ASSESSING THE FUNCTIONAL ASYMMETRY OF THE BACILLUS SUBTILIS

MUTL HOMODIMER

ASSESSING THE FUNCTIONAL ASYMMETRY OF THE BACILLUS SUBTILIS

MUTL HOMODIMER

By LINDA LIU, B.Sc. (Honours)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the

Requirements for the Degree of Master of Science

McMaster University © Copyright by Linda Liu, July 2017

McMaster University MASTER OF SCIENCE (2017) Hamilton, Ontario (Biochemistry)

TITLE: Assessing the functional asymmetry of the *Bacillus subtilis* MutL homodimer

AUTHOR: Linda Liu, B.Sc. (Queen's University)

SUPERVISOR: Dr. Alba Guarné

NUMBER OF PAGES: xii, 74

ABSTRACT

DNA mismatch repair corrects base-base mismatches and small insertion/deletion loops generated during normal DNA replication. If left unrepaired, these errors become permanent mutations and can lead to increased susceptibility to cancer. In most prokaryotes and all eukaryotes, the mismatch repair protein MutL is a sequence-unspecific endonuclease that plays an essential role in the strand discrimination step of this pathway. Prokaryotic MutL forms homodimers with two endonuclease sites, whereas eukaryotic MutL homologs form heterodimers with a single active site. To elucidate the mechanistic differences between prokaryotic and eukaryotic MutL, we tested whether both endonuclease sites are necessary for prokaryotic MutL nicking activity. MutL interaction with the processivity clamp is required to stimulate endonuclease activity. Therefore, we also tested whether both subunits of the MutL dimer needed to interact with the processivity clamp. To this end, we engineered a system to independently manipulate each protomer of the homodimer. We demonstrated that prokaryotic MutL is regulated by the processivity clamp to act in a similar manner to eukaryotic MutL with only one functional site contributing to the endonuclease activity. We also devised a strategy to stabilize the transient interactions between MutL, the β-clamp, and DNA through disulfide bridge crosslinking and heterobifunctional crosslinking. Stabilizing transient protein-protein and protein-DNA interactions will help optimize future structural studies in obtaining the ternary complex for mechanistic insights to the MutL endonuclease activity and regulation imposed by the β-clamp.

ACKNOWLEDGEMENTS

I would first like to thank my family for their unwavering support. To my brother, who taught me that that the only person standing in your way is yourself and to never shy away from the unknown. To my parents, thank you for always encouraging me to pursue my dreams and frequently made calls to make sure I was well fed and sleeping properly.

I would like to thank my committee members, Dr. Yingfu Li and Dr. Giuseppe Melacini, for taking the time out of their busy schedules to give me helpful advice and guidance. I would like to thank the members of the Guarné lab, past and present. I would have never made it past some hurdles without their meaningful discussions and feedback on my work. I will thoroughly miss those late nights in the lab, full of crazy antics, much laughter and music blasting on the side. Finally, I would very much like to thank my supervisor, Dr. Alba Guarné who has always been a source of motivation and encouragement. You are an awe-inspiring woman in science with unwavering confidence and I am fortunate to have been able to learn from you for the past two years.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

CHAPTER 2 MATERIALS AND METHODS

CHAPTER 3 RESULTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

α	Alpha
ATP	Adenosine triphosphate
bp	Base pair
β	Beta
BSA	Bovine serum albumin
^{o}C	Degrees celsius
CMCB	Centre for Microbial Chemical Biology
CTD	C-terminal domain
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside-triphosphate
DtxR	Diphtheria toxin repressor
DTT	Dithiothreitol
DLS	Dynamic light scattering
EDTA	Ethylenediaminetetraacetic acid
γ	Gamma
GHKL	Gyrase, Hsp90, Histidine Kinase, MutL ATPase family
HNPCC	Hereditary Non-polyposis Colorectal Cancer
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
h	Hour
IDLs	Insertion/deletion loops
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	Kilo-Dalton
β-Me	β-Mercaptoethanol
MLH1	MutL homolog 1
MLH3	MutL homolog 3
min	Minutes
MOBIX	McMaster Institute for Molecular Biology and Biotechnology
MW	Molecular weight
M	Molecular weight marker
MSH2	MutS protein homolog 2
MSH3	MutS protein homolog 3
MSH6	MutS protein homolog 6
NTD	N-terminal domain
OD_{600}	Optical density measured at 600 nm
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PMS2	Postmeiotic segregation increased 2
PMSF	Phenylmethylsulfonyl fluoride
pol	Polymerase

RNA	Ribonucleic acid
SMCC	Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate
SSB	Single strand binding protein
SAXS	Small angle X-ray scattering
s	Seconds
SDS	Sodium dodecyl sulfate
PAGE	Polyacrylamide gel electrophoresis
UV	Ultraviolet
v/v	Volume per volume

DECLARATION OF ACADEMIC ACHIEVEMENT

MutL chimeras with different linker lengths were cloned, over-expressed, and initially purified by Julia Cai. In collaboration with Mary Carmen Ortiz Castro, we optimized the expression and purification of the *Bacillus subtilis* MutL C-terminal domain heterodimer. I performed endonuclease assays using fluorescently labeled substrates to observe the necessity of each active site and β-clamp binding motif on the nicking activity of the chimera heterodimer. I formed the β-CTD complex through disulfide crosslinking according to the crosslinking protocol from Pillon et al., (2015). I determined the same disulfide crosslinking strategy would not work for the β-DNA complex formation and devised a strategy of complex formation through heterobifunctional crosslinking.

CHAPTER 1

INTRODUCTION

1.1) DNA replication

The ability to duplicate DNA and pass down genetic information is essential for the proliferation of life. Cells must ensure the entire genome is replicated with high fidelity and within the time constraints of the cell cycle. The challenge of replicating each DNA template strand efficiently and accurately falls on a single multi-protein complex called the replisome that is fundamentally conserved across bacteria, archaea, and eukaryotes (Table 1.1) (Yao and O'Donnell, 2010). Proteins assembled at the replication origin to form the replisome encompass a wide range of functions including DNA unwinding and synthesis functions as well as processivity, protective, and scaffolding factors. The tightly coordinated enzymatic activities of the replisome allow for semiconservative DNA replication that results in two identical daughter copies, each composed of a parental template strand and an antiparallel daughter strand (Meselson and Stahl, 1958).

Table 1.1 Components of the replisome in different organisms

Component	Bacteria (*Escherichia coli*)	Eukaryote (Human)
Helicase	DnaB	MCM2-7
Clamp loader	γ/τ complex	RFC
Processivity/sliding clamp	β clamp	PCNA
Replicating polymerase	Pol III	pol ε (on leading strand) and pol δ (on lagging strand)
Primase	DnaG	pol α
Single strand binding protein	SSB	RPA
Other	Unknown	GINS, Cdc45

Beginning at the replication origin, helicase is loaded on by initiation factors. Activation of the helicase unwinds duplex DNA and form a double-strand/single-strand junction that keeps the two strands separated, establishing the replication fork. The exposed single stranded DNA is stabilized by single-strand binding (SSB) protein to remove secondary structures which may impede replication and protect the strands from damage (Wold and Kelly, 1988). Before elongation of the nascent strand can begin, primase must first synthesize a short RNA (bacterial) or RNA-DNA hybrid (eukaryotic) primer to serve as a starting point for DNA synthesis. The primer provides a 3'-hydroxyl for polymerase to incorporate free deoxyribonucleotide triphosphates (dNTPs) to the growing strand (Hubscher et al., 2002).

DNA polymerase uses the parent strand as a template to synthesize the complementary daughter strand in a 5' to 3' direction. Due to the antiparallel nature of DNA, the replication machinery uses different mechanisms to simultaneously synthesize the leading and lagging strands. Recent studies have shown that the leading and lagging polymerases are not coordinated together and function autonomously to replicate their respective strands (Graham et al., 2017). DNA synthesis occurs in irregular bursts and pauses which are independent of each strand but on average the leading strand and lagging strand synthesis proceeds at similar rates to complete replication at the same time (Graham et al., 2017).

Replication of the leading strand is coordinated by the clamp loader, processivity clamp, and DNA polymerase. The ring-shaped processivity clamp is loaded onto DNA at the primer-template junction by the clamp loader which is attached to the moving helicase

(Kelch et al., 2012). The processivity clamp creates a topological link with the DNA template and tethers the replicative polymerase to confer high processivity during DNA elongation (Prelich et al., 1987). The replicative polymerase moves in the direction of DNA unwinding and synthesizes the nascent strand in a continuous manner.

The lagging strand is synthesized discontinuously in stretches of DNA called Okazaki fragments, ranging in length from $1 - 2$ kilobases in prokaryotes or $100 - 200$ bases in eukaryotes (Okazaki et al., 1968). Primase must first catalyze the synthesis of a primer at the 5' end of each Okazaki fragment to serve as a platform for polymerase to bind onto. Once the clamp loader loads the processivity clamp onto the primer, the replicative polymerase can associate with the clamp to begin synthesizing the fragment (Tsurimoto and Stillman, 1991). Each time an Okazaki fragment is complete, the polymerase releases the clamp and subsequently moves upstream to associate with a new clamp on the next primer to begin synthesizing another fragment (Stukenberg et al., 1994). This process continues until the entire lagging strand is copied. Okazaki fragment maturation occurs when the RNA primer is replaced with DNA and the gaps between Okazaki fragments are sealed by a ligase to form a continuous DNA chain (Beattie and Bell, 2011; Waga et al., 1994).

The core of *Escherichia coli (E.coli)* Pol III, the major replicating DNA polymerase (Pol), is composed of three subunits: α, ε, and θ. The α subunit is responsible for the 5' polymerase activity, the ε subunit is responsible for the 3' exonuclease proofreading activity, and the θ subunit helps stabilize the ε subunit (Fijalkowska et al, 2012). By itself, Pol III replicates DNA slowly (~20 nucleotides per second) and is only

weakly processive (<10 bases per binding event) (Maki and Kornberg, 1985). Coupled to the processivity clamp, Pol III drastically increases its catalytic rate (~750 nucleotides per second) and is highly processive (>50 kilobases per binding event) (Stukenburg et al., 1991).

The processivity clamp is best understood for its role in DNA replication. It forms a topological link to DNA through its central cavity, allowing for non-sequence specific enzymes such as DNA polymerase, endonucleases, ligases, and helicases to associate with their DNA substrate. (Georgescu, 2008). Once the processivity clamp is loaded onto DNA, it can freely slide along the duplex in an ATP-independent manner, hence it is also known as the sliding clamp (Stukenberg et al., 1991). In prokaryotes, the sliding clamp is the β subunit of DNA polymerase III holoenzyme (β-clamp), and its eukaryotic homolog is the proliferating cell nuclear antigen (PCNA). Despite no obvious sequence homology between the prokaryotic β-clamp and eukaryotic PCNA, structurally they are almost superimposable (Gulbis et al., 1996). These structurally conserved ring shaped clamps are arranged in a head-to-tail manner, creating two distinct faces with different electrostatic properties (Krishna et al., 1994). Most clamp-binding partners interact with the clamp via the hydrophobic pockets located on the C-terminus. In prokaryotes, the β-clamp binding motif (QLxLF) is poorly conserved and absent in numerous β-interacting proteins (Dalrymple et al., 2001). On the other hand, the eukaryotic PCNA binding motif (QxxLxxFF) is strictly defined and present in most PCNA-binding proteins (Moldovan et al., 2007). Access to the clamp provides high processivity and faithful access to newly synthesized DNA for a variety of cellular processes. Besides playing a critical role in

recruitment and association of replicative polymerase during DNA replication, the processivity clamp is also crucial in other cellular functions such as translesion DNA synthesis, DNA repair, and DNA transposition (Gomez et al., 2014; Parks et al., 2008; Moldovan et al., 2007; Lopez de Saro and O'Donnell, 2001).

1.2) Mutations naturally occur during DNA replication

DNA is subjected to change from the constant bombardment of exogenous (environmental sources) and endogenous (cellular metabolic processes) factors. This includes formation of bulky adducts or pyrimidine dimers induced by UV light and 8-oxoguanine lesions or strand breaks caused by reactive oxygen species (Yoon et al., 2000; Loon et al., 2010). Although repair mechanisms quickly target these lesions for repair, DNA lesions which persist may remain during replication. The high fidelity required for DNA replication renders the replication polymerase unable to process the structurally altered template DNA, resulting in fork stalling. The processivity clamp orchestrates polymerase switching by replacing the replicative polymerase with a specialized polymerase used to traverse the damage before the replicative DNA polymerase resumes its function (Heltzel et al., 2012). Translesion synthesis polymerases have low fidelity and lacks the 3' exonuclease proofreading activity (Heltzel et al., 2012; Lehmann, 2003). Although this mechanism is used to avoid deleterious replication fork collapse, it is a process that is intrinsically error-prone and a source of DNA damage.

Despite the high accuracy of the DNA replication mechanism, mutations also arise because of polymerase misincorporating nucleotides at an error rate of $10^{-4} - 10^{-5}$

(Kunkel, 2004). The first line of defense in correcting these mistakes is polymerase itself through its intrinsic proofreading exonuclease (McCulloch and Kunkel, 2008). Polymerase performs its proofreading function when the incorporated mismatch causes difficulty for strand extension and allows for the primer terminus to separate from the template and move into the 3' exonuclease active site for removal (Reha-Krantz, 2010). Depending on the sequence context and the polymerase, this can improve replication fidelity up to 100-fold (Reha-Krantz, 2010). Regions of DNA with small repetitive sequences are prone to insertion/deletion loops (IDLs) resulting from daughter or template strand slippage. These may be harder to identify and remove through polymerase proofreading as the end of the primer terminus is properly annealed and prepped for elongation.

Unlike DNA damage which offer a distinct chemically modified structure, mismatches offer a unique form of DNA damage as they consist entirely of undamaged bases that only transiently remains mismatched while the two strands of the duplex are annealed (Jiricny, 2013). This situation is dangerous because uncorrected DNA errors may be used as templates for subsequent replication events, permanently engraving the error into the cell's genetic identity. As errors only arise in the newly synthesized strand, repair must quickly occur before the replication process is complete to distinguish between the template and nascent strand. This implies that DNA repair mechanisms must be closely coupled with DNA replication. Indeed, several proteins shared between the two processes such as the processivity clamp have been implicated in coordinating several DNA repair factors in response to DNA damage (Kunkel and Erie, 2015).

1.3) Replication errors lead to cancer

If replicative errors are not corrected by the intrinsic polymerase proofreading activity, the second line of defense in correcting replicative errors is by the highly conserved mismatch repair (MMR) mechanism (Figure 1.1). MMR plays a major role in maintaining genome stability through correcting base-base mismatches and small IDLs generated by DNA polymerase during replication (Kunkel and Erie, 2005). This process improves replication fidelity by 50 – 1000 fold (Iyer et al., 2006).
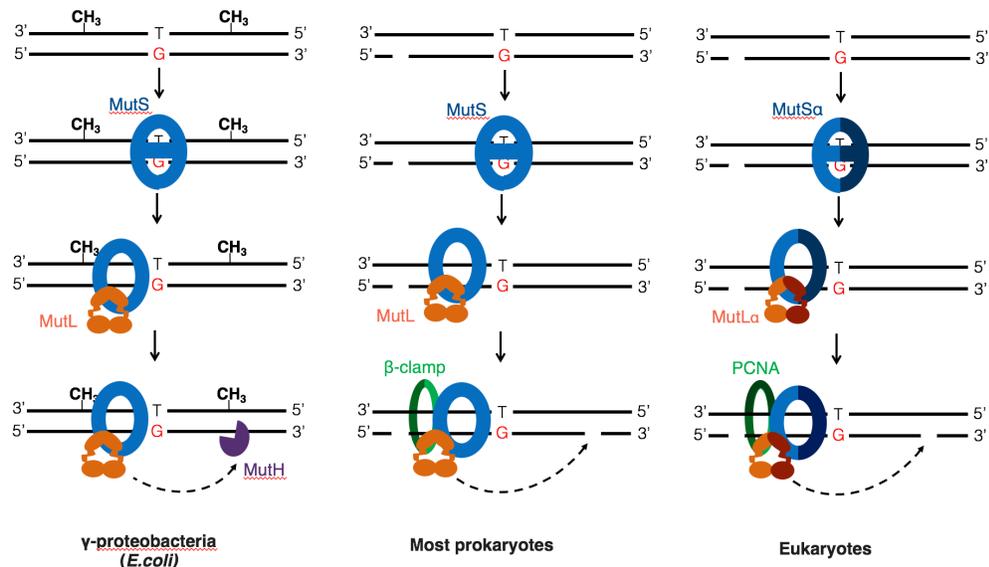
Figure 1.1 Early steps of mismatch repair in different organisms. Mechanism of mismatch recognition, strand discrimination, and nascent strand cleavage. Methyl-directed MMR occurs in a subset of γ-proteobacteria such as *E. coli*. Nick directed MMR in most prokaryotes and all eukaryotes. The mismatch is shown in red.

The necessity of human MMR for genomic stability was highlighted as impaired MMR gene function and expression dramatically increases the susceptibility to tumor development. Lynch syndrome, also known as hereditary nonpolyposis colon cancer (HNPCC), is a hereditary cancer syndrome with genetic defects in one or more MMR

genes (Lynch et al., 2009). It is characterized by frequent mutations in simple repetitive sequences known as microsatellite instability (Thibodeau et al, 1993). Patients with Lynch syndrome have an 80% chance of developing colorectal cancer, in comparison the risk of developing colorectal cancer in the general population is only 2% (Dinh et al., 2011). Women with Lynch syndrome also have a 71% chance of developing endometrial cancer, while the general population only has a risk of 1.5% (Dinh et al., 2011). As well, there is an increased susceptibility to gastric cancer, ovarian cancer, small bowel cancer, urinary tract cancer, and brain cancer (Vasen et al., 2007). MMR defects have also been shown to cause 10-15% of sporadic cases of colon cancer (Peltomäki, 2001).

