# Conformations of MutS in DNA mismatch repair

Flora Sulamith Groothuizen

ISBN: 978-90-75575-43-9

Publisher: Het Nederlands Kanker Instituut - Antoni van Leeuwenhoek Ziekenhuis

Printing: Offpage, www.offpage.nl

Cover: Model for initiation of DNA mismatch repair represented as a movie, based on snapshots of different conformations of MutS.



This thesis was printed with financial support from Erasmus University Rotterdam and the Netherlands Cancer Insitute (Antoni van Leeuwenhoek).

**Conformations of MutS in DNA mismatch repair**

Conformaties van MutS in DNA mismatch herstel

**Proefschrift**

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 18 februari 2015 om 13:30 uur
door

**Flora Sulamith Groothuizen**
geboren te Ede

## Promotiecommissie

| | |
|---|---|
| Promotor: | Prof.dr. T.K. Sixma |
| Overige leden: | Prof.dr. C.L. Wyman |
| | Prof.dr. H. te Riele |
| | Prof.dr. J.H.J. Hoeijmakers |

## Table of contents

# 1

# Introduction

# 1 Genomic stability and DNA mismatch repair

## The genome

Every cell contains deoxyribonucleic acids (DNA), long macromolecules that hold the genetic information. This DNA codes for the proteins that need to be made for the cell and, if multicellular, for the whole organism to grow and maintain itself. Although the nucleotide building blocks of DNA are identical, the exact order of the nucleotides in the chains is different for every organism, and even for every individual.

In 1953, the general structure of the DNA chain was published (1). Since this structure showed a double helix of two DNA strands with always the same nucleotide bases paired opposite each other (G/C and A/T), it immediately suggested a copying mechanism for the DNA. Indeed it has been long confirmed that at every cell division, the DNA strands of the genome are duplicated by separating the two strands in the DNA helix, and building new complementary nucleotides to pair with each nucleotide in the parent strands.

The DNA polymerases perform the task of placing new nucleotides opposite their complementary partner, making sure that both daughter cells will contain the same genetic information as the parent cell. This is a great undertaking since cells contain between half a million nucleotides (some mycoplasma bacteria (2)) to 3 billion (humans) or even over 100 billion (some plants) (3). With such numbers, making errors is inevitable. To reduce the chance of making errors, the generic DNA replication polymerases themselves contain a 'proofreading' mechanism. In this proofreading, the polymerase checks for correct base paring after a nucleotide has been built, and removes the nucleotide if incorrect, which reduces the error rate 10-100 fold (4). Depending on the conditions and the organism, the polymerase thus only leaves an error in the DNA about once every 10 million base pairs.

To reduce the number of remaining errors, referred to as DNA mismatches, almost all cellular organisms contain a DNA mismatch repair (MMR) system. This system comprises several proteins that together recognize mispaired nucleotides and make sure that part of the new DNA strand, containing the error, is removed and resynthesized. This reduces the number of errors 100-1000 fold (4,5). In cells that are deficient in MMR, there is a bigger chance that errors end up being permanently incorporated in genes that are essential for cell maintenance. This may result in cell death, which usually will not directly harm a multicellular organism. However, if genes are affected that control cell development, stress response or cell division, this could result in tumorigenic growth. Therefore, mismatch repair deficiency increases the risk of developing cancer. In humans, loss of a single allele of one of the mismatch repair genes predisposes to hereditary non-polyposis colon cancer (HNPCC), also known as Lynch syndrome (6).

**DNA mismatch repair**

The basic elements of the DNA mismatch repair mechanism have been conserved from prokaryotes to higher eukaryotes. MMR is able to correct most base-base mispairs, but also small insertions or deletions loops (IDLs) (7-10). Defective MMR results in a mutator phenotype. To understand how the general MMR system works, the bacteria *Escherichia coli* is widely used as a model organism.

In *E. coli*, MMR involves the following cascade: first, the MutS protein recognizes and binds a mismatch in DNA. MutS then binds adenosine triphosphate (ATP) and undergoes a conformational change to form a sliding clamp on DNA (11,12). MutL specifically binds to this MutS sliding clamp (13-15), and then activates the endonuclease MutH (16). MutH is able to discriminate between daughter and parent DNA strands by the absence of adenine methylation at d(GATC) sites within newly synthesized DNA, and functions to make a nick in the daughter strand. MutL also activates the helicase UvrD (17), which unwinds the DNA starting from the nick, while exonucleases remove the unwound strand, thereby eliminating the newly incorporated mispaired nucleotide. The DNA can then be resynthesized by the DNA polymerase III complex, after which ligase seals the nick (8).

The eukaryotic MMR system follows essentially the same initiation steps, except that no MutH protein is present. The nicking step in eukaryotic cells is either performed by an endonuclease activity in MutL homologs (18), or nicks present in Okazaki fragments are used as a starting point for repair (9,19). Association of the MMR proteins with the replication sliding clamp PCNA is thought to direct strand discrimination (20,21).

The MMR system is essential to maintain genomic stability, and its molecular mechanism and efficiency have fascinated many scientists. The separate MMR steps, the movements of the molecular machines and the communication of the recognition of an error with the strand discrimination are extensively studied. Yet several aspects of the MMR cascade remain unclear and the existing models differ in mechanistic aspects, such as whether MutS leaves the mismatch, whether looping of DNA is necessary, or how many MutL molecules are required (11,15,22,23).

**Conservation of MMR proteins**

Homologs of MutS and MutL proteins have been found in almost all cellular organisms, including prokaryotes, archaeabacteria, yeast, plants and mammals. Eukaryotic MMR is more complex than the prokaryotic pathway, involving different sets of homologs that are specialized in different types of MMR.

Prokaryotic MutS proteins act as homodimers on DNA, but the specific mismatch recognition is performed by one of the two subunits, making this an asymmetric dimer (24-26). In contrast to the prokaryotic homodimer, the eukaryotic MutS homologs form heterodimers, illustrating a need for asymmetric association. There are at least eight

eukaryotic homologs for MutS: MSH1 (MutS homologue 1), MSH2, MSH3, MSH4, MSH5, MSH6, MSH7, and MSH8, although not all are involved in MMR (27). The complex of MSH2 and MSH6, known as MutSα, is the main player in eukaryotic MMR. It is able to recognize DNA mismatches and single base insertions and deletions (28,29). The complex of MSH2 and MSH3 forms a second mismatch repair heterodimer, MutSβ. This heterodimer is involved in recognizing and repairing IDLs from 2 to 16 nucleotides (30). Neither hMSH6 nor hMSH3 are stable in the absence of the smaller heterodimeric partner hMSH2 (28). MSH1 is found in plants and yeast, and suppresses homeologous mitochondrial DNA exchange (31). A complex formed by MSH4 and MSH5 is found to function in meiotic recombination (27), which is further described below. MSH7 is present in plants and is thought to have evolutionary diverged from MSH6 (32). It also forms a heterodimer with MSH2, known as MutSγ, but this complex preferentially binds to different subsets of mismatches than MutSα (33). Finally, the homolog MSH8 was identified in Euglenozoa, but its function is not characterized at this time (27).

Also for MutL, several different eukaryotic homologs are known: MLH1 (MutL homologue 1), MLH2, MLH3, PMS1 (post-meiotic segregation 1) and PMS2. As is the case for MutS proteins, the eukaryotic MutL homologs form functional heterodimers while the prokaryotic MutL protein acts as a homodimer. MutLα (human MLH1/PMS2 or yeast MLH1/PMS1) is the main MutL protein complex that functions in MMR (34,35), although MutLγ (MLH1/MLH3) has a partially redundant function (36,37). The function of a third heterodimer, MutLβ (human MLH1/PMS1 or yeast MLH1/MLH2), is not fully understood, but it may enhance the function of MutLα in MMR (38,39). No eukaryotic homolog is known for MutH, the endonuclease that is essential for nicking in *E. coli* MMR. Instead, it is suggested that nicking of the DNA is performed by the endonuclease activity of MutL homologs, which resides in the C-terminal domains of PMS2 or MLH3 in humans (18).

**Involvement in other cellular mechanisms**

MMR is directly coupled to replication via interactions of MutS and MutL proteins with the sliding clamps of the replication machinery (β-clamp in prokaryotes and PCNA in eukaryotes) (20,40,41). This ensures that replication can immediately be followed by removal of any incorporated errors. Besides preventing the incorporation of replication errors, MMR also functions to inhibit recombination between non-identical DNA strands (homeologous recombination) (42,43). MMR is then activated by mismatches that arise in regions of recombination intermediates. Furthermore, MMR proteins are found to guard the genome by involvement in cell cycle regulation and signalling functions, among which chromosome pairing during meiosis (44), and DNA damage signalling, especially during G2 phase, leading to apoptosis (45-51).

In contrast, the presence of MMR in cells can also have a mutagenic effect. MMR recognizes and repairs small IDLs and usually has a preventive role in triplet repeat expansion, in which repeat tracts in DNA are increased in size which can cause several common neurodegenerative diseases. However, the expansion of longer repeat tracts is in some cases dependent on the presence of functional MMR proteins (52). Similarly, MMR is involved in promoting hypermutation during the formation of the repertoire of immunoglobulins in B lymphocytes (53,54).

In recombination before the first division of meiosis, MutS homologs MSH4 and MSH5 are known to promote crossover. While MutSα can prevent initiation of recombination by binding to heteroduplex regions in D-loops (55), the MSH4/MSH5 heterodimer binds to Holliday junctions in meiosis to stabilize and preserve recombination intermediates (56). This is essential for proper chromosomal segregation in the first meiotic division, while MSH4/MSH5 may have additional functions in maintaining genome stability (57). The MSH4/MSH5 complex is predicted to be structurally similar to known MutS dimers, but with a larger cavity for DNA (58). Similar to mismatch-repair MutS proteins, the MSH4/MSH5 complex undergoes a ATP-driven conformational change upon recognition of its substrate, forming a sliding clamp that embraces two homologous DNA duplexes. MSH4 interacts with MLH1 and MLH3 during meiosis (59,60) and MutLγ binds to Holliday junctions too (61), while lack of MLH1 or MLH3 results in aberrant meiosis (62,63). This suggests that also in this system, the MutS proteins work together with MutL homologs.

Although the diverse functions of the homologous proteins indicate a complex involvement in eukaryotic cellular mechanisms, the conservation and structural similarities of the homologs suggest that similar mechanisms are utilized. Understanding of the basic mechanism of MutS and MutL proteins can therefore be extrapolated to multiple systems.

## Functional states of MutS

### Structural organization of MutS

MutS proteins of different species are structurally very similar. They form dimers with each subunit consisting of multiple domains (Figure 1A). MutS proteins belong to the ABC (ATP-binding cassette) family of ATPases, of which the members use ATP to regulate their activity (64). The main dimerization interface in MutS proteins is formed by the ATPase domains, thus creating two composite active sites for ATP.

Two different domains are involved in binding of DNA. The clamp domains in the dimer embrace the DNA helix, and are connected to the core of the protein by helices in the lever domains. The mismatch recognition is performed by the N-terminal mismatch-

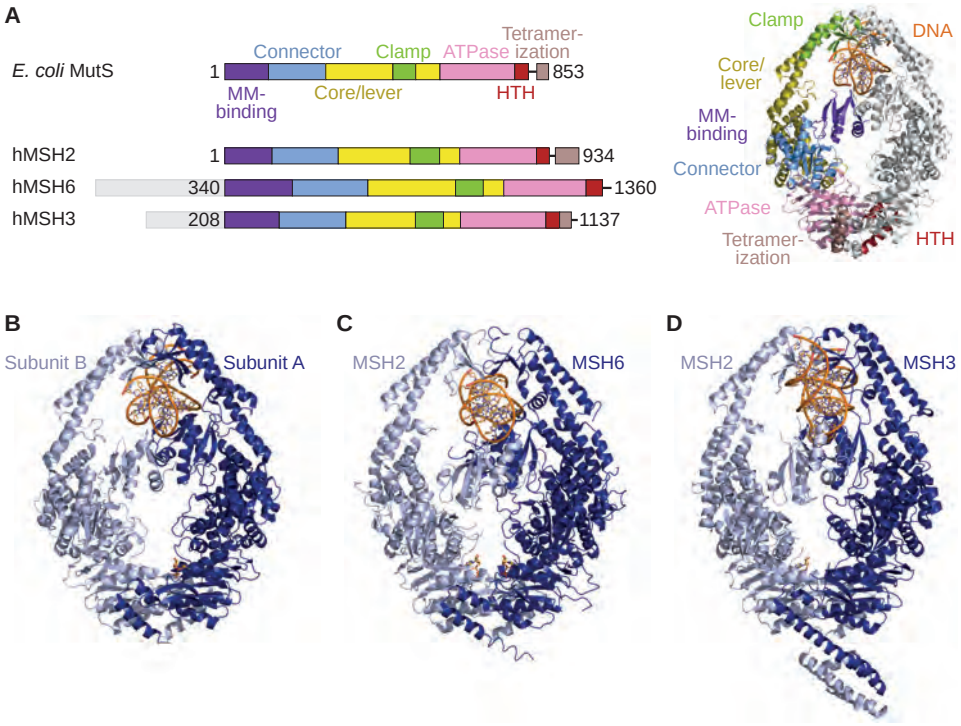



**Figure 1.** MutS proteins. **A)** *E. coli* MutS and human homologs and their domains. Domains are colored as in the schematic representation in the crystal structure of full-length *E. coli* MutS (PDB entry 3ZLJ). **B)** Crystal structure of *E. coli* MutS bound to a GT mismatch (PDB entry 1E3M). **C)** Crystal structure of human MutSα bound to a GT mismatch (PDB entry 2O8B). **D)** Crystal structure of human MutSβ bound to an IDL of 3 bases (PDB entry 3THX).

binding domains of only one subunit in the prokaryotic homodimer (24,25), or specifically by MSH6 in MutSα (65) and predominantly by MSH3 in MutSβ (66).

The mismatch-binding domain is attached to the connector domain. This domain is important for interaction with MutL proteins (67), as are the core and ATPase domains ((68), Figure 3C). Towards the C-terminal end of the dimeric association, there is a helix-turn-helix domain which is required for dimer stability and interacts with the ATPase domain of the opposite subunit (69).

**MutS cycles through different states.**

Conformational changes in MutS allow for switching between the different steps in its mechanism (Figure 2). These steps are linked to its nucleotide status, in which MutS uses ATP to validate mismatch recognition. In solution, MutS homodimers and heterodimers hydrolyze ATP in an asymmetric manner (70-72), which is important for mismatch

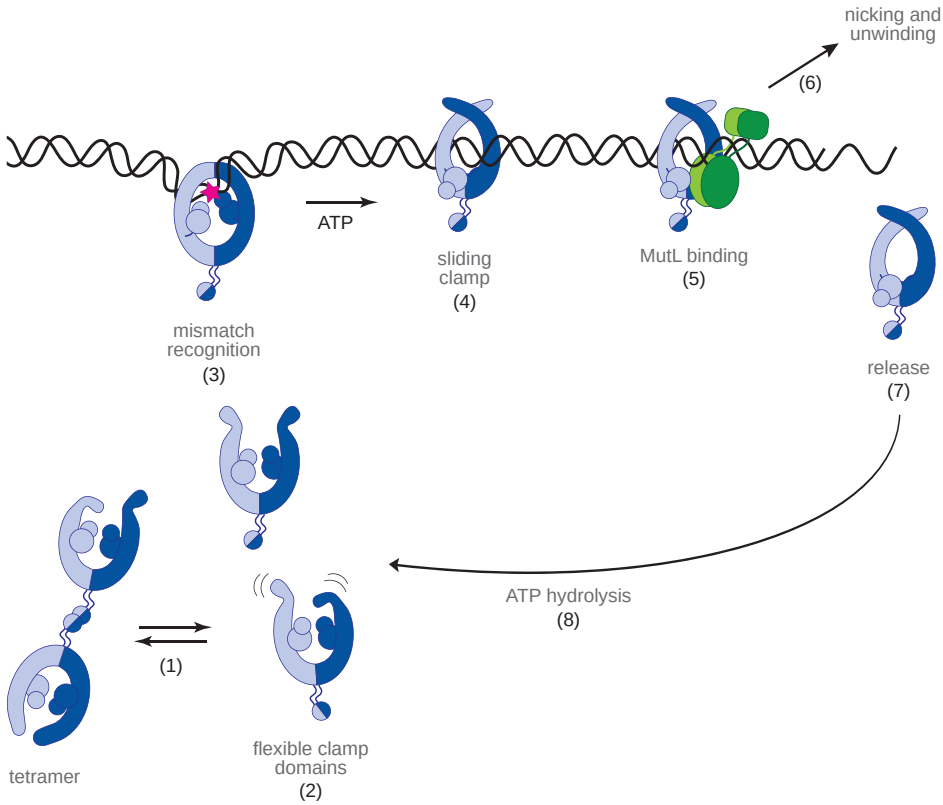



**Figure 2.** Schematic representation of the predominant states of the MutS cycle. (1) In solution, MutS is in equilibrium between tetramers and dimers. (2) The clamp domains are flexible in solution, which allows for initial binding of a DNA helix. (3) A MutS dimer recognizes a DNA mismatch (depicted as pink star), kinking the DNA at the site of the mismatch. (4) After mismatch recognition, MutS forms an ATP-induced sliding clamp on DNA. (5) The MutS sliding clamp binds via two interfaces to MutL, which loads MutL onto DNA. (6) MutL promotes downstream events (endonuclease and helicase activities). (7) MutS releases DNA at a ssDNA region. (8) MutS hydrolyzes ATP and becomes available again for new cycles.

recognition. Once a mismatch has been recognized, the ADP in MutS is rapidly released and exchanged for ATP (73-75), and ATPase activity is suppressed (71,73,76,77). Both binding sites may then become occupied by ATP, which is thought to induce a conformational change in MutS that signals for repair by activating MutL (71,74-76).

The different states of MutS and their structural features will be discussed below.

### MutS in dimer-tetramer equilibrium

*E. coli* MutS can transiently tetramerize through interactions between C-terminal domains of MutS dimers (78,79). Because of the unstable nature of the tetramerization, *E. coli*

MutS exists in equilibrium of dimers and tetramers in solution. The exact function of the tetrameric assembly is unknown, but it has been suggested to be important for the antirecombination function of MutS (80). For MMR, tetramerization is not required, as observed by functional assays with MutS point mutants that do not tetramerize (81,82). Truncation of the tetramerization domain results in a mutator phenotype (83), but this is probably due to the decreased dimer stability of the truncation variant. The tetramerization of *E. coli* MutS results in prolonged DNA binding by the protein, since the two dimers within a tetramer can simultaneously bind DNA, making release a slow event (79).

For the eukaryotic MutS homologs, tetramerization has not been reported, but C-terminus of MutSβ shows a very similar double helix-loop-helix fold as the C-terminal domain of the *E. coli* protein (66), and it is possible that this domain can support tetramerization as well. A difference between these proteins, however, is the way that the C-terminal domain is linked to the rest of the protein. In *E. coli* MutS, the linker region is flexible, which allows 'bending over' of the dimers within the tetramer (79). In the human protein, the C-terminal domains are linked by a α-helix of MSH3 as shown in the crystal structure, suggesting a more rigid conformation.

### Pre-DNA binding state

DNA-free crystal structures of MutS proteins show disorder of the DNA clamp domains (25) or kinking of the helices in the lever domain into different directions (Figure 3A, Chapter 3). This indicates that, in the absence of DNA, there is conformational flexibility of the clamp domains that can embrace the DNA helix. Such conformational flexibility

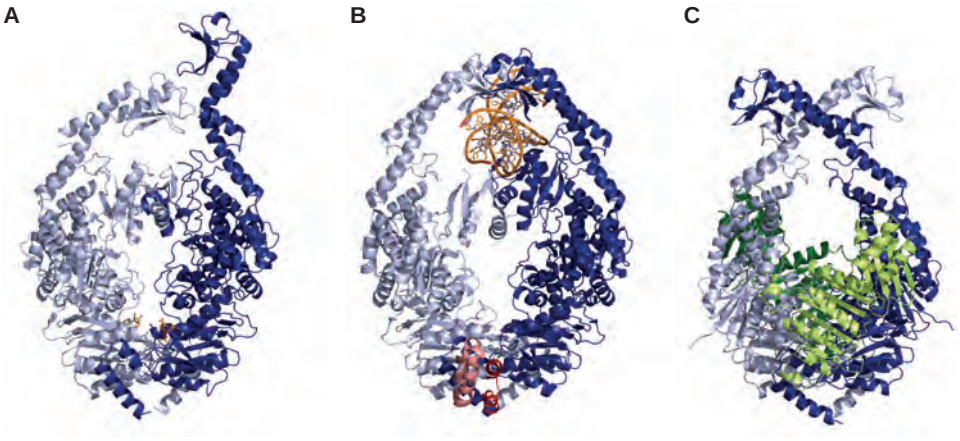



**Figure 3.** Crystal structures of *E. coli* MutS in different conformations. **A)** DNA-free MutS (Chapter 3), **B)** mismatch-bound MutS with its tetramerization domains shown in salmon and red (PDB entry 3ZLJ, Chapter 2), **C)** the MutS sliding clamp bound to the N-terminal domain of MutL (green) (Chapter 4).

suggests a way of 'opening up' to allow DNA to enter the MutS dimer. The opening of the clamp domains would then also function to release from DNA if no mismatch was recognized by the mismatch-binding domains.

The kinking occurs in the same two helices (at residues 441-443 and 515-517) in both subunits of the DNA-free *E. coli* dimer (Figure 3A), which suggests that this is a specific hinge point for this movement. This region already displayed high B-factors (indicating some degree of disorder) in previous crystal structures of DNA-bound MutS. This hypothesis is supported by preliminary assays in which mutations in these specific helical regions, attempting to stabilize (prevent kinking) or destabilize (promote kinking) them, influence DNA binding kinetics (Chapter 3). This shows the need for a balance between a stable helical fold and the possibility to kink.

**Mismatch-recognition state**

Several crystal structures of MutS or homolog recognizing DNA mismatches have been published (Figure 1, Figure 3B): *E. coli* MutS bound to DNA mismatches (24,26,79,84), *Thermus aquaticus* MutS bound to DNA containing an unpaired thymidine (25), the human MutSα bound to different DNA substrates (65) and human MutSβ bound to different DNA substrates (66). All these structures show similar arrangement of MutS dimers, in which the clamp domains embrace the DNA duplex and one of the mismatch binding domains contacts the DNA at the site of the mismatch (Figure 1, Figure 3B).

During mismatch recognition, MutS proteins kink the DNA at the site of the mismatch; in *E. coli* by an angle of approximately 60 degrees (24). It is thought that this is a way to sample the reduced helical stability due to distorted base pairing of the DNA (85). In *E. coli*, the asymmetry of mismatch recognition initially results in asymmetry of the ATPase domains in which the mismatch-contacting subunit binds ADP (24,75).

**Sliding-clamp state**

Upon mismatch recognition, MutS exchanges its ADP for ATP and undergoes a conformational change to form a long-lived sliding clamp on DNA (12,86,87). This can be shown in assays in which MutS releases fast from DNA ends in the
presence of ATP, whereas MutS remains stably bound to the DNA when the ends are blocked (79,88). It is generally accepted that the sliding clamp is a state in which MutS signals for repair.

The sliding clamp conformation of MutS is revealed for the first time in the crystal structure where it is in complex with the N-terminal domain of MutL (Figure 3C, Chapter 4). In this conformation, the subunits in the dimer are tilted across each other, hinging from the ATPase domains (Figure 3C). At the same time, the mismatch-binding and connector domains have rotated outward. The conformational change pushes the DNA

downward into a new channel, and MutS will be bound as a loose ring around the DNA, explaining how MutS can diffuse on the DNA helix.

**Release and recycling**

As MutS is an ATPase protein, it will hydrolyze its bound ATP over time. The two MutS subunits have ATPase domains with alternating ATP hydrolysis activity (70,89). In the absence of DNA, MutS hydrolyzes ATP and does not rapidly rebind new ATP (70).

After mismatch recognition and signalling for repair, it is possible that the MutS sliding clamp releases DNA at single-stranded regions (12) where replication is not complete yet, or where removal of the new strand has already started. MutS will then hydrolyze its bound ATP, and revert to its pre-recognition state in which it is available again for new cycles of mismatch recognition. If the original DNA mismatch is not repaired yet, MutS may bind to it again. This allows for a model where MutS proteins have multiple loading starting from the mismatch, and activate downstream processors until the error has been removed. Such rebinding would also be able to account for directionality of the system towards the mismatch, guided by repeated MutS localization.

**Comparison of MutS with other ABC proteins**

The ABC superfamily of ATPases, to which MutS belongs, contains proteins that are involved in important cellular processes such as membrane transport, DNA repair and chromosome condensation. ABC proteins contain several conserved motifs in the nucleotide binding domains over which they dimerize forming two composite ATPase domains. The basic mechanism in these proteins results from ATP-binding-induced compacting of the dimers, which can propagate as a movement in the rest of the protein (Figure 4). ATPase activity is then thought to reset the proteins for further cycles (64,90).

Transmembrane ABC transporters actively transfer substrates across cellular membranes. These transporters undergo large conformational changes in which the ATP-bound state opens up one side of the protein to take up (importers) or release (exporters) the substrate, while ATP hydrolysis and nucleotide release allows the substrate to release (importers) or enter (exporters) from the other side (91,92). The MR complex (Mre11 nuclease and Rad50 ABC ATPase), which is a sensor for DNA double-strand breaks, uses ATP binding by the two Rad50 subunits to modulate its structure. This creates a clamp conformation with increased binding to DNA ends (93). In SMC proteins, the function of ATP binding by the head domains is not fully understood. It has been proposed to allow closure of an SMC dimer as a ring around multiple DNA strands (94), but this ring formation around DNA may instead require transient opening the other dimerization interface of SMC dimers (95).

The cycle of MutS proteins is comparable to that of ABC transporters: ATP-induced

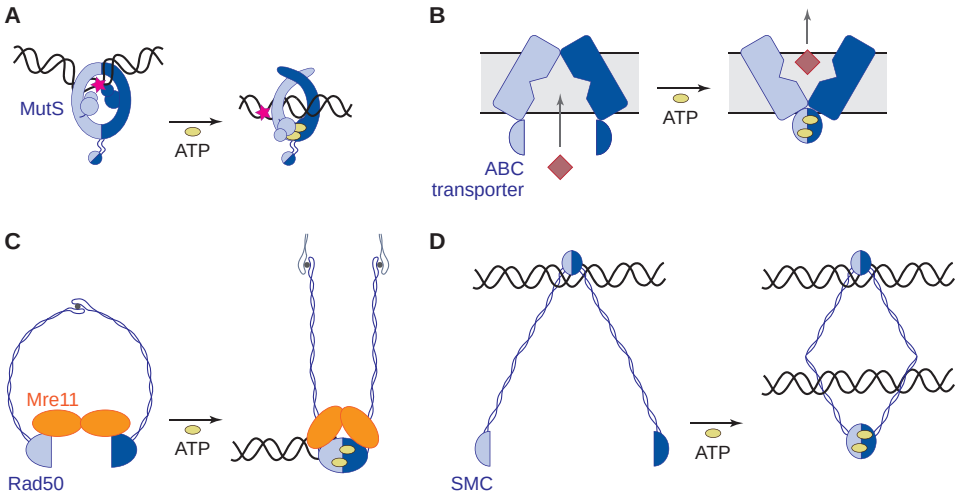



**Figure 4.** Schematic representation of ATP-driven motions in different ABC proteins. **A)** ATP binding induces a hinge motion that translocates mismatched DNA to a new channel in MutS proteins. **B)** ATP binding by ABC transporters opens up the protein on the opposite side, thus transporting substrate across a membrane (exporter visualized). **C)** ATP binding by Rad50/Mre11 modulates the protein structure to increase binding to DNA ends (93). **D)** Model in which ATP binding induces dimerization of the head domains of SMC proteins, possibly entrapping DNA in this way (94).

compaction of the ATPase domains transmits a hinge-like motion to the rest of the protein, thus actively translocating it substrate DNA to the sliding clamp channel. Different from ATP transporters, however, hydrolysis of ATP is not known to destabilize the dimerization of the ATPase domains such that it opens up the protein from the other side. The flexible linkers to the dimerized C-terminal domains in *E. coli* MutS would also make it difficult for the dimer to completely open up on that side.

## MutL proteins

### Structural organization of MutL

MutL is essential for MMR and is activated by the MutS sliding clamp, linking this MutS signal to downstream repair effectors (8,10). *E. coli* MutL activates the endonuclease MutH to nick the newly synthesized strand, while most other MutL proteins contain an endonuclease motif. This endonuclease activity resides in PMS2 or MLH3 for the human homologs (18). The nick that is thus made is required to unwind and remove part of the new DNA strand, for which *E. coli* MutL activates the helicase UvrD (17).

MutL proteins form dimers that contain flexible linkers separating the C-terminal from the N-terminal domains (Figure 5). The primary dimerization interface is formed by the C-terminal domains, while the N-termini contain ATP-binding sites that become more structured and dimerize in the presence of ATP (96,97). These N-terminal ATP-binding domains are structurally similar to the ATPase domains of members of the GHKL family (98). Upon ATP binding, the N-terminal domains of MutL proteins have been found to physically interact with the C-terminal domains, potentially altering the endonuclease function in these domains (97,99). Since MutL contains weak ATPase activity, the protein is thought to go through a cycle of dimerization and becoming more compact, until ATP is hydrolyzed and the nucleotide is released (97,100) (Figure 6A).

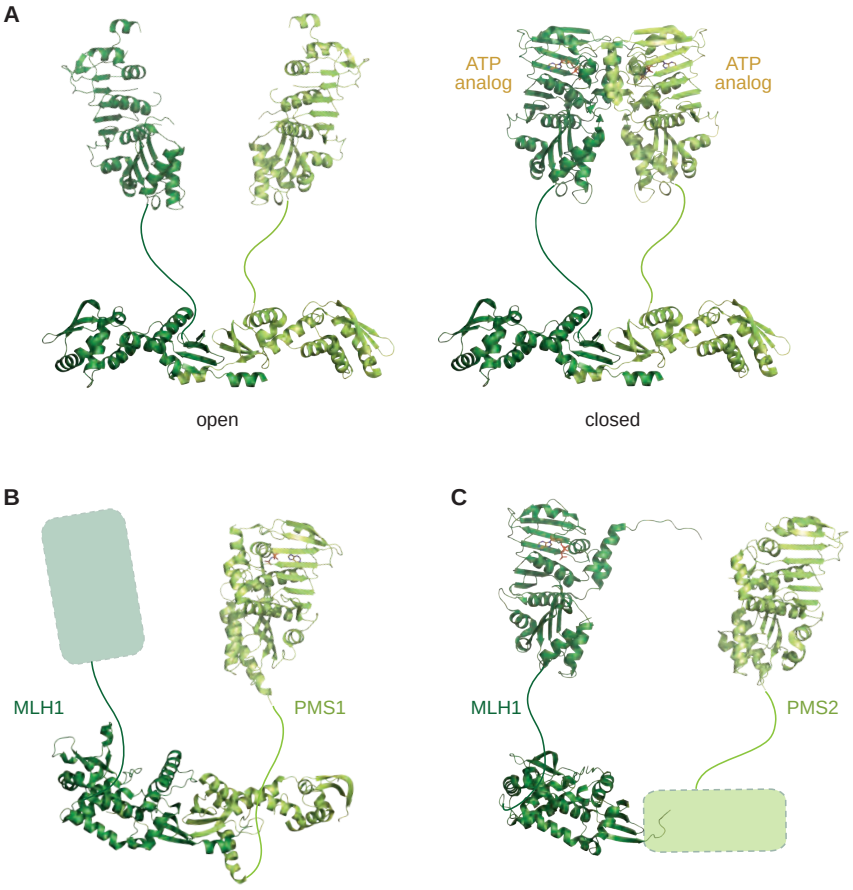



**Figure 5.** Structural models for full-length MutL proteins of three organisms. **A)** The *E. coli* MutL homodimer with open (without nucleotide, PDB entries 1X9Z and 1BKN) and closed (AMP-PNP bound, PDB entries 1X9Z and 1NHJ) N-terminal domains. **B)** Yeast MutLα (PDB entries 4E4W and 3H4L). **C)** Human MutLα (PDB entries 3RBN, 3NA3, and 1H7S).

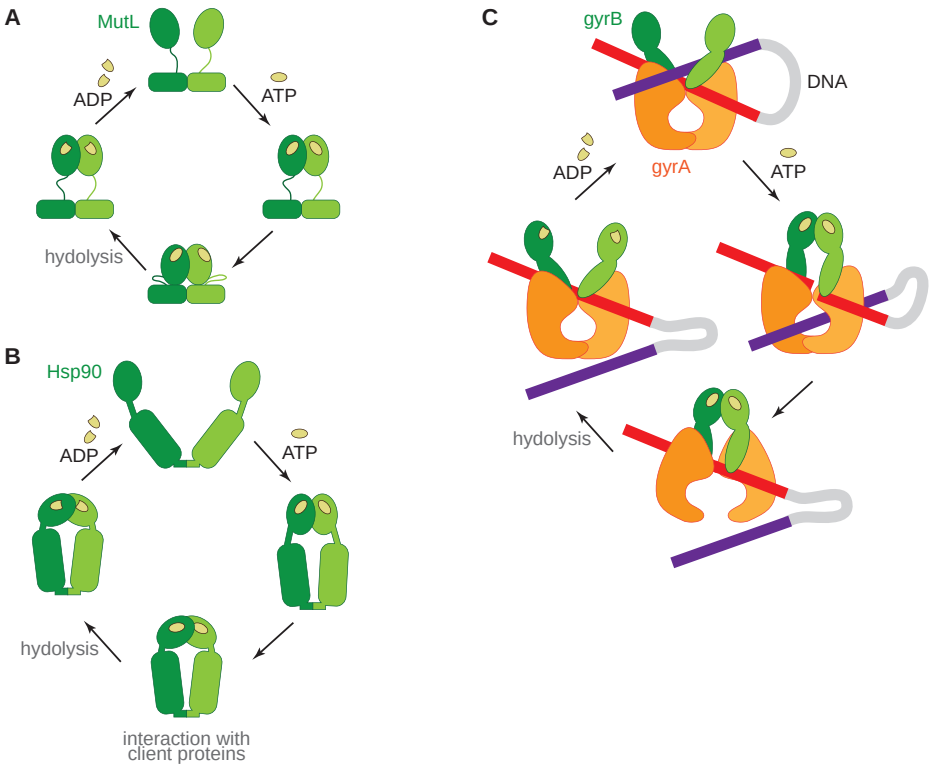



**Figure 6.** Schematic representation of ATP-driven conformational changes in different members of the GHKL-family. **A)** ATP binding induces dimerization of the N-terminal domains of MutL and compaction of the dimer until hydrolysis takes place. **B)** ATP binding by Hsp90 induces dimerization of the N-terminal domains and facilitates interaction with client proteins. **C)** ATP-induced conformational changes in the gyrB subunit of DNA gyrase allow for strand passage through the gyrA subunit.

The dimerization of the N-terminal domains forms a potential DNA-binding groove, and mutations in this groove can reduce the (weak) DNA-binding ability of MutL (101,102). The closing of the N-terminal domains also results in the formation of a central channel between the N-terminal and C-terminal domains, which could allow MutL to enclose a DNA helix. This DNA binding may also involve specific interactions of the flexible linker regions (103). It is mainly the N-terminal domain that interacts with MutS, in which MLH1 is involved in the eukaryotic heterodimers (35,104). The crystal structure of *E. coli* MutS in complex with the N-terminal domain of MutL suggests that MutL is loaded onto DNA by binding to the MutS sliding clamp, and this is required for activation of MutH (Chapter 4).

### Comparison of MutL with other GHKL proteins

The GHKL (DNA gyrase, Hsp90, histidine kinase, MutL) superfamily comprises proteins

with diverse functions that share the nucleotide-binding 'Bergerat' fold (105). These proteins contain several conserved motifs in their ATPase domains. Most members of the GHKL family form dimers in which ATP binding and/or hydrolysis induces large conformational changes due to transient dimerization of the ATPase domains (Figure 6).

The MutL cycle can be compared to that of the chaperone Hsp90, which is also a constitutive dimer, with N-terminal domains that dimerize upon ATP binding (106). Stabilization and conformational changes of the N-terminal domains of Hsp90 due to nucleotide binding allow binding to client proteins. Similar regulation is suggested for binding of N-terminal domain of MutL for binding to MutS or downstream effectors in MMR (107). In the case of DNA gyrase, the dimerization of ATPase domains occurs in the gyrB subunit and this motion is transmitted to the gyrA subunit, which allows for passage of a DNA strand, thus adjusting supercoiling of DNA (108).

## Outline of this thesis

The work presented in this thesis is aimed to understand the nature of the different conformations of MutS, and in particular the manner in which these conformations allow MutS to function in the context of DNA and MutL in the initial steps of MMR.

In *Chapter 2* we describe the dimer-tetramer equilibrium of full-length MutS. We uncouple these two oligomerization states to study them separately. Moreover, we use a stable dimer variant of MutS to measure mismatch recognition specificity, DNA binding kinetics, and sliding-clamp formation.

In *Chapter 3* we report a crystal structure of DNA-free MutS. We propose that this structure represents a state that precedes DNA binding, and show that site-specific kinking in helices of the lever domains allows DNA to enter the MutS dimer.

In *Chapter 4* we present a crystal structure of the MutS in the sliding clamp state, in complex with the N-terminal domain of MutL. The structure shows large conformational changes in MutS and suggests relocation of the DNA upon mismatch recognition and ATP binding, which we validate using biophysical assays. This chapter explains how the MutS sliding clamp can specifically activate MutL to promote downstream repair.

In *Chapter 5* we summarize the highlights of Chapters 2-4 and discuss their implications in the context of the MMR mechanism.

## References

1. J.D. Watson and F.H.C. Crick (1953) Molecular structure of nucleic acids. *Nature*, **171**, 737–738.
2. C.J. Su and J.B. Baseman (2009) Genome size of *Mycoplasma genitalium*. *J. Bacteriol.*, **172**, 4705–

4707.

3. T.R. Gregory, J.A. Nicol, H. Tamm, B. Kullman, K. Kullman, I.J. Leitch, B.G. Murray, D.F. Kapraun, J. Greilhuber and M.D. Bennett (2007) Eukaryotic genome size databases. *Nucleic Acids Res.*, **35**, D332–D338.
4. S.D. McCulloch and T.A. Kunkel (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.*, **18**, 148–161.
5. R.R. Iyer, A. Pluciennik, V. Burdett and P.L. Modrich (2006) DNA mismatch repair: functions and mechanisms. *Chem. Rev.*, **106**, 302–323.
6. H.T. Lynch and A. de la Chapelle (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.*, **36**, 801–818.
7. P. Modrich and R. Lahue (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.*, **65**, 101–133.
8. T.A. Kunkel and D.A. Erie (2005) DNA mismatch repair. *Annu. Rev. Biochem.*, **74**, 681–710.
9. J. Jiricny (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell. Biol.*, **7**, 335–346.
10. J. Jiricny (2013) Postreplicative mismatch repair. *Cold Spring Harbor Perspect. Biol.*, **5**, a012633.
11. D.J. Allen, A. Makhov, M. Grilley, J. Taylor, R. Thresher, P. Modrich and J.D. Griffith (1997) MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.*, **16**, 4467–4476.
12. C. Jeong, W. Cho, K. Song, C. Cook, T. Yoon, C. Ban, R. Fishel and J. Lee (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat Struct Mol Biol*, **18**, 379–385.
13. M. Grilley, K.M. Welsh, S. Su and P. Modrich (1989) Isolation and characterization of the *Escherichia coli mutL* gene product. *J. Biol. Chem.*, **264**, 1000–1004.
14. K. Drotschmann, A. Aronshtam, H.J. Fritz and M.G. Marinus (1998) The *Escherichia coli* MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic Acids Res.*, **26**, 948–953.
15. S. Acharya, P.L. Foster, P. Brooks and R. Fishel (2003) The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair. *Mol. Cell*, **12**, 233–246.
16. K.G. Au, K. Welsh and P. Modrich (1992) Initiation of methyl-directed mismatch repair. *J. Biol. Chem.*, **267**, 12142–12148.
17. M. Yamaguchi, V. Dao and P. Modrich (1998) MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J. Biol. Chem.*, **273**, 9197–9201.
18. F.A. Kadyrov, L. Dzantiev, N. Constantin and P. Modrich (2006) Endonucleolytic function of MutLα in human mismatch repair. *Cell*, **126**, 297–306.
19. S.A.N. McElhinny, G.E. Kisslin and T.A. Kunkel (2010) Differential correction of lagging-strand replication errors made by DNA polymerases α and δ. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21070–21075.
20. H.E. Kleczkowska, G. Marra, T. Lettieri and J. Jiricny (2001) hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. *Genes Dev.*, **15**, 724–736.
21. S.D. Lee and E. Alani (2006) Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *J. Mol. Biol.*, **355**, 175–184.
22. G. Li (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res.*, **18**, 85–98.
23. M. Elez, M. Radman and I. Matic (2012) Stoichiometry of MutS and MutL at unrepaired mismatches *in vivo* suggests a mechanism of repair. *Nucleic Acids Res.*, **40**, 3929–3938.
24. M.H. Lamers, A. Perrakis, J.H. Enzlin, H.H.K. Winterwerp, N. de Wind and T.K. Sixma (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature*, **407**, 711–717.

25. G. Obmolova, C. Ban, P. Hsieh and W. Yang (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703–710.
26. G. Natrajan, M.H. Lamers, J.H. Enzlin, H.H.K. Winterwerp, A. Perrakis and T.K. Sixma (2003) Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.*, **31**, 4814–4821.
27. P. Sachadyn (2010) Conservation and diversity of MutS proteins. *Mutation Res.*, **694**, 20–30.
28. S. Acharya, T. Wilson, S. Gradia, M.F. Kane, S. Guerrette, G.T. Marsischky, R.D. Kolodner and R. Fishel (1996) hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 13629–13634.
29. E. Alani (1996) The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol. Cell. Biol.*, **16**, 5604–5615.
30. Y. Habraken, P. Sung, L. Prakash and S. Prakash (1996) Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Curr. Opin. Genet. Dev.*, **6**, 1185–1187.
31. Y.Z. Xu, M.P. Arrieta-Montiel, K.S. Virdi, W.B.M. de Paula, J.R. Widhalm, G.J. Basset, J.I. Davila, T.E. Elthon, C.G. Elowsky, S.J. Sato, T.E. Clemente and S.A. Mackenzie (2011) MutS HOMOLOG1 is a nucleoid protein that alters mitochondrial and plastid properties and [lant response to high light. *Plant Cell*, **23**, 3428–3441.
32. S.M. Tam, S. Samipak, A. Britt and R.T. Chetelat (2009) Characterization and comparative sequence analysis of the DNA mismatch repair MSH2 and MSH7 genes from tomato. *Genetica*, **137**, 341–354.
33. S.Y. Wu (2003) Dissimilar mispair-recognition spectra of *Arabidopsis* DNA-mismatch-repair proteins MSH2·MSH6 (MutSα) and MSH2·MSH7 (MutSγ). *Nucleic Acids Res.*, **31**, 6027–6034.
34. G. Plotz, J. Raedle, A. Brieger, J. Trojan and S. Zeuzem (2002) hMutSα forms an ATP-dependent complex with hMutLα and hMutLβ on DNA. *Nucleic Acids Res.*, **30**, 711–718.
35. G. Plotz, J. Readle, A. Brieger, J. Trojan and S. Zeuzen (2003) N-terminus of hMLH1 confers interaction of hMutLα and hMutLβ with hMutSα. *Nucleic Acids Res.*, **31**, 3217–3226.
36. E. Cannavo, G. Marra, J. Sabates-Bellver, M. Menigatti, S.M. Lipkin, F. Fischer, P. Cejka and J. Jiricny (2005) Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res.*, **65**, 10759–10766.
37. M.K. Korhonen, E. Vuorenmaa and M. Nyström (2008) The first functional study of *MLH3* mutations found in cancer patients. *Genes Chromosom. Cancer*, **47**, 803–809.
38. M. Raschle, G. Marra, M. Nyström-Lahti, P. Schar and J. Jiricny (1999) Identification of hMutLβ, a heterodimer of hMLH1 and hPMS1. *J. Biol. Chem.*, **274**, 32368–32375.
39. C.S. Campbell, H. Hombauer, A. Srivatsan, N. Bowen, K. Gries, A. Desai, C.D. Putnam and R.D. Kolodner (2014) Mlh2 Is an accessory factor for DNA mismatch repair in *Saccharomyces cerevisiae*. *PLoS Genet*, **10**, e1004327.
40. F.J. López de Saro and M. O'Donnell (2001) Interaction of the β sliding clamp with MutS, ligase and DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 8376–8380.
41. F.J. López de Saro, M.G. Marinus, P. Modrich and M. O'Donnell (2006) The β sliding clamp binds to multiple sites within MutL and MutS. *J. Biol. Chem.*, **281**, 14340–14349.
42. A. Datta, A. Adjiri, L. New, G.F. Crouse and S. Jinks-Robertson (1996) Mitotic crossovers between diverged sequences are regulated by mismatch repair proteins in *Saccaromyces cerevisiae*. *Mol. Cell. Biol.*, **16**, 1085–1093.
43. K.C. Tham, N. Hermans, H.H.K. Winterwerp, M.M. Cox, C. Wyman, R. Kanaar and J.H.G.

Lebbink (2013) Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. *Mol. Cell*, **51**, 326–337.

