Covalent and kinetic trapping repair machinery for structural and functional studies

Dissertation

zur Erlangung des Akademischen Grades
des Doktors der Naturwissenschaften
-Dr.rer.nat.-
des Fachbereiches Biologie und Chemie, FB08
der Justus-Liebig-Universität Gießen

von

Vladislav Kunetsky
Giessen
June 2021

This dissertation was carried out at the Institute for Biochemistry in the Department of Biology and Chemistry (FB08) of the Justus Liebig University Giessen, from August 2017 to June 2021 under the direction of Apl. Prof. Dr. Peter Friedhoff. The work was funded by Marie Skłodowska-Curie actions "DNAREPAIRMAN" (H2020-MSCA-ITN-2016).

Primary Referee:
Apl. Prof. Dr. Peter Friedhoff
Institut für Biochemie, FB08
Justus-Liebig-Universität
Heinrich-Buff-Ring 17
35392 Giessen

Secondary Referee:
Prof. Dr. Mark Dillingham
School of Biochemistry
University of Bristol
University Walk
Bristol BS8 1TD, UK

Abstract

The principal attribute of living matter is genetic information used for reproduction. Before the cell can divide, all genetic information must be reproduced without errors leading to mutations. To decrease the number of errors, a DNA mismatch repair (MMR) system must identify and repair the error. Due to the enormous length of the genome, this process must be fast and efficient. Errors can initiate apoptosis, cause tumors in higher organisms, or lead to the fixation of mutations in a population.

The topic of research in these theses is the evolutionarily conserved mismatch DNA repair system from the model organism Escherichia coli, namely the initial stage of the process, which involves proteins such as MutS, MutL, MutH, and UvrD. The purpose of this work was also the development of new methods, such as single-cysteine site-specific cross-linking and Förster Resonance Energy Transfer (FRET)-based methods, to study these proteins.

The primary player in such a system is MutS, which is responsible for recognizing mismatches, working with incredible accuracy. Scanning millions of nucleotides, it is able to find a single mismatch and cause a cascade reaction that will start the process of DNA repair. In this study, we investigated and reviewed in detail the mechanisms and conformational changes occurring in this protein, which are necessary for its accurate and correct operation. The structural information and site-specific single cysteine approach result in a very productive pipeline. The data obtained during this work supported the sliding clamp hypothesis of the active state of MutS and also helped to characterize several other novel transient states. In particular, ATP-induced rotation of the connector domain has been shown, as well as the inability of MutS to recruit MutL when the mismatch domain is cross-linked to DNA.

The same approach was applied to another pair of proteins important for the mismatch pathway, namely, MutL and MutH cross-linking. The site-specific cross-linking with the proper design of single-cysteine variants can be used for obtaining active complexes. Furthermore, this method in the long term can give significant results for structural studies of the active MutL-MutH complex, which has not been obtained before. This strategy has previously demonstrated its power in obtaining a very important new state of MutS, namely the sliding clamp.

Developing FRET-based functional methods for the binding of MutL to the DNA and MutL recruitment by MutS to the DNA, we obtained a universal and novel technology that can be used in the future to studies the kinetics and other activities of DNA interacting proteins. The fundamental knowledge acquired during conformational changes studies in combination with the fluorescent methods described above led us to the establishment of a novel, simple and robust method for checking the quality of DNA, and in particular for observing mismatches and other damages using the accumulation of fluorescent MutS.

Erklärung:

Ich erkläre: Ich habe die vorgelegte Dissertation selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Ich stimme einer evtl. Überprüfung meiner Dissertation durch eine Antiplagiat-Software zu. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Declaration:

“I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.”

(Vladislav Kunetsky)
Gießen, June 2021

Acknowledgments

I would like to express my sincere gratitude to my scientific advisor Apl. Professor Dr. Peter Friedhoff for scientific inspiration, recommendations, and valuable suggestions, as well as constructive criticism of my work. In particular, I would like to note the special atmosphere created in the laboratory, in which I was able to effectively and flexibly conduct my scientific work.

Special thanks to Gwen and the working group of Prof. Dr. Mark Dillingham and Prof. Dr. Nigel Savery for the warm welcome, teamwork, and support during my stay in Bristol.

I am very much obliged to Prof. Dr. Katja Sträßer for organizing the work of the laboratory and the institute at the highest level and for participating in many exciting projects.

I would like to thank my excellent colleague Olha, a member of our small "MMR" team, for fruitful discussions, tips, helping in the experiments, and for the emotional support during this time.

I sincerely express my heartfelt thanks to Mrs. Heike Büngen for teaching me new experimental techniques and incredible help in preparing for experiments.

Many thanks to the secretariat, in particular to Karina, who helped me with many bureaucratic and organizational problems.

Special thanks to Maya and the working group of Prof. Dr. Elena Kubareva for teamwork and fruitful results.

I would like to express my gratitude to all the members of the ITN DNARepairman consortium for this amazing project, which has connected many scientists in Europe. Thank you for this wonderful time, for the vast and invaluable experience I received.

In particular, I would like to thank Prof. Dr. Titia K. Sixma, Dr. Joyce Lebbink, Prof. Dr. Meindert Lamers, Dr. Rafael Fernández Leiro for the opportunity to participate in an important and interesting project.

I also express my gratitude to Anna, Prof. Dr. Marcin K. Chmielewski, and the collective of company "FutureSynthesis" for a warm welcome during my secondment in Poznan.

I acknowledged the EU Horizon 2020 Marie Skłodowska Curie grant [722433] funding support.

List of publication

This thesis is based on the following articles:

1. DNA mismatch/damage detection using FRET-based assay monitoring the loading of multiple MutS.
   Kunetsky, V., Storozhuk, O., Brouwer, G., Laffeber, C., Dillingham, M., Lebbink, J., Friedhoff, P. [in preparation]

2. The selection process of licensing a DNA mismatch for repair.
   Fernandez-Leiro, R., Bhairosing-Kok, D., **Kunetsky, V.**, Laffeber, C., Winterwerp, H., Groothuizen, F., Fish, A., Lebbink, J., Friedhoff, P., Sixma, T., Lamers, M. Nat Struct Mol Biol (2021).

3. Probing the DNA-binding center of the MutL protein from the *Escherichia coli* mismatch repair system via crosslinking and Förster resonance energy transfer.
   Monakhova, M., Ryazanova, A., **Kunetsky, V.,**, Li, P., Shilkin, E., Kisil, O., Rao, D., Oretskaya, T., Friedhoff, P., Kubareva, E. Biochimie, Volumes 171–172, 2020, Pages 43-54.

