Mechanistic Dissection of DNA Repair Pathways in *Bacillus subtilis*
by

Justin Stephen Lenhart

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Molecular, Cellular and Developmental Biology)
in the University of Michigan
2014

Doctoral Committee:

Assistant Professor Lyle A. Simmons, Chair
Professor James Bardwell
Associate Professor Matthew R. Chapman
Associate Professor Gyorgyi Csankovszki
Associate Professor Mats E.D. Ljungman

Justin Stephen Lenhart ©

2014

Dedication

To Adrienne Laurel Lenhart:

I vow to take your hand, on this day, and never relinquish it
To walk forward, through day and night, never to take a step apart
Overcoming obstacles and grief, while accomplishing goals and encouraging aspirations
And to one day share a more perfect love, destined only for you.

Since I declared this vow to you, I have been as happy as I could have ever surmised. Your love has given me the strength to accomplish all that is included in this document. As important as earning this doctorate is for me, this accomplishment pales to the importance of joining my life with yours on July 23rd, 2011. Wherever life takes us, I will be happy knowing that I have a partner to both enjoy successes and weather hardships with, and more importantly, a companion to experience all that life has to offer. Thank you for being there always.

Love Forever
Justin (BCB)

To Lyle A. Simmons:

Thank you. For the better part of five years, I have enjoyed coming to work everyday. I especially would like to thank you for your mentorship and the intellectual freedom granted during the completion of my thesis. The opportunities you provided: going to conferences, reviewing manuscripts, writing and submitting my manuscripts to revered journals, have truly prepared me for the next step of my career. I want you to know that whatever I accomplish in the future, you are largely responsible. Go Blue!

Justin

Acknowledgements

I am truly indebted to the following people and organizations for support that made this document possible.

The University of Michigan

The Department of Molecular, Cellular and Developmental Biology

Dr. Lyle A. Simmons

My Thesis Committee

Jeremy Schroeder

Members of the Simmons Lab

Nicholas Bolz

The Rackham Predoctoral Fellowship Program

Dr. Titus Franzmann

Wendy Will Case Cancer Fund

National Science Foundation grant MCB1050948 awarded to L.A.S.

MCDB start-up funds awarded to L.A.S.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Acronyms:

MMR=Mismatch Repair

IDL=Insertion/Deletion Loops

EMSA=Electrophoretic Mobility Shift Assay

ADP=Adenosine Diphosphate

ATP=Adenosine Triphosphate

ATPγS=Adenosine 5′ (3 thiotriphosphate) tetralithium salt

AMPPNP= Adenosine 5′-(β,γ-imido) triphosphate

dNTP=deoxyribonucleotide

dNMP=deoxyribonucleoside monophosphate

rNTP=ribonucleotide

rNMP=ribonucleoside monophosphate

RER=Ribonucleotide Excision Repair

GFP=Green Fluorescent Protein

mCherry/GFP=Monomeric Cherry/GFP

2-AP=2-aminopurine

IPTG=Isopropyl β–D-1-thiogalactopyranoside

DNA=Deoxyribonucleic acid

PIP=PCNA interacting peptide/protein

PCNA=Proliferating Cell Nuclear Antigen

BS3=bis(sulfosuccinimidyl)suberate

SSB=Single-stranded DNA binding Protein

DSB=Double Strand Break

DSG=Daughter Strand Gap

MMC=Mitomycin C

UV=Ultraviolet light

HPURa=6-(*p*-Hydroxyphenylazo-uracil

Abstract

All cells must accurately copy and maintain their DNA to ensure the faithful transmission of their genetic material to the next generation. Organisms ranging from bacteria to humans contain a suite of DNA repair pathways dedicated to the specific recognition and repair of the myriad of damaged or incorrect bases that occur throughout the lifetime of a cell. Here, I study the mechanisms of *Bacillus subtilis* mismatch repair and homologous recombination that maintain DNA integrity in the complex environment of a living cell.

I. Mismatch repair increases the fidelity of DNA replication by identifying and correcting replication errors. Although Mismatch repair has been thoroughly studied *in vitro*, little is known about how its central components, MutS and MutL, identify replication errors and orchestrate their repair within a living cell. Here, I investigate how the *B. subtilis* processivity clamp DnaN aids mismatch detection by MutS *in vivo*. I found that DnaN serves as an essential platform for mismatch detection, targeting the MutS search for mismatches to nascent DNA.

II. Following mismatch detection, MutS recruits MutL to the mismatch by an unknown mechanism. I identified a discrete site composed of two adjacent phenylalanine residues on MutS that binds MutL. Disruption of this site renders MutS defective in binding MutL *in vitro* and *in vivo*, while eliminating mismatch repair *in vivo*. Analysis of MutS repair complexes defective in MutL recruitment revealed a continuous loading response by MutS, revealing an intermediate step in the repair process.

III. The recombinase RecA is required for homologous recombination and stabilization of stalled replication forks in many bacteria, yet the polypeptides important for RecA loading remains unclear in organisms lacking RecBCD. Here,

I find that RecA loading is dependent on the presence of RecOR; proteins that associate with the SSB DNA maintenance hub, ensuring that RecA loading localizes to active replication forks. Furthermore, we find that RecF is not required for RecA loading; however, provides an enhancement in either RecA filament nucleation or elongation.

Chapter I

Background on Prokaryotic Mismatch Repair.

Chapter composed by JSL.

Introduction

Mismatch repair (MMR) is a DNA repair pathway primarily involved in correcting errors formed by DNA polymerases. Polymerase errors deviate from Watson-Crick base-pairing rules, including incorrect base pairings, termed mismatches, as well as extrahelical nucleotides termed insertion/deletion loops (IDLs). In the presence of mismatch repair, the chromosomes genetic information is maintained with extraordinary precision (~$1*10^{-3}$ mutation per genome per generation), yet in its absence, the fidelity of chromosome replication can be reduced up to a 1000 fold. MMR defects that reduce fidelity are demonstrated by genome instability, and in mammals, represents a key hallmark in the transition of a normal cell to a cancer cell. To facilitate mismatch correction, MMR must perform two basic functions: 1) identify replication errors within the complicated 3-Dimensional context of a dynamic chromosome and 2) determine which base of the mispair is incorrect. Within this chapter, we discuss the current understanding of the MMR pathway and how it maintains guard against mutation and genome instability. We will focus on recent *in vivo* advancements in prokaryotes lacking the MutH/Dam-containing mismatch repair pathway.

Background: Mismatch repair in *Escherichia coli*.

The paradigm for prokaryotic mismatch repair (MMR) has been the Gram-negative bacterium *Escherichia coli*. Work in this model system includes extensive biochemical characterization of all MMR components, including a successful reconstitution of the pathway *in vitro* (1)(See (2, 3) for review). Importantly, *E. coli* and a

few closely related bacteria represent systems where the MMR strand-discrimination signal are known: methylation of adenine in GATC sequences by the Dam methylase. At the start of the cell cycle, these GATC sites are fully methylated; however upon replication, most GATC sites exist in a transient hemimethylated state for >2 min (4). This brief hemimethylated state is exploited by MMR to correct the base located in the unmethylated strand, and hence the newly replicated daughter strand.

E. coli MMR is initiated upon the identification of a replication error by the homodimer sensor protein MutS (**Figure 1A**) (5). Upon mismatch detection, MutS recruits the homodimer MutL through a yet unknown mechanism. One model suggests that the MutS•MutL binary complex will rapidly dissociates from the mismatch and diffuse away along the DNA (**Figure 1B**). A third principal component, the restriction endonuclease-like protein MutH, follows transiently behind the ongoing replication fork through association with the hemi-methylated GATC sites (6). The diffusing MutS•MutL binary complex physically interacts with and activates MutH, which upon activation will specifically nick the unmethylated strand of the hemimethylated GATC site (7). The combination of the Dam methylase and MutH activities provides a signal directing the MMR pathway to the template strand in a process termed strand discrimination (8). At this newly incised nick, MutL loads and stimulates the activity of UvrD, a helicase with 3´-5´ polarity, on either the continuous strand or the incised strand, depending on whether the MutH-directed nick exists 5´ or 3´ to the mismatch respectively (**Figure 1C**) (9-11). Helicase loading ensures that helicase unwinding occurs toward the detected mismatch, removing the replication error, while the unwound mismatch containing strand is degraded by one of several single-stranded exonucleases (ExoI, RecJ, ExoVII, ExoX), producing a tract of SSB-bound ssDNA (12). In the final step, the DNA Polymerase III holoenzyme replicates over the gap employing the ssDNA as a template and DNA ligase seals the nick in the sugar-phosphate backbone (**Figure 1D**). In total, the repair of replication errors *in vitro* and *in vivo* relies on a minimum set of protein activities: MutS, MutL, MutH, Dam, UvrD, ExoI, the Pol III holoenzyme, LigA, and SSB (1).

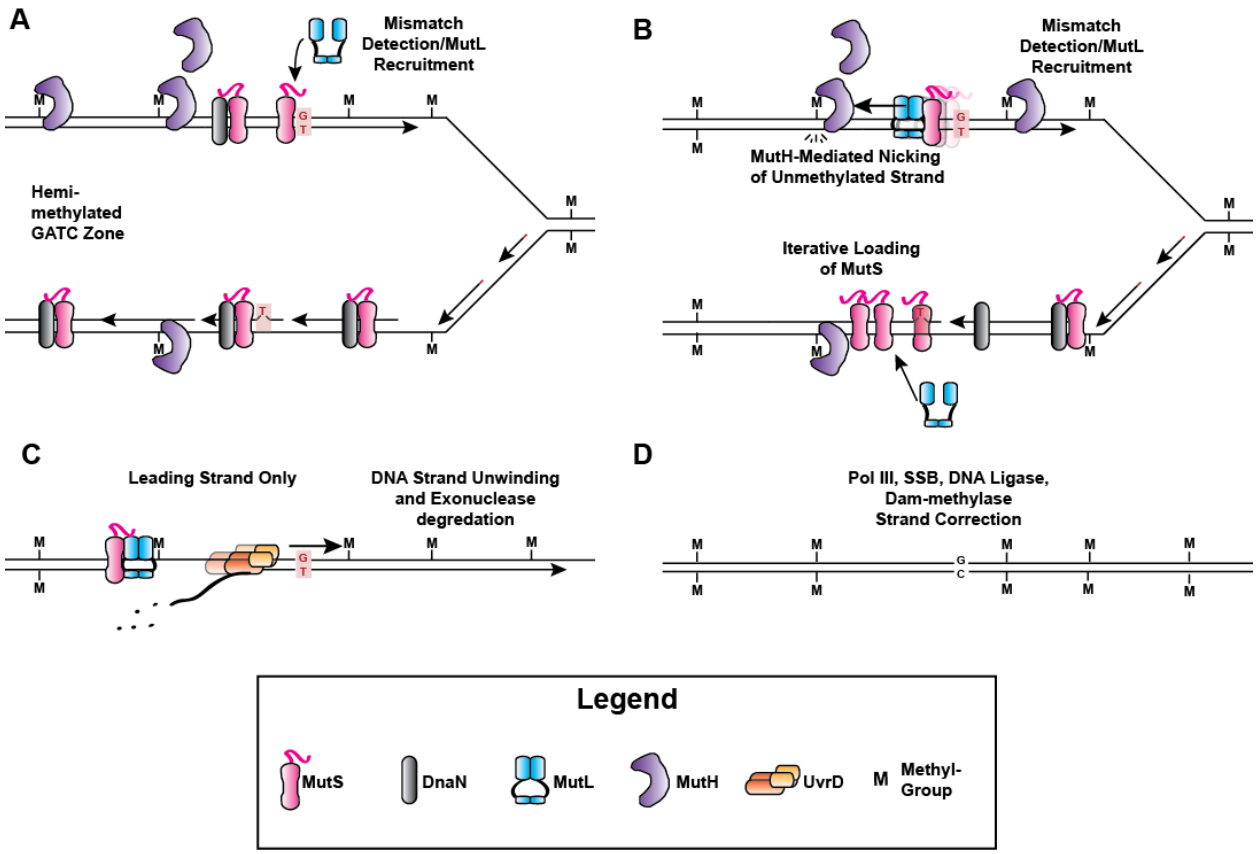


Figure 1. Model of *Escherichia coli* Methyl-Directed Repair.
A. A replication error is detected by MutS within a narrow window (≤ 2 minutes) after the progression of the replication fork where GATC sites are hemimethylated. To assist with the detection of the mismatch, the DnaN clamp positions MutS on nascent DNA. MutL is recruited by mismatch-bound MutS. **B.** The MutS•MutL binary complex diffuses away from the mismatch, colliding with and activating MutH. MutH is positioned at hemimethylated GATC sites and upon activation, will nick the unmethylated strand. **C.** MutL recruits the UvrD helicase to the nick, where it will unwind the unmethylated strand through the mismatch. Simultaneously, exonucleases will degrade the strand. **D.** SSB bound ssDNA surrounding the recently removed mismatch is replicated by the Pol III holoenzyme. Ligase will seal the remaining gap, completing repair.

The *E. coli* model system has traditionally represented the prokaryotic paradigm for MMR studies; however, most prokaryotes and all eukaryotes rely on a MutH and methylation-independent pathway. Most prokaryotic and eukaryotic MutL homologues (MLHs) contain a heavily conserved endonuclease active site, where extensive conservation is apparent among distantly related organisms from human to Arabidopsis to the Gram-positive bacterium *Bacillus subtilis*. Despite this extensive evolutionary conservation of MLHs, *E. coli* MutL lacks an intrinsic endonuclease activity that defines eukaryotic, and even most prokaryotic, MMR pathways (13). Although *E. coli* has served as the MMR paradigm since the discovery of mutator alleles in the 1960's, it

represents the exception rather than the rule for the study of *in vivo* MMR mechanisms in most prokaryotes (3).

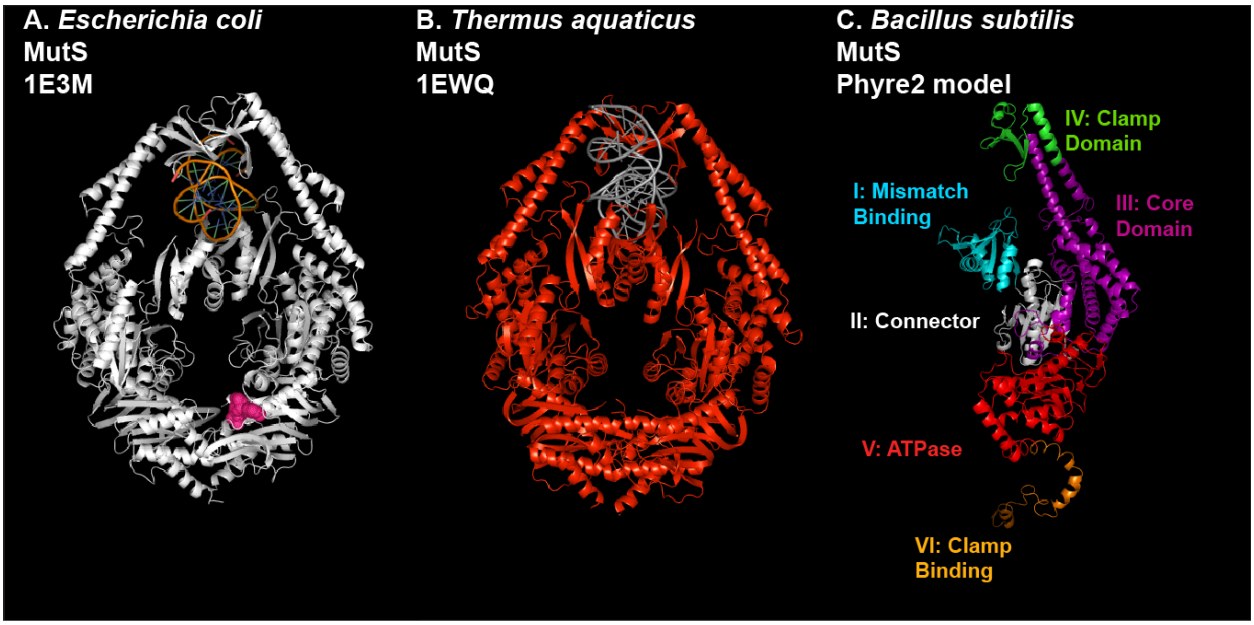


Figure 2 MutS: The hands of a guardian.
A. The crystal structure of *Escherichia coli* MutS dimer bound to a G•T mismatch (PDB 1E3M). ADP in Protomer A is represented by a dot structure in pink. **B.** The crystal structure of the *Thermus aquaticus* MutS dimer bound to a G•T mismatch (PDB 1EWQ). **C.** A Phyre2 model of the *Bacillus subtilis* MutS monomer. The domain organization has been labeled and indicated by the corresponding colors. The processivity clamp-binding domain (Domain VI), which is not present in MutS crystal structures in (**A**) and (**B**) was included in the the *B. subtilis* model, and is pictured in orange.

Biochemical characterization of MutS

In the *E. coli*, *Thermus aquaticus*, *Saccharomyces cerevisiae*, and human model systems, there has been extensive *in vitro* biochemical characterization of MutS homologues. These studies have revealed extensive mechanistic and biochemical conservation from mismatch detection through MutS sliding clamp formation in these systems. Here, important biochemical conclusions of this work will be discussed and used to understand MMR in an *in vivo* context.

Crystal structures of the *E. coli* and *T. aquaticus* MutS dimer, as well as the human MutSα heterodimer, has revealed that MutS homologues form a characteristic structure that resembles a pair of praying hands, with a DNA substrate being inspected between the thumb and the forefingers (**Figure 2**) (14-16). In the absence of DNA, the main dimer contacts of MutS homologues exist within Domain V (the wrist in our

analogy and the ATPase domain). In the *T. aquaticus* crystal structure, in the absence of DNA, the thumbs and fingers (Domains I and IV) are not resolved, illustrating a dynamic and mobile behavior of these domains without DNA. The addition of DNA stabilizes the mismatch bound praying hand structure by making additional dimer contacts between the thumbs (Domain I) and the forefingers (Domain IV) (**Figure 2**). Additionally upon DNA binding, structural changes occur in the ATPase active site (Domain V), suggesting that mismatch binding by MutS ultimately influences the ATPase cycle, and subsequent behavior of the MutS dimer during repair.

During association with DNA, MutS homodimers make both low affinity contacts, as well as specific interactions used for interrogating nucleobases for replication errors. Minor DNA contacts between the thumbs (Domain I) and the forefingers (Domain IV) position MutS on complementary DNA sequences, forming numerous salt-bridges with the sugar-phosphate backbone (14, 15). These interactions facilitate rotational movement in accordance with the helical contour (corkscrew-like rotation) of the DNA during the mismatch search. Furthermore, a second and more extensive DNA contact is made in Domain I of a single protomer (protomer A). Here, a phenylalanine in a highly conserved GxFxE motif specifically interrogates base pairs by inserting into the helix and stacking with the nitrogenous bases (in this case, a mismatch) (14, 15, 18). All specific contacts outside of this motif are made only with the sugar-phosphate backbone, ensuring unrestrained movement over complementary DNA, while preventing neighboring sequence context from influencing mismatch detection. Upon the phenylalanine stacking with the mismatched base, the DNA contour at the mismatch becomes bent (~60˚ from linear) (14, 15, 19). Little to no bending is observed by MutS on a complementary DNA substrate, suggesting that the degree of bending acts as a critical biochemical signal of mismatch recognition (19). Upon dissociation of MutS from the mismatch, all MutS-mediated bending ceases (19).

The ATPase site of MutS directs each step of MMR through ATP-mediated conformational changes, and importantly, signals for mismatch excision upon identification of a mismatch. The ATPase cycle has been shown to be conserved in all MutS homologues, as the ATP active sites (characteristic of ATP-binding cassette [ABC] ATPases) and the Walker A and B motifs that compose them, are heavily

conserved (17). In the Apo form, evidence suggests that MutS resides in an open structure, preventing DNA association, suggesting that partial occupancy of one of the ATPase sites with a nucleotide is required for DNA association (15) (**Figure 3**). The crystal structure of *E. coli* MutS bound to a mismatch revealed that the ATPase active sites were not super-imposable, identifying an asymmetry that is used to drive mismatch repair via the nucleotide cycle (14). During the search for replication errors, the nucleotide binding site of protomer A (the protomer interrogating the DNA) is occupied by ADP (14, 15). Upon mismatch detection, MutS exchanges ADP for ATP in a defined ordered transition, which after a brief pause will facilitate an extensive conformational change that signals for MutL recruitment in preparation for mismatch excision.

During the ADP to ATP transition, conformational changes in MutS domain I releases DNA, while Domain IV maintains dimer contacts, forming a MutS sliding DNA clamp (7, 20-24) (**Figure 3**). Specifically, this conformational change elicits a contraction of the ATPase domains, which causes a retraction of domain I towards the outside of the connector domain (Domain III) (26). The contraction of the ATPase active site was captured upon diffusing ATP into preformed crystals and elucidating the structure (25), whereas the retraction of domain I towards Domain III can be captured using chemical crosslinking (26). MutS, now devoid of any specific DNA contacts due to an increased solvent channel gauge, diffuses freely along the DNA. This diffusion occurs at a faster velocity than during the initial mismatch search, since rotational movements have been lost in the sliding clamp conformation (20, 21). This loss of constant register with the sugar-phosphate backbone is further illustrated by observed increases in the diffusion coefficient in both an ionic-strength-dependent, as well as a flow rate-dependent (due to laminar flow) manner (20). While in this state, ATP hydrolysis is suppressed, maintaining a prolonged complex ($t_{1/2} \geq 268 \pm 62$ s) that is passive to further mismatch detection (20, 21).

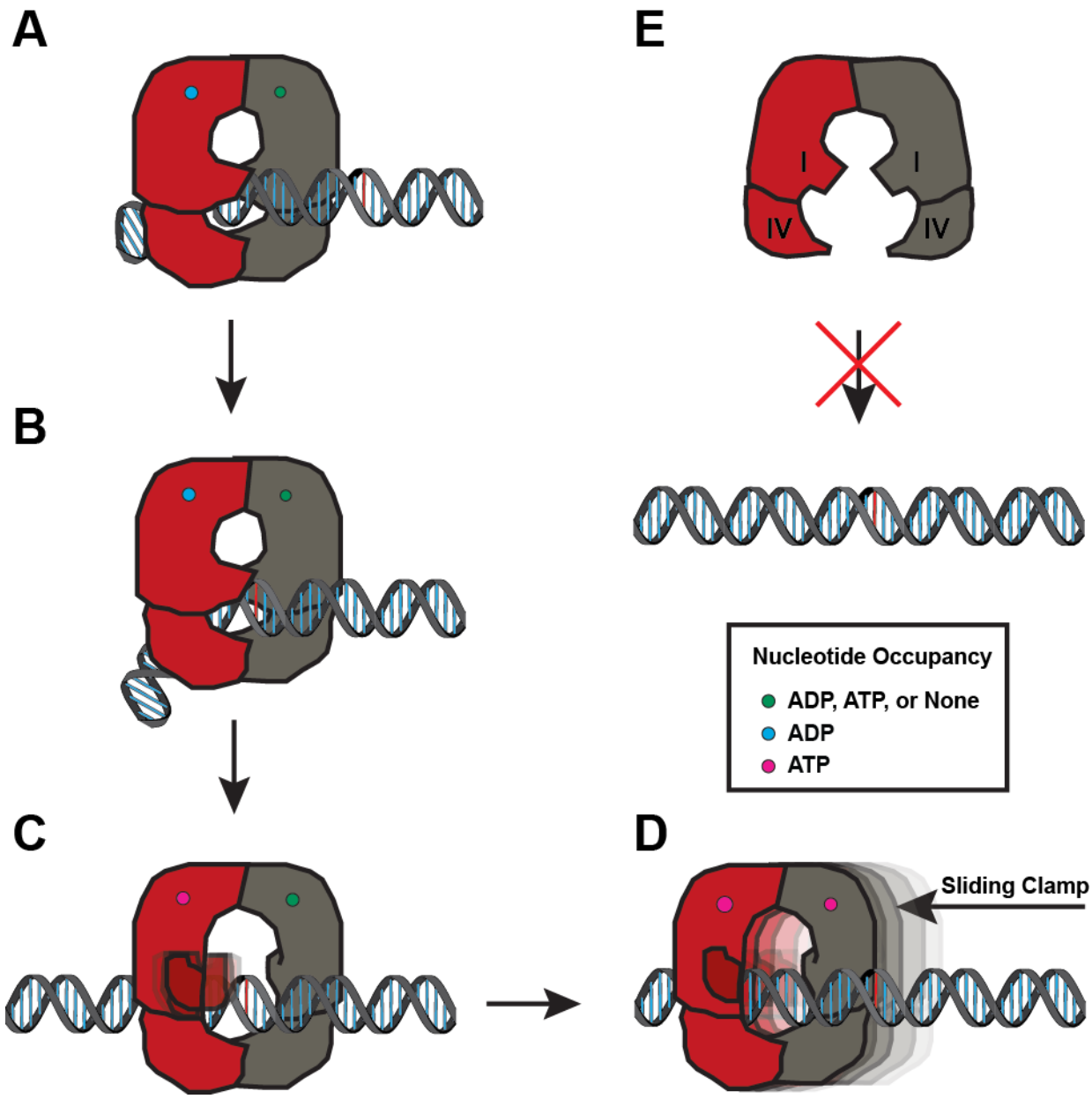


Figure 3. Biochemical schematic of mismatch detection by MutS.
The MutS dimer is pictured with each protomer assigned a unique color: protomer A is red while protomer B is grey. Protomer A represents the protomer that interrogates the DNA helix with the conserved phenylalanine. Predicted nucleotide occupancy is indicated based on the legend (Inset). **A.** MutS is scanning for mismatches on complementary DNA. Protomer A is predominantly in the ADP bound form. **B.** Upon detection of a mismatch, MutS pauses for a brief moment before signaling the need to initiate mismatch repair. **C.** After a pause of a few seconds, ATP occupancy in the active site of protomer A and possibly protomer B induces the formation of a sliding clamp via extensive conformational changes. **D.** The sliding clamp diffuses along the DNA refractory for further mismatch association to prevent re-binding to the same mismatch. Sliding clamp formation exposes the mismatch for further MutS binding events, facilitating iterative loading. **E.** The MutS dimer is unable to associate with DNA absent nucleotide occupancy. Domains I and IV are highly mobile, opening up the structure.

Replication coupling of mismatch repair

Despite the extensive studies elucidating the biochemical mechanisms of mismatch detection by MutS homologues on *in vitro* substrates, much less is known about how mismatch detection is performed *in vivo*. The main challenge posed to replication-error detection by MutS *in vivo* is the low frequency of their formation (~1 error per 10^8 bases replicated) (27, 28). Furthermore, a dynamic chromosome further obfuscates mismatch detection of these rare mismatches through DNA supercoiling, chromosome compaction, and a myriad of DNA-binding proteins. Furthermore, mismatch repair assemblies must engage in a competition with active processes such a transcription that could disrupt repair. Despite these challenges, early assertions were made that MutS identified errors via one-dimensional (1D) scanning of the whole genome in between replication cycles; however recent findings reveal that most mismatch detection is physically coupled to active replication forks, with subsequent repair occurring on replication proximal DNA (72). For example, *Saccharomyces cerevisiae* MMR is proficient only when MutSα is available during S-phase, and is conversely ineffective when MutSα is only made available during G2 or M phase after ongoing replication has ceased. This study illustrates temporal coupling of mismatch detection with active replication (29).

Spatial replication-coupled repair was also suggested due to the effect of strand breaks on the efficiency of MMR. Early biochemical studies using either cell extracts or reconstituted enzymatic systems observed that circular heteroduplex (mismatch containing) substrates were repaired efficiently only when a mismatch proximal nick (≤ 1kb away) was present (30-32). Furthermore, a nick near the mismatch is sufficient for strand discrimination by designating the *cis* strand as the strand targeted for nucleotide removal and repair. These nicks may represent strand termini that form during DNA replication, asymmetrically distributed to the lagging strand. In favor of strand termini directing repair, genetic evidence shows that MMR displayed a strand bias *in vivo*; where the lagging strand is repaired more efficiently than the leading strand (33-35). These observations suggest that either topological landmarks and/or protein factors

associated with DNA termini may direct efficient MMR and the strand discrimination step.

Replication coupling of mismatch detection *in vivo*.

One possible source of MMR coupling to replication is DnaN and PCNA (Proliferating Cell Nuclear Antigen), representing the prokaryotic and eukaryotic processivity clamps, respectively. DnaN (β-clamp) and PCNA are loaded at DNA termini by clamp loader complexes and primarily serve to tether replicative polymerases to DNA, increasing processive replication. Moreover, these clamps act as an organization scaffold for the replisome, as many proteins including MutS homologues are recruited to nascent DNA via interaction with processivity clamps (36-39). The majority of these interactions occur on the C-terminal face of the DnaN clamp in a cleft between domains 2 and 3, designated as the hydrophobic pocket. Many proteins have a dedicated clamp-binding motif (QL[SD]LF in prokaryotes; QxxLxxFF in eukaryotes) required for binding processivity clamps. The specific interaction between MutS homologues and processivity clamps is direct but weak, and occurs through the clamp-binding motif found within the unstructured C-terminal domain of MutS (40). Disruption of the clamp-binding motif or deletion of the C-terminal domain of MutS is sufficient to eliminate binding to DnaN *in vitro* (38, 39). However, attempts to discern an *in vivo* function for this interaction were complicated by observations that plasmid-borne *mutS* alleles defective in DnaN interaction successfully complemented a *ΔmutS* allele, failing to produce an elevated mutation rate, which led to the suggestion that interaction between MutS and DnaN does not contribute to mismatch repair *in vivo* (14, 38, 41). In *B. subtilis*, a clamp-binding motif mutant (MutS-^{806}AAAAA^{810}) at its native locus produced a minor increase in mutation rate (39). However, truncation of the amino acids that comprise the unstructured clamp binding domain of MutS (Domain VI-producing MutS800) in both *E. coli* and *B. subtilis* eliminates repair, revealing a dependence on yet-to-be identified binding interfaces within this domain (39, 42, 43). This increase of mutation rate is completely suppressed via over-expression of *mutS800*, revealing a dependence on the DnaN•MutS interaction for mismatch detection at native MutS concentrations (~100 nM *in vivo*), and perhaps explaining why disruption of the motif in

E. coli, *B. subtilis*, and *Pseudomonas aeruginosa* fails to produce elevated rates of mutagenesis at high expression levels (≥500 nM) ((38, 41-43) and Chapter 2).

The important role that processivity clamps play in *in vivo* mismatch detection was elucidated in part through cytological studies. In both prokaryotes and eukaryotes, DnaN and PCNA forms large clamp assemblies termed “clamp-zones” immediately following the progressing replication forks (44, 45) (**Figure 4A**). In *B. subtilis*, these clamp zones have been shown to contain as many as 200 DNA-bound DnaN clamp dimers as clamp loading and unloading rates reach an equilibrium (45). Importantly, these clamps are unevenly distributed in favor of the lagging strand due to Okazaki fragment synthesis and maturation; an asymmetry that may facilitate the increased MMR efficiency on the lagging strand. Initially, MutS is targeted to nascent DNA via its interaction with DnaN, allowing it to maintain a critical spatial and temporal relationship with the replication forks ((43) and Chapter 2). This occurs at a step prior to mismatch detection as MutSF30A, a mutant defective for mismatch detection, localizes to the clamp zone. This localization optimizes mismatch detection efficiency in part by restricting the search for mismatches to newly replicated DNA. When the MutS•DnaN interaction is disrupted, eliminating MutS enrichment on nascent DNA, 90% of MMR capacity is lost, emphasizing the importance of spatial-temporal coupling between mismatch detection and the replication forks on MMR (Chapter 2).

Alternatively, the clamp may increase processive DNA scanning of nascent DNA, much the same as the clamp imparts on the replicative DNA polymerases. In the absence of the DnaN clamp interaction, more MutS is needed to search DNA for mismatches using one-dimensional scanning. On short, end-blocked DNA substrates, as well long, 15.3 kb λ-based DNA curtains, DNA sampling times of MutS without the processivity clamp were ~1 s, with diffusion distances spanning between 700-1000 bps of naked DNA. Once a mismatch has been detected, MutS undergoes conformational changes that may facilitate dissociation from the clamp (46). Importantly, *S. cerevisiae* MutSα also forms replication-coupled foci preceding mismatch detection; however, active MMR repair centers are unable to be discerned due to low stoichiometry of MutSα during repair (47). Overall, coupling mismatch detection to replication forks by

processivity clamps increases the efficiency of mismatch detection, and may provide proximity to strand discrimination signals (**See Below**) to complete the repair process.

MutL recruitment to a mismatch.

Upon mismatch detection by MutS, MutL is recruited forming a DNA bound binary complex. Early cytological studies of *B. subtilis mutS-gfp* showed that MutS-GFP formed foci in response to mismatch detection (39, 43, 48). For visualization on an epifluorescence microscope, a minimum of 8 constrained fluorophores must distinctly localize to form a focus; visualization of MutS foci therefore suggested that during active repair within a living cell, numerous MutS dimers (directly quantitated as ≤12) must be involved (49). This observation suggested that MutS loads iteratively at a single mismatch, producing a transient increase in the local DNA-bound MutS concentration, which may aid in the recruitment of MutL (**Figure 4B**). This finding is consistent with observed repetitive loading of *E. coli* MutS in EMSA assays (7). When the downstream step of MutL recruitment is lost, MutS repair centers became brighter, signifying the continued loading (quantified as ≤ 30 dimers) of MutS on mismatch proximal DNA ((49) and Chapter 3). This exacerbated loading even continued upon successful MutL recruitment if the MutL endonuclease active site was disrupted, suggesting that only DNA incision and mismatch removal halts MutS repetitive loading at a mismatch, and not simply successful recruitment of MutL ((49) and Chapter 3).

Two possible mechanisms explain MutL recruitment by iterative loading of MutS: MutL binds MutS 1) upon sliding clamp formation or 2) exclusively when MutS is currently bound to a mismatch. Work with Surface Plasmon Resonance (SPR) found that MutS and MutL form stable complexes on a substrate containing both G/T mismatch and blocked ends in the presence of the non-hydrolysable ATP analogue ATPγS (7, 24). In a complementary approach, a gel retardation assay found that the MutS•MutL binary complex was refractory to an excess of G/T competitor DNA when incubated with ATPγS and not ATP (7). This supports the model where MutS•MutL interaction on DNA, despite sliding clamp formation by MutS forms a stable complex. Furthermore, both of these approaches failed to observe a MutS•MutL complex when MutS was bound at a mismatch in the presence of ADP (7). An alternate possibility for

in vivo iterative loading is that MutL must bind MutS bound at the mismatch, which has been observed to last ~3s *in vitro* (21, 24). Direct *in vitro* imaging failed to observe MutLα association with the MutSα sliding clamp after its departure from the mismatch (21), suggesting that MutL homologues (at least human) must bind MutS homologues while simultaneously bound to a mismatch (21, 24). Therefore, iterative loading would increase the opportunity for MutL to bind MutS by increasing the duration of mismatch occupancy by MutS (24, 49-51).

MutL recruitment is orchestrated through a physical association with MutS. A conserved motif found in the core domain (Phe 319, Phe 320) of MutS, that when disrupted, eliminates interaction *in vitro* and eliminates repair *in vivo* ((49) and Chapter 3). This motif is conserved within a significant number of Gram-positive bacteria, including many nosocomial pathogens (49). Interestingly in *E. coli*, a dual Glutamine motif is found on the opposite side of the MutS monomer from the *B. subtilis* binding interface in the connector domain (51). In both organisms, this places one MutL binding interface on each face of MutS, supporting interactions when diffusing either 5′ or 3′ to the mismatch. It is attractive to hypothesize that these MutL binding interfaces may only be exposed for interactions upon mismatch detection and occupancy of the ATPase site with ATP, providing a logical 'on-switch' for MutL binding and initiation of the downstream events of repair. Unfortunately, stoichiometries of MutL at MutS repair centers are unknown in *B. subtilis*; however, in cytological studies in *S. cerevisiae*, a functional Pms1 fluorescent fusion is speculated to localize to MutSα repair centers in supra-stoichiometries relative to MutSα (47, 48). Whether *B. subtilis* MMR is similar to *S. cerevisiae* remains unknown.

MutL endonuclease activity and activation.

In organisms with methylation-independent MMR, a highly conserved endonuclease active site was found within the C-terminal domain of MutL homologues (13). Discovery of the endonuclease activity in human MutLα subunit PMS2 explained how EXO1, a 5′-3′ exonuclease important for human MMR, could paradoxically perform bi-directional excision of a mismatch initiated from a nick either 5′ or 3′ to the mismatch *in vitro*. During repair, MutLα endonuclease activity is directed along the strand

containing the pre-existing nick, with the goal of nicking upstream (5′) of the mismatch. This provides a 5′ loading site for EXO1, which can then mediate digestion of the mismatch containing strand, and thus bidirectional repair (13).

A Mn^{2+} coordinating motif $DQHA(X)_2E(X)_4E$ composes a weak endonuclease active site in the C-terminal domain of PMS2 that nicks mismatch proximal DNA (13, 52). In *B. subtilis*, a homologous Mn^{2+} dependent endonuclease activity has been shown to exist in its C-terminal domain, and is found to be essential for MMR *in vivo* (53). A crystal structure of the C-terminal domain of *B. subtilis* MutL revealed two distinct metal coordinating sites within the endonuclease active site. The main site, which coordinates the catalytic Mn^{2+} ion, is primarily composed of the $^{462}DQHA(X)_2E(X)_4E$ with the Asp462 and His464 residues predicted to define the catalytic residues (53). Additional conserved motifs ^{604}CPHGRP and ^{572}SCK, and a third ^{623}FKR motif originated from the C-terminus of other protomer in the MutL dimer, assisted with Mn^{2+} coordination (54). A second catalytic site was identified adjacent to the Mn^{2+} binding site that coordinated Zn^{2+}. The E468 residue from the $^{462}DQHA(X)_2E(X)_4E$ site and ^{604}CPHCRP motif stabilized the Zn^{2+} ion. Although Zn^{2+} isn't necessary to facilitate *in vitro* nicking, disruption of Zn^{2+} coordination eliminated mismatch repair *in vivo*, suggesting an essential structural role in the coordination of the catalytic Mn^{2+} ion during endonuclease-mediated catalysis (53).

Regulation of the endonuclease activity of MutL is an important consideration, because aberrant nicking could result in loss of viability and increased genome instability. Part of this regulation is dependent on the intrinsically weak nature of the endonuclease activity, which suggests a licensing event that must occur to activate MutL at the proper time (53). A view of the electrostatic surface potential surrounding the active site reveals a sheath of negative charged residues guarding against DNA docking within the site until licensing occurs (53). Consequently, the C-terminal domain is unable to bind DNA, because it relies on dimerization of the N-terminal domain of MutL after association with newly replicated DNA. Additionally, conformational changes of MutL facilitated by its nucleotide bound state may indirectly or directly modulate endonuclease activity. Previous biochemical studies of *S. cerevisiae* MutLα using atomic force microscopy showed that at an ATP concentration that supports both Mlh1

and Pms1 binding of ATP (5 mM), MutLα formed a condensed conformation, which could draw the endonuclease site towards bound DNA, overcoming the repulsive forces of the electrostatically charged surface surrounding the active site (55).

Other protein factors may further license MutL endonuclease activity. Human MutLα endonuclease activity is stimulated by the processivity clamp PCNA (56). In *B. subtilis*, the C-terminal domain of MutL contains a conserved, albeit degenerate DnaN clamp binding motif. Disruption of this motif eliminates interaction with DnaN *in vitro*, and abrogates mismatch repair *in vivo* (53, 57*)*. Importantly in *E. coli*, disruption of an analogous motif fails to produce a significant increase in mutation rate, demonstrating that organisms employing methyl-directed repair rely less on the MutL•DnaN interaction than organisms that employ a MutL-contained endonuclease activity (57). Interestingly while loaded onto DNA, the clamp maintains a specific orientation, with its single-C-terminal face positioned to interact with the replicative polymerases. Furthermore, the clamp is tilted on DNA (22˚ off perpendicular) upon being loaded by the clamp loader complex, potentially dictating strand orientation ((58) **Figure 4C**). One idea is that strand discrimination can be determined *in vitro* by PCNAs loaded orientation (56). Overall, both intrinsic acitivites of MutL, as well as the processivity clamp can regulate endonuclease nicking. Yet, the complete mechanism remains unclear.

Ribonucleotides as strand discrimination signals.

During DNA replication, rNTPs (ribonucleoside triphophates) are occasionally incorporated in place of their corresponding dNTPs (deoxyribonucleoside triphosphates) since their intracellular concentrations are 10-100 fold greater (59-61). *In vitro*, the rate of rNTP incorporation by the *E. coli* Pol III holoenzyme is1 rNMP for every 2.3 kb of replicated DNA, which *in vivo* would correspond to ~2,000 rNMPs embedded per chromosome per replication cycle (61). dNMPs are less reactive (~100,000 fold less), and subsequently more stable than rNMPS, which when incorporated into DNA, produce both single stranded and deleterious double stranded breaks via alkali-promoted transesterification (62). Furthermore, rNTPs and template rNMPs slow the rate of progression of the replicative polymerase *in vitro* (61). Despite these detriments, their incorporation could provide a benefit to the cell: their removal would produce a nick

in the newly replicated stand, marking the nascent DNA strand for MMR (63, 64). During Ribonucleotide Excision Repair (RER), RNase HII (RnhB in *B. subtilis*) incises the rNMP embedded in DNA 5′ of the ribonucleoside, generating a nick with 3′-hydroxyl and 5′-phospho-ribonucleotide ends (65). Preceding further processing, these RER-dependent nicks appear to be used by the MMR pathway to direct repair to the newly replicated DNA. In *B. subtilis*, deletion of *rnhB* causes a ~2.4 fold increase in mutation rate (compared to the 60 fold increase of *ΔmutSL* strain) (61). This elevated mutation rate is further increased when combined with a deletion of another RNase H enzyme: the RNase HI RnhC, causing a 5 fold increase in the spontaneous mutation rate (61). A deletion of *rnhC* alone does not affect MMR efficiency. In *S. cerevisiae*, MMR is more dependent on incorporated rNMPs for strand discrimination signals in the leading strand, whereas on the lagging strand where there are significantly more 3′ and 5′ DNA termini and DNA-bound PCNA, there is little effect on MMR (63, 64). This data suggests that incorporated rNMPs act as strand discrimination signals, but most likely on the leading strand in *B. subtilis*, accounting for the relatively minor mutator phenotype. Importantly, deletion of *rnhB* in *E. coli*, the only RNase H2 enzyme, fails to increase the mutation rate, arguing that RER predominantly contributes as a strand discrimination signal in methylation-independent organisms (61). Moving forward, two directions should be pursued in prokaryotes: 1) reconstitution of rNMP-directed repair *in vitro* and 2) comprehensive analysis of the contribution of RER on spontaneous mutation rate, as well as its distribution between leading and lagging strands, *in vivo*. These directions are necessary to ultimately validate that the observed increase in spontaneous mutation rate is due to a MMR defect and not the rNMP mispairing during replication. However, this is unlikely since base pairing is affected by alterations to the nitrogenous bases, not the sugar subunit. Since the discovery of rNTP incorporation in DNA synthesis in 2010, many reports have been published on RER, yet to date, both the RER mechanism and its crosstalk on MMR in bacterial systems remains unclear.

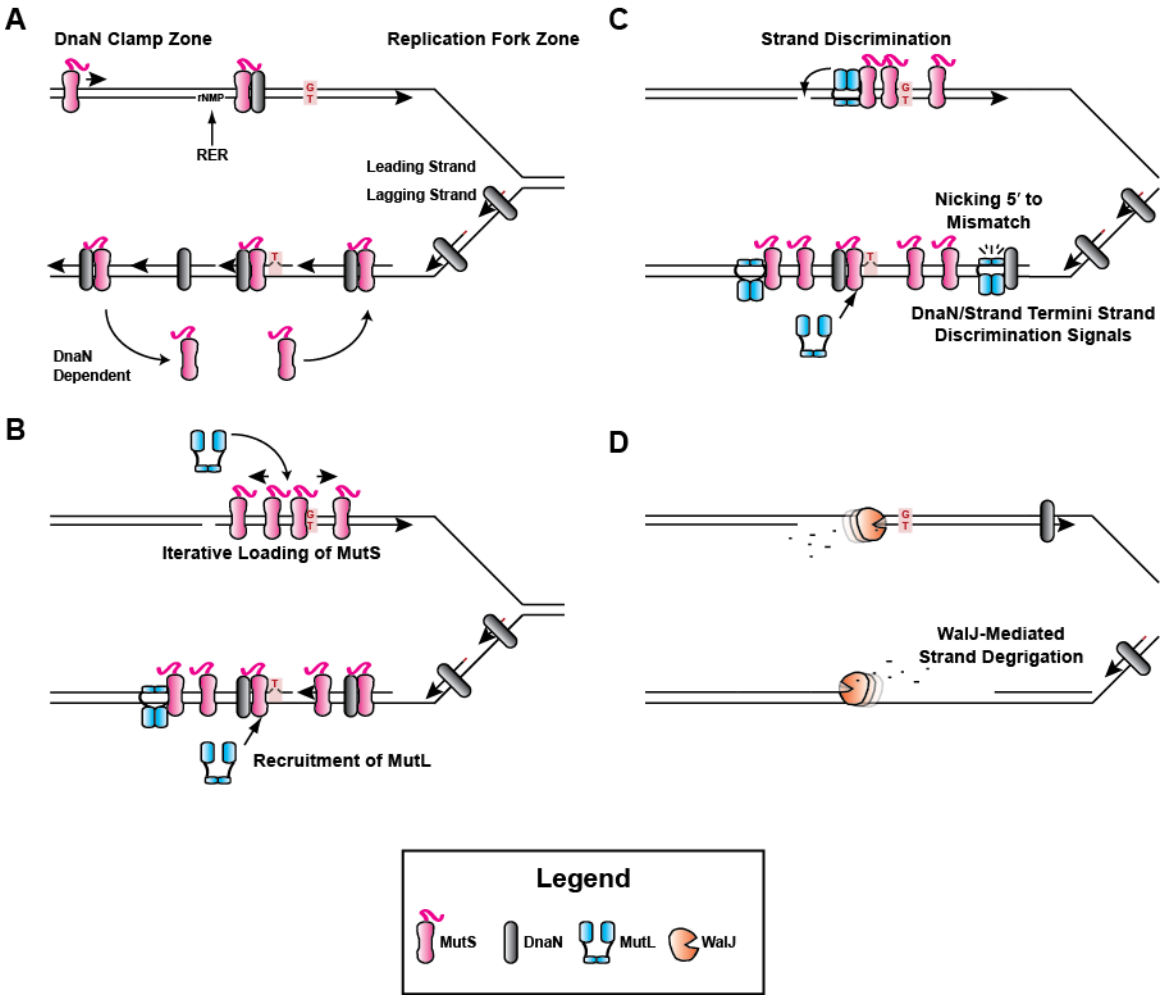


Figure 4. Model of Mismatch Repair in Methyl-Independent Systems.
A. Replication coupling of mismatch detection is facilitated by a DnaN clamp zone. **B.** Upon mismatch detection, the ATPase cycle of MutS facilitates loading and subsequent release of MutS from the mismatch, producing a local increase in mismatch proximal MutS in a process termed iterative loading. MutL is recruited to mismatch proximal DNA by MutS. **C.** MutL endonuclease activity is regulated via strand discrimination signals. **D.** WalJ loads/is loaded at a MutL processed nick 5´ to the mismatch. WalJ degrades the mismatch containing strand in the 5´ - 3´ polarity. After strand removal by WalJ, error free replication over the mismatch will result in the correction of the error.

Strand Removal and Resynthesis

Upon mismatch detection, the newly replicated strand containing the base pairing error needs to be removed to facilitate its correction. This process is initiated after an endonuclease nicks the newly replicated strand upstream of a mismatch. This nicked strand is either unwound or displaced towards the mismatch, allowing for an exonuclease to degrade the mismatch containing DNA tract. In *E. coli* and other

MutH/Dam containing bacteria, the UvrD helicase is required to unwind the mismatch-containing strand, followed by strand degradation by one of several redundant exonucleases (1). In *Bacillus anthracis*, a *uvrD* mutant fails to produce an elevation in spontaneous mutation frequency, arguing against a role in MMR in this bacterium (66). In *B. anthracis*, disruption of *recD2*, a member of Superfamily 1 helicases for which UvrD is a member, results in an increase in mutagenesis consistent with a loss of MMR (66). However, despite an extensive phylogenetic distribution of *recD2*, *B. subtilis* RecD2 appears to function outside of MMR, engaging in the repair of other types of DNA damage (67). Therefore, the extent of RecD2 helicase function in bacterial MMR requires further study. Furthermore, further study is needed to determine whether a dedicated helicase exists in other MutH/Dam-independent bacterial species. One possibility is that these bacterial MMR systems resemble that of eukaryotes, where a helicase has not been identified to function in MMR. For mismatch removal, eukaryotes either rely on direct strand digestion by the EXO1 exonuclease or strand displacement by polymerase δ loaded at a 3′ terminus upstream of the identified mismatch, followed by nucleolytic processing by Fen1 or a single-stranded nuclease (68). One possibility is that a newly discovered exonuclease, YycJ, functions in an analogous role to eukaryotic EXO1. Disruption of *yycJ* produces a spontaneous mutation rate equivalent to a MMR deficient strain in both *B. anthracis* and *B. subtilis* (69). Biochemical studies of YycJ show that it is a Mn^{2+} dependent 5′-3′ exonuclease (70) (**Figure 4D**). After removal of the mismatch containing tract, a polymerase extends the upstream 3′ terminus resynthesizing the gap generated a single nick, which is then sealed by DNA ligase to complete the repair process.