44. C. Rudolph, C. Kunz, S. Parisi, E. Lehmann, E. Hartsuiker, B. Fartmann, W. Kramer, J. Kohli and O. Fleck (1999) The *msh2* gene of *Schizosaccharomyces pombe* is involved in mismatch repair, mating-type switching, and meiotic chromosome organization. *Mol. Cell. Biol.*, **19**, 241–250.
45. M.T. Hawn, A. Umar, J.M. Carethers, G. Marra, T.A. Kunkel, C.R. Boland and M. Koi (1995) Evidence for a connection between the mismatch repair system and the $G_2$ cell cycle checkpoint. *Cancer Res.*, **55**, 3721–3725.
46. M.J. Hickman and L.D. Samson (1999) Role of DNA mismatch repair and p53 in signaling induction of apoptosis by alkylating agents. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 10764–10769.
47. D.R. Duckett, S.M. Bronstein, Y. Taya and P. Modrich (1999) hMutSα- and hMutLα-dependent phosphorylation of p53 in response to DNA methylator damage. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 12384–12388.
48. B.S. Zhou and S.J. elledge (2000) The DNA damage response: putting checkpoints in perspective. *Nature*, **408**, 433–439.
49. E.J. Brown and D. Baltimore (2003) Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. *Genes Dev.*, **17**, 615–628.
50. K.D. Brown, A. Rathi, R. Kamath, D.I. Beardsley, Q. Zhan, J.L. Mannino and R. Baskaran (2003) The mismatch repair system is required for S-phase checkpoint activation. *Nat. Genet.*, **33**, 80–84.
51. L. Stojic, R. Brun and J. Jiricny (2004) Mismatch repair and DNA damage signalling. *DNA Repair*, **3**, 1091–1101.
52. A. Jaworski, W.A. Rosche, R. Gellibolian, S. Kang, M. Shimizu, R.P. Bowater, R.R. Sinden and R.D. Wells (1995) Mismatch repair in *Escherichia coli* enhances instability of $(CTG)_n$ triplet repeats from human hereditary diseases. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 11019–11023.
53. S. Schanz, D. Castor, F. Fischer and J. Jiricny (2009) Interference of mismatch and base excision repair during the processing of adjacent U/G mispairs may play a key role in somatic hypermutation. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 5593–5598.
54. R. Chahwan, W. Edelmann, M.D. Scharff and S. Roa (2011) Mismatch-mediated error prone repair at the immunoglobulin genes. *Biomed. Pharmacother.*, **65**, 529–536.
55. M. Honda, Y. Okuno, S.R. Hengel, J.V. Martín-López, C.P. Cook, R. Amunugama, R.J. Soukup, S. Subramanyanm, R. Fishel and M. Spies (2014) Mismatch repair protein hMSH2-hMSH6 recognizes mismatches and forms sliding clamps within a D-loop intermediate. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E316–E325.
56. T. Snowden, S. Acharya, C. Butz, M. Berardini and R. Fishel (2004) hMSH4-hMSH5 recognizes Holliday junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Mol. Cell*, **15**, 437–451.
57. N. Clark, X. Wu and C. Her (2013) MutS homologues hMSH4 and hMSH5: genetic variations, functions, and implications in human diseases. *Curr. Genomics*, **14**, 81–90.
58. R. Rakshambikai, N. Srinivasan and K.T. Nishant (2013) Structural insights into *Saccharomyces cerevisiae* Msh4–Msh5 complex function using homology modeling. *PloS One*, **8**, e78753.
59. S. Santucci-Darmanin, D. Walpita, F. Lespinasse, C. Desnuelle, T. Ashley and V. Paquis-Flucklinger (2000) MSH4 acts in conjunction with MLH1 during mammalian meiosis. *FASEB J.*, **14**, 1539–1547.
60. S. Santucci-Darmanin, S. Neyton, F. Lespinasse, A. Saunières, P. Gaudray and V. Paquis-Flucklinger (2002) The DNA mismatch-repair MLH3 protein interacts with MSH4 in meiotic cells, supporting a role for this MutL homolog in mammalian meiotic recombination. *Hum. Mol. Gen.*, **11**, 1697–

1706.

61. L. Ranjha, R. Anand and P. Cejka (2014) The *Saccharomyces cerevisiae* Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. *J. Biol. Chem.*, **289**, 5674–5686.
62. S.M. Lipkin, P.B. Moens, V. Wang, M. Lenzi, D. Shanmugarajah, A. Gilgeous, J. Thomas, J. Cheng, J.W. Touchman, E.D. Green, P. Schwartzberg, F.S. Collins and P.E. Cohen (2002) Meiotic arrest and aneuploidy in MLH3-deficient mice. *Nat. Genet.*, **31**, 385–390.
63. R. Kan, X. Sun, N.K. Kolas, E. Avdievich, B. Kneitz, W. Edelmann and P.E. Cohen (2008) Comparative analysis of meiotic progression in female mice bearing mutations in genes of the DNA mismatch repair pathway. *Biol. Reprod.*, **78**, 462–471.
64. K. Hopfner and J.A. Tainer (2003) Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. *Curr. Opin. Struct. Biol.*, **13**, 249–255.
65. J.J. Warren, T.J. Pohlhaus, A. Changela, R.R. Iyer, P.L. Modrich and L.S. Beese (2007) Structure of the human MutSα DNA lesion recognition complex. *Mol. Cell*, **26**, 579–592.
66. S. Gupta, M. Gellert and W. Yang (2011) Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat Struct Mol Biol*, **19**, 72–78.
67. M.L. Mendillo, V.V. Hargreaves, J.W. Jamison, A.O. Mo, S. Li, C.D. Putnam, V.L. Woods Jr and R.D. Kolodner (2009) A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 22223–22228.
68. J.S. Lenhart, M.C. Pillon, A. Guarné and L.A. Simmons (2013) Trapping and visualizing intermediate steps in the mismatch repair pathway *in vivo*. *Mol. Microbiol.*, **90**, 680–698.
69. I. Biswas, G. Obmolova, M. Takahashi, A. Herr, M.A. Newman, W. Yang and P. Hsieh (2001) Disruption of the helix-u-turn-helix motif of MutS protein: loss of subunit dimerization, mismatch binding and ATP hydrolysis. *J. Mol. Biol.*, **305**, 805–816.
70. M.H. Lamers, H.H.K. Winterwerp and T.K. Sixma (2003) The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J.*, **22**, 746–756.
71. E. Antony and M.M. Hingorani (2003) Mismatch recognition-coupled stabilization of Msh2-Msh6 in an ATP-bound state at the initiation of DNA repair. *Biochemistry*, **42**, 7682–7693.
72. E. Antony and M.M. Hingorani (2004) Asymmetric ATP binding and hydrolysis activity of the *Thermus aquaticus* MutS dimer is key to modulation of its interactions with mismatched DNA. *Biochemistry*, **43**, 13115–13128.
73. E. Jacobs-Palmer and M.M. Hingorani (2007) The effects of nucleotides on MutS-DNA binding kinetics clarify the role of MutS ATPase activity in mismatch repair. *J. Mol. Biol.*, **366**, 1087–1098.
74. C.D. Heinen, J.L. Cyr, C. Cook, N. Punja, M. Sakato, R.A. Forties, J.M. Lopez, M.M. Hingorani and R. Fishel (2011) hMSH2 controls ATP processing by hMSH2-hMSH6. *J. Biol. Chem.*, **286**, 40287–40295.
75. M.C. Monti, S.X. Cohen, A. Fish, H.H.K. Winterwerp, A. Barendregt, P. Friedhoff, A. Perrakis, A.J.R. Heck, T.K. Sixma, R.H.H. van den Heuvel and J.H.G. Lebbink (2011) Native mass spectrometry provides direct evidence for DNA mismatch-induced regulation of asymmetric nucleotide binding in mismatch repair protein MutS. *Nucleic Acids Res.*, **39**, 8052–8064.
76. E. Antony, S. Khubchandani, S. Chen and M.M. Hingorani (2006) Contribution of Msh2 and Msh6 subunits to the asymmetric ATPase and DNA mismatch binding activities of *Saccharomyces cerevisiae* Msh2-Msh6 mismatch repair protein. *DNA Repair*, **5**, 153–162.
77. D.J. Mazur, M.L. Mendillo and R.D. Kolodner (2006) Inhibition of Msh6 ATPase activity by mispaired DNA induces a Msh2(ATP)-Msh6(ATP) state capable of hydrolysis-independent movement along DNA. *Mol. Cell*, **22**, 39–49.
78. K.P. Bjornson, L.J. Blackwell, H. Sage, C. Baitinger, D. Allen and P. Modrich (2003) Assembly and

molecular activities of the MutS tetramer. *J. Biol. Chem.*, **278**, 34667–34673.

79. F.S. Groothuizen, A. Fish, M.V. Petoukhov, A. Reumer, L. Manelyte, H.H.K. Winterwerp, M.G. Marinus, J.H.G. Lebbink, D.I. Svergun, P. Friedhoff and T.K. Sixma (2013) Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp formation. *Nucleic Acids Res.*, **41**, 8166–8181.
80. M.A. Calmann, A. Nowosielska and M.G. Marinus (2005) Separation of mutation avoidance and antirecombination functions in an *Escherichia coli mutS* mutant. *Nucleic Acids Res.*, **33**, 1193–1200.
81. L. Manelyte, C. Urbanke, L. Giron-Monzon and P. Friedhoff (2006) Structural and functional analysis of the MutS C-terminal tetramerization domain. *Nucleic Acids Res.*, **34**, 5270–5279.
82. M.L. Mendillo, C.D. Putnam and R.D. Kolodner (2007) *Escherichia coli* MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair *in vivo*. *J. Biol. Chem.*, **282**, 16345–16354.
83. M.A. Calmann, A. Nowosielska and M.G. Marinus (2005) The MutS C terminus is essential for mismatch repair activity in vivo. *J. Bacteriol.*, **187**, 6577–6579.
84. M.H. Lamers, D. Georgijevic, J.H. Lebbink, H.H.K. Winterwerp, B. Agianian, N. de Wind and T.K. Sixma (2004) ATP increases the affinity between MutS ATPase domains: implications for ATP hydrolysis and conformational changes. *J. Biol. Chem.*, **279**, 43879–438885.
85. T.K. Sixma (2001) DNA mismatch repair: MutS structures bound to mismatches. *Curr. Opin. Struct. Biol.*, **11**, 47–52.
86. S. Gradia, D. Subramanian, T. Wilson, S. Acharya, A. Makhov, J. Griffith and R. Fishel (1999) hMSH2–hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell*, **3**, 255–261.
87. W. Cho, C. Jeong, D. Kim, M. Chang, K. Song, J. Hanne, C. Ban, R. Fishel and J. Lee (2012) ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure*, **20**, 1–11.
88. M.J. Schofield, S. Nayak, T.H. Scott, C. Du and P. Hsieh (2001) Interaction of *Escherichia coli* MutS and MutL at a DNA mismatch. *J. Biol. Chem.*, **276**, 28291–28299.
89. K. Drotschmann, W. Yang and T.A. Kunkel (2002) Evidence for sequential action of two ATPase active sites in yeast Msh2-Msh6. *DNA Repair*, **1**, 743–753.
90. J.H.G. Lebbink and T.K. Sixma (2005) Variations on the ABC. *Structure*, **13**, 498–500.
91. P.M. Jones, M.L. O'Mara and A.M. George (2009) ABC transporters: a riddle wrapped in a mystery inside an enigma. *Trends Biochem. Sci.*, **34**, 520–531.
92. C.A. Shintre, A.C.W. Pike, Q. Li, J.I. Kim, A.J. Barr, S. Goubin, L. Shrestha, J. Yang, G. Berridge, J. Ross, P.J. Stansfeld, M.S.P. Sansom, A.M. Edwards, C. Bountra, B.D. Marsden, F. von Delft, A.N. Bullock, O. Gileadi, N.A. Burgess-Brown and E.P. Carpenter (2013) Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 9710–9715.
93. K. Lammens, D.J. Bemeleit, C. Möckel, E. Clausing, A. Schele, S. Hartung, C.B. Schiller, M. Lucas, C. Angermüller, J. Söding, K. Sträßer and K. Hopfner (2011) The Mre11:Rad50 structure shows an ATP-dependent molecular clamp in DNA double-strand break repair. *Cell*, **145**, 54–66.
94. M. Hirano and T. Hirano (2006) Opening closed arms: long-distance activation of SMC ATPase by hinge-DNA interactions. *Mol. Cell*, **21**, 175–186.
95. S. Gruber, P. Arumugam, Y. Katou, D. Kuglitsch, W. Helmhart, K. Shirahige and K. Nasmyth (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell*, **127**, 523–537.
96. C. Ban and W. Yang (1998) Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, **95**, 541–552.

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

98. S.H. Jun, T.G. Kim and C. Ban (2006) DNA mismatch repair system. Classical and fresh roles. *FEBS J.*, **273**, 1609–1619.

99. T. Yamamoto, H. Iino, K. Kim, S. Kuramitsu and K. Fukui (2011) Evidence for ATP-dependent structural rearrangement of nuclease catalytic site in DNA mismatch repair endonuclease MutL. *J. Biol. Chem.*, **286**, 42337–42348.

100. W. Yang (2000) Structure and function of mismatch repair proteins. *Mutation Res.*, **460**, 245–256.

101. A. Robertson, S.R. Pattishall and S.W. Matson (2006) The DNA binding activity of MutL is required for methyl-directed mismatch repair in *Escherichia coli*. *J. Biol. Chem.*, **281**, 8399–8408.

102. A.N. Schorzman, L. Perera, J.M. Cutalo-Patterson, L.C. Pedersen, L.G. Pedersen, T.A. Kunkel and K.B. Tomer (2011) Modeling of the DNA-binding site of yeast Pms1 by mass spectrometry. *DNA Repair*, **10**, 454–465.

103. A.J. Plys, M.V. Rogacheva, E.C. Greene and E. Alani (2012) The unstructured linker arms of Mlh1–Pms1 are important for interactions with DNA during mismatch repair. *J. Mol. Biol.*, **422**, 192–203.

104. I. Winkler, A.D. Marx, D. Lariviere, R. Heinze, M. Cristóvão, A. Reumer, U. Curth, T.K. Sixma and P. Friedhoff (2011) Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in *Escherichia coli*. *J. Biol. Chem.*, **286**, 17326–17337.

105. R. Dutta and M. Inouye (2000) GHKL, an emergent ATPase/kinase superfamily. *TIBS*, **25**, 24–28.

106. S.K. Wandinger, K. Richter and J. Buchner (2008) The Hsp90 chaperone machinery. *J. Biol. Chem.*, **283**, 18473–18477.

107. Y.Y. Polosina and C.G. Cupples (2009) MutL: conducting the cell's response to mismatched and misaligned DNA. *Bioessays*, **32**, 51–59.

108. A. Gubaev and D. Klostermeier (2014) The mechanism of negative DNA supercoiling: A cascade of DNA-induced conformational changes prepares gyrase for strand passage. *DNA Repair*, **20**, 130–141.

# 1

# 2

# Using stable MutS dimers and tetramers to quantitatively analyze mismatch recognition and sliding clamp formation

Flora S. Groothuizen[1*], Alexander Fish[1*], Maxim V. Petoukhov[2], Annet Reumer[1], Laura Manelyte[3], Herrie H.K. Winterwerp[1], Martin G. Marinus[4], Joyce H.G. Lebbink[5], Dmitri I. Svergun[2], Peter Friedhoff[3] and Titia K. Sixma[1]

[1] Division of Biochemistry and CGC.nl, Netherlands Cancer Institute, Amsterdam, the Netherlands [2] European Molecular Biology Laboratory, Hamburg Outstation, Hamburg, Germany [3] Institute for Biochemistry, Justus-Liebig-University, Giessen, Germany [4] Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, USA [5] Department of Cell Biology and Genetics, Cancer Genomics Center and Department of Radiation Oncology, Erasmus Medical Center, Rotterdam, the Netherlands * Equal contribution

*Nucleic Acids Research*, **41**, 8166–8181 (2013)

## Abstract

The process of DNA mismatch repair is initiated when MutS recognizes mismatched DNA bases and starts the repair cascade. The *Escherichia coli* MutS protein exists in an equilibrium between dimers and tetramers which has compromised biophysical analysis. To uncouple these states we have generated stable dimers and tetramers respectively. These proteins allowed kinetic analysis of DNA recognition and structural analysis of the full-length protein by X-ray crystallography and small angle X-ray scattering. Our structural data reveal that the tetramerization domains are flexible with respect to the body of the protein, resulting in mostly extended structures. Tetrameric MutS has a slow dissociation from DNA, which can be due to occasional bending over and binding DNA in its two binding sites. In contrast, the dimer dissociation is faster, primarily dependent on a combination of the type of mismatch and the flanking sequence. In the presence of ATP, we could distinguish two kinetic groups: DNA sequences where MutS forms sliding clamps and those where sliding clamps are not formed efficiently. Interestingly, this inability to undergo a conformational change rather than mismatch affinity is correlated with mismatch repair.



## Introduction

The DNA mismatch repair (MMR) system is important to maintain genomic stability. Mismatch repair protein MutS recognizes misincorporated nucleotides and this starts a cascade of events involving the recruitment of several proteins. This eventually results in the removal of the mismatch and resynthesis of the new DNA strand (1). Loss of MutS leads to a mutator phenotype in bacteria, and in humans mutations in MutS homologs can result in hereditary non-polyposis colon cancer, also known as Lynch syndrome (2,3). To understand the molecular impact of such mutations, the bacterial mismatch repair system is a widely used model.

In bacterial mismatch repair, MutS binds DNA in a homodimeric form. Within this dimer there is an asymmetric association of the subunits when bound to heteroduplex DNA, as has been shown in crystal structures (4-6). After mismatch recognition, MutS exchanges adenosine diphosphate (ADP) for adenosine triphosphate (ATP) and undergoes a conformational change that allows sliding on DNA (7).

Each subunit of MutS consists of 853 amino acids in *Escherichia coli*. The currently known crystal structures of MutS bound to mismatches, however, lack a 53 amino acids long C-terminal domain. The structure of this domain has partly been elucidated fused to maltose binding protein (MBP) (8). It has been found to be important for mismatch repair (9), although this dependence could be due to decreased stability of the dimer upon

truncation of the C-terminal domain (8).

By interactions via these C-terminal domains, *E. coli* MutS dimers can form tetramers ($K_d$ of 50 nM) (10), which has also been shown for other bacterial MutS homologs (11-13). Because of the unstable nature of the tetramerization, however, full-length MutS exists in equilibrium between dimers and tetramers in solution under physiological conditions in cells (~200 nM MutS monomers) (14). The biological function of the tetrameric form of MutS is still under discussion (8,15), but tetramerization is not essential for mismatch repair (8,16). It may, however, be important for other functions of MutS, such as its role in anti-recombination (16). The MutS tetramer is thus still a topic of interest (8,10,11,13,15-18).

The *E. coli* mismatch repair system is better understood and easier to reconstitute *in vitro* than eukaryotic mismatch repair. However, for the main heterodimeric human mismatch-recognition proteins (MutSα and MutSβ) tetramerization has not been reported, even though a similar double helix-loop-helix fold was observed for the C-termini of MutSβ as in the crystal structure of the dimerized C-termini of *E. coli* MutS (19).

The added variable of the dimer-tetramer equilibrium of wild-type *E. coli* MutS also complicates *in vitro* kinetic analysis of DNA binding. The DNA-binding kinetics of a single mismatch recognition unit of MutS have therefore never been studied quantitatively, instead it has only been done for the wild-type protein that still forms tetramers, where cooperative binding had to be taken into account (20). Uncoupling the dimer and the tetramer of MutS can help to make this system less complex.

Here, we achieve uncoupling of dimer and tetramer by site-specific point mutations in the C-terminal tetramerization domain and chemical crosslinking techniques respectively. Using point mutations to prevent tetramerization rather than truncating the C-terminal domain, the stability of the dimer is not compromised (8,17). This enabled us to study both dimer and tetramer independently.

We used the stabilized full-length dimer and tetramer for structural studies to understand how the domains are organized within the proteins. Moreover, with our methods the kinetics of DNA-binding by the dimer and tetramer could be compared. This will be important for future studies to explain their relevance in DNA recognition. Elimination of the dimer-tetramer equilibrium enabled us to fit mismatch-recognition kinetics of single DNA-binding units of MutS. We used this possibility to quantitatively investigate recognition of different mismatches and to investigate sliding clamp formation by MutS.

## Materials and methods



### Wild-type and mutant MutS proteins

Full-length dimer mutants P839E or D835R were created in the *mutS* gene in vector pET-3d (4). To obtain single-cysteine MutS R848C protein, the six native cysteines in the *mutS* gene in vector pET-3d were mutated (C93A, C235S, C239A, C297S, C569S, C711V) followed by introduction of the R848C mutation. For single-cysteine $His_6$-MutS N162C, the N162C mutation was introduced into cysteine-free $His_6$-MutS in vector pET15b, which has been described previously (17,21). All mutations were introduced using the QuikChange Site-Directed or Multi Site-Directed Mutagenesis Kits (Stratagene) and appropriate primer sequences (obtained from Invitrogen), following the manufacturers protocol.

Wild-type and mutant MutS proteins were expressed and purified as described (4,6,17), except that in the final gel filtration buffer KCl was used instead of NaCl.

### Crosslinking of single-cysteine MutS 848C

Reducing agent was removed from purified single-cysteine MutS R848C by loading it on a 5 mL HiTrap desalting column (GE) that was pre-equilibrated with buffer containing no reducing agent [25 mM Hepes (pH 7.5), 300 mM KCl, 5 mM $MgCl_2$, 10% glycerol], and eluted fractions containing the protein peak were collected. ATP was subsequently added to the protein to a final concentration of 1 mM, after which 11-bis-maleimidotriethyleneglycol ($BM(PEG)_3$, Pierce) dissolved in dimethyl sulfoxide (DMSO) was added to a final concentration of 250 mM. The protein was then incubated on ice for one hour, after which excess crosslinker was quenched by adding dithiothreitol (DTT) to a final concentration of 5 mM. The tetrameric protein was purified by size-exclusion chromatography on a Superdex 200 column in buffer A [25 mM Hepes (pH 7.5), 150 mM KCl, 5% glycerol]. Peak fractions were analyzed on SDS-PAGE, concentrated, and flash-frozen until further use. See Supplementary Figure S1 for size-exclusion chromatography profile and SDS-PAGE analysis of crosslinked product.

To isolate DNA-bound complexes for small-angle X-ray scattering (SAXS) studies, crosslinked protein was incubated with excess 21-bp or 60-bp DNA (obtained from Invitrogen; 21-bp: AGCTGCCAGGCACCAGTGTCA annealed with TGACACTGGTGCTTGGCAGCT, 60-bp: TGAAGCTTAGCTTAGGATCATCGAGGA TCGAGCTCGGTGCAATTCAGCGGTACCCAATTC annealed with GAATTGGGTA CCGCTGAATTGCACCGAGCTTGATCCTCGATGATCCTAAGCTAAGCTTCA) for 25 minutes on ice, and purified by size-exclusion chromatography on a Superdex 200 column in buffer A. Peak fractions containing the protein-DNA complexes were concentrated to multiple concentrations and flash-frozen. Protein:DNA ratios in the

complex were calculated using the ratio of absorption at 260 and 280 nm and absorption properties at both wavelengths of the individual components.

**Size-exclusion chromatography and multi-angle laser light scattering analysis**

For each protein sample, 2 mg was injected onto a Superdex 200 10/30 column in buffer [25 mM Hepes (pH 7.5), 150 mM KCl, 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol]. Elution profiles were monitored at 280 nm. Protein was subjected in-line to multi-angle laser light scattering (MALLS) measurements in a Mini-Dawn light scattering detector (Wyatt Technology) upon elution from the column. Data were analyzed using the Astra software (Wyatt Technology).

**Crystallization, data processing and refinement**

For crystallization, 100 mM MutS D835R (monomer concentration) was combined with 50 mM 21-bp DNA containing a mismatch (same 21-bp sequence as in the SAXS DNA binding studies), and 100 mm ADP. Crystals were grown using hanging drop vapor diffusion from a well solution of 25 mM Tris pH 8, 750 mM NaCl, 12% PEG 6000, 10 mM $MgCl_2$. Microseeding was used to improve crystal quality. Before data collection, crystals were transferred to a cryobuffer consisting of the mother liquor supplemented with 30% glycerol and flash cooled in liquid nitrogen.

Diffraction data was collected at beamline ID14-4 at the ESRF in Grenoble, France. Data reduction was performed using XDS (22) and Scala (23) in the CCP4 suite (24). The structure of the C-terminally truncated MutS:DNA complex (PDB entry 1E3M) (4) was used as a search model for structure solution using Phaser (25). Refinement jobs were carried out using REFMAC5 (26). PDB_REDO (27) was used to optimize refinement parameters for REFMAC5. During the refinement process, the structure of the dimer of the C-terminal 33 residues of MutS (8) was used as an initial model to fit the density for the C-terminal domain using the program Coot (28). Most of the structure could be modeled confidently, but electron density for the mismatch binding domain in subunit B and the C-terminal domains (823-853) is relatively weak and residues 658-669 and 749-757 of the ATPase domain of subunit A (the mismatch-contacting subunit), residues 1-25, 55-74 and 95-106 of the mismatch-binding domain of subunit B, residues 801-822 in both subunits, and six bases are missing. We did not observe clear density for nucleotides in either of the two subunits of the crystallized protein, and the positions of the P-loops resembled empty nucleotide binding sites. Nevertheless, some residual difference density is present in the nucleotide-binding site of the mismatch-contacting subunit, suggesting a small fraction of ADP-bound protein in the crystal. Coordinates of the refined model of MutS D835R have been deposited in the Protein Data Bank with entry code 3ZLJ. For crystallographic statistics, see Table 1. Figures were generated using PyMOL (*http://www.pymol.org*).

**Table 1.** Crystallographic data collection and refinement statistics.

| | |
|---|---|
| **Data collection** | |
| λ (Å) | 0.976 |
| Resolution range (Å) | 47.24-3.1 (3.27-3.1) |
| Completeness (%) | 99.9 (100.0) |
| I/σ(I) | 9.9 (2.4) |
| $R_{merge}$ (%) | 11.9 (61.7) |
| Space group | P $2_1$ |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 110.29, 91.15, 112.86 |
| α , β, γ (°) | 90.00, 101.79, 90.00 |
| Total no. of observations | 175019 (25663) |
| Total no. of unique reflections | 40031 (5825) |
| Multiplicity | 4.4 (4.4) |
| Wilson's B-factor (Å$^2$) | 61.9 |
| **Refinement** | |
| No. of atoms (protein+DNA) | 13234 |
| Average B-factor (Å$^2$) | 61.8 |
| $R_{free}$ reflections | 2006 |
| $R_{work}$ (%) | 22.73 |
| $R_{free}$ (%) | 26.28 |
| r.m.s.Z(bond) | 0.365 |
| r.m.s.Z(angle) | 0.462 |
| Ramachandran statistics [a] (preferred/allowed/outliers) | 1549/18/0 |

Numbers within brackets refer to the highest resolution shell.
[a] Calculated using MolProbity (29).

### SAXS measurements and analysis

SAXS measurements were performed at beamlines P12 and X33 (30) at EMBL Hamburg. Samples of MutS D835R were prepared in buffer containing 25 mM Hepes (pH 7.5), 250 mM KCl, 5% glycerol. MutS tetramer samples were prepared in buffer A as described above. The samples were thawed and centrifuged at high speed for 1 min just before measurement. Samples were exposed to X-rays in a measuring cell cooled to 10 °C. Data were analyzed using the ATSAS software package (31): data processing was performed using PRIMUS (32) where the Guinier plots were used to assess $R_g$ values and data quality at low-angles (Supplementary Figure S2D), after which GNOM (33) was used to generate distance

distribution plots. GNOM results were used as input for DAMMIF (34) to generate 10 independent *ab initio* models for both the dimer and the DNA-free tetramer, which were subsequently averaged using DAMAVER (35). For SAXS statistics, see Table 2. Figures were generated using PyMOL (*http://www.pymol.org*).

**Crosslinking of single-cysteine mutant D162C**

Single-cysteine $His_6$-MutS D162C was incubated for 10 minutes on ice at 10 μM concentration in buffer [20 mM Hepes (pH 7.5), 5 mM $MgCl_2$, 125 mM KCl and 1 mM ADP]. M4M crosslinker (1,4-butanediyl-bismethanethiosulfonate; Toronto Research Chemicals) or M17M crosslinker (3,6,9,12,15-pentaoxaheptadecane-1,17-diyl-bismethanethiosulfonate; Toronto Research Chemicals) was added in 5-fold molar excess over the protein and incubated for 20 min on ice. The extent of crosslinking was monitored by 6% SDS-PAGE after staining with colloidal Coomassie. The gel was imaged with a video documentation system (BioRad).

**Surface plasmon resonance measurements**

Surface plasmon resonance (SPR) measurements were performed in a Biacore T200 system at 25 °C. Unless otherwise indicated, DNA for the SPR measurements (obtained from Sigma) contained a 21-base pair long duplex (see Supplementary Table S1 for the full range of DNA sequences measured for MutS binding) with a single-stranded DNA (ssDNA) overhang consisting of 20 thymidines ($(dT)_{20}$). The ssDNA end was biotinylated for coupling to a Biacore streptavidin chip. The double-stranded DNA (dsDNA) end had a fluorescein moiety attached and by flowing over anti-fluorescein antibody (Invitrogen) after immobilization the DNA end was blocked. DNA was immobilized on a Biacore streptavidin chip to a maximum total signal of 7.0 RU.

For the mismatch variation experiments and the sequence context variation experiments, MutS protein was premixed with an equal volume of SPR buffer [25 mM Hepes (pH 7.5), 150 mM KCl, 5 mM $MgCl_2$, 0.05% TWEEN-20] or SPR buffer containing 2 mM ATP just before injection, and flown over the chip in SPR buffer with or without 1 mM ATP respectively. Protein flow was maintained for 120 s, after which only SPR buffer was flown over for 240 s. Between injections of different concentrations of protein, the chip was regenerated with 0.05 % SDS. Measurements were performed *in duplo*.

For the sliding clamp dissociation experiments the DNA constructs for a GT mismatch or T insertion were immobilized, but the ends were not blocked with antibody. In all these measurements, 200 nM of MutS D835R was injected for 120 s to achieve maximum binding, after which buffer with varying ATP concentration was injected for 120 s to observe dissociation.

Initial kinetic fitting (as shown in Figure 4) was performed using the Biacore

**Table 2.** SAXS data collection and statistics.

| | MutS D835R | Crosslinked MutS tetramer | MutS tetramer + 21-bp DNA | MutS tetramer + 60-bp DNA |
|---|---|---|---|---|
| **Data-collection parameters** | | | | |
| Beam line | P12 | X33 | X33 | X33 |
| Beam geometry | 0.2 x 0.12 $mm^2$ | 2 x 0.6 $mm^2$ | 2 x 0.6 $mm^2$ | 2 x 0.6 $mm^2$ |
| Wavelength (nm) | 0.124 | 0.15 | 0.15 | 0.15 |
| *s* range ($nm^{-1}$) | 0.07-4.4 | 0.06-6.0 | 0.06-6.0 | 0.06-6.0 |
| Exposure time (s) | 1 | 120 | 120 | 120 |
| Concentration range (mg/mL) | 0.4-11.7 | 0.7-3.9 | 0.5-6.6 | 0.5-3.9 |
| Temperature (K) | 283 | 283 | 283 | 283 |
| **Structural parameters** | | | | |
| $I(0)$ (relative) [from $P(r)$] | 10300 ± 200 | 260 ± 5 | 270 ± 5 | 490 ± 10 |
| $R_g$ (nm) [from $P(r)$] | 4.6 ± 0.1 | 7.9 ± 0.1 | 8.2 ± 0.1 | 11.8 ± 0.3 |
| $I(0)$ (relative) [from Guinier] | 10400 ± 100 | 270 ± 10 | 270 ± 10 | 510 ± 10 |
| $R_g$ (nm) [from Guinier] | 4.7 ± 0.1 | 7.8 ± 0.4 | 7.8 ± 0.3 | 11.5 ± 0.4 |
| $D_{max}$ (nm) | 11.5 ± 0.5 | 27 ± 1 | 27 ± 1 | 43 ± 2 |
| Porod volume estimate ($nm^3$) | 307 | 702 | 707 | 1271 |
| Dry volume calculated from sequence ($nm^3$) | 217 | 460 | | |
| **Molecular-mass determination** | | | | |
| $I(0)$ (relative) for BSA | 4300 ± 100 | 60 ± 2 | 60 ± 2 | 60 ± 2 |
| Molecular mass $M_r$ [from $I(0)$] (kDa) | 160 ± 10 | 340 ± 30 | 310 ± 30 | 580 ± 50 |
| Calculated monomeric $M_r$ from sequence (kDa) | 95 | 95 | | |

T200 Evaluation Software version 1.0. Values for $K_d^{app}$ were determined by non-linear fitting using Graphpad Prism 4 (36) with a model for single-site binding: Response = $Max_{Response} \cdot [MutS] / (K_d^{app} + [MutS])$, where MutS concentration was expressed as monomers. Values for $K_d$ and $k_{off}$ were determined using EvilFit (37,38).

For kinetic fitting for DNA binding by the tetramer, the following input for the Biacore T200 Evaluation Software was used:

$$A + B \underset{k_{d1}}{\overset{k_{a1}}{\rightleftharpoons}} AB \quad (1) \qquad AB \underset{k_{d2}}{\overset{k_{a2}}{\rightleftharpoons}} A^*B \quad (2) \qquad A^*B + B \underset{k_{d3}}{\overset{k_{a3}}{\rightleftharpoons}} A^*B_2 \quad (3)$$

where $A$ = free MutS in solution; $B$ = unbound DNA; $AB$ = MutS bound to DNA, $A^*B$ = MutS bound to DNA after a conformational change; $A^*B_2$ – MutS bound to two DNA binding sites; $k_{a1}$, $k_{a2}$, $k_{a3}$, $k_{d1}$, $k_{d2}$, $k_{d3}$ are kinetic constants. $A_{tot}$ (total amount of MutS based on concentration) and $B_{tot}$ (maximum binding capacity of the surface) were also defined. This explicit second binding site model is applied to account for the fact that bending over does not change the mass as detected by SPR.

## Results

### The dimer and tetramer of MutS can be stabilized separately

To study the dimer of MutS, the formation of tetramers has to be prevented. Although this can be achieved by truncating the C-terminal domains, this will also compromise the stability of the dimer (39). A previously reported crystal structure of the C-terminal domains of MutS showed a tetramerization interface formed by crystallographic symmetry (8). Mutations in this interface can perturb interactions between full-length dimers. It has been reported previously that a mutation at D835 to R, resulting in a charge inversion (Figure 1A), will prevent tetramerization of MutS without jeopardizing the stability of the dimer (17). This has also been described for an R840 to E mutation (8). We hypothesized that mutating P839 to E, introducing a negative charge (Figure 1A), will have a similar effect.

We verified the oligomerization states of the different mutants using MALLS analysis in line with size-exclusion chromatography (Figure 1C). In this measurement, wild-type full-length MutS shows a large peak for the dimer-tetramer equilibrium, with an apparent weight of 324 kDa (one monomer is 95 kDa) and a small peak corresponding to a monomer fraction. MutS D835R and P839E show molecular weights of ~198 and 189 kDa respectively (Figure 1C), indicating that these mutants do not tetramerize. A decline in measured molecular weight toward the right ends of the peaks indicates minor

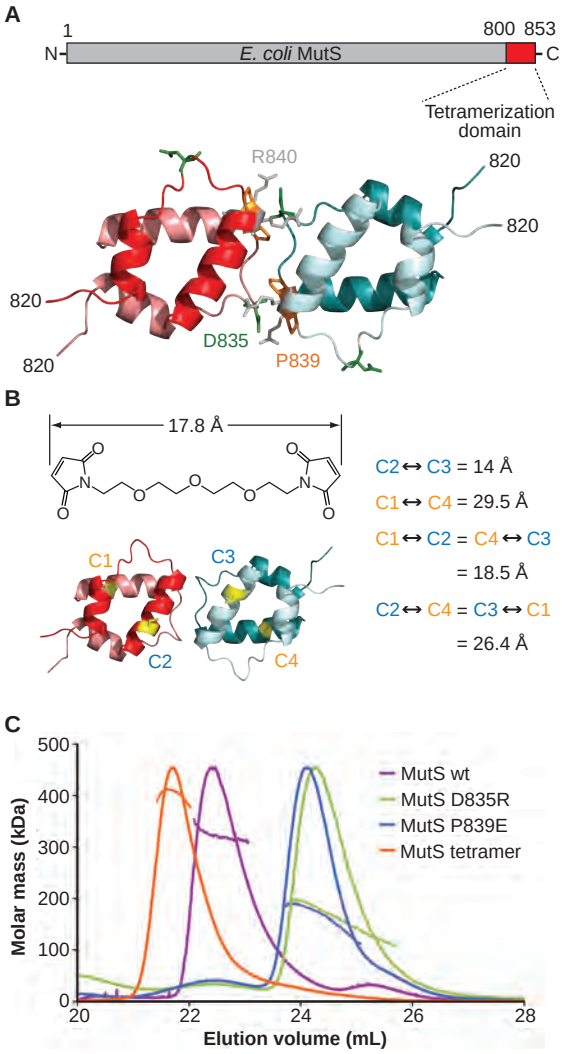



**Figure 1.** Stabilization of MutS dimer and tetramer. **A)** Tetramerization of MutS can be prevented by changing the local charge in the tetramerization interface via mutation of residues in the C-terminal domains. These residues are indicated in the crystal structure of the tetramerized C-terminal domains (from PDB entry 2OK2): D835 (green), P839 (orange) and R840 (gray). **B)** The molecular structure of the flexible molecule 1,11-bis-maleimidotriethyleneglycol (BM(PEG)$_3$) has a maximum dimension of ~18 Å. Introduced cysteines at position 848 were mapped on the crystal structure of the tetramerized C-terminal domains. The residues were mutated to cysteines in PyMOL (*http://www.pymol.org*) in preferred rotamer positions and are shown in yellow with the labels C1-4. Theoretical distances between sulphur atoms of these cysteines are indicated. **C)** Normalized UV profiles and MALLS signal from size-exclusion chromatography of the full-length MutS proteins indicate apparent molecular weights of 324 kDa for wild-type MutS (purple), 410 kDa for the crosslinked tetramer (orange), 189 kDa for MutS P839E (blue), and 198 kDa for MutS D835R (green). The molecular weight of one full-length MutS monomer is 95 kDa.

dissociation into monomers, as was also observed in atomic force microscopy (AFM) studies with wild-type MutS, indicating that this is an intrinsic property of MutS (15). Nevertheless, it is much less pronounced for these full-length MutS point mutants than for the C-terminal truncation mutant ΔC800 (39). In this way, relatively stable dimers can be made via different single point mutations predicted from the published tetramer interface. We chose to use the D835R mutant for further characterization.

To stabilize the tetrameric form of MutS, chemical crosslinking was used. The residues at position 848 in the tetramerization domains are solvent-exposed and therefore appeared to be a good option for crosslinking. For this purpose, a cysteine was introduced at this position in a cysteine-free construct of full-length MutS, resulting in single-cysteine MutS R848C (SC-MutS R848C). Cysteine-free MutS is active in mismatch repair, as has been shown previously (8,17). The SC-MutS R848C mutant is fully active in mismatch repair *in vitro* (Supplementary Table S2). Its mutation rate *in vivo* is similar to wild-type with a slight elevation, owing to fluctuation in the assay or a minor defect (Supplementary Table S2). When allowing the SC-MutS R848C to chemically react with $BM(PEG)_3$, a cross-linker that contains a flexible PEG spacer and two reactive maleimides, the C-terminal domains of two adjoining dimers can be coupled together irreversibly by forming stable thioether linkages with the cysteine residues. The position of the cysteines is such that only one possible crosslink can be made with $BM(PEG)_3$ within each tetramer (Figure 1B), resulting in 50% crosslinking (Supplementary Figure S1). Size-exclusion chromatography and MALLS analysis of this purified crosslinked product indicates a molecular weight of ~410 kDa, corresponding to four MutS subunits (Figure 1C), indicating that indeed the tetramer of MutS has been stabilized.

**The C-terminal domain is mobile within the MutS dimer**

The dimer-tetramer equilibrium of MutS complicates structural studies of the full-length protein. In previously reported structures of *E. coli* MutS, the C-terminal 53 residues had been truncated, which excluded the tetramerization and facilitated crystallization (4,6). The tetramerization domain has been crystallized outside the context of MutS, fused to MBP (8). With our current dimer mutants the dimer-tetramer equilibrium of full-length MutS has been eliminated. We were able to crystallize full-length MutS D835R bound to a mismatch and determine the structure at a resolution of 3.1 Å (Figure 2A).

The full-length MutS protein crystallized in space group $P2_1$. The first 800 residues superpose well with previous MutS structures that were crystallized in a different crystal lattice, with an rmsd of 0.9 Å for the Cα atoms of PDB entry 1E3M. The crystal structure is almost complete except for a few missing loops (see 'Materials and methods' section). Although we crystallized with a shorter DNA duplex [21-bp instead of 30-bp DNA (4,6,39-42)] we could resolve one more base (36 of 42 possible).

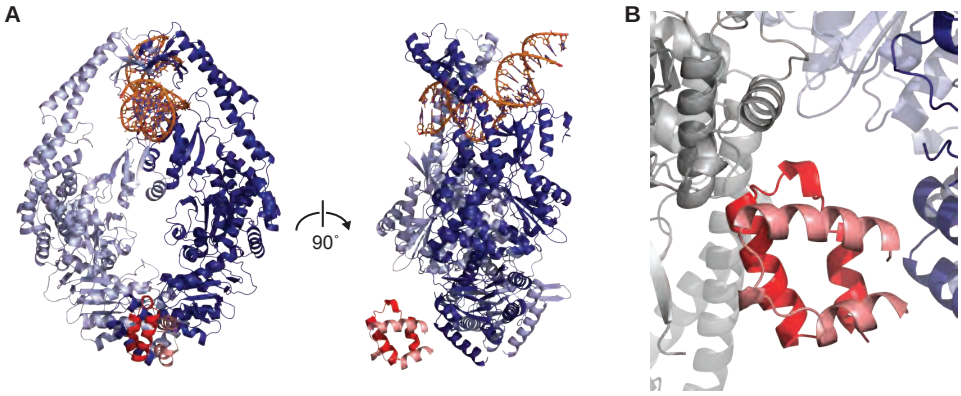



**Figure 2.** Crystal structure of full-length dimeric MutS D835R. **A)** Front and side view of a cartoon representation of the dimerized C-terminal domains (red and salmon) adjacent to the ATPase domains of the rest of the dimer (dark blue and light blue). DNA is shown in orange. **B)** The position of the dimerized C-terminal domains (red and salmon) is stabilized by crystal contacts with an adjacent MutS dimer (gray).

The two nucleotide-binding sites of the crystallized MutS dimer are mostly in the empty state, even though ADP was present in the crystallization mixture. This is the first *E. coli* MutS crystal structure of a nucleotide-free state, although a *Thermus aquaticus* MutS-DNA structure without nucleotide has been described (5). It has been shown by native mass spectrometry that mismatch binding regulates asymmetric nucleotide binding in *E. coli* MutS, but a nucleotide-free mismatch-bound state was also observed (43). The P-loop of subunit A is displaced compared with an ADP-bound MutS structure (4), resulting in similar empty nucleotide binding sites for subunits A and B.

This similarity between the two subunits makes the dimer more symmetrical than in ADP-bound structures, except that the N-terminal mismatch-binding domain (residues 2-115) of subunit B is in a different position than in subunit A since it does not contact the DNA mismatch. This domain is affected by partial crystallographic disorder in subunit B and together with the connector domain it moved slightly inward compared with the existing nucleotide-bound structures, probably due to the different crystal contacts.

Our structure includes the dimerized C-terminal domains (residues 823-853) that were truncated in previous MutS structures. Electron density for the C-terminal domains is not well defined, but we could build the two-layer helix-loop-helix fold as observed in the MBP fusion-protein structure (8), showing that this previously determined structure is present in the context of MutS. The 22 amino acids linking the C-terminal domains to the rest of the dimer could not be resolved in density, probably owing to intrinsic disorder. The chains of the two C-terminal domains can therefore not be allocated to their corresponding N-terminal subunits. It is, however, evident to which dimer in the crystal the C-terminal

domains belong, as the distance to other dimers in the crystal is larger than the ~68 Å that can be spanned by 22 stretched-out residues (>75 Å to residue 800 of the nearest neighbor).

In the crystal structure the dimerized C-terminal domains are positioned adjacent to the ATPase domains (Figure 2A). The C-termini appear to have taken the space that was occupied by residues 749-757 of subunit A as observed in other MutS structures, and these residues are not visible in our structure. Both subunits in the dimerized C-terminal domains contact helices in the connector domain of an adjacent dimer in the crystal, which probably stabilizes their position (Figure 2B).

The stable MutS D835R dimer was also used for solution X-ray scattering (SAXS) analysis. This technique gives information on the shape of the protein and can therefore give an indication of the organization of the C-terminal domains in the full-length dimer mutant in solution. The SAXS data were used for *ab initio* modeling to obtain information about the full-length dimer shape. Lack of DNA in the measured MutS protein is expected to allow flexibility of the DNA clamp domains of MutS. This was indeed reflected in the *ab initio* models by diversion between the 10 different modeling runs. However, the overall organization of the models remained similar. Averaging of the models provided an envelope that has space for the dimerized C-terminal domains in line with the DNA-binding clamp in the dimer (Figure 3A, Supplementary Figure S2A,B), thereby extending the longest axis of the dimer.

The SAXS model is different from what was observed in the crystal structure, where the domains are positioned adjacent to the dimer. Apparently, the dimerized C-terminal domains are mobile with respect to the rest of the dimer, but in solution there appears to be a most-occupied area extending the long axis of the dimer (Supplementary Figure S2B).

**The tetramer of MutS is predominantly extended in solution but can still bend over**

To get further insight into the function of the MutS tetramer, we analyzed the conformation of the tetramer in solution. Attempts to analyze tetramer conformation have been made previously for the MBP-C-terminal MutS fusion protein or for the wild-type MutS protein that is a mixture of dimeric and tetrameric states (8), but not for stable tetramers. Our crosslinked tetramer of MutS gave an opportunity to obtain more information on the shape of this oligomerization state in solution using SAXS analysis.