Other papers by the author:

1. Cryo-EM structures reveal how ATP and DNA binding in MutS coordinate the sequential steps of DNA mismatch repair.
   Borsellini, A., **Kunetsky, V.**, Friedhoff, P., Lamers, M. [submitted]

2. From parts lists to functional significance—RNA–protein interactions in gene regulation.
   Kilchert, C., Sträßer, K., **Kunetsky, V**., & Änkö, M. L. (2020) Wiley Interdisciplinary Reviews: RNA, 11(3), e1582.

3. Dynamic mRNP Remodeling in Response to Internal and External Stimuli.
   Zarnack, K., Balasubramanian, S., Gantier, M. P., **Kunetsky, V.**, Kracht, M., Schmitz, M. L., & Sträßer, K. (2020) Biomolecules, 10(9), 1310.

Other activities

Poster presentations:

1. September 2018, 11^{th} Annual GGL Conference, Giessen, Germany, "Covalent and kinetic trapping repair machinery for structural and functional studies"

Oral presentations:

1. January 2018, DNARepairMan First Annual Meeting, Netherlands Cancer Institute, Amsterdam, Netherlands
2. November 2018, 3^{rd} Marie Curie Workshop, "Protein labelling and cross-linking, and fluorescence approaches", Giessen, Germany
3. November 2018, DNARepairMan Second Annual Meeting, Justus Liebig University, Giessen, Germany
4. July 2019, DNARepairMan Scientific Meeting, Leiden University Medical Center, Leiden, Netherlands
5. September 2019, 12^{th} Annual GGL Conference 2019, "Covalent and kinetic trapping repair machinery for structural and functional studies", Giessen, Germany
6. May 2020, 3^{rd} Annual DNARepairMan Meeting, Online
7. September 2020, First Digital GGL Annual Conference 2020, "Covalent trapping mismatch activated long-lived signaling clamp state of Muts", 3^{rd} place winner, Giessen, Germany
8. November 2020, 4^{th} Annual DNARepairMan Meeting, Online

Animations:

1. June 2018, "The DNA hero story", in a callobration with the Willem de Kooning Academy, Rotterdam, Netherlands
2. July 2019, "DNA mismatch repair systems in *E. coli*"
3. September 2019, "Conformational changes in MutS"
4. January 2021, "Microscopic molecular machines & our future", in a callobration with the Willem de Kooning Academy, Rotterdam, Netherlands

Secondments/visits:

1. January - February 2020, University of Bristol, Bristol, England
2. March - May 2020, FutureSynthesis company, Poznan; Poland

Courses given:

1. January - April 2021, "Scientific Illustrations and Animations", Justus Liebig University, Giessen, Germany

Contents

List of Figures

List of Tables

Chapter 1

Introduction

DNA repair systems

DNA is a key component of life, it encodes the genetic information of every organism. The genome is constantly under attack by intrinsic and extrinsic factors that can alter DNA chemical structure and damage encoded information. Environmental factors, products of normal cellular metabolism and spontaneously breaking chemical bonds can generate alteration in genome structure[1]. Genetic stability plays a significant role in preventing oncogenesis. To preserve organism genome evolution created various DNA repair systems. The systems in mammalian cells which are known by today represent nucleotide- and base-excision repair, homologous recombination, end joining, telomere metabolism, and mismatch repair[2]. Base-excision repair (BER) targets small chemical modifications of bases and corrects single-nucleotide damages caused by oxidation, deamination, and alkylation, concerned with damage of endogenous origin and uses two pathways: short and long patches[3]. Nucleotide-excision repair (NER) is involved in repairing single-stranded DNA damages by using an intact complementary chain as a matrix. The mechanism includes two subpathways: global genome NER (GG-NER) and transcription-coupled repair (TC-NER) (Figure 1.1)[4]. Unlike the BER, NER is designed for larger DNA damage, such as pyrimidine dimers, formed in DNA by ultraviolet radiation or other exogenous sources[5]. Homologous recombination (HR) and end joining (EJ) resolve the problem of double-strand breaks caused by ionizing radiation, free radicals, chemicals, and during replication of a single strand breaks[6]. The mismatch repair system differs from all other systems since it does not repair the damage, but rather recognizes and removes incorrectly inserted nucleotides[7]. Secondly, the system is highly conserved. The system in Escherichia coli (*E. coli*) serves as a helpful model for biological research, while the MMR system is also present in eukaryotic cells. The relevance of a mechanistic understanding of DNA repair pathways has been recognized by the award of the 2015 Nobel Chemistry Prize to three pioneers in the field of DNA repair.

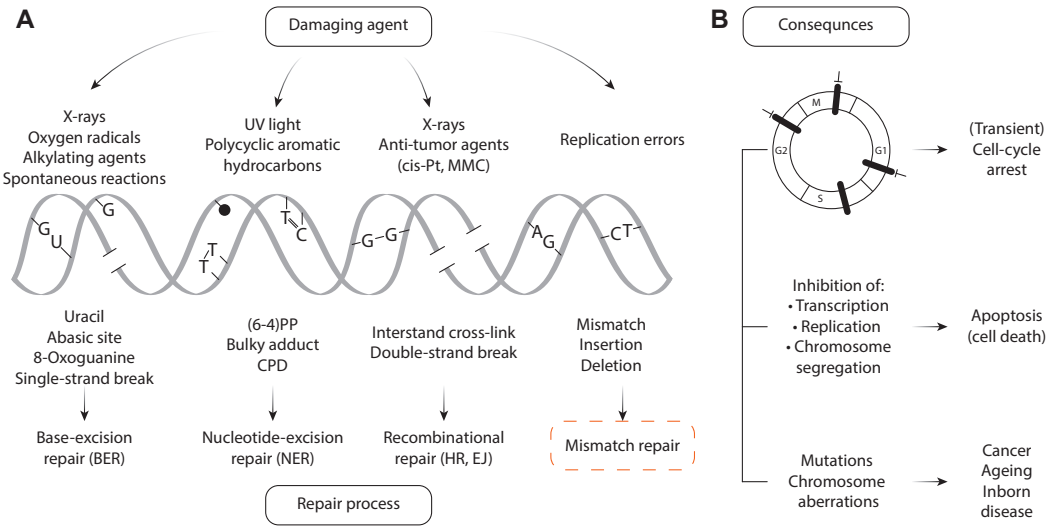


Figure 1.1: Common repair systems and consequences. **(A)** DNA damaging agents (top), DNA lesions (middle), and repair systems (bottom). **(B)** Damages lead to transient arrest (top) and inhibition of the metabolic mechanisms of DNA (center). Long-term consequences lead to point mutations that can affect individual genes or cause chromosomal abnormalities, which in turn affect several genes (below) and thus can cause biological changes (cancer, aging). The figure was adapted from[2].