Conclusion

Since the development of *B. subtilis* as an experimental system, important advancements have been made during the study of the *in vivo* architecture and mechanistic steps of MMR in organisms lacking MutH and Dam methylase. Advancements in replication coupling of mismatch repair, the crystallization of the first MutL endonuclease domain, iterative loading, and identifying DNA embedded ribonucleotides as strand discrimination signals have been made. Yet much still remains

unanswered and in need of further study. 1) How is the MutL endonuclease activity regulated and directed. Does MutL nick DNA at strand discrimination signals, or does it rely on the orientation of MutS and its subsequent interaction during recruitment? 2) How does MMR co-opt rNMP removal to facilitate/direct repair to the correct strand? 3) Upon strand nicking, how does the mismatch-containing tract get removed? Advancements here can trigger new and exciting studies, just as the initial discovery of the *mutS* in *E. coli* by Siegel and Bryson in 1964 triggered 5 decades of intense study of Mismatch Repair in organisms from bacteria to humans (71).

References

1 Lahue, R. S., Au, K. G. & Modrich, P. DNA mismatch correction in a defined system. *Science* **245**, 160-164 (1989).

2 Jiricny, J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* **7**, 335-346, (2006).

3 Jiricny, J. Postreplicative mismatch repair. *Cold Spring Harb Perspect Biol* **5**, a012633, doi:10.1101/cshperspect.a012633 (2013).

4 Campbell, J. L. & Kleckner, N. E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* **62**, 967-979 (1990).

5 Su, S. S. & Modrich, P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 5057-5061 (1986).

6 Waldminghaus, T., Weigel, C. & Skarstad, K. Replication fork movement and methylation govern SeqA binding to the Escherichia coli chromosome. *Nucleic Acids Res* **40**, 5465-5476 (2012).

7 Acharya, S., Foster, P. L., Brooks, P. & Fishel, R. The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Molecular cell* **12**, 233-246 (2003).

8 Glickman, B. W. & Radman, M. Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 1063-1067 (1980).

9 Dao, V. & Modrich, P. Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. *J Biol Chem* **273**, 9202-9207 (1998).

10 Hall, M. C., Jordan, J. R. & Matson, S. W. Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD. *The EMBO journal* **17**, 1535-1541, (1998).

11 Yamaguchi, M., Dao, V. & Modrich, P. MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J Biol Chem* **273**, 9197-9201 (1998).

12 Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S. T. & Modrich, P. In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 6765-6770, (2001).

13 Kadyrov, F. A., Dzantiev, L., Constantin, N. & Modrich, P. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297-308, (2006).

14 Lamers, M. H. *et al.* The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* **407**, 711-717, doi:10.1038/35037523 (2000).

15 Obmolova, G., Ban, C., Hsieh, P. & Yang, W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407**, 703-710, (2000).

16 Warren, J. J. *et al.* Structure of the human MutSalpha DNA lesion recognition complex. *Molecular cell* **26**, 579-592, (2007).

17 Sachadyn, P. Conservation and diversity of MutS proteins. *Mutat Res* **694**, 20-30 (2010).

18 Malkov, V. A., Biswas, I., Camerini-Otero, R. D. & Hsieh, P. Photocross-linking of the NH2-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA. *J Biol Chem* **272**, 23811-23817 (1997).

19 Hura, G. L. *et al.* DNA conformations in mismatch repair probed in solution by X-ray scattering from gold nanocrystals. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 17308-17313 (2013)

20 Cho, W. K. *et al.* ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure* **20**, 1264-1274 (2012)

21 Gorman, J. *et al.* 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**, E3074-3083 (2012).

22 Gradia, S. *et al.* hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Molecular cell* **3**, 255-261 (1999).

23 Jeong, C. *et al.* MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat Struct Mol Biol* **18**, 379-385 (2011).

24 Mendillo, M. L., Mazur, D. J. & Kolodner, R. D. Analysis of the interaction between the Saccharomyces cerevisiae MSH2-MSH6 and MLH1-PMS1 complexes with DNA using a reversible DNA end-blocking system. *J Biol Chem* **280**, 22245-22257 (2005).

25 Lamers, M. H. *et al.* ATP increases the affinity between MutS ATPase domains. Implications for ATP hydrolysis and conformational changes. *J Biol Chem* **279**, 43879-43885 (2004).

26 Winkler, I. *et al.* Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in Escherichia coli. *J Biol Chem* **286**, 17326-17337 (2011).

27 Lee, H., Popodi, E., Tang, H. & Foster, P. L. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2774-2783 (2012)

28 Zanders, S. *et al.* Detection of heterozygous mutations in the genome of mismatch repair defective diploid yeast using a Bayesian approach. *Genetics* **186**, 493-503 (2010)

29 Hombauer, H., Srivatsan, A., Putnam, C. D. & Kolodner, R. D. Mismatch repair, but not heteroduplex rejection, is temporally coupled to DNA replication. *Science* **334**, 1713-1716 (2011).

30 Dzantiev, L. *et al.* A defined human system that supports bidirectional mismatch-provoked excision. *Molecular cell* **15**, 31-41 (2004)

31 Fang, W. H. & Modrich, P. Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. *J Biol Chem* **268**, 11838-11844 (1993).

32 Thomas, D. C., Roberts, J. D. & Kunkel, T. A. Heteroduplex repair in extracts of human HeLa cells. *J Biol Chem* **266**, 3744-3751 (1991).

33 Fijalkowska, I. J., Jonczyk, P., Tkaczyk, M. M., Bialoskorska, M. & Schaaper, R. M. Unequal fidelity of leading strand and lagging strand DNA replication on the Escherichia coli chromosome. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 10020-10025 (1998).

34 Pavlov, Y. I., Mian, I. M. & Kunkel, T. A. Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol* **13**, 744-748 (2003).

35 Pavlov, Y. I., Newlon, C. S. & Kunkel, T. A. Yeast origins establish a strand bias for replicational mutagenesis. *Molecular cell* **10**, 207-213 (2002).

36 Gu, L., Hong, Y., McCulloch, S., Watanabe, H. & Li, G. M. ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. *Nucleic Acids Res* **26**, 1173-1178 (1998).

37 Kleczkowska, H. E., Marra, G., Lettieri, T. & Jiricny, J. hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. *Genes Dev* **15**, 724-736 (2001).

38 Lopez de Saro, F. J., Marinus, M. G., Modrich, P. & O'Donnell, M. The beta sliding clamp binds to multiple sites within MutL and MutS. *J Biol Chem* **281**, 14340-14349 (2006).

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

40 Dalrymple, B. P., Kongsuwan, K., Wijffels, G., Dixon, N. E. & Jennings, P. A. 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 (2001).

41 Monti, M. R., Miguel, V., Borgogno, M. V. & Argarana, C. E. Functional analysis of the interaction between the mismatch repair protein MutS and the replication processivity factor beta clamp in Pseudomonas aeruginosa. *DNA repair* **11**, 463-469 (2012).

42 Calmann, M. A., Nowosielska, A. & Marinus, M. G. The MutS C terminus is essential for mismatch repair activity in vivo. *Journal of bacteriology* **187**, 6577-6579 (2005).

43 Lenhart, J. S., Sharma, A., Hingorani, M. M. & Simmons, L. A. DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular microbiology* **87**, 553-568 (2013).

44 Shibahara, K. & Stillman, B. Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* **96**, 575-585 (1999).

45 Su'etsugu, M. & Errington, J. The replicase sliding clamp dynamically accumulates behind progressing replication forks in Bacillus subtilis cells. *Molecular cell* **41**, 720-732 (2011).

46 Lau, P. J. & Kolodner, R. D. Transfer of the MSH2.MSH6 complex from proliferating cell nuclear antigen to mispaired bases in DNA. *J Biol Chem* **278**, 14-17 (2003).

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

48 Smith, B. T., Grossman, A. D. & Walker, G. C. Visualization of mismatch repair in bacterial cells. *Molecular cell* **8**, 1197-1206 (2001).

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

50 Junop, M. S., Obmolova, G., Rausch, K., Hsieh, P. & Yang, W. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Molecular cell* **7**, 1-12 (2001).

51 Mendillo, M. L. *et al.* A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 22223-22228 (2009).

52 Gueneau, E. *et al.* Structure of the MutLalpha C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. *Nat Struct Mol Biol* **20**, 461-468 (2013).

53 Pillon, M. C. *et al.* Structure of the endonuclease domain of MutL: unlicensed to cut. *Molecular cell* **39**, 145-151 (2010).

54 Kosinski, J., Plotz, G., Guarne, A., Bujnicki, J. M. & Friedhoff, P. The PMS2 subunit of human MutLalpha contains a metal ion binding domain of the iron-dependent repressor protein family. *Journal of molecular biology* **382**, 610-627 (2008).

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

56 Pluciennik, A. *et al.* PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc Natl Acad Sci U S A* **107**, 16066-16071 (2010).

57 Pillon, M. C., Miller, J. H. & Guarne, A. The endonuclease domain of MutL interacts with the beta sliding clamp. *DNA repair* **10**, 87-93 (2011).

58 Georgescu, R. E. *et al.* Structure of a sliding clamp on DNA. *Cell* **132**, 43-54 (2008).

59 Nick McElhinny, S. A. *et al.* Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* **6**, 774-781 (2010).

60 Nick McElhinny, S. A. *et al.* Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4949-4954 (2010).

61 Yao, N. Y., Schroeder, J. W., Yurieva, O., Simmons, L. A. & O'Donnell, M. E. 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 (2013).

62 Li, Y. F., Breaker, R.R. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group. *J. Am. Chem. Soc* **121**, 5364-5372 (1999).

63 Ghodgaonkar, M. M. *et al.* Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Molecular cell* **50**, 323-332 (2013).

64 Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B. & Kunkel, T. A. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Molecular cell* **50**, 437-443 (2013).

65 Sparks, J. L. *et al.* RNase H2-initiated ribonucleotide excision repair. *Molecular cell* **47**, 980-986 (2012).

66 Yang, H., Yung, M., Sikavi, C. & Miller, J. H. The role of Bacillus anthracis RecD2 helicase in DNA mismatch repair. *DNA repair* **10**, 1121-1130 (2011).

67 Walsh B.W., B. S. A., Wessel S.R.,SchroederJ.W., Keck J.L., Simmons L.A. RecD2 helicase limits replication fork stress in Bacillus subtilis. *J. Bacteriol* (2014). (In Press)

68 Kadyrova, L. Y., Blanko, E. R. & Kadyrov, F. A. CAF-I-dependent control of degradation of the discontinuous strands during mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 2753-2758 (2011).

69 Yang, H. *et al.* Papillation in Bacillus anthracis colonies: a tool for finding new mutators. *Molecular microbiology* **79**, 1276-1293 (2011).

70 Yang, H. *et al.* Evidence that YycJ is a novel 5'-3' double-stranded DNA exonuclease acting in Bacillus anthracis mismatch repair. *DNA repair* **12**, 334-346 (2013).

71 Siegel E.C., B., V. Selection of resistant strains of Escherichia coli by antibiotics and antibacterial agents: role of normal and mutator strains. *J. Bacteriol* **94**, 38-47 (1964).

72 Claverys J.P., Lacks S.A. Heteroduplex deoxyribonucleic acid base mismatch repair in bacteria. *Microbiol. Rev.* **50**, 133-165 (1986)

Chapter II

DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication

Justin S. Lenhart[1], Anushi Sharma[2], Manju M. Hingorani[2], and Lyle A. Simmons[1*],

[1]Molecular, Cellular, and Developmental Biology, University of Michigan. Ann Arbor, MI

[2]Department of Molecular Biology and Biochemistry, Wesleyan University. Middletown, CT

Author contributions: All work except Figure 5C was performed by JSL. Figure 5C was performed by AS. JSL, LAS, and MMH designed the research. All of the authors contributed to the writing of this chapter.

Manuscript is published: Lenhart, J. S., Sharma, A., Hingorani, M. M. & Simmons, L. A. DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular microbiology* **87**, 553-568 (2013).

Mailing address: 4140 Kraus Natural Science building Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109. Phone: (734) 647-2016. E-mail: lasimm@umich.edu

Abstract

Mismatch Repair (MMR) increases the fidelity of DNA replication by identifying and correcting replication errors. Processivity clamps are vital components of DNA replication and mismatch repair, yet the mechanism and extent to which they participate in MMR remains unclear. We investigated the role of the *Bacillus subtilis* processivity clamp DnaN, and found that it serves as a platform for mismatch detection and coupling of repair to DNA replication. By visualizing functional MutS fluorescent fusions *in vivo*, we find that MutS forms foci independent of mismatch detection at sites of replication (i.e., the replisome). These MutS foci are directed to the replisome by DnaN clamp zones that aid mismatch detection by targeting the search to nascent DNA. Following mismatch detection, MutS disengages from the replisome, facilitating repair. We tested the functional importance of DnaN-mediated mismatch detection for MMR, and found that it accounts for 90% of repair. This high dependence on DnaN can be bypassed by increasing MutS concentration within the cell, indicating a secondary mode of detection *in vivo* whereby MutS directly finds mismatches without associating with the replisome. Overall, our results provide new insight into the mechanism by which DnaN couples mismatch recognition to DNA replication in living cells.

Introduction

Mismatch repair (MMR) increases the fidelity of DNA replication by identifying and correcting errors made by the replicative DNA polymerase [for review (2, 3)]. Upon detection of an error, MMR orchestrates its removal and accurate resynthesis of the surrounding DNA, ultimately increasing the fidelity of DNA replication by several hundred-fold [for review (2-4)]. Due to this important role in maintaining genome stability, MMR is found in all domains of life, with a high degree of conservation, specifically among MutS and MutL proteins (5). In bacteria, deletion of either *mutS* or *mutL* homologs leads to an increased mutation rate (6-10). This mutator phenotype is known to accelerate acquisition of multidrug resistant strains in hospital settings, while also enabling increased survival of bacterial pathogens in harsh environments, including growth inside the lungs of cystic fibrosis patients [(11-17) and for review (18, 19)]. MMR defects in eukaryotes are characterized by hypermutability and microsatellite instability; both of which have been linked to an increased predisposition for spontaneous tumorigenesis, as well as inherited conditions such as Lynch syndrome and Turcot syndrome (20-23).

In bacteria, the MutS homodimer initiates MMR by detecting base-base mismatches or small insertion/deletion loops (IDLs) (24). In eukaryotes, base-base mismatches and small IDLs (1 or 2 extrahelical nucleotides) in DNA are primarily recognized by Msh2-Msh6 (MutSα), while larger IDLs (1-15 extrahelical nucleotides) are recognized by Msh2-Msh3 (MutSβ) heterodimers (25-28). In all systems, following mismatch or IDL detection by a MutS homolog, MutL (MutLα or MutLβ in eukaryotes) is recruited to the site of the mismatch in a reaction that requires ATP (29). Following this step, MutL is hypothesized to facilitate removal of the mismatch by coordinating numerous DNA transactions including endonuclease nicking, helicase driven unwinding, and excision of the segment containing the misincorporated base(s) (30).

Since replicative DNA polymerases have high fidelity, base pairing errors occur at a low frequency of one in 10^7-10^8 correctly paired bases [(31) and for review (32, 33)]. In addition to the challenges posed by the low rate or error formation, base mispairs may also be obscured by DNA supercoiling, compaction, and protein binding. MutS must also contend with other active processes on the DNA, including

transcription, when searching for mismatches and IDLs in DNA [for review on chromosome organization (34)]. Given these challenges, it has been proposed that MutS is coupled to DNA replication forks in order to facilitate efficient mismatch detection where mismatches are newly formed, and where the DNA is more likely to be free of protein impediments (35, 36). In support of this model, cytological studies conducted in *B. subtilis, S. cerevisiae*, and human cells have shown that prokaryotic MutS-GFP and eukaryotic MutSα (Msh6-mCherry) form foci that are often coincident with DNA replication foci *in vivo* (35-38). Furthermore, in *B. subtilis*, MutS and mismatches were shown to alter localization of an essential DNA polymerase (39). These results suggest that MutS is spatially coordinated with active replisomes in these systems. In addition to possible spatial coupling between MMR and DNA replication, *S. cerevisiae* MMR was shown to be defective when Msh6 was unavailable during S phase (DNA replication), supporting the importance of temporal coupling of MutSα to DNA replication in eukaryotes (40).

Studies in various model organisms indicate that DNA processivity clamps function in MMR (35, 37, 38, 41-44). Processivity clamps exist as either a homodimer in bacteria (DnaN) or a homotrimer in archaea and eukaryotes (PCNA) (45-47). These clamps are loaded onto the 3′ termini of DNA by the clamp loader complex [e.g. (48-50)], and once loaded, DnaN and PCNA confer processive activity to replicative polymerases by tethering the polymerase to the DNA template (51, 52). MutS homologs contain a conserved DnaN clamp-binding motif, or PIP box [PCNA Interacting Protein] in eukaryotes, that mediates interactions between MutS proteins and their cognate processivity clamps (35, 37, 38, 41-44, 53). Studies in *B. subtilis* showed that deletion of the unstructured region of MutS (MutS800) containing a putative DnaN clamp-binding motif, reduced interaction with DnaN, yet MutS800 maintained the ability to preferentially bind mismatched DNA *in vitro* (35). *In vivo*, the *mutS800* allele eliminated functional MMR, and when translationally fused to *gfp*, failed to form foci demonstrating that although proficient in mismatch detection, MutS800 was defective for forming repair complexes *in vivo* (35). Recent work in *S. cerevisiae* demonstrates that PCNA-associated MutSα accounts for 10-15% of MMR *in vivo*, and that Msh6-GFP (MutSα) foci are dependent upon interaction with PCNA through the Msh6 PIP box (37).

Processivity clamps are also proposed to function in downstream steps of MMR, such as facilitating activation of endonuclease activity in MutL homologs (54-57) and in re-synthesis of the gap in DNA following strand excision [for review (58, 59)]. While it is clear that interactions between MutS and processivity clamps play a role in MMR, important questions remain about their significance and the mechanism(s) by which clamps might influence MMR *in vivo*. Three models have been used to explain the role of processivity clamps in MMR: 1) that clamps stabilize MutS at a mismatch (35, 41, 60), 2) clamps recruit MutS to sites of DNA replication (37, 38), or 3) that clamps are required for DNA synthesis and do not have an earlier role during MMR (61).

In vivo studies of DNA replication in *B. subtilis* show that the DnaN processivity clamp exists in a "clamp zone" immediately following the progressing replication forks. DnaN clamps are retained on nascent DNA during Okazaki fragment maturation and accumulate until a steady state level is reached between actively loaded and unloaded clamps (1). Because DnaN clamp zones trail the replication fork, these zones have the potential to serve as platforms that maintain the spatial and temporal relationship between mismatch recognition and active replication forks. In this work, we used several separation-of-function MutS mutants that are defective in either mismatch detection or DnaN binding to determine when and where during repair the MutS•DnaN interaction is mechanistically significant in live cells. Using functional *mutS-gfp* fusions expressed from the *mutS* native locus, we report that DnaN clamp zones position MutS at newly replicated DNA prior to, and independent of, mismatch binding. After mismatch detection, MutS no longer remains coincident with the replication machinery, instead localizing to sites of repair. Importantly, ~90% of MMR *in vivo* is initiated through DnaN-clamp-zones, revealing a heavy reliance by MutS on the clamp during the initial steps of repair. We used the MutS800 mutant to uncouple MMR from DnaN and found that this mutant could account for only ~10% of *in vivo* repair. Remarkably, we were able to restore DnaN-independent MMR to wild-type levels by increasing the cellular levels of MutS800, illustrating that the functional significance of the DnaN•MutS interaction lies in maximizing the efficiency of mismatch detection *in vivo*. Ultimately, by having MutS bind to DnaN clamp zones that closely trail replication forks, MMR and DNA replication

become tightly coupled, allowing for efficient mismatch detection, MutL activation and subsequent repair in *B. subtilis* cells.

Results

B. subtilis MutSF30A is mismatch repair deficient due to loss of mismatch binding specificity.

In order to determine if mismatch binding is necessary for MutS localization, we monitored MutSF30A, which has a mutation that should abolish mismatch recognition (62). During high affinity interaction between MutS and mismatched DNA, the phenylalanine residue in the conserved GXFY$(X)_5$DA motif stacks with the mismatched or unpaired base (**Figure 5A**) (63, 64). Substitution of phenylalanine to alanine eliminates mismatch detection *in vitro* and functional MMR *in vivo* in several organisms (62, 65, 66). This mutation does not disrupt the ATPase mechanism, indicating that MutSF30A activities other than mismatch binding are unaffected (67).

We tested the corresponding *mutSF30A* mutation for the ability to support both mismatch binding *in vitro* and functional repair in *B. subtilis*. We purified *B. subtilis* MutSF30A using standard chromatography techniques without the use of an affinity tag (**Figure 5B**). We found that purified MutS binds a T-bulge DNA substrate (containing an extrahelical thymidine) selectively and with a K_d of 24 nM, while MutSF30A shows little binding to either a T-bulge or a homoduplex DNA substrate precluding us from calculating a K_d (**Figure 5C**). Furthermore, we verified that the F30A mutation did not have an adverse affect on MutS binding to DnaN. An immunodot blot analysis shows comparable retention of DnaN by MutS and MutSF30A (**Figure 5D**).

MutSF30A function was also tested *in vivo* by introducing an unmarked *mutSF30A* allele at the native *mutS* locus by allelic replacement (see “Experimental Procedures”). Immunoblot analysis confirmed that the mutant MutS protein, as well as the downstream gene product MutL, accumulated to wild-type levels *in vivo* (**Figure 5E**). Using spontaneous rifampicin resistance as an indicator for mutation rate, we found that the *mutSF30A* allele conferred a mutation rate of 155.4×10^{-9} mutations/generation, significantly higher than the mutation rate of the wild type strain, which was 1.82×10^{-9} mutations/generation (**Table 1**). The *mutSF30A* mutation rate

was indistinguishable from a strain with the *mutSL::spec* allele, which eliminates all *in vivo* MMR, showing an ~85 fold increase in mutagenesis to 154.5 x 10^{-9} mutations/generation. With these data we conclude that the *B. subtilis mutSF30A* allele is MMR defective *in vivo* due to a loss in mismatch binding specificity.

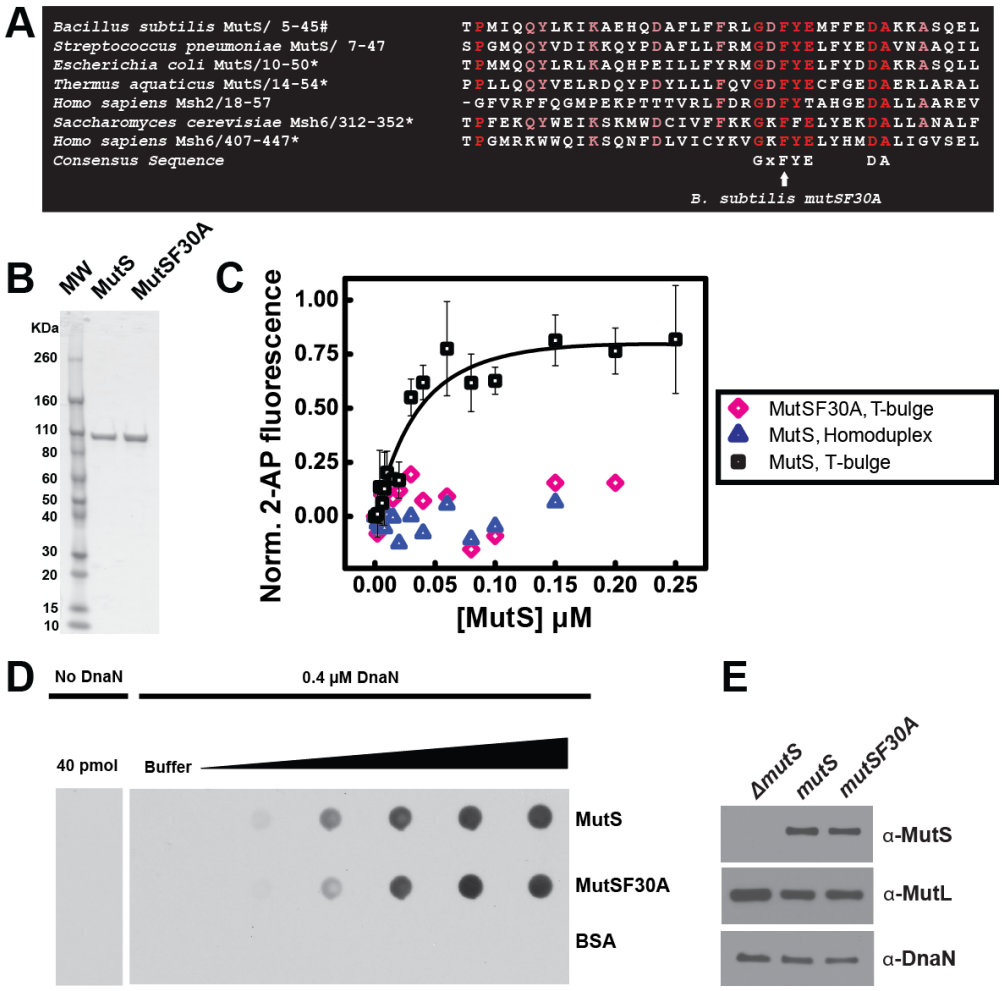


Figure 5. MutSF30A is unable to bind mismatches *in vitro*.
(A) The following conserved motif necessary for mismatch binding (GXFYXXXXXDA) is found in a wide range of eukaryotic and prokaryotic MutS homologs. Indicates (*) organisms previously demonstrated that substitution of the conserved phenylalanine residue (F) eliminates mismatch binding *in vitro* and prevents MMR *in vivo*. **(B)** 1 μg of purified MutS and MutSF30A protein electrophoresed on a 4-15% SDS-PAGE gradient gel. **(C)** *In vitro* binding of MutS and MutSF30A to T-bulge substrate. Legend: black squares and pink diamonds show MutS and MutSF30A interaction with T-bulge containing DNA, respectively, and blue triangles show MutS interaction with homoduplex DNA. **(D)** An immunodot blot (far western) analysis was performed to monitor interaction between MutS or MutSF30A with DnaN. Purified MutS and MutSF30A were blotted onto a nitrocellulose membrane over a range of 0.625 pmol to 40 pmols of dimer. Purified DnaN was incubated with the membrane and probed with affinity-purified antisera against DnaN in a 1:500 dilution. Shown (left most blot) is the purified antisera control against purified MutS and MutSF30A. Shown (right most blot) is the retention of DnaN by MutS and MutSF30A as described in "Experimental Procedures." **(E)** *In vivo* steady state levels of MutS, MutL, and DnaN. A total of 5 μg of cell extract derived from the indicated strain was electrophoresed and immunoblotted in the indicated lanes.

Table 1. *mutSF30A* is defective for mismatch repair *in vivo*.

Relevant genotype	No. of cultures	Mutation rate (10^{-9} mutations/generation)±[95%CI]	Relative mutation rate	Total MMR activity (%)
PY79 (wild type)	51	1.82 [1.14-2.37]	1	100%
mutSL::spc	23	154.4 [146.6-162.2]	84.9	0%
mutSF30A	24	155.9 [147.5-163.3]	85.5	-0.7%
dnaN5	30	39.2 [33.7-44.7]	21.6	75.5%

Mismatch repair proficiency and analysis of the mutation rate of the *mutSF30A* strain compared to wild type cells and MMR deficient cells. The bracketed values represent the lower and upper bounds of the 95% confidence limit.

MutSF30A-GFP forms foci on DNA independent of mismatch binding.

After demonstrating that *B. subtilis* MutSF30A is defective for mismatch binding, we sought to determine whether mismatch binding is a prerequisite for localization of MutS into discrete foci *in vivo*. We first determined that the *mutS-gfp* native locus allele exhibits ~90% of wild-type MMR activity, indicating that the *gfp* fusion has little impact on MutS function *in vivo* (**Table 6:Ap**). MutS-GFP foci were detected in only ~9% of untreated, exponentially growing cells, whereas treatment with the mismatch forming agent 2-aminopurine (2-AP) resulted in >45% of cells with MutS-GFP foci (**Figure 6A and C**). A similar increase in MutS-GFP foci occurred following introduction of a DNA polymerase mutant, *polC mut-1* defective in proofreading (herein referred to as $polC^{exo-}$), which substantially increases the frequency of errors during DNA replication (68). MutS-GFP focus formation was observed in ~25% of cells when $polC^{exo-}$ was the sole source of the replicative DNA polymerase in the cell (**Figure 6C**). These results demonstrate that MutS-GFP focus formation responds to natural mismatches formed by normal bases during DNA replication.

To test if mismatch binding was a prerequisite for MutS localization, we built a *mutSF30A-gfp* reporter allele at its native locus. Strikingly, MutSF30A-GFP formed foci during exponential growth in ~6% of cells (**Figure 6A and C**). As a control, we determined via immunoblot that the cellular level of MutS and MutSF30A-GFP were indistinguishable (**Figure 6B**). When MutSF30A-GFP cells were challenged with 2-AP, we did not observe an increase in foci above ~6% (**Figure 6C**) supporting our data that MutSF30A is defective in mismatch recognition. Furthermore, MutL-GFP focus formation in *mutSF30A* cells was not stimulated by 2-AP treatment, indicating that MutS

must be able to bind a mismatch in order to efficiently activate the downstream steps of repair *in vivo*, including recruitment of MutL (**Appendix 1 Results and Figure 24**). Thus, MutSF30A fails to detect and respond to 2-AP formed mismatches *in vivo*, consistent with the above data showing loss of mismatch binding *in vitro* (**Figure 5**).

Interestingly, we did observe a small, though statistically significant difference in the percent of cells with foci between untreated *mutSF30A-gfp* and *mutS-gfp* cells (**Figure 6C and Figure 23:App**). This result is not explained by differences in binding of MutS or MutSF30A to DnaN since both proteins bind DnaN equally. Because we expect a small subset of cells to undergo MMR in the functional *mutS-gfp* background, we suggest that the slightly greater percentage of cells with MutS-GFP foci relative to the MutSF30A-GFP in untreated cells represents MutS-GFP foci engaged in repair. We conclude that MutS forms two types of foci: one licensed by mismatch detection and one that is mismatch-detection independent. Together our results demonstrate that MutS forms foci on DNA independent of, or prior to, mismatch binding in live cells.

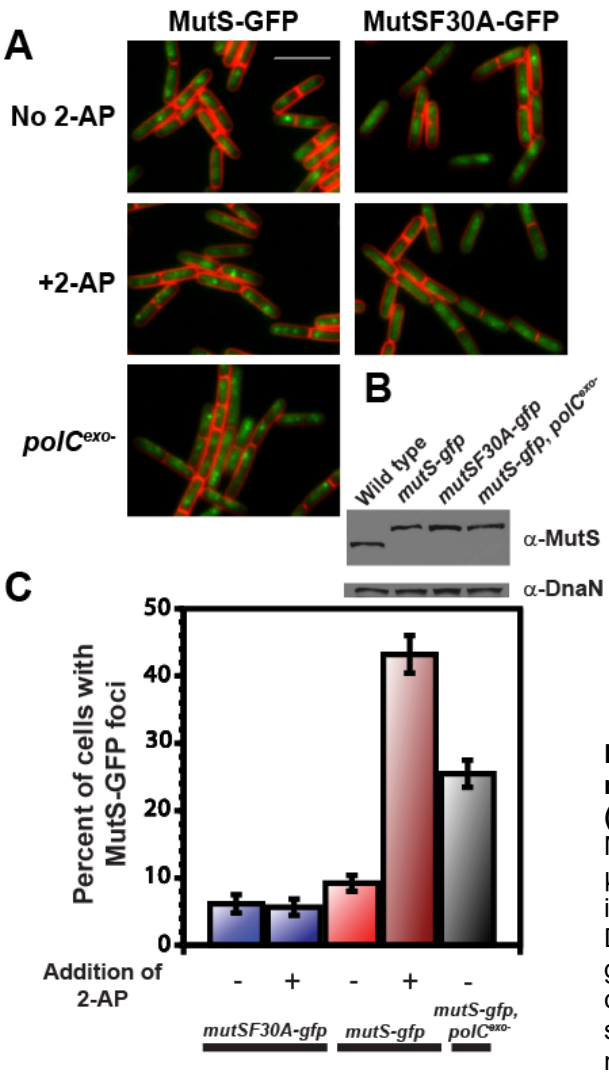
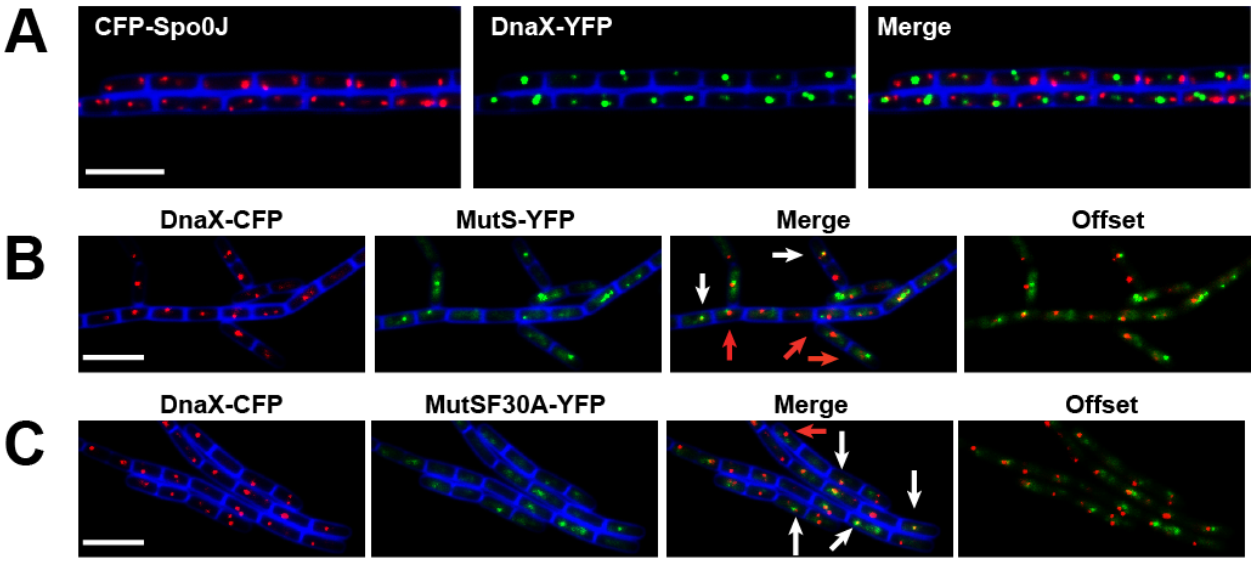


Figure 6. MutS-GFP forms foci independently of mismatch formation.
(A) Representative images of MutS-GFP or MutSF30A-GFP foci in cells with or without 600 μg/mL of 2-AP. The white bar is 4 μm. **(B)** An immunoblot of the indicated MutS derivative and DnaN as a loading control. **(C)** Bar graph of the groups represented in (A) showing the percent of cells with MutS-GFP foci. Total number of cells scored for each condition was from left to right: n=1234, 1410, 2380, 1222, and 1797.



D

Percent colocalization to DnaX

Fusion	Addition of 2-AP	% colocalization (+/- 95% CI)	No. of cells analyzed	
MutS-YFP	-	55.6% (7.7)	160	*
	+	35.5% (5.6)	279	
MutSF30A-YFP	-	73.4% (6.7)	169	** ***
	+	70.1% (8.7)	107	
MutL-GFP	-	58.5% (8.5)	130	#
	+	49.5% (7.1)	190	
CFP-Spo0J	-	12.8% (3.8)	297	

Figure 7. MutS-GFP colocalizes with the replisome prior to mismatch detection. **(A)** Representative images of CFP-Spo0J and DnaX-YFP colocalization with the replisome. **(B)** Representative images of colocalization between MutS-YFP with DnaX-CFP following 2-AP treatment. White arrows denote MutS-YFP foci that colocalize with the DnaX-CFP, whereas red arrows denote MutS-YFP foci that do not colocalize with DnaX-CFP **(C)** Representative images of colocalization between MutSF30A-YFP with DnaX-CFP following 2-AP treatment. The vital membrane stain TMA-DPH is shown in blue. The white bar is 4 μm. **(D)** Scoring of colocalization of MutS-YFP and MutL-GFP at the replisome in presence and absence of 2-AP, p-values are one-tailed; p= * 2.03×10^{-5}; **2.77; ***4.77×10^{-10}; # 0.0568

MutS is staged at active replisomes prior to mismatch recognition.

We investigated the localization dynamics of MutS foci before and after mismatch detection in order to better understand the spatial-temporal coupling of MutS to the replisome. We define the replisome as replication associated proteins (replicative polymerases, clamp loader components, processivity clamp, etc.) that localize as discrete foci at the replication forks *in vivo*. In *B. subtilis*, the replisome occupies

characteristic subcellular positions denoting the site of DNA synthesis (69-71). Immediately after replication initiation from *oriC*, origin regions and replicated DNA translocate away from the replisome toward the opposite cell poles (72, 73). Previously, it was shown that *B. subtilis* MutS-YFP colocalizes with the replisome in ~48% of cells (36). As shown above, the percent of cells with MutS-GFP foci increase following 2-AP treatment (**Figure 6**) (35, 36, 39), indicating that more repair complexes are formed. As DNA replication progresses, the newly replicated chromosomal DNA moves toward the cell poles (72), presumably taking replication errors away from the replisome. Thus, we hypothesized that the MutS-GFP foci would initially associate with mismatches in DNA at or near the replisome, and as repair and replication continued, the mismatch•MutS complex would move toward the cell poles, reducing colocalization with the replisome.

Initially, we tested our ability to spatially resolve replicated DNA from the replisome by monitoring the colocalization of CFP-Spo0J with DnaX-YFP. We grew cells slowly so that most cells would have one or two replisome foci during this analysis (**Figure 25:App**) (see "Experimental Procedures"). Spo0J, which localizes to and helps organize the origin of replication (*oriC*), should only colocalize with the clamp loader protein DnaX during replication of the origin region, and then translocate away from the replisome to the cell pole (74, 75). We found that in cells containing a single DnaX-mYFP focus, only 12.8% of replisome foci colocalized with CFP-Spo0J (n=297). Furthermore, inspection of these cells shows that most single replisome cells contain the expected origin-replisome-origin localization pattern along their longitudinal axis (**Figure 7**).

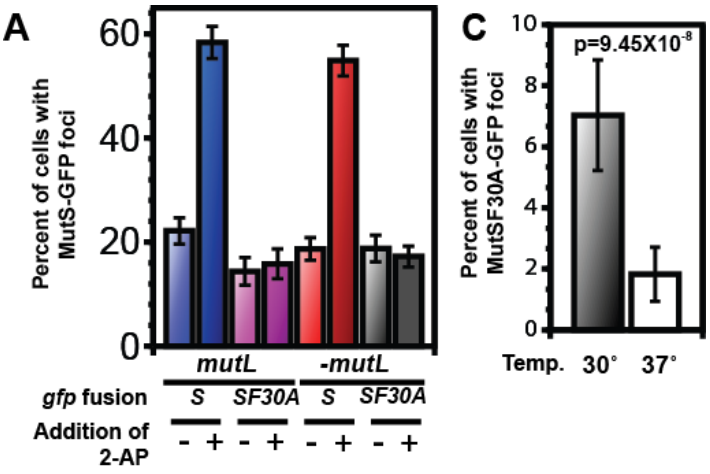
To test our hypothesis that MutS moves away from the replisome after mismatch binding, we introduced functional *dnaX-cfp* and *mutS-yfp* alleles (>90% MMR activity) (**Table 6:App**) into *B. subtilis* cells, with both fusions placed at their native locus and under control of their native promoters. During exponential growth, MutS-YFP foci colocalized with the replisome in ~56% of cells containing at least one DnaX-CFP focus and one MutS-YFP focus (**Figure 7B**). When cells were treated with 2-AP to form mismatches, we observed a significant decrease in colocalization to ~35% ($p = 2.03 \times 10^{-5}$) (**Figure 7B and D**). These data support the hypothesis that mismatch recognition by MutS-YFP reduces colocalization with the replisome.

When the same experiment was performed with *mutSF30A-yfp*, we observed that for the foci that form MutSF30A-YFP foci colocalize with the replisome ~73% of the time in the absence of 2-AP challenge (**Figure 7C and D**). When this strain was treated with 2-AP, there was no significant statistical difference in the position of MutSF30A-YFP foci compared with the untreated group (~70.1% colocalized: p=0.277). Thus, lacking the ability to detect mismatches in DNA, MutSF30A-YFP remains colocalized with the replisome. These results lead us to conclude that MutS-GFP foci, when not bound to mismatches, are staged near the active replisome, possibly due to physical coupling with a replication protein. Subsequently, upon encountering a mismatch, MutS disengages from the advancing replisome and remains behind on nascent DNA to direct the remaining steps in repair. This result provides insight into how mismatch recognition affects the dynamic association of MutS with the replisome *in vivo*.

Based on this model, we hypothesized that if MutS is positioned on newly replicated DNA through interaction with a replisome protein, then increasing expression of MutS or MutSF30A should increase the number of mismatch-independent foci by promoting this interaction *in vivo*. To this end, we constructed an in frame *mutS* deletion that maintains transcriptional control of *mutL* from its native promoter (**Figure 5D & Figure26:App**). We then expressed MutS-GFP or MutSF30A-GFP from an ectopic locus driven by an IPTG regulated promoter (P_{spac}). The *ΔmutS, amyE::*P_{spac}*mutS-gfp* strain was 88.7% functional compared to *ΔmutS, amyE::*P_{spac}*mutS* (**Table 6:App**). When either *mutS-gfp* or *mutSF30A-gfp* was ectopically expressed, we observed a 2-3 fold increase in the percent of untreated cells with foci (**Figure 8A, compare with Figure 6C**). This result was not affected by the presence or absence of *mutL*.

We then asked if increased expression of *mutS-gfp* and *mutSF30A-gfp* and the associated mismatch independent foci correlated with colocalization with the replisome marker *dnaX-mcfp*. We found that ectopic expression caused an increase in colocalization to ~65%, (**Figure 8B**). When these cells were challenged with 2-AP, we expect a decrease in colocalization and indeed found ~41% were colocalized following 2-AP challenge (**Figure 8B**). Ectopic expression of MutSF30A increased the percent of cells with foci, and colocalization of MutSF30A, which only forms mismatch-independent foci, remained at ~70% upon ectopic expression. These results show that increased

expression of MutS increases the percentage of cells with foci (**Figure 8A**), supporting the hypothesis that MutS is positioned at the replisome via a binding partner prior to mismatch identification.



B

Colocalization following ectopic expressing MutS or MutSF30A-GFP

Genotype	Addition of 2-AP	Percent of colocalization (95% CI)
ΔmutS, amyE:P$_{spac}$mutS-gfp, dnaX - mCherry	-	65.1 (9.1)
	+	40.8 (6.9)
ΔmutS, amyE:P$_{spac}$mutSF30A-gfp, dnaX - mCherry	-	68.7 (7.1)
	+	70.6 (7.7)

Figure 8. Elevated expression of MutS-GFP or MutSF30A-GFP increases replisome-associated foci. **(A)** Elevated expression of *mutS-gfp* and *mutSF30A-gfp* increases the percentage of cells with MutS foci. The first four groups contain a *ΔmutS* deletion and the expressed GFP fusion represents the sole copy of *mutS* within the cell. The second four groups represent a deletion of the entire *mutSL* locus with *mutS-* or *mutSF30A-GFP* expressed ectopically from a P_{spac} promoter **(B)** Elevated expression of *mutS-gfp* causes an increase in colocalization of MutS foci to DnaX-mCherry foci. Total number of cells scored is 106-196. **(C)** MutSF30A-GFP foci expressed from the *mutS* native locus form in a higher percentage of cells at a lower temperature for DNA replication in a *dnaN5* background (30 ºC).

It should be noted that in our colocalization experiments we used DnaX as a replisome marker instead of DnaN, which binds MutS *in vitro*, because the *dnaN-mcfp* fusion maintains an elevated mutation frequency (25.3 x 10^{-9} mutations/ generation) whereas *dnaX-mcfp* is wild type for mutation rate (**data not shown**). We determined that the smaller DnaX-mCherry foci colocalizes with DnaN-GFP foci in ~89% of cells, establishing DnaX as an appropriate substitute for DnaN in this analysis (**Figure 27:App**).

MutSF30A-GFP foci are positioned at the replisome by DnaN prior to mismatch binding.

The ability of MutSF30A-GFP to assemble into foci without mismatch identification suggests that MutS may be positioned at the replisome as a mechanism to spatially target newly formed mismatches in DNA. Furthermore, colocalization of MutSF30A with the replisome and MutSF30A binding to DnaN *in vitro* (**Figure 5E**) suggests that MutSF30A localization is dependent on an interaction with a protein component of the replisome. In *B. subtilis,* DnaN forms large clamp assemblies termed “clamp zones” that form behind progressing replication forks (1). DnaN clamp zones contain ~200 accumulated clamps as clamp loading and unloading rates achieve equilibrium (1). We hypothesized that a clamp zone facilitates the formation of mismatch-independent foci by recruiting MutS to the replisome via contact with DnaN. To test this hypothesis, we took advantage of the *dnaN5* allele, which exhibits an increase in mutation frequency due to partial loss of MMR [(35, 76) **and Table 6:App**]. The *dnaN5* allele exhibits a temperature sensitive defect in MMR, which leads to a significant decrease in MutS-GFP focus formation at 37^{o}C relative to 30^{o}C. We determined that DnaN5 functions normally in DNA replication by measuring replication *in vivo* and found that *dnaN5* is wild type for DNA synthesis and growth rate at both 30^{o}C and 37^{o}C (**Figure 28:App**). We introduced *dnaN5* into a strain bearing the *mutSF30A-gfp* allele at its native locus and scored the number of MutSF30A-GFP foci at 30°C and 37°C. At 30°C, we found that ~7% of cells contain a MutSF30A-GFP focus, results consistent with that of a *dnaN*$^{+}$ strain (**Figure 8C**). In contrast, at 37°C, the percentage of cells with MutSF30A-GFP foci decreased to <2% ($p=9.45 \times 10^{-8}$). Two prominent models have been used to explain the role of DnaN in mismatch repair. The first model suggests that MutS is stabilized by the replication clamp after mismatch recognition (35, 41, 60). The second model supported from studies in human cell culture and *S. cerevisiae* predicts that MutS is recruited to sites of replication prior to mismatch binding (37, 38). Our data show that interaction with DnaN is critical for formation of the mismatch-independent MutS foci *in vivo.* We further interpret these results to mean that DnaN clamp zones stage MutS immediately behind the advancing replication forks *in vivo* supporting the model that MutS is recruited to sites of replication before mismatch binding.

Table 2. DnaN allows for efficient mismatch repair *in vivo*.

Relevant genotype	No. of cultures	Mutation rate (10^{-9} mutations/generation)±[95%CI]	Relative mutation rate	Total MMR activity (%)
ΔmutS, amyE::P_{spac}mutS	25	3.09 [1.35-4.68]	1	100
ΔmutS	28	159.3 [152.0-166.6]	51.5	0
mutS800	25	144.5 [134.9-154.2]	46.7	9.5
ΔmutS, amyE::P_{spac}mutS800	23	7.10 [4.39-9.74]	2.3	97.4

The *ΔmutS* designation indicates an in frame deletion of *mutS*, maintaining a functional *mutL* at its native locus (see "Materials and Methods"). The *mutS800,* allele was expressed from its native locus with *mutL* expressed ectopically from *amyE* using 1 mM IPTG. Brackets enclose the lower bounds and upper bounds of the 95% confidence limits. Percent MMR activity was determined using the following equation: [(R.M.R.null - R.M.R.strain)/(R.M.R.null-R.M.R.wild type)]•100. RMR=relative mutation rate. Mutations per culture (m) are as follows: *ΔmutS, amyE::P_{spac}mutS* (1.78); *ΔmutS* (104.5); *mutS800* (64.2); *ΔmutS, amyE::P_{spac}mutS800* (5.0).

DnaN clamp zones increase efficiency of mismatch detection by targeting MutS to nascent DNA.

In defined *in vitro* MMR systems, purified MutS and MutSα can detect a mismatch in DNA without the need for a processivity clamp [e.g. (77, 78)]. We hypothesized that the association of MutS with DnaN might be necessary *in vivo* in order to restrict the search for mismatches to nascent DNA, making mismatch detection more efficient relative to MutS identifying a mismatch independent of DnaN binding. To test this key hypothesis, we took advantage of the *mutS800* allele, which lacks a C-terminal tether and is defective in DnaN binding but proficient for mismatch identification (35). When the *B. subtilis mutS800* allele was expressed from its native promoter, only ~9.5% of MMR activity is observed *in vivo* (**Table 2**). MMR activity of the *mutS800* allele recovers to 97.4% when this mutant protein is overexpressed from an IPTG driven P_{spac} promoter from an ectopic locus in a strain lacking the native *mutS* locus (**Table 2**). Immunoblotting shows that the P_{spac}*mutS800* protein level was 4-fold higher than the level produced from the *mutS800* allele located at the *mutS* native locus (**Figure**

29:App). This result supports the hypothesis that increasing MutS800 concentration can compensate for the loss of interaction with DnaN and restore efficient MMR activity.

