Mismatch repair in Gram-positive bacteria

Justin S. Lenhart^{1,2}, Monica C. Pillon³, Alba Guarné^{3*}, Julie S. Biteen^{2*} and Lyle A. Simmons^{1*}

¹Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109. ²Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109. ³Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario L8S 4K1, Canada

*Corresponding authors

LAS: Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109-1055, United States. Phone: (734) 647-2016, Fax: (734) 647-0881
E-mail: lasimm@umich.edu

JSB: Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055, United States. Phone: (734) 647-1135; Fax: (734) 647-1179 E-mail: jsbiteen@umich.edu

AG: Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario L8S 4K1, Canada Tel.: 905-525-9140 (ext.26394); Fax: 905-522-9033; E-mail: guarnea@mcmaster.ca.

Abstract

DNA mismatch repair (MMR) is responsible for correcting errors formed during DNA replication. DNA polymerase errors include base mismatches and extra helical nucleotides referred to as insertion and deletion loops. In bacteria, MMR increases the fidelity of the chromosomal DNA replication pathway approximately 100-fold. MMR defects in bacteria reduce replication fidelity and have the potential to affect fitness. In mammals, MMR defects are characterized by an increase in mutation rate and by microsatellite instability. In this review, we discuss current advances in understanding how MMR functions in bacteria lacking the MutH and Dam methylase-dependent MMR pathway.

Mismatch repair in *Escherichia coli*

The Gram-negative bacterium *Escherichia coli* has served as the traditional model for the study of bacterial MMR. Experiments in this model system include the extensive biochemical characterization of all MMR components, including a successful reconstitution of the pathway *in vitro* [1](see for review [2, 3]).

Upon the identification of a replication error, *E. coli* MMR is initiated by the mismatch sensing protein MutS (**Figure 1A**) [4]. There are a number of models that have been proposed to describe the initial phase of mismatch recognition and the transmission of this signal to downstream proteins in the pathway [for review [5]]. Briefly, the “translocation model”, suggests that energy released by ATP hydrolysis drives translocation from the mismatch stimulating formation of a loop [6]. The “sliding clamp or molecular switch” model, proposes that upon mismatch binding, ADP is exchanged for ATP, eliciting a conformational change in MutS converting it to a sliding clamp allowing for diffusion along the DNA helix [7, 8]. The third model “static transactivation” suggests that mismatch bound MutS remains bound to the mismatch and facilitates communication with downstream events through DNA looping [9]. Although each model has experimental support, based on considerable *in vitro* and the *in vivo* observations of MMR complexes, we favor, the molecular “sliding clamp model” where MutS converts to a sliding clamp and diffuses along the DNA in search of MutL.

Based on the molecular switch model, after mismatch detection, MutS converts to a sliding clamp followed by recruitment of MutL through an interaction with residues Q211 and Q212 located in the connector domain of MutS [10, 11, 7] (**Figure 1B**). A third principal component of MMR in *E. coli* is the restriction endonuclease-like protein MutH. MutH follows behind the ongoing replication fork through association with the hemi-methylated d(GATC) sites [12]. The MutS•MutL binary complex forms allowing for MutL to activate MutH [13]. Upon activation, MutH specifically nicks the unmethylated strand of the hemimethylated d(GATC) site [7]. The *E. coli* combination of the Dam methylase and MutH activities provides a signal that directs the MMR pathway to the nascent strand in a process termed strand discrimination [14, 1]. MutL loads the UvrD helicase at the newly incised nick on either the continuous strand or the incised strand, depending on whether the MutH-directed nick exists 5' or 3' to the mismatch, allowing for bidirectional excision (**Figure 1**) [15]. Helicase loading ensures that helicase movement occurs toward the detected mismatch by unwinding the strand with the replication error, followed by degradation by one of several single-stranded exonucleases (ExoI, RecJ, ExoVII, ExoX) [16]. In the final correction step, DNA polymerase III holoenzyme replicates the ssDNA gap and DNA ligase seals the nick in the sugar-phosphate backbone (**Figure 1D**) (see for review [17]). 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, exonucleases (RecJ, Exo I, VII, X), the Pol III holoenzyme, DNA ligase, and SSB [1] (Table 1).

Mismatch repair in MutH-independent bacteria

In considering bacterial MMR, it is important to mention that *E. coli* and a few closely related gamma proteobacteria use a MMR pathway where the strand-discrimination signal is known: methylation of adenine in d(GATC) sequences by the Dam methylase marks the template strand [18]. At the start of the cell cycle, d(GATC) sites are fully methylated; however upon replication, most d(GATC) sites exist in a transient hemimethylated state for >2 min [19]. As described above, the brief hemimethylated state is exploited by MutH to incise the DNA and target mismatch

correction to the base located in the unmethylated strand, representing the newly synthesized strand.

E. coli has traditionally served as the bacterial model for MMR studies; however, most prokaryotes and all eukaryotes rely on a MutH- and Dam methylation-independent pathway (see for review [20]). In place of the MutH endonuclease activity, most prokaryotic and eukaryotic MutL homologs (MLHs) contain a highly conserved endonuclease active site, where extensive conservation is apparent even among distantly related organisms including human, *Arabidopsis* and the Gram-positive bacterium *Bacillus subtilis* [21, 22]. Even *Pseudomonas* species, which have Dam-methylase, but lack MutH contain a MutL-endonuclease active homolog [23]. Despite this extensive evolutionary conservation of MLHs, *E. coli* MutL lacks the intrinsic endonuclease activity that defines eukaryotic, and even most prokaryotic, MMR pathways [21]. Therefore, though *E. coli* has certainly served as an important organism for studying the mechanistic steps of MMR since the discovery of mutator alleles [24, 25], *E. coli* uses a methylation-directed MMR pathway, which is rare in biology.