Determining if a patient is affected with Lynch syndrome is important so that they and their relatives can take appropriate preventative measures to improve the health outcomes among carriers. With the information obtained from genetic screening, further development in understanding the mechanistic information of the MMR pathway and the function of the specific genes altered in cases of Lynch syndrome and other cancers is critical in developing methods of diagnosis, treatment, and prevention (Peltomäki, 2003)

1.4) Mechanism of DNA mismatch repair

Much of our understanding about the highly conserved MMR mechanism first arose from the well characterized reconstituted system in *E.coli*. MMR can be broken into three distinct steps: the first is recognition and binding to the mismatch, followed by strand removal of the segment with the error and finally error-free re-synthesis of the strand and ligation completes repair. In *E.coli*, initiation of this mechanism requires three

proteins: MutS, MutL and MutH. First, MutS recognizes the mismatch or IDLs (Modrich and Lahue, 1996). It then recruits the molecular matchmaker MutL to subsequently interact with downstream repair proteins (Sancar and Hearst, 1993). MutL then recruits and activates the sequence and methylation specific endonuclease MutH. Immediately after synthesis, deoxyadenine methylase has yet to methylate the adenine at guanine-adenine-thymine-cytosine (GATC) sites on the newly replicated strand hence the duplex DNA is in a transient hemi-methylated state (Barras and Marinus, 1989). The hemimethylated DNA acts as a strand discrimination signal to differentiate the parental and nascent strands for MutH which only nicks DNA 5' of the dG of unmethylated d(GATC) sites (Geier and Modrich, 1979). The gap created by MutH acts as an entry point for downstream repair proteins. This nick can be made either 3' or 5' of the mismatch as the excision-resynthesis mechanism of MMR is bidirectional, orchestrated by exonucleases and the replication machinery (Schmutte et al., 2001; Kadyrov et al., 2009).

Although homologues of MutS and MutL have been found in almost all organisms, homologues of MutH are absent in most prokaryotes and all eukaryotes. The use of hemimethylated DNA as a strand discrimination signal has also not been conserved. The answer to what has replaced MutH in nicking the daughter strand comes from the discovery that MutL in MutH-less MMR systems harbours a latent endonuclease activity (Kadyrov et al., 2006). Unlike MutH, MutL is an endonuclease that is independent of both DNA sequence and structure (Kadyrov et al., 2006; Fukui et al., 2008). The ability of MutL to cleave DNA at any point on the erroneous strand and the

absence of a strand discrimination signal suggests that MutL must rely on other MMR proteins to regulate its nicking activity. The processivity clamp, which plays a role in replication, is also critical for the earlier steps in MMR.

The DNA bound processivity clamp can maintain the temporal and spatial organization of its interacting binding partners with the replication fork. In the presence of a mismatch, MutS is recruited to the damage site on the nascent strand through specific interactions with the processivity clamp (Lopez de Saro et al., 2006; Simmons et al., 2008). The *B. subtilis* MutS-β-clamp complex was proposed to stabilize MutS at the mismatch through repetitive loading of MutS in the early stages of MMR (Simmons et al., 2008). The clamp is also required to activate the latent endonuclease of MutL. In *B. subtilis* MutL, disruption of this β-clamp binding motif completely abrogates endonuclease activity (Pillon et al., 2010). *E. coli* MutL has an analogous β-clamp binding motif on its CTD however mutations in this region resulted in a mild mutator phenotype suggesting this β-clamp binding motif is only critical in MutH-less MMR systems (Pillon et al., 2011).

Once MutL nicks the nascent strand, subsequent coordination of the exonuclease reaction for excision and the replication machinery for re-synthesis completes the MMR mechanism. As these later steps are well defined, focus has been put onto the earlier MMR steps. This includes understanding how error recognition and signaling to direct MMR towards the nascent strand are intimately coordinated with DNA replication.

1.5) MutS recognizes the mismatch or insertion/deletion loops

Initiation of the MMR mechanism occurs when MutS recognizes an error in the replicating strand. In prokaryotes, MutS functions as a homodimer that recognizes both mismatches and small IDLs. Conversely, eukaryotic MutS functions as a heterodimer composed of two of the three identified MutS paralogs (MSH2, MSH3, and MSH6). MutSα (MSH2-MSH6) recognizes both single base mismatches as well as small IDLs of 1 or 2 nucleotides. MutSβ recognizes larger IDLs, containing up to 16 excess nucleotides (McCulloch et al., 2003). In these heterodimers, error recognition is mainly performed by the MSH6 or MSH3 subunit (Warren et al., 2007, Gupta et al., 2011). MutS is a structurally conserved dynamic structure which cycles between states depending on the repair step (Figure 1.2) (Gupta et al., 2012; Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). MutS regulates these states by hydrolyzing ATP in an asymmetric manner (Lamers et al., 2003). In the pre-recognition state, MutS is in an ADP bound state and can slide along and scan DNA due to its conformational flexibility. This allows duplex DNA to enter the dimer's clamp domain and be released if no mismatch was recognized by the mismatch-binding domains (Obmolova et al., 2000).
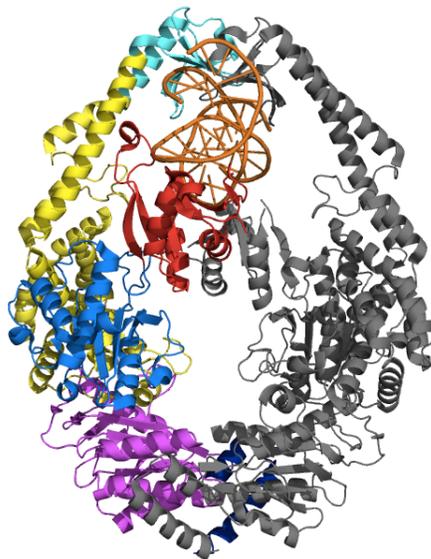
Figure 1.2. Crystal structure of *E.coli* MutS in complex with a G/T mismatched heteroduplex (PDB ID: 1E3M, (Lamars et al., 2000)). G/T mismatch is in orange. Domains of one MutS protomer is coloured. The mismatch binding domain is shown in red. The connector domain is shown in blue. The core/lever domain is shown in yellow. The clamp domain is shown in cyan. The ATPase domain is shown in purple. The helix-turn-helix domain is shown in dark blue. The G/T mismatched heteroduplex is coloured in orange.

Upon mismatch recognition, one MutS mismatch binding domain interacts with the DNA backbone while the other mediate binding with the mismatch through a Phe-X-Glu motif (Lamers et al., 2000; Obmolova et al., 2000). MutS makes a 60° kink to the DNA by inserting the phenylalanine into the minor groove of the mismatch and stacks on top of the unmatched base (Lamers, et al., 2000). The glutamic acid residue is found to form a hydrogen bond with the mismatch (Natrajan et al., 2003). This is thought to sample the reduced stability of the distorted helix and ensure proper mismatch recognition (Sixma, 2001).

MutS undergoes a conformational change into a mobile clamp upon exchanging its ADP for ATP (Figure 1.3) (Gradia et al., 1997; Gradia et al., 1999; Jeong et al., 2011). The mismatch-binding and connector domains, which played a role in mismatch recognition, rotate outwards and the DNA duplex is pushed down into a new channel which loosely encircles the DNA (Groothuizen et al., 2015). This effectively releases the interaction with the mismatch and allows for MutS to freely slide along the DNA helix in search for MutL (Jeong et al., 2011; Gorman et al., 2012). The conformational change in MutS also creates a new binding site composed of one ATPase domain and one connector domain that interacts with and orients MutL to be loaded onto the DNA running through the MutS channel (Groothuizen et al., 2015).
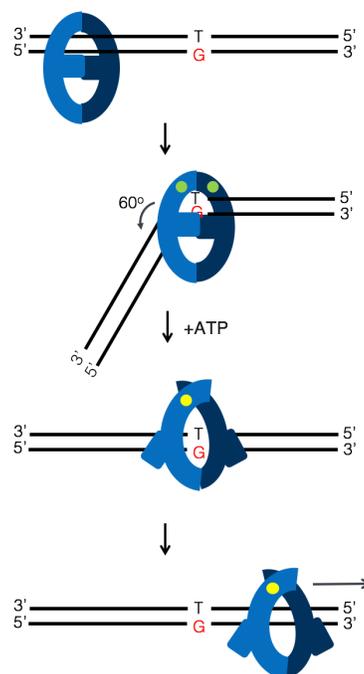
Figure 1.3. Mechanism of mismatch recognition by MutS. MutS ensures proper mismatch recognition by inducing a 60° kink within the DNA. Upon ATP binding, MutS undergoes conformational change into a sliding clamp which releases the mismatch and proceeds to travel along the duplex to recruit MutL.

The mechanism of how mismatch recognition by MutS results in MutL strand specific nicking on the newly synthesized strand remains elusive. Attempts to understand this pathway have resulted in several different models for MutS-MutL complex formation. One model suggests that the MutS-MutL complex forms a sliding clamp to diffuse along DNA in search for the processivity clamp to help induce nicking activity (Gorman et al., 2012). Movement away from the mismatch would allow for several rounds of MutS binding and amplification of the MMR initiation signal until the error has been corrected. Through single molecule fluorescence studies, *Thermus aquaticus* MutL was determined to trap MutS at the mismatch before it forms a sliding clamp (Qiu et al., 2015). This agrees with the other proposed models suggesting the MutS-MutL complex remains at the error followed by looping of DNA or assembly of multiple MutL along DNA until it reaches a strand discrimination signal (Iyer et al., 2006; Kunkel and Erie, 2005; Hombauer et al., 2011; Elez et al., 2012). Preventing MutS from freely sliding along the DNA duplex would ensure that the nicking activity occurs within the vicinity of the mismatch and prevent excess daughter strand excision which can destabilize the genome.

1.6) Strand discrimination

Arguably the most critical aspect of the MMR mechanism is strand discrimination. The endonuclease proteins must correctly determine which strand to nick since the ensuing gap acts as an entry or termination point for excision. As only the erroneous daughter strand requires repair, without tight regulation, indiscriminate nicking

would turn this pathway into a mutagenic process. Biochemical and structural information have suggested that replication timing and strand discontinuities such as pre-existing nicks generated during replication may direct the MMR mechanism.

Strand discrimination signals to direct incision towards the newly synthesized strand are thought to exist in the form of a pre-existing nicks or gaps (Figure 1.4) (Kunkel and Erie, 2005; Kadyrov et al., 2006). On the lagging strand, these signals can be from the discontinuities between Okazaki fragments (Pavlob et al., 2003). The leading strand can use the 3' ends at the replication fork but other means may be used to generate more gaps along the strand as the replication synthesis is continuous. Transient strand breaks introduced by RNase-H2 during removal of misincorporated ribonucleotides may also direct MMR (Ghodgaonkar et al., 2013; Lujan et al., 2013). So far, these nicks are formed as a natural product of DNA replication which guarantees that the gaps will be located on the nascent strand. However, they are unlikely to act as major strand discrimination signals as incorporation of ribonucleotides only occur on average once every 6-8 kb during replication which is much longer than the MMR excision tracts (Reijins et al., 2012). Also, loss of RNase H2 catalytic functions only resulted in a weak mutator phenotype, indicating its expendability (Yao et al., 2013). Nicks generated by MutL itself may also be used as a strand discrimination signal. Introduction of multiple nicks has been shown to increase repair efficiency in both *E. coli* and humans (Hermans et al., 2016). By having multiple ways of generating nicks along the nascent strand, it ensures that timely repair can occur before nick ligation erases the strand discrimination signal.
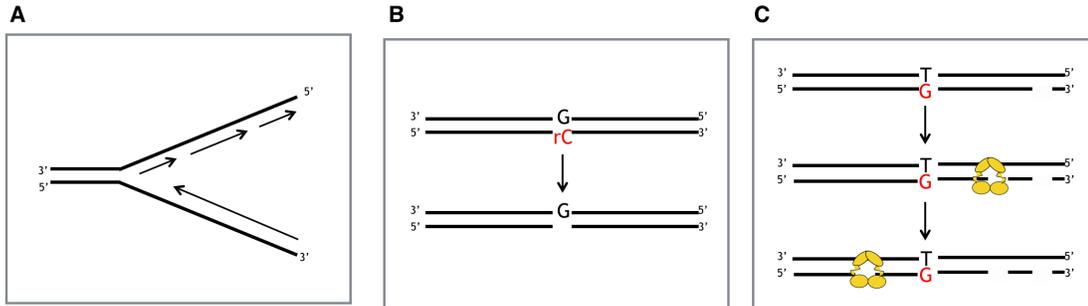
Figure 1.4 Three ways pre-existing nicks can be introduced to the nascent strand. A) As the discontinuities between Okazaki fragments on the lagging strand or at the replication fork on the leading strand. B) From the removal of misincorporated ribonucleotides shown in red. C) From nicks generated by MutL shown in yellow. Mismatch is shown in red.

These pre-existing nicks have been shown to direct eukaryotic MMR *in vitro*. Only a single exonuclease (EXO1) with an obligate 5' to 3' polarity has been implicated in eukaryotic MMR. Yet when the human MMR mechanism was reconstituted *in vitro* from MutSα, MutLα, EXO1, RFC and RPA, it was still capable of mismatch provoked bidirectional excision (Dzantiev et al., 2004). It was revealed that MutLα harboured a latent endonuclease that introduced an additional nick to the original 3'-nicked strand, located 5' to the mismatch (Kadyrov et al., 2006). This allowed for EXO1 to enter and excise DNA. This discovery opened a new set of questions. How did MutL discriminate between strands and selectively nick the nascent strand?

MutS interaction with MutL ensures endonuclease activity occurs only after a mismatch or IDL is recognized (Lenhart et al., 2013). Although MutS also plays a role in stimulating endonuclease activity of MutL by directly loading MutL onto DNA, it does not impose a strand bias (Groothuizen et al., 2015). It was then discovered that the sliding clamp, PCNA, activated MutLα and caused strand discrimination (Pluciennik et al.,

2010). An ATP-dependent reaction by clamp loaders such as γ-complex in prokaryotes or RFC in eukaryotes is required to open the clamp and load it onto DNA (Indiani and O'Donnell, 2006). The pre-existing nicks allow for preferential loading of the clamp onto a 3' double strand-single strand junction with a fixed orientation (Yao et al., 2000). This implies that the fixed geometry of the clamp can dictate the orientation and alignment of its binding partners. PCNA can impose a specific alignment to MutL in relation to the nascent strand (Pluciennik et al., 2010). It is thought that the same clamp-mediated strand discrimination process also occurs for prokaryotic MutL with endonuclease function.

1.7) MutL is multifunctional in DNA mismatch repair

MutL mediates interactions with several mismatch repair proteins and coordinates the earlier mismatch recognition step to the later strand discrimination, excision, and resynthesis steps.

Only one MutL homolog has been identified in prokaryotes which functions as a homodimer. On the other hand, four paralogs of MutL have been identified in eukaryotes with formation of three functional MutL heterodimers, each serving a specific, albeit sometimes redundant, role. MutLα, which is formed by the association of MLH1 and PMS2 (in humans) or MLH1 and PMS1 (in *S. cerevisiae*) is the primary MutL heterodimer required for mismatch repair. MutLγ (MLH1-MLH3) can partially compensate for the lack of MutLα in vitro (Cannavo et al., 2005) but is mainly involved in the resolution of recombination intermediates during meiosis (Wang et al., 1999; Zakharyevich et al., 2010). The role of MutLβ (MLH1-PMS1 in humans, MLH1-PMS2

in *S. cerevisiae*) remains unknown (Raschle et al., 1999). In all cases, MutL is composed of an N-terminal ATPase domain and a C-terminal dimerization domain tethered together by a flexible linker of varying lengths (Guarné et al., 2004). The N-terminal domain (NTD) of MutL (MLH1 in eukaryotic heterodimer) interacts with MutS to ensure endonuclease activity occurs only in the presence of a mismatch. The C-terminal domain contains the endonuclease activity found in several prokaryotes such as *Aquifex aeolicus* (Fukui et al., 2008), *Neisseria gonorrhoeae* (Duppatla et al., 2009), *Thermus thermophiles* (Mauris and Evans, 2009), *Bacillus subtilis* (Pillon et al., 2010), *Pseudomonas aeruginosa* (Correa et al., 2013), *Thermus aquaticus* (Qiu et al., 2015), and in eukaryotes such as humans (Kadyrov et al., 2006), and *S. cerevisiae* (Kadyrov et al., 2007; Gueneau et al., 2013).

