Characterizing the effects and applications of Methylation dependent restriction endonucleases on genomic DNA

A Thesis

submitted to
Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme
by

Lakshmi Sriram
20171047



Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, India.

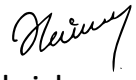
April, 2022

Supervisor: Prof. Saikrishnan Kayarat

Lakshmi Sriram
All rights reserved

Certificate

This is to certify that this dissertation entitled "**Characterizing the effects and applications of Methylation dependent restriction endonucleases on genomic DNA"** towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Lakshmi Sriram at Indian Institute of Science Education and Research under the supervision of Prof. Saikrishnan Kayarat, Professor, Department of Biology, during the academic year 2021-2022.



Prof. Saikrishnan Kayarat

Committee:

Prof. Saikrishnan Kayarat
Department of Biology
IISER Pune

Dr. Ramanujam Srinivasan
School of Biological Sciences
NISER

Declaration

I hereby declare that the matter embodied in the report entitled **"Characterizing the effects and applications of Methylation dependent restriction endonucleases on genomic DNA"** are the results of the work carried out by me at the Department of Biology, Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Saikrishnan Kayarat, and the same has not been submitted elsewhere for any other degree.



Lakshmi Sriram



Prof. Saikrishnan Kayarat

Date: 13th May 2022

Acknowledgments

As my time at IISER draws to a close, it strikes me that the past five years at IISER have been the best years of my life. Building a wonderful community of friends and chosen family, and learning more about science, both, in classes and by first-hand experience in laboratories, has been an incredible experience. Being able to work in a lab for a year and writing an entire thesis feels like a fitting culmination to this.

I am grateful to my supervisor, Prof. Saikrishnan Kayarat, whose style of mentoring was everything I'd hoped for in a supervisor. His belief in me, and his ability to trust that I had done my background work and knew what I was talking about only encouraged me and made me enjoy my experience working in his lab more. Frequent discussions with him helped shape this project and allowed it to evolve into its final form. I thank my expert, Dr. Ramanujam Srinivasan, for his input to this project.

I'd like to thank Vinayak, who has been the best PhD mentor one could wish for. Discussing experiments and troubleshooting problems with him made me appreciate the scientific process, and helped tremendously when frustrations about inevitable experimental failures arose.

I could not imagine what this past year of lab work would have been like without Abhinav. Having his company and constant support in lab has been incredible, and has made this so much easier. SK and G3 lab members have been an integral part of working in lab, and I thank them for wonderful conversations over Chai. I would like to acknowledge IISER Pune and DST-Inspire.

To Tathagata, Praniti and Shraddha, I cannot express my gratitude for providing me a shoulder to lean on when I needed it. I did not think it was possible to love friends this much. Rishabh, thanks for always having my back and being there for me, whether it be when I came back from lab at 2 a.m some nights or when I whined about how my experiment had failed yet again. There are countless friends whose company and laughter have evoked immeasurable joy and happiness in me, and I will keep those memories close to my heart.

And lastly, I'd like to thank my beautiful family. You show me time and again how wonderful, loving and supportive you are, and I have no words to express how lucky I feel to have born into this family. Amma and appa, I couldn't have done this without you. Sundu, your kindness and sense of humour make me feel happier than you'd know. Radha, Keshu and Chitti, thanks for being by my side. Chittaps, the excitement you had about me moving to live with you after graduating and pursuing research motivates me. Shravan, atthai and uncle, thanks for attempting to understand my project. Viji aunty, you've been an amazing and inspiring person. And lastly, to my paatis and thatha, who are the cutest grandparents and support me pursuing my passion, even if they don't really understand it.

This thesis is dedicated to my dearest Chittaps.

Table of Contents

DECLARATION **3**

ACKNOWLEDGMENTS **4**

ABSTRACT **9**

CHAPTER 1: INTRODUCTION **10**

1.1 Bacterial Defence Mechanisms and Restriction-Modification Systems 10

1.2 Type IV Restriction Modification Systems 10

1.3 SauUSI 11

1.4 McrBC 13

1.4.1 McrB 14

1.4.2 McrC 14

1.4.3 Endonuclease activity of McrBC 15

1.5 Applications of Type IV restriction enzymes 15

1.6. Scope of the thesis 17

CHAPTER 2: MATERIALS AND METHODS **18**

2.1 Bacterial Genomic DNA extraction 18

2.2 HEK 293 Genomic DNA extraction 18

2.3 SauUSI substrate 19

2.4 McrBC substrate 19

2.5 SauUSI purification 20

2.6 McrB expression and purification 20

2.7 McrC expression and purification 21

2.8 McrBC complex formation 22

2.9 Ethanol precipitation 22

2.10 Isopropanol precipitation 22

2.11 Cleavage assays 23

2.12 AMPure XP based size selection: 24

2.13 PCR amplification 24

CHAPTER 3: RESULTS 26

3.1 Extraction of Human and bacterial genomic DNA (gDNA) 26

3.2 Shredding of methylated bacterial genomic DNA with SauUSI 26

3.3 Shredding of Human genomic DNA with SauUSI 29

3.4 Optimising HEK gDNA cleavage using McrBC 29

3.5. Optimisation of buffers 32

3.6 Optimising combination of enzymes 34

3.7 Confirming the presence of DNA via PCR and optimization of PCR 37

3.8 Post-cleavage assay cleanup of DNA 42

CHAPTER 4: DISCUSSION 45

REFERENCES 48

List of Tables:

Table 1. List of primers used 24
Table 2. Constituents of PCR reactions 25

List of Figures

Figure 1. Primary domain architecture of SauUSI 12
Figure 2. Components of McrBC system 13
Figure 3. Extraction of genomic DNA from NEBTurbo, DH5α and HEK293 cells 26
Figure 4. Optimization of SauUSI concentration 27
Figure 5. Shredding of methylated bacterial genomic DNA by SauUSI and comparison to BstNI 28
Figure 6. Optimization of SauUSI incubation period 29
Figure 7. Optimization of McrBC concentration 30
Figure 8. Cleavage of unmethylated DNA by 200 nM McrBC 31
Figure 9. Optimisation of McrBC incubation period 32
Figure 10. Optimisation of buffer for combination reaction 33
Figure 11. HEK gDNA – Optimised enzyme concentrations 36
Figure 12. NEBTurbo gDNA– Optimised enzyme concentrations 37
Figure 13. Effects of NEB4 Buffer on PCR Amplification 36
Figure 14. Comparison between effects of Ethanol and Isopropanol precipitation on PCR Amplification 40
Figure 15. Comparison between effects of NEB4 and Ethanol precipitation on PCR Amplification 41
Figure 16. Size-selection using AMPure XP beads 43

Abstract

Type IV Restriction Endonucleases are modification-based restriction enzymes which are typically methylation dependent. The Type IV restriction enzymes SauUSI and McrBC are ATP and GTP dependent respectively and specifically cleave DNA after recognising sequences containing methylated cytosine residues. Methylated Cytosine residues are commonly found in the genomic DNA of many organisms. In Prokaryotes, methylation of genomic DNA is common as part of Restriction-modification systems, and methylation of eukaryotic genomic DNA is an epigenetic modification that can influence the rates of transcription of genes.

While SauUSI is a shredding enzyme that cleaves DNA in between two recognition sites into multiple small fragments, McrBC doesn't shred DNA, but still cleaves methylated genomic DNA into small fragments due to it short, 2 base pair long recognition site. This ensures that treatment of methylated genomic DNA with either of these enzymes, or a combination of both, would cause the DNA to cleave into fragments of a short length.

Here, we aimed to characterise and optimise the use of these Type IV restriction enzymes to selectively enrich unmethylated DNA in a mixture of methylated and unmethylated DNA for use in multiple applications, with the goal of developing a kit that could be used for this purpose in multiple applications in the future.

Chapter 1: Introduction

1.1 Bacterial Defence Mechanisms and Restriction-Modification Systems

Bacteria in the external environment are constantly under threat from bacteriophage infections and from other mobile genetic elements such as plasmids. Infection of bacteria with bacteriophages is a parasitic relationship that results in lysis of bacterial cells, following which the newly propagated phages are released into the external environment. [Westra et al., 2012] Bacteria possess various strategies to overcome untoward incidents such as bacteriophage attack of which, the 'innate immunity' and 'adaptive immunity' predominate. The CRISPR/cas system of bacteria is the primary adaptive immunity mechanism bacteria possess.

Restriction modification systems provide an important innate immunity mechanism of protecting the cell against foreign DNA. Restriction-modification systems consist of restriction endonucleases (REs) that cleave foreign DNA based on specific recognition sites, and cognate modification systems that modify those very recognition sites in the genome of the organism itself to prevent self-digestion. These modification systems are methyltransferases that confer protection from cleavage by methylating Cytosine or Adenine residues within the recognition site [Tock et al., 2005]. Thus, REs recognises non-self DNA by the lack of such modified residues in the recognition sites. Endonucleolytic cleavage of the foreign DNA renders it inactive for further replication, thus protecting the host bacterium.

Restriction systems are primarily classified into modification independent restriction systems and modification dependent restriction systems. Further classification is based on co-factor requirements, primary structure of the protein, position of cleavage etc.

1.2 Type IV Restriction Modification Systems

The evolutionary arms race between host (bacteria) and pathogen – phages or Mobile Genetic Elements (MGEs) – has led to the evolution of modification dependent restriction endonucleases that lack cognate methyltransferases. Type IV restriction enzymes (REs) are modification-dependent restriction endonucleases. Examples of such modifications are

methylated, hydroxymethylated or glucosyl-hydroxymethylated residues in DNA. The 5th Carbon on the Cytosine residue is often modified, and these residues are methylated cytosines (5mC), hydroxymethylated cytosines (5hmC), or glucosyl-hydroxymethylated residues. The 4th Carbon atom on the Cytosine could also be modified (4mC). Typically, these enzymes are methylation-dependent and cleave DNA at or in between specific recognition sequences containing methylated residues. SauUSI and McrBC are two such methylation-dependent restriction enzymes.

Type IV restriction modification systems were discovered in the 1950s and 1960s while studying the propagation of bacteriophages in different hosts. Scientists discovered that T-even phages that were grown in the bacteria *Shigella dysenteria* were unable to propagate in *Escherichia coli (E.coli)* B cells. They discovered that phage DNA contained hydroxymethylated cytosine residues which were then glucosylated, but the phage particles propagated in *S.dysenteria* lacked the glucosylation on the hydroxymethylated residues. Further experiments led scientists to hypothesise that the lack of glucosyl groups on the hydroxymethylated cytosine residues was causing the *E.coli* to restrict the propagation of phages, and this system must be absent from *S.dysenteria.* [Luria et al., 1970] This was later confirmed to be the case, and the genes responsible for the restriction of these phages in *E.coli* were termed *rglA* (restricts glucoseless phages) and *rglB* genes. These genes were present in *E.coli* but absent from *S. dysenteria*. In the 1980s, scientists began studying other modifications, including methylated cytosine residues. The *mcrA* and m*crB* genes were studied as a system with 5mC-specific functions, and were later found to be identical to *rglA* and *rglB* genes. [Loenen et al., 2014] Later genetic and sequential analysis showed that the *rglB* locus actually contained two genes, *mcrB* and *mcrC.* [Dila et al., 1990]

The ability of Type IV restriction endonucleases to recognise the difference between C, m5C, hm5C and other modifications can be exploited to study epigenetic phenomena in organisms. Some studies have used McrBC as a tool to study the same.

1.3 SauUSI

SauUSI is a methylation-dependent type IV restriction enzyme found in *Staphylococcus aureus.* It serves as a barrier to Horizontal Gene Transfer in the organism. It has been shown

to be responsible for thwarting the uptake of foreign DNA, including genes conferring Vancomycin Resistance. [Tumuluri et al., 2021]

SauUSI was discovered in clinical *S.aureus* strains when a Type I restriction modification system was mutated. Even on the mutation of this Type I RM system, hypothesised as the main barrier to HGT in *S. Aureus*, these strains failed to be transformed with plasmids from *E.coli*. A new restriction system was hypothesised to be the barrier against this Horizontal Gene Transfer (HGT). Initially, it was assumed that this yet unknown restriction endonuclease was a Type III restriction endonuclease. However, a cognate methylase gene was not found near the gene encoding this restriction endonuclease, which was very unusual for a Type III Restriction Endonuclease. Later experiments showed that the Restriction Endonuclease responsible was SauUSI, a type IV Restriction Endonuclease that selectively recognised a recognition sequence containing 5-methylated cytosine residues. [Xu et al., 2011]

SauUSI is a single polypeptide protein consisting of a N-terminal Phospholipsae D (PLD) nuclease, followed by a helicase-like ATPase domain belonging to the Superfamily 2 (SF2) of helicases, and a C-terminal SRA target recognition domain. A coupler domain connects the ATPase domain and the Target Recognition Domain.

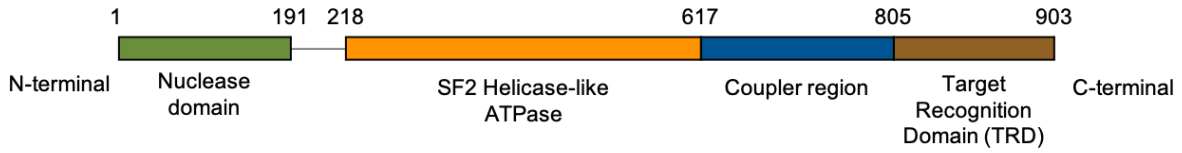


Figure 1. Primary domain architecture of SauUSI. The region between the Nuclease domain and the SF2 Helicase-like domain represents an unstructured linker.