Therefore, to understand how MMR functions in bacteria that lack MutH and Dam, other experimentally tractable bacterial systems, including *Streptococcus pneumonia* and *Bacillus* species, were developed to study MMR offering key insights into how MMR functions in bacteria using an MMR pathway more widely conserved across biology[20, 26].

Overview of the MMR reaction in *Bacillus* species

With the exception of MutS and MutL, the remaining core MMR proteins in *Bacillus* are different from *E. coli*. Furthermore even the biochemistry of MutS and MutL shows several key differences. In Table 1, we list the core set of proteins known to be MMR components in *Bacillus* compiled from recent work on *B. subtilis* and *B. anthracis* [27, 28, 22, 29-34]. Like most bacteria and all eukaryotes, *Bacillus* species lack MutH and Dam methylase orthologs. The MMR reaction has not been reconstituted *in vitro* for a *Bacillus* species, however based on reconstitution of the MMR reaction in *E. coli*, human, and yeast [35, 1, 36], as well as the information

known about *Bacillus* MMR components, we can make some predictions about the potential mechanism.

In the first step, MutS is targeted to identify mispairs through interaction with the replication processivity clamp DnaN (β -clamp) [28, 30, 37]. Following mismatch identification, MutS recruits MutL to mismatch-proximal DNA [38]. MutL is then licensed through interaction with DnaN (β -clamp) to nick the nascent strand [22, 29]. The mechanism that orients MutL to the nascent strand is unknown; however, we expect MutL endonuclease activity is preferential for a strand with a pre-existing nick, as has been shown for the eukaryotic MutL homologs [21, 39, 40]. The source of the pre-existing strand nick is also not known, although removal of misincorporated ribonucleotides on the leading strand and discontinuities provided by Okazaki fragments on the lagging strand may play a role [34]. After MutL-mediated incision, the WalJ exonuclease is responsible for excision in the 5' to 3' direction [32]. Whether or not a helicase plays an analogous role to *E. coli* UvrD is unclear, but the RecD2 helicase provides a major role in MMR for *Bacillus anthracis* [41, 33]. In the final two steps, we speculate that PolC or DnaE is responsible for resynthesis, while we expect DNA ligase to seal the phosphodiester backbone to complete repair. Below, we describe in more detail how studies within *Bacillus* have led to a better understanding of MMR mechanisms in MutH-independent organisms.

Coupling mismatch repair to DNA replication in *B. subtilis*

Despite the extensive studies that elucidate the biochemical mechanisms of mismatch detection by MutS homologs 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 every 30,000,000 bps synthesized [42]. Clearly, a dynamic chromosome further complicates mismatch repair by obscuring mismatch detection through DNA supercoiling, chromosome compaction, and myriad of DNA-binding proteins associated with the nucleoid. Furthermore, mismatch repair assemblies must compete with active processes such as transcription that could potentially disrupt or impair MMR efficiency. To

circumvent these challenges, several *in vivo* studies suggest that MMR is coupled to DNA replication in both bacteria and eukaryotes. In this review, we focus on the bacterial studies and direct readers interested in how MMR is coupled to DNA replication in eukaryotes to the following recent reviews [43, 3].

A major source of coupling between MMR and DNA replication is through DnaN (β -clamp) and PCNA (Proliferating Cell Nuclear Antigen): the bacterial and eukaryotic processivity clamps, respectively. DnaN (β -clamp) is loaded at DNA termini by clamp loader complexes and primarily serves to tether replicative polymerases to DNA, increasing the overall processivity of the replication process [44]. Moreover, DnaN (β -clamp) acts as an organization scaffold for the replisome, therefore many proteins including MutS are recruited to nascent DNA via interactions with DnaN (β -clamp) [45, 30, 28]. 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 (see for review [46]). Various proteins have a dedicated clamp-binding motif (QL[SD]LF) that is critical for binding β -clamp [47].

The interaction between MutS and processivity clamp is direct but weak, and it occurs through the clamp-binding motif found within the unstructured C-terminal domain of bacterial MutS [48, 30]. Disruption of the clamp-binding motif or deletion of the MutS C-terminal domain is sufficient to eliminate binding to DnaN (β -clamp) *in vitro* [48, 30]. In *B. subtilis*, a clamp-binding motif mutant (MutS-⁸⁰⁶AAAAA⁸¹⁰) at its native locus produces a minor increase in mutation rate [30]. 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 of MMR on yet-to-be identified binding interfaces within this domain [49, 28]. The resulting increase in mutation rate is completely suppressed by ectopic expression of *mutS800*. This effect reveals that mismatch detection depends on the DnaN•MutS interaction at native MutS concentrations (~ 100 nM *in vivo*), and perhaps explains why disruption of the motif in *E. coli*, *B. subtilis*, or *Pseudomonas aeruginosa* fails to produce elevated rates of mutagenesis at high expression levels (≥ 500 nM) [49, 28, 50]. In addition, unlike the *B. subtilis* protein, *E. coli* MutS800 is less stable *in vivo*, which likely contributes to the phenotype associated with this variant [49].

