Binding of the regulatory domain of MutL to the sliding β-clamp is species specific

Ahmad W. Almawi[1], Michelle K. Scotland[2,3,8], Justin R. Randall[5], Linda Liu[1], Heather K. Martin[5], Lauralicia Sacre[6], Yao Shen[6], Monica C. Pillon[1,7], Lyle A. Simmons[5], Mark D. Sutton[2,3,4], Alba Guarné[1,6*]

[1]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON, Canada. [2]Department of Biochemistry and [3]Witebsky Center for Microbial Pathogenesis and Immunology, and [4]Genetics, Genomics and Bioinformatics Program. The Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY, USA. [5]Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor MI, USA. [6]Department of Biochemistry, McGill University, Montreal, QC, Canada. [7]Current address: National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, NC, USA. [8]Current address: Cornell University, Ithaca, NY, USA.

[*]Correspondence to: alba.guarne@mcgill.ca

Running title: Crystal structures of the regulatory domain of MutL bound to the β-clamp

Abstract

The β-clamp is a protein hub central to DNA replication and fork management. Proteins interacting with the β-clamp harbor a conserved clamp-binding motif that is often found in extended regions. Therefore, clamp interactions have –almost exclusively– been studied using short peptides recapitulating the binding motif. This approach has revealed the molecular determinants that mediate the binding but cannot describe how proteins with clamp-binding motifs embedded in structured domains are recognized. The mismatch repair protein MutL has an internal clamp-binding motif, but its interaction with the β-clamp has different roles depending on the organism. In *Bacillus subtilis*, the interaction stimulates the endonuclease activity of MutL and it is critical for DNA mismatch repair. Conversely, disrupting the interaction between *Escherichia coli* MutL and the β-clamp only causes a mild mutator phenotype. Here, we determined the structures of the regulatory domains of *E. coli* and *B. subtilis* MutL bound to their respective β-clamps. The structures reveal different binding modes consistent with the binding to the β-clamp being a two-step process. Functional characterization indicates that, within the regulatory domain, only the clamp binding motif is required for the interaction between the two proteins. However, additional motifs beyond the regulatory domain may stabilize the interaction. We propose a model for the activation of the endonuclease activity of MutL in organisms lacking methyl-directed mismatch repair.

Keywords: X-ray crystallography, weak protein-protein interactions, DNA mismatch repair, sliding β-clamp, MutL nuclease activity.

1. INTRODUCTION

The sliding β-clamp, and its eukaryotic counterpart PCNA, are ring-shaped structures that encircle DNA and tether their binding partners to DNA. The main role of these clamps is to increase the processivity of DNA polymerase, but they also have critical roles regulating DNA replication, polymerase switching, and DNA mismatch repair (1-4). Initiation of DNA mismatch repair depends on the coordinated actions of two proteins –MutS and MutL– both of which interact with the β-clamp (5-7). MutS recognizes mismatches and small insertion/deletion loops that have escaped DNA polymerase proofreading, and recruits MutL to mark the newly synthesized strand for repair either directly (methyl-independent repair) or indirectly (methyl-directed repair). *Escherichia coli* (*E. coli*) has a devoted nuclease that recognizes transiently hemimethylated DNA and nicks the unmethylated strand, effectively discriminating the new from the template strand. However, most prokaryotes –and all eukaryotes– lack this gene. Instead, MutL homologs from these organisms harbor their own nuclease activity that is stimulated by the interaction with the β-clamp or its eukaryotic counterpart PCNA (8-11). Accordingly, disruption of the clamp-binding motif in *B. subtilis* MutL causes a strong mutator phenotype, whereas disruption of this motif in *E. coli* MutL –which does not have nuclease activity– causes a less severe mutator phenotype (6,10). Sliding clamps bind DNA with a defined orientation. Therefore, proteins interacting with the β-clamp also bind DNA in a specific orientation. In the case of MutL, this asymmetry determines the strand of the DNA duplex that gets cut (12).

All clamp-binding partners contain a conserved linear motif known as PCNA-interacting protein (PIP) box or clamp-binding motif (CBM) that binds a hydrophobic groove on the sliding clamp (**Figure 1A**). PIP boxes have a strict QxxLxxFF consensus sequence, where the conserved leucine and phenylalanine residues form a 3_{10} helix that defines a "three-forked plug" interaction with PCNA (13). Only one PIP-box lacking the two aromatic residues has been described to date (14), underscoring the importance of these two residues. Clamp-binding motifs are shorter and have a general QL(D/S)LF

consensus sequence (**Supplementary Figure 1**). This consensus sequence tolerates greater variability than the PIP-box consensus, a feature that often challenges the identification of CBMs (15). PIP boxes and CBMs are normally found in flexible terminal regions and, therefore, most structural information comes from complexes of PCNA or the β-clamp bound to short peptides derived from binding partners (16-19). These studies have shown that all clamp-binding partners share a characteristic bidentate interaction, where the conserved glutamine and leucine residues of the motif are bound to adjacent pockets at the C-terminus of the β-clamp. They have also shown how the loops surrounding these two pockets are flexible and close in to define the walls of the groove upon binding of the peptide. These studies, however, fail to explain how the β-clamp accommodates multiple binding partners or how additional surfaces of the β-clamp regulate partner switching (1,2). Studies with PCNA bound to longer fragments of their respective binding partners have resulted in structures that maintain the clamp-binding partners in locked-down, inactive conformations (16,20,21), thereby providing information about the peptide-protein and protein-protein interactions that enable partner switching. However, structural information of active β-clamp or PCNA complexes is still very limited (22).

Compared to terminal CBMs, internal clamp-binding motifs within structured domains are poorly characterized. *E. coli* DNA polymerase V (UmuC) has an internal CBM located between its little finger and C-terminal domains, but since it is found within the linker connecting the two domains its interaction with the β-clamp was studied structurally using a short peptide (17). The α catalytic and ε proofreading subunits of the *E. coli* DNA polymerase III also contain internal CBMs (23-25). Interaction of *E. coli* Pol IIIα with the β-clamp was also studied structurally using a short peptide (18). Therefore, as in previous analyses, these structures do not provide information as to how the surrounding domains in the partner affect its binding to the β-clamp. The clamp-binding motif of MutL is located within a structured region of the endonuclease domain (26). Although the clamp-binding motif of MutL resides in a surface exposed loop of the endonuclease domain, binding to the β-clamp must impose significant

rearrangements to avoid steric hindrance. The interaction between MutL and the sliding β-clamp is conserved from bacteria to humans (11,26), therefore the same binding motif mediates the interaction with either the β-clamp and PCNA (**Figure 1A**). To understand how MutL homologs interact with the β-clamp, we have determined the crystal structures of *B. subtilis* and *E. coli* β-clamp bound to the regulatory domains of their respective MutL homologs. We find that *B. subtilis* MutL interacts with the β-clamp through the characteristic bidentate interaction, whereas *E. coli* MutL forms a monodentate interaction. Comparison of the two structures reveals a trade-off between conformational changes within the regulatory domain and the formation of a bidentate interaction with the β-clamp. We propose a model describing how conformational changes in the regulatory domain of *B. subtilis* MutL upon binding to the β-clamp align the nuclease site with the central cavity of the β-clamp. This model unveils the role of the conserved GQ motif found in MutL homologues with endonuclease activity.

2. MATERIALS AND METHODS

2.1. Design of the clamp-MutL fusion proteins

The fragment of *E. coli* MutL encoding residues 471-574 (regulatory domain, $MutL^{RGD}$) was subcloned in a modified pET15b expression vector containing an N-terminal His_6-tag removable with tobacco etch virus (TEV) protease site using the BamHI/BlpI restriction sites (pAG8902). Full-length *E. coli* β-clamp was subsequently subcloned using the NdeI/BamHI restriction sites to create the *E. coli* clamp-$MutL^{RGD}$ fusion (pAG8903). A SalI restriction site was engineered between the two protein fragments to introduce a glycine/serine-rich linker (SGASG). Oligonucleotides encoding the linker flanked by SalI/BamHI sites were purchased from Integrated DNA Technologies and ligated into to generate the final clamp-SGASG-$MutL^{RGD}$ fusion expression plasmid (pAG8918). The *B. subtilis* fusion expression plasmid (pAG9021) was generated analogously and included full-length β-clamp and residues 482-574 from *B. subtilis* MutL

(MutL^{RGD}) corresponding to its regulatory domain joined by a three-amino acid linker (VDS). The identity of all plasmids was confirmed by DNA sequencing (MOBIX, McMaster University).

2.2. Protein expression and purification

The *E. coli* and *B. subtilis* clamp-MutL^{RGD} fusions were produced in BL21 (DE3) cells supplemented with a plasmid encoding rare tRNAs. Cells were grown in Luria-Bertani media to an OD_{600} ~ 0.7 and protein expression was induced by addition of 1.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation (1,000 x *g* for 15 min) after overnight incubation at 16 °C. Cell pellets were re-suspended in 20 mM TRIS-HCl pH 8, 500 mM NaCl, 1.4 mM β-mercaptoethanol, 5 % glycerol and lysed by sonication. Lysates were clarified by centrifugation at 39,000 x *g* for 40 minutes. The supernatant was loaded onto a Nickel-chelating affinity column (GE Healthcare) pre-equilibrated with 20 mM TRIS-HCl pH 8, 500 mM NaCl, 1.4 mM β-mercaptoethanol, 5 % glycerol. The column was washed with 63 mM imidazole and the His-tag fusions were eluted with 195 mM imidazole. The fractions containing protein were pooled, and the salt concentration was adjusted to 150 mM NaCl prior to loading them onto a Q-Sepharose ion exchange column (GE Healthcare) equilibrated with 20 mM TRIS-HCl pH 7.5, 150 mM NaCl, 1.4 mM β-mercaptoethanol, 5% glycerol. The protein was eluted off the column using a linear gradient to 500 mM NaCl (clamp-MutL^{RGD} fusions elute at ~350 mM NaCl). The protein was further purified through size exclusion chromatography using a Superdex200 (S200) 10/300 GL size-exclusion column (GE Healthcare) equilibrated with 20 mM TRIS-HCl pH 7.5, 150 mM NaCl, 1.4 mM β-mercaptoethanol, 5% glycerol. The eluted samples were concentrated to 20 mg/mL and protein concentration was calculated using the Beer-Lambert equation with an extinction coefficient of 29,450 $M^{-1}cm^{-1}$ and 24,710 $M^{-1}cm^{-1}$ for the *E. coli* and *B. subtilis* clamp-MutL^{RGD} fusions, respectively.

2.3. Crystallization and structure determination

Crystals of the *E. coli* clamp-MutL^{RGD} fusion grew in 100 mM BIS-TRIS pH 5.5, and 2 M ammonium sulfate and were cryo-protected by addition of 8 % glycerol to the mother liquor. Crystals of the *B. subtilis* clamp-MutL^{RGD} fusion grew in 100 mM HEPES pH 7.5, 25 % PEG 3350 (v/v), and 0.2 M ammonium sulfate, were cryo-protected by addition of 12 % ethylene glycol. Complete data sets were collected at the O8ID-1 and O8B1-1 beam lines of the Canadian Light Source and were processed with XDS (**Table 1**) (27). The structures were determined by molecular replacement using the structures of the individual components as search models (PDBs: 4K3M and 1X9Z for the *E. coli* fusion and PDBs: 4TR6 and 3KDK for the *B. subtilis* fusion). The initial models were refined by iterative cycles of manual model building in Coot and refinement in PHENIX (28,29). Quantitative analysis of the interfaces of both structures was done using the online PISA server (30). Figures showing molecular structures were generated using PyMOL.

