Investigation of the link between DNA replication and mismatch repair in

*Thermus aquaticus*

By

Shreya Sawant

Faculty Advisor: Professor Manju Hingorani

A Dissertation submitted to the Faculty of Wesleyan University in partial
fulfillment of the requirements for the Degree of Masters of Arts with the
Department of Molecular Biology and Biochemistry.

Middletown, Connecticut    July 2016
Abstract:

Misincorporation of nucleotide bases during DNA replication that escapes proofreading by the polymerase results in the formation of either unpaired or mispaired bases. The Mismatch Repair (MMR) pathway targets these mismatched base pairs for correction and causes ~1000 fold increase in the fidelity of DNA replication. Inactivation of MMR genes causes predisposition to tumorigenesis and cancer. This pathway is well conserved across prokaryotes and eukaryotes and is best characterized in *E.coli*. A mismatch is recognized by MutS, which recruits MutL to the mismatch. In *E.coli* the MutS-MutL complex recruits MutH to generate a nick in the newly synthesized strand. However, most organisms do not have MutH homologues. Human MutL (hMLH1-PMS2), *Bacillus subtilis* and *Thermus thermophilus* MutL have been reported to have endonuclease activity. MutL nicks DNA in the vicinity of the mismatch followed by its excision and resynthesis.

Temporal coupling of MMR with DNA replication has been suggested by studies reporting that the presence of MutS in S-phase decreases the frequency of mutation. Additionally, interactions between MutS and MutL with the DNA replication protein β clamp further support this hypothesis. This study aims to investigate how interactions between MutS and β clamp may define the linkage between MMR and DNA replication in the *Thermus aquaticus* model system, which is relatively simple to work with and also comparable to a eukaryotic system. The goal of this project was to reconstitute a minimal *in vitro* DNA replication and repair system and monitor interactions between the proteins to
understand how they might influence the timing and/or efficiency of DNA mismatch repair.
Table of Contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>List of figures</td>
<td>5</td>
</tr>
<tr>
<td>(I) Introduction</td>
<td>8</td>
</tr>
<tr>
<td>(II) Research Objective</td>
<td>32</td>
</tr>
<tr>
<td>(III) Materials and Methods</td>
<td>33</td>
</tr>
<tr>
<td>(IV) Results and Discussion</td>
<td>49</td>
</tr>
<tr>
<td>(V) Conclusion and Future Direction</td>
<td>68</td>
</tr>
<tr>
<td>(VI) Addendum: Additional projects on the kinetic mechanism</td>
<td>70</td>
</tr>
<tr>
<td>of <em>Escherichia coli</em> Heptosyltransferase I</td>
<td></td>
</tr>
<tr>
<td>(VII) Appendices</td>
<td>84</td>
</tr>
<tr>
<td>(VIII) Reference</td>
<td>113</td>
</tr>
</tbody>
</table>
**List of Figures**

<table>
<thead>
<tr>
<th>Figure no</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. i. a.</td>
<td>DNA mismatch Repair</td>
<td>16</td>
</tr>
<tr>
<td>I. i. b.</td>
<td>Distribution of MMR proteins</td>
<td>18</td>
</tr>
<tr>
<td>I. i. c.</td>
<td><em>T. aq</em> MutS crystal structure</td>
<td>23</td>
</tr>
<tr>
<td>I. i. d.</td>
<td>C terminal region of <em>T. aq</em> MutS</td>
<td>23</td>
</tr>
<tr>
<td>I. i. e.</td>
<td>Crystal structure of clamp</td>
<td>26</td>
</tr>
<tr>
<td>I. i. f.</td>
<td><em>T. aq</em> β clamp monomer</td>
<td>29</td>
</tr>
<tr>
<td>I. i. g.</td>
<td>Proposed <em>T. aq</em> β clamp mutants</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td><strong>III. Materials and Methods</strong></td>
<td></td>
</tr>
<tr>
<td>III. i. a.</td>
<td>Forked DNA substrate for β clamp loading</td>
<td>44</td>
</tr>
<tr>
<td>III. i. b.</td>
<td>Magnetic bead clamp loading assay</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td><strong>IV. Results</strong></td>
<td></td>
</tr>
<tr>
<td>IV. i. a.</td>
<td><em>T. aq</em> MutS C42A/A764C purification and labeling with Alexa Fluor 488</td>
<td>52</td>
</tr>
<tr>
<td>IV. i. b.</td>
<td>$K_D$ of <em>T. aq</em> MutS C42A/A764C to 2Ap+T</td>
<td>52</td>
</tr>
<tr>
<td>IV. i. c.</td>
<td>Emission scan of AF-488 <em>T. aq</em> MutS C42A/A764C</td>
<td>53</td>
</tr>
<tr>
<td>IV. i. d.</td>
<td>Photobleaching of AF-488 <em>T. aq</em> MutS C42A/A764C</td>
<td>53</td>
</tr>
<tr>
<td>IV. i. e.</td>
<td>Emission scan of AF-488 <em>T. aq</em> MutS C42A/A764C in the presence of nucleotides</td>
<td>54</td>
</tr>
<tr>
<td>IV. i. f.</td>
<td>Equilibrium binding of <em>T. aq</em> MutS C42A/A764C with β</td>
<td>54</td>
</tr>
<tr>
<td>Figure no</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>IV. ii. a.</td>
<td>Isolation of clamp and loader genes from <em>T.aq</em> genomic DNA</td>
<td>60</td>
</tr>
<tr>
<td>IV. ii. b.</td>
<td>Construction of tagged and untagged plasmids of β clamp and clamp loader genes</td>
<td>61</td>
</tr>
<tr>
<td>IV. ii. c.</td>
<td>Purification and labeling of <em>T.aq</em> β clamp</td>
<td>61</td>
</tr>
<tr>
<td>IV. ii. d.</td>
<td>Purification of <em>T.aq</em> clamp loader subunits</td>
<td>63</td>
</tr>
<tr>
<td>IV. ii. e.</td>
<td>Excitation and emission scan of MDCC labeled <em>T.aq</em> β clamp D282C mutant.</td>
<td>64</td>
</tr>
<tr>
<td>IV. ii. f.</td>
<td>Excitation and emission scan of MDCC labeled <em>T.aq</em> β clamp G178C mutant.</td>
<td>64</td>
</tr>
<tr>
<td>IV. ii. g.</td>
<td>Equilibrium binding of MDCC β clamp G178C to MutS</td>
<td>65</td>
</tr>
<tr>
<td>IV. ii. h.</td>
<td>Equilibrium binding of MDCC β clamp G178C to MutS in the presence of DNA.</td>
<td>65</td>
</tr>
<tr>
<td>IV. ii. i.</td>
<td>Equilibrium binding of WT <em>T.aq</em> MutS and WT <em>T.aq</em> β clamp by GMSA.</td>
<td>66</td>
</tr>
<tr>
<td>IV. ii. j.</td>
<td>Screen of conditions to test <em>T.aq</em> MutS and <em>T.aq</em> β clamp interaction.</td>
<td>66</td>
</tr>
<tr>
<td>IV. ii. k.</td>
<td>Screen of conditions to test WT <em>T.aq</em> MutS and WT <em>T.aq</em> β clamp interaction in the presence of WT <em>T.aq</em> MutL.</td>
<td>67</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure no</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI: Addendum: Additional projects on the kinetic mechanism of <em>Escherichia coli</em> Heptosyltransferase I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI. a. i</td>
<td>Crystal structure of Hep I enzyme and it’s proposed conformational change</td>
<td>75</td>
</tr>
<tr>
<td>VI. a. ii</td>
<td>The Hep I reaction mechanism</td>
<td>75</td>
</tr>
<tr>
<td>VI. a. iii</td>
<td>Emission spectra of Hep I with OLDA</td>
<td>76</td>
</tr>
<tr>
<td>VI. a. iv</td>
<td>Structure of Hep I depicting tryptophan’s</td>
<td>76</td>
</tr>
<tr>
<td>VI. a. v</td>
<td>Stopped flow traces of Hep I and substrate</td>
<td>77</td>
</tr>
<tr>
<td>VI. a. vi</td>
<td>Titration of Hep I versus ODLA</td>
<td>77</td>
</tr>
<tr>
<td>VI. a. vii</td>
<td>Kinetics of ODLA binding and associated change in Hep I conformation</td>
<td>78</td>
</tr>
<tr>
<td>VI. a. viii</td>
<td>Scheme of two-step binding of ODLA to Hep I</td>
<td>79</td>
</tr>
<tr>
<td>VI. b. i</td>
<td>Experimental setup of the experiment to determine the ADP release rate of Hep I.</td>
<td>82</td>
</tr>
<tr>
<td>VI. b. ii</td>
<td>ADP release from Hep I resulting from catalysis.</td>
<td>82</td>
</tr>
<tr>
<td>VI. b. iii</td>
<td>Autodocking simulation of ODLA with HepI</td>
<td>83</td>
</tr>
</tbody>
</table>
I: Introduction

(i) **Background and Significance:**

The DNA replisome is a dynamic multi-component complex that employs DNA polymerase to carry out rapid leading and lagging strand DNA replication. Replicosome from bacteria to humans contain conserved proteins that perform a variety of functions. The *E. coli* replicase is well characterized and consists of ten subunits that can be categorized into three major components; (a) Pol III core (α, ε, θ): contains DNA polymerase and 3’ to 5’ exonuclease proofreading activity (b) β clamp: confers processivity to the polymerase by tethering the Pol III core onto DNA (c) γ complex clamp loader (γ, δ, δ’, χ, ψ): assembles β clamps onto DNA at primed sites and helps organize all the proteins involved in leading and lagging strand synthesis. In eukaryotes, the three components of the replisome are DNA polymerase (δ polymerase for the leading strand and ε for the lagging strand), the Proliferating Cell Nuclear Antigen (PCNA) sliding clamp, and Replication Factor C (RFC) clamp loader.

Replicative DNA polymerases are responsible for duplicating entire genomes with high fidelity. Nevertheless, replication errors still occur at a frequency of 1 in $10^5 – 10^6$ nucleotides incorporated into DNA (1). The DNA mismatch repair (MMR) pathway is the primary mechanism for correcting base-base mismatches and insertion/deletion loops (IDL) thereby decreasing the error frequency to 1 in $10^8$- $10^9$ nucleotides incorporated into DNA. The MMR pathway is highly conserved from bacteria to humans. In humans, defects in the MMR pathway are closely associated with an increased risk of developing cancer.
such as hereditary nonpolyposis colorectal carcinoma (HNPCC) \(^2\) and sporadic cancers. The function of MMR proteins is not limited to recognizing and correcting mismatches. In bacteria, MMR proteins participate in an anti-recombination mechanism preventing recombination between non-identical DNA sequences \(^3\). In eukaryotes, the MMR proteins have been shown to function in meiosis \(^4\) and are important in maintaining microsatellites \(^5\). MMR proteins are also reported to contribute to several other cellular processes including apoptosis \(^6,7\) and DNA damage checkpoint activation \(^8\).

The central and most intensively studied role for MMR proteins is in the repair of mismatches and IDLs introduced principally as polymerase errors. The MMR pathway is best characterized in *Escherichia coli*, where MutS (the first protein in the MMR pathway) detects the mismatch and then recruits MutL to the mismatch. In *E.coli* the MutS-MutL-DNA complex recruits MutH to generate a nick in the newly synthesized strand \(^52,53\). The eukaryotic homologs of MutS are Msh2-Msh6 (detects mismatches and single base insertion) and Msh2-Msh3 (detects 2-15 insertion-deletion loops). MutH nicks the newly synthesized DNA at hemimethylated GATC sites. UvrD helicase is then loaded at the nick and the error-containing strand is unwound and degraded by exonucleases. The single strand gap is filled in by DNA polymerase and sealed by ligase to complete the repair process. The hemimethylated sites generated by replication of methylated chromosomal DNA provide a temporary post-replication window for strand discrimination that closes when Dam-methylase modifies the newly synthesized strand. Homologues of *E.coli* MutS and MutL exist in the majority of organisms, but homologs of MutH are not found in many organisms, including many bacteria.
and eukaryotes. Therefore in organisms lacking MutH, MutL itself has endonuclease activity (e.g. eukaryotic human hMLH1-hPMS2 \(^{10}\), and bacterial \textit{Bacillus subtilis} \(^{11}\) and \textit{Thermus thermophilus MutL} \(^{12}\).

The formation of a MutS-DNA-MutL ternary complex in the presence of ATP has been established \(^{13}\). Mendillo et.al \(^{14}\) reported the connector domain of MutS interacts with MutL, and Plotz et.al \(^{15}\) reported the N-terminal ATPase domain of MutL is in proximity to MutS. Recently, the crystal structure of MutS-MutL complex by site specific crosslinking was solved and showed the sliding clamp conformation of MutS to promote MutL loading onto DNA \(^{108}\). The β clamp from the replisome is reported \(^{16}\) to bind to both MutS and MutL. Pillon et.al \(^{11}\) reported the C terminal endonuclease domain of MutL interacts with the β clamp while Lopez de Saro et.al \(^{16}\) reported that both the N terminal and C terminal of MutS interact with the β clamp; however, only the MutS N-terminal interaction with β clamp is thought to be important \textit{in vivo}. In eukaryotes, it has been reported that Msh2-Msh6, Msh2-Msh3 \(^{17}\) and Mlh-Pms1 \(^{18}\) interact with the PCNA clamp. The interactions between β clamp/PCNA and MutS, MutL / Msh2 - Msh6, Mlh1 - Pms2 suggests a temporal coupling of the MMR pathway to DNA replication.

The mismatch search and recognition process employed by MutS is complex. MMR is capable of detecting a single mispaired base against a background of \(\sim 10^7\) Watson-Crick base pairs. Moreover the mispaired bases must be detected before transient markers identifying the parental and newly synthesized strand are lost and before the DNA is packaged. To account for such efficiency, MutS may work in concert with additional factors to enable it to detect
mismatches \textit{in vivo}. This hypothesis is strongly supported in a recent report by Hombauer, H. et.al \cite{19}, who showed that the availability of Msh2-Msh6 during the S phase of the cell cycle is proficient in suppressing mutations due to replication errors. Iyer et al. \cite{20} reported that MMR \textit{in vitro} requires single strand breaks in the DNA (which appear necessary to activate MutL endonuclease), suggesting that MMR might be targeted to strand breaks in the nascent DNA strands during DNA replication. Recently, Liao et al. \cite{100} characterized the real time dynamic interaction between MutS and components of the replisome in live \textit{Bacillus subtilis} cells using super-resolution microscopy and single molecule tracking. The study identified a heterogeneous population of MutS that moves back and forth from the replisome either rapidly scanning the nucleoid or is transiently associated with it in an ATP dependent manner. It has been speculated that coupling of MMR and DNA replication pathways would be advantageous as it would efficiently localize MMR proteins to where mismatches and IDLs are formed, enable strand discrimination, and facilitate completion of MMR before the DNA is packaged.

MutS-β clamp complex is proposed to be a key intermediate in the mismatch repair pathway \cite{32,20,54,55}. The N terminal region of \textit{S.cerevisiae} Msh6 contains a conserved region \([Qxx(L/I)xxFF]\) which has been implicated in assisting Msh2-Msh6 and PCNA complex formation \cite{42}. The significance of this interaction however remains controversial. In human MutSa, a variant of Msh6 which lacks 77 amino acids in the N-terminal domain, including the conserved region, is severely compromised in its ability to support mismatch repair in \textit{vitro} \cite{56}. However, in yeast, alanine substitutions of the conserved region of Msh6
result in an insignificant increase in mutability (57). Furthermore, biochemical analysis has indicated that PCNA enhances the affinity of Msh2-Msh6 for a mismatch (58), but a subsequent study showed that although PCNA, Msh2-Msh6 and DNA form a stable ternary complex, binding to the mismatch leads to disruption of the Msh2-Msh6-PCNA interaction (54). The binding sites of this ternary complex remain under investigation, and its dynamics are unknown. Therefore the functional significance of clamp and MutS interaction is unresolved.

The model system of choice for my project is Thermus aquaticus (T.aq) because it is a non-MutH dependent system and therefore the findings are applicable to eukaryotic organisms. The crystal structure of T.aq MutS is solved (30) and the kinetic mechanism of MutS with respect to nucleotide binding, hydrolysis, and phosphate release has been determined (40,59). The kinetics of MutS binding and release from mismatched DNA under the effect of nucleotides has also been determined (40,59). Therefore, the affinity and dynamics of MutS-β clamp and MutS-β clamp DNA complex (+/- nucleotides) can be compared to previous findings to determine the role of MutS-β clamp interaction in the MMR mechanism. Furthermore, the dynamics of MutS-β clamp and MutS-β clamp-DNA complex can also be studied by single molecule FRET in an ongoing collaboration with Keith Weninger from North Carolina State University. In the future, to understand the steps succeeding mismatch detection by MutS, we can also extend the study by adding MutL to the system to reconstitute the entire MMR machinery in vitro and get a complete picture of the MMR mechanism and its links to replication in both early and late stages of the pathway.
MutS-β clamp interaction has been studied in vitro by surface plasmon resonance \(^{(60)}\), equilibrium gel filtration \(^{(48)}\) and gel mobility shift assays \(^{(16,58)}\) to determine the binding affinity \((K_D)\) of the complex. However this measure of MutS-β interaction does not provide any information about when the complex is assembled/disassembled in the repair reaction, and what role mismatched DNA and nucleotides play in these dynamics. It is therefore necessary to measure the kinetic parameters defining the interactions between MutS, β clamp and DNA in solution in order to understand their role in the MMR mechanism.