When *mutS800-gfp* is expressed from the *mutS* native locus, the protein is defective in forming foci in response to mismatches *in vivo* (35). Since *mutS800* restores MMR activity when ectopically expressed, we asked if *mutS800-gfp* also forms foci following ectopic expression. Upon visualizing ectopically expressed MutS800-GFP in a *ΔmutS* background, we found that MutS800-GFP formed occasional foci in untreated cells (**Figure 9A**); however, focus intensity was barely above the elevated background fluorescence and only formed in ~3.3% of cells. Upon 2-AP treatment, the percent of cells with faint foci substantially increased to 14.1%, indicating that like MutS-GFP (**Figure 6**), MutS800-GFP focus formation is responsive to an increase in mismatches in DNA. To verify this observation, we asked if mismatch binding by MutS800 was important for focus formation. To answer this question, *mutSF30A800-gfp* was placed under the control of an IPTG driven P_{spac} promoter and inserted at an ectopic locus in a Δ*mutS* background. This allele is defective in both DnaN clamp binding and mismatch detection. We predicted that overexpressed MutSF30A800 would fail to localize into foci if the observed localization of ectopically expressed *mutS800* was dependent on mismatches and independent of DnaN. Indeed, we found that MutSF30A800 was blocked for focus formation (**Figure 9B**) (<1% in both 2-AP treated and untreated samples). We conclude that when the DnaN tether on MutS is removed, mismatch binding becomes obligatory for focus formation *in vivo* (**Figure 9A, B, C**). We further verified this observation by inserting the *dnaN5* allele into the *ΔmutS, amyE::*P_{spac}*mutS-gfp* background. At both 30°C and 37°C, MutS800-GFP formed foci in ~14% of cells, consistent with our results for the *dnaN*$^{+}$ allele (**Figure 8C**) and further confirming bypass of the DnaN role in MMR following overexpression (**Figure 9C, D**).

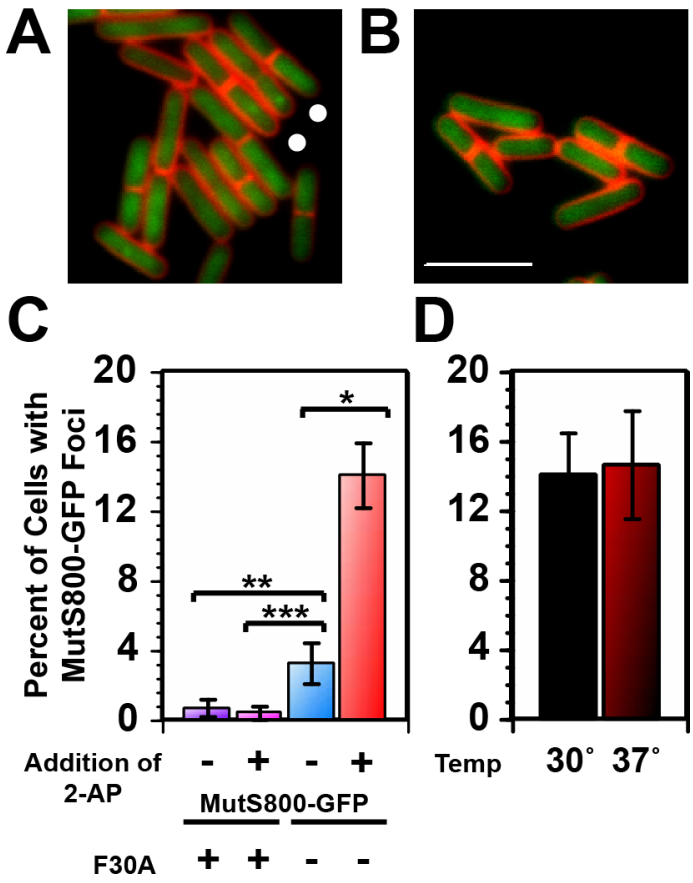


Figure 9. Mismatch detection by MutS800-GFP induces focus formation at nascent DNA when ectopically expressed.
(A) MutS800-GFP foci form in response to mismatches independent of DnaN (faint foci indicated by white circles) **(B)** MutSF30A800 fails to form foci. The vital membrane stain TMA-DPH is shown in red. The white scale bar is 4 μm. **(C)** Bar graph of ectopically expressed MutS800-GFP and MutSF30A800-GFP foci with and without 2AP. From left to right, total scored cells: 1114, 1154, 883, 1343. P-values are 1-tailed: *p=2.69x10^{-17}, **p=1.20x10^{-5}, ***p=3.22x10^{-7}. **(D)** Bar graph of ectopically expressed MutS800-GFP foci within the *dnaN5* background revealed no statistical difference at 30 ºC and 37°C (p=0.38).

Since overexpression of *mutS800* restores MMR to near wild type levels, bypassing the need for DnaN (**Table 2**), we asked where ectopically expressed MutS800 foci form *in vivo.* To address this question, we visualized and scored the subcellular location of ectopically expressed MutS800-GFP in comparison to natively expressed MutS-GFP. Upon scoring focus positions in the cell relative to the closest pole, we found that MutS800-GFP foci formed in the same subcellular positions as MutS-GFP foci (**Figure 10A**). Moreover, following 2-AP challenge, MutS800-GFP colocalizes with DnaX-mCherry to almost the same extent as wild-type MutS-GFP (**Figure 10B**, 29.9±6.28% for MutS800-GFP and 35.5±5.6 for MutS-GFP: P=0.099). Finally, the vast majority of foci that do not colocalize with the replisome are replisome proximal (**Figure 10C**). These results indicate that ectopically expressed MutS800-GFP also localizes to replisome proximal DNA for mismatch detection and initiation of repair; however, as shown previously, higher amounts of this mutant protein are required to achieve the same level of MMR as wild-type MutS (**Table 2**). Thus, MMR in *B. subtilis* is initiated and occurs predominantly in replisome proximal regions of DNA, and association of MutS with DnaN increases the

efficiency of mismatch detection and repair by targeting MutS to replisome proximal DNA. When DnaN-mediated mismatch detection is bypassed as with MutS800 following overexpression, we found that MutS800 still forms foci in replisome proximal DNA *in vivo* (see discussion). Together, our results show that in order to initiate MMR, MutS must localize to the replisome or at least replisome proximal DNA in *B. subtilis*.

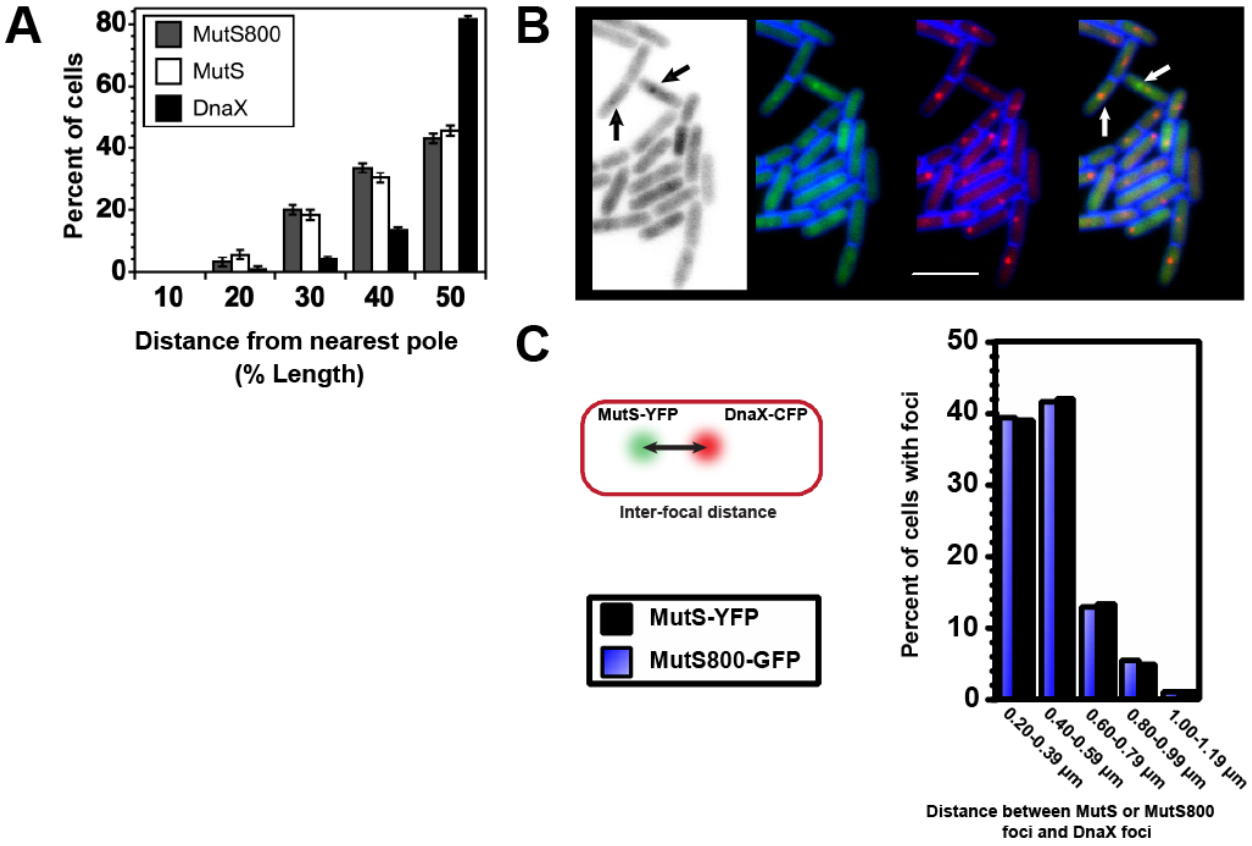


Figure 10. DnaN-independent focus formation of MutS800-GFP localizes to the same subcellular position as MutS-YFP.
(A) Position of MutS-GFP, ectopically expressed MutS800-GFP, and DnaX-GFP foci scored relative to cell length (n=125 cells for each group). **(B)** 29.9+6.3% (n=204) of the MutS800-GFP colocalizes with DnaX-mCherry. These results are not statistically different with p=0.099 when compared with MutS-YFP colocalization with DnaX-CFP 35.5+5.6 shown in Figure 3. The left most image is the negative image, MutS800-GFP, DnaX-mCherry, and a merge. The membrane is stained with TMA-DPH and is shown in blue. The scale bar represents 4 µm. **(C)** We measured the inter-focal distance (IFD) between MutS-YFP and MutS800-GFP foci that failed to colocalize with DnaX-mCherry foci. No IFDs were measured less than 0.2 µm. We measured 94 MutS800-GFP, DnaX-mCherry and 105 MutS-YFP, DnaX-CFP focal pairs that failed to colocalize. Each bar represents the percentage of cells with IFD between the indicated distance.

Discussion

Three current models are used to explain the role of processivity clamps in mismatch repair: that clamps stabilize MutS at a mismatch (35, 41, 60), clamps recruit MutS to sites of DNA replication (37, 38), or that clamps are required for DNA synthesis and do not have an earlier role during MMR (61). We have shown in this study that the *B. subtilis* processivity clamp, DnaN, facilitates ~90% of MMR and targets MutS to replisome proximal DNA prior to mismatch binding. A DnaN clamp zone forms in the wake of active replication forks (1), providing a platform for MutS to maintain a critical spatial and temporal relationship with the replisome (**Figure 11**). We propose that DnaN-mediated targeting of MutS to nascent DNA allows for efficient mismatch detection by allowing for MutS to target newly formed errors.

MutS homologs spanning bacteria to humans exhibit the near-ubiquitous presence of a clamp-binding motif, suggesting that association with processivity clamps is important for MMR (79). It has been known for decades that MutS is able to detect mismatches without accessory factors *in vitro* [e.g. (24, 26, 80)]. Nevertheless MutS800, which lacks the DnaN clamp-binding tether, is largely inactive for MMR while under control of its native promoter, retaining less than 10% activity. Interestingly, the same *mutS800* allele restored MMR activity to 97% *in vivo* upon over expression.

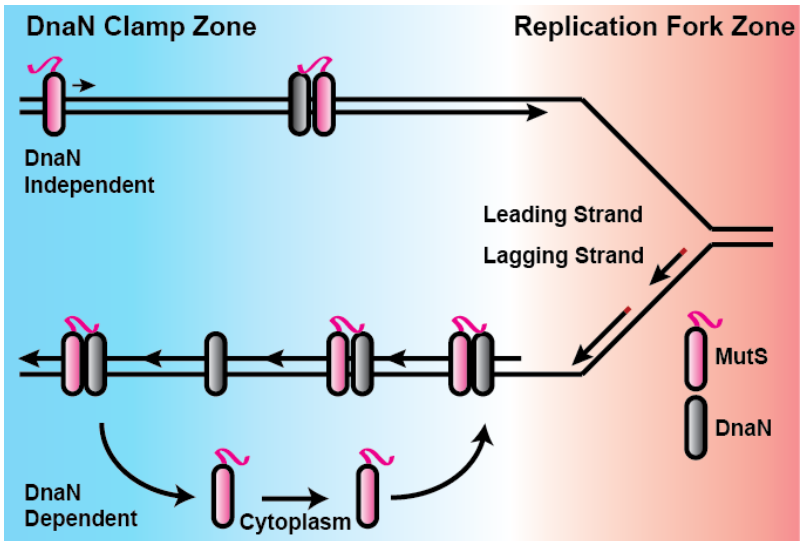


Figure 11. Model for temporal coupling of MutS to DNA replication.
MutS relies on a DnaN clamp zone- to target MutS to nascent DNA. DnaN-dependent mode of mismatch detection represents 90% of repair. This figure is adapted from the clamp zone model (1).

Our finding that MutS800 is capable of mismatch detection in the absence of DnaN *in vivo* led us to speculate that the protein could find mismatches and initiate repair at chromosomal locations distal to the replisome. To the contrary, we found that ectopically expressed MutS800-GFP formed foci at virtually identical subcellular positions as MutS-GFP, while MutSF30A800-GFP, which cannot bind mismatches or DnaN, was completely defective for focus formation. This result demonstrates that MutS800 binds mismatches in a replisome-proximal position like wild type MutS; however, a high concentration of the mutant protein is required to restore mismatch detection efficiency and compensate for the loss of interaction with DnaN. It was shown previously that *B. subtilis* MutL binds DnaN, and disruption of this contact causes complete loss of MMR *in vivo* (55, 56). Overexpression of MutL mutants defective in binding DnaN fail to bypass the need for interaction with DnaN in MMR (55, 56). We propose that MutS800 must identify mismatches in replisome proximal DNA to enable the downstream steps of repair, which include MutL recruitment and activation of its endonuclease activity. Current data suggests that these steps are dependent on interaction with DnaN, and may therefore require MutS to bind mismatches in the DnaN clamp zone in order to complete the downstream steps of MMR (55, 56).

Consistent with our findings in *B. subtilis*, in *E. coli* when *mutS800* is expressed from its native promoter it confers an MMR defect, but is close to wild type for MMR when overexpressed from a plasmid (63, 81). Similarly, the equivalent mutant in *P. aeruginosa*, *mutS798*, complements a *mutS* deficient strain when overexpressed from a plasmid (53). Thus, in these systems, MutS is also capable of operating independent of DnaN since both *P. aeruginosa* MutS798 and *E. coli* MutS800 fail to bind their cognate DnaN clamps (43, 53, 81). We propose that in these bacteria, when MutS is present at wild type levels, interaction with DnaN is important for targeting MMR to nascent DNA. When the protein is overexpressed, the requirement for DnaN is bypassed due to the increased likelihood of MutS directly binding mismatches in nascent DNA without DnaN association.

The model of a DnaN clamp zone that facilitates spatial and temporal coupling of mismatch detection with replication is an intriguing one, especially when considering the

conservation of processivity clamp-binding motifs in MutS homologs (41, 79). Consistent with the clamp zone model, fluorophore-labeled processivity clamps form foci *in vivo* in bacteria and eukaryotes [e.g. (37, 38)], suggesting that substantial local concentrations of clamps are present in organisms other than *B. subtilis*. The observation that a PCNA clamp zone may exist also agrees well with the higher proficiency of MMR on the lagging strand relative to the leading strand in *S. cerevisiae* (82). Other *B. subtilis* studies have shown that DnaN clamps are competent for protein recruitment to nascent DNA *in vivo*, as fluorophore-labeled peptides encoding a DnaN-binding motif are sufficient for forming replisome-localized foci (1, 35). A similar finding was reported for *S. cerevisiae msh6 305-1242Δ*, the unstructured region of *msh6* that contains the PIP motif for binding PCNA (^{27}-**Q**SS**L**LS**FF**-^{34}), when expressed from the native *msh6* promoter (37). Furthermore, in human cell culture, overexpression of either MSH6 or MSH3 (which contain PIP motifs ^{4}**Q**ST**L**YS**FF**-^{11} and ^{21}-**Q**AV**L**SR**FF**-^{28}, respectively) resulted in colocalization of MutSα and MutSβ with both PCNA and BrdU stain (38). Truncating the N-terminal 77 residues of MSH6 eliminated both MSH6 binding to PCNA *in vitro* and focus formation *in vivo*, indicating that localization of human MutS homologs to nascent DNA is dependent on interaction with PCNA (38). Collectively, these results support the hypothesis that a PCNA clamp zone present at nascent DNA facilitates mismatch detection *in vivo*. Our data further agrees well with the observation that MMR must occur concurrently with DNA replication in *S. cerevisiae* (40). It was recently reported in *S. cerevisiae*, that PCNA-dependent MMR accounts for only 10-15% of MMR (37), whereas DnaN-dependent MMR in *B. subtilis* accounts for ~90% of MMR by MutS. This is a notable difference in the orchestration of MMR between these organisms *in vivo*.

Another important finding from our study is that MutS localizes near the replisome independent of mismatch identification through the DnaN clamp zone. This conclusion is based on the observation that MutSF30A also forms foci that colocalize with the replisome, despite its inability to bind mismatches. Moreover, the MMR compromised *dnaN5* mutant nearly abolishes MutSF30A-GFP localization in *B. subtilis*, indicating that MutSF30A foci are dependent on interaction with DnaN. Based on these data, we propose that MutS is coupled with the progressing replication fork prior to

mismatch identification. An additional important finding is that after detecting a mismatch, MutS appears to detach from the replisome (DnaN) and remains at the mismatch site to conduct repair. This conclusion is based on the observation that MutS-YFP colocalizes less frequently with the replisome in 2-AP treated cells even though the percentage of cells with MutS foci increase following 2-AP treatment. Moreover, overexpression of MutS-YFP increases colocalization to the replisome in untreated cells, but there is little increase in colocalization following 2-AP treatment. Finally, the frequency of MutSF30A colocalization with the replisome is unchanged following 2-AP treatment, again, supporting the hypothesis that binding to mismatches causes release of MutS from the replisome. Overall, this study shows that MutS foci represent assemblies undergoing active repair of replication errors with two distinctive steps clearly identified: DnaN-coupled targeting of MutS to nascent DNA and release of MutS from the replisome following mismatch recognition.

Experimental Procedures

Bacteriological methods

All *B. subtilis* strains (**Table 7:App**) are isogenic derivatives of PY79 and grown according to standard procedures (83). To determine relative mutation rate, *B. subtilis* cells were grown at 37°C to OD_{600} of ~1.2, concentrated, and resuspended in 100 μL of 0.85% saline. A portion of the cells was serial diluted (10^{-6}) and plated onto LB agar plates to determine total the viable count within the culture. The remaining resuspension was plated onto LB agar plates supplemented with 150 μg/mL rifampin. Mutation rate analysis was performed using MSS Maximum Likelihood Method with the 95% confidence interval, and statistical significance assessed using a one- or two-tailed T-test (84-86).

Epifluorescence microscopy

Cells were prepared for live cell imaging essentially as described (87-89). Briefly, strains were inoculated to a starting OD_{600} in $S7_{50}$ minimal media supplemented with 2% D-glucose. Cells were grown past 3 doublings to an OD_{600} of 0.4-0.5 and were split: one control culture and one culture challenged with 2-aminopurine to a final concentration of

600 µg/mL for one hour. Cell membranes were visualized using the fluorescent probe TMA-DPH at a working concentration of 10 µM. Replisome fusions were imaged with 0.5-1.0s exposures while MMR fusion proteins were imaged at 1.2-2.5s exposures. Scoring of cells as containing a MutS focus is outlined in **Figure 30**. The average cell focus encompasses ~4% of the cell area with an average intensity 2-fold greater than background. Colocalization and localization experiments were conducted as above except cells were grown in $S7_{50}$ minimal media supplemented with 1% L-arabinose. These conditions were used to produce cells with predominantly 1 DnaX-GFP focus per cell (avg. is 1.72 foci per cell with 39.1% of cells containing one focus) (**Figure 27**). Image capture of both fusion proteins during colocalization experiments was performed in immediate succession and timed <2s of total capture to minimize any intracellular movement of either the MutS or replisomal fusions. For temperature release experiments, cells were grown and treated as above, but the prepared slide was incubated for 15 minutes at indicated temperature. Upon removal from the temperature-regulated chamber, slides were imaged for five minutes immediately upon removal.

Statistical analysis

Bar graphs are presented with error bars representing the 95% confidence interval, and statistical significance was determined using a one- or two-tailed T-test.

Immunoblotting

B. subtilis whole cell extracts were obtained basically as described (90). Briefly, mid-exponential phase cultures were centrifuged and lysed by sonication (20 Hz), resuspended in lysis buffer [10 mM Tris-HCl (pH 7.0), 0.5 mM EDTA, 1 mM AEBSF, and 1X Protease Inhibitor Cocktail (Thermo Scientific)], and protein concentration was determined using Pierce BCA Protein Assay Kit (Thermo Scientific). Equal amounts of total protein were applied to each lane on a 4-15% gradient gel, and protein levels were probed with purified antisera: α-MutS (MI-1042), α-MutL (MI-1044), and α-DnaN (MI-1039). Immunoblots were developed as described previously (91).

Immunodot blotting

Immunodot blotting was performed as previously described (39, 92). Briefly, indicated proteins were immobilized onto a nitrocellulose membrane with the assistance of a Bio-dot microfiltration apparatus (Bio Rad). The membrane was incubated in blocking buffer (5% dry non-fat milk, 17.4 mM Na_2HPO_4, 2.6 mM NaH_2PO_4, 150 mM NaCl, 0.05% Tween-20) at 22˚C for one hour. All subsequent washes and incubations took place in blocking buffer. After blocking, the membrane was incubated with 0.4 μM DnaN in blocking buffer for 3 hours at 22˚C. The blot was subsequently washed three times and then incubated with affinity purified α-DnaN antisera overnight at 4˚C. The blot was removed from primary antibody (MI 1038) and washed three times at 22˚C and placed in secondary antisera (1:2000 α-Rabbit) for 2 hours at 22˚C. The blot was washed 3 more times, followed by a wash in PBS (17.4 mM Na_2HPO_4, 2.6 mM NaH_2PO_4, 150 mM NaCl, 0.05% Tween-20) to remove excess milk proteins. The membrane was developed with chemiluminescence (Super-Signal Altra, Pierce) and expose to film as described (39).

***B. subtilis* MutS–DNA interactions at equilibrium**

Procedures used to purify untagged MutS and MutSF30A are detailed within the supplemental methods.

The mismatched DNA substrates for the MutS-DNA binding assay were prepared by annealing 37 nt 2-aminopurine (2-AP) labeled +T strand (5´ - TAA AGG AAG TCG TCT AT2-Ap *T*AT GGT ATG ACT AAG TGT A - 3') with 36 nt (5' - T ACA CTT AGT CAT ACC AT TAT AGA CGA CTT CCT TTA - 3') or with 37 nt (5' - T ACA CTT AGT CAT ACC AT*G* TAT AGA CGA CTT CCT TTA - 3') strands to yield 2-AP(+T) and 2-AP(GT) duplexes, respectively. The matched substrate, 2-AP(GC) was prepared by annealing 37 nt 2-AP labeled strand (5' - TAA AGG AAG TCG TCT AT2Ap CAT GGT ATG ACT AAG TGT A - 3') with a 37 nt strand (5' - T ACA CTT AGT CAT ACC ATG TAT AGA CGA CTT CCT TTA - 3'). The strands were heated to 95 °C, followed by slowly cooling to room temperature to obtain annealed duplex DNAs.

DNA binding was measured on a FluoroMax-3 fluorimeter (Jobin-Yvon Horiba Group; Edison, NJ). Titrations of 0.02 μM 2-AP(+T), 2-AP(GC) and 2-AP(GT) duplex

DNAs with 0 – 0.4 μM MutS were performed in 3 ml quartz cuvettes in DNA binding buffer (20 mM Tris-HCl, pH 7.6, 50 mM NaCl, 5 mM $MgCl_2$) at 25 °C. MutS was added incrementally to the sample, and fluorescence intensity was measured after mixing for 30 seconds (λ_{EX} = 315 nm and λ_{EM} = 375 nm). The data were corrected for intrinsic MutS fluorescence by subtracting data from parallel experiments with unlabeled DNA. Fluorescence intensity was plotted versus MutS concentration and the apparent dissociation constant (K_D) for the interaction was obtained by fitting the data to a quadratic equation:

$$[D\cdot M]=F_0+(F_{max}-F_0)\{[K_D+[D_t]+[M_t])-[(K_D+[D_t]+[M_t])^2-4[D_t][M_t]]^{1/2}]/2[D_t]\}$$

where $D\cdot M$ is the fraction of MutS•DNA, F_0 is 2-AP(+T) fluorescence in the absence of protein and F_{max} is maximal fluorescence, and D_t and M_t are total molar concentrations of DNA and MutS, respectively. The data were fit by non-linear regression using KaleidaGraph (Synergy Software)

Supplemental Material

Supplemental Material accompanying Chapter II is found in Appendix I: Supplemental Methods and Results, **Table 6 and 7**, and **Figures 23-30**.

Acknowledgements

We thank Dr. David Rudner (Harvard Medical School) and Dr. Heath Murray (Newcastle University) for strains. We would like to thank members of the Simmons lab for critical reading of this manuscript. This work was supported by grants from the Wendy Will Case Cancer Fund to L.A.S. and grants from the National Science Foundation to L.A.S. (MCB 1050948) and M.M.H. (MCB 1022203)

References

1. Su'etsugu M & Errington J (2011) The replicase sliding clamp dynamically accumulates behind progressing replication forks in Bacillus subtilis cells. *Molecular cell* 41(6):720-732.
2. Kunkel TA & Erie DA (2005) DNA mismatch repair. *Annu Rev Biochem* 74:681-710.
3. Schofield MJ & Hsieh P (2003) DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* 57:579-608.

4. Iyer RR, Pluciennik A, Burdett V, & Modrich PL (2006) DNA mismatch repair: functions and mechanisms. *Chem Rev* 106(2):302-323.

5. Culligan KM, Meyer-Gauen G, Lyons-Weiler J, & Hays JB (2000) Evolutionary origin, diversification and specialization of eukaryotic MutS homolog mismatch repair proteins. *Nucleic Acids Res* 28(2):463-471.

6. Cox EC, Degnen GE, & Scheppe ML (1972) Mutator gene studies in Escherichia coli: the mutS gene. *Genetics* 72(4):551-567.

7. Ginetti F, Perego M, Albertini AM, & Galizzi A (1996) Bacillus subtilis mutS mutL operon: identification, nucleotide sequence and mutagenesis. *Microbiology* 142 (Pt 8):2021-2029.

8. Prudhomme M, Martin B, Mejean V, & Claverys JP (1989) Nucleotide sequence of the *Streptococcus pneumoniae hexB* mismatch repair gene: Homology of *HexB* to *MutL* of *Salmonella typhimurium* and to *PMS1* of *Saccharomyces cerevisiae*. *J. Bacteriol.* 171(10):5332-5338.

9. Cooper LA, Simmons LA, & Mobley HL (2012) Involvement of Mismatch Repair in the Reciprocal Control of Motility and Adherence of Uropathogenic Escherichia coli. *Infect Immun* 80(6):1969-1979.

10. Davies BW*, et al.* (2011) DNA damage and reactive nitrogen species are barriers to Vibrio cholerae colonization of the infant mouse intestine. *PLoS Pathog* 7(2):e1001295.

11. Prunier AL*, et al.* (2003) High rate of macrolide resistance in Staphylococcus aureus strains from patients with cystic fibrosis reveals high proportions of hypermutable strains. *J Infect Dis* 187(11):1709-1716.

12. Roman F, Canton R, Perez-Vazquez M, Baquero F, & Campos J (2004) Dynamics of long-term colonization of respiratory tract by Haemophilus influenzae in cystic fibrosis patients shows a marked increase in hypermutable strains. *J Clin Microbiol* 42(4):1450-1459.

13. Watson ME, Jr., Burns JL, & Smith AL (2004) Hypermutable Haemophilus influenzae with mutations in mutS are found in cystic fibrosis sputum. *Microbiology* 150(Pt 9):2947-2958.

14. Turrientes MC*, et al.* (2010) Polymorphic mutation frequencies of clinical and environmental Stenotrophomonas maltophilia populations. *Appl Environ Microbiol* 76(6):1746-1758.

15. Denamur E*, et al.* (2002) High frequency of mutator strains among human uropathogenic Escherichia coli isolates. *Journal of bacteriology* 184(2):605-609.

16. Oliver A, Canton R, Campo P, Baquero F, & Blazquez J (2000) High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. *Science* 288(5469):1251-1254.

17. Mena A*, et al.* (2008) Genetic adaptation of Pseudomonas aeruginosa to the airways of cystic fibrosis patients is catalyzed by hypermutation. *Journal of bacteriology* 190(24):7910-7917.

18. Oliver A & Mena A (2010) Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clin Microbiol Infect* 16(7):798-808.

19. Chopra I, O'Neill AJ, & Miller K (2003) The role of mutators in the emergence of antibiotic-resistant bacteria. *Drug Resist Updat* 6(3):137-145.

20. Hamilton SR*, et al.* (1995) The molecular basis of Turcot's syndrome. *N Engl J Med* 332(13):839-847.

21. Nystrom-Lahti M*, et al.* (2002) Functional analysis of MLH1 mutations linked to hereditary nonpolyposis colon cancer. *Genes Chromosomes Cancer* 33(2):160-167.

22. Peltomaki P (2005) Lynch syndrome genes. *Fam Cancer* 4(3):227-232.

23. Fishel R*, et al.* (1993) The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis cancer. *Cell* 75:1027-1038.

24. Su SS & Modrich P (1986) Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proceedings of the National Academy of Sciences of the United States of America* 83(14):5057-5061.

25. Alani E, Chi NW, & Kolodner R (1995) The Saccharomyces cerevisiae Msh2 protein specifically binds to duplex oligonucleotides containing mismatched DNA base pairs and insertions. *Genes & development* 9(2):234-247.

26. Prolla TA, Pang Q, Alani E, Kolodner RD, & Liskay RM (1994) MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science* 265(5175):1091-1093.

27. Habraken Y, Sung P, Prakash L, & Prakash S (1996) Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Current biology : CB* 6(9):1185-1187.

28. Palombo F*, et al.* (1996) hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current biology : CB* 6(9):1181-1184.

29. Schofield MJ, Nayak S, Scott TH, Du C, & Hsieh P (2001) Interaction of Escherichia coli MutS and MutL at a DNA mismatch. *J Biol Chem* 276(30):28291-28299.

30. Lahue RS, Au KG, & Modrich P (1989) DNA mismatch correction in a defined system. *Science* 245(4914):160-164.

31. Schaaper RM (1993) Base selection, proofreading, and mismatch repair during DNA replication in Escherichia coli. *The Journal of Biological Chemistry* 268(32):23762-23765.

32. Kunkel TA (1992) DNA replication fidelity. *J. Biol. Chem.* 267:18251-18254.

33. Kunkel TA & Bebenek K (2000) DNA replication fidelity. *Annu Rev Biochem* 69:497-529.

34. Jackson D, Wang X, & Rudner DZ (2012) Spatio-temporal organization of replication in bacteria and eukaryotes (nucleoids and nuclei). *Cold Spring Harbor perspectives in biology* 4(8).

35. Simmons LA, Davies BW, Grossman AD, & Walker GC (2008) Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Molecular cell* 29(3):291-301.

36. Smith BT, Grossman AD, & Walker GC (2001) Visualization of mismatch repair in bacterial cells. *Mol. Cell* 8(6):1197-1206.

37. Hombauer H, Campbell CS, Smith CE, Desai A, & Kolodner RD (2011) Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 147(5):1040-1053.

38. Kleczkowska HE, Marra G, Lettieri T, & Jiricny J (2001) hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. *Genes Dev* 15(6):724-736.

39. Klocko AD*, et al.* (2011) Mismatch repair causes the dynamic release of an essential DNA polymerase from the replication fork. *Molecular microbiology* 82(3):648-663.

40. Hombauer H, Srivatsan A, Putnam CD, & Kolodner RD (2011) Mismatch repair, but not heteroduplex rejection, is temporally coupled to DNA replication. *Science* 334(6063):1713-1716.

41. Flores-Rozas H, Clark D, & Kolodner RD (2000) Proliferating cell nuclear antigen and Msh2p-Msh6p interact to form an active mispair recognition complex. *Nat Genet* 26(3):375-378.

42. Lee SD & Alani E (2006) Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *Journal of molecular biology* 355(2):175-184.

43. Lopez de Saro FJ, Marinus MG, Modrich P, & O'Donnell M (2006) The beta sliding clamp binds to multiple sites within MutL and MutS. *J Biol Chem* 281(20):14340-14349.

44. Shell SS, Putnam CD, & Kolodner RD (2007) The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to PCNA. *Molecular cell* 26(4):565-578.

45. Kong X-P, Onrust R, O'Donnell M, & Kuriyan J (1992) Three-dimensional structure of the b subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* 69:425-437.

46. Krishna TSR, Kong X-P, Gary S, Burgers PM, & Kuriyan J (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79:1233-1243.

47. Matsumiya S, Ishino Y, & Morikawa K (2001) Crystal structure of an archaeal DNA sliding clamp: proliferating cell nuclear antigen from Pyrococcus furiosus. *Protein Sci* 10(1):17-23.

48. Georgescu RE*, et al.* (2008) Structure of a sliding clamp on DNA. *Cell* 132(1):43-54.

49. Jeruzalmi D, O'Donnell M, & Kuriyan J (2001) Crystal structure of the processivity clamp loader gamma (gamma) complex of E. coli DNA polymerase III. *Cell* 106(4):429-441.

50. Bowman GD, O'Donnell M, & Kuriyan J (2004) Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* 429(6993):724-730.

51. Stukenberg PT, Studwell-Vaughn PS, & O'Donnell M (1991) Mechanism of the sliding clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* 266:11328-11334.

52. Huang CC, Hearst JE, & Alberts BM (1981) Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. *The Journal of Biological Chemistry* 256(8):4087-4094.

53. Monti MR, Miguel V, Borgogno MV, & Argarana CE (2012) Functional analysis of the interaction between the mismatch repair protein MutS and the replication processivity factor beta clamp in Pseudomonas aeruginosa. *DNA Repair (Amst)* 11:463-469.

54. Kadyrov FA, Dzantiev L, Constantin N, & Modrich P (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* 126(2):297-308.

55. Pillon MC*, et al.* (2010) Structure of the endonuclease domain of MutL: unlicensed to cut. *Molecular cell* 39(1):145-151.

56. Pillon MC, Miller JH, & Guarne A (2011) The endonuclease domain of MutL interacts with the beta sliding clamp. *DNA repair* 10(1):87-93.

57. Pluciennik A*, et al.* (2010) PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc Natl Acad Sci U S A* 107(37):16066-16071.

58. Larrea AA, Lujan SA, & Kunkel TA (2010) SnapShot: DNA mismatch repair. *Cell* 141(4):730 e731.

59. Lenhart JS, Schroeder JW, Walsh BW, & Simmons LA (2012) DNA Repair and Genome Maintenance in Bacillus subtilis. *Microbiology and molecular biology reviews : MMBR* 76(3):530-564.

60. Lau PJ & Kolodner RD (2003) Transfer of the MSH2.MSH6 complex from proliferating cell nuclear antigen to mispaired bases in DNA. *The Journal of Biological Chemistry* 278(1):14-17.

61. Pluciennik A, Burdett V, Lukianova O, O'Donnell M, & Modrich P (2009) Involvement of the beta clamp in methyl-directed mismatch repair in vitro. *J Biol Chem* 284(47):32782-32791.

62. Malkov VA, Biswas I, Camerini-Otero RD, & Hsieh P (1997) Photocross-linking of the NH2-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA. *J Biol Chem* 272(38):23811-23817.

63. Lamers MH*, et al.* (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* 407(6805):711-717.

64. 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(6805):703-710.

65. Bowers J, Tran PT, Liskay RM, & Alani E (2000) Analysis of yeast MSH2-MSH6 suggests that the initiation of mismatch repair can be separated into discrete steps. *Journal of molecular biology* 302(2):327-338.

66. Schofield MJ*, et al.* (2001) The Phe-X-Glu DNA binding motif of MutS. The role of hydrogen bonding in mismatch recognition. *J Biol Chem* 276(49):45505-45508.

67. Jacobs-Palmer E & Hingorani MM (2007) The effects of nucleotides on MutS-DNA binding kinetics clarify the role of MutS ATPase activity in mismatch repair. *Journal of molecular biology* 366(4):1087-1098.

68. Sanjanwala B & Ganesan AT (1991) Genetic structure and domains of DNA polymerase III of Bacillus subtilis. *Mol Gen Genet* 226(3):467-472.

69. Berkmen MB & Grossman AD (2006) Spatial and temporal organization of the Bacillus subtilis replication cycle. *Molecular microbiology* 62(1):57-71.

70. Lemon KP & Grossman AD (1998) Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282:1516-1519.

71. Migocki MD, Lewis PJ, Wake RG, & Harry EJ (2004) The midcell replication factory in Bacillus subtilis is highly mobile: implications for coordinating chromosome replication with other cell cycle events. *Molecular microbiology* 54(2):452-463.

72. Webb CD*, et al.* (1997) Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of B. subtilis. *Cell* 88(5):667-674.

73. Teleman AA, Graumann PL, Lin DC, Grossman AD, & Losick R (1998) Chromosome arrangement within a bacterium. *Current biology : CB* 8(20):1102-1109.

74. Gruber S & Errington J (2009) Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in B. subtilis. *Cell* 137(4):685-696.

75. Sullivan NL, Marquis KA, & Rudner DZ (2009) Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 137(4):697-707.

76. Dupes NM*, et al.* (2010) Mutations in the Bacillus subtilis beta clamp that separate its roles in DNA replication from mismatch repair. *Journal of bacteriology* 192(13):3452-3463.

77. Tessmer I*, et al.* (2008) Mechanism of MutS searching for DNA mismatches and signaling repair. *J Biol Chem* 283(52):36646-36654.

78. Zhai J & Hingorani MM (2010) Saccharomyces cerevisiae Msh2-Msh6 DNA binding kinetics reveal a mechanism of targeting sites for DNA mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America* 107(2):680-685.

79. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, & Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc. Natl. Acad. Sci. U.S.A.* 98(20):11627-11632.

80. Acharya S*, et al.* (1996) hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proceedings of the National Academy of Sciences of the United States of America* 93(24):13629-13634.

81. Calmann MA, Nowosielska A, & Marinus MG (2005) The MutS C terminus is essential for mismatch repair activity in vivo. *Journal of bacteriology* 187(18):6577-6579.

82. Pavlov YI, Mian IM, & Kunkel TA (2003) Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol* 13(9):744-748.

83. Hardwood CR & Cutting SM (1990) *Molecular Biological Methods for Bacillus* (John Wiley & Sons, Chichester).

84. Foster PL (2006) Methods for determining spontaneous mutation rates. *Methods Enzymol* 409:195-213.

85. Hall BM, Ma CX, Liang P, & Singh KK (2009) Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics* 25(12):1564-1565.

86. Bolz NJ, Lenhart JS, Weindorf SC, & Simmons LA (2012) Residues in the N-terminal domain of MutL required for mismatch repair in Bacillus subtilis. *Journal of Bacteriology* 194(19):5361-5367.

87. Klocko AD, Crafton KM, Walsh BW, Lenhart JS, & Simmons LA (2010) Imaging mismatch repair and cellular responses to DNA damage in Bacillus subtilis. *J Vis Exp* (36):1-4.

88. Simmons LA*, et al.* (2009) Comparison of responses to double-strand breaks between Escherichia coli and Bacillus subtilis reveals different requirements for SOS induction. *Journal of bacteriology* 191(4):1152-1161.

89. Simmons LA, Grossman AD, & Walker GC (2007) Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* 104(4):1360-1365.

90. Rokop ME, Auchtung JM, & Grossman AD (2004) Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of Bacillus subtilis. *Molecular microbiology* 52(6):1757-1767.

91. Simmons LA & Kaguni JM (2003) The DnaAcos allele of Escherichia coli: hyperactive initiation is caused by substitution of A184V and Y271H, resulting in

defective ATP binding and aberrant DNA replication control. *Molecular microbiology* 47(3):755-765.

92. Walsh BW, Lenhart JS, Schroeder JW, & Simmons LA (2012) Far Western blotting as a rapid and efficient method for detecting interactions between DNA replication and DNA repair proteins. *Methods Mol Biol* 922:161-168.

Chapter III

Trapping and visualizing intermediate steps in the mismatch repair pathway *in vivo*

Justin S. Lenhart[1§], Monica C. Pillon[2§], Alba Guarné[2] and Lyle A. Simmons[1*]

[1] Department of Molecular, Cellular, and Developmental Biology
University of Michigan, 830 North University Ave, Ann Arbor, MI 48109-1048
[2] Department of Biochemistry and Biomedical Sciences, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1

Author contributions: Figures 12A and B, 15C, 16, 17 and 18, as well as Table 3 and 4 were performed by JSL. Figures 13 and 14 were performed by MCP. Modeling found in Figures 12C, 15A and B, and 19 was performed by AG. All authors contributed to the design of the research and composition of this chapter.

[§]These authors contributed equally

Manuscript is published: Lenhart, J. S., Pillon, M. C., Guarne, A. & Simmons, L. A. Trapping and visualizing intermediate steps in the mismatch repair pathway in vivo. *Molecular microbiology* **90**, 680-698 (2013).

Abstract

During mismatch repair, MutS is responsible for mismatch detection and the recruitment of MutL to the mismatch through a mechanism that is unknown in most organisms. Here, we identified a discrete site on MutS that is occupied by MutL in *Bacillus subtilis*. The MutL binding site is composed of two adjacent phenylalanine residues located laterally in an exposed loop of MutS. Disruption of this site renders MutS defective in binding MutL *in vitro* and *in vivo,* while also eliminating mismatch repair. Analysis of MutS repair complexes *in vivo* shows that MutS mutants defective in interaction with MutL are “trapped” in a repetitive loading response. Furthermore, these mutant MutS repair complexes persist on DNA away from the DNA polymerase, suggesting that MutS remains loaded on mismatch proximal DNA awaiting arrival of MutL. We also provide evidence that MutS and MutL interact independent of mismatch binding by MutS *in vivo* and *in vitro,* suggesting that MutL can transiently probe MutS to determine if MutS is mismatch bound. Together, these data provide insights into the mechanism that MutS employs to recruit MutL, and the consequences that ensue when MutL recruitment is blocked.

Introduction

Mismatch repair (MMR) is a highly conserved pathway responsible for identifying and correcting DNA polymerase errors, which substantially improves the overall fidelity of genome replication (1-4). Defects in bacterial MutS or MutL cause a substantial increase in mutation rate (5-9), while inactivation of the eukaryotic homologues, MutSα and MutLα, causes an increase in mutation rate and microsatellite instability (10, 11). In humans, disruption of MMR can lead to the development of sporadic cancers, as well as hereditary cancers such as Lynch and Turcot syndromes (12-15). In prokaryotes, disruption of MMR can lead to an increased possibility of generating mutations that confer antibiotic resistance, and has been linked to antibiotic resistant strains of nosocomial human pathogens (16-19).

In bacteria, the pathway and the mechanisms underlying MMR are best understood in the MutH and Dam containing bacterium *Escherichia coli*. In *E. coli*, MMR is initiated upon the recognition of single base mismatches or insertion/deletion loops (IDLs) by the mismatch binding protein MutS [for review (1, 3, 20)]. While scanning for replication errors, MutS exists in an ADP bound state (21, 22). Following mismatch recognition, a prominent model is that MutS exchanges ADP for ATP, converting MutS to a sliding clamp causing MutS to diffuse away from the mismatch along the DNA in search of MutL (23, 24). After arrival, MutL then performs several tasks necessary to facilitate removal of the strand bearing the mismatch (25-27).

The initial steps of MMR have been thoroughly studied and elucidated in the Gram-positive bacterium *Bacillus subtilis,* an organism lacking the Dam and MutH dependent pathway [for review (3)]. Preceding mismatch detection, MutS is targeted to newly replicated DNA through interaction with the DNA replication processivity clamp DnaN (3, 28-30). DnaN, a critical component of the pathway, is required for 90% of MMR in *B. subtilis* (31). During Okazaki fragment maturation, DnaN accumulates behind the progressing replication forks forming a transient DnaN clamp zone that facilitates coupling between mismatch detection and concurrent DNA replication (31). Within this zone, MutS detects mismatches and initiates the downstream steps of repair, which includes MutL recruitment. The mechanism used by MutS to recruit MutL is unknown *in vivo* and *in vitro* for *B. subtilis*.

Even though MutS and MutL have been extensively characterized at the biochemical and genetic level, their binding interface and the mechanism used to recruit MutL is poorly understood in most organisms. In *E. coli*, an effort employing hydrogen/deuterium exchange mass spectrometry identified a MutL docking site on MutS composed of two adjacent glutamines (residues 211 and 212) found within the MutS connector domain (32). However, this site is not conserved in Gram-positive bacteria (Figure S1), suggesting a separate uncharacterized interface that facilitates binding in other organisms. To our knowledge, no other sites have been identified in any bacterium lacking the Dam/MutH-dependent MMR, and the effect of MutS mutants defective for MutL interaction have not been tested on repair intermediates *in vivo* for any organism. Therefore, very little is known about MutL recruitment, yet this step represents the second step in one of the most important pathways for maintaining high fidelity replication in organisms from bacteria to humans.

Here we define the MutS•MutL interface in *B. subtilis*. We show that MutS binds the N-terminal domain of MutL via two adjacent phenylalanine residues, F319 and F320. Substitution of these phenylalanines to serine eliminates crosslinking of MutS to the N-terminal domain of MutL *in vitro* while also eliminating MMR *in vivo*. Importantly, these substitutions do not seem to affect other biochemical properties of MutS, including dimerization, ATPase activity, and mismatch binding. Furthermore, using single cell fluorescence microscopy, we show that MutS mutants defective in MutL interaction form repair centers that increase in both frequency within the cell population, as well as overall fluorescence intensity. These data provide *in vivo* evidence for *in vitro* models proposing that MutS loads repetitively at a mismatch. Our work also defines a regulatory role for MutL in limiting or preventing additional MutS dimers from loading at a mismatch. We show that repetitive loading of MutS is repressed following excision of the mismatch, which requires not only MutL recruitment but also endonuclease directed nicking of the DNA. We also provide evidence against the paradigm that MutL requires MutS bound to a mismatch to initiate interaction. We show that within living cells and with purified components, we can selectively crosslink MutS to MutL in the absence of a mismatch, suggesting a mechanism where MutL can transiently probe MutS to determine if MutS is indeed mismatch bound, and if so, license repair. Together, our data provide new

insight into the MutL recruitment mechanism, the physiological consequences that result from MutS mutants unable to bind and recruit MutL, and we describe a model where MutL can transiently probe MutS before initiating the second step of MMR.

Results

The *E. coli* MutL binding interface is not conserved in *B. subtilis.*

The MutS•MutL interface has previously been characterized in the Gram-negative bacterium *E. coli* (32). The interface is found within the connector domain of MutS, and centers around a double glutamine motif (Q211 and Q212) (32). Disruption of this site causes a loss of mismatch repair *in vivo* and has been shown to eliminate interaction with MutL on a mismatched DNA substrate *in vitro* (32). Initially, we asked if the *E. coli* MutL binding motif was conserved in the Gram-positive bacterium *B. subtilis*; however, a sequence alignment revealed that the connector domain motif is not conserved, and the surrounding amino acid sequence is highly variable (**Figure 31:App**). Further examination of a *B. subtilis* MutS model shows that although the amino acid sequence in the connector domain is not conserved, the secondary and tertiary structure of the connector domain is conserved with that of *E. coli* MutS (**Figure 31 B and C: App**). Therefore, we mutated four residues, ^{205}VTII (*mutS Patch Ec*), which directly align with the *E. coli* ^{211}QQ motif, and occupy the corresponding location in a *B. subtilis* MutS model. Mutation of ^{205}VTII to ^{205}ASAA has no effect on MMR *in vivo*, conferring a mutation rate identical to the wild type control ($2.47X10^{-9}$ mutations /generation [0.95-3.82]) (**Table 3, last row**). With this result, we conclude that the MutL binding site on *B. subtilis* MutS is distinct from the site identified for *E. coli* MutS.

MutL binds several surface exposed peptides on MutS.

We previously showed that a direct interaction between MutS and MutL can be detected in *B. subtilis* without a DNA substrate using a far Western blot (33). To verify the direct MutS•MutL interaction, we performed a far Western blot to compare the binding of MutL to MutS and another known binding partner, the replication processivity clamp DnaN (**Figure 12A**) (29). We found that MutS retained MutL and DnaN on the nitrocellulose membrane during the binding reaction.