The important role of processivity clamps during *in vivo* mismatch detection was elucidated in part through cytological studies. In bacteria, DnaN (β -clamp) forms large clamp assemblies termed clamp-zones, which immediately follow the progressing replication forks [51, 28, 52] (**Figure 2A**). In *B. subtilis*, clamp zones have been shown to contain 30-60 DNA-bound DnaN clamp dimers as clamp loading and unloading rates reach an equilibrium [51]. Importantly, DnaN (β -clamp) clamps are unevenly distributed in favor of the lagging strand due to Okazaki fragment synthesis and maturation; this asymmetry may facilitate the increased MMR efficiency that has been observed on the lagging strand, presumably due to the higher occurrence of 5' and 3' termini of the Okazaki fragments [53]. Initially, MutS is targeted to nascent DNA via its interaction with DnaN (β -clamp), allowing MutS to maintain a critical spatial and temporal relationship with replication forks [28]. MutS targeting seems to occur at a step prior to mismatch detection because MutSF30A, a mutant defective for mismatch detection, localizes to the clamp zone in a subset of cells [28]. MutS localization is speculated to optimize 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 newly replicated DNA, 90% of MMR capacity is lost, emphasizing the importance of a spatio-temporal coupling between mismatch detection and the replication forks on MMR in Gram-positive bacteria [28].

Studies demonstrating the importance of MutS and/or MutL interaction with DnaN (β -clamp) are not limited to *B. subtilis*. The pathogen *Bacillus anthracis* has two *dnaN* genes [37]. Simultaneous deletion of both genes is lethal, while deletion of just *dnaNI* results in a complete MMR deficiency [37]. This study demonstrates that the role of DnaN (β -clamp) is upstream of DNA synthesis because both *dnaN* genes function in DNA replication but only *dnaNI* functions in MMR. The existence of two *dnaN* genes was identified in several *Bacillus* species including *B. cereus* and *B. thuringiensis* [37]. The study of *dnaNI* in *B. anthracis* provides strong genetic evidence that *dnaN* plays a critical role at a step in MMR preceding DNA synthesis in Gram-positive bacteria or at least in *Firmicutes*. Based on studies of DnaN (β -clamp) in *B. subtilis*, we speculate that the role of *dnaNI* in MMR is to help target the MutS search for mismatches to the replisome and to license MutL-incision after mismatch detection (see below).

MutL recruitment to a mismatch

Upon mismatch detection by MutS, MutL is recruited forming a DNA-bound binary complex [54]. Cytological studies of a functional *B. subtilis mutS-gfp* showed that immediately preceding MutL recruitment, MutS-GFP forms foci in response to mismatch detection [28, 30, 55]. The number of MutS-GFP molecules in a focus was quantified and determined to be ~8 MutS dimers in a focus, suggesting that numerous MutS dimers are involved in active repair within a living cell [38]. This observation suggested that MutS loads iteratively at a single mismatch, producing a transient increase in the local DNA-bound MutS concentration resulting in the efficient recruitment of MutL (**Figure 2B**). As mentioned above, MutS homologs have been shown to form a hydrolysis-independent sliding clamp after mismatch recognition [7, 8]. The basic model stipulates that multiple MutS dimers can load at a mismatch and then diffuse along the DNA in search of MutL and/or strand discrimination features [7, 8]. The cell biological findings of MutS-GFP foci are consistent with the biochemical observations of repetitive loading of *E. coli* MutS and sliding clamp conversion *in vitro* [7]. When the downstream MutL recruitment step is lost, MutS repair centers became considerably brighter, demonstrating a continued loading (quantified as ≤ 30 dimers) of MutS on mismatch-proximal DNA [38]. Exacerbated MutS loading 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 [38].

Two possible mechanisms can explain MutL recruitment by iterative loading of MutS: either MutL binds MutS upon sliding clamp formation, or MutS binds MutL only when bound to a mismatch, forming a tertiary complex. Surface Plasmon Resonance (SPR) studies found that MutS and MutL form stable complexes on a substrate containing both a G/T mismatch and blocked ends in the presence of the slow hydrolysable ATP analog, ATP γ S [7, 56]. Using EMSA, the MutS•MutL binary complex was refractory to an excess of G/T competitor DNA when incubated with ATP γ S, but not when ATP was used [7]. These data supports a model where MutS•MutL interact on DNA, despite sliding clamp formation by MutS. 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, an event which has been observed to last ~3s *in vitro* [56, 57]. Direct *in vitro* imaging failed to observe an association of MutL α with the MutS α sliding clamp after its departure from the mismatch [57], suggesting that MutL homologs must bind MutS homologs while simultaneously bound to a mismatch [57]. Regardless of the specific model, iterative loading would increase the opportunity for MutL to find and bind MutS by increasing the duration of the mismatch-proximal occupancy by MutS.

MutL recruitment is orchestrated through a physical association with MutS [10, 11, 38]. In *E. coli*, a tandem glutamine-binding site was found in the connector domain of MutS to be important for MutL binding [11, 10]. Strikingly, these *E. coli* residues are not conserved in Gram-positive bacteria [38, 11]. To determine the MutL binding site on *B. subtilis* MutS, a peptide array screening approach identified a highly conserved motif found in the core domain (Phe 319, Phe 320) of *B. subtilis* MutS. Disruption of this sequence eliminates both interaction *in vitro* and repair *in vivo* [38]. The FF motif is conserved across many Gram-positive bacteria, including a number of nosocomial pathogens [38]. In *E. coli* and *B. subtilis*, the identified MutL binding interface is found on each face of the MutS dimer, supporting interactions that may occur whether diffusing along the DNA either 5' or 3' to the mismatch [38]. It is attractive to hypothesize that the MutL-binding interfaces of MutS 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 repair events. In support of this hypothesis, ATP binding and sliding clamp formation induce a conformational state more competent for MutL interaction and recruitment *in vitro* [58].