My research aims to investigate the linkage between DNA replication and mismatch repair pathways by purifying and reconstituting the proteins involved in DNA replication (initially β clamp, Υ complex) and mismatch repair (MutS, MutL), and preparing DNA substrates to analyze the effects of interactions between these proteins on MutS and MutL activity in MMR.

(ii) The Mut System:

In 1966, the Ephrussi-Taylor laboratory \(^{(21)}\) observed differences in transformation efficiency of donor DNA in *Streptococcus pneumoniae* and correlated them with certain genetic markers of the donor DNA. They found that the double stranded DNA once taken up by the cell was degraded to a single stranded DNA, which bound to the complementary portion of chromosomal DNA in the cell by physical displacement to form a heteroduplex species. They measured transformation efficiency across various strains of *Streptococcus pneumoniae* and classified them into high efficiency (HE) and low efficiency (LE). The HE class was defined as having the ratio of DNA marker integrated into the amount of genomic DNA
taken up by the cell of 1. The LE markers were 5-20 fold weaker in transformation efficiency compare to the HE. Based on this observation they proposed a cellular mechanism that could recognize and subsequently repair base-base mismatch DNA.

In 1973, the Tiraby-Fox (22) laboratory observed that while most of their Streptococcus pneumoniae stock strains could also be classified into HE and LE categories, some strains incorporated LE markers at the same efficiency as HE markers and they called this population of strains as Hex⁻. They then proposed that the Hex⁻ strains could distinguish between HE and LE markers as a result of cellular machinery in those strains that could discern between matched and mismatched DNA. This machinery would excise the mismatch strand and use the complementary DNA to resynthesize the degraded strand and eliminate the error. Work done by Guild and Shoemaker (23) further supported the hypothesized repair mechanism when they showed that LE marker strains that survived were homozygous to the original genome.

The first Hex genes were cloned and characterized by Lacks et.al (24) in 1986. HexA and HexB genes were transformed into B. subtilis and they could recognize DNA strand breaks and specifically excise markers in the 5’ to 3’ direction. Based on this observation they proposed that one gene could be involved in DNA error recognition and the other in strand excision. While the Hex system of and S.pneumoniae and B. subtilis first suggested a mechanism that maintained post DNA replication genome integrity, it was the discovery of E.coli mismatch repair that transformed the field, and MMR was soon recognized as a conserved pathway from prokaryotes to eukaryotes. Currently there are two types
of mismatch repair mechanisms (a) Methyl directed mismatch repair in *E.coli* and (b) MMR eukaryotes and most bacteria.

(ii) (a) Overview of methyl directed mismatch repair in *E.coli*

The mismatch repair (MMR) system has been best understood in *E.coli*. In this system a mismatched base is recognized by a MutS homodimer, following which a MutL homodimer interacts with the MutS-DNA complex. The MutH restriction endonuclease is activated by MutL following which MutH nicks the unmethylated strand at a nearby hemimethylated GATC site to introduce a point of entry for the UvrD helicase and exonuclease to excise and degrade the error-containing strand. The resulting single strand is stabilized by single strand binding proteins and serves as the template for DNA resynthesis by DNA polymerase III and ligase. The absence of a methylation mark on the newly synthesized strand serves as a signal for strand discrimination. Although homologues of *E.coli* MutS and MutL exist in most organisms, no homologue of *E.coli* MutH has been identified in the majority of organisms.
Figure I (i) a

**DNA Mismatch Repair**

**Mismatch repair pathway in *Escherichia coli***:

- **MutS**
  - ATPase.
  - Identifies mismatches, insertion/deletions, damaged bases in DNA.

- **MutL**
  - ATPase.
  - Stabilizes MutS-DNA interaction.
  - Endonuclease in eukaryotes.

- **MutH**
  - Sequence specific endonuclease.
  - Identifies hemi-methylated sites and nicks the unmethylated nascent strand.

- **GT**
  - Helicase and Exonuclease.
  - DNA Replication Proteins.

- **GC**

Step 1: The MutS ATPase scans DNA and identifies mismatches.
Step 2: Once the mismatch is detected, MutL is recruited to stabilize the MutS-DNA complex.
Step 3: MutH sequence specific endonuclease nicks the strand containing the error.
Step 4: Helicase unwinds the error containing strand, exonuclease degrades it, DNA polymerase re-synthesizes the strand and ligase seals the ends.
(ii) (b) Overview of eukaryotic mismatch repair

In eukaryotes, MMR is initiated by heterodimeric ATPase proteins MutSα or MutSβ that are composed of two MutS homologues, Msh2-Msh6 and Msh2-Msh3. MutSα recognizes base-base mismatches and 1 nucleotide IDLs, while MutSβ can recognize IDLs larger than 2 nucleotides. Once bound to the mismatch MutSα kinks the duplex DNA and recruits MutLα (a heterodimeric ATPase made up of two MutL α homologues, Mlh1/Pms2 or Pms1 in yeast) to form a ternary protein complex \(^{(25)}\). The latent endonuclease activity of MutLα is activated on interactions with MutSα at a mismatch and makes nicks in DNA followed by ExoI strand excision. The DNA polymerase δ then resynthesizes the strand and ligase seals the nick to complete the reaction \(^{(26,27)}\). In the absence of a MutH homolog, strand discontinuity appears to serve as a signal to initiate MMR. In vivo, MMR initiates at the end termini of Okazaki fragments of lagging strand and at nicks generated in the leading strand by the mismatch activated Mlh1/Pms3 endonuclease. Recent work done by Ghodgaonkar \(^{(28)}\) shows that a single ribonucleotide in the vicinity of a mismatch can also act as an initiation site for MMR in human cell extracts. MMR activation in this system is dependent on RNAse H2, which excises single ribonucleotides from DNA-DNA duplexes and therefore provides the nick required for the activation of MutL endonucleolytic activity.
### Table 1: Distribution of MMR proteins.

<table>
<thead>
<tr>
<th>Molecular function</th>
<th><em>Thermus thermophilus</em></th>
<th><em>Escherichia coli</em></th>
<th><em>Saccharomyces cerevisiae</em></th>
<th><em>Homo sapiens</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>β-clamp clamp-loader</td>
<td></td>
<td></td>
<td>RFC</td>
<td>RFC</td>
</tr>
<tr>
<td>Strand incision</td>
<td>MutL</td>
<td>MutH</td>
<td>MutLα (MLH1/PMS1) MutLβ (MLH1/MLH3)</td>
<td>MutLα (MLH1/PMS1) MutLβ (MLH1/MLH3)</td>
</tr>
<tr>
<td>Match making</td>
<td>MutL</td>
<td>MutLα</td>
<td>MutLβ (MLH1/PMS1) MutLγ (MLH1/MLH3)</td>
<td>MutLβ (MLH1/PMS1) MutLγ (MLH1/MLH3)</td>
</tr>
<tr>
<td>Strand excision (single-stranded DNA-binding)</td>
<td>SSB</td>
<td>SSB</td>
<td>RPA</td>
<td>RPA</td>
</tr>
<tr>
<td>Strand excision (exonuclease)</td>
<td>RecJ</td>
<td>Exol</td>
<td>EXO1*</td>
<td>EXO1*</td>
</tr>
<tr>
<td>Strand excision (helicase)</td>
<td>UvrD DNA polymerase III</td>
<td>DNA polymerase III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repair synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The involvement of bacterial clamp and clamp-loader in the strand incision reaction has not yet been confirmed. It is demonstrated that the endonuclease motif in MLH3 is responsible for in vivo MMR [83]; however, the endonuclease activity of MutLγ has not yet been confirmed biochemically. In yeast and human, EXO1 has the 5’-flap endonuclease activity in addition to 5’-3’ exonuclease activity.*

Reference:
(ii) (c) Structure and function of MutS

Significant insights into the structure-function relationship of MutS was provided by the crystal structure of MutS in the apo form and bound to mismatched DNA. *Thermus aquaticus* MutS crystal structure shows that this protein binds to a 21-base pair heteroduplex containing an unpaired thymidine (+T) nucleotide in a clamp-like homodimer (30, 29). Each subunit can be divided into five structurally distinct domains. Domains I and IV are called the mismatch DNA binding and clamp domains and they interact with the sugar-phosphate backbone of the DNA. Domain IV is 40 Å long and comprises mainly of α-helices that encompass the DNA upon mismatch binding. Domain I of subunit 1 provide residues that make mismatch specific contacts Phe39 and Glu41 and are highly conserved and crucial for binding to the IDL with high affinity. The Phe39 forms aromatic stacking interactions with the unpaired thymidine whereas the Glu41 donates a hydrogen bond to the nucleotide. The subunit 2 also has these residues in the DNA binding domain yet they do not make the same interactions with the mismatch. This implies there is a functional asymmetry between the two monomers (30). Once stabilized on the mismatch, MutS kinks the duplex helix approximately 60° towards the major groove. The C terminal domain V contains conserved Walker A and Walker B motifs for ATP binding and hydrolysis. Residues from both subunits contribute to form this composite ATPase binding domain that is imperative to ATP hydrolysis and dimerization. An allosteric relationship has been proposed between the C-terminal composite ATPase domain and the DNA mismatch binding domain that is critical for its function.
Studying the MutS-β clamp interaction with respect to MutS.

The MutS-β clamp interaction interface with respect to MutS is best characterized in the *E.coli* (which in a MutH endonuclease dependent system). The N terminal region of MutS (which is involved in mismatch binding) was mapped for β clamp interaction. Peptides designed for the N terminal region spanning residues 12 PMMQQYLRLK 21 were found to be important for β clamp interaction \(^{(16)}\). Alanine scanning with the peptides showed Q15, L18, and R19 severely reduced β binding, while M13, M14, and Q16 show intermediate reduction in binding. Peptide derivatives do not represent the tertiary structure of MutS, therefore alanine scanning was done to generate a double (Q15 L18) mutant and quadruple mutant (Q15, Q16, L18, R19) of MutS designated MutS\(^N\). DNA binding and ATPase activity of MutS\(^N\) was indistinguishable from wild type MutS, therefore alanine substitutions preserved the functionality of MutS \(^{(16)}\). Gel mobility shift assays of double and MutS\(^N\) mutant with \(^{32}\)P labeled β clamp did not show a decrease in binding, but *in vivo*, MutS\(^N\) could not complement a mismatch repair deficient strain \(^{(16)}\). Therefore, the interaction between the N terminal region of *E.coli* MutS and β clamp is not necessary in solution, but is important for its function in *vivo*.

In *Bacillus subtilis* (which is a non-MutH dependent system like *T.aq*) peptide scanning approach was used to map the N terminal region upto residue 223 encompassing domain I (mismatch binding domain) and domain II (connector domain). A peptide derivative containing the conserved N terminal β-binding region: 9 QQYL 12 (identified in *E.coli* MutS), failed to bind Myc tagged β clamp \(^{(32)}\). Peptides derived from domain I [43 QELEITLTSR 52] and domain II
[214 EDITKTLRL 223] of B. subtilis MutS showed strong interaction with the β clamp. Since only a peptide array approach was used to study the N terminal region of MutS, it is uncertain which regions specifically interact with the β clamp.

The C terminal region of MutS is involved in dimerization and nucleotide binding. To identify the C terminal region of E. coli MutS that could interact with the β clamp, MutS sequence was aligned with other bacterial species to identify a putative β clamp-binding region [812 QMSLL 816] (33). This β clamp-binding motif has been also identified in Bacillus subtilis MutS [806 QLSFF 810] and Pseudomonas aeruginosa MutS [816 QSDLF 820] (32, 34). In E. coli a peptide derived from the last 30 residues present in the C terminal region containing the putative β clamp binding motif could bind β clamp in SPR assay (35). Gel filtration analysis of MutS containing a deletion of the β–binding motif behaves like wild type MutS, therefore this deletion mutant did not alter the structure of MutS (16). A 53 amino acid C terminal deletion or a deletion of the β–binding motif in MutS failed to bind $^{32}$P labelled β clamp in gel mobility shift assays (16). Additionally, a quintuple mutant of the β clamp binding motif [Q812G, M813S, S814G, L815S, L816G] also did not bind the β clamp in far western assays (16), but in vivo a deletion mutant of the β clamp binding motif could complement a repair deficient strain (16). Single alanine mutations of the β clamp binding motif identified M813 and L815 to be important for β clamp binding by far western assays (32). This suggests that although transient interactions between MutS N terminal region and β clamp is possible, the C terminal region of MutS dominates the interaction
between MutS and β clamp. Therefore the C terminal region of MutS is a good position to introduce a fluorophore to monitor MutS-β clamp interaction.

The interaction interface between MutS and β clamp in _T.aq_ is not known. However, since this interface has been studied in _E.coli, Bacillus subtilis_ and _Pseudomonas aeruginosa_, a multiple sequence alignment of _T.aq_ with the aforementioned systems was done to identify the potential region(s) of interaction as shown in Appendix IV (c). The putative C terminal β clamp binding motif present in _E.coli, B.subtilis_ and _P.aeruginosa_ is absent in _T.aq_ MutS. However, peptide array scan of the C terminal region in _B.subtilis_ MutS showed residues [832 VIDAFKSLNI 841] downstream of the β clamp binding motif to be able to bind β clamp in far western assays (33). This region in _T.aq_ MutS corresponds to [773 VLERLLALD 782]. 43 amino acids at the C terminal end is truncated in the _T.aq_ MutS crystal structure (30, 36). Based on the preliminary characterization of MutS-β clamp interaction interface done in the aforementioned systems, _T.aq_ MutS C-terminal region is expected to interact with the β clamp. Although the putative β binding motif is not present in _T.aq_ MutS, it may contain a β clamp binding region within its last 43 amino acids. Based on the crystal structure of _T.aq_ MutS PDB ID: 1EWQ (Figure I. i. c), residue A764 is the most appropriate site to introduce a fluorescent label as it is proximal to the proposed β clamp interaction region and is further from the nucleotide binding or dimerization region of _T.aq_ MutS (Figure I. i. d). Therefore _T.aq_ MutS A764C can be labeled with a fluorophore to report MutS-β clamp interaction without disrupting the functionality of MutS.
Figure I. i (c)

Crystal structure of *Thermus aquaticus* MutS (PDB ID 1EWQ)

Figure I. i (d)

A764 residue in C terminal region of *T.aq* MutS is chosen for labeling.
(iii) The DNA Replisome:

DNA replication is performed by the replisome complex that coordinates the actions of many proteins at the replication fork. This complex is responsible for copying the entire genomic DNA allowing for the passage of hereditary information from parental cell to daughter cell with high fidelity. The proteins that constitute to the replisome complex in *E.coli* and other prokaryotes are the following:

**DNA polymerase III:** comprising of α, ε, and θ subunits.
- α subunit: is encoded by the *dnaE* gene and has the polymerase activity.
- ε subunit: is encoded by *dnaQ* and has 3'→5' exonuclease activity.
- θ subunit: is encoded by *holE* gene and stimulates the ε subunit proofreading ability.

**β clamp:** is encoded by the *dnaN* gene and acts as a sliding DNA clamp that tethers the DNA polymerase to the replication fork thereby increasing its processivity. In 1992 the first crystal structure of the *E.coli* sliding clamp was solved \(^{(101)}\). The sliding clamps from all domains of life have a similar architecture comprised of six domains arranged in a circle and the chain-folding topologies of each domain are the same (Figure 1. i. e) \(^{(102)}\). β clamp is a circular homodimer with the 2 subunits aligned head-to-tail forming a ring with two structurally distinct faces, one is the C-terminal face and serves as a platform for interaction with other proteins including the DNA polymerase and the clamp loader and the other is the N-terminal face. Each β subunit consists of three independent domains connected by a long loop referred to as the interdomain connecting loop (IDCL), which protrudes from the C-terminal face. Although the amino acid sequence of
each domain is quite different, all fold into similar 3-D structures with identical chain topologies; two adjacent α helices flanked by two, four-stranded antiparallel β sheets. Therefore the β clamp has pseudo six-fold symmetry (102). Twelve α helices line the inner surface of the ring while the β sheets form a continuous outer surface. The 35Å central channel of the clamp is large enough to accommodate double helical nucleic acid. Although the β clamp has a net negative charge, a strong positive electrostatic potential lines the inner surface of the ring where water molecules mediate protein-DNA interactions allowing the β clamp to move freely along DNA in a sequence-independent manner.