The SAXS data for the tetramer indicate a radius of gyration of 7.8 nm with an overall maximum dimension ($D_{max}$) of 26.5 nm, implying an extended conformation (Figure 3B). *Ab initio* modeling generated an envelope that can fit two MutS dimers (Figure 3A, Supplementary Figure S2C). As the C-termini of two dimers were covalently linked in our stable tetramer, this means that the DNA binding clamps are facing away from each other in this model. These data reinforce the observation that the C-terminal domains are not

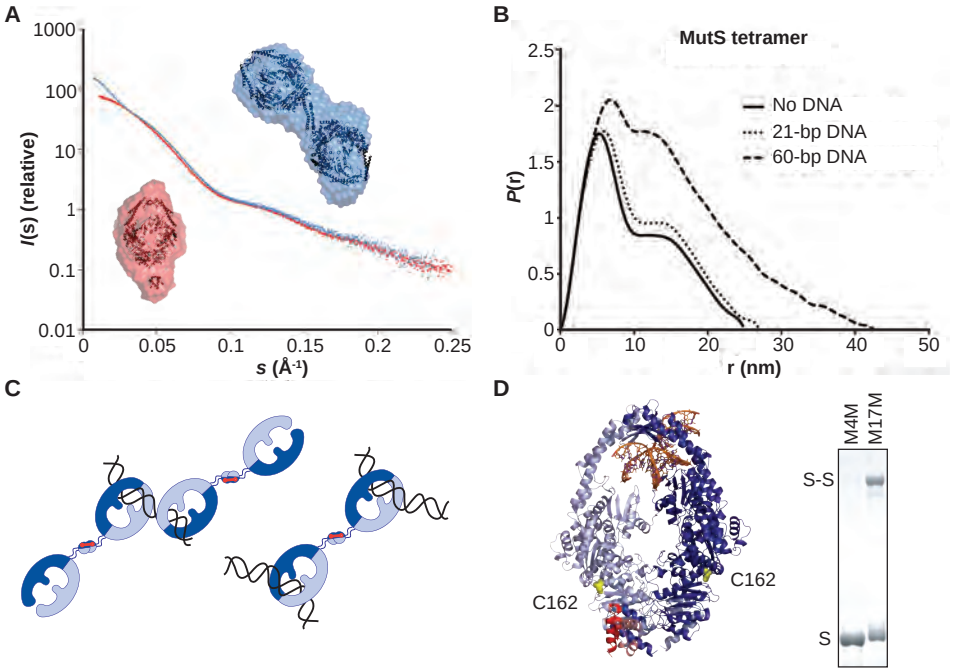



**Figure 3.** Conformations of the stable dimer and tetramer of MutS. **A)** SAXS curves and *ab initio* models of the MutS D835R dimer mutant (red) and the crosslinked tetramer (blue). The scattering curves are displaced in logarithmic scale for better visualization. For the dimer envelope, the dimerized core of MutS (first 800 residues from our structure without DNA), and the dimerized C-terminal domains (residues 823-853 from our structure) are superposed onto the envelope and represented as black cartoons. For the tetramer envelope, two dimerized MutS cores and the tetramerized C-terminal domains (derived from PDB entry 2OK2) are superposed onto the envelope. **B)** Distance distribution plot for the crosslinked tetramer of MutS: unbound (solid curve), bound to 21-bp DNA (dotted curve) and bound to 60-bp DNA (dashed curve). **C)** Two possible models of the crosslinked tetramer of MutS bound to 60-bp DNA, which would result in a larger $D_{max}$ values than the tetramer bound to 21-bp DNA. **D)** Position 162 is indicated on our full-length MutS structure as yellow spheres. Using M17M crosslinker, but not M4M crosslinker, two MutS dimers can be covalently linked through cysteines at position 162 as visualized on SDS-PAGE.

adjacent to the ATPase domain, but extend the long axis of the molecule.

As observed for the MutS D835R dimer, flexibility of the DNA clamp domains resulted in some diversion of the models between runs. This also biased the averaging of different models to result in an envelope in which one dimer appears somewhat larger than the other. This can be explained by assuming that overall, the flexible unbound DNA clamp domains of the dimers in a tetramer do not occupy the same conformational volume for both dimers at the same time, resulting in an asymmetric model.

We wondered whether the tetramer would bend over and bind a single DNA molecule with both DNA-binding dimers. To investigate this, the tetramer was bound to either

21-bp DNA containing a mismatch, too short to accommodate double binding, or 60-bp DNA containing a mismatch. We then performed SAXS measurements of the DNA-bound complexes.

When incubating the tetramer with excess 21-bp DNA and purifying the resulting 1:1.8 (tetramer:DNA) complex, the SAXS measurement resulted in a curve with more distinct features than unbound tetrameric MutS (Supplementary Figure S2G). This is probably the result of reduced flexibility in the clamp-domains when bound to DNA. The corresponding distance probability *P*(r) plot reflects this with better-defined features (Figure 3B), whereas there is no significant change in the $D_{max}$ (27 nm), indicating that the tetramer is still extended when both dimers bind a short strand of DNA.

Purification with size-exclusion chromatography after incubation with excess 60-bp DNA resulted in a 1:1.2 (tetramer:DNA) ratio, suggesting a mixture of complexes of different ratios. Although the 60-bp DNA strand would have been long enough to accommodate bending over of the tetramer to bind the DNA in two of its dimers, instead the SAXS measurement indicated that $D_{max}$ was increased to 42.5 nm (Figure 3B, Supplementary Figure S2G). This suggests binding of multiple tetramers on DNA or two DNA strands in each tetramer in our experiment (Figure 3C). In either case the data suggest that the tetramer remained predominantly extended when bound to DNA.

In the crystal structure of the full-length dimer, the position of the C-termini suggested that the tetramer could have other conformations, using the flexible linker of the C-termini. To investigate whether occasional bending of the tetramer happens at all, a crosslinking experiment was performed for full-length $His_6$-MutS with a single cysteine at position 162 (Figure 3D). When incubating this protein with the reagent M4M, which can span ~8 Å, no crosslinking occurred. In contrast, when using M17M, which can span ~22 Å, a species was captured that has the connector domains linked together. This indicates that within tetramers of MutS, the two dimers can move within 22 Å of each other. We wanted to know whether the SAXS results would allow for such bending, in the light of the extended radius. In SAXS, a molecule with large intermolecular distances such as the extended form, can dominate the scattering and there may be a mixture of conformations. We used an ensemble optimization method in which an ensemble of different random conformations is obtained and the set that fits the scattering curve is selected (Supplementary Figures 2E-F). This analysis suggests that the SAXS data allow for co-existence of a minor population of a more bent-over form next to the major extended population. We conclude that bending over of the tetramer is possible, but probably occurs relatively rarely in solution.

**The MutS tetramer dissociates slower from DNA than the MutS dimer**

We compared DNA-binding kinetics between the dimer and the tetramer state, using surface plasmon resonance (SPR) analysis. To minimize effects of homoduplex binding,

short DNA duplexes (21 bp) with a GT mismatch at position 11 were used. These were immobilized to a Biacore chip via a 20-base long ssDNA linker to allow for some spacing from the chip surface. The dsDNA ends had fluorescein moieties attached, which were bound by anti-fluorescein to obtain blocked ends and therefore exclude end dissociation. Dimeric MutS D835R does not bind to the ssDNA linker (Supplementary Figure S3A), and end-dissociation does not play a large role as binding kinetics are similar for end-blocked and unblocked DNA (Supplementary Figure S3C). Using this setup, we analyzed dimer and tetramer DNA-binding kinetics separately.

The SPR measurements showed that the D835R dimer mutant dissociates completely from the DNA (Figure 4B), whereas part of the crosslinked tetramer releases relatively slowly (Figure 4C). The binding profile for wild-type MutS appears to be a combination of both dimer and tetramer binding kinetics (Figure 4A). A superposition of the three kinetic profiles clearly shows this difference (Supplementary Figure S3D).

The very slow dissociation of the tetramer in our experiments is intriguing. When combining the crosslinked tetramer with excess 21-bp DNA for our SAXS samples, the calculated ratio for the purified complex was close to two DNA strands per tetramer. This indicates that the tetramer can be saturated with DNA in which each dimer binds a strand. This finding is in agreement with what has been found for the wild-type protein (43), although an earlier report had indicated otherwise (10). Such bivalent binding of DNA by the tetramers may explain the slow dissociation.

The slow dissociation rate of the tetramer was more pronounced when binding to 100-bp DNA or 42-bp DNA than to 21-bp DNA (Supplementary Figure S3E), but the dissociation rate did not change within the range of immobilization levels of the DNA that we used in these experiments (Supplementary Figure S3F). This suggests that in our SPR experiments, binding of the same DNA strand in the two binding sites via bending over of the tetramer or flexibility of the DNA (44) may play a role here, while at our immobilization levels 'bridging' between two immobilized DNA strands does not happen. When we use DNA that is too short for bending over (21 bp without ssDNA linker), the tetramer fully dissociates (Supplementary Figure S3G), showing that aggregation on the chip does not occur under our experimental conditions. Therefore, we conclude that the slow dissociation is due to bivalent binding of the same DNA strand by the MutS tetramer.

**Kinetic analysis of MutS binding to DNA is possible when dimer and tetramer are uncoupled**

Wild-type MutS can bind DNA as dimers or as tetramers. Reports differ on whether binding to heteroduplex DNA shifts the dimer-tetramer equilibrium to either side (10,17,43,45), but we observed both the DNA-bound dimer and the DNA-bound tetramer at MutS concentrations up to 300 nM in an electrophoretic mobility shift assay (Supplementary

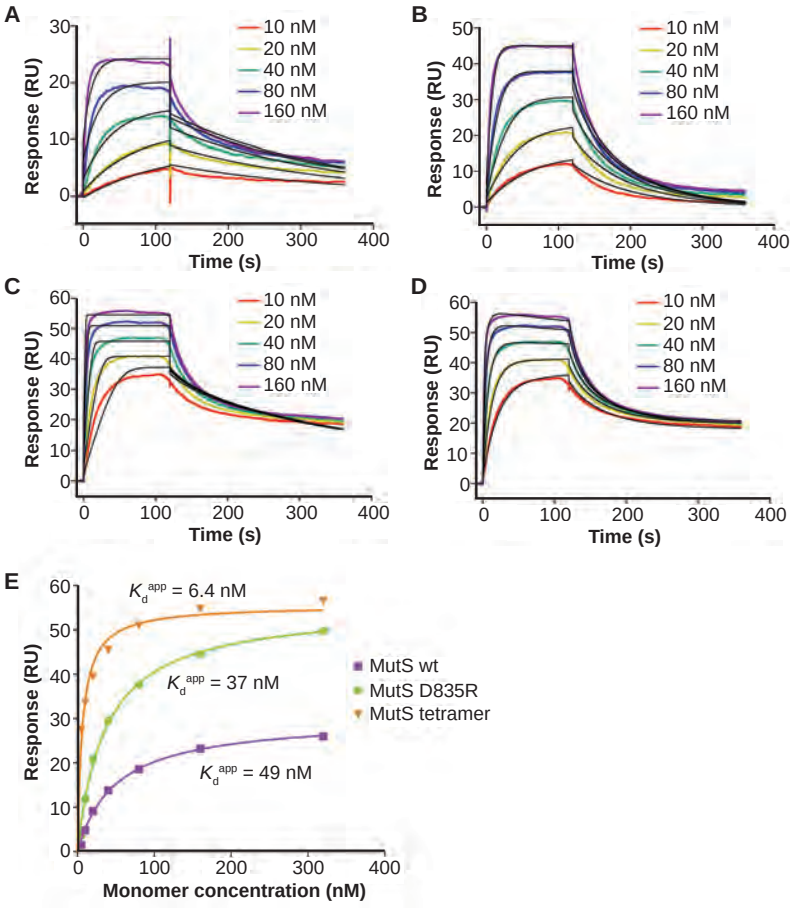



**Figure 4.** DNA binding by the MutS dimer and tetramer. SPR measurements of binding **A)** wild-type MutS, **B)** dimeric MutS D835R or **C)** the crosslinked tetramer of MutS to a GT mismatch. Different protein concentrations are represented by different colors (legend refers to monomer concentrations) and black lines indicate fitted kinetics using a one-phase binding model (fitted with Biacore T200 Evaluation Software). **D)** The kinetics of the tetramer of MutS fitted using a model that takes into account bivalent binding and a conformational change, represented in black lines (fitted with Biacore T200 Evaluation Software as specified in the 'Materials and methods' section). **E)** Determination of $K_d^{app}$ for the dimer and tetramer using SPR signal at equilibrium for different protein concentrations binding to the DNA [fitted with Graphpad Prism (36)].

Figure S3B). As the contribution of the dimer and tetramer to DNA binding can be different for every protein concentration, kinetic fitting of SPR assays is made increasingly complex, as illustrated by bad overlap of the model with the SPR data when using a one-phase binding model for fitting (Figure 4A).

Elimination of the tetramer from the equilibrium resolves the complications of wild-type protein, as a single DNA-binding unit is a dimer of MutS. Therefore, DNA binding

by dimeric MutS D835R can be fitted using one-phase binding kinetics as shown by good overlap of the fitted model with the data curves (Figure 4B).

Such a simple binding mode is clearly not the case for the tetramer of MutS (Figure 4C). However, tetramer kinetics can be fitted with a model that takes into account the presence of two DNA-binding sites and a conformational change that needs to occur in order to bring the second site toward the same strand of DNA as described in the 'Materials and methods' section (Figure 4D).

The tetramer displays stronger overall affinity for DNA than wild-type and dimeric MutS, as seen from the apparent dissociation constants ($K_d^{app}$) determined using the response at equilibrium binding for different protein concentrations as specified in the 'Materials and methods' section (Figure 4E). This is in-line with the bivalent binding by the tetramer. The $K_d^{app}$ values for wild-type and dimeric MutS, however, are comparable. Such similarity between the binding affinities for wild-type and dimeric MutS is in agreement with a previous report (17). The dissociation from DNA for wild-type MutS is slower than for the dimer (Figures 4A and 4B), probably owing to partial tetramerization which enables bivalent DNA binding. At the same time, the time needed for tetramerization also slows down the association of wild-type MutS on DNA, explaining the similar overall affinities.

**MutS binds different mismatches with different affinities, greatly influenced by flanking sequences**

To investigate differences in kinetics, we measured binding of dimeric MutS D835R to all possible single mismatched bases and up to four G or C insertion loops (Table 3, Figure 6A, Supplementary Figure S4). Mismatches at position 11 in the 21-bp dsDNA were varied while the flanking sequences remained constant (see Supplementary Table S1). As controls, measurements of binding to non-mismatched sequences were included.

In our assays, the DNA on the Biacore chip is not static, but may differ slightly in orientation over the whole chip surface. Therefore, when the kinetic parameters are defined without assuming homogeneous binding, more accurate kinetic fitting of mismatch binding by the MutS dimer can be achieved. This can be done using the program EvilFit (37,38) which determines the presence of populations with similar values for dissociation rates ($k_{off}$) and dissociation constants ($K_d$) (Figure 5A and Figure 5C). From the heat map it can be derived that binding of dimeric MutS D835R to a GT mismatch can be represented as a sum of single binding events within one population, while minor background noise can be observed as gray areas.

The possibility of kinetic analysis for dimeric MutS binding to DNA using SPR enabled us to quantitatively investigate DNA binding by a single mismatch-recognition unit of MutS for different mismatches. Values for $k_{off}$ higher than 0.4 per second, however, were too fast to be determined. Consequently, as kinetic $K_d$ determination using fitting dependent

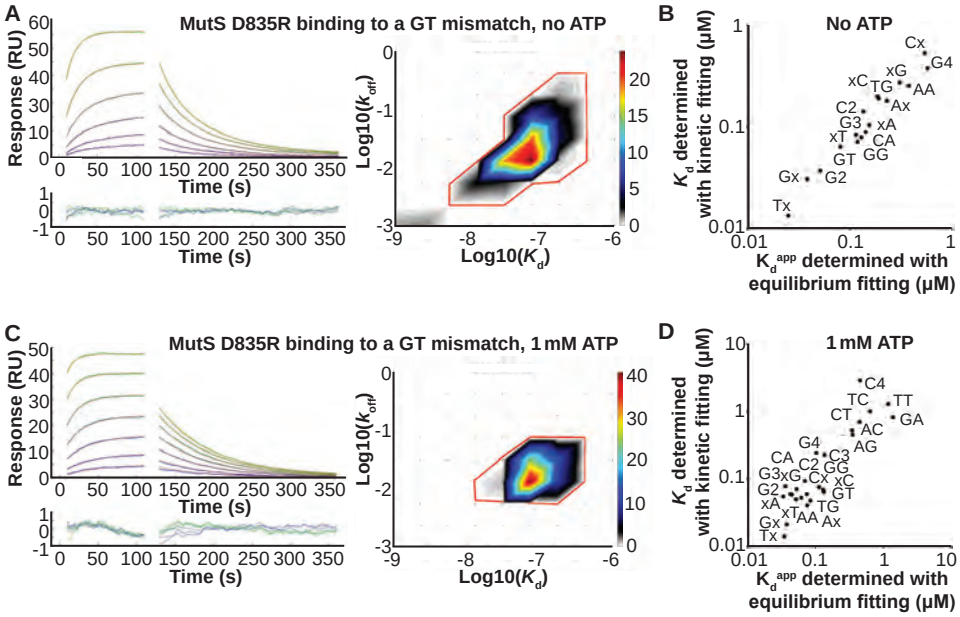



**Figure 5.** Kinetic fitting of the mismatch binding with EvilFit software (37,38). **A)** Example of fitted DNA-binding kinetics of dimeric MutS D835R in the absence of ATP. In the left graph, data are represented in a gradient of green to blue colors for multiple protein concentrations, and red lines indicate fitted model curves. Injection peaks were removed from the data points; thus they would not be fitted. Residuals of the fits are plotted below the graph. Right panel shows a heat map of the *k*off and $K_d$ distribution. A red line is drawn around the area for which a weighted averaged was taken for determination of the kinetic parameters. **B)** Correlation between the $K_d$ values for different mismatches that could be determined with EvilFit, and the $K_d^{app}$ values that were determined using equilibrium binding (fitted with Graphpad Prism (36) as specified in the 'Materials and methods' section). Values shown are averages of two binding experiments in the absence of ATP. Data points are labeled with mismatches for which binding was measured. **C)** Example of fitted DNA-binding kinetics of dimeric MutS D835R in the presence of 1 mM ATP. **D)** Correlation between the $K_d$ values for different mismatches that could be determined with EvilFit, and the $K_d^{app}$ values that were determined using equilibrium binding. Values shown are averages of two binding experiments in the presence of 1 mM ATP.

on simultaneous estimation of $k_{off}$ values, the $K_d$ could not be determined with this method for very fast release of MutS. Therefore, we estimated apparent affinities $K_d^{app}$ using the equilibrium binding (see 'Materials and methods' section), which could be calculated for every mismatch. The $K_d^{app}$ determined with this method correlates well with those $K_d$ values that could be determined using kinetic fitting with EvilFit (Figure 5B and Figure 5D). In this way, we compared kinetics for binding to all the different DNA mismatches.

The experiment showed that for different mismatches, the $K_d^{app}$ and $k_{off}$ values vary greatly: values of 25 nM – 2.3 mM are observed for $K_d^{app}$ and values of 0.0062 to higher than 0.4 $s^{-1}$ are observed for the dissociation rates. As expected, the homoduplexes had weaker

**Table 3.** Quantitative measurements of binding of dimeric MutS D835R to end-blocked 21-bp DNA of different sequences, as determined with SPR.

| Mismatch | $K_d^{app}$ - no ATP | | $K_d^{app}$ - 1 mM ATP | | $k_{off}$ - no ATP | | $k_{off}$ - 1 mM ATP | |
|---|---|---|---|---|---|---|---|---|
| | in μM | (SD) | in μM | (SD) | in $s^{-1}$ | (SD) | in $s^{-1}$ | (SD) |
| AT* | 1.44 | (0.18) | 2.52 | (0.76) | >0.4 | | >0.4 | |
| TA* | 2.17 | (0.15) | 1.90 | (0.16) | >0.4 | | >0.4 | |
| GC* | 1.85 | (0.71) | 1.80 | (0.13) | >0.4 | | >0.4 | |
| CG* | 1.90 | (0.59) | 1.58 | (0.40) | >0.4 | | >0.4 | |
| GA | 2.01 | (1.2) | 1.17 | (0.47) | >0.4 | | 0.0153 | (0.0014) |
| AG | 1.60 | (0.067) | 0.35 | (0.12) | >0.4 | | 0.0215 | (0.0018) |
| CA | 0.115 | (0.040) | 0.061 | (0.017) | 0.0353 | (0.0057) | 0.0162 | (0.0021) |
| AC | 1.31 | (0.39) | 0.344 | (0.087) | >0.4 | | 0.0188 | (0.0038) |
| GT | 0.0809 | (0.0047) | 0.131 | (0.021) | 0.0233 | (0.0033) | 0.0173 | (0.0024) |
| TG | 0.231 | (0.084) | 0.0837 | (0.051) | 0.119 | (0.038) | 0.0175 | (0.0020) |
| CT | 0.622 | (0.15) | 0.443 | (0.053) | >0.4 | | 0.0186 | (0.0028) |
| TC | 1.47 | (0.39) | 0.634 | (0.10) | >0.4 | | 0.0191 | (0.0011) |
| TT | 2.32 | (1.1) | 1.38 | (0.47) | >0.4 | | 0.0162 | (0.0046) |
| AA | 0.377 | (0.18) | 0.0511 | (0.019) | 0.180 | (0.11) | 0.0171 | (0.0024) |
| CC | 1.50 | (0.15) | 1.52 | (0.23) | >0.4 | | >0.4 | |
| GG | 0.119 | (0.026) | 0.0742 | (0.015) | 0.0298 | (0.020) | 0.0165 | (0.0018) |
| Tx | 0.0249 | (0.0037) | 0.0345 | (0.011) | 0.00620 | (0.0010) | 0.0119 | (0.0017) |
| xT | 0.155 | (0.069) | 0.0445 | (0.014) | 0.0564 | (0.014) | 0.0198 | (0.0062) |
| Ax | 0.187 | (0.11) | 0.0753 | (0.046) | 0.116 | (0.055) | 0.0168 | (0.0021) |
| xA | 0.143 | (0.070) | 0.0423 | (0.0033) | 0.0626 | (0.018) | 0.0178 | (0.0035) |
| Cx | 0.542 | (0.10) | 0.128 | (0.090) | 0.346 | (0.011) | 0.0169 | (0.000071) |
| xC | 0.192 | (0.079) | 0.112 | (0.044) | 0.0999 | (0.043) | 0.0162 | (0.0014) |
| Gx | 0.0382 | (0.0010) | 0.0375 | (0.012) | 0.0186 | (0.0038) | 0.0138 | (0.0016) |
| xG | 0.309 | (0.17) | 0.0523 | (0.019) | 0.241 | (0.032) | 0.0182 | (0.0035) |
| C2 | 0.136 | (0.012) | 0.0695 | (0.013) | 0.122 | (0.073) | 0.0174 | (0.0033) |
| C3 | 0.288 | (0.11) | 0.102 | (0.041) | >0.4 | | 0.0243 | (0.0089) |
| C4 | 1.07 | (0.29) | 0.450 | (0.069) | >0.4 | | 0.0254 | (0.0024) |
| G2 | 0.0514 | (0.010) | 0.0358 | (0.0030) | 0.0297 | (0.0045) | 0.0188 | (0.00057) |
| G3 | 0.130 | (0.068) | 0.0332 | (0.0032) | 0.118 | (0.034) | 0.0178 | (0.0028) |
| G4 | 0.578 | (0.12) | 0.135 | (0.028) | 0.244 | (0.11) | 0.0205 | (0.0074) |

*(Continued on next page)*

*(Table 3 continued from previous page)*

| Sequence context | $K_d^{app}$ - no ATP | | $K_d^{app}$ - 1 mM ATP | | $k_{off}$ - no ATP | | $k_{off}$ - 1 mM ATP | |
|---|---|---|---|---|---|---|---|---|
| | in μM | (SD) | in μM | (SD) | in $s^{-1}$ | (SD) | in $s^{-1}$ | (SD) |
| GT-1 | 0.0809 | (0.0047) | 0.131 | (0.021) | 0.0233 | (0.0033) | 0.0173 | (0.0024) |
| GT-2 | 0.231 | (0.084) | 0.0837 | (0.051) | 0.119 | (0.038) | 0.0175 | (0.0020) |
| GT-3 | 0.278 | (0.13) | 0.0616 | (0.014) | 0.0999 | (0.0030) | 0.0134 | (0.0029) |
| GT-4 | 0.0962 | (0.042) | 0.0380 | (0.011) | 0.0440 | (0.0095) | 0.0168 | (0.0021) |
| GT-5 | 0.0480 | (0.012) | 0.0387 | (0.0047) | 0.0219 | (0.0018) | 0.0111 | (0.00064) |
| GT-6 | 0.0550 | (0.0023) | 0.0327 | (0.0061) | 0.0118 | (0.0047) | 0.0100 | (0.0015) |
| GT-7 | 0.0400 | (0.0035) | 0.0464 | (0.0072) | 0.0183 | (0.022) | 0.0105 | (0.0011) |

$K_d^{app}$ and $k_{off}$ values in the absence of ATP or in the presence of 1 mM ATP were determined as described in the 'Materials and methods' section and expressed in monomers of MutS. Values represent averages of two measurements and standard deviations are shown within brackets. The asterisk indicates normal (Watson-Crick) base pairing in the same sequence context, an 'x' indicates a single insertion of the indicated base, and the numbers 2-4 indicate two to four base insertions in the sequence. Binding to a GT mismatch was measured in the context of different flanking sequences. See Supplementary Table S1 for full duplex sequences.

$K_d^{app}$ values than most of the mismatched sequences. The strongest affinity determined in the absence of ATP was for binding a single thymidine insertion, and MutS had only very weak affinity for TT, AG or CC mismatches, which were in the $K_d^{app}$ range of homoduplex binding. These mismatches have been shown to have low affinities in other experiments as well (39,46,47). Stronger affinities correlated with slower dissociation rates (Figure 6A, left graph), indicating that MutS binds tighter and longer to certain DNA mismatches than to others, which is in agreement with previous results on MutS and MutSα (20,48).

For all mismatches, both possible directions were analyzed: for example, binding to both GT and TG were measured. Even within such pairs of mismatches differences were observed for $K_d^{app}$ (81 nM for GT; 231 nM for TG) and for $k_{off}$ (0.023 $s^{-1}$ for GT; 0.12 $s^{-1}$ for TG). It appears that the context of the mismatch within the DNA duplex has an effect on binding by MutS, as our tested duplexes were not palindromic around the mismatches and MutS binds to GT mismatches in a preferred orientation (49). To verify that sequence context influences mismatch binding, a range of sequences was tested in which the mismatch was kept constant and the positions of the flanking sequences were varied. Indeed, for all seven duplex sequences tested, $K_d^{app}$ and $k_{off}$ values differed (Table 3 and Figure 6B).

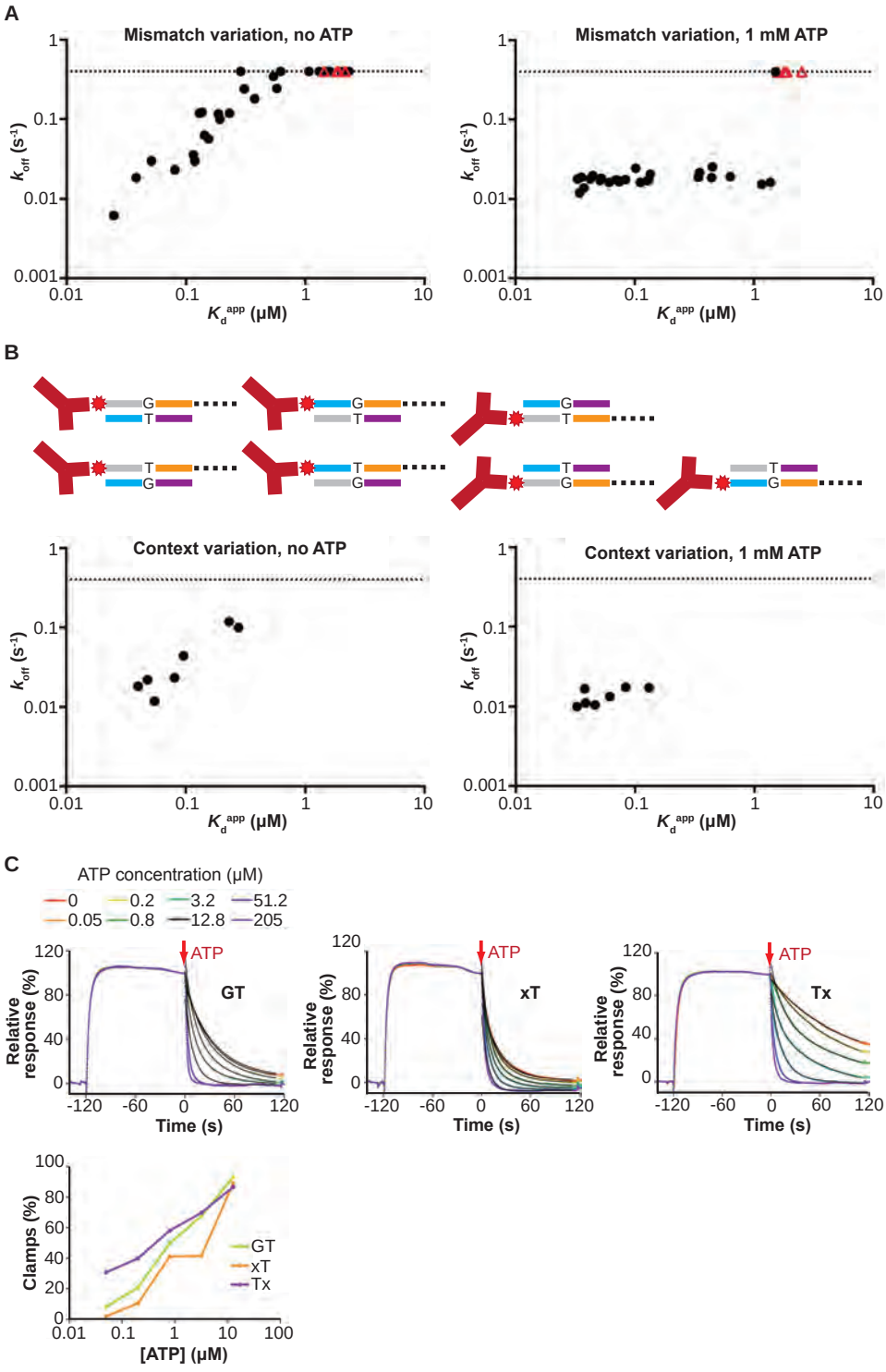
Moreover, we observed a difference in the binding preference for the mismatches when changing their direction in the DNA sequence. The order of affinities of the MutS dimer

for single-base mismatches is as follows: Tx > Gx > GT > CA ~ GG > Ax > AA > Cx > CT >> CC > GA > TT (where an x indicates an insertion of the indicated base), whereas in opposite direction of these mismatches in the same DNA duplex, the order of affinities is: GG > xA ~ xT > xC > TG > xG > AA >> AC > TC ~ CC > AG > TT. This means that sequence context is of greater influence on some mismatches than on others, explaining differences between the orders of affinities found in different experiments (20,46,50). The change in affinity when changing direction of the mispair in the DNA sequence is very pronounced for a single G insertion or for a CA mismatch: 8-fold and 11-fold respectively. Such a large difference for CA compared with AC was not observed in a previous report (47), indicating that the effect is dependent on the flanking sequences that are used.

**In the presence of ATP MutS efficiently forms sliding clamps except when binding homoduplex or a CC mismatch**

We investigated the effect of mismatch variation on DNA binding in the presence of ATP. This is representative for the *in vivo* situation, where ATP is present and MutS would then be able to undergo a conformational change to form sliding clamps on DNA, the next step in MMR (7,39,51). Similar to the mismatch-binding experiments without nucleotide, we used end-blocked DNA so that sliding clamps could not slide off the DNA ends which would make their off-rates too fast to determine. In the presence of ATP, the dissociation rates of MutS D835R were lower than without nucleotide and very similar (between 0.012 $s^{-1}$ and 0.025 $s^{-1}$) for almost all mismatches, whereas for homoduplex sequences MutS still showed a fast dissociation rate (>0.4 $s^{-1}$) (Figure 6A, right graph; Table 3). This indicates that our assay is mismatch specific, where the fast dissociation rate indicates

**Figure 6.** Mismatch binding and sliding clamp formation by dimeric MutS. **A)** Quantitative measurements of binding of dimeric MutS D835R to 21-bp DNA of different sequences (see Supplementary Table S1) as determined with SPR. Measured values for $k_{off}$ [determined with Evilfit (37,38)] are plotted against $K_d^{app}$ [determined with Graphpad Prism (36) as specified in the 'Materials and methods' section] in the absence (left graph) and the presence (right graph) of ATP for different mismatches. Values for $k_{off}$ larger than 0.4 $s^{-1}$ could not be determined and were plotted on the dotted line at 0.4 $s^{-1}$. Red triangles indicate measurements for binding homoduplex sequences. **B)** Variation of flanking DNA sequences around a GT mismatch is schematically represented by colors. To achieve blocked DNA ends, anti-fluorescein antibodies were bound to fluorescein moieties (both represented in red) that were coupled to the DNA strands. Single-stranded DNA linkers attaching the DNA duplex to the surface of the chip are indicated by dashed lines. Measured values for $k_{off}$ are plotted against $K_d^{app}$ in the absence (left graph) and the presence (right graph) of ATP for binding to a GT mismatch in the context of different flanking sequences. **C)** MutS sliding clamp formation was investigated for three mismatches by binding dimeric MutS D835R to unblocked DNA, and subsequently releasing the MutS with buffer with increasing ATP concentrations at the time point indicated by the red arrow (top three graphs). Using fixed kinetics for the slow release observed in the absence of ATP, the contribution of a faster release corresponding to the percentage of sliding clamps formed could be estimated for each ATP concentration (bottom graph). ▶


A
Mismatch variation, no ATP
Mismatch variation, 1 mM ATP
$k_{off}$ ($s^{-1}$)
$K_d^{app}$ (μM)
B
Context variation, no ATP
Context variation, 1 mM ATP
C
ATP concentration (μM)
0
0.05
0.2
0.8
3.2
12.8
51.2
205
ATP
GT
xT
Tx
Relative response (%)
Time (s)
Clamps (%)
[ATP] (μM)

direct dissociation and the slow dissociation rates indicate sliding clamp formation and then dissociation, thus reflecting the stable nature of the sliding clamps (7,51). In our experiment, these slow dissociation rates of sliding clamps would dominate over any faster release by direct dissociation. Strikingly however, although we observed that most mismatches induce formation of the stable MutS sliding clamp, the CC mismatch released MutS with a dissociation rate similar to a homoduplex. This indicates that a sliding clamp is hardly formed when MutS binds this mismatch.

Affinities of MutS for the different mismatches that all form a sliding clamp still varied over 40 fold in the presence of ATP (ranging 33 nM - 1.4 mM, Table 3). This indicates that the efficiency of the binding and changes toward the clamp state still differs, even though the sliding clamps formed are all equally stable. The effect of ATP on $K_d^{app}$ values is also greater for some mismatches than for others, resulting in a different order of affinities than in the absence of ATP. Still, T or G insertions can be very good substrates, and TT and GA mismatches are bad substrates with affinities in the same range as homoduplex binding.

To investigate the efficiency of sliding clamp formation after binding mismatches, we performed an SPR experiment in which MutS was bound to mismatch-DNA without blocked ends in the absence of ATP, and was subsequently released by flowing over buffer containing different concentrations of ATP (Figure 6C). In the case of low concentrations of ATP, two components play a role: a slow dissociation rate for ATP-independent release (as in Figure 7A, yellow arrow), and a faster dissociation due to sliding clamps releasing from the free DNA ends (Figure 7A, green arrow). The fast dissociation was more pronounced with increasing ATP concentrations, as can be seen in the graphs. Using fixed kinetic parameters for the slow ATP-independent release, the contribution of the fast dissociation could be estimated for each ATP concentration, which would be representative of the efficiency of sliding clamp formation (Figure 6C, bottom graph). This revealed that for the mismatches tested, efficiency of MutS sliding clamp formation varied at low ATP concentrations. Kinetics at 51.2 mM and 205 mM ATP were too fast to be fitted, but already at ATP concentrations higher than 12.8 mM, >86% of the MutS dimers are released as sliding clamps from the three mismatches tested. Thus, except for CC, at physiological ATP concentrations the dissociation from DNA is dominated by sliding clamp formation rather than the mismatch and its environment.

## Discussion

In this research, mutagenesis and chemical tools were used to stabilize the dimeric and tetrameric states of *E. coli* MutS. The full-length MutS mutants D835R and P839E abolished tetramerization but formed more stable dimers than the C-terminally truncated

MutS (8) and are thus more representative for DNA mismatch repair. The stabilization of the dimer and tetramer greatly reduced the complexity caused by a dimer-tetramer equilibrium, which allowed for structural analysis of both states independently. Our new insights in tetramer conformations provided understanding of its DNA-binding kinetics. Using the stable MutS dimer, we analyzed kinetics of binding to different mismatches and ATP-dependent sliding clamp formation as summarized in Figure 7.

We used solution scattering and crystallography to obtain structural information on the full-length dimeric MutS protein. Combined, these techniques are powerful tools to provide complete models for shape and conformation of proteins, as has been described in several reviews (52,53). The SAXS data indicated that in solution, the C-termini of

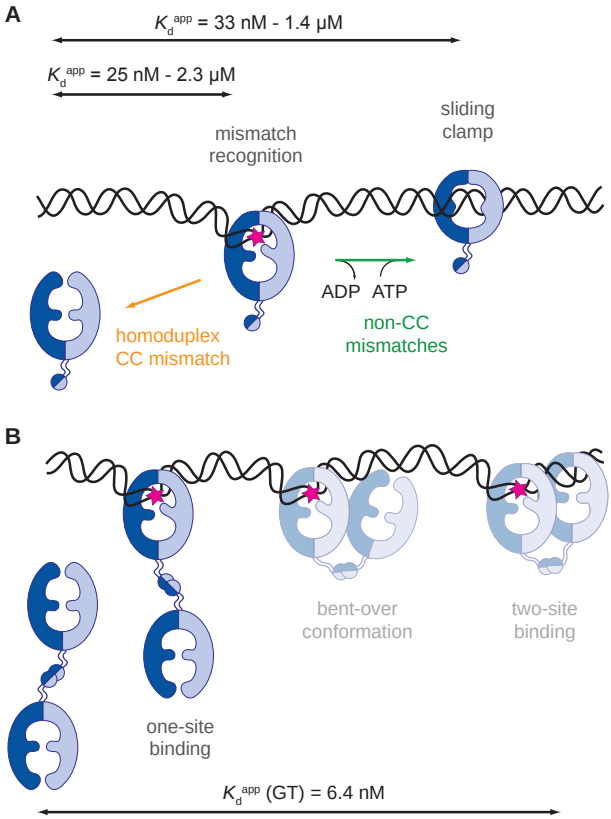



**Figure 7.** States of MutS. **A)** MutS recognizes a mismatch (pink star) as a dimer. Except when binding homoduplex or a CC mismatch, MutS then exchanges its ADP for ATP and undergoes a conformational change to form a sliding clamp on DNA. Apparent dissociation constants as determined with SPR are indicated. **B)** MutS tetramers can bind DNA in each of the two dimers by occasionally bending over. The apparent dissociation constant as determined with SPR of the MutS tetramer for DNA with a GT mismatch is indicated.

full-length MutS occupy space in line with the DNA-clamp domains, thus extending the dimer. The crystal structure of full-length MutS showed an alternate position of the C-terminal domains stabilized by crystal contacts, indicating that there is mobility of the C-terminal domains with respect to the rest of the dimer. Nonetheless, the use of stable full-length dimer mutants opens possibilities to address structural questions that cannot be answered when working with the C-terminally truncated MutS dimer. One such question regards the interaction with the β-clamp which has been reported to involve residues within the C-terminus (54) and has recently been shown to be important for the efficiency of MMR *in vivo* (55).

SAXS analysis of the stable crosslinked tetramer indicated a predominantly extended shape, with the two dimers facing away from each other. This confirms SAXS assays by Mendillo *et al.* (8), who used the C-terminal domain alone, as fusion proteins or wild-type MutS where the tetramers were mixed with dimeric MutS. We observed that when bound to 21-bp DNA, more features could be observed in the SAXS curve, probably caused by stabilization of the DNA-binding clamp domains. The resulting distance distribution plot of our measurement with the tetramer-21-bp DNA complex is in good agreement with the predicted distance distribution plot for an extended arrangement based on the crystal structure of MutS bound to DNA (8).

Although no single model adequately describes the tetramer bound to 60-bp DNA, the SAXS data indicate an increase in the maximum dimension to almost twice the length of a single tetramer. This suggests that the tetramer remains mostly extended when bound to DNA, as bending over to bind the DNA in both sites would instead have reduced the maximum dimension of the complex. This is in agreement with AFM studies, in which tetramers of wild-type MutS showed one of its dimers contacting the DNA while overall shape appeared to be extended (45). Nevertheless, detailed ensemble optimization method analysis of our SAXS data suggests the presence of smaller population of free tetrameric MutS in a bent-over conformation. As crosslinks could be made between residues 162 within tetramers, it is likely that the MutS tetramer occasionally bends over.

In our SPR experiments, bending over of the two dimers within the MutS tetramer and the subsequent possibility for bivalent binding of the DNA can explain its slow dissociation from DNA (as illustrated in Figure 7B). Although such bending over is only a rare event in solution, as shown in the SAXS analysis, in the SPR experiments we enrich for the bent-over tetramers as this two-site DNA-bound species binds stronger. The extra parameters in the kinetic model for tetramer binding made the fit possible, but we are hesitant to interpret the exact meaning of the rate constants of the steps involved, as they can be further modulated by other factors that may play a role, such as bending of DNA (44). Nevertheless, our more detailed analysis of the tetramer shape begins to explain the unexpected kinetics of this form on DNA.

Such complex kinetics was not observed for a single DNA-binding unit of MutS, the dimer. Therefore, the dimeric MutS mutants are of great value to investigate factors that influence efficiency in recognition or sliding clamp formation. As has been reported before (20), we find that MutS recognizes specific mismatches better than others with differences in $K_d^{app}$. The fact that affinities diverge is interesting, as the binding mode of MutS is similar for all mismatches (6). It could mean that some mismatches adapt to the protein-bound state more efficiently than others. The correlation between slow dissociation rates and strong affinities suggests that differences in dissociation rates contribute to the differential mismatch affinities, which is in agreement with what has been found by Huang and Crothers (20) and what has been reported for human MutS homologs (48).

In our experiments, the flanking sequences of the mismatch influenced recognition by the MutS dimer, as had been indicated previously for wild-type MutS (47,50) and has also been described for MutSα (48). Since we saw that the influence of flanking sequences is larger on some mismatches than for others, as is the influence of ATP, comparisons between affinities for mismatches should always be made with care.

In the presence of ATP, mismatch-bound MutS can undergo a conformational change to form a sliding clamp on DNA (7,51). The dissociation rates of the sliding clamps from end-blocked DNA is ten times faster in our experiments than what has been reported previously (7,42,56). This is probably due to partial dissociation of the sliding clamps at the ssDNA linkers in our SPR setup (7), but such dissociation would be consistent for all mismatches. We saw that at low ATP concentrations, the efficiency of clamp formation can differ, but at concentrations above 12.8 mM ATP concentration, almost all of the MutS protein released as sliding clamps from the mismatches that we tested in this manner. This means that at cellular ATP concentrations [0.5-3 mM in *E. coli* (57)], sliding-clamp formation is not a limiting factor to initiate subsequent steps.

As for the CC mismatch ATP-independent release was already faster than 0.4 $s^{-1}$, the same ATP-dependent dissociation experiment could not be performed. However, in the experiments with excess ATP and end-blocked DNA, only for this mismatch no stable sliding clamp formation was observed (as illustrated in Figure 7A). This may explain the inefficient repair of the CC mismatch (46). This lack of sliding clamp formation can not be explained solely by the low affinity of MutS for this mismatch, as a similar low affinity was measured in the presence of ATP for binding a TT mismatch for which clamp formation was not impaired. It has, however, been reported that MutS still kinks DNA when binding a CC mismatch (20), indicating that MutS does bind CC differently than homoduplex. MutS in solution shows fast ATP hydrolysis but slow nucleotide exchange (40); therefore, in our experiment, MutS does not contain ATP in initial binding. After mismatch recognition, MutS readily binds ATP followed by the conformational change to a sliding clamp. As the recognition steps are not abnormal for the CC mismatch, it may be that either subsequent

ATP binding or the conformational change is not efficient.

Concluding, we have investigated both the full-length dimer of MutS and the tetramer of MutS separately. The bacterial mismatch repair system is easy to modify and isolate for *in vitro* assays, allowing for ways to uncouple functions of MutS. Using the stable dimer, we find that it is not mismatch binding, but the ability to undergo the correct conformational change upon ATP binding that is correlated to mismatch repair efficiency. Again this indicates the importance of proper mismatch verification and signaling (41,58) to achieve specificity during this complicated DNA repair process.

## Accession numbers

The coordinates of the crystal structure of MutS D835R have been deposited in the PDB with entry code 3ZLJ.

## Acknowledgements

The authors thank Prakash Rucktooa and Robbie Joosten for assistance with crystallographic data processing and structure refinement, members of the Sixma and Perrakis laboratories for discussions on the project and Anastassis Perrakis for critical reading of the manuscript. They thank the beamline scientists from ID14-4 at ESRF for technical support during data collection.

## Funding

European Community's Seventh Framework Programme MM2M [HEALTH-F4-2008-223545] and BioStruct-X [FP7-INFRASTRUCTURES-2011-1-283570]; The Netherlands Organisation for Scientific Research - Chemical Sciences [711.011.011 and 700.58.428]; and funding of the Centre for Biomedical Genetics. Funding for open access charge: Institute funding.

*Conflict of interest statement.* None declared.

## References


1. T.A. Kunkel and D.A. Erie (2005) DNA mismatch repair. *Annu. Rev. Biochem.*, **74**, 681–710.
2. H.T. Lynch and A. de la Chapelle (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.*, **36**, 801–818.
3. J. Jiricny (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell. Biol.*, **7**, 335–346.
4. M.H. Lamers, A. Perrakis, J.H. Enzlin, H.H.K. Winterwerp, N. de Wind and T.K. Sixma (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature*, **407**, 711–717.
5. G. Obmolova, C. Ban, P. Hsieh and W. Yang (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703–710.
6. G. Natrajan, M.H. Lamers, J.H. Enzlin, H.H.K. Winterwerp, A. Perrakis and T.K. Sixma (2003) Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.*, **31**, 4814–4821.
7. C. Jeong, W. Cho, K. Song, C. Cook, T. Yoon, C. Ban, R. Fishel and J. Lee (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat Struct Mol Biol*, **18**, 379–385.
8. M.L. Mendillo, C.D. Putnam and R.D. Kolodner (2007) *Escherichia coli* MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair *in vivo*. *J. Biol. Chem.*, **282**, 16345–16354.
9. M.A. Calmann, A. Nowosielska and M.G. Marinus (2005) The MutS C terminus is essential for mismatch repair activity in vivo. *J. Bacteriol.*, **187**, 6577–6579.
10. K.P. Bjornson, L.J. Blackwell, H. Sage, C. Baitinger, D. Allen and P. Modrich (2003) Assembly and molecular activities of the MutS tetramer. *J. Biol. Chem.*, **278**, 34667–34673.
11. I. Biswas, C. Ban, K.G. Fleming, J. Qin, J.W. Lary, D.A. Yphantis, W. Yang and P. Hsieh (1999) Oligomerization of a MutS mismatch repair protein from *Thermus aquaticus*. *J. Biol. Chem.*, **274**, 23673–23678.
12. S. Takamatsu, R. Kato and S. Kuramitsu (1996) Mismatch DNA recognition protein from an extremely thermophilic bacterium, *Thermus thermophilus* HB8. *Nucleic Acids Res.*, **24**, 640–647.
13. V. Miguel, M.R. Monti and C.E. Argaraña (2008) The role of MutS oligomers on *Pseudomonas activityeruginosa* Mismatch Repair System activity. *DNA Repair*, **7**, 1799–1808.
14. G. Feng, H. Tiffany Tsui and M.E. Winkler (1996) Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. *J. Bacteriol.*, **178**, 2388–2396.
15. Y. Jiang and P.E. Marszalek (2011) Atomic force microscopy captures MutS tetramers initiating DNA mismatch repair. *EMBO J.*, **30**, 2881–2893.
16. M.A. Calmann, A. Nowosielska and M.G. Marinus (2005) Separation of mutation avoidance and antirecombination functions in an *Escherichia coli mutS* mutant. *Nucleic Acids Res.*, **33**, 1193–1200.
17. L. Manelyte, C. Urbanke, L. Giron-Monzon and P. Friedhoff (2006) Structural and functional analysis of the MutS C-terminal tetramerization domain. *Nucleic Acids Res.*, **34**, 5270–5279.
18. V. Miguel, R.J. Pezza and C.E. Argaraña (2007) The C-terminal region of *Escherichia coli* MutS and protein oligomerization. *Biochem. Biophys. Res. Commun.*, **360**, 412–417.
19. S. Gupta, M. Gellert and W. Yang (2011) Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat Struct Mol Biol*, **19**, 72–78.
20. S.N. Huang and D.M. Crothers (2008) The role of nucleotide cofactor binding in cooperativity and specificity of MutS recognition. *J. Mol. Biol.*, **384**, 31–47.