The biological function of the DNA mismatch repair system

The term *mismatch repair* was introduced in 1964 in the studies of nucleotide bromination and as an explanation of gene conversion in genetic recombination[8]. The first evidence that mismatches initiate a repair reaction was shown using bacterial transformation experiments[9,10]. The artificial heteroduplex DNA with mismatches transformed in *E. coli* cells elicits a repair reaction[11]. All these experiments were carried out as part of the study of gene recombination. But it has also been hypothesized that this system may be responsible for identifying and removing errors done during replication. Using cell extract of *E. coli* it was demonstrated that repair by this system is associated with adenine methylation at GATC sequence[12]. Deletion of any of four mutator genes (MutS, MutL, MutH, or UvrD) increases mutation rate in *E. coli* 50-100 fold[13,14]. It was confirmed the results that the MMR pathway depends on four mutator genes MutS, MutL, MutH, and UvrD. At the same time, it requires also adenosine triphosphate (ATP), and reaction is prevented when both DNA strands are methylated[15]. MMR system is able to repair all types of mismatches except C/C[16]. The process was reconstructed in vitro first in 1989[17]. In subsequent investigations with a modified early substrate, it was shown that a mismatch and one methylated GATC site are enough for DNA repair. Changing also the position of mismatches and GATC site on circular DNA resulted in confirmation that the mismatch repair system supports bidirectional excision[18]. All these four proteins MutS, MutL, MutH, and UvrD are not sufficient to carry out a complete methyl-directed mismatch repair in vitro. Later were found the missing and essential enzymes involved in this process in vitro: single-strand binding proteins, exonuclease I, DNA polymerase holoenzyme III, and DNA ligase[17]. Since

early experiments showed bidirectional exonuclease activity of the process, missing endonucleases were found: ExoVII or RecJ with 5' to 3' excision and ExoI or ExoX with 5' to 3' activity[19,20]. DNA polymerase III interacts with MutS and MutL and does not affect activation MutH in an MM-dependent manner. Beyond the β-clamp role in chromosomal replication, MutS and MutL also interact with β-clamp[21,22].

The DNA Mismatch repair system corrects mistakes done by polymerase that escape proofreading activity and rise replication fidelity by 100-1000 fold[23]. Mismatch repair system has two main roles: it recognizes errors done during replication by polymerase and also discriminates newly synthesized strand with a mistake from mother strand using methylation of GATC site[24]. Despite the different accuracy of polymerases, the system corrects all errors and is able to balance the replication fidelity of the leading and lagging strands[25]. Dam methylase follows the replication fork with a lag of about a minute, so the daughter chain remains temporarily unmethylated, providing a time window for mismatch repair pathway[26,27]. Mutations in four human DNA Mismatch repair genes (MSH2, MLH1, PMS2, PMS1) cause hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome[28].

Comparison of MMR system in *E. coli* and eukaryotic organisms

Prokaryotes

The core mismatch repair process in *E. coli* involves at least seven proteins (MutS, MutL, MutH, UvrD, one of several exonucleases, sliding clamp, DNA polymerase III holoenzyme, and DNA ligase). Despite having the structures of MutS and MutL, the dynamic and flexible nature of the MutS–MutL complex has left open many puzzles regarding the mechanism of MMR. One of the major obstacles hindering the structural analyses is the highly dynamic nature of this complex[29].

The process starts with the evolutionary conserved MutS searching and recognizing mismatches in the genome. The binding of MutS to DNA causes bending of heteroduplex DNA (Figure 1.2). Followed by ATP-induced conformational changes MutS recruits another conserved ATPase MutL[30–32]. The strand discrimination and error removal are believed to be coordinated by this MutS/MutL/DNA complex. Upon ATP binding by MutL conformational changes take place and it results in activation of MutH endonuclease activity[33], as well as activation of UvrD helicase activity[34]. In *E. coli* cells daughter strand remains transiently unmethylated until adenines will be methylated by Dam-methyltransferase methylases[26,27]. It gives a transient opportunity for MMR to incise the daughter strand[35]. UvrD begins unwinding towards mismatch (MM) from the nick caused by activated MutH[36]. The purpose of ExoI is still under discussion[37]; recently MMR-pathway in vitro was reconstructed in an exonuclease-independent manner[38].

Eukaryotes

Firstly, in comparison with a bacterial system where MutS and MutL proteins are homodimers, proteins in eukaryotic systems represent heterodimeric proteins[39]. Secondly, MMR in *E. coli* system was already well studied, in eukaryotic cells, many