2.4. Analysis of the mutation frequency of *E. coli* MutL variants

E. coli strains MG1655 and JW4128-1 were obtained from the *E. coli* Genetic Stock Center. The Δ*mutL720*::*kan* allele from strain JW4128-1 was transduced into MG1655 using P1*vir* (31), resulting in strain MKS108. Cultures of strain MKS108 bearing either pET15b or a pET15b derivative expressing the indicated MutL protein were grown at 37^{o} for 12 hours with aeration in LB medium (10 g/l Difco tryptone, 5 g/l Difco yeast extract, 10 g/l NaCl) supplemented with ampicillin (150 μg/ml) and kanamycin (40 μg/ml). Saturated cultures were serially diluted in 0.8% saline and 100 μl of the 10^{-6} dilution ($n \geq 15$) was spread onto LB agar plates to determine culture titers. Three-hundred μl of the same undiluted cultures (n=20) were spread onto LB agar plates supplemented with rifampicin (50 μg/ml) to identify spontaneous mutations. Plates were incubated at 37^{o} for 12 h prior to counting colonies and the spontaneous mutation frequency of each strain was determined by dividing the number of colony forming units on plates containing rifampicin by the average number of colony forming units on LB lacking rifampicin (10,32). The 95% confidence intervals were calculated as described (32).

163

164 **2.5. Analysis of the mutation frequency of *B. subtilis mutL* variants**

165 Each strain was created by integration of each *mutL* mutant into the *amyE* locus of the Δ*mutL* strain of
166 *B. subtilis* for ectopic expression with IPTG. A stock for each strain was frozen and stored at -80°C. The
167 wild type PY79 and isogenic *ΔmutL* strains were struck out on LB agar; while the $mutL^{+}$, $mutL^{HK}$,
168 $mutL^{CBM}$ were struck on LB agar supplemented with spectinomycin (100 µg/ml) and IPTG (1 mM) for
169 overnight growth. Both plates were incubated overnight at 30°C. The following day, individual colonies
170 were picked and cultured in 2 ml LB and 200 µM IPTG rotating at 37°C for 2-3 hours to OD_{600} of 1-1.2.
171 One ml from each tube was also pelleted at 11,000 x g and the supernatant was removed. This pellet was
172 resuspended in 100 µl of 0.85% saline solution and 100 µl was plated on LB agar supplemented with
173 rifampicin (100 µg/ml). The cell suspension was then diluted 10^{-6} and 100 µl of each strain was plated
174 on LB agar to determine the total number of viable cells. Plates were then incubated at 30°C and each
175 plate was scored for colony forming units. LB plates were scored if they contained at least 70 colonies.
176 Mutation frequency analysis was conducted as described previously (33). In total, 11-15 independent
177 cultures were used per strain. The mutation frequency given is relative to wild type PY79 and is consistent
178 with prior studies (26,34).

179

180 **2.6. Cysteine-crosslinking complex formation**

181 Variants of *B. subtilis* MutL and β-clamp were generated as described earlier (10). In brief, the *B. subtilis*
182 β-clamp does not have any surface exposed cysteine residues, therefore we added a Ser379–Cys380
183 dipeptide at the extreme C-terminus of the protein (pAG 8807). The single-Cys variant of the
184 endonuclease domain $MutL^{CTD}$ (pAG 8803; residues 433–627) included point mutations E485C, C531S,
185 C573S and C604S. Both variants were produced and purified as described earlier (10). The single-Cys
186 variant of the regulatory domain of MutL (pAG9148; $MutL^{RGD}$-E485C/ C531S/C573S) was generated

187 by site-directed mutagenesis and produced as described elsewhere (6). To form the complexes between
188 the β-clamp and the endonuclease or regulatory domains of *B. subtilis* MutL, the β-clamp was incubated
189 with either the endonuclease or the regulatory domain at a 1:1 ratio to a final concentration of 20 μM.
190 The samples (1–2 ml) were dialyzed in 20 mM Tris pH 7.6, 150 mM KCl, 10 mM DTT, 10% glycerol
191 for 2 hours at 4°C. The mixture was then dialyzed into buffer supplemented with 5 mM DTT for 1 h,
192 followed by 1 h in dialysis buffer without DTT. The sample was then left in dialysis buffer without DTT.
193 Complex formation was monitored by resolving samples collected at different time points on denaturing
194 polyacrylamide gradient gels stained with Coomassie Brilliant Blue. We quantified the intensities of the
195 crosslinked clamp+CTD and clamp+RGD species in ImageJ and normalized the values using the
196 intensity of the band corresponding to free clamp on each experiment to account for loading differences
197 between lanes. The ratio I(clamp+RGD)/I(clamp) was 0.9±0.01 across all time points. The ratio
198 I(clamp+CTD)/I(clamp) was 1.0 at day 1, increased to 1.18±0.1 at day 2 and stayed the same at day 3,
199 indicating that the reaction was complete by day 2.
200 The Cys-crosslinked clamp-MutL^{CTD} complex has been previously characterized by small angle
201 X-ray scattering (10). To assess whether the conformational changes predicted in our model resemble
202 the structure of the *B. subtilis* Cys-crosslinked clamp-MutL^{CTD} complex in solution, we calculated the
203 theoretical scattering profiles for the two potential conformations of the complex using CRYSOL (35).
204
205 **2.7. Nuclease assays**
206 Nuclease assays were performed as described previously with minor modifications (10). A 195 base pairs
207 DNA substrate was amplified from pUC19 (nucleotides 378-572) using primers
208 $^{5'}$d(AGTTAGCTCACTCATTAGGCACCCCAGGC) and 6-carboxyfluorescein-$^{5'}$d(TGTAAAACGA
209 CGGCCAGTGAATTCGAGCTCGG). MutL^{CTD} variants carrying point mutations in the endonuclease
210 motif (E468K), the ^{443}GQ motif (G443K and Q444E), and the clamp-binding motif (^{487}QEMIV^{491} →

^{487}AEMAA^{491}) were generated by side-directed mutagenesis. All the variants were purified as described elsewhere (10). Each MutL^{CTD} variant (1.2 μM) was then incubated with the 195 bp linear DNA substrate (10 nM) in the absence and presence of equimolar amounts of the β-clamp in reaction buffer (20 mM Tris pH 7.6, 30 mM KCl, 1 mM $MnCl_2$, 1 mM $MgCl_2$, 152 pM $Zn(O_2CCH_3)_2$, 0.05 mg/mL BSA, 4% glycerol). The reactions were incubated at 37ºC for 2 hours and quenched by the addition of 25 mM EDTA and 1mg/mL proteinase K followed by incubation at 55ºC for 20 min. The digestion products were resolved by gel electrophoresis in 8% denaturing polyacrylamide gels. Gels were visualized using a Typhoon Trio+ (GE Healthcare).

3. RESULTS

3.1. Stabilization of the MutL-clamp complex

E. coli and *B. subtilis* MutL form weak, yet specific, interactions with their respective sliding β-clamps (6,10). The dimerization domain of MutL is organized into two independently folded regions connected by an α-helix that, in the case of *B. subtilis* MutL, harbors the endonuclease motif (26). The N- and C-terminal ends of the domain of MutL define the dimerization interface of the protein and the intervening region defines an independently folded subdomain (**Figure 1**). This subdomain is often referred to as the regulatory subdomain because it mediates the interaction with the sliding β-clamp. The clamp-binding motif of MutL is located at the N-terminus of the regulatory subdomain (**Figure 1A-B**). Since the sliding β-clamp interacts with its binding partners through a conserved groove located at its C-terminus, we stabilized the interaction by connecting the two polypeptide chains with a short linker – an approach that has been successfully used to stabilize other weak protein-protein interactions for crystallographic studies (36-39). The *E. coli* and *B. subtilis* fusions could be purified to homogeneity and eluted from a size exclusion chromatography column at retention volumes consistent with the expected mass for each complex (**Figure 1C-D**).

3.2. The regulatory domain of *E. coli* MutL forms a monodentate interaction with the β-clamp

Crystals of the *E. coli* β-clamp fused to the regulatory domain of MutL ($MutL^{RGD}$, residues 471-574) diffracted to 2.1 Å (**Table 1**). The asymmetric unit contained one dimer of the fusion (**Figure 2A**), where both the clamp dimer and the two MutL regulatory subdomains had similar overall structures to the individual proteins. Superimposition of the regulatory subdomain from the original structure of the endonuclease domain of MutL onto the $MutL^{RGD}$ portion of the fusion results in a root mean square deviation (rmsd) of 0.495 Å (518 atoms), whereas superimposition of the structure of the β-clamp bound to a canonical β-binding peptide onto the β-clamp monomer yields an rmsd of 0.628 Å (2,211 atoms). For both protomers of the ring, the clamp-binding motif of MutL (^{482}**Q**PL**L**IP^{487}) sits atop the binding site of the β-clamp in identical orientations and shows well-defined electron density (**Supplementary Figure 1B**). The linker joining both halves of the fusion, however, is disordered in both protomers of the dimer, suggesting that the association between the two proteins is held together by the interactions between clamp-binding motif of MutL and the β-clamp rather than constraints imposed by the linker.

Studies using short peptides have shown that canonical clamp-binding motifs contact two sites on the *E. coli* β clamp referred to as subsites 1 and 2 **(Supplementary Figure 1)** (19). The conserved Leu485 of the clamp-binding motif occupies the conserved hydrophobic pocket known as subsite 1 (**Figure 2B and Supplementary Figure 1**). As in other clamp binding motifs, the following residue, Ile486, is also inside this hydrophobic pocket (**Figure 2B**). In contrast to other clamp-binding motifs, however, the conserved Gln482 does not occupy subsite 2. Instead, a sulfate ion from the crystallization solution occupies this pocket and recapitulates the interactions normally mediated by the amide moiety of the conserved glutamine (**Figure 2B-C**). The ion forms a hydrogen-bond network involving the side chains of Asn320 from the β-clamp and Arg531 and Gln532 from MutL, as well as water-mediated hydrogen bonds with the main chain of His175, Asn320, Met362 and Pro363 (**Figure 2C and**

Supplementary Figure 2). Additionally, the side chain of Leu528 from MutL is sandwiched between the side chains of Phe278 and Met364 from the β-clamp and closes subsite 2. While the interactions in subsite 2 are not conserved, the overall width of the binding groove is similar to other structures of the β-clamp bound to clamp-binding motifs (**Supplementary Figure 2**). When the conserved glutamine of the clamp-binding motif occupies subsite 2, as it does in the case of Pol II (**Supplementary Figure 2**), its side chain forms a hydrogen bond with the carbonyl group of Met362, effectively restricting the width of the groove. Here the sulfate ion forms a water-mediated hydrogen bond with the same group, and the guanidinium group of Arg531 further constrains the width of the pocket by forming a hydrogen-bond with the carbonyl group of Pro363 (**Figure 2C-D**). Additionally, the side chain of Gln482 forms a bidentate hydrogen bond with the main chain of Arg365 from the β-clamp (**Figure 2D**).