1.7.1) MutL N-terminal domain

The MutL NTD has high sequence conservation and is composed of an ATPase domain with a characteristic fold of the Gyrase/Hsp 90/Histidine Kinase/MutL (GHKL) superfamily and a DNA binding groove (Ban et al., 1999). Atomic force microscopy (AFM) of human and yeast MutLα have revealed the presence of four distinct conformations, regulated by a cycle of ATP binding, hydrolysis, and ADP release (Figure 1.5) (Sacho et al., 2008). The open ("extended") state describes a dimerized CTD connected to the NTD by elongated linkers, and the compact ("condensed") state is where the NTD interacts or is folded onto the CTD. The "semi-condensed" state is similar to the "condensed" state however the NTD is not interacting with the CTD. The "one-

armed" state shows only a single NTD is compact against the CTD dimer (Sacho et al., 2008). These different conformations are likely a result of the different ATP affinities exhibited by the two subunits, allowing MutL ATP binding and hydrolysis to occur sequentially or in an alternating manner. Although both MLH1 and PMS1 are both ATPases, MLH1 has a higher intrinsic ATP binding affinity (Hall et al., 2002). This supports the functional asymmetry of eukaryotic MutLα.
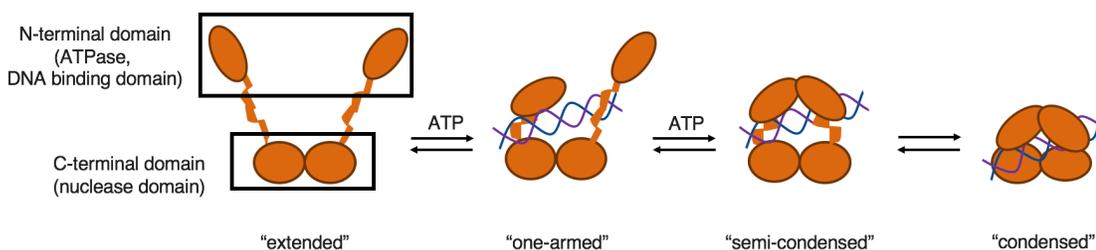
Figure 1.5 MutLα undergoes conformational changes upon ATP and nucleotide binding.

Nucleotide binding stimulates ATP hydrolysis and, in turn, coordinates the dynamic equilibrium between these different conformations (Ban et al., 1999; Sacho et al., 2008). Dimerization of the NTD has been suggested to encircle DNA and through compaction, bring the DNA towards the CTD which contains the endonuclease site (Kadyrov et al., 2006; Pillon et al., 2010). Mutations in MLH1 strongly reduced DNA binding while mutations in PMS1 did not exhibit reduced DNA binding (Hall et al., 2003). This suggests that the two independent DNA binding sites can bring together and facilitate communication between two different duplex DNA molecules.

Variation among linker lengths, amino acid substitution/deletion permissibility, and sequences are also prevalent between MutL homologs. So far, the role of the linker

has only been studies in yeast. MLH1 has a linker length of ~1v50 amino acids and PMS1 has a linker length of ~220 amino acids. Linker deletions on MLH1 have a stronger mutator phenotype, while deletions made in the PMS1 linker had a larger effect on MutLα DNA binding activity (Guarné and Charbonnier, 2015). The asymmetry observed in MutLα likely plays a part in regulating MutL function and interactions within the MMR pathway.

Prokaryotic MutL functions as a homodimer, implying the ATPase domain, DNA binding sites, and linkers on both subunits are identical. Whether similar ATP induced functional and structural asymmetry is observed in prokaryotic MutL homodimers with endonuclease activity remains unclear.

1.7.2) MutL C-terminal domain

Although eukaryotic and prokaryotic MutL CTD homologs share limited sequence identity, it is structurally conserved, composing of a dimerization subdomain connected by a helix to the regulatory subdomain (Figure 1.6) (Guarné et al., 2004; Pillon et al., 2010). The dimerization interface consists of hydrophobic residues from a four-strand β-sheet which facilitates the formation of constitutive dimers. The yeast heterodimerization interface is twice as large and facilitates additional interactions (Gueneau et al., 2013). Including the four-stranded β-sheet, the interface also consists of the last 12 residues of PMS1 extending into the linker and regulatory domains of MLH1 and the last 14 residues of MLH1 extending and reaching the PMS1 metal-binding site (Gueneau et al., 2013).
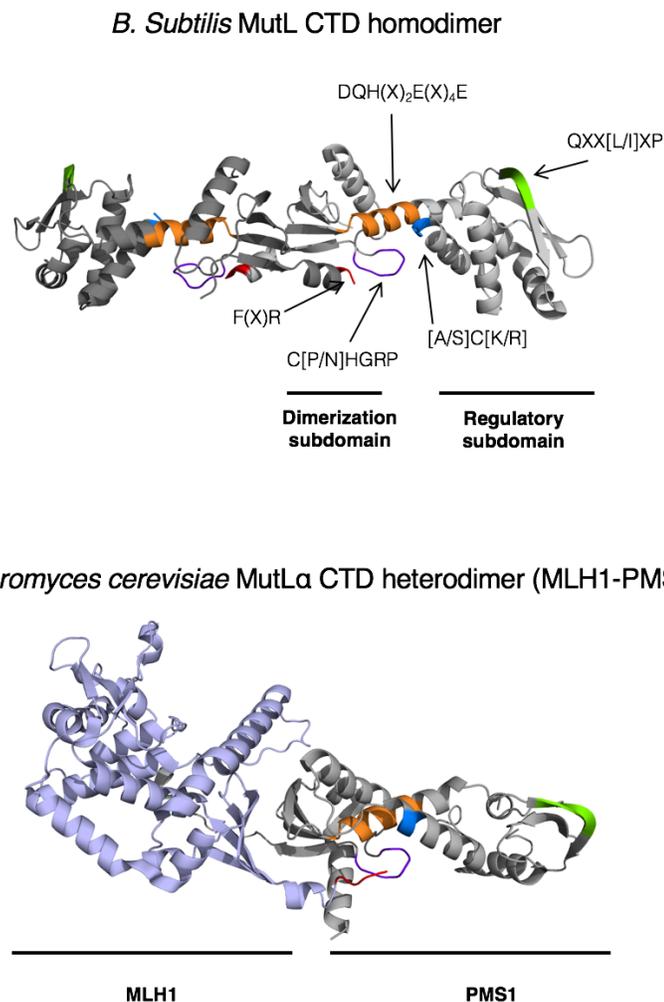
Figure 1.6 Crystal structures of MutL C-terminal domains. *B.subtilis* MutL C-terminal domain (PDB ID: 3KDK, (Pillon et al., 2010) and *Saccharomyces cerevisiae* MutLα C-terminal domain (PDB ID: 4FMO, (Gueneau et al., 2013)). The four motifs which make up the endonuclease active site, $DQH(X)_2E(X)_4E$, [A/S]C[K/R], C[P/N]HGRP, and F(X)R are coloured in orange, blue, purple, and red respectively. The processivity clamp (β-clamp in prokaryotes, PCNA in eukaryotes) binding motif, QXX[L/I]XP, is coloured in green.

The MutL endonuclease site resides in the junction between the dimerization and external subdomains. Despite the low sequence homology, the active site in MutH-less organisms such as *B. subtilis* MutL, *S. cerevisiae* PMS1, and human PMS2 share the

same endonuclease and metal-binding motifs (Kadyrov et al., 2006; Pillon et al., 2010; Gueneau et al., 2013). The catalytic site consists of the conserved endonuclease motif $DQH(X)_2E(X)_4E$ and motifs C[P/N]HGRP, [A/S]C[K/R] and F(X)R which constitute a metal-binding site (Kosinski et al., 2008). The major difference between the prokaryotic and eukaryotic structures is the F(X)R motif.

Eukaryotes have a stricter consensus sequence for the F(X)R motif (FERC). As seen in yeast MutLα, the MLH1 FERC motif extends across the dimerization interface and directly participates in the formation of the PMS1 endonuclease site. The cysteine at the end of the FERC motif may act to increase the affinity for the two zinc ions as it is positioned to chelate the two zinc ions found in the active site. Indeed, tight Zn^{2+} binding was observed in yeast MutLα as the crystal structure of the C-terminal domain contained the metal ion despite its absence in the purification and crystallization conditions (Gueneau et al., 2013). Conversely, the structure of *B. subtilis* MutL showed a disordered FXR motif which does not extend as far into the other subunit as its eukaryotic counterpart nor participate in the endonuclease site. Although two zinc ions were also observed in the *B. Subtilis* MutL C-terminal domain, the zinc ions must be supplied in the crystallization conditions and only one was Zn^{2+} ion was fully occupied in the electron density map (Pillon et al., 2010).

The different affinities for zinc in eukaryotic and prokaryotic MutL dimers reflect structural and functional differences in the endonuclease sites. In yeast MutLα, both Zn^{2+} sites were fully occupied and coordinated by conserved residues, reminiscent of a two-metal ion catalysis mechanism (Yang et al., 2008). These two metal binding sites may

define a catalytic site necessary for endonuclease activity (Gueneau et al., 2013). The conserved aspartate residue in the $DQH(X)_2E(X)_4E$ motif has an N-capping mechanism which stabilizes the N-terminus helix containing the motif. On the other hand, *B. subtilis* MutL only one zinc ion was fully occupied. The second partially occupied zinc ion may define a catalytic role which can be replaced by other metal ions in coordination to the conserved aspartate residue in the $DQH(X)_2E(X)_4E$ motif (Pillon et al., 2010). Based on the zinc binding induced conformation changes and similarities to the regulatory metal-binding site found in iron dependent repressors from the DtxR/MntR family, the zinc-binding site in prokaryotic MutL is proposed to have a structural role instead of a catalytic role (Kosinski et al., 2008; Pillon et al., 2010).

It is not clear whether the different interpretations of roles within the endonuclease sites are unique to each prokaryotic and eukaryotic MutL dimer. The catalytic mechanism for MutL endonuclease activity is still unknown and cannot be predicted because it does not resemble any other known nuclease. Instead, based on sequence and structural similarities, the endonuclease site resembles the iron dependent transcription repressors from the DtxR/MntR family which do not have endonuclease functionality (Kosinski et al., 2008). Understanding of the mechanistic similarities and differences between prokaryotic and eukaryotic MutL endonuclease activity awaits a structure in coordination with DNA.

1.7.3) Regulation of the MutL endonuclease activity

MutL endonuclease activity is intrinsically weak and unspecific. The CTD which encompasses the endonuclease site does not bind DNA (Pillon et al., 2010; Gueneau et al., 2013). Requiring other DNA binding domains to bring the nascent strand towards the endonuclease site for cleavage reflects a powerful regulatory mechanism to prevent indiscriminate nicking. Tight regulation is crucial to ensure DNA cleavage occurs only in the presence of a mismatch/IDLs and is directed to the newly synthesized daughter strand.

ATP binding may regulate the progression of the MMR mechanism by communicating the presence of a mismatch to the endonuclease domain of MutL. Human and yeast MutLα as well as *B. subtilis* MutL endonuclease activity is stimulated by ATP. ATP-dependent MutL dimerization and compaction can encircle DNA and bring the substrate from the NTD towards the active site on the CTD, facilitating cross-talk between the two domains (Kadyrov et al., 2006; Kadyrov et al., 2007; Sacho et al., 2008, Pillon et al., 2010).

MutL by itself cannot discriminate between the template and nascent strand. Studies have shown that the interaction with the replication sliding clamp, β-clamp in prokaryotes and PCNA in eukaryotes, stimulates the endonuclease activity and determines strand discrimination (Lee and Alani, 2006; Pluciennik et al., 2010). Initial studies showed that only in the presence of PCNA does MutLα nicking activity occur on the strand with a pre-existing nick (Kadyrov et al., 2006). Recent studies have shown that the conserved QXX[L/I]XP motif found on the external surface of human PMS2 and yeast PMS1 is essential for direct interaction with PCNA and PCNA-dependent activation

of the MutLα endonuclease nicking activity (Genschel et al., 2017). Similarly in prokaryotic MutL, the conserved QXX[L/I]XP motif found on the external surface of *B. subtilis* MutL CTD also mediates interaction with the β-clamp (Pillon et al., 2011). The clamp may orient and directly thread DNA onto the endonuclease site as the MutL CTD itself does not bind DNA (Pillon et al., 2015). However, the molecular mechanism behind MutL endonuclease activity stimulation via clamp regulation remains unknown.

1.8 Thesis objective

Prokaryotic MutL homodimers have two endonuclease active sites and two clamp binding motifs as opposed to eukaryotic MutLα heterodimers which only have one of each. The main objective of this thesis was to determine if both endonuclease sites are necessary for MutL nicking activity and whether both subunits of the MutL dimer needs to interact with the processivity clamp.

The MutL endonuclease site does not resemble any known nuclease, hence the molecular mechanism remains elusive. My second goal in this thesis was to design an approach for future structural investigation of a ternary complex, MutL bound to DNA while interacting with the clamp, using different crosslinking methods to stabilize weak protein-protein and protein-nucleotide interactions.

CHAPTER 2

MATERIALS AND METHODS

2.1) Cloning of the *B. subtilis* MutL and SKN1-MutL variants

B. subtilis MutL C-terminal domain (CTD; pAG 8188, residues 433-627) and full length MutL (FL; pAG 8220, residues 1 – 627) were cloned as described earlier (Pillon et al., 2010). The SKN1-(linker length)-fusion proteins were created by connecting the *Caenorhabditis elegans* SKN1 DNA binding domain (residues 450 – 533) to the N-terminal end of the MutL C-terminal domain with an 8 amino acid (GSASKSEF) linker (L8) as part of an undergraduate thesis project by Julia Cai in the Guarné lab. Variants of the CTD and L8 with a point mutation at the zinc binding site (E468K) and/or lacking the β-clamp binding motif (^{487}QEMIVP —> ^{487}AEMAAP) were generated using the Q5-Site directed mutagenesis Kit (New England Biolabs (NEB)). Primers were designed according to the NEBaseChanger.neb.com software (Table 2.1). DNA oligomers were purchased through Integrated DNA Technologies (IDT). All mutants were verified by DNA sequencing (MOBIX, McMaster University).

Table 2.1 Primers used to generate L8-CTD mutants

Name	Use	Sequence
ag1953	E468K	$5`$CGCCGCCCAAAAACGTATTAA$3`$
ag1954		$5`$TGCTGGTCGATAATATATAGGC$3`$
ag2056	487QEMIVP—> 487AEMAAP	$5`$GGCAGCACCGCTGACGTTCCACTAC$3`$
ag2057		$5`$ATCTCTGCCACCTCAGGATCAACCTC$3`$

2.2) Protein expression and solubility of *B. subtilis* MutL heterodimers

Expression and solubility of L8-CTD heterodimers were assayed as previously described (Rashev et al., 2017), with minor modifications. Plasmids encoding for the MutL C-terminal domain (pAG 8188) and L8 (pAG 8887) plasmids were transformed together into *E. coli* BL21 (DE3) cells and grown to an OD_{600} of 0.7 (sample '-'). Expression was induced by the addition of isopropyl β-D-1-thiogalactopryanoside (IPTG) at a final concentration of 0.5 mM. Cultures were grown at 16 ^{o}C, 25 ^{o}C, and 37 ^{o}C for 16 h, 5 h, and 3 h respectively (sample '+'). To test for solubility, cells were harvested by centrifugation and resuspended in lysis buffer (20 mM TRIS pH 8.0, 1.4 mM β-mercaptoethanol). Cells were lysed with 1mg/mL of lysozyme and left on ice for 30 min. Followed by the addition of 90 mM KCl, 10 mM $MgCl_2$, and 0.05% LDAO and left on ice for 15 min. 20 units of Dnase I were added and incubated at room temperature for 15 min to sheer chromosomal DNA. Soluble protein in the lysate was isolated by

centrifugation (sample 's'). Samples '-', '+', and 's' were resolved in 12% SDS-polyacrylamide gels to assess the amount of soluble versus total protein expressed.

2.3) Expression and purification of L8-CTD, FL, and CTD

Expression and purification of the his-tagged *B. subtilis* L8-CTD heterodimer was optimized in collaboration with another graduate student in the Guarné lab (Mary Carmen Ortiz Castro) and described in Ortiz Castro (2016). Expression and purification of the additional heterodimer combinations (Table 2.2) followed the same protocol.

Full length MutL (pAG 8220) and MutL C-terminal domain (pAG 8188) were overproduced in *E. coli* BL21 Star (DE3) cells and *E.coli* BL21 (DE3) cells respectively. Cells were grown to an OD_{600} of 0.7 at 37 ^{o}C before expression was induced by the addition of IPTG at a final concentration of 1 mM, followed by incubation at 37 ^{o}C for 3 h with agitation on an orbital shaker. FL and CTD were purified as described earlier by Pillon et al., (2010), with minor modifications. Histidine-tags on FL and CTD remained. Protein were concentrated with a 10 kDa MW cut-off concentrator (Vivaspin) in storage buffer (20 mM Tris pH 8.0, 5 mM DTT, 100 mM KCl, and 5% glycerol (v/v)). Glycerol concentration was increased from 5% to 25% before the protein were frozen in small volume aliquots and stored at -80 ^{o}C.

Table 2.2 L8-CTD heterodimer combinations

Name	L8 plasmid	Mutations	CTD plasmid	Mutations	L8-CTD
L8-CTD	pAG 8887	N/A	pAG 8188	N/A	
L8-CTD^I	pAG 8887	N/A	pAG 8238	E468K	
L8^I-CTD	pAG 8989	E468K	pAG 8188	N/A	
L8^I-CTD^I	pAG 8989	E468K	pAG 8238	E468K	
L8^{Iø}-CTD	pAG 8991	^{487}QEMIVP —> ^{487}AEMAAP, E468K	pAG 8188	N/A	
L8^ø-CTD^I	pAG 8987	^{487}QEMIVP—> ^{487}AEMAAP	pAG 8238	E468K	
L8^ø-CTD	pAG 8987	^{487}QEMIVP—> ^{487}AEMAAP	pAG 8188	N/A	
L8-CTD^{Iø}	pAG 8887	N/A	pAG 9093	487QEMIVP —> AEMAAP, E468K	
L8^I-CTD^ø	pAG 8989	E468K	pAG 8350	487QEMIVP—> AEMAAP	
L8-CTD^ø	pAG 8887	N/A	pAG 8350	487QEMIVP—> AEMAAP	

2.4) Expression and purification of *B. subtilis* β-clamp

 B. subtilis β-clamp was transformed into *E. coli* BL21 (DE3) *recA⁻ (BLR)* cells.

Cells were grown to OD$_{600}$ of 0.7 at 37°C and protein expression was induced by adding

IPTG to a final concentration of 0.5 mM. Culture was incubated with agitation on an orbital shaker at 37°C for 3 h. Cells were harvested by centrifugation and stored at -80°C. His-tagged *B. subtilis* β-clamp (pAG 8337) was purified as described earlier with minor modifications (Pillon et al., 2011). A final purification step using size exclusion chromatography was added. Protein (500 μL) was loaded into a Superdex-200 (GE Healthcare) in buffer containing 20 mM Tris pH 7.6, 5 mM DTT, 100 mM KCl, and 5% glycerol (v/v). Fractions containing the β-clamp were collected and concentrated using a 30 kDa MW cut-off centricon (Vivaspin). Glycerol concentration was increased from 5% to 25% before the protein were frozen in small volume aliquots and stored at -80 $^{\circ}$C.