21. I. Winkler, A.D. Marx, D. Lariviere, R. Heinze, M. Cristóvão, A. Reumer, U. Curth, T.K. Sixma and P. Friedhoff (2011) Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in *Escherichia coli*. *J. Biol. Chem.*, **286**, 17326–17337.
22. W. Kabsch (2010) XDS. *Acta Crystallogr.*, **D66**, 125–132.
23. Scaling and assessment of data quality (2005) Scaling and assessment of data quality. **62**, 72–82.
24. M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A.J. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A.A. Vagin and K.S. Wilson (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr.*, **D67**, 235–242.
25. A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni and R.J. Read (2007) Phaser crystallographic software. *J. appl. Crystallogr.*, **40**, 658–674.
26. G.N. Murshudov, A.A. Vagin and E.J. Dodson (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.*, **D53**, 240–255.
27. R.P. Joosten, J. Salzemann, V. Bloch, H. Stockinger, A.C. Berglund, C. Blanchet, E. Bongcam-Rudloff, C. Combet, A.L. Da Costa, G. Deleage, M. Diarena, R. Fabbretti, G. Fettahi, V. Flegel, A. Gisel, V. Kasam, T. Kervinen, E. Korpelainen, K. Matilla, M. Pagni, M. Reichstadt, V. Breton, I.J. Tickle and G. Vriend (2009) PDB_REDO: automated re-refinement of X-ray structure models in the PDB. *J. appl. Crystallogr.*, **42**, 376–384.
28. P. Emsley and K.D. Cowtan (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr.*, **D60**, 2126–2132.
29. V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson and D.C. Richardson (2009) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.*, **D66**, 12–21.
30. C.E. Blanchet, A.V. Zozulya, A.G. Kikhney, D. Franke, P.V. Konarev, W.F. Shang, R. Klaering, B. Robrahn, C. Hermes, F. Cipriani, D.I. Svergun and M. Roessle (2012) Instrumental setup for high-throughput small- and wide-angle solution scattering at the X33 beamline of EMBL Hamburg. *J. appl. Crystallogr.*, **45**, 489–495.
31. M.V. Petoukhov, D. Franke, A.V. Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, C. Gorba, H.D.T. Mertens, P.V. Konarev and D.I. Svergun (2012) New developments in the *ATSAS* program package for small-angle scattering data analysis. *J. appl. Crystallogr.*, **45**, 342–350.
32. P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch and D.I. Svergun (2003) *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *J. appl. Crystallogr.*, **36**, 1277–1282.
33. D.I. Svergun (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. appl. Crystallogr.*, **25**, 495–503.
34. D. Franke and D.I. Svergun (2009) *DAMMIF*, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. appl. Crystallogr.*, **42**, 342–346.
35. V.V. Volkov and D.I. Svergun (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. appl. Crystallogr.*, **36**, 860–864.
36. H. Motulsky (1999) Analyzing data with GraphPad prism. GraphPad Software Inc.
37. J. Svitel, A. Andrea Balbo, R.A. Mariuzza, N.R. Gonzales and P. Schuck (2003) Combined affinity and rate constant distributions of ligand populations from experimental surface binding kinetics and equilibria. *Biophys. J.*, **84**, 4062–4077.
38. J. Svitel, H. Boukari, D. Van Ryk, R.C. Willson and P. Schuck (2007) Probing the functional heterogeneity of surface binding sites by analysis of experimental binding traces and the effect of mass transport limitation. *Biophys. J.*, **92**, 1742–1758.

39. M.H. Lamers, D. Georgijevic, J.H. Lebbink, H.H.K. Winterwerp, B. Agianian, N. de Wind and T.K. Sixma (2004) ATP increases the affinity between MutS ATPase domains: implications for ATP hydrolysis and conformational changes. *J. Biol. Chem.*, **279**, 43879–438885.
40. M.H. Lamers, H.H.K. Winterwerp and T.K. Sixma (2003) The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J.*, **22**, 746–756.
41. J.H.G. Lebbink, D. Georgijevic, G. Natrajan, A. Fish, H.H.K. Winterwerp, T.K. Sixma and N. de Wind (2006) Dual role of MutS glutamate 38 in DNA mismatch discrimination and in the authorization of repair. *EMBO J.*, **25**, 409–419.
42. J.H.G. Lebbink, A. Fish, A. Reumer, G. Natrajan, H.H.K. Winterwerp and T.K. Sixma (2010) Magnesium coordination controls the molecular switch function of DNA mismatch repair protein MutS. *J. Biol. Chem.*, **285**, 13131–13141.
43. M.C. Monti, S.X. Cohen, A. Fish, H.H.K. Winterwerp, A. Barendregt, P. Friedhoff, A. Perrakis, A.J.R. Heck, T.K. Sixma, R.H.H. van den Heuvel and J.H.G. Lebbink (2011) Native mass spectrometry provides direct evidence for DNA mismatch-induced regulation of asymmetric nucleotide binding in mismatch repair protein MutS. *Nucleic Acids Res.*, **39**, 8052–8064.
44. R. Vafabakhsh and T. Ha (2012) Extreme bendability of DNA less than 100 base pairs Long revealed by single-molecule cyclization. *Science*, **337**, 1097–1101.
45. H. Wang, Y. Yang, M.J. Schofield, C. Du, Y. Fridman, S.D. Lee, E.D. Larson, J.T. Drummond, E. Alani, P. Hsieh and D.A. Erie (2003) DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 14822–14827.
46. S. Su, R.S. Lahue, K.G. Au and P. Modrich (1987) Mispair specificity of methyl-directed DNA mismatch correction *in vitro*. *J. Biol. Chem.*, **263**, 6829–6835.
47. A. Joshi and B.J. Rao (2001) MutS recognition: multiple mismatches and sequence context effects. *J. Biosci.*, **26**, 595–606.
48. A. Mazurek, C.N. Johnson, M.W. Germann and R. Fishel (2009) Sequence context effect for hMSH2-hMSH6 mismatch-dependent activation. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 4177–4182.
49. M. Cristóvão, E. Sisamakis, M.M. Hingorani, A.D. Marx, C.P. Jung, P.J. Rothwell, C.A.M. Seidel and P. Friedhoff (2012) Single-molecule multiparameter fluorescence spectroscopy reveals directional MutS binding to mismatched bases in DNA. *Nucleic Acids Res.*, **40**, 5448–5464.
50. J. Brown, T. Brown and K.R. Fox (2001) Affinity of mismatch-binding protein MutS for heteroduplexes containing different mismatches. *Biochem. J.*, **354**, 627–633.
51. I. Biswas and R. Vijayvargia (2000) Heteroduplex DNA and ATP induced conformational changes of a MutS mismatch repair protein from *Thermus aquaticus*. *Biochem. J.*, **347**, 881–886.
52. M.V. Petoukhov and D.I. Svergun (2013) Applications of small-angle X-ray scattering to biomacromolecular solutions. *Int. J. Biochem. Cell Biol.*, **45**, 429–437.
53. C.E. Blanchet and D.I. Svergun (2013) Small-angle X-Ray scattering on biological macromolecules and nanocomposites in solution. *Annu. Rev. Biochem.*, **64**, 37–54.
54. F.J. López de Saro, M.G. Marinus, P. Modrich and M. O'Donnell (2006) The β sliding clamp binds to multiple sites within MutL and MutS. *J. Biol. Chem.*, **281**, 14340–14349.
55. J.S. Lenhart, A. Sharma, M.M. Hingorani and L.A. Simmons (2012) DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Mol. Microbiol.*, **87**, 553–568.
56. W. Cho, C. Jeong, D. Kim, M. Chang, K. Song, J. Hanne, C. Ban, R. Fishel and J. Lee (2012) ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure*, **20**, 1–11.
57. R.D. Lasko and D.I.C. Wang (1996) On-line monitoring of intracelIuIar ATP concentration in *Escherichia coIi* fermentations. *Biotechnol. Bioeng.*, **52**, 364–372.

58. M.S. Junop, G. Obmolova, K. Rausch, P. Hsieh and W. Yang (2001) Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Mol. Cell*, **7**, 1–12.
59. T. Loh, K.C. Murphy and M.G. Marinus (2001) Mutational Analysis of the MutH Protein from *Escherichia coli*. *J. Biol. Chem.*, **276**, 12113–12119.
60. G. Feng and M.E. Winkler (1995) Single-step purifications of His6-MutH, His6-MutL and His6-MutS repair proteins of escherichia coli K-12. *BioTechniques*, **6**, 956–965.
61. Y. Xiao, C. Jung, A.D. Marx, I. Winkler, C. Wyman, J.H.G. Lebbink, P. Friedhoff and M. Cristovao (2011) Generation of DNA nanocircles containing mismatched bases. *BioTechniques*, **51**, 259–265.



## Supplementary materials and methods

### Complementation mutator assay

Cells lacking a functional chromosomal *mutS* gene show a mutator phenotype, which is analyzed by the frequency of rifampicin-resistant clones arising from unrepaired polymerase errors in the *rpoB* gene (59). Single colonies of *mutS* deficient TX2929 cells transformed with vector control or plasmids carrying the indicated gene were grown overnight at 37°C in 3 ml LB cultures containing 100 μg/ml ampicillin. Five independent single colonies were used per plasmid. Aliquots of 50 μl of the undiluted or $10^{-6}$ diluted culture were plated on LB-agar plates containing 25 μg/mL ampicillin with or without 100 μg/mL rifampicin. Colonies were counted after o/n incubation at 37°C.

### *In vitro* DNA mismatch repair activity assay

DNA mismatch repair activity was assessed by MutH activation. For this assay, $His_6$-tagged MutS, MutL and MutH were expressed and purified as described (60). 400 nM MutS, 1 mM MutL and 200 nM MutH were used to cleave 10 nM of GT-mismatch containing DNA of 484 base pairs [generated as described (61)] in 10 mM Tris-HCl pH 7.9, 5 mM $MgCl_2$, 1 mM ATP, 50 μg/ml BSA and 125 mM KCl at 37°C for 1 to 30 min. MutH endonuclease activity was scored by the appearance of cleaved products and analyzed by 6% PAAG electrophoresis.

### Fluorescence polarization (FP) competition assay

FP measurements were done *in duplo* at room temperature, in a 96-well plate with 50 ml of sample per well in buffer [25 mM Hepes (pH 7.5), 150 mM KCl, 5 mM $MgCl_2$, 0.05% TWEEN-20]. Each sample contained 0.5 nM of TAMRA-labeled 21-bp DNA with a GT mismatch (same 21-bp sequence as in the SAXS experiments) and 32 nM of MutS

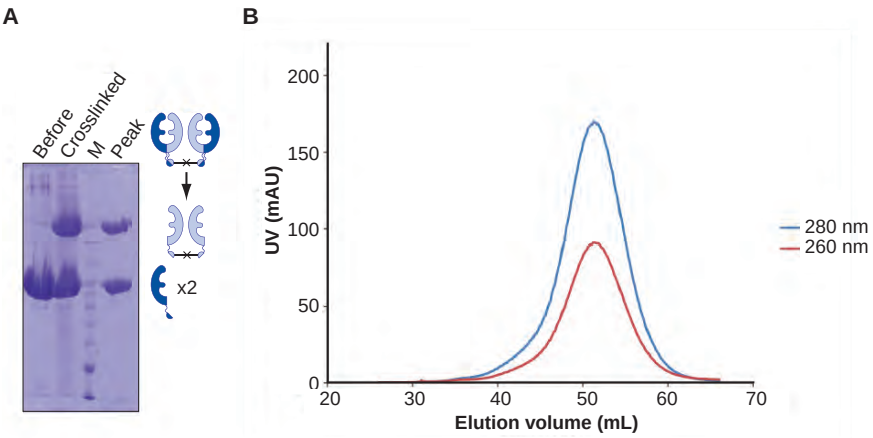
D835R. Competition was assessed by presence of increasing concentrations of unlabeled DNA with the same sequence; either 21-bp duplex or 21-bp duplex with a $(dT)_{20}$ ssDNA overhang. After 10 minutes of equilibration at room temperature, FP was measured using a PheraStar Plate Reader (BMG Inc.).

**Electrophoretic mobility shift assay**

A radioactively labeled 41-bp DNA duplex containing a mismatch was obtained as follows. Strand A (ATAGGACGCTGACACTGGTGCTTGGCAGCTTCTAATTCGAT) was incubated with [γ -$^{32}$P]-ATP and T4 PNK enzyme (New England BioLabs) in T4 PNK buffer for 40 minutes at 37 °C. The T4 PNK was subsequently inactivated by boiling the sample for 20 minutes. The labeled strand was then annealed with strand B (ATCGAATTAGAAGCTGCCAGGCACCAGTGTCAGCGTCCTAT) by combining the two strands in a 4:5 ratio in annealing buffer [25 mM Hepes (pH 7.5), 150 mM KCl, 10 mM $MgCl_2$] and incubating them together for 20 minutes at 90 °C. The sample was then slowly cooled to room temperature. Excess [γ -$^{32}$P]-ATP was removed from the labeled DNA using a G50 spin column (GE).

1 nM of $^{32}$P-labeled DNA was incubated with varying concentrations of MutS protein (wild type, D835R mutant or crosslinked tetramer) in buffer [25 mM Hepes (pH 7.5), 125 mM KCl, 50 mg/mL BSA, 0.5 mM ADP] in a total volume of 20 mL for 15 minutes. 4 mL of loading buffer (50% glycerol, 20 mM EDTA) was then added to the samples. 10 mL of each sample was loaded on a 4% acrylamide gel in TAE buffer and run for 1.5 h at 60 V at 4 °C. The gel was subsequently dried for 1 h and exposed to a phospho-imager plate for 18 h, after which the plate was scanned in a FLA-3000 phospho-imager scanner (Fujifilm).

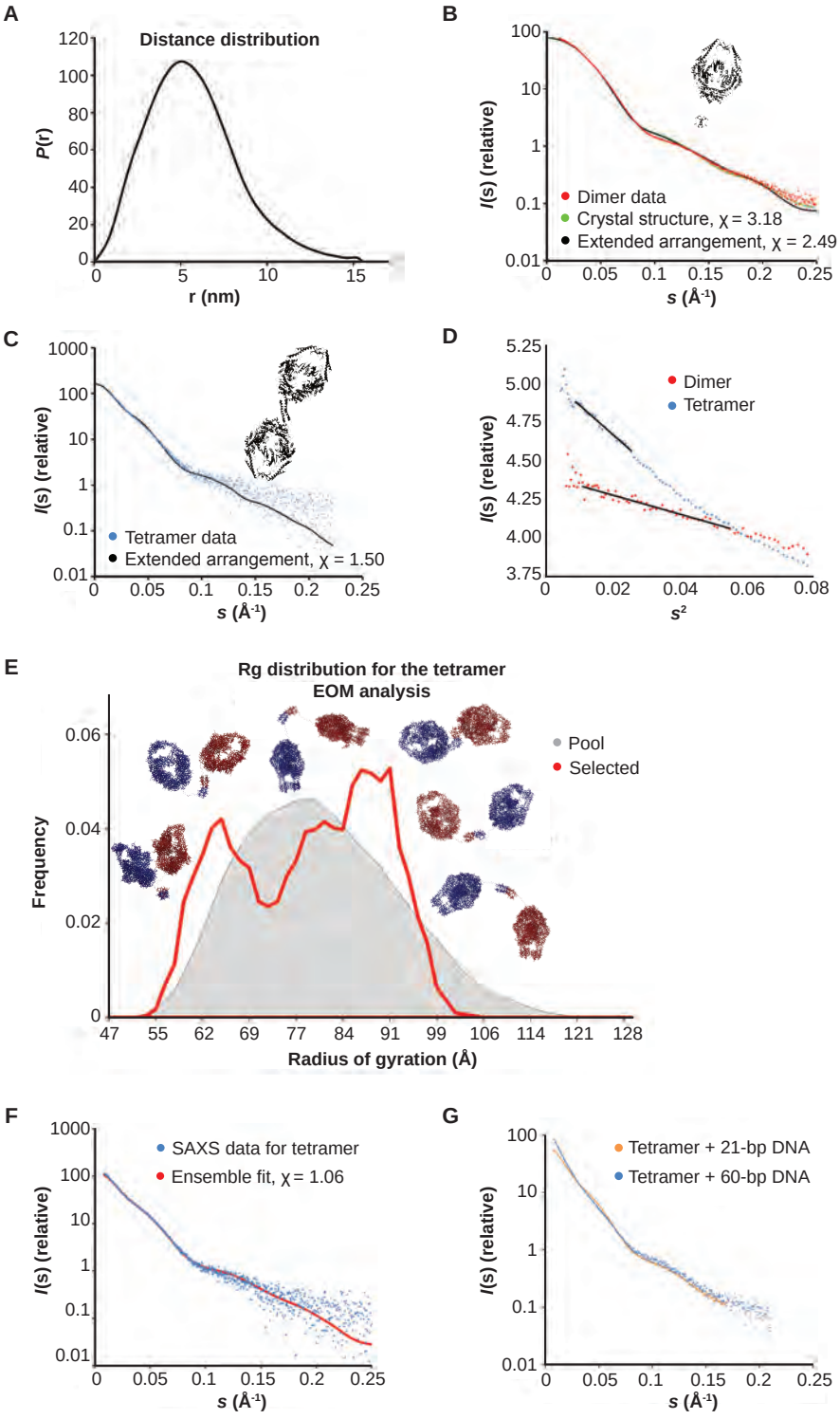
## Supplementary figures and tables

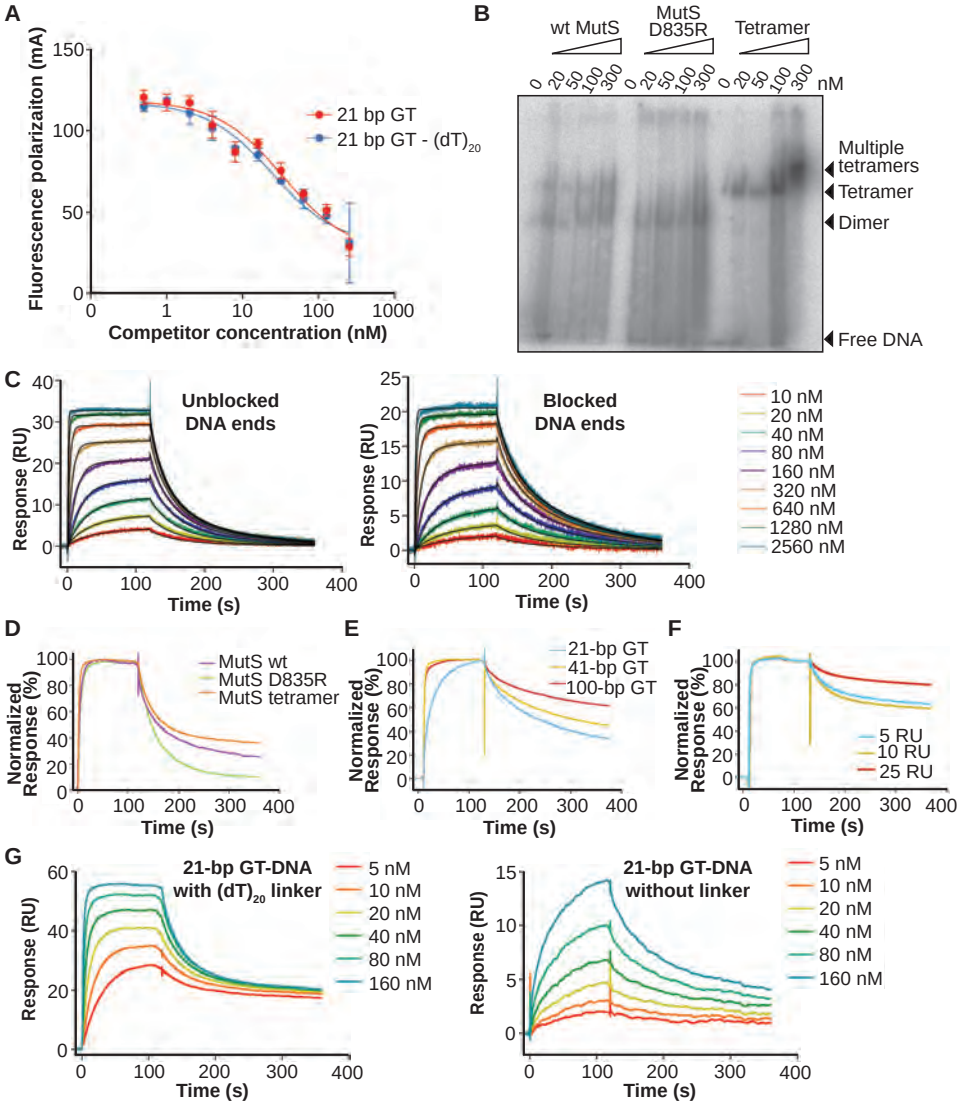



**Supplementary Figure S1.** Formation of obligate MutS tetramers by crosslinking of the single-cysteine MutS R848C with 1,11-bis-maleimido-triethyleneglycol. **A)** After crosslinking, two bands can be observed with SDS-PAGE analysis (protein is visualized with Coomassie staining), corresponding to a tetramer in which C-terminal domains of two dimers are irreversibly coupled (see schematic representation next to the gel). **B)** The stable tetramer runs as a single peak in size-exclusion chromatography (Superdex 200 16/60 column). The slight shoulder at lower elution volume, corresponding to higher-molecular weight molecules, results from a small amount of nonspecific crosslinks and was not taken with the rest of the peak for further experiments.

**Supplementary Figure S2.** SAXS analysis for the MutS dimer and tetramer. The scattering curves are displaced in logarithmic scale for better visualization. **A)** Distance distribution plot for the SAXS data for the MutS D835R dimer mutant. **B)** Predicted curves for crystal structure of MutS D835R (green) and the dimerized core of MutS and the dimerized C-terminal domains as fitted in the SAXS envelope (black) based on scattering for the MutS D835R dimer mutant (red). **C)** Predicted curve (black) for two MutS cores and the tetramerized C-terminal domains as fitted in the SAXS envelope based on scattering for the crosslinked MutS tetramer (blue). **D)** Guinier plots for the SAXS data of the dimer (red) and the tetramer (blue). **E)** For EOM analysis of the SAXS data of the crosslinked tetramer of MutS, 10000 conformations were generated based on the crystal structure of the dimerized residues 1-800 (from our MutS D835R structure) linked flexibly by 22 dummy atoms to tetramerized C-terminal domains (derived from PDB entry 2OK2). The distribution of different conformations over $R_g$ is plotted. Gray curve: pool of generated conformations. Red curve: selected conformations that together describe the scattering curve best. Conformations representative of the major peak from the EOM analysis are shown. **F)** Fit (red line) of the EOM selected conformations to the SAXS data (blue dots). **G)** SAXS data curves for the crosslinked MutS tetramer bound to 21-bp DNA (orange) and 60-bp DNA (blue).

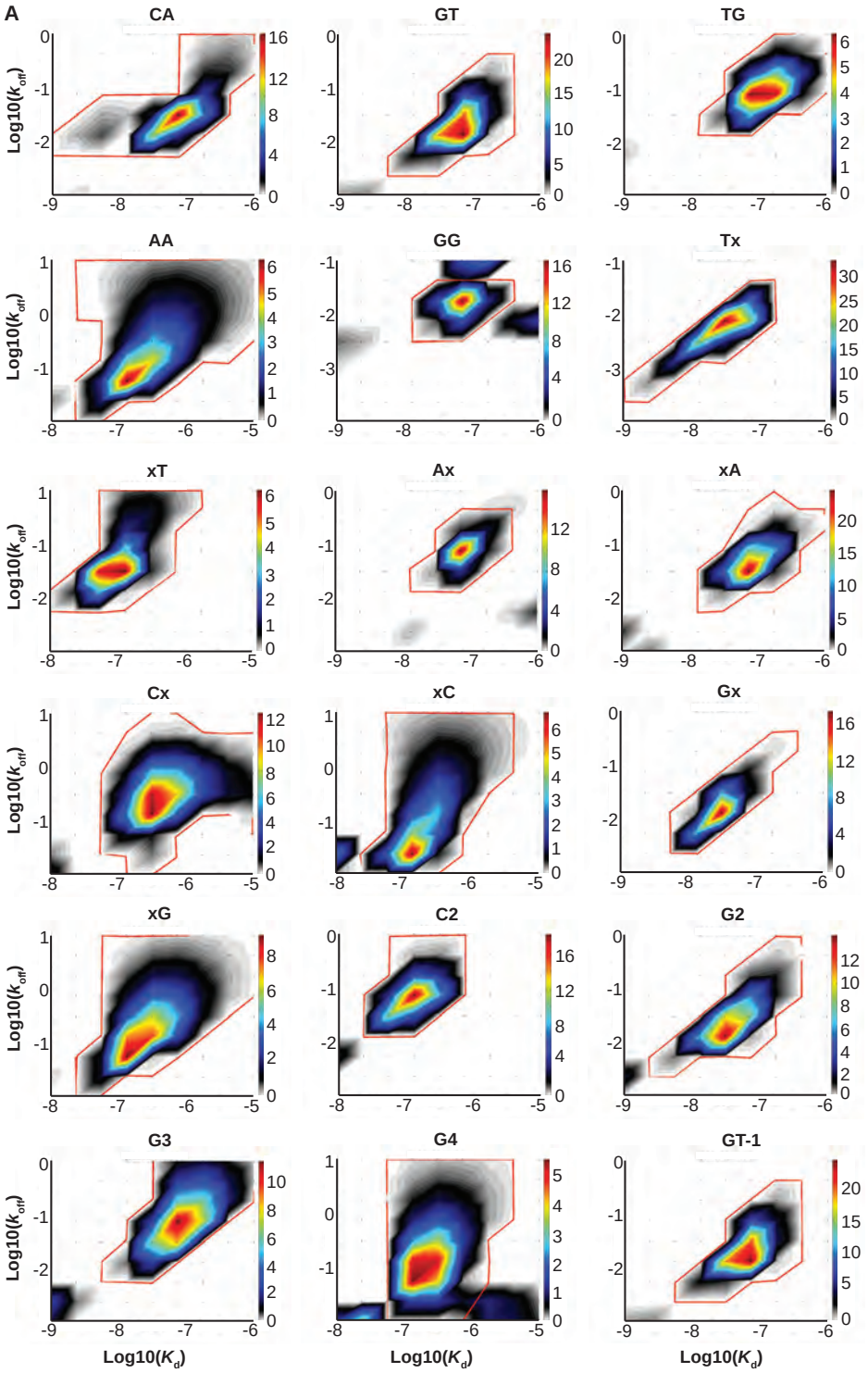
▶

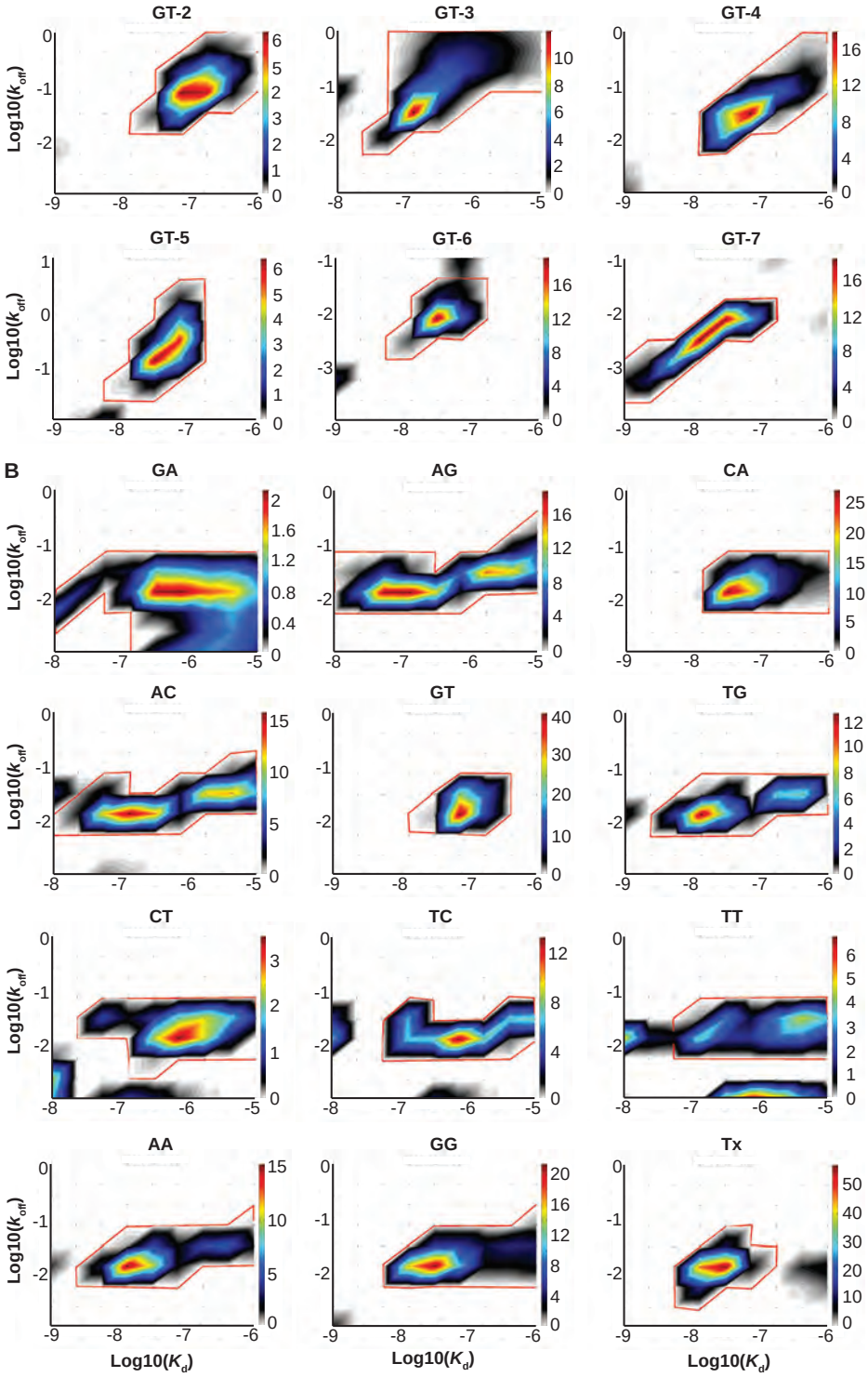

A
Distance distribution
P(r)
r (nm)
B
I(s) (relative)
Dimer data
Crystal structure, χ = 3.18
Extended arrangement, χ = 2.49
s (Å-1)
C
I(s) (relative)
Tetramer data
Extended arrangement, χ = 1.50
s (Å-1)
D
I(s) (relative)
Dimer
Tetramer
s2
E
Rg distribution for the tetramer
EOM analysis
Pool
Selected
Frequency
Radius of gyration (Å)
F
I(s) (relative)
SAXS data for tetramer
Ensemble fit, χ = 1.06
s (Å-1)
G
I(s) (relative)
Tetramer + 21-bp DNA
Tetramer + 60-bp DNA
s (Å-1)

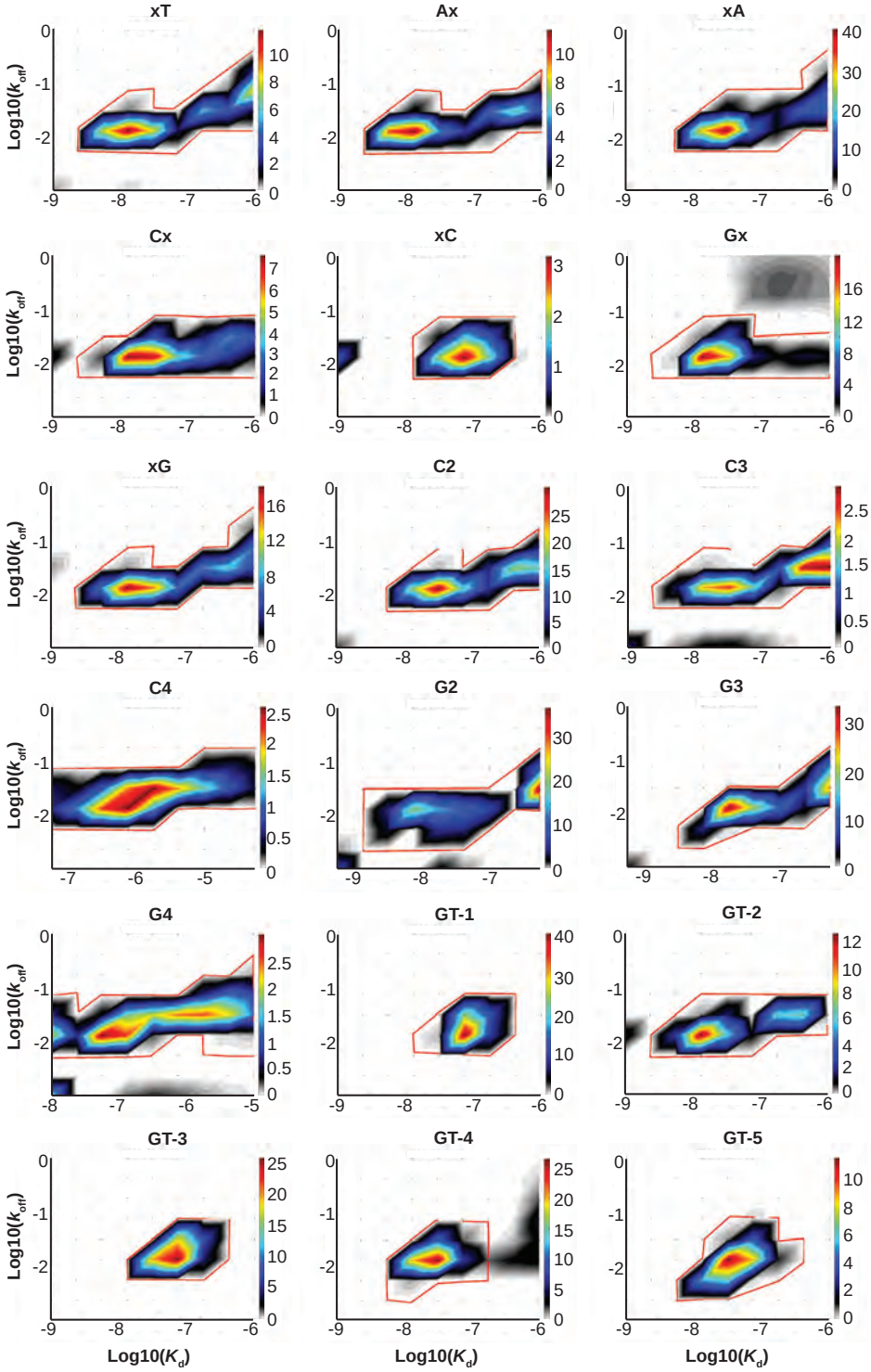


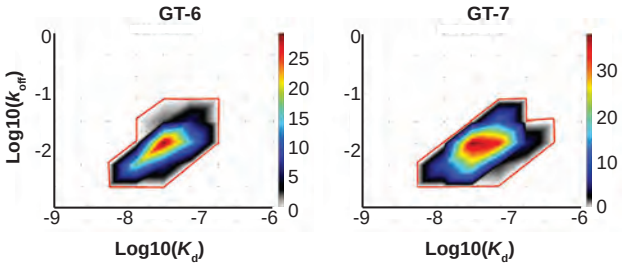

**Supplementary Figure S3.** DNA binding by the different MutS oligomers. **A)** Fluorescence polarization experiment with the MutS D835R dimer. When MutS is bound to 21-bp DNA with a GT mismatch and a 5' fluorescent TAMRA label, same-sequence 21-bp DNA can compete for binding as well as a 21-bp duplex with a $(dT)_{20}$ ssDNA overhang, suggesting no extra binding of MutS to the ssDNA. **B)** An EMSA assay shows shifts of radioactively labeled 41-bp heteroduplexes of DNA due to binding of wild-type MutS, dimeric MutS D835R or crosslinked tetrameric MutS (numbers refer to monomer concentrations in nM). First and second shift correspond dimer and tetramer binding respectively, whereas the highest shift indicates multiple binding by the tetramer at high protein concentration. When using wild type MutS, shifts corresponding to both dimer and tetramer species are observed at each protein concentration. **C)** SPR experiments of the MutS D835R dimer binding to unblocked and end-blocked DNA. Multiple concentrations of protein were flown over a chip with DNA with a GT mismatch (Supplementary Table S1). The dimer shows similar kinetic profiles in SPR measurements when binding to DNA with unblocked ▶

▶ ends (top graph) or DNA with blocked ends (bottom graph). In both cases, kinetics can be fitted using a one-phase binding model. **D)** Normalized superposition of SPR measurements of 160 nM (monomer concentrations) of wild type MutS, dimeric MutS D835R or the crosslinked MutS tetramer binding to DNA with a GT mismatch (Supplementary Table S1). Differences in kinetics can be observed between the oligomers: tetrameric MutS (orange curve) has a slower dissociation from the DNA, while dimeric MutS D835R (green curve) comes off fast. Wild type MutS (purple curve) appears to have a profile that is an intermediate of dimer and tetramer dissociations, suggesting a combination of these two kinetic profiles. **E)** Normalized SPR measurements show binding of 320 nM (monomer concentration) of the tetramer of MutS to different lengths of DNA with a GT mismatch (21-bp, 41-bp or 100-bp) immobilized on the chip without ssDNA linker. The slow dissociation of the tetramer is more pronounced with increasing DNA length. The peak at 120 s is due to the switch to the buffer injection step. **F)** Normalized SPR measurements show binding of 320 nM (monomer concentration) of the tetramer of MutS to DNA with a GT mismatch (Supplementary Table S1) immobilized with varying amounts (5, 10 or 25 RU). Different amounts of immobilized DNA correspond to different DNA densities on the chip surface. If the stronger binding of the tetramer observed in our assays was due to binding of two DNA strands in each of its two DNA-binding sites, this effect would be greater in the case of higher DNA density. We only see an increased effect when immobilizing as much as 25 RU, but not at the ~7 RU range that was used in all standard SPR experiments. **G)** SPR measurements of different concentrations of the tetramer of MutS binding to 21-bp DNA with a GT mismatch, either directly immobilized to the chip (right graph) of via a $(dT)_{20}$ linker (left graph). In the case of no ssDNA linker, the tetramer shows continuing dissociation, indicating that the slow tetramer dissociation in the assay with ssDNA linker is not due to aggregation under the experimental conditions.


A
CA
GT
TG
AA
GG
Tx
xT
Ax
xA
Cx
xC
Gx
xG
C2
G2
G3
G4
GT-1
Log10($k_{off}$)
Log10($K_d$)


GT-2
GT-3
GT-4
GT-5
GT-6
GT-7
B
GA
AG
CA
AC
GT
TG
CT
TC
TT
AA
GG
Tx
Log10($k_{off}$)
Log10($K_d$)


xT
Ax
xA
Cx
xC
Gx
xG
C2
C3
C4
G2
G3
G4
GT-1
GT-2
GT-3
GT-4
GT-5
Log10($k_{off}$)
Log10($K_d$)




**Supplementary Figure S4.** Examples of EvilFit heat maps for determination of kinetic parameters of MutS binding. Areas for which a weighted average was taken to obtain parameters are lined in red. **A)** Measurements in the absence of ATP, **B)** measurements in the presence of 1 mM ATP.

**Supplementary Table S1.** Sequences used in the SPR screen for binding of MutS to different mismatches.

| | |
|---|---|
| **Standard DNA sequence** | Flu 5′ AGTCGCCAGG * ACCAGTGTCA $(dT)_{20}$ 3′<br>3′ TCAGCGGTCC * TGGTCACAGT 5′ |
| **GT-1** | Flu 5′ AGTCGCCAGG G ACCAGTGTCA $(dT)_{20}$ 3′<br>3′ TCAGCGGTCC T TGGTCACAGT 5′ |
| **GT-2** | Flu 5′ AGTCGCCAGG T ACCAGTGTCA $(dT)_{20}$ 3′<br>3′ TCAGCGGTCC G TGGTCACAGT 5′ |
| **GT-3** | Flu 5′ TCAGCGGTCC G ACCAGTGTCA $(dT)_{20}$ 3′<br>3′ AGTCGCCAGG T TGGTCACAGT 5′ |
| **GT-4** | Flu 5′ TCAGCGGTCC T ACCAGTGTCA $(dT)_{20}$ 3′<br>3′ AGTCGCCAGG G TGGTCACAGT 5′ |
| **GT-5** | 5′ TCAGCGGTCC G TGGTCACAGT 3′<br>Flu 3′ AGTCGCCAGG T ACCAGTGTCA $(dT)_{20}$ 5′ |
| **GT-6** | 5′ TCAGCGGTCC T TGGTCACAGT 3′<br>Flu 3′ AGTCGCCAGG G ACCAGTGTCA $(dT)_{20}$ 5′ |
| **GT-7** | 5′ AGTCGCCAGG T TGGTCACAGT 3′<br>Flu 3′ TCAGCGGTCC G ACCAGTGTCA $(dT)_{20}$ 5′ |

Via a biotin-label attached at the end of the $(dT)_{20}$ ssDNA linker, the DNA was immobilized on a Biacore streptavidin chip. 'Flu' indicates a fluorescein moiety coupled to the indicated strand. The asterisks indicate positions of bases that were varied. In the case of insertions larger than one base, the extra bases were placed at the position of the asterisk in the bottom strand. For GT-1 to GT-7, colors of the flanking sequences correspond to the schematic representation in Figure 6B.

**Supplementary Table S2.** Activities of MutS variants *in vivo* and *in vitro*.

| Variant[a] | *In vivo* | | *In vitro*[d] | |
|---|---|---|---|---|
| | Median[b] | Range[c] | Relative activity (%) | SD (%) |
| pET-15b vector | 230 | 142-250 | - | - |
| WT MutS | 2 | 0-15 | 100 | 1 |
| Cysteine-free MutS | 3 | 1-11 | 98 | 7 |
| Single-cysteine MutS R848C | 10 | 1-59 | 113 | 3 |

[a] *E. coli* TX2929 (*mutS*$^-$) was transformed with the plasmid pET-15b (vector control), wild type *mutS* or *mutS* mutants as indicated. For details see 'Supplementary materials and methods'.

[b] Median number of rifampicin-resistant clones as counted on individual plates, which arise by spontaneous mutation in the *rpoB* gene. Cells lacking a functional MutS protein have a higher frequency of mutation.

[c] Minimum and maximum number of rifampicin-resitant clones on individual plates.

[d] DNA cleavage by MutS-activated MutH. 400 nM MutS, 1 mM MutL and 200 nM MutH were used to cleave 10 nM 484-bp GT mismatch containing DNA.

# 2

# 3

# Kinking of the coiled coil in the MutS lever domain allows DNA binding and release

Flora S. Groothuizen[1*], Alexander Fish[1*], Shreya D. Dharadhar[1], Herrie H.K. Winterwerp[1] and Titia K. Sixma[1]

[1] Division of Biochemistry and CGC.nl, Netherlands Cancer Institute, Amsterdam, the Netherlands * Equal contribution

*In preparation*

## Abstract

Mispaired bases in DNA are recognized by the DNA mismatch repair protein MutS, which then signals for repair of the error. During mismatch recognition, dimers of MutS embrace the DNA helix with their clamp domains. Previous studies showed generic flexibility of these domains of MutS in solution. Here we present a novel crystal structure of DNA-free MutS. In this structure, the clamp domains are repositioned due to kinking at specific sites in the helices in the lever domains, suggesting a defined hinge point for movement of the clamps. This hinging could allow for DNA to enter the dimer for initial binding, implying that this is a state of MutS that precedes DNA binding. We have created mutants aiming to perturb or stabilize the coiled coil where the kinking takes place. Preliminary DNA-binding assays with these mutants indicate that a balance between helical folding and kinking at this region is important for DNA binding and release of MutS.



## Introduction

The DNA mismatch repair (MMR) pathway is responsible for maintaining genetic information by correcting base-substitution and insertion-deletion mismatches generated during DNA replication (1,2). MMR deficiency results in a mutator phenotype, and in humans it predisposes to cancer, referred to as HNPCC or Lynch syndrome (3). MutS is the protein that is essential for the detection of the DNA mismatch, and signals for repair by forming a sliding clamp that activates MutL (4,5). The MutS and MutL proteins have been evolutionarily conserved from bacteria to mammals.

MutS proteins form constitutive dimers, while some prokaryotic MutS homologs can also tetramerize through their C-terminal domain, but this is not required for MMR (6-8). In eukaryotic cells, the MutS homologs that are active in MMR form heterodimers (MSH2/MSH6 or MSH2/MSH3) (9). In both prokaryotes and eukaryotes, MutS proteins are asymmetric dimers during mismatch recognition, where only one of the subunits of the dimer recognizes the DNA mismatch through its mismatch-binding domain. Both subunits in MutS protein dimers, however, embrace the DNA helix with their clamp domains (10-13).

The binding of the MutS clamp domains around the DNA helix is expected to be a general feature, also for binding to DNA without a mismatch. However, without interaction of the mismatch-binding domain with a DNA mismatch, the MutS dimer has only low affinity for DNA and releases without initiation of repair (8). A crystal structure of DNA-free *Thermus aquaticus* MutS indicated disorder of large portions of MutS, including the clamp domains, in the absence of DNA, while the dimer itself was kept intact (11). Conformational freedom

of the clamp domains was also observed in SAXS studies with the DNA-free protein (8). Both studies indicates that in the absence of DNA, the clamp domains of MutS dimers are flexible, but how such flexibility is achieved is unclear.

Here we present a novel crystal structure of a DNA-free *Escherichia coli* MutS dimer, in which the clamp domains are displaced compared to the mismatch-bound state. In this structure, the conformational freedom originates from hinging over specific regions in the lever domains that connect the clamp domains to the rest of the protein. These lever domains extend from within the core domains of MutS and are each composed of two helices that form an anti-parallel coiled-coil structure. Using mutagenesis and DNA binding studies, we have investigated whether kinking of the helices in this coiled coil is required for DNA binding by MutS.