**Clamp loader complex**: is a multi subunit ATPase protein complex that binds to β clamp and loads it onto primer-template junctions at the replication fork. The *E.coli* clamp loader is comprised of three different subunits, γ/τ subunit (encoded by dnaX gene), δ subunit (encoded by holA) and δ’ subunits (encoded by holB) which assemble into a heteropentamer with stoichiometry δ δ’(γ/τ)3 (103, 104, 105).

Two other subunits Χ and Ψ are associated with the clamp loader but these subunits are not required for clamp loader assembly nor its clamp loading activity. The τ subunit is the full-length product and when present within the clamp loader binds the DNA polymerase III core and helicase via its carboxyl terminus. However the carboxyl terminus of the τ subunit is not required for clamp loading and its truncation by a translational frameshift produces γ subunit. (106, 107)
Figure 1. i (e) Crystal structure of sliding clamp.

Sliding clamps from different organisms.

A. *E. coli* β clamp

B. T4 bacteriophage gp45

C. *Pyrococcus furiosus* PCNA

D. *S. cerevisae* PCNA

E. *Homo sapiens* PCNA


For each organism the top and side views of the 3-D crystal structure are shown on the left and alignment of the globular domains in the middle. The electrostatic surface potential for each is shown on the right with red and blue representing negative and positive electrostatic potential.
**Studying the MutS-β clamp interaction with respect to β clamp.**

Preliminary characterization of the MutS-β clamp interaction interface with respect to the β clamp has been done in the *E.coli* model system\(^{(32)}\). A kinase protection assay of a C terminal kinase tagged β clamp revealed that MutS binds to this face of the β clamp\(^{(32)}\). The β dimer also contains two hydrophobic patches that serve as binding surfaces for many proteins involved in DNA replication and repair (Figure I. i. f). In one study aimed at defining the binding surface of MutS on β clamp, mutations were made in the hydrophobic patch region [G174, H175, R176, V247, V360, M362] and on a loop present near the dimer interface of the β clamp [Q299]. Surface plasmon resonance was used to test β clamp double mutant R176A/L177A and single mutant Q299A for binding immobilized MutS. The double mutant showed 10-fold reduction in binding to MutS and single mutant showed 20-fold reduction in binding affinity\(^{(32)}\). Thus, residue Q299 in *E.coli* β clamp is a key residue for interaction with MutS.

Q299 labeled with pyrene has been previously used to study *E.coli* Υ complex clamp loader-β clamp interaction\(^{(37, 38)}\). β clamp labeled with a pyrene fluorophore at position 299 showed a two fold increase in signal upon addition of Υ complex. Presumably the Υ complex - β clamp interaction alters the environment of the fluorophore thereby changing the intensity of the signal. Based on this information, I decided to introduce a fluorophore in *T.aq* β clamp at or near a position analogous to *E.coli* Q299 to monitor MutS-β clamp interaction. However, one concern is that since the mutation of Q299 to alanine diminished MutS-β clamp binding as described above\(^{(32)}\).
The multiple sequence alignment of *T.aq* and *E.coli* β clamp is shown in Appendix IV (a). The protein sequence of β clamp is not well conserved (25% sequence similarity). The structure of *T.aq* β clamp was generated with I-TASSER program and aligned with *E.coli* β clamp [PDB ID: 3BEP]. By structural alignment, position D304 in *T.aq* β clamp was found to be analogous to residue Q299 in *E.coli* β clamp. Therefore *T.aq* β clamp D304C could be labeled to monitor MutS-β clamp interaction. However, as noted above, mutating *T.aq* β clamp to generate D304C could reduce the binding affinity of MutS to β clamp. Therefore residue D282 in the adjacent loop may be labeled to study MutS-β clamp interaction (Figure I. i. g). Alternatively, the hydrophobic patch residues (G178C and Y179) could also be labeling sites as the adjacent residues R180 and L181 correspond to R176 and L177 in *E.coli* β clamp and have been shown to interact with MutS (32).
Figure I. i (f)

Structure of *T.aq* β clamp generated by TASSER I program.

Figure I. i (g)

Based on literature *T.aq* β clamp mutants proposed for labeling.
(iv) Approach to study the MutS-β clamp interaction.

In order to study the MutS-β clamp interaction in a minimal *in vitro* reconstituted system, the genes encoding β clamp and clamp loader (δδ’γτ) need to be isolated from *Thermus aquaticus* YT-1 genomic DNA, cloned into expression vectors, expressed in *E.coli* and then purified by chromatographic techniques. *T.aq* MutS has been purified and characterized in our laboratory. Gene isolation, protein purification and labeling of *T.aq* β clamp and clamp loader subunits followed by assay development to study MutS-β clamp and MutS-DNA interaction is described in the material and methods section. In order to mimic the *in vivo* conditions of MutS - β clamp interaction that occur at the replisome, the β clamp must be loaded onto a DNA substrate that allows for β clamp loading with long half-life of the β clamp-DNA complex. The DNA substrate must contain a primer-template junction to allow the clamp loader to load the β clamp. The substrate should also prevent the β clamp from sliding off and should be long enough to enable MutS to bind the DNA and interact with the clamp. The minimum length of the DNA substrate can be estimated from the crystal structures of MutS, clamp and clamp loader on DNA.

Crystal structure of the bacteriophage T4 clamp and clamp loader complex bound to primer-template DNA shows that 10 bp of the double helix are within the interior of the clamp loader and 10 bp are within the sliding clamp \(^{(41)}\). The single strand template in this co-complex is 10 bases in length. A similar DNA footprint of the *Pyrococcus furiosus* clamp and clamp loader complex was observed by electron microscopy \(^{(50)}\). The crystal structure of *T.aq* MutS bound to a +T bulge shows that MutS interacts with the phosphate backbone with a foot
print of 15 nucleotides (6 nucleotides on one side of the +T bulge and 9 on the other side) \(^{(4)}\).

In human, mouse and yeast, Msh6 has a consensus PCNA binding motif [Qxx(I/L/M)xxF(F/Y)] present in the N terminal region (NTR). This consensus PCNA binding motif is not present in *T. aq* MutS, but *E. coli* MutS interacts with the \(\beta\) clamp by both the N and C terminal domains \(^{(32)}\). Small angle X-ray scattering (SAXS) study done with yeast Msh6-NTR with PCNA revealed that the Msh6 NTR is natively disordered and forms an extended tether between Msh6 and PCNA \(^{(42)}\). However, a similar study done with human MutS\(\alpha\) and PCNA revealed that they interact in an elongated conformation in which the two proteins associate in an end-to-end manner that does not involve an extended unstructured tether \(^{(43)}\). These two studies contradict each other and the binding sites and dynamics of the MutS and \(\beta\) clamp interaction is still debatable. The features of the proposed bi fork DNA substrate is described in the material and methods section III (iv).
(II) **Research Objective.**

My project aims to understand the molecular mechanism of how DNA mismatch repair is coupled to DNA replication. Misincorporation of nucleotide bases during DNA replication which that escapes DNA polymerase proofreading is detected by the Mismatch Repair (MMR) pathway which corrects these mismatches causing ~1000 fold increase in the fidelity of DNA replication. The search for post replicative errors is challenging because the DNA polymerase has a low error rate \([1 \text{ in } 10^6-10^7 \text{ nucleotides}]\), the errors must be detected and repaired quickly before transient markers identifying the newly synthesized strand is lost and errors must be detected before the DNA is packaged.

Hombauer *et al.* \(^{(31)}\) reported that the presence of MutS in S-phase decreases the frequency of mutation, indicating temporal coupling of MMR with DNA replication. Other studies have reported \(^{(32)}\) interactions between the MutS and MutL mismatch repair proteins and the β clamp replication protein. This proposed study aims to understand how interactions between MutS and β clamp may define the linkage between MMR and DNA replication in a non-MutH dependent model system *Thermus aquaticus* by measuring parameters like binding affinity \((K_D)\), association rates \((k_{on})\) and dissociation rates \((k_{off})\) between MutS-β clamp and MutS-DNA. The first step is to isolate and characterize the DNA replication protein needed for this study namely β clamp and clamp loader complex, followed by studying MutS-β clamp interaction with β clamp loaded on the DNA to mimic the conditions of the replication fork.
III: Materials and Methods

III. i. (a) Purify and label a MutS construct that can report MutS-β clamp complex and test its DNA binding activity.

*T.aq* MutS is proposed to interact with β clamp via its C-terminal region. The mutation A764C is introduced at this end as it is beyond the nucleotide binding and dimerization region of MutS. Therefore, labeled *T.aq* MutS C42A/A764C could report β clamp binding without disrupting the functionality of MutS. However, this must be experimentally tested by conducting a DNA binding and/or ATPase assay of this construct.

*Approach:*

- *T.aq* MutS C42A/A764C: The mutation A764C was introduced into *T.aq* MutS C42A pET3a plasmid by the Quik Change Site-Directed Mutagenesis Kit (Stratagene). Primers are listed in Appendix IV (c). Wild type *T.aq* MutS has a cysteine at position 42 which has been mutated to alanine in our wild type reference plasmid. This enables us to introduce cysteine residues at specific sites in *T.aq* MutS for labelling. *T.aq* MutS C42A/A764C was purified using the standard protocol (Q-Sepharose column and Ammonium Sulphate precipitation) and labeled with Alexa Fluor 488 (AF-488).

- DNA binding of *T.aq* MutS C42A/A764C: MutS binding to DNA has been previously characterized in our laboratory by using 2-Aminopurine (2-Ap). 2-Ap is an isomer of adenine that forms a base pair with thymine and emits fluorescence when $\lambda_{ex} = 315$nm. 2-Ap is popularly used to study nucleic acid structure and protein-nucleic acid interactions as it is functionally equivalent to adenine and shows fluorescence sensitivity to base stacking. It also has a reasonable signal-to-noise ratio due to low absorbance of radiation by DNA.
and protein at 315nm. In our DNA substrate, 2-Ap is placed adjacent to a +T mismatch on a 37mer duplex DNA. The conserved Phe-X-Glu motif of MutS recognizes the +T mismatch. The phenylalanine residue stacks with the thymine ring, thereby unstacking the adjacent 2-Ap which results in an increase in fluorescence signal.

**Experimental setup:** MutS-DNA interaction at equilibrium was measured on a FluoroMax-3 fluorometer (Jobin-Yvon Horiba group: Edison, NJ). 2-Ap+T 37mer DNA (0.01µM) was titrated with or without *T.aq* MutS C42A/A764C dimer (0-0.2µM) in 20mM HEPES pH 7.7, 50mM NaCl, 5mM MgCl₂ (DNA binding buffer) at 40°C after 30 sec mixing. λᵫ = 315nm and λₑm = 375nm. Background fluorescence from the buffer and MutS was subtracted from the raw data. The data were corrected for intrinsic MutS fluorescence by subtracting data from parallel experiments with unlabeled DNA. Fluorescence intensity was plotted versus MutS concentration and the apparent dissociation constant (*K*_D) for the interaction was obtained by fitting the data to the quadratic equation:

\[
D \cdot M = F₀ + 0.5(F_{max} - F₀) \left\{ (K_D + D_t + M_t) - (K_D + D_t + M_t)^2 - 4 D_t M_t \right\}^{1/2}
\]

D·M = Fraction of MutS•DNA

F₀ = 2-AP+T fluorescence in the absence of protein

F_{max} = Maximal fluorescence

Dₜ and Mₜ = Total molar concentrations of DNA and MutS respectively.

The data were fit by non-linear regression using KaleidaGraph (Synergy Software).
III. i. (b) Study the interaction between MutS and β clamp using MutS as a reporter.

*T.aq* MutS C42A/A764C is proposed to interact with β clamp to form a complex. However this needs to be experimentally validated in solution.

**Approach:** An emission scan was performed on a FluoroMax-3 fluorometer. If a change of signal was observed upon mixing AF-488 *T.aq* MutS C42A/A764C with unlabeled β clamp, this construct can be potentially used to measure MutS-β clamp interaction.

**Experimental setup:**

- **Emission scan:** AF-488 *T.aq* MutS C42A/A764C (5nM) was mixed with unlabeled wild type β clamp (0.2µM, 0.5µM, 1µM) in Hepes NaOH pH 7.5 (20mM), NaCl (50mM) and MgCl₂ (5mM). Each reaction was excited at λₜₘₐₓ = 495nm and emission was collected from λₜₜₘₜ = 505nm to 605nm in individual cuvettes. The instrument was set up as follows: slit width (Ex/Em = 2nm/2nm), Increment (1nm), Integration time (1s), temperature (40°C) and incubation time (1min).

Measure $K_D$ of MutS-β clamp interaction by AF-488 MutS C42A/A764C: MutS-β clamp interaction at equilibrium was measured on a FluoroMax-3 fluorometer. AF-488 *T.aq* MutS C42A/A764C (50nM) was titrated against wild type β clamp (0-1µM) in 200µl reaction mixture containing HEPES pH 7.5 (20mM), NaCl (50mM) and MgCl₂ (5mM). Each reaction was observed in a separate cuvette after 1min incubation at 40°C. Raw data was corrected for photo bleaching and the apparent dissociation constant ($K_D$) was determined as described in IV.i.a.
III. ii. (a) Isolation of $\beta$ clamp and clamp loader subunits from *Thermus aquaticus* genome.

(a) Cloning: The genomic sequence of *T.aq* strain YT-1 is not publicly available. Primers to isolate the genes encoding the clamp loader subunits were designed for *T.aq* strain Y51MC23 (NCBI Reference ID: NZ_ABVK02000022.1). The gene encoding the $\beta$ clamp [dnaN], $\delta$ [holA], $\delta'$ [holB] and $\Upsilon$ [dnaX] were isolated from genomic DNA. Preparation of genomic is described in Appendix I (i). Primers used to isolate $\beta$ clamp and clamp loader genes from genomic DNA are listed in Appendix I (ii), and primers used to sequence the isolated genes are listed in Appendix I (iii).

The PCR protocol used to isolate the genes is as follows:

*PCR set up*

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phusion GC Buffer</td>
<td>5X</td>
<td>1X</td>
</tr>
<tr>
<td>dNTP mix</td>
<td>10mM</td>
<td>200μM</td>
</tr>
<tr>
<td>Forward primer</td>
<td>10μM</td>
<td>0.5μM</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>10μM</td>
<td>0.5μM</td>
</tr>
<tr>
<td><em>T.aq</em> genomic DNA</td>
<td>150ng</td>
<td>50ng</td>
</tr>
<tr>
<td>Phusion DNA polymerase (NEB)</td>
<td>2000U/ml</td>
<td>10U</td>
</tr>
<tr>
<td>dH$_2$O</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>100ul</td>
<td></td>
</tr>
</tbody>
</table>

PCR cycle set up for dnaN:

1. 98°C for 3min
2. 98°C for 10sec
(3) 62.5°C for 30sec [for holA=65°C and holB=50°C]
(4) 72°C for 45sec [for holA = 30sec]
(5) Goto step 2 for 30X [for holA=35X]
(6) 72°C for 5mins
(7) 15°C forever
(8) END

(b) PCR product was isolated and purified by the QIAquick Gel Extraction Kit (QIAGEN). PCR product and the vectors (pET11a and pET16b) were restriction digested as follows:

- **Setup of restriction digestion**

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>1µg</td>
</tr>
<tr>
<td>10X Buffer 3</td>
<td>5µl</td>
</tr>
<tr>
<td>Nde I</td>
<td>10U</td>
</tr>
<tr>
<td>Bam HI</td>
<td>10U</td>
</tr>
<tr>
<td>BSA</td>
<td>100µg</td>
</tr>
<tr>
<td>H₂O</td>
<td>uptil 50µl</td>
</tr>
<tr>
<td>Total</td>
<td>50ul</td>
</tr>
</tbody>
</table>

Incubate at 37°C for 3 hours and purify by GenElute PCR clean up kit. Ligation reaction was set up as follows:

- **Ligation reaction**:

<table>
<thead>
<tr>
<th>Component</th>
<th>20ul reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X T4 DNA Ligase Buffer</td>
<td>2ul</td>
</tr>
<tr>
<td>Component</td>
<td>Amount/Volume</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Vector DNA</td>
<td>50ng</td>
</tr>
<tr>
<td>Insert DNA</td>
<td>50ng</td>
</tr>
<tr>
<td>Nuclease free water</td>
<td>To 20ul</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>1ul</td>
</tr>
</tbody>
</table>

Optimal ligation ratio = 1:5

Thaw and resuspend ligase buffer at room temperature. Assemble the components on ice and add T4 DNA ligase last. Gently mix the reaction by pipetting up and down then microcentrifuge briefly. Incubate at 16°C overnight. Transform 10ul of ligation mixture into 50ul of DH5α cells and select using LB Ampicillin resistance plates. Select 5 colonies and isolate plasmid with the QIAprep Spin miniprep kit (Qiagen). To test if the target gene was inserted into the plasmid, the following restriction digest was done.

