,K., Linder,P. and Schrenzel,J. (2010) A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical Staphylococcus aureus strains. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 11954—11958.

2. Xu,S.-Y., Corvaglia,A.R., Chan,S.-H., Zheng,Y. and Linder,P. (2011) A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300. *Nucleic Acids Res.*, **39**, 5597–5610.

3. Monk,I.R., Shah,I.M., Xu,M., Tan,M.-W. and Foster,T.J. (2012) Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *MBio*, **3**.

4. Tumuluri,V.S. and Saikrishnan,K. (2022) Heterologous Expression and High Degree Purification of the Restriction Endonuclease SauUSI. *Bio-protocol*, **12**, e4275.

5. Tumuluri,V.S., Rajgor,V., Xu,S.-Y., Chouhan,O.P. and Saikrishnan,K. (2021) Mechanism of DNA cleavage by the endonuclease SauUSI: a major barrier to horizontal gene transfer and antibiotic resistance in Staphylococcus aureus. *Nucleic Acids Res.*, **49**, 2161–2178.

LIST OF PUBLICATIONS

- **Vinayak Sadasivam Tumuluri**, Vrunda Rajgor, Shuang-Yong Xu, Om Prakash Chouhan, Kayarat Saikrishnan, Mechanism of DNA cleavage by the endonuclease SauUSI: a major barrier to horizontal gene transfer and antibiotic resistance in *Staphylococcus aureus*, *Nucleic Acids Research*, Volume 49, Issue 4, 26 February 2021, Pages 2161–2178, https://doi.org/10.1093/nar/gkab042
- **Tumuluri, V. S**. and Saikrishnan, K. (2022). Heterologous Expression and High Degree Purification of the Restriction Endonuclease SauUSI. *Bio-protocol* 12(1): e4275. DOI: 10.21769/BioProtoc.4275.

PERMISSIONS

Case #01641659 - Request to reprint paper as part of thesis External Inbox ×

customercare@copyright.com <customercare@copyright.com> Jun 30, 2022, 4:04 AM

to me

Dear Dr. Vinayak Sadasivam Tumuluri,

Thank you for contacting Copyright Clearance Center (CCC). We act on behalf of copyright owners in providing permissions to our customers. Permission availability can vary depending on a rightsholder and also on the permission type. My name is Lex and I hope you are having an excellent day and that this message finds you in great health.

I understand you would like permission to reuse content from "Mechanism of DNA cleavage by the endonuclease SauUSI: a major barrier to horizontal gene transfer and antibiotic resistance in Staphylococcus aureus". This is an **open-access** title distributed under the terms of the **Attribution 4.0 International (CC BY 4.0)** license. **You are not required to obtain permission to reuse this article** and you are free to:

- **Share** — copy and redistribute the material in any medium or format
- **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

As long as you give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

Reservation of Rights.

a. Ownership. Author reserves ownership of the Article and all of Author's other proprietary rights with respect to the Article, including the copyright and any patent rights Author may have in any inventions described in the Article, subject to the rights granted to Bio-protocol in Section 1. Bio-protocol shall be the holder of title for the purpose of copyright registration.

b. Other Rights. Author retains the following non-exclusive rights to use the Article without further permission after the publication of the Article by Bio-protocol, provided that Author will reference (and, where applicable, require any third party to reference) the appropriate citation of the publication by Bio-protocol of the Article when exercising any of the rights retained by the Author in this subsection 2.b.:

   i. Reprint the Article in print collections of Author's own writings;
   ii. Reprint the Article in print format for inclusion in a thesis or dissertation that the Author writes for a degree granted by an educational institution;
   iii. Present the Article orally;
   iv. Reproduce the Article for use in courses that the Author is teaching (if the Author is employed by an academic institution, that institution may also reproduce the Article for course teaching at that institution);
   v. Distribute photocopies of the Article to colleagues for non-commercial purposes only;
   vi. Reuse figures or tables in the Article for future publications written by the Author; and

Author agrees to obtain prior consent from Bio-protocol for any uses not expressly authorized in this Section 2.