In canonical clamp-binding motifs, the residue following the conserved glutamine is also conserved (**Figure 1A**), and it occupies a shallow pocket opposite to Gln-binding pocket further stabilizing the bidentate interaction between the β-clamp and its binding partners. The second position of the motif is not conserved in MutL homologues (**Figure 1A**). In *E. coli* MutL, Pro483 follows the conserved Gln482 (**Supplementary Figures 1 and 2**). Although the conformation of the main chain is similar to other clamp-binding motifs, the restricted phi/psi angles of Pro483 and its smaller size likely favor the monodentate interaction with the β-clamp. We cannot tell whether the sulfate ion forces Gln482 out of subsite 2 or fills the empty pocket, but we have shown that substitution of *E. coli* MutL residue Gln482 with Ala does not affect mismatch repair function *in vivo*, whereas substitution of Leu485 with Ala causes a mutator phenotype (6). This structure supports the idea that Leu485 drives the interaction between *E. coli* MutL and the β-clamp.

3.3. MutL residues interacting with subsite 2 are dispensable for the interaction with the β-clamp

Given the roles of Arg531, Gln532 and Leu528 at stabilizing the conformation of subsite 2 in the structure, we checked whether point mutations on these residues affected mismatch repair activity *in vivo*. We generated a double-mutant of MutL (MutL-R531S/Q532A) unable to mediate the electrostatic interactions with the sulfate ion, as well as a triple variant (MutL-L528A/R531S/Q532A) in which both, electrostatic and van der Waals interactions, had been abrogated. We then measured the frequency of spontaneous mutation of *rpoB* to rifampicin resistance (Rif^R) for the wild-type and the variant *mutL* strains. Both variants displayed similar mutation frequencies to wild-type *mutL* (**Figure 2E**), indicating that the interactions mediated by these three residues in subsite 2 are not necessary for mismatch repair activity *in vivo*. We have previously shown that disruption of the MutL-clamp interaction causes a moderate mutator phenotype in *E. coli* (6), therefore we concluded that the Arg531, Gln532 and Leu528 residues of MutL are not necessary for the functional interaction with the β-clamp during mismatch repair.

3.4. The regulatory domain of *B. subtilis* MutL forms a bidentate interaction with the β-clamp

In contrast the *E. coli* MutL, the interaction between *B. subtilis* MutL and the β-clamp is essential for mismatch repair activity *in vivo* (26). Therefore, we sought to determine whether *B. subtilis* MutL also forms a monodentate interaction with the β-clamp. Crystals of the *B. subtilis* β-clamp fused to the regulatory domain of MutL ($MutL^{RGD}$, residues 482-574) contained two β-clamp dimers in the asymmetric unit. The two protomers of the β-clamp dimer were virtually identical to each other (rmsd of 0.438 Å over 2,165 atoms) and to the *B. subtilis* β-clamp on its own (PDB ID 4TR6). The two regulatory domains of MutL on each β-clamp dimer interacted with the β-clamp through the canonical bidentate interaction observed for other clamp-interacting partners (**Figure 3A-B and Supplementary Figures 1 and 2**). The side chains of the conserved isoleucine and glutamine residues within the motif (^{487}**Q**EM**I**VP^{492}) occupy subsites 1 and 2 of the β-clamp. The presence of the linker does not determine

the interaction because the linker connecting the *B. subtilis* β-clamp and MutL halves of the fusion is only visible in one of the protomers of the dimer, but both copies share the same binding mode **(Supplementary Figure 3)**.

However, binding to the β-clamp caused conformational changes onto the regulatory domain of MutL. Superimposition of the domain onto the original structure of the endonuclease domain of *B. subtilis* MutL (PDB ID 3KDK) results in rmsd >1.5 Å, primarily caused by changes on the relative orientation of the clamp-binding motif and helices αB-αD (**Figure 3A**). The β5-αB loop, as well as part of helix αD, is disordered in both protomers. In the protomer with the ordered linker, the last turn of helix αB, the loop connecting this helix to β6, and helix αD are also disordered (**Figure 3A, shown in blue**). The last two turns of helix αD, however, are visible in the protomer with the disordered linker (**Figure 3A-B and Supplementary Figure 3**). This is the first structure of *B. subtilis* β-clamp bound to one of its binding partners, but kinetic studies have shown that *B. subtilis* β-clamp binds a model clamp-binding peptide with similar affinity to *E. coli* β-clamp (40). This suggests that bidentate binding is also the canonical binding mode for *B. subtilis* clamp-binding partners. This structure also reveals that formation of this bidentate interaction causes additional conformational changes when the clamp-binding motif is embedded within a structured domain.

3.5. No additional residues in the regulatory domain are required for the interaction of *B. subtilis* MutL and β-clamp

The induced flexibility of the *B. subtilis* MutL regulatory domain allows the ^{487}**Q**EM**I**VP^{492} motif to reach into the hydrophobic groove of the binding β-clamp. This, in turn, brings the surrounding loops of the regulatory subdomain close to the surface of the β-clamp. The loop connecting the β7' strand to the αC helix wraps around subsite 2 defining an intricate hydrogen-bond network involving the side chains of His532 and Lys538 from MutL and the side chains of His182, Ser331 and Tyr334 from the β-

clamp, as well as the main chain carbonyl of Trp535 (MutL) and the amino group of Ala287 (β-clamp) (**Figure 3C**). To test the relevance of this hydrogen-bond network, we generated a strain of *mutL* (*mutL*^{HK}) where His532 and Lys538 had been mutated to alanine residues. Each *mutL* allele was expressed from an ectopic chromosomal locus with a concentration of IPTG that yields *mutL* expression to a level indistinguishable from the wild type *mutL* (41). The Δ*mutL*/*mutL*^{HK} strain had similar mutation frequency to a Δ*mutL* strain complemented with wild-type MutL (Δ*mutL*/*mutL*$^{+}$), in contrast to the strain with a mutated clamp-binding motif (Δ*mutL*/*mutL*^{CBM}) that had equivalent mutation frequency to the Δ*mutL* strain lacking a complementing allele (**Figure 3D**). These results indicate that the changes on the relative orientation of helices αB-αD enhance the shape and charge complementarity with the β-clamp, but the functional interaction of MutL and the β-clamp remains exclusively mediated by the clamp-binding domain.

3.6. Re-orientation of helices αB-αD aligns the endonuclease motif with the β-clamp channel

Superimposition of the endonuclease domain of MutL (PDB ID 3KDK) using the clamp-binding motif as reference resulted in significant clashes between the distal protomer of the MutL dimer and the β-clamp (**Figure 4A and Supplementary movie 1**). However, superimposition of the endonuclease domain taking into account the re-orientation of helices αB-αD upon binding to the β-clamp resolved most of the clashes (**Figure 4B and Supplementary movie 1**). The endonuclease domain of MutL crosslinked to the β-clamp has been characterized using small angle X-ray scattering (10), therefore we compared the two conformations of the model to that of the MutL^{CTD}-clamp complex in solution. To this end, we calculated the theoretical scattering curves for the two potential conformations using CRYSOL (35). We then compared them to the experimental scattering curve of the crosslinked complex resulting in on chi^2 values of 11.22 (conformation on Figure 4A) and 3.28 (conformation on Figure 4B). These chi^2 values indicate that the superimposition of the endonuclease domain taking into account the re-

orientation of helices αB-αD closely resembles the conformation of the complex in solution (**Figure 4B**). Interestingly, the dimerization domain of *E. coli* MutL could be directly superimposed onto the structure of the fusion without major clashes (**Supplementary Figure 4A**), suggesting that the conformational changes of the regulatory subdomain in *B. subtilis* MutL may be important for its nuclease activity.

The conformation of the model depicted in Figure 4B aligns the nuclease motif of the interacting MutL protomer with the central cavity of the β-clamp. This is in good agreement with data showing that *B. subtilis* MutL only requires the endonuclease site from the protomer bound to the β-clamp for nuclease activity (42), and reinforces the idea that the role of PCNA and the β-clamp at this step is threading DNA onto the nuclease activity site (12).

3.7. The conserved GQ motif helps align the MutL dimer with the central cavity of the β-clamp

MutL homologs harboring an endonuclease motif also harbor four additional motifs (9). The structure of the endonuclease domain of *B. subtilis* MutL revealed that three of these motifs (^{572}SCK, ^{604}CPHGRP, ^{623}FKR), together with the endonuclease motif (^{462}DQHAAQERIKYE), define the endonuclease site and coordinate the two zinc metal ions necessary for MutL activity (26). The fourth motif (^{443}GQ) is located at the dimerization interface of MutL and could not be assigned a specific function from the structure of the endonuclease domain of *B. subtilis* MutL. The interaction between MutL and the β-clamp places this motif in close proximity of the α1-β2 (Ala22-Thr33), β3- β4 (Ser49-Asp52) and α4-β11 (Ser156-Gly166) loops of the β-clamp and, therefore, the conserved small and polar residues at these positions may prevent steric hindrance.

To test this possibility, we generated variants of the endonuclease domain of *B. subtilis* MutL (BsMutL^{CTD}) including point mutations on the ^{443}GQ motif and tested their ability to nick a linear DNA substrate. We have previously shown that the nuclease activity of *B. subtilis* MutL is stimulated by the

β-clamp (10). Therefore, we tested whether these variants were able to degrade a DNA substrate (**Figure 5A**). As expected, BsMutL^{CTD} was able to degrade the DNA substrate in the presence of the β-clamp, but variants of BsMutL^{CTD} unable to bind the β-clamp or with a point mutation on the endonuclease motif showed no nuclease activity (**Figure 5A**). The BsMutL^{CTD}-G443K and BsMutL^{CTD}-Q444E variants had residual nuclease activity (**Figure 5A**), indicating that the integrity of the ^{443}GQ motif is important for the nuclease activity. Based on our model and the endonuclease defects associated with mutations in the ^{443}GQ motif, we predicted that additional surfaces beyond the regulatory domain of *B. subtilis* MutL may stabilize the interaction with the β-clamp.

Using the single-cysteine variants of the β-clamp, as well as the endonuclease and regulatory domains of *B. subtilis* MutL that we had previously developed for the small-angle X-ray scattering characterization (10), we tested whether the endonuclease dimer (MutL^{CTD}) interacted more readily with the β-clamp than the regulatory domain (MutL^{RGD}). Incubation of the regulatory and endonuclease domains of MutL with the β-clamp in the absence of reducing agents resulted in the formation of crosslinked species at 63 kDa and 75 kDa, consistent with the interaction of both domains with the β-clamp (**Figure 5B**). Unexpectedly, the 75 kDa (clamp+MutL^{CTD}) and 63 kDa (clamp+MutL^{RGD}) species accumulated to similar extents (**Figure 5B**). We had previously shown that only one of the two protomers of the endonuclease domain dimer can interact with the clamp dimer, because binding partially blocks the second binding site of the β-clamp (10). Therefore, complex formation results in crosslinking gels that have similar amounts of complex (clamp+MutL^{CTD}) and free β-clamp (**Figure 5B**). Conversely, the regulatory domain is a monomer and two regulatory domains could simultaneously bind to the β-clamp ring. However, quantification of the crosslinking gels also showed equal amounts of the 63 kDa species and free β-clamp, and this ratio persisted over time (**Figure 5B**). Therefore, either the regulatory domain has lower affinity for the β-clamp than the endonuclease domain or binding to one site triggers an allosteric change that prevents binding to the second site. Although we cannot rule out the contribution

of allosteric effects, we have previously observed that MutL variants with lower affinity for the β-clamp tend to form higher-order non-specific crosslinks in this assay (10). The reactions containing the regulatory domain of MutL yield clearly defined higher-order crosslinked species (**Figure 5B**), thus we favor the idea that the endonuclease domain has higher affinity than the regulatory domain for the β-clamp and that the conserved GQ motif may contribute to strengthening the interaction.