- **Restriction digest to check for insert:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount/Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid</td>
<td>300ng</td>
</tr>
<tr>
<td>10X Buffer</td>
<td>2.5ul</td>
</tr>
<tr>
<td>Enzyme</td>
<td>0.5ul</td>
</tr>
<tr>
<td>dH2O</td>
<td>Uptil 25ul</td>
</tr>
</tbody>
</table>

(c) Sequencing: Constructs were sequenced using T7 promotor and terminator primers. Genomic DNA was also sequenced using primers designed with the reference PubMed deposited sequence. The sequencing result of all the isolated genes is listed in Appendix I (iv). The sequence of the construct and the difference w.r.t the Pubmed deposited sequence is described in Appendix I (v).
III. ii (b) Making β clamp mutants that can report MutS-β clamp and MutS-β clamp-DNA complex in solution

*T.aq* β clamp is proposed to interact with MutS via residues in the hydrophobic pocket (G178, Y179) and/or on a loop at the inter-dimer interface (D282). To test this hypothesis, mutation G178C and D282C was introduced into *T.aq* β clamp pET11a plasmid by QuikChange Site-Directed Mutagenesis Kit (Stratagene). Primers used to make the mutants are described in Appendix IV (ii).

III (iii) β clamp and clamp loader expression and purification

(i) *Preparation of Biomass:* 1ul of plasmid were transformed into BL21 cells using standard protocol and plated on LB Ampicillin and incubated at 37°C for 15 hours. Single colony was picked and resuspended in 0.5ml LB and seeded into 1Lt LB media. Cells were grown at 37°C, 220 rpm until O.D at 600nm = 0.6 and induce with 1mM IPTG for 3hours at 37°C. Spin down cells at 5000 rpm for 20mins and resuspended in 20mM Tris pH7.5, 20% sucrose at store at -80°C. Resuspension buffer: biomass ratio is 10ml : 1Liter Biomass.

(ii) Purification of *T.aq* β clamp wild type and mutants was done using a combination of Ammonium Sulphate precipitation, DEAE Sepharose, Q column and P6DG gel filtration column as described in Appendix II (i). The protein was labeled with MDCC [7-Diethylamino-3-(((2-Maleimidyl) ethyl) amino) carbonyl) coumarin] as described in Appendix II (ii).

(iii) Purification of *T.aq* clamp loader subunits δ, δ’ and Υτ is described in Appendix III (a) (b) and (c).
III (iv) Study the interaction between MutS and β clamp using β clamp as a reporter.

Experimental validation of the proposed interaction between MDCC labeled *T.aq* β clamp G178C and D282C with MutS (+/- DNA) was done by conducting an excitation and emission scan on a FluoroMax-3 fluorometer. If a change in signal was observed upon mixing *T.aq* β clamp with unlabeled wild type MutS, these constructs can potentially be used to measure MutS-β clamp interaction.

*Experimental setup:*

- **Excitation scan:** MDCC labeled *T.aq* β clamp (50nM) was mixed with *T.aq* MutS (0.5μM) in 0.2ml reaction mixture containing Hepes NaOH pH 7.5 (20mM), NaCl (50mM) and MgCl₂ (5mM). Each reaction was monitored in individual cuvettes. Excitation was varied from λ<sub>ex</sub> = 390nm - 464nm and emission was collected at λ<sub>em</sub> = 474nm. The instrument was set up as follows: slit width (Ex/Em = 2nm/2nm), Increment (1nm), Integration time (1s), temperature (40°C) and incubation time (1min).

- **Emission scan:** The condition for emission scan were the same as excitation scan except excitation was fixed at λ<sub>ex</sub> =425nm and emission was collected from λ<sub>em</sub> = 440nm-550nm.

- **K<sub>D</sub> of MutS-β clamp interaction measured by β clamp:** MutS-β clamp interaction at equilibrium was measured on a FluoroMax-3 fluorometer. MDCC labeled *T.aq* β clamp (50nM) was titrated against MutS (0-10μM) in 0.2ml binding buffer. Each reaction was observed in a separate cuvette at λ<sub>ex</sub> = 425nm, and λ<sub>em</sub> = 465nm after 1min incubation at 40°C. Raw data was corrected for photo bleaching and the apparent dissociation constant (K<sub>D</sub>) was determined. To
determine the effect of DNA on complex formation, the experiment was also conducted in the presence of 37mer duplex DNA (0.5mM) containing either AT (homoduplex) or +T (heteroduplex).

- **Equilibrium binding of \(T.aq\) MutS and \(\beta\) by gel mobility shift assay (GMSA):** To determine \(K_D\) of wild type MutS and \(\beta\) clamp, MutS (50nM) was mixed with different concentration of \(\beta\) clamp (0-10uM) in 15\(\mu\)l reaction volume in containing 20mM TrisCl pH 7.5, EDTA 0.1mM, Glycerol 4%, BSA 50ug/ml, 100mM NaCl, 5mM DTT. The reaction was incubated at 40°C for 5 min. 10ul reaction was applied to 4% Native PAGE. After loading the samples, the gel was initially run at 180V for 1min, followed by electrophoresis in 1X TBE buffer at 19mA for 120min at room temp. The gel was then stained by SyproRuby and visualized on the Typhoon imager.

**III (v) Study \(T.aq\) \(\beta\) clamp - MutS interaction with \(\beta\) clamp loaded onto DNA:**

In the *Thermus aquaticus* model system, \(\beta\) clamp loaded onto a primer-template DNA substrate may be a stringent requirement to observe functional interaction between MutS and \(\beta\) clamp. \(\beta\) clamp loaded on DNA may interact differently with MutS (e.g. there may be an effect on MutS based on the orientation of \(\beta\) clamp loading with respect to the primer-template junction). \(\beta\) clamp is loaded onto the primer template junction with the C-terminal facing the junction. Kinase protection assays of C-terminal tagged \(\beta\) clamp titrated with MutS suggests that MutS binds the C-terminal face of \(\beta\) clamp\(^{(32)}\). To try and mimic a physiologically relevant complex between MutS and \(\beta\) clamp, we should study this interaction when \(\beta\) clamp is loaded onto DNA. To load the \(\beta\) clamp on a DNA
substrate, the clamp loader complex must be isolated as described in Appendix III and reconstituted as follows:

• **Reconstitution of the clamp loader complex**: The clamp loader complex will be assembled in two stages based on a protocol developed for the *E.coli* clamp loader. Equimolar quantities of δ and δ’ will be pre-incubated for 40 minutes at 15°C. This will be followed by the addition of the Υ subunit to the δ: δ’ complex and incubation will be continued for another 40 minutes at 15°C to a final ratio of 1.2:1:1 subunits. The mixture will then be loaded on a Fast flow Q Sepharose column and washed with 100 mM NaCl. The complex will be eluted with a 100 mM to 400 mM NaCl gradient. Aliquots of the eluted fraction will be analyzed on a 12% SDS PAGE gel to detect fractions containing the fully formed clamp loader complex.

**III (vi) Design of a DNA substrate that allows for β clamp loading with a long half-life of the β clamp-DNA complex.**

**Approach**: A forked DNA substrate has been shown to be suitable for clamp loading \(^{(44)}\). Salient features of the proposed fork-DNA substrate (Figure III. i. a.) are as follows:

• Primer-template junction: The clamp loader can load the β clamp on 3’ primer-template junction with high efficiency \(^{(45)}\).

• One end of the linear substrate is modified with biotin. A streptavidin coated magnetic bead will be attached at this end by coupling with biotin.

• The magnetic bead will be used to pull down β clamp loaded onto DNA.

• Clamp sliding off the DNA is prevented by the 3’ and 5’ fork.
• *Taq* MutS binds +T mismatch with high affinity (40). A 37 bp core sequence has been well characterized in our laboratory and will be placed at varying distances from the primer-template junction to scan for an optimum distance that enables β clamp-MutS interaction. The first +T insertion will be placed 26bp from the 3’ primer-template junction. The position of the mismatch will be varied upto 50 bp from the junction if necessary. The orientation of the +T, relative to the 3’ junction will also be tested by flipping the 37 bp core sequence.

• To monitor MutS-DNA interaction in the presence of β clamp, 6 methyl isoxanthopterin (6MI) fluorophore will be placed adjacent to +T.

Schematic describing the construction of the bi-forked DNA substrate is shown in Appendix V.
Figure III (i) a
Schematic of the proposed bi forded DNA substrate for clamp loading.
III (vii) Development of an assay to measure clamp loading.

*Rationale:* To study MutS-β clamp interaction on DNA, β clamp must be loaded onto DNA. The efficiency of clamp loading will be determined by a magnetic-bead pull down assay (Figure III. i. b).

*Approach:* Biotin labeled DNA substrate will be coupled with Dynabeads M-280 Streptavidin (Invitrogen). The β clamp-DNA complex will be isolated by placing the reaction mixture in DynaMag Spin magnetic particle concentrator (Invitrogen) and analyzed by SDS-PAGE.

*Experimental setup:* The clamp loading reaction containing β clamp, DNA and clamp loader will be initiated by the addition of ATP. Loaded β clamp-DNA complex will be isolated from the reaction mixture by placing it in a magnetic concentrator and removing the supernatant. The pellet will be washed with binding buffer to remove excess ATP and free β clamp, followed by treatment with 1% SDS to release the loaded β clamp. The eluent will be analyzed on SDS PAGE gel and stained with SybrRuby or InVision His stain to visualize loaded β clamp.

*Control:* The reaction will be set up in the absence of ATP or clamp loader complex. To maximize the loading assay, excess DNA and clamp loader can be used or the incubation time can be increased.
**Expected results:**

Lane 1: Marker

Lane 2: Load contains the reaction mixture.

Lane 3: In the absence of the clamp loader, $\beta$ clamp is detected in the reaction mixture.

Lane 4: In the presence of the clamp loader, assuming there is 100% efficiency, $\beta$ clamp is loaded on the DNA substrate and no $\beta$ clamp is detected in reaction mixture.

Lane 5: Elution contains $\beta$ loaded onto DNA upon treating with SDS.
III (viii) Determine the half-life of the β clamp-DNA complex

The β clamp-DNA complex will dissociate over time. To assess the stability of the complex, I have to determine its half-life ($t_{1/2}$). This result will also give us information about the fraction of β clamp-DNA complex we can expect to be present during our experiments. The *E.coli* β clamp and yeast PCNA have reported half-lives of ~1 hour and 24 minutes \(^{(46)}\).

**Approach:** Half-life of β on DNA will be measured by monitoring the dissociation of β clamp from β clamp-DNA complex over time.

**Experimental setup:** Following the clamp loading assay, the pellet containing β clamp-DNA complex will be re-suspended in solution and incubated over time. At different intervals, aliquots of the reaction will be spun down and analyzed for β clamp remaining on DNA or released into solution, as described above. The β clamp in these fractions will be quantified and plotted versus time to get an observed rate ($k_{obs}$), and half-life $t_{1/2} = 0.693/k$.

III (ix) Isolation of β clamp-DNA complex for experiments with MutS.

To study the dynamics of the MutS-β clamp interaction on DNA (described in Aim I) the β clamp-DNA complex must be isolated.

**Approach:** Clamp loading will be performed in the presence of His-tag Dynabeads (Novex) coated with cobalt-based Immobilized Metal Affinity Chromatography (IMAC) that can selectively bind His-tagged proteins.

**Experimental setup:** The clamp loading assay will be carried out as described in Aim II (c) with His-tag Dynabeads instead of streptavidin coupled Dynabeads. Pulled down *His*β-DNA complex will be separated from the Dynabeads by
washing the pellet with 300 mM Imidazole. The supernatant (containing Hisβ-DNA complex) will be run through a PD10 desalting column to remove Imidazole.
IV. Results and Discussion:

(i) To monitor the MutS - β clamp interaction w.r.t labeled MutS:

(a) Purification of *Thermus aquaticus* MutS C42A/A764C: *T.aq* MutS C42A/A764C obtained was > 95% pure and the labeling efficiency with AF-488 was ~98% as seen in Figure IV. i. a.

(b) Functional analysis of *Thermus aquaticus* MutS C42A/A764C: *T.aq* MutS C42A/A764C is functional and binds +T DNA with wild type affinity (15nM-30nM range). Therefore, I proceeded to use Alexa Fluo 488 labeled *T.aq* MutS C42A/A764C to report MutS-β clamp interaction as shown in figure IV. i. b.

(c) Emission scan of *Thermus aquaticus* MutS C42A/A764C: Buffer (purple) and β clamp (blue) alone samples were close to zero indicating there is no background fluorescence in the reaction mixture. AF-488 *T.aq* MutS C42A/A764C gave a signal at ~ 5*10^-5cps (green). Upon addition of 0.2µM and 1µM β clamp, I observed a decrease in signal (black and grey) relative to AF-488 MutS alone. However upon addition of 0.5µM β clamp I observed an increase in signal (red). λ emission maximum was consistent at 518nm as seen in Figure IV. i. c.

The decrease in signal observed at 0.2µM and 1µM of β clamp could indicate either interaction between MutS and β clamp or photo bleaching. The increase in signal at 1µM β clamp could be an anomaly as I just have 3 data points for comparison. Therefore I proceeded to determine if Alexa Fluo-488 *T.aq* MutS C42A/A764C was undergoing photo bleaching.
(d) Photo bleaching of AF-488 *Thermus aquaticus* MutS C42A/A764C:

AF-488 *T.aq* MutS C42A/A764C (50nM) was mixed with wild type β clamp (1μM) in 200μl reaction containing HEPES pH 7.5 (20mM), NaCl (50mM) and MgCl₂ (5mM). Consecutive reading was taken after incubating the sample at 40°C while stirring for 2mins, followed for 30sec without stirring. Over 20 exposures I did not see significant photo bleaching of AF-488 *T.aq* MutS C42A/A764C (AF-488 signal dropped from 1.225X10⁶ to 1.2X10⁶ i.e < 2% decrease in signal) as seen in Figure IV. i. d. Therefore I can proceed to determine the equilibrium binding constant between MutS and β clamp with AF-488 *T.aq* MutS C42A/A764C.

(e) Emission scan of AF-488 *Thermus aquaticus* MutS C42A/ A764C in the presence on various nucleotides: Since MutS is an ATPase, emission scan was also performed in the presence of ATP, ADP and ATPγS. In the presence of ATP, AF-488 *T.aq* MutS C42A/ A764C undergoes a 1.5X decrease in fluorescence signal when compared to no nucleotides, ADP and ATPγS (a non-hydrolysable analogue of ATP). as seen in Figure IV. i. e. Therefore *T.aq* MutS C42A/ A764C can be used to report ATP binding and hydrolysis.

To monitor *T.aq* MutS ATP binding and hydrolysis, our laboratory has previously generated a cysteine mutant of *T.aq* MutS in the mismatch binding domain I [*T.aq* MutS C42A/M88C] and labeled it with the 2- (4’- (iodoacetamido) anilino) naphthalene-6-sulfonic acid [IAANS] fluorophore. Even though position 88 in domain I (mismatch binding domain) is distant from the C terminal nucleotide binding region, it has been shown that the mismatch binding
and ATP binding domains of MutS are coupled [32]. Since *T.aq* MutS C42A/A764C is at the C terminal and near the nucleotide binding region, it can only report nucleotide binding and hydrolysis.

*Discussion:* AF-488 *T.aq* MutS C42A/ A764C cannot be used to monitor the MutS-β clamp complex. However this construct can be used to report ATP binding and hydrolysis as discussed above.

(f) To measure *K*<sub>D</sub> of MutS-β clamp interaction measured by MutS reporter:
AF-488 *T.aq* MutS C42A/ A764C did not show a concentration dependent change in signal upon the additional of unlabeled β clamp as seen in Figure IV. i. f. The signal scattered around 1.39×10<sup>6</sup> cps even at high concentration of β clamp.
Figure IV (i) a. *T.aq* MutS C42A/A764C was purified with Q sepharose and PD10 columns and labeled with Alexa Fluor 488 dye.

![Purification of T.aq MutS C42A/A764C](image)

**Figure IV (i) a. Purification of T.aq MutS C42A/A764C**

Q Sepharose column: Ladder, Pellet, Lysate, heat treated pellet, Load, FT, Wash, Fractions (1-72)

Ammonium sulphate precipitation.

*T.aq* MutS ~ 90kDa

**Labelling of T.aq MutS C42A/A764C with Alexa Fluor 488**

1: Before concentration. 2: After concentration 3: Filtrate 4: Control, Free dye.

Visualized on UV illuminator. Stained with Coomassie blue dye

Figure IV (i) b. Functionality of *T.aq* MutS C42A/A764C was determined by measuring it’s binding affinity to mismatched DNA substrate.

![Taq MutS C42A/A764C binding to 2AP +T DNA](image)

**Taq MutS C42A/A764C binding to 2AP +T DNA**

**Reaction conditions:**
- Buffer: 20mM HEPES pH 7.7, 50mM NaCl, 5mM MgCl2,
- 0.01μM 2Ap+T DNA,
- 0-0.2μM *Taq* MutS C42A/A764C, 30sec incubation, 40°C

MutS interaction with 2AP+T DNA

+T insertion

**2-Aminopurine**

**MutS Phenylalanine**

![Graph](image)

**Graph**

$K_D = 28\text{nM}$
Figure IV (i) c. An emission scan of *T.aq* MutS C42A/A764C with β clamp was done to determine it could report MutS- β interaction.