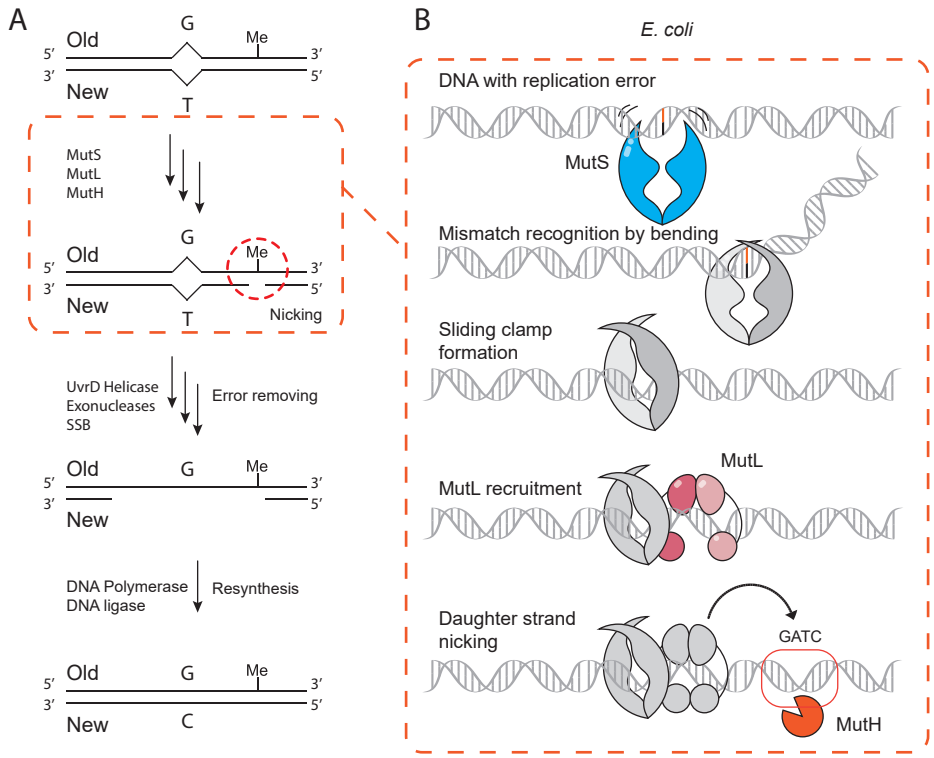


Figure 1.2: **(A)** Overview of the general mechanism of MMR pathway. **(B)** Strand discrimination in *E. coli* pathway.

aspects remain unclear. And finally, only *E. coli* and several other gram-negative bacteria use the Dam-methylation mechanism for strand discrimination, such mechanisms do not exist for eukaryotic cells and most prokaryotes, i.e. lack of MutH[40]. Instead of this pathway MutL or MutLα during evolution obtained a latent endonuclease activity[41,42], which is stimulated by interaction with β-clamp (or in eukaryotic cells with PCNA)[43–45]. Recognition of the nascent strand in eukaryotes seems to occur due to the orientation of the PCNA/MutLα complex[43]. In eukaryotic cells, pre-existing strand differentiation such as strand breaks which appear during removal of misincoporated ribonucleotides may also serve as a marker for nascent strand and help orient MMR system[46,47].

MMR proteins

MutS

MutS is an essential protein in mismatch repair systems that specifically binds to the mismatch DNA and small insertion/deletion loops[48,49]. MutS homologous are evolutionary conserved and presented in all three kingdoms of life (in archaea only in Euryarchaeota)[50,51]. MutS homologs eukaryotes are presented as MSH1-MSH7, mammalian cells contain only MSH2-MSH6 homologs. MutSα and MutSβ participate in MMR (Figure 1.3C,D and E).

MutS is a weak ATPase and it belongs to the ABC (ATP binding cassette) family of ATPase such as Rad50, UvrA, and structural maintenance of chromosomes (SMC) proteins which undergo large conformational changes and use ATP to con-

E. coli	Function	Homologs	Function
MutS	Binds mismathes	MutSα MSH2-MSH6	base-base/small IDL recognition; sliding clamp
		MutSβ MSH-MSH3	small IDL/ large IDL (12 nt) recognition;
MutL	MutH tether, Activates UvrD Coordinates multiple steps	Mutlα MLH1-PMS2	Molecular matchmaker in MMR; PCNA-activated endonuclease;
		Mutlβ MLH1-PMS1	Unknown function
		Mutlγ MLH1- MLH3	Molecular matchmaker in meiosis
MutH	Hemimethylated-GATC endonuclease	-	-
UvrD Helicase II	3'to 5' helicase	-	-
SSB	Binds ssDNA regions	RPA	ssDNA binding/protection,
ExoI, ExoX	ssDNA-dependent 3'to 5' exonuclease	ExoI	Excision of dsDNA
RecJ	ssDNA-dependent 5'to 3' exonuclease	3' exo of Pol δ	Excision of ssDNA
ExoVII	ssDNA-dependent 5'to 3' also 3' to 5' exonuclease	3' exo of Pol ϵ	Synergistic mutator with Exo1 mutant
DNA pol III	Accurate resynthesis of DNA	DNA pol δ	Accurate repair synthesis
β-clamp	Processivity subunit	PCNA	Processivity subunit (sliding clamp)
DNA ligase A	Seals nicks after completion of DNA synthesis	DNA ligase 1	Seals nicks after completion of DNA synthesis

Table 1.1: Comparison of MMR protein functions in *E. coli* and eukaryotic organisms

trol their activity, behaving as chemo-mechanical engines[52]. MutS is a homodimer protein, two monomers form an oval form dimer around the DNA[53,54]. MutS composed of 6 domains (Figure 1.3A). MutS exist in equilibrium between dimers and tetramer via C-terminal domain, whereby the dimeric form alone is able to initiate MMR process[55–57]. The most conservative domain in MutS homologous represents a helix-turn-helix region which is related to Walker A nucleotide-binding site and is important for adenine and magnesium binding[58] (Figure 1.3B). Dimerization of the protein takes place in this ATPase domain. Mutations in Walker A nucleotide-binding site of MSH2 and MSH6 cause a complete MMR defect in vivo[59].

Before MutS recognizes a mismatch, it searches for a high number of matched nucleotides without starting the DNA repair pathway. MutS has a flexible lever and clamp domains[53,62], which possibly can move freely in the absence of DNA or nucleotides (Figure 1.4A). One possible mechanism of the DNA scanning by MutS for mismatches is one-dimensional diffusion (Figure 1.4B)[63,64]. Historically the *E. coli* MutS gene product in "footprinting" experiments showed an ability to specifically bind to a DNA region with single base-base mismatches[48]. Protein has a high binding affinity to the ends of DNA, as well as to mismatch[65]. Detection of the mismatch leads to the DNA kinking by MutS (Figure 1.4C)[54,66]. Two domains are