MutL endonuclease activity and activation

In organisms with methylation-independent MMR, a highly conserved endonuclease active site is found within the C-terminal domain of MutL homologs [21]. Discovery of the endonuclease activity in human MutL α subunit PMS2 explained how EXO1, a 5'-3' exonuclease important for human MMR, could paradoxically perform bidirectional excision of a mismatch initiated from a nick either 5' or 3' to the mismatch

281 *in vitro* [21]. During repair, MutL α endonuclease activity is directed along the strand
282 containing the pre-existing nick, with the goal of nicking upstream (5') of the mismatch.
283 This provides a 5' loading site for EXO1, which can then mediate digestion of the
284 mismatch containing strand, allowing for bidirectional repair [21].

285 A conserved metal-binding motif DQHA(X)₂E(X)₄E located in the C-terminal
286 domain of human PMS2 is necessary for the endonuclease activity of MutL α [21, 22]. In
287 *B. subtilis*, this motif is identical to the human PMS2 homolog and its integrity is
288 important for endonuclease activity *in vitro* and for MMR function *in vivo* [22]. The
289 presence of the MutL endonuclease active site is a hallmark of MutH-independent MMR
290 in bacteria. *In vitro*, the endonuclease activity of PMS2 and *B. subtilis* MutL can be
291 recapitulated in the absence of a mismatch or in the absence of other mismatch repair
292 proteins when MutL is bound to Mn²⁺ [21, 22]. The crystal structure of the C-terminal
293 domain of *B. subtilis* MutL revealed an elongated endonuclease site that can
294 accommodate at least two distinct metal ions. The presumed catalytic site is defined by
295 the N-terminal end of the ⁴⁶²DQHA(X)₂E(X)₄E motif, in particular by the Asp462 and
296 His464 residues [22]. A second metal-binding site is defined by Glu468 in the
297 ⁴⁶²DQHA(X)₂E(X)₄E motif, and there are additional conserved motifs, ⁶⁰⁴CPHGRP and
298 ⁵⁷²SCK [40]. In eukaryotic MutL α , a third conserved motif at the C-terminus of MLH1
299 (⁷⁶⁶FERC) also assists with metal coordination [59]. This second metal-binding site
300 coordinates a Zn²⁺ metal ion and is presumed to stabilize the catalytic conformation of
301 the domain [40]. Human PMS2 [22] and yeast PMS1 [60] are also Zn²⁺-metal binding
302 proteins, reinforcing the idea that Zn²⁺-binding at the C-terminal domain of MutL is a
303 universal requirement for endonuclease function. Although Zn²⁺ does not support DNA
304 nicking in the absence of Mn²⁺, it does stimulate the Mn²⁺-dependent DNA nicking
305 activity of *B. subtilis* MutL, and disruption of Zn²⁺ coordination eliminates mismatch
306 repair *in vivo* [22].

307 Regulation of the MutL endonuclease activity is important because aberrant
308 nicking could result in decreased genome stability and the potential for loss in viability.
309 MutL-dependent incisions are regulated in part by the intrinsically weak nature of the
310 endonuclease activity, which suggests a licensing event that must occur to activate MutL
311 at the proper time [22]. A view of the electrostatic surface potential surrounding the

active site reveals a sheath of negatively charged residues that guard against DNA docking within the site until licensing occurs [22].

Other protein factors may further license MutL endonuclease activity. In *B. subtilis*, the C-terminal domain of MutL contains a conserved, albeit degenerate, DnaN clamp binding motif (⁴⁸⁷QEMIV). Disruption of this motif eliminates MutL•DnaN interaction *in vitro*, and abrogates mismatch repair *in vivo* [22, 29]. Importantly, in *E. coli*, disruption of the analogous motif (⁴⁸²QPLLI) in plasmid-borne *mutL* failed to produce a significant increase in mutation rate, suggesting that organisms employing methyl-directed repair rely less on the MutL•DnaN (β -clamp) interaction than organisms employing a MutL-contained endonuclease activity [29]. 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 aiding in strand orientation [40]. One idea is that strand discrimination could be determined *in vitro* by the loaded orientation and tilt of PCNA and DnaN (β -clamp) [40]. Overall, the intrinsic activities of both MutL and the processivity clamp can regulate endonuclease nicking. Although it seems clear that the processivity clamp provides a significant role in regulating MutL incision, the mechanism of regulation is not known for the bacterial proteins. Also, it is unclear if the bacterial MutL, which contains two endonuclease active sites per dimer, is regulated differently from the eukaryotic MutL homolog, which contains one active site per dimer. In Figure 3, we present a model for the early steps of MMR and endonuclease activation.

Strand discrimination signals

For organisms lacking a methylation signal and the Dam/MutH proteins, the process of strand discrimination is still unknown. Some recent evidence suggests that removal of misincorporated ribonucleotide monophosphates (rNMPs) from DNA could provide a signal to direct MMR to the nascent strand. We discuss this possibility below.