4. DISCUSSION

The structures of the β-clamp bound to the regulatory domains of MutL are the first describing how the β-clamp recognizes internal clamp-binding motifs embedded within structured domains. *B. subtilis* MutL has the characteristic bidentate interaction with its β-clamp, whereas *E. coli* MutL forms a monodentate interaction with its β-clamp. Previous studies suggested that linear clamp-binding motifs bind to the *E. coli* β-clamp in a two-step process that engages the hydrophobic subsite first, and this is followed by the orientation and binding of flanking residues to subsite 2 (43). Therefore, the structure of the *E. coli* clamp-MutL fusion appears to visualize the first binding step. The interaction between *E. coli* MutL and the β-clamp is important, but not essential, for DNA mismatch repair *in vivo* (6). In fact, point mutations on the conserved Gln that normally occupies subsite 2 do not cause a mutator phenotype and, thus, we presume that *E. coli* MutL may not need to engage the second subsite. Conversely, *B. subtilis* MutL depends on the interaction with the β-clamp for its endonuclease activity. In contrast to *E. coli* MutL, the regulatory domain of *B. subtilis* MutL forms the canonical bidentate interaction with the β-clamp. However, the interaction is associated with conformational changes within the regulatory domain. The movement of helices αB-αD upon binding to the β-clamp affords steric and electrostatic complementarity of the surface of MutL surrounding the clamp-binding motif and help align of the endonuclease motif with the central cavity of the β-clamp (**Figure 4 and Supplementary movie 1**).

Therefore, the differences between the two structures likely recapitulate the relevance of the interaction in each species.

The β-clamp has been identified as a potential antimicrobial target (44). Despite the structural conservation of the binding sites across species, different organisms recognize clamp-binding motifs with different affinities (40). The two interacting modes observed in the structures of the *E. coli* and *B. subtilis* β-clamps fused to the regulatory domain of MutL suggests that despite the conservation of the motif and the binding subsites, interactions beyond the clamp-binding motif may modulate the interaction. This seems to be the case for the conserved GQ motif found in the dimerization domain of *B. subtilis* MutL (**Figures 4 and 5**). Work with short peptides cannot recapitulate these additional interactions and, therefore, structural characterization of larger clamp-bound complexes remains a priority to explore the true potential of the β-clamp as a drug target.

PCNA- and clamp-binding motifs often contain one or two aromatic residues at the end of the motif (**Figure 1A**). Their presence was considered the necessary signature to determine binding specificity (13). However, the binding motifs for MLH1 (MIP motif) and Rev1 (RIR motif) also feature adjacent aromatic residues, suggesting that binding specificity for PCNA is determined by features of the motif beyond those aromatic residues. Indeed, MutL homologs lack the pair of aromatic residues at the end of the binding motif, but they bind specifically to their cognate clamps. Additionally, the clamp-binding motif found in MutL homologs includes a conserved proline residue – a common feature in internal clamp binding motifs. This proline helps expose the clamp binding motif to the solvent, but conformational changes facilitate the canonical bidentate interaction between *B. subtilis* MutL and the β-clamp.

Docking of the nuclease domain of *Saccharomyces cerevisiae* MutLα (PDB ID 4E4W) onto the structure of PCNA does not result in steric clashes between MLH1 of the dimer, but the endonuclease motif of PMS1 is not aligned with the central cavity of PCNA (**Supplementary Figure 4B**). PCNA

449 enhances the endonuclease activity of human and yeast MutLα and provides directionality to the nicking
450 reaction (8,12,45). Therefore, flexibility of the regulatory domain of yeast PMS1 and human PMS2 may
451 also be necessary to align the endonuclease motif with the central cavity of PCNA. We predict that
452 conformational changes within the domains harboring the clamp binding motif may be a general feature
453 for the interactions of all MutL homologs with endonuclease activity, as well as other proteins containing
454 internal clamp-binding motifs.

455 With the structures of the β-clamp and PCNA bound to DNA now available (46,47), it is tempting
456 to speculate how MutL recognizes the newly synthesized strand of the DNA duplex. Both PCNA and the
457 β-clamp slide on DNA and, therefore, the crystal structures lack information about the dynamics of the
458 interaction. Understanding the dynamics of the clamp-DNA interaction will be paramount to elucidate
459 how the β-clamp – and PCNA – present the newly synthesized strand to MutL proteins.
460

461 **FUNDING**

462 This work was supported by grants from the Canadian Institutes of Health Research (MOP-
463 67189) and the Natural Sciences and Engineering Research Council of Canada (288295) to AG, NIH
464 R01 GM066094 to MDS and NSF grants MCB1050948 and MCB1714539 to LAS. JRR was supported
465 in part by NIH Cellular Biotechnology Training Grant (T32 GM008353) and a pre-doctoral fellowship
466 from the Rackham Graduate School at the University of Michigan.
467

468 **ACKNOWLEDGEMENTS**

469 We are thankful to Dr. Martin Schmeing for help preparing Supplementary Movie 1 and to former
470 and current members of the Guarné laboratory for stimulating discussions.
471

472 **AUTHOR CONTRIBUTIONS**

AWA, MCP, AG, MKS, MDS, JRR, LAS designed the experiments and analyzed data. AWA, LL, MKS, JRR, HKM, YS, LS performed experiments and prepared figures. AG wrote the manuscript with input from all authors.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest.

REFERENCES

1. Heltzel, J.M., Maul, R.W., Scouten Ponticelli, S.K. and Sutton, M.D. (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc Natl Acad Sci U S A*, **106**, 12664-12669.
2. Kath, J.E., Chang, S., Scotland, M.K., Wilbertz, J.H., Jergic, S., Dixon, N.E., Sutton, M.D. and Loparo, J.J. (2016) Exchange between Escherichia coli polymerases II and III on a processivity clamp. *Nucleic Acids Res*, **44**, 1681-1690.
3. Lopez de Saro, F.J. and O'Donnell, M. (2001) Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. *Proc Natl Acad Sci U S A*, **98**, 8376-8380.
4. Moldovan, G.L., Pfander, B. and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell*, **129**, 665-679.
5. Lopez de Saro, F.J., Marinus, M.G., Modrich, P. and O'Donnell, M. (2006) The beta sliding clamp binds to multiple sites within MutL and MutS. *J Biol Chem*, **281**, 14340-14349.
6. Pillon, M.C., Miller, J.H. and Guarne, A. (2011) The endonuclease domain of MutL interacts with the beta sliding clamp. *DNA Repair (Amst)*, **10**, 87-93.
7. Simmons, L.A., Davies, B.W., Grossman, A.D. and Walker, G.C. (2008) Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Mol Cell*, **29**, 291-301.

8. Kadyrov, F.A., Dzantiev, L., Constantin, N. and Modrich, P. (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell*, **126**, 297-308.
9. Kosinski, J., Plotz, G., Guarné, A., Bujnicki, J.M. and Friedhoff, P. (2008) The PMS2 subunit of human MutLalpha contains a metal ion binding domain of the iron-dependent repressor protein family. *J Mol Biol*, **382**, 610-627.
10. Pillon, M.C., Babu, V.M., Randall, J.R., Cai, J., Simmons, L.A., Sutton, M.D. and Guarne, A. (2015) The sliding clamp tethers the endonuclease domain of MutL to DNA. *Nucleic Acids Res*, **43**, 10746-10759.
11. Genschel, J., Kadyrova, L.Y., Iyer, R.R., Dahal, B.K., Kadyrov, F.A. and Modrich, P. (2017) Interaction of proliferating cell nuclear antigen with PMS2 is required for MutLalpha activation and function in mismatch repair. *Proc Natl Acad Sci U S A*, **114**, 4930-4935.
12. Pluciennik, A., Dzantiev, L., Iyer, R.R., Constantin, N., Kadyrov, F.A. and Modrich, P. (2010) PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc Natl Acad Sci U S A*, **107**, 16066-16071.
13. Boehm, E.M. and Washington, M.T. (2016) R.I.P. to the PIP: PCNA-binding motif no longer considered specific: PIP motifs and other related sequences are not distinct entities and can bind multiple proteins involved in genome maintenance. *Bioessays*, **38**, 1117-1122.
14. Armstrong, A.A., Mohideen, F. and Lima, C.D. (2012) Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. *Nature*, **483**, 59-63.
15. Dalrymple, B.P., Kongsuwan, K., Wijffels, G., Dixon, N.E. and Jennings, P.A. (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc Natl Acad Sci U S A*, **98**, 11627-11632.
16. Bunting, K.A., Roe, S.M. and Pearl, L.H. (2003) Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the beta-clamp. *EMBO J*, **22**, 5883-5892.

17. Patoli, A.A., Winter, J.A. and Bunting, K.A. (2013) The UmuC subunit of the E. coli DNA polymerase V shows a unique interaction with the beta-clamp processivity factor. *BMC Struct Biol*, **13**, 12.

18. Georgescu, R.E., Yurieva, O., Kim, S.S., Kuriyan, J., Kong, X.P. and O'Donnell, M. (2008) Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. *Proc Natl Acad Sci U S A*, **105**, 11116-11121.

19. Burnouf, D.Y., Olieric, V., Wagner, J., Fujii, S., Reinbolt, J., Fuchs, R.P. and Dumas, P. (2004) Structural and biochemical analysis of sliding clamp/ligand interactions suggest a competition between replicative and translesion DNA polymerases. *J Mol Biol*, **335**, 1187-1197.

20. Xing, G., Kirouac, K., Shin, Y.J., Bell, S.D. and Ling, H. (2009) Structural insight into recruitment of translesion DNA polymerase Dpo4 to sliding clamp PCNA. *Mol Microbiol*, **71**, 678-691.

21. Sakurai, S., Kitano, K., Yamaguchi, H., Hamada, K., Okada, K., Fukuda, K., Uchida, M., Ohtsuka, E., Morioka, H. and Hakoshima, T. (2005) Structural basis for recruitment of human flap endonuclease 1 to PCNA. *EMBO J*, **24**, 683-693.

22. Fernandez-Leiro, R., Conrad, J., Scheres, S.H. and Lamers, M.H. (2015) cryo-EM structures of the E. coli replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and tau. *Elife*, **4**.

23. Dohrmann, P.R. and McHenry, C.S. (2005) A bipartite polymerase-processivity factor interaction: only the internal beta binding site of the alpha subunit is required for processive replication by the DNA polymerase III holoenzyme. *J Mol Biol*, **350**, 228-239.

24. Jergic, S., Horan, N.P., Elshenawy, M.M., Mason, C.E., Urathamakul, T., Ozawa, K., Robinson, A., Goudsmits, J.M., Wang, Y., Pan, X. *et al.* (2013) A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J*, **32**, 1322-1333.