## Materials and methods

### Proteins

All mutations were introduced in the *mutS* gene in vector pET-3d, using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and appropriate primer sequences (obtained from IDT), following the manufacturers protocol.

Wild-type (WT) and mutant MutS proteins were expressed and purified as described (10,14), except that in the final gel filtration buffer KCl was used instead of NaCl, and for mutants MutS-G and MutS-FG the lysis buffer contained 10% glycerol and an increased salt concentration of 400 mM.

### Crystallography

Crystallization of full-length DNA-free MutS P839E was performed using MutS at a concentration of 50 mM mixed with 100 μm ADP. The protein was crystallized using vapor diffusion with 3-8% dioxane, 1.4-1.7 mM $(NH_4)_2SO_4$, and 100 mM Hepes pH 7.0. The crystal was transferred to mother liquor supplemented with 30% glycerol before flash cooling it in liquid nitrogen.

Crystallographic data was collected at ESRF beamline ID14-1 and was processed using iMosflm (15) and Scala (16). The initial structure was solved using molecular replacement in Phaser (17) with part of chain A of PDB entry 1WB9 as search model. Structure refinement was performed using Buster (18). The C-terminal tetramerization domains are not resolved in density, probably because they are flexible (8) and their position is not stabilized by crystal contacts in this crystal form. It should be noted that structure refinement is ongoing; see Table 1 for current crystallographic statistics.

**Table 1.** Crystallographic data collection and refinement statistics.

| | |
|---|---|
| **Data collection** | |
| λ (Å) | 0.934 |
| Resolution range (Å) | 71.6102.85 (2.92-2.85) |
| Completeness (%) | 98.5 (98.2) |
| I/σ(I) | 5.9 (2.0) |
| $R_{merge}$ (%) | 21.8 (74.4) |
| Space group | P $2_1$ $2_1$ $2_1$ |
| Cell dimensions (Å) | 113.38, 113.53, 158.90 |
| Total no. of observations | 195304 (28145) |
| Total no. of unique reflections | 47557 (6833) |
| Multiplicity | 4.1 (4.1) |
| Wilson's B-factor (Å$^2$) | 53.1 |
| **Refinement** | |
| No. of atoms (protein/ligand/solvent) | 11740/146/48 |
| Average B-factor (Å$^2$) | 54.2 |
| $R_{free}$ reflections | 2357 |
| $R_{work}$ (%) | 19.3 |
| $R_{free}$ (%) | 25.1 |
| Bond r.m.s.d. (Å) | 0.010 |
| Angle r.m.s.d. (°) | 1.20 |

Numbers within brackets refer to the highest resolution shell.

### Protein stability measurements

Protein stability was assessed using an Optim1000 machine (Avacta Analytical). WT or mutant MutS proteins were diluted to 0.25 mg/mL and subjected to a temperature gradient to determine melting temperatures ($T_m$), which was read out by changes in tryptophan fluorescence.

### Protein Equilibrium DNA binding

Fluorescence polarization measurements to assess DNA-binding affinities of WT and mutant MutS proteins were performed in buffer with 25 mM Hepes pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 10% Glycerol, and 1 mg/ml BSA. A concentration of 0.5 nM of 5' labeled TAMRA-21-bp DNA with a mismatch at position 9 (5'-TAMRA-AGCTGCCAGG CACCAGTGTCA annealed with TGACACTGGTGCTTGGCAGCT) was used as fluorescent probe. MutS proteins were serial diluted in black flat-bottomed 384 well

plates (Corning) in 30 mL volumes. The plate was equilibrated at RT for 5 minutes, after which polarization of the TAMRA label was read out in a PHERAstar FS machine (BMG Labtech) with a 540/590 (excitation/emission) FP module. $K_d^{app}$ values were determined using nonlinear regression fitting with Graphpad Prism (19), with a model for single-site binding that corrects for depletion of the protein:

$$FP = FP^{free} + (FP^{bound} - FP^{free}) \times \frac{c + K_d^{app} + [S] - \sqrt{(-c - K_d^{app} - [S])^2 - 4c \times [S]}}{2c}$$

where $c$ is the concentration of the DNA probe, and MutS concentration $[S]$ is expressed as monomers.



**DNA-binding kinetics**

Kinetics of binding to 21-bp DNA containing a GT mismatch, which was attached to a streptavidin chip via a biotin-conjugated $(dT)_{20}$ linker, were determined using surface plasmon resonance (SPR). The measurements were performed in a Biacore T200 system (GE Healthcare) at 25 °C with the same setup and determination of kinetic parameters as described previously (8).

## Results

**Crystal structure of DNA-free MutS**

Full-length *E. coli* MutS was crystallized in the absence of DNA and its structure was solved at a resolution of 2.85 Å (Figure 1A). In contrast to the previously published DNA-free *Taq* MutS crystal structure (11), however, the clamp domains are resolved in density in this new structure. The arrangement of the subunits in the dimer is slightly more compact than mismatch-bound structures of MutS, probably allowed due to the absence of DNA between the subunits. The two mismatch-binding domains are each in different positions. This is consistent with the variable positions of mismatch-binding domains of the subunits that do not contact the mismatch in other crystal structures (8,11,20).

Our crystal structure of the MutS dimer contains two ADP nucleotides in its ATP-binding domains. Although MutS has asymmetric ATPase activity within dimers (21), the two-ADP state is the most abundant nucleotide state found for DNA-free MutS in native mass-spectrometry measurements (22). The ATPase domains are therefore symmetric in this DNA-free structure, while mismatch recognition induces asymmetry in the two ATP binding sites of MutS homodimers (10,20),

The position of the clamp domains in this DNA-free crystal structure is different than in DNA-bound MutS structures (8,10,14), confirming that in the absence of DNA, there

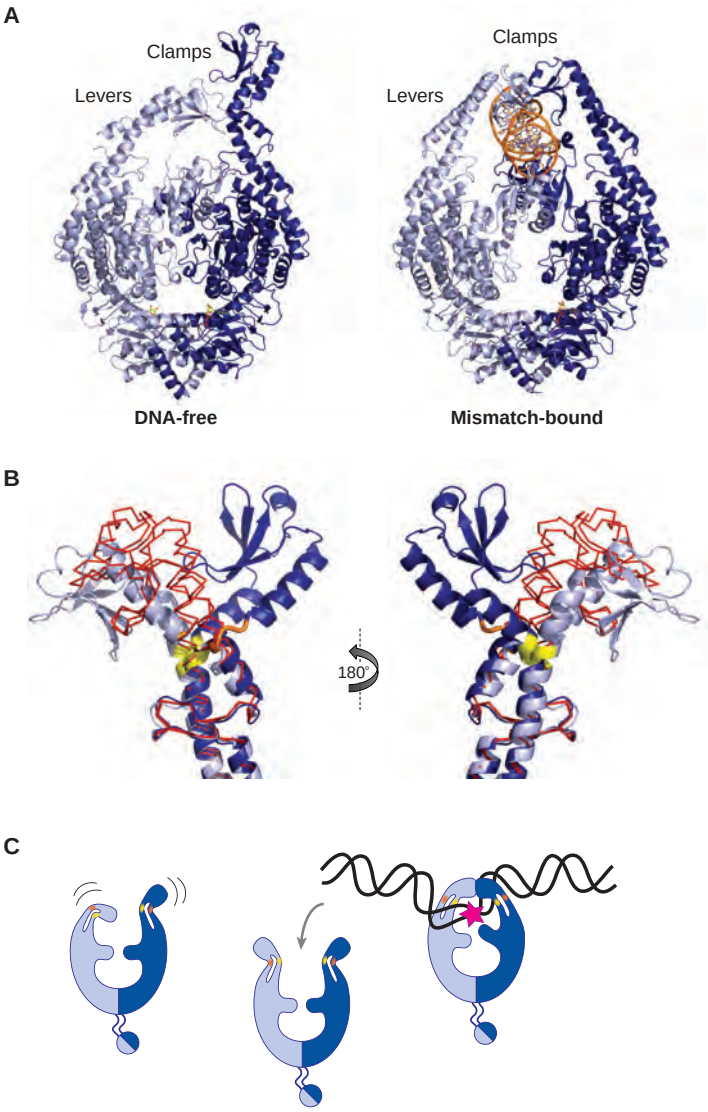



**Figure 1.** DNA-free crystal structure of MutS with differently positioned clamp domains. **A)** Crystal structure of mismatch-bound (PDB entry 1E3M) and DNA-free *E. coli* MutS. The two subunits are shown as dark blue and light blue cartoons, and DNA is shown in orange. **B)** Superposition to show the kinking that reorients the clamp domains of both subunits in our structure (dark blue and light blue cartoons) with both subunits of mismatch-bound MutS (shown as red ribbons). Residues involved in the kinking of the helices are shown in orange (441-443) and yellow (515-517). **C)** Schematic model in which occasional kinking of the lever-domain helices allows the MutS dimer to 'open up' so that DNA, possibly containing a mismatch (pink star), can be bound.

is conformational freedom of these domains (Figure 1B). One of the clamp domains is positioned closer to the core of the dimer (inward), while the other has moved outward. Strikingly, both motions originate from hinging over the same two regions (residues 441-443 and 515-517) in the lever domains of MutS.

The hinge points in the lever domains in our DNA-free MutS structure involve kinking of the two helices of this domain. Interestingly, relatively high B-factors in the mismatch-bound MutS structures already suggested some degree of disorder in these parts of the helices. The presence of these defined hinge points suggests a way for MutS dimers to 'open up' and allow a DNA helix to enter the DNA-binding site (Figure 1C), while DNA could also be released in this manner if no sliding clamp is formed. We therefore hypothesize that this crystal structure represents a state of MutS that precedes DNA binding.

**Mutations to influence the kinking in the lever domain**

We wanted to verify that the new crystal structure with specific hinge regions represents a state that precedes binding of DNA, by studying the effects of mutations in these regions on DNA binding. The lever domains of MutS consist mainly of two helices that form a left-handed antiparallel coiled coil arrangement (23) (Figure 2A-B). Both helices in this coiled coil are kinked in the DNA-free structure, and we hypothesized that the ability to kink is inversely related to the stability of the coiled coil. Using this hypothesis, we designed mutants aimed to stabilize or perturb the coiled coil, to further investigate the importance of the kinking that we observe.

Coiled-coil structures are common structural motifs in which helices wrap around each other to form a super helix. While the effect of sequence variation on coiled-coil stability is not fully understood, it is generally accepted that hydrophobic residues such as leucines and alanines at the *a* and *d* positions of the heptad repeats facilitate dimer interaction. Similarly, charged residues such as glutamine or lysine at the *e* and *g* positions facilitate interhelical electrostatic interactions (24).

We designed and purified several mutants aiming to change the helical stability of the lever domains and labeled them alphabetically with A to G (Table 2). Two variants aimed to promote helical folds at the kinking regions (MutS-A and MutS-B). To promote coiled-coil stability, mutant MutS-C contains hydrophobic residues commonly found at position *a* and *d* in the heptad repeat of coiled coils, while MutS-CD also has additional charged residues at position *e* and *g* to allow for more ionic interactions (Figure 2B). Since the clamp domains of the two subunits in the MutS dimer are close to each other in the DNA-bound state (Figure 2C), repulsion between arginines in these domains could assist in the 'opening up' of the clamp domains. Therefore, we also designed a mutant in which these charges are removed (MutS-E). Finally, to perturb the helical fold, we introduced Pro/Gly/

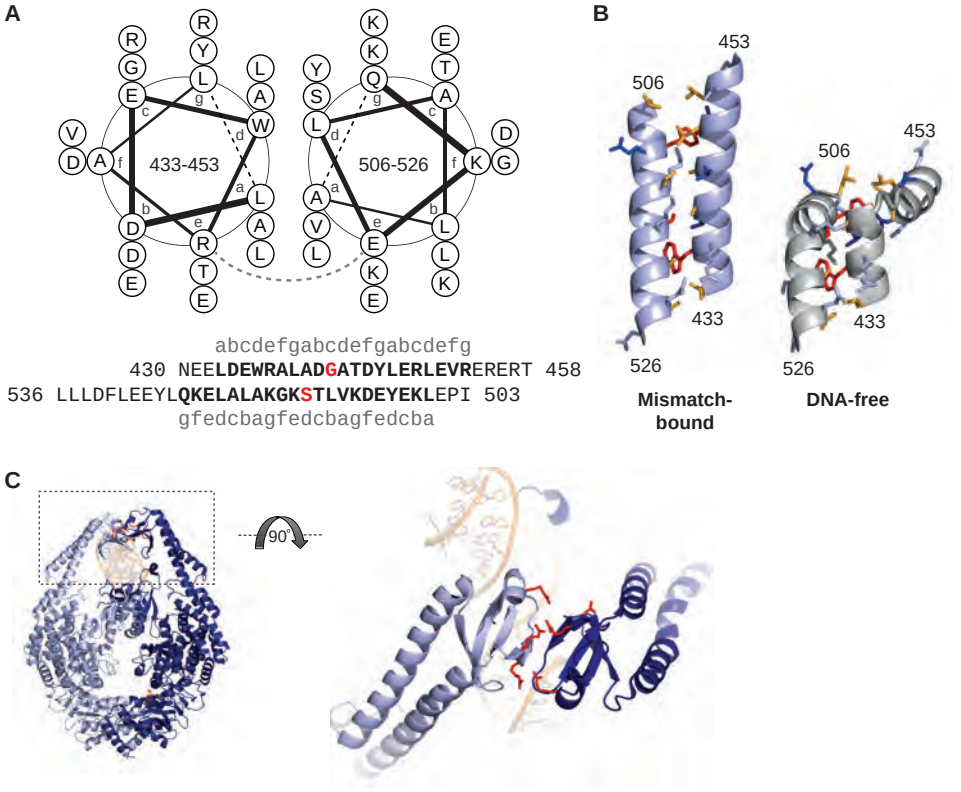



**Figure 2.** The coiled coil in the lever domain of MutS. **A)** Helical wheel representation of the coiled coil in the lever domain of WT MutS. Sequences of the helices and their corresponding heptad assignment are written below. The residues around which the kinking takes place are indicated in red. **B)** Structure of the coiled coil of the lever domain as present in PDB entry 1E3M and chain A of our DNA-free structure. Residues involved in the hydrophobic core of the coiled coil (*a* and *d* in the heptad) are colored orange, or red for the residues that have been mutated in this study. Residues that could contribute to stability due to ionic interactions (*e* and *g* in the heptad) are shown as light blue sticks, and dark blue for the residues that have been mutated in this study. **C)** Clamp domains of DNA-bound MutS (PDB entry 1E3M). The arginines that have been mutated in this study are shown as red sticks.

Pro motifs at the kinking regions of each of the helices separately (MutS-F and MutS-G) or in both helices (MutS-FG).

Since mutations affecting secondary structure can influence overall protein stability, the protein variants were subjected to thermostability measurements and compared to WT protein (Table 2). Except for MutS-B, all variants had slightly lower melting temperatures than WT MutS ($T_m$ = 45.5 °C). MutS-G showed a $T_m$ that was 5 °C lower than WT and also required additional salt and glycerol in the lysis step of purification, indicating that its stability is slightly compromised. Assays performed with some of the mutant proteins should therefore be interpreted with caution.

**Table 2.** Properties of the MutS variants used in this study.

| **Protein** | | | **$T_m$ (°C)** | **$K_d^{app}$ in FP (nM)** | **$K_d^{app}$ in SPR (nM)** | **$k_{off}$ in SPR ($s^{-1}$)** |
|---|---|---|---|---|---|---|
| WT MutS | | | 45.5 | 3.2 ± 0.5 | 69 ± 4 | 0.025 |
| **Aimed to promote closed state** | **Mutations** | **Description** | **$T_m$ (°C)** | **$K_d^{app}$ in FP (nM)** | **$K_d^{app}$ in SPR (nM)** | **$k_{off}$ in SPR ($s^{-1}$)** |
| MutS-A | D441A/G442A/A443V/ Y509F | Promote helical fold in *h1*, increase *cc* packing with A443V/Y509F | 42.5 | 0.95 ± 0.2 | 74 ± 6 | 0.026 |
| MutS-B | S516A | Promote helical fold at kink in *h2* | 45.5 | 2.5 ± 0.3 | 73 ± 10 | 0.030 |
| MutS-C | W436L/Y509L/S516A | Promote quality of hydrophobic core of *cc* | 44.5 | 2.5 ± 0.3 | 76 ± 20 | 0.016 |
| MutS-CD | W436L/L439E/T444E/ T446E/Y509L/E510K/S516A | Promote quality of hydrophobic core and salt bridges of *cc* | 44.5 | 1.2 ± 0.2 | 108 ± 23 | 0.0057 |
| MutS-E | R479A/R491A/R492A | Charge removal in clamp domains | 43.5 | 840 ± 60 | 200 ± 38 | 0.030 |
| **Aimed to promote kinking** | **Mutations** | **Description** | **$T_m$ (°C)** | **$K_d^{app}$ in FP (nM)** | **$K_d^{app}$ in SPR (nM)** | **$k_{off}$ in SPR ($s^{-1}$)** |
| MutS-F | D441P/A443P | Pro/Gly/Pro motif in *h1* | 42.5 | 22 ± 0.7 | 220 ± 20 | 0.080 |
| MutS-G | T515P/S516G/T517P | Pro/Gly/Pro motif in *h2* | 40.5 | >670 | 900 ± 90 | 0.33 |
| MutS-FG | D441P/A443P/T515P/ S516G/T517P | Pro/Gly/Pro motif in *h1* and *h2* | 42.5 | >4600 | *N/A* | *N/A* |

*h1* = helix 1 (kinks around residue 442); *h2* = helix2 (kinks around residue 516); *cc* = coiled coil; standard errors of fitting are indicated after the ± sign; *N/A*: kinetic parameters could not be determined.

**DNA binding is affected by mutations in the kinking regions**

To investigate whether the mutations in the coiled coil influenced DNA binding abilities of the proteins, we measured equilibrium binding using fluorescence polarization (FP) with TAMRA-labeled DNA containing a GT mismatch. Interestingly, the MutS proteins with mutations aimed to increase the stability of the helices showed similar or stronger binding than WT protein (Table 2, Figure 3A). However, MutS-E, with charges removed from the clamp domains, showed weaker binding ($K_d^{app}$ = 840 ± 60 nM) than WT MutS ($K_d^{app}$ = 3.2 ± 0.5 nM). The mutants that were designed aiming to destabilize the helical fold at the kinking regions also bound weaker to DNA, where mutating residues 515-517 (MutS-G, $K_d^{app}$ > 670 nM) appeared to have a larger effect than mutating residues 441-443 (MutS-F, $K_d^{app}$ = 22 ± 0.7 nM).

We wondered whether these differences in affinities for DNA originated from changes in binding kinetics due to the mutations in the lever domains. Mutants that did not show differences in overall affinity could also have altered DNA-binding kinetics. Therefore, we performed surface plasmon resonance (SPR) assays, in which DNA binding kinetics can be assessed. Since the mutants were created in context of the WT protein, determination of kinetics parameters are complicated by the dimer-tetramer equilibrium (8), but the approximation of kinetic parameters will still be useful to compare the variants to WT protein. Because the setup of SPR measurements is different, the values of affinities differ in order of magnitude from the FP measurements, but variations between mutants remain mostly consistent.

Just as in the FP assays, MutS-A, MutS-B and MutS-C had similar overall affinities for mismatched-DNA as WT MutS (Table 2, Figure 3B), although MutS-A and MutS-C showed some aspecific binding to the chip in the assay. For MutS-A and MutS-B, $k_{off}$ values (0.026 $s^{-1}$ and 0.030 $s^{-1}$ respectively) are also comparable to that of WT (0.025 $s^{-1}$), indicating that the mutations did not influence DNA binding kinetics. MutS-C, however, which does not differ from WT in affinity for DNA, appears to release somewhat slower from DNA ($k_{off}$ = 0.016 $s^{-1}$), while very slow release ($k_{off}$ = 0.0057 $s^{-1}$) and slow association are observed in the profiles of MutS-CD (Figure 3B). The weaker DNA binding by MutS-E, also observed in FP, does not result from changes in release rate ($k_{off}$ = 0.030 $s^{-1}$) but this mutant mostly shows slower association on DNA (Figure 3B).

The mutants that have a Pro/Gly/Pro motif introduced in one of the two helices of the lever domain (MutS-F and MutS-G) release faster from DNA than WT ($k_{off}$ = 0.080 $s^{-1}$ and 0.33 $s^{-1}$ respectively). The overall affinity is also much weaker than for WT MutS, showing that these variants do not compensate with a faster on-rate. For the combination mutant MutS-FG, binding was to weak to assess kinetic parameters with SPR.

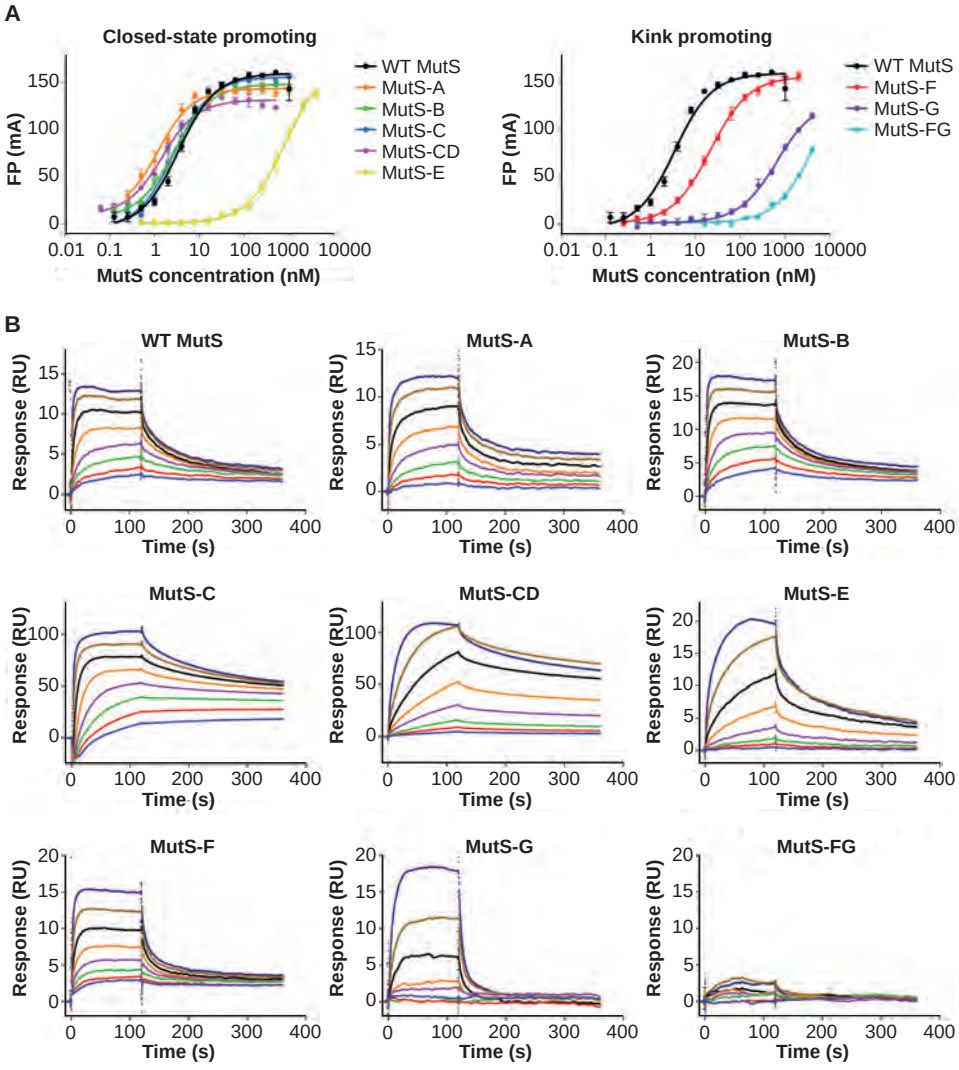



**Figure 3.** DNA binding by the MutS variants. **A)** Equilibrium binding measurements using TAMRA-labeled 21-bp DNA containing a GT mismatch. Data points are averages between three measurements and error bars represent SEM. **B)** Kinetic profiles of the MutS variants binding to 21-bp DNA containing a GT mismatch as measured with SPR. Colors represent different protein concentrations, ranging from 5 nM to 640 nM.

## Discussion

MutS is known to possess generic conformational freedom of its clamp domains in solution, as observed in SAXS studies (8) and in a previous crystal structure of DNA-free *Taq* MutS (11). However, the nature of this disorder remained unclear. In this work, we have shown that defined regions in helices in the lever domains of MutS can allow for movement of the clamp domains, as observed in the new DNA-free crystal structure. Our initial measurements indicate that perturbation of the helical fold in these regions influences DNA binding properties of MutS, therefore we hypothesize that the state of MutS in this new crystal structure is one that precedes DNA binding.

The symmetry of the ATPase domains in this structure suggests that MutS homodimers are essentially symmetrical in the absence of a DNA mismatch, although flexible in certain regions. Similar symmetry was observed in the crystal structure of DNA-free Taq MutS, which had no nucleotide bound (11). Mismatch recognition then initially induces asymmetry in the nucleotide status (10,20). Interestingly, when validation of the mismatch results in binding of two ATP nucleotides (22), this asymmetry is lost again when the sliding clamp conformation is adopted (Chapter 4).

While some mutations had little effect on DNA binding, those mutations that aimed to improve the stability of the coiled coil of the MutS lever domain resulted in slower kinetics. This could indicate that the 'opening up' of the dimer to allow DNA entering and release indeed requires some potential to distort the coiled-coil fold.

MutS-E, which had several positive charges in the clamp domains removed, showed weaker binding to DNA. This variant was designed aiming to change the charge repulsion between clamp domains within the dimer, so that it would be in the closed state more often than opening up. Indeed this could explain the reduced binding, but the removal of the positive charges are likely to also affect initial binding to the negatively charged backbone DNA itself. That could explain why the off-rate is not affected by the mutations, but the low affinity is mainly due to slow association. This should be further investigated before drawing any conclusions on the exact effects.

Introducing Pro/Gly/Pro motifs at the kinking sites (MutS-F, MutS-G, MutS-FG), which aimed to destabilize the helical folds in those regions, had clear effects on DNA binding . Mutations in either of the two helices resulted in weaker binding to DNA, mainly due to faster release. In contrast to mutations that aim to improve packing of the coiled coil (MutS-C, MutS-CD), the effect of the Pro/Gly/Pro motif on release is not balanced out by a comparable effect on association. Possibly, the effect on the helical stability is too big to form a stable MutS:DNA complex. The effect of introducing the Pro/Gly/Pro motif is smaller for residues 441-443 than for residues 515-517, in-line with the smaller sequence change since residue 442 is already a glycine in the native protein. However, the

variants with the Pro/Gly/Pro motif all showed reduced thermostability which may have contributed to the effects on DNA binding.

It is clear that mutating the hinge regions can affect DNA binding. However, it would be interesting to validate that the mutations that we created have the predicted effect on the helical stability of the kinking regions. This could be done using Förster resonance energy transfer (FRET) experiments in which the clamp domains of a single-cysteine MutS variant (25,26) are fluorescently labeled (e.g. at position 456). If we can normalize for the labeling efficiency, the amount of intramolecular FRET (for heterodimers of MutS with two different labels) would be an indication of the amount of protein that is in the closed state.

In conclusion, we have found unexpected defined hinge points in the middle of helices in the lever domains of MutS. It suggests a manner in which this protein can subtly achieve flexibility before adopting the more ordered and probably favorable DNA-bound state. Some mutations in the hinging region have devastating effects on DNA binding properties of MutS, suggesting that there is a fine balance of the helical stability of the lever domains for proper DNA binding and release.

## References


1. T.A. Kunkel and D.A. Erie (2005) DNA mismatch repair. *Annu. Rev. Biochem.*, **74**, 681–710.
2. J. Jiricny (2013) Postreplicative mismatch repair. *Cold Spring Harbor Perspect. Biol.*, **5**, a012633.
3. H.T. Lynch and A. de la Chapelle (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.*, **36**, 801–818.
4. M. Yamaguchi, V. Dao and P. Modrich (1998) MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J. Biol. Chem.*, **273**, 9197–9201.
5. S. Acharya, P.L. Foster, P. Brooks and R. Fishel (2003) The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair. *Mol. Cell*, **12**, 233–246.
6. M.A. Calmann, A. Nowosielska and M.G. Marinus (2005) Separation of mutation avoidance and antirecombination functions in an *Escherichia coli mutS* mutant. *Nucleic Acids Res.*, **33**, 1193–1200.
7. M.L. Mendillo, V.V. Hargreaves, J.W. Jamison, A.O. Mo, S. Li, C.D. Putnam, V.L. Woods Jr and R.D. Kolodner (2009) A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 22223–22228.
8. F.S. Groothuizen, A. Fish, M.V. Petoukhov, A. Reumer, L. Manelyte, H.H.K. Winterwerp, M.G. Marinus, J.H.G. Lebbink, D.I. Svergun, P. Friedhoff and T.K. Sixma (2013) Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp formation. *Nucleic Acids Res.*, **41**, 8166–8181.
9. S. Acharya, T. Wilson, S. Gradia, M.F. Kane, S. Guerrette, G.T. Marsischky, R.D. Kolodner and R. Fishel (1996) hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 13629–13634.
10. M.H. Lamers, A. Perrakis, J.H. Enzlin, H.H.K. Winterwerp, N. de Wind and T.K. Sixma (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature*, **407**,

711–717.

11. G. Obmolova, C. Ban, P. Hsieh and W. Yang (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703–710.
12. J.J. Warren, T.J. Pohlhaus, A. Changela, R.R. Iyer, P.L. Modrich and L.S. Beese (2007) Structure of the human MutSα DNA lesion recognition complex. *Mol. Cell*, **26**, 579–592.
13. S. Gupta, M. Gellert and W. Yang (2011) Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat Struct Mol Biol*, **19**, 72–78.
14. G. Natrajan, M.H. Lamers, J.H. Enzlin, H.H.K. Winterwerp, A. Perrakis and T.K. Sixma (2003) Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.*, **31**, 4814–4821.
15. T.G.G. Battye, L. Kontogiannis, O. Johnson and H.R. Powell (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr.*, **D67**, 271–281.
16. P.R. Evans (2006) Scaling and assessment of data quality. *Acta Crystallogr.*, **D67**, 72–82.
17. A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni and R.J. Read (2007) Phaser crystallographic software. *J. appl. Crystallogr.*, **40**, 658–674.
18. G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, O.S. Smart, W. Paciorek, P. Roversi, A. Sharff, C. Vonrheinm and T.O. Womack (2011) BUSTER Cambridge, United Kingdom: Global phasing Ltd.
19. H. Motulsky (1999) Analyzing data with GraphPad prism. GraphPad Software Inc.
20. M.H. Lamers, D. Georgijevic, J.H. Lebbink, H.H.K. Winterwerp, B. Agianian, N. de Wind and T.K. Sixma (2004) ATP increases the affinity between MutS ATPase domains: implications for ATP hydrolysis and conformational changes. *J. Biol. Chem.*, **279**, 43879–438885.
21. M.H. Lamers, H.H.K. Winterwerp and T.K. Sixma (2003) The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J.*, **22**, 746–756.
22. M.C. Monti, S.X. Cohen, A. Fish, H.H.K. Winterwerp, A. Barendregt, P. Friedhoff, A. Perrakis, A.J.R. Heck, T.K. Sixma, R.H.H. van den Heuvel and J.H.G. Lebbink (2011) Native mass spectrometry provides direct evidence for DNA mismatch-induced regulation of asymmetric nucleotide binding in mismatch repair protein MutS. *Nucleic Acids Res.*, **39**, 8052–8064.
23. A. Lupas (1996) Coiled coils: new structures and new functions. *Trends Biochem. Sci.*, **21**, 375–382.
24. K.M. Arndt, J.N. Pelletier, K.M. Müller, A. Plückthun and T. Alber (2002) Comparison of in vivo selection and rational design of heterodimeric coiled coils. *Structure*, **10**, 1235–1248.
25. L. Giron-Monzon, L. Manelyte, R. Ahrends, D. Kirsch, B. Spengler and P. Friedhoff (2004) Mapping protein-protein interactions between MutL and MutH by cross-linking. *J. Biol. Chem.*, **279**, 49338–49345.
26. I. Winkler, A.D. Marx, D. Lariviere, R. Heinze, M. Cristóvão, A. Reumer, U. Curth, T.K. Sixma and P. Friedhoff (2011) Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in *Escherichia coli*. *J. Biol. Chem.*, **286**, 17326–17337.

# 4

# MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA

Flora S. Groothuizen[1], Ines Winkler[2], Michele Cristóvão[2], Alexander Fish[1], Herrie H.K. Winterwerp[1], Annet Reumer[1], Andreas D. Marx[2], Nicolaas Hermans[3], Robert A. Nicholls[4], Garib N. Murshudov[4], Joyce H.G. Lebbink[3,5], Peter Friedhoff[2] and Titia K. Sixma[1]

[1] Division of Biochemistry and CGC.nl, Netherlands Cancer Institute, Amsterdam, the Netherlands [2] Institute for Biochemistry, Justus-Liebig-University, Giessen, Germany [3] Department of Genetics, Cancer Genomics Netherlands, Erasmus Medical Center, Rotterdam, the Netherlands [4] Structural Studies Division, MRC Laboratory of Molecular Biology, Cambridge, United Kingdom [5] Department of Radiation Oncology, Erasmus Medical Center, Rotterdam, the Netherlands

*Submitted*

## Abstract

To avoid mutations in the genome, DNA replication is generally followed by DNA mismatch repair (MMR). MMR starts when a MutS homolog recognizes a mismatch and undergoes an ATP-dependent transformation to an elusive sliding clamp state. How this transient state promotes MutL homolog recruitment and activation of repair is unclear. Here we present a crystal structure of the MutS/MutL complex based on cysteine crosslinking and examine how large conformational changes lead to activation of MutL. The structure captures MutS in the sliding clamp conformation, where tilting of the MutS subunits across each other pushes DNA into a new channel, and reorientation of the connector domain creates an interface for MutL with both MutS subunits. Our work explains how the sliding clamp promotes loading of MutL onto DNA, to activate downstream effectors. We thus elucidate a crucial mechanism that ensures that MMR is initiated only after detection of a DNA mismatch.



## Introduction

To enable the correct and complete transfer of genetic information during cell division, DNA polymerases efficiently replicate the genome by pairing nucleotide bases opposite their complementary template base. However, despite the polymerase proofreading ability (1), incorrect nucleotides are occasionally incorporated into the new DNA strand, resulting in mutations when left uncorrected. To reduce the number of such mismatches and maintain genomic stability, replication is followed by DNA mismatch repair (MMR) in almost all cellular organisms (2). Initiation of this MMR system is evolutionarily conserved, although in eukaryotes heterodimeric homologs replace the bacterial homodimeric components. Defects in MMR result in a mutator phenotype and in humans in predisposition for cancer, known as Lynch syndrome or HNPCC (3).

MMR is initiated when a MutS homolog binds to a mismatch. In this mismatch recognition step, the MutS dimer kinks the DNA at the site of the mismatch and stacks a phenylalanine onto the mispaired base (4-6). Upon ATP binding MutS releases the mismatch (7) and travels as a 'sliding clamp' along the DNA helix (8-10), and only this specific state of MutS is recognized by MutL or its homologs (11-14).

MutL proteins are constitutive dimers through their C-terminal domains, while the N-terminal ATPase domains reorganize and dimerize upon ATP binding (11,15,16). Once recruited by the MutS sliding clamp, the MutL homologs activate downstream repair. This includes the nicking of the newly replicated strand by a nuclease, which is either part of the MutL C-terminal domain (17), or a separate protein such as MutH in *Escherichia coli*.

MutL also activates UvrD in bacteria to unwind the DNA (18), after which the new DNA strand can be removed and re-replicated (2).

As loss of MutS homologs (MSH2, MSH3 and MSH6 in humans) or MutL homologs (MLH1 and PMS2 in humans) leads to mutator and/or cancer phenotypes, these proteins evidently have critical roles in mismatch repair and it is therefore important to understand their exact mechanism. Despite extensive studies (8,19-21), it is unclear how MutS achieves the sliding-clamp state, how this promotes MutL recognition and why this results in activation of the MutL protein.



Here, we trap the transient complex between MutS and MutL to resolve a crystal structure of the MutS sliding clamp bound to MutL. This is, to our knowledge, the first time that not only this MutS conformation but also the complex between MutS and MutL could be observed. We show how rearrangements in MutS promote interactions from both MutS subunits with a single MutL N-terminal domain, and how this domain is then positioned to load onto DNA running through a novel channel in the MutS dimer. We use biophysical methods to analyze the transient states and mechanistically understand the specificity and effect of MutL binding to MutS, and functional assays to address how this affects MMR initiation.

## Materials and methods

### Proteins

MutS mutants were created in the *mutS* gene in vector pET-3D (4,22-24) or vector pET15b (23,24) (for His-tagged MutS constructs in FRET assays). MutL mutants were generated in the *mutL* gene in plasmid pTX418 (15,25). Single-cysteine MutS and MutL constructs were obtained as described (22,26). Mutant and WT MutS and MutL proteins were purified as described previously (4,23), except that in the buffers KCl was used instead of NaCl (final gel filtration buffer for MutS: 25 mM Hepes pH 7.5, 150 mM KCl, 1 mM dithiothreitol (DTT); for MutL: 20 mM Tris pH 8.0, 0.5 M KCl, 10% glycerol, 1 mM DTT).

MutH was purified as follows: *E. coli* BL21(DE3) cells were transformed with MutH expression plasmid pTX417 (25) and plated onto LB agar with 50 μg/ml carbenicillin. A colony was picked and cells were grown in LB with 50 μg/ml carbenicillin at 37°C to OD600 ~0.6 and induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside for 4 hours. Cells were harvested and resuspended in binding buffer (25 mM Tris pH 8.0, 300 mM KCl, 10 mM imidazole, 0.2 mM DTT) with 1 mM PMSF and protease inhibitors (Roche Diagnostics) and lysed by sonication. The cleared supernatant was incubated with Talon resin (Clonetech Laboratories) for 30 minutes on ice. Beads were washed using binding

buffer with 1 M KCl, and MutH was eluted with 250 mM imidazole in binding buffer. The His-tag was removed by cleavage with Thrombin protease (~5 units thrombin/mg MutH; GE Healthcare) while dialyzing against 20 mM Tris pH 8.0, 100 mM KCl, 0.2 mM DTT for 2 hours at 22 °C followed by overnight incubation at 4 °C. The mixture was brought to 20 mM imidazole, incubated with Talon beads to remove uncleaved protein, and loaded onto a heparin column equilibrated in buffer A (25 mM Tris pH 8.0, 0.1 M KCl, 1 mM DTT). MutH was eluted using a gradient of 0.1-1.0 M KCl in buffer A, pooled and diluted 2-fold with buffer A and loaded onto a MonoQ column equilibrated with buffer A. MutH was eluted using the same gradient, pooled and dialyzed overnight against 25 mM MES pH 5.5, 150 mM KCl, 1 mM DTT. MutH was loaded onto a MonoS column equilibrated with 25 mM MES pH 5.5, 0.1 M KCl, 1 mM DTT and eluted using a 0.1-1.0 M KCl gradient. Peak fractions were pooled, concentrated using Centriprep 10 and loaded onto a Superdex 75 column equilibrated with 25 mM Tris pH 8.0, 250 mM KCl, 1 mM DTT. Peak fractions were pooled, concentrated, flash frozen in 25 mM Tris pH 8.0, 250 mM KCl, 1 mM DTT, 50% glycerol and stored at -80 °C.



**Small-scale protein crosslinking**

Single cysteine MutS$^{\Delta C800}$ and His-tagged MutL$^{LN40}$ proteins were reduced with 10 mM DTT for 20 minutes and O/N dialyzed into buffer B (25 mM Hepes pH 7.5, 400 mM KCl, 5 mM $MgCl_2$, 10% glycerol) at 4 °C, to remove DTT. MutS$^{\Delta C800}$ (0.57 mM) was incubated 100-bp DNA containing a GT mismatch (AAACAGGCTTAGGC TGGAGCTGAAGCTTAGCTTAGGATCATCGAGGATCGAGCTCGGTGCAAT TCAGCGGTACCCAATTCGCCCTATAGGCATCCAGGTT annealed with AACCTGGA TGCCTATAGGGCGAATTGGGTACCGCTGAATTGCACCGAGCTTGATCCT CGATGATCCTAAGCTAAGCTTCAGCTCCAGCCTAAGCCTGTTT, 0.29 mM) for 25 minutes on ice in buffer C (25 mM Hepes pH 7.5, 125 mM KCl, 5 mM $MgCl_2$). MutL$^{LN40}$ (4 mM) was incubated with 5 mM ATP for 25 minutes on ice. MutS$^{\Delta C800}$/DNA and MutL$^{LN40}$/ATP samples were then combined to final protein concentrations 0.4 mM (DNA concentration 0.2 mM) and additional ATP was added to a final concentration of 1 mM. Samples were then incubated for 10 minutes at RT, after which they were adjusted to 37 °C for 2 minutes. Crosslinker (BMOE or $BM(PEO)_3$, Pierce, dissolved to 0.5 mM in DMSO) was added to a final concentration of 50 mM and samples were incubated for exactly 2 minutes at 37 °C. Reactions were stopped by adding protein loading buffer with DTT and crosslinking was assessed on SDS-PAGE gels stained with Coomassie.

**MutS$^{\Delta C800}$/MutL$^{LN40}$ complex purification**

To obtain crystallizable amounts of crosslinked MutS$^{\Delta C800}$/MutL$^{LN40}$ complex, equimolar amounts of MutS$^{\Delta C800}$ D246C and His-tagged MutL$^{LN40}$ N131C (or with additional

arginine mutations) were reduced and dialyzed separately, as described above. $MutL^{LN40}$ was diluted to 2 mM in buffer D (25 mM Hepes pH 7.5, 400 mM KCl, 10% glycerol) and incubated with a 5-fold molar excess of $BM(PEO)_3$ (from 50 mM solution in DMSO) for 10 minutes at 4 °C. The low $MutL^{LN40}$ concentration prevented the formation of $MutL^{LN40}$-$MutL^{LN40}$ crosslinks, while the excess crosslinker ensured each $MutL^{LN40}$ to react with one maleimid group so that the other reactive side of the crosslinker remained available. The $MutL^{LN40}$ was then bound to Talon beads and the beads were subsequently washed with 20 column volumes of buffer D and 20 column volumes of buffer E (25 mM Hepes pH 7.5, 150 mM KCl, 10% glycerol, 5 mM imidazole) to remove excess crosslinker. $MutS^{\Delta C800}$ was incubated for 10 minutes with equimolar amounts of 30-bp DNA with a GT mismatch at position 9 (AGCTGCCAGGCACCAGTGTCAGCGT CCTAT annealed with ATAGGACGCTGACACTGGTGCTTGGCAGCT) in buffer C. The DNA-bound $MutS^{\Delta C800}$ was then added to the Talon-bound $MutL^{LN40}$, and 30-fold excess ATP was immediately added after which everything was incubated to crosslink for 1 h at 4 °C. The beads were then washed with 10 column volumes buffer E to remove $MutS^{\Delta C800}$-$MutS^{\Delta C800}$ crosslinks, after which the protein was eluted in buffer E with 300 mM imidazole and DTT was added to quench excess crosslinker. The protein was bound to a 5 mL heparin column (GE) and eluted with a 0.1-1 M KCl gradient, which removed the DNA from the protein. The elution was subsequently concentrated and purified with size-exclusion chromatography in buffer B containing 1 mM DTT, for which two S200 16/60 columns were coupled resulting in one long column. The $MutS^{\Delta C800}$/$MutL^{LN40}$ protein peak was then concentrated, after which the $MutS^{\Delta C800}$ concentration was estimated using ε = 95,238 and the whole process (including DTT incubation and dialysis) was repeated to obtain $S_2L_2$ complexes. The resulting protein (5-10% final yield) was concentrated to 80-90 mM (expressed in MutS monomer concentrations; ε = 94,660) and flash-frozen until further use.

**Crystallization and structure solution**

For crystallization, 50 mM $MutS^{\Delta C800}$/$MutL^{LN40}$ complex was incubated with 25 mM DNA containing a GT mismatch (27-bp: TGCCAGGCACCAGTGTCAGCGTCCTAT annealed with ATAGGACGCTGACACTGGTGCTTGGCA or 100-bp, same sequence as above) for 25 minutes on ice. AMP-PNP was subsequently added to a concentration of 1 mM and the protein was crystallized at 4 °C using vapor diffusion in 9-12% PEG-8000, 100 mM Tris pH 7.0, 200 mM $MgCl_2$, and 80-450 mM sodium malonate. Microseeding was used to increase crystal nucleation. Crystals were cryoprotected in mother liquor supplemented with 25% ethylene glycol and 100 mM KCl before flash-cooling in liquid nitrogen. Diffraction data were collected at 100 K at beamline ID-29 at the ESRF or beamline PX-III at the SLS.

Crystallographic data were processed with XDS (27) or iMOSFLM and scaled in Aimless

from the CCP4 suite (28). Crystal structures were solved in consecutive steps of finding domains using Phaser (28). Several search models were used, but best results were obtained with domains from chain A of PDB entry 1W7A as search models for $MutS^{\Delta C800}$ and chain A from PDB entry 1BKN for $MutL^{LN40}$, while clear density for residues 150-164 of $MutL^{LN40}$ allowed building as in PDB entry 1NHJ. The search process was improved by going back and forth between the different datasets to find missing domains. In one crystal form a second conformer has a more compact conformation, probably due to crystal packing, but as it has identical interfaces with $MutL^{LN40}$ we focussed on the main conformation throughout this paper. For refinement, rigid body refinement was performed for the separate domains, followed by refinement with external restraints generated in ProSMART (29) to chain A of the PDB_REDO (30) optimized structure of entry 1W7A and chain A of the PDB_REDO optimized structure of entry 1BKN. Finally, jelly-body refinement was performed. All refinements were performed using REFMAC5 (31,32). Refinement and data collection statistics can be found in Supplementary Table S2. Figures were generated with MacPyMOL (*http://www.pymol.org*), interpolations between conformations were created with LSQMAN (33) and electrostatic surface with CCP4mg (28). Protein interface areas were calculated using PISA (28), for which the missing loop of residues 126-131 of $MutL^{LN40}$ in interface 2 was modeled as in PDB entry 1NHJ.