In its active form, SauUSI exists as a dimer in solution. SauUSI recognises the site 5'-S^{5m}CNGS-3' (where S = G or C and 5mC can be 5-methylcytosine or 5-hydroxymethylcytosine) and cuts only after the recognition of these sites. It cuts around 15-20 bp from these sites. Initial evidence showed that cleavage by SauUSI required two target sites, and the DNA would be cut close to one of the target sites. Further research has shown, however, that single-site cleavage also occurs, albeit at a lower efficiency. Models have hypothesised that two-site or multiple-site cleavage occurs when a SauUSI dimer translocates from a free target site towards a second target site with a stationary SauUSI

dimer. Cleavage occurs near the second target site on convergence of the two SauUSI dimers. At a higher concentration of the enzyme, it has been shown to cleave the DNA in between two such target sites into fragments of short lengths; SauUSI is considered a 'shredder'. These models explain this phenomenon as a result of multiple translocating enzymes that get stalled in between the target sites. On convergence of these, cleavage occurs. [Tumuluri et al., 2021]

1.4 McrBC

McrBC is a GTP dependent Type IV restriction endonuclease from *Escherichia coli.* McrBC was the first restriction modification system ever discovered, and was studied in T phages. This system recognises modified cytosines in the genome; the 'Mcr' name stands for Modified Cytosine Restriction and is representative of this. [Raleigh et al., 1989]

The genes responsible for encoding the McrB and McrC proteins are the *mcrB* and *mcrC* genes respectively. McrB and McrC form the active endonuclease McrBC complex. McrBC is a functional tetradecamer consisting of two hexamers of McrB complexed with a dimer of McrC. McrB is a GTPase activated by the endonuclease McrC, and in turn, the energy provided by the GTPase activity of McrB powers the endonuclease activity of McrC.

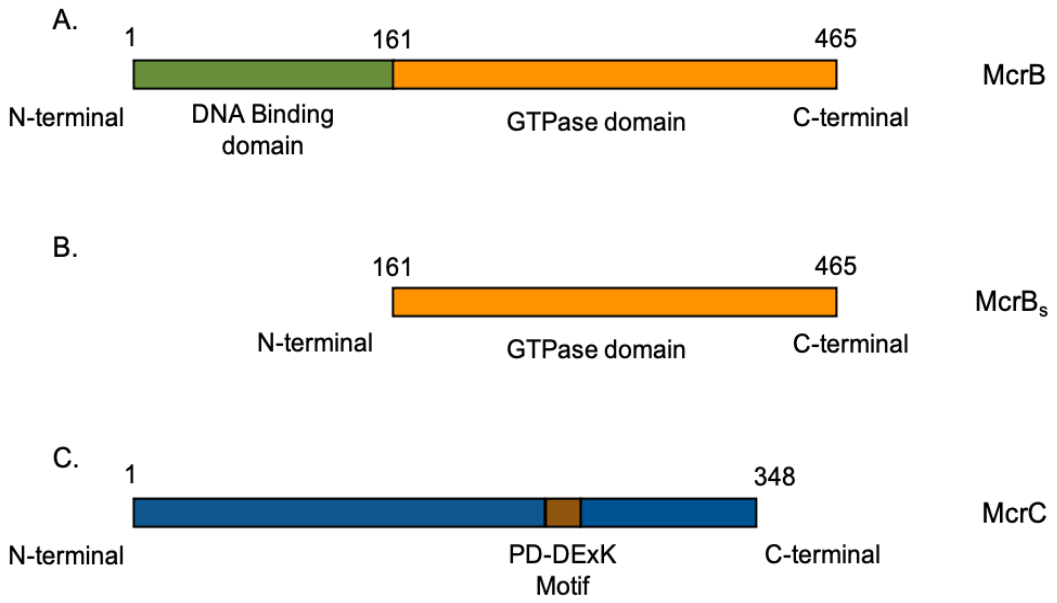


Figure 2 The components of the McrBC system. (A) The mcrB gene codes for the full-length GTPase McrB. (B) mcrB also codes for a shorter McrBs protein that modulates the activity of the McrBC complex. (C) The McrC protein contains a conserved motif found in the active site of Endonucleases.

Although McrB is part of the family of AAA+ proteins that utilise the energy from NTP

hydrolysis for various cellular activities, McrB is unique amongst these proteins as it is the only member of the AAA+ family that hydrolyses GTP instead of ATP. [Nirwan et al., 2019]

1.4.1 McrB

McrB is a GTPase that hexamerises in solution in the presence of GTP. Its GTPase activity is very low and is stimulated ~30 times by McrC. [Nirwan et al. 2020]

It has a size of 54 kDa and its N-terminal sequence of 161 amino acids is responsible for recognising and biding to the target site R^mC, where R stands for one of the purine residues Adenine or Guanine.

The *mcrB* gene also produces a smaller protein termed $McrB_s$ that lacks this N-terminal and thus shows no DNA binding activity. [Daniel Panne et al., 2001]. $McrB_s$ is expressed from the *mcrB* gene via initiation of translation at an alternate in-frame codon. It has been shown to modulate the activity of the McrBC complex. It is not directly involved in the cleavage of DNA.

The N terminal of McrB has been shown to contain a DNA binding fold that is adapted for the recognition of 5mC nucleotides. While studying the structure of this N terminal in complex with methylated DNA, it was shown that the 5mC base in the recognition site is flipped out from the DNA duplex and is in the structure of this N terminal in a complex along with methylated DNA, research has shown that the 5mC base is flipped out from the DNA duplex and is positioned within a binding pocket. [Sukackaite et al., 2012]

1.4.2 McrC

McrC is an endonuclease protein that forms the McrBC complex along with McrB. It has a size of 39 kDa. It was initially hypothesised that McrB was responsible for the endonuclease activity of McrBC and McrC was responsible to recognise some additional modifications such as 5hmC and 4-methylcytosine residues. Initial research claimed that McrC had no nuclease activity. [Dila et al., 1990] Later work showed that McrC stimulated the GTPase activity of McrC. It was later shown that the amino acid sequence of McrC had a conserved sequence motif. This conserved motif was similar to motifs in the active sites of other Restriction Endonucleases. [Piper and Pingoud, 2002]

1.4.3 Endonuclease activity of McrBC

The McrBC complex is formed on the binding of two McrC promoters to 2 hexamers of McrB. The McrBC complex recognises and binds to the consensus methylated sequence R^mC, where R represents either of the purine residues, A or G. The modified cytosine residue can be 5mC, 5hmC, or 4mC. [Nirwan et al., 2020] McrBC cleaves any DNA that has two of these recognition sites separated by 40-3000 bp, but optimal distance between the two recognition sites is 55-103 bp. [Panne at al., 1999]

The prevalent hypothesis explains that on binding of the McrBC complexes to the recognition sites, the GTPase activity of McrB powers the translocation of the complexes on DNA. The complexes remain bound to the recognition site, but the GTP-powered translocation causes looping of the DNA, which causes the convergence of the two complexes, following which cleavage occurs around 30 bp from one of the recognition sites. [Panne at al., 1999]

1.5 Applications of Type IV restriction enzymes

Methylation dependent restriction endonucleases can be used in a plethora of applications.

The Type IIM restriction endonuclease DpnI digests methylated and hemimethyalted DNA with a recognition site containing an Adenine residue methylated at the 6th carbon atom (6mA). DpnI is often used for experiments involving site-directed mutagenesis. PCR products with mutations in them can be created from template plasmids containing methylated Adenine residues. DpnI can then be used to digest the template DNA, only leaving behind the mutated plasmids [Liu et al., 2008]. Vectors with a single DpnI digestion site can also be linearized by DpnI digestion.

SauUSI has been shown to display shredding, where the methylated DNA is broken down into short fragments of DNA. Although McrBC shows no such shredding activity, the recognition site of McrBC is a sequence only two nucleotides long (R^mC), and a methylated genome contains many such sites. These two factors imply that any methylated genomic DNA would not just be cleaved into distinct large fragments on treatment with these

enzymes, as would be apparent via distinct bands on an agarose gel, but is instead cleaved into multiple small fragments an average of 150 bp long.

McrBC has been used to study methylation patterns in human DNA. Researchers have used McrBC to degrade methylated human DNA in order to study epigenetic modifications via methylation involved in cancer. [Ordway et al., 2006]

The modification-dependent cleavage by SauUSI and McrBC means that in any mixture of methylated and unmethylated DNA, only methylated DNA - containing the particular methylated residues corresponding to the recognition sites of the enzymes - will be degraded on treatment with the enzymes. The methylated residues in the genome of a particular organism will depend on the methyltransferase enzymes present in the organism.

Any sort of kit that is able to selectively enrich unmethylated DNA in a mix of methylated and unmethylated DNA would be useful in multiple applications. The characterisation of methylation patterns in metagenomes is one such application.

Another application for which such a kit that can be used would be to selectively enrich plasmodium DNA from the blood sample of a person infected with Malaria. Obtaining field isolates and sequencing them is an important step that is required to study the genetic diversity of *Plasmodium* parasites in epidemiological settings and understand the spread of drug resistance amongst *Plasmodium* isolates. A huge problem faced while sequencing field *Plasmodium* isolates is the large amounts of Human genomic DNA contamination from white blood cells. Even if the number of white blood cells in an isolate is lower that the number of parasites in the sample, the size of the human genome and its diploidy mean that a single WBC contains DNA equivalent to around 200 early blood stage *Plasmodium* parasites [Auburn et al., 2011]. While sequencing DNA extracted from these samples, most of the sequencing readouts obtained correspond to human genomic DNA instead of *Plasmodium* DNA. The use of Type IV restriction endonucleases to selectively enrich *Plasmodium* DNA is not a novel idea on its own, and has been done in the past with enzymes such as MspJ1, with mixed results [Oyola et al., 2013], [Cowell et al., 2017]. A recent study that was published in 2021 checked the efficacy of using McrBC to enrich the amounts and quality of *Plasmodium* DNA for Whole Genome Sequencing (WGS) from dried

blood spots. Dried Blood spots have been used as an easy method to collect small sample volumes from infected patients, but have small amounts of DNA as a result of the small samples. [Teyssier et al., 2021] They were able to improve the efficacy of sequencing using a Tween-chlorox method of extraction following digestion with McrBC, and showed that digestion with McrBC showed a significant improvement compared to samples that were extracted using the Tween-chlorox method but not digested with McrBC.

1.6. Scope of the thesis

SauUSI and McrBC are both Type IV restriction endonucleases that can be used to shear methylated DNA into small fragments. The main objective of this thesis was to optimise the use of these enzymes - whether the combination of these enzymes provides better activity, their concentrations in solution, the buffers used, and the duration of activity of these enzymes - to shear methylated human DNA completely while confirming that unmethylated DNA stays intact on treatment with these enzymes.

The intact unmethylated DNA must then be separated from the mixture of shredded methylated DNA and intact unmethylated genomic DNA. We aimed to do this via a method of size selection of the DNA that would result in the removal of the short fragments of methylated DNA from the sample. On the recommendation of Dr. Krishnapal Karmodiya, it was decided that AMPure XP® beads could be used for the same purpose. The plan was to show the selective enrichment of unmethylated DNA from a mixture of methylated and unmethylated DNA, even when present at low concentrations. This enrichment could be demonstrated via NGS readouts or PCR confirmations of shredding of the methylated DNA. The PCR confirmations of the selective enrichment would include amplification of sequences specific to the genomes involved.

This project was done in collaboration with the lab of Dr. Krishnapal Karmodiya at IISER Pune. This laboratory works on *Plasmodium falciparum* and was interested in the development of an enzymatic system that would be able to selectively enrich concentrations of *Plasmodium* DNA before sequencing field isolates.

This project could further downstream lead to the development of a kit that could be used in a variety of applications to separate methylated and unmethylated DNA.

Chapter 2: Materials and Methods

2.1 Bacterial Genomic DNA extraction

DH5α and NEB Turbo genomic DNA were extracted using the QIAGEN® 100/G Genomic-tips. 5 ml bacterial cultures in LB media were grown overnight at 37° C. The cultures were then pelleted down by centrifugation for 5 minutes at 8,000 RPM. The pellets were resuspended in 3.5 ml of Lysis Buffer B1 (50 mM TrisCl, pH 8.0; 50 mM EDTA, pH 8.0; 0.5% Triton X-100; 0.5% Tween®-20) containing 0.2 mg/ml RNAse A. 100 µl of Proteinase K solution (20 mg/ml) and 80 µl of Lysozyme stock solution (100 mg/ml) were added and the samples were incubated at 37° C for 30-45 minutes. 1.2 ml of Buffer B2 (3 M guanidine HCl; 20% Tween-20) was added and the mixed samples were added to the QIAGEN® Genomic-tip columns that had been equilibrated with 4 ml of QBT Buffer (750 mM NaCl; 50 mM MOPS, pH 7.0; 0.15% Triton X-100;15% isopropanol). 2x7.5 ml QC buffer (1.0 M NaCl; 50 mM MOPS, pH 7.0; 15% isopropanol) was used to wash the column twice. The DNA was then eluted with 5 ml of buffer QF (1.25 M NaCl; 50 mM Tris-Cl, pH 8.5; 15% isopropanol) warmed to 50°C. 0.7 volumes room temperature isopropanol was used to precipitate the DNA via centrifugation at 10,000 g at 4°C for 15 minutes. 500 µl of cold 70% ethanol was used to wash the DNA in each 2 ml tube, and the samples were centrifuged at 10,000 g at 4° C for 10 minutes. The supernatant was removed and the samples were air-dried to remove any residual ethanol. The DNA was then resuspended in 50 µl of buffer EB (10 mM Tris-Cl) by shaking at 300 rpm at 55° C for 2 hours. The concentration of the DNA was measured using the Nanodrop™ 2000 C and the samples were visualised via electrophoresis using a 1% Agarose gel.