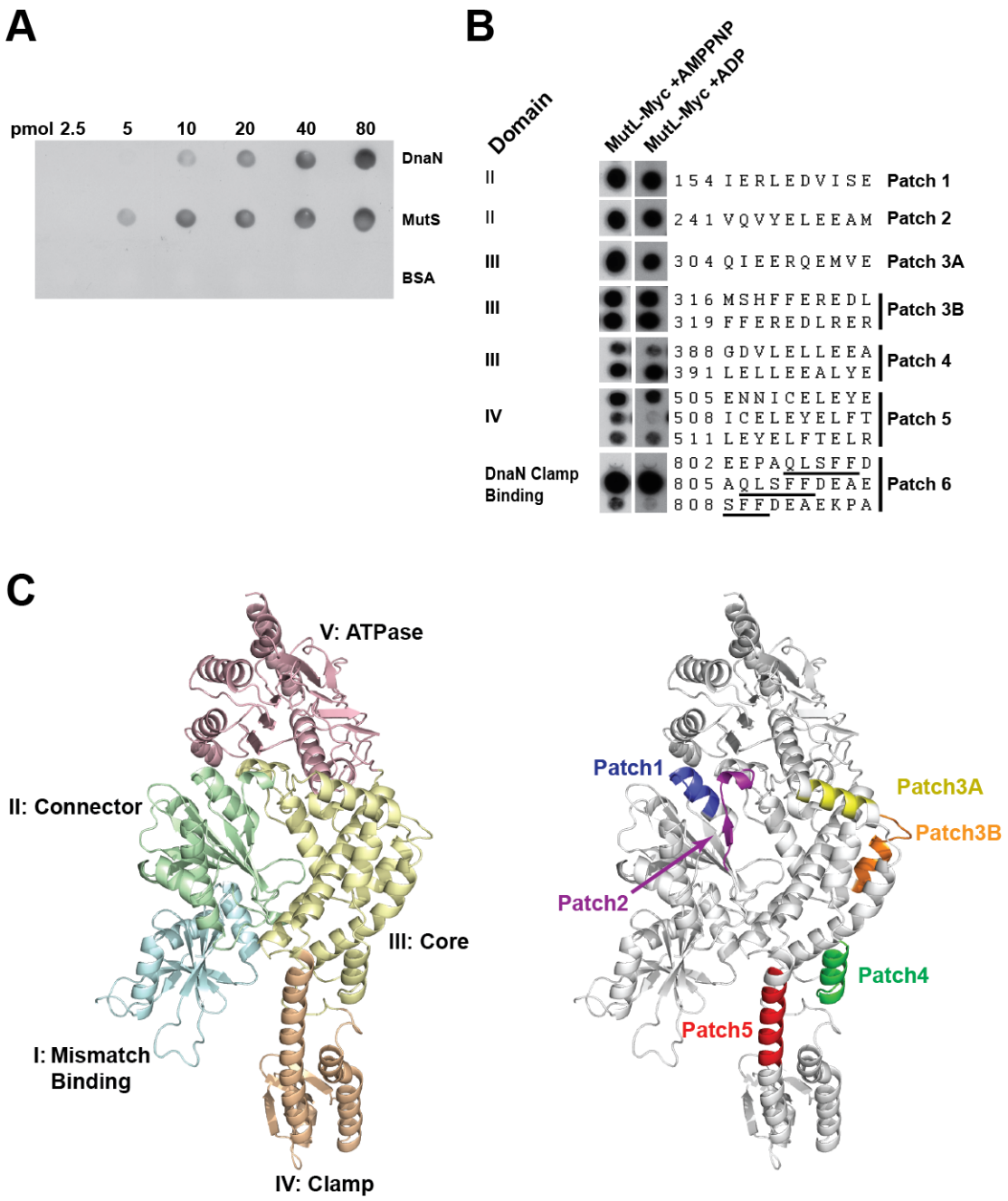


Figure 12 *Bacillus subtilis* MutL binds surface exposed peptides in MutS.
A) A far Western blot using MutL to probe for interaction with purified MutS, DnaN, and BSA. Equal amounts of the indicated protein monomer was applied via a dot blot apparatus and probed with 0.4 μM purified MutL. **B)** Screening of a MutS peptide array library with MutL-Myc. MutL-Myc was incubated with 0.5 mM of either ADP or AMPPNP during incubation with the peptide array. MutL-Myc bound peptides were detected with α-Myc antibodies. Indicated position of positive peptides on the array, as well as the amino acid sequence, is shown adjacent to the MutL-Myc bound peptides. **C)** *B. subtilis* MutS was modeled using the SWISS-MODEL server (59). Both monomers of the model are shown as a ribbon diagram, with either the five functional domains of MutS (left panel) or the surface exposed peptides identified in the peptide array (right panel) color-coded and labeled according to their representative patch definition.

To identify candidate residues in MutS that may be important for MutL binding, we employed a peptide array library, which functions analogously to the far Western, using peptides in place of purified proteins. We screened a peptide array library composed of peptides representing the entire MutS primary structure. The MutS peptide library consisted of 10mer peptides offset by 3 residues, providing 3-fold coverage of the entire sequence of MutS. We determined the association of MutL bearing a single C-terminal Myc tag with the MutS peptide array library in the presence of ADP and the non-hydrolysable ATP analog adenosine 5′-(β,γ-imido) triphosphate (AMPPNP). We used ADP and AMPPNP to determine if the nucleotide bound state altered the putative MutL binding sites on MutS, because it has been previously shown that MutL undergoes substantial conformation changes during ATP binding and hydrolysis (34).

We found that MutL•AMPPNP bound to 18 of 292 total peptides screened while MutL•ADP showed a nearly identical pattern and bound to 16 of the 18 peptides identified with MutL•AMPPNP (Figure 1B, showing only surface exposed peptides**)**. Peptides 508 and 808 did not retain MutL•ADP binding (**Figure 12B**). The data further shows that MutL-ADP bound at least one peptide in groups of overlapping peptides, suggesting that interaction within these regions occurred regardless of the nucleotide cofactor, and that overall, the nucleotide composition of MutL doesn't affect the specific MutS peptides bound. Further analysis of the amino acid composition of the MutS peptides bound by MutL-Myc revealed an enrichment of glutamic acid and phenylalanine residues, suggesting a preferred amino acid target on the MutS peptide array (**Figure 32: App**).

To determine the location of each putative MutS binding peptides, we modeled *B. subtilis* MutS based on the *E. coli* and *T. aquaticus* structures (35, 36). In doing so, we found that most peptides (13 in the AMPPNP group and 11 in the ADP group) were surface exposed and located on the outer rim of the MutS dimer (**Figure 12B and C**). Based on the location of the surface exposed peptides, we defined six unique regions (identified as patch 1-6) composed of single or multiple MutL-Myc bound peptides, which could facilitate an interaction between MutS and MutL (**Figure 12C**). Since the peptides spanning residues 802-817 (patch 6) are absent from the crystal structures of *E. coli* and *T. aquaticus* MutS (35, 36), we were unable to include them in the model.

Interestingly, patch 6, overlaps with the site known to bind DnaN, referred to as the DnaN clamp-binding motif (**Figure 12B,** ^{806}QLSFF) (29, 37). In *B. subtilis* mutation of this region does reduce MutL recruitment into foci although the mutant *mutS* still retains almost all MMR activity *in vivo* suggesting this region is not critical for binding MutL (29).

Table 3. Mutation rate analysis of *mutS* patch variants

Genotype	*mutS* variant	Number of cultures	Mutation rate (10-9 mutations/generation)± [95% CI]	Relative mutation rate (% MMR activity)
Wild-Type (PY79)	$mutS^{WT}$	24	3.30 [1.44-5.00]	1 (100%)
mutL::spec	$mutS^{WT}$	18	159.9 [152.5-167.2]*	48.5 (0%)
mutS Patch 1	*E155S, R156S, L157A, E158S*	19	4.50 [2.23-6.64]	1.36 (99.2%)
mutS Patch 2	*E245S, E247S, E248S*	24	4.28 [2.10-6.34]	1.30 (99.4%)
mutS Patch 3A	*E306S, E307S, E310S*	25	4.19 [2.18-6.11]	1.27 (99.4%)
mutS Patch 3B	*F320S, E321S, R322S, E323S*	26	78.2 [72.2-84.2]*	23.7 (52.1%)
mutS Patch 4	*E392S, E395S, E396S*	20	5.60 [2.59-8.40]	1.70 (98.5%)
mutS Patch 5	*E510S, E512S, E514S*	20	7.24 [4.45-9.94]	2.20 (97.5%)
mutS Patch 6A	*Q806A, L807A, F809A, F810A*	23	8.83 [6.03-11.58]*	2.68 (96.5%)
mutS Patch 6B	*D811S, E812S, E814S*	20	3.03 [1.37-4.57]	0.92 (101%)
mutS Patch Ec	*V205A, T206S, I207A, I208A*	18	2.47 [0.95-3.82]	0.03 (101%)

All *mutS* variants were constructed using allelic replacement (see "Experimental Procedures"), which maintains the *mutS* variant gene at its normal genetic locus and under the control of its native promoter. The downstream *mutL* gene remains intact. Brackets enclose the lower bounds and upper bounds respectively of the 95% confidence limits. Percent MMR activity was determined using the following equation: [(R.M.R.null-R.M.R.strain)/(R.M.R.null-R.M.R.wild type)]•100. RMF=relative mutation rate. Relative mutation rate was obtained by dividing the mutation rate of each strain by that obtained for the wild type control.

Substitution of surface exposed residues within the putative MutL interaction sites on MutS causes defects in MMR.

The peptide array analysis identified sites on MutS that could potentially mediate a direct interaction with MutL, thus we began by introducing three to four amino acid

substitutions in residues both conserved and surface exposed within each “patch” to determine the effect on repair (**Table 8** in appendix for summary of substitutions). Each mutant *mutS* patch allele encoding a set of missense mutations was used to replace the wild type allele at the native *mutS* locus by allelic exchange as described (31). For each mutant allele, we determined the mutation rate by measuring the rate of spontaneous rifampin resistant colony formation as an indicator for mutagenesis and MMR dysfunction [(28, 30, 31, 38) and “Experimental Procedures”]. Patch mutants 1, 2, 3A, 4, 5 and 6B conferred a statistically equivalent mutation rate to wild type *mutS*, showing no effect on the MMR pathway *in vivo* (**Table 3**). Patch mutant 6A, which contains the DnaN clamp-binding motif, showed a slight but significant increase in mutation rate at $8.83X10^{-9}$ mutations/generation (**Table 3**) as we previously reported (29). Interestingly, we found that the four missense mutations introduced into patch 3B caused a significant increase in mutation rate ($78.2X10^{-9}$ mutations/generation), resulting in this mutant retaining only 50% of MMR activity *in vivo*. With these data, we conclude that patch mutant 3B, which includes the *F320S, E321S, R322S*, and *E323S* missense mutations, causes a significant defect in the MMR pathway in *B. subtilis* (**Table 3**). Hereafter, we refer to the patch mutant 3B as MutS3B.

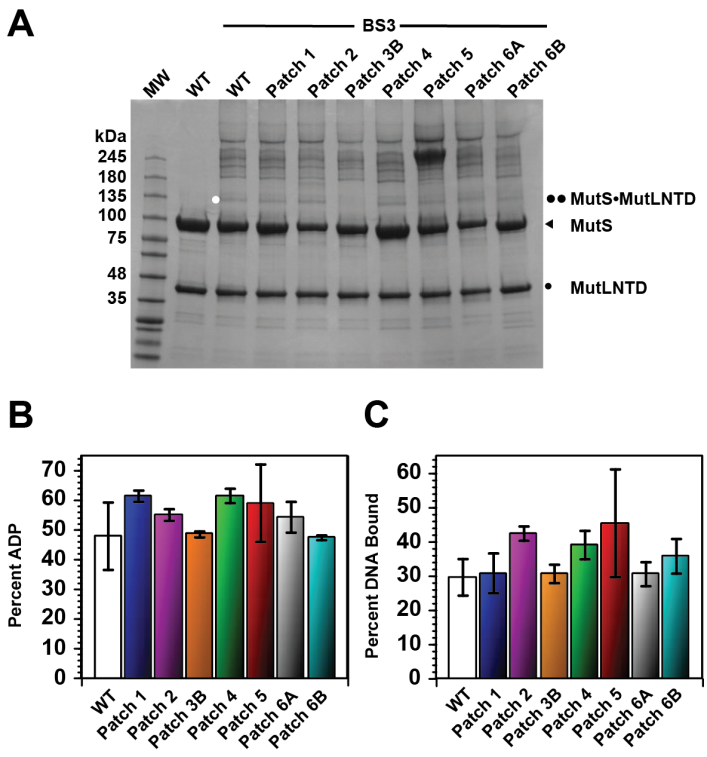


Figure 13. Purified MutS3B fails to crosslink with the N-terminal domain of MutL.
A) Crosslinking of MutS or MutS variants to the N-terminal domain of MutL with a 90 bp DNA substrate containing a centrally located G/T mismatch (Mis90). Mixtures of each protein, 10 mM ATP, and the G/T DNA substrate were incubated with the crosslinker BS3. Protein complexes were then resolved on a 4-15% gradient SDS polyacrylamide gel. The bands corresponding to the MutS and MutL-NTD monomers, as well as the MutS•MutL-NTD complex are labeled. All MutS variants show similar ATPase activity **B)** and DNA binding to the G/T mismatched DNA substrate **C)** to wild-type MutS. Bar diagrams present the average of three independent measurements and the error bars correspond to the standard errors of the mean (SEM=s/√n, where s is the standard deviation and n the sample size).

MutS3B is defective for interaction with MutL.

The MutS•MutL interaction has been previously monitored using chemical crosslinking (24). The work by Winkler and co-workers demonstrated that the MutS•MutL interaction only requires the N-terminal domain of MutL (MutL-NTD) and it is enhanced in the presence of ATP and a heteroduplex. Therefore, we purified N-terminal His_6 tagged variants of *B. subtilis* MutS and MutL-NTD and screened for interaction defects using this approach. Incubation of MutS and the chemical crosslinker bis(sulfosuccinimidyl)suberate (BS3) in the presence of ATP and a 90 base-pair G/T mismatch DNA substrate (Mis90) resulted in the formation of several high molecular weight species. Conversely, incubation of MutL-NTD with BS3 predominantly yielded monomers, as expected due to the absence of the dimerization domain of the proteins. Incubation of MutS with MutL-NTD and BS3 in the presence of ATP and Mis90, yielded a new species that was not present when either protein was incubated with BS3 and corresponded to the molecular weight of the MutS•MutL-NTD complex (**Figure 13A**). We excised this band, and using LC MS/MS, verified the presence of both MutS and

MutL-NTD as the sole components of this band (data not shown). Interestingly, we do observe some interaction between MutS and MutL-NTD in the presence of a 90 bp DNA homoduplex in place of Mis90, showing that the MutS•MutL-NTD interaction is not strictly dependent on the presence of a mismatched substrate (**Figure 33:App**).

We subsequently tested whether any of the MutS patch variants abrogated the interaction with MutL-NTD. We found that all MutS variants formed a MutS•MutL-NTD complex except for the MutS3B variant (F320S, E321S, R322S and E323S) (**Figure 13A**). We note that the MutS5 variant (including the E510S, E512S and R514S mutations) showed a very prominent band of a molecular weight consistent with formation of a MutS tetramer. In fact, this prominent species was present in the crosslinking reaction of all MutS variants when MutL-NTD was not present (data not shown), but disappeared upon incubation with MutL-NTD. Since our goal was to probe for the formation of a MutS•MutL complex, and MutS5 retained the interaction with MutL-NTD, we did not characterize this variant further. The crosslinking defect of MutS3B agrees well with the mutation rate analysis showing that *patch 3B* lost 50% of MMR activity *in vivo* (**Table 3**). All "patch" variants of MutS behave similar and have similar mismatch binding and ATPase activities compared to wild-type MutS (**Figure 13B and C**), implying that the reduced MMR activity of the MutS3B variant is unlikely due to improper folding or attenuation of other critical biochemical activities. Furthermore, all of the MutS variants eluted from a gel filtration column similarly to the wild type protein (**Data Not Shown**) and formed dimers in solution as measured by dynamic light scattering (**Figure 34:App**). Collectively, we show that the MMR defect associated with the MutS3B variant is due to the impaired interaction with MutL rather than loss of some other biochemical activity of MutS. Furthermore, we conclude that residues changed in the MutS3B (F320, E321, R322, and E323) variant are important for direct interaction between *B. subtilis* MutS and MutL-NTD.

Residues F319 and F320 define the MutL binding site on MutS.

Since *mutS3B* contains four successive missense mutations, we replaced the native *mutS* gene in *B. subtilis* with alleles encoding each of the single missense mutations that comprise *mutS3B* using allelic exchange in order to further define the

functional residues important for MutL interaction. We also included amino acids S317, H318, and F319 in this analysis due to their adjacent position in MutS relative to the peptide identified in the array and because each residue is predicted to be surface exposed. We found that mutation of F319S and F320S separately reduced MMR activity below 50% *in vivo* (**Table 4**). In addition, we found that *mutSE323S* had the most striking effect of all the single missense mutations on MMR as this allele supported only 7% of MMR activity *in vivo*. It should be noted that the effect observed in the *mutSE323S* mutant far exceeds that of the *mutS3B* mutant, which reduced MMR to 50% of wild type level. We suggest that the E321S and R322S substitutions may partially suppress the defect caused by E323S on its own. All other substitutions examined confer a mutation rate indistinguishable from wild type (**Table 4**). We did not pursue E323S for a role in MutL binding because this single mutant blocks MutS localization on its own and may have a folding defect. We describe the effects of this mutant later within this manuscript.

Table 4. Mutation rate analysis of missense mutations in and near *mutS3B.*

Genotype	Mutation rate (10-9 mutations/generation)± [95% CI]	Fold increase in mutation rate	% MMR activity
Wild-Type (PY79)	3.30 [1.44-5.00]	1	100
mutL::spec	159.9 [152.5-167.2]	48.5	0
mutSS317A	2.62 [0.94-4.09]*	0.63	100.4
mutSH318S	2.36 [0.79-3.69]*	0.55	100.6
mutSF319S	105.1 [97.9-112.4]	31.9	35.0
mutSF320S	94.8 [87.3-102.4]	28.8	41.5
mutSE321S	4.02 [1.83-6.06]*	1.22	99.5
mutSR322S	2.46 [0.92-3.78]*	0.75	100.5
mutSE323S	148.7 [140.1-157.2]#	45.1	7.2
mutSF319SF320S	156.0 [148.9-163.1]#	47.3	2.5
mutSF319SF320S, amyE::Pspac mutL	134.0 [127.3-140.5]#	40.6	16.6

All *mutS* variants were constructed using allelic replacement (see "Experimental Procedures"), which maintains the *mutS* variant gene at its normal genetic locus and under the control of its native promoter. The downstream *mutL* gene remains intact. Brackets enclose the lower bounds and upper bounds respectively of the 95% confidence limits. Percent MMR activity was determined using the following equation: [(R.M.R.null-R.M.R.strain)/(R.M.R.null-R.M.R.wild type)]•100 RMF=relative mutation rate. Relative mutation rate was obtained by dividing the mutation rate of the strain by that of wild type. The symbols * and # indicates that the mutation rate is statistically equivalent to that of the wild type and MMR deficient strains respectively. For expression of *mutL*, 1 mM IPTG was added to the media during growth.

Because *mutSF319S* and *mutSF320S* showed significant and substantial defects in MMR, we combined these missense mutations to measure the effect on MMR *in vivo*. The resulting *mutSF319SF320S* allele showed a mutation rate ($156X10^{-9}$ mutations/ generation) indistinguishable from a strain lacking *mutL* (*mutL::spc*) function (**Table 4**). Immunoblot analysis verified that MutSF319SF320S, as well as MutS variants containing each individual mutation, accumulate to the same steady state levels as wild-type MutS *in vivo* (**Figure 14A**). These results show that *mutSF319SF320S* phenocopies loss of *mutL* function in *B. subtilis* supporting the hypothesis that *mutSF319SF320S* is defective in MutL interaction. In addition we asked if overexpression of *mutL* could suppress the increased mutation rate caused by MutSF319SF320S. We expressed *mutL* using an IPTG inducible promoter from and ectopic locus and recovered only ~16 of MMR (**Table 4**, last row). This experiment further supports our conclusion that the MutSF319SF320S variant is substantially impaired for MutL interaction *in vivo*.

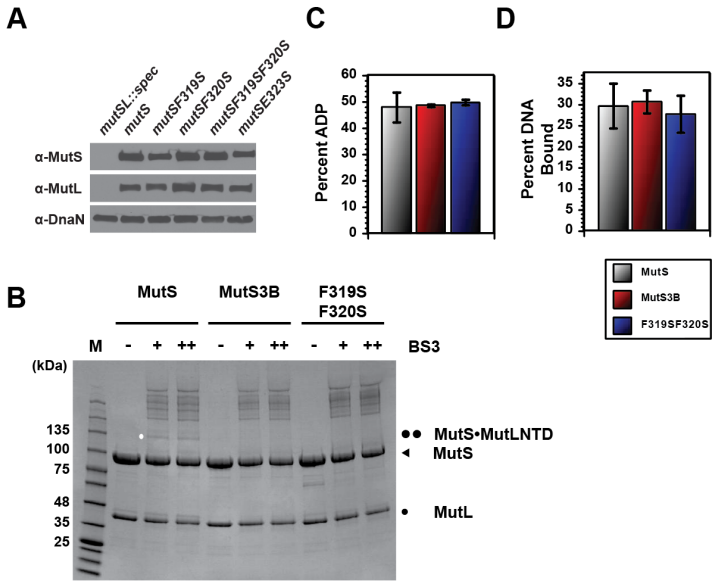


Figure 14. A distinct di-phenylalanine binding site within and around MutS3B defines the MutL binding interface. **A)** Immunoblot analysis indicated proteins from the soluble fraction of cell lysates. 50 μg of soluble fraction was probed for MutS, MutL and DnaN. **B)** Complex formation of MutS, MutS3B, and MutSF319SF320S to the N-terminal domain of MutL was assayed on a 90 bp DNA substrate containing a centrally located G/T mismatch using crosslinking analysis. Reactions contained 10 μM MutS variants, 20 μM MutL-NTD, protein, 10 mM ATP, and 10 μM of the G/T DNA substrate were incubated with the hydrophilic crosslinker BS3 (**+**=0.8 mM and **++**=1.6 mM, respectively). The products were then separated on a 4-15% gradient SDS-PAGE. The bands corresponding to the MutS and MutL-NTD monomers, as well as the MutS•MutL-NTD complex are labeled. The biochemical activity of purified MutS, MutS3B, and MutSF319SF320S were tested for **C)** ATPase activity and **D)** DNA binding to the G/T DNA substrate. Bar diagrams present the average of three independent measurements and the error bars correspond to the SEM.

In order to determine if residues F319 and F320 of MutS define a MutL binding site, we purified MutSF319SF320S and tested its ability to interact with MutL-NTD using chemical crosslinking (**Figure 14B)**. Like MutS3B, MutSF319SF320S fails to crosslink

with MutL-NTD, indicating that the mutation of the phenylalanine pair is sufficient to eliminate interaction between MutS and MutL *in vitro*. We also verified that these substitutions were wild type for other biochemical activities of MutS. MutSF319SF320S maintained wild type levels of ATPase activity, binding to mismatched DNA substrate, and dimer formation (**Figure 14C and D**, and **Data Not Shown**), suggesting that loss of *in vivo* MMR in the *mutSF319SF320S* background is attributed to loss of binding to MutL.

MutSF319SF320S defines a highly conserved MutL binding site on MutS in Gram-positive bacteria.

We asked if the MutL binding site on MutS is conserved in other organisms. The MutS residues important for MutL binding, F319 and F320, model to the outer rim of MutS and reside in the loop connecting helices α4 and α5 of the core domain (**Figure 15A**). Importantly, both residues appear solvent exposed, and available for interaction with MutL based on our structural model (**Figure 15A**). In human MSH2 and the Gram-negative bacteria *E. coli* and *T. aquaticus* MutS, the site appears to be structurally conserved, despite the limited sequence conservation (**Figure 15B**). Based on previous results (32), this interface is not the sole binding interface for *E. coli* MutL, but may however function as a secondary site located on the opposite side of the MutS face. Importantly, helix α4 is part of the allosteric transmitter proposed to connect the ATP- and DNA-binding sites of MutS (36), and hence, F319 and F320 pose an attractive mechanism to relay the nucleotide- and mismatch-bound state of MutS to MutL. Interestingly, we do find that the di-phenylalanine site is conserved in several eukaryotic proteins known to bind MutL homolog Mlh1 (39) including MutSβ (**Figure 35:App**), as well as in anchoring interaction between mammalian Rev1 and Polκ (40-42) (**see discussion**).

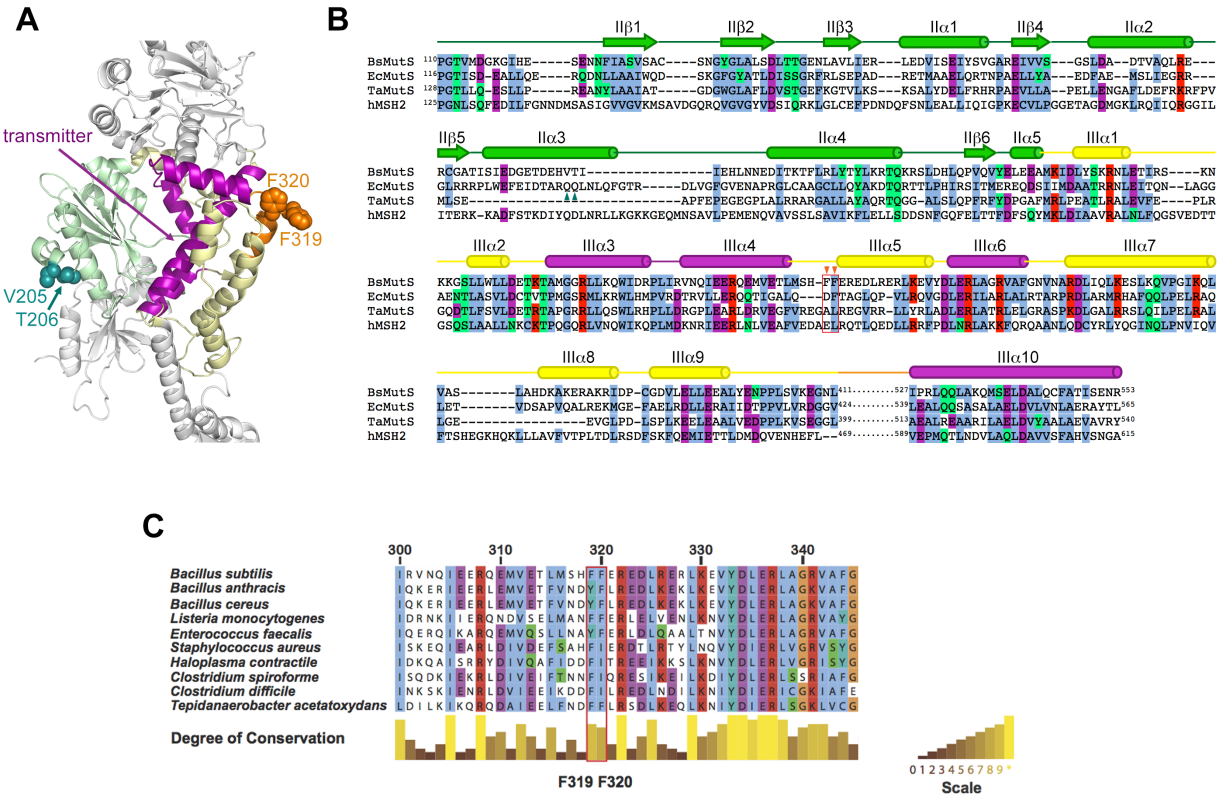
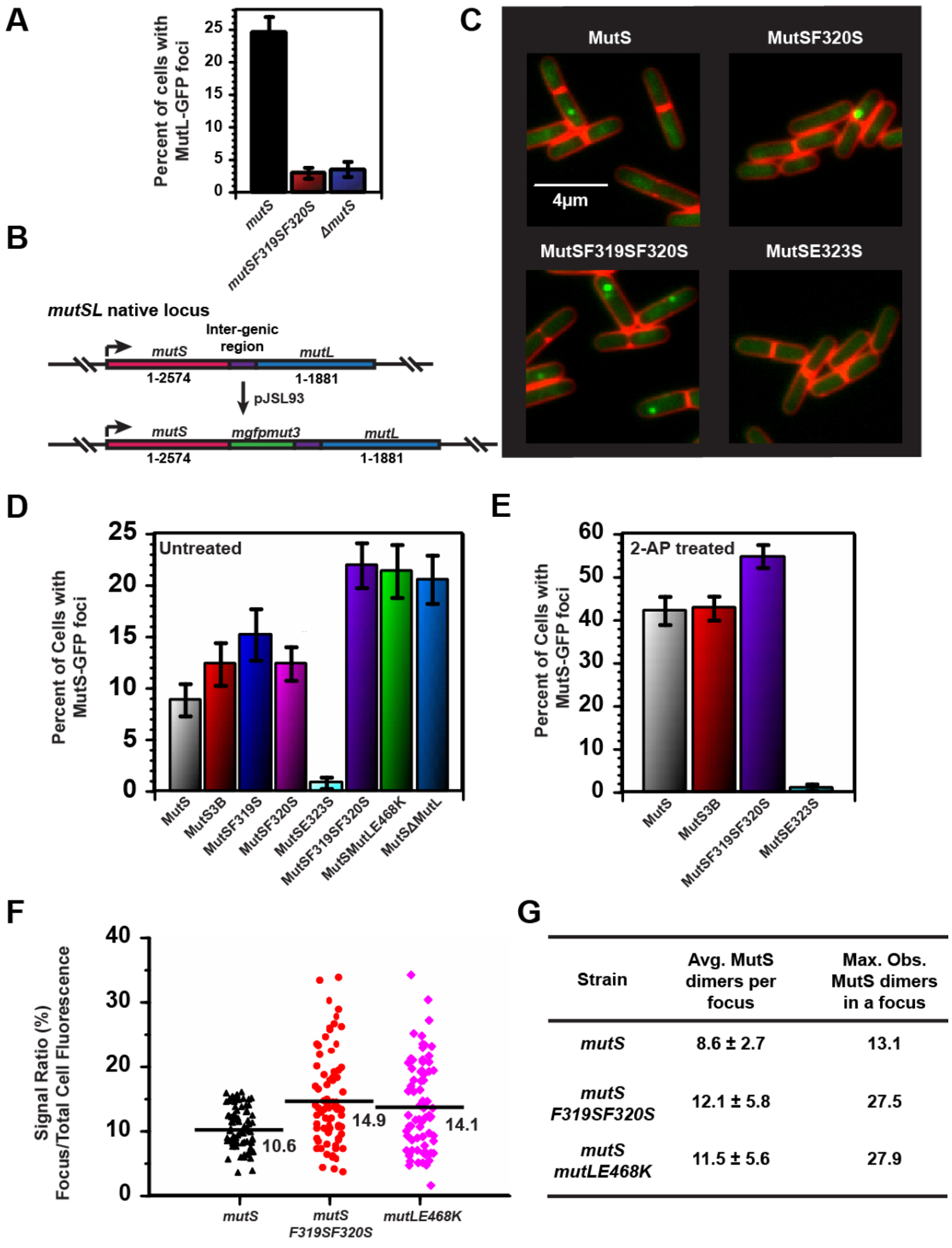


Figure 15. The di-phenylalanine site is conserved in MutS homologs.
A) Ribbon diagram of the connector (light green) and core (light yellow) domains of MutS. The side chains of the di-phenylalanine motif are shown in orange, those of the QQ motif are shown in teal (32) and the structural elements of the transmitter proposed by Obmolova and co-workers belonging to the core domain are colored in purple (36). **B)** Structure based sequence alignment of *B. subtilis* MutS and other MutS homologs for which the three-dimensional structure are known. Conserved hydrophobic (blue), polar (green), positive- (purple) and negative-charged (red) residues are highlighted. The secondary structure elements are color-coded for domains II, domain III and transmitter as in **A**. The location of the QQ and FF motifs is indicated with teal and orange carets, respectively. **C)** Sequence alignment of MutS from Gram-positive bacteria shows conservation of F319 and F320.

When we align MutS sequences from Gram-positive bacteria, many of which cause serious health concerns including *Staphylococcus aureus* and *Listeria monocytogenes*, we find that these residues are highly conserved in *mutS* homologs (**Figure 15C**). In some Gram-positive bacteria, a few accepted substitutions are tolerated at these positions, such as the aromatic residue tyrosine or the hydrophobic residue isoleucine (**Figure 15C**). Based on our results, we suggest that mutation of these conserved residues could eliminate MMR function in related pathogenic bacteria, increasing mutagenesis and altering antibiotic susceptibility and persistence within the host environment.

mutSF319SF320S* is defective for recruitment of MutL *in vivo.

It has been previously shown that MutL-GFP forms foci in response to spontaneous or 2-aminopurine (2-AP) formed mismatches detected by MutS, providing an *in vivo* assay to monitor MutL recruitment in response to mismatch detection by MutS (31, 43). A caveat with this assay is that the *mutL-gfp* allele is nearly defective for MMR as measured by mutation rate (43), however focus formation of MutL-GFP is dependent on *mutS,* providing a single cell assay for MutL-GFP recruitment in live cells (31, 43). We asked if MutSF319SF320S was able to recruit MutL-GFP into foci in cells grown with 2-AP. In a background with the native *mutS* gene, we observed MutL-GFP repair centers in ~25% of cells (**Figure 16A**). We found that cells with *mutSF319SF320S* or the Δ*mutS* allele did not support MutL-GFP focus formation, as MutL-GFP repair centers only formed in ~3% of the cell population in both genetic backgrounds. Furthermore, we also found that MutSF319SF320S is defective for recruitment of MutL-GFP in response to mismatch detection *in vivo*, supporting the *in vitro* experiments showing that *mutSF319SF320S* is defective for interaction with MutL.



Strain	Avg. MutS dimers per focus	Max. Obs. MutS dimers in a focus
mutS	8.6 ± 2.7	13.1
mutS F319SF320S	12.1 ± 5.8	27.5
mutS mutLE468K	11.5 ± 5.6	27.9

Figure 16. MutS mutants defective for MutL interaction form persistent complexes *in vivo*. Fluorescent single cell microscopy of MutL-GFP repair centers responding to mismatch formation in a *mutS, mutSF319SF320S* or *ΔmutS* background (n=1320, 1559, and 988 cells scored). Cells were treated with 600 μg of 2-AP and incubated for 1 hour prior to imaging. 95% confidence intervals are shown. **B)** Shown is a schematic for cloning an unmarked in frame fusion of *mutS* to *mutS-22-mgfpmut3* (*mutS-gfp*), while maintaining expression of the *mutL* gene downstream ("Experimental Procedures"). **C)** Representative micrographs of the indicated MutS-GFP fusion proteins. The cell membrane was imaged using the vital membrane stain TMA-DPH, which was pseudo-colored red. **D)** Shown is a bar graph of the percent of cells with each MutS-GFP fusion untreated during exponential growth (n=1276, 957, 796, 1568, 1008, 1382, and 1148 respectively refers to the number of cells scored for each strain). Error bars represent 95% confidence intervals (CI). All groups are statistically significant with respect to MutS, including MutS Patch 3B (P= 0.0038) and MutSF320S (P= 0.0013). The strain with *ΔmutL* has a *mutS-mgfpmut2*

fusion. **E)** The percent of cells with the indicated MutS-GFP fusion following challenge with 2-AP (n=879, 1212, 711, and 725 are the number of cells scored respectively) the error bars represent 95% CI. **F)** Focus intensity was determined by normalizing total signal of the repair centers to the total cell fluorescence. A total of 75 MutS-GFP foci were analyzed for each group. All foci examined were in cells statistically equivalent in regards to area, length, and average intensity of cellular fluorescence. **G)** Using the number of MutS molecules per cell (Figure S6) and the average MutS repair center fluorescent intensity F), we were able to determine the average number of MutS dimers per repair center ± standard deviation, as well as the highest observed number of MutS dimers within repair centers scored.

MutSF319SF320S forms large repair complexes *in vivo*, supporting a model for persistent loading.

With a MutS variant defective in recruitment of MutL, we can now uncouple mismatch binding from functional repair and "trap" repair intermediates that would normally be resolved during repair. To observe mismatch repair intermediates, we fused *mutS* to a monomeric *gfpmut3* variant (*gfpmut3* referred to herein as *gfp*) since *gfpmut3* represents the most monomeric derivative of GFP, providing the least invasive method for observing protein localization in living bacterial cells (44). We constructed a native locus *mutS-gfp* strain by allelic exchange in order to maintain expression of the downstream gene *mutL* under its native promoter (**Figure 16B**). The *mutS-gfp* background maintained ~85% of MMR activity (mutation rate $2.56X10^{-8}$ [2.0-3.1]), providing a functional fusion to observe active repair in real time. Upon mismatch detection, MutS-GFP forms complexes in response to mismatches in order to orchestrate repair. The *mutS-gfp* strain forms repair centers in ~9% of cells within the population during exponential growth, and repair center formation is stimulated to ~42% of cells by addition of 2-AP to the growth media (**Figure 16C, D, and E**) (28-31). Thus, using *B. subtilis*, we can bridge biochemical and genetic data to understand how disruption of MutL recruitment by MutS alters repair center dynamics *in vivo*, providing important mechanistic insight into intermediate steps.

We subsequently fused *gfp* to *mutS3B*, *mutSF319S*, *mutSF320S*, *mutSF319SF320S*, and *mutSE323S* and found that all strains except for *mutSE323S-gfp* formed repair complexes *in vivo* (**Figure 16C**). Interestingly, MutSE323S-GFP, which was defective for repair *in vivo*, was also completely defective for focus formation suggesting that although this protein accumulates *in vivo* (**Figure 14A**), the E323S mutation appears to cause some defect other than blocking MutL interaction, since it failed to form a repair complex. The MutSE323S variant was not amenable to

recombinant expression and purification and therefore we did not further pursue characterization of this variant (**Data Not Shown**).

MutSF319S-GFP, MutSF320S-GFP, and MutSF319SF320S-GFP all formed foci in a higher percentage of untreated cells than the MutS-GFP control. We hypothesized that the increase in focus formation is a consequence of an increase in the duration of repair center formation due to unproductive repair caused by a failure to properly signal for MutL. Another possibility is that there is an increase in mismatch detection, however, we ruled out this possibility by showing that a deletion of *mutL* downstream of *mutS-gfp* causes the same effect by increasing MutS-GFP repair centers *in vivo* (**Figure 16D**). Furthermore, since the error rate of the replication process in the absence of functional MMR is one mispair every two rounds of replication the likelihood of closely spaced errors is extremely low [(31, 38) and **Table 3**]. Time-lapse imaging of repair center formation and resolution would be preferred to support our hypothesis, but is not feasible due to long exposure times of the MutS-GFP fusions and rapid photobleaching dynamics (**Data Not Shown**). MutSF319S-GFP, MutSF320S-GFP, and MutS3B-GFP formed repair complexes in a nearly indistinguishable percentage of cells (12-13% of the population) (**Figure 16D**). The double mutant, MutSF319SF320S-GFP, shows an increase in the percentage of cells with MutS-GFP foci above our measurements for each of the single variants (**Figure 16D**). Furthermore, 2-AP treatment elicited an increase in the percentage of cells with MutSF319SF320S-GFP, showing that this variant still binds mismatches and initiates repair, further supporting our *in vitro* results that mismatch binding is unaffected (**Figure 16E** and **Figures 14D**). Ultimately, loss of MutL recruitment by MutS causes a corresponding increase in the percentage of cells with MutS repair complexes.

We also asked if MutS repair center formation is not only affected by MutL recruitment, but also by the next step of repair--incision. To do so, we asked if MutS-GFP repair centers accumulate in cells where MutL endonuclease nicking is prevented using the *mutLE468K* allele (45). Nicking by MutL is a required step for repair and we have previously shown that the E468K substitution eliminates MutL endonuclease activity *in vitro* and MMR activity *in vivo* (45). Indeed, the percentage of cells with MutS-GFP foci increased in the *mutLE468K* background to levels observed in both the

mutSF319SF320S and the *ΔmutL* backgrounds, indicating that if MutL-directed nicking is prevented, MutS-GFP foci persist when the next step of repair is blocked (**Figure 16D**). With these data we argue that MutL recruitment is not sufficient to halt MutS loading *in vivo* per se, but that timely repair of the mismatch is required to prevent further loading.

We also found that a proportion of repair centers exhibited high fluorescence intensity in backgrounds defective for MutL recruitment and MutL endonuclease activity (**Figure 16C, F and G**). We quantified percent focus intensity relative to whole cell fluorescence intensity. In doing so, we found that many repair centers associated with MutSF319SF320S had elevated focus intensities relative to a MutL recruitment proficient MutS-GFP strain (**Figure 16F**). These data suggest that more MutS protomers are present in a focus for MutS mutants defective in MutL interaction or in strains where MutL function has been eliminated by blocking incision (*mutLE468K*) (**Figure 16F**). We also analyzed MutSF319SF320S foci in cells where *mutL* expression was induced and observed no difference in percent of cells with foci or focus intensity (**Data Not Shown**).

We quantified the number of MutS dimers found within *B. subtilis* under the exact conditions used during live cell imaging, and found that in *B. subtilis* steady state levels of MutS are ~80 dimers per cell (100 nM) (**Figure 36:App**). Using these data, we determined that the mean number of MutS dimers in a repair center was ~8.5 (this corresponds to 17 GFP moieties) (**Figure 16G**). Both the MutL recruitment and endonuclease-deficient backgrounds contained a higher mean number of MutS-GFP dimers per repair center (12.1 and 11.5 respectively) (**Figure 16G**). The increase in repetitively loaded MutS-GFP dimers is more pronounced in the broad distribution of individual intensity measurements of the repair centers, with as many as >3 fold (~30 MutS-GFP dimers; ~35% of cellular MutS) more molecules in the highest intensity MutS complexes observed in repair deficient strains. These observations support a model where MutS can load iteratively at a mismatch, increasing the local concentration of MutS. We propose that iterative MutS loading aids in efficient MutL recruitment to the mismatch, providing *in vivo* support for *in vitro* observations (23). We also find it interesting that we quantify 8.6±2.7 MutS dimers per focus and in *S. cerevisiae* the

number of Msh6 dimers per focus was determined to be 10.8±4.4 (46). Therefore, the stoichiometry of MutS within a focus is remarkable similar between these two organisms.

MutSF319SF320S repair centers localize away from the replisome.

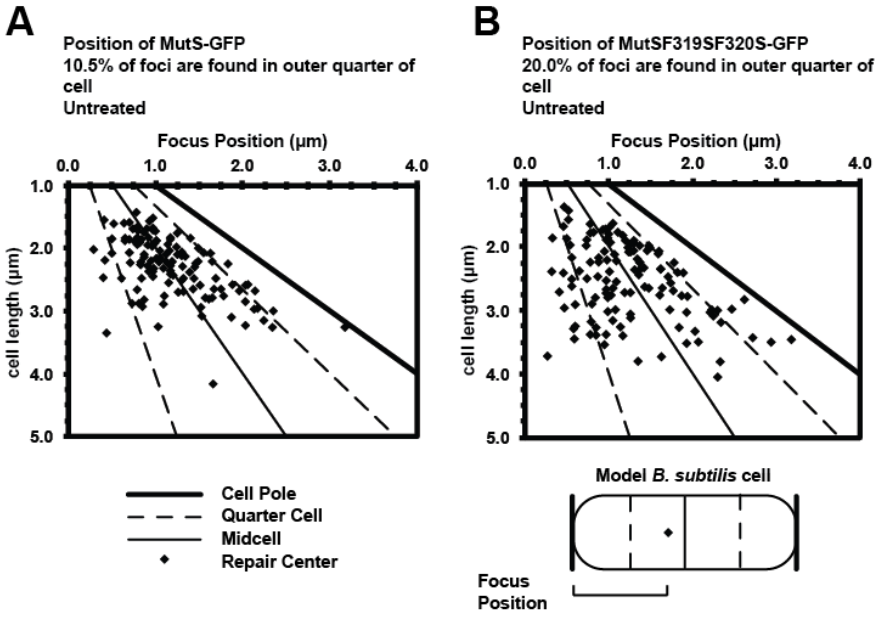
During DNA replication, chromosomal DNA is replicated within an organized replisome (47-49). Here, the replisome is defined as replication associated proteins that localize as discrete foci *in vivo*. Within *B. subtilis*, replisomes maintain a well-characterized subcellular position (47-49). Once replication is initiated from the single origin of replication (*ori*), two sets of replication forks are often contained within a single replisome predominantly found at midcell (47-49). Once replicated, the daughter chromosomes begin to translocate to the cell poles, taking mismatched DNA away from the centrally located replisome. We have previously shown that MutS foci colocalize to the replisome preceding mismatch detection and are released following mismatch binding (31). Therefore, we asked if localization of MutS repair complexes is altered when MutS is broken for MutL recruitment.

In order to test if the MutSF319SF320S repair centers persist at the site of mismatch identification, we monitored their position during DNA replication in minimal medium under slow growth conditions (~ 123 min. doubling time). Slow growth maintains approximately half of the cell population with a single replisome focus (~52% of cells). We first determined the distance of the MutS and MutSF319SF320S repair centers relative to the cell poles (**Figure 17A**). MutS-GFP repair centers maintain a mostly midcell position with 48.4% found within the middle 10% of the cell. Only 10.5% of these repair centers occupy a distal position within the outer quarters of the cell. Relative to the distribution of MutS-GFP, MutSF319SF320S-GFP position was more dispersed, as only 27.2% of repair centers were found within the middle 10% of the cell (**Figure 17B**). About 2 fold more MutSF319SF320S repair centers (20%) were found in the distal quarter of the cell. These data support the hypothesis that upon identifying a mismatch at the replisome, the assembled MutS repair center is maintained at the site of the mismatch for extended periods of time, causing migration away from the replisome as DNA synthesis continues, an effect more pronounced when MutL recruitment is blocked.

To further test this hypothesis, we examined colocalization between the MutS-GFP repair centers and replisomes during the same slow growth conditions described above. Colocalization between MutS-GFP and DnaX-mCherry (a component of the processivity clamp loader complex) was performed and scored as described (31). During exponential growth, MutS-GFP forms repair complexes that colocalize with the replisome in about 51% of cells. When we stimulate mismatch formation by adding 2-AP to the media, we found a decrease in colocalization to ~35% (p=0.00052), consistent with previous results (**Figure 17C**) (31). During exponential growth, MutSF319SF320S repair complexes colocalize with the replisome in 38% of the population; a significant decrease compared to MutS-GFP during exponential growth (p=0.0090). Upon treatment with 2-AP, only 29% of repair complexes colocalize with the replisome. With these results we conclude that when MutS-GFP is unable to recruit MutL to the site of a mismatch, repetitive loading of MutS-GFP at the mismatch will continue, resulting in a brighter and more persistent MutS-mismatch complex, which migrates away from the replisome as replication continues.

MutL crosslinks with MutS independent of mismatch detection *in vivo* and *in vitro*.

An outstanding problem in MMR is how MutL senses when MutS is mismatch bound to initiate downstream steps of repair. Previously, we showed that the *mutSF30A* allele, supports formation of more MutL-GFP repair centers than are observed in the *ΔmutS* background (31). MutSF30A is a variant that is unable to distinguish mismatched DNA from complementary DNA (31). This observation is interesting because it suggests that MutS can interact with MutL, in the absence of mismatch binding *in vivo*, even though the interaction is reduced (31). Here, we directly test the hypothesis that MutL can transiently probe MutS for the appropriate conformational change to initiate MMR. To test this hypothesis, we used immunoprecipitation (IP) targeting MutS to co-IP any proteins associated with MutS *in vivo*. Since the MutS•MutL interaction is transient in nature, we employed the use of the thiol-cleavable, membrane permeable crosslinker Dithiobis[succinimidyl propionate] (DSP) to crosslink MutS•MutL complexes formed in growing cells (**Figure 18**).



C

Colocalization of MutS-GFP foci with the Replisome (DnaX-mCherry)

Fusion	Addition of 2-AP	% colocalization (+/- 95% CI)	No. of cells analyzed	
MutS-GFP	-	51.3% (7.8)	158	*
	+	35.0% (5.8)	260	#
MutSF319SF320S-GFP	-	38.2% (7.4)	165	**
	+	29.4 % (6.4)	197	

Figure 17. MutSF319SF320S foci persist on DNA away from the replisome in the absence of MutL recruitment.
The position of repair centers for **A)** MutS-GFP and **B)** MutSF319SF320S-GFP within each cell was plotted by the coordinates (cell length, distance to pole). Solid black line indicates midcell, whereas dashed lines indicates the quarter cell positions. The thick black line indicates the cell end. n=125 **C)** Table indicating colocalization values for MutS-GFP with DnaX-mCherry. The number of cells scored is indicated (n). p-values: *=0.00052, **=0.040, #=0.0090, and difference between the 2-AP treatment groups =0.105.

The IP was accomplished under normal growth conditions in the absence of 2-AP to test for association in the absence of active MMR. Using this procedure, we were able to IP ~10.0% of the intracellular MutS. Importantly, we were able to capture the MutS•MutL interaction in the wild type strain, yet failed to IP MutL in the *ΔmutS* strain, validating the requirement of MutS for successful co-IP of MutL. We were able to detect a MutL band (0.02% of input) in the IP lane from the wild type strain. The low amount of MutL recovered in the wild type strain is likely because we are precipitating only 10.0% of intracellular MutS, as well as we expect only 9% of cells to have ongoing MMR as determined by the assembly of active MutS-GFP repair centers. In agreement with our

in vitro data and the MutL-GFP microscopy (**Table 4** and **Figure 16A**), we recovered low amounts of MutL in the IP lane from the *mutSF319SF320S* lysate (<0.001% of the input), confirming that MutSF319SF320S is compromised for interaction with MutL *in vivo* (**Figure 18**). In Figure 7, we also present error measurement from three independent IP experiments. In the other experiments performed we did not recover any detectable amount of MutL in the MutSF319SF320S lysate further supporting our conclusion that this mutant does not interact with MutL (**data not shown**).