**MutS conformational changes**

To look at changes within MutS dimers, we used MutS D835R dimer (23,26) variants that do not form tetramers, with single cysteines in positions R449C (His-tagged), D246C, S798C, or A336C. The proteins were labeled with Alexa Fluor 488 or Alexa Fluor 594 maleimide (Invitrogen) according to the manufacturers instruction. Excessive dye was removed using Zeba Spin Desalting columns (Thermo Scientific) and the degree of labeling determined from the absorbance spectra recorded from 220-700 nm (nanodrop) according to the manufactures instructions.

Clamp-domain crossover movement and connector domain movement within MutS dimers were measured using Förster resonance energy transfer (FRET) in which fluorescence emission spectra were recorded with excitation at either 485 nm (5 nm slit width) for FRET or 590 nm (5 nm slit width) for direct acceptor measurements. FRET was determined by the ratio between signal at 485 and 620 nm while direct acceptor was determined by the ratio between signal at 590 and 620 nm and after correction for spectral crosstalk the ratio FRET/acceptor was calculated, and normalized for unbound protein. Heterodimers of single-cysteine MutS variants labeled with Alexa Fluor 488 and Alexa Fluor 594, respectively, were allowed to form by mixing 200 nM of each protein and incubation at 22 °C for at least 30 min in the absence of ADP in buffer F (25 mM Hepes pH 7.2, 150 mM KCl and 5 mM $MgCl_2$) supplemented with 0.05% TWEEN-20. Next, 200 nM

of 59-bp DNA with a GT mismatch (TGAAGCTTAGCTTAGGATCATCGAGGATCGA GCTCGGTGCAATTCAGCGGTACCCAATT annealed with AATTGGGTACCGCT GAATTGCACCGAGCTTGATCCTCGATGATCCTAAGCTAAGCTTCA, with blocked ends as described above) was added, followed by addition of 1 mM ATP.

MutS-DNA FRET was measured in a Hitachi Fluorescence spectrofluorimeter F-4500 (Program FL Solutions 2.0). Fluorescence emission spectra (600 to 700 nm) were recorded with excitation at either 435 nm (5 nm slit width) for FRET or 590 nm (5 nm slit width) for direct acceptor measurements. FRET was determined by the ratio between signal at 435 and 617 nm while direct acceptor was determined by the ratio between signal at 590 and 617 nm and after correction for spectral crosstalk the ratio FRET/acceptor was calculated. We used 30-bp DNA with or without a GT mismatch (AATTGCACCGAGCTTGATCCTCGATGATCC annealed with complementary strand or GGATCATCGAGGATCGAGCTCGGTGCAATT), where the T-containing strand had 5' and 3' digoxigenin labels so that both DNA ends were blocked with anti-digoxigenin Fab fragments (Roche). 100 nM of the DNA with 6 μM SYTOX Blue (Invitrogen) was mixed with 200 nM MutS variants labeled with Alexa Fluor 594 in buffer F, after which ATP was added to 1 mM to induce the conformational change in MutS.



### *In vivo* MMR complementation

Spontaneous mutation rates were assessed using acquired rifampicin resistance. Strains KR1517 (*mutS*, as in (34)) or GM4250 (gift from M. Marinus, (35)) (*mutL*) were transformed with empty vector or plasmid containing WT or mutant MutS or His-MutL genes, and plated on LB/agar plates with 50 mg/mL carbenicillin and 30 mg/mL kanamycin. After overnight incubation at 37 °C, single colonies were picked and grown in 10 mL LB with antibiotics to $OD_{600}$ ~1.0. Next, 0.33 * $10^8$ or 1 * $10^8$ cells were plated on LB plates with antibiotics and 0.1 mg/mL rifampicin. Plates were O/N incubated at 37 °C and rifampicin resistant colonies were counted. Mutation rates and 95% confidence intervals were determined using Fluctuation AnaLysis CalculatOR with the MSS maximum-likelihood method (*http://www.mitochondria.org/protocols/FALCOR.html*).

### DNA binding kinetics

SPR experiments for binding $MutS^{\Delta C800}$ D246C or crosslinked $MutS^{\Delta C800}$/$MutL^{LN40}$ complex to DNA were performed in a Biacore T200 system (GE) as described (26). The experiments were performed in SPR buffer containing 25 mM Hepes pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 0.05% TWEEN-20 and 1 mM ATP, using immobilized biotinylated 100-bp DNA (same sequence as above) with a fluorescein moiety at the other end.

**MutL-MutS binding assay**

Full-length $His_6$-MutL binding to the full-length MutS sliding clamp was assessed using a two-step SPR assay. The resulting graphs are not strictly affinity curves, as changes in MutS stability on DNA contribute to the observed response, but serve to assess the effect of mutations. The SPR buffer was supplemented with 20% glycerol to ensure MutL stability. Before every measurement, anti-fluorescein Fab fragment (Invitrogen) was injected to block the fluorescein-labeled DNA (100-bp, see above) ends. MutS sliding clamps were captured on the end-blocked DNA by injecting 200 nM WT or mutant MutS protein (in buffer with 1mM ATP) for 120 s. Then WT or mutant MutL protein (in buffer with 1 mM ATP) was injected for 120 s, followed by dissociation with buffer only. This was repeated with varying concentrations of MutL.

**DNA binding by MutL**

Fluorescence polarization measurements to assess DNA-binding of $MutL^{LN40}$ mutants were performed in low-salt FP buffer with 25 mM Hepes pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTT and 0.05% TWEEN-20. For full length MutL, the buffer was supplemented with 10% glycerol. A concentration of 0.5 nM of 5' labeled TAMRA-41-bp DNA (ATAGGACGCTGACACTGGTGCTTGGCAGCTTCTAATTCGAT annealed with complementary strand) was used. MutL proteins were serial diluted in black 96-well microplates (PerkinElmer) in 100 mL volumes. Polarization of the TAMRA label was read out in a PHERAstar FS machine (BMG Labtech) with an 540/590 (excitation/emission) FP module.

**DNA kinking assays**

Stopped-flow assays to assess DNA binding and kinking were performed in buffer containing 25 mM Hepes pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1mM DTT, 0.05% TWEEN-20 and 10 mM ADP, with or without 1 mM ATP. One syringe contained 50 nM of 45-bp DNA with or without a GT mismatch (GTCATCCTCG[T*]C TCAAGCTGCCAGGCACCAGTGTCAGCGTCCTAT annealed with complementary strand or ATAGGACGC[T*]GACACTGGTGCTTGGCAGCT TGAGACGAGGATGAC) which was either labeled with Alexa Fluor 594 at position 11 in the top strand and Alexa Fluor 488 at position 10 in the bottom strand (indicated by T*), or with 5'-labeled with TAMRA in the top strand. The other syringe contained 200 nM $MutS^{\Delta C800}$ D246C or crosslinked $MutS^{\Delta C800}$/$MutL^{LN40}$ complex. For assays with double-labeled DNA, donor fluorophores were excited at 473 nm and measured using filters 540IB+540IK, while acceptor fluorophores were measured at the same time using an OG590 filter. For experiments with TAMRA-labeled DNA, the fluorophore was excited at 545 nm and OG540 filters were used for read-out. Samples were co-injected in a KinetAsyst SF-61DX2 stopped-flow machine

(TgK Scientific) fitted with R10699 photomultiplier tubes (Hamamatsu) at 15 °C and measured for 100 s, which was repeated 5-10 times and averages were calculated.

**MutH activation assay**

Circular DNA containing a single GT mismatch and 12 hemi-methylated GATC sites was prepared via primer extension on single stranded DNA from a derivative of pGEM-13Zf (gift from J. Jiricny) as described (36) with the exception that closed circular DNA was purified from gel using a Wizard gel purification kit (Promega). To enable quantification, an Alexa Fluor 647 labeled oligo (IBA GmbH, Göttingen, Germany) was used: CCAGACGTCTGTCGACGTTGGGAAGCT[T*]GAGTATTCTATAGTGTCACCT, where the G is nucleotide forming a GT mismatch and the T* is the Alexa Fluor 647 labeled nucleotide. Nicking reactions contained 25 mM Hepes KOH pH 7.5, 150 mM KCl, 0.1 mg/ml BSA, 5 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 0.5 nM circular DNA, 200 nM MutS, 200 nM WT MutL, single-cysteine MutL N131C, MutL N131C R266E or MutL N131C R162E/R266E/R316E and 100 nM MutH as well as two-fold dilutions thereof. Control reactions contained either no protein or 200 nM MutS and 100 nM MutH. Reactions were incubated for 5 minutes at 37°C and stopped with an equal volume of 20% glycerol, 1% SDS and 50 mM EDTA. Samples were analyzed on 0.8% agarose gels supplemented with 1 μg/ml ethidium bromide, run in 1x TAE. Conversion of covalently closed circles into nicked product was visualized using the fluorescence of the Alexa Fluor 647 label using a Typhoon Trio Imager (GE Healthcare) with excitation at 633 nm and emission filter 670BP30.

**ATPase assay**

ATPase activity of WT MutS and MutS P595A/I597A/M759D was measured by coupling ATP hydrolysis to oxidation of NADH as in (34). MutS protein (5 mM) was mixed with 3.125-500 mM ATP and hydrolysis was measured in a spectrophotometer during 5 minutes.

**$MutL^{LN40}$ dimerization assay**

Crosslinked $MutS^{\Delta C800}$/$MutL^{LN40}$ complex (1 mg/mL) was incubated for 5 minutes on ice with equimolar amounts of 100-bp DNA containing a GT mismatch (sequence as in main text). $MutL^{LN40}$ (2 mg/mL) was incubated with the $MutS^{\Delta C800}$/$MutL^{LN40}$/DNA complex or with DNA only, and 1 mM AMP-PNP as described (15). Samples were injected onto a S200 5/150 column in buffer containing 20 mM Tris pH 8.0, 150 mM KCl, 0.1 mM EDTA, 5 mM $MgCl_2$, 1 mM DTT and 5% glycerol. Eluted fractions were analyzed on SDS-PAGE stained with Coomassie.

## Results

### Structure of the MutS/MutL complex

To trap the *E. coli* MutS/MutL complex we used site-specific chemical crosslinking of single-cysteine variants of MutS and MutL. First all cysteines in MutS and MutL were replaced and functionality of the resulting protein was confirmed (22-24). Then single cysteines were introduced to find positions where crosslinking was dependent on sliding clamp formation. MutS D246C crosslinks specifically to MutL N131C only when a DNA mismatch and a nucleotide are present ((24), Figure 1A, Supplementary Figure S1A), indicating that a complex relevant for MMR is trapped.

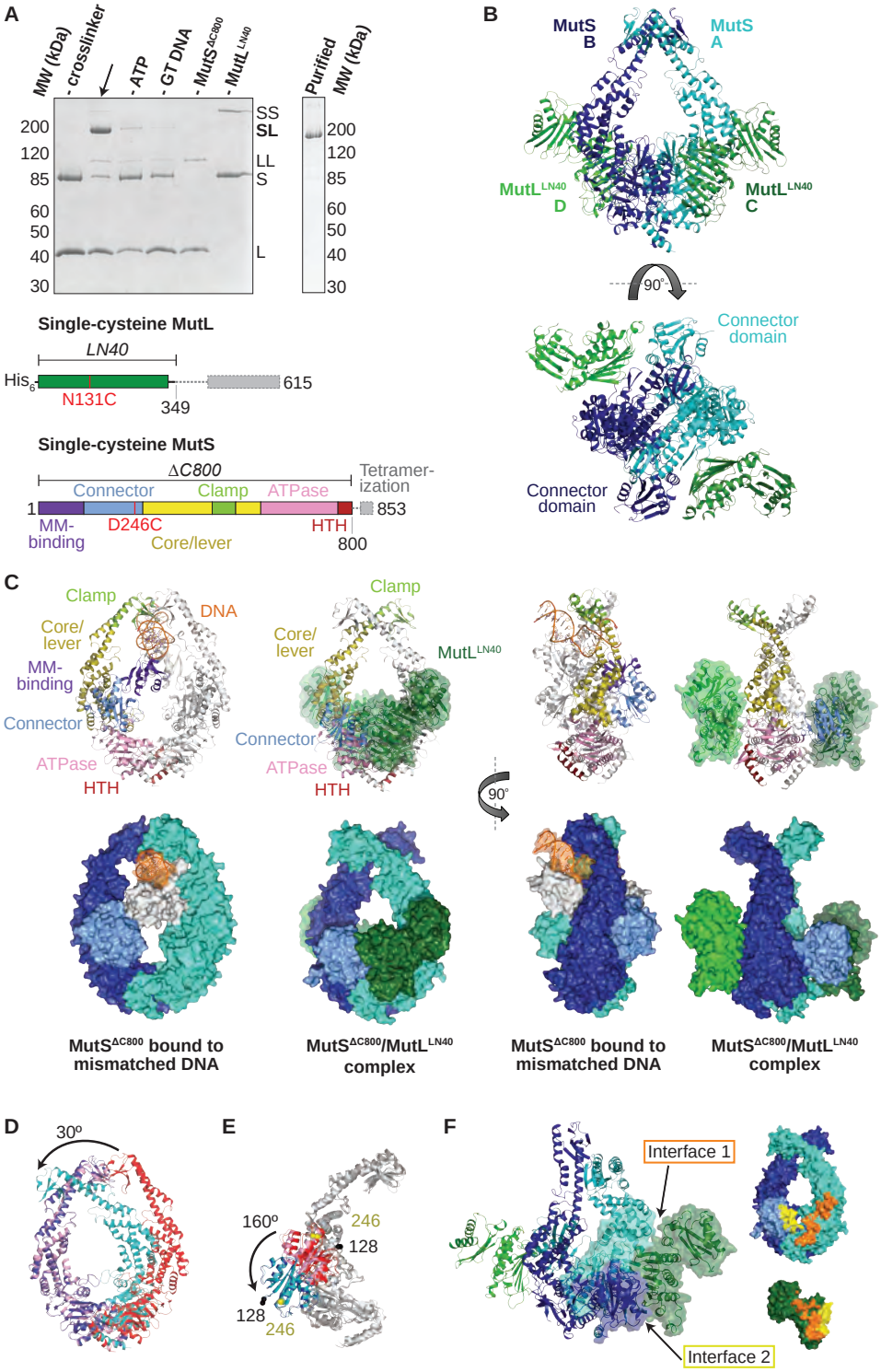


For structural studies, we scaled up the reaction and removed C-terminal domains from MutS and MutL (Figure 1A), to capture the complex between MutS$^{\Delta C800}$ D246C (which we will refer to as MutS$^{\Delta C800}$) and the 40 kDa N-terminal LN40 domain (15) of MutL N131C (which we will refer to as MutL$^{LN40}$). The proteins were crosslinked in the presence of mismatched DNA and ATP, followed by purification to obtain the protein in two successive cycles. This generated a complex where each MutS$^{\Delta C800}$ subunit in the dimer binds to a MutL$^{LN40}$ monomer (Figure 1A, Supplementary Figure S1B,C), which was sufficiently homogeneous and stable to allow crystallization.

We crystallized the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex in the presence of DNA containing a GT mismatch and the non-hydrolyzable ATP analog AMP-PNP (adenylyl-imidodiphosphate). The complex crystallized in several different space groups, diffracting to resolutions from 7.6 to 4.7 Å. In all crystal forms, we could elucidate the same structure of the protein complex (Figure 1B), using parts of higher-resolution MutS$^{\Delta C800}$ and MutL$^{LN40}$ structures for molecular replacement.

The crystal structure shows a novel conformation of MutS, in which the subunits in the dimer are tilted across each other by ~30°, compared to the mismatch recognition state (Figure 1C,D). The subunits are tilted as a rigid body, but the C-terminal HTH domains hinging around residues 765-766, move with the opposite subunit, maintaining their role

**Figure 1.** Crystal structure of the crosslinked MutS$^{\Delta C800}$/MutL$^{LN40}$ complex. **A**) DNA and ATP-dependent crosslinking of MutS$^{\Delta C800}$ D246C (S) and MutL$^{LN40}$ N131C (L) and large-scale purification. Constructs and domain definitions are shown. **B**) Crystal structure of the trapped transient complex of MutS$^{\Delta C800}$ dimer (blue/cyan) with MutL$^{LN40}$ (green). **C**) Comparison between MutS$^{\Delta C800}$ in mismatch-recognition state (PDB entry 1E3M) and the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex, with MutS subunit B colored as in A). **D**) The dimer subunits (blue/cyan) tilt across each other (connector and mismatch-binding domains not shown for clarity) compared to the mismatch-bound state (red/pink). **E**) The connector domain (blue/cyan) rotates around residues 265-266 compared to the mismatch-bound state (red/pink) relative to other domains. Reorientation of residues 128 and 246 indicated. **F**) Each MutL$^{LN40}$ subunit (green) interacts via two interfaces (orange/yellow) with the MutS$^{\Delta C800}$ dimer (blue/cyan). ▶


A
MW (kDa)
- crosslinker
- ATP
- GT DNA
- MutS^ΔC800
- MutL^LN40
Purified
MW (kDa)
200
120
85
60
50
40
30
SS
SL
LL
S
L
Single-cysteine MutL
LN40
His6
N131C
349
615
Single-cysteine MutS
ΔC800
Tetramer-ization
Connector
Clamp
ATPase
1
853
MM-binding
D246C
Core/lever
HTH
800
B
MutS B
MutS A
MutL^LN40 D
MutL^LN40 C
90°
Connector domain
Connector domain
C
Clamp
DNA
Core/ lever
MM-binding
Connector
ATPase
HTH
Clamp
Core/ lever
MutL^LN40
Connector
ATPase
HTH
90°
MutS^ΔC800 bound to mismatched DNA
MutS^ΔC800/MutL^LN40 complex
MutS^ΔC800 bound to mismatched DNA
MutS^ΔC800/MutL^LN40 complex
D
30°
E
160°
246
128
128
246
F
Interface 1
Interface 2

in stabilizing MutS dimers (37). Meanwhile, the connector domains are rotated by ~160° around residues 265-266, which moves these domains out of the center of the molecule and packs them against the ATPase domains (Figure 1C,E). The mismatch-binding domain could not be resolved in the density, probably because it is flexible in this state. While the mismatch recognition state of MutS is asymmetric (4), this MutL$^{LN40}$-bound conformation shows a more symmetrical MutS$^{\Delta C800}$ dimer.

The MutL$^{LN40}$ interaction with MutS$^{\Delta C800}$ involves two interfaces (Figure 1F). The first interface is formed by the largest β-sheet of the ATPase domain of MutL$^{LN40}$, and the ATPase and core domains of one subunit of MutS$^{\Delta C800}$. The second interface involves the side of this same β-sheet and a looped-out helix of MutL$^{LN40}$, and the newly positioned connector domain of the other MutS$^{\Delta C800}$ subunit. Each MutL$^{LN40}$ monomer is therefore interacting with both subunits in the MutS$^{\Delta C800}$ dimer.



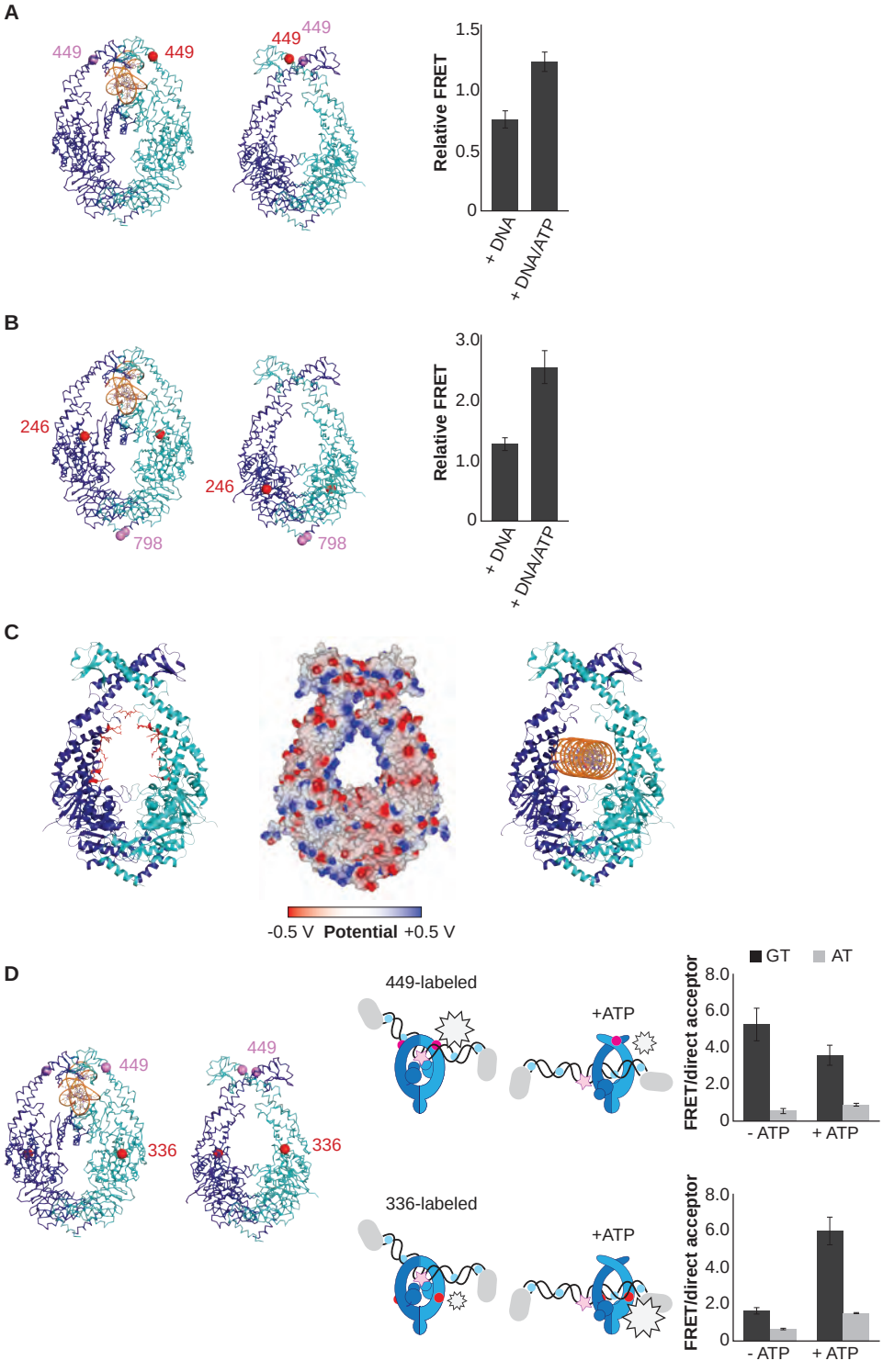
**Conformation of the MutS sliding clamp**

The novel conformation of MutS in our crystal structure reveals a rearrangement of the subunits in the MutS$^{\Delta C800}$ dimer, tilting around the interface formed by the two ATPase sites (Figure 1D). The tilting creates a new MutS dimer interface of ~500 Å$^2$ where the clamp domains cross over, partially from interactions between the helices themselves (200 Å$^2$), the rest from the ends of the clamp domains with the helices.

We observe nucleotide density in the ATP binding sites of both subunits in the MutS$^{\Delta C800}$ dimer (Supplementary Figure S2A), and since we crystallized the protein with AMP-PNP we modeled these nucleotides in the density. This type of ATP-induced tilting and increased packing of ATPase domains is more often observed upon ATP binding in ABC ATPases, such as ATP transporters, SMCs and RAD50 (38), but although we previously predicted a tilting motion (34), the extent of the motion and the crossing of the DNA clamp domains of MutS was unexpected.

To validate the relevance of the observed conformational changes for the MMR process, MutS proteins with a single cysteine at position 449 were site-specifically labeled with two different Alexa fluorophores and combined into heterodimers by random subunit exchange (Figure 2A). When labeled protein was bound to end-blocked DNA containing

**Figure 2.** The structure of the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex reveals the MutS sliding clamp conformation. **A**) FRET within MutS dimers (normalized for unbound protein) reveals residues 449 coming closer together upon ATP addition. Bars depict mean +/- SD, n=3. **B**) FRET assay agrees with residue 246 on the connector domain of MutS moving towards residue 798 upon ATP addition after mismatch recognition. **C**) Mismatch and ATP-induced conformational changes open a channel lined by positively charged residues (left: arginines and lysines as red sticks, middle: electrostatic surface), which would fit a DNA helix (right). **D**) FRET assay agrees with movement of DNA away from residues 449 in MutS, while approaching residues 336 upon ATP addition. ▶


A
449
449
449
449
Relative FRET
1.5
1.0
0.5
0
+ DNA
+ DNA/ATP
B
246
798
246
798
Relative FRET
3.0
2.0
1.0
0
+ DNA
+ DNA/ATP
C
-0.5 V Potential +0.5 V
D
449
336
449
336
449-labeled
+ATP
336-labeled
+ATP
GT
AT
FRET/direct acceptor
8.0
6.0
4.0
2.0
0
- ATP
+ ATP
FRET/direct acceptor
8.0
6.0
4.0
2.0
0
- ATP
+ ATP

a GT mismatch, FRET increased upon ATP addition. This indicates that ATP-induced sliding clamp formation moves these residues toward each other, in line with the shorter distance in the new conformation (from 50 Å in the mismatch-recognition state to 43 Å in the MutL$^{LN40}$-bound structure).

The new position of the connector domain brings it closer to the ATPase domain (Figure 1E). To analyze this movement we combined two single-cysteine variants of MutS, labeled in the connector domain (residue 246) and the ATPase domain (residue 798) respectively, into heterodimers, and measured the FRET signal between these sites upon sliding clamp formation (Figure 2B). Indeed, after ATP addition the FRET increased, indicating that these residues come closer together. As this is measured in the absence of MutL it suggests that after mismatch binding, ATP is sufficient to induce movement of the connector domain away from the mismatch-recognition position.



Although the complex was crystallized in the presence of DNA containing a mismatch, the DNA is not visible in the structure. This could be due to smearing out of the electron density over multiple positions or the DNA may not be present in the crystal, both indicating that the mismatch has been released, as expected for the ATP-bound state of MutS.

The subunit tilting has occluded the original DNA binding site, but because the connector and mismatch-binding domains have moved, a large channel (~35 Å wide) in MutS has become accessible, which could easily accommodate a DNA duplex (20 Å diameter). The new channel is lined by conserved lysines and arginines (Figure 2C, Supplementary Figure S2C), which can govern nonspecific contacts with the negative backbone of DNA, as expected for the MutS sliding clamp state (19). Moreover, packing of molecules in all crystal forms results in alignment of these channels between symmetry mates or even within the asymmetric unit (Supplementary Figure S2B). We therefore hypothesize that the DNA is pushed down to this channel during the ATP-induced conformational changes of MutS after mismatch recognition.

To test whether DNA moves down into the new channel, we analyzed FRET signals between fluorescently labeled DNA (end-blocked) and specific sites in single-cysteine MutS variants (Figure 2D). After addition of ATP, DNA moves away from residues 449 at the DNA-clamp position (FRET/acceptor ratio reduction ~1.5 fold), while an increase in FRET/acceptor ratio (>3.6 fold) was observed when MutS was labeled at position 336. Since the connector domain moves down itself, no substantial change in FRET/acceptor ratio is observed between residue 246 and DNA (Supplementary Figure S2D). Combined, these FRET data are in agreement with repositioning of the DNA towards the channel created by the new conformation.

Based on these validations, we conclude that the observed MutS conformation in our crystal structure is induced by ATP after mismatch recognition. Since the new position of

the DNA would allow MutS to fit as a loose ring around the DNA duplex (with a channel size similar to that of PCNA (39)), consistent with free movement over DNA (19), we propose that this is the MutS sliding clamp conformation.

**Orientation of MutL$^{LN40}$ on MutS**

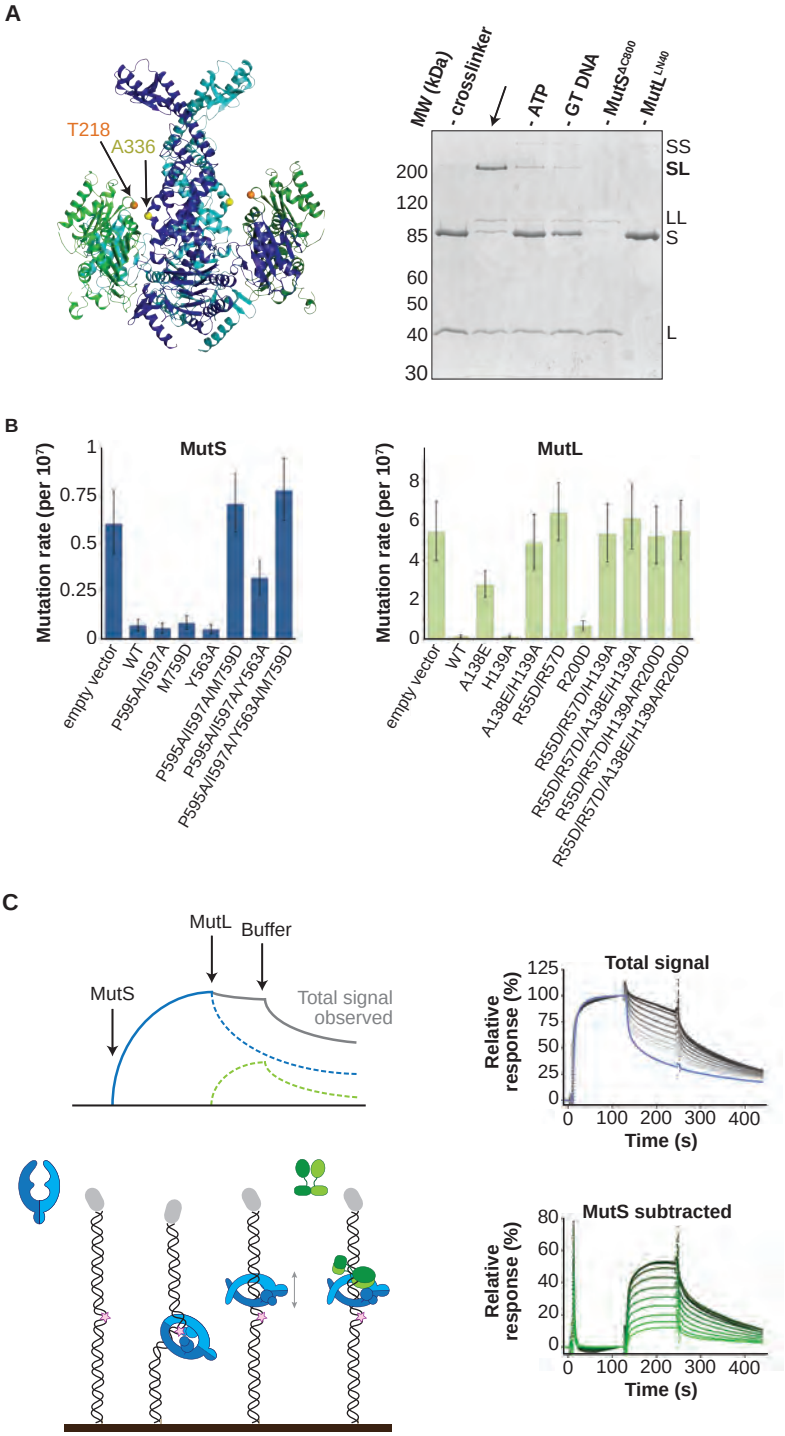
In the structure MutL$^{LN40}$ makes two interfaces with MutS$^{\Delta C800}$. Interface 1 orients MutL$^{LN40}$ on the ATPase and core domains of MutS. Recently, a loop in *Bacillus subtilis* MutS was found to be essential for MutL interaction (40). Although the equivalent loop is shorter in *E. coli* MutS and the explicit residues (F319/F320) are missing, the corresponding region is located within the ~590 Å$^2$ interface (interface 1) with MutL$^{LN40}$.



We validated the observed interaction at interface 1 by a crosslinking experiment with a short crosslinker. We created single-cysteine mutants MutS$^{\Delta C800}$ A336C and MutL$^{LN40}$ T218C (Figure 3A), which are located ~7.4 Å apart in the structure, and then showed that we could crosslink them efficiently with a short cysteine-specific crosslinker (8 Å, BMOE), dependent on the presence of both mismatched DNA and ATP. Only background crosslinking occurred when using MutS$^{\Delta C800}$ D246C (connector domain) with MutL$^{LN40}$ T218C (interface 1) under these conditions (Supplementary Figure S3A), indicating that the crosslinking between MutS$^{\Delta C800}$ A336C and MutL$^{LN40}$ T218C is specific.

To further verify interface 1 between MutS and MutL, we tested whether mutations in the interface affected MMR activity *in vivo*, in a complementation assay with MutS or MutL deficient cells (Figure 3B,F; Supplementary Table S1). We found several mutants of MutL (A138E, A138E/H139A, R55D/R57D, or combinations) and a triple mutant in MutS (P595A/I597A/M759D) that could not complement loss of wild type (WT) protein. We purified the mutants that impaired MMR and characterized their defects. The MutS triple mutant has a slight defect in ATPase activity but this does not impair its sliding clamp formation (Supplementary Figure S3B,C), and other mutants with similar ATPase effects (e.g. MutS F596A) can almost fully reconstitute MMR (41), suggesting that the *in vivo* effect we observe is due to the perturbed interface with MutL.

To assess the effect of these mutations on binding of MutL to the transient MutS sliding clamp we designed a two-stage assay using Surface Plasmon Resonance (SPR). We first formed and trapped MutS sliding clamps on 100-bp end-blocked DNA in the presence of ATP. Next, MutL was injected, which could then bind to these MutS clamps. By subtraction of the MutS signal, the contribution of MutL could be evaluated for the different mutants (Figure 3C), since MutL alone shows no DNA binding under these conditions (Supplementary Figure S3D). Indeed the interface 1 mutants that were deficient for MMR conferred a deficiency in binding (Figure 3D).


A
T218
A336
MW (kDa)
- crosslinker
- ATP
- GT DNA
- MutS^ΔC800
- MutL^LN40
200
120
85
60
50
40
30
SS
SL
LL
S
L
B
MutS
MutL
Mutation rate (per 10^-7)
1
0.75
0.5
0.25
0
8
6
4
2
0
empty vector
WT
P595A/I597A
M759D
Y563A
P595A/I597A/M759D
P595A/I597A/Y563A
P595A/I597A/Y563A/M759D
A138E
H139A
A138E/H139A
R55D/R57D
R200D
R55D/R57D/H139A
R55D/R57D/A138E/H139A
R55D/R57D/H139A/R200D
R55D/R57D/A138E/H139A/R200D
C
MutL
Buffer
MutS
Total signal observed
Total signal
Relative response (%)
125
100
75
50
25
0
0
100
200
300
400
Time (s)
MutS subtracted
80
60
40
20
0

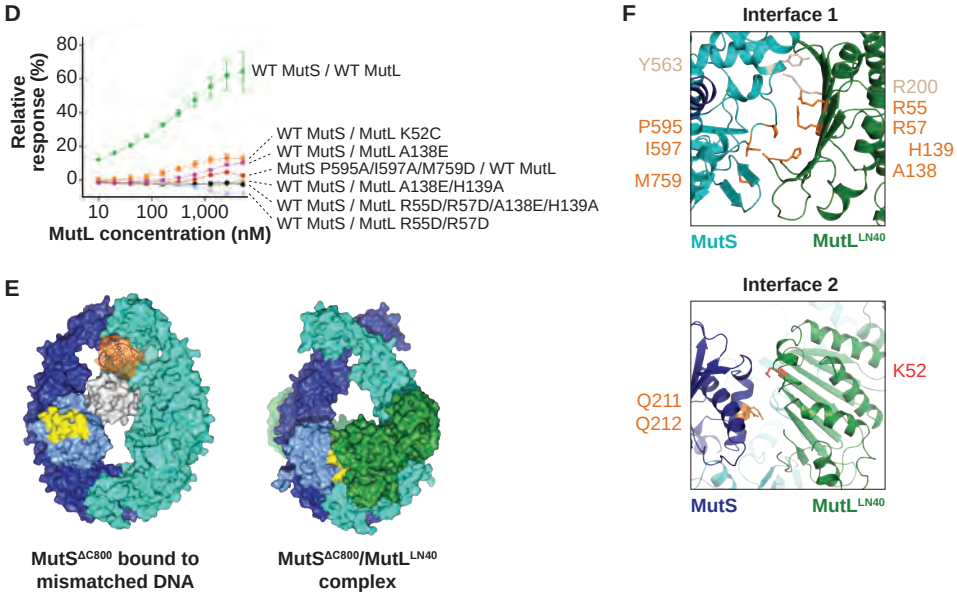



**Figure 3.** Interaction of the MutS$^{\Delta C800}$ sliding clamp with MutL$^{LN40}$. **A**) Crosslinking occurs between MutS$^{\Delta C800}$ A336C and MutL$^{LN40}$ T218C using BMOE (right panel), as suggested by the structure (left panel). **B**) Spontaneous mutation rates after complementing MutS or MutL-deficient cells with the indicated mutants. Error bars represent 95% confidence intervals. **C**) SPR assay to measure MutL binding to pre-formed MutS sliding clamps on end-blocked DNA. MutL contribution (green dotted line) is approached by subtracting MutS-only contribution (blue line) from the total signal (solid line). Data normalized to maximum MutS response. **D**) MutL and MutS mutants with deficiency in MMR show reduced MutS/MutL complex formation in SPR. Error bars represent SD for averages between two experiments. **E**) The yellow patch of MutS$^{\Delta C800}$ interacts with MutL$^{LN40}$ in the new conformation after rearrangement of the connector domain. **F**) Residues in MutS$^{\Delta C800}$/MutL$^{LN40}$ interfaces. Red: full MMR deficiency upon mutation; orange: deficiency upon combination; white: mild effect.

## MutS sliding clamp recognition by MutL

The rearrangement of the connector domain creates a second interface with MutL$^{LN40}$ (interface 2, Figure 3E). Previous deuterium exchange experiments (42) indicated that the connector domain interacts with MutL, particularly via MutS glutamines 211 and 212. Indeed in our structure these residues are buried within this ~670 Å$^2$ interface with MutL$^{LN40}$ (Figure 3F). Interestingly, the deuterium exchange experiments identified a second region on the MutS surface that was protected upon MutL interaction in the ATPase domain (residues 673-686). In the complex structure this region is masked by the new position of the MutS$^{\Delta C800}$ connector domain (Supplementary Figure S3E).

The connector domain is already displaced from the mismatch-recognition position after ATP binding (Figure 2B), and MutL$^{LN40}$ interaction may stabilize the position of the connector domain that we see in the crystal structure. At the resolution of our structure,

there is no clear electron density for the connecting crosslinker that we used to stabilize the complex, and the crosslinked residue 131C on $MutL^{LN40}$ could not be modeled. However, the distance between Cα atoms of crosslinked residue 246C in $MutS^{\Delta C800}$ and residue 132 in $MutL^{LN40}$ is ~15.5 Å, suggesting that the exact positioning of the connector domain is not caused by strain through the crosslinker (which is 18 Å long).

On the MutL side of interface 2, residue K52 of $MutL^{LN40}$ is involved in the interaction with the connector domain of $MutS^{\Delta C800}$ (Figure 3F). This explains the previously reported unexpected mutator phenotype of MutL K52C (22). To confirm its role in the interface we measured the binding of MutL K52C to the MutS sliding clamp in our SPR assay (Figure 3D). Indeed, the binding of this mutant is reduced compared to WT MutL.



The conformational change of MutS and the accompanying connector domain movement positions the two interaction surfaces such that they become simultaneously available for binding to the N-terminal domain of MutL (Figure 1F). Perturbing either interface 1 or interface 2 impairs MutL binding and MMR (Figure 3F). This explains the specificity of MutL for the MutS sliding clamp, which has never been understood before.

MutL proteins dimerize through the C-terminal LC20 domains while the LN40 domains are monomeric in isolation, but can form unstable dimers after ADP or ATP binding or stable dimers when incubated with AMP-PNP (15,43). Our crosslinked protein crystallizes as $MutS^{\Delta C800}$ dimers bound to $MutL^{LN40}$ monomers, and does not show the $MutL^{LN40}$ dimer arrangement through crystal contacts. However, the crystal structure sterically allows MutL dimerization (Supplementary Figure S3F), and in analytical gel filtration, $MutL^{LN40}$ coelutes with the $S_2/L_2$ complex after incubation with DNA and AMP-PNP (Supplementary Figure S3G).

The stoichiometry of the MutS/MutL complex *in vivo* is a topic of interest (44). To obtain crystallizable complexes, $MutL^{LN40}$ was bound to each $MutS^{\Delta C800}$ subunit in our experiments, but in MMR a symmetric complex may not be necessary. Indeed the asymmetry of the eukaryotic MMR proteins suggest that this is not required and that a single MutL homolog dimer will bind to one MSH2/MSH6 or MSH2/MSH3 heterodimer. Literature suggests that interface 2 will be made with MSH2 (42), implying that interface 1 will be with MSH6. The observed $MutL^{LN40}$ protein would then correlate with the MLH1 subunit (45) (Supplementary Figure S3H).

**Binding to MutS positions MutL on DNA**

MutL and homologs have weak DNA-binding ability (41,43), which is only clearly observed in low salt conditions, and retention of MutS on DNA upon MutL interaction has been observed (13,46). Although different from the proposed DNA orientation in the crystal (Supplementary Figure S2B), a model can be constructed in which the DNA running through the channel in the MutS sliding clamp is simultaneously bound by the proposed

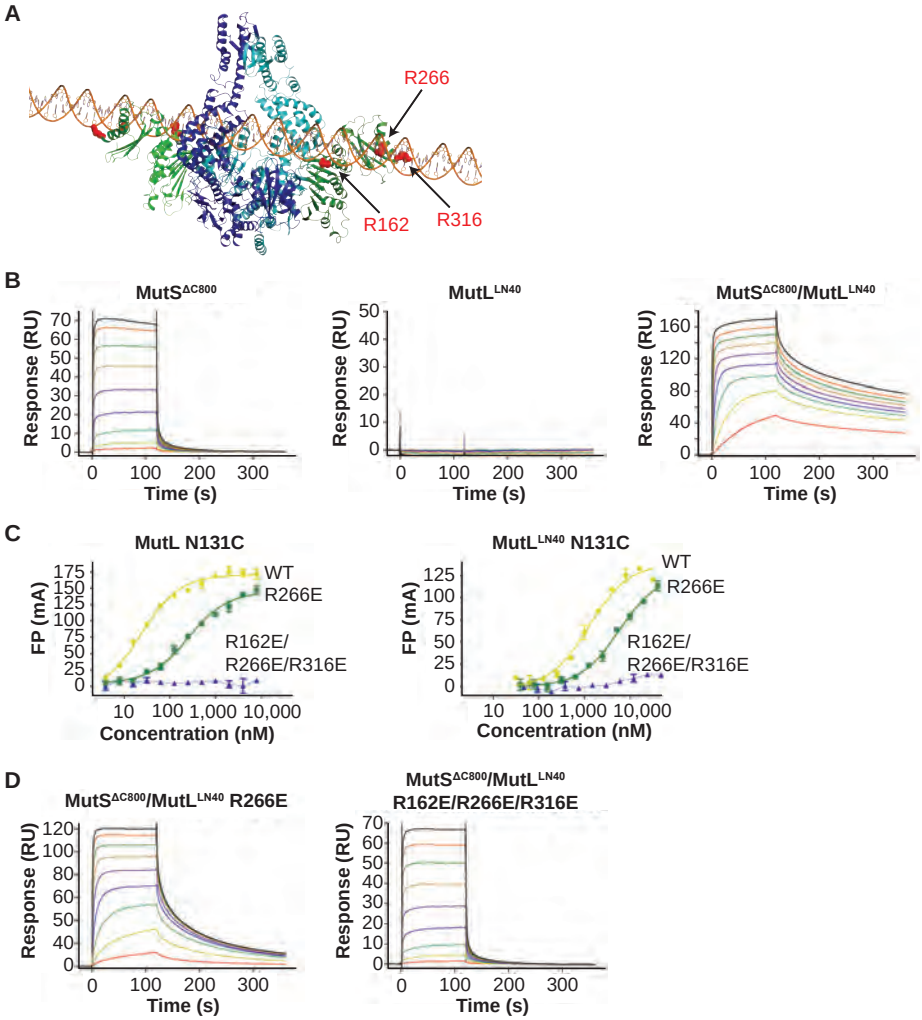
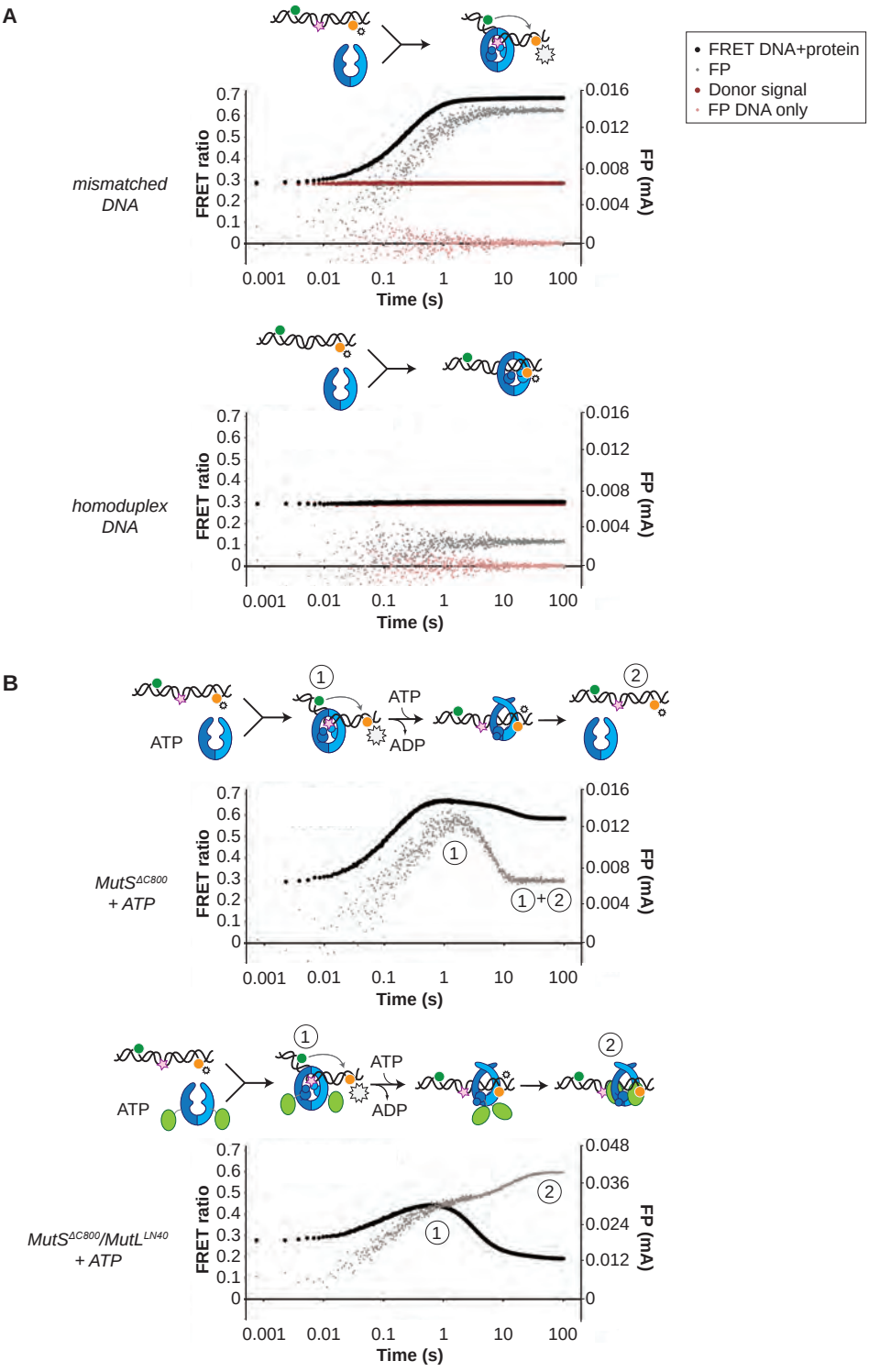



**Figure 4**. The MutS sliding clamp positions MutL onto DNA. **A**) Model of DNA binding by the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex. Three arginines in the MutL$^{LN40}$ DNA-binding groove are shown as red spheres. **B**) In the presence of ATP, MutS$^{\Delta C800}$ has a fast off-rate from 100-bp DNA and MutL$^{LN40}$ alone does not bind DNA under physiological salt conditions (150 mM KCl), while the crosslinked MutS$^{\Delta C800}$/MutL$^{LN40}$ complex releases slowly from DNA. **C**) Mutations in the DNA-binding groove of MutL reduce its DNA-binding ability (observed in low salt, 50 mM KCl) and **D**) affect release rates of the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex in physiological salt conditions.