![Emission scan of AF-488 T.aq MutS C42A/A764C](image)

Figure IV (i) d. Photobleaching of *T.aq* MutS C42A/A764C was done to determine the stability of the Alexa Fluor 488 dye.

![Photobleaching of AF-488 T.aq MutS C42A/A764C](image)
Figure IV (i) e. An emission scan of *T.aq* MutS C42A/A764C with various nucleotides showed that it could report ATP binding and hydrolysis.

![Emission scan in presence of nucleotides](image)

**Emission scan in presence of nucleotides:**
Hepes NaOH pH 7.5 - 20 mM, NaCl - 50 mM, MgCl2 - 5 mM
*T.aq* MutS C42A/A764C MDCC = 50 nM,
ADP, ATP, ATPγS = 0.5 mM
Excitation = 425 nm, Emission = 440 nm-550 nm.

Ex $\lambda_{max}$ = 465 nm

Figure IV (i) f. Equilibrium binding of *T.aq* MutS C42A/A764C with β clamp could not report MutS–β interaction.

![Equilibrium binding of MutS C42A/A764C with β clamp](image)

**Equilibrium binding of MutS C42A/A764C with β clamp**
Reaction conditions:
Hepes NaOH pH 7.7-20 mM, NaCl=50 mM, MgCl2 = 5 mM.
*T.aq* MutS C42A/A764C MDCC = 50 nM, β WT=14.5 uM. T=40C.
Excitation = 425 nm, Emission = 465 nm.
(ii) To monitor the MutS - β clamp interaction w.r.t labeled β clamp:

(a) Isolation of β clamp and clamp loader genes from *T.aq* genomic DNA: Genes encoding the β clamp (dnaN), δ subunit (holA), δ’ subunit (holB) and Υτ (dnaX) were successfully isolated from *Thermus aquaticus* strain YT-1 as shown in Figure IV. ii. a. The genes were then sub-cloned in vectors pET11a and pET16b to obtain untagged and His-tag versions of the proteins and shown in Figure IV. ii. b.

(b) Purification and labeling of *T.aq* β clamp WT and mutants: *T.aq* β clamp was purified by a combination of ammonium sulphate precipitation, DEAE column, Q column and PD6 gel filtration columns. Elution profile of *T.aq* β clamp is shown in Figure IV (ii) c and the protein obtained was >95% pure. *T.aq* β clamp mutants G178C and D282C were labeled with MDCC and further purified on the PD6 gel filtration column to remove excess dye. The maximum labeling efficiency obtained was 97%.

(c) Purification of *T.aq* clamp loader: *T.aq* δ subunit was purified by Nickel Sepharose fast flow and Q Sepharose fast flow. *T.aq* δ’ and Υτ subunit was purified only by Nickel Sepharose fast flow. *T.aq* δ and δ’ were >95% pure. Protocol to obtain pure *T.aq* Υτ is still to be optimized. The protein purification profiles of all the subunits is shown in Figure IV (ii) d.
(d) Excitation and emission scan of MDCC labeled \(T.aq\) β clamp D282C:

The excitation scan of \(T.aq\) β clamp D282C indicated that there was no background fluorescence and MDCC \(T.aq\) β clamp D282C could be excited at \(\lambda_{\text{max}} = 425\text{nm}\). However, the emission scan of \(T.aq\) β clamp D282C did not show a change of signal in the presence of unlabeled MutS (compare red trace: MDCC D282C β clamp to green trace: MutS with MDCC D282C β clamp) as seen in Figure IV (ii) e. Therefore, \(T.aq\) β clamp D282C is not a suitable construct to monitor MutS-β clamp interaction.

(e) Excitation and emission scan of MDCC labeled \(T.aq\) β clamp G178C:

The excitation scan of \(T.aq\) β clamp G178C indicated that there was no background fluorescence and MDCC \(T.aq\) β clamp G178C could be excited at \(\lambda_{\text{max}} = 425\text{nm}\). In the emission scan, \(T.aq\) β clamp G178C showed a 1.2 fold increase of signal in the presence of unlabeled MutS (compare red trace: MDCC G178C β clamp to green trace: MutS with MDCC G178C β clamp) as seen in Figure IV (ii) f. This indicates that \(T.aq\) β clamp G178C could be potentially used to report MutS-β clamp interaction. Therefore I proceeded to use this construct to determine the of \(K_D\) MutS-β clamp in solution.

(f) \(K_D\) of MutS-β clamp interaction reported by MDCC \(T.aq\) β clamp G178C:

Literature survey: \textit{E.coli} MutS binds to β clamp with a \(K_D\) value of \(~250\text{ nM}\) as determined by native PAGE assay \(^{16}\). Msh2-Msh6 and PCNA interaction in yeast was monitored by SPR and a \(K_D\) of 2.5 nM \(^{43}\) and 6 nM \(^{10}\) was reported. In
humans, MutSα-PCNA interaction was studied by equilibrium gel filtration and a $K_D \sim 700$ nM was reported \cite{35}.

$T.aq$ β clamp G178C shows a concentration dependent change in signal upon the addition of unlabeled MutS. However, the curve does not saturate even at high concentration of MutS (10µM) and the estimated $K_D$ is 9µM (+/- 3µM) as seen in Figure IV (ii) g. The observed high $K_D$ could be due to the following caveats of the study:

- The binding conditions are not optimal.
- Since MutS is an ATPase and a DNA binding protein, the presence of nucleotides and/or DNA may be required for MutS-β clamp interaction.
- In $T.aq$, β clamp loaded on DNA may be a stringent requirement for interaction with MutS.
- The putative β clamp binding motif is absent in $T.aq$ MutS, therefore β clamp may not interact at all and the signal observed may be due to non-specific binding seen at higher concentrations of protein.

The second caveat was tested by measuring equilibrium binding constant of MutS-β clamp in the presence of homoduplex and heteroduplex DNA.

\(^{(g)}\) $K_D$ of MutS-β clamp interaction in the presence of DNA reported by MDCC labeled β clamp G178C mutant: In the presence of DNA, $T.aq$ β clamp G178C showed a concentration dependent change in signal upon the additional of unlabeled MutS. However, the curve still did not saturate and the estimated $K_D$ for homoduplex (AT 37mer) DNA is 7.3µM +/- 1.6µM [Figure IV (ii) h], comparable within error to MutS-β clamp in solution. In the presence of heteroduplex DNA
the $K_D$ was 3µM +/- 2.6 µM. The observed $K_D$ in the various conditions (3µM-9µM) is much greater than the $K_D$ reported for *E.coli* MutS-β clamp binding by GMSA (~250nM) \(^{(16)}\). Therefore, it is possible that either the mutation G178C is disrupting the interaction between MutS and β clamp accounting for the observed high $K_D$, or *T.aq* MutS and β clamp do not interact at all and what we observe is non-specific binding at high concentrations of protein. To resolve this issue, I determined the $K_D$ of wild type *T.aq* MutS and β clamp by GMSA. Alternatively, if placing a fluorophore in the hydrophobic pocket is interrupting the binding then we can end label β clamp with pyrene fluorophore and measure MutS binding by fluorescence anisotropy.

(h) Equilibrium binding of *T.aq* MutS and β by gel mobility shift assay (GMSA): As seen in Figure IV (ii) i. no distinct super-shifted band relative to free MutS and free β clamp was observed, indicating that wild type *T.aq* MutS and β clamp did not interact under these experimental conditions. The band observed at higher β clamp concentrations may be due to non-specific binding. It is possible that these conditions are not optimized to enable MutS-β clamp interaction. MutS may interact with β clamp in the presence of DNA and/or nucleotides. Therefore a series of conditions were screened to determine the optimal condition.

(i) Screen for optimum conditions to measure $K_D$ of MutS and β clamp by GMSA: MutS-β clamp interaction was tested under various conditions like the presence of DNA, nucleotides and *T.aq* MutL. The reactions were set up as described in materials and methods but with the following modifications:
Reactants = β clamp (6µM), MutS (0.5µM), ATP (1mM), ADP (1mM), +T 37mer DNA (1µM), AT 37mer DNA (1 µM), and T.aq MutL (0.5 µM).

Reaction conditions = Incubated at 50°C for 5 min.

Again, as seen in Figure IV (ii) j and k, no distinct super shifted band was observed relative to free MutS and free β clamp, indicating that wild type T.aq MutS and β did not interact under all the conditions tested.

Wild type T.aq MutS and β clamp did not show interaction in all experimental conditions tested. It is possible that GMSA is a stringent assay for T.aq, or the incubation temperature needs to be higher (~70°C) to enable interaction. It is also possible that MutS interacts with β clamp only when it is loaded onto DNA. These possibilities need to be further tested.
Isolation of $\beta$ clamp and clamp loader genes from *T.aq* genomic DNA.

The genes dnaN ($\beta$), holA ($\delta$) and holB ($\delta'$) were directly cloned N terminal to C terminal out of the *Thermus aquaticus* strain YT-1 genome. dnaX ($\gamma\tau$) was first cloned from N terminal to +400bp after the stop codon into pBluescript vector and then sub-cloned into the expression pET vector.
Figure IV (ii) b

Construction of plasmids to express tagged and untagged clamp and clamp loader.

Figure IV (ii) c

Purification and labeling of *T. aq* β clamp by ammonium sulphate precipitation, DADE column, Q sepharose, PD6 gel filtration column and labeling by Alexa Fluor 488.
Figure IV (ii) d

Purification of *T.aq* clamp loader subunits by ion exchange and affinity chromatography
Figure IV (ii) e

Excitation and emission scan of MDCC labeled T.aq β clamp D282C mutant.

Figure IV (ii) f

Excitation and emission scan of MDCC labeled T.aq β clamp G178C mutant.
Figure IV (ii) g. Equilibrium binding experiment to measure β G178C MDCC and MutS interaction.

![Equilibrium binding of β G178C MDCC to MutS.](image)

$K_D \sim 9\mu M (\pm 3\mu M)$

Figure IV (ii) h. Equilibrium binding experiment to measure β G178C MDCC and MutS interaction in the presence of DNA.

![Equilibrium binding of β G178C MDCC to MutS in the presence of DNA.](image)

$K_D \sim 3.4\mu M (\pm 2.6\mu M)$

$K_D \sim 7.3\mu M (\pm 1.6\mu M)$
Figure IV (ii) i. Equilibrium binding experiment to measure β and MutS interaction by GMSA.

Figure IV (ii) j. Presence of DNA, nucleotide tested to detect β and MutS interaction by GMSA.
Figure IV (ii) k. Presence of MutL, DNA and nucleotide tested to detect β and MutS interaction by GMSA.
V. Conclusions and future directions

This research reports the reconstitution of the components of the first minimal \textit{in vitro} non-MutH dependent DNA replication and repair in the \textit{Thermus aquaticus} model system that is most similar to a eukaryotic model system. By labeling MutS, β clamp and DNA independently we had hoped to be able to dissect the molecular level details of the interaction between MutS - β clamp and MutS – DNA in the presence of loaded β clamp by employing both bulk and single molecule methods to accurately mimic \textit{in vivo} conditions.

Based of available literature, MutS was labeled at position A764C at the C-terminal with Alexa-488 dye. Labeling of this position unfortunately did not report β clamp binding. However MutS A764C could report ATP binding and hydrolysis as it is near the ATP binding site. To monitor the MutS- β clamp interaction by labeling β clamp, two constructs were tested (D282C and G178C). MDCC labeled β clamp D282C could not report MutS binding but β clamp G178C showed a change in the emission scan. Based on this result, in solution assays to measure the binding constant between MutS and β clamp was done but it reported a very weak $K_D$. It is possible that ensemble fluorometer measurements may not be ideal to determine this type of protein-protein interaction and we may have to explore single molecule techniques to successfully monitor this interaction and determine the equilibrium binding constant. Additionally, the assay conditions may not be optimal to measure this interaction. It is possible that GMSA is a stringent assay for \textit{T.aq}, or the incubation temperature needs to be higher (~70°C) to enable interaction. It is also possible that MutS interacts with β clamp only when it is loaded onto DNA. First it is advisable to test if WT MutS
and β clamp interact by methods such as SPR or gel filtration. Once that is established the site specific labeling to explore the protein-protein interactions can be carried forward. A FRET assay between labeled MutS and β clamp can also be done to see interaction. Just AF 488 labeled MutS may report β clamp binding by anisotropy, which has still not been explored.

The presence of nucleotides and DNA could effect MutS since it is an ATPase, or the loading of the β clamp on to the DNA substrate could give the β clamp directionality with respect to the replication fork and therefore effect it’s interaction interface with MutS. These possibilities need to be further tested.
VI. Addendum

Additional projects on the kinetic mechanism of *Escherichia coli*

Heptosyltransferase I

(i) Background and Significance:

The emergence of multi-drug resistant bacteria due to the widespread use and misuse of antimicrobial agents has made it necessary to develop new drug targets. Gram-negative bacteria have an outer lipopolysaccharide layer (LPS) that is important for biofilm formation. The LPS layer significantly hampers the uptake of hydrophobic antibiotics increasing the survival of bacteria sheltered in biofilms. A new approach is being developed that can inhibit biofilm formation and increase the efficacy of existing antimicrobial agents \(^{(61,62)}\). Biofilm inhibitors can also be used to treat persistent infections in immune compromised patients \(^{(63)}\).

Heptosyltransferase I (Hep I) belong to the Glycosyltransferases structural family B (GT-B) and catalyzes the transfer of the first heptose moiety in the core oligosaccharide of LPS. GT-B enzymes are characterized as having two \(\beta/\alpha/\beta\) Rossman-like domains connected by a linker region and are predicted to undergo conformational change from open to closed states for catalysis to occur (Figure VI. a. i) \(^{(64)}\). The *Escherichia coli* Hep I has been expressed and structurally characterized (PDB 2H1H), but complete kinetic and mechanistic analysis of this enzyme has yet to be performed \(^{(65-67)}\).

Hep I catalyzes the transfer of \(L\)-glycero-\(D\)-manno-heptose from ADP-\(\beta\)-L-glycero-\(D\)-manno-heptose (APD-Hep) to the first 3-deoxy-\(\alpha\)-D-oct-2-ulpopyranosonic acid (KDO) of KDO\(_2\)-LipidA by the formation of an \(\alpha\) \((1\rightarrow5)\) linkage (Figure VI. a. ii) \(^{(61-68)}\). The N terminal domain of Hep I binds the KDO\(_2\)-
LipidA and the C terminal domain binds ADP-Hep. The crystal structures of apo HepI, and HepI with ADP and ADP-L-glycero-D-manno-Heptose analogue in the donor binding site are available (PDB codes: 2H1F, 2H1H and 2GT1). The structures of HepI with Lipid A or both substrates have not been determined. In all crystal structures HepI adopts an open conformation (Figure VI. a. i), which shows that the catalytic residue Asp13 that extracts a proton from the Lipid A substrate to generate an oxyanionic nucleophile is more than 8 Å away from the anomeric carbon of ADP-Hep. Based on the crystal structure and MD simulation, Hep I will need to adopt a closed conformation in order to bring the substrates into proximity and exclude the water from the active site to bring Asp13 into a catalytically competent position.

This project aims to understand the protein dynamics and kinetic mechanism of Hep I in collaboration with Prof. Erika Taylor’s laboratory. While the Taylor group is characterizing the catalytic mechanism of Hep I, the Hingorani group will elucidate the kinetic pathway of Hep I. By understanding the mechanism of how Hep I catalyzes the addition of KDO₂-LipidA in the presence of ADP we expect to enhance drug discovery against this class of bacterial enzymes.

(ii) Specific Aim:
(a) Characterize the binding of substrates to HepI:

The binding of Hep I to KDO₂ was being studied in the Taylor lab with a simplified lipid substrate, O-deacylated lipid A (ODLA), as it is soluble in water and is catalyzed with similar efficiency as the native substrate. Hep I contains eight tryptophan residues and fluorescence spectra of Hep I (λ = 290nm) in the
presence and absence of ODLA and ADP was done to determine if intrinsic tryptophan can be used to monitor conformation change upon substrate binding. Apo Hep I and Hep I with ADP-Hep have identical fluorescence spectra under equilibrium conditions (Figure VI. a. iii). In contrast, HepI spectra in the presence of ODLA showed a significant blue shift (with or without ADP-Hep) indicating movement of one or more of the tryptophans into a relatively non-polar environment on binding this substrate. These tryptophans are likely to be one of the four surface exposed residues in the N-terminal domain (W47, W62, W66 or W116) all of which are predicted to undergo movements of greater than 7 Å from their locations in the open apo enzyme structure to the closed conformation. (Figure VI. a. iv).

Stopped flow experiments were done to determine the kinetics of this conformation change in HepI and its place in the reaction mechanism. HepI was rapidly mixed with the substrates and fluorescence was monitored over time (note: a 350 nm cut-on filter was used and hence there is an apparent decrease in fluorescence intensity on ODLA binding to Hep I). Consistent with the results in Figure VI. a. iii, the signal did not change when HepI was mixed with buffer or ADP-Hep alone but with ODLA there is a biphasic decrease in signal which yields a fast rate of 21s-1 and a slower rate of 5s-1 at 1.5μM ODLA when fit to a double exponential function (Figure VI. a. v).