We then tested whether mismatch detection was necessary to facilitate MutS•MutL interaction *in vivo*, speculating that MutL may frequently probe MutS for the appropriate protein conformation, signaled by mismatch detection. To test this, we IPed MutSF30A: a MutS variant capable of DNA binding, yet incapable of discriminating mismatched DNA from complementary DNA (31). When MutSF30A was immunoprecipitated, we found that we successfully captured MutL (~0.007% of input). This result shows that the MutS•MutL interaction may dynamically occur independent from mismatch identification *in vivo*, suggesting that MutL is capable of transiently "checking" to determine if MutS is mismatch bound before licensing downstream repair events. Similar observations have been seen in *S. cerevisiae in vitro* showing that MutSα interaction with MutLα is not entirely mispair dependent (50).

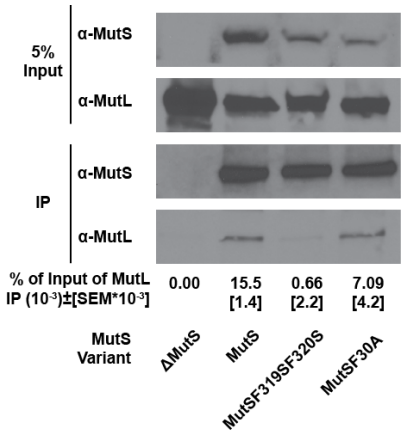


Figure 18. MutS crosslinks with MutL in the absence of mismatch detection *in vivo*.

Co-immunoprecipitation of MutS and MutL in the indicated backgrounds with affinity purified polyclonal antibodies against MutS. MutS and MutL levels were probed for using antiserum directed against MutS and for MutL. Band intensity was determined by using ImageJ quantitation software (See Experimental Procedures). Relative IP MutL levels reflect absolute band intensity per lane normalized to the wild type MutS lane. The error (SEM) were calculated from 3 independent experiments.

Discussion

Here, we have identified a conserved MutL binding site on MutS in the Gram-positive bacterium *B. subtilis*. Using peptide array mapping, extensive mutagenesis, single-cell fluorescence microscopy and *in vitro* crosslinking studies, we have identified residues found within the core domain, important for MMR *in vivo* and interaction with MutL-NTD *in vitro*. This site was further refined to a discrete MutL docking site composed of adjacent phenylalanine residues F319 and F320. Substitution of both phenylalanines to serine completely eliminates MMR *in vivo* and is defective for crosslinking to MutL-NTD *in vitro*. We also show that purified MutSF319SF320S is similar to wild type MutS for dimerization, ATPase activity, and binding to mismatched DNA substrates. We can therefore attribute the loss of MMR *in vivo* to a failure in MutL binding and recruitment. To our knowledge this effort defines the first MutL binding site on MutS in a bacterial organism lacking a methylation-directed MMR pathway.

Importantly, the di-phenylalanine motif that we identified in *B. subtilis* MutS to mediate interaction with MutL appears to be conserved and is part of a larger S[X]FF motif known to mediate MutL interaction with eukaryotic proteins. In *S. cerevisiae,* the S[X]FF motif was shown to be important for interaction between eukaryotic MutL homolog (Mlh1) and several Mlh1 bindings partners including Exo 1, BLM and Sgs1 proteins (39). Furthermore, a S[X]FF motif was also shown to mediate interaction between MutSβ (Msh2-Msh3) and MutLβ (Mlh1-Pms2), for the human proteins (42). In addition, a di-phenylalanine motif has been shown to be critical for interaction between mammalian translesion polymerases Rev1 and pol κ (40, 41). Our results in consideration with those above, show that adjacent phenylalanine residues play important roles in mediating protein interactions in a wide-variety of organisms.

Previous analysis of the *E. coli* MutS•MutL interaction identified residues in the mismatch recognition and connector domains involved in this interaction (24, 32). However, the connector domain of *E. coli* MutS on its own only weakly interacts with MutL, suggesting that additional surfaces on MutS may be involved in this interaction. While the mismatch recognition and connector domains are in close proximity to the core domain, the residues identified previously in *E. coli* MutS (Q211 and Q212) and patch 3B (F319 and F320) reside in opposite faces of the monomer and are separated

by the allosteric transmitter that connects the mismatch- and ATP-binding domains (**Figure 19**). It is conceivable that the different techniques used in our study and that by Mendillo and co-workers may have revealed distinct anchoring points of the MutS•MutL interface. If true, the MutS•MutL complex could adopt two distinct architectures. MutL could interact with both protomers of the MutS dimer to form a productive complex (arrow ii in **Figure 19**), thereby implying a mechanism to "check" MutS for mismatch binding through contacts with the mismatch-binding domain that would support the distance restraints reported by Winkler and co-workers (24).

Alternatively, MutL could interact with a single protomer of the MutS dimer (arrow i in **Figure 19**). This model poses an attractive mechanism to sense the mismatch and nucleotide binding states of MutS. The mismatch- and the nucleotide-binding domains of MutS are connected by a transmitter helix that runs along the outer rim of the MutS protomer (**Figure 19**) (36). Therefore, if MutL binds this face of MutS, the transmitter helix is probably a central feature of the interaction interface. This model also supports the established idea that only one of the MutS protomers mediates the interaction with MutL (32, 51, 52). Interestingly, mutation of patch 3A, which is part of the MutS transmitter, does not affect mismatch repair *in vivo* (**Table 3** and **Figure 1**). However, the MutS3A variant could not be over-expressed recombinantly in *E. coli*, implying a stability defect that could result from mutation of the transmitter. This, in turn, implies that MutL senses a different region of the transmitter, potentially α-helix 10 located in the C-terminus of the core domain (**Figures 15** and **19**).

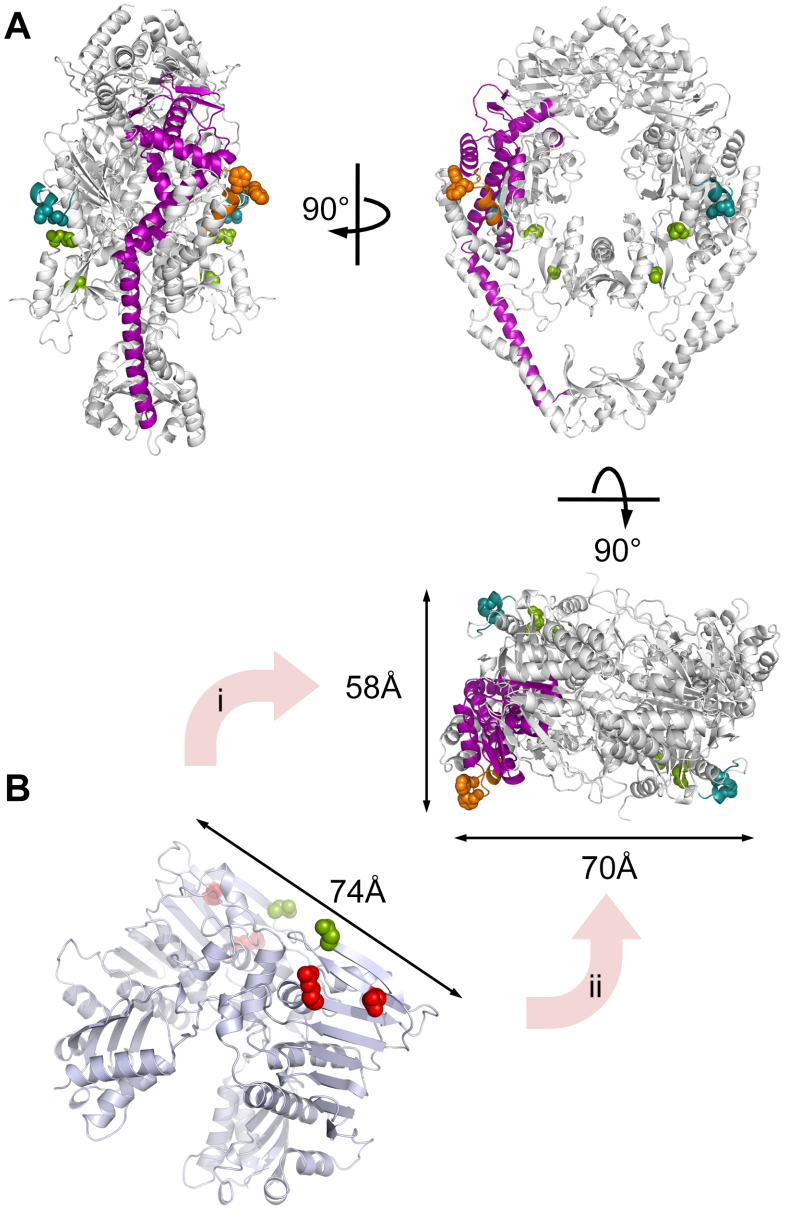


Figure 19. Potential interfaces of the MutS•MutL complex.
A) Orthogonal views of the *B. subtilis* MutS dimer shown as a ribbon diagram with the residues deemed important for the interaction with MutL shown in orange (FF motif, this work), teal (QQ motif, {32}) or green (distance constraints identified by crosslinking, {24}) The transmitter region of MutS is highlighted in purple. **B)** Ribbon diagram of the *E. coli* MutL-NTD dimer shown as a ribbon diagram with the residues identified in crosslinking studies shown in green (24) and additional residues deemed important for the interaction with MutS shown in red (60). Dimensions of the MutS and MutL dimer surfaces are indicated in angstroms (Å) and the two potential surfaces of MutS that MutL could recognize are indicated with pink arrows and labeled i and ii, respectively.

Upon identifying the binding interface, we extended our study to investigate the dynamic nature of MutS repair complexes *in vivo*. We present data showing that disruption of MutL recruitment causes repetitive MutS loading in response to mismatch formation. Since MutL recruitment is blocked we interpret this to mean that a MutS intermediate is “trapped” because the downstream step is prevented and we find that MutS foci persist with the number of MutS dimers per focus increased. These results support a model for repetitive loading by MutS in response to mismatch formation. Furthermore, even upon successful recruitment of MutL to a mismatch, we show that loss of function due to disruption of the endonuclease active site phenocopies a *ΔmutL* allele, supporting the hypothesis that not only does MutS loading occur independent from MutL recruitment, but that endonuclease directed nicking, and presumably excision of the mismatch, is a critical feature to disassemble MutS complexes. As more dimers of MutS load onto the mismatch proximal DNA, more MutS is available to recruit MutL. In support of this hypothesis, real-time *in vitro* imaging of MutSα and MutLα on DNA curtains revealed that the interaction requires a mismatch, yet interaction between MutSα and MutLα may occur after MutSα formed the ATP hydrolysis-dependent sliding clamp (53). Therefore, even mismatch-dissociated MutSα dimers can still facilitate a MutLα interaction, in essence amplifying a signal for MutL recruitment and for the advancement of repair. Our experiments represent *in vivo* data supporting repetitive loading of MutS at a mismatch, supporting previous *in vitro* experiments showing repetitive loading using circular mismatch containing substrates (23, 54). In addition to providing evidence for repetitive loading *in vivo*, we also provide evidence that mismatch excision is an important step in disassembling MutS repair complexes.

After excision of the mismatch, MutS loading is halted and the already bound MutS dimers will dissociate from the DNA in a timely manner, leading to disassembly of the repair center. *In vitro*, single molecule imaging reveals that after lesion recognition, the newly formed MutSα sliding clamp will remain on DNA with a lifetime of $t_{1/2}$≥198±23.4s (53). If loading is restricted to nascent DNA, then defective repair centers should persist on DNA flanking mismatches for up to 10 minutes after initial mismatch recognition. Furthermore, as DNA replication continues, newly replicated DNA moves farther away from the replisome, taking newly formed mismatches with it. In support of

the hypothesis that MutS repair centers defective for MutL recruitment persist longer on DNA surrounding the mismatch, the distribution of these MutS centers are located farther away from the replisome than repair centers engaging in active repair. Moreover, in exponentially growing cells, less repair centers colocalize with predominantly midcell replication centers. These two observations support the hypothesis that unproductive MutS repair centers persist on mismatch proximal DNA.

In *B. subtilis*, a model is emerging for the early steps of mismatch repair *in vivo*. We propose that MutS is positioned at the replisome preceding mismatch detection by a DnaN clamp zone that results from Okazaki fragment maturation (**Figure 37:App**) (31). MutS binds free DnaN clamps via a DnaN-binding motif (^{806}**QLSFF**) found in the unstructured C–terminal clamp-binding domain. MutS is able to find ~90% of mismatches through a DnaN coupled mechanism. Once MutS detects a mismatch, we propose that MutS (**Figure 37B**) loads repetitively at the mismatch, producing numerous DNA-bound MutS dimers (**Figure 37C**). We propose that repetitive MutS loading facilitates efficient MutL recruitment by increasing the local concentration of DNA bound MutS dimers surrounding the mismatch. MutS diffusing away from a mismatch with MutL may also help MutL identify strand discontinuities necessary to direct incision to the nascent strand (**Figure 37D**). Finally, the data presented here support the model that MutL-incision is necessary for disassembly of MutS complexes suggesting that mismatch excision is important for preventing further MutS loading.

Overall, this work describes the interaction between the core domain of MutS and MutL both *in vitro* and *in vivo*, and the implications of this interaction for the recruitment and activation of MutL at MutS repair centers, providing insight into the intermediate steps of mismatch repair in live cells.

Experimental Procedures

Bacteriological methods

B. subtilis strains were grown according to established procedures (28). Briefly, strains were grown in Luria-Bertani (LB) medium or defined $S7_{50}$ minimal medium. Unless otherwise stated antibiotics were used when appropriate with the following concentrations: 100 µg/mL spectinomycin (*spc*), 5 µg/mL chloramphenicol (*cam*), 5

μg/mL tetracycline (*tet*), 0.5 μg/ml erythromycin and 12.5 μg/mL lincomycin (*mls*), 5 μg/mL kanamycin (*kan*), 150 μg/mL rifampin (*rif*).

Peptide array analysis

The MutS peptide array was synthesized at the Massachusetts Institute of Technology Biopolymers laboratory, (Cambridge, MA). The synthesized MutS peptides provided 1x coverage spanned the entire amino acid sequence of MutS by overlapping 10 mer peptides offset by 3 residues. The final array consisted of 282 spotted peptides. The peptide array was activated by wetting with 100% ethanol, followed by 3 successive washes in Tris-buffered saline + Tween 20 (TBS-T) (50 mM Tris-Cl pH 7.5, 150 mM NaCl and 0.05% Tween 20) (pH 7.6) for 5 minutes to remove excess ethanol. The array was then blocked overnight in TBS-T and 10% milk solids at 4˚C. The following day, the array was washed in TBS-T, and then incubated in 56 nM MutL-myc in protein incubating solution (40 mM HEPES-KOH pH 7.6, 23 mM KCl, 1 mM $MgSO_4$, 1 mM DTT, 0.5 mg/mL BSA, 2% glycerol, and 0.5 mM of either AMPPNP or ADP) for 15 hours at 4˚C with gentle rocking. The next day, the array was washed for 30 minutes total with 3 washes each of the following buffers in order: TBS-T, TBS-T with 500 mM NaCl, TBS-T+0.5% Triton X-100, and TBS-T. The array was then incubated with 1:5000 α-myc antibody in TBS-T+5% milk for one hour at room temperature. The wash series was repeated, followed by incubation with 1:2000 anti-mouse in TBST+5% milk for 1 hour at room temperature. After antibody incubation, one more wash series was performed and the array was exposed using Pierce SuperSignal. Exposure time course of 2 minutes, 5 minutes, and overnight were obtained to identify bound peptides.

False-positive peptides were removed by comparing to a negative control (no myc-tagged protein exposure). Finally, peptides that were surface exposed based on a structure-guided sequence alignment of MutS homologs were deemed putative MutL binding peptides.

Strains and Plasmids

All *B. subtilis* strains used are derivatives of PY79 and are described in supplemental Table 1. Plasmids created for use in this study are as follow:

B. subtilis mutS and *mutL* expressing plasmids MutS (pAG 8483; residues 1-858) was amplified from *B. subtilis* strain 168 genomic DNA and ligated into pET-15b (Novagen) using restriction sites NdeI and BamHI. MutS variants Patch 1 (pAG 8561; E155S, R156S, L157A, E158S), Patch 2 (pAG 8674; E245S, E247S, E248S), Patch 3B (pAG 8635; F320S, E321S, R322S, E323S), Patch 4 (pAG 8634; E392S, E395S, E396S), Patch 5 (pAG 8616; E510S, E512S, E514S), Patch 6A (pAG 8646; Q806A, L807A, F809A, F810A), and Patch 6B (pAG 8535; D811S, E812S, E814S) were generated using overlap PCR and ligated into pET-15b using NdeI and BamHI. *B. subtilis* MutL N-terminal domain (MutL-NTD) (pAG 8286; residues 1-339) was amplified and ligated into pProEX HTa (Invitrogen) using NcoI and XhoI. All mutants were verified by DNA sequencing (MOBIX, McMaster University).

Purification of his_6MutS

B. subtilis MutS variants were overproduced in BL21 (DE3) pRARE or BL21 (DE3) pRARE pLysS cells (Invitrogen) and induced with 1 mM IPTG for 5 hours at 25°C. Cells were resuspended in buffer A (20 mM Tris pH 8.0, 0.5 M NaCl, 30 mM imidazole, 1.4 mM 2-mercaptoethanol, and 5% glycerol), lysed by sonication, and clarified by centrifugation at 39,000 g. The soluble fraction was purified over a nickel-chelating column equilibrated with buffer A and eluted with 240 mM imidazole. MutS was then injected onto an ion exchange column (Q-Sepharose, GE Healthcare) equilibrated with buffer B (20 mM Tris pH 8.0, 5 mM EDTA, 2.8 mM 2-mercaptoethanol, 100 mM NaCl, and 5% glycerol) and eluted using a linear gradient to 400 mM NaCl. MutS was injected into a gel filtration column (Superdex-200, GE Healthcare) equilibrated with crosslinking buffer (20 mM Hepes pH 7.5, 100 mM NaCl, 5 mM DTT, and 5% glycerol). Protein concentration was measured at 280 nm.

Purification of the MutL N-terminal domain

B. subtilis MutL-NTD was overexpressed in BL21 Star (DE3) cells (Invitrogen) with 0.5 mM IPTG for 5 hours at 25°C. MutL-NTD was purified using a nickel chelating column equilibrated with buffer A (pH 9.0) and eluted using 240 mM imidazole. MutL-NTD was then injected into a sizing column (Superdex-200, GE Healthcare) equilibrated

with crosslinking buffer. Protein concentration was measured by absorbance at 280 nm.

Spontaneous mutation rate analysis

Fluctuation analysis was performed essentially as described (31, 38). We inoculated 3 mL of LB with a single colony, and grew at 37˚C until an OD_{600} of ~1.2. At that point, 1 mL of culture was pelleted and resuspended in 100 µL of saline. A portion of this resuspension was further diluted to 10^{-6}, and plated onto LB plates in order to enumerate the total viable cells with incubation overnight at 30°C to ensure the plates with viable cells did not over grow. The original resuspension was plated on LB supplemented with 150 µg/mL rifampin plates overnight at 37°C in order to determine the number of spontaneous mutations causing rifampin resistance. After performing a minimum of 15 independent cultures, the mutation rate was determined using the MSS Maximum Likelihood Method using the publicly available FALCOR tool at http://www.mitochondria.org/protocols/FALCOR.html. 95% confidence intervals were determined and percent mismatch repair activity, was determined using the following equation:

[(RMR null – RMR strain)/(RMR null – RMR wild type)]•100 where RMR = relative mutation rate (55).

Chemical crosslinking

B. subtilis MutS variants (20 μM), 20 μM Mis90, and 20 mM ATP were pre-incubated on ice for 1 hour. MutL-NTD (40 μM) was then added with equal volume to the MutS•ATP•DNA reaction and incubated for 30 minutes at 4°C. Reactions were then incubated with 0.8-1.6 mM bis (sulfosuccinimidyl) suberate (Sigma, BS^3) for 30 minutes at 22°C. Reactions (10 μL) were quenched with 30 mM Tris pH 7.5 for 15 minutes at 22°C and separated on a 4-15% SDS gradient gel (BioRad) and stained with Coomassie Blue.

ATPase Assay

ATP hydrolysis assays were performed as previously described (56) with minor modifications. ATPase activity was measured with 0.3 μM MutS and 5 mM $MgCl_2$ in reaction buffer (20 mM Tris pH 8.0, 90 mM KCl, 1 mM DTT, 1 mg/ml BSA, and 5% glycerol). Reactions (15 μL) were initiated by the addition of 1 mM α-^{32}P-labeled ATP and incubated for 1 hour at 22°C. Reactions were stopped with 25 mM EDTA and hydrolyzed product was detected by thin-layer chromatography using 750 mM KH_2PO_4 for running buffer. ATPase activity was measured in triplicates for each MutS variant.

DNA binding- Mis90 is a 90 base pair DNA substrate harboring a G/T base mismatch (5' gaaaacctgtattt**t**cagggcaggcctattggaattcaacatatgaagtcgacgcagctggcggccgcttctagaggatcc ctcgagaag 3' annealed to 5' gcttctcgagggatcctctagaagcggccgccagctgcgtcgacttcatatgttgaattccaataggcctgccctg**g**aaata caggtttt 3'). MutS (600 pmol) was incubated with equimolar Mis90 in binding buffer (10 mM Hepes pH 7.5, 70 mM KCl, 2 mM DTT, 5 mM $MgCl_2$, 1 mg/ml BSA, and 15% glycerol) for 1 hour on ice. Reactions (15 mL) were resolved on a 6% TBE gel and stained with ethidium bromide. Bands were quantified using ImageJ (http://rsbweb.nih.gov/ij/). DNA binding activity was measured in triplicates for each MutS variant.

Live cell microscopy

Cultures for imaging were prepared as described previously (28, 30, 31). Briefly, strains for imaging were inoculated in pre-warmed $S7_{50}$ minimal media supplemented with either 1% L-arabinose or 2% D-glucose at a starting OD_{600} of 0.05. Cells were grown past three doublings to an OD_{600} of 0.4-0.5 and imaged. To treat cultures with the mismatch-forming drug 2-aminopurine, we split the cultures and added a mock treatment to one and 600 μg/mL 2-aminopurine to the other followed by growth for an additional hour. Cell membranes were visualized with the fluorescent dye TMA-DPH at a working concentration of 10 μM (31). MutS fluorescent fusions were captured with a 1.2 second exposure. Colocalization experiments were conducted with L-arabinose as the sole carbon source, where all other experiments used D-glucose.

In vivo crosslinking/co-immunoprecipitation

B. subtilis cultures were inoculated in LB at a starting OD_{600} of 0.05 and grown at 37°C to an OD_{600} of 0.7. Cells were pelleted, washed twice with crosslinking buffer (40 mM HEPES pH 7.4, 500 mM sucrose, 2 mM $MgCl_2$, 150 mM NaCl, 0.02% Tween-20) to remove LB, and resuspended in 1.75 mL of crosslinking buffer. To crosslink intracellular protein complexes, 0.5 mM of Dithiobis[succinimidyl propionate] (DSP) was added to the growing cells and crosslinking occurred for 30 minutes at room temperature on a rotisserie. Cultures were quenched by adding Tris-HCl (pH 7.5) to a final concentration of 20 mM, and incubated an additional 30 minutes at room temperature on a rotisserie. After quenching, cells were lysed via sonication. Lysates were cleared of debris by centrifugation for 30 minutes at 4° C at 14,000 rpm. Lysates were then concentrated to 50 µL, resuspended in crosslinking buffer supplemented with 1X protease inhibitor cocktail and 0.5 mM EDTA to a final volume of 500 µL. A 5% input fraction was pulled from the final volume. The 5% input and the rest of the prepared lysate were incubated overnight on a rotisserie at 4°C. The IP fraction was incubated with 50 µL equilibrated magnetic beads bound with affinity purified α-MutS antisera (MI-1042). Beads were prepared according to protocol. In the morning, the lysates were washed 5X for 5 minutes each with crosslinking buffer on a room temperature rotisserie. The antibodies were eluted from the magnetic beads by a 10 minute incubation in 900 µL of antibody stripping buffer (5 mM Glycine pH 2.4, 150 mM NaCl). The IP fraction was concentrated by TCA precipitation, and resuspended in 1X western loading dye. IP and Input fractions were electrophoresed on the same gel (4-15% gradient gel). Quantitative analysis of the resulting bands was conducted in ImageJ. The numbers represent the statistical mean of 3 independent experiments with the background subtracted from the JSL281 strain. Relative numbers were determined relative to JSL364 (PY79 wild type strain).

Western and Far Western Blotting

B. subtilis whole-cell extracts were obtained by centrifuging 25 mL of mid-exponential cultures, followed by resuspension in lysis buffer [10 mM Tris-HCl (pH 7.0), 0.5 mM EDTA, 1 mM AEBSF, 1X Protease Inhibitor cocktail) followed by 3 rounds of

sonication (20 Hz, 45 second duration) on ice as described (31). After sonication, SDS was added to a final concentration of 1% and non-soluble cellular debris and whole cells were removed by centrifugation at 4˚C. The lysate was divided into one-time use samples and stored at -20˚C. Total protein concentration of prepared soluble lysates was determined using Pierce BCA Protein Assay Kit (Thermo Scientific). Equal amounts of total protein were applied to each lane on a 4-15% gradient gel followed by transfer to a nitrocellulose membrane (57, 58). Protein levels were determined by using primary antisera: α-MutS (MI-1042), α-MutL (MI-1044), and α-DnaN (MI-1038).

Immunodot blotting was performed as described (30). Briefly, equal molar amounts of the indicated proteins were immobilized onto a nitrocellulose membrane with the assistance of a Bio-dot microfiltration apparatus (Bio Rad). The membrane was incubated in blocking buffer (5% milk solids, 17.4 mM Na_2HPO_4, 2.6 mM NaH_2PO_4, 150 mM NaCl, 0.05% Tween-20, 0.5 mM ATP, 4 mM $MgSO_4$) at 22˚C for one hour. All subsequent washes and incubations took place in blocking buffer. After blocking, the membrane was incubated with 0.4 μM MutL in blocking buffer for 3 hours at 22˚C. The blot was subsequently washed three times and then incubated in affinity purified α-MutL antisera overnight at 4˚C. In the morning, the blot was removed from primary antibody and washed three times at 22˚C and placed in secondary antisera (1:2000 α-Rabbit) for 2 hours at 22˚C. The blot was washed 3 more times, followed by a wash in PBS (17.4 mM Na_2HPO_4, 2.6 mM NaH_2PO_4, 150 mM NaCl, 0.05% Tween-20) to remove excess milk solids and exposed.

Acknowledgements

We thank members of the Simmons and Guarné laboratories for helpful discussions and comments on this manuscript. We would also like to thank Nicholas Bolz for assistance in early cloning and mutagenesis assays. This work was supported by grant MCB1050948 from the National Science Foundation to L.A.S. and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to A.G. In addition, M.C.P was supported in part by a doctoral scholarship from NSERC and J.S.L was supported in part by a predoctoral fellowship from the Rackham Graduate School at the University of Michigan.

Supplemental Material

Supplemental Material accompanying this chapter is found in Appendix II: Supplemnetal procedures, **Table 8 and 9**, and **Figures 31-37**.

References

1. Iyer RR, Pluciennik A, Burdett V, & Modrich PL (2006) DNA mismatch repair: functions and mechanisms. *Chem Rev* 106(2):302-323.
2. Kunkel TA & Erie DA (2005) DNA mismatch repair. *Annu Rev Biochem* 74:681-710.
3. Lenhart JS, Schroeder JW, Walsh BW, & Simmons LA (2012) DNA Repair and Genome Maintenance in Bacillus subtilis. *Microbiology and molecular biology reviews : MMBR* 76(3):530-564.
4. Schofield MJ & Hsieh P (2003) DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* 57:579-608.
5. Cooper LA, Simmons LA, & Mobley HL (2012) Involvement of Mismatch Repair in the Reciprocal Control of Motility and Adherence of Uropathogenic Escherichia coli. *Infect Immun* 80(6):1969-1979.
6. Cox EC, Degnen GE, & Scheppe ML (1972) Mutator gene studies in Escherichia coli: the mutS gene. *Genetics* 72(4):551-567.
7. Davies BW*, et al.* (2011) DNA damage and reactive nitrogen species are barriers to Vibrio cholerae colonization of the infant mouse intestine. *PLoS Pathog* 7(2):e1001295.
8. Ginetti F, Perego M, Albertini AM, & Galizzi A (1996) Bacillus subtilis mutS mutL operon: identification, nucleotide sequence and mutagenesis. *Microbiology* 142 (Pt 8):2021-2029.
9. Prudhomme M, Martin B, Mejean V, & Claverys JP (1989) Nucleotide sequence of the *Streptococcus pneumoniae hexB* mismatch repair gene: Homology of *HexB* to *MutL* of *Salmonella typhimurium* and to *PMS1* of *Saccharomyces cerevisiae*. *J. Bacteriol.* 171(10):5332-5338.
10. Fishel R, Ewel A, Lee S, Lescoe MK, & Griffith J (1994) Binding of mismatched microsatellite DNA sequences by the human MSH2 protein. *Science* 266(5189):1403-1405.
11. Umar A*, et al.* (1994) Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. *The Journal of Biological Chemistry* 269(20):14367-14370.
12. Fishel R*, et al.* (1993) The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis cancer. *Cell* 75:1027-1038.

13. Hamilton SR*, et al.* (1995) The molecular basis of Turcot's syndrome. *N Engl J Med* 332(13):839-847.

14. Nystrom-Lahti M*, et al.* (2002) Functional analysis of MLH1 mutations linked to hereditary nonpolyposis colon cancer. *Genes Chromosomes Cancer* 33(2):160-167.

15. Peltomaki P (2005) Lynch syndrome genes. *Fam Cancer* 4(3):227-232.

16. Klein E, Smith DL, & Laxminarayan R (2007) Hospitalizations and deaths caused by methicillin-resistant Staphylococcus aureus, United States, 1999-2005. *Emerging infectious diseases* 13(12):1840-1846.

17. Klevens RM*, et al.* (2007) Invasive methicillin-resistant Staphylococcus aureus infections in the United States. *Jama* 298(15):1763-1771.

18. Kluytmans J, van Belkum A, & Verbrugh H (1997) Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. *Clin Microbiol Rev* 10(3):505-520.

19. Lowy FD (1998) Staphylococcus aureus infections. *The New England journal of medicine* 339(8):520-532.

20. Larrea AA, Lujan SA, & Kunkel TA (2010) SnapShot: DNA mismatch repair. *Cell* 141(4):730 e731.

21. Blackwell LJ, Bjornson KP, Allen DJ, & Modrich P (2001) Distinct MutS DNA-binding modes that are differentially modulated by ATP binding and hydrolysis. *The Journal of Biological Chemistry* 276(36):34339-34347.

22. Bjornson KP, Allen DJ, & Modrich P (2000) Modulation of MutS ATP hydrolysis by DNA cofactors. *Biochemistry* 39(11):3176-3183.

23. Acharya S, Foster PL, Brooks P, & Fishel R (2003) The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Molecular cell* 12(1):233-246.

24. Winkler I*, et al.* (2011) Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in Escherichia coli. *The Journal of Biological Chemistry* 286(19):17326-17337.

25. Ban C, Junop M, & Yang W (1999) Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell* 97(1):85-97.

26. Ban C & Yang W (1998) Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell* 95(4):541-552.

27. Guarne A*, et al.* (2004) Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *The EMBO journal* 23(21):4134-4145.

28. Dupes NM*, et al.* (2010) Mutations in the Bacillus subtilis beta clamp that separate its roles in DNA replication from mismatch repair. *Journal of bacteriology* 192(13):3452-3463.

29. Simmons LA, Davies BW, Grossman AD, & Walker GC (2008) Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Molecular cell* 29(3):291-301.

30. Klocko AD*, et al.* (2011) Mismatch repair causes the dynamic release of an essential DNA polymerase from the replication fork. *Molecular microbiology* 82(3):648-663.

31. Lenhart JS, Sharma A, Hingorani MM, & Simmons LA (2013) DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular microbiology* 87(3):553-568.

32. Mendillo ML*, et al.* (2009) A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proceedings of the National Academy of Sciences of the United States of America* 106(52):22223-22228.

33. Klocko AD, Crafton KM, Walsh BW, Lenhart JS, & Simmons LA (2010) Imaging mismatch repair and cellular responses to DNA damage in Bacillus subtilis. *J Vis Exp* (36):1-4.

34. Sacho EJ, Kadyrov FA, Modrich P, Kunkel TA, & Erie DA (2008) Direct visualization of asymmetric adenine-nucleotide-induced conformational changes in MutL alpha. *Molecular Cell* 29(1):112-121.

35. Lamers MH*, et al.* (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* 407(6805):711-717.

36. 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(6805):703-710.

37. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, & Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc. Natl. Acad. Sci. U.S.A.* 98(20):11627-11632.

38. Bolz NJ, Lenhart JS, Weindorf SC, & Simmons LA (2012) Residues in the N-terminal domain of MutL required for mismatch repair in Bacillus subtilis. *Journal of Bacteriology* 194(19):5361-5367.

39. Dherin C*, et al.* (2009) Characterization of a highly conserved binding site of Mlh1 required for exonuclease I-dependent mismatch repair. *Molecular and Cellular Biology* 29(3):907-918.

40. Wojtaszek J*, et al.* (2012) Structural basis of Rev1-mediated assembly of a quaternary vertebrate translesion polymerase complex consisting of Rev1, heterodimeric polymerase (Pol) zeta, and Pol kappa. *The Journal of Biological Chemistry* 287(40):33836-33846.

41. Wojtaszek J*, et al.* (2012) Multifaceted recognition of vertebrate Rev1 by translesion polymerases zeta and kappa. *The Journal of Biological Chemistry* 287(31):26400-26408.

42. Iyer RR*, et al.* (2010) MutLalpha and proliferating cell nuclear antigen share binding sites on MutSbeta. *The Journal of Biological Chemistry* 285(15):11730-11739.

43. Smith BT, Grossman AD, & Walker GC (2001) Visualization of mismatch repair in bacterial cells. *Mol. Cell* 8(6):1197-1206.

44. Landgraf D, Okumus B, Chien P, Baker TA, & Paulsson J (2012) Segregation of molecules at cell division reveals native protein localization. *Nature methods* 9(5):480-482.

45. Pillon MC*, et al.* (2010) Structure of the endonuclease domain of MutL: unlicensed to cut. *Molecular cell* 39(1):145-151.

46. Hombauer H, Campbell CS, Smith CE, Desai A, & Kolodner RD (2011) Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 147(5):1040-1053.

47. Berkmen MB & Grossman AD (2006) Spatial and temporal organization of the Bacillus subtilis replication cycle. *Molecular microbiology* 62(1):57-71.

48. Lemon KP & Grossman AD (1998) Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282:1516-1519.

49. Lemon KP & Grossman AD (2000) Movement of replicating DNA through a stationary replisome. *Mol. Cell* 6(6):1321-1330.

50. Kijas AW, Studamire B, & Alani E (2003) Msh2 separation of function mutations confer defects in the initiation steps of mismatch repair. *Journal of Molecular Biology* 331(1):123-138.

51. Habraken Y, Sung P, Prakash L, & Prakash S (1997) Enhancement of MSH2-MSH3-mediated mismatch recognition by the yeast MLH1-PMS1 complex. *Current biology : CB* 7(10):790-793.

52. Prolla TA, Pang Q, Alani E, Kolodner RD, & Liskay RM (1994) MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science* 265(5175):1091-1093.

53. Gorman J*, et al.* (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):E3074-3083.

54. Gradia S*, et al.* (1999) hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Molecular Cell* 3(2):255-261.

55. Hall BM, Ma CX, Liang P, & Singh KK (2009) Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics* 25(12):1564-1565.

56. Junop MS, Yang W, Funchain P, Clendenin W, & Miller JH (2003) In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair (Amst)* 2(4):387-405.

57. Simmons LA*, et al.* (2009) Comparison of responses to double-strand breaks between Escherichia coli and Bacillus subtilis reveals different requirements for SOS induction. *Journal of bacteriology* 191(4):1152-1161.

58. Simmons LA & Kaguni JM (2003) The DnaAcos allele of Escherichia coli: hyperactive initiation is caused by substitution of A184V and Y271H, resulting in defective ATP binding and aberrant DNA replication control. *Molecular microbiology* 47(3):755-765.

59. Arnold K, Bordoli L, Kopp J, & Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22(2):195-201.

60. Plotz G*, et al.* (2006) Mutations in the MutSalpha interaction interface of MLH1 can abolish DNA mismatch repair. *Nucleic Acids Research* 34(22):6574-6586.

Chapter IV

RecO and RecR are necessary for RecA loading in response to DNA damage and replication fork stress in *Bacillus subtilis*

Justin S. Lenhart [1], Eileen R. Brandes [1], Jeremy W. Schroeder[1], Roderick J. Sorenson[2], Hollis D. Showalter[2] and Lyle A. Simmons[1*]

[1]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109-1048

[2]Vahlteich Medicinal Chemistry Core, College of Pharmacy, University of Michigan, Ann Arbor, MI 48109-1048

Author contributions: Figures 20 and 22B-D were performed by JSL. Figure 22A and B was performed by ERB. Figure 21B was performed by JWS. Figure 21A was performed by RJS. JSL, LAS, JWS, and HDS designed the research. All authors contributed to the writing of this chapter.

In Revision: Journal of Bacteriology

*Corresponding author

Mailing address: 4140 Kraus Natural Science building Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109. Phone: (734) 647-2016. E-mail: lasimm@umich.edu

Abstract

The recombinase RecA is required for homologous recombination and stabilization of stalled replication forks in many bacteria. In *Escherichia coli,* the RecFOR and RecBCD pathways serve to load RecA, and the choice between these two pathways depends on the type of damage encountered. Using RecA-GFP filament assembly as a proxy to study RecA loading *in vivo*, we determined which recombinase mediator proteins are necessary in bacteria lacking the canonical RecBCD pathway. We find in *Bacillus subtilis* that the rapid localization of RecA is dependent on the RecOR pathway in response to all types of damage examined. This finding excludes AddAB, the RecBCD functional homologue, as an alternate RecA loading pathway since RecOR is necessary for RecA loading to occur. Furthermore, we find that RecF is not required for RecOR-mediated loading of RecA *in vivo*, yet the presence of RecF increases the efficiency of either RecA filament nucleation or elongation. Additionally, we tested the role of the single-stranded DNA binding protein (SSB) during RecA loading, as SSB has two opposing roles: it represses RecA loading by competing for ssDNA, yet recruits the RecA mediator protein RecO to tracts of single-stranded DNA. Truncation of the C-terminal tail of SSB, which interacts with RecO, eliminates RecA filament formation in response to DNA damage. We suggest that this mechanism ensures that RecA is loaded only when RecO is present, reducing aberrant filament formation. Overall, we provide novel insight into the mechanism that establishes RecA repair filaments in live cells.

Introduction

DNA repair pathways are critical for genome maintenance in all living cells [for review (1)]. Both prokaryotic and eukaryotic organisms are constantly exposed to endogenous and exogenous sources of DNA damage that compromise genome integrity, where a single unrepaired double-stranded break (DSB) is lethal. In response to clastogenic events, DNA damage encountered by replication forks result in the recruitment of DNA tolerance and repair pathways, headlined by the multi-faceted recombinase RecA (RAD51 in eukaryotes) to stabilize stalled forks and facilitate strand exchange during homologous recombination [for review (2-4)].

In vitro, RecA loads onto naked DNA in two kinetic steps: 1) nucleation and 2) filament elongation. When DNA is pre-coated with SSB, RecA loading is inhibited as SSB outcompetes RecA for ssDNA (5). Recombinase mediator proteins (RMPs) are biochemically and genetically defined proteins that facilitate RecA loading by reducing the kinetic barriers of nucleation by sliding/unwrapping SSB from ssDNA (6). Nucleation is complete when a RecA filament reaches a critical mass of 2-5 RecA monomers, establishing position on ssDNA by SSB occlusion, supporting future filament elongation (7, 8). Elongation of RecA nucleates occurs either by spontaneous or RMPs mediated addition of RecA monomers bi-directionally, with the majority of growth in the 5′-3′ polarity (7, 8).

The mechanisms *E. coli* employs to both tolerate and repair DNA damage has been extensively studied. DSBs are produced when either both strands of DNA get severed or when the replication fork replicates over a single-stranded nick, producing a free DNA end. In response to DSBs, *E. coli* loads RecA via two major pathways: RecBCD and RecFOR [for review (3, 9-11)]. Most DSBs (95-99%) are processed by the RecBCD pathway, and when deleted, results in a substantial loss of viability (~70%) and increased sensitivity to DNA damaging agents (12-15). During DSB repair, double-stranded ends are bound by the RecBCD helicase-nuclease complex, which after association, unwinds and digests both strands simultaneously (16, 17). Upon encountering a χ site (a RecBCD regulatory motif 5′-GCTGGTGG-3′), RecBCD nuclease activity is attenuated in the 3′-5′ polarity, while endonucleolytic activity is switched primarily to the 5′-3′ polarity, generating a 3´ssDNA extension (17). Concurrent

with end processing, RecBCD physically loads RecA onto the 3´ssDNA extension, which nucleates RecA filament formation (18, 19). This newly generated RecA nucleoprotein filament is now capable of synapsis formation and homology search, ultimately identifying a homologous template to facilitate repair of the DSB by error-free homologous recombination.

The second loading pathway in *E. coli* is the RecFOR pathway, and is predominantly charged with RecA loading during daughter strand gap repair (20). A daughter strand gap is produced when the replication fork tries to replicate over a damaged base in the template strand, which produces ssDNA as replication restarts on the other side of the lesion. In this pathway, RecO and RecR form a complex (RecOR) that is required for RecA loading onto SSB coated ssDNA *in vitro* (20, 21). RecOR, which does not interact with RecA, is thought to locally displace SSB along the DNA. The third component, RecF, has been shown to bind to ss/dsDNA junctions, making it a critical component for RecA loading on daughter strand gaps, yet its exact biochemical function and whether it forms a physical complex with RecOR, is uncertain (21). Biochemically, RecOR and RecFOR function by accelerating RecA nucleation, while RecOR has been shown to further enhance the rate of filament elongation (6, 22, 23). The RecA loading substrates of RecFOR can be expanded to sites of DSBs when RecBCD and the single-stranded exonucleases ExoVIII (*sbcA*) and/or ExoI (*sbcB*) function are disrupted (24). Therefore, *E. coli* RecFOR is capable of all RecA loading activities in the cell when ssDNA persists at break sites in this special situation.

The RecBCD pathway is found in 39.0% of bacteria, while the homologous AddAB pathway is found in 52.6% of bacteria; both pathways are mutually exclusive and covering most taxonomic groups (25). In *B. subtilis*, the AddAB helicase-nuclease enzyme is required for double-stranded end processing (26, 27). Like RecBCD, AddAB end processing generates a 3´ extension suitable for RecA binding and polymerization into a nucleoprotein filament (26, 27). In contrast to RecBCD, *B. subtilis* AddAB has not been shown to physically bind or directly load RecA. The *B. subtilis* genome also encodes the RecFOR proteins (28). A notable difference is that *B. subtilis* RecO is necessary and sufficient to load RecA onto SSB-coated ssDNA *in vitro* (21, 29). Interestingly, prior work showed that RecA-GFP foci were largely *recO* and *recR*-

independent and defects in AddAB did not substantially reduce RecA-GFP foci *in vivo* (30, 31). Therefore in *B. subtilis*, the process responsible for loading RecA in response to DNA breaks, strand gaps, and damage-independent replication fork arrest remains unknown.

Here, we investigate candidate RMPs for their ability to load RecA and establish repair centers *in vivo*. We employed several exogenous damaging agents in equitoxic doses to elicit various forms of DNA damage repair and monitor the cells ability to form RecA-GFP foci, a proxy for RecA loading. Strikingly, we found that the RMPs RecO and RecR are necessary for RecA to load and organize into repair foci/filaments in response to endogenous damage and all sources of exogenous DNA damage and fork arrest tested. These results support an exclusive role for RecOR in RecA filament nucleation on ssDNA in *B. subtilis* and show that AddAB is unlikely to direct RecA loading. RecF, another RMP, was shown to delay RecA filament formation when function was eliminated, suggesting a role in RecA nucleation, enhancement of extension, or possibly both for DSBs and DSGs. We also show that truncation of the C-terminal 35 amino acids from SSB prevents the DNA damage-dependent increase in assembly of RecA repair centers, supporting a model where SSB recruits the RecOR complex to damaged DNA at the replication fork for nucleation and assembly of RecA repair centers. With these results, we propose that RecOR represents the major RecA loading pathway in *B. subtilis* in response to DNA damage, damage-independent fork arrest, and double-stranded breaks.

Results

RecA colocalizes to the replisome in response to DNA damage.

Previous work in *E. coli* and *B. subtilis* examined the ability of RecA-GFP to organize into foci (14, 15, 30, 32-37). Prior experiments have either tested RecA-GFP as the only source of RecA *in vivo* (14, 32, 33, 38) or tested GFP-RecA expressed ectopically with or without the native *recA* allele intact (30, 36). Here, we examined RecA-GFP as the only source of RecA in the cell, expressed at its native locus and under control of its native promoter (33). The fusion allele is fully functional at low levels of damage (**Figure 38: App** and **Appendix III Results**). Unless otherwise indicated, the

experiments described below were performed under conditions of DNA damage where the *recA-GFP* allele was fully functional (either the dose is greater than the LD_{90} or equitoxic with PY79).

A

Colocalization to DnaX-mCherry

Condition	% Colocalization (+/- 95% CI)	Number of Cells Analyzed
Endogenous (untreated)	74.8 (8.4)	103
40 nM Mitomycin C	65.0 (6.7)	197
UV 40 Joules/m ^{2}	84.3 (5.8)	153
400 nM Phleomycin	67.9 (6.3)	212
40 µM HPUra	93.3 (3.3)	223

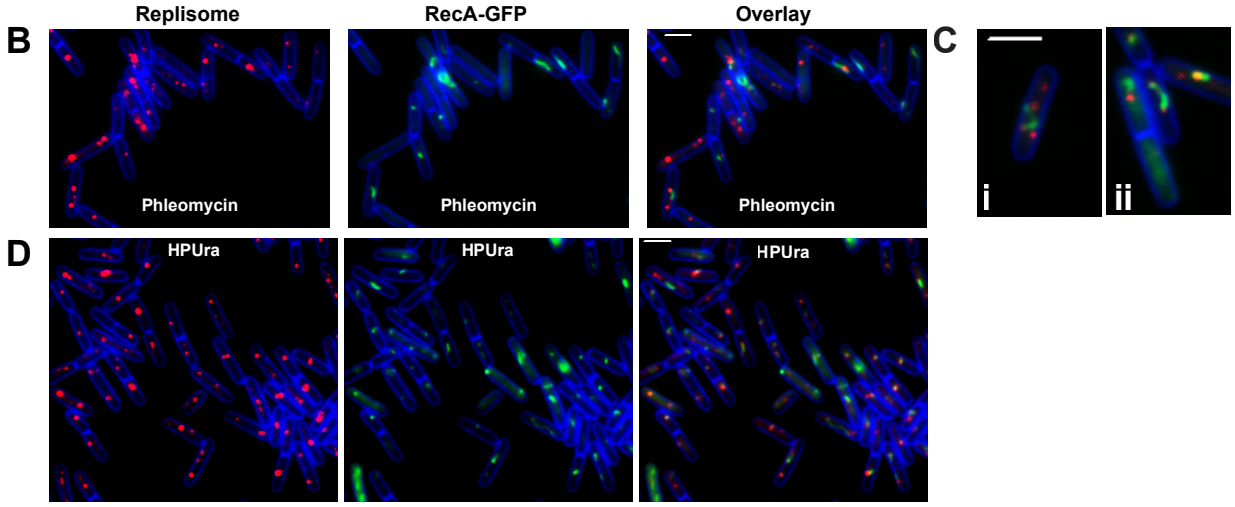


Figure 20. RecA colocalizes to the replisome.
(A) Percentage of DnaX-mCherry foci colocalized with RecA. **(B-D)** Representative micrographs of RecA-GFP and DnaX-mCherry in cells challenged with phleomycin (**B-C**) and (**D**) HPUra. The white scale bar represents 4 µm. The membranes were stained with TMA-DPH (Blue). (**C**) Shows two examples of RecA-GFP bridging replisome foci following phleomycin treatment.

One current model suggests that RecA predominantly localizes away from the replisome in response to DNA damage and that replication is not required for repair center formation (39). An alternate model predicts that replication is required for RecA-GFP to form foci (33) and that these foci form at midcell, a localization consistent with the site of ongoing replication (14, 33, 36). To differentiate between these two models, we constructed a strain with fluorescent fusions to RecA and the replisome marker DnaX as the only source of each of these proteins in the cell. Imaging of this strain

showed that in untreated cells, ~75% of DnaX (replisome) foci colocalized with a RecA focus with colocalization occurring at RecA foci or at the end of long RecA protrusions termed filaments. This observation, in consideration with prior published work (33), leads us to suggest that RecA initiates filament formation to establish repair centers at the replication fork, followed by a search for homologous DNA (**Figure 20A**).

With this result, we asked if different forms of DNA damage cause RecA to differentially localize to or away from the replisome. Various agents were used to elicit unique DNA damage or replication stresses. Following treatment with mitomycin C (MMC), an agent that forms mostly bulky monoadducts and interstrand crosslinks (40), ~65% of replisome foci colocalized with RecA filaments/repair centers (**Figure 20A**). The drug phleomycin, which is a single stranded and double stranded DNA break inducing peptide, predominantly requires functional end-processing. A strain lacking all end-processing, *ΔrecJ, addA::erm*, is strongly deficient in RecA filament formation in response to phleomycin treatment (**Table 5**). Similarly to MMC treatment, challenge of cells with phleomycin showed ~68% of RecA colocalized with replisomes (**Figure 20A** and **B**). Both MMC and phleomycin-induced damage are expected to require homologous recombination (41). The majority of RecA foci that assembled in response to MMC and phleomycin colocalized with the replisome, supporting earlier work that replication fork progression is important for RecA repair center assembly in response to a site-specific DSBs (33).