544 25. Toste Rego, A., Holding, A.N., Kent, H. and Lamers, M.H. (2013) Architecture of the Pol III-clamp-
545 exonuclease complex reveals key roles of the exonuclease subunit in processive DNA synthesis and
546 repair. *EMBO J*, **32**, 1334-1343.
547 26. Pillon, M.C., Lorenowicz, J.J., Uckelmann, M., Klocko, A.D., Mitchell, R.R., Chung, Y.S., Modrich,
548 P., Walker, G.C., Simmons, L.A., Friedhoff, P. *et al.* (2010) Structure of the endonuclease domain
549 of MutL: unlicensed to cut. *Mol Cell*, **39**, 145-151.
550 27. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta*
551 *Crystallogr D Biol Crystallogr*, **66**, 133-144.
552 28. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta*
553 *Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
554 29. Afonine, P.V., Grosse-Kunstleve, R.W. and Adams, P.D. (2005) phenix.refine. *CCP4 Newsletter*,
555 **42**, contribution 8.
556 30. Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state.
557 *J Mol Biol*, **372**, 774-797.
558 31. Sutton, M.D. (2004) The Escherichia coli dnaN159 mutant displays altered DNA polymerase usage
559 and chronic SOS induction. *J Bacteriol*, **186**, 6738-6748.
560 32. Dixon, W.J. and Massey, F.J. (1969) *Introduction to Statistical Analysis*. McGraw-Hill, New York.
561 33. Dupes, N.M., Walsh, B.W., Klocko, A.D., Lenhart, J.S., Peterson, H.L., Gessert, D.A., Pavlick, C.E.
562 and Simmons, L.A. (2010) Mutations in the Bacillus subtilis beta clamp that separate its roles in
563 DNA replication from mismatch repair. *J Bacteriol*, **192**, 3452-3463.
564 34. Lenhart, J.S., Pillon, M.C., Guarne, A. and Simmons, L.A. (2013) Trapping and visualizing
565 intermediate steps in the mismatch repair pathway in vivo. *Mol Microbiol*, **90**, 680-698.
566 35. Svergun, D., Barberato, C. and Koch, M.H.J. (1995) CRYSOL - A program to evaluate x-ray
567 solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr*, **28**,
568 768-773.

36. Almawi, A.W., Matthews, L.A., Larasati, Myrox, P., Boulton, S., Lai, C., Moraes, T., Melacini, G., Ghirlando, R., Duncker, B.P. *et al.* (2016) 'AND' logic gates at work: Crystal structure of Rad53 bound to Dbf4 and Cdc7. *Sci Rep*, **6**, 34237.
37. Williams, S.J., Sohn, K.H., Wan, L., Bernoux, M., Sarris, P.F., Segonzac, C., Ve, T., Ma, Y., Saucet, S.B., Ericsson, D.J. *et al.* (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science*, **344**, 299-303.
38. Antczak, A.J., Tsubota, T., Kaufman, P.D. and Berger, J.M. (2006) Structure of the yeast histone H3-ASF1 interaction: implications for chaperone mechanism, species-specific interactions, and epigenetics. *BMC Struct Biol*, **6**, 26.
39. Kingston, R.L., Hamel, D.J., Gay, L.S., Dahlquist, F.W. and Matthews, B.W. (2004) Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc Natl Acad Sci U S A*, **101**, 8301-8306.
40. Wolff, P., Amal, I., Olieric, V., Chaloin, O., Gygli, G., Ennifar, E., Lorber, B., Guichard, G., Wagner, J., Dejaegere, A. *et al.* (2014) Differential modes of peptide binding onto replicative sliding clamps from various bacterial origins. *J Med Chem*, **57**, 7565-7576.
41. Bolz, N.J., Lenhart, J.S., Weindorf, S.C. and Simmons, L.A. (2012) Residues in the N-terminal domain of MutL required for mismatch repair in Bacillus subtilis. *J Bacteriol*, **194**, 5361-5367.
42. Liu, L., Ortiz Castro, M.C., Rodriguez Gonzalez, J., Pillon, M.C. and Guarne, A. (2019) The endonuclease domain of Bacillus subtilis MutL is functionally asymmetric. *DNA Repair (Amst)*, **73**, 1-6.
43. Yin, Z., Kelso, M.J., Beck, J.L. and Oakley, A.J. (2013) Structural and thermodynamic dissection of linear motif recognition by the E. coli sliding clamp. *J Med Chem*, **56**, 8665-8673.
44. Kling, A., Lukat, P., Almeida, D.V., Bauer, A., Fontaine, E., Sordello, S., Zaburannyi, N., Herrmann, J., Wenzel, S.C., Konig, C. *et al.* (2015) Antibiotics. Targeting DnaN for tuberculosis therapy using novel griselimycins. *Science*, **348**, 1106-1112.

45. Kadyrov, F.A., Holmes, S.F., Arana, M.E., Lukianova, O.A., O'Donnell, M., Kunkel, T.A. and Modrich, P. (2007) Saccharomyces cerevisiae MutLa is a mismatch repair endonuclease. *J Biol Chem*.

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

47. De March, M., Merino, N., Barrera-Vilarmau, S., Crehuet, R., Onesti, S., Blanco, F.J. and De Biasio, A. (2017) Structural basis of human PCNA sliding on DNA. *Nat Commun*, **8**, 13935.

602 **Table 1: Data collection and refinement statistics**

	E. coli β-MutL fusion	*B. subtilis* β-MutL fusion
Data Collection		
Beamline	08ID-1	08B1-1
Wavelength	0.97936	0.98010
Space Group	$P2_12_12_1$	P1
Cell dimensions a, b, c (Å) α, β, γ (°)	78.4, 103.2, 141.6 90, 90, 90	59.0, 83.9, 128.3 80.3, 83.6, 90
Resolution (Å)[#]	46.0-2.07 (2.12-2.07)	48.2-2.34 (2.42-2.34)
Completeness (%)[#]	97.5 (99.8)	97.1 (97.1)
CC½ (%)	99.8 (35.6)	99.6 (32.7)
I/σ(I)	11.5 (1.2)	13.4 (0.95)
Redundancy	4.4 (4.5)	2.2 (1.7)
Refinement		
Resolution (Å)	46 – 2.07	48.2 – 2.34
No. Reflections	53,370	98,638
R_{work} / R_{free} (%)	19.0 / 22.5	21.6 / 25.0
Atoms refined (no H)	7,224	13,850
Solvent atoms	207	195
Rmsd in bonds (Å)	0.004	0.004
Rmsd in angles (°)	0.96	0.73
Mean B values (Å)	54	94.5
Ramachandran Plot (%)		
Favored	97.4	95.5
Outliers	0.1	0.4

603 [#]Data in the highest resolution shell is shown in parentheses.
604

605

606 **FIGURES AND FIGURE LEGENDS**

607 **Figure 1: Stabilization of the clamp-MutL complex. (A)** Sequence conservation of the canonical
608 clamp-binding motif (top), MutL clamp- and PCNA-binding motif (center), and PIP-box (bottom). The
609 Gln and bulky hydrophobic residue, conserved in all motifs, are highlighted in purple. Conserved
610 residues unique to each motif highlighted in green, yellow, and cyan. **(B)** Ribbon representation of the
611 endonuclease domain of *B. subtilis* MutL (433-627). The beginning of the regulatory domain (RGD) is
612 mark with a red arrow. The endonuclease motif is colored purple and the zinc metal ions are shown as
613 green spheres. The clamp-binding motif (CBM) is colored in cyan with the side chains of the conserved
614 Gln and Leu residues shown as sticks. **(C)** Construction of the *E. coli* and *B. subtilis* clamp-MutL^{RGD}
615 fusions shown in blue and red, respectively. Numbers correspond to the residues from each protein
616 included in the fusion (Ecclamp (1-366), EcMutL^{RGD} (471-574), Bsclamp (1-378), BsMutL^{RGD} (482-
617 574). **(D)** Size exclusion chromatography column profiles of the *E. coli* and *B. subtilis* clamp-MutL^{RGD}
618 fusions. Elution volumes of reference proteins are: 1. thyroglobulin (669 kDa), 2. ferritin (440 kDa), 3.
619 catalase (232 kDa), 4. aldolase (158 kDa), 5. albumin (67 kDa), 6. ovoalbumin (43 kDa), 7.
620 Chymotrypsinogen A (25 kDa), and 8. ribonuclease A (13.7 kDa). The vertical lines mark the elution
621 volumes of the Bsclamp (13.1 mL), the Ecclamp (13.7 mL), the BsMutL^{RGD} (17.2 mL) and the
622 EcMutL^{RGD} (17.3 mL).

623

624 **Figure 2: Interaction between *E. coli* MutL^{RGD} and the β-clamp. (A)** Crystal structure of the *E. coli*
625 clamp-MutL^{RGD} fusion with the β-clamp ring shown as a semi-transparent surface and the two regulatory
626 domains of MutL bound to the ring shown as grey ribbons. **(B)** Detail of the interaction between the β-
627 binding motif of *E. coli* MutL^{RGD} and the conserved subsites 1 (red) and 2 (cyan) on the surface of the
628 β-clamp. The conserved residues of the β-binding motif are shown, as well as the sulfate ion occupying
629 subsite 2, as color-coded. This view is orthogonal to panel **(A)**. **(C)** Hydrogen-bond network stabilizing

630 the interaction of the sulfate ion found in subsite 2. Residues from MutL^{RGD} and the β-clamp are,
631 respectively, shown in grey and pale-yellow, and hydrogen-bonds are shown as dashed lines. Water
632 molecules are shown as red spheres. The 2Fo-Fc electron density map for the sulfate ion is shown as a
633 gold mesh contoured at 1.2 σ. **(D)** Detail of the interaction of the conserved Gln482 of MutL^{RGD} and the
634 C-terminal strand of the β-clamp. **(E)** Mutation frequency associated with the MutL-R531S/Q532A (RQ)
635 and MutL-L528A/R531S/Q532A (LRQ) variants compared to a *mutL*-deficient *E. coli* strain
636 complemented with an empty vector (empty) or MutL wild-type (WT). The error bars represent the upper
637 and lower bound 95% confidence intervals (32).
638
639 **Figure 3: Interaction between *B. subtilis* MutL and the β-clamp. (A)** Opposite views of the regulatory
640 domains of MutL from the structure of the *B. subtilis* clamp-MuL fusion (green and blue) superimposed
641 onto the structure of MutL^{RGD} (grey, PDB ID 3KDK). **(B)** Detail of the crystal structure of the *B. subtilis*
642 clamp-MutL fusion with the β-clamp ring shown as a semi-transparent surface and the regulatory domain
643 of MutL shown as a green ribbon. The conserved residues of the β-binding motif are shown as color-
644 coded sticks and labeled. The conserved β-clamp binding pockets are colored in red (subsite 1) and blue
645 (subsite 2). **(C)** Detail of the residues of the β7-αC loop contributing to the stabilization of subsite 2. **(D)**
646 Mutation frequency of a (wild type PY79) and isogenic Δ*mutL* strains, as well as the Δ*mutL* strain
647 complemented with wild-type MutL (Δ*mutL*/*mutL*$^{+}$), MutL-H532A/K538A (Δ*mutL*/*mutL*^{HK}) or MutL-
648 CBM (Δ*mutL*/*mutL*^{CBM}). The complementing alleles were expressed from an ectopic chromosomal locus
649 with 200 μM IPTG. Error bars indicate the standard error of the mean.
650
651 **Figure 4: Conformational changes of the MutL^{RGD} upon clamp-binding align the endonuclease**
652 **motif of MutL.** Ribbon diagram of the endonuclease domain dimer of *B. subtilis* MutL (pink)
653 superimposed onto the *B. subtilis* clamp-MutL fusion (pale yellow and green) using either the clamp-

binding motif **(A)** or the position of helices αB-αD **(B)** as reference. Re-orientation of helices αB-αD aligns the endonuclease motif of MutL with the central cavity of the β-clamp and prevents clashes with the second protomer of the dimer (red circle). **(C)** Theoretical solution scattering curve of the model shown in panel **(A)**, superimposed onto the experimental solution scattering curve of the Cys-crosslinked complex between the *B. subtilis* β-clamp and the endonuclease domain of *B. subtilis* MutL. **(D)** Theoretical solution scattering curve of the model shown in panel **(B)**, superimposed onto the experimental solution scattering curve of the Cys-crosslinked complex between the *B. subtilis* β-clamp and the endonuclease domain of *B. subtilis* MutL.