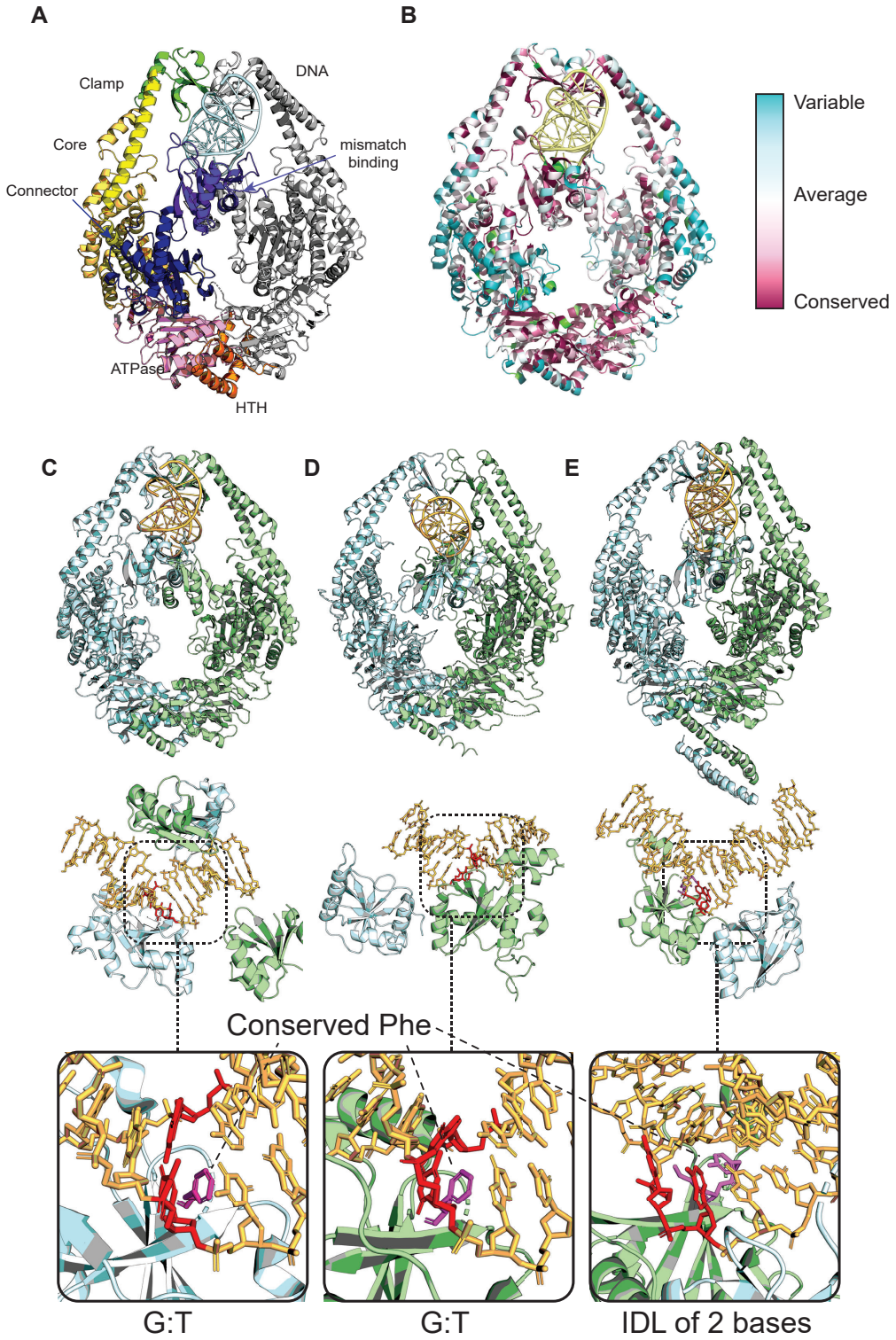


Figure 1.3: **(A)** MutS domains (pdb id: 1e3m): (the mismatch-recognition domain (2-115) is colored dark violet, the connector domain (116-266) dark blue, the core domain (267-443) orange, the levers (504-567) yellow, the clamp domain (444-503) green, the ATPase domain (568-765) pink, and the helix-turn-helix (HTH) domain (766-800) orange). **(B)** MutS domains and the evolutionary conservation of amino acid positions in MutS. Comparison of the structures of *E. coli* MutS (**C**, pdb 1e3m[54]), MutSα (**D**, pdb 2O8B[60]) with mismatched DNA, and MutSβ (**E**, pdb 3THY[61]) with an insertion-deletion loop (IDL) of 2 bases. Created using PyMOL.

involved in MutS mismatch DNA interaction[67]. Clamp domains are closed above the DNA after mismatch recognition and the mismatch domain of only one monomer interacts with a mismatch. The region Phe-X-Glu involved in DNA interaction is evolutionary conserved[68]. The interaction with MM involves highly conserved phenylalanine residue. It wedges into DNA by stacking interaction via aromatic rings with thymine in G/T mismatch[54]. Mutations in Phe36 lead to failure recognition of insertion/deletion mismatch or a G/T base pair mismatch[69]. Regardless of the type of mismatch (except for C/C), DNA is bent by 60 degrees[68].