2.2 HEK 293 Genomic DNA extraction

HEK293 pellet was obtained from the lab of Dr. Mayurika Lahiri. HEK293 cells were cultured at 37°C and 5% CO_2 concentration. The media used to culture the cells was a high glucose DMEM (Dulbecco's Modified Eagle Medium) that contained 10% FBS (fetal bovine serum) and 1x penicillin-streptomycin (pen-strep). 5-6 million cells were pelleted down at 1,000 RPM for 5 minutes. The Genomic DNA was extracted using the QIAGEN® 100/G genomic tips. The pellet was washed with Phosphate Buffered Saline (PBS) and was resuspended in cold PBS to a final concentration of 10^7 cells/ml. 1 Volume ice-cold C1 buffer (1.28 M sucrose; 40 mM Tris-Cl, pH 7.5; 20 mM $MgCl_2$; 4% Triton X-100) and 3 volumes of

ice-cold ethanol were added to the sample and mixed, following which the sample was centrifuged at 1300xg for 15 minutes at 4°C. 1 ml of ice-cold C1 buffer and 3 ml of ice-cold distilled water were added, followed by 5 ml of G2 buffer (800 mM guanidine HCl; 30 mM Tris-Cl, pH 8.0; 30 mM EDTA, pH 8.0; 5% Tween-20; 0.5% Triton X-100). 95 µl of Proteinase K solution (20 mg/ml) was then added. The sample was then incubated at 50°C for around 45 minutes. The sample was then loaded onto a QIAGEN® Genomic-tip column that had been equilibrated with 4 ml of buffer QBT. The same protocol as for bacterial genomic DNA extraction was then followed. The concentration of the DNA was measured using the Nanodrop™ 2000 C and the samples were visualised via electrophoresis using a 1% Agarose gel.

2.3 SauUSI substrate

As controls to verify the activity of SauUSI, a linear substrate, SauS was used. This substrate was a linearised form of a pHis17 plasmid. The plasmid was methylated in vivo by culturing it in the dcm+ *Escherichia coli* strain DH5α. Dcm is a bacterial methyltransferase that recognises the DNA sequence 5'-CCAGG-3' and methylated the first Cytosine residue in the sequence, thereby converting the sequence to 5'-C^{5m}CAGG-3'. This methylated sequence is a subset of the recognition site of SauUSI, 5' -S^{5m}CNGS-3'.

The methylated plasmids were obtained via transformation into DH5α E. coli. The plasmid was a derivative of the pHis17 vector, which is ampicillin resistant. After transformation, plates were incubated for around 12 hours at 37°C, following which a single colony was picked and inoculated in LB containing 0.1 mg/ml of Ampicillin. The cells were pelleted following growth of the culture overnight. QIAPrep® miniprep spin protocol was used to obtain the plasmids. The SauS substrate was obtained by linearizing this plasmid. NcoI-HF was used to linearize the plasmid for 2 hours following NEB protocol. The enzyme was then denatured by heating the reaction to 65°C for 20 minutes.

2.4 McrBC substrate

As a control to verify the activity of McrBC, a 1127-bp substrate termed RmC was used. Previous work in the laboratory had included cloning a version of the McrB protein lacking its N terminal, termed $McrB_s$, into a pHIS17 vector with a T7 promoter and T7 terminator. This 1127 bp substrate was initially amplified from pHIS$McrB_s$ plasmid by PCR. The sequence amplified was flanked by T7 promoter and terminator regions, so primers for these regions

– T7-Forward (5'-TAATACGACTCACTATAGGG-3') and T7-Reverse (5'-GCTAGTTATTGCTCAGCGG-3') – were used to amplify this region. This substrate had randomly methylated Cytosine residues in its sequence due to the use of 5-methylcytosine instead of deoxycytosine in the dNTP mix.

To amplify the amounts of this substrate to use as a control, a 50 µl PCR reaction was set up using small amounts of the substrate as template DNA. Just as the previous PCR, dNTPs containing 5-methylcytosine were used to randomly methylate certain cytosines in the substrate. Confirmation of amplification was done by visualising the product on a 1% Agarose gel. QIAGEN® PCR Purification kit was used to purify the amplified PCR product.

2.5 SauUSI purification

Existing SauUSI stocks from the laboratory were used for all the cleavage assays. The *SauUSIR* gene was cloned via restriction-free cloning in a pHis17 (ampicillin resistant) vector and was followed by a BamH1 site and a C-terminal 6xHis tag, and flanked upstream by an NdeI site. Sequencing of the positive clone confirmed the presence of the SauUSIR gene.

For the sake of purification, SauUSI was expressed in a 2 litre culture of *E.coli* BL21 (AI) cells. Three-column purification strategy at 4°C was used to purify SauUSI. The pellet was obtained and lysed via sonication. The three columns used for the purification were the GE HisTrap™ HP 5 ml column, a GE MonoQ™ 10/100 GL column and a GE Superdex 200™ increase 10/300 GL column, in that order. The protein was eluted in Buffer B100 (100 mM NaCl, 50 mM Tris-HCl, pH 8; 1 mM DTT) following the third column. The purified protein was analysed on a 10% SDS-PAGE gel. After concentrating the protein, the stocks were flash frozen using liquid nitrogen for storage at -80°C.

For the purpose of cleavage assays, SauUSI was diluted along with buffer B100 to create substocks with concentrations of 1 µM.

2.6 McrB expression and purification

Existing McrB stocks from the laboratory were used for all the cleavage assays. The *mcrB* gene was cloned with a 6xHis tag into a pHis17 vector. McrB was expressed in a 2 litre culture of *E.coli* BL21 (AI) cells. Once a secondary culture reached an OD of 0.3 at 600 nm,

the culture was incubated at 18°C and 0.06% Arabinose was added to the culture to induce expression of McrB. The culture was then grown further overnight. The culture was pelleted via centrifugation and was lysed in a sonicator following the addition of 0.04% CHAPS and 50 ml of Lysis buffer. The sample was then ultracentrifuged and the supernatant was added to a 5 ml Ni-NTA column that had been equilibrated with Buffer A (500 mM NaCl; 50 mM Tris-Cl, pH 8; 25 mM imidazole). The protein was eluted with buffer A and buffer B (500 mM NaCl; 50 mM Tris-Cl, pH 8; 500 mM Imidazole) via a step gradient. The purest fraction was dialysed with dialysis buffer at 4°C for 2 hours. It was then loaded onto an 8 ml MonoQ™ column 10/100 GL column that had been equilibrated with Buffer B50 (50 mM NaCl; 50 mM Tris-Cl, pH 8; 1 mM EDTA; 1 mM dithiothreitol (DTT)). The sample was eluted using B50 and B1000 (1000 mM NaCl; 50 mM Tris-Cl, pH 8; 1 mM EDTA; 1 mM DTT) in a linear gradient and collected in 2 ml fractions of 20 column volumes. Pure fractions were collected, concentrated and then washed with storage buffer (100 mM NaCl, 10 mM Tris-Cl pH 7.4 and 1 mM DTT). Bradford reagent was used to measure the concentration of the protein. The sample was washed until the A260/A280 ratio was confirmed to be less than 0.5, representing removal of dNTPs. Samples were flash frozen and stored at -80°C.

2.7 McrC expression and purification

Existing McrC stocks from the laboratory were used for all the cleavage assays. The *mcrC* gene was cloned into a pHis17 vector. McrB was expressed in a 2 litre culture of *E.coli* BL21 (AI) cells. Once a secondary culture reached an OD of 0.3 at 600 nm, the culture was incubated at 18°C and 0.06% Arabinose was added to the culture to induce expression of McrC. The culture was then grown further overnight. The culture was pelleted via centrifugation and was lysed in a sonicator following the addition of 0.04% CHAPS and 50 ml of Lysis buffer. The sample was then ultracentrifuged and the supernatant was added to a 5 ml Ni-NTA column equilibrated with Buffer A. The protein was eluted with buffer A and buffer B via a step gradient. The purest fraction was dialysed with dialysis buffer at Room temperature for 45 minutes. It was then loaded onto an 8 ml MonoS column 10/100 GL column that had been equilibrated with Buffer B50. 2 ml fractions were eluted using B50 and B1000 in a linear gradient. Pure fractions were collected, concentrated and washed with storage buffer. Bradford reagent was used to measure the concentration of the protein. Samples were flash frozen and stored at -80°C.

2.8 McrBC complex formation

Stocks of McrB and McrC proteins were thawed, centrifuged, and transferred to fresh tubes. McrB was centrifuged at 15,000 RPM for 10 min at 4°C. McrC was centrifuged at 14,500 RPM for 8 minutes at Room Temperature. This was done to remove any protein that had precipitated. Concentrations of the two proteins were measured using Nanodrop™ 2000 C following denaturation with 6M Guanidinium Chloride for 10 minutes. Concentrations were measured using an extension coefficient of 13.

McrB and McrC were used in a 4:1 Molar ratio to form the complex as follows. Cold McrB and TMDK buffer (10 mM Tris-Cl, pH 8; 5mM $MgCl_2$; 1 mM DTT; 50 mM KCl) were mixed together and allowed to come to room temperature for 2-3 minutes. Following the addition of McrC, the complex was incubated at room temperature for 2-3 minutes and then kept on ice until use.

For a majority of assays, 50 µl of a 416.66 nM concentration of McrBC was prepared for cleavage assays by mixing 5 µM and 1.25 µM of McrB and McrC respectively. The remaining volume was made up using TMDK buffer.

2.9 Ethanol precipitation

The samples were made up to a volume of 50 µl with autoclaved nuclease-free water. 0.1 volumes of 3M Sodium Acetate and 3 volumes of ice-cold 100% ethanol were added to the samples, which were then flash-frozen in liquid Nitrogen. After being stored at -80°C for 1 hour, the samples were centrifuged at 13,000 RPM for 30 minutes at 4°C. The supernatant was then discarded and 500 µl cold 70% ethanol was carefully added to wash the pellet along the sides of the tubes. The samples were centrifuged at 13,000 RPM for 10 minutes at 4°C. The supernatant was removed and the samples were air-dried to remove any residual ethanol. The DNA was then resuspended in 10 µl of nuclease-free water by shaking at 300 rpm at 55 °C for around 45 minutes.

2.10 Isopropanol precipitation

A protocol similar to that of ethanol precipitation was followed for isopropanol precipitation, with the difference of addition of cold isopropanol instead of ethanol prior to flash freezing.

2.11 Cleavage assays

Cleavage assays were performed with SauUSI and McrBC. A SauUSI substrate (SauS) was used as a control for the activity of SauUSI, and an McrBC substrate (RmC) was used as a control for the activity of McrBC. HEK293 and NEBTurbo genomic DNA were used as methylated and unmethylated substrates respectively to optimise the activities of these enzymes.

Multiple variables in the cleavage assays needed to be tested and optimised. Concentrations of the enzymes, cleavage buffers, incubation periods and combinations of the enzymes.

To optimise the activity of SauUSI and McrBC, multiple cleavage buffers were tested out. Previously, the enzymes were shown to have optimal activity in CutSmart and NEB4 buffers for SauUSI, and in TMDK buffer for McrBC.

To optimise concentrations of the enzymes, multiple cleavage assays were set up with differing concentrations of the enzymes, as mentioned in the figures.

For the purpose of these experiments, combined activities of SauUSI and McrBC were tested. Thus, cleavage assays for SauUSI, McrBC and combinations of SauUSI and McrBC were tried out with various buffers - NEB4 (50 mM potassium acetate; 20 mM Tris-acetate; 10 mM magnesium acetate; 1 mM DTT, pH 7.9), TMDK (10 mM Tris-Cl, pH 8; 5mM $MgCl_2$; 1 mM DTT; 50 mM KCl) and CutSmart (50 mM potassium acetate; 20 mM Tris-acetate; 10 mM magnesium acetate; 1 mM DTT, pH 7.9; 1 mg/ml BSA).

Cleavage assays were set up in 10 µl of reaction if individually done for SauUSI or McrBC, and 20 µl if set up for combinations of the two enzymes. The constituents of the reactions were 1x buffer, 2 mM ATP and/or 2 mM GTP, varying concentration of enzymes, and 120-150 ng of DNA. The reactions were incubated at 37°C for varying amounts of time, and the enzymes were denatured by heating the reactions to 65°C for 20 minutes. 1x Purple Gel loading dye was added, and the cleavage fragments were visualised and analysed on a 1% agarose gel.

2.12 AMPure XP based size selection:

In order to remove the sheared small fragments of methylated DNA left behind in the reaction mixtures following the cleavage assays, AMPure XP beads were used.