Early studies of the Gram-positive bacterium *Streptococcus pneumoniae* by Sanford Lacks and coworkers suggested that nicks present in the DNA strands could be used to direct MMR in organisms lacking methylation direction [61]. The question then

became, what process(es) preferentially generates strand nicks in the newly synthesized strand serving to direct mismatch repair? One possibility is that nicks in the newly synthesized strand result from the removal of misincorporated ribonucleotides during DNA replication. During genome replication, rNTPs (ribonucleoside triphosphates) are occasionally used and incorporated in place of their corresponding dNTPs (deoxyribonucleoside triphosphates) [62, 34]. Ribonucleotide misincorporation is largely a consequence of rNTP intracellular concentrations ranging from 10-100 fold greater than those of their dNTP counterparts [63]. *In vitro*, the rate of rNTP incorporation by the *E. coli* Pol III holoenzyme is 1 rNMP for every 2.3 kb of replicated DNA, which *in vivo* would correspond to ~2,000 rNMPs embedded per chromosome per replication cycle [34]. dNMPs are less reactive and subsequently more stable than rNMPs, which when incorporated into DNA, produce both single stranded and deleterious double stranded breaks via alkali-promoted transesterification [64]. Furthermore, rNTPs and template rNMPs slow the rate of progression of the replicative polymerase *in vitro* [34]. Despite these potential problems, rNMP incorporation could provide a benefit to the cell: their removal would produce a nick in the newly replicated strand, possibly marking the strand for MMR [65, 66].

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 [67]. Preceding further processing, these RER-dependent nicks have the potential to be used by the MMR pathway to direct repair MutL incision to the newly replicated strand. In *B. subtilis*, deletion of *rnhB* causes a ~2.4 fold increase in mutation rate (compared to a 60 fold increase for the Δ mutSL strain) [34]. The Δ rnhB elevated mutation rate is further increased when combined with a null mutation in another RNase H enzyme, the RNase HIII (RnhC) [34]. The resulting *rnhB/rnhC* pair caused a 5-fold increase in the spontaneous mutation rate, but also showed a striking decrease in growth rate and viability [34]. Importantly, deletion of *rnhB* in *E. coli*, the only RNase HII enzyme, fails to increase the mutation rate, arguing that if RER has a minor role in contributing to strand discrimination, it does so in bacteria lacking a methyl-directed pathway [34]. Moving forward, it will be important to reconstitute MMR *in vitro* and to determine if a single rNMP can direct repair using the

bacterial proteins. It will also be important to perform a comprehensive analysis of the contribution of RER to the spontaneous mutation rate, as well as determine the mutation spectrum and the molecular basis for mutagenesis observed in the absence of RNase HIII in *B. subtilis*.

Strand removal and resynthesis

Upon mismatch detection, the newly replicated strand containing the incorrect nucleotide needs to be removed [1]. This process is initiated after an endonuclease nicks the newly replicated strand in the vicinity of the mismatch. The nicked strand is either unwound or displaced towards the mismatch, allowing an exonuclease to degrade the mismatch containing strand [16, 68]. In *E. coli* and other MutH and Dam containing bacteria, the UvrD helicase is required to unwind the mismatch-containing strand, followed by strand degradation by one of several redundant exonucleases (see for review [17]). In *Bacillus anthracis*, a *uvrD* mutant failed to produce an elevated spontaneous mutation frequency arguing against a role for UvrD in MMR for *B. anthracis* [31, 33]. Interestingly, in *B. anthracis*, disruption of *recD2*, a member of the same Superfamily 1 family helicases as UvrD, resulted in a striking increase in mutagenesis consistent with a loss of MMR [31, 33]. Despite an extensive phylogenetic distribution of *recD2* among bacteria, *B. subtilis* RecD2 functions outside of MMR, engaging in the repair of other types of DNA damage [41]. Therefore, the extent to which RecD2 helicase participates in bacterial MMR requires further study and appears to be organism specific.

Further study is also needed to determine whether a dedicated helicase exists in other MutH or Dam-independent bacterial species. One possibility is that bacteria lacking MutH or Dam contain MMR systems more closely resemble the machinery found in eukaryotes where a helicase has yet to be identified to provide a role in strand removal. For mismatch removal, eukaryotes either rely on direct strand digestion by EXO1, strand displacement by polymerase δ , or 3' to 5' exonuclease activity likely performed by polymerase δ (see for review [43]). In bacteria, one possibility is that a newly discovered exonuclease, YycJ (WalJ), functions in an analogous role as eukaryotic EXO1. Disruption of *walJ* produces a spontaneous mutation rate equivalent to a MMR-deficient strain in both *B. anthracis* and *B. subtilis* [32]. Biochemical studies of WalJ show that it

is a Mn²⁺ dependent 5'-3' exonuclease [32] (**Figure 2D**). Interestingly, deletion of *walJ* also causes sensitivity to beta lactam antibiotics, suggesting a role for WalJ in cell wall homeostasis, as the name suggests [69]. Further analysis of *walJ* will be required to understand the role of this protein in MMR and in cell wall related functions in Gram-positive bacteria.