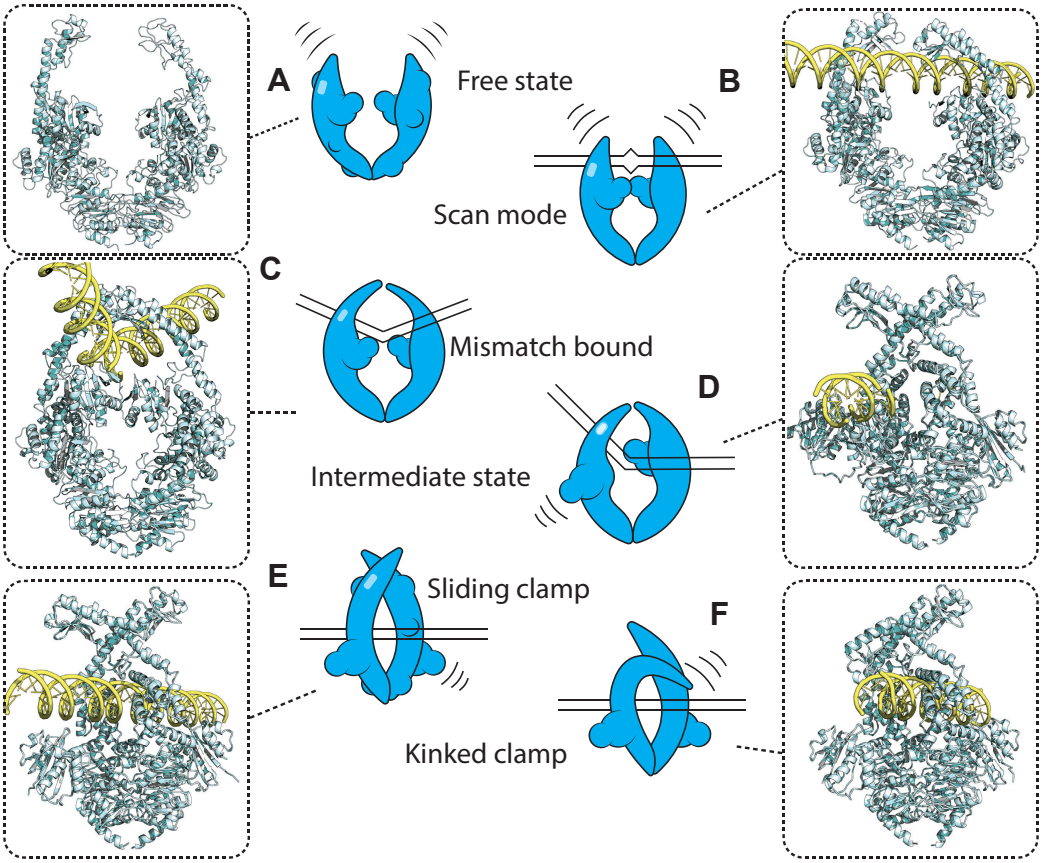


Figure 1.4: Schematic representation of MutS conformational changes and corresponding structures. **(A)** The free state of a protein from *Neisseria gonorrhoeae* with flexible and dynamic domains, which is possible in the absence of nucleotides and DNA (pdb 5yk4)[70]. **(B)** Before launching the cascade of repair reactions, MutS scans DNA bases for the presence of a mismatch (pdb 7AI5)[71]. **(C)** After finding the DNA mismatch, MutS bends the DNA 60 degrees. In this state, both mismatch domains interact with DNA (pdb 1E3M)[54]. **(D)** The intermediate state in which one of the mismatch domains is turned 180 degrees and the second mismatch domain is still interacting with DNA (pdb 7AI7)[71]. **(E)** The sliding clamp state, in which both domains are turned and the connector domains act as the recruitment site of the MutL, and the DNA is lowered into the center of the protein (pbd 5AKB)[55]. **(F)** The kinked clamp state, which resembles the sliding clamp state except for the lever domain that bends and interacts with the DNA (pdb 7AIC)[71].

Upon ATP binding MutS releases the MM[72,73] and MutS forms so-called "sliding clamp" (Figure 1.4E)[32,66] with unusual stability on DNA ($\approx$ 600 seconds[63] that can dissociate in the presence of ATP from unblocked linear heteroduplexes (or with one end blocked) within 15 seconds. When both ends are blocked MutS-DNA complexes

exist with half-lives of 30 min[74]. MutS can be released from DNA at the site of an adjacent single-strand[63]. The releasing mechanism to the date remains unclear. After forming a sliding clamp state ATP hydrolysis is strongly reduced[75,76]. During its conversion from a free state to a sliding clamp, MutS undergoes at least three conformational states[66,67]. The nature of these transient states was investigated in this work. One such state is the intermediate state obtained using cryo-electron microscopy at the end of the oligonucleotide (Figure 1.4D)[71]. The peculiarity of such a state is the intermediate position of the connector domain, which corresponds to the position of this domain in the sliding clamp state. At the same time, the second mismatch domain continues to interact with DNA as in the mismatch state. The sliding clamp ATP-bound state randomly diffuses along with DNA with a contact for MutL interaction[77]. The structure of MutS in sliding clamp looks like a "doughnut" with a channel inside where DNA can be loaded[78]. To obtain the structure of the complex a site-specific cross-linking approach was used and it showed the interface between two proteins[79]. The size and charge of this channel of sliding clamp suggest that it could be a place for DNA in a sliding state mode[78] and it was confirmed by cross-linking experiments and Cryo-Em data[71]. At the same time, large conformational changes to the mismatch and connector domains occur. The connector domains are rotated 180 and create an interface for binding to the N-terminal domain of MutL, while mismatch becomes flexible and is not observed in the structure[78]. The role of ATP hydrolysis by MutS is still unclear, but it is essential for MutL recruitment. Also interesting is a state similar to a sliding clamp, but with a kinked domain interacting with DNA (Figure 1.4F)[71]. The scan mode, intermediate state, kinked clamp structures and sliding clamp with DNA are new findings that are described in more detail in the following chapters.

In eukaryotic organisms, mispaired bases in DNA are recognized by a heterodimer protein MutSα (MSH2-MSH6) and MutSβ (MSH2-MSH3) complexes which have high amino acid sequence similarity with *E. coli* MutS proteins (Figure 1.3C,D and E)[23,80]. MutSα or MSH2-MSH6 targets single base mismatches or one nucleotide IDLs. A series of crystal structures of human MutSa shows that protein similarity in recognition of difference lesions[60]. MutSβ or MSH2-MSH3 recognizes primarily IDLs[81]. MutSα or MSH2-MSH6 complex in the fluorescence anisotropy experiments binds to sites of various mismatches with a similar rate ($K_{ON} \approx 10^7$ M^{-1} s^{-1}), but its selectivity is manifested in differing dissociation rates for mismatch types (for G/T $k_{off} \approx 0.013$ s^{-1})[82]. In the case of MutSα, MSH6 monomer interacts with DNA[60]. In contrast to the MutS and MutSα, MutSβ binds three phosphates in an insertion-deletion loop in the mismatch binding domain of MSH3[61]. After recognition of mismatch and binding of ATP, MutS homologs undergo conformational changes to a sliding clamp state[83,84]. This activated form of MutS (MSH2) can recruit next protein partners MutL homologs (MLH1, PMS1[85]). MSH2 can also interact with PCNA, but the role of this interaction remains unclear[86].

MutL

MutL homologous, similarly to MutS, are evolutionary conserved and presented in all three kingdoms of life (in archaea only in Euryarchaeota) (Figure 1.5B)[51]. Gram-positive bacteria and Euryarchaeota contain only one MutL gene, while in other cells exist 4 homologous MLH1 (MutL homolog 1), PMS2 (postmeiotic segregation increased 2), MLH3, and PMS1. Three heterodimers can be formed MutLα (MLH1-PMS2), MutLβ (MLH1-PMS1), and MutLγ (MLH1-MLH3). In MMR participate

only MutLα and MutLγ[87].