Based on data conveying the length of DNA fragments that are eluted using different concentrations of AMPure XP beads, a 0.6x concentration of beads was used.

Constituents of each 10 µl reaction were transferred to 0.2 ml tubes, and 6 µl of AMPure XP beads were added after vortexing the beads thoroughly. The beads were mixed with the samples via pipetting. The tubes were then placed in a magnetic stand and incubated until the solution became visibly clear and the beads stuck to the inside of the tubes. The supernatant of the solution was carefully discarded and 200 µl of freshly prepared 70% Ethanol was used to wash the beads, and the supernatant was discarded. This wash step was repeated a second time, and the samples were taken off the magnetic stand. The beads were mixed with 40 µl of nuclease-free water and incubated at room temperature for 2-3 minutes. The sampled were then transferred back onto the magnetic stand. After roughly 10 minutes, the supernatant was collected and vacuum centrifuged to remove any residual ethanol, and resuspended in 10 µl of nuclease-free water. The DNA was dissolved in the water by shaking at 65°C for 2 hours.

2.13 PCR amplification

To confirm the presence of HEK and NEBTurbo genomic DNA in the reaction mixtures following the cleavage assays, PCR amplifications of genes or gene fragments specific to Human genomic DNA or *E.coli* genomic DNA were performed.

Name of primer	Sequence of primer
Hexim1_F	CCAAGAAGAAGCGGCATTGG
Hexim1_R	ACTGCGTGGTGTTATAGGGC
FtsZ_F	ATGTTTGAACCAATGGAACTTAC
Ftsz_R	ATCAGCTTGCTTACGCAGGAATG
FtsZ_part_F	CTGAGATGGGCTACGCAATGTTCGGTTCTGGCGTGGCGAGCGGTG

Table 1. List of primers used

For Human genomic DNA, a set of primers corresponding to a 150 bp fragment from the Hexim-1 gene, a gene having an important role in development and differentiation [Michels et al., 2018] were used. For NEB Turbo genomic DNA, a set of primers corresponding to a 1.1 kb long *E.coli* gene, FtsZ, were used. The sequence of the primers is listed in Table 1.

PCR to confirm the presence of HEK gDNA was performed at an annealing temperature of 60°C, with an extension time of 15 seconds. 35 cycles of denaturation, annealing and extension were used.

PCR to confirm the presence of NEBTurbo gDNA was performed at an annealing temperature of 58°C, with an extension time of 1 minute. 35 cycles of denaturation, annealing and extension were used.

PCR was performed downstream of the cleavage assays to check for the presence of gDNA. The entire constituents of cleavage assays following heat inactivation of the restriction enzymes were directly used for PCR reactions.

To check the difference in template amounts using protocols of ethanol and isopropanol precipitation following cleavage assays, the gDNA resuspended in 10 µl of nuclease-free water (as described in ethanol precipitation and isopropanol precipitation protocols) was used as the template for those PCR reactions.

Constituent	Amount added / 50 µl PCR	Final concentration
10x pfu buffer	5 µl	1x
dNTPs (2.5 mM each)	2 µl	100 µM each
20 µM Forward Primer	1 µl	400 nM
20 µM Reverse primer	1 µl	400 nM
Pfu	0.8 µl	1 unit/50 µl

Table 2. Constituents of PCR reactions

10 µl of each PCR product was mixed with 1x Gel purple loading dye and visualised on a 1% agarose gel.

Chapter 3: Results

3.1 Extraction of Human and bacterial genomic DNA (gDNA)

Using the protocols mentioned in the materials and methods section, genomic DNA was extracted from *E.coli* DH5α strain, *E.coli* NEBTurbo strain, and HEK293 cells. Concentration of the genomic DNA was measured using the Nanodrop2000^{TM} and stocks were made of ~150 ng/µl. The genomic DNA obtained was visualized on 1% Agarose gels, and the bands on the gel were seen to be free of visible contamination.

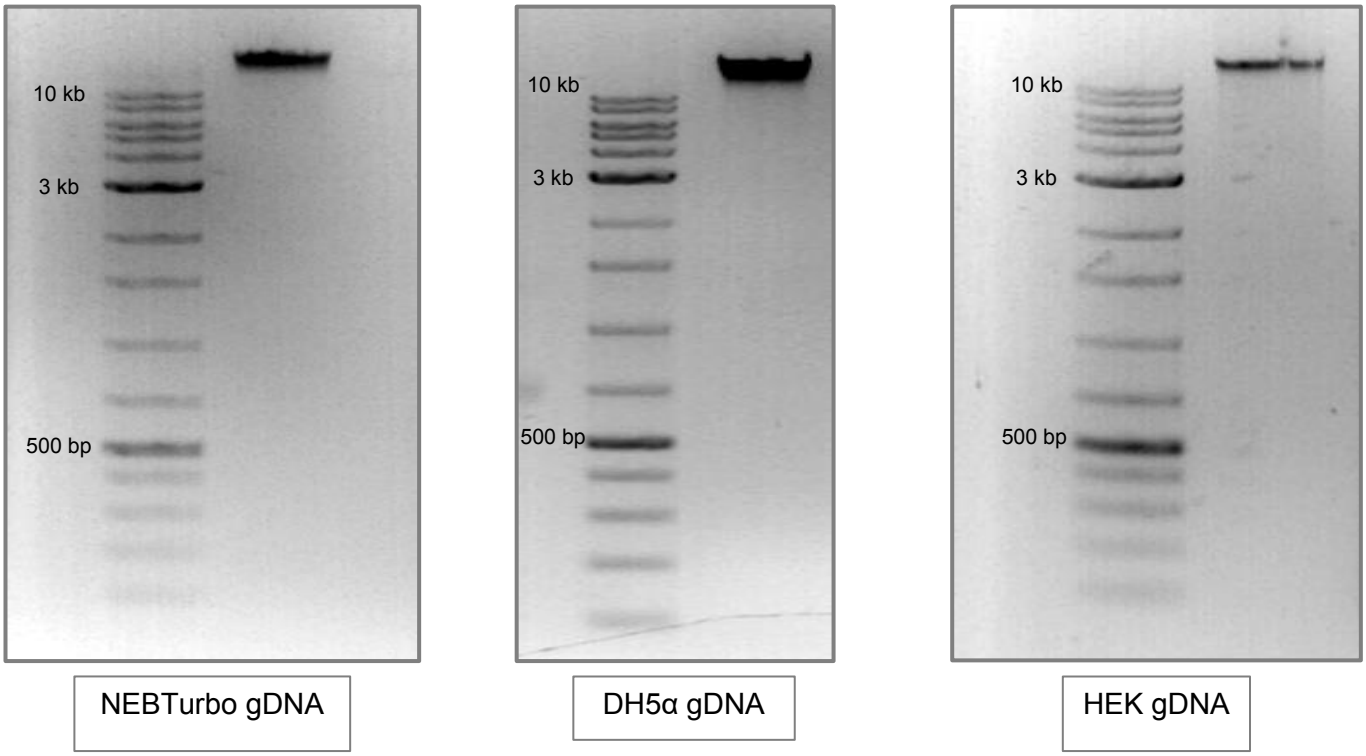


Figure 3.Genomic DNA was extracted from HEK293 cells and E.coli NEBTurbo and DH5α strains. 150 ng of DNA was loaded on a 1% agarose gel.

3.2 Shredding of methylated bacterial genomic DNA with SauUSI

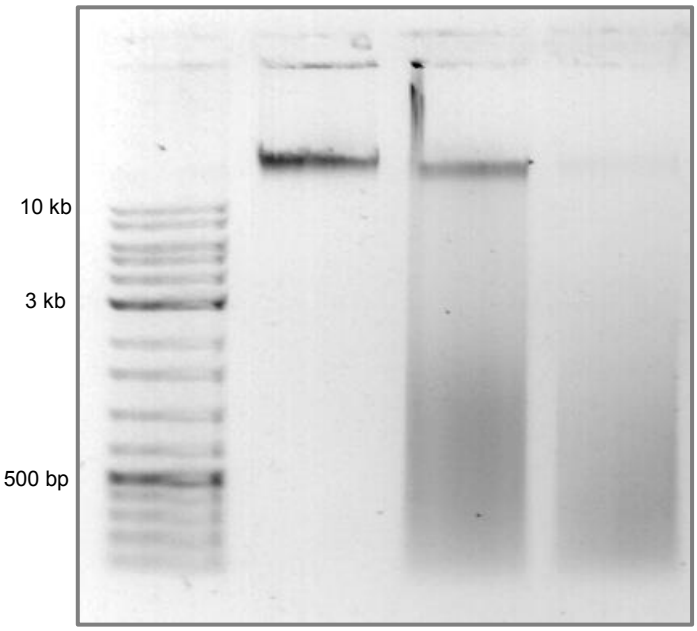
As an initial proof of the shredding effects of SauUSI on methylated genomic DNA, DH5α and NEBTurbo genomic DNA was treated with SauUSI.

DH5α and NEBTurbo genomic DNA differ in their methylation status. DH5α is a dcm+ strain, while NEBTurbo is a dcm- strain. Dcm is a bacterial methyltransferase enzyme that is

responsible for methylating cytosine residues. It methylates Cytosine residues within the recognition site 5'-CCAGG-3', converting it to 5'-C^{5m}CAGG-3'. This sequence is a subset of SauUSI's recognition site 5' -S^{5m}CNGS-3'. The absence of dcm methyltransferase and any other methyltransferase that would create methylated sequences susceptible to cleavage by SauUSI in NEBTurbo should ensure the NEBTurbo genomic DNA is not shred by SauUSI.

Previous data from the laboratory had shown that at concentrations of SauUSI that are 400 nM or higher, there is an inhibitory effect on the activity of the enzyme. Due to this, we set up reactions where DH5α genomic DNA was treated with two different concentrations of SauUSI – 50 and 200 nM.

We observed that 200 nM of SauUSI cleaved the DNA completely, as visible in Figure 4.



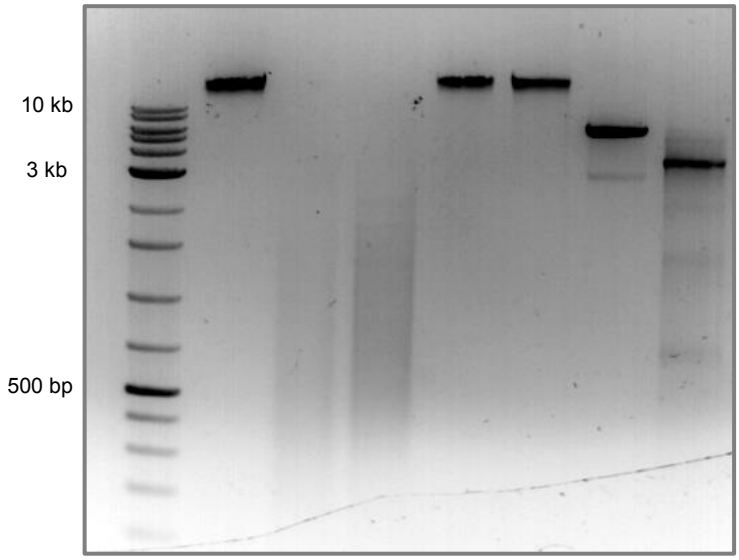
DH5α gDNA	+	+	+
NEB4	+	+	+
SauUSI	–	50 nM	200 nM
ATP (2 mM)	–	+	+

Figure 4. Cleavage pattern observed on incubation of DH5α genomic DNA with 50 and 200 Nm of SauUSI for 1 hour. Samples were visualized on a 1% Agarose gel.

In order to visualize the shredding of DNA by SauUSI as compared to the restriction pattern that would be obtained due to cleavage only at the restriction sites, we digested DH5α genomic DNA with BstN1, a Type II restriction endonuclease. BstNI recognizes the target site 5'-CCWGG-3' (dcm site) and hence can be used to compare cleavage of SauUSI (which recognizes dcm sites as a subset of its target recognition repertoire) on genomic

DNA. Although BstNI recognizes identical sites to SauUSI and cleaves at the target sites, the bands produced by BstNI are of higher molecular weight. This indicates that SauUSI has the ability to shred DNA to smaller molecular weights, as shown earlier with two-site and three-site substrates in Tumuluri et al. 2021.

Cleavage assays were set up using 200 nM SauUSI and BstN1 for DH5α genomic DNA and 200 nM SauUSI on NEBTurbo genomic DNA (Figure 4). Buffer NEB4 was used for SauUSI and NEB3.1 buffer was used for BstN1. As a control to confirm the activity of SauUSI, 100 nM of SauUSI was used with the SauUSI substrate.