We further tested co-localization of RecA during repair of daughter strand gaps using UV. Following challenge with UV, daughter strand gaps and replication-stalling thymine-thymine dimers represent the major lesion produced. Daughter strand gaps on the lagging strand are easily bypassed during Okazaki fragment synthesis; however, upon encountering damage on the leading strand, DNA synthesis between the two replication forks are momentarily uncoupled, producing a stretch of ssDNA. Here, we employed a dose of UV where the leading strand replication fork would encounter a lesion on average every 8.5 s (42-44). Within 5 minutes, we found that RecA foci formed immediately, and ~84% of replisomes were colocalized with RecA (**Figure 20A**). With these results, we conclude that the majority of RecA repair centers localize to the

replisome in response to endogenous and exogenous sources of DNA damage, including daughter strand gaps and DNA breaks.

Interestingly, we occasionally observed RecA filaments bridging two replisomes in untreated cells, as well as following both phleomycin and MMC treatments (**Figure 20C**). In *E. coli*, RecA filaments have been shown to pair with distant sister loci during repair (37). Likewise, we suggest that these events represent RecA coated ssDNA filaments probing for sister loci. We hypothesize that in these cells undergoing multi-fork replication (2 or more replisomes), the same loci is being replicated at roughly the same time, which provides the RecA filament with a sister template that is clear of protein occlusions, making an ideal template for homologous recombination repair.

RecA colocalizes to the replisome in response to damage-independent fork arrest.

The compound HPUra has been widely used as a tool to rapidly block replication fork progression in *B. subtilis* and other Gram-positive bacteria [e.g. (34, 45-50)]. HPUra is a replication-specific class III DNA polymerase inhibitor which reversibly blocks replication within minutes (51). To study the RecA response to DNA damage-independent replication fork arrest, we chose to use this compound because it has been so well characterized, with the limitation that HPUra is not commercially available.

We began by synthesizing HPUra (see Appendix III). From prior literature, the hydrazine congener H_2-HPUra (compound **4**) is described as the compound used to inhibit DNA synthesis [e.g. (45, 48-50)]. We show the scheme for the attempted synthesis of compound **4** and related congeners in **Figure 21A.** Briefly, we found that HPUra (compound **3**) is readily synthesized as described (50). However, numerous attempts to reduce HPUra to H_2-HPUra using sodium dithionite as described (48), or slight variations thereof, did not provide compound **4** but only recovered compound **3**. This led us to the conclusion that either compound **3** is active as an inhibitor of DNA synthesis or compound **3** is reduced within the cell to produce compound **4**. A detailed description of the synthesis and structure determination for **3** and congeners is provided in Appendix III (**data not shown**). To test compound **3** for DNA synthesis inhibition, we synchronized replication initiation using a temperature sensitive mutant of *dnaB* as

described (33). We extracted untreated cells for analysis at 20 and 60 minutes post release and determined the position of replication forks by quantitative illumina sequencing (**Figure 21B**). Cells treated with HPUra were treated at 20 minutes post release and genomic DNA was isolated at 60 minutes post release. We show that replication forks in untreated cells progress throughout the 60-minute time course, while DNA replication in the HPUra treated cells arrests immediately upon HPUra addition (20 min after release). Therefore, we show that although H_2-HPUra is likely only transiently present during its attempted synthesis, being unstable and readily oxidizing to HPUra (compound **3**), HPUra provides the desired effect of rapidly arresting DNA synthesis (**Figure 21B**).

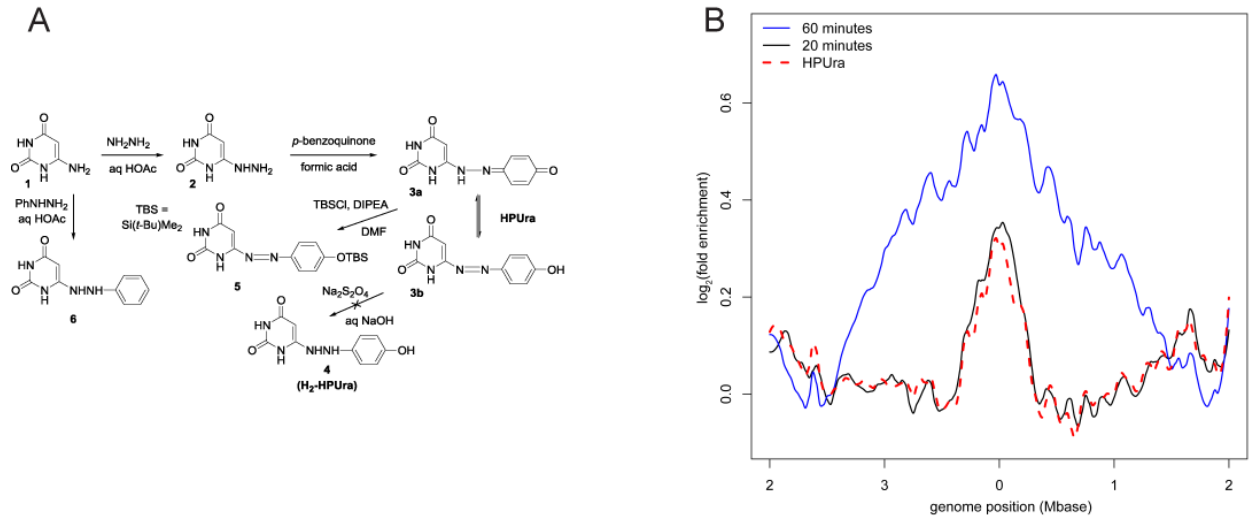


Figure 21. HPUra immediately stops DNA synthesis.
(A) Shows the synthesis scheme for HPUra (compound 3). **(B)** Log_2 (fold-enrichment) of read coverage in 1000 nucleotide wide windows is plotted versus PY79 genome position. The plot is centered on the origin of replication. Because the HPUra stock was dissolved in 50 mM KOH, the 60 minute control cells were treated with an equal volume of 50 mM KOH at 20 minutes with cells harvested at 60 minutes

It has been shown previously that treatment of *B. subtilis* cells with HPUra blocks DNA synthesis and triggers replication stress as RecA-GFP forms foci in most cells (52). We found that at a 40 nM concentration of HPUra, RecA-GFP formed foci in most cells as rapidly as we could prepare the sample for inspection by microscopy, as 76% of cells contained RecA-GFP foci at 140 seconds post treatment (n= 74). At five minutes post treatment, ~94% of cells had RecA-GFP foci (**Figure 20A and B**). We performed colocalization with RecA and DnaX and found that ~93% of replisome foci (n=181) were

colocalized with RecA following addition of HPUra to the growth medium (**Figure 20D**). These results show that damage-independent fork arrest causes RecA to localize to nearly all replisomes in *B. subtilis.* This work also clarifies the method for HPUra synthesis.

RecA-GFP focus formation is dependent on *recO* and *recR*.

Having established that RecA-GFP localizes to the replisome in response to various sources of damage, as well as damage-independent replication fork arrest, we asked which proteins are required for RecA loading *in vivo*. We predicted that *recO* might contribute to RecA-GFP focus formation as *B. subtilis* since it is sufficient to nucleate formation of RecA onto SSB coated ssDNA *in vitro* (29); however RecO was previously reported to only moderately affect RecA thread formation, and not foci formation, *in vivo* (30). In the absence of *recO*, RecA-GFP formed foci in less than 2.1% of cells untreated or following exogenous DNA damage (DSB and DSG repair) (**Table 5**). In addition, we found *recO* to be necessary for damage-independent RecA-GFP filament formation in response to fork arrest. Ectopic expression of *recO* complements the *recO* mutant, restoring RecA-GFP foci formation upon MMC and Phleomycin treatment (**Figure 41B:App**). These results could be explained by release of the GFP moiety via proteolytic cleavage from RecA in a *recO* mutant, however we found that RecA-GFP remained intact in the *recO::cat* background (**Figure 41A:App**). Together, these results show that regardless of the type of damage, RecO is necessary for assembly of RecA repair centers.

Table 5. RecOR are necessary for RecA loading *in vivo*.

	DNA Damage Assault				
Genotypes	**Endogenous Stresses/Damage**	**40 nM Mitomycin C**	**0.4 µM Phleomycin**	**40 µM HPUra**	**40 J m^{-2} UV**
recA23mGFPmut2	12.4% [2.6%]	31.9% [3.8%]	47.2% [3.5%]	94.1% [2.0%]	77.6% [4.9%]
ΔrecF	9.2% [2.1%]	7.4% [2.1%]	8.5% [2.1%]	8.2% [2.0%]	9.0% [2.3%]
ΔrecO	0.7% [0.7%]	1.5% [0.9%]	2.1% [1.3%]	<0.5%	0.6% [0.9%]
ΔrecR	2.5% [1.1%]	4.5% [1.4%]	0.7% [0.8%]	1.9% [1.2%]	0.3% [0.7%]
ΔrecJ, addA::erm	3.0% [1.0%]	25.6% [3.2%]	*10.6% [2.2%]	60.7% [3.6%]	44.9% [3.8%]

All backgrounds have the native locus *recA* fused to monomeric *GFPmut2* via a flexible 23 amino acid linker. All cultures were grown to mid-log before chemical/UV treatment. Phleomycin and Mitomycin-C were treated for 30 minutes while UV and HPUra treatment lasted 5 minutes. Upon the end of treatment, immediate imaging within a 2 minute window was undertaken. The bracketed number indicates the 95% confidence interval. * A separate lab stock of Phleomycin was used at a dose of 0.3 μM, which was

separately tested to verify equitoxic damaging dose (>LD_{90}).

Although RecO and RecR have not been shown to form a complex in *B. subtilis* as they do in *E. coli*, they have been shown to function in the same genetic pathway (20, 21, 29, 48). We performed the same experiment in a strain where the *recR* gene was replaced by a *cat* cassette and obtained nearly identical results to those obtained in *recO* deficient cells **(Table 5)**. All conditions examined in this background showed that RecA-GFP formed foci in 4.5% of cells or less. These results show that *recO* and *recR* are necessary for RecA to form foci in response to DNA damage and damage-independent fork arrest. This result was not due to proteolytic cleavage of RecA-GFP, and this phenotype was complemented by ectopic *recR* expression **(Figure 41B, Data Not Shown)**. We interpret these results to mean that the RecOR pathway is necessary for RecA loading *in vivo*.

RecF is the third component of the RecFOR pathway, which in *E. coli* is dedicated for the repair of DSGs *in vivo*. This damage is the product of a single-stranded gap following UV or mitomycin C exposure. To test the role of RecF in RecA-GFP repair center formation, we replaced the *recF* locus with a *cat* marker (*recF::cat*) and found that the percentage of cells with RecA-GFP foci was mostly unaffected in untreated cells lacking *recF* (wild type cells untreated: 12.4%±2.6. *recF::cat* untreated: 9.2%±2.1). However, we find that in *recF::cat* cells, RecA-GFP lost the ability to form foci in response to exogenously introduced DSBs and DSGs, as well as damage-independent replication fork arrest (**Table 5**). This illustrates that unlike *E. coli* RecF, *B. subtilis* RecF functions in all RecA loading responses and not just DSG repair.

RecF affects the efficiency of RecA focus assembly.

We were concerned that the defects we observed in RecA-GFP foci assembly in the *recOR* and *recF* deficient backgrounds might reflect poor efficiency of RecA loading at lower doses of DNA damage. Thus, we repeated the experiments with near saturating levels of DNA damage for an hour, ensuring most cells had a single RecA focus per cell (**Figure 22A and Figure 39:App**). At most, RecA generally forms 1 focus per nucleoid under extensive damaging conditions (33). Here, we found that *recO* and *recR* were still necessary for assembly of RecA-GFP foci, even with more time to

support filament formation **(Figure 22B and Figure 40:App)**. Interestingly, RecA-GFP foci were partially restored in the *recF::cat* strain **(Figure 22B)**. The foci were predominantly circular foci or short filaments, representative of early stages of RecA filament assembly, suggesting that RecF may be important for reducing the lag time of RecA-GFP nucleation *in vivo.* This result shows that RecF function is distinct from that of RecOR, which is most likely important for both initial RecA nucleation and filament growth, as suggested by *in vitro* experiments (7).

In total, we conclude that *recOR* is necessary for RecA-GFP focus formation in response to daughter strand gaps (UV, MMC), damage-independent fork arrest (HPUra), and DSBs (phleomycin). Furthermore, we conclude that RecF contributes to efficient RecA-GFP repair center formation for all types of damage, but is not required, for RecA-GFP repair center formation.

SSB C-terminus contributes to the DNA damage-induced assembly of RecA-GFP foci.

In both *B. subtilis* and *E. coli,* RecO has been identified as a binding partner of SSB (20, 53, 54). Because we show above that RecOR is necessary for RecA to assemble into foci, we asked if SSB contributes to this response. In *E. coli*, the SSB C-terminus mediates interaction with its binding partners, with the last two residues (PF), being required for these interactions [for review (54)]. Therefore, the PF residues are essential and mutation of just the proline (*ssb113*) causes temperature sensitive growth and UV sensitivity (55-57). Such a strong phenotype limits the ability to study *E. coli* SSB binding partners in an *in vivo* context. Interestingly, *ssb* alleles in *B. subtilis* encoding truncations of the C-terminal 6 or 35 amino acids are viable (53, 58), providing an experimental platform for understanding the effect of the SSB C-terminus on DNA replication and repair *in vivo* (53, 58). To this end, we integrated the *ssbΔ35* allele, which uses an IPTG regulated promoter to drive gene expression downstream and the corresponding wild type control (*ssb3*$^{+}$) into an isogenic strain background carrying the *recA-gfp* allele as described (53, 58). In the *ssb*3^{+} strain, RecA-GFP foci formed in ~15% of cells, which is similar to a native *ssb*$^{+}$ control [**Table 5** and (33, 38) (33, 38)]. Challenge with MMC resulted in RecA-GFP foci formation in ~35% of cells,

demonstrating damage-inducible assembly of RecA-GFP foci in the *ssb3*$^{+}$ control **(Figure 22C)**. In the *ssbΔ35* background, we observed a similar percentage of untreated cells with RecA-GFP foci (~20%) and following treatment with MMC (~22%). This indicates a failure in the damage-inducible response of RecA in the *ssbΔ35* strain **(Figure 4C and Figure 42:App)**. Despite the loss of RecO recruitment, we still observe elevated levels of RecA foci in both the treated and untreated group. In *E. coli*, loss of the C-terminal tails leads to a loss of ssDNA binding stability (59). We hypothesize that, likewise, the 35 residue truncation of the C-terminal tail of *B. subtilis* SSB may cause a decrease in binding stability to ssDNA, allowing spontaneous RecA nucleation at elevated rates (**See Discussion**). Overall, these results show that the DNA damage-induced localization of RecA is dependent on the SSB C-terminus *in vivo*.

Discussion

Identification of necessary recombinase mediator proteins in *Bacillus subtilis*.

In *B. subtilis*, the proteins required to elicit formation of RecA-GFP foci have remained unknown. We show that RecOR are necessary for the assembly of RecA-mediated DNA repair foci in live *B. subtilis* cells to both endogenous and exogenous sources of DNA damage, as well as HPUra, a drug that causes DNA damage-independent replication fork arrest (52). In *E. coli*, DSBs are initially processed by RecBCD to provide the 3´ ssDNA segment, followed by RecA loading, making most spontaneous RecA-GFP foci in rich medium *recB*-dependent (14, 18). In *B. subtilis*, AddAB provides a 3´ extension (26), but AddAB has not been shown to load RecA biochemically. Our results with phleomycin, which requires DNA end processing to produce DSBs, shows that RecOR is still necessary for assembly of RecA repair centers on AddAB processed ends. This observation makes it unlikely that AddAB can physically load RecA after generating a 3′ DNA segment. We suggest that SSB binds the 3´ extensions to protect the DNA end, followed by recruitment of RecOR through interaction with the SSB C-terminus **(Figure 22D)**.

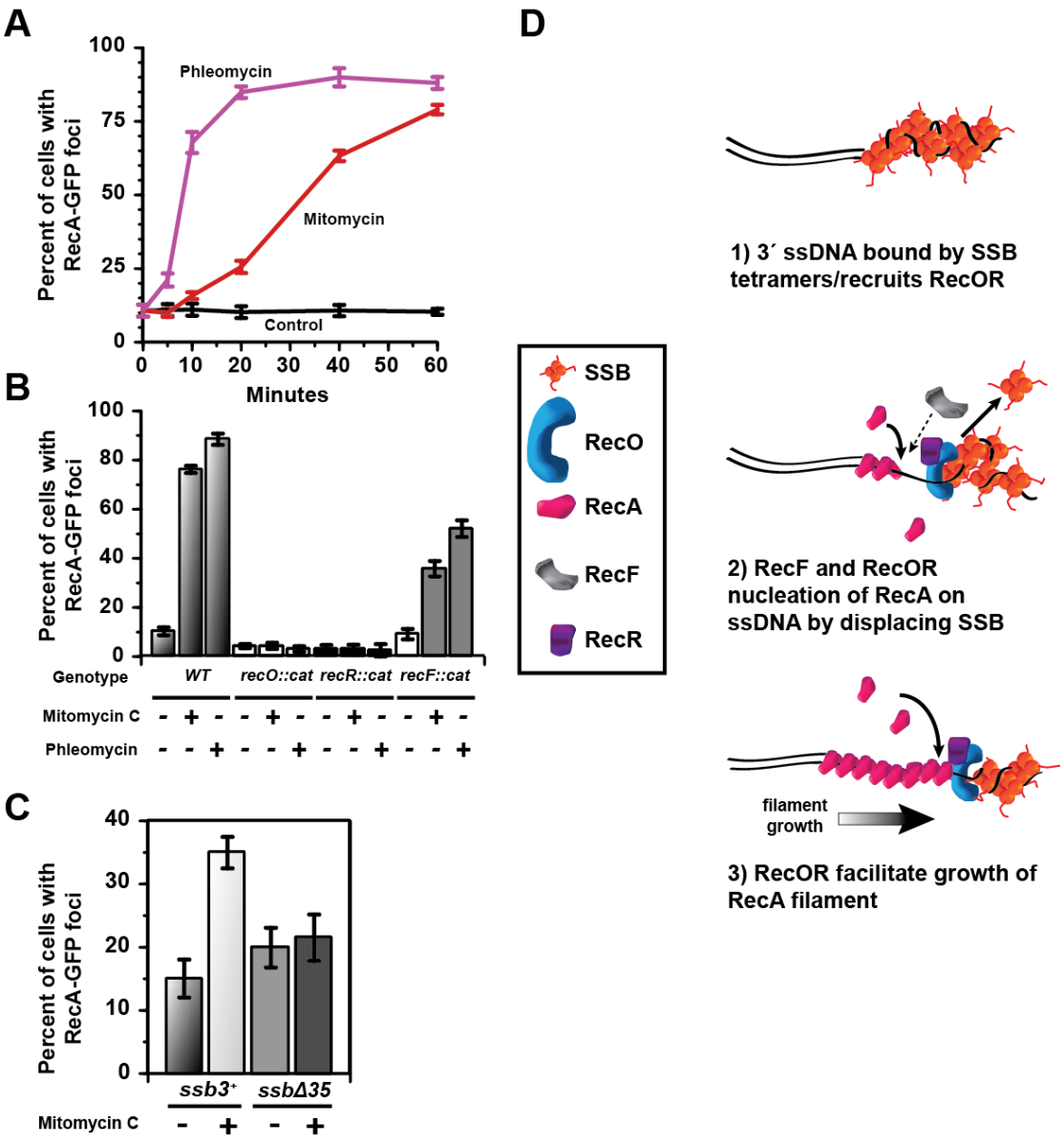


Figure 22. RecF and SSB are important for efficient assembly of RecA foci.
(A) Shown is a time course experiment representing the percentage of cells with RecA-GFP foci untreated (black), challenged with 100 ng/mL mitomycin C (red) or 3 μM phleomycin (purple) over 60 minutes. Each treatment was scored in triplicate at the time points indicated. Error bars reflect the 95% confidence interval. **(B)** Shown is a bar graph for the percentage of cells with RecA-GFP foci in the indicated backgrounds untreated or following challenge with 100 ng/mL mitomycin C or 3 μM phleomycin. **(C)** Bar graph quantifying the percentage of cells with RecA-GFP foci following treatment with MMC. The number of cells quantified are from left to right are 545, 1401, 622, and 479. The difference in the percentage of cells with foci is significant for *ssb3*$^{+}$ when comparing MMC treated to untreated groups with p=2.25X10^{-18}. The differences in the percentage of cells with RecA-GFP foci in the *ssbD35* strain are not significant when comparing MMC (p=0.26) to untreated using a one-tailed p-value. Error bars represent the 95% confidence interval. **(D)** A model for RecOR-dependent RecA loading and establishment of RecA repair centers in *B. subtilis* at the site of a DSB.

It has been published that GFP-RecA is partly unaffected for focus formation in *recO* or *recR* deficient cells, as cells deficient for RecO function still showed GFP-RecA foci in ~25% of cells challenged with 50 ng/mL MMC (30). The differences between the two studies are the RecA GFP fusions. We used a *recA-gfp* fusion located at the native locus as the sole source of RecA, whereas the other study expressed GFP-RecA ectopically from an inducible promoter (where the native locus *recA* was deleted). We suggest that ectopic expression of GFP-RecA may have increased the levels of RecA to a point where RecA was able to bypass the requirement for *recO* and *recR* for repair center assembly. *B. subtilis* RecA can self-assemble on SSB coated ssDNA *in vitro* although the reaction is accelerated with RecO (29). Taking this prior work into consideration with our study, we suggest that *recOR* is necessary for RecA to load *in vivo* when RecA is expressed at its native levels. Importantly, many more bacteria use AddAB for end processing of DSBs than RecBCD (25). Therefore, we speculate that RecOR may also facilitate RecA loading in other AddAB containing bacteria as well.

RecF has been placed with the *recOR* pathway by epistasis (28, 53). We therefore tested the effect of a *recF* null allele on the ability of RecA to assemble into foci. We found that RecF was important, but not necessary, for RecA to rapidly organize into filaments in response to DNA damage. Because we found at high concentrations of damage that RecA formed repair filaments consistent with early stages of filament formation, we hypothesize that RecF may be important for reducing the lag time of RecA-GFP nucleation *in vivo.* This result shows that RecF function is distinct from that of RecOR, which is most likely important for both initial RecA nucleation and filament growth.

Because we found a complete dependency on RecO for assembly of RecA repair centers, we tested a role for SSB since RecO is an SSB binding partner in both *E. coli* and *B. subtilis* (20, 53). SSB contains a C-terminal tail, which binds to and interacts with several proteins, recruiting them to sites of excess ssDNA (53, 54, 60, 61). In *B. subtilis,* it has been shown that RecO binds SSB, and that focus formation by ectopically expressed GFP-RecO relies on the SSB C-terminus (53). The data we present here shows that the SSB C-terminal 35 amino acids are important for the DNA damage-dependent localization of RecA. In *E. coli,* it has been shown that deletion of the SSB C-

terminal PF is lethal since these two residues mediate protein-protein interactions important for essential replication functions and repair (55, 56). Thus, how can *B. subtilis* tolerate truncation of the C-terminal 35 amino acids, including the ultimate PF motif? In contrast to *E. coli*, *B. subtilis* SSB contains three PF repeats in the C-terminal 38 amino acids and the Δ35 truncation would leave the sequence PFG on the C-terminus **(Figure 42:App)**. We suggest that the remaining PF motif remaining in the Δ35 truncation allows for viability and recruitment of a limited number of SSB binding partners. It was shown previously that *ssbΔ6* and *ssbΔ35* alleles are sensitive to DNA damage, including MMC, suggesting the need for all three PF motifs to facilitate repair (53). Our results here show that the *ssbΔ35* allele eliminates RecA-GFP assembly in a DNA damage-dependent manner. However, we do observe an elevated level of RecA-GFP foci in *ssbΔ35*. We suggest this may be due to a reduced level of SSB DNA binding, as purified SSB variants lacking the C-terminal tail fails to wrap the full amount of SSB at physiologic NaCl levels (55 nucleotides occluded in SSB variant vs. 65 nucleotides of WT SSB) (59). Since RecA binds a minimal dsDNA unit of 3 nucleotides, this disruption may be sufficient to support spontaneous nucleation, which can then support filament growth of RecA (8).

HPUra-dependent inhibition of DNA replication.

6-(*p*-Hydroxyphenylazo-uracil **(**HPUra) is a PolC specific inhibitor that has been widely used in *B. subtilis* and related Gram-positive bacteria to block DNA replication [e.g. (34, 45-50)]. HPUra has been a very important tool for investigating the involvement of DNA replication in a number of different DNA transactions *in vivo* [e.g. (34, 46, 47, 52)]. Prior work suggested that the hydrazine congener (H_2-HPUra) is the active form that inhibits DNA synthesis (45, 48-50). We were unable to synthesize H_2-HPUra and could only obtain the oxidized form (HPUra). We cannot exclude the possibility that HPUra is reduced *in vivo* to form H_2-HPUra, however, an HPUra tautomer (compound 3a) could form the proper hydrogen bonds to pair with dCMP *in vivo* in the absence of reduction to H_2-HPUra. With these results, we suggest that the penultimate compound (HPUra), which is readily synthesized as described [(48) and

Supplemental Information], is the appropriate compound for rapidly inhibiting DNA synthesis in *B. subtilis* and perhaps other Gram-positive bacteria.

We also show that HPUra is indeed both a potent inhibitor of DNA synthesis and triggers replication stress in live cells (Figures 1 and 2). We speculate that the cause of HPUra-induced replication stress is due to an uncoupling of the replicative polymerase PolC from the replicative helicase, DnaC. We suggest that a blockage to PolC movement and continued unwinding of the DNA by the helicase DnaC (different from *E. coli* DnaC) ahead of the blocked fork would form excess ssDNA, providing a substrate for RecA binding. Interestingly, it was shown that induction of the stringent response with arginine hydroxymate, which stops replication elongation via inhibition of primase activity, fails to form RecA-GFP filaments (52). We suggest that by directly inhibiting primase, the Primase-DnaC complex is maintained, causing DnaC to remain at the blocked fork and preventing formation of excess ssDNA and RecA-GFP recruitment.

Material and Methods

Bacteriological methods.

The bacterial plasmids and strains are listed in **Table 10 and 11** respectively. The following concentrations of compounds were used: 5 µg/mL chloramphenicol (*cat*), 100 µg/mL spectinomycin (*spc*), 12.5 µg/mL tetracycline (*tet*), 0.5 µg/mL erythromycin (*erm*), 100 ng/mL MMC, 3 µM phleomycin, 5 µM IPTG and 0.125 xylose for *recO* and *recR* (62, 63).

Live cell microscopy.

Microscopy of live cells was completed as described [(64-66) for review (67)]. Briefly, a starting culture was inoculated in 1X defined $S7_{50}$ minimal media supplemented with 2% glucose at a starting OD_{600} of 0.05 and allowed to grow for three doublings (66, 67). Cultures were split and one culture was challenged with the indicated damaging agent, while the other was left untreated as a control and grown for the indicated times. Following incubation, 200 µL aliquots of cells were stained with the vital membrane stain FM 4-64 or 10µM TMA-DPH and placed onto 1% agarose pads made with 1X spizizen salts. Cells were imaged with an Olympus BX61 microscope using an Olympus 100X

oil immersion 1.45 NA TIRFM objective lens (67). Imaging for each strain was performed independently a minimum of three times.

Compound synthesis.

All small molecule compounds used in this study were synthesized by the Vahlteich Medicinal Chemistry Core, University of Michigan. 6-(*p*-Hydroxyphenylazo)-uracil (**3**, HPUra) was synthesized by known methods (50) along with the novel silylated ether analogue **5** and the known analogue, 6-(phenylhydrazino)-uracil, **6** (50). Complete experimental details are provided in Supplemental Information. Stock solutions (20 mM) of assayed compounds were prepared in dimethyl sulfoxide (DMSO) or 50 mM KOH in the case of HPUra, stored in the dark, and kept for no more than 4 weeks. The compounds were added at 162 μM except when indicated otherwise. DMSO or KOH were used as controls and the same volume as used for the stock solutions was added.

Mitomycin C survival assay.

In general, sensitivity to DNA damage was determined as described (38, 68, 69). Briefly, strains were grown with the appropriate antibiotic overnight at 30°C. Single isolates were grown in 6 mL of LB in the dark to a final OD_{600} of 0.4-0.5. Five 1 mL aliquots of culture were centrifuged for 2 minutes at 10,000 rpm and the supernatants were removed. Cells were resuspended in 1 mL of 0.85% saline containing various MMC concentrations from 0 to 40 nM in the dark. After 30 minutes, cells were pelleted, damaging agents were aspirated, and cells were resuspended in 100 μL 0.85% saline. The suspension was serial diluted, followed by plating on LB agar plates and grown overnight at 30 °C. Percent survival was calculated by taking the mean survival of an indicated treatment relative to the mean survival of untreated cultures.

Replication inhibition assays.

Cultures were grown in $S7_{50}$ minimal medium at 30°C to an OD_{600} of 0.35. Using the *dnaB134* allele, replication initiation was inhibited while ongoing replication was completed by shifting cultures to 45°C for one hour. Upon downshift to 30°C, aliquots of culture were removed at specific time points and added to an equal volume of ice-cold

methanol prior to centrifugation and genomic DNA purification. When used, HPUra, RJS-7-061 or RJS-7-070 were added to cultures at a final concentration of 162 μM 20 minutes after temperature downshift with cells harvested at 60 minutes. Genomic DNA was sequenced on the Illumina HiSeq-2000 platform, generating 50 base single-end reads. Library preparation and sequencing was performed by the University of Michigan DNA Sequencing Core. Sequence data were aligned to the PY79 reference genome (Accession number CP006881, (70)) using bwa aln with default parameters followed by bwa samse with the "-n" parameter set to 1 (71). For coverage calculation, the genome was split into 1000 nucleotide wide windows and coverage of each window was normalized per million reads mapped to the entire genome. Fold-enrichment for each sample was calculated by dividing normalized coverage for each window by the normalized coverage of the same window for genomic DNA harvested immediately after temperature downshift. Resulting enrichment values were then offset to achieve equal baseline between samples.

Acknowledgements

We would like to thank members of the Simmons lab for their critical reading of this manuscript. We thank Dr. C. Lee and Dr. Alan D. Grossman (Massachusetts Institute of Technology) for the gift of a *recA* overexpression plasmid. This work was supported by a grant from the National Science Foundation (MCB1050948) to L.A.S. In addition, E.R.B. was supported in part by an undergraduate summer fellowship from the Honors College at the University of Michigan. J.S.L. was supported by a predoctoral fellowship from the Rackham Graduate School at the University of Michigan. J.W.S. was supported in part by the NIH Genetics Training Grant (T32GM007544). H.D.S. R.J.S. acknowledges generous support by the University of Michigan College of Pharmacy Ella and Hans Vahlteich Research Fund.

SUPPLEMENTAL MATERIAL

The Supplemental Material found in Appendix III includes a description of the chemistry and biological methods, **Table 10 and 11**, and **Figures 38-42**.

References:

1. Friedberg EC*, et al.* (2006) *DNA Repair and Mutagenesis: Second Edition* (American Society for Microbiology, Washington, DC).
2. Simmons LA, Foti JJ, Cohen SE, & Walker GC (2008) The SOS Regulatory Network. in *EcoSal-Escherichia coli and Salmonella: cellular and molecular biology*, eds Böck A, Curtiss III R, Kaper JB, Karp PD, Neidhardt FC, Nyström T, Slauch JM, & C.L. S (ASM Press, Washington, D.C.).
3. Cox MM (2007) Regulation of bacterial RecA protein function. *Crit Rev Biochem Mol Biol* 42(1):41-63.
4. Little JW & Mount DW (1982) The SOS regulatory system of *Escherichia coli*. *Cell* 29:11-22.
5. Kowalczykowski SC & Krupp RA (1987) Effects of Escherichia coli SSB protein on the single-stranded DNA-dependent ATPase activity of Escherichia coli RecA protein. Evidence that SSB protein facilitates the binding of RecA protein to regions of secondary structure within single-stranded DNA. *Journal of molecular biology* 193(1):97-113.
6. Umezu K, Chi N-W, & Kolodner RD (1993) Biochemical interactions of the *Escherichia coli* RecF, RecO, and RecR proteins with RecA protein and single-stranded binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 90:3875-3879.
7. Bell JC, Plank JL, Dombrowski CC, & Kowalczykowski SC (2012) Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA. *Nature* 491(7423):274-278.
8. Joo C*, et al.* (2006) Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* 126(3):515-527.
9. Clark AJ & Sandler SJ (1994) Homologous genetic recombination: the pieces begin to fall into place. *Crit. Rev. Microbiol.* 20(2):125-142.
10. Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, & Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* 58(3):401-465.
11. Lenhart JS, Schroeder JW, Walsh BW, & Simmons LA (2012) DNA Repair and Genome Maintenance in Bacillus subtilis. *Microbiol Mol Biol Rev* 76(3):530-564.
12. Capaldo-Kimball F & Barbour SD (1971) Involvement of recombination genes in growth and viability of Escherichia coli K-12. *Journal of bacteriology* 106(1):204-212.
13. Emmerson PT (1968) Recombination deficient mutants of Escherichia coli K12 that map between thy A and argA. *Genetics* 60(1):19-30.
14. Renzette N*, et al.* (2005) Localization of RecA in Escherichia coli K-12 using RecA-GFP. *Molecular microbiology* 57(4):1074-1085.

15. Centore RC, Lestini R, & Sandler SJ (2008) XthA (Exonuclease III) regulates loading of RecA onto DNA substrates in log phase Escherichia coli cells. *Molecular microbiology* 67(1):88-101.

16. Dillingham MS & Kowalczykowski SC (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiology and molecular biology reviews : MMBR* 72(4):642-671, Table of Contents.

17. Spies M, Amitani I, Baskin RJ, & Kowalczykowski SC (2007) RecBCD enzyme switches lead motor subunits in response to chi recognition. *Cell* 131(4):694-705.

18. Arnold DA & Kowalczykowski SC (2000) Facilitated loading of RecA protein is essential to recombination by RecBCD enzyme. *J. Biol. Chem.* 275(16):12261-12265.

19. Spies M & Kowalczykowski SC (2006) The RecA binding locus of RecBCD is a general domain for recruitment of DNA strand exchange proteins. *Molecular cell* 21(4):573-580.

20. Umezu K & Kolodner RD (1994) Protein interactions in genetic recombination in *Escherichia coli*. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J. Biol. Chem.* 269(47):30005-30013.

21. Morimatsu K & Kowalczykowski SC (2003) RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Molecular cell* 11(5):1337-1347.

22. Morimatsu K, Wu Y, & Kowalczykowski SC (2012) RecFOR proteins target RecA protein to a DNA gap with either DNA or RNA at the 5' terminus: implication for repair of stalled replication forks. *The Journal of Biological Chemistry* 287(42):35621-35630.

23. Ryzhikov M, Koroleva O, Postnov D, Tran A, & Korolev S (2011) Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein. *Nucleic Acids Res* 39(14):6305-6314.

24. West SC (1992) Enzymes and molecular mechanisms of genetic recombination. *Annu Rev Biochem* 61:603-640.

25. Cromie GA (2009) Phylogenetic ubiquity and shuffling of the bacterial RecBCD and AddAB recombination complexes. *Journal of bacteriology* 191(16):5076-5084.

26. Chedin F, Ehrlich SD, & Kowalczykowski SC (2000) The Bacillus subtilis AddAB helicase/nuclease is regulated by its cognate Chi sequence in vitro. *Journal of molecular biology* 298(1):7-20.

27. Saikrishnan K*, et al.* (2012) Insights into Chi recognition from the structure of an AddAB-type helicase-nuclease complex. *The EMBO journal*.

28. Fernandez S, Kobayashi Y, Ogasawara N, & Alonso JC (1999) Analysis of the Bacillus subtilis recO gene: RecO forms part of the RecFLOR function. *Mol Gen Genet* 261(3):567-573.

29. Manfredi C, Carrasco B, Ayora S, & Alonso JC (2008) Bacillus subtilis RecO nucleates RecA onto SsbA-coated single-stranded DNA. *J Biol Chem* 283(36):24837-24847.

30. Kidane D & Graumann PL (2005) Dynamic formation of RecA filaments at DNA double strand break repair centers in live cells. *J Cell Biol* 170(3):357-366.

31. Sanchez H, Kidane D, Castillo Cozar M, Graumann PL, & Alonso JC (2006) Recruitment of Bacillus subtilis RecN to DNA double-strand breaks in the absence of DNA end processing. *Journal of bacteriology* 188(2):353-360.

32. Renzette N, Gumlaw N, & Sandler SJ (2007) DinI and RecX modulate RecA-DNA structures in Escherichia coli K-12. *Molecular microbiology* 63(1):103-115.

33. Simmons LA, Grossman AD, & Walker GC (2007) Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* 104(4):1360-1365.

34. Bernard R, Marquis KA, & Rudner DZ (2010) Nucleoid occlusion prevents cell division during replication fork arrest in Bacillus subtilis. *Molecular microbiology* 78(4):866-882.

35. Renzette N & Sandler SJ (2008) Requirements for ATP binding and hydrolysis in RecA function in Escherichia coli. *Molecular microbiology* 67(6):1347-1359.

36. Meile JC, Wu LJ, Ehrlich SD, Errington J, & Noirot P (2006) Systematic localisation of proteins fused to the green fluorescent protein in Bacillus subtilis: identification of new proteins at the DNA replication factory. *Proteomics* 6(7):2135-2146.

37. Lesterlin C, Ball G, Schermelleh L, & Sherratt DJ (2013) RecA bundles mediate homology pairing between distant sisters during DNA break repair. *Nature* 10.1038/nature12868.

38. Simmons LA*, et al.* (2009) Comparison of responses to double-strand breaks between Escherichia coli and Bacillus subtilis reveals different requirements for SOS induction. *Journal of bacteriology* 191(4):1152-1161.

39. Kidane D, Sanchez H, Alonso JC, & Graumann PL (2004) Visualization of DNA double-strand break repair in live bacteria reveals dynamic recruitment of Bacillus subtilis RecF, RecO and RecN proteins to distinct sites on the nucleoids. *Molecular microbiology* 52(6):1627-1639.

40. Dronkert ML & Kanaar R (2001) Repair of DNA interstrand cross-links. *Mutat Res* 486(4):217-247.

41. Sleigh MJ (1976) The mechanism of DNA breakage by phleomycin in vitro. *Nucleic Acids Res* 3(4):891-901.

42. Courcelle CT, Chow KH, Casey A, & Courcelle J (2006) Nascent DNA processing by RecJ favors lesion repair over translesion synthesis at arrested replication forks in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* 103(24):9154-9159.

43. Sassanfar M & Roberts JW (1990) Nature of the SOS-inducing signal in Escherichia coli. The involvement of DNA replication. *Journal of molecular biology* 212(1):79-96.

44. Sedgwick SG (1975) Genetic and kinetic evidence for different types of postreplication repair in Escherichia coli B. *Journal of bacteriology* 123(1):154-161.

45. Gass KB, Low RL, & Cozzarelli NR (1973) Inhibition of a DNA polymerase from Bacillus subtilis by hydroxyphenylazopyrimidines. *Proceedings of the National Academy of Sciences of the United States of America* 70(1):103-107.

46. Goranov AI, Katz L, Breier AM, Burge CB, & Grossman AD (2005) A transcriptional response to replication status mediated by the conserved bacterial replication protein DnaA. *Proceedings of the National Academy of Sciences of the United States of America* 102(36):12932-12937.

47. Goranov AI, Kuester-Schoeck E, Wang JD, & Grossman AD (2006) Characterization of the global transcriptional responses to different types of DNA damage and disruption of replication in Bacillus subtilis. *Journal of bacteriology* 188(15):5595-5605.

48. Mackenzie JM, Neville MM, Wright GE, & Brown NC (1973) Hydroxyphenylazopyrimidines: characterization of the active forms and their inhibitory action on a DNA polymerase from Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* 70(2):512-516.

49. Neville MM & Brown NC (1972) Inhibition of a discrete bacterial DNA polymerase by 6-(p-hydroxyphenylazo)-uracil and 6-(p-hydroxyphenylazo-)-isocytosine. *Nat New Biol* 240(98):80-82.

50. Wright GE & Brown NC (1974) Synthesis of 6-(phenylhydrazino)uracils and their inhibition of a replication-specific deoxyribonucleic acid polymerase. *J Med Chem* 17(12):1277-1282.

51. Brown NC (1970) 6-(p-hydroxyphenylazo)-uracil: a selective inhibitor of host DNA replication in phage-infected Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* 67(3):1454-1461.

52. Wang JD, Sanders GM, & Grossman AD (2007) Nutritional control of elongation of DNA replication by (p)ppGpp. *Cell* 128(5):865-875.

53. Costes A, Lecointe F, McGovern S, Quevillon-Cheruel S, & Polard P (2010) The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS genetics* 6(12):e1001238.

54. Shereda RD, Kozlov AG, Lohman TM, Cox MM, & Keck JL (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* 43(5):289-318.

55. Chase JW, L'Italien JJ, Murphy JB, Spicer EK, & Williams KR (1984) Characterization of the Escherichia coli SSB-113 mutant single-stranded DNA-binding protein. Cloning of the gene, DNA and protein sequence analysis, high

pressure liquid chromatography peptide mapping, and DNA-binding studies. *The Journal of Biological Chemistry* 259(2):805-814.

56. Curth U, Genschel J, Urbanke C, & Greipel J (1996) In vitro and in vivo function of the C-terminus of Escherichia coli single-stranded DNA binding protein. *Nucleic Acids Research* 24(14):2706-2711.

57. Johnson BF (1977) Genetic mapping of the lexC-113 mutation. *Molecular & general genetics : MGG* 157(1):91-97.

58. Lecointe F*, et al.* (2007) Anticipating chromosomal replication fork arrest: SSB targets repair DNA helicases to active forks. *The EMBO journal* 26(19):4239-4251.

59. Antony E*, et al.* (2013) Multiple C-terminal tails within a single E. coli SSB homotetramer coordinate DNA replication and repair. *Journal of molecular biology* 425(23):4802-4819.

60. Lu D & Keck JL (2008) Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proceedings of the National Academy of Sciences of the United States of America* 105(27):9169-9174.

61. Shereda RD, Bernstein DA, & Keck JL (2007) A central role for SSB in Escherichia coli RecQ DNA helicase function. *J Biol Chem* 282(26):19247-19258.

62. Hardwood CR & Cutting SM (1990) *Molecular Biological Methods for Bacillus* (John Wiley & Sons, Chichester).

63. Smith BT, Grossman AD, & Walker GC (2001) Visualization of mismatch repair in bacterial cells. *Mol. Cell* 8(6):1197-1206.

64. Lemon KP & Grossman AD (2000) Movement of replicating DNA through a stationary replisome. *Mol. Cell* 6(6):1321-1330.

65. Dupes NM*, et al.* (2010) Mutations in the Bacillus subtilis beta clamp that separate its roles in DNA replication from mismatch repair. *Journal of bacteriology* 192(13):3452-3463.

66. Klocko AD*, et al.* (2011) Mismatch repair causes the dynamic release of an essential DNA polymerase from the replication fork. *Molecular microbiology* 82(3):648-663.

67. Klocko AD, Crafton KM, Walsh BW, Lenhart JS, & Simmons LA (2010) Imaging mismatch repair and cellular responses to DNA damage in Bacillus subtilis. *J Vis Exp* 36(36):1-4.

68. Davies BW*, et al.* (2011) DNA damage and reactive nitrogen species are barriers to Vibrio cholerae colonization of the infant mouse intestine. *PLoS Pathog* 7(2):e1001295.

69. Kobayashi H, Simmons LA, Yuan DS, Broughton WJ, & Walker GC (2008) Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-

living form and during chronic infection of Sinorhizobium meliloti. *Molecular microbiology* 67(2):350-363.

70. Schroeder JW & Simmons LA (2013) Complete Genome Sequence of Bacillus subtilis Strain PY79. *Genome Announc* 1(6).

71. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754-1760.

Chapter V

Summary and Perspectives

JSL composed this chapter.

Cytological studies in *Bacillus subtilis* have identified the progressing replication fork as a target site for the initiation of mismatch repair (MMR), daughter strand gap repair, and homologous recombination (1-3). Over the last few years, significant progress has been made towards understanding how the replication fork's architecture facilitates the identification of various forms of DNA damage and base pairing errors: helix distorting IDLs (insertion/deletion loops), non-helix distorting Watson-Crick mismatches, and various forms of DNA breaks that leads to replication fork arrest/collapse. Hence, replication coupling appears to serve as the first step in a choreographed response dedicated to the rapid and efficient repair of various lesions.

In mismatch repair, error detection by MutS is physically tied to the DnaN processivity clamp (1, 2). After directing processive replication by the replicative polymerases, DnaN continues to encircle DNA, passively diffusing along its length. Numerous rounds of clamp release from recycling polymerases on the leading and lagging strands producing an enriched zone of DNA-bound DnaN clamps, forming an ideal scaffolding to direct mismatch detection by MutS to nascent DNA (4, 5). This coupling ensures mismatch detection on replication fork-proximal DNA, which is cleared of inhibitory protein inclusions, as well as adjacent to the ongoing replication forks. Additional downstream steps, such as strand discrimination and MutL-mediated endonuclease cleavage, further require an intimate connection with the replication forks, possibly via the DnaN-clamp zone itself.

To date, the key unresolved question in the MMR field is what comprises the

mechanism of strand discrimination. A mismatch itself possesses no information implicating one of its bases as originating from the parental (template) strand. Therefore, upon mismatch detection, the MMR apparatus must determine from outside the mismatch which strand is the parental strand containing the correct genetic information, and subsequently direct repair towards the daughter strand. This entails excision of the DNA surrounding the mismatch on the newly replicated strand. Studies to date have offered several clues as to how strand discrimination is accomplished: 1) MMR occurs preferentially on the lagging strand, 2) a single-strand nick directs strand excision to the mismatched base topologically contiguous to the nick, 3) MutL and MutLα have clamp binding motifs located adjacent to the endonuclease active site, 4) PCNA and the RFC clamp loader complex enhances MutLα endonuclease activity, and 5) PCNA loading onto a covalently closed circular DNA substrate containing a mismatch is sufficient to license MutLα endonuclease activity, although with absolute loss of strand preference(6-13). These clues suggest that processivity clamps, via licensing endonuclease activity, function as the strand determinant during repair.

Strand breaks function as loading sites of processivity clamps, but once loaded, how could they relay information to the MMR machinery? Interestingly, the monomers of the clamp face are non-equivalent, as clamps appear tilted at various angles when loaded onto DNA. A computational modeling study of PCNA on DNA predicted that a ~20° tilt formed to optimize basic amino acid contacts of the PCNA inner ring to the negatively charged sugar-phosphate backbone on DNA (14). Additionally, a crystal structure of the *E. coli* clamp bound to DNA found the clamp skewed from the perpendicular axis by 22° (~20-40˚ for PCNA)(15). Yet further studies confirmed the presence of a tilt. A DNA-bound PCNA co-crystal structure was elucidated (albeit at low density) with a ~40° tilt (16). Finally, a single-molecule electron microscopy study of *Pyrococcus furiosus* PCNA also found a tilt on PCNA (17). Despite this evidence, it remains speculative that a tilt exists, yet if present on DNA loaded clamps, may reflect the loading orientation onto the primer termini by the clamp loader, thereby providing intrinsic assignment of parental and daughter strands. Mechanistically, the tilt of the clamp may present only one monomer to engage the MutL endonuclease via a clamp-binding motif, licensing strand nicking to only one strand: the newly replicated daughter.

Importantly, the PCNA clamp has been shown capable of propagating strand polarity information, as it tracks along the rotational pitch of the double helix (corkscrew-like motion), and therefore, its loading orientation would persist as it diffuses down the DNA(18). Although current data predisposes us to consider processivity clamps as the carrier of information begetting strand discrimination, our notions of this role remain highly speculative.

When *B. subtilis* cells are exposed to a sufficient large dose of ionizing radiation, potentially lethal double-strand breaks (DSBs) are formed. Cytological studies surprisingly found that these cells fail to produce RecA-GFP filaments in the absence of concurrent replication, despite the presence of potentially lethal lesions (3). This observation was extended to various forms of DNA damage: endonuclease-mediated DSBs, Mitomycin C induced inter- and intra-strand crosslinks, and UV-induced pyrimidine dimers. When replication is initiated, RecA filaments are generated that further localize to the vicinity of the replisome, arguing that the replication forks license or couple RecA loading to lesions encountered. This indeed appears to be the case, as Rec(F)OR facilitates RecA loading via association with the C-terminal tail of SSB. Importantly, other recombination/replication restart proteins physically bind SSB via its C-terminal tail: RecJ, RecS, RecQ, YrrC, RecG, SbcC, PriA, among others, suggesting an intimate coupling of recombination repair and replication restart with the replication fork and SSB (19, 20).