A concentration dependence of ODLA binding to HepI over the range of 1-100μM showed that the fast rate increased from 1-5μM and then became independent of ODLA concentration (Figure VI. a. vi). Is it to be noted that at ODLA concentrations above 20μM there is some signal interference likely due to
light scattering caused by the lipid. The rate of the second phase however remained constant at \( \sim 5 \text{ s}^{-1} \) and then increased slightly at > 10\( \mu \text{M} \) ODLA. The plot of these rates verses ODLA concentration (Figure VI. a. vii) suggests that the fast rate shows hyperbolic dependence on ODLA concentration saturating at \( k_{\text{fast}} = 80 +/− 7 \text{ s}^{-1} \) with a dissociation constant \( K_{d} \) of 3 +/− 0.7\( \mu \text{M} \). This suggests a two-step binding mechanism in which formation of the initial collision complex between Hep I and ODLA is followed by rate limiting isomerization to form the final Hep I-ODLA complex (Figure VI. a. viii). Since the \( K_{d} \) value is comparable previously reported ODLA \( K_m \) of 0.9\( \mu \text{M} \), we assumed that ODLA binding to Hep I is in rapid equilibrium in this two-step model. \( K_{d} \) is equivalent to \( K_{D1} \), the saturating rate is \( k_2 + k_{−2} \) and the y intercept is \( k_{−2} \) (estimated at < 5 s-1). The apparent bimolecular rate constant for ODLA binding can be estimated as \( k_{\text{on}} \sim 2.5 \times 10^{7} \text{M}^{-1} \text{s}^{-1} (k_2/K_{D1}) \). The overall dissociation constant is estimated as 0.2\( \mu \text{M} \) \( (K_{D1}k_{−2}/k_2) \), which is comparable to the \( K_D \) of 0.6 +/-0.1\( \mu \text{M} \) obtained from the hyperbolic dependence of the signal amplitude on ODLA concentration.

The second slower change in intrinsic fluorescence cannot be interpreted at present. The rate remains nearly constant (\(\sim 5 \text{ s}^{-1}\)) until higher ODLA concentrations (and potentially higher error in signal), this phase may reflect subsequent conformational change in the enzyme (Figure VI. a. vii). It is possible that this signal may also originate from a Hep I subpopulation or an off-pathway species. Further investigation of the reaction mechanism is needed to resolve this question.

Similarly, experiments done in the presence of 50 \( \mu \text{M} \) ADP-Hep showed the same hyperbolic dependence of the fast rate on ODLA concentration and
almost the same kinetic parameters for the binding mechanism (Figure VI. a. vii; $K_{1/2}$ of 3.9 +/- 1μM and $k_{\text{fast}}$ of 87 +/- 9 s$^{-1}$). This confirms that lipid is responsible for inducing the observed conformation change in the enzyme. Similar experiments performed with the catalytically inactive mutant Hep I D13A, with and without ADP-Hep, also yields the same results (Figure VI. a. vii; $K_{1/2}$ of 2.6 +/- 0.7μM and $k_{\text{fast}}$ of 93 +/- 8 s$^{-1}$, with ADP-Hep; $K_{1/2}$ of 3.6 +/- 0.9μM and $k_{\text{max}}$ of 97 +/- 9 s$^{-1}$ without ADP-Hep). Therefore ODLA induced change in Hep I occurs prior to chemistry. While limiting formation of the enzyme-substrate complex, this isomerization is fast relative to $k_{\text{cat}}$ (0.3 s$^{-1}$), implying that other step(s), including conformational dynamics before or after chemistry limit catalytic turnover.
Figure VI. a. i

Crystal Structure of Hep I enzyme and it’s proposed conformation change.

Figure VI. a. ii

The Hep I reaction mechanism.
Figure VI. a. iii

Emission spectra of HepI with ODLA ± ADP-Hep are blue shifted when compared with apo HepI (neither substrate contributes significantly to the signal).

The gray dashed line indicates the 350 nm cut-on filter boundary.

---

Figure VI. a. iv

HepI open structure in tan, with tryptophan residues denoted in green, superimposed with a structural model of closed HepI in grey, with tryptophans in magenta.
Figure VI. a. v

Stopped-flow traces show rapid, biphasic decrease in HepI fluorescence when mixed with ODLA alone (1.5 μM) or with ODLA plus ADP-Hep (1.5 μM and 50 μM, respectively), but no change with ADP-Hep alone.

Figure VI. a. vi

Representative traces from a titration of HepI with ODLA. The data were fit with a double exponential function to determine the rate constants for Figure VI. a. vii.
Kinetics of ODLA binding and associated change in HepI conformation

(A) Observed rates from WT HepI titrated with ODLA,

(B) WT HepI pre-incubated with 50 μM ADP-Hep and titrated with ODLA, and

(C) HepI D13A pre-incubated with 50 μM ADP-Hep titrated with ODLA.

(fast phase: blue; slow phase: red).
Figure VI. a. viii.

Scheme of two-step binding of ODLA to HepI

\[ E + ODLA \rightleftharpoons E \cdot ODLA \rightleftharpoons *E \cdot ODLA \]

\[ K_{D1} = 3\mu M \quad k_2 = 75s^{-1} \quad k_2 = <5s^{-1} \]
(b) Characterize the release of ADP after catalysis from Hep I:
We know that binding of ODLA to Hep I occurs by a two-step mechanism, and the ODLA induced conformational change in Hep I occurs prior to chemistry. But the steps following catalysis are not known.

I aim to characterize one of the steps following catalysis, which is ADP release. ADP can be detected in real-time with the ParM ADP sensor \(^{68}\). 7-Diethylamino-3-(((2-Maleimidyl) ethyl) amino) carbonyl) coumarin labeled ParM protein (MDCC-ParM) can bind ADP at a fast rate \((k_{on} = 0.65 \mu M^{-1}s^{-1})\) & high affinity \((K_D = 0.46 \mu M)\) and undergoes >3.5X increase in fluorescence intensity \(^{68}\). I plan to measure ADP release from Hep I after catalysis to determine if this step is the rate limiting step of Hep I reaction.

(iii) Preliminary Data:

Experimental setup: In an initial experiment, the Hep I reaction was performed on the KinTek stopped-flow instrument. ADP release from Hep I was measured by change in fluorescence of MDCC-ParM upon ADP binding at \(\lambda_{ex} = 424 \text{ nm}\) and \(\lambda_{em} = 450 \text{ nm}\). A calibration curve relating the MDCC-ParM fluorescence signal to ADP concentration was generated prior to the experiment. The setup of the experiment is shown in (Figure VI. b. i).

(iv) Result: ADP release from Hep I resulting from catalysis was observed (Figure VI. b. ii.). A near stoichiometric amount of ADP (~3 \(\mu M\)) was released in the burst phase (at a rate of ~4 \(sec^{-1}\)) and the second phase is a steady-state rate of 0.4 sec\(^{-1}\), which is comparable to the previously determined \(k_{cat}\) of 0.9 sec\(^{-1}\) \(^{61}\). The experimental conditions still need to be optimized and the next experiment would be to test Hept I mutants that are deficient in substrate binding and catalysis.
(v) *Future experiments:*

The substrate induced conformation change of Hep I binding to ODLA was studied by intrinsic tryptophan fluorescence as a blue shift was observed \(^{(62)}\). In preliminary experiments performed by Joy Cote in the Taylor lab and computational modeling performed by Carlos A. Ramirez-Mondragon from the Gao lab at University of Minnesota, residue W62 on the N3 loop of Hep I N terminal domain has been identified as responsible for the observed fluorescence (Figure VI. b. iii).
Figure VI. b. i.

Experimental setup of the experiment to determine the ADP release rate of HepI by MDCC-ParM.

Figure VI. b. ii.

ADP release from Hep I resulting from catalysis
Autodocking simulation of fully deacylated lipid A (ODLA) with Hep I

This residue is sandwiched between positively charge residues and has a sequence conservation of >80% across species (Figure VI. a. vi). Autodocking experiments of fully deacylated lipid A (ODLA) with Hep I predict that residues R60 and R120 bind to the negatively charged phosphate on the lipid substrate. Molecular dynamic simulations suggest that only when DLA is bound to the N terminal domain, the ADP-Hep is able to stably bind the C terminal domain, otherwise it is immediately released and catalysis is prevented. The nest step is to investigate the hypothesis of an ordered binding mechanism of Hep I by testing substrate binding induced conformational changes and ADP release from Hep I N3 loop mutants.
VII: Appendices

Appendix I:

(a) Preparation of *Thermus aquaticus* genomic DNA strain YT-1.

The preparation of genomic DNA from *Thermus aquaticus* was performed suing a modification of the method proposed by Hoffman and Winston (1987) for the preparation of yeast DNA. This technique is outlines below:

- Grow 10ml of culture to saturation in the appropriate media.
- Collect the cells by centrifugation. Remove the supernatant.
- Add 200μl of TE and vortex until cells are in solution. Transfer to 1.5ml microfuge tube.
- Add 200μl of lysis buffer (2% Triton X-100, 1% SDS, 100mM NaCl, 10mM Tris-Cl (pH = 8), 1mM EDTA).
- Add 200μl phenol:chloroform.
- Add 0.3g of acid-washed glass beads. Alternately, freeze-thaw 5 times.
- Vortex 3-5 minutes.
- Microfuge for 10 minutes and collect the aqueous layer, which contains the DNA.
- Subsequent purification by ethanol precipitation is strongly recommended.

Reference:


(b) Primers used to isolate *T. aqug* clamp and loader genes.

| β: DnaN Forward primer (NdeI) | 5'-CCCCTATGCGCTAACCCTTC-3' |
| β: DnaN Reverse primer (BamHI) | 5'-GGGATCTTATTACACCAGAGGGG-3' |
| δ: holA Forward primer (NdeI) | 5'-GGGACCATATGGTGCCTTCA-3' |
| δ: holA Reverse primer (BamHI) | 5'-TGGATCCCCATCTACCCGGGAGAC-3' |
| δ': holB Forward primer (NdeI) | 5'-GGTTTCCATATGCTTACCCC-3' |
| δ': holB Reverse primer (BamHI) | 5'-GAGGATCTACGACTCTAGCCAC-3' |
| γτ: dnaX Forward primer (NdeI) | 5’-AATCATATGCAGCCTCTCTACCCGACT3’ |
| γτ: dnaX Reverse primer (BamHI) | 5’-GGTTATGAGGATCTCAGCGGTCAGCCCTC3’ |

(c) Sequencing results of *Thermus aquaticus* strain YT-1 β clamp and clamp loader genes:

(i) Sequence of *Thermus aquaticus* strain YT-1 β clamp (dnaN gene) sequence.

Nucleotide sequence:

```
ATGCGCGTAACCCTTCCAAAAACCTGTGCGCCGAGGATCGCCCTC
CTGGAAACGGGTGCTACCCCCCTCTAGAAGCTCAGACCACCCCCCTTCTTCTAC
CTGGGCGCTGCGCCGGAAAGGACGCCTGTCCTTTCGCGACGCAAC
GGCGAGGTGCGGACCTGCGGCTCGTTCCGCGTGAGCAGCTGCGGCGAG
AGGGCGGTACCTGATCCCCCTCTAGCCCTTCTTCGCTGCTTGCAGCAAG
CTTCCCCCGCGAGACGGTGAGCTAGCTGCGGAAAGCGAGCTGGGCC
TGCGCTGGGGCGGCTCGGGCCCTACCCGAGGGGAGGTGGGGAAGGAGGCC
GCTACCCCGAGCTCCTCTTCACCCGGCCCTTGAGGAGGGAGGAGGCC
TCACCAGCGGCGACCAGTTCCTGCGCGAGGAGCTACACCGGCGCCCTT
CCCAGGTGCCTACCGCGCTAGCAGACGAGGATCAGCGGCGCCATCTCC
```
GGGGGGTGCAGCTGGAGTTCTCCGACCGGGCTTTGGCCGCCTGGCCT
CGGATGGCTACCGCCTGGCCCTCTTTGACCTGGAGAGGCCCCAGCCCT
TCACCAAGGAGTTCCCCCCTCAAAAGCGGCCCTTGAGCGTGGAACCCCTTC
CGGGAGGCCCCTGAAGCGGTCCTCGGTCTCGGCCGACAAGCAAGAACC
TCGGGGTGGACCTCTCTCTGGAGGAGGGGCGGTTCCTGCCTTCTCCGCGA
GGGGGACTACGGCAAGGGGCAGGAGGAGATTTCGTTTCCCTGGAGG
GGGCTCCTATGAGCCTGGCCTACAACGCCGCTACCTCCTCCTCGAGGCC
GGGGGCCTTTCGGCCTGCCGCTACCTCCTCGAGGCC
AGCTGGGCTTCTCGGGGCCCACCA
GCCCCACCCTGGTGCGGCCCTGGAGGAGGGGGGCTACCAGGCGGTG
GTGGTGCCCCTCCGGGCTTCTCGGGGCCCACCA
Protein sequence:
MRVTLPKNLLAERIALLERVIPSRSSNPLLTYLGLAAEKDLAVLFSGN
DLEVRLPAQLEGSEQYRLIPSQPPFQLVRSLPGETVELSLGSELALASGRFTT
RLALAPMEYGPELFPPDLEGGGEEAFQTQRAQFLAEELQRALSQVRAA
EYRAIFRQVQLFSDRGFRAVSDGYRLALFDLERPQPFKAKVVPARSV
DELVRVLRSAEVEVVLALGPGTLGLAVRQEAGALRMAVRLMEGEFPDYE
RVIPKEFPLKAALDVEPFREALKRVVLADKQNHRVDLFLEVELGRVLS
GDDYKGQEEIFVSLAGPMASYNARYLLEALAPLSGRAELGFSGPTSP
LVRPLEEGGQAVVVPLRV = 371 amino acids
No. of cysteine = 0; No. of Tryptophan = 0; No. of Tyrosine = 11
Theoretical pI/Mw: 4.97 (40.88kDa) as determined by Expasy program.

(http://web.expasy.org/compute_pi/)

(ii) Sequence of *Thermus aquaticus* strain YT-1 δ (holA gene) sequence.

Nucleotide sequence:

ATGGTGATCGCCTTCACCGGAGACCCCTTCTTGCCCAGGAGGCCCTTTTCTGGGCTACCCGTGACACCCACCAGGAGGCCCTGGCCGGAGGCCCTGGCCCCGGGGCTTTTCGGGGGCGGGGGGCCATGCTGGACCTGAGGGAGGTTTCCGAGGGGGAGTGGAAGGCCCTGAAGCCTATCCTGGAGGGCGTTCCCGAAGGCGTCCCCGTCCTCC

TCCTAGACCACCAGCCAAGCCCTTCCCCCCGGCCCGCTTCTTACCGTAGCCGGGAGCGGGGACTTACCCACCCCCAAGGGGAAGGACCTCGTGAGGACACCTTGAGAACCGGGCCAGGCGCCTGGGCCTCAGGCTCCCCTCCGGCG

GTGGCCCAGTACCTGCGGCGGCTCTGAGGCGGACCTTGGAGGCGGCTGGCCGCTGGAGATGGCGGTCCTAAGCCTGGCGGTTGACCGGCGGGTCAGCCCCGGGTAG

GC

Protein sequence:
MVIAFTGDFLAREALLEEALLLGLEEPTRFDPTPEALAEALAPGLFGGGG
AMLDLREVSEGGEWALKPILEGVPVEGPVLLLDPRPSRAPFSRER
RLPTPKGDKLVLHLENRRARGLRLPSGVAQYLAALADREREL
EKLLALLPPPPLTLREDVAMVVALKKPPITGFDLVRALAVLEKPNKEALARKRL
KEEGEPKLLLGAFAWQFGLSAYFLQQNPRPKEALARLGAHPYA
AKKALEEARRLSEEALRQALTTLIEAERRAKGKDPWLALEMAYVLSLA
VDRRVSPG = 299 amino acids
No. of cysteine = 0; No. of Tryptophan = 3; No. of Tyrosine = 4
Theoretical pI/Mw: 8.68 / 32925.45 (32.9kDa)

(iii) *Thermus aquaticus* strain YT-1 δ' (holB gene) sequence.