DNA binding grooves of the MutL$^{LN40}$ subunits (47) (Figure 4A, Supplementary Figure S4A). While such DNA binding may require additional conformational changes of MutL, it suggests a mechanism where MutS loads MutL onto DNA.

We tested for MutL$^{LN40}$ loading onto DNA in the context of the MutS/MutL complex


A
FRET DNA+protein
FP
Donor signal
FP DNA only
mismatched DNA
homoduplex DNA
FRET ratio
FP (mA)
Time (s)
B
ATP
ADP
MutS^ΔC800 + ATP
MutS^ΔC800/MutL^LN40 + ATP

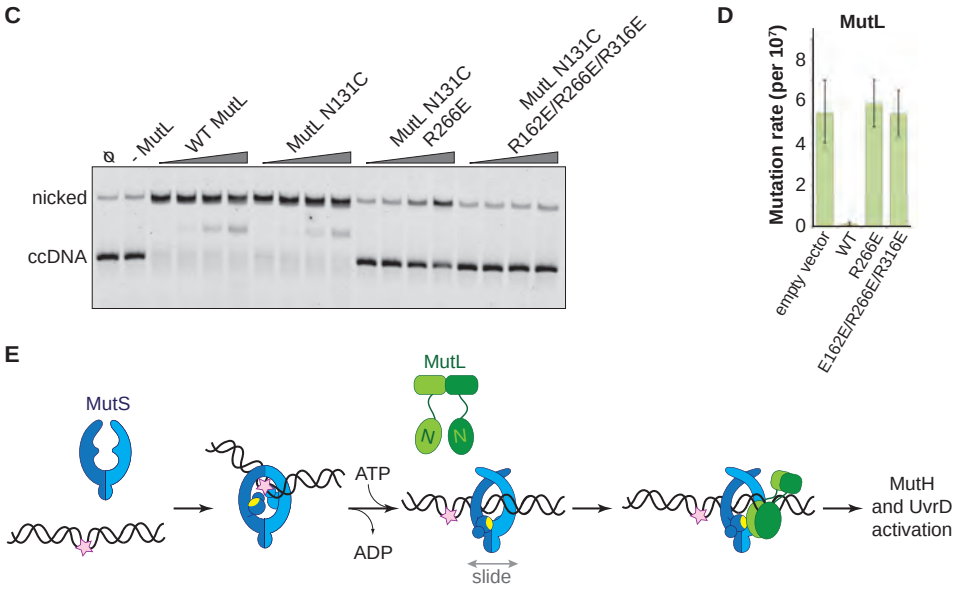



**Figure 5.** Implications for DNA mismatch repair initiation. **A**) Stopped-flow FRET and FP assay shows kinking of 45-bp DNA by MutS$^{\Delta C800}$ binding only if there is a mismatch. Magnitude of FRET events are indicated by stars in the cartoon. **B**) While MutS$^{\Delta C800}$ initially kinks the DNA and subsequently releases in the presence of ATP, the MutS$^{\Delta C800}$/MutL$^{LN40}$ shows a secondary FP event without kinking the DNA. **C**) Nicking assay of mismatch containing closed circular DNA (ccDNA) shows that WT or single-cysteine MutL can activate MutH, while mutations in the DNA-binding groove of MutL strongly impair the activation. **D**) Spontaneous mutation rates after complementing MutL-deficient cells shows that the DNA-binding ability of MutL is essential for MMR *in vivo*. Error bars represent 95% confidence intervals. **E**) Model for MMR initiation. After MutS undergoes an ATP-induced conformational change to allow binding of both subunits to one MutL molecule, MutL N-termini can interact and possibly dimerize, to be loaded onto DNA where MutL can activate downstream effectors.

in an SPR assay, comparing MutS$^{\Delta C800}$ alone with MutS$^{\Delta C800}$ crosslinked to MutL$^{LN40}$ when it is flowed over 100-bp DNA with a GT mismatch in the presence of ATP (Figure 4B). MutS$^{\Delta C800}$ alone displays fast release from the DNA due to ATP-dependent sliding-clamp formation (26), as shown by the effect of blocking the end of the DNA (Supplementary Figure S4E). The presence of crosslinked MutL$^{LN40}$ greatly reduces the rate of release, suggesting additional DNA binding. The magnitude of the signal in response units on a 41-bp oligomer shows that a single MutS$^{\Delta C800}$/MutL$^{LN40}$ complex is sufficient for this effect (Supplementary Figure S4B,C). This delay in release from DNA is also observed when using a mixture of WT MutL and MutS$^{\Delta C800}$, although to a lesser extent (Supplementary Figure S4D). The remaining slow release of the crosslinked complex is not affected by blocking of the free DNA end by antibody (Supplementary Figure S4E) indicating that the constitutive interaction with crosslinked MutL$^{LN40}$ completely stops MutS$^{\Delta C800}$ dissociation from DNA ends.

To validate that the slower release from DNA is indeed due to MutL$^{LN40}$ binding to DNA, we made point mutants of the MutL$^{LN40}$ protein and crosslinked them to MutS. Mutation R266E reduces DNA binding by MutL ((41,48), Figure 4B), most pronounced in full-length context. This mutation also reduces the ability of crosslinked MutL$^{LN40}$ to retain the MutS$^{\Delta C800}$ sliding clamp on DNA (Figure 4d, Supplementary Figure S4B,F). When introducing two additional mutations (R162E and R316E) in the MutL$^{LN40}$ DNA binding site as suggested by the crystal structure (Figure 4A), DNA binding is completely abolished (Figure 4C) and the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex releases as fast as MutS$^{\Delta C800}$ alone (Figure 4D, Supplementary Figure S4B,F). This indicates that MutL binds DNA when interacting with the MutS sliding clamp.



**MutL is loaded onto DNA after MutS releases the mismatch**

To assess whether the loading of MutL$^{LN40}$ onto DNA is kinetically distinct from MutS mismatch recognition, we set up an assay to separate events. We read out mismatch recognition (4-6) by the kinking of DNA, which can be assessed using 45-bp heteroduplex DNA labeled with Alexa fluorophores on each side of the mismatch ((49), Figure 5A, Supplementary Figure S5A), in a stopped-flow set up. In parallel we follow DNA interaction using fluorescence polarization (FP) of TAMRA-labeled DNA with the same sequence. This shows that the kinking is concurrent with DNA binding by MutS$^{\Delta C800}$, while kinking is not observed when homoduplex is used (Figure 5A, Supplementary Figure S5A). When the assay is performed in the presence of ATP, MutS$^{\Delta C800}$ binds and kinks the DNA but subsequently releases due to sliding clamp formation, after which an equilibrium is reached between rebinding and release (Figure 5B, Supplementary Figure S5B).

In the presence of the crosslinked complex we observed a two-step sequence of events (Figure 5B). The first increase in FP is consistent with mismatch recognition by MutS$^{\Delta C800}$, simultaneous with an increase in FRET due to kinking of the DNA. A second event increases FP even more but reduces the FRET signal to below starting value (Supplementary Figure S5B). This can be explained by release of the mismatch (unkinking) and sliding clamp formation. Now, however, the complex does not slide off the DNA but instead the MutL$^{LN40}$ is docked onto the DNA to keep the complex bound, as observed in the SPR assays (Figure 4B) and by the increase in FP (Figure 5B). At this time, since DNA has been pushed to the new channel, the DNA is not kinked any more but kept relatively rigid by the MutL$^{LN40}$ binding. This, and interaction of the fluorophore itself with bound protein, can explain the lowered FRET. The effect is also present to lesser extent when using a mixture of MutS$^{\Delta C800}$ with WT MutL in this setup (Supplementary Figure S5C). The result indicates that MutL$^{LN40}$ loading occurs after mismatch recognition and sliding clamp formation by MutS$^{\Delta C800}$.

**MutL loading on DNA is essential in MMR**

Since we observed that upon sliding clamp formation, MutS loads MutL onto DNA, we wondered whether this DNA loading step is essential for MMR. Indeed we observed a correlation with the DNA binding ability of MutL for MutH activation (Figure 5C). Moreover, the DNA-binding mutants of MutL impair *in vivo* MMR ((48), Figure 5D, Supplementary Table S1), indicating that loading of MutL onto DNA after mismatch recognition is essential for MMR.

Taken together, our data reveal how the large conformational changes within MutS after mismatch recognition promote MMR activation. In the mismatch and ATP activated state (10) MutS pushes DNA into a new channel, which allows sliding of the protein over DNA. The new state with the clamps crossed over the DNA explains the stability of the MutS sliding clamp on DNA (46), as electrostatic interactions with DNA may stabilize the new clamp conformation. The conformational change pushes the connector domain away from the center and on top of the ATPase domains, to provide a second interface for the MutL protein that binds to the opposing MutS subunit, while DNA in the new MutS channel can also contribute to MutL binding. This loads the N-terminal domains of MutL onto the DNA and the MutL binding delays the sliding of MutS (Figure 5E). The loading step of MutL onto DNA is required for MutH activation and nicking, while UvrD loading and activation at this nick (18) would follow similar validation. In this way, the requirement of the MutS conformational change for full MutL interaction is a sophisticated validation mechanism, which presumably is conserved in the eukaryotic homologs. It ensures that repair is only initiated when necessary, and due to the MMR system DNA replication can be completed with few errors incorporated in the genome.

In conclusion, we have used single-cysteine mutants and chemical crosslinking to trap and analyze a relevant MMR intermediate state that has long been elusive. This sliding clamp state of MutS bound to a MutL domain is highly informative. It corresponds to a reaction intermediate that occurs during a series of conformational changes triggered by mismatch recognition, and explains why specifically this conformation of MutS is able to recruit MutL. A stabilized asymmetric ATP state of MutS on the mismatch will precede this symmetric ATP state (10,19,50), and conformational changes of MutL that involve ATP binding (12-14,43) and its C-terminal domains (16) will follow the step that has been visualized here. The presented combination of structural and biophysical methods provides a powerful approach to resolve conformational changes within large and transient protein complexes that form and act during biologically relevant processes.

## Acknowledgements

We are grateful for contributions of group members, to Randy Read for assistance with molecular replacement, and to Lea Geissert, Miguel Keidel, Gaelle Cyriale Ngatcheu Famou, Rosine Djamfa and Matthias Trohart for assistance with FRET assays. We thank Jacques Neefjes, Hein te Riele and Thijn Brummelkamp for critical reading of the manuscript. This research was funded by European Community's Seventh Framework Programme mismatch2model HEALTH-F4-2008-223545, the Centre for Biomedical Genetics, and NWO-CW ECHO 711.011.011 (to T.S.) and VIDI 700.58.428 (to J.L.). The authors declare to have no competing interests.



## Author Contributions

F.G. performed and designed molecular biology, biophysical, structural and *in vivo* assays. A.F contributed and designed SPR and stopped-flow biophysics. H.W. contributed to biochemical and *in vivo* assays. N.H and J.L. contributed the MutH activation assay. I.W. and A.R. contributed initial MutS/MutL crosslinking. R.N. and G.M. contributed to structure refinement. A.M. and M.C. set up FRET experiments with labeled MutS and stopped-flow experiments. P.F. designed all tools and supervised FRET experiments. T.S. supervised the study with contributions from J.L. and P.F. F.G. wrote the manuscript with contributions from T.S, A.F., J.L. and P.F.

## References

1. P.J. Rothwell and G. Waksman (2005) Structure and mechanism of DNA polymerases. *Adv. Protein Chem.*, **71**, 401–440.
2. T.A. Kunkel and D.A. Erie (2005) DNA mismatch repair. *Annu. Rev. Biochem.*, **74**, 681–710.
3. H.T. Lynch and A. de la Chapelle (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.*, **36**, 801–818.
4. M.H. Lamers, A. Perrakis, J.H. Enzlin, H.H.K. Winterwerp, N. de Wind and T.K. Sixma (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature*, **407**, 711–717.
5. G. Obmolova, C. Ban, P. Hsieh and W. Yang (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703–710.
6. J.J. Warren, T.J. Pohlhaus, A. Changela, R.R. Iyer, P.L. Modrich and L.S. Beese (2007) Structure of the human MutSα DNA lesion recognition complex. *Mol. Cell*, **26**, 579–592.
7. D.J. Allen, A. Makhov, M. Grilley, J. Taylor, R. Thresher, P. Modrich and J.D. Griffith (1997) MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.*, **16**, 4467–4476.
8. S. Gradia, D. Subramanian, T. Wilson, S. Acharya, A. Makhov, J. Griffith and R. Fishel (1999)

hMSH2–hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell*, **3**, 255–261.

9. J. Gorman, A. Chowdhury, J.A. Surtees, J. Shimada, D.R. Reichman, E. Alani and E.C. Greene (2007) Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Mol. Cell*, **28**, 359–370.
10. C. Jeong, W. Cho, K. Song, C. Cook, T. Yoon, C. Ban, R. Fishel and J. Lee (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat Struct Mol Biol*, **18**, 379–385.
11. M. Grilley, K.M. Welsh, S. Su and P. Modrich (1989) Isolation and characterization of the *Escherichia coli mutL* gene product. *J. Biol. Chem.*, **264**, 1000–1004.
12. T.A. Prolla, Q. Pang, E. Alani, R.D. Kolodner and R.M. Liskay (1994) MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science*, **265**, 1091–1093.
13. K. Drotschmann, A. Aronshtam, H.J. Fritz and M.G. Marinus (1998) The *Escherichia coli* MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic Acids Res.*, **26**, 948–953.
14. S. Acharya, P.L. Foster, P. Brooks and R. Fishel (2003) The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair. *Mol. Cell*, **12**, 233–246.
15. C. Ban and W. Yang (1998) Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, **95**, 541–552.
16. A. Guarné, S. Ramon-Maiques, E.M. Wolff, R. Ghirlando, X. Hu, J.H. Miller and W. Yang (2004) Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J.*, **23**, 4134–4145.
17. F.A. Kadyrov, L. Dzantiev, N. Constantin and P. Modrich (2006) Endonucleolytic function of MutLα in human mismatch repair. *Cell*, **126**, 297–306.
18. M. Yamaguchi, V. Dao and P. Modrich (1998) MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J. Biol. Chem.*, **273**, 9197–9201.
19. W. Cho, C. Jeong, D. Kim, M. Chang, K. Song, J. Hanne, C. Ban, R. Fishel and J. Lee (2012) ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure*, **20**, 1–11.
20. M.L. Mendillo, D.J. Mazur and R.D. Kolodner (2005) Analysis of the interaction between the S*accharomyces cerevisiae* MSH2-MSH6 and MLH1-PMS1 complexes with DNA using a reversible DNA end-blocking system. *J. Biol. Chem.*, **280**, 22245–22257.
21. R. Qiu, V.C. DeRocco, C. Harris, A. Sharma, M.M. Hingorani, D.A. Erie and K.R. Weninger (2012) Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling. *EMBO J.*, **31**, 2528–2540.
22. L. Giron-Monzon, L. Manelyte, R. Ahrends, D. Kirsch, B. Spengler and P. Friedhoff (2004) Mapping protein-protein interactions between MutL and MutH by cross-linking. *J. Biol. Chem.*, **279**, 49338–49345.
23. L. Manelyte, C. Urbanke, L. Giron-Monzon and P. Friedhoff (2006) Structural and functional analysis of the MutS C-terminal tetramerization domain. *Nucleic Acids Res.*, **34**, 5270–5279.
24. I. Winkler, A.D. Marx, D. Lariviere, R. Heinze, M. Cristóvão, A. Reumer, U. Curth, T.K. Sixma and P. Friedhoff (2011) Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in *Escherichia coli*. *J. Biol. Chem.*, **286**, 17326–17337.
25. G. Feng and M.E. Winkler (1995) Single-step purifications of His6-MutH, His6-MutL and His6-MutS repair proteins of escherichia coli K-12. *BioTechniques*, **6**, 956–965.
26. F.S. Groothuizen, A. Fish, M.V. Petoukhov, A. Reumer, L. Manelyte, H.H.K. Winterwerp, M.G. Marinus, J.H.G. Lebbink, D.I. Svergun, P. Friedhoff and T.K. Sixma (2013) Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp

formation. *Nucleic Acids Res.*, **41**, 8166–8181.

27. W. Kabsch (2010) XDS. *Acta Crystallogr.*, **D66**, 125–132.
28. M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A.J. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A.A. Vagin and K.S. Wilson (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr.*, **D67**, 235–242.
29. R.A. Nicholls, F. Long and G.N. Murshudov (2012) Low resolution refinement tools in REFMAC5. *Acta Crystallogr.*, **D68**, 404–417.
30. R.P. Joosten, J. Salzemann, V. Bloch, H. Stockinger, A.C. Berglund, C. Blanchet, E. Bongcam-Rudloff, C. Combet, A.L. Da Costa, G. Deleage, M. Diarena, R. Fabbretti, G. Fettahi, V. Flegel, A. Gisel, V. Kasam, T. Kervinen, E. Korpelainen, K. Matilla, M. Pagni, M. Reichstadt, V. Breton, I.J. Tickle and G. Vriend (2009) PDB_REDO: automated re-refinement of X-ray structure models in the PDB. *J. appl. Crystallogr.*, **42**, 376–384.
31. G.N. Murshudov, A.A. Vagin and E.J. Dodson (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.*, **D53**, 240–255.
32. G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long and A.A. Vagin (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr.*, **D67**, 355–367.
33. G.J. Kleywegt and T.A. Jones (1994) A super position. *CCP4/ESF-EACBM Newslett. on Protein Crystallogr.*, **31**, 9–14.
34. M.H. Lamers, D. Georgijevic, J.H. Lebbink, H.H.K. Winterwerp, B. Agianian, N. de Wind and T.K. Sixma (2004) ATP increases the affinity between MutS ATPase domains: implications for ATP hydrolysis and conformational changes. *J. Biol. Chem.*, **279**, 43879–438885.
35. A. Aronshtam and M.G. Marinus (1996) Dominant negative mutator mutations in the *mutL* gene of *Escherichia coli*. *Nucleic Acids Res.*, **24**, 2498–2504.
36. K. Baerenfaller, F. Fischer and J. Jiricny (2006) Characterization of the "mismatch repairosome" and its role in the processing of modified nucleosides *in vitro*. *Method. Enzymol.*, **408**, 285–303.
37. I. Biswas, G. Obmolova, M. Takahashi, A. Herr, M.A. Newman, W. Yang and P. Hsieh (2001) Disruption of the helix-u-turn-helix motif of MutS protein: loss of subunit dimerization, mismatch binding and ATP hydrolysis. *J. Mol. Biol.*, **305**, 805–816.
38. K. Hopfner and J.A. Tainer (2003) Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. *Curr. Opin. Struct. Biol.*, **13**, 249–255.
39. T.S.R. Krishna, X. Kong, S. Gary, P.M. Burgers and J. Kuriyan (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*, **79**, 1233–1244.
40. J.S. Lenhart, M.C. Pillon, A. Guarné and L.A. Simmons (2013) Trapping and visualizing intermediate steps in the mismatch repair pathway *in vivo*. *Mol. Microbiol.*, **90**, 680–698.
41. M.S. Junop, W. Yang, P. Funchain, W. Clendenin and J.H. Miller (2003) In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair*, **2**, 387–405.
42. M.L. Mendillo, V.V. Hargreaves, J.W. Jamison, A.O. Mo, S. Li, C.D. Putnam, V.L. Woods Jr and R.D. Kolodner (2009) A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 22223–22228.
43. C. Ban, M. Junop and W. Yang (1999) Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell*, **97**, 85–97.
44. M. Elez, M. Radman and I. Matic (2012) Stoichiometry of MutS and MutL at unrepaired mismatches *in vivo* suggests a mechanism of repair. *Nucleic Acids Res.*, **40**, 3929–3938.

45. G. Plotz, J. Readle, A. Brieger, J. Trojan and S. Zeuzen (2003) N-terminus of hMLH1 confers interaction of hMutLα and hMutLβ with hMutSα. *Nucleic Acids Res.*, **31**, 3217–3226.

46. M.J. Schofield, S. Nayak, T.H. Scott, C. Du and P. Hsieh (2001) Interaction of *Escherichia coli* MutS and MutL at a DNA mismatch. *J. Biol. Chem.*, **276**, 28291–28299.

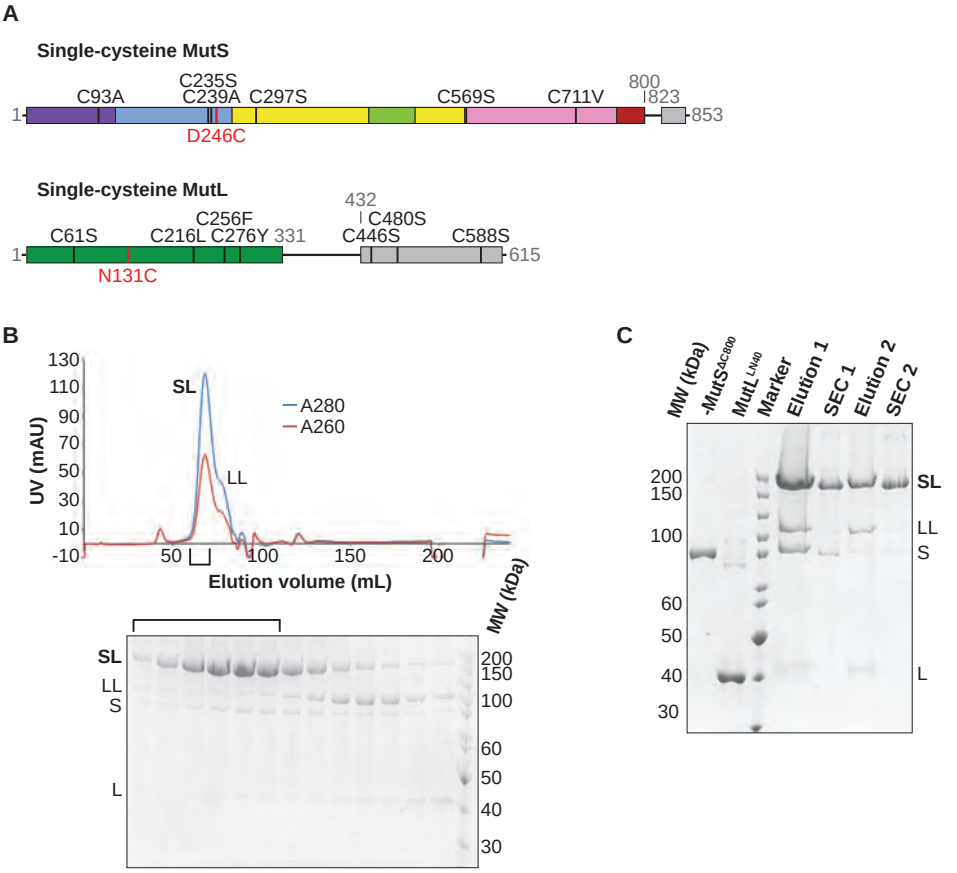
47. A.N. Schorzman, L. Perera, J.M. Cutalo-Patterson, L.C. Pedersen, L.G. Pedersen, T.A. Kunkel and K.B. Tomer (2011) Modeling of the DNA-binding site of yeast Pms1 by mass spectrometry. *DNA Repair*, **10**, 454–465.

48. A. Robertson, S.R. Pattishall and S.W. Matson (2006) The DNA binding activity of MutL is required for methyl-directed mismatch repair in *Escherichia coli*. *J. Biol. Chem.*, **281**, 8399–8408.

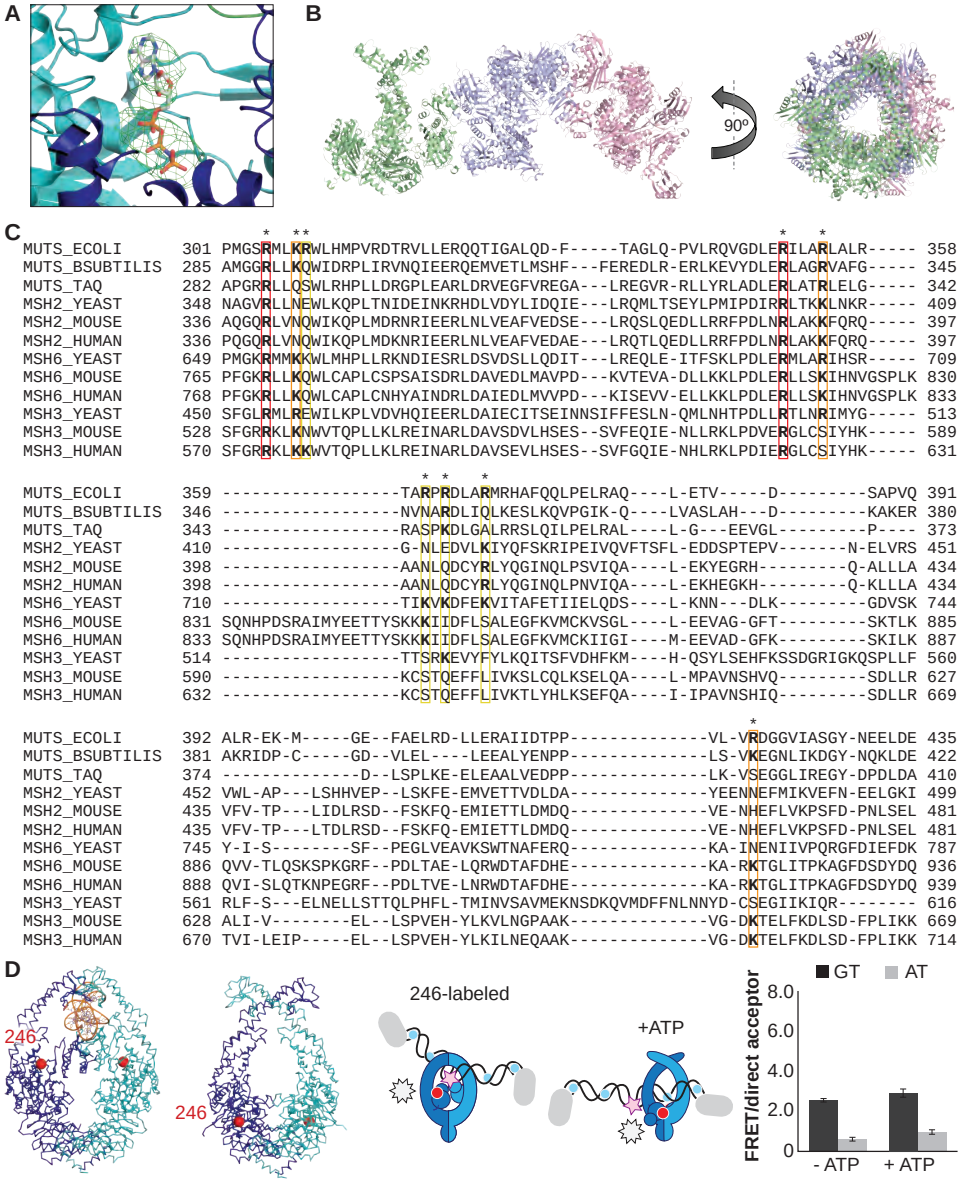
49. M. Cristóvão, E. Sisamakis, M.M. Hingorani, A.D. Marx, C.P. Jung, P.J. Rothwell, C.A.M. Seidel and P. Friedhoff (2012) Single-molecule multiparameter fluorescence spectroscopy reveals directional MutS binding to mismatched bases in DNA. *Nucleic Acids Res.*, **40**, 5448–5464.

50. M.C. Monti, S.X. Cohen, A. Fish, H.H.K. Winterwerp, A. Barendregt, P. Friedhoff, A. Perrakis, A.J.R. Heck, T.K. Sixma, R.H.H. van den Heuvel and J.H.G. Lebbink (2011) Native mass spectrometry provides direct evidence for DNA mismatch-induced regulation of asymmetric nucleotide binding in mismatch repair protein MutS. *Nucleic Acids Res.*, **39**, 8052–8064.

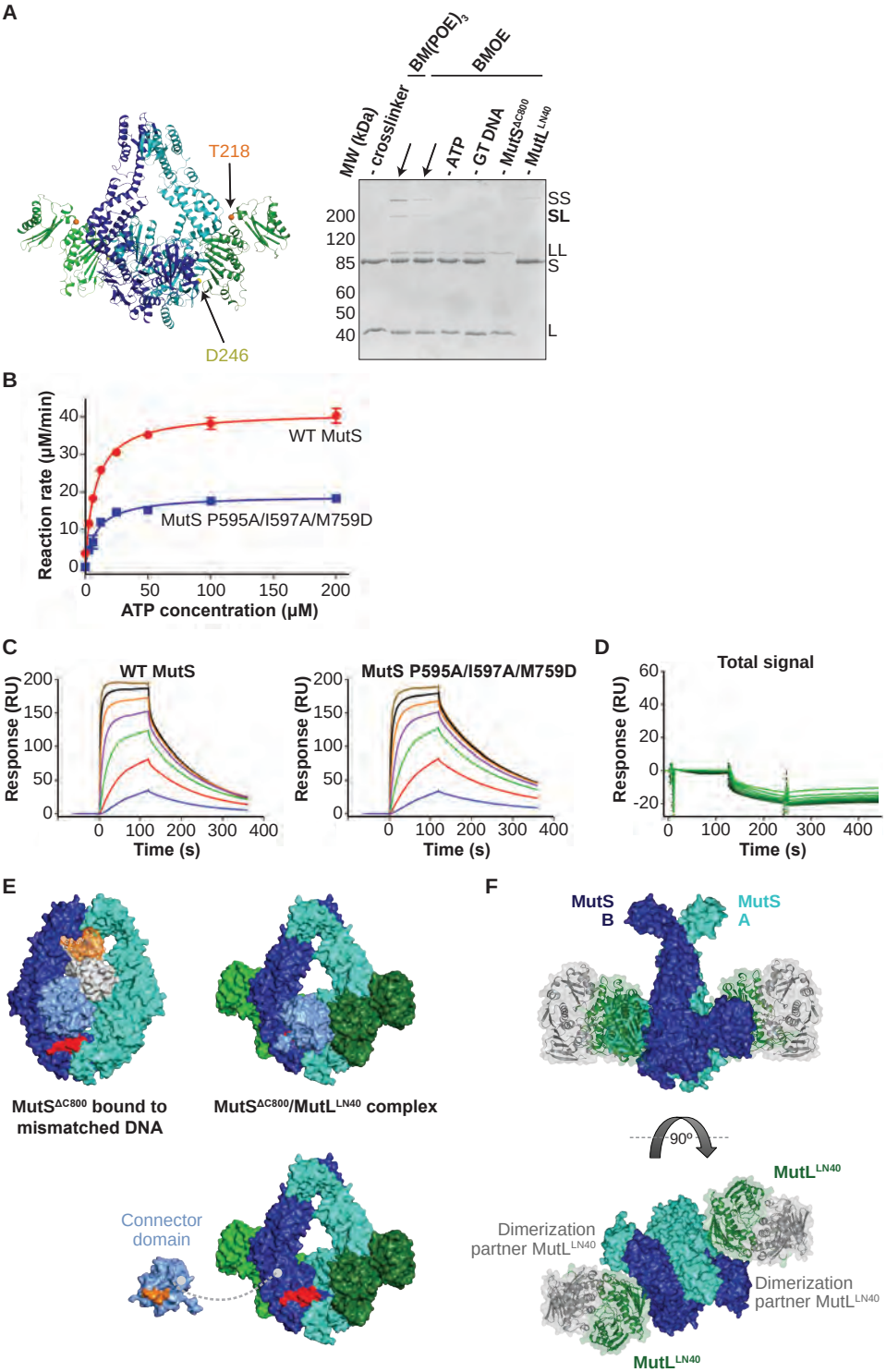
# Supplementary figures and tables

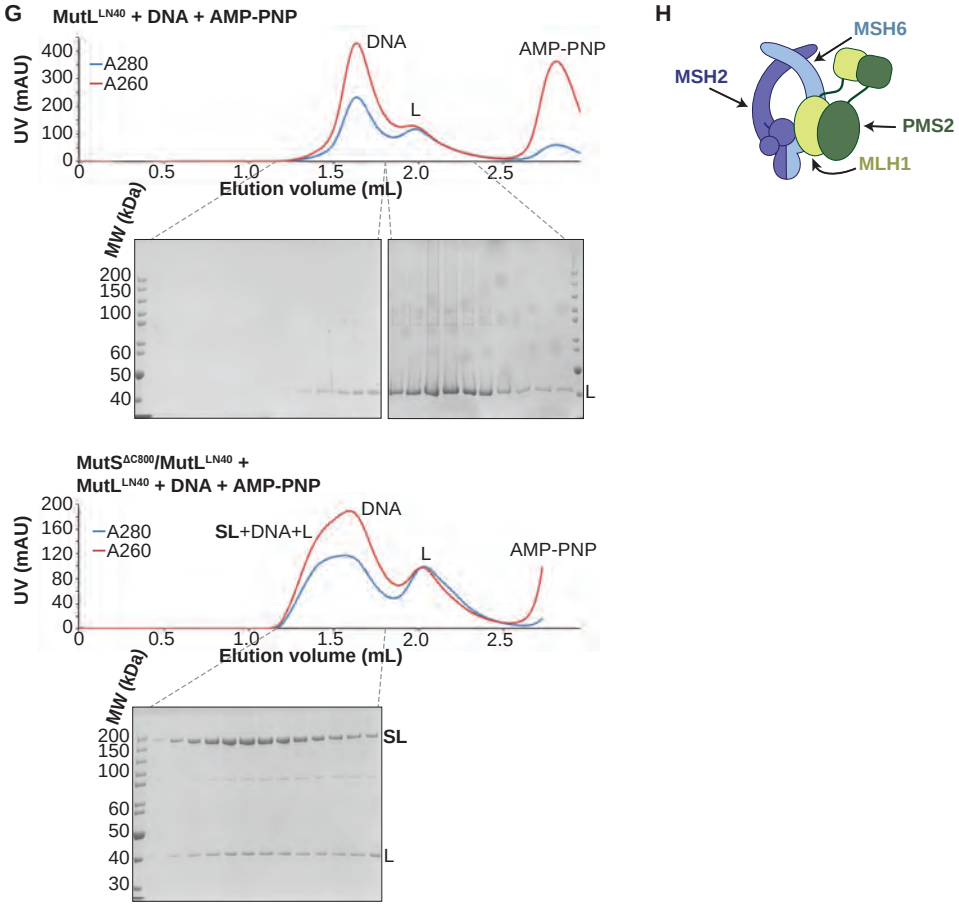



**Supplementary Figure S1.** Crosslinking and purification of the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex. **A**) Single-cysteine MutS D246C and single-cysteine MutL N131C constructs with replaced and introduced cysteine positions are shown, and colored according to domain definitions in main text Figure 1. **B**) Final size-exclusion chromatography profile and corresponding SDS-PAGE gel for the purification of the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex (SL). Pooled fractions are indicated. **C**) Two rounds of crosslinking and purification of MutS$^{\Delta C800}$ D246C and MutL$^{LN40}$ N131C result in almost all MutS$^{\Delta C800}$ subunits crosslinked to MutL$^{LN40}$, as shown on SDS-PAGE (elutions from Talon beads and size-exclusion chromatography [SEC] are indicated).

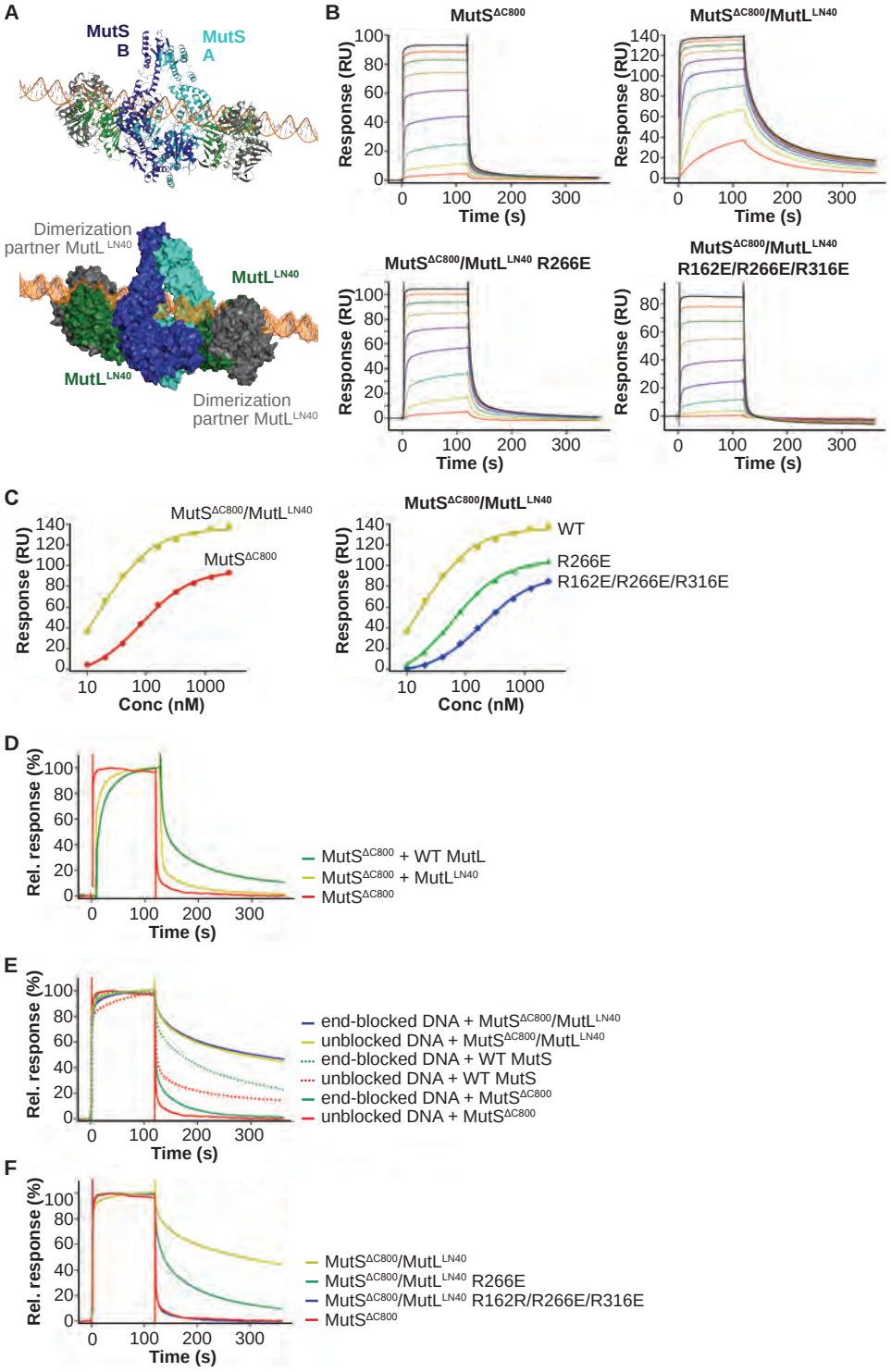



**Supplementary Figure S2.** ATP-analog and DNA in the crystal structure. **A**) Difference density map for AMP-PNP (mFo-dFc at 2.8σ) after refinement without the nucleotide is shown for MutS subunit A in the 4.7 Å crystal structure. **B**) The asymmetric unit in the $P2_1$ crystal form that diffracted to 7.6 Å, which was crystallized with 100-bp DNA aligns the channels in the three $MutS^{\Delta C800}$ dimers (green, violet and pink cartoon representations) such that a DNA strand would fit through all complexes simultaneously. **C**) Sequence conservation of the positively charged residues (indicated by asterisks) in the DNA channel of the MutS sliding clamp. Red: positive charge fully conserved; orange: conserved in >50% of the species shown; yellow: conserved in <50% of the species shown. **D**) No change in FRET between labeled residue 246 on MutS and labeled DNA is observed upon addition of ATP, in agreement with reorientation of the connector domain concomitant with repositioning of the DNA to the new channel.


A
T218
D246
BM(POE)3
BMOE
MW (kDa)
- crosslinker
- ATP
- GT DNA
- MutSΔC800
- MutL LN40
200
120
85
60
50
40
SS
SL
LL
S
L
B
Reaction rate (μM/min)
WT MutS
MutS P595A/I597A/M759D
ATP concentration (μM)
C
WT MutS
MutS P595A/I597A/M759D
Response (RU)
Time (s)
D
Total signal
Response (RU)
Time (s)
E
MutSΔC800 bound to mismatched DNA
MutSΔC800/MutLLN40 complex
Connector domain
F
MutS B
MutS A
90°
MutLLN40
Dimerization partner MutLLN40
Dimerization partner MutLLN40
MutLLN40

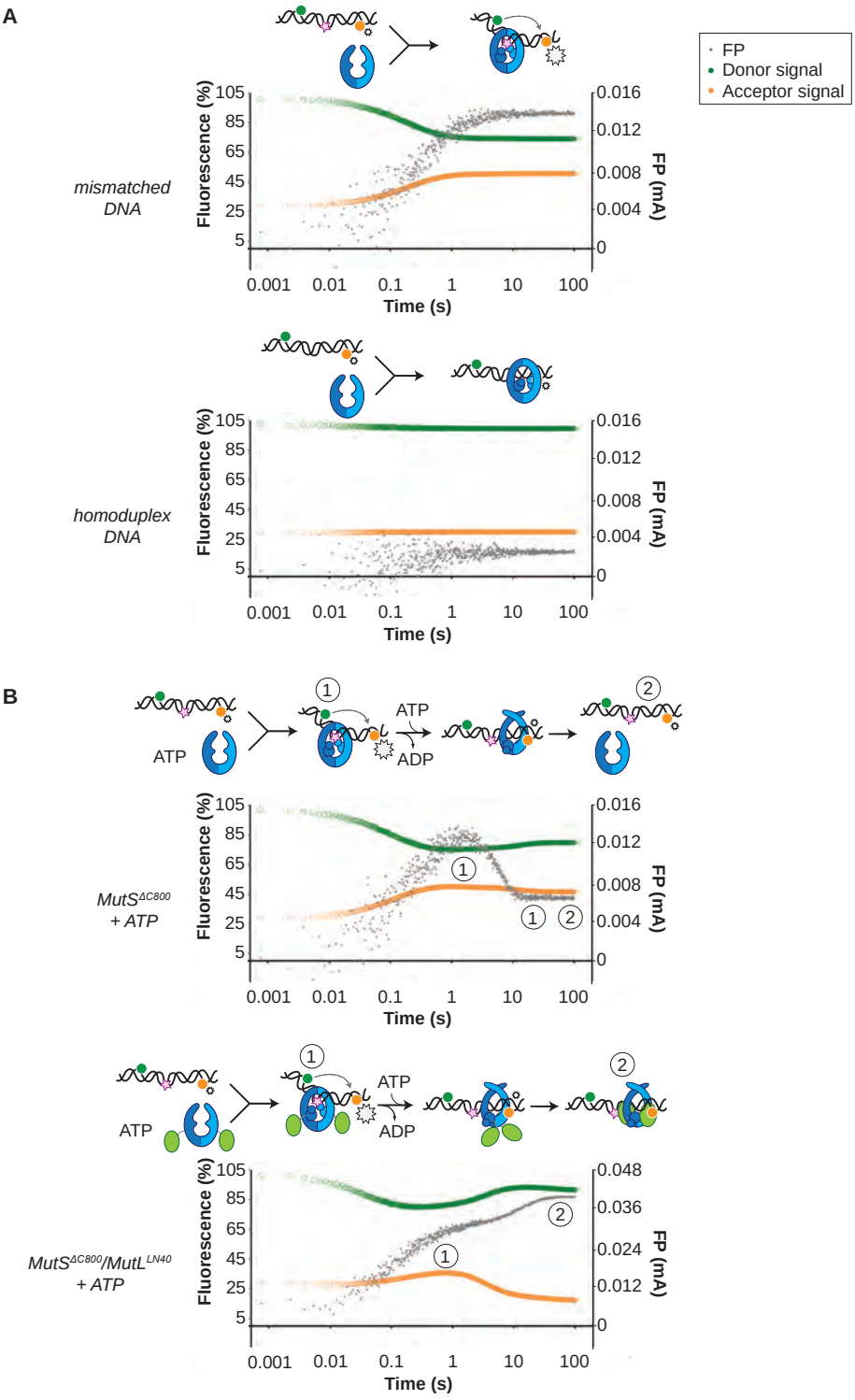


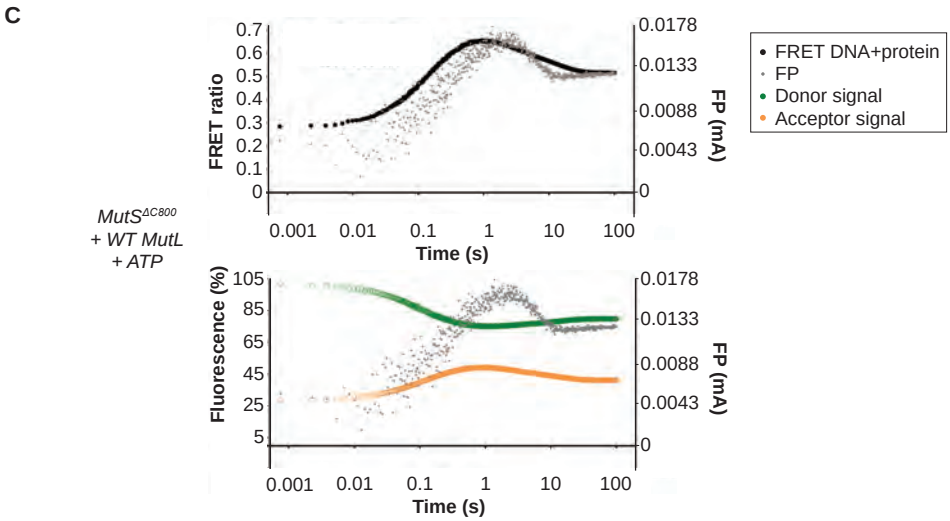

**Supplementary Figure S3.** MutS-MutL interaction. **A**) MutS$^{\Delta C800}$ D246C and MutL$^{LN40}$ T218C do not crosslink efficiently with either a short (BMOE, 8 Å) or a long (BM(PEO)$_3$, 18 Å) crosslinker, as e.g. seen by lack of MutS and MutL depletion. **B**) MutS P595A/I597A/M759D shows ATPase activity with similar $K_m$ (9.5 ± 1 mM) as WT MutS (8.1 ± 0.7 mM) but differs in $K_{cat}$ (mutant: 3.8 ± 0.1 min$^{-1}$; WT: 8.3 ± 0.2 min$^{-1}$). Data points are averages between two measurements and error bars indicate standard deviations. **C**) MutS P595A/I597A/M759D shows similar sliding clamp formation as WT MutS (assay on end-blocked 21-bp DNA with dT$_{20}$ linker (26)). **D**) Without preformed MutS sliding clamps, there is only little binding of MutL to DNA. **E**) The region in the ATPase domain that is solvent exposed during mismatch recognition but showed reduced deuterium exchange upon MutL binding (red, residues 673-686) (42) is buried by the connector domain (light blue surface; orange patch is close to residues 673-686). **F**) The association with the MutS$^{\Delta C800}$ sliding clamp (blue/cyan) does not sterically hinder potential dimerization by the MutL$^{LN40}$ domains (dimer modeled in green and grey as present in PDB entry 1NHJ). **G**) MutL$^{LN40}$ (L) coelutes with crosslinked MutS$^{\Delta C800}$/MutL$^{LN40}$ complex (SL) from size-exclusion chromatography (right), after incubation with 100-bp DNA with a GT mismatch and AMP-PNP, indicating that MutL can still dimerize in this complex. **H**) Model for MSH2/MSH6 interaction with MLH1/PMS2, in which the N-terminus of MLH1 simultaneously binds to the connector domain of MSH2 and the APTase and core domains of MSH6.