2.5) Dynamic light scattering

Dynamic light scattering was performed using a Zetasizer Nano S (Malvern Instruments). All measurements were taken using a 12 μL quartz cell (ZEN2112) at 4 $^{\circ}$C. Size distribution of the samples was calculated based on the correlation function provided by the Zetasizer Nano S software.

2.6) Generation of linear 195 bp substrates

Linear DNA substrate (195 base pairs) was amplified using the pUC19 vector (Invitrogen). The first 195 bp substrate (Sub1) was generated using the forward 5' end labeled 5'6-carboxyfluorescein d(TGTAAAACGACGGCCAGTGAATTCGAGCTCGG) primer and reverse 5'(AGTTAGCTCACTCATTAGGCACCCCAGGC) primer (BioBasic Inc.). This amplified region 378-572 contained an SKN1 site ([465]GTCAT). The

second 195 bp substrate (sub2) did not contain an SKN1 site and was generated using the forward 5' end labeled 5' 6-carboxyfluorescein-d(GCGTTTCTGGGTGAGCAAAA) primer and the reverse 5'-(GAAATGTGCGCGGAACCC) primer (BioBasic Inc) which amplified region 469 – 663. PCR reaction mixture (50μL) contained 300 ng pUC19 template, 0.5 μM forward primer, 0.5 μM reverse primer,0.4 mM dNTPs, 1x PFU buffer, and 1 unit/μL PFU enzyme. The cycle conditions were programmed for an initial denaturation at 95°C for 5 min, followed by 20 cycles of 95°C for 30 s, 55°C for sub1 (52°C for sub2) for 1 min and 72°C for 35 s, and the final extension period at 72°C for 20 min. PCR fragments were loaded onto a 2% agarose gel and ran for 45 min at 100 V. Bands were excised under UV light (320 nm) and gel extracted using the QIAEX II agarose gel extraction kit (QIAGEN).

2.7) Endonuclease assay with L8-CTD, FL, and CTD

Endonuclease assays were performed as described previously (Pluciennik et al. 2010, Pillon et al. 2015) with minor modifications. MutL variants (240 nM) were incubated with 195 bp linear DNA substrate (10 nM) in the absence and presence of of β (240 – 480 nM). Reactions were incubated at 37°C for 1 h in reaction buffer (20 mM Tris pH 7.6, 30 mM KCl, 1 mM MnCl$_2$, 1 mM MgCl$_2$, 152 pM Zn(O$_2$CCH$_3$)$_2$, 0.05 mg/mL BSA, 4% glycerol). Reaction was terminated by the addition of 25 mM EDTA and 1mg/mL Proteinase K and incubation at 55 $^{\circ}$C for 20 min. Immediately afterwards, 2x-loading dye (90% formamide, 0.025% bromophenol blue, 0.025% xylene cyanol, 0.5 mM EDTA, 10% glycerol) was added to the reaction and incubated at 95 $^{\circ}$C for 5 min.

Digestion products were analyzed by electrophoresis through an 8% polyacrylamide (8 M Urea) gels in 0.5x TRIS-borate-EDTA buffer and visualized using the Typhoon Trio+ (GE Healthcare, CMCB McMaster University).

2.8) Cloning of the MutL and β-clamp cysteine variants for crosslinking

Cysteine modified *B. subtilis* MutL containing mutations C69S, C424S, E485C and C531S (CTD^{cys}; pAG 8803; residues 433-627) and *B. subtilis* β with a S379-C380 dipeptide at the C-terminus of the protein (pAG 8803) were generated as described in Pillon et al., (2015). Modification of *B. subtilis* β with a S379-C380 dipeptide to include a point mutation (C178A) and extend the C-terminus end to Ser379-Glu380-Ser381-Glu382-Cys383, β^{cysL} (pAG 9048; residues 1-380), was generated using the Q5-Site directed mutagenesis Kit (New England Biolabs). Cysteine variant of the *B. subtilis* β-clamp with mutations C178A and D218C, β^{cysD} (pAG 9027; residues 1-380) was also generated by using the Q5-Site directed mutagenesis Kit (New England Biolabs). Primers were designed according to the software NEBaseChanger.neb.com (Table 2.3). DNA oligomers were purchased through Integrated DNA Technologies (IDT). Constructs were verified by sequencing of DNA fragments through MOBIX Facility at McMaster University.

Table 2.3 Primers used to generate MutL and β-clamp cysteine variants

Name	Use	Sequence
ag3036	Ser379-Glu380-Ser381-Glu382-Cys383	5` AGGATGCTAAGGATCCGGCTGC 3`
ag3037		5` GAACCGCTATAGGTTCTGACAGGAAGG 3`
ag2016	D218C	5` CAAGATTTTATGTGACAACCAGGAACTTGTAG 3`
ag2017		5` CTGAGTTCAGTTAAACTTTTTC 3`
ag2090	C178A	5` TGAATTATTAGCCACTGCAACGGATAG 3`
ag2091		5` CTTTGCTCCACTTTCCAG 3`

2.9) Expression and purification of the MutL and β-clamp cysteine variants

B. subtilis β variants were overproduced in *E. coli* BL21 Star (DE3) pRARE pLysS. Cells were grown to OD_{600} of 0.7 at 37°C and protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM. Culture was incubated with agitation on an orbital shaker at 25°C for 5 h. Cells were harvested by centrifugation and stored at -80°C. The cysteine modified *B. subtilis* MutL was expressed similarly to the β variants except they were transformed into *E. coli* BL21 (DE3) cells and after IPTG induction, the culture was incubated with agitation on an orbital shaker at 37°C for 3 h. MutL CTD^{Cys} and β variants were purified as described by the purification of BsMutL and Bsβ in Pillon et al. (2010) with some modifications. Proteins were further purified by size exclusion chromatography equilibrated with 20 mM Tris pH 7.6, 150 mM NaCl, 10 mM DTT, and 5% glycerol before complex formation. In the situation where the β^{cysD}-DNA complex was formed using the heterobifunctional crosslinker succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), β^{cysD} was further purified by size

exclusion chromatography equilibrated with 20 mM HEPES pH 7.3, 150 mM NaCl, 1 mM EDTA, and 5% glycerol.

2.10) CTD^{cys}- β^{cysL} complex formation

To form the CTD^{cys}- β^{cysL} complex, β^{cysL} was incubated with CTD^{cys} at a 1:1 ratio. The sample was dialyzed against 500 mL of dialysis buffer A (20 mM Tris pH 7.6, 150 mM NaCl, 50 nM ZnCl$_2$, 10 mM DTT, 5% glycerol) for 2 h at 4 ^{o}C. The mixture was transferred into 500 mL of dialysis buffer B (same as A but with 5 mM DTT) for 1h, followed by a 1 h dialysis in 500 mL dialysis buffer C (same as A but with 0 mM DTT), and lastly transferred to 500 mL of fresh dialysis buffer C to be left overnight. Complex formation was monitored over time by resolving samples on 9% denaturing gels stained with Coomassie Brilliant Blue.

2.11) Generation of 19/19 and 15/19mer substrates

Substrates used in the endonuclease assays were either the 195 bp (Sub1) or a 19/19mer substrate with a 4 base 3' and 5' overhang. The shorter substrate was generated by annealing a 5' end labeled 5' 6-carboxyfluorescein-d(TTTTCCCATCGATCGGTAT) 19mer oligomer to a 19 mer 5'(CCCATCGATCGGTAT) oligomer (BioBasic Inc). Oligomers were annealed at a 1.2:1 (fluorescently labeled oligo:non-labeled oligo) ratio in annealing buffer (10 mM Tris pH 7.6 and 50 mM NaCl). Annealing reaction mixture was placed in a 1 L boiling water for 5 min and let to cool overnight. Generating the thiol modified 15/19mer substrate with a 4 base 3' overhang (hence called "S-15/19mer") with

a 5'(Thiol modified S-S C6/CCCATCGATCGGTAT) 15 mer oligomer and a 19 mer 5'(CCCATCGATCGGTAT) oligomer (BioBasic Inc) was done in the same manner as the 19/19 mer substrate. The amino modified 15/19 mer substrate with a 4 base 3' overhang (hence called "Am-15/19 mer) was generated by annealing a 5'(amino modified C6/CCCATCGATCGGTAT) 15 mer oligomer with a 19 mer 5'(CCCATCGATCGGTAT) oligomer (BioBasic Inc) in a similar manner to the 19/19 mer substrate except the annealing buffer used consisted of 10 mM HEPES pH 7.5 and 50 mM NaCl.

2.12) β^{cysD}-DNA complex formation

Before complex formation, the S-15/19mer substrate was incubated in 10 mM DTT for 30 min at room temperature to deprotect the thiol group. To form the β^{cysD}-DNA complex, β^{cysD} was incubated with the substrate at a 1:1.2 (β:DNA) ratio. Subsequent dialysis steps to remove DTT were identical to CTD^{cys}- β^{cysL} complex formation. Complex formation was monitored over time by resolving samples on 11% denaturing gels stained with Coomassie Brilliant Blue.

2.13) β^{cysD}-DNA complex formation with heterobifunctional crosslinkers

Heterobifunctional maleimide crosslinking of protein to DNA was adopted from Tram et al., (2016). To form the β^{cysD}-DNA complex, the conjugator succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC) was resuspended in DMSO and reacted with the Am-15/19 mer at a 10:1 ratio. This was done by adding 10μL of the Am-

15/19 mer substrate to 140 μL ultrapure water and 40 μL 10x PBS, followed by the addition of 80.5 μL SMCC and 159.5 μL DMSO. The mixture was vortexed and centrifuged briefly using a benchtop centrifuge before incubating at 37 ^{o}C for 1 h. The addition of 200 μL 1x PBS, 60 μL of 3 M NaOAc (pH 5.2) and 1.25 mL cold ethanol to the mixture allowed for the precipitation of DNA and SMCC-DNA conjugates while the excess SMCC remains soluble. The reaction was left in -20 ^{o}C overnight (16 h). Precipitated DNA was isolated by centrifugation at 20, 000 g for 20 min at 4°C. The supernatant was removed and dried under vacuum. The dried conjugate was resuspended in 400 μL of β^{cysD} at a ratio of 5:1 (DNA: β^{cysD}) and incubated at 4°C overnight (16 h). The crude conjugate was concentrated in a 30 kDa MW cutoff 0.5 mL centrifugal filter column (Amicon) by centrifugation at 14, 000 g for 5 min. Column was inverted into a collection tube and centrifuged at 1, 000 g for 3 min. Collected β^{cysD}-SMCC-DNA product was stored at 4 ^{o}C and resolved on pre-cast gradient gels (4-15%) (BioRad Inc.) stained with Coomassie Brilliant Blue.

CHAPTER 3

RESULTS

3.1) Engineering the *B. subtilis* MutL heterodimer

To understand the contribution of each active site and β-clamp binding motif on the *B. subtilis* MutL endonuclease activity, we needed to create a system where the CTD homodimer subunits could be individually manipulated. A *B. subtilis* MutL heterodimer was formed by the addition of a *C. elegans* SKN1 DNA binding domain (SKN1) to the N-terminus of one CTD subunit. SKN1 is a monomer that binds to DNA with high affinity (1 nm K_d) and sequence specificity ((G/A)TCAT) (Rupert et al. 1998). The NTD of full length MutL was not chosen for formation of the heterodimer because it binds DNA weakly and with no sequence specificity. By anchoring the substrate at a known sequence, we may be able to determine where MutL may preferentially cleave.

A former undergraduate thesis student (Julia Cai) generated variants of the SKN1-(linker)-CTD chimera with different glycine-serine rich linker lengths of 8, 15, 21, 25, 30, and 32 amino acids. The chimera with a linker length of 32 amino acids (L32) was readily purified, stable during concentration, and ready to use for biochemical assays (Ortiz Castro, 2016). An electrophoretic mobility shift assay was used to show that L32 binds specifically to DNA substrates with the SKN1 site. However, problems arose while characterizing the endonuclease activity of the L32 homodimer as, in comparison to the CTD homodimer, there was less observable nicking activity using substrates with or without the SKN1 site. This was attributed to the sequestration of the substrate away from

the endonuclease site by the positively charged nature of the SKN1 domain. To avoid sequestering the DNA, a shorter linker must be used to generate the chimera.

Initial assessment of the chimera's expression and solubility as a homodimer found that variants with a linker length shorter than 25 amino acids were unstable and insoluble after cell lysis. As seen in the small-scale expression and solubility of L8 (Figure 3.1A), the appearance of an additional band at 36.7 kDa in the induced culture lane, not observed in the pre-induction lane, indicated L8 expression. However, the band was not observed in the supernatant after the cells were lysed, suggesting that the L8 was insoluble. We hypothesized the chimera homodimers with shorter linkers were unstable because the proximity of the highly positively charged SKN1 domains caused electrostatic repulsion. Longer linker lengths (>30 amino acids) were then required for stable homodimers to allow enough SKN1 domain separation for proper CTD dimerization. This suggests that eliminating one of the SKN1 domains and effectively forming a CTD heterodimer would allow for the use of shorter linkers. To assess this speculation, expression and solubility of L8 was assessed in the presence of the CTD (Fig. 3.1B). The protein band corresponding to L8 (36.7 kDa) was observed after induction, as with the band corresponding to the CTD which runs a bit higher than its 22.6 kDa molecular weight. Both L8 and the CTD remained soluble in the supernatant post-cell lysis. As L8 presumably does not form soluble homodimers, the presence of L8 in the soluble fraction suggests it is interacting with the CTD as a heterodimer.
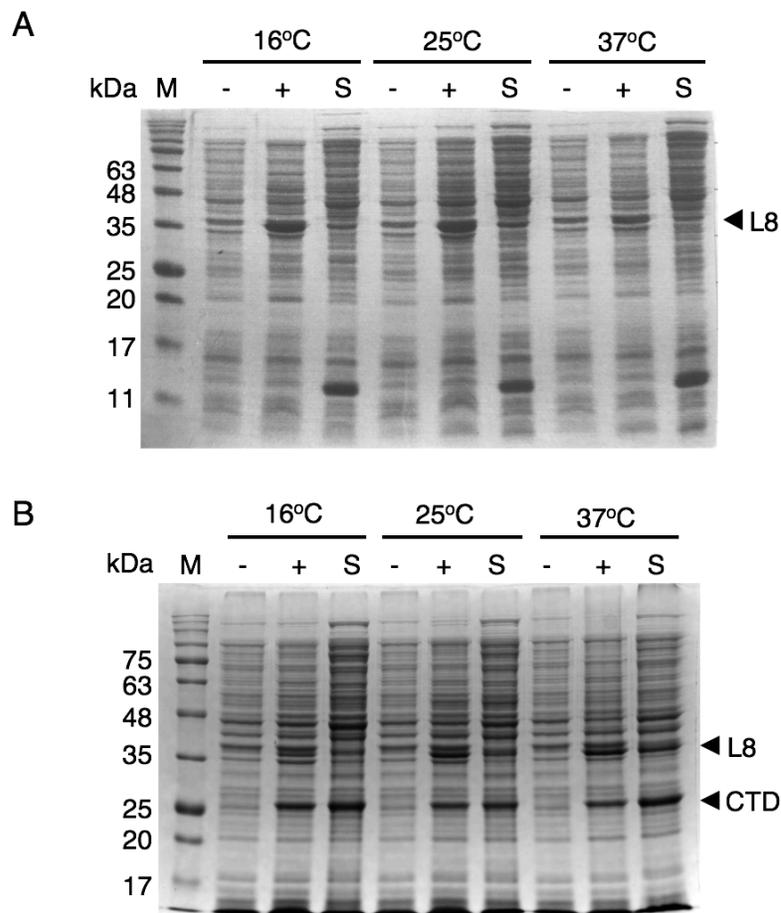
Fig 3.1 Expression and solubility of L8 and L8-CTD in BL21 (DE3). (A) Solubility assay for L8 in BL21 (DE3) at three growth temperatures. SDS-polyacrylamide gel shows (from left to right): 1) molecular weight markers (M); 2-4) protein content of the pre-induction (-), post-induction (+), and soluble (s) fractions of the cultures grown at 16 ^{o}C for 16 h; 5-7) protein content of the pre-induction (-), post-induction (+), and soluble (s) fractions of the cultures grown at 25 ^{o}C for 5 h; 8-10) protein content of the pre-induction (-), post-induction (+), and soluble (s) fractions of the cultures grown at 37 ^{o}C for 3 h. (B) Solubility assay for L8 and CTD co-expressed in BL21 (DE3) at three growth temperatures. Gel is loaded: 1) Molecular weight markers (M, in kDa); 2-4) protein content of the pre-induction (-), post-induction (+), and soluble (s) fractions of the cultures grown at 16 ^{o}C for 16 h; 5-7) protein content of the pre-induction (-), post-induction (+), and soluble (s) fractions of the cultures grown at 25 ^{o}C for 5 h; 8-10) protein content of the pre-induction (-), post-induction (+), and soluble (s) fractions of the cultures grown at 37 ^{o}C for 3 h.