Nucleotide sequence:
ATGGCCTCTACGCCCCGCCTCACCCTTGGGGAATAATCGGCCACGAGGCC
ATTCTGGAGCTTCTGCCCAGGCTCAAGACCACCACCTCCTTCTCCTCGG
GGCCCGAGGGGGGTTGGCAGGCGGCTCAGCAGCCCTCTTGAGGCCACCCCGAC
GCCTCAACCGGGGCTTCCCCCGCCCCCTCCTTCTGGAGGCCACCCCGAC
CTCCTGGAGATCGGGCCCAAGGAGCAGGGGCTGAGGGGGGAGGCCGAG
GGTGCCGCTTCGAGGAGGCTGGGAGCCCTCTTTTGAACGTTTCCAGACCA
CCCGGGAGCGGGTGCAAGGTGGCCATCCTCAGCCCGCGGGCCACCTCCT
CAAGCAGGCGCGCCCCAAACGGCCTTCTTGGAAGCTCCTGAGGCCCTCC
CAGCTACGGGCGCATCAGTCGCCCCAGCGGGGAGACCCTCCT
CGAGGAAGGGCTGCGCCCTGACCAGGAACCAGGACCCTTCTCGCCT
ACGCCGCCGGGGCCCCAGGCCGCCCTCCTCAAGGCCCTTTCTGACCCCC
TGGCCTTCAGAAGGCTCAAGATGGCGAGGAGGGTTTTGCAAAGC
CCCCCTCTGAGCGCCTGGGCCTTTTGGAGGAGCTTTTGGCCGAGGAG
GAGGGGTTTTACCTCCTCCCGCCTCTCCTCCTAAAAACCGCCCCCCCGGCC
CTCCTGGGCCCTTGGATAGGGCCCGGGAGGCCCTGGAGGCCTATGTGAGC
CCCAACCTCGTGCCTGGCCCGCCTGGCCCTTAGACTTAGAGCTATGA

810 bases. The start codon has been changed from GTG (seen in strain Y51MC23) to ATG in my construct.

Protein sequence:
MALRPPHPWGIIGHEAILELLPRLKTATLLFSGPEGVGRRVARWYLGS
LNRGGPFPFLAHPDLEIGPKERGLRGEAEVRLEEEVEPLFDWFQHSPRE
RVKVALDAAHLLTEAAANALLKELLEEPSYGRILLAPIPSGETLLPLTLS
RALEVAFGPVPEERLRALTQDPDLLAYAAGAPGRRLKALSDPLAFQER
LKMARRVLQSPPLERLLGLEELLAEEEGFYLLRALLQNPRALALLDR
AREALEAYVSPNLVLARLALDLEL = 269 AMINO ACIDS
No. of cysteine = 0; No. of Tryptophan = 3; No. of Tyrosine = 5
Theoretical pI/Mw: 5.86/ 29914.97 (29.9kDa)

(iv) Sequence of Thermus aquaticus strain YT-1 Υτ (dnaX gene) sequence.

Nucleotide sequence:
ATGGGCGCCCTCTACCGGACTTTTCGCCCCCCTCACCCTCCAGGAGGTG
GTGGGCGCCAGGAGCACTGAAAGGAGGCCCTCTGCGGCATACGGGA
AGGGCGCTTGGCCCGACCCTACTCTTCCTTCGGCCCCCGGGGTGGG
CAAGACACCACCACCGCCCGCTCCTGGGCAATGGCCGTTGGGTGCCAGG
GGCCCGAGCGCCCTTGCGGGGTCTGCCCCCACTGCCAGGCGGTGCAGA
AGGGCGCCCACCCCCGACGTGTTGGAGATTGACGAGCCGCCGACGAAACG
TCCTACGCCCACAAAAGGTCTTCTCATCCCTGGAGAAGGCCCCACGCTGCT
CTCAAAAAGCGCCTTCACGCCCCTCCTCAAGGACGAGGAAGGCCACCCG
GCCCCACGCTCCTCCTCGTCCACCAACCGCAGGAGGAGTCCTTTTGGGCC
CCCAACCATCCTCTCCCGGCTCACGACTACCGCTTCCGCCGCCGCTTAG
GGAGGAGGAGATCAGCCTGAAGCTCTCGGCGATCTCTGGAGAAATGG
GTAGGGAGGCAGGAGGAGGCCCTTCCTCCTCATACTCCAGCTTTTCCG
ACGGGCGGATGCGGAGCAGGCCGAAAGCCTCTGGAGGCGCTTTTTGCTCC
TCGAGGCGCCCTCACCGGAAACCGAGGTGAGGAGGAGCTTTTG GG CTC
CCCAAAAGGAGGCCCTCGGCCGCGCTGGCCAGGGCCCTCGCCCTGGGAA
AGGTC
CCCGAGGCCTAAGGAGGCCAGGCGGCTTTACGGCCTGGGC
TTCGCCCCGAGGAGCCTGGTGCCCAGGAGGCTCATGGAGGACGCTCTCA
GGGA
GGGGCTTTACGCGCCTACGGCCTCGGTGGGCGCCAGGCTTTCCCGCCTG
GCCCGAGGAGCTTCCAGGCCCTCACGCGCCGTGGAGGCAAGGCCTGC
GCCTTGAGCCCGCGCTCGAGACCTTCTGGCCCTGGAGGCCCTCCTCAACCTCCTCAACCTGGCCCCCA
CCGGGCCCACCCCGAGCC
CGCCTCGGCCTAGGCCCGGTCCCCCGAGTTTGACCCCCACGGCCCGGC
CAGGCCCCAAGGGCGCCCCCGAGGCCGGGGAAGGCAAGGACGCTGCGG
AGGGGTGCGCGGGCTCTCGTGGAGCTTTGAAGGCCACCCCTTAAGGGCC
TCGTGCGGGAGGCCAGGCCCCACCTGGAGGGGAAGACCCTCGTCCTC
CGCTTCCCCGAGAGCAAGCCCTTCCACCAACAGAAGGCCCA
GGAGGCCACCTCCTCCCCCTGGGCGCCGGCCAGTTCGGGGTGGA
AGG
AGCTGGCCTTCGTCCTGGAAAAAAAAAGCTGAAGTGGGCTAGCCCC
CCACCCCCAACCAGCCGTTCCCCCAAGAGAAGCCCTCCCCCGGTG
GCGGCTCCCGGCCCGAACGGAGCCTCCCTCTGGAAGACCCCTCAGG
GAAGCGGGAGAGGAGGATTCTCTAG = 1.502 Kb

The start codon has been changed from GTG (seen in strain Y51MC23) to ATG in my construct.

Dark Blue = Shine-Dalgarano sequence [AGGAG] before frame-shift signal
Light Blue = Shine-Dalgarano sequence after frame-shift signal.
Green = Stop signal for τ protein; Purple = Stop signal for Υ protein.
Red = frame-shift site.

Protein sequence:
MGALYRTFRPLTFQEVVGQEHKEPLLRAIREGRLAQAYLFSGPRGVGK
TTTARLLAMAVGCQGPERPCGVCPHCQAVQKGAHPDVVEIDAASNSV
EDVRELRERILLAPLTKVILDEAHMLSKSAFNLKLEEEPVPHVLF
VFATTEPERMPPTILSVQHYRFRRRLREEEIALKLRLLEEMGREAEEALL
LIARLSDGAMRDAESLLERLLLLLEGPLTRTVQVAAALGLPPKEALRALAEA
LLLGKVPEALKKEARRLYGLGFAORSLVGGGLMEVLREGLYAAYGLGARL
PAGPEELLQALTRLDEALERLSRRSDLANLAALLNLNLAPTGPTETPPEP
ASASGPVFEPDPHAPARPAPPAPEGAEQDLAEGWRAFLALKPRLRAFVR
EARPHLEGKTLVLRFPESKAFHHKAAEKEQKAHLLPRLARAQFGVEELAFVL
EKKSLSGASPPPPTPKPVPPREAPPVAPPPEPPEPPEPPEPPEPPEPPEPPE

= 500 amino acids.

No. of cysteine = 4; No. of Tryptophan = 2; No. of Tyrosine = 6

Underline = ATP binding site. Red = putative Zn^{2+} finger near ATP binding site.
(iv) Primers used to sequence genomic DNA.

<table>
<thead>
<tr>
<th>Primer Set</th>
<th>Genomic Region</th>
<th>Forward Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T.aq</em> β: dnaN Genomic Seq_1</td>
<td>5'-GAGGAAGCAACTGGAAACGC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> β: dnaN Genomic Seq_2</td>
<td>5'-GCTCCTGGCCGGAC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> β: dnaN Genomic Seq_3</td>
<td>5'-GGCCCTTTATTACTACGAC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> δ: holA Genomic Seq_1</td>
<td>5'-CGCCCTCTAAAGCCCCCGCC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> δ: holA Genomic Seq_2</td>
<td>5'-GGGAAGAGACAGCTCGCC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> δ: holA Genomic Seq_3</td>
<td>5'-CCGCTTTCAAATACGCC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> δ': holB Genomic Seq_1</td>
<td>5'-CGTCCAGAGGCTCTA-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> δ': holB Genomic Seq_2</td>
<td>5'-GCGCATCATCTCATCGC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> δ': holB Genomic Seq_3</td>
<td>5'-CGGAGGACAGGCTCG-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> Υτ: dnaX Genomic Seq_1</td>
<td>5'-CCCAGCAGACGGCC-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> Υτ: dnaX Genomic Seq_2</td>
<td>5'-CAGGAGGAGGGCTCG-3'</td>
<td></td>
</tr>
<tr>
<td><em>T.aq</em> Υτ: dnaX Genomic Seq_3</td>
<td>5'-GGAAGCTCGCAACAGCCC-3'</td>
<td></td>
</tr>
<tr>
<td>T7 promotor sequence</td>
<td>5'-TAATACGACTCACTATAGGG-3'</td>
<td></td>
</tr>
<tr>
<td>T7 terminator sequence</td>
<td>5'-GCTAGTTATTGCTAGCGG-3'</td>
<td></td>
</tr>
</tbody>
</table>
(v) Difference between *Thermus aquaticus* strain YT-1 sequenced genes and PubMed deposited sequences:

The isolated genes from *Thermus aquaticus* strain YT-1 were compared to the Pubmed deposited sequence for *Thermus aquaticus* strain Y51MC23 and the discrepancies in the sequencing results for each gene were noted.

- **β clamp:** Pubmed has 3 deposited sequences for *Thermus aquaticus* strain Y51MC23 β clamp. Deposited sequence no. 2 was used as a reference for analysis. The comparison showed that in *Thermus aquaticus* strain YT-1 the β clamp has a substitution at amino acid position 126 from Aspartic acid to Glutamic acid.

<table>
<thead>
<tr>
<th>Amino acid no.</th>
<th>T.aq strain Y51MC23</th>
<th>T.aq strain YT-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>126</td>
<td>D</td>
<td>E</td>
</tr>
</tbody>
</table>

Reference gene: DNA polymerase III β clamp [*Thermus aquaticus* Y51MC23]

Accession: EED10041.1 GI: 218243512.

- **Clamp loader subunit δ:** The following amino acid substitutions were observed in the sequencing results of *Thermus aquaticus* strain Y51MC23 versus strain YT-1 for δ:

<table>
<thead>
<tr>
<th>Amino acid no.</th>
<th>T.aq strain Y51MC23</th>
<th>T.aq strain YT-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>101</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>162</td>
<td>E</td>
<td>A</td>
</tr>
</tbody>
</table>

- Clamp loader subunit δ': The following amino acid substitutions were observed in the sequencing results of *Thermus aquaticus* strain Y51MC23 versus strain YT-1 for δ':

<table>
<thead>
<tr>
<th>Amino acid no.</th>
<th><em>T.aq</em> strain Y51MC23</th>
<th><em>T.aq</em> strain YT-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>101</td>
<td>F</td>
<td>V</td>
</tr>
<tr>
<td>120</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>173</td>
<td>K</td>
<td>R</td>
</tr>
<tr>
<td>185</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>209</td>
<td>S</td>
<td>P</td>
</tr>
<tr>
<td>232</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>262</td>
<td>H</td>
<td>R</td>
</tr>
</tbody>
</table>


- Clamp loader subunit Υτ: The following amino acid substitutions were observed in the sequencing results of *Thermus aquaticus* strain Y51MC23 versus strain YT-1 for Υτ:

<table>
<thead>
<tr>
<th>Amino acid no.</th>
<th><em>T.aq</em> strain Y51MC23</th>
<th><em>T.aq</em> strain YT-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>Q</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>E</td>
<td>K</td>
</tr>
<tr>
<td>355</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>360</td>
<td>F</td>
<td>S</td>
</tr>
<tr>
<td>372</td>
<td>P</td>
<td>L</td>
</tr>
<tr>
<td>373</td>
<td>E</td>
<td>K</td>
</tr>
</tbody>
</table>

Appendix II:

(a) Purification protocol of *T.aq* β clamp

Day: 1

- Thaw biomass and add lysis buffer (30% of culture volume), DTT (5mM), PMSF (1mM) and PIC (2µl/ml) and Lyzosyme (3mg/Liter culture).
- Stir lysate on rotor setting#4 for 1 h in cold room and transfer lysate to a 50 ml falcon tube.
- Pass lysate through the Emulsiflex 3X to lyze the cells.
- Heat lysate 70°C for 10mins. Spin, F21S rotor, 17k, 1 h, 4°C.
- Decant supernatant into a graduated cylinder and bring up volume to 100ml with lysis buffer + 5mM DTT+ 1mM PMSF + 1µl/ml PIC.

Ammonium Sulfate Fractionation

0 – 40%

- Pour 100 ml supernatant in a 250 ml glass beaker set in an ice bath.
- Weigh out 22.6g (NH₄)₂SO₄ [40%] and break up any lumps. Add to the lysate in portions until completely dissolved and keep stirring slowly for a total of 20 mins. Transfer to 50 ml oakridge tubes and spin in F21S rotor at 17k for 30 min at 4°C. Decant supernatant to a 250 ml glass beaker set in an ice bath.

40 – 75%

- Weigh out 18.7g (NH₄)₂SO₄ [40%-70%] and add to the lysate in portions until completely dissolved and keep stirring slowly for a total of 30 mins. Transfer to 50 ml oakridge tubes and spin in F21S rotor at 17k for 1 hour at 4°C. Decant supernatant to a 250 ml glass beaker set in an ice bath.
• PCNA is fractionated in pellet in this step. Pour 5 ml AD₀ in each tube, resuspend pellets by pipeting up and down, and collect to a 50 ml falcon tube. Rinse all tubes with another 5 ml AD₀. Total volume is ~25 ml.

• Dialyze the crude protein in a dialysis tubing (MWCO 6 – 8000, 3.3 ml/cm) in a 1Lt cylinder containing 450 ml AD₀ in cold room O/N. After O/N dialysis, salt concentration is reduced to ~100 mM.

Day: 2

• Prepare DEAE column (binding capacity 8-10mg/ml) by flushing out EtOH with plenty of H₂O and equilibrate with AD₀+100mM NaCl for 20X column volume.

• Take out the crude protein from the dialysis bag and mix with an equal volume of AD₀. Load the diluted protein onto the column @ 0.5ml/min while stirring the load slowly at #7 to prevent flocculation during loading.

• Wash with AD₀+120mM NaCl for 1 Column Volume.

• Elute with 10X Column volume with the gradient AD₀+130mM-AD₀+260mM NaCl @ 0.5 ml/min. Collect fractions (1.4 ml or ~52 drops/tube). Run 12 % SDS gels to analyze the fractions. The peak should be at 180mM-220mM NaCl.

• Pool purest fractions containing PCNA, dialyze in a dialysis tubing (MWCO 6 – 8000, 3.3 ml/cm) in a 1 L cylinder containing 500 ml AD₀+100mM NaCl in cold room O/N.

Day: 3
• Take out the protein from the dialysis bag. Check protein concentration by standard Bradford and determine volume of resin accordingly.

• Binding capacity of Q column is 8-10mg/ml. Use 10X resin excess to amount of protein. Remove EtOH from resin by washing it with excess dH2O and equilibrate with AD₀+100mM NaCl. Load the protein onto a Q Sepahrose column @ 0.3 ml/min with load shaking on a rotator #7.

• Wash with AD₀+230mM NaCl for 2.5X column volume and elution is done with the gradient AD₀+230mM- AD₀+400mM NaCl @ 0.3 ml/min. Collect fractions (0.3 ml or ~13 drops/tube). Run 12 % SDS gels to analyze the fractions. The peak should be at 300mM-360mM NaCl.

• Pool purest fractions containing PCNA. For labeling with MDCC; Dialyze the pooled protein in a dialysis tubing (MWCO 6–8000, 3.3 ml/cm) in a 1 L cylinder containing 1 L Labeling Buffer in cold room O/N. Presoak Centriplus10 in LLB.

Buffers:

(1) Lysis Buffer: Tris HCl pH7.5=35mM, EDTA=1mM, Sucrose=10%w/v.

(2) AD₀ Tris Cl pH7.5 (20mM), EDTA (0.5mM), Glycerol [10% (v/v)], DTT (5mM)

(3) 2X AD₀ Tris Cl pH7.5 (40mM), EDTA (1mM), Glycerol [20% (v/v)], DTT (10mM)

(4) Labeling Buffer: Hepes NaCl pH 7.5 (30mM), NaCl (150mM) EDTA (0.5mM).
(b): Labeling of *T. aq* β clamp with MDCC

Day: 1

a. Pack P6DG gel filtration column.