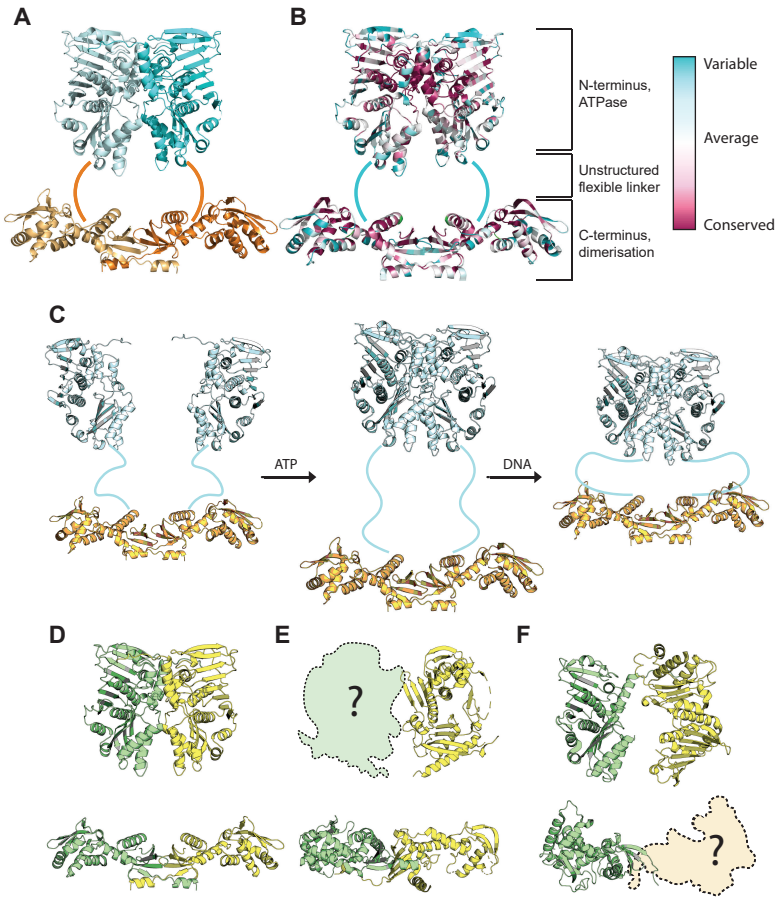


Figure 1.5: **(A)** Structures of MutL CTD and NTD (pdb id: 1B63 and 1X9Z). **(B)** The evolutionary conservation of amino acid positions in MutL. **(C)** Possible mechanism of conformational changes in MutL upon binding to nucleotides and DNA. Comparison of the structures of **(D)** *E. coli* MutL (pdb id: 1B63 and 1X9Z), **(E)** *Saccharomyces* MutLα (pdb id: 4E4W and 3H4L) and **(F)** human MutLα (pdb id: 4P7A and 1H7S)

MutL protein forms a homodimer that contains a flexible linker separating two conserved N-terminal (NTD) and C-terminal domains (CTD) (Figure 1.5A)[88,89]. To date, the full structure of full-length MutL has not been reported. C-terminal domain participates in dimerization and has two subdomains[89], while N-terminal domains tend to dimerization in the presence of ATP and are responsible for DNA binding and ATP hydrolysis[90]. For human homologous structure was also solved for both domains separately (Figure 1.3D and E)[91,92]. The protein dimerization interface of the C-terminal domain for *E. coli* was refined using cross-linking and bioinformatics analysis methods[93]. The N-terminal has an ATP binding domain which is structurally similar to the ATPase domain of members of the GHKL family such as DNA **G**yrase, chaperone heat shock protein (**H**SP) 90, Histidine **K**inase and Mut**L**[52].

ATP binding induces large conformational changes of the protein[90,94]. The binding of ATP or adenylyl-imidodiphosphate (ADPNP) (non-hydrolyzable ATP-analog) in size-exclusion chromatography experiments leads to the compact state of MutL dimer which eluates faster than open state[90,95]. MutL can adopt open or condensed

forms and it is possible to distinguish four different MutL states: "extended", "one-arm", "semi-condensed", and "condensed"[96]. Mutation in the ATP binding site leads to losing the ability to perform the DNA repair functions[42,44]. The protein is thought to undergo conformational changes in the ATP-driven cycle.

The process of MutL binding to DNA is still not enough investigated. Presumably, the protein closes N-terminal domains and compacts onto DNA in the presence of ATP. (Figure 1.3C). MutL independently can binds sequence non-specifically to double-stranded DNA (dsDNA) physiological conditions[90,95]. MutL prefers binds single-stranded DNA (ssDNA) regions, while less to 3' overhangs and dsDNA[97]. The process of binding to DNA stimulates ATPase activity[95]. MutL also interacts with DNA being recruited by MutS and forms the ternary complex of MutS/MutL/DNA in mismatch dependent manner. N-terminal ATPase domain of MutL interacts with MutS connector domain[74,98,99]. A possible DNA binding interface is between the ATP-binding site and the α–β sandwich subdomains of the MutL N-terminal domain. Mutations in this region decrease DNA binding ability and reduce ATPase binding activity[97,100]. MutL is thought to wrap around DNA closing the remaining unbound N-terminal domain with another N-terminal domain activating a second sliding clamp[77]. The interaction of MutL with DNA takes place more efficiently in mismatch dependent manner in the presence of Mg^{2+}[101].

organism	homologue	subunit	domain	PDB code	Reference
E. coli	MutL	-	C-terminal	1X9Z	99
		-	N-terminal	1B63	106
Saccharomyces cerevisiae	MutLα	Mlh1/Pms1	C-terminal	4E4W	101
		Pms1	N-terminal	3H4L	112
Homo sapiens	MutLα	Mlh1	N-terminal	4P7A	102
		PMS2	N-terminal	1H7S	100

Table 1.2: Structures of MutL homologous

MutL interacts not only with MutS but also is involved in the process with other proteins: MutH endonuclease, UvrD helicase, DNA polymerase II, β sliding clamp, exonuclease EXOI, and DNA polymerase[102]. The protein interacts and stimulates MutH latent nuclease activity at physiological conditions in a mismatch-dependent manner[35,103]. MutL can activate MutH independently at low ionic strength and it requires ATP binding[90,104]. MutL-MutH interaction sites were mapped by site-specific cross-linking[105–107]. In eukaryotic cells, the PMS2 subunit of MutLα has a latent endonuclease activity[41].

MutH

MutH enzyme is a monomeric state-specific latent d(GATC) endonuclease that belongs to the type II family of restriction enzymes such as Scu3AI, HinCII, and EcoRv[52]. It can bind and nick hemimethylated and unmethylated GATC sites (Figure 1.6)[35]. Its endonuclease activity can be activated in a mismatch, MutS, MutL, and ATP dependent manner[108]. Namely, MutH interacts with MutL sliding clamp state on the MM region creating a stable diffusion-mediated complex MutS/L/H complex[77]. It increases the ability of MutH to interact with DNA by at least 1000 fold[77]. Several crystallized structures with and without DNA were obtained[88].