DH5α gDNA	+	+	+	–	–	–	–
NEBTurbo gDNA	–	–	–	+	+	–	–
SauS	–	–	–	–	–	+	+
Buffer	NEB4	NEB4	NEB3.1	NEB4	NEB4	NEB4	NEB4
SauUSI	–	+ 200 nM	–	–	+	–	+ 100 nM
2 mM ATP	–	+	–	–	+	–	+
BstN1	–	–	+	–	–	–	–

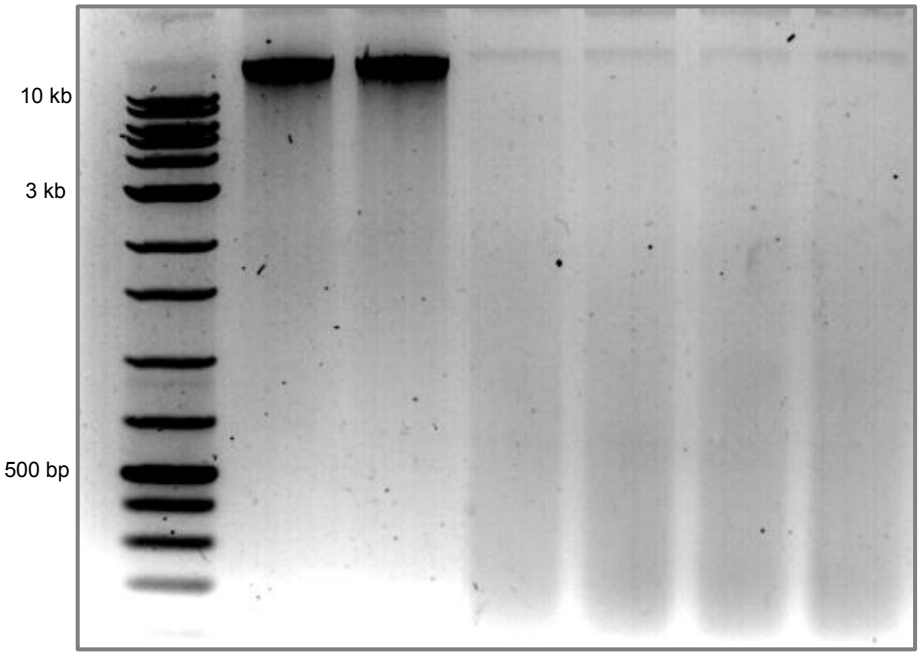
Figure 5. DH5α and NEBTurbo genomic DNA were treated with 200 nM of SauUSI. Methylated DH5α genomic DNA is shredded, whereas unmethylated NEBTurbo genomic DNA is intact. DH5α gDNA treated with BstNI (Type II restriction endonuclease) shows distinct bands on the gel.

Our experiment showed that methylated bacterial genomic DNA did indeed get shredded by SauUSI, whereas unmethylated NEBTurbo genomic DNA did not. Treatment of DH5α genomic DNA with BstN1 led to cleavage of the genomic DNA. The sheer number of

BstN1 sites in the genome of DH5α causes the DNA to cleave into small fragments, but distinct bands are visible in Lane 3. This is in stark contrast to Lane 2, which confirms the shredding activity of SauUSI.

3.3 Shredding of Human genomic DNA with SauUSI

Having already determined that 200 nM of SauUSI showed us optimal cleavage with methylated DH5α genomic DNA, we set up cleavage assays with 200 nM of SauUSI incubated for different durations. As observed in Figure 6, there did not seem to be much difference in cleavage of incubated DNA when incubated for 1 hour as compared to 4 hours.



HEK gDNA	+	+	+	+	+	+
NEB4	+	+	+	+	+	+
SauUSI (200 nM)	–	+	+	+	+	+
ATP (2 mM)	–	–	+	+	+	+
Incubation period	4 h	4 h	1 h	2 h	3 h	4 h

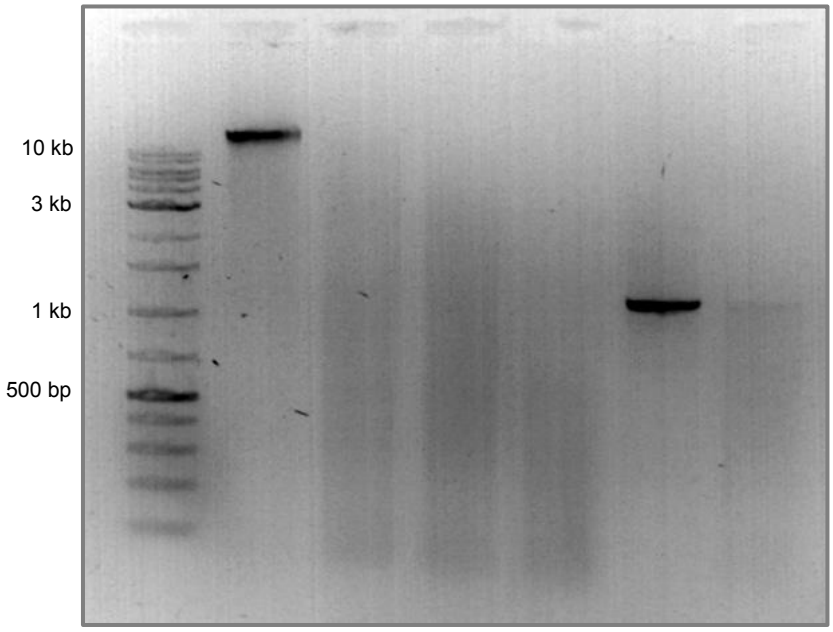
Figure 6. Treatment of HEK293 gDNA with 200 nM of SauUSI in NEB4 buffer for varied incubation periods (1 hour, 2 hours, 3 hours and 4 hours). Samples were visualized on a 1% Agarose gel.

3.4 Optimising HEK gDNA cleavage using McrBC

McrBC, another methylation dependent restriction endonuclease, was used to test the cleavage of Human genomic DNA. McrBC is a tetradecameric enzyme consisting of two

hexamers of McrB complexed with a dimer of the endonuclease McrC. Existing stocks of McrB and McrC from the laboratory were used for all of these experiments.

We set up reactions of HEK Genomic DNA with varying concentrations of McrBC in order to find the most optimal concentration of McrBC for cleavage of HEK DNA. Reactions were set up with 50 nM, 100 nM and 200 nM of McrBC incubated at 37°C for 2 hours. Previous experimentation in the lab had confirmed cleavage of McrBC substrate RmC with 50 nM of the protein. As seen in Figure 7, 200 nM was seen to have the most efficient shredding of the DNA. Cleavage of McrBC substrate RmC with 50 nM of McrBC was used as the positive control.

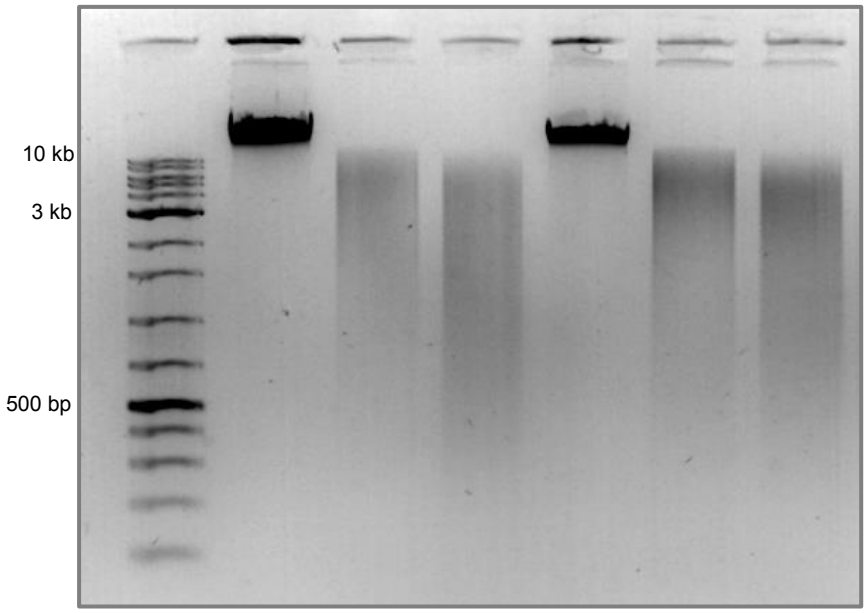


HEK gDNA	+	+	+	+	–	–
RmC	–	–	–	–	+	+
TMDK Buffer	+	+	+	+	+	+
McrBC	–	50 nM	100 nM	200 nM	–	50 nM
2 mM GTP	–	+	+	+	–	+

Figure 7. Treatment of HEK gDNA with 50 nM, 100 nM and 200 nM of McrBC for 2 hours. The McrBC substrate RmC was used as a control for the activity of McrBC. Samples were visualized on a 1% Agarose gel.

We decided that 200 nM of McrBC when incubated with genomic DNA displayed the most efficient shredding. However, when setting up a control without GTP in an experiment later, we saw that there was enzymatic cleavage even in the absence of GTP, although McrBC is an enzyme dependent on GTP for cleavage.

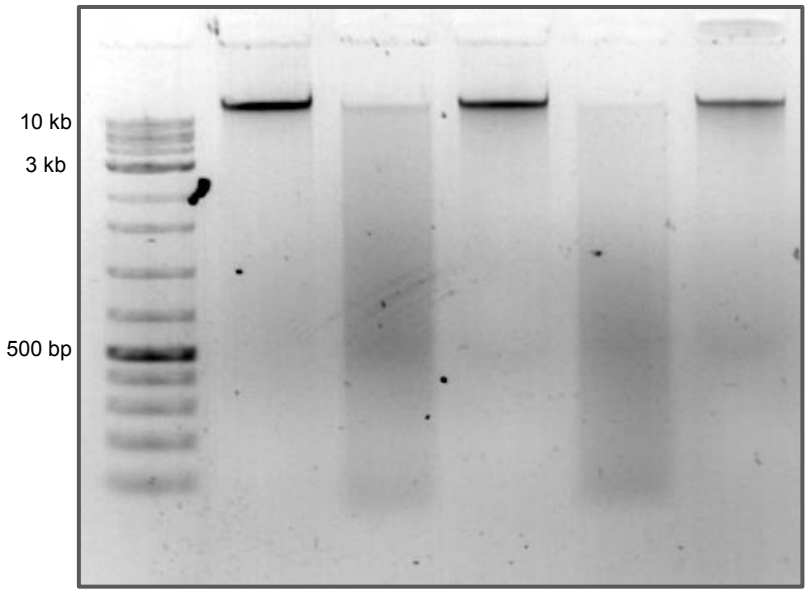
200 nM of McrBC even showed us cleavage in NEBTurbo® genomic DNA, which is unmethylated. We speculate that this observation is due to a non-specific nuclease contamination that could have inadvertently eluted out with our desired protein (data not shown).



DH5α gDNA	+	+	+	–	–	–
NEBTurbo gDNA	–	–	–	+	+	+
CutSmart buffer	+	+	+	+	+	+
McrBC 200 nM	–	+	+	–	+	+
2 mM GTP	–	+	–	–	+	–

Figure 8. Methylated DH5α gDNA and Unmethylated NEBTurbo gDNA were incubated with 200 nM McrBC for 2 hours. Surprisingly, both methylated and unmethylated DNA showed cleavage. Methylated DNA was cleaved in the absence of GTP as well. Samples were visualized on a 1% Agarose gel.

We reduced the concentration of McrBC to 50 nM in an attempt to dilute out any contaminating nuclease. We visually observed that this low concentration of McrBC, when incubated with methylated DNA for 1 hour, effectively shredded the DNA, while the control reaction containing the DNA but no GTP showed no cleavage. Further, increasing the time of the reaction to 2 hours seemed to cause a partial smear of the parent DNA band in the negative control, indicating non-specific activity of McrBC or other contaminating nucleases upon prolonged incubation. Thus, we fixed 50 nM of McrBC and an incubation period of 1 hour for cleavage reactions with McrBC.



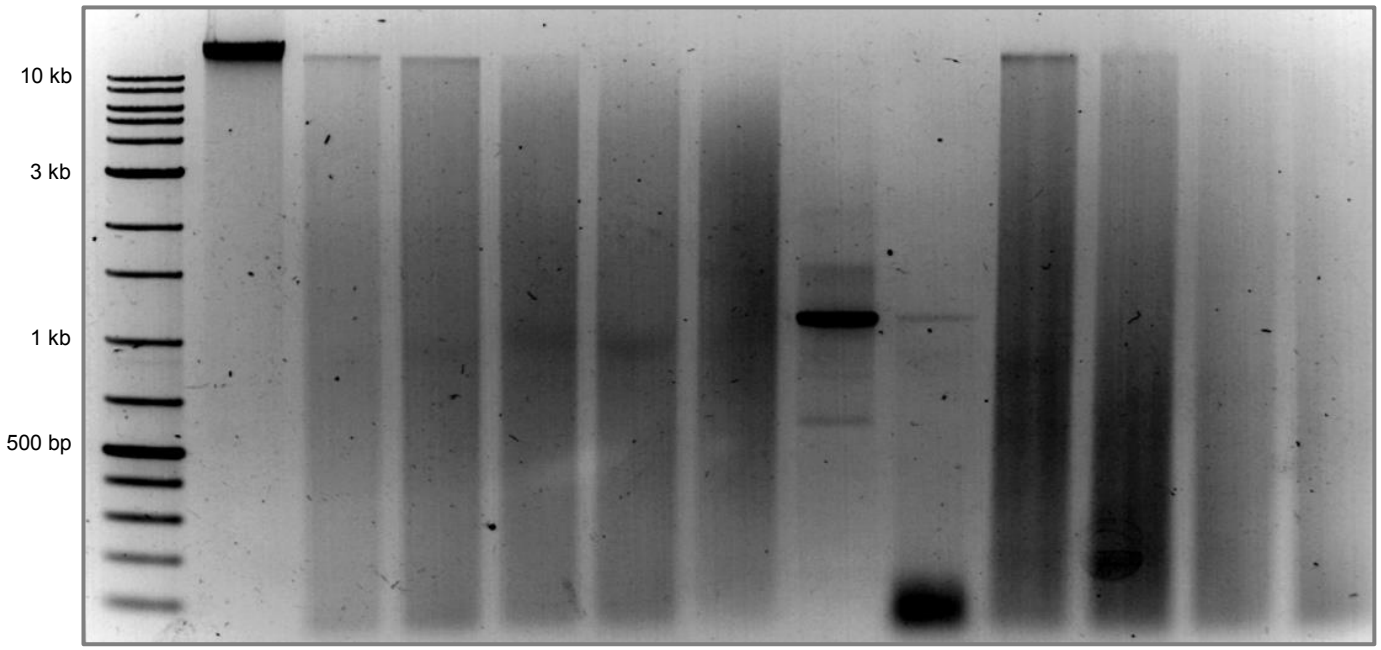
HEK gDNA	+	+	+	+	+
CutSmart	+	+	+	+	+
McrBC (50 nM)	–	+	+	+	+
2 mM GTP	–	+	–	+	–
Incubation period	2 hrs	1 hr	1 hr	2 hr	2 hr

Figure 9. HEK gDNA was incubated with 50 nM of McrBC in the presence and absence of GTP for 1 hour and 2 hours. Samples were visualized on a 1% Agarose gel.