1. Rehydrate 2.5 g of BioRad P6DG (1 g = 5 ml) in H2O for more than 4 hrs at room temp. Pour off the H2O and replace with labeling buffer (30 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM EDTA) @ 0.5 ml/min for more than 3 times.

2. Degas the resin to remove any air bubbles for 15-20 mins. Only swirl, no stir bar needed.

3. Pour the gel filtration column (50 cm x 1 cm) and equilibrate with about 40 ml labeling buffer.

b. Labeling reaction.

1. Concentrate the protein if necessary, to make the reaction volume to 500 µL.

2. Weigh MDCC (MW: 383.4) 5 mg in 521.7 µL of DMF to a final concentration of 25 mM. (ε of MDCC at 436 nm = 50,000 M⁻¹ cm⁻¹)

3. Make 100 mM TCEP (MW: 286.64) solution. (Weigh 2.87 mg in 100 µL of 0.2 mM TrisHCl pH 8)

4. Treat β clamp with TCEP to reduce formed disulfide bond. Add TCEP to β clamp for a final conc of 1 mM. Incubate at 37 C for 5 min.

5. Slowly add MDCC (1.5 µ moles) to treated β clamp solution. Incubate in the dark, wrapped with foil, on a roller drum for > 4 hrs (or O/N at 4°C).

Day: 2

c. Purification of labeled PCNA-MDCC.

1. Free dye removal via concentrator: Use the concentrator to remove the excess free dye the labeling reaction mixture. Add 2.5 mL labeling buffer to 0.5 mL
labeling reaction, and concentrate to ~ 0.5 mL (7.5 rmp for 7mins). Repeat for more than 8 times.

2. Run labeling buffer through the packed gel filtration column until the top layer of resin is barely dry by gravity flow.

3. Load labeled β clamp, let run through the column until the top layer of the resin is barely dry by gravity flow. Then add labeling buffer to the top for elution. (Keep a sample as “Load”)

4. Collect 1 mL fractions with gravity flow elution.

5. When finally see labeled β clamp in yellow color, try to collect in 1-2 fractions. The free dye is expected to elute later in a separate peak.

6. Verify the fractions on SDS-PAGE. Take a picture under UV or use typhoon scanner for an image of fluorescence. Stain the gel in coomassie blue later for an image of proteins.

7. Pool the fractions, and dialyze in dialysis buffer O/N. Aliquot the pool and store at -80C.

Dialysis buffer: Hepes pH 7.5 (30mM), EDTA (0.5 mM), NaCl (150 mM), DTT (2mM), Glycerol (10%).

8. Quantify the labeling efficiency by measuring the protein conc and the dye conc.
Appendix III:

(a) *T. aq* δ subunit protein purification protocol.

Day: 1

*Preparation of Lysate:* 200ml biomass is resuspended in 100ml 1X Lysis Buffer (50% dilution). Add lysozyme (3mg/Lt), PIC, and PMSF (1mM). Mix the lysate on a rotator for 30 mins. Pour lysate into thin 15ml falcon tubes and freeze in liquid N2 for 2 mins and thaw at 37°C for 10 mins 3X. Dounce for ~80X and centrifuge the lysate 23,000g (14,000 rpm in SS34 rotor) for 1 hour 4°C. Collect the lysate in 50ml falcon tube and store 20ul sample of the lysate and pellet to be run on the gel.

*Preparation of the column:* Binding capacity of Ni Sepharose fast flow column = 40mg of His tagged protein/ml of resin. Estimate yield of protein from the solubility test and prepare resin amount accordingly. The NiSFF resin is exchanged in water and equilibrated in wash buffer.

*Running the column:*

- Load the lysate at 0.6ml/min and collect the flow through.
- Wash the column with the Wash buffer (120mM Imidiazole) for 20 times column volume at 1ml/min. Wash the column 1 times column volume with low salt buffer (150mM Imidiazole) 0.5ml/min.
- Elution gradient from 150mM-500mM Imidiazole at 0.5ml/min. Total volume of gradient = 10 times column volume (0.5*CV of low salt and 0.5*CV of high salt).
- Collect 20ul samples of flow through, wash, and elution and add 5ul of 5X Loading dye to it. Heat the samples and resolve it on 12% SDS PAGE gel.
• Determine the peak of the gradient with the least contaminating bands.

Pool and dialyze the protein in Tris HCl pH7.5 (50mM), NaCl (40mM), Glycerol (5%).

*Buffer preparation:*

(1X) Lysis Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock conc</th>
<th>Final conc</th>
<th>I add</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidazole</td>
<td>2M</td>
<td>5mM</td>
<td>0.5ml</td>
<td>0.025ml (25ul)</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>25ml</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>4ml</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td></td>
<td>170.5ml</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>200ml</td>
</tr>
</tbody>
</table>

Wash Buffer 1

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidazole</td>
<td>2M</td>
<td>120mM</td>
<td>0.6ml (300ul)</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>8.25ml</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>
Wash Buffer 2 / Low Imidiazole Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidiazole</td>
<td>2M</td>
<td>150mM</td>
<td>0.75ml (750ul)</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>7.8ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>

Elution Buffer / High Imidiazole Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidiazole</td>
<td>2M</td>
<td>500mM</td>
<td>2.5ml</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>6.05ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>

Day: 2

- The calculated pI of *T.aq* holA is 8.6 therefore a cationic column (SP sepharose) can be used to purify holA. The binding capacity of S Sepharose column is 120mg/ml. Prepare resin by washing out EtOH out with excess dH2O and equilibrate with Tris HCl pH 7.5 (20mM), NaCl (40mM) for 10X column volume.
- Load sample on the SP sepharose column at the rate of 0.5ml/min.
• Wash the column with low salt buffer (150mM NaCl) at 0.5ml/min for 20 times column volume.
• Elute the protein with a gradient (150mM – 500mM NaCl) at 0.5ml/min. Peak is seen at 245mM-335mM NaCl.
• Collect 20ul samples of flow through, wash, and elution and add 5ul of 5X Loading dye to it.
• Heat the samples and resolve it on 12% SDS PAGE gel.
• Determine the peak of the gradient with the least contaminating bands. If the protein is still not clean ammonium sulphate precipitation could be tried.
(McHenry protocol suggested 40%) Dialyze the protein in NaCl = 150mM, Glycerol 5%, Tris pH 7.5 = 20mM
• Store an aliquot to determine protein concentration. Freeze in liquid nitrogen and store in -80C.

Low Salt Buffer:

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>150mM</td>
<td>0.375ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>200ul</td>
</tr>
<tr>
<td>Glycerol</td>
<td>100%</td>
<td>5%</td>
<td>0.5ml</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>8.925ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>
Elution Buffer / High Salt Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>200ul</td>
</tr>
<tr>
<td>Glycerol</td>
<td>100%</td>
<td>5%</td>
<td>0.5ml</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>7.8ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>

(b) \textit{T}\textsubscript{aq} $\delta'$ subunit protein purification protocol.

Day: 1

\textit{Preparation of Lysate and Ni sepharose column}: Is same as that for \textit{T}\textsubscript{aq} $\delta$.

\textit{Running the column}:

• Load the lysate at 0.6ml/min and collect the flow through.
• Wash the column with the Wash buffer (180mM Imidiazole) for 20 times column volume at 1ml/min. Wash the column 1 times column volume with low salt buffer (150mM Imidiazole) 0.5ml/min.
• Elution gradient from 180mM-350mM Imidiazole at 0.5ml/min. Total volume of gradient = 10 times column volume (0.5*CV of low salt and 0.5*CV of high salt).
• Collect 20ul samples of flow through, wash, and elution and add 5ul of 5X Loading dye to it. Heat the samples and resolve it on 12% SDS PAGE gel.
• Determine the peak of the gradient with the least contaminating bands.

Pool and dialyze the protein in Tris HCl pH 7.5 (50mM), NaCl (40mM), Glycerol (5%).
Buffer preparation: (1X) Lysis Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock conc</th>
<th>Final conc</th>
<th>I add</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidiazole</td>
<td>2M</td>
<td>5mM</td>
<td>0.5ml</td>
<td>0.025ml (25ul)</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>25ml</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>4ml</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td></td>
<td>170.5ml 8.525ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>200ml 10ml</td>
</tr>
</tbody>
</table>

Wash Buffer 1

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidiazole</td>
<td>2M</td>
<td>180mM</td>
<td>0.9ml</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>0.2ml</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>7.65ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>

Wash Buffer 2 / Low Imidiazole Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidiazole</td>
<td>2M</td>
<td>350mM</td>
<td>0.75ml (750ul)</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>7.8ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>
Elution Buffer / High Imidazole Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Final</th>
<th>I add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidazole</td>
<td>2M</td>
<td>500mM</td>
<td>1.75ml</td>
</tr>
<tr>
<td>NaCl</td>
<td>4M</td>
<td>500mM</td>
<td>1.25ml</td>
</tr>
<tr>
<td>Tris HCl pH 7.5</td>
<td>1M</td>
<td>20mM</td>
<td>0.2ml (200ul)</td>
</tr>
<tr>
<td>dH2O</td>
<td></td>
<td></td>
<td>6.8ml</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>10ml</td>
</tr>
</tbody>
</table>

\((c)\) \textit{T.aq} \textit{Y}τ subunit protein purification protocol.

- Thaw biomass in a falcon tube by placing it in a beaker of water and add 45 ml binding buffer.
- Pack 5ml Nickel column. (Remove EtOH, pack, wash with 15ml water then equilibrate in binding buffer. Wash the pump and fraction collector first with water and then with 10mM Imidazole buffer.
- Prepare PMSF by adding 2 granules of Phenylmethylsulphonyl fluoride via a 100µl pipette tip to 10ml of isopropanal and shake well. Add 300µl of PMSF to the 60ml cell lysate solution (16ml biomass + 45ml binding buffer).
- Add 0.4g/ml of lysozyme to the cell lysate. Place stir bar and stir for 45mins at enough speed that the cells are mixed well without frothing.
- Divide the 60ml lysate equally into two 50ml falcon tubes. Lyse the cells by freezing them in liquid N2 for 1 minute followed by thawing at 37°C for 8-10 minutes and repeat 3X.
• Spin the cells at 16,000 rpm at 4°C for 1 hour. Carefully pour out the supernatant without disturbing the cell pellet.

• *Ammonium Sulphate Precipitation*: To get a final of 40% saturation, add 11.3g NH$_2$SO$_4$ to the cell lysate. Put a stir bar in the beaker and add the salt occasionally over 0.5hr. Ensure the cells are mixed well but there is no frothing observed. When done adding the salt, stir for an additional 0.5hr. Spin the cells at 14,000 rpm for 0.5 hr.

• Save the supernatant. Resuspend the pellet in 10mM Imidiazole (30ml total lysate sample to be loaded).

• Load the 30ml sample on the Nickel column@0.5ml/min. Then wash the column with 20ml of 20mM Imidiazole @0.5ml/min. Collect the two flow through separately.

• Prepare and start the gradient (20mM-200mM Imidiazole). Collect 0.8ml fractions. Prepare and run the fractions on a 10% SDS gel to detect $\Upsilon\tau$ subunits.
Appendix IV:

(a) Multiple sequence alignment of β clamp

Program: CLUSTAL Omega (1.2.0) multiple sequence alignment

Blue = residues composing the hydrophobic patch in E.coli β clamp.

Green = residues composing the hydrophobic patch in T.aq β clamp.

Red = E.coli β clamp mutants tested for MutS binding.

Purple = D304 in T.aq β clamp is analogous to Q299 in E.coli β clamp.

(b) Primers used to make β clamp mutants (5’ → 3’)

(i) T.aq β clamp G178C

Forward: CGGGCCGTGCCCTCGGATTGCTACCGCCTGGCCCTCTTTG
Reverse: CAAAGAGGGCCAGGCGGTAGCAATCCGAGGCCACGGCCCG

(ii) T.aq β clamp D282C

Forward: CGGGTCTCGGTCCTG
Reverse: CCACCCGATGGTTCTGCTTGCAGGCCAGGACCG

109
(c) Multiple sequence alignment of MutS

Program: CLUSTAL Omega (1.2.0) multiple sequence alignment

Bold = conserved β clamp binding region identified in the N and C terminal regions of MutS.

Red = alanine scanning of N terminal peptide residues that showed severe reduction in MutS-β clamp binding in *E.coli* MutS.

Green = alanine scanning of C terminal conserved residues that showed severe reduction in MutS-β clamp binding in *E.coli* MutS.

Highlight=peptide derived from *B.subtilis* MutS that interacts with β

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taq</td>
<td>MEGMLKGEQPGPLPPLLQYVEILRDQYFDDLYLLFQVQVYECFGDEAELARALGLVLTH</td>
</tr>
<tr>
<td>Bsubtilis</td>
<td>--------</td>
</tr>
<tr>
<td>Ecoli</td>
<td>--------</td>
</tr>
<tr>
<td>Paeruginosa</td>
<td>--------</td>
</tr>
<tr>
<td>Taq</td>
<td>KT-SKDFTTMAGIPLRAFEAYARLKLVMGFLAVQVPEPAEEAGLVRREVQTLT</td>
</tr>
<tr>
<td>Bsubtilis</td>
<td>--------</td>
</tr>
<tr>
<td>Ecoli</td>
<td>--------</td>
</tr>
<tr>
<td>Paeruginosa</td>
<td>--------</td>
</tr>
<tr>
<td>Taq</td>
<td>VLLAPELLENAGAFD----ERKRFPMULSPE---------------FEPGEGEGPLA---</td>
</tr>
<tr>
<td>Bsubtilis</td>
<td>--------</td>
</tr>
<tr>
<td>Ecoli</td>
<td>--------</td>
</tr>
<tr>
<td>Paeruginosa</td>
<td>--------</td>
</tr>
<tr>
<td>Taq</td>
<td>RRAARGALLAYAQRTQGGAL-SLQPFRFYPDGAFLMRPEATLRALEVFPELR---QDTLF</td>
</tr>
<tr>
<td>Bsubtilis</td>
<td>--------</td>
</tr>
<tr>
<td>Ecoli</td>
<td>--------</td>
</tr>
<tr>
<td>Paeruginosa</td>
<td>--------</td>
</tr>
<tr>
<td>Taq</td>
<td>SVLDETRATAPRPILSSWLRHPDLLGRLPRLAIRLDRVEGFVREGARQRVRLRRYLADLE</td>
</tr>
<tr>
<td>Bsubtilis</td>
<td>--------</td>
</tr>
<tr>
<td>Ecoli</td>
<td>--------</td>
</tr>
<tr>
<td>Paeruginosa</td>
<td>--------</td>
</tr>
</tbody>
</table>

110
(d) Primers used to make *T. aq* MutS C42A/A764C mutant (5’ → 3’)

Forward: CTCTCCAGGCCCATGCGCTGCAGCAGGGGAAGGGCCTTTTG

Reverse: CCAAGGCCCCCTCCCGCGCTGCAGCCATGGCCTGGAGGAGG
Appendix V:

III) Construction of bi-forked DNA substrate

**Step 1:** 37bp ssDNA strand containing the core DNA sequence will be annealed to a 90bp ssDNA strand containing the complementary DNA sequence. The mixture will be resolved on native PAGE and the product will be recovered by gel extraction.

**Step 2:** The product from step 1 will be annealed with a 27mer ssDNA. 17 of the 24 bases will be complementary to the 5’ overhang of the product obtained from step 1. The annealed product will generate a 3’ primer template region that allows for clamp loading and a 17 base fork region that prevents the β clamp sliding off the DNA.

**Step 3:** The product obtained from step 2 will be annealed with a 26mer ssDNA. 16 bases of the 26 bases will be complementary to the 3’ overhang of the product obtained from step 2. The annealed product will generate a 3 fork region that prevents the β clamp sliding off the DNA. The biotin tag can be coupled with streptavidin beads to isolate β clamp-DNA complex.
(VIII) Reference:


(34) Monti MR, Miguel V, Borgogno MV, Aragarana CE. Functional analysis of the interaction between the mismatch repair protein MutS and the repication
processivity factor β clamp in Pseudomonas aeruginosa. DNA Repair 2012 May 1;11(5):463-9.


(42) Shell SS, Putnam CD, Kolodner RD. The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to PCNA. Mol Cell, 2007 May 26;26(4):565-78.


(48) Pillon MC, Miller JH, Guarne A. The endonuclease domain of MutL interacts wit the beta sliding clamp. DNA Repair. 2011 Jan 2;10(1):87-93.
(49) Fukui K, Nishida M, Nakagawa N, Masui R, Kuramitsu S. Bound nuclotide controls the endonuclease activity of mismatch repair enzyme MutL. J Biol Chem. 2008 May 2;283(18):12136-45.


(59) Antony E, Hingorani MM. Asymmetric ATP binding and hydrolysis activity of the Thermus aquaticus MutS dimer is key to modulation of its interactions with mismatch DNA. Biochemistry 2004 Oct 19;43(41):13115-28.


(71)  Mark Hedglin, Ravindra Kumar and Stephen J. Benkovic. Replication clamps and clamp loaders. CSHL 2013. doi: 10.1101/cshperspect.a010165