3.2) Purification of the MutL L8-CTD heterodimer

During cell lysis, rapid precipitation of the L8 protein was observed. The remaining soluble protein were first purified using nickel affinity chromatography because both L8 and the CTD contained N-terminal His$_6$ tags (Fig 3.2). The fractions under the apex of the peak were pooled, followed by further purification using an ionic exchange (S-sepharose) chromatography (Fig 3.3). The theoretical pI of the SKN-1 domain and the CTD, as determined by using the Expasy Protparam tool, are 10.68 and 5.49 respectively. Under buffer conditions at pH 7.6 used throughout the purification, the CTD has a net negative charge and the SKN1 domain has a net positive charge. The CTD did not bind to the column and flowed through during protein loading. An elution gradient from 200 – 800 mM KCl resulted in two major peaks. Subsequent SDS-PAGE analysis showed a 1:1 ratio of L8 to CTD for peak 1, suggesting this peak corresponded to the L8-CTD heterodimer, while peak 2 corresponded to only L8. DLS measurements performed on the L8 heterodimer from peak 1 confirmed sample homogeneity and stability. Attempts to concentrate the L8 protein from peak 2 resulted in protein aggregation, underlining its instability.
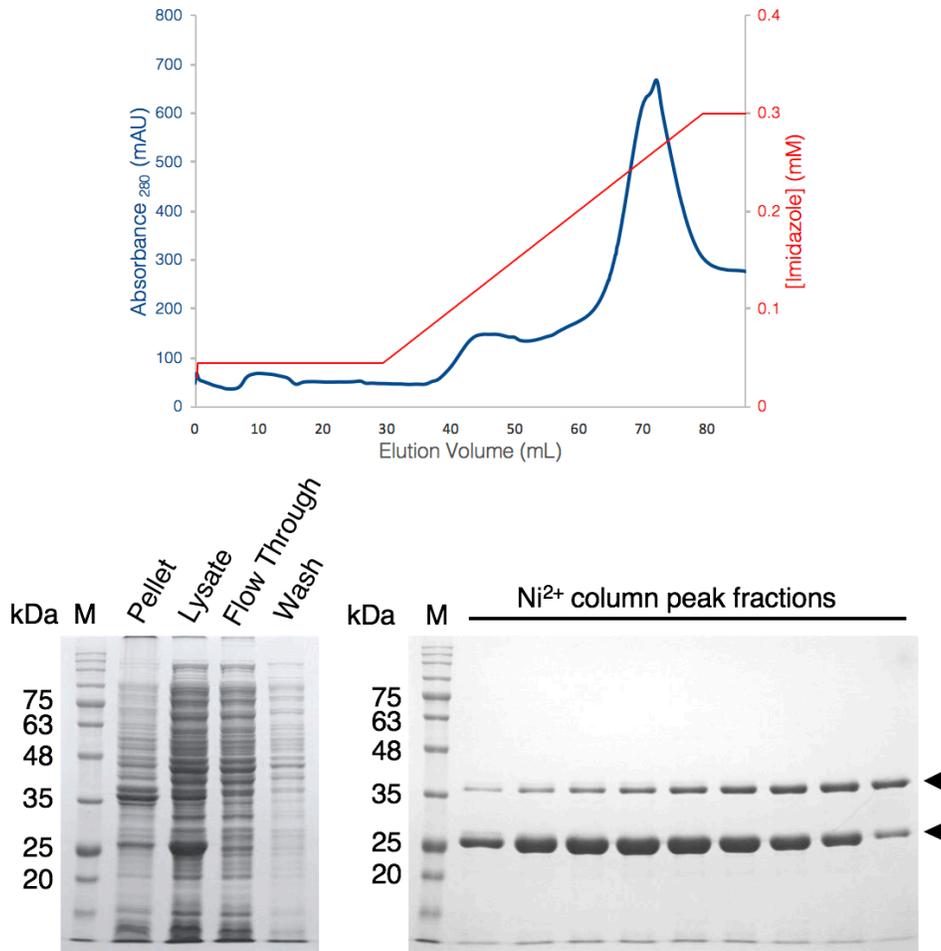
Fig 3.2 Purification of the L8-CTD heterodimer with a nickel-affinity column. Initial step to purify the L8-CTD heterodimer using nickel-affinity chromatography. Elution profile of the three his-tagged MutL dimer species (CTD homodimer, L8 homodimer, and CTD-L8 heterodimer). The blue line represents the absorbance at 280 nm (mAu) and the red line represents the concentration of imidazole (mM). First 11% SDS-polyacrylamide gel shows (from left to right): 1) molecular weight marker (M); 2) resuspended cell debris pellet (pellet); 3) lysate loaded into the column (lysate); 4) protein that flowed through the column after loading (flow through); 5) protein from the first wash (wash). Second 11% SDS-polyacrylamide gel shows (from left to right): 1) molecular weight marker (M); 2-9) fractions under the apex of the peak contained a mixture of CTD and L8 homodimers with L8-CTD heterodimer (Ni^{2+} column peak fractions) and were pooled for subsequent ion exchange chromatography.
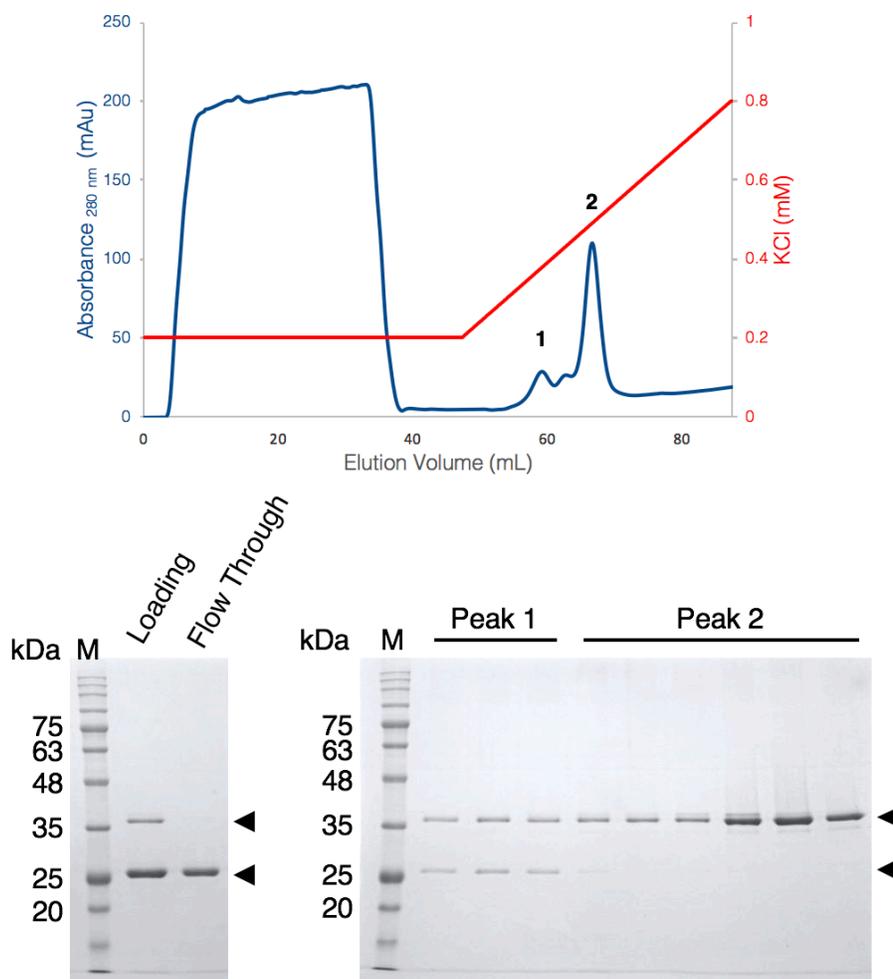
Fig 3.3 Purification of the L8-CTD heterodimer with an S-Sepharose column. Elution profile of the His-tagged L8-CTD heterodimer from an S-Sepharose chromatography column showing separation from the CTD and L8 homodimers. The blue line represents the absorbance at 280 nm and the red line represents the concentration of KCl (mM). The first 11% SDS-polyacrylamide gel shows (from left to right): 1) molecular weight marker (M); 2) pooled fractions from the nickel column after filtration (loading); 3) protein that flowed through the column after loading (flow through). The second 11% SDS-polyacrylamide gel shows (from left to right): 1) molecular weight marker (M); 2-4) L8-CTD heterodimer (peak 1); 5-6) L8 alone (peak 2).

3.3) The L8-CTD heterodimer has endonuclease activity

Once the L8-CTD heterodimer was obtained, the MutL endonuclease functionality was evaluated in comparison to the FL and CTD homodimers (Fig 3.4). Endonuclease activity of the CTD in the presence of β using a 200 bp linear substrate was previous observed in Pillon et al., (2015). Initial endonuclease activities were tested with a substrate of similar length (195bp) generated from pUC19 with an SKN1 site (GTCAT) in the middle. Consistent with previous studies, L8-CTD heterodimer had endonuclease activity in the presence of the β-clamp. It has comparable activity to the FL homodimer and greater endonuclease activity than the CTD homodimer. Both the FL homodimer and the L8-CTD heterodimer seem to accumulate a greater amount of short DNA fragments due to the contributions of both the DNA binding domain and interaction with the β-clamp (Fig 3.4, band boxed in red) compared to the CTD homodimer which only interacts with the β-clamp.
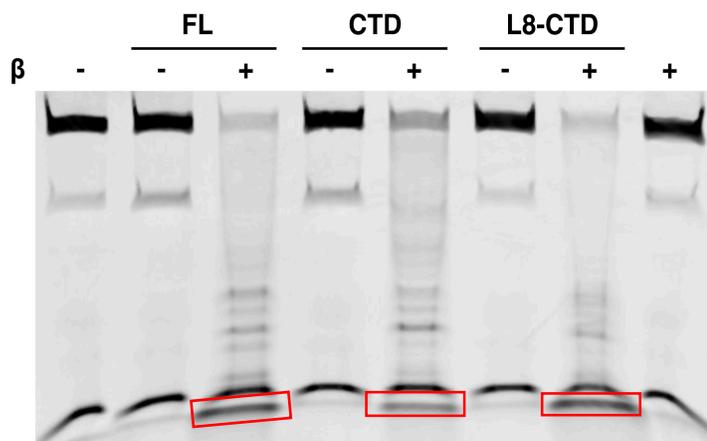
Fig 3.4 Endonuclease activity of L8-CTD heterodimer to ensure functionality. Cleavage products were separated through an 8%, 8M urea denaturing polyacrylamide gel. 1) 5' fluorescein labeled 195 bp linear DNA substrate alone; 2-7) Full length MutL, MutL C-terminal domain, and L8-CTD heterodimer (240 nM) were incubated with a 5' fluorescein labeled 195 bp linear substrate (10 nM) in the absence and presence of equimolar β-clamp; 8) β-clamp (240 nM) incubated with 5' fluorescein labeled 195 bp linear substrate (10 nM).

Since the SKN1 DNA-binding domain on the L8 chimera has sequence specificity to the GTCAT site in the middle of the substrate, we expected the pattern of cleavage products to differ when compared to the FL and CTD. In a DNA binding assay by another graduate in the lab, specific DNA binding by the SKN1 domain was only observed at protein concentrations of 10-80 nM (Ortiz Castro, 2016). Unfortunately, due to the transient interaction between MutL and the β-clamp, a much higher concentration of both proteins is required to ensure complex formation and observable nicking activity (Pillon et al., 2015). A second 195 bp linear substrate was generated from a different region of pUC19 without a (G/A)TCAT site to compare it with the first substrate (Fig 3.5). Between substrates, the major species produced by the nicking activity were different.

Between MutL dimers, the major species were the same. This implies the DNA binding domain does not seem to effect where MutL cleaves DNA but mainly contributes to increase the local concentration of substrate around the endonuclease site. The β-clamp may have a greater effect on dictating where MutL cleaves the substrate.
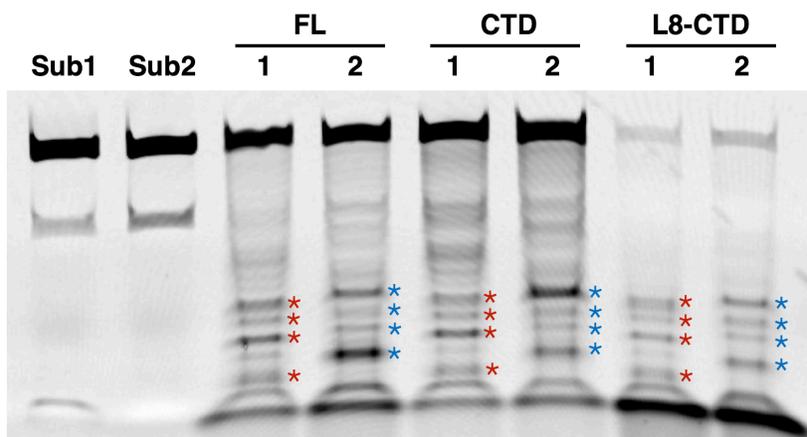
Fig 3.5 Endonuclease activity of MutL variants with two different substrates. Cleavage products were separated through an 8%, 8M urea denaturing polyacrylamide gel. 1-2) Two different 195 bp linear DNA substrates (Sub1 and Sub2) alone; 3-8) Full length MutL, MutL C-terminal domain, and L8-CTD heterodimer (240 nM) were incubated with 5' fluorescein labeled 195 bp linear substrate (Sub 1 or Sub2; 10 nM) in the presence of equimolar β-clamp. The red stars and blue stars correspond to the major species of Sub1 and Sub2 respectively.

3.4) Only one active site is required for BsMutL CTD endonuclease activity

After confirming the validity of the L8-CTD heterodimer, we then proceeded to selectively mutate each MutL C-terminal domain subunit while keeping both β-clamp binding motifs available to determine if *B. subtilis* MutL CTD required both functional active sites for endonuclease activity. L8-CTD heterodimer variants with only one functional active site were generated by introducing the point mutation E468K on the C-

terminal domain which abrogates endonuclease activity (Pillon et al. 2010). Regardless of which subunit possessed the mutation, the L8-CTD heterodimer still retained endonuclease activity even with only one functional active site (Fig 3.6A). The L8-CTD variant with the functional active site on the L8 subunit seemed to have greater endonuclease activity than having the functional active site on the CTD subunit, as indicated by the major accumulation of short cleavage products. The SKN1 may have a stimulatory effect on the CTD subunit it is linked to by increasing the local concentration of substrate for interaction with the β-clamp. Additionally, the SKN1 domain may also be sequestering away substrate from interacting with the β-clamp and the endonuclease site on the other subunit.