3.5. Optimisation of buffers

Once we established the enzyme concentration and incubation time necessary for optimal cleavage of Human genomic DNA with SauUSI and McrBC, it was imperative to test the cleavage pattern when both these enzymes acted in tandem. To this end, we decided to optimize a common cleavage buffer for both these enzymes. Previous independent studies in the laboratory have shown that SauUSI works efficiently in the presence of NEB4 (50 mM potassium acetate; 20 mM Tris-acetate; 10 mM magnesium acetate; 1 mM DTT, pH 7.9) and CutSmart (50 mM potassium acetate; 20 mM Tris-acetate; 10 mM magnesium acetate; 1 mM DTT, pH 7.9; 1 mg/ml BSA) buffer whereas McrBC works efficiently in the presence of TMDK (10 mM Tris-Cl pH 8, 50 mM KCl, 5 mM MgCl2, 1 mM DTT).

Both the NEB4 and CutSmart buffer seemed to provide efficient cleavage with both enzymes. NEB4 and CutSmart are identical in composition apart from the 1 mg/ml BSA present in CutSmart. Most restriction endonucleases that show optimum activity in one of these buffers can also be used efficiently with the other buffer. We decided to proceed with the NEB4 buffer.



DNA	HEK	HEK	HEK	HEK	HEK	HEK	RmC	RmC	HEK	HEK	HEK	HEK
NEB4	+	–	–	–	+	–	–	–	+	+	+	–
CutSmart	–	–	–	–	–	+	–	–	–	–	–	+
TMDK	–	+	+	+	–	–	+	+	–	–	–	–
McrBC	–	50 nM	100 nM	200 nM	200 nM	200 nM	–	50 nM	–	–	–	–
GTP	–	+	+	+	+	+	–	+	–	–	–	—
SauUSI	–	–	–	–	–	–	–	–	50 nM	100 nM	200 nM	200 nM
ATP	–	–	–	–	–	–	–	–	+	+	+	+

Figure 10. In order to optimize a buffer that could be used for combinations of enzymes, the activity of SauUSI and McrBC to cleave HEK gDNA was tested in different buffers. The activity of 200 nM of McrBC was compared in TMDK, NEB4 and CutSmart buffers. The activity of SauUSI was tested in NEB4 and CutSmart buffers. RmC was used as a control to check the activity of McrBC. Samples were visualized on a 1% Agarose gel.

3.6 Optimising combination of enzymes

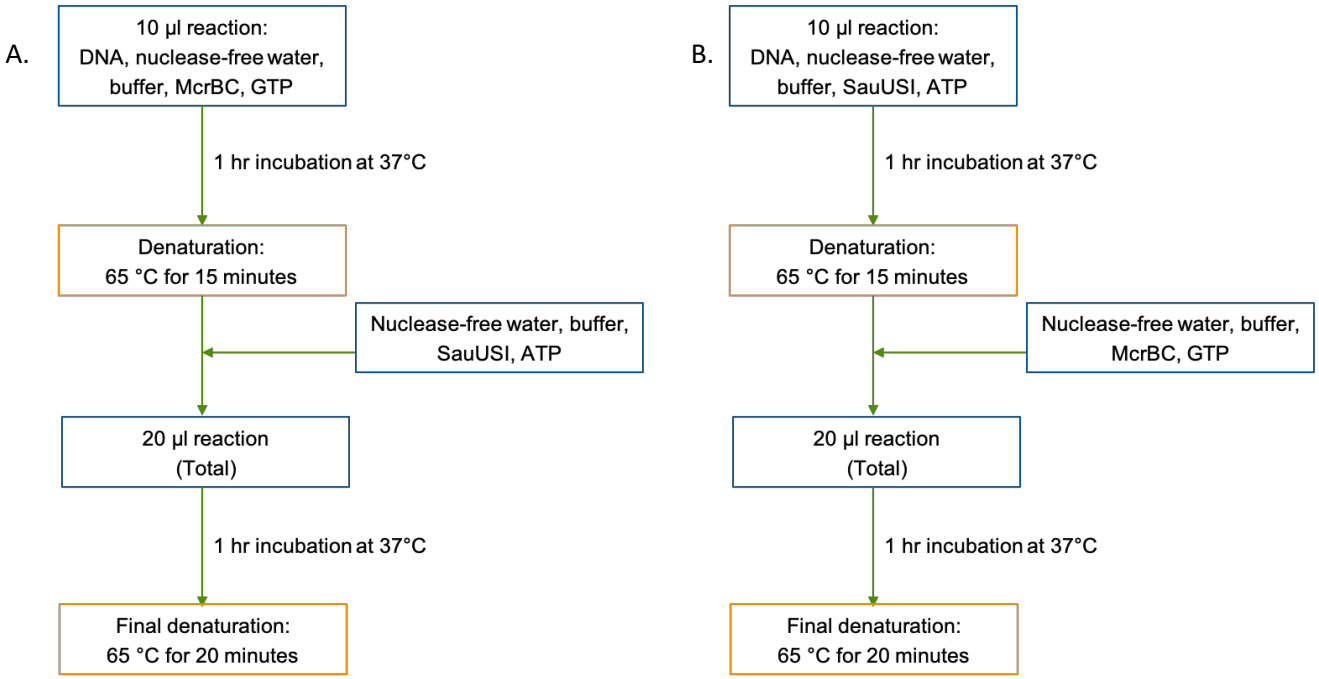
In order to best selectively enrich unmethylated DNA from a mixture of methylated and unmethylated DNA, the shredding of the methylated DNA must be optimised in order for the maximum amount of DNA to be shred into small fragments. Having already established the efficacy of SauUSI and McrBC at shredding HEK DNA and optimising the parameters of these reactions, we aimed to see whether treating methylated Human genomic DNA with a combination of the two enzymes would further improve the shredding observed. These combination reactions were tested in NEB4 buffer.

While testing out combinations of the reactions, we decided to proceed with a method in which one of the Type IV restriction endonucleases was added first and incubated for an hour, following which, it was denatured by heating the reaction at 65°C for 15 minutes. The second restriction endonuclease was then added. Denaturing one enzyme before adding the next enzyme can prevent the potential occluding of methylation sites as both these enzymes can have shared target sites on Human genomic DNA. Since the recognition site for SauUSI is 5'-S^{5m}CNGS-3', where S = G or C, and the recognition site for McrBC is R^{m}C, where R = A or G, any site with the sequence 5'-G^{5m}CNGS-3' could be bound to by either of the restriction enzymes. Thus, addition of the second enzyme without the denaturation of the first enzyme could prevent the second restriction endonuclease from binding to those sites, reducing the effect of the addition of the second enzyme.

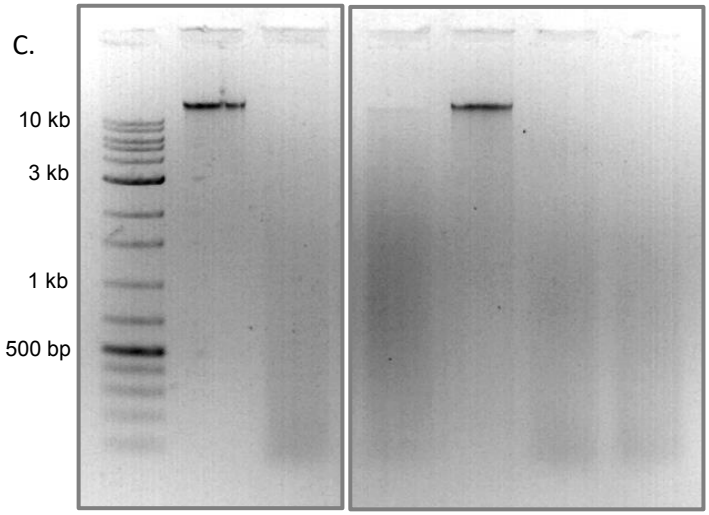
Due to this, we proceeded with denaturation-based combination assays. Initial combination assays were set up using 200 nM of SauUSI and 200 nM of McrBC and incubated for 1 hour + 1 hour, but due to the reasons mentioned in section 3.4, we then set up combination assays using 200 nM of SauUSI and 50 nM of McrBC.

The results of these combination assays on HEK and NEBTurbo gDNA can be seen in Figures 11 and 12 respectively. As visible in Figure 11, HEK gDNA shows the most shredding when treated with 200 nM of SauUSI and in the two combination reactions, named MI,SII and SI,MII. MI, SII refers to the combination reaction in which the genomic DNA is initially incubated with 50 nM of McrBC for 1 hour, following which McrBC is denatured and 200 nM of SauUSI is added to the sample (while doubling the reaction volume) and incubated for 1

hour. SI, MII refers to the combination reaction in which the genomic DNA is initially incubated with 200 nM of SauUSI for 1 hour, following which the SauUSI is denatured and 50 nM of McrBC is added to the sample (while doubling the reaction volume) and incubated for 1 hour. Figures 11 (B) and (C) schematically represent the same.



C.



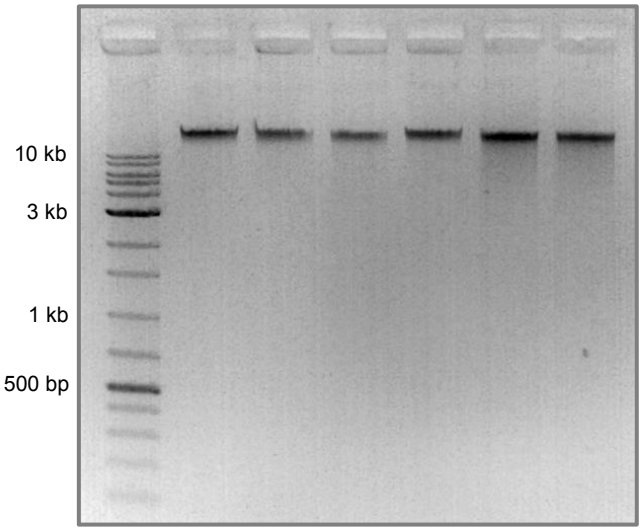
HEK gDNA	+	+	+	+	+	+
SauUSI (200 nM) 1 hour	–	+	–	–	–	–
ATP (2mM)	–	+	–	–	–	–
McrBC (50 nM) 1 hour	–	–	+	+	–	–
GTP (2 mM)	–	–	+	–	–	–
MI, SII*	–	–	–	–	+	–
SI, MII**	–	–	–	–	–	+

Figure 11. (A) Protocol in which sample was incubated with 50 nM of McrBC for 1 hour, following which protein was denatured and 200 nM SauUSI was added and sample was incubated for another hour.

(B) Protocol in which sample was incubated with 200 nM of SauUSI for 1 hour, following which protein was denatured and 50 nM McrBC was added and sample was incubated for another hour.

(C) HEK gDNA was incubated with 50 nM of McrBC, 200 nM of SauUSI and combinations of the two enzymes that were set up following the protocols shown in (A) and (B). MISII corresponds to the protocol shown in (A), and SIMII corresponds to the protocol shown in (B). Samples were visualized on a 1% Agarose gel.

We also incubated NEBTurbo genomic DNA with the same enzymes and combinations of enzymes to confirm that unmethylated DNA wasn't cleaved by the enzymes. Figure 12 confirms that unmethylated gDNA isn't cleaved by McrBC or SauUSI, or either of the two combinations of the enzymes tested.



NEBTurbo gDNA	+	+	+	+	+	+
SauUSI (200 nM) 1 hour	–		–	–	–	–
ATP (2mM)	–	+	–	–	–	–
McrBC (50 nM) 1 hour	–	–	+	+	–	–
GTP (2 mM)	–	–	+	–	–	–
MI, SII*	–	–	–	–	+	–
SI, MII**	–	–	–	–	–	+

Figure 12. NEBTurbo gDNA was incubated with 50 nM of McrBC for 1 hour, 200 nM of SauUSI for 1 hour, and combinations of the two enzymes that were set up following the protocol shown in Figure 11 (A) and (B). MISII corresponds to the protocol represented in 11 (A), and SIMII corresponds to the protocol represented in 11 (B). Samples were visualized on a 1% Agarose gel.