Lesions that facilitate replication fork collapse are intrinsically tied to replication as the source of ssDNA, the intimacy between DSBs and replication remains surprising and raises an additional question: Is there a spatial and temporal relationship between DNA end-processing and DNA replication? As in *Escherichia coli*, *B. subtilis* has two parallel pathways for end-processing: the RecBCD-like AddAB complex, as well as the RecQ helicases (RecQ and RecS) and the RecJ 5′-3′ exonuclease (21). All components of the second pathway associate with the C-terminus of SSB, and when fused to GFP and overexpressed, constitutively colocalize with the replisome (19). This data suggests a coupling between end-processing using the RecJQS pathway, with an application of for processing or repairing of regressed replication forks. Yet neither AddA nor AddB has been shown to bind SSB. Additionally, deletion of either AddA or RecJ fails to

produce a significant increase in sensitivity to MMC, suggesting redundancy for end-processing machinery in *B. subtilis*. AddAB is a highly processive helicase/nuclease, as it is able unwind and degrade dsDNA to the next χ site. If RecA loading occurs strictly at the replication fork, it seems logical to suspect that rapid and extensive strand degradation would occur at the replication fork, with the added benefit to the cell of a readily available homologous locus on the newly replicated sister chromosome. Yet, no evidence suggests a coupling exists of the AddAB associating with the replication fork or its protein scaffolds. How the cell balances end processing, accompanied with extensive stand degradation, and loss of genetic information with repair to ensure viability remains unclear.

References

1. Lenhart JS, Sharma A, Hingorani MM, & Simmons LA (2013) DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular microbiology* 87(3):553-568.
2. Simmons LA, Davies BW, Grossman AD, & Walker GC (2008) Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Molecular cell* 29(3):291-301.
3. Simmons LA, Grossman AD, & Walker GC (2007) Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* 104(4):1360-1365.
4. Shibahara K & Stillman B (1999) Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 96(4):575-585.
5. Su'etsugu M & Errington J (2011) The replicase sliding clamp dynamically accumulates behind progressing replication forks in Bacillus subtilis cells. *Molecular cell* 41(6):720-732.
6. Pavlov YI, Mian IM, & Kunkel TA (2003) Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol* 13(9):744-748.
7. Pillon MC*, et al.* (2010) Structure of the endonuclease domain of MutL: unlicensed to cut. *Molecular cell* 39(1):145-151.
8. Gueneau E*, et al.* (2013) Structure of the MutLalpha C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. *Nat Struct Mol Biol* 20(4):461-468.
9. Kadyrov FA, Dzantiev L, Constantin N, & Modrich P (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* 126(2):297-308.

10. Pluciennik A*, et al.* (2010) PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America* 107(37):16066-16071.

11. Thomas DC, Roberts JD, & Kunkel TA (1991) Heteroduplex repair in extracts of human HeLa cells. *J Biol Chem* 266(6):3744-3751.

12. Dzantiev L*, et al.* (2004) A defined human system that supports bidirectional mismatch-provoked excision. *Molecular cell* 15(1):31-41.

13. Fang WH & Modrich P (1993) Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. *J Biol Chem* 268(16):11838-11844.

14. Ivanov I, Chapados BR, McCammon JA, & Tainer JA (2006) Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications. *Nucleic Acids Res* 34(20):6023-6033.

15. Georgescu RE*, et al.* (2008) Structure of a sliding clamp on DNA. *Cell* 132(1):43-54.

16. McNally R, Bowman GD, Goedken ER, O'Donnell M, & Kuriyan J (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC structural biology* 10:3.

17. Mayanagi K*, et al.* (2011) Architecture of the DNA polymerase B-proliferating cell nuclear antigen (PCNA)-DNA ternary complex. *Proceedings of the National Academy of Sciences of the United States of America* 108(5):1845-1849.

18. Kochaniak AB*, et al.* (2009) Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J Biol Chem* 284(26):17700-17710.

19. Costes A, Lecointe F, McGovern S, Quevillon-Cheruel S, & Polard P (2010) The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS genetics* 6(12):e1001238.

20. Lecointe F*, et al.* (2007) Anticipating chromosomal replication fork arrest: SSB targets repair DNA helicases to active forks. *The EMBO journal* 26(19):4239-4251.

21. Lenhart JS, Schroeder JW, Walsh BW, & Simmons LA (2012) DNA repair and genome maintenance in Bacillus subtilis. *Microbiology and molecular biology reviews : MMBR* 76(3):530-564.

Appendix I
DnaN clamp zone provides a platform for spatiotemporal coupling of mismatch detection to DNA replication

Justin S. Lenhart[1], Anushi Sharma[2], Manju M. Hingorani[2], and Lyle A. Simmons[1*],

[1]Molecular, Cellular, and Developmental Biology, University of Michigan. Ann Arbor, MI

[2]Department of Molecular Biology and Biochemistry, Wesleyan University. Middletown, CT

Author contributions: All work was performed by JSL. JSL and LAS designed the research. All authors contributed to the writing of this appendix.

Appendix is published: Lenhart, J. S., Sharma, A., Hingorani, M. M. & Simmons, L. A. DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular microbiology* **87**, 553-568 (2013).

Supplemental Results

Mismatch detection by MutS is necessary for MutL-GFP focus formation. To further test the effects of the *mutSF30A* allele on MMR, we fused *mutL* to *gfp* to determine if the MutSF30A complex, which forms foci independent of mismatch binding, could elicit the downstream step of MutL-GFP focus formation. When native *mutS* is upstream of *mutL-gfp*, we found that ~16% of cells form MutL-GFP foci untreated (Fig S2). Challenge of MutL-GFP cells with 2-AP or placement into a genetic background containing a proofreading deficient *polC*$^{exo-}$ allele, we found an increase in the percent of cells with foci to ~36% and ~33% of cells respectively (Fig S2). This result shows that mismatch detection by MutS stimulates MutL to form foci *in vivo*. In contrast, a strain with a clean deletion of the *mutS* coding region showed a striking reduction in the percent of cells with MutL-GFP foci (<3% of cells) and in the absence of mutS, MutL-GFP foci were not stimulated by treatment with 2-AP (Fig S2). We asked if mismatch binding by MutS was required to elicit MutL-GFP foci. We show that in a *mutSF30A* background, MutL-GFP foci formed in ~11% of cells. These foci were qualitatively dim relative to MutL-GFP foci that form in a *mutS*$^{+}$ strain. Furthermore, we found that in the *mutSF30A* strain, 2-AP failed to stimulate an increase in MutL-GFP focus formation when compared with wild type *mutS*. With these results we conclude that MutL-GFP focus formation is primarily driven by the formation of a binary complex between MutS and a mismatch, since *mutSF30A* is substantially reduced for MutL-GFP recruitment and unresponsive to 2-AP treatment (Fig S2).

MutL-GFP colocalizes with the replication apparatus. We also characterized the ability of MutL-mCherry to colocalize with the replisome. We constructed a strain possessing both *dnaX-gfp* and *mutL-mCherry* alleles, where both were integrated at their native locus and expressed from their native promoters. In growing cells, MutL-mCherry colocalized with the replisome at levels comparable to MutS (~59%, n=130) (Fig 3D). This is different from results obtained in *S. cerevisiae*, where the Mlh1-PMS1 foci are not coincident with replication centers (1-3). It should be noted however, that MutL foci in both organisms are dependent on MutS. This suggests that in *B. subtilis*, MutL may also be positioned at the replisome waiting for MutS to detect a mismatch. When challenged with 2-AP, MutL-mCherry was reduced for colocalized with the

replisome to ~49% (p=0.0568). Because MutL-GFP foci behave similarly to MutS-GFP foci when treated with 2-AP, we speculate that MutL may be poised with MutS at the replisome waiting for mismatch identification. We further speculate that upon the initiation of repair, MutL associates with the repair complex and moves away from the replisome in a complex that can be observed by fluorescence microscopy.

Materials and Methods

MutS and MutSF30A protein purification

MutS and MutSF30A were overexpressed from pET11t-*mutS* and -*mutSF30A* respectively in *E. coli* $BL21_{DE3}$ using standard procedures (4). Cell pellets were resuspended in resuspension buffer (20 mM Tris HCl (pH 7.6), 500 mM NaCl, 4 mM DTT, 20 mM spermidine trihydrochloride, and 0.5 mM EDTA) and lysed by passage through a French pressure cell (3 passes at a setting of 15,000 PSI). Cells debris was pelleted at 15,000 rpm for 30 min at 4˚C. All subsequent protein purification steps were performed on ice or at 4˚C. Solid ammonium sulfate was added in order to achieve 30% saturation of the supernatant. Precipitated protein was removed by centrifugation at 15,000 rpm for 30 minutes. The supernatant was collected and ammonium sulfate was added to 40% saturation, which preferentially precipitated MutS and MutSF30A. Precipitated protein was removed via centrifugation at 15,000 rpm for 30 min and flash frozen in liquid N_2 and stored at -80˚C.

The pellet was thawed on ice and then resuspended in buffer QA (20 mM Tris HCl (pH 7.6), 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, and 10% glycerol). Once the pellet was resuspended, the sample was desalted using HiPrep 26/10 desalting column (GE Healthcare, Uppsala, Sweden) pre-equilibrated in buffer QA. This material was further purified with a HiTrap Q HP column (GE Healthcare, Uppsala, Sweden) pre-equilibrated in Buffer QA. MutS and MutSF30A were eluted with a 20 column volume gradient (from 100 mM NaCl to 600 mM NaCl). MutS and MutSF30A containing fractions were pooled and once again desalted with a HiPrep 26/10 desalting column pre-equilibrated in Buffer HA1 (5 mM potassium phosphate (pH 7.4), 30 mM NaCl, 1mM DTT, and 10% glycerol). The desalted sample was purified with a hydroxyapatite Bio-Scale Mini CHT Type 1 column (Bio-Rad) and eluted with a phosphate gradient (5 mM-400 mM) over 20 column

volumes. Protein samples were concentrated with a Vivaspin 20 concentrator (GE Healthcare-28-9323-62) with a 50 kDa molecular weight cutoff and then desalted with a HiPrep 26/10 column following using the following storage buffer (20 mM Tris HCl (pH 7.6), 200 mM NaCl, 4 mM DTT, and 10% glycerol). Samples were again concentrated with a Vivaspin 20 concentrator and aliquoted into small usable samples and flash frozen in liquid N_2. MutS and MutSF30A behaved identically throughout the purification process, indicating similar biochemical characteristics. Absorbance spectra were obtained using a 50-Bio UV Spectrophotometer (Varian, Palo Alto, CA) in both the natured state (storage buffer) and denatured state (storage buffer with 6M Guanidine HCl). No light scattering was observed from 300-600 nm in the native conditions, indicating no aggregation within the sample. In addition, no significant absorbance was detected at 250 nm indicating no nucleotides present. All protein concentrations were determined using extinction coefficients derived at ExPASy Proteomics Server (http://expasy.org/). DnaN was purified as described previously (5).

Construction of an unmarked-in frame *mutSF30A* mutant allele

In order to introduce the *mutSF30A* allele into *B. subtilis*, we used the pMiniMAD2 vector as described (9). pMiniMAD2 contains an erythromycin cassette, as well as a temperature sensitive origin of replication. Plasmid pJSL44 was introduced into PY79 via Campbell-type integration at the restrictive temperature (37˚C) for replication, selecting for integration of the plasmid into the chromosome by *mls* selection. After the initial Campbell type integration, pJSL44 was evicted as described (9). To rid the strain of the integrated plasmid, transformed colonies were grown in 3 ml of LB broth at a permissive temperature for plasmid replication (22°C) for 14 h, diluted 30-fold in fresh LB broth and incubated at 22°C for another 8 h. This process was continued over three successive days, and then serial dilutions were plated on LB agar at 37°C. Colonies were then colony purified on LB at the restrictive temperature and tested for *mls* resistance. Strains that were sensitive to *mls* selection were screened for increased spontaneous mutagenesis, consistent with a mismatch repair defect. Colonies with a high increase in mutagenesis were subject to sequencing of the *mutS* genetic locus to verify the presence of the *F30A* change and confirm that no other nucleotide changes were present.

Construction of an unmarked-in frame *ΔmutS* mutant allele

The *ΔmutS* allele was built in a similar way to the unmarked-in frame *mutSF30A* mutant allele. Differences are that a 500 bp region upstream of *mutS* was amplified with primers JSL 156/157 (insert A) and a 500 bp region downstream of *mutS* was amplified with pJSL 158/159 (insert B). Furthermore, we amplified the pMiniMAD2 vector (previously cut with KpnI in order to linearize the plasmid) with pJS282/283 to get a ~6.3 kB linear PCR product. These DNA fragments all maintained 20-25 bp of sequence homology between adjacently targeted segments. The two inserts, along with the prepared pMiniMAD2 vector were fused using sequence and ligation-independent cloning (SLIC) (10). The *ΔmutS* was integrated into the chromosome as described above for the *mutSF30A* mutant allele. The final strain was *ΔmutS,* as well as containing a deletion of the *mutSL* intergenic region. The resulting strain placed the 4-1881 bases of *mutL* behind the start codon (bases 1-3) of *mutS*, as shown in SI Fig 4.

Thymidine incorporation to monitor DNA synthesis

B. subtilis cells were grown in $S7_{50}$ minimal media supplemented with 2% Glucose at 37˚ C from an initial starting inoculum of 0.050. When the cells reached mid-exponential phase, 1.8 mL of culture was pulse labeled with 45 µL of ^{3}H-thymidine (50 Ci/mmol; 1 mCi/mL) and incubated at 37˚ C. At the appropriate time points (30, 120, 300 and 600s), 0.4 mL of culture was removed and mixed into 3 mL of ice cold 10% TCA and incubated on ice for ≥30 min. Samples were filtered on glass microfibre filters (GF/C, Whatman) via vacuum. Each filter was washed three times with 5 mL ice-cold 10% TCA, followed by three additional washes with 5 mL of -20˚C 70% ethanol. Filters were dried under a heat lamp, placed into scintillation fluid, and incorporation of ^{3}H-thymidine into nucleic acids via replication was elucidated using scintillation counting.

Plasmid construction

pJSL30 was constructed in order to place *mutS-gfp* expression under the control of the P_{spac} promoter at the ectopic *amyE* locus. *mutS* was amplified off of *B. subtilis* (PY79) chromosomal DNA using the primers oJSL81 and oJSL82. The insert was gel purified, digested with SalI and XhoI, PCR purified, and ligated into pBS226 with the original *mutL* insert removed.

pJSL35 was constructed to knock out the mismatch detection capability of pJSL30 my introducing *mutSF30A*. To do so, pJSL30 was quick-changed using standard protocol (Zheng, Baumann, Reymond, 2004) to produce the F30A mutation using the primers oJSL125 and oJSL126.

pJSL36 was constructed to generate the *mutSF30A* mutant allele at the native locus without the addition of any exogenous nucleotide sequence. The sequence spanning from 457 bases upstream of the *mutS* +1 base to 570 bases downstream was amplified off of the *B. subtilis* PY79 chromosomal DNA using the primers oJSL123 and oJSL124. The fragment was gel extracted, digested with BamHI and KpnI, and ligated into pMiniMAD2. pMiniMAD2 contains an erythromycin cassette, as well as a temperature sensitive origin of replication (37˚).

pJSL40 was constructed to place codon optimized (C.o.) and monomeric CFP downstream of *dnaX* for native locus integration. *C.o. cfp* was amplified off of pDR200 (Doan, Marquis, Rudner. 2005) with primers oJSL127 and oJSL128. The fragment was gel purified, cut with restriction enzymes SphI and XhoI, and ligated into pKL147 with the *gfp* insert removed.

pJSL44 was constructed to knock out the mismatch detection capability of MutS at its native locus. pJSL44 is pJSL36 with the F30A mutation introduced via quick-change. The primers used for the quick-change reaction were oJSL125 and oJSL126.

pJSL45 was constructed to over-express MutSF30A without any added tags. The F30A directed change was introduced into *pET11t-mutS* using oJSL125 and oJSL126.

pJSL57 (See supplemental methods)

pJSL58 was constructed in order to place *mutS800* under the control of a P_{spac} promoter at the *amyE* ectopic locus. We amplified *mutS800* off of the *B. subtilis* chromosome using primers oJSL178 and oJSL179. We gel purified and digested the PCR fragment with SphI and NheI, followed by a PCR clean-up step. Finally, we ligated the insert together with opened pDR110.

pJSL63 was constructed to place *mutS800-gfp* under the control of the P_{spac} promoter at the *amyE* ectopic locus. We amplified *mutS800* off of the *B. subtilis* chromosome using primers oJSL81 and oJSL133. We digestion and amplified and gel purified fragment with XhoI and SalI. This digested fragment was ligated into opened pBS226.

pJSL64 was constructed to place *mutSF30A800-gfp* under the control of the P_{spac} promoter at the *amyE* ectopic locus. This construct is derived from pJSL63, where the F30A substitution was quick-changed into the plasmid using primers oJSL125 and pJSL126.

pJSL67 was constructed for localizing DnaX within the cell using the mCherry fluorophore. I amplified *dnaX-mCherry* off of pJSL52 (*pKL147-dnaX-mCherry*) using primers oJSL170 and 171. Using SLIC (see methods), I placed the insert into pBGSC6 that was amplified using primers JSL168 and oJSL169.

All constructs were sequenced prior to use (University of Michigan core sequencing facility).

Table 6. Functional characterization of fusion alleles used within Chapter II.

Relevant genotype	No. of cultures	Mutation rate (10^{-9} mutations/generation)± [95% CI]	Relative mutation rate (% MMR activity)[a]
PY79 (wild type)*	51	1.82 [1.14-2.37]	1 (100%)
*mutSL::spec**	23	154.4 [146.6-162.2]	84.9 (0%)
mutL::mutL-GFP (spc)	25	139.8 [130.7-148.8]	76.9 (9.5%)
mutS::mutS-GFP (spec), amyE::P_{spac}-mutL (cm)	23	17.0 [12.5-21.5]	9.4 (89.0%)
mutS::mutS-YFP (cm) , amyE::P_{spac}-mutL (mls), dnaX::dnaX-coCFP (spc)	23	15.2 [10.5-19.9]	8.4 (90.2%)
ΔmutS, amyE:P_{spac} mutS (kan)	25	3.09 [1.35-4.68]	1 (100%)
ΔmutS, amyE:P_{spac} mutS-GFP (cm)	23	20.8 [16.0-25.6]	6.7 (88.7%)
ΔmutS, amyE:P_{spac} mutS800-GFP (cm)	30	133.3 [125.7-140.9]	43.1 (16.6%)
ΔmutS, amyE:P_{spac} mutS800-GFP (cm), dnaX::dnaX-mCherry (spc)	19	127.5 [120.8-134.2]	41.3 (20.2%)

* Data also found within Table 1.

Calculations based on data obtained in Table II using isogenic *ΔmutS* backgrounds (JSL292 and JSL281). Brackets enclose the lower bounds and upper bounds of the 95% confidence limits

[a] % MMR activity was determined by the following equation: [(R.M.F.null - R.M.F.strain)/(R.M.F.null-R.M.F.wild-type)]•100. The designation “co” represents the codon optimized version of CFP (11).

Table 7. Bacillus subtilis strains used in Chapter II.

Strain	Relevant Genotype	Source or reference
JSL1	PY79-Prototroph, SPβ˚	(12)
LAS45	*mutS::mutS800-gfpmut2 (spc), amyE::Ps$_{pac}$-mutL (cm)*	(2)
LAS86	*mutS::mutS800 (spc), amyE::P$_{spac}$-mutL (cm*	(2)
LAS440	*mutS::mutS-gfpmut2 (spc), amyE::P$_{spac}$-mutL (cm)*	(3)
LAS397	*mutL::mutL-gfp (spc)*	(3)
LAS392	*amyE::P$_{spac}$-mutS-mgfpmut2 (cm)*	(2)
NMD11	*dnaN5(G73R), spoIIIJ::kan*	
JSL171	*amyE::P$_{spac}$-mutS-mgfpmut2 (cm)*	
JSL196	*amyE::P$_{spac}$-mutSF30A-mgfpmut2 (cm)*	
JSL201	*mut-1[polC G430E, S621N] (cm), mutL::mutL-gfpmut2(spc)*	This work and (13)
JSL202	*mut-1[polC G430E, S621N] (cm), mutS::mutS-mgfpmut2 (spc), amyE::P$_{spac}$-mutL (mls)*	
JSL204	*mutSF30A*	
JSL214	*mutSL::spec, amyE::P$_{spac}$-mutSF30A-gfpmut2 (cm)*	
JSL217	*mutS::mutSF30A-mgfpmut2 (spc), amyE::P$_{spac}$-mutL (mls)*	
JSL219	*mutSL::spc, amyE::P$_{spac}$-mutS-F30A-gfpmut2 (cm)*	
JSL230	*mutSF30A::mutS-yfp (cm), amyE::P$_{spac}$-mutL (mls)*	
JSL234	*mutS::mutS-yfp (cm), amyE::P$_{spac}$-mutL (mls), dnaX::dnaX-CFP(opt)(spc)*	
BKM1725	*dnaX::dnaX-yfp (spc),pelB::Psoj (opt-rbs)-cfp(d)-spo0J (ΔparS) (cm)*	(14)
JSL259	*dnaN5, SpoIIIJ::kan, mutSF30A::mutS-mGFPmut2, amyE::P$_{spac}$-mutL (mls)*	
JSL281	*ΔmutS*	
JSL292	*ΔmutS, amyE::P$_{spac}$-mutS(spc)*	
JSL295	*ΔmutS, amyE::P$_{spac}$-mutS-gfp mut2(cm)*	
JSL297	*ΔmutS, amyE::P$_{spac}$-mutS800 (spc)*	

Table S2 continued

Strain	Relevant Genotype	Source or reference
JSL297	*ΔmutS, amyE::P_{spac}-mutS800 (spc)*	
JSL298	*ΔmutS, amyE::P_{spac}-mutSF30A-gfpmut2 (cm)*	
JSL307	*ΔmutS, amyE::P_{spac}-mutS-gfpmut2 (cm), dnaX::dnaX-mCherry (spc)*	
JSL308	*mutL::mutL-gfpmut2 (spc), dnaX:dnaX-mCherry (cm)*	
JSL309	*mutSF30A, mutL::mutL-gfpmut2 (spc)*	
JSL311	*dnaN::dnaN-mgfpmut2 (spc), dnaX::dnaX-mCherry (cm)*	
JSL315	*ΔmutS, amyE::P_{spac}-mutSF30A-gfpmut2 (cm), dnaX::dnaX-mCherry*	
JSL328	*ΔmutS, amyE::P_{spac}-mutS800-GFP (cm)*	
JSL330	*ΔmutS, amyE::P_{spac}-mutSF30A800-GFP (cm)*	
JSL336	*ΔmutS, amyE::P_{spac}-mutS800-GFP (cm), dnaX::dnaX-mCherry (spc)*	

All strains are derivates of PY79.

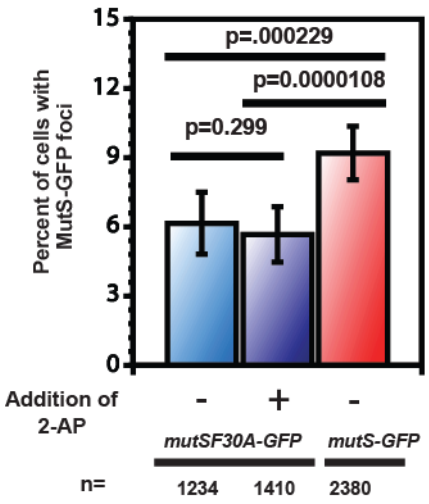


Figure 23. Response of MutS-GFP to the intrinsic error rate *in vivo.*
Shown is a bar graph of the percent of cells with foci for MutSF30A-GFP with (+) and without (-) 2-AP compared to cells with MutS-GFP foci in the absence of 2-AP challenge. The p-values indicate the difference between the percent of cells with foci between each group shown. MutSF30A-GFP and MutS-GFP untreated are statistically significant. The data presented here is the same as in Figure 2C.

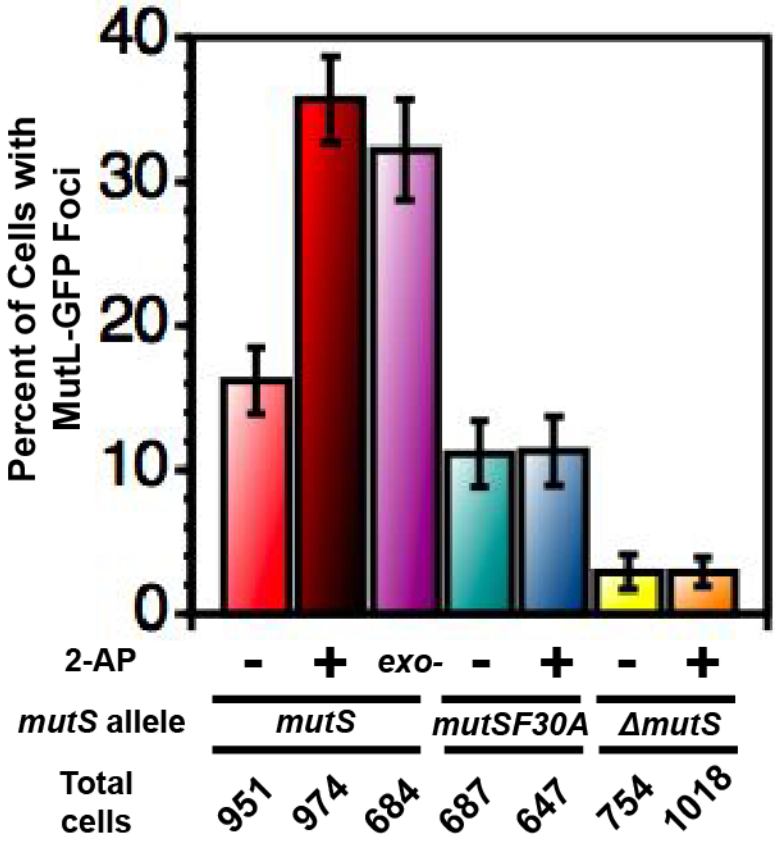


Figure 24. MutL-GFP form foci dependent on MutS and mismatch detection.
Different *mutS* alleles were placed upstream of *mutL-gfp* to determine their effects on MutL-GFP focus formation *in vivo*. The percentage of cells with MutL-GFP foci is shown and represented as a bar graph. The number of cells for each sample scored is indicated as well as each condition. Error bars denote the 95% confidence intervals.

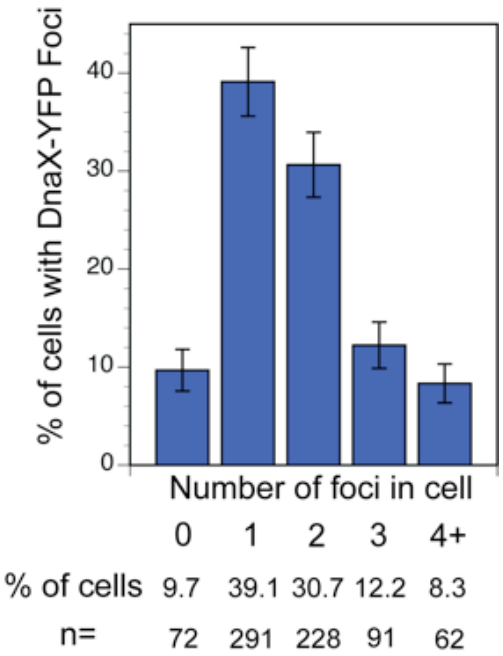


Figure 25. Scoring of DnaX-GFP foci in live cells.
We scored the percentage of cells with DnaX-GFP foci in $S7_{50}$ defined minimal medium supplemented with 1% L-arabinose. The percentage of cells with 0, 1, 2, 3, and 4 or greater foci are shown. The number of cells (n) in each population are also indicated. The error bars reflect the 95% confidence interval. A total of 744 cells were scored over two independent experiments.

Figure 26. Schematic representation of the in frame deletion of *mutS* from the *mutSL* operon. This construct (JSL281) represents a deletion of *mutS* that maintains *mutL* under the control of its native promoter. The start codon of *mutS* (ATG) replaces the start codon of *mutL* (GTG) in order to ensure *mutL* expression. Steady state levels of MutL in this background were ~2-3 fold higher than that observed when *mutS*$^{+}$ is upstream (**Figure 1D**).

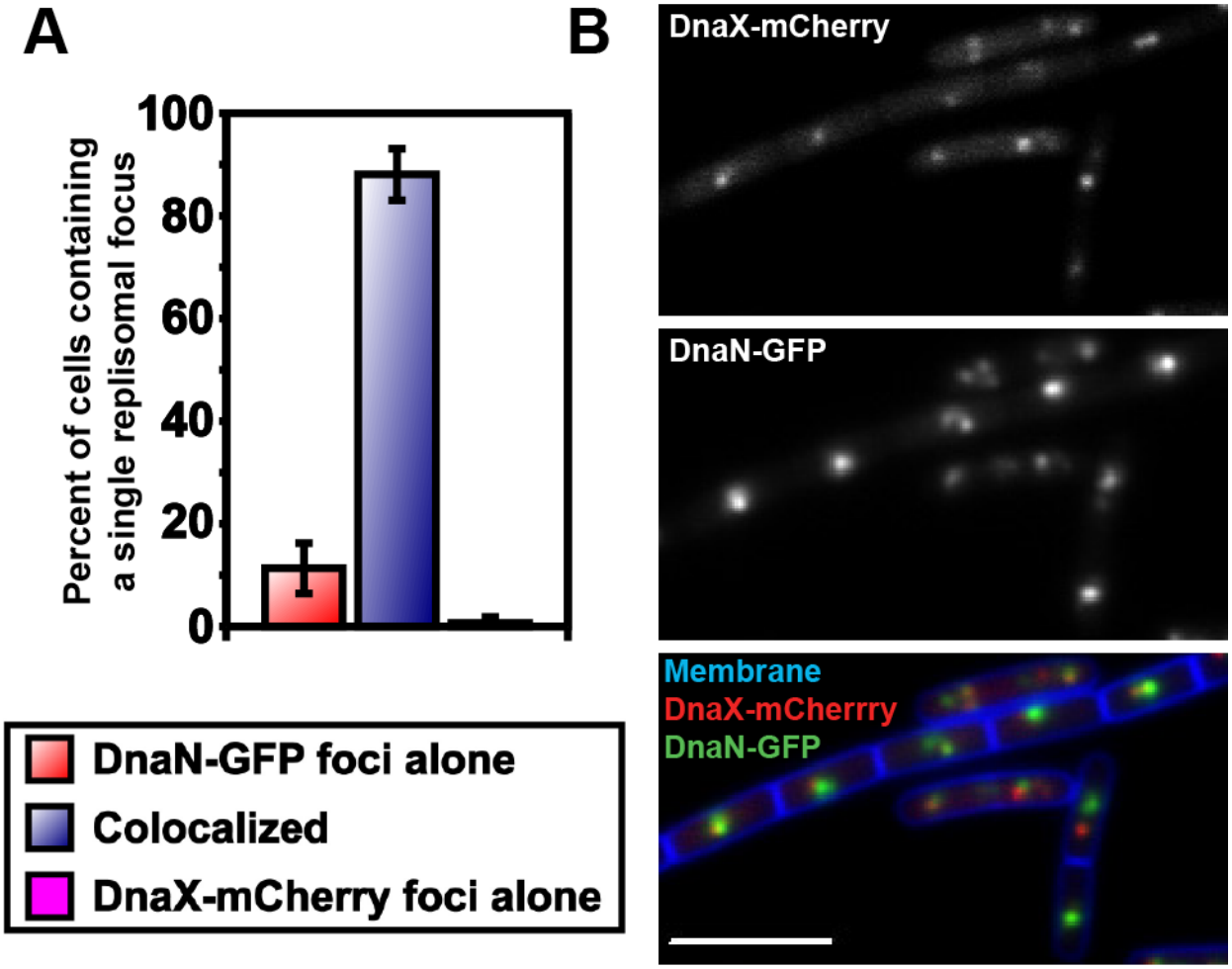


Figure 27. Colocalization of DnaX-mCherry foci with DnaN-GFP foci.
(**A**) The bar graph shows the percentage of cells with a single replisomal focus, visualized with DnaX-mCherry or DnaN-GFP. (**B**) Shown are representative images of DnaX-mCherry (1), DnaN-GFP (2), and the merge of both images with a membrane stain (3). The scale bar indicates 4 μm.

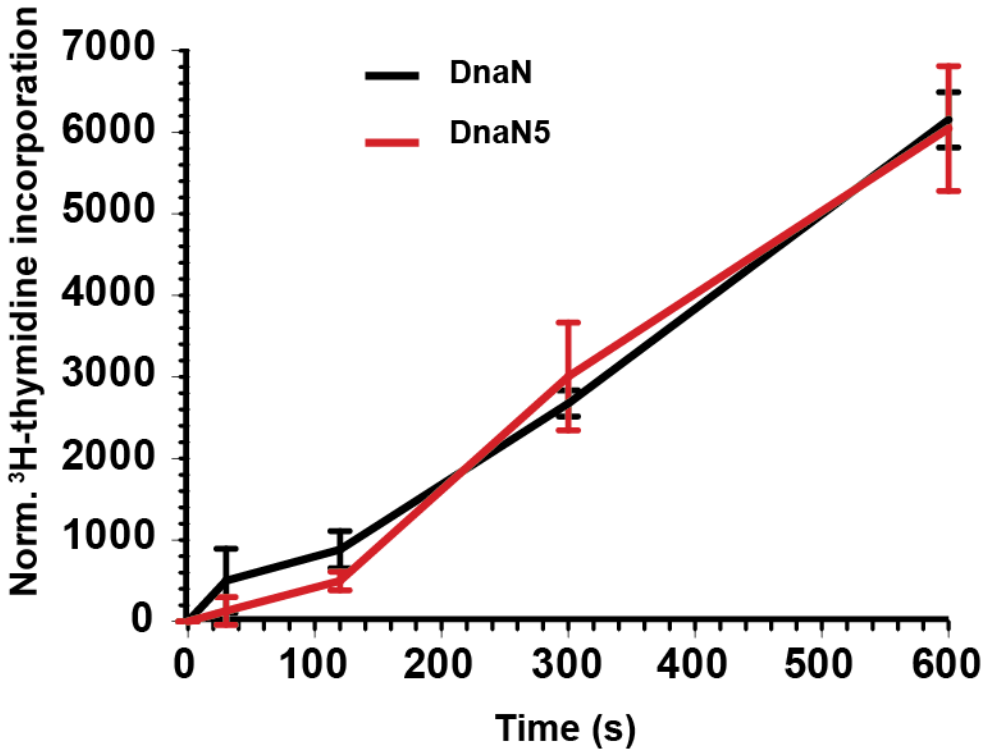


Figure 28. DnaN5 supports DNA replication to wild type levels at 37°C.
The rate of DNA synthesis was measured by monitoring ^{3}H-thymidine incorporation into acid insoluble material. Each data point represents the mean of duplicate samples from 4 independent experiments. The error bars represent the standard error of the mean. ^{3}H-thymidine incorporation was normalized to OD_{600} of the samples at each time point tested.

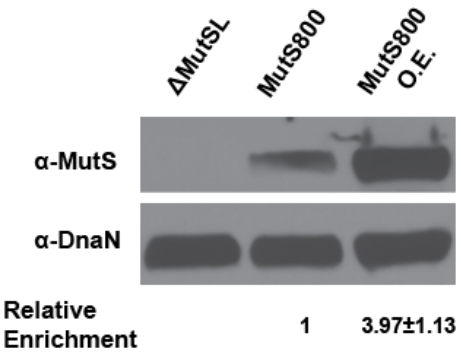


Figure 29. Quantifying ectopic expression of MutS800.
Relative protein levels of MutS800 were determined by performing immunoblot analysis followed by quantification using Image J software (http://rsb.info.nih.gov/ij/). Total band intensities were determined by subtracting the background signal from the *ΔmutSL* lane followed by normalization to the total protein load determined from analysis of the DnaN band intensity. The relative enrichment was determined from 3 independent samples with the standard deviation shown.

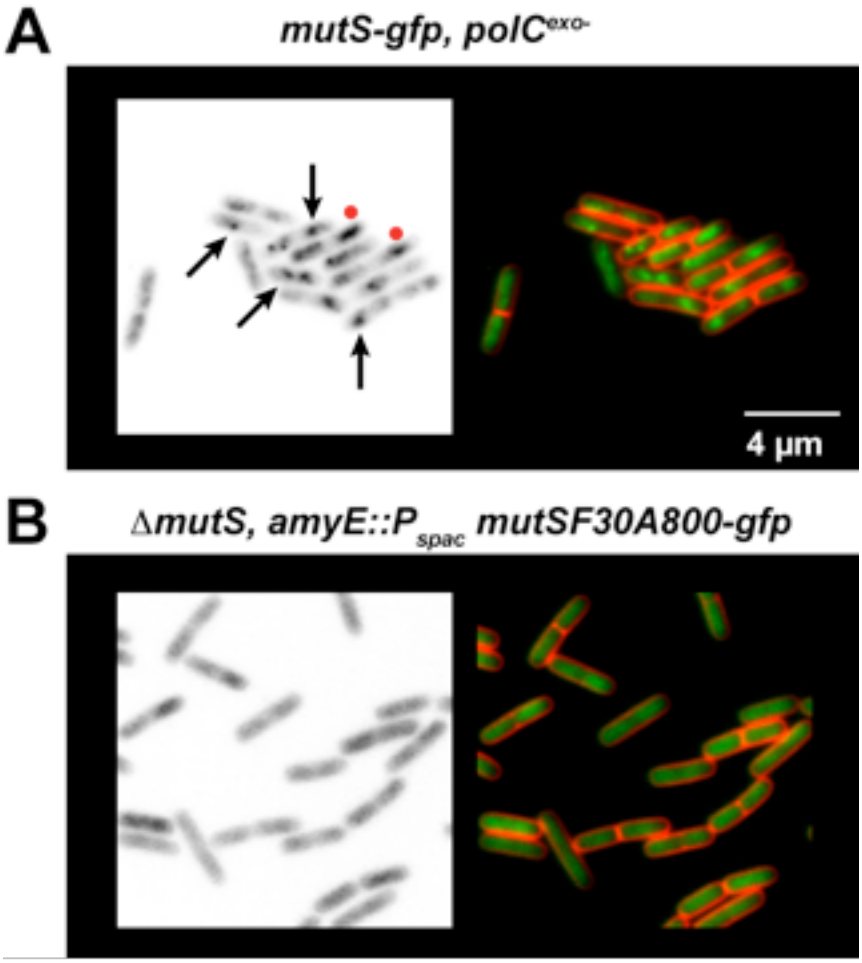


Figure 30. Method for scoring MutS-GFP foci in live *B. subtilis* cells.
MutS-GFP and its derivatives form visible repair complexes *in vivo* termed foci. Repair complexes were scored as foci when visualized as having an elevated fluorescence signal occupying a discrete area above the diffuse fluorescence associated with the nucleoid. Shown in (**A**) and (**B**) are representative images of MutS-GFP or a mutant form defective in forming foci. The images are divided into an inverted monochromatic image of the GFP signal (left) and a colored image showing the GFP image in green overlaid with the membrane imaged with the vital membrane stain TMA-DPH (DAPI channel), pseudo-colored red. For every experiment in this work, foci were scored using corresponding inverted monochrome images that were compared with the colored image. Shown in (**A**) is a representative image of cells with MutS-GFP foci that have formed in response to natural mismatches produced by the *polC*$^{exo-}$ allele. Black arrows indicate examples of scored foci while the red circle denotes elevated nucleoid fluorescence, which does not represent MutS foci and was excluded from the calculation. (**B**) Shown are cells with MutSF30A800-GFP, a mutant which fails to form repair complexes.

References

1. Hombauer H, Campbell CS, Smith CE, Desai A, & Kolodner RD (2011) Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 147(5):1040-1053.
2. Simmons LA, Davies BW, Grossman AD, & Walker GC (2008) Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Mol Cell* 29(3):291-301.
3. Smith BT, Grossman AD, & Walker GC (2001) Visualization of mismatch repair in bacterial cells. *Mol. Cell* 8(6):1197-1206.
4. Klocko AD, Crafton KM, Walsh BW, Lenhart JS, & Simmons LA (2010) Imaging mismatch repair and cellular responses to DNA damage in Bacillus subtilis. *J Vis Exp* (36).
5. Klocko AD*, et al.* (2011) Mismatch repair causes the dynamic release of an essential DNA polymerase from the replication fork. *Mol Microbiol* 82(3):648-663.
6. Hall BM, Ma CX, Liang P, & Singh KK (2009) Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics* 25(12):1564-1565.
7. Rokop ME, Auchtung JM, & Grossman AD (2004) Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of Bacillus subtilis. *Mol Microbiol* 52(6):1757-1767.
8. Simmons LA & Kaguni JM (2003) The DnaAcos allele of Escherichia coli: hyperactive initiation is caused by substitution of A184V and Y271H, resulting in defective ATP binding and aberrant DNA replication control. *Mol Microbiol* 47(3):755-765.
9. Patrick JE & Kearns DB (2008) MinJ (YvjD) is a topological determinant of cell division in Bacillus subtilis. *Mol Microbiol* 70(5):1166-1179.
10. Li MZ & Elledge SJ (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods* 4(3):251-256.
11. Doan T, Marquis KA, & Rudner DZ (2005) Subcellular localization of a sporulation membrane protein is achieved through a network of interactions along and across the septum. *Mol Microbiol* 55(6):1767-1781.
12. Youngman P, Perkins JB, & Losick R (1984) Construction of a cloning site near one end of Tn917 into which foreign DNA may be inserted without affecting transposition in Bacillus subtilis or expression of the transposon-borne erm gene. *Plasmid* 12(1):1-9.
13. Sanjanwala B & Ganesan AT (1991) Genetic structure and domains of DNA polymerase III of Bacillus subtilis. *Mol Gen Genet* 226(3):467-472.
14. Sullivan NL, Marquis KA, & Rudner DZ (2009) Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 137(4):697-707.

Appendix II
Supplemental Information for Trapping and visualizing intermediate steps in the mismatch repair pathway *in vivo*

Justin S. Lenhart[1§], Monica C. Pillon[2§], Alba Guarné[2] and Lyle A. Simmons[1*]

[1] Department of Molecular, Cellular, and Developmental Biology
University of Michigan, 830 North University Ave, Ann Arbor, MI 48109-1048
[2] Department of Biochemistry and Biomedical Sciences, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1

Author contributions: All work except Figures 33 and 34 was performed by JSL. Figures 33 and 34 were performed by MCP. JSL, LAS, MCP, and AG designed the research. All authors contributed to the writing of this appendix.

Appendix is published: Lenhart, J. S., Pillon, M. C., Guarne, A. & Simmons, L. A. Trapping and visualizing intermediate steps in the mismatch repair pathway in vivo. *Molecular microbiology* **90**, 680-698 (2013).

The supplemental information contains “Experimental Procedures”, 7 figures, and 2 tables.

Experimental Procedures

Quantitative Western blotting (LiCOR) analysis.

Whole cell lysates were prepared from independent JSL364 (wild-type) and JSL281 (*ΔmutS*) strains in 6 mL cultures ($S7_{50}$ minimal media supplemented with 2% D-Glucose) grown at 30˚C and harvested at a normalized OD_{600} of 0.5, while simultaneously plated for viables (10^{-6} dilution, see *Spontaneous mutation rate analysis* in the main text*)*. Cells were pelleted and incubated in 1 mL of lysis buffer (10 mM Tris HCl [pH 7.5], 1 mM EDTA, 10 mM $MgCl_2$, 1 mM AEBSF, 0.5 mg/mL lysozyme, and 0.1mg/mL DNase I) and incubated at 37˚C for 10 minutes. After incubation, SDS was added to a final concentration of 1% to lyse cells. Cells were heated for 5 minutes at 100˚C and lysates concentrated to a known final volume in a 10 kDa concentrator column (Amicon Ultra Centrifugal Filters, Millipore).

Immunodot blotting was performed essentially as described {Klocko, 2011 #3902} and as described in “Material and Methods” in the main text. Briefly, whole cell lysates were immobilized onto a nitrocellulose membrane via wet transfer using the mini Trans-Blot electrophoresis transfer cell in transfer buffer without SDS (Bio Rad). The membrane was incubated in blocking buffer (5% milk solids, 17.4 mM Na_2HPO_4, 2.6 mM NaH_2PO_4, 150 mM NaCl) at 22˚C for one hour. All subsequent washes and incubations took place in blocking buffer. After blocking, the membrane was incubated with primary antisera α-MutS (MI 1042) in blocking buffer (minus tween-20) overnight at 4˚C with constant agitation. The next morning, the blot was washed three times for 15 minutes each in blocking buffer supplemented with 0.05% Tween-20. After washing, the blots were then incubated in the dark in 1:15,000 Odyssey Goat anti-Rabbit IR Dye 800CW (926-32211, LiCOR Biosciences) at 22˚C for 2 hours in blocking buffer. All subsequent steps were performed in the dark. The blot was then washed 3 more times in blocking buffer with 0.05% Tween-20, followed by a wash in PBST (17.4 mM Na_2HPO_4, 2.6 mM NaH_2PO_4, 150 mM NaCl, 0.05% Tween-20) to remove excess milk solids. Membranes were dried for 2 hours followed by exposure using an Odyssey CLx Infrared Imaging System (LiCOR, Lincoln, Nebraska). All data analysis and band quantifications were performed using the Odyssey CLx software.

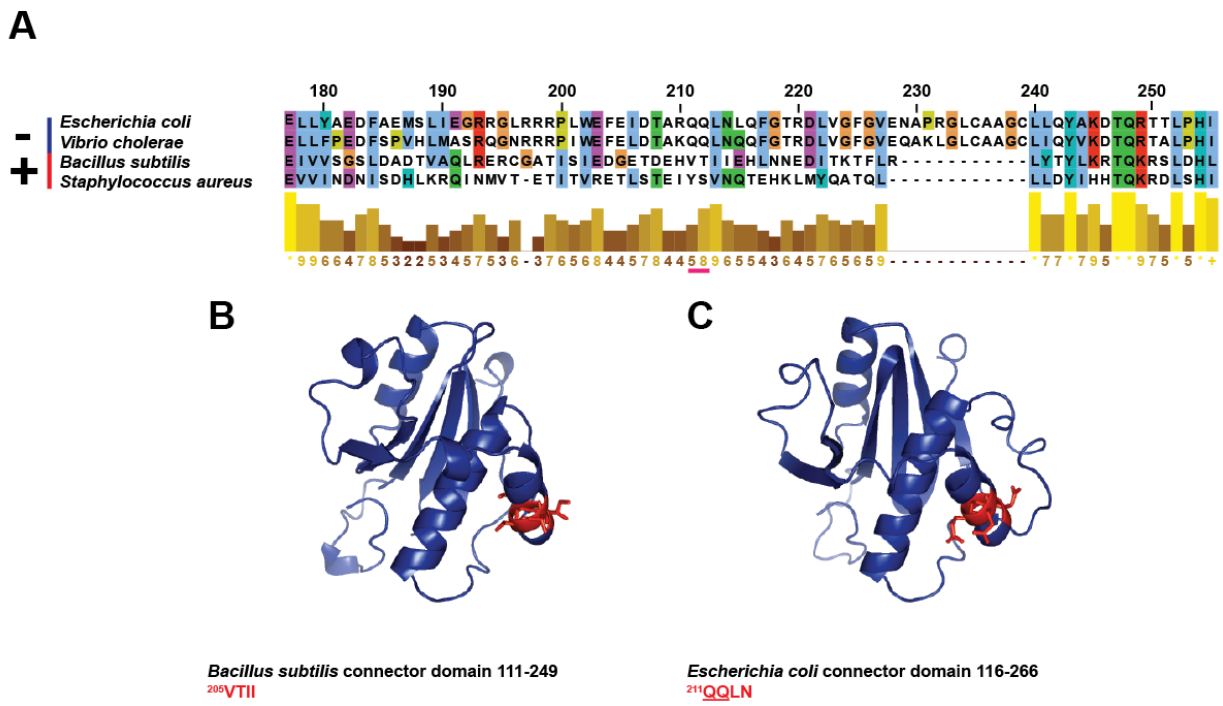


Figure 31. The *E. coli* MutS di-glutamine (Q211 and Q212) binding site for MutL is not conserved in *B. subtilis* MutS.
A) A sequence alignment directly comparing the MutS amino acid sequence of the Gram-negative bacteria *E. coli* and *V. cholerae* to the MutS amino acid sequence of the Gram-positive bacteria *B. subtilis* and *S. aureus*. The region surrounding the di-glutamine MutL docking site (underlined in pink) is shown. The alignment was generated using the http://www.ebi.ac.uk/Tools/msa/clustalw2/ server. The residue numerical designations shown above the alignment are relative to the *E. coli* amino acid sequence. Protein structure models of the B) *B. subtilis* (Phyre2 server model) and the C) *E. coli* connector domain (PDB file 1E3M). Shown in red are the residues corresponding to either the ^{205}VTII site in *B. subtilis* or the ^{211}QQLN of *E. coli*. D) A sequence alignment generated employing the http://www.ebi.ac.uk/Tools/msa/clustalo/ server to visualize the conservation of the QQ site in *E. coli* shown in both Gram-positive and negative bacteria. The order within the alignment is based on a phylogenetic organization of aligned MutS homologs using a Neighbour-joining tree without distance corrections.