Conclusions

Since the development of *Bacillus* as an experimental system for understanding MMR, much has been discovered concerning the mechanism of MMR in bacteria lacking MutH and Dam methylase. Experiments in *B. subtilis* and *B. anthracis* demonstrate the importance of DnaN (β -clamp) during the early steps in MMR including targeting of MutS to mismatches [30, 37], interaction with MutL [29] and spatiotemporal coupling of MMR to DNA replication [28]. Crystalization of the C-terminal domain of *B. subtilis* MutL yielded the first MutL endonuclease domain structure providing important implications for the regulation of MutL endonuclease activity in bacteria [22]. The search for mutator alleles in *B. anthracis* identified a helicase (RecD2) and exonuclease (WalJ) important for MMR [31-33].

Though many new discoveries have been made over the last 10 years, important gaps remain in our understanding of MMR in MutH-less bacteria. Further study is needed to address questions including: What mechanism(s) activate and regulate MutL endonuclease activity? Does MutL identify strand discrimination signals and do ribonucleotides contribute to strand identification? Upon strand nicking, how is WalJ regulated and loaded to direct excision? Advances in these areas will stimulate new and exciting discoveries and uncover the mechanistic steps of an MMR process that is likely far more representative among bacteria than the more-well known model *E. coli*.

Acknowledgements

This review summarizes current advances in understanding MMR in Gram-positive bacteria. Our research would not be possible without the work of the many dedicated scientists studying MMR across bacteria and we apologize for citations we were unable

to include due to space limitations. This work was supported by National Science Foundation grant MCB1050948 to LAS, by a Burroughs Wellcome Fund Career Award at the Scientific Interface to JSB and by the Natural Sciences and Engineering Research Council to AG. JSL was supported in part by a Rackham Dissertation Award.

Figure Legends

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 d(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 d(GATC) sites, and upon activation, MutH nicks the unmethylated strand. (C) Mismatch repair occurs on both the leading and lagging strands. In (C) we show correction on the leading strand where MutL recruits the UvrD helicase to the nick, where it unwinds the unmethylated strand through the mismatch. Simultaneously, exonucleases degrade the strand. (D) SSB bound ssDNA surrounding the recently removed mismatch is replicated by the Pol III holoenzyme. Ligase seals the remaining nick, completing repair.

Figure 2. Model of mismatch repair in MutH-independent systems

(A) Replication coupling of mismatch detection is facilitated by a DnaN clamp zone. (B) Upon mismatch detection, the MutS ATPase cycle 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 at a MutL processed nick 5' to the mismatch. WalJ degrades the mismatch-containing strand in the 5' - 3' direction. After strand removal by WalJ, error-free replication corrects the error.

Figure 3. Model for MutL Recruitment and Endonuclease activation in most bacteria

(A) MutS is targeted to search DNA through interaction with the β -clamp (DnaN). **(B)** Following mismatch recognition, MutS recruits MutL to the mismatch and converts into a sliding clamp that searches for a strand discrimination signal. **(C)** MutL interacts with the β -clamp (DnaN), likely positioning and stimulating the endonuclease activity of MutL to incise the DNA.

472 **Table 1. Comparison between the core MMR proteins in *E. coli* and *Bacillus* species**

<i>E. coli</i>	<i>Bacillus</i>	Function
MutS	MutS	Mismatch recognition protein
β -clamp	β -clamp	β -clamp (DnaN) is the replication processivity factor that binds MutS and MutL and aids in MMR.
MutL	MutL (endonuclease)	MutL is important for coordinating protein interactions during repair. In most bacteria, except <i>E. coli</i> and closely related gamma proteobacteria MutL has a conserved endonuclease active site.
MutH	None	MutH is a methyl-directed endonuclease. In <i>Bacillus</i> and most organisms, MutL provides endonuclease activity.
Dam	None	Methylates adenosine (m6A) to discriminate between template and nascent strand during repair.
UvrD	RecD2/none	RecD2 is critical for mismatch repair in <i>B. anthracis</i> , but not in <i>B. subtilis</i> .
RecJ, Exo I, VII and X	WalJ	Exonucleases are involved in excision of mismatched strand. WalJ is a 5' to 3' exonuclease.

References

- [1] Lahue RS, Au KG, Modrich P. DNA mismatch correction in a defined system. *Science* 1989;245:160-4.
- [2] Jiricny J. Replication errors: cha(lle)nging the genome. *EMBO J* 1998;17:6427-36.
- [3] Jiricny J. Postreplicative mismatch repair. *Cold Spring Harbor perspectives in biology* 2013;5:a012633.
- [4] Su SS, Modrich P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proc Natl Acad Sci U S A* 1986;83:5057-61.
- [5] Li GM. New insights and challenges in mismatch repair: getting over the chromatin hurdle. *DNA Repair (Amst)* 2014;19:48-54.
- [6] Allen DJ, Makhov A, Grilley M, Taylor J, Thresher R, Modrich P, et al. MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J* 1997;16:4467-76.
- [7] Acharya S, Foster PL, Brooks P, Fishel R. The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Mol Cell* 2003;12:233-46.
- [8] Gradia S, Subramanian D, Wilson T, Acharya S, Makhov A, Griffith J, et al. hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Molecular Cell* 1999;3:255-61.
- [9] Junop MS, 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. *Mol Cell* 2001;7:1-12.
- [10] Winkler I, Marx AD, Lariviere D, Heinze RJ, Cristovao M, Reumer A, et al. Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in Escherichia coli. *The Journal of Biological Chemistry* 2011;286:17326-37.
- [11] Mendillo ML, Hargreaves VV, Jamison JW, Mo AO, Li S, Putnam CD, 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* 2009;106:22223-8.
- [12] Waldminghaus T, Weigel C, Skarstad K. Replication fork movement and methylation govern SeqA binding to the Escherichia coli chromosome. *Nucleic Acids Res* 2012;40:5465-76.
- [13] Spampinato C, Modrich P. The MutL ATPase is required for mismatch repair. *J Biol Chem* 2000;275:9863-9.
- [14] Glickman BW, Radman M. Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. *Proc Natl Acad Sci U S A* 1980;77:1063-7.
- [15] Dao V, Modrich P. Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. *J Biol Chem* 1998;273:9202-7.
- [16] Burdett V, Baitinger C, Viswanathan M, Lovett ST, 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* 2001;98:6765-70.