Figure 5: Motifs beyond the regulatory subdomain of *B. subtilis* MutL are important for nuclease activity and binding to the β-clamp. (A) Nuclease activity assay for $MutL^{CTD}$ (WT), and variants of the domain with point mutations in the ^{443}GQ motif (G443K and Q444E), the nuclease motif (E468K), and the clamp-binding motif (^{487}QEMIV→ ^{487}AEMAA, $CBM^{(-)}$) on a fluorescently-labeled 195 bp linear DNA (10 nM). The nuclease activity is dependent on the presence (+) and ability to interact with the β-clamp. **(B)** Single-cysteine variants of the β-clamp, as well as the endonuclease ($MutL^{CTD}$) and regulatory ($MutL^{RGD}$) domains of *B. subtilis* MutL were purified and equimolar mixtures of either clamp-$MutL^{CTD}$ or clamp-$MutL^{RGD}$ were incubated in the absence of reducing agents. Samples withdrawn from the reaction at the indicated time points were resolved on denaturing gels in the absence of β-mercaptoethanol. Quantification of the bands corresponding to the crosslinked clamp+$MutL^{CTD}$ or clamp+$MutL^{RGD}$ species with respect to free clamp is shown below the gel.

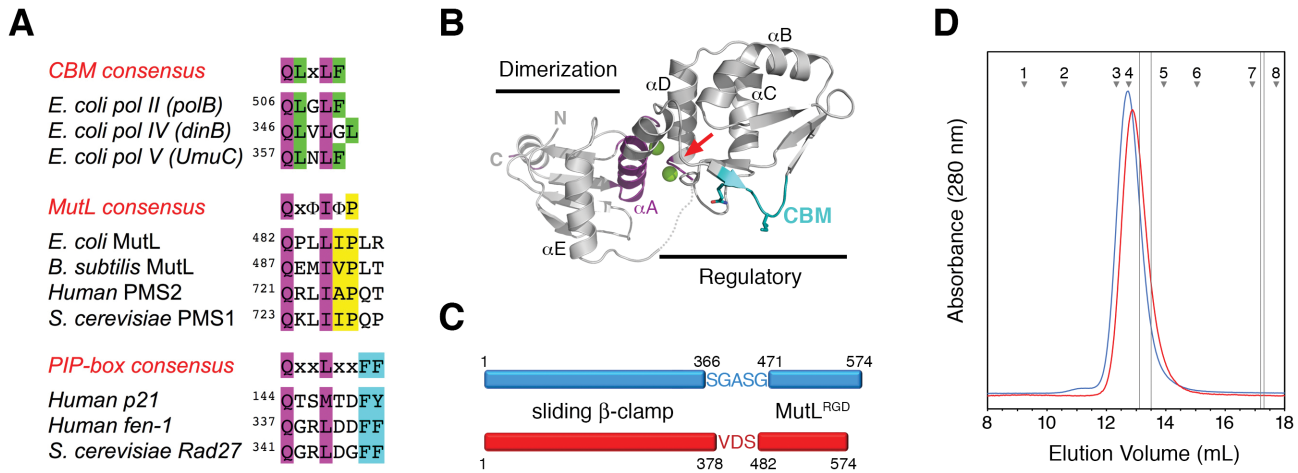
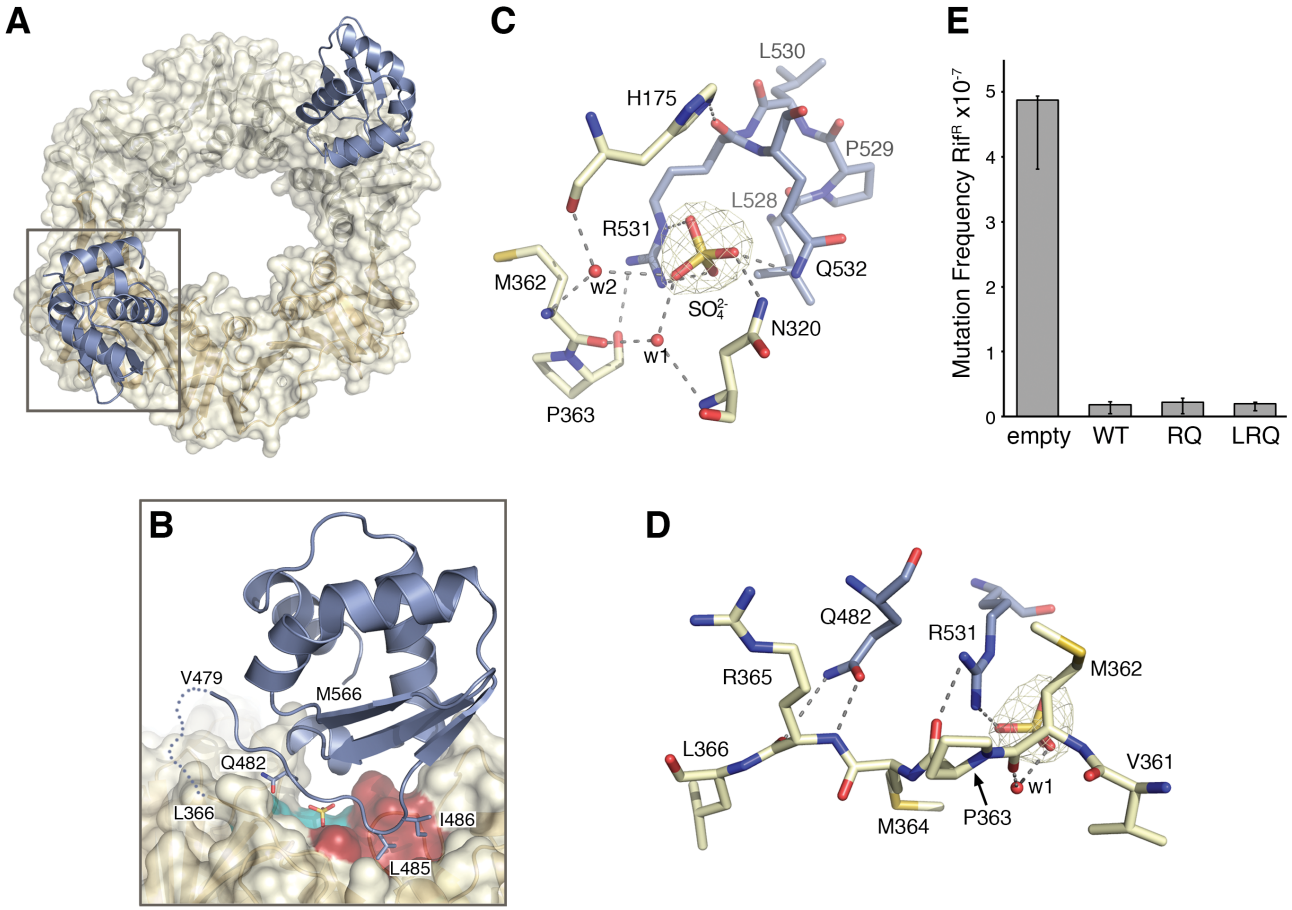

A
CBM consensus QLxLF
E. coli pol II (polB) 506 QLGLF
E. coli pol IV (dinB) 346 QLVLGL
E. coli pol V (UmuC) 357 QLNLF
MutL consensus QxΦIΦP
E. coli MutL 482 QPLLIPLR
B. subtilis MutL 487 QEMIVPLT
Human PMS2 721 QRLIAPQT
S. cerevisiae PMS1 723 QKLIIPQP
PIP-box consensus QxxLxxFF
Human p21 144 QTSMTDFY
Human fen-1 337 QGRLDDFF
S. cerevisiae Rad27 341 QGRLDGFF
B
Dimerization
αB
αD
αC
N
C
αA
CBM
αE
Regulatory
C
1
366
471
574
SGASG
sliding β-clamp
MutL RGD
VDS
1
378
482
574
D
1
2
3 4
5
6
7
8
Absorbance (280 nm)
8
10
12
14
16
18
Elution Volume (mL)

Figure 1.