3.5) The β-clamp stimulates nicking activity of one endonuclease site

Although the MutL CTD homodimer has two β-clamp binding motifs, studies have shown that only one protomer on the β-clamp interacts with one monomer of BsMutL CTD homodimer while in complex (Pillon et al. 2015). In the previous endonuclease assay, the β-clamp can freely interact with either protomer of the L8-CTD heterodimer. We were interested to see whether the location of the β-clamp, in relation to the functional active site, would alter the endonuclease activity (Fig 3.6B). We first used variants of L8-CTD lacking the β-clamp binding motif (^{487}QEMIVP ⟶ ^{487}AEMAAP) and a nonfunctional active site (E468K) on the same protomer (L8^{Iø}-CTD/ L8-CTD^{Iø}). In this heterodimer, the β-clamp is interacting proximal to the functional active site and we see endonuclease activity comparable to the control which has both functional active sites

but only one β-clamp binding motif (L8$^{\varnothing}$-CTD, L8-CTD$^{\varnothing}$). Next, to have the β-clamp interact distally to the functional active site, we used the same mutations however they are no longer on the same protomer (L8$^{\varnothing}$-CTD^{I}, L8^{I}-CTD$^{\varnothing}$). Having the β-clamp interact with the protomer without the functional active site seemed to reduce or eliminate the observed endonuclease activity. The weak endonuclease activity seen in the distal heterodimer with the functional active site on the L8 may be attributed to effects from the SKN1 domain to bring DNA close to the active site. However, it is the β-clamp that plays the major role in regulating the endonuclease activity of each active site in the MutL homodimer.
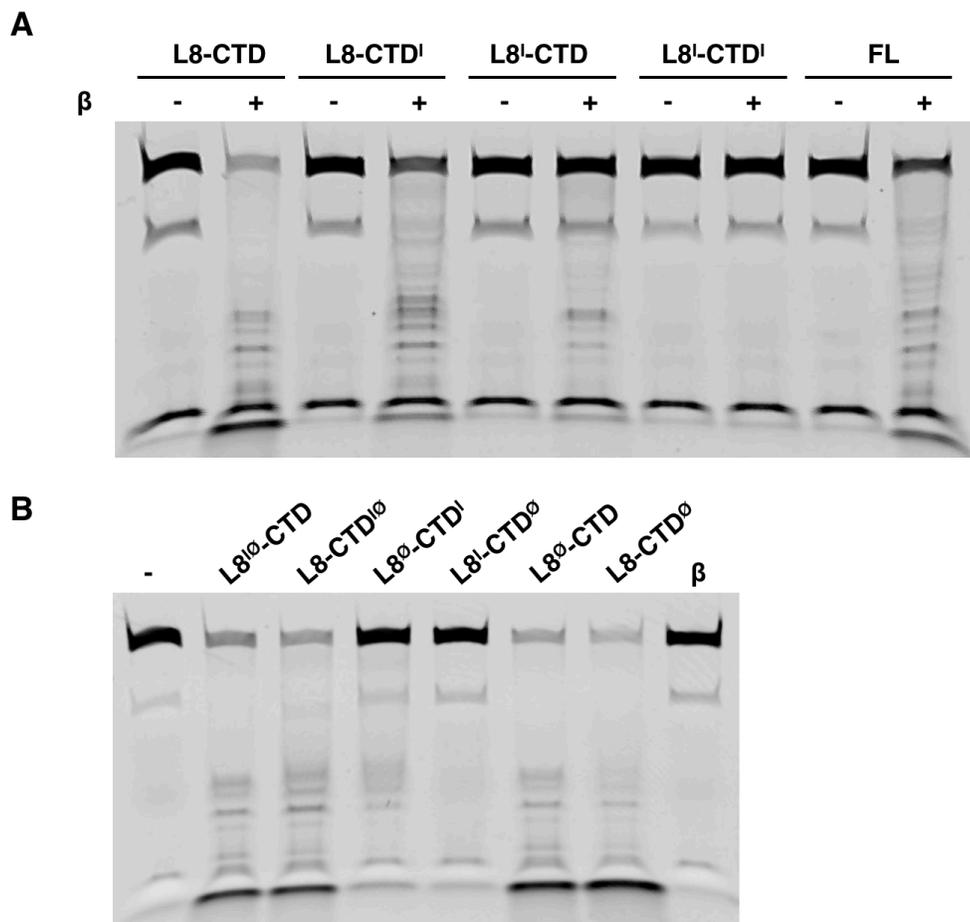
Fig 3.6. The β-clamp stimulates the proximal endonuclease active site. Cleavage products were separated through an 8%, 8M urea denaturing polyacrylamide gel. (A) Determining endonuclease activity of L8-CTD with one functional and one inactive (I) active site. 1-2) L8-CTD with two functional active sites; 3-6) L8-CTD^I and L8^I-CTD with only one functional active site; 7-8) L8^I-CTD^I with no functional active site; 9-10) Full length MutL for activity comparison (240 nM). MutL species were incubated with 5' fluorescein labeled 195 bp linear substrate (10 nM) in the absence and presence of equimolar β-clamp. (B) Endonuclease activity of L8-CTD heterodimers with a β-clamp binding motif mutation (ø) on the L8 or CTD subunit were assessed. 1) Linear 195 bp DNA substrate alone; 2-3) L8^{Iø}-CTD and L8-CTD^{Iø} heterodimers with functional active sites closest from the interacting β-clamp; 4-5) L8^ø-CTD^I and L8^I-CTD^ø with functional active sites furthest from the interacting β-clamp; 6-7) L8^ø-CTD and L8-CTD^ø with both functional active sites (240 nM). All MutL heterodimer variants were incubated with 5' fluorescein labeled 195 bp linear substrate (10 nM) in the presence of equimolar β-clamp; Lane 8) β-clamp alone with substrate.

Only one subunit of the *B. subtilis* MutL C-terminal domain needs to be interacting with the β-clamp for endonuclease activity. To ensure that mutation of one β-clamp binding motif did not affect binding affinity to the MutL C-terminal domain, the endonuclease activity of each β-clamp mutant L8-CTD heterodimer variant was assessed with a 1:1 and 1:2, L8-CTD variant:β molar ratio (Fig 3.7). L8-CTD heterodimer variants' endonuclease activity did not change much between having equimolar or two times excess of β.
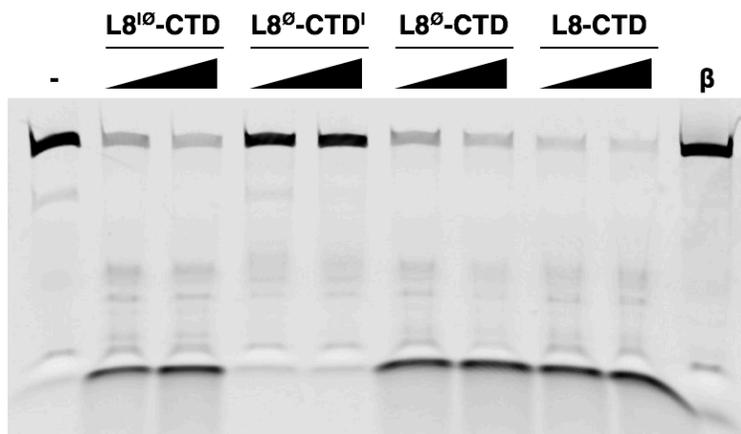
Fig 3.7 Mutating one β-clamp binding motif does not affect β binding affinity. Endonuclease activity of L8-CTD heterodimer variants with a β-clamp binding motif mutation on the L8 subunit at increasing concentrations of β. Cleavage products were separated through an 8%, 8M urea denaturing polyacrylamide gel. 1) Linear 195 bp DNA substrate alone; 2-3) $L8^{I-}$-CTD with a functional active site proximal to the interacting β-clamp (240nM); 4-5) $L8^{-}$-CTD^{I} with a functional active sites distal to the interacting β-clamp (480 nM); 6-7) $L8^{-}$CTD with both functional active sites (480 nM). MutL heterodimer variants were incubated with 5' fluorescein labeled 195 bp linear substrate (10 nM) in the presence of equimolar or double the β-clamp (240 – 480 nM); Lane 8) β-clamp alone with substrate (480 nM).

3.6) β-DNA complex formation through disulfide crosslinking is not efficient

Rather than relying on the protein's innate interactions, we used different crosslinking methods to generate stable binary complexes. Either with β and DNA or β with CTD because we know approximately where these two components interact with each other. The third component can be subsequently added to the binary complex during future studies in solving a crystal structure of the ternary complex (CTD-DNA- β).

In the presence of linear DNA, the processivity clamp will simply associate with DNA on one end and then slide off the other end. To overcome this problem, we covalently tethered the β-clamp to DNA using a disulfide crosslinking method. A cysteine residue was generated on the surface of the clamp, close to the central cavity for disulfide mediated crosslinking with a thiol-modified DNA substrate. Crosslinking DNA to β will hopefully prevent the substrate from going through the clamp in reverse orientation due to crystal packing as found in the structure of *E. coli* β bound to a 10/14-mer primed site (Georgescu et al, 2008). It will also help ensure the presence of DNA within the complex after the addition of the MutL CTD.

A cysteine variant of *B. subtilis* β-clamp with mutation D218C was generated to covalently crosslink with a 5' thiol modified 15/19mer primed substrate through the formation of a disulfide bridge when the two components interact (Figure 3.8). β^{D128C} was incubated with DNA substrate in the absence of reducing agents. The sample was resolved through an SDS-polyacrylamide gel and the appearance of a minor new species at ~65 kDa was observed after day 4 of incubation. The new species had a molecular weight consistent with the β^{D128C} monomer with DNA complex. The major species which

accumulated during the 4 day incubation at $\sim$ 45 kDa could not be the β^{D128C} forming a crosslinked complex with a contaminant as the size is smaller than the β^{D128C} itself. Addition of a reducing agent causes both the major and minor bands to disappear, confirming that formation of these species was mediated by disulfide bond linkage.

Surprisingly, the native cysteine residue, thought to be buried within the β-clamp structure, also participated in the crosslinking reaction. We reason that the β-clamp is likely in equilibrium between the open and closed state which would allow for intramolecular crosslinking to occur between the native cysteine residue and the mutated cysteine residue. Subsequent C178A mutation was made to β^{D128C} ($\beta^{D128C, C178A}$). After incubating this new β-clamp cysteine variant with DNA for 4 days, the species at $\sim$45 kDa had disappeared. However, the efficiency of forming the β^{cysD}-DNA complex remained the same even in the absence of the native cysteine residue. Either the location of the mutated cysteine is not optimal for the formation of disulfide bonds with DNA or the thiol-modified DNA substrates may be dimerization together and reducing the supply of free substrates. Instead of using disulfide bonds, heterobifunctional crosslinkers which possess different reactive groups at either end can be used to minimize unwanted self-conjugation.
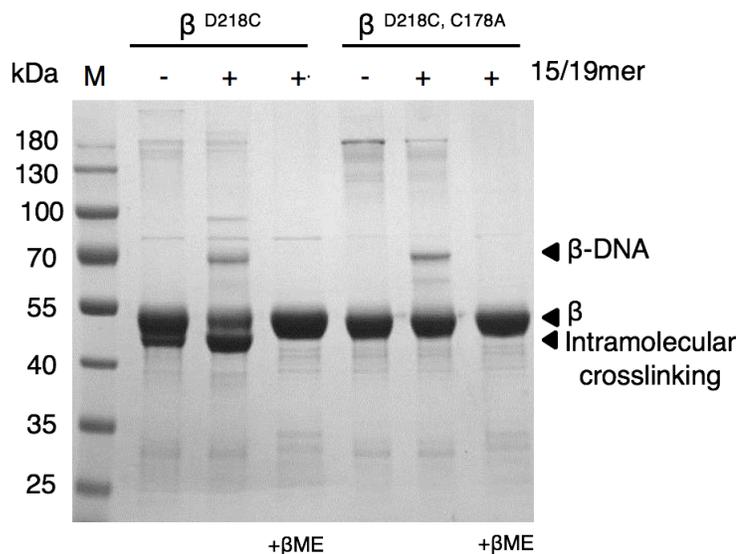
Figure 3.8 β^{cys} and DNA disulfide crosslinked complex formation. B^{cys}-15/19mer substrate complex formation on Day 4 of crosslinking reaction. 1) molecular weight marker (M); Cysteine modified *B. subtilis* β^{D218C} and $\beta^{D218C, C178A}$ were purified and complex formation was monitored 2,5) without DNA substrate and 3-4,6-7) with DNA substrate at a 1:1.2 (β:DNA) ratio. Samples were resolved on a 9% SDS PAGE denaturing gel in the absence and presence of β-mercaptoethanol.

3.7) β-DNA complex formation through heterobifunctional crosslinking

Instead of directly crosslinking the β-clamp to DNA, indirect complex formation can occur using a crosslinker that acts as a spacer arm. Heterobifunctional crosslinkers have two different reactive groups. Succinimidyl 4-(-N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) has an amine-reactive NHS-ester group on one end and a sulfhydrl reactive group on the other end. Amino-modified DNA was used as the ends do not react with each other, unlike thiol-modified DNA.

Amino- modified 15/19mer primed substrate were incubated with SMCC at a 1:10 ratio to form a DNA-SMCC conjugate followed by incubation with β^{cysD} at a 5:1, DNA: β^{cysD}, ratio (Figure 3.9). The sample was resolved through an SDS-polyacrylamide gel

and the appearance of a new species at ~65 kDa was observed. The new species had a molecular weight consistent with the β^{cysD} monomer with DNA complex. Ideally, the 1:1 ratio of β^{cysD} monomer to DNA-SMCC-β^{cysD} conjugate would confer to a β^{cysD} dimer with only one crosslinked DNA which would be able to thread through the central cavity of the sliding clamp. However, there may also be the presence of dimers with two conjugate DNA substrates, dimers with no substrates, and excess DNA. As well, the DNA substrate may have multiple interacting states with the β^{cysD} such as within the cavity or on the outer surface of the clamp. Future studies should focus on isolating β-clamp dimers conjugated to one DNA substrate.
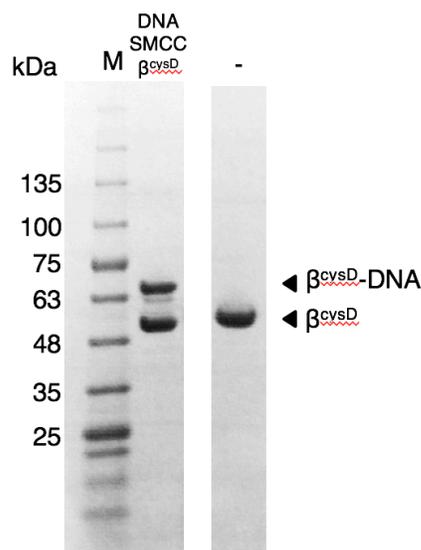
Figure 3.9 β^{cysD}-DNA heterobifunctional crosslinked complex formation. 1) molecular weight marker (M); 2) β^{cysD}-15/19mer substrate complex crosslinking reaction with a 1:10 DNA to succinimidyl 4-(-N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) ratio followed by reaction with β^{cysD} in a 5:1 (DNA: β^{cysD}) ratio; 3) β^{cysD} only lane taken from the same gel where the intervening lanes were cropped out. Samples were resolved on a pre-case gradient SDS PAGE denaturing gel (4-15%) in the absence of β-mercaptoethanol.

3.8) β- CTD complex formation through disulfide crosslinking

β-clamp with a cysteine residue and an extended C-terminal end (β^{cysL}) was incubated with a cysteine modified CTD (CTD^{cys}) in the absence of reducing agents to facilitate disulfide bond formation and crosslinking. Complex formation was resolved on an SDS-polyacrylamide gel and the appearance of a species at ~75 kDa was observed over a 6-day incubation period (Figure 3.10). This ~75 kDa molecular weight is consistent with a β^{cysL} monomer conjugated to a CTD^{cys} monomer. Formation of a higher MW species at around 100 kDa suggests a β^{cysL} –β^{cysL} crosslinked complex. Complex formation stopped around day 6.

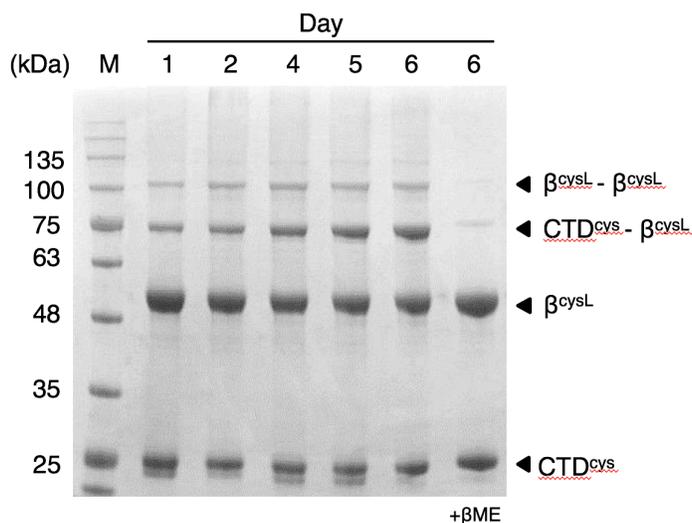**Figure 3.10 β^{cysL}-CTD^{cys} crosslinked complex formation.** Cysteine modified *B. subtilis* β and CTD were purified and equimolar amounts were incubated together in the absence of reducing agent. 1) molecular weight marker (M); 2-6) Samples withdrawn from the reaction at the indicated time points in the absence of βME ;7) Day 6 sample in the presence of βME. Samples were resolved on a 9% SDS PAGE denaturing gel.

3.9) β-CTD crosslinked complex has endonuclease activity

Endonuclease activity of the β^{cysL} –CTD^{cys} complex with a 195 bp substrate was assessed to verify that the crosslinking reaction did not eliminate the β-clamp stimulated endonuclease activity of MutL. First, we compared the endonuclease activity of β^{cysL} – CTD^{cys} crosslinked complex after 7 days of incubation with freshly thawed β^{cysL} and CTD^{cys} protein (Day 0) which have not yet had the chance to crosslink together (Figure 3.11A). Endonuclease activity of the β^{cysL} –CTD^{cys} crosslinked complex was lower than the non-crosslinked protein and was only reported in the higher concentrations of β^{cysL} – CTD^{cys} crosslinked complex. The lower nicking activity may be due to the presence of non-specific crosslinked complexs (CTD^{cys}-CTD^{cys} and β^{cysL}-β^{cysL}) found over the course of forming the CTD^{cys}-β^{cysL} complex may reduce the concentration of functional CTD^{cys}. We also can't rule out the possibility that the nicking activity may be caused or enhanced by any remaining free CTD^{cys} and β^{cysL} in the reaction. However, by day 6, the majority of CTD^{cys} and β^{cysL} seem to have formed a complex (specific or nonspecific).