- [17] Schofield MJ, Hsieh P. DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* 2003;57:579-608.
- [18] Barras F, Marinus MG. The great GATC: DNA methylation in *E. coli*. *Trends Genet* 1989;5:139-43.
- [19] Campbell JL, Kleckner N. *E. coli* oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* 1990;62:967-79.
- [20] Lenhart JS, Schroeder JW, Walsh BW, Simmons LA. DNA Repair and Genome Maintenance in *Bacillus subtilis*. *Microbiology and molecular biology reviews : MMBR* 2012;76:530-64.
- [21] Kadyrov FA, Dzantiev L, Constantin N, Modrich P. Endonucleolytic function of MutL α in human mismatch repair. *Cell* 2006;126:297-308.
- [22] Pillon MC, Lorenowicz JJ, Uckelmann M, Klocko AD, Mitchell RR, Chung YS, et al. Structure of the endonuclease domain of MutL: unlicensed to cut. *Mol Cell* 2010;39:145-51.
- [23] Correa EM, De Tullio L, Velez PS, Martina MA, Argarana CE, Barra JL. Analysis of DNA structure and sequence requirements for *Pseudomonas aeruginosa* MutL endonuclease activity. *J Biochem* 2013;154:505-11.
- [24] Cox EC, Degnen GE, Scheppe ML. Mutator gene studies in *Escherichia coli*: the mutS gene. *Genetics* 1972;72:551-67.
- [25] Treffers HP, Spinelli V, Belser NO. A Factor (or Mutator Gene) Influencing Mutation Rates in *Escherichia Coli*. *Proc Natl Acad Sci U S A* 1954;40:1064-71.
- [26] Claverys JP, Lacks SA. Heteroduplex deoxyribonucleic acid base mismatch repair in bacteria. *Microbiol Rev* 1986;50:133-65.
- [27] Dupes NM, Walsh BW, Klocko AD, Lenhart JS, Peterson HL, Gessert DA, et al. Mutations in the *Bacillus subtilis* beta clamp that separate its roles in DNA replication from mismatch repair. *J Bacteriol* 2010;192:3452-63.
- [28] Lenhart JS, Sharma A, Hingorani MM, Simmons LA. DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular microbiology* 2013;87:553-68.
- [29] Pillon MC, Miller JH, Guarne A. The endonuclease domain of MutL interacts with the beta sliding clamp. *DNA Repair (Amst)* 2011;10:87-93.
- [30] Simmons LA, Davies BW, Grossman AD, Walker GC. Beta clamp directs localization of mismatch repair in *Bacillus subtilis*. *Mol Cell* 2008;29:291-301.
- [31] Yang H, Sikavi C, Tran K, McGillivray SM, Nizet V, Yung M, et al. Papillation in *Bacillus anthracis* colonies: a tool for finding new mutators. *Mol Microbiol* 2011;79:1276-93.
- [32] Yang H, Yung M, Li L, Hoch JA, Ryan CM, Kar UK, et al. Evidence that YycJ is a novel 5'-3' double-stranded DNA exonuclease acting in *Bacillus anthracis* mismatch repair. *DNA repair* 2013;12:334-46.
- [33] Yang H, Yung M, Sikavi C, Miller JH. The role of *Bacillus anthracis* RecD2 helicase in DNA mismatch repair. *DNA Repair (Amst)* 2011;10:1121-30.
- [34] Yao NY, Schroeder JW, Yurieva O, Simmons LA, O'Donnell ME. Cost of rNTP/dNTP pool imbalance at the replication fork. *Proceedings of the National Academy of Sciences of the United States of America* 2013;110:12942-7.