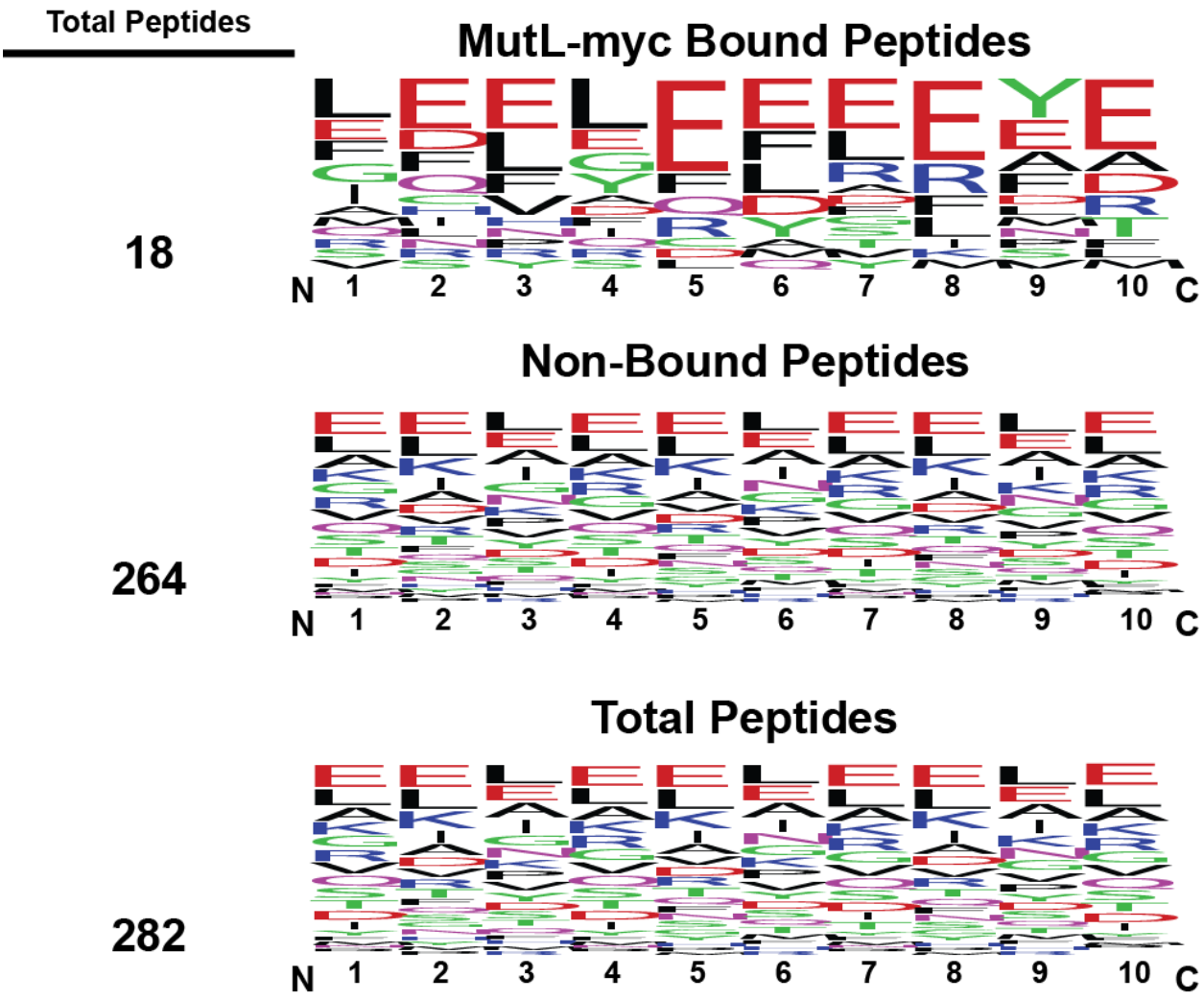


Figure 32. Amino acid composition of MutS peptides recognized by MutL.
Analysis of the amino acid composition of MutL interaction peptides, MutL non-bound peptides, and total MutS peptides in the peptide array. Results visualized with WebLogo 3.1.

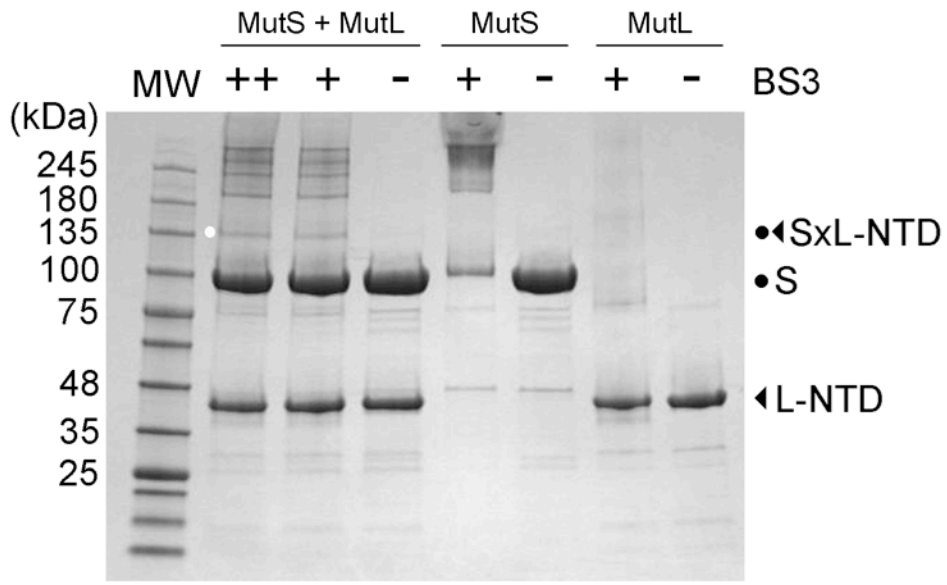


Figure 33. Crosslinking of wild type MutS to the N-terminal domain of MutL with homoduplex DNA. Mixtures of each protein, 10 mM ATP, and a 90 base-pair homoduplex DNA were incubated with the chemical crosslinker BS3 (+=0.8 mM and ++=1.6 mM, respectively). Protein complexes were then separated on a 4-15% gradient SDS polyacrylamide gel. The bands corresponding to the MutS and MutL-NTD monomers, as well as the MutS•MutL-NTD complex are labeled.

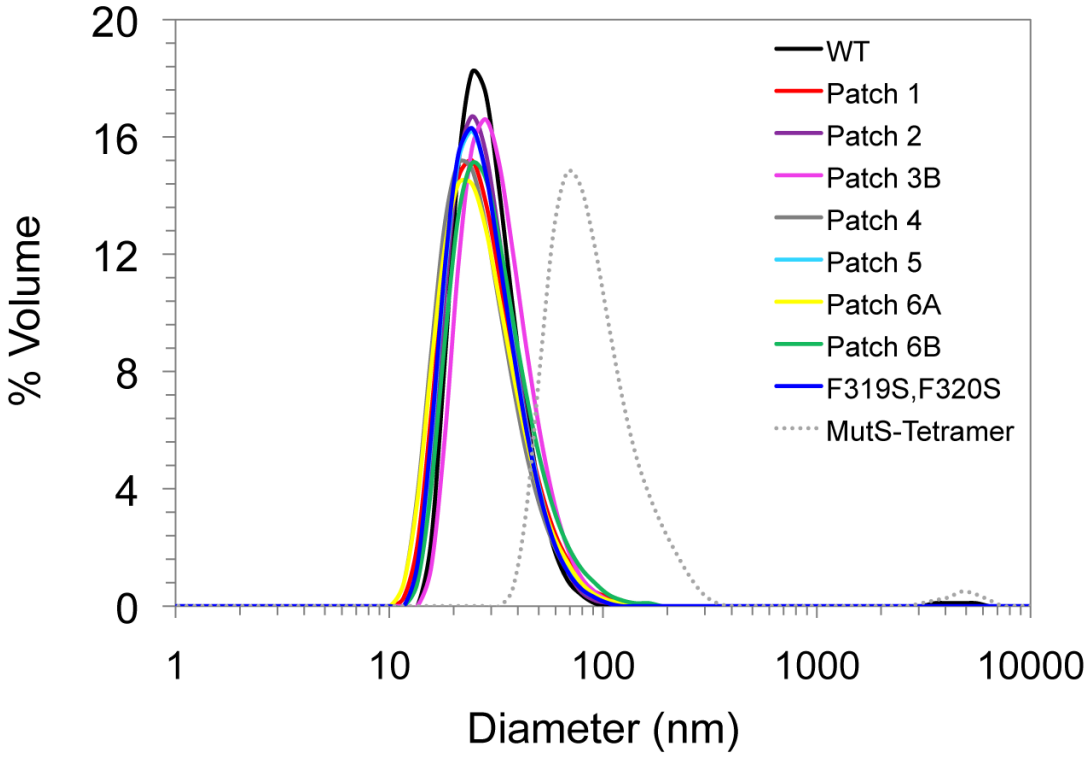


Figure 34. Characterization of the particle size distribution of *B. subtilis* MutS variants at 10 µM (dimer) and 220 µM (tetramer) concentrations reveal similar oligomeric status.
Samples were centrifuged at 15,700 x g for 10 minutes at 4˚C and measured using a Zetasizer Nano S (Malvern Instruments) with a 4 mW He-Ne laser at 633 nm. All measurements were taken using a 12 µL quartz cell (ZEN2112) at 4°C.

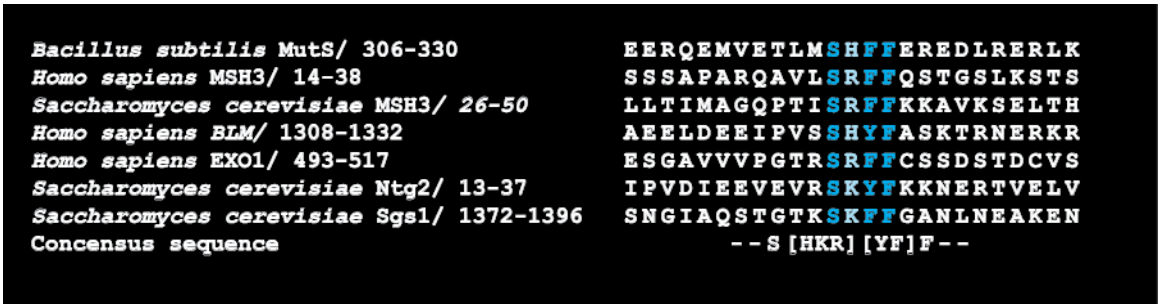
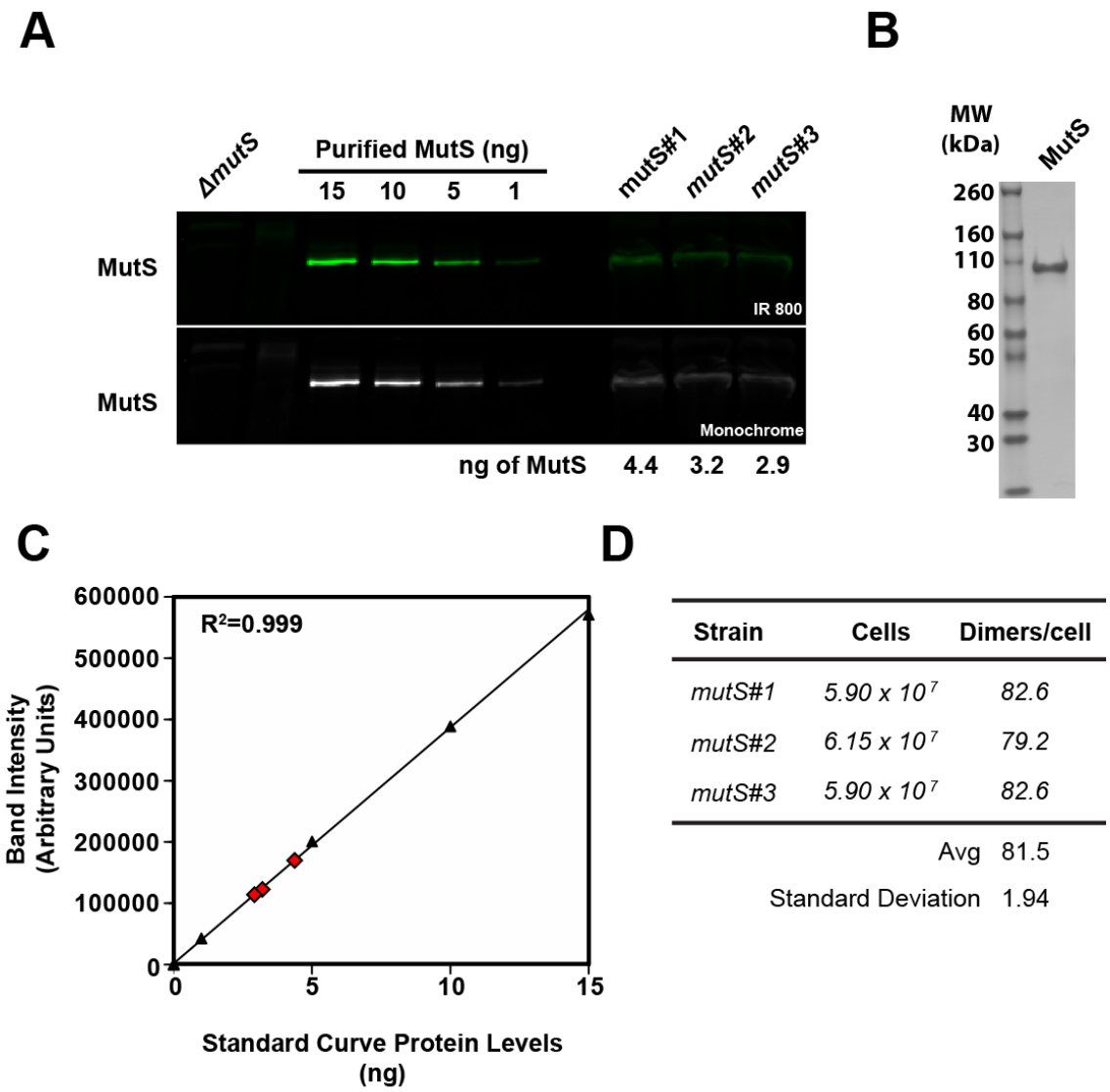


Figure 35. Overlay of the *B. subtilis* MutL binding site on MutS with the Mlh1 binding site on its binding partners.

Shown is an overlay of the *B. subtilis* MutL binding site on MutS with the MIP box (Mlh1 Interacting Protein box_[R/K]-S-[H/R/K]-[Y/F]-F) reveals a conserved serine followed by the di-phenylalanine shared between the *B. subtilis* MutL binding site on MutS and Mlh1 binding partners in *S. cerevisiae*. The blue text in the overlay represents heavily conserved residues found within the MIP box, with the darkest blue representing the most conserved residues based on an eukaryotic Exo I alignment {Dherin, 2009 #4473}.



Strain	Cells	Dimers/cell
mutS#1	*5.90 x 10⁷*	*82.6*
mutS#2	*6.15 x 10⁷*	*79.2*
mutS#3	*5.90 x 10⁷*	*82.6*
	Avg	81.5
	Standard Deviation	1.94

Figure 36. Determination of the absolute number of MutS molecules in *B. subtilis*.
A) The steady state levels of wild type MutS from whole cell extracts grown in the same conditions as those used for live cell microscopy were compared to a protein standard using purified MutS to determine the number of MutS molecules per cell. Band intensity was determined using LI-COR quantitative Western analysis technology. **B)** Image of 2 µg of purified MutS used to construct the protein standard curve found in **A** on a 4-20% gradient gel. **C)** Standard curve of purified MutS protein is pictured in **A**. Red squares indicate the coordinates of total MutS found within the extracts in A. **D)** A fraction of the culture used to make the whole cell extract was used to determine the number of viable cells by plating on LB agar. The cellular content of MutS within the whole cell extract was determined by adjusting the total ng of MutS identified in **A** and **C** by normalizing the amount of MutS to g mol^{-1} using the molecular weight of MutS (97394 Da), followed by multiplying by Avogadro's number (6.022×10^{23}) to obtain total molecules within the extract. This amount was divided by the total viable count of the culture to obtain the number of MutS monomers $cell^{-1}$. This number was further divided by 2 to obtain the number of dimers per cell. Determination of molarity (M) was based on *B. subtilis* cell lengths and widths under standard imaging conditions (avg. length=2.63 µm, avg. width=0.83 µm, V_{cyl}=~5.32×10^{-15} L, M=113 nM).

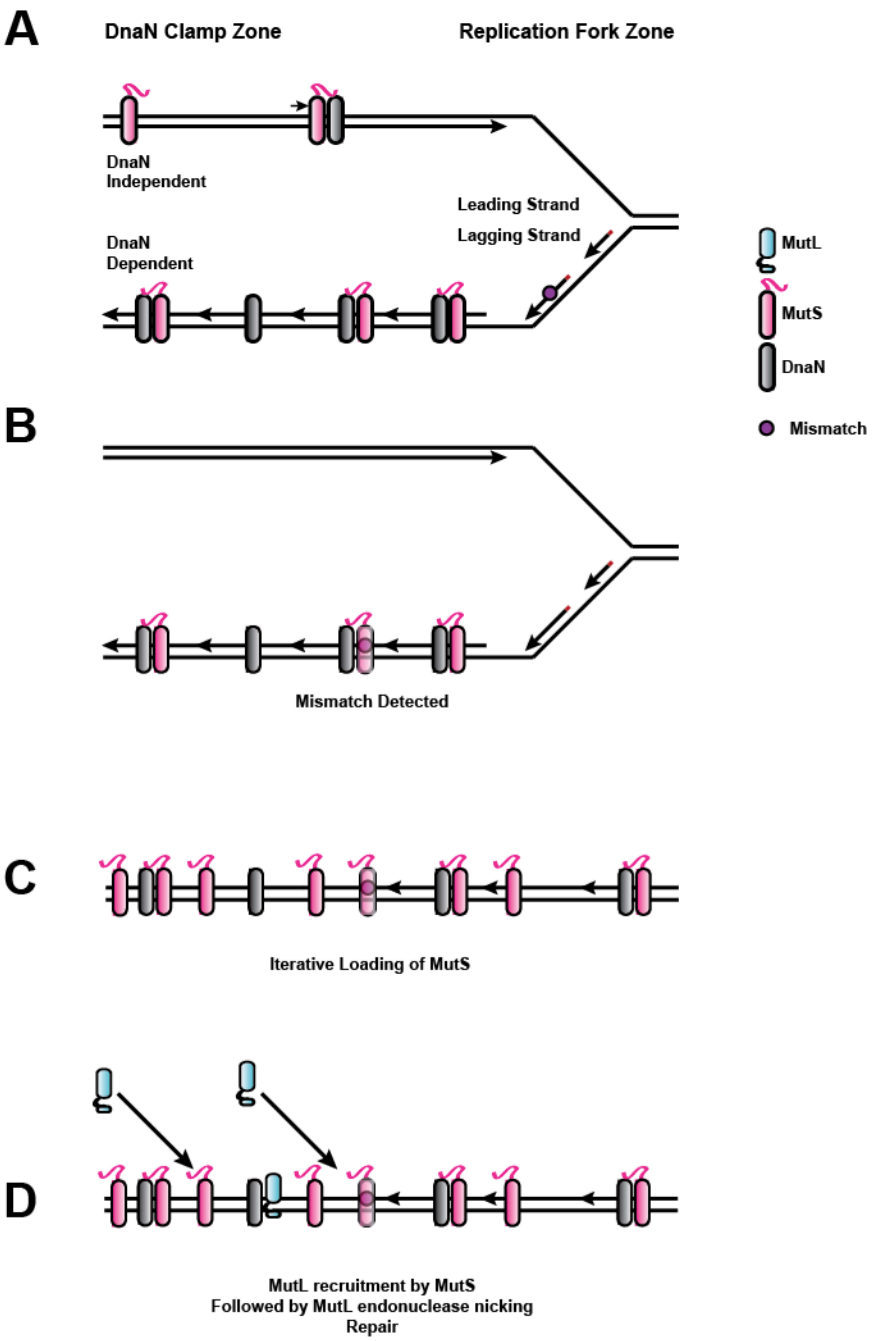


Figure 37. A model of the initial steps of *B. subtilis* mismatch repair.
A) DnaN clamp zones direct MutS to newly replicated DNA to enhance mismatch detection. B) Mismatch detection by MutS. C) Iterative loading of MutS occurs at the site of the mismatch, which D) facilitates recruitment of MutL and endonuclease activation.

Table 8. Individual amino acid substitutions that comprise each MutS patch variant.

Patch designation	Residue substitutions	Domain
WT	None	
Patch 1	E155S, R156S, L157A, E158S	Connector
Patch 2	E245S, E247S, E248S	Connector
Patch 3a	E306S, E307S, E310S	Core
Patch 3b	F320S, E321S, R322S, E323S	Core
Patch 4	E392S, E395S, E396S	Core
Patch 5	E510S, E512S, E514S	Clamp
Patch 6a	Q806A, L807A, F809A, F810A	DnaN clamp binding
Patch 6b	D811S, E812S, E814S	DnaN clamp binding

Each amino acid substitution tested in the MutS patch mutants are listed and the domain location for each patch mutant is indicated based on the structural determination of bacterial MutS homologs {Lamers, 2000 #4022;Obmolova, 2000 #4023}.

Table 9. *Bacillus subtilis* strains used in Chapter III.

Strain	Relevant Genotype	Source or Reference
JSL364	PY79-Prototroph, SPβ˚	(Youngman et al., 1984)
LAS393	*mutL::spec*	(Smith et al., 2001)
LAS435	*mutS::mutS-23-mgfpmut2-spec, mutL*$^{-}$	
JSL161	*mutL::mutL-23-mgfpmut2-spec*	
JSL305	*ΔmutSmutL-23-mgfpmut2-spec*	
JSL342	*mutS Patch 3B-1*	
JSL345	*mutS Patch 4-1*	
JSL346	*mutS Patch 5-6*	
JSL355	*mutS Patch 1-14*	
JSL372	*mutS Patch 3a*	
JSL377	*mutS Patch 2*	
JSL380	*mutSF320S (2-1)*	
JSL382	*mutS E321S*	
JSL386	*mutS Patch 6a-C2*	
JSL395	*mutS Patch 6b-3*	
JSL400	*mutS-22-mgfpmut3mutL*$^{+}$	
JSL402	*mutSF319SF320S*	
JSL414	*mutSF320S-22-mgfpmut3mutL*$^{+}$	
JSL416	*mutS R322S-1*	
JSL419	*mutSE323S-1*	
JSL424	*mutSF319S*	
JSL425	*mutSF319SF320S-22-mgfpmut3mutL*$^{+}$	

JSL438	*mutSF319SF320SmutL::mutL-23mgfpmut2-spec*
JSL440	*mutS-22mgfpmut3mutL::mutLE468K-cm*
JSL450	*mutSF319S-22mgfpmut3mutL*$^+$
JSL453	*mutSE323S-22mgfomut3mutL*$^+$
JSL455	*mutSF319SF320S-22mgfpmut3mutL*$^+$*, dnaX::dnaX-23mCherry-spec*
JSL460	*mutS-22mgfpmut3mutL*$^+$*, dnaX::dnaX-23mCherry-spec*
JSL467	*mutS Patch 3B-22mgfpmut3mutL*$^+$
JSL469	*mutSV206A, T207S, I208A, I209AmutL*$^+$
JSL471	*mutSS317A*
JSL473	*mutSH318S*

All strains listed are derivatives of PY79.

References

Dherin, C., E. Gueneau, M. Francin, M. Nunez, S. Miron, S. E. Liberti, L. J. Rasmussen, S. Zinn-Justin, B. Gilquin, J. B. Charbonnier & S. Boiteux, (2009) Characterization of a highly conserved binding site of Mlh1 required for exonuclease I-dependent mismatch repair. *Molecular and Cellular Biology* **29**: 907-918.

Klocko, A. D., J. W. Schroeder, B. W. Walsh, J. S. Lenhart, M. L. Evans & L. A. Simmons, (2011) Mismatch repair causes the dynamic release of an essential DNA polymerase from the replication fork. *Mol Microbiol* **82**: 648-663.

Lamers, M. H., A. Perrakis, J. H. Enzlin, H. H. Winterwerp, N. de Wind & T. K. Sixma, (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* **407**: 711-717.

Obmolova, G., C. Ban, P. Hsieh & W. Yang, (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407**: 703-710.

Appendix III
RecO and RecR are necessary for RecA loading in response to DNA damage and replication fork stress in Bacillus subtilis

Justin S. Lenhart[1], Eileen R. Brandes[1], Jeremy W. Schroeder[1], Roderick J. Sorenson[2], Hollis D. Showalter[2] and Lyle A. Simmons[1*]

[1]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI

[2]Vahlteich Medicinal Chemistry Core, College of Pharmacy, University of Michigan, Ann Arbor, MI

Author contributions: Figure 38 was performed by JSL. Figures 39, 40, and 41 were performed by ERB. Compound synthesis was done by HDS and RJS. JSL, LAS, ERB, and HDS designed the research. All authors contributed to the writing of this appendix.

Appendix is in revision: *Journal of Bacteriology*

CHEMISTRY EXPERIMENTAL SECTION

Synthesis of HPUra and congeners

The scheme shown in **Figure 2A** in the main text was utilized for the synthesis of HPUra and analogs. Condensation of 5-aminouracil (**1**) with hydrazine (1) or phenylhydrazine (2) led to the adducts **2** and **6**, respectively, in good yield. Reaction of **2** with *p*-benzoquinone proceeded cleanly to yield HPUra (**3**) as a brilliant orange solid. It was identical analytically and spectroscopically to a reference sample obtained from external sources. Compound **3** was further characterized as a *tert*-butyldimethylsilyl (TBS) ether derivative **5** (~ 87:13 mixture of isomers with the major one as assigned). The attempted synthesis of the hydrazino congener **4** (H_2-HPUra) by reduction of **3** was conducted by procedures described previously (2, 3, 4). Numerous attempts at sodium dithionite, following the literature procedure (3) or slight variations thereof, did not provide **4** but only recovered **3**. Any quenching of solution color was always transient, with a return to orange or pale orange upon acidification on workup. Other known methods of reducing diazenes were investigated including zinc in refluxing acetic acid (5), hydrazine hydrate in ethanol (6), and stannous chloride. Neither a color discharge nor reaction was observed for any of these except stannous chloride, which led to degradation of **3**. Since we were unable to produce the reduced form, we tested the oxidized form and found that addition of HPUra (compound **3**) to cells caused a rapid arrest to DNA synthesis *in vivo*. We found that compound 3 indeed blocked replication fork progression **(Figure 2B)**. Therefore, we conclude that compound **3** (HPUra) cannot be reduced to compound **4** (H_2-HPURra).

General chemical methods. All starting materials were obtained from commercial suppliers and were used without further purification. Reactions were performed under a blanket of nitrogen unless specified otherwise. Melting points were determined in open capillary tubes on a Laboratory Devices Mel-Temp apparatus and are uncorrected. ^{1}H and ^{13}C NMR spectra were recorded on a Varian 300 MHz or 400 MHz instrument utilizing d_6-DMSO or D_2O/NaOD as solvent. Chemical shifts are reported relative to the residual solvent peak in δ (ppm). Mass spectra were recorded on a Micromass LCT time-of-flight instrument utilizing the electrospray ionization mode. Thin-layer

chromatography (TLC) was performed on silica gel GHLF plates (250 microns) purchased from Analtech.

6-Hydrazinylpyrimidine-2,4(1*H*,3*H*)-dione (2). A stirred suspension of 6-aminouracil (**1**; 5 g, 39.3 mmol), 62% hydrazine monohydrate (9.85 mL, 197 mmol), acetic acid (5 mL), and water (30 mL) was heated at reflux for 6 h and then cooled. The precipitate was collected by filtration, washed successively with water, ethanol, ether, and then dried to leave 4.58 g (82%) of **2** as an off-white solid, mp 279° (dec); TLC (4:1 chloroform / methanol) showed complete absence of starting material.

6-((4-Hydroxyphenyl)diazenyl)pyrimidine-2,4(1*H*,3*H*)-dione (3). A suspension of 6-hydrazinylpyrimidine-2,4(1*H*,3*H*)-dione (**2**; 12.7 g, 89 mmol), benzoquinone (9.66g, 89 mmol), and formic acid (68.5 mL) was maintained at room temperature for 5 min, and then heated at 50° C for 5-8 min. After cooling the orange solids were collected by filtration, and then washed successively with formic acid, water, ethanol, ether, and vacuum dried to leave 17.4 g (84%) of **3**, mp >260°C (dec);TLC R_f 0.67 (4:1 chloroform / methanol).

2-((*tert*-Butyldimethylsilyl)oxy)-6-((4-hydroxyphenyl)diazenyl)pyrimidin-4(3*H*)-one (5). A mixture of **3** (46 mg, 0.2 mmol), *tert*-butyldimethylsilyl chloride (179 mg, 1.19 mmol), diisopropylethylamine (DIPEA; 0.31 mL, 1.78 mmol), and DMF (0.8 mL) under nitrogen was heated at 50 °C for 20 h. A deep-red solution formed within a few minutes. After cooling, the mixture was diluted with 0.5 mL 2-propanol and placed in the refrigerator to initiate precipitation. The precipitated solids were collected by filtration, rinsed with 2-propanol, and dried to leave 40 mg (58%) of **5** as an orange solid, mp 260°C;TLC R_f 0.24 (1:1 ethyl acetate / hexanes).

6-(2-Phenylhydrazinyl)pyrimidine-2,4(1*H*,3*H*)-dione (6). A mixture of 6-aminouracil (**1**; 0.5 g, 3.9 mmol), phenylhydrazine (0.78 mL, 0.851g, 7.87 mmol), acetic acid (0.45 mL, 7.91 mmol), and water (15 mL) was heated at reflux for 3 h. After cooling, the precipitated solids were collected by filtration, washed well with water and then ethanol, and dried to give 0.72 g (84%) of product as cream-colored solid; TLC R_f 0.68 (3:1 acetonitrile / 0.2M aq. ammonium chloride).

BIOLOGY METHODS SECTION

Plasmids used in this study. Unless otherwise indicated, all plasmids used in this study were constructed using standard cloning procedures (7). All primers used for plasmid construction are available upon request.

pEB1 was constructed for integration and expression of *recO* from a P_{spac} promoter at the *amyE* locus. The 765 nucleotide *recO* coding sequence was PCR amplified using primers oEB1 and oEB2. The PCR product was digested with SphI and HindIII, the same enzymes used to digest pDR66 (8). Plasmid pEB1 was then constructed by ligation of the *recO* coding region with double digested pDR66 using the same enzymes.

pEB20 was constructed by SLIC (9) of a partial fragment of *recR* into pBGSC6 for the disruption of the *recR* gene. The fragmented *recR* coding region was PCR amplified using primers oEB78 and oEB79.

pEB21 was constructed by SLIC of the entire *recR* coding region into pJS101 for integration at the *amyE* locus under a xylose inducible promoter. The *recR* coding region was PCR amplified using primers oEB83 and oEB84. The pJS101 vector was amplified using primers oJS431 and oJS432.

pJSL112 was constructed by Gibson assembly. A ~1 kB region corresponding to the immediate upstream and downstream regions of *recF* was amplified by PCR with primers oJSL304/oJSL305 and oJSL306/oJSL307, respectively. The chloramphenicol cassette was amplified from pGEM with primers oJSL294 and oJSL295. A basic plasmid backbone containing the *ori* and the ampicillin cassette was amplified from pDR111 using oJSL302 and oJSL303. All four DNA fragments were assembled in a single Gibson Assembly reaction at 50˚ C for 1 hour using standard Gibson protocol.

All constructed clones were sequenced prior to use by the University of Michigan core sequencing facility (http://seqcore.brcf.med.umich.edu/).

Immunoblot analysis. Immunoblot analysis was done essentially as described (11). Briefly, strains were grown in 10 mL defined $S7_{50}$ minimal media containing 2% glucose, with appropriate antibiotics to mid exponential phase, and concentrated by centrifugation. Cells were resuspended using 300 µL lysis buffer [10 mM Tris-HCL (pH 7.0), 1X Protease Inhibitor, 0.5 mM EDTA, 1 mM 4-(2-aminoethyl)-benzenesulfonyl

fluoride (AEBSF), 1 mg/mL lysozyme, 1 U/mL DNase I in ddH_2O] as previously described (12) and lysed via sonication for 15 seconds twice with a 1 minute pause in-between, followed by the addition of SDS to 1%. Proteins were separated on a 10% SDS-PAGE followed by transfer to a nitrocellulose membrane (Whatman) using 1X transfer buffer (24 mM Tris, 192 mM glycine, pH 8.2, 15% methanol) overnight at 15 volts as described (13, 14). The membrane was blocked with 5% non-fat milk in TBS+0.02% Tween 20. 1:1000 primary antibody was added to fresh blocking solution, and incubated with the membrane overnight on an orbital shaker at 4 °C. The nitrocellulose membrane was washed 3X in 1X TBS-Tween 20 (0.02%) followed by 1:1000 dilution of goat-anti-rabbit-HRP conjugated secondary antibody in 5% milk/TBS-Tween on an orbital shaker at room temperature for 2 hours. The nitrocellulose membrane was washed again 3X with 1X TBS-tween (0.02%) followed by incubation with 2 mL SuperSignal West Pico Luminol/ Enhancer Solution and 2 mL SuperSignal West Pico Stable Peroxide Solution (Thermo Scientific). Blots were exposed to film (BioExpress) for 1 minute prior to developing.

BIOLOGY RESULTS

We also established the time dependency of formation of RecA foci before performing the experiments below. We found that the DNA break inducing peptide phleomycin caused a linear increase in the percentage of cells with foci beginning to form 7.5 minutes after chemical treatment, with 1.2% of cells gaining RecA-GFP foci/minute (R^2=0.983) over a 60 minute time course **(Figure 38)**. Mitomycin C (MMC), which upon being imported into the cell is reduced to produce its toxic form, begins producing a linear increase in the percentage of cells with RecA-GFP foci at a similar rate (1.2% of cells gaining foci/minute) starting at 20 minutes past treatment (R^2=0.998). At 30 minutes past treatment, both MMC and phleomycin cause production of RecA filaments.

Table 10. List of Plasmids for Chapter IV

Plasmids	Characteristics	Source
pEB1	*recO* in pDR66	This work
pEB21	*recR* in pJS101	This work
pEB22	*recO* in pJS102	This work
pCm::Tet	*pCM::Tet*	BGCS
pCM::Er	*pCM::Er*	BGCS
pJSL112	*recF::cm*	This work

Table 11. List of *Bacillus subtilis* strains for Chapter IV.

Strains	Genotype	Source
LAS508	PY79 (wild type)	{Youngman, 1984 #3434}
LAS40	*recA-23mgfpA206K(spc^r)* (referenced below as *recA-gfp*)	{Simmons, 2007 #2929}
ERB1	*amyE::P$_{spac}$-recO(cm^r)*	This work
ERB2	*amyE::P$_{spac}$-recO(tet^r)*	This work
ERB5	*amyE::P$_{spac}$-recO (tet^r); recO::cat (cm^r)*	This work
ERB5	*amyE::P$_{spac}$-recO(tet^r) ; recO::cat (cm^r); recA-gfp (spc^r)*	This work
ERB63	*recR::cat(cm^r)*	This work
ERB64	*recR::cat(cm^r), recA-gfp(spc^r)*	This work
ERB69	*amyE::P$_{xyl}$-recR(mls^r)*	This work
ERB70	*amyE::P$_{xyl}$-recR(mls^r); recR::cat(cm^r)*	This work
ERB76	*amyE::P$_{xyl}$-recR(mls^r); recR::cat; recA-gfp(spc^r)*	This work
ERB77	*dnaB134(ts) (tet^r); recO::cat(cm^r); recA-gfp(spc^r)*	This work
ERB78	*recO::cat (cm^r)*	This work
ERB85	*recO::cat(cm^r); recA-gfp(spc^r)*	This work
JSL484	*recA-gfp, dnaX::dnaX-mCherry(cm^r)*	This work
JSL496	*recF::cat(cm^r); recA-gfp(spc^r)*	This work
JSL635	*ΔrecJ, addA::erm, recA-gfp(spc^r)*	This work

SUPPLEMENTAL FIGURES

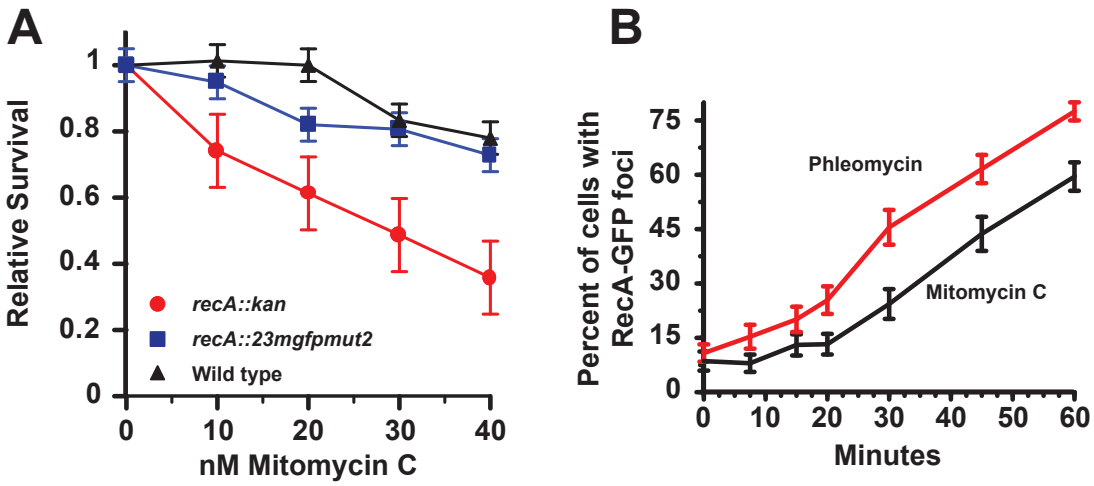


Figure 38. RecA-GFP focus formation in response to DNA damage.
(A) Shown is a survival curve of cells challenged with concentrations of mitomycin C ranging from 0-40 nM for 30 minutes. The error bars represent the standard error of the mean **(B)** Shown is a time course experiment representing the percentage of cells with RecA-GFP foci at 40 nM mitomycin C (black) or 400 nM phleomycin (red) over 60 minutes. Each treatment was scored in triplicate at the time points indicated. Error bars reflect the 95% confidence interval.

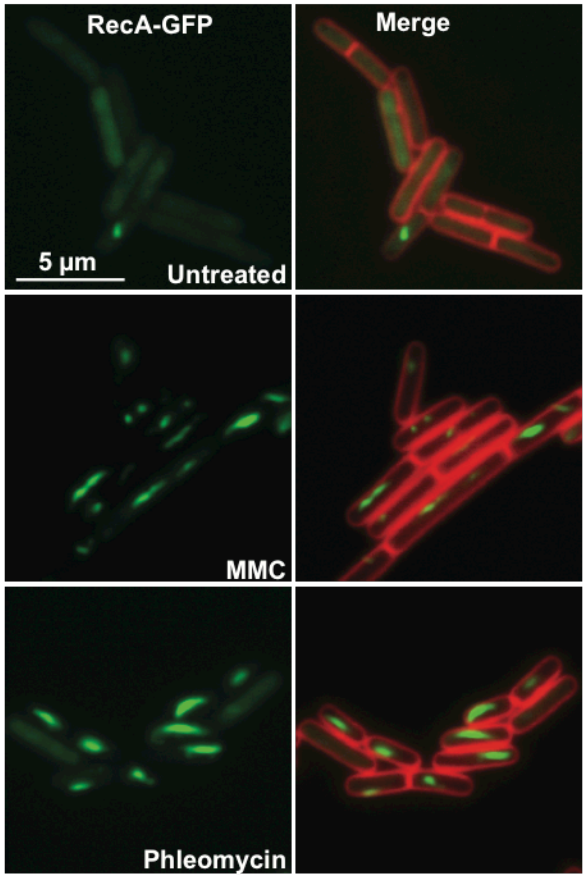


Figure 39. RecA-GFP focus formation in response to DNA damage.
Shown are representative images of cells with the *recA-gfp* allele that were untreated, or challenged with 100 ng/ml MMC or 3 µM phleomycin. The membrane is stained with TMA-DPH and is pseudocolored red.

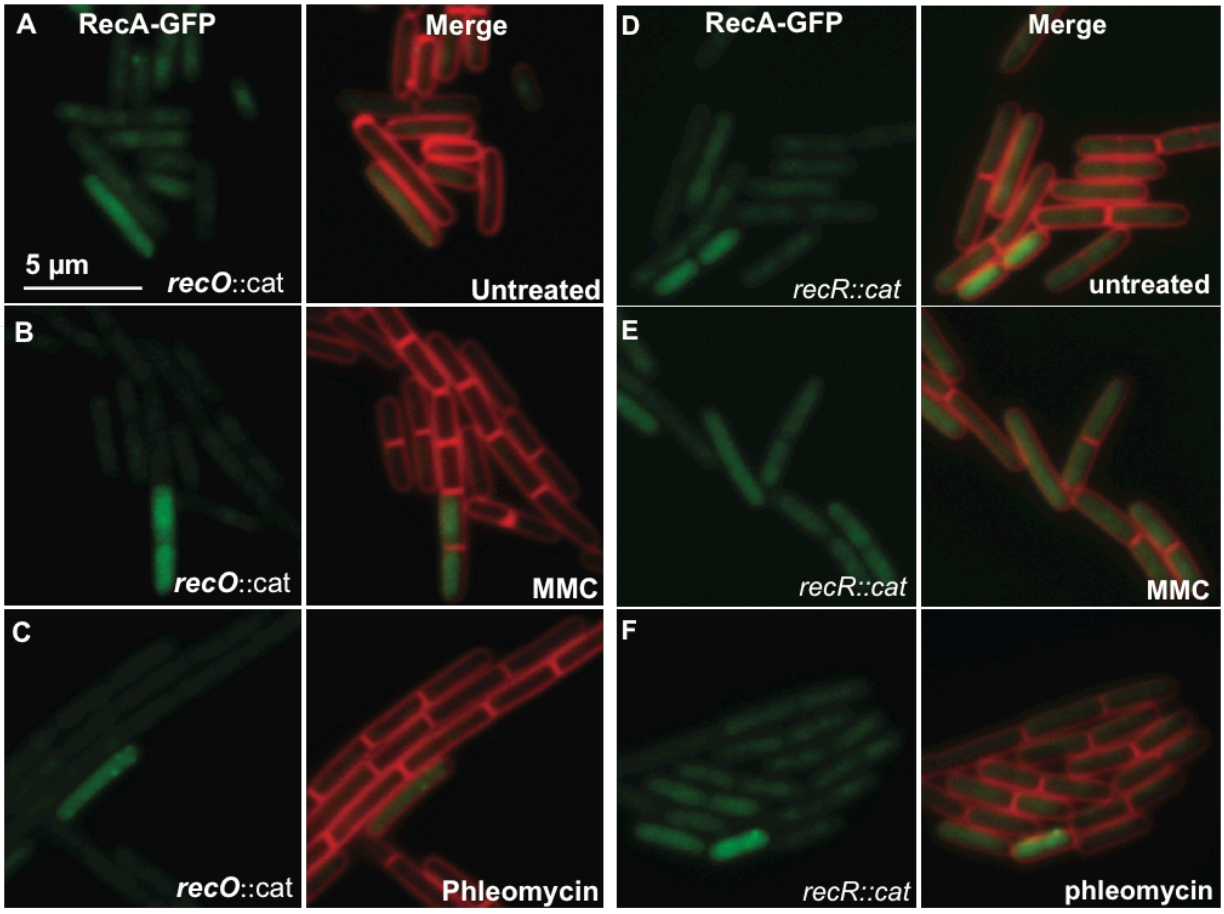


Figure 40. The *recO* and *recR* genes are necessary for RecA-GFP focus formation.
Shown are representative images of a *recA-gfp* allele in a *recO::cat* (**A-C**) or *recR::cat* background (**D-F**). Each strain was untreated, or challenged with mitomycin C (100 ng/ml) or phleomycin (3 μM) as indicated. Membranes were stained with TMA-DPH and pseudocolored red.

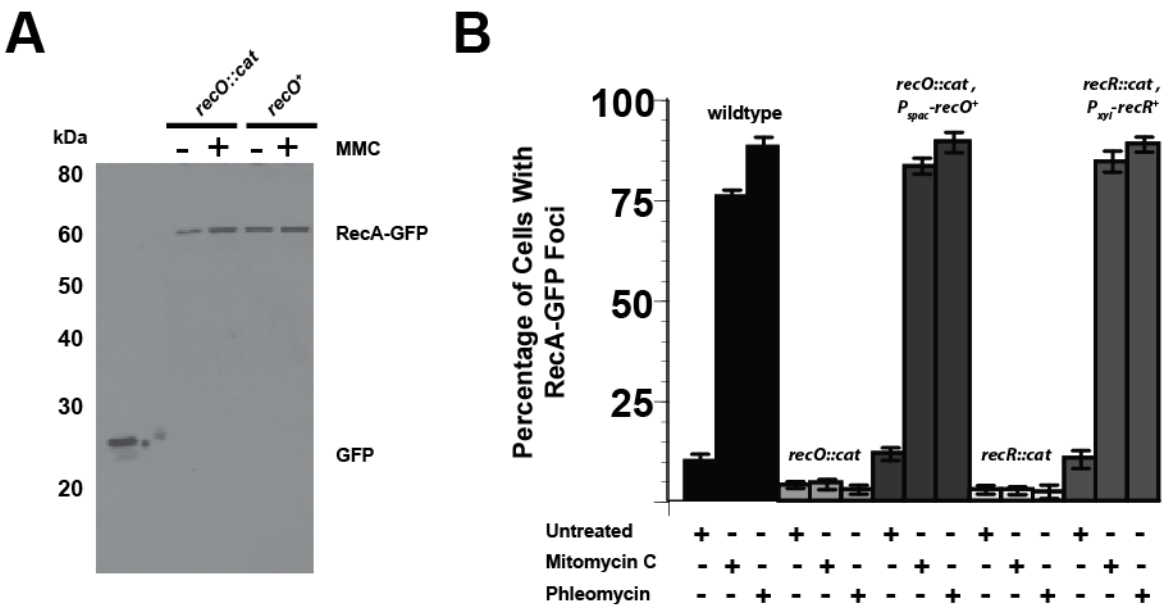


Figure 41.
(**A**) Immunoblot of RecA-GFP in cells that were wild type or contained the *recO::cat* allele. Samples treated with mitomycin C are indicated (+). (**B**) A bar graph quantifying the percentage of cells with RecA-GFP foci under the given condition. The error bars represent the 95% confidence interval. For each condition at least 850 cells were scored from at least two independent experiments.

Figure 42. Sequence alignment of the C-terminal residues of *B. subtilis* and *E. coli* SSB. Shown is the primary structure of the C-terminal 57 amino acids of *B. subtilis* SSB aligned with *E. coli*. The red residues correspond to the PF motif and the underlined region denotes the portion missing in the *ssbΔ35* allele {Costes, 2010 #3427}.

REFERENCES

1. **Kirkor W, Kozlowska E, Grudzinska P.** 1969. Pyrimidine bases. II. Condensation of 2,6-dihydroxy-4-hydrazinepyrimidine with aldehydes. Rocz. Chem **43:**519-523.
2. **Wright GE, Brown NC.** 1974. Synthesis of 6-(phenylhydrazino)uracils and their inhibition of a replication-specific deoxyribonucleic acid polymerase. J Med Chem **17:**1277-1282.
3. **Mackenzie JM, Neville MM, Wright GE, Brown NC.** 1973. Hydroxyphenylazopyrimidines: characterization of the active forms and their inhibitory action on a DNA polymerase from Bacillus subtilis. Proceedings of the National Academy of Sciences of the United States of America **70:**512-516.
4. 1961. Pyrimidine derivates patent GB876601A.
5. **Ismagilov R, Moskva V, Lebedev Y, Mosunova L.** 2001. Synthesis and chemical transformations of β-(3,5-di-tert-butyl-4-hydroxyphenyl)propionic hydrazide. Russ. J. Gen. Chem. **71:931-933**.
6. **Loetchutinat C, Chau F, Mankhetkorn S.** 2003. Synthesis and evaluation of 5-Aryl-3-(4-hydroxyphenyl)-1,3,4-oxadiazole-2-(3H)-thiones as P-glycoprotein inhibitors. Chem. Pharm. Bull **51:728-730**.
7. **Sambrook J, Fritsch EF, Maniatis T.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York.
8. **Ireton K, Rudner DZ, Siranosian KJ, Grossman AD.** 1993. Integration of multiple developmental signals in Bacillus subtilis through the Spo0A transcription factor. Genes Dev **7:**283-294.
9. **Li MZ, Elledge SJ.** 2007. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods **4:**251-256.
10. **Lovett Jr. CM, Roberts JW.** 1985. Purification of a RecA analogue from *Bacillus subtilis*. J. Biol. Chem. **260:**3305-3313.
11. **Bolz NJ, Lenhart JS, Weindorf SC, Simmons LA.** 2012. Residues in the N-terminal domain of MutL required for mismatch repair in Bacillus subtilis. Journal of Bacteriology **194:**5361-5367.

12. **Rokop ME, Auchtung JM, Grossman AD.** 2004. Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of Bacillus subtilis. Mol Microbiol **52:**1757-1767.

13. **Simmons LA, Kaguni JM.** 2003. The DnaAcos allele of Escherichia coli: hyperactive initiation is caused by substitution of A184V and Y271H, resulting in defective ATP binding and aberrant DNA replication control. Mol Microbiol **47:**755-765.

14. **Cooper LA, Simmons LA, Mobley HL.** 2012. Involvement of Mismatch Repair in the Reciprocal Control of Motility and Adherence of Uropathogenic Escherichia coli. Infect Immun **80:**1969-1979.

15. **Youngman P, Perkins JB, Losick R.** 1984. Construction of a cloning site near one end of Tn917 into which foreign DNA may be inserted without affecting transposition in Bacillus subtilis or expression of the transposon-borne erm gene. Plasmid **12:**1-9.

16. **Simmons LA, Grossman AD, Walker GC.** 2007. Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. Proc Natl Acad Sci U S A **104:**1360-1365.

17. **Costes A, Lecointe F, McGovern S, Quevillon-Cheruel S, Polard P.** 2010. The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. PLoS Genet **6:**e1001238.