676 F



677

678

679

680 Figure 2.

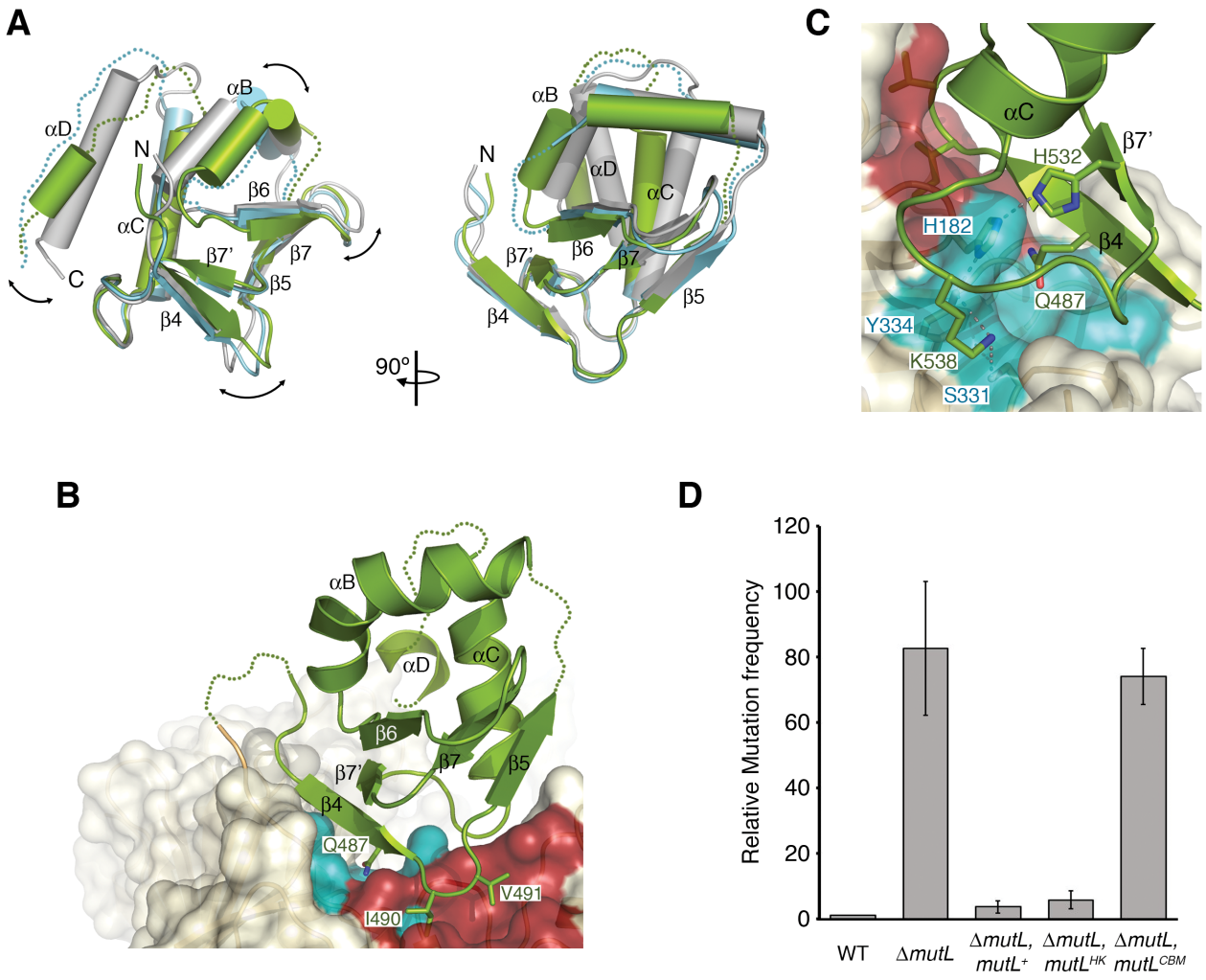

A
αB
αD
N
β6
αC
β7'
β7
β5
C
β4
90°
αB
N
αD
αC
β7'
β6
β7
β5
β4
B
αB
αC
αD
β6
β7
β5
β7'
β4
Q487
V491
I490
C
αC
β7'
H532
H182
β4
Q487
Y334
K538
S331
D
Relative Mutation frequency
120
100
80
60
40
20
0
WT
ΔmutL
ΔmutL, mutL+
ΔmutL, mutLHK
ΔmutL, mutLCBM

Figure 3.

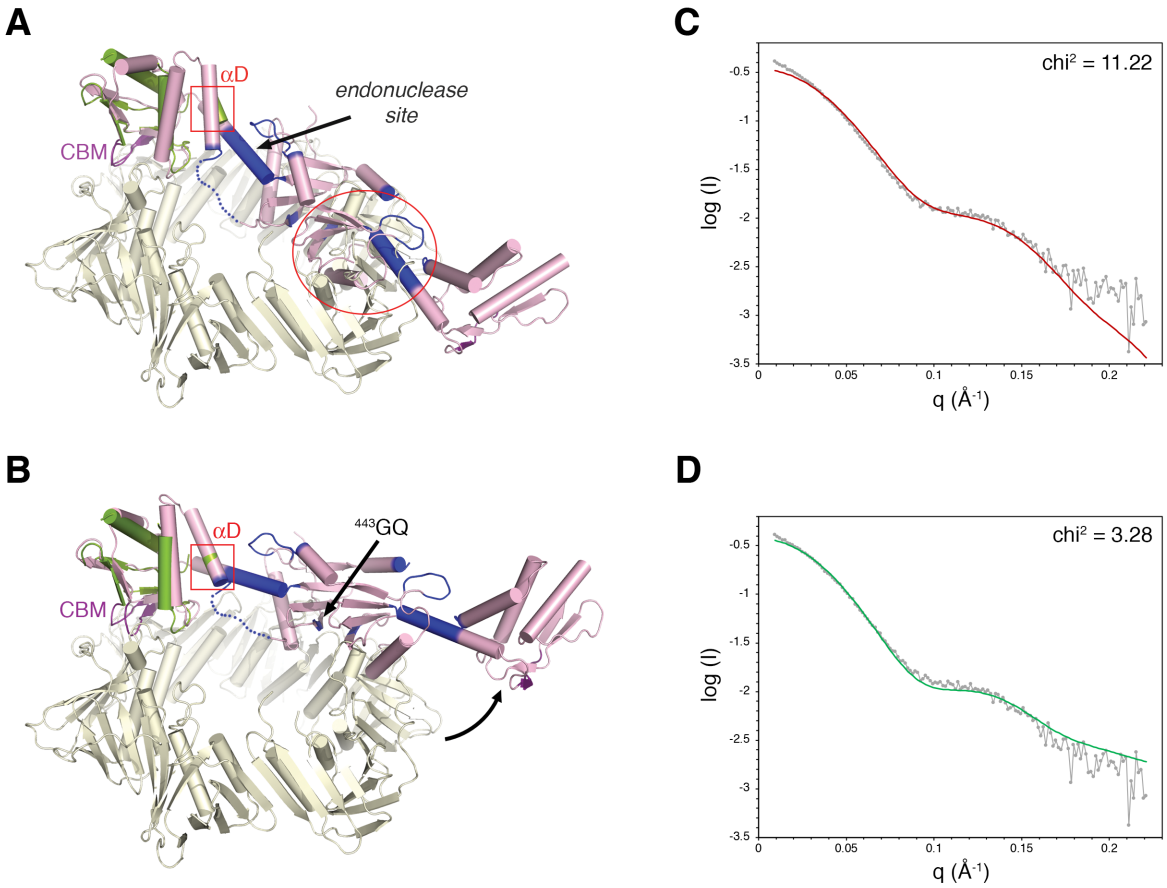
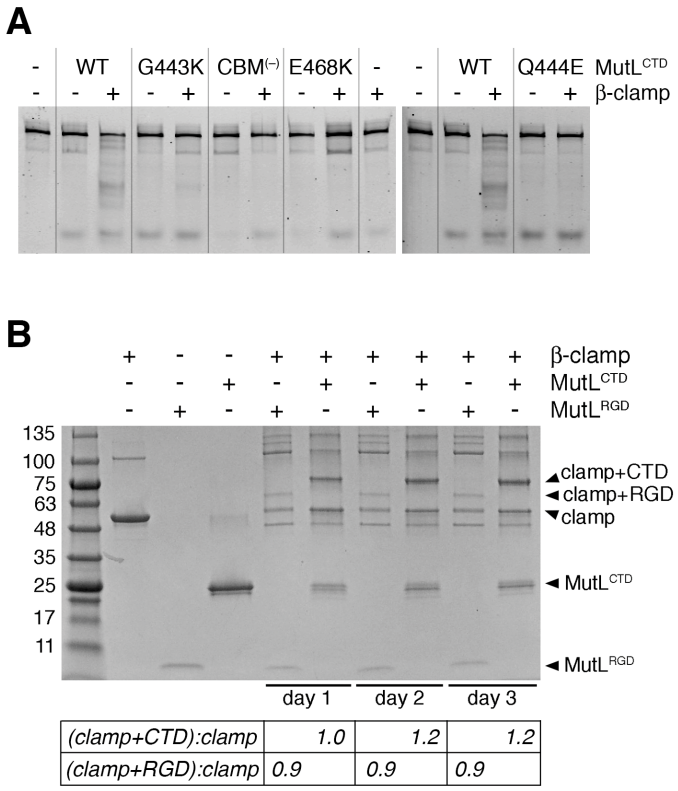

A
αD
endonuclease
site
CBM
B
αD
443GQ
CBM
C
chi² = 11.22
log (I)
q (Å-1)
D
chi² = 3.28
log (I)
q (Å-1)

Figure 4.



(clamp+CTD):clamp	*1.0*	*1.2*	*1.2*
(clamp+RGD):clamp	*0.9*	*0.9*	*0.9*

Figure 5.

Binding of MutL to the sliding β-clamp is species specific

Ahmad W. Almawi, Michelle K. Scotland, Justin R. Randall, Linda Liu, Heather K. Martin, Lauralicia Sacre, Yao Shen, Monica C. Pillon, Lyle A. Simmons, Mark D. Sutton, Alba Guarné*

Supplementary Material:

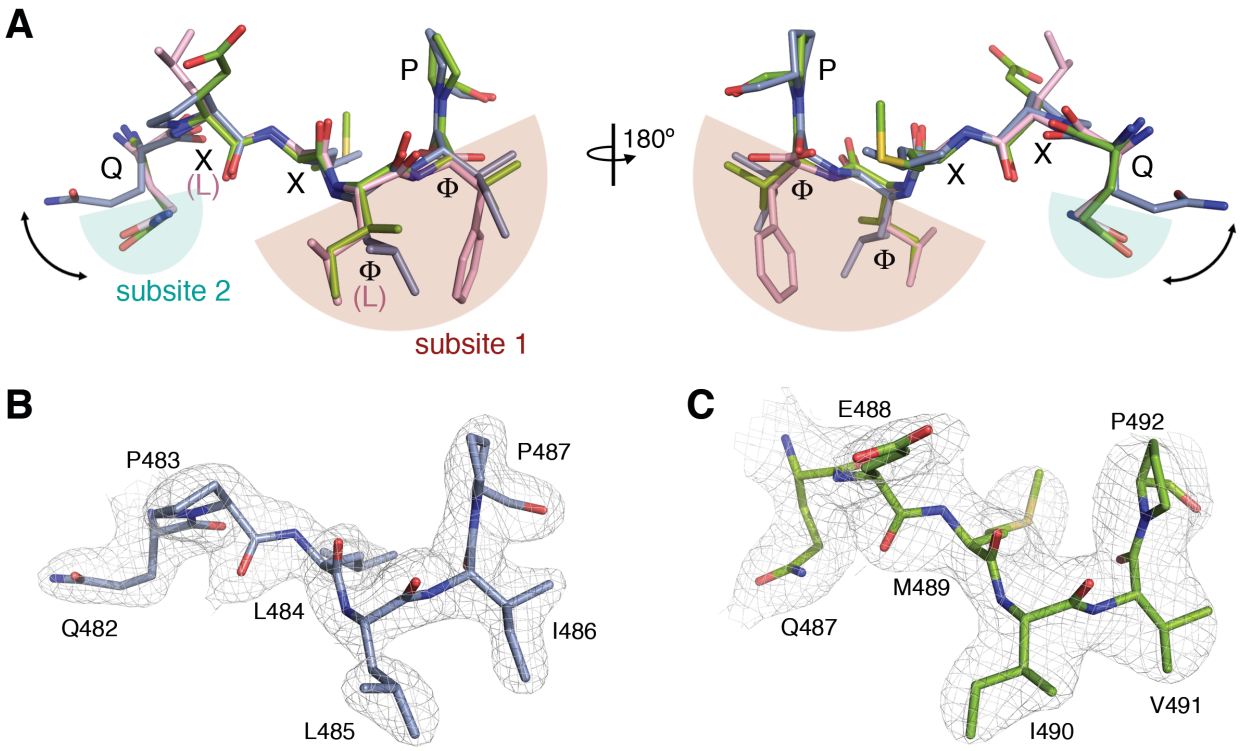
Supplementary Figure 1: Canonical interaction between clamp-binding motifs and the β-clamp.

Supplementary Figure 2: Interaction between MutL and the β-clamp.