- [35] Bowen N, Smith CE, Srivatsan A, Willcox S, Griffith JD, Kolodner RD. Reconstitution of long and short patch mismatch repair reactions using *Saccharomyces cerevisiae* proteins. *Proc Natl Acad Sci U S A* 2013;110:18472-7.
- [36] Dzantiev L, Constantin N, Genschel J, Iyer RR, Burgers PM, Modrich P. A defined human system that supports bidirectional mismatch-provoked excision. *Mol Cell* 2004;15:31-41.
- [37] Yang H, Miller JH. Deletion of *dnaN1* generates a mutator phenotype in *Bacillus anthracis*. *DNA Repair (Amst)* 2008;7:507-14.
- [38] Lenhart JS, Pillon MC, Guarne A, Simmons LA. Trapping and visualizing intermediate steps in the mismatch repair pathway in vivo. *Molecular microbiology* 2013;90:680-98.
- [39] Kadyrov FA, Holmes SF, Arana ME, Lukianova OA, O'Donnell M, Kunkel TA, et al. *Saccharomyces cerevisiae* MutLalpha is a mismatch repair endonuclease. *J Biol Chem* 2007;282:37181-90.
- [40] Pluciennik A, Dzantiev L, Iyer RR, Constantin N, Kadyrov FA, Modrich P. PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc Natl Acad Sci U S A* 2010;107:16066-71.
- [41] Walsh BW, Bolz SA, Wessel SR, Schroeder JW, Keck JL, Simmons LA. RecD2 helicase limits replication fork stress in *Bacillus subtilis*. *J Bacteriol* 2014;196:1359-68.
- [42] Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci U S A* 2012;109:E2774-83.
- [43] Reyes GX, Schmidt TT, Kolodner RD, Hombauer H. New insights into the mechanism of DNA mismatch repair. *Chromosoma* 2015.
- [44] Huang CC, Hearst JE, Alberts BM. 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* 1981;256:4087-94.
- [45] Hombauer H, Campbell CS, Smith CE, Desai A, Kolodner RD. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 2011;147:1040-53.
- [46] Sutton MD. Coordinating DNA polymerase traffic during high and low fidelity synthesis. *Biochim Biophys Acta* 2009;1804:1167-79.
- [47] Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA. A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc. Natl. Acad. Sci. U.S.A.* 2001;98:11627-32.
- [48] Lopez de Saro FJ, Marinus MG, Modrich P, O'Donnell M. The beta sliding clamp binds to multiple sites within MutL and MutS. *J Biol Chem* 2006;281:14340-9.
- [49] Calmann MA, Nowosielska A, Marinus MG. The MutS C terminus is essential for mismatch repair activity in vivo. *J Bacteriol* 2005;187:6577-9.
- [50] Monti MR, Miguel V, Borgogno MV, Argarana CE. Functional analysis of the interaction between the mismatch repair protein MutS and the replication processivity factor beta clamp in *Pseudomonas aeruginosa*. *DNA Repair (Amst)* 2012;11:463-9.
- [51] Su'etsugu M, Errington J. The replicase sliding clamp dynamically accumulates behind progressing replication forks in *Bacillus subtilis* cells. *Mol Cell* 2011;41:720-32.

- [52] Moolman MC, Krishnan ST, Kerssemakers JW, van den Berg A, Tulinski P, Depken M, et al. Slow unloading leads to DNA-bound beta2-sliding clamp accumulation in live *Escherichia coli* cells. *Nat Commun* 2014;5:5820.
- [53] Pavlov YI, Mian IM, Kunkel TA. Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol* 2003;13:744-8.
- [54] Galio L, Bouquet C, Brooks P. ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. *Nucleic Acids Res* 1999;27:2325-31.
- [55] Smith BT, Grossman AD, Walker GC. Visualization of mismatch repair in bacterial cells. *Mol. Cell* 2001;8:1197-206.
- [56] Mendillo ML, Mazur DJ, Kolodner RD. 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* 2005;280:22245-57.
- [57] Gorman J, Wang F, Redding S, Plys AJ, Fazio T, Wind S, 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* 2012;109:E3074-83.
- [58] Lebbink JH, Fish A, Reumer A, Natrajan G, Winterwerp HH, Sixma TK. Magnesium coordination controls the molecular switch function of DNA mismatch repair protein MutS. *The Journal of Biological Chemistry* 2010;285:13131-41.
- [59] Kosinski J, Plotz G, Guarne A, Bujnicki JM, 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* 2008;382:610-27.
- [60] Gueneau E, Dherin C, Legrand P, Tellier-Lebegue C, Gilquin B, Bonnesoeur P, et al. Structure of the MutLalpha C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. *Nat Struct Mol Biol* 2013;20:461-8.
- [61] Lacks SA, Dunn JJ, Greenberg B. Identification of base mismatches recognized by the heteroduplex-DNA-repair system of *Streptococcus pneumoniae*. *Cell* 1982;31:327-36.
- [62] Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundstrom EB, Burgers PM, et al. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107:4949-54.
- [63] Buckstein MH, He J, Rubin H. Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *Journal of Bacteriology* 2008;190:718-26.
- [64] Li Y, Breaker RR. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2' hydroxyl group. *J. Am. Chem. Soc* 1999;121:5364-72.
- [65] Ghodgaonkar MM, Lazzaro F, Olivera-Pimentel M, Artola-Boran M, Cejka P, Reijns MA, et al. Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell* 2013;50:323-32.
- [66] Lujan SA, Williams JS, Clausen AR, Clark AB, Kunkel TA. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Molecular Cell* 2013;50:437-43.
- [67] Ohtani N, Haruki M, Morikawa M, Crouch RJ, Itaya M, Kanaya S. Identification of the genes encoding Mn²⁺-dependent RNase HII and Mg²⁺-dependent RNase HIII from *Bacillus subtilis*: classification of RNases H into three families. *Biochemistry* 1999;38:605-18.

- [68] Viswanathan M, Burdett V, Baitinger C, Modrich P, Lovett ST. Redundant exonuclease involvement in *Escherichia coli* methyl-directed mismatch repair. *J Biol Chem* 2001;276:31053-8.
- [69] Biller SJ, Wayne KJ, Winkler ME, Burkholder WF. The putative hydrolase YycJ (WalJ) affects the coordination of cell division with DNA replication in *Bacillus subtilis* and may play a conserved role in cell wall metabolism. *J Bacteriol* 2011;193:896-908.

Figure 1

Figure 2

A

Mismatch detection

B

MutL recruitment
(MutS imposed mismatch-dependency)

C

β-clamp stimulates DNA nicking of MutL

Figure 3.