Next, we wanted to see if endonuclease activity was observable using a shorter, 19/19mer primed, substrate (Figure 3.11B). Nicking activity was observed using freshly thawed β^{cysL} and CTD^{cys} protein (β^{cysL} + CTD^{cys} , Day 0). Little to no activity was seen using β^{cysL} and CTD^{cys} protein which have been separately left at 4°C for 4 days (β^{cysL} + CTD^{cys} ,Day 4). This confirms that, at least after 4 days, there are no remaining free CTD^{cys} and β^{cysL} remaining that could contribute to the observable endonuclease activity of CTD^{cys}-β^{cysL}. However, no activity was observed with the β^{cysL} –CTD^{cys} crosslinked complex after 4 days (β^{cysL} - CTD^{cys} ,Day 4). A shorter substrate was used in this assay

which may have also impacted the observable activity. The added rigidity from crosslinking the complex may have impeded proper coordination of β^{cysL}, CTD^{cys} and DNA, preventing endonuclease activity.
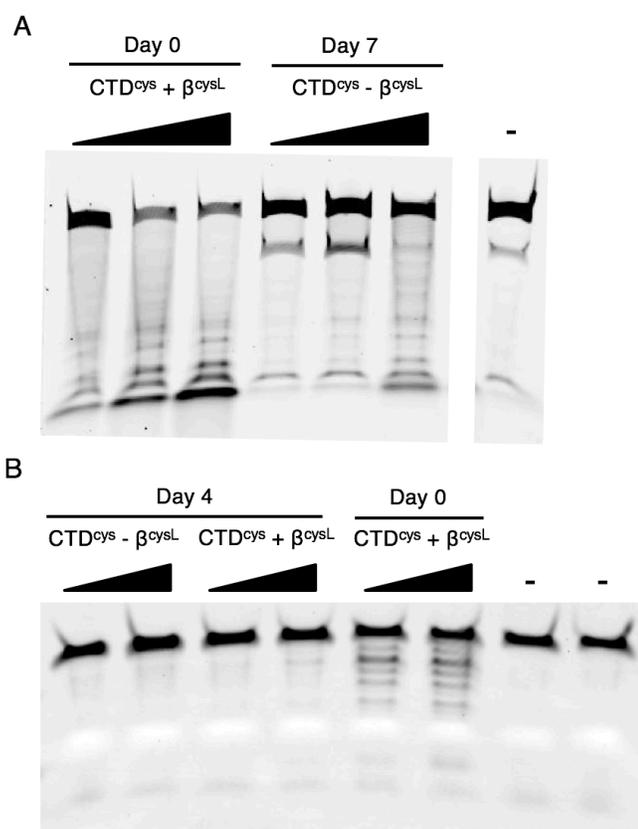
Figure 3.11 Endonuclease activity of β^{cysL}-CTD^{cys} complex. (A)195 bp linear substrate (10 nM) incubated with 1-3) Day 0, freshly thawed, *B. subtilis* CTD^{cys} mixed with equimolar β^{cysL} and 4-6) 7 day crosslinked CTD^{cys}- β^{cysL} (1 – 3.2 μM); 7) DNA only lane taken from the same gel where the intervening lanes were cropped out. (B) 19/19mer substrate with 4 base overhangs incubated with 1-2) Day 4 crosslinked *B. subtilis* CTD^{cys}-β^{cysL} (1.3 – 2.6 μM); 3-4) Day 4 CTD^{cys} mixed with equimolar Day 4 β^{cysL}; 5-6) Day 0, newly thawed, CTD^{cys} mixed with equimolar Day 0 β^{cysL}.

CHAPTER 4

DISCUSSION

We have shown that prokaryotic MutL homodimers primarily use only one of its two active sites for endonuclease activity which is regulated and stimulated by the interacting proximal β-clamp. This parallels the endonuclease activity observed in eukaryotic MutL heterodimers where both the active site and PCNA interacting motif are located on the same protomer (yPMS1/hPMS2) (Kadyrov et al., 2006; Kosinski et al., 2008; Gueneau et al., 2013). This reinforces the idea that MutH-less organisms share a common endonuclease mechanism.

Although the MutL C-terminal domain has low sequence conservation among species, the motifs composing the endonuclease active site in *B. subtilis* MutL are identical to the sites found in *S. cerevisiae* MutL homolog PMS1 and in human MutL homolog PMS2 (Kadyrov et al., 2006; Pillon et al., 2010; Gueneau et al., 2013). Incidentally in both eukaryotes and prokaryotes, the subunit that binds to the processivity clamp is the one that dominates cleavage activity. It was unsurprising that the β-clamp only interacts with one MutL subunit as the structural model of the *B. subtilis* MutL in complex with the β-clamp shows only one MutL subunit binds to the β-clamp ring (Pillon et al. 2015). The presence of one interacting β-clamp on the CTD occludes the binding of a second β-clamp due to steric hindrance (Pillon et al., 2010). Recent studies have found the conserved QXX[L/I]XP motif, which was identified as the β-clamp binding motif in *B. subtilis* MutL, to also function as an essential motif on both *S. cerevisiae* PMS1 and human PMS2 for PCNA interaction and endonuclease stimulation (Kosinski et al. 2008;

Pillon, et al., 2011; Genschel et al., 2017). This implicitly suggests that the molecular mechanism of the MutL endonuclease activity in eukaryotes and prokaryotes would be universal.

Asymmetry of the MutL homodimer for nicking activity is reminiscent of the asymmetry of MutS homodimer in its mismatch recognition state during DNA repair. Both homodimers have been replaced in the eukaryotic system with multiple MutS and MutL homologs which combine to form heterodimers. The functional diversification and specialization of MMR genes to recognize specific types of DNA errors and acquire roles in meiotic recombination may have contributed to the evolution of eukaryotes.

It is enticing to speculate that metal ion binding may also play a role in maintaining the functional asymmetry in MutL homodimers. *B. subtilis* MutL endonuclease activity is manganese dependent and the addition of zinc strongly stimulates nicking activity however both metal ions bind weakly. The crystal structure of the *B. subtilis* C-terminal domain did not contain Zn^{2+} unless the ion was supplemented in the crystallization conditions and Mn^{2+} was not found despite its importance (Pillon et al., 2010). This may allow for situations were only one protomer is bound to Zn^{2+} and Mn^{2+} thus allowing only one endonuclease active site of the dimer to be in the active form.

As previous studies have shown, *B. subtilis* MutL C-terminal domain does not exhibit endonuclease activity because it does not bind DNA (Pillon et al., 2010). This DNA binding defect acts as a powerful regulatory mechanism to prevent indiscriminate nicking of DNA. The β-clamp greatly stimulates MutL nicking activity, presumably by bypassing the DNA binding defect (Pillon et al., 2015). Similarly, we see the same effect

with human PCNA on MutLα (Kadyrov et al., 2006; Kadyrov et al., 2007). The engineered *B. subtilis* MutL heterodimer allowed us to test whether enhancing nucleotide binding to MutL by fusing a domain with high DNA binding affinity to one CTD subunit would induce nicking activity. Despite anchoring the DNA close to the CTD, there was barely any observable nicking activity without the presence of the β-clamp. Any nicked product in the absence of the β-clamp may be a result of collisions between the DNA substrate and the CTD active site due to the high concentration of enzyme within the reaction. This result shows that simply bringing the DNA substrate within the vicinity of the active site is insufficient in stimulating MutL.

The effects of the SKN1 DNA binding domain on the *B. subtilis* MutL heterodimer were most noticeable when comparing the L8-CTD heterodimer with the CTD homodimer. A large accumulation of short DNA degradation products was observed in the presence of the DNA binding domain. However, it was not possible to see SKN1 specific binding of DNA for the substrate with the GTCAT site because at the concentration of MutL variants used in endonuclease assays (240 nM) for observable nicking activity, the SKN1 domain had extensive non-specific DNA binding activity (Ortiz Castro, 2016). This explains the appearance of a similar pattern of specific degradation products between MutL variants with and without the SKN1 domain.

The individual effects of SKN1 on each subunit of the L8-CTD heterodimer were masked when both active sites and β-clamp binding motifs were functional. Minor effects were noticeable when comparing heterodimers with active site mutations on either the CTD or the L8. When comparing reciprocal variants, the heterodimer with a functional

active site on the L8 subunit has more nicking activity than when the functional active site is on the CTD subunit. We expected that only the MutL variants with the β-clamp interacting proximal to the functional active site would have endonuclease activity. However, nicking activity was observed for the L8$^{\text{ø}}$-CTD$^{\text{I}}$ variant where the endonuclease site from the L8 subunit remained active and the β-clamp interacted in a distal manner on the CTD subunit. Conversely, as the reciprocal variant, L8$^{\text{I}}$-CTD$^{\text{ø}}$, exhibited no endonuclease activity, we attributed the observed nicking activity on the distal heterodimer variant to be an effect of the SKN1 domain sequestering DNA.

The observation of nicking activity with the distal heterodimer L8$^{\text{ø}}$-CTD$^{\text{I}}$ gave confirmation that the heterodimer was not in equilibrium with its respective homodimers (L8$^{\text{ø}}$- L8$^{\text{ø}}$ and CTD$^{\text{I}}$- CTD$^{\text{I}}$). Neither the homodimers would have exhibited observable endonuclease activity without β-clamp interaction or functional active sites. As well, the instability of the L8 homodimers would have resulted in rapid precipitation of half the protein within the reaction, which in this case did not occur.

The SKN1 domain may function similarly to the MutL N-terminal domain by increasing the frequency of MutL bound to DNA (Ban and Yang, 1998; Hall et al., 2003; Junop et al., 2003; Pillon et al., 2015). However, it is the β-clamp that remains as the dominant factor which stimulates the MutL endonuclease activity. It has been proposed that the β-clamp stimulates endonuclease activity by coordinating the proper relative orientation between DNA and the active site through its interaction with the β-clamp binding motif of MutL CTD (Pillon et al., 2015). Threading DNA directly into the active site by PCNA is also the presumed mechanism for stimulating the 5' flap endonuclease

Fen-1 (Craggs et al., 2014). The β-clamp may also neutralize the negatively charged surface around the active site which prevents DNA from unintentionally entering (Pillon et al., 2010; Fukui et al. 2016).

The processivity clamp interacts transiently with MutL to promote nicking activity (Pluciennik et al., 2010; Pillon et al, 2011; Pillon et al., 2015). Different strengths of protein-protein interactions with the β-clamp have been correlated with variability in the binding motif sequence (Pillon et al., 2011; Maga and Hubscher, 2003; Yin et al., 2013; Rolef et al., 2009). In comparison, the PCNA interacting motif has a strictly defined sequence however hydrophobic packing has been attributed to different interaction strengths (Bruning and Shamoo, 2004). The processivity clamp plays an essential role in ensuring faithful and complete DNA replication by orchestrating the coordination of multiple enzymes and regulatory factors in a wide variety of DNA processing mechanisms such as DNA replication and DNA damage response (Moldovan et al., 2007). With so many binding partners, the clamp utilizes both strong and weak interactions to strictly regulate protein enzymatic functions occur during the proper cellular event. The K_d value for the MutL- β-clamp interaction is within the range of weak protein-protein affinities ($K_d > 10^{-4}$). In this case, the transient β-clamp-MutL interaction is thought to act as a regulatory mechanism to prevent excess nicking of the nascent strand.

MutL is a sequence unspecific endonuclease, capable of nicking DNA at any point of the newly synthesized nascent strand. Different resulting nicking patterns were observed when comparing two sequentially different 195 bp linear substrates. One possibility may be attributed to secondary structures generated on the substrates during

annealing of the two oligomeric strands. Once the β-clamp encounters a secondary structure, the clamp may stall which gives enough time for MutL interaction and stimulation of the endonuclease activity. MutL may nick past the secondary structure, releasing the DNA for subsequent clamp loading and MutL nicking. Accumulation of specific DNA fragment lengths was unlikely sequence specific as the 195 bp substrate is not palindromic yet the major products observed when the duplex was labeled either on the top or bottom strand were the same (Ortiz Castro, 2016).

The molecular mechanism of how the β-clamp regulates MutL nicking activity remains elusive. Physical interaction between MutL and the β-clamp allows for complex formation through disulfide crosslinking. Pillon et al., (2015), showed that cysteine modified variants of MutL and the β-clamp were both active *in vivo* and *in vitro*. Although activity of the crosslinked product was observed with the 195 bp substrate, none was seen with the 19/19mer substrate with 4 base overhangs. However, non-crosslinked cysteine modified MutL and the β-clamp could cleave the shorter substrate. The added rigidity from the crosslinking may have prevented proper MutL-β-clamp coordination, resulting in the 19/19mer substrate being too short to thread through both the β-clamp central cavity and the MutL active site. This finding reinforces the importance of the interaction between MutL and the β-clamp, despite the transient nature of the complex, and emphasizes the need to understand the regulatory mechanism imposed by the β-clamp.

The strategy of using disulfide crosslinking for complex formation of the β-clamp with DNA was unsuccessful because of the low efficiency of β-clamp-DNA complex

formation. This might occur if the thiol-modified DNA self-conjugate into dimers, reducing the amount of free substrate available for crosslinking with the β-clamp. Heterobifunctional crosslinkers used to indirectly connect a cysteine modified β-clamp with an amino-modified DNA substrate were observed to form a complex of one DNA substrate to one β-clamp ring. Once a single complex species can be isolated, this work can lead to determine if the β-clamp is still functional to stimulate MutL endonuclease activity to cleave the crosslinked DNA.

4.1) Conclusion and future direction

Engineering the MutL chimer to assemble prokaryotic MutL heterodimers allowed for individual mutations of each subunit. Beyond mismatch repair, this may be used for biochemical characterization of other proteins which are obligatory dimers, but function asymmetrically. Despite being in different species, the endonuclease domains in both prokaryotic MutL which do not rely on MutH and d(GATC) methylation for MMR and eukaryotic MutL are composed of the same conserved motifs. We have shown that the *B. Subtilis* MutL C-terminal domain acts similarly to eukaryotic MutL heterodimers where only one active site is stimulated by the processivity clamp for nicking activity. The DNA binding defect of the CTD cannot be overcome by the addition of a DNA binding domain. Endonuclease activity remains β-clamp dependent as the clamp may directly orient DNA within the active site. This work brings further validation that future biochemical characterization of prokaryotic MutL may be applied to eukaryotic MutL.

The long-term goal of this project is to obtain a mechanistic understanding of the

MutL endonuclease activity through structural studies of the ternary complex containing the MutL CTD, the β-clamp, and DNA. Recent studies have utilized crosslinking methods in conjunction with X-ray crystallography or SAXS to study the structural organization of transient protein partners (Pillon et al., 2015; Groothuizen et al., 2015). To observe a ternary complex, crosslinked complexes (CTD-β or β-DNA) with the addition of the third component (DNA or CTD) may be used for future crystallization experiments. By obtaining a detailed view of the spatial organization of this complex, an accurate model of how MutL and the β-clamp orients on DNA will provide insight to how these transient interactions respond in mismatch repair. The model can be further probed through structure-guided mutagenesis and biochemical characterization of the resulting mutants. This lays the foundation for future work which will unravel a novel endonuclease mechanism and bring a better understanding to the evolutionarily conserved mismatch repair pathway.

REFERENCES

Ban, C., Junop, M., & Yang, W. (1999). Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell*, 97, 85–97.

Ban, C., & Yang, W. (1998). Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, 95, 541–552.

Barras, F., & Marinus, M.G. (1989). The great GATC: DNA Methylation in E.coli. *Trends in Genetics,* 5(5), 139-143.

Bruning, J.B. & Shamoo, Y. (2004). Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and ap endonuclease-1. *Structure*, 12, 2209–2219.

Correa, E.M.E., De Tullio, L., Velez, P.S., Martina, M.A., Argarana, C.E., & Barra, J.L. (2013). Analysis of DNA structure and sequence requirements for *Pseudomonas aeruginosa* MutL endonuclease activity. *Journal of Biochemistry*, 154(6), 505–511.

Craggs, T.D., Hutton, R.D., Brenlla, A., White, M.F., & Penedo, J.C. (2014). Single-molecule characterization of Fen1 and Fen1/PCNA complexes acting on flap substrates. *Nucleic Acids Research,* 42(3), 1857–1872.

Dalrymple, B.P., Kongsuwan, K., Wijffels, G., Dixon, N.E., & Jennings, P. (2001). A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 11627–11632.

Dinh, T.A., Rosner, B.I., Atwood, J.C., Boland, J.R., Syngal, S., & Vasen, H.F. (2011). Health benefits and cost-effectiveness of primary genetic screening for Lynch syndrome in the general population. *Cancer Prevention Research*, 4, 9–22.

Duppatla, V., Bodda, C., Urbanke, C., Friedhoff, P., & Rao, D.N. (2009). The C-terminal domain is sufficient for endonuclease activity of *Neisseria gonorrhoeae* MutL. *Biochemical Journal*, 423, 265–277.

Dzantiev, L., Constantin, N., Genschel, J., Iyer, R.R., Burgers, P.M., & Modrich, P. (2004). A defined human system that supports bidirectional mismatch-provoked excision. *Molecular Cell,* 15, 31-41.

Elez, M., Radman, M., & Matic, I. (2012). Stoichiometry of MutS and MutL at unrepaired mismatches in vivo suggests a mechanism of repair. *Nucleic Acids Research*, 40, 3929–3938.

Fijalkowska, I.J., Schaaper, R.M., & Jonczyk, P. (2012). DNA replication fidelity in Escherichia coli: a multi-DNA polymerase affair. *FEMS Microbiology Reviews*, 36, 1105-1121.

Fukui, K., Baba, S., Kumasaka, T., & Yano, T. (2016). Structural features and functional dependency on β-clamp define distinct sub-families of bacterial mismatch repair endonuclease MutL. *Journal of Biological Chemistry,* 291(33),16990-17000.

Fukui, K., Nishida, M., Nakagawa, N., Masui, R., & Kuramitsu, S. (2008). Bound nucleotide controls the endonuclease activity of mismatch repair enzyme MutL. *Journal of Biological Chemistry*, 283(18), 12136–45.

Geier, G.E., & Modrich, P. (1979). Recognition squence of the dam methylase of Escherichia coli K12 and mode of cleavage of Dpn I endonuclease. *Journal of Biological Chemistry*, 254(4), 1408-1413.

Genschel, J., Kadyrova, L.Y., Lyer, R.R., Dahal, B.K., Kadyrov, F.A., & Modrich, P. (2017). Interaction of proliferating cell nuclear antigen with PMS2 is required for MutLα activation and function in mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America*, 114(19), 4930-4935.

Ghodgaonkar, M.M., Lazzaro, F., Olivera-Pimentel, M., Artola-Boran, M., Cejka, P., Reijns, M.A., Jackson, A.P., Plevani, P., Muzi-Falconi, M., & Jiricny, J. (2013). Ribonucleotides Misincorporated into DNA Act as Strand-Discrimination Signals in Eukaryotic Mismatch Repair. *Molecular cell,* 50, 323-332.