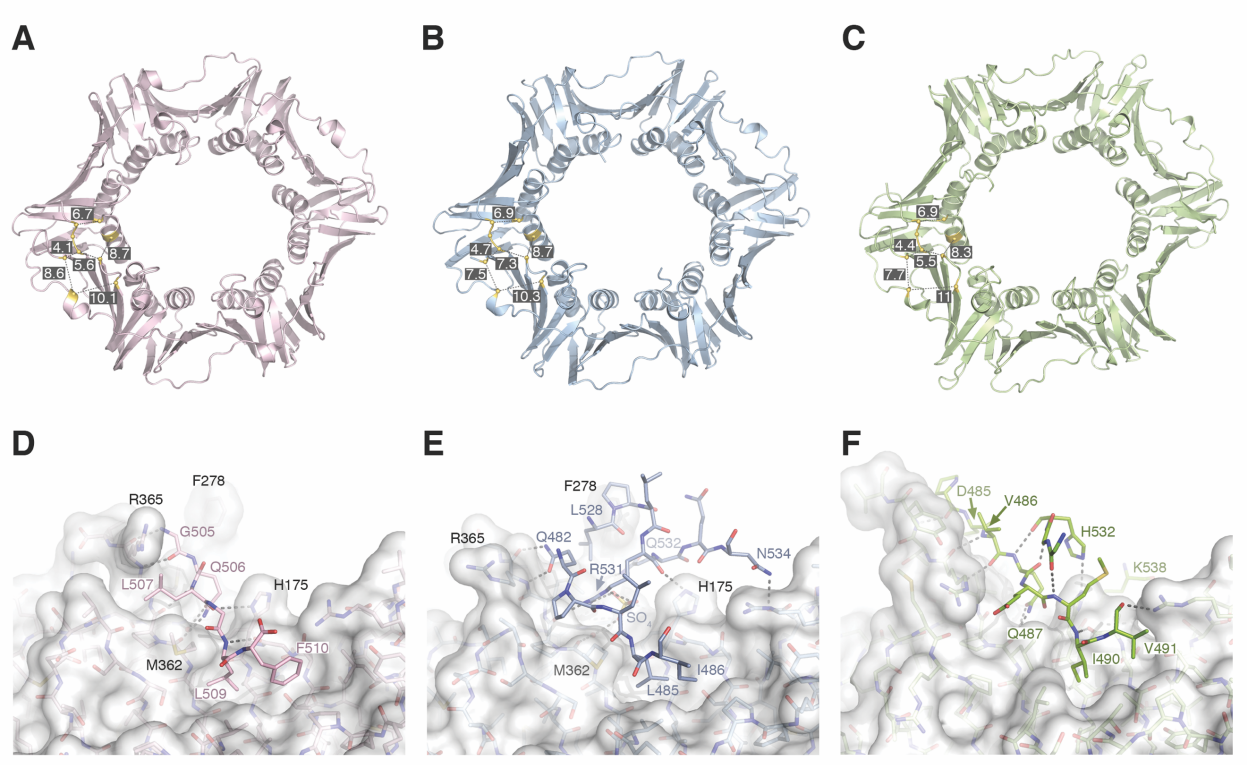
Supplementary Figure 3: Interaction between B. subtilis MutL and the β-clamp.

Supplementary Figure 4: Model of the interaction between MutL homologues and their clamps.

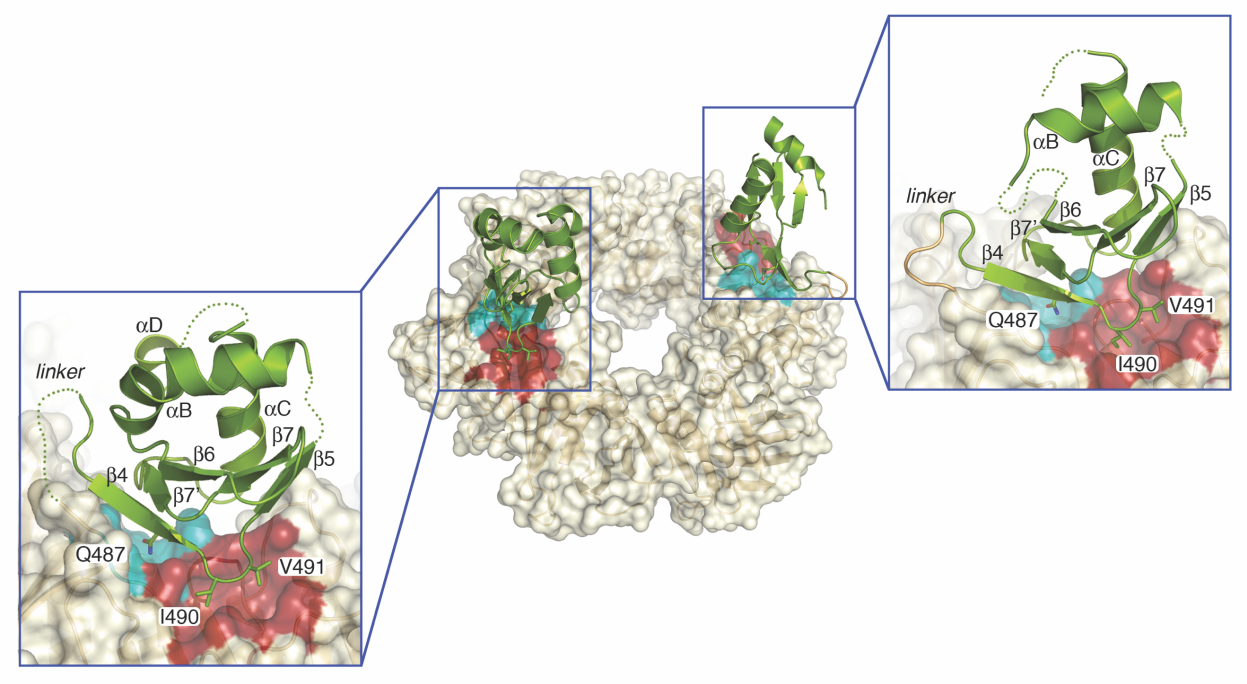
Supplementary Movie 1: MutL-clamp complex.



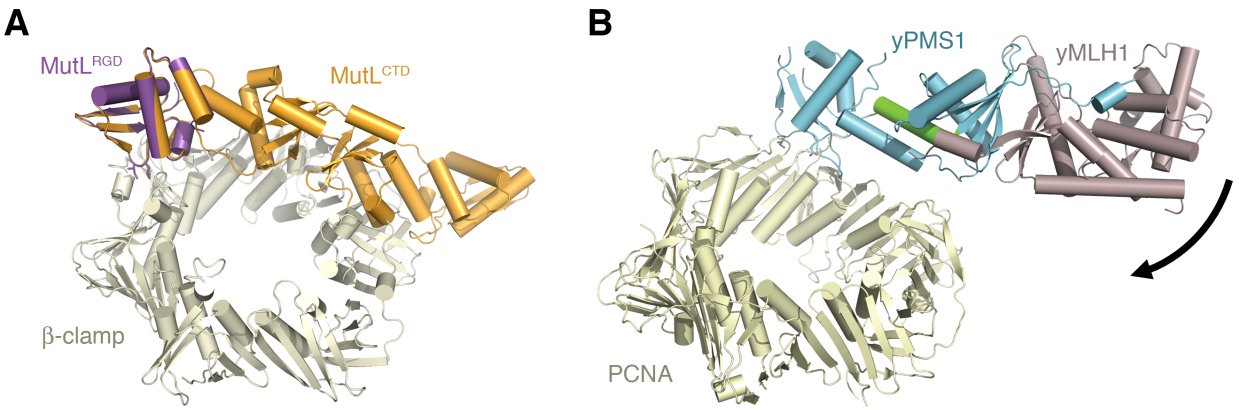
Supplementary Figure 1: Canonical interaction between clamp-binding motifs and the β-clamp. (A) Opposite views of the clamp-binding motifs (CBM) of *E. coli* MutL (grey) and *B. subtilis* MutL (green) superimposed onto the clamp-binding motif of *E. coli* polymerase II (pink). The residues in the CBM of MutL (Q-X-X-Φ- Φ -P, where X is any residue and Φ is a bulky hydrophobic residue) are labeled. Additional conserved position in canonical clamp-binding motifs are indicated in parenthesis. The movement of the conserved Gln out of subsite 1 is indicated with a double head arrow. **(B)** Omit map of the clamp-binding motif of *E. coli* MutL contoured at 1.5σ. **(C)** Omit map of the clamp-binding motif of *B. subtilis* MutL contoured at 1.5σ. The left panel in **(A)**, is shown on a similar orientation to panels **(B)** and **(C)**.



Supplementary Figure 2: Interaction between MutL and the β-clamp. (A) Top view of the *E. coli* β-clamp from the structure bound to a pol II peptide (PDB: 3D1E). **(B)** Top view of the *E. coli* β-clamp from the structure in complex with the regulatory domain of MutL. **(C)** Top view of the *B. subtilis* β-clamp from the structure in complex with the regulatory domain of MutL. Numbers correspond to Cα-Cα distances (in angstroms) between residues delimiting the clamp-binding groove. **(D)** Detailed interactions between the *E. coli* β-clamp bound to the CBM of pol II. **(E)** Interactions mediating the *E. coli* clamp-MutL interaction. **(F)** Detailed interactions between the *B. subtilis* β-clamp and *B. subtilis* MutL. Residues mediating hydrogen-bond interactions are labeled. Note that Phe278 (*E. coli* β-clamp) is conserved in the *B. subtilis* β-clamp, but the loop containing the equivalent residue is disordered in the crystal structure of the *B. subtilis* clamp-MutL^{RGD} fusion.



Supplementary Figure 3: Interaction between *B. subtilis* MutL and the β-clamp. Crystal structure of the *B. subtilis* clamp-MutL fusion with the β-clamp ring shown as a semi-transparent surface and the two regulatory domains of MutL shown as green ribbons (centre). Detail of the interaction between the β-binding motif on each protomer of *B. subtilis* MutL and the conserved subsites 1 (red) and 2 (light blue) on the β-clamp.



Supplementary Figure 4: Model of the interaction of MutL homologues and their clamps. **(A)** Ribbon diagram of the dimerization domain of *E. coli* MutL (gold) superimposed onto the structure of the *E. coli* clamp-MutL^{CTD} fusion (pale yellow and purple). **(B)** Ribbon diagram of the endonuclease domain dimer of *S. cerevisiae* MutLα (yPMS1 (blue), yMLH1 (taupe)) superimposed onto human PCNA (pale yellow) using the PIP box of yPMS1 to guide the superimposition. The helix harboring the nuclease motif of yPMS1 is colored green.

Supplementary Movie 1: MutL-clamp complex. Morphing movie depicting the overall movement imposed on the MutL endonuclease dimer (green) upon binding to the β-clamp (cream). The initial point depicts the superimposition of the MutL endonuclease dimer onto the structure of the *B. subtilis* clamp-MutL fusion using the clamp-binding motif to anchor the superimposition. The final point depicts the superimposition taking into consideration the shifts of helices αB-αD from the regulatory subdomain of MutL in the fusion structure. The clamp binding motif is colored red and the conserved motifs defining the endonuclease site are shown in purple.