It is not possible to visibly identify which enzymatic cleavage setup results in the cleavage of HEK gDNA into the smallest fragments using an agarose gel. 200 nM SauUSI and the two combination reactions all show fragmentation of DNA that visibly look similar. PCR based amplification could be a semi-quantitative method of optimising the system among these three options.

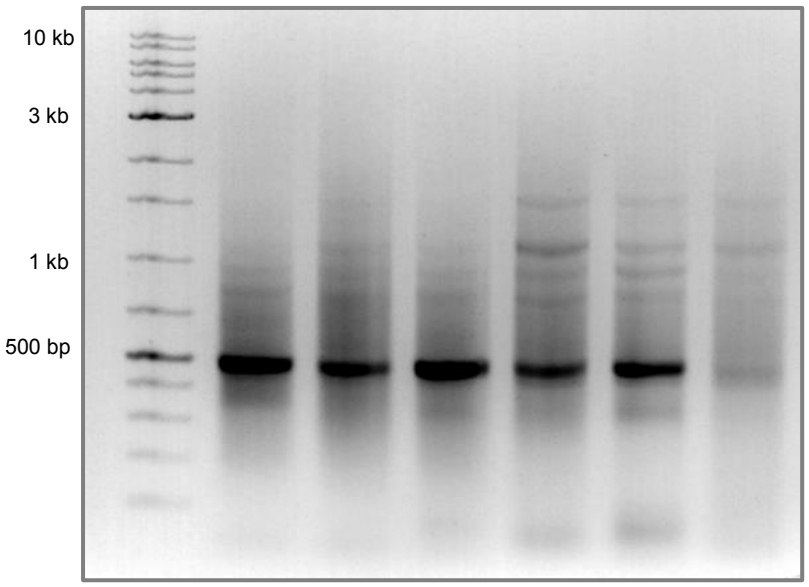
3.7 Confirming the presence of DNA via PCR and optimization of PCR

In order to provide evidence that unmethylated NEBTurbo gDNA could be selectively enriched from a mixture of HEK and NEBTurbo gDNA, the reduction in concentration of HEK gDNA could be shown via a reduction in amplification of a human-specific gene post cleavage with the Type IV restriction endonucleases. This could also be used on HEK gDNA after incubation with the optimised enzymes to test which reaction cleaved the DNA into the smallest fragments.

In order to optimize these PCRs, we picked a 150 bp fragment from the Hexim-1 gene particular to human genomic DNA, and the 1.1kb FtsZ gene from *E.coli*. Primers for these genes were obtained from existing stocks from our laboratory and the laboratory of Dr. Gayathri Pananghat respectively.

These PCRs would also provide a method of distinguishing between the multiple different optimized enzymatic reactions. As mentioned in Section 3.6, 200 nM of SauUSI and combination reactions of 50 nM of McrBC and 200 nM of SauUSI visually resulted in shredding of the methylated DNA. It is difficult to visually distinguish which system among these is best suited for the shredding of methylated DNA. Since these PCRs would be performed post-cleavage assays, there was a concern that the remnants of buffers from the cleavage assays would interfere with the efficacy of the PCR reactions set up. To confirm this, PCR reactions containing small amounts of NEB4 buffer were set up, corresponding to the buffer concentrations that would be present in the PCR reactions if half of the constituents of each 10 µl enzymatic reaction was amplified using PCR. Figure 13 shows the result of this experiment, confirming that even 0.5 µl of 10x NEB4 buffer in a 50 µl PCR reaction seemed to have an adverse effect on the amount of amplification occurring in the PCR reaction.

For this reaction, primers corresponding to a 500 bp fragment from the *E.coli* gene FtsZ were used (Fts_part_F and FtsZ_R); the sequences of the primers are listed in Table 1.



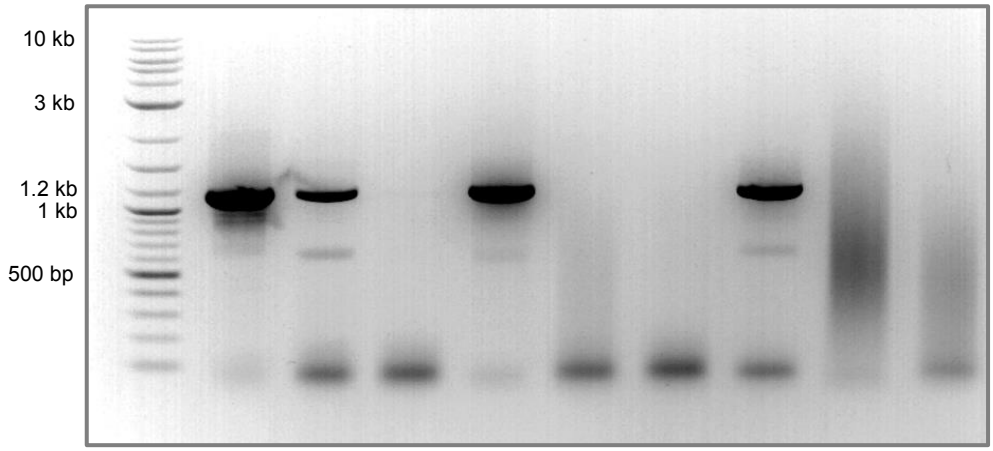
NEBTurbo gDNA	75 ng	75 ng	15 ng	15 ng	7.5 ng	7.5 ng
NEB4 amounts (50 µl reaction)	–	0.5 µl	–	0.5 µl	–	0.5 µl

Figure 13. Effects of NEB4 buffer on PCR amplification of a 500 bp fragment from E.coli gene FtsZ.

Due to effect of NEB4 buffer on PCR amplification, ethanol and isopropanol precipitation were tested as methods to precipitate the DNA in the given sample and remove any salts from the buffer. Varying concentrations of NEBTurbo genomic DNA were used - 150 ng, 15 ng and 1.5 ng. While testing out the efficacy of these methods at precipitating DNA, ethanol and isopropanol-precipitated DNA was used to set up PCR reactions where a 1.1 kb bp *E.coli* specific gene, FtsZ, was amplified (using primers FtsZ_F and FtsZ_R).

On comparing the PCR products on an Agarose gel, it was clear that the ethanol and isopropanol precipitation methods were reducing the amount of DNA present in the sample. HEK gDNA (control) showed us amplification of the fragment at 150 ng, 15 ng and 1.5 ng of template DNA. However, Ethanol precipitated HEK gDNA that was PCR amplified showed us amplification only at 150 ng and not at lower concentrations. Isopropanol precipitation did not seem to precipitate enough DNA to be amplified at all. (Figure 14)

Due to this, we decided to proceed with ethanol precipitation and test it out again. The yield from PCR was low once more, and this made us wonder if there was any advantage in attempting to remove the low salt concentrations at all.

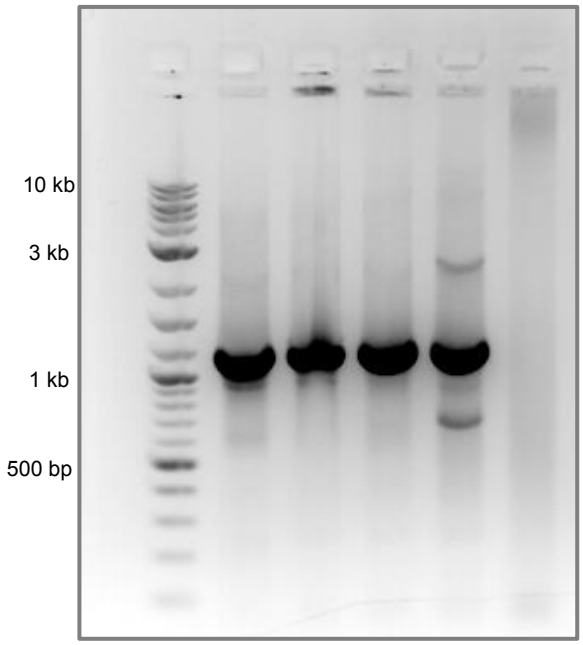


NEBTurbo gDNA amounts	150 ng	150 ng	150 ng	15 ng	15 ng	15 ng	1.5 ng	1.5 ng	1.5 ng
Method of precipitation (Eth – Ethanol, Iso – Isopropanol)	–	Eth	Iso	–	Eth	Iso	–	Eth	Iso

Figure 14. Differing amounts of NEBTurbo genomic DNA were precipitated using Ethanol and Isopropanol precipitation, following which the FtsZ gene was amplified using PCR. The products from PCR were visualized directly on a 1% Agarose gel.

For this, we set up PCR reactions with 150 ng of our DNA as a control, 150 ng of DNA with varying amounts of NEB4 buffer, and ethanol precipitated DNA. This experiment told us that the reduction in efficiency of PCR amplification with NEB4 buffer was less than the reduction in amplification from the loss in template DNA that occurred during ethanol precipitation, as there seemed to be no amplification from this precipitated DNA at all. (Figure 15)

Even though the exact same protocol was followed each time, we speculate that the small amounts of DNA, and the inability to clearly distinguish a pellet, might have resulted in the loss of DNA.



NEBTurbo gDNA	150 ng	150 ng	150 ng	150 ng	150 ng – Ethanol precipitated
NEB4 amounts (50 µl reaction)	–	0.2 µl	0.5 µl	1 µl	–

Figure 15. The E.coli FtsZ gene was amplified via PCR from 150 ng of NEBTurbo gDNA with varying NEB4 concentrations. This was compared to 150 ng of NEBTurbo gDNA that was precipitated using Ethanol precipitation

When a mixture of methylated and unmethylated DNA is incubated with Type IV restriction endonucleases, the methylated DNA is cleaved into small fragments. In order to use PCR as a semi-quantitative measure of the uncleaved methylated DNA, the shredded DNA must be removed from the sample.

While deciding on the method of post-cleavage assay clean-up of the shredded DNA, AMPure XP beads were selected as the method to remove the shredded fragments while preserving as much of the unmethylated DNA as possible. The protocol for AMPure XP beads has an ethanol wash step to wash any salts that remain in the solution. Due to this, we decided that the DNA precipitation step was unnecessary, and wouldn't be a needed step post cleavage assay.

Thus, we were able to test and optimise the PCR reactions that would provide a method of confirming the shredding of methylated DNA. Once a mixture of methylated and unmethylated DNA is treated with an optimized system of enzymes, PCR can be used to amplify a sequence specific to the methylated genomic DNA. If this methylated DNA is cleaved into small fragments, the amount of amplification of this sequence will be reduced as compared to DNA that has not been shred.

3.8 Post-cleavage assay cleanup of DNA

The cleavage assays shred methylated DNA in the sample. In order to develop a system which truly can separate methylated and unmethylated DNA post the shredding of methylated DNA, the shredded fragments need to be removed from the sample. Any quantitative method of showing the selective enrichment of unmethylated DNA from a mixture of methylated and unmethylated DNA, such as PCR or Next Generation Sequencing, would require the removal of these fragments. Next Generation Sequencing, for example, requires that the DNA to be sequenced is cleaved into fragments of around 200-300 bp. The

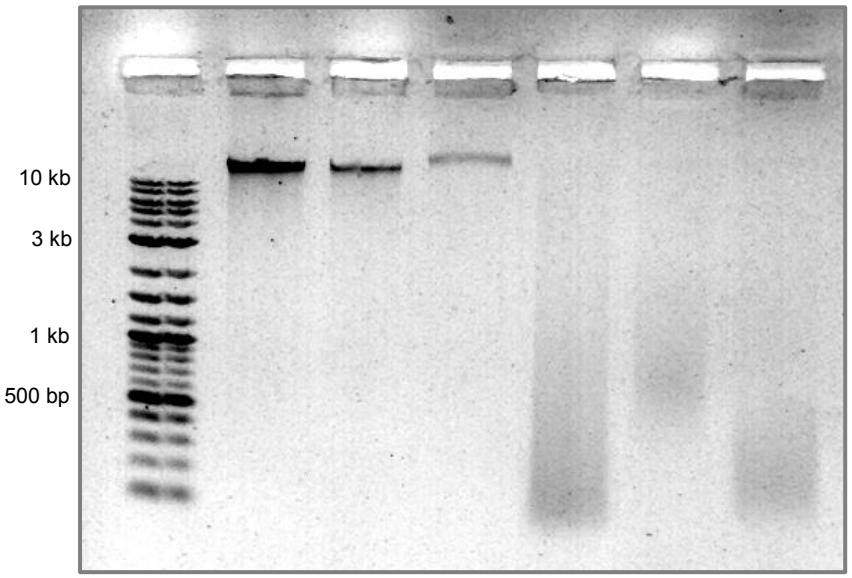
For this purpose, we decided to use AMPure XP® beads to size-select and bind to larger fragments of DNA, which were then eluted. AMPure XP contains paramagnetic beads along with an optimised buffer that selectively allows DNA fragments of size 100 bp and larger to bind to the paramagnetic beads. This DNA is later eluted.

Although some of the shredded DNA was of a small size (<100 bp), we aimed to remove as much of the shredded product as available. We thus aimed to see if we could increase the cut-off size of the AMPure XP bands to selectively bind to DNA only larger than ~500 bp.

Improved protocols that were used to enhance downstream Illumina sequencing displayed the increase in cut-off size of fragments bound to DNA on decreasing the amounts of AMPure XP beads used. [Quail et al., 2013] We decided to use a concentration of 0.6x the amount of the total sample.