Gorman, J., Wang, F., Redding, S., Plys, A. J., Fazio, T., Wind, S., Alani, E.E., & Greene, E. C. (2012). Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America*, 109(45), E3074–83.

Gradia, S., Acharya, S., & Fishel, R. (1997). The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch. *Cell*, 91, 995–1005.

Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J., & Fishel, R. (1999). hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Molecular Cell,* 3, 255–261.

Groothuizen, F.S., Winkler, I., Cristóvão, M., Fish, A., Winterwerp, H.H.K., Reumer, A., Marx, A.D., Hermans, N., Nicholls, R.A., Murshudov, G.N., Lebbink, J.H.G., Friedhoff, P., & Sixma, T.K. (2015). MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *eLife*, 4, 1–24.

Georgescu, R.E., Kim, S.S., Yurieva, O., Kuriyan, J., Kong, X.P., & O'Donnell, M. (2008). Structure of a sliding clamp on DNA. *Cell*, 132, 43-54.

Gomez, M.J., Diaz-Maldonado, H., Gonzalez-Tortuero, E., & Lopez de Saro, F.J. (2014). Chromosomal replication dynamics and interaction with the beta sliding clamp determine orientation of bacterial transposable elements. *Genome biology and evolution*, 6, 727-740.

Graham, J.E., Marians, K.J., & Kowalczykowski, S.C. (2017). Independent and stochastic action of DNA polymerase in the replisome. *Cell*, 169, 1201-1213.

Guarné, A., & Charbonnier, J.B. (2015). Insights from a decade of biophysical studies on MutL: Roles in strand discrimination and mismatch removal. *Progress in Biophysics and Molecular Biology*, 117, 149–156.

Guarné, A., Ramon-Maiques, S., Wolff, E. M., Ghirlando, R., Hu, X., Miller, J. H., & Yang, W. (2004). Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO Journal*, *23*(21), 4134–4145.

Gueneau, E., Dherin, C., Legrand, P., Tellier-Lebegue, C., Gilquin, B., Bonnesoeur, P., Londino, F., Quemener, C., Le Du, M. H., Marquez, J.A., Moutiez, M., Gondry, M., Boiteux, S., & Charbonnier, J.B. (2013). Structure of the MutLα C-terminal domain reveals how MLH1 contributes to PMS1 endonuclease site. *Nature Structural & Molecular Biology*, 20(4), 461–8.

Gulbis, J.M., Kelman, A., Hurwitz, J., O'Donnell, M., & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA, *Cell*, 87, 297–306.

Gupta, S., Gellert, M., & Yang, W. (2012). Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nature Structural & Molecular Biology*, 19(1), 72–8.

Hall, M.C., Shcherbakova, P.V., & Kunkel, T.A. (2002). Differential ATP binding and intrinsic ATP hydrolysis by amino-terminal domains of the yeast MLH1 and PMS1 proteins. *Journal of Biological Chemistry*, 277(5), 3673–3679.

Heltzel, J.M., Maul, R.W., Scouten Ponticelli, S.K., & Sutton, M.D. (2009). A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 12664-12669.

Hermans, N., Laffeber, C., Cristovao, M., Artola-Baron, M., Mardenborough, Y., Ikpa, P. Jaddoe, A., Winterwerp, H., Wyman, C., Jiricny, J., Kanaae, R., Friedhoff, P., & Lennink, J. (2016). Dual daughter strand incision is processive and increases the efficiency of DNA mismatch repair. *Nucleic Acids Research,* 44, 6770-6786.

Hombauer, H., Campbell, C.S., Smith, C.E., Desai, A., & Kolodner, R.D. (2011). Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell*, 147(5), 1040–1053.

Hubscher, U., Maga, G., & Spadari, S. (2002). Eukaryotic DNA polymerases. *Annual Review of Biochemistry*, 71, 133–163.

Indiani, C. & O'Donnell, M. (2006) The replication clamp-loading machine at work in the three domains of life. *Nature Reviews Molecular Cell Biology*, 7, 751–761.

Iyer, R.R., Pluciennik, A., Burdett, V., & Modrich, P.L. (2006). DNA mismatch repair: Functions and mechanisms. *Chemical Reviews*, *106*, 302–323.

Jeong, C., Cho, W.K., Song, K.M., Cook, C., Yoon, T.-Y., Ban, C., Fishel, R., & Lee, J.B. (2011). MutS switches between two fundamentally distinct clamps during mismatch repair. *Nature Structural & Molecular Biology*, 18(3), 379–385.

Jiricny, J. (2013). Postreplicative mismatch repair. *Cold Spring Harbor Perspectives in Biology*, 5, 1–23.

Junop, M.S., Yang, W., Funchain, P., Clendenin, W., & Miller, J.H. (2003). In vitro and in vivo studies of MutS, MutL and MutH mutants: Correlation of mismatch repair and DNA recombination. *DNA Repair*, 2, 387–405.

Kadyrov, F.A., Dzantiev, L., Constantin, N., & Modrich, P. (2006). Endonucleolytic Function of MutLα in Human Mismatch Repair. *Cell*, 126, 297–308.

Kadyrov, F.A., Holmes, S.F., Arana, M.E., Lukianova, O.A., O'Donnell, M., Kunkel, T.A., & Modrich, P. (2007). *Saccharomyces cerevisiae* MutLα is a mismatch repair endonuclease. *Journal of Biological Chemistry*, 282(51), 37181–37190.

Kelch, B.A., Makino, D. L., O'Donnell, M., & Kuriyan, J. (2012). Clamp loader ATPases and the evolution of DNA replication machinery. *BMC Biology*, 10, 34.

Kosinski, J., Plotz, G., Guarné, A., Bujnicki, J.M., & Friedhoff, P. (2008). The PMS2 subunit of human MutLα contains a metal ion binding domain of the iron-dependent repressor protein family. *Journal of Molecular Biology*, 382, 610–627.

Krishna, T.S., Kong, X.P., Gary, S., Burgers, P.M., & Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell,* 79(7),1233-43.

Kunkel, T.A. (2004). DNA replication fidelity. *Journal of Biological Chemistry,* 279,16895–16898.

Kunkel, T.A, & Erie, D.A. (2005). DNA mismatch repair. *Annual Review of Biochemistry*, 74, 681–710.

Lamers, M.H., Perrakis, A., Enzlin, J.H., Winterwerp, H.H., de Wind, N., & Sixma, T.K. (2000). The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*, 407, 711–7.

Lamers, M.H., Winterwerp, H.H.K., & Sixma, T.K. (2003). The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO Journal,* 22,746–56.

Lee, S.D., & Alani, E. (2006). Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *Journal of Molecular Biology*, 355, 175-184.

Lehmann, A.R. (2003). Replication of damaged DNA. *Cell cycle (Georgetown, Tex.)* 2, 300-302.

Lenhart, J.S., Pillon, M.C., Guarné, A., & Simmons, L.A. (2013). Trapping and visualizing intermediate steps in the mismatch repair pathway in vivo. *Molecular Microbiology*, 90, 680–698.

Lopez de Saro, F.J., & O'Donnell, M. (2001). Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. *Proceedings of the National Academy of Sciences of the United States of America,* 98, 8376–80.

Lopez de Saro, Marinus, M.G., Modrich, P., & O'Donnell, M. (2006). The β Sliding Clamp Binds to Multiple Sites within MutL and MutS. *The Journal of Biological Chemistry,* 281, 14340-14349.

Lujan, S.A., Williams, J.S., Clausen, A.R., Clark, A.B., & Kunkel, T.A. (2013). Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Molecular Cell,* 50, 437-433.

Lynch, H.T., Lynch, J.F., & Attard, T.A. (2009). Diagnosis and management of hereditary colorectal cancer syndromes: Lynch syndrome as a model. *Canadian Medical Association journal,* 181, 273- 280.

Maga, G. & Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *Journal of Cell Science*, 116, 3051–3060.

Maki, H. & Kornberg, H. (1985). The polymerase subunit of DNA polymerase III of *Escherichia coli*. II. purification of the alpha subunit, devoid of nuclease activities. *The Journal of Biological Chemistry,* 260, 12987–12992.

Mauris, J., & Evans, T.C. (2009). Adenosine triphosphate stimulates *Aquifex aeolicus* MutL endonuclease activity. *PLoS ONE*, *4*(9), e7175.

Meselson, M., & Stahl, F.W. (1958). The Replication of DNA in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 44, 671–682.

McCulloch, S.D, Gu, L., & Li, G.M. (2003). Nick-dependent and -independent processing of large DNA loops in human cells. *Journal of Biological Chemistry,* 278(50), 50803–50809.

Modrich, P., & Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual Review of Biochemistry*, 65, 101–133.

Moldovan, G.L., Pfander, B., & Jentsch, S. (2007). PCNA, the maestro of the replication fork. *Cell*, 129(4), 665-79.

Natrajan, G., Lamers, M.H., Enzlin, J.H., Winterwerp, H.H.K., Perrakis, A., & Sixma, T.K. (2003). Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: A common recognition mode for diverse substrates. *Nucleic Acids Research*, 31(16), 4814–4821.

Obmolova, G., Ban, C., Hsieh, P., & Yang, W. (2000). Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, 407, 703–10.

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., & Sugino, A. (1968). Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proceedings of the National Academy of Sciences of the United States of America*, 59, 598–605.

Ortiz Castro, M. (2016). Characterizing the interactions of ATP and DNA with the MutL Mismatch Repair protein. (Master's thesis, McMaster University). Retrieved from http://hdl.handle.net/11375/20279.

Parks, A.R., Li, Z., Shi, Q., Owens, R.M., Jin, M.M., & Peters, J.E. (2009). Transposition into replicating DNA occurs through interaction with the processivity factor. *Cell*, 138, 685-695.

Pavlov, Y., Mian, I., & Kunkel, T.A. (2003). Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Current Biology*, 13, 744–748.

Peltomäki, P. (2001). Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Human Molecular Genetics,* 10(7), 735-740.

Peltomäki, P. (2003). Role of DNA mismatch repair defects in the pathogenesis of human cancer. *Journal of Clinical Oncology*, 21(6), 1174–1179.

Pillon, M.C., Babu, V.M.P., Randall, J.R., Cai, J., Simmons, L.A., Sutton, M.D., & Guarné, A. (2015). The sliding clamp tethers the endonuclease domain of MutL to DNA. *Nucleic Acids Research*, 9140(9), 1-14.

Pillon, M.C., Lorenowicz, J.J., Uckelmann, M., Klocko, A.D., Mitchell, R.R., Chung, Y.S., Modrich, P., Walker, G.C., Simmons, L.A., Friedhoff, P., & Guarné, A. (2010). Structure of the endonuclease domain of MutL: unlicensed to cut. *Molecular Cell*, 39, 145–151.

Pillon, M.C., Miller, J.H., & Guarné, A. (2011). The endonuclease domain of MutL interacts with the β sliding clamp. *DNA Repair*, 10(1), 87–93.

Pluciennik, A., Dzantiev, L., Iyer, R. R., Constantin, N., Kadyrov, F.A, & Modrich, P. (2010). PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 16066–16071.

Prelich, G, Tan, C.K., Kostura, M., Matthews, M.B., So, A.G., Downey, K.M., & Stillman, B. (1987). Functional identity of proliferating cell nuclear antigen and a DNA polymerase- delta auxiliary protein. *Nature.* 236, 517-520.

Qiu, R., Sakato, M., Sacho, E. J., Wilkins, H., Zhang, X., Modrich, P., Hingorani, M.M., Erie, D.A., & Weninger, K.R. (2015). MutL traps MutS at a DNA mismatch. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 10914–9.

Räschle, M., Marra, G., Nyström-Lahti, M., Schär, P., & Jiricny, J. (1999). Identification of hMutLβ, a heterodimer of hMLH1 and hPMS1. *Journal of Biological Chemistry*, 274(45), 32368–32375.

Rashev M., Surtees, J.A., & Guarné, A. (2017). Large-scale production of recombinant Saw1 in Escherichia coli. *Protein Expression and Purification*, 133, *75-80.*

Reha-Krantz, L. J. (2010). DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1804(5), 1049–1063.

Reijns, M.A.M., Rabe, B., Rigby, R.E., Mill, P., Astell, K.R., Lettice, L.A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., Devenney, P. S., Sexton, D., Grimes, G., Holt, I.J., Hill, R.E., Taylor, M.S., Lawson, K.A., Dorin, J.R., & Jackson, A.P. (2012). Enzymatic Removal of Ribonucleotides from DNA Is Essential for Mammalian Genome Integrity and Development. *Cell*, 149, 5.

Rolef Ben-Shahar, T., Castillo, A.G., Osborne, M.J., Borden, K.L., Kornblatt, J. & Verreault, A. (2009). Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. *Molecular and Cellular Biology.*, 29, 6353–6365.

Rupert, P., Daughdrill, G., Bowerman, B., & Matthews, B. (1998). A new DNA-binding motif in the SKN-1 binding domain-DNA complex. *Nature Structural Biology*, 5, 484-491.

Sacho, E.J., Kadyrov, F.A., Modrich, P., Kunkel, T.A., & Erie, D.A. (2008). Direct visualization of asymmetric adenine nucleotide-induced conformational changes in MutLα. *Molecular Cell*, 29, 112–121.

Sancar, A. & Hearst, J.E. (1993). Molecular matchmakers. *Science,* 1259, 1415–1420

Schmutte, C., Sadoff, M.M., Shim, K.S., Acharya, S., & Fishel, R. (2001). The interaction of DNA mismatch repair proteins with human exonuclease I. *Journal of Biological Chemistry*, 276(35), 33011–33018.

Simmons, L.A., Davies, B.W., Grossman, A.D., & Walker, G.C. (2008). Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Molecular Cell,* 29, 291-301.

Sixma, T.K. (2001). DNA mismatch repair: MutS structures bound to mismatches. *Current Opinion in Structural Biology,* 11, 47-52.

Stukenberg, P.T., Studwell-Vaughan, P.S. & O'Donnell, M. (1991). Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *Journal of Biological Chemistry* ,266, 11328–11334.

Stukenberg, P.T., Turner, J., & O'Donnell, M. (1994). An explanation for lagging strand replication: Polymerase hopping among DNA sliding clamps. *Cell*, 78, 877–887.

Thibodeau, S., Bren, G., and Schaid, D. (1993). Microsatellite instability in cancer of the proximal colon. *Science,* 260, 816–819.

Tram, K., Manochehry, S., Feng Q., Chang, D., Salena, B., & Li, Y. (2016). Colorimetric Detection of Bacteria Using Litmus Test. *Journal of Visualized Experiments*, (115), e54546.

Tsurimoto, T., & Stillman, B. (1991). Replication factors required for SV40 DNA replication *in vitro*. II. DNA structure specific recognition primer template junction by eukaryotic DNA polymerases and their accessory factors. *Journal of Biological Chemistry*, 266, 1950-1960.

Van Loon, B., Markkanen, E., & Hübscher, U. (2010). Oxygen as a friend and enemy: How to combat the mutational potential of 8-oxo-guanine. *DNA Repair,* 9, 604–616.

Vasen, H.F., Moslein, G., Alonso, A., Bernstein, I., Bertario, L., Blanco, I., Burn, J., Capella, G., Engel, C., Frayling, I.*,* Friedl, W., Hes, F.J., Hodgson, S., Mecklin, J.P., Moller, P., Nagengast, F., Parc, Y., Renkonen-Sinisalo, L., Sampson, J.R., Stormorken, A., & Wijnen, J. (2007). Guidelines for the clinical management of Lynch syndrome (hereditary non-polyposis cancer). *Journal of medical genetics*, 44, 353-362.

Wang, T.F., Kleckner, N., & Hunter, N. (1999). Functional specificity of MutL homologs in yeast: Evidence for three MLH1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. *Proceedings of the National Academy of Sciences*, 96, 13914–13919.

Warren, J.J., Pohlhaus, T.J., Changela, A., Iyer, R.R., Modrich, P.L., & Beese, L. (2007). Structure of the human MutSα DNA lesion recognition complex. *Molecular Cell*, *26*, 579–592.

Wold, M.S., & Kelly, T. (1988). Purification and characterization of replication protein A a cellular protein required for in vitro replication of simian virus 40 DNA. *Proceedings of the National Academy of Sciences of the United States of America,* 82, 2523-2537.

Yao, N.Y., & O'Donnell, M. (2010). SnapShot: The Replisome. *Cell*, 141(6), 1088–1088.

Yao, N.Y., Schroeder, J.W., Yurieva, O., Simmons, L.A., & O'Donnell, M.E. (2013). Cost of rNTP/dNTP pool imbalance at the replication fork. *Proceedings of the National Academy of Sciences of the United States of America,*110,12942–12947.

Yin, Z., Kelso, M.J., Beck, J.L. & Oakley, A.J. (2013). Structural and thermodynamic dissection of linear motif recognition by the E. coli   sliding clamp. *Journal of Medicinal Chemistry*, 56, 8665–8673.

Yoon, J.H., Lee, C.S., Oconnor, T.R., Yasui, A. & Pfeifer, G. P. (2000). The DNA damage spectrum produced by simulated sunlight, *Journal of Molecular Biology,* 299, 681–693.

Zakharyevich, K., Ma, Y., Tang, S., Hwang, P. Y. H., Boiteux, S., & Hunter, N. (2010). Temporally and biochemically distinct activities of Exo1 during Meiosis: Double- strand break resection and resolution of double Holliday Junctions. *Molecular Cell*, 40(6), 1001–1015.