A
MutS B
MutS A
Dimerization partner MutL LN40
MutL LN40
MutL LN40
Dimerization partner MutL LN40
B
MutS ΔC800
MutS ΔC800/MutL LN40
MutS ΔC800/MutL LN40 R266E
MutS ΔC800/MutL LN40 R162E/R266E/R316E
Response (RU)
Time (s)
C
MutS ΔC800/MutL LN40
MutS ΔC800
WT
R266E
R162E/R266E/R316E
Response (RU)
Conc (nM)
D
Rel. response (%)
Time (s)
MutS ΔC800 + WT MutL
MutS ΔC800 + MutL LN40
MutS ΔC800
E
Rel. response (%)
Time (s)
end-blocked DNA + MutS ΔC800/MutL LN40
unblocked DNA + MutS ΔC800/MutL LN40
end-blocked DNA + WT MutS
unblocked DNA + WT MutS
end-blocked DNA + MutS ΔC800
unblocked DNA + MutS ΔC800
F
Rel. response (%)
Time (s)
MutS ΔC800/MutL LN40
MutS ΔC800/MutL LN40 R266E
MutS ΔC800/MutL LN40 R162R/R266E/R316E
MutS ΔC800

**Supplementary Figure S4.** DNA binding by the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex. **A**) The model for MutS$^{\Delta C800}$/MutL$^{LN40}$ complex on DNA (orange) sterically allows for LN40 dimerization (dimer modeled in green and grey as present in PDB entry 1NHJ). **B**) Analysis as in Figure 4B, but using shorter, 41-bp mismatched DNA (ATAGGACGCTGACACTGGTGCTTGGCAGCTTCTAATTCGAT annealed with ATCGAATTAGAA GCTGCCAGGCACCAGTGTCAGCGTCCTAT) shows that crosslinked LN40 slows down the MutS sliding clamp on with a GT mismatch, but point mutations in the MutL$^{LN40}$ DNA-binding groove can abolish this. **C**) Absolute response from the assay on 41-bp DNA fitted (single-site binding mode) to estimate maximum signal. Comparison shows that maximum response for MutS$^{\Delta C800}$/MutL$^{LN40}$ binding (150 RU) is 45% higher than for MutS$^{\Delta C800}$ alone (104 RU). Since the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex is 43% larger than MutS$^{\Delta C800}$ dimer alone, this indicates that in this assay a single MutS$^{\Delta C800}$/MutL$^{LN40}$ complex binds the DNA duplex. **D**) Co-injection of WT MutL results in slower release from DNA than MutS$^{\Delta C800}$ alone, while uncrosslinked MutL$^{LN40}$ only has a minor effect (traces normalized to maximum response). **E**) While end-blocks on the DNA slow down MutS release, there is no effect on the already slow MutS$^{\Delta C800}$/MutL$^{LN40}$ release. This is more pronounced for WT MutS, which forms more stable dimers than MutS$^{\Delta C800}$. Crosslinked MutL$^{LN40}$ with mutations in the DNA-binding groove is deficient in the ability to retain MutS$^{\Delta C800}$ on DNA. **F**) Normalized overlay of the 1280 nM traces from Figure 4B,C.

◀


A
FP
Donor signal
Acceptor signal
mismatched DNA
homoduplex DNA
Fluorescence (%)
FP (mA)
Time (s)
B
ATP
ADP
MutS^ΔC800 + ATP
MutS^ΔC800/MutL^LN40 + ATP




**Supplementary Figure S5.** DNA kinking by MutS$^{\Delta C800}$ and MutS$^{\Delta C800}$/MutL$^{LN40}$. **A**) Stopped-flow FRET and FP assay shows kinking of 45-bp DNA by MutS binding only if there is a mismatch. Separate traces for the fluorophores are shown (orange: acceptor; green: donor fluorophore; grey: FP). Size of FRET events are indicated by stars. **B**) While MutS$^{\Delta C800}$ initially kinks the DNA and subsequently releases in the presence of ATP, the MutS$^{\Delta C800}$/MutL$^{LN40}$ complex remains bound to unkinked DNA. **C**) Using WT MutL (400 nM) mixed with the MutS$^{\Delta C800}$ in this assay results in more binding at equilibrium than MutS$^{\Delta C800}$ alone, while there is less FRET than for MutS$^{\Delta C800}$ alone. Under the graph with the FRET ratio (black: FRET; grey: FP), separate traces for the fluorophores are shown as in A).

**Supplementary Table S1.** Mutator assay with acquired rifampicin resistance.

| Protein | Mutations per $10^7$ | (95% confidence interval) |
|---|---|---|
| **MutS variant (MutL interface)** | | |
| *Empty vector* | 0.601 | (0.446-0.772) |
| WT MutS | 0.0686 | (0.0408-0.101) |
| MutS P595A/I597A | 0.0545 | (0.0310-0.0826) |
| MutS M759D | 0.0819 | (0.0490-0.121) |
| MutS Y563A | 0.0488 | (0.0272-0.0749) |
| MutS P595A/I597A/M759D | 0.704 | (0.556-0.864) |
| MutS Y563A/P595A/I597A | 0.317 | (0.233-0.411) |
| MutS Y563A/P595A/I597A/M759D | 0.773 | (0.618-0.941) |
| **MutL variant (MutS interface)** | | |
| *Empty vector* | 5.43 | (4.00-7.00) |
| WT His-MutL | 0.121 | (0.0542-0.206) |
| His-MutL A138E | 2.76 | (2.12-3.46) |
| His-MutL H139A | 0.103 | (0.0439-0.179) |
| His-MutL A138E/H139A | 4.87 | (3.55-6.33) |
| His-MutL R55D/R57D | 6.41 | (4.99-7.95) |
| His-MutL R200D | 0.663 | (0.432-0.932) |
| His-MutL R55D/R57D/H139A | 5.33 | (3.93-6.89) |
| His-MutL R55D/R57D/A138E/H139A | 6.13 | (4.58-7.84) |
| His-MutL R55D/R57D/H139A/R200D | 5.22 | (3.84-6.76) |
| His-MutL R55D/R57D/A138E/H139A/R200D | 5.48 | (4.04-7.06) |
| **MutL variant (DNA binding)** | | |
| His-MutL R266E | 5.87 | (4.78-7.04) |
| His-MutL R162E/R266E/R316E | 5.39 | (4.37-6.49) |

Mutation rates and 95% confidence intervals were determined using the Fluctuation AnaLysis CalculatOR (*http://www.mitochondria.org/protocols/FALCOR.html*) using the MSS-MLE method. For MutS, at least 24 independent colonies were picked; for MutL, at least 12 independent colonies were picked.

**Supplementary Table S2.** Crystallographic data collection and refinement statistics.

| | Crystal form 1<br>27-bp DNA | Crystal form 2<br>27-bp DNA | Crystal form 3<br>100-bp DNA |
|---|---|---|---|
| **Data collection** | | | |
| Space group | C2 | C2 | $P2_1$ |
| Cell dimensions | | | |
| a, b, c (Å) | 165.9, 188.5, 200.4 | 380.6, 126.5, 243.3 | 192.6, 109.4, 277.5 |
| α,β,γ (°) | 90.0, 94.8, 90.0 | 90.0, 91.4, 90.0 | 90.0, 90.0, 90.0 |
| Resolution (Å)* | 82.7-4.71 (4.96-4.71) | 49.94-6.6 (7.13-6.6) | 49.3-7.6 (8.5-7.6) |
| $R_{merge}$ | 19.4 (79.7) | 21.3 (80.1) | 16.8 (91.9) |
| I/σ(I) | 2.5 (1.0) | 3.4 (1.1) | 4.3 (1.0) |
| Completeness (%) | 97.3 (98.0) | 96.8 (97.7) | 81.3 (82.5) |
| Redundancy | 2.4 (2.4) | 2.9 (3.0) | 2.3 (2.2) |
| **Refinement** | | | |
| Resolution (Å) | 4.7 | 6.6 | 7.6 |
| No. reflections | 31052 | 21305 | 11763 |
| $R_{work}/R_{free}$ | 35.15/36.75 | 28.34/33.20 | 27.79/34.28 |
| No. atoms | 21915 | 45054 | 45054 |
| Protein | 21822 | 44868 | 44868 |
| Ligand/ion | 93 | 186 | 186 |
| Water | 0 | 0 | 0 |
| B-factors | | | |
| Protein | 242.2 | 206.7 | 161.8 |
| Ligand/ion | 260.2 | 178.5 | 126.2 |
| Water | *n/a* | *n/a* | *n/a* |
| R.m.s deviations | | | |
| Bond lengths (Å) | 0.0156 | 0.0145 | 0.0113 |
| r.m.s.Z(bonds) | 0.795 | 0.727 | 0.558 |
| Bond angles (°) | 1.8624 | 1.8045 | 1.5694 |
| r.m.s.Z(angles) | 0.854 | 0.827 | 0.709 |

*Highest resolution shell is shown in parentheses.

# 5

# General discussion

Before cells divide, their whole genome needs to be replicated. To avoid permanent incorporation of mutations, the DNA mismatch repair (MMR) system has evolved to detect and remove base pairing errors. While this is not an easy task in the context of billions of base pairs, MMR is strikingly efficient (1,2).

Central to this efficiency lies the initial recognition of the mismatch by MutS proteins. In this research, we have zoomed in on the *Escherichia coli* variant of this protein to understand how it can both sense the error and signal for repair. The structural information gained by crystallography and single-cysteine tools has been indispensable to understand the molecular mechanism behind the activity of MutS. At the same time, we gained crucial information about how this signalling is transferred to the downstream effectors, starting with MutL.



**MutS proteins as regulators of repair initiation**

The various states and conformations of MutS proteins illustrate the complex manner in which they function. In the previous chapters, we aimed to shed light on some of these elusive states, to better understand how initiation of MMR is regulated. The state of *E. coli* MutS is difficult to define when in solution: it is in equilibrium between dimers and tetramers (Chapter 2), and its clamp domains are mobile (Chapter 3). We found that both of these properties affect mismatch recognition.

Until now, only the prokaryotic homodimers of MutS are known to exist in an equilibrium between dimers and tetramers (3-6). As the tetramer has been implicated in antirecombination (7,8), this may be a way for these prokaryotic proteins to expand their versatility, whereas their eukaryotic homologs have diversified to each perform specialized functions (9). Other explanations include stabilization of the dimeric state of the eukaryotic homologs by partner proteins.

Before a mismatch has been found, MutS must sample the DNA, which is thought to initially involve embracing the helix by the clamp domains. The presence of a mismatch can then be detected by the mismatch-binding domain, or the protein will release again. We observed that a delicate balance between kinking and a helical fold of the lever domains controls DNA association and dissociation (Chapter 3). Comparable 'opening up' of the dimers is expected to take place in the structurally similar eukaryotic MutS homologs.

After mismatch recognition, MutS binds ATP, which induces a conformational change to form a sliding clamp on DNA (10-12) (Chapter 4). This occurs only when a mismatch is recognized, with the exception of a CC mismatch (13,14). In this way, ATP is used to validate that indeed a mismatch has been found. This switching system implies a way to ensure that repair is only initiated when necessary. For the eukaryotic mismatch repair proteins MutSα and MutSβ, ATP-induced sliding clamp formation is also observed (15,16). It is expected that those sliding clamps are structurally similar to that of *E. coli* MutS.

As for the MMR MutS proteins, the homologous meiotic MSH4/MSH5 heterodimer also undergoes an ATP-induced conformational change (17). In this case, however, the resulting sliding clamp is thought to enclose two DNA helices, thus stabilizing recombination intermediates. While the cavity for DNA in the MSH4/MSH5 dimer is expected to be larger than for MMR MutS proteins (18), a similar conformational change involving subunit tilting would not easily leave space for two DNA helices while still preserving the ability to slide along the DNA. Little is known about the exact mechanism of this heterodimer, and potentially there are two dimers that each enclose one DNA strand and interact in a manner similar to the *E. coli* MutS tetramer.

The transitions between the various states of MutS result from different levels of structural regulation: the equilibrium between dimers and tetramers involves transient binding between C-termini; the flexibility that allows an 'open' state arises from partial unfolding at specific helical sites; and the conformational change to form a sliding clamp is specifically induced by ATP binding, resulting in increased packing of the ATPase domains and whole-subunit tilting.



Future research may shed more light on the specific transition between the mismatch-recognition state and the sliding-clamp state of MutS. While the tilting of the subunits may push out the connector and mismatch-binding domains, the outward rotation of a single connector domain could instead precede the full subunit tilting. Capturing intermediate states for structural studies, or FRET assays combined with single-molecule studies may be able to answer such questions (19).

Deficiency in MMR predisposes to cancer (20), but it is not always straight forward to diagnose such predisposition when missense mutations in MMR genes are the cause of nonfunctional proteins. Moreover, simple mapping of variants on static structures does not always explain pathogenicity either. It could be that certain residues are important for a different state of the protein (or transitioning toward it), or for protein-protein interactions. Indeed pathogenic mutations are found in many different regions of MutS homologs (21), and possibly MutS proteins are susceptible to perturbations precisely because they have to transition between several states. Full mechanistic understanding of this complex system is therefore crucial to understand and predict the impact of variations in MutS proteins.

**Activation of MutL in MMR**

While recognition and signalling of a DNA mismatch is essential, reading out this signal to initiate repair is equally important. In MMR, MutL performs this function by promoting nicking and unwinding of the newly replicated strand. MutL will initiate this only upon recognition of the sliding-clamp conformation of MutS (22-25). How MutL can so specifically recognize this MutS state was observed in the crystal structure of the MutS/MutL complex, in which it became clear that the two binding interfaces can only be

simultaneously bound by MutL when MutS is in the sliding-clamp state, and both interfaces are essential for complex formation (Chapter 4). This mode of interaction and validation can probably be directly extrapolated to the eukaryotic MutS and MutL homologs. In this case, the heterodimeric nature of the eukaryotic proteins will more clearly impose asymmetry on the interaction.

Due to the repositioned DNA in the channel of the MutS sliding clamp, MutL is loaded onto the DNA by the sliding clamp (Chapter 4), whereas MutL alone has low DNA binding affinity (26,27). In this binding, the MutL-DNA interaction itself may contribute to MutS/MutL/DNA complex formation. MMR is only initiated when MutL is loaded onto DNA, thus transferring the validation of a mismatch by MutS to MutL.



This explains why MutL will only act after MutS has switched to the sliding-clamp conformation, but does not reveal how MutL in turn initiates repair. Possibly, the localization of MutL towards DNA is the first essential step, but the presence of a mismatch still needs to be communicated to a strand discrimination site. In *E. coli*, his could mean that the whole MutS/MutL complex has to travel towards a GATC site, where the endonuclease MutH can make a nick (28). While crosslinked MutS/MutL does not travel as freely on DNA as the MutS sliding clamp alone (Chapter 4), dynamic binding of MutL to the MutS sliding clamp could virtually result in sliding of the whole complex.

Additionally, MutL can undergo ATP-induced conformational changes, which include structuring and dimerization of the N-terminal domains, and compaction of the dimer (29-31). The exact effect and timing of these slow conformational changes is not clear, but they may be required for activation of MutH and UvrD. The dimeric assembly of the N-terminal domain of MutL is compatible with its interaction with the MutS sliding clamp and DNA (Chapter 4). This leaves the possibility for a model where the dimerization of the MutL N-termini occurs during MutS interaction. This could either place the MutL C-termini on the same side of the DNA (close to the C-terminus of MutS), or on the other side of the DNA, thus enclosing the helix between the flexible linkers of MutL. The possibility of enclosing DNA has been suggested and observed for MutL proteins previously (26,32-35). If then the MutL N-termini interact with the C-termini as suggested (31), this compaction would in the MutL variants with endonuclease activity direct these nicking domains towards the DNA. β-clamp or PCNA interaction by the C-terminal domains (36-38) could in this case be responsible for the strand discrimination. How exactly MutH (in *E. coli*) or helicase activities are activated, however, is not answered by this model and should be the focus of future research.

In Chapter 1, we discussed the well-regulated initiation of DNA mismatch repair and the many cellular functions to which MMR proteins are coupled. Interactions with the replication machinery not only allows direct coupling of repair, but also links the MMR

proteins to the numerous signalling pathways that are associated with replication. Moreover, the specialized eukaryotic variants and their interacting proteins result in many layers of cellular regulation. Added to the intrinsic conformational versatility of MutS and MutL proteins that we described in this research, it is not surprising that a wide variety of functions are attributed to these proteins.

## References


1. R.R. Iyer, A. Pluciennik, V. Burdett and P.L. Modrich (2006) DNA mismatch repair: functions and mechanisms. *Chem. Rev.*, **106**, 302–323.
2. S.D. McCulloch and T.A. Kunkel (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.*, **18**, 148–161.
3. S. Takamatsu, R. Kato and S. Kuramitsu (1996) Mismatch DNA recognition protein from an extremely thermophilic bacterium, *Thermus thermophilus* HB8. *Nucleic Acids Res.*, **24**, 640–647.
4. I. Biswas, C. Ban, K.G. Fleming, J. Qin, J.W. Lary, D.A. Yphantis, W. Yang and P. Hsieh (1999) Oligomerization of a MutS mismatch repair protein from *Thermus aquaticus*. *J. Biol. Chem.*, **274**, 23673–23678.
5. K.P. Bjornson, L.J. Blackwell, H. Sage, C. Baitinger, D. Allen and P. Modrich (2003) Assembly and molecular activities of the MutS tetramer. *J. Biol. Chem.*, **278**, 34667–34673.
6. V. Miguel, M.R. Monti and C.E. Argaraña (2008) The role of MutS oligomers on *Pseudomonas activityeruginosa* Mismatch Repair System activity. *DNA Repair*, **7**, 1799–1808.
7. M.A. Calmann, A. Nowosielska and M.G. Marinus (2005) Separation of mutation avoidance and antirecombination functions in an *Escherichia coli mutS* mutant. *Nucleic Acids Res.*, **33**, 1193–1200.
8. K.C. Tham, N. Hermans, H.H.K. Winterwerp, M.M. Cox, C. Wyman, R. Kanaar and J.H.G. Lebbink (2013) Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. *Mol. Cell*, **51**, 326–337.
9. B.D. Harfe and S. Jinks-Robertson (2000) DNA mismatch repair and genetic instability. *Annu Rev Genet.*, **34**, 359–399.
10. C. Jeong, W. Cho, K. Song, C. Cook, T. Yoon, C. Ban, R. Fishel and J. Lee (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat Struct Mol Biol*, **18**, 379–385.
11. W. Cho, C. Jeong, D. Kim, M. Chang, K. Song, J. Hanne, C. Ban, R. Fishel and J. Lee (2012) ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure*, **20**, 1–11.
12. R. Qiu, V.C. DeRocco, C. Harris, A. Sharma, M.M. Hingorani, D.A. Erie and K.R. Weninger (2012) Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling. *EMBO J.*, **31**, 2528–2540.
13. S. Su, R.S. Lahue, K.G. Au and P. Modrich (1987) Mispair specificity of methyl-directed DNA mismatch correction *in vitro*. *J. Biol. Chem.*, **263**, 6829–6835.
14. F.S. Groothuizen, A. Fish, M.V. Petoukhov, A. Reumer, L. Manelyte, H.H.K. Winterwerp, M.G. Marinus, J.H.G. Lebbink, D.I. Svergun, P. Friedhoff and T.K. Sixma (2013) Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp formation. *Nucleic Acids Res.*, **41**, 8166–8181.
15. S. Gradia, D. Subramanian, T. Wilson, S. Acharya, A. Makhov, J. Griffith and R. Fishel (1999)

hMSH2–hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell*, **3**, 255–261.

16. S. Gupta, M. Gellert and W. Yang (2011) Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat Struct Mol Biol*, **19**, 72–78.
17. T. Snowden, S. Acharya, C. Butz, M. Berardini and R. Fishel (2004) hMSH4-hMSH5 recognizes Holliday junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Mol. Cell*, **15**, 437–451.
18. R. Rakshambikai, N. Srinivasan and K.T. Nishant (2013) Structural insights into *Saccharomyces cerevisiae* Msh4–Msh5 complex function using homology modeling. *PloS One*, **8**, e78753.
19. D.A. Erie and K.R. Weninger (2014) Single molecule studies of DNA mismatch repair. *DNA Repair*, **20**, 71–81.
20. H.T. Lynch and A. de la Chapelle (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.*, **36**, 801–818.
21. M. Drost, A. Lützen, S. van Hees, D. Ferreira, F. Calléja, J.B.M. Zonneveld, F.C. Nielsen, L.J. Rasmussen and N. de Wind (2013) Genetic screens to identify pathogenic gene variants in the common cancer predisposition Lynch syndrome. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 9403–9408.
22. M. Grilley, K.M. Welsh, S. Su and P. Modrich (1989) Isolation and characterization of the *Escherichia coli mutL* gene product. *J. Biol. Chem.*, **264**, 1000–1004.
23. T.A. Prolla, Q. Pang, E. Alani, R.D. Kolodner and R.M. Liskay (1994) MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science*, **265**, 1091–1093.
24. K. Drotschmann, A. Aronshtam, H.J. Fritz and M.G. Marinus (1998) The *Escherichia coli* MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic Acids Res.*, **26**, 948–953.
25. S. Acharya, P.L. Foster, P. Brooks and R. Fishel (2003) The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair. *Mol. Cell*, **12**, 233–246.
26. C. Ban, M. Junop and W. Yang (1999) Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell*, **97**, 85–97.
27. M.S. Junop, W. Yang, P. Funchain, W. Clendenin and J.H. Miller (2003) In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair*, **2**, 387–405.
28. T.A. Kunkel and D.A. Erie (2005) DNA mismatch repair. *Annu. Rev. Biochem.*, **74**, 681–710.
29. C. Ban and W. Yang (1998) Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, **95**, 541–552.
30. E.J. Sacho, F.A. Kadyrov, P. Modrich, T.A. Kunkel and D.A. Erie (2008) Direct visualization of asymmetric adenine nucleotide-induced conformational changes in MutLα. *Mol. Cell*, **29**, 112–121.
31. T. Yamamoto, H. Iino, K. Kim, S. Kuramitsu and K. Fukui (2011) Evidence for ATP-dependent structural rearrangement of nuclease catalytic site in DNA mismatch repair endonuclease MutL. *J. Biol. Chem.*, **286**, 42337–42348.
32. M.C. Hall, P.V. Shcherbakova, J.M. Fortune, C.H. Borchers, J.M. Dial, K.B. Tomer and T.A. Kunkel (2003) DNA binding by yeast Mlh1 and Pms1: implications for DNA mismatch repair. *Nucleic Acids Res.*, **31**, 2025–2034.
33. A. Guarné, S. Ramon-Maiques, E.M. Wolff, R. Ghirlando, X. Hu, J.H. Miller and W. Yang (2004) Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J.*, **23**, 4134–4145.
34. J. Kosinski, I. Steindorf, J.M. Bujnicki, L. Giron-Monzon and P. Friedhoff (2005) Analysis of the quaternary structure of the MutL C-terminal domain. *J. Mol. Biol.*, **351**, 895–909.

35. J. Gorman, A.J. Plys, M. Visnapuu, E. Alani and E.C. Greene (2010) Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nat Struct Mol Biol*, **17**, 932–938.

36. S.D. Lee and E. Alani (2006) Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *J. Mol. Biol.*, **355**, 175–184.

37. A. Pluciennik, L. Dzantiev, R.R. Iyer, N. Constantin, F.A. Kadyrov and P. Modrich (2010) PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 16066–16071.

38. M.C. Pillon, J.H. Miller and A. Guarné (2010) The endonuclease domain of MutL interacts with the β sliding clamp. *DNA Repair*, **10**, 87–93.

# Summary
# Samenvatting
# Curriculum vitae
# List of publications
# Dankwoord

## Summary

To ensure proper functioning of new cells, the genetic information of the parent cell needs to be copied accurately. The large number of base pairs in most genomes make this a difficult task, in which creating errors is inevitable. To resolve this, almost all cells contain a DNA mismatch repair (MMR) system, which efficiently removes most of the errors. MMR is therefore essential to maintain a stable genome, and defects in MMR can lead to cancer in humans. The MMR system requires multiple successive steps for initiation, which are the focus of the research in this thesis.

In **Chapter 1**, we give an overview of MMR and explain how this system prevents incorporation of replication errors. We zoom in on the two key proteins that initiate the repair: MutS and MutL. Although these proteins are well-conserved throughout evolution, eukaryotes contain multiple homologs that are specialized in different types of error repair and are linked to several cellular systems. To understand these complex systems, it is helpful to study the basic MMR mechanism in *Escherichia coli*. We discuss the multiple states that MutS adopts during mismatch repair, in which the detection of a mismatch is validated by ATP and signalled by changing to a sliding clamp conformation. MutL activates downstream effectors upon binding to the MutS sliding clamp, while MutL proteins can also use ATP to change conformation. These different layers of regulation emphasize the complex manner in which MutS and MutL control initiation of DNA mismatch repair.

**Chapter 2** focusses on the dimeric and tetrameric assemblies of *E. coli* MutS. The equilibrium between these oligomeric states complicates kinetic analysis of DNA binding by MutS. Using point mutations and crosslinking in the tetramerization domains, we stabilized the dimer and tetramer of MutS separately. We found that the tetramer of MutS binds stronger to DNA, probably due to binding of DNA by both dimers within the tetramer. DNA-binding experiments with stable MutS dimers showed that affinities differ greatly, depending on the mismatch and the flanking DNA sequence. Moreover, we observed that MutS did not efficiently form an ATP-induced sliding clamp after binding a CC mismatch, explaining why this mismatch is not repaired efficiently.

In **Chapter 3**, we discuss the flexibility of the MutS clamp domains, which allows the MutS dimer to associate onto DNA and dissociate from it. We present a DNA-free crystal structure of a MutS dimer, which shows kinking in specific regions, allowing for movement of the clamp domains. Using mutagenesis, we validated that the ability to kink these regions is important for DNA binding by MutS.

In **Chapter 4**, we shed light on the signalling from mismatch recognition to downstream repair events. We crosslinked MutS to the N-terminal domain of MutL, and crystallized the stable complex. In this structure, MutS is in the sliding clamp conformation, which we validate using FRET. The ATP-induced conformational change in MutS pushes the DNA into a new channel, and MutS forms a loose ring around the DNA helix. At the same time, the connector domain has rotated outward and is available for MutL, so that MutL can bind to a MutS dimer via two interfaces. The structure of the complex suggests that MutS loads MutL onto DNA where it can activate downstream effectors, which we validate using mutagenesis.

In **Chapter 5**, we discuss the findings of Chapters 2-4 and place them in the context of the general MMR knowledge. While the different states of MutS allow subtle regulation of MMR initiation, future work may include studies of the transitions between these states. Moreover, conformational changes in MutL after binding to MutS may be necessary to activate the downstream effectors, of which more understanding is required. Finally, the conformational versatility of these proteins helps us to understand how their diverse cellular functions can be achieved.

## Samenvatting

Om het functioneren van nieuwe cellen te waarborgen, moet de genetische informatie van de ouder-cel nauwkeurig worden gekopieerd. Het grote aantal basenparen in het genoom maakt dit een lastige taak, waarbij het maken van fouten (zogenaamde 'mismatches') onvermijdelijk is. Om dit op te lossen bevatten bijna alle cellen een systeem voor DNA mismatch herstel (*DNA mismatch repair*; MMR), dat efficiënt de meeste fouten verwijdert. Omdat MMR belangrijk is om een stabiel genoom te behouden, is het duidelijk dat een niet-functioneel MMR systeem bij mensen kan leiden tot kanker. De activatie van MMR vereist meerdere opeenvolgende stappen die de focus van het onderzoek in dit proefschrift zijn.

In **Hoofdstuk 1** geven we een overzicht van MMR en leggen we uit hoe dit systeem replicatie-fouten herstelt. We richten ons op de twee belangrijkste eiwitten die het herstel starten: MutS en MutL. Hoewel deze eiwitten evolutionair zijn geconserveerd bevatten eukaryote organismes meerdere homologen, die gespecialiseerd zijn in verschillende soorten herstel en die gekoppeld zijn aan verscheidene cellulaire functies. Om deze complexe systemen te begrijpen is het nuttig om het basale MMR mechanisme van *Escherichia coli* te bestuderen. We bespreken de verschillende conformaties die MutS aanneemt tijdens MMR, waarbij de detectie van een mismatch wordt gevalideerd door ATP te binden en als signaal wordt doorgegeven wanneer MutS zich tot een ring om DNA vormt (de *sliding-clamp* conformatie). MutL bindt aan deze MutS *sliding-clamp* en activeert dan andere eiwitten, terwijl MutL zelf ook ATP kan gebruiken om van conformatie te veranderen. Deze verschillende lagen van regulering benadrukken de complexe wijze waarop MutS en MutL het starten van DNA mismatch herstel dirigeren.

**Hoofdstuk 2** richt zich op de dimeer- en tetrameer-samenstelling van *E. coli* MutS. De overgang tussen deze oligomere toestanden bemoeilijkt kinetische analyse van DNA-binding door MutS. Met behulp van mutaties en chemische verbindingen in de tetramerisatie domeinen stabiliseerden we de dimeer en de tetrameer van MutS afzonderlijk. We kwamen erachter dat de tetrameer van MutS sterker aan DNA bindt, waarschijnlijk doordat beide dimeren in de tetrameer aan DNA kunnen binden. In DNA-bindingsexperimenten met stabiele MutS dimeren bleek dat de bindings-affiniteit sterk kan verschillen, afhankelijk van de mismatch en de DNA-sequentie rondom de mismatch. Bovendien zagen we dat MutS niet goed in staat is om de ATP-geïnduceerde *sliding-clamp* conformatie aan te nemen in het geval van een CC mismatch, wat verklaart waarom deze mismatch niet efficiënt wordt gerepareerd.

In **Hoofdstuk 3** bespreken we de flexibiliteit van de MutS *clamp*-domeinen, die het mogelijk maakt dat de MutS dimeer aan DNA bindt en er weer vanaf gaat. We presenteren de kristalstructuur van een MutS dimeer zonder DNA, die knikken in bepaalde gebieden laat zien, waardoor beweging van de *clamp*-domeinen mogelijk is. Met behulp van mutaties in die gebieden bevestigden we dat deze knikken belangrijk zijn voor DNA-binding door MutS.

In **Hoofdstuk 4** onderzoeken we hoe mismatch-herkenning wordt omgezet tot een signaal dat aangeeft dat herstel moet starten. We maakten een chemische verbinding tussen MutS en het N-terminale domein van MutL en kristalliseerden dit stabiele complex. In de kristalstructuur is MutS in de *sliding-clamp* conformatie, wat we konden bevestigen met fluorescentie-experimenten. De door ATP geïnduceerde conformatieverandering in MutS duwt het DNA naar een nieuw kanaal, waardoor MutS dan een losse ring rond de DNA helix vormt. Tegelijkertijd is het *connector*-domein naar buiten gedraaid, waar het beschikbaar is voor MutL, zodat MutL kan binden aan de MutS dimeer via twee bindingsvlakken. De kristalstructuur van het complex suggereert dat MutL door MutS op DNA geladen wordt en dan het herstel kan opstarten, wat we bevestigden door mutaties in MutL te maken.

In **hoofdstuk 5** bespreken we de bevindingen van de hoofdstukken 2-4 en plaatsen we ze in de context van wat er al bekend is rond MMR. De verschillende toestanden van MutS maken het mogelijk om de initiatie van MMR op subtiele wijze te reguleren. Toekomstig onderzoek kan zich richten op de overgang tussen deze toestanden. Bovendien kunnen conformatieveranderingen in MutL na binding aan MutS noodzakelijk zijn voor het starten van herstel; een onderwerp dat nog meer inzicht vereist. Tot slot heeft de conformationele veelzijdigheid van deze eiwitten ons wat beter laten begrijpen hoe ze hun verscheidene cellulaire functies kunnen uitoefenen.

## Curriculum vitae

Flora Sulamith Groothuizen werd geboren op 27 december 1984 te Ede. Ze groeide op in Zoetermeer, waar ze in 2003 bij het Oranje Nassau College haar gymnasiumdiploma behaalde. Aansluitend startte zij een bacheloropleiding Bio-Farmaceutische Wetenschappen met een minor in Scheikunde aan de Universiteit Leiden. Na deze opleiding afgerond te hebben ging ze vijf maanden naar India, Nepal en Pakistan om vrijwilligerswerk te doen en rond te reizen. In 2007 begon ze aan de onderzoeksmaster Biomolecular Sciences aan de Vrije Universiteit in Amsterdam. Als onderdeel van deze master deed ze een stage in de groep van prof.dr. Titia Sixma aan het Nederlands Kanker Instituut te Amsterdam. Hier werkte zij aan de structurele en functionele analyse van eiwitten uit het polycomb complex. Een tweede stage deed ze in de groep van prof.dr. Alan Mark aan de University of Queensland in Brisbane in Australië. Daar deed ze computationeel onderzoek naar de activatie van de prolactine-receptor. Ze sloot haar master vervolgens *cum laude* af in de specialisaties Biological Physics en Biological Chemistry.

In november 2009 begon Flora als onderzoeker in opleiding op de afdeling Biochemie van het Nederlands Kanker Instituut, onder begeleiding van prof.dr. Titia Sixma. De resultaten van dit onderzoek staan in dit proefschrift beschreven.

&

## List of publications

F.S. Groothuizen*, A. Fish*, S.D. Dharadhar, H.H.K. Winterwerp, T.K. Sixma. Kinking of the coiled coil in the MutS lever domain allows DNA binding and release. *In preparation.*
* equal contribution

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

F.S. Groothuizen*, A. Fish*, M.V. Petoukhov, A. Reumer, L. Manelyte, H.H.K. Winterwerp, M.G. Marinus, J.H.G. Lebbink, D.I. Svergun, P. Friedhoff and T.K. Sixma (2013) Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp formation. *Nucleic Acids Research*, **41**, 8166-8181
* equal contribution

F.S. Groothuizen, D. Poger and A.E. Mark (2010) Activating the prolactin receptor: effect of the ligand on the conformation of the extracellular domain. *Journal of Chemical Theory and Computation*, **6**, 3274–3283

# Dankwoord

Het is af! En dat voelt goed! Het was een lange weg met veel ups en downs maar nu heb ik dit proefschrift eindelijk in handen. Natuurlijk had ik dit niet kunnen doen zonder alle mensen om me heen!

Allereerst wil ik uiteraard **Titia** bedanken. Je gaf me de kans om in de kristallografie te werken, wat voor mij de perfecte keuze bleek. Ik wist van tevoren niet dat ik het zo'n prachtig vakgebied zou vinden, waarin zoveel antwoorden voor fundamentele vragen te vinden zijn. Bedankt voor de vrijheid die je me gaf om verschillende dingen uit te proberen, terwijl je ook precies de vinger wist te leggen op de punten waar verbetering nodig was. Je zorgt hier voor een omgeving met veel mogelijkheden om meer te leren en je stimuleerde me om naar interessante congressen en cursussen te gaan. Ik ben blij dat ik me daardoor op meerdere vlakken heb kunnen ontwikkelen.

I want to thank **Tassos** for contributing so much crystallographic and biochemical knowledge to the lab, and for the encouraging words when I needed them. Thank you for being in my PhD committee, together with **Hein** en **Heinz**, in which you all had great constructive input during our meetings.

I would like to thank all the people who have been involved in the MM2M network. It was amazing to discuss questions about the complex mismatch-repair system with people who understand exactly what you are talking about. Hierbij bedank ik in het bijzonder **Joyce Lebbink**, die een drijvende factor was voor deze bijeenkomsten. Je enorme kennis van de literatuur en je kritische inzichten maakten ons werk altijd weer beter. I also want to thank **Peter Friedhoff** for our wonderful collaboration. A large portion of my PhD was based on the tools developed in your lab. Thank you for the many detailed and fruitful scientific discussions that we had. My thanks also go to **Terence Strick**, **Josef Jiricny** and **Kim Sneppen** for all scientific discussions and ideas during these meetings. The MM2M meetings were a lot of fun too, especially thanks to **Ines**, **Andi**, **Klaas**, **KC**, **Michele** and **Evan**.

I also want to thank our other collaborators for their contributions to several parts of this work: **Dmitri Svergun**, **Maxim Petoukhov**, **Martin Marinus**, **Garib Murshudov** and **Rob Nicholls**.

Het was geweldig om deze jaren te kunnen werken met alle collega's in het Sixma-en-Perrakis-lab. Zoals gebruikelijk in de wetenschap gingen er mensen weg en kwamen er

nieuwe voor in de plaats, waardoor dit een lange lijst is.

**Herrie**, jij bent al zo lang betrokken bij het MutS project en je was onmisbaar bij mijn werk! Dankjewel voor de enorme hoeveelheden eiwit die je steeds weer maakte (nee, 'slechts' 5 mg is toch echt teveel om weg te gooien)! De todo-lijstjes waren soms niet aan te slepen omdat je zo snel prepte en ik ben blij dat we steeds beter op elkaar ingespeeld waren. Ik hoop dat je nog lang geniet van het klussen aan je huis in Hongarije (maar wanneer is dat zwembad nou af?)! **Sasha**, you were closely involved with the MutS project and we did some great work together! Thank you for always being there to discuss biophysical questions, and especially for the endless trouble shooting that you did on the Biacore. **Pim**, jij bent de spil van dit lab! Bij jou kon ik altijd terecht met lab- of persoonlijke zaken en dat was echt heel fijn. Dankjewel voor de vele gezellige lunches met de meest uiteenlopende gespreksonderwerpen. **Caroline**, alles loopt op rolletjes als jij het regelt! Wat hebben we veel dewars naar de synchrotron verstuurd en ondanks je zorgen ging dat eigenlijk gewoon altijd goed! Bedankt voor alle dingen die we met elkaar konden bespreken. **Danny**, we zijn op dezelfde dag begonnen, waardoor we samen alles over kristallografie konden uitzoeken. Helaas zijn we niet in het jaar van de kristallografie gepromoveerd maar we halen wel allebei de eindstreep met mooie structuren. Jij bent er ook bijna! Maar eerst alvast in die zaal staan als mijn paranimf. **Robbert**, bedankt voor alle humor en gezelligheid die je op ons kantoor bracht en dat je straks als paranimf naast me staat! Verlies de moed niet; met jouw inzet zal je vast gauw de vruchten (of kristallen) plukken van je enorme kloneer-werk! **Francesca**, you helped me to take the first steps in this lab, and we made a lot of good memories! Your enthusiasm for science has been an inspiration to me. Thank you for always reminding me to stay positive when I encountered problems on the way. I wish you all the happiness in the world together with Robert and Emma! **Judith**, van jou leerde ik om te relativeren binnen dit werk. Je zorgde voor een motiverende sfeer op kantoor en ik vond het heel leuk om je paranimf te zijn. **Shreya**, I was very lucky to have such a good student. You really want to get to the bottom of things, and I am sure that this will get you some great results in your PhD! **Michael**, thanks for the discussions about experiments and so many other topics, which somehow always took place close to your bench. I hope to see your thesis soon as well! **Tatjana**, thank you for your endless ability to help everyone in the lab. I think that countless more crystals will grow in plates that you set up, and I wish you many future dives together with Carsten. **Elli**, thanks for the nice time in the office, for the many fun hours that we spent at the SLS, and for all the chats about cats and balcony gardening! **Patrick**, je bent altijd bereid om iedereen te helpen en je bent steeds weer op de hoogte van nieuwe mogelijkheden in het lab. Dankjewel voor de positieve sfeer die jij brengt op de woensdagmeetings. **Marcello**, thanks for all the discussions about the peculiar aspects of academic science and the process of submitting papers. I wish you all the best in your next postdoc. **Doreth**, succes met de volgende stappen in het mismatch-

repair project! **Nassos**, good luck in the States and with the rest of you PhD! I'm sure you'll do great! **Robbie**, dankjewel voor de vele suggesties voor refinements en de leuke gesprekken over computationele aspecten. **Willem-Jan**, je bent nu zowel celbioloog als kristallograaf dus dat kan alleen maar goed komen in jouw carrière! **Thanga**, I hope that the many protein complexes that you produce will give you some new insights! **Magda,** I wish you and your family all the best in the future. **Yoshi**, good luck with all the NMR experiments and the rest of your postdoc. **Andrea**, I hope you have a great time these years in Amsterdam. **Prakash**, thanks for patiently explaining me how to process crystallographic data, and for all the help in other ways. I will always have good memories of Mauritius, and I wish you and Caroline all the best. **Bernat**, you were such a great and positive factor in the lab, always willing to help out. I hope you will manage to work through the frustrations of a PhD as well! **Bas**, bedankt voor de korte maar leuke tijd dat je in het lab was. Maar hou wel die urea ver van mijn eiwit hè! **Mark**, de begintijd van mijn promotie was extra vrolijk met jou in het lab. Wie weet zien we elkaar nog in New York! **Alex**, dank je voor de vele gezellige lunches op Pims kantoor. **Annet**, veel van dit werk is gebaseerd op jouw voorwerk en daardoor lijkt het alsof we langer samen hebben gewerkt dan die paar maanden. Het allerbeste! **Rick**, I wish you all the best for the future. **John**, veel geluk met alle toekomstige reizen naar je familie! **Krista**, ik wens je het allerbeste. Ook mijn stagestudent **Niels** wil ik bedanken voor de hulp bij het enige niet-mismatch-herstel-project waar ik even aan heb gewerkt. I want to thank several other former lab members for the good times in the lab: **Jens**, **Christophe**, **Mariano**, **Leonie**, **Azusa**, **Dene**, and **Andrea**.

The past few years we shared the division with other groups as well. **Thijn**, **Lisa** (bedankt voor de gezellige treinritjes en veel geluk met je nieuwe gezinnetje!), **Jacqueline** (ja, mijn eiwit is de mooiste!), **Vincent** (bedankt dat we de spanning om onze papers konden delen!) **Joppe**, **Lucas**, **Matthijs**, **Markus**, **Ferdy** and **Sjoerd**: thanks for sitting through all the crystallographic talks! It was great to learn so much about the challenges of genetics and virus biology.

**Sach**, **Nicholas** (I enjoyed our discussions about travel plans!) , **Frank**, **Steven**, **Anas**, **Chris**, thanks for the nice time and the interesting talks that we had from your group!

Ook mijn vrienden en familie hielpen me om de eindstreep te halen. **Effi** en **Bianca**, ik ben heel blij dat we altijd wel weer tijd vonden om samen de dingen in ons leven door te spreken. Ook al vonden jullie het maar een raar wereldje waarin ik werkte, ik heb er altijd veel aan gehad om alles hierover bij jullie kwijt te kunnen. **Anne**, ook als er lange tijd tussen zit is het altijd leuk als we elkaar weer spreken en het was fijn dat jij altijd begrip had voor werkstress. Ik hoop dat we elkaar niet uit het oog verliezen en dat jouw toekomst leuk werk biedt, afgewisseld met mooie reizen!

En dan is er nog onze BFW-groep met **Louise** (wanneer kom je weer logeren?), **Annelieke** (gaan we gauw weer naar een Leids festivalletje?), **Petra** (dankjewel voor al je lieve kaartjes!), **Henriette** (het alleberste aan jou, Elise en Hakiem!), **Leonie** (veel geluk met je eigen plekje in Beverwijk!) en **Jeroen** (meneer de UD!). Tijdens onze studie was het allemaal begonnen en gelukkig was het daarna niet voorbij met onze feestjes, etentjes en pubquizzen. Ik hoop dat er nog vele zullen volgen!

**Mieke** en **Johan**, bedankt voor altijd een warm welkom en voor de interesse in mijn onderzoek. **Sven** en **Pru**, de heerlijke veganistische maaltijden die we deelden waren altijd een lichtpuntje in de drukke tijden tijdens mijn promotie.

**Dineke**, dankjewel dat je er altijd voor ons bent. Laten we gauw weer eens langs de Waal gaan wandelen. **Jolijn** en **Matthijs**, het is fijn dat wij er met z'n drieën samen voor staan als dat nodig is en ik ben blij dat we elkaar de laatste tijd wat vaker zien.
Lieve **mam** en **pap**, de basis voor alles is een goed thuis, dat ik van jullie kreeg.

Tenslotte had ik dit alles nooit kunnen volbrengen zonder **Coen**, mijn steun in alles ♡. Het is altijd fijn om thuis te kunnen komen bij jou en Pindi en ik kijk uit naar de avonturen die onze toekomst brengt!