To initially check the working of the AMPure XP® system, we set up a test 10 µl cleavage assays using 150 ng of HEK293 DNA. 200 nM of SauUSI in NEB4 buffer was used to shred HEK293 DNA and the reaction was incubated at 37°C for 2 hours. A 10 µl control of 150 ng HEK293 DNA along with NEB4 buffer and Nuclease-free water was also incubated.

The control and test sample were processed using AMPure XP beads as mentioned in the protocol and both the elute (supposed to have the larger fragments) and the flowthrough were loaded on a 1% Agarose gel.



	HEK gDNA	HEK gDNA	HEK gDNA	HEK gDNA + 200 nM SauUSI	HEK gDNA + 200 nM SauUSI	HEK gDNA + 200 nM SauUSI
Original sample	+	–	–	+	–	–
Eluted sample	–	+	–	–	+	–
Flowthrough	–	–	+	–	–	+

Figure 16. Size-selection of cleaved gDNA was obtained using AMPure XP beads. Using this size selection downstream of cleavage assays can remove the shredded methylated DNA from the sample, resulting in the selective enrichment of any unmethylated DNA present in the sample. This would also allow for the use of quantitative methods, such as PCR or NGS, to compare and quantify the amounts of methylated DNA that is shred.

Figure 16 displays that the AMPure XP size selection resulted in a loss in the amount of DNA, as visible by the difference in intensities between the bands in lanes 2 and 3. The treatment of the test sample with the beads verifies that a size-dependent selection of the shredded products has taken place. The cut-off seems to be around the 500 bp size, effectively removing most of the shredded DNA from the sample.

This experiment confirms that majority of the shred DNA which is <=500bp can be eliminated using AMPure XP bead technology.

Thus, methylated gDNA can be cleaved with Type IV restriction Endonucleases, followed by usage of AMPure XP beads to remove small DNA fragments.

Chapter 4: Discussion

The main aim of this project was to characterise and optimise the use of two restriction enzymes, SauUSI and McrBC, to successfully cleave methylated genomic DNA into small fragments, and utilise these optimisations to selectively enrich unmethylated DNA from any mixture of methylated and unmethylated DNA.

We initially confirmed the shredding activity of SauUSI on bacterial genomic DNA. The *E.coli* strain DH5α is methylated as it is a dcm+ strain, whereas the *E.coli* strain NEBTurbo is unmethylated as it is a dcm- strain. We were able to demonstrate the methylated DH5α DNA showed shredding on treatment with SauUSI, but NEBTurbo gDNA remained intact. Comparing the restriction patterns obtained on treatment of DH5α gDNA with SauUSI and BstNI, a Type II restriction endonuclease, demonstrated the shredding behaviour of SauUSI.

The ability of SauUSI and McrBC to shred methylated human (HEK293) DNA into small fragments was initially verified. Different concentrations of SauUSI and McrBC were used in cleavage assays to find the optimum concentrations that visibly caused cleavage of HEK genomic DNA (gDNA). After fixing optimal concentrations of SauUSI and McrBC, we attempted to try combinations of the two enzymes to see whether that resulted in cleavage of the genomic DNA into smaller fragments.

During the process of optimisation of these enzymes and while setting up combinations of the enzymes, there were challenges. At a concentration of 200 nM, the McrBC complex prepared seemed to display non-specific nuclease activity, we hypothesize that this is due to low amounts of a contaminating nuclease in the protein stock. At a lower concentration of 50 nM of protein, this activity wasn't seen, and we speculate that this potentially contaminating nuclease was diluted and its activity was minimal at this concentration.

On visualising the optimised reactions on Agarose gels, we observed complete shredding of human genomic DNA when using 200 nM SauUSI, and when using combination

reactions consisting of 200 nM SauUSI and 50 nM McrBC added one after the other. The two types of combination reactions, one with McrBC added first, and the other with SauUSI added first (the second protein was added following denaturation of the first protein) both showed a similar visual result. Genomic DNA just incubated with 50 nM of McrBC didn't seem to shred as well as DNA incubated with just 200 nM SauUSI or either of the combination reactions.

We were able to show that NEBTurbo® genomic DNA remained uncut on treatment with these enzymes, as the NEBTurbo strain lacks the methyltransferase dcm that would methylate its genomic DNA.

One of the main applications of an optimised system like this would be to remove contaminant human genomic DNA from field isolates of *Plasmodium falciparum.* Whole Genome Sequencing (WGS) of DNA extracted from the blood of infected patients results in a huge percentage of reads from sequencing corresponding to the human genome. In order to truly confirm which of our optimised enzymatic reactions caused the most shredding of HEK DNA, the products of the cleavage assays could be sequenced via NGS. Before trying out NGS, we decided to use PCR as a simple method of visualising the shredding of methylated genomic DNA following cleavage assays. We obtained primers specific to NEBTurbo and HEK genomic DNA, and verified the amplification of sequences specific to the two genomes. As we were concerned about the influence of remnants of buffers from the cleavage reactions on the efficacy of PCR reactions, we initially set up PCRs of just genomic DNA along with NEB4 buffer. A reduction in the efficiency of PCR confirmed that the buffer was adversely affecting the PCR. Due to this, we tested out ethanol and isopropanol precipitations on DNA, followed by PCR. Such precipitation would get rid of any salts from the buffers in solution. This, however, severely reduced the yield of DNA. We observed that the presence of trace amounts of buffer had a less adverse effect on PCR than the reduction in yield on precipitating the DNA from the sample.

Before carrying out PCR reactions with the constituents of cleavage reactions, any shredded DNA would have to be removed from the sample. In order to selectively remove all the shredded DNA, the AMPure XP system was optimised to selectively remove fragments with a size of <500 bp. This system was initially tested out with just HEK DNA

treated with 200 nM of SauUSI and showed a removal of shredded HEK DNA with a size of less than 500 bp. PCR reactions remain to be set up to confirm the degradation of HEK gDNA. As an added bonus, the AMPure XP protocol has an ethanol washing step that would remove any salts.

Future direction:

To actually show the selective enrichment of unmethylated DNA from a mixed sample of methylated and unmethylated DNA, mixtures of methylated and unmethylated DNA need to be created.

In order to use this system to separate *Plasmodium* gDNA from Human genomic DNA, or to separate methylated genomic DNA from unmethylated genomic DNA, it would be useful to generate mixtures of methylated and unmethylated DNA that are similar to those that would be obtained in such situations; the ratio of methylated and unmethylated DNA in such mixtures should reflect that in real-world applications.

These mixtures would then undergo cleavage with SauUSI and McrBC using these optimised enzymatic conditions, and then be subjected to size-selection using AMPure XP beads. The resultant DNA from the reaction mixtures could then be used to set up PCR reactions where a fragment specific to the methylated DNA would be used as a marker of the shredding of methylated DNA.

It would be of interest to take this project forward and specifically try these optimised reactions on mixtures of Human DNA and *Plasmodium falciparum* DNA that correspond to different parasitemia levels. Whole Genome Sequencing or Next Generation Sequencing could be used as a definitive, quantitative method of checking the efficacy of such a system.

References

Auburn, S., Campino, S., Clark, T.G., Djimde, A.A., Zongo, I., Pinches, R., Manske, M., Mangano, V., Alcock, D., Anastasi, E., et al. (2011). An Effective Method to Purify Plasmodium falciparum DNA Directly from Clinical Blood Samples for Whole Genome High-Throughput Sequencing. PLoS ONE *6*, e22213. https://doi.org/10.1371/journal.pone.0022213.

Cowell, A.N., Loy, D.E., Sundararaman, S.A., Valdivia, H., Fisch, K., Lescano, A.G., Baldeviano, G.C., Durand, S., Gerbasi, V., Sutherland, C.J., et al. (2017). Selective Whole-Genome Amplification Is a Robust Method That Enables Scalable Whole-Genome Sequencing of *Plasmodium vivax* from Unprocessed Clinical Samples. MBio *8*, e02257-16. https://doi.org/10.1128/mBio.02257-16.

Dila, D., Sutherland, E., Moran, L., Slatko, B., and Raleigh, E.A. (1990). Genetic and sequence organization of the mcrBC locus of Escherichia coli K-12. J Bacteriol *172*, 4888–4900. https://doi.org/10.1128/jb.172.9.4888-4900.1990.

Liu, H., and Naismith, J.H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol *8*, 91. https://doi.org/10.1186/1472-6750-8-91.

Loenen, W.A.M., Dryden, D.T.F., Raleigh, E.A., Wilson, G.G., and Murray, N.E. (2014). Highlights of the DNA cutters: a short history of the restriction enzymes. Nucleic Acids Research *42*, 3–19. https://doi.org/10.1093/nar/gkt990.

Nirwan, N., Singh, P., Mishra, G.G., Johnson, C.M., Szczelkun, M.D., Inoue, K., Vinothkumar, K.R., and Saikrishnan, K. (2019a). Hexameric assembly of the AAA+ protein McrB is necessary for GTPase activity. Nucleic Acids Research *47*, 868–882. https://doi.org/10.1093/nar/gky1170.

Nirwan, N., Itoh, Y., Singh, P., Bandyopadhyay, S., Vinothkumar, K.R., Amunts, A., and Saikrishnan, K. (2019b). Structure-based mechanism for activation of the AAA+ GTPase McrB by the endonuclease McrC. Nat Commun *10*, 3058. https://doi.org/10.1038/s41467-019-11084-1.

Ordway, J.M., Bedell, J.A., Citek, R.W., Nunberg, A., Garrido, A., Kendall, R., Stevens, J.R., Cao, D., Doerge, R.W., Korshunova, Y., et al. (2006). Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. Carcinogenesis *27*, 2409–2423. https://doi.org/10.1093/carcin/bgl161.

Oyola, S.O., Gu, Y., Manske, M., Otto, T.D., O'Brien, J., Alcock, D., MacInnis, B., Berriman, M., Newbold, C.I., Kwiatkowski, D.P., et al. (2013). Efficient Depletion of Host DNA Contamination in Malaria Clinical Sequencing. J Clin Microbiol *51*, 745–751. https://doi.org/10.1128/JCM.02507-12.

Panne, D. (2001). The McrBC restriction endonuclease assembles into a ring structure in the presence of G nucleotides. The EMBO Journal *20*, 3210–3217. https://doi.org/10.1093/emboj/20.12.3210.

Panne, D., Raleigh, E.A., and Bickle, T.A. (1999). The McrBC endonuclease translocates DNA in a reaction dependent on GTP hydrolysis 1 1Edited by J. Karn. Journal of Molecular Biology *290*, 49–60. https://doi.org/10.1006/jmbi.1999.2894.

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. https://doi.org/10.1021/bi0156862.

Pieper, U., Groll, D.H., Wünsch, S., Gast, F.-U., Speck, C., Mücke, N., and Pingoud, A. (2002). The GTP-Dependent Restriction Enzyme McrBC from E. coli forms High-Molecular Mass Complexes with DNA and Produces a Cleavage Pattern with a Characteristic 10-Base Pair Repeat. Biochemistry *41*, 5245–5254. https://doi.org/10.1021/bi015687u.

Quail, M.A., Swerdlow, H., and Turner, D.J. (2009). Improved Protocols for the Illumina Genome Analyzer Sequencing System. Current Protocols in Human Genetics *62*. https://doi.org/10.1002/0471142905.hg1802s62.

Raleigh, E.A., Trimarchi, R., and Revel, H. (1989). Genetic and physical mapping of the mcrA (rglA) and mcrB (rglB) loci of Escherichia coli K-12. Genetics *122*, 279–296. https://doi.org/10.1093/genetics/122.2.279.

Revel, H.R., and Luria, S.E. (1970). DNA-GLUCOSYLATION IN T-EVEN PHAGE: GENETIC DETERMINATION AND ROLE IN PHAGE-HOST INTERACTION. Annu. Rev. Genet. *4*, 177–192. https://doi.org/10.1146/annurev.ge.04.120170.001141.

Sukackaite, R., Grazulis, S., Tamulaitis, G., and Siksnys, V. (2012). The recognition domain of the methyl-specific endonuclease McrBC flips out 5-methylcytosine. Nucleic Acids Research *40*, 7552–7562. https://doi.org/10.1093/nar/gks332.

Teyssier, N.B., Chen, A., Duarte, E.M., Sit, R., Greenhouse, B., and Tessema, S.K. (2021). Optimization of whole-genome sequencing of Plasmodium falciparum from low-density dried blood spot samples. Malar J *20*, 116. https://doi.org/10.1186/s12936-021-03630-4.

Tock, M.R., and Dryden, D.T. (2005). The biology of restriction and anti-restriction. Current Opinion in Microbiology *8*, 466–472. https://doi.org/10.1016/j.mib.2005.06.003.

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 Research *49*, 2161–2178. https://doi.org/10.1093/nar/gkab042.

Westra, E.R., Swarts, D.C., Staals, R.H.J., Jore, M.M., Brouns, S.J.J., and van der Oost, J. (2012). The CRISPRs, They Are A-Changin': How Prokaryotes Generate Adaptive Immunity. Annu. Rev. Genet. *46*, 311–339. https://doi.org/10.1146/annurev-genet-110711-155447.

Xu, S., 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 Research *39*, 5597–5610. https://doi.org/10.1093/nar/gkr098.