UvrD (DNA helicase II)

UvrD is an important component in the hemimethylated MMR pathway in *E. coli*[40]. UvrD is DNA helicase/translocate and it is involved MMR pathway[7] as well in

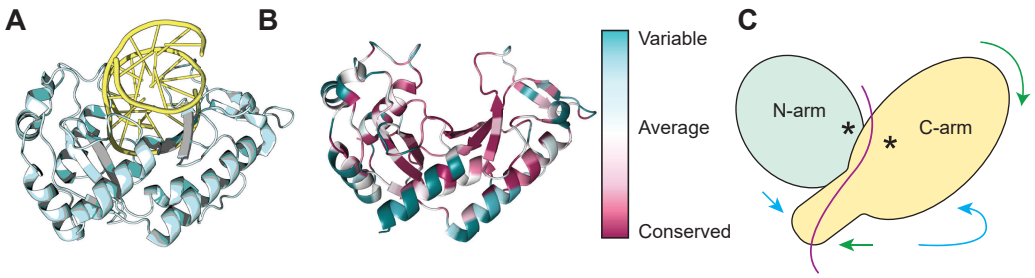


Figure 1.6: **(A)** MutH structure (pdb id: 2aoq). **(B)** The evolutionary conservation of amino acid positions in MutH. **(C)** Simplified diagram of the molecular motion of the C-arm relative to the N-arm (denoted by arrows). When the conformational state is open, binding to DNA occurs. The site of endonuclease activity is indicated by a star[90].

many other systems including replication[109,110], recombination[111], nucleotide excision replication[112]. UvrD gene was characterized independently from mismatch repair system as DNA helicase II, which unwinds DNA strands in ATP dependent manner[113]. MutL stimulates UvrD in a mismatch-dependent manner[34,114,115]. UvrD can be loaded by MutS and MutL at the MutH strand break towards mismatch[36,114]. UvrD DNA unwinding activity in vitro is inefficient and can be observed only at low salt[116]. UvrD unwinds in 3'→5' directionality[116] and can be in the form of the monomer and dimer[117,118], showing low processivity (20 bp)[119]. Unwinding in an MMR-dependent manner is more efficient on a substrate with 2 nicks flanking the mismatch, as compared to substrates containing a single nick or two nicks on the same of the mismatch[120].

Structural studies and fluorescence experiments demonstrated the rotation of flexible domain 2B in the range of 130 to 160°(Figure 1.7)[121–123]. Using optical tweezers with confocal fluorescence microscopy it was shown that UvrD unwinding activity takes place when protein is in a relatively close conformational state[119]. Another evidence for active "close" state is cross-linking of similar structural helicase Rep which resulted in superhelicase[124]. At the same time, the role of the 2B domain remains unclear, since the removal of this domain in structural similar Rep helicase does not lower the helicase activity of the protein, but rather makes Rep helicase more active[125,126].

The signal pathway between MM and GATC sites

To date, it is known that the mismatch repair system in *E. coli* includes eleven proteins mentioned above and two DNA sites (GATC and MM). Despite years of research the coupling mechanism of mismatch recognition with strand discrimination (interaction between these MM and GATC distal sites) remains unclear and exists many models[7,127,128]. It was shown that the repair mechanism is bidirectional[20] and the distance between GATC site and MM shouldn't exceed 1000 bp[17]. It was accepted that MutS recruits MutL to the DNA[32], but the interaction between MM and GATC sites is the subject of much debate. Following modern concepts of system, the mechanisms can be divided according to the principle of interaction between these two sites i.e. activation of MutH on the cis and trance pathways. In the case of transactivation, the signal comes from the mismatch to the GATC site by looping

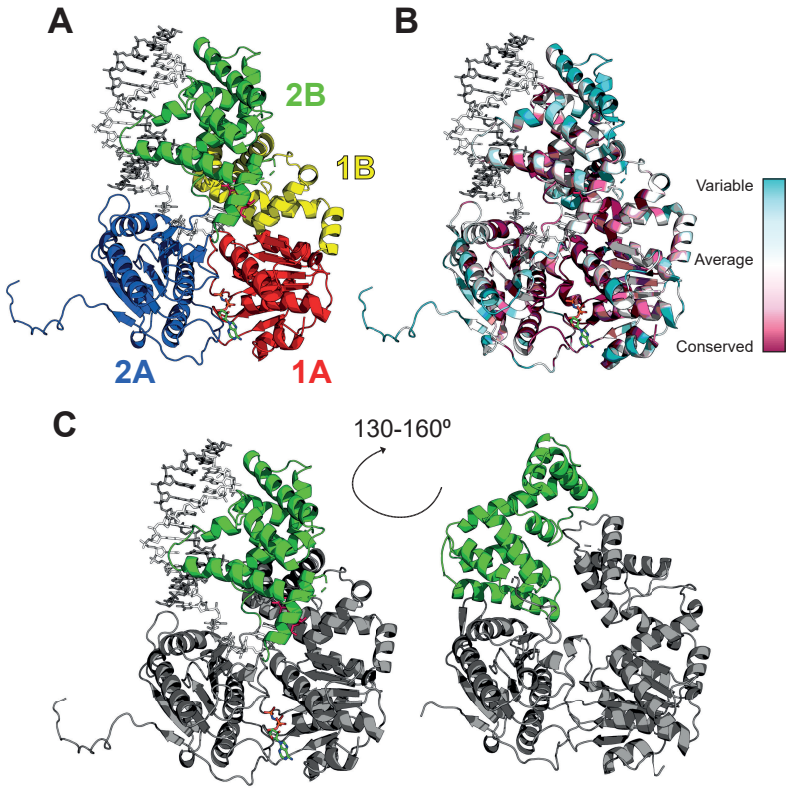


Figure 1.7: **(A)** UvrD domains in the crystal structure (pdb id: 2is4). **(B)** Evolutionary conservation of amino acid positions in UvrD. **(C)** Conformational changes in UvrD (pdb id: 2is4 and 3lfu). Structural studies and fluorescence experiments demonstrated the rotation of flexible domain 2B in the range of 130 to 160°[121–123]. Using optical tweezers with confocal fluorescence microscopy it was shown that UvrD unwinding activity takes place when protein is in a relatively close conformational state[119]. Another evidence for active "close" state is cross-linking of similar structural helicase Rep which resulted in superhelicase[124]. At the same time, the role of the 2B domain remains unclear, since the removal of this domain in structural similar Rep helicase does not lower the helicase activity of the protein, but rather makes Rep helicase more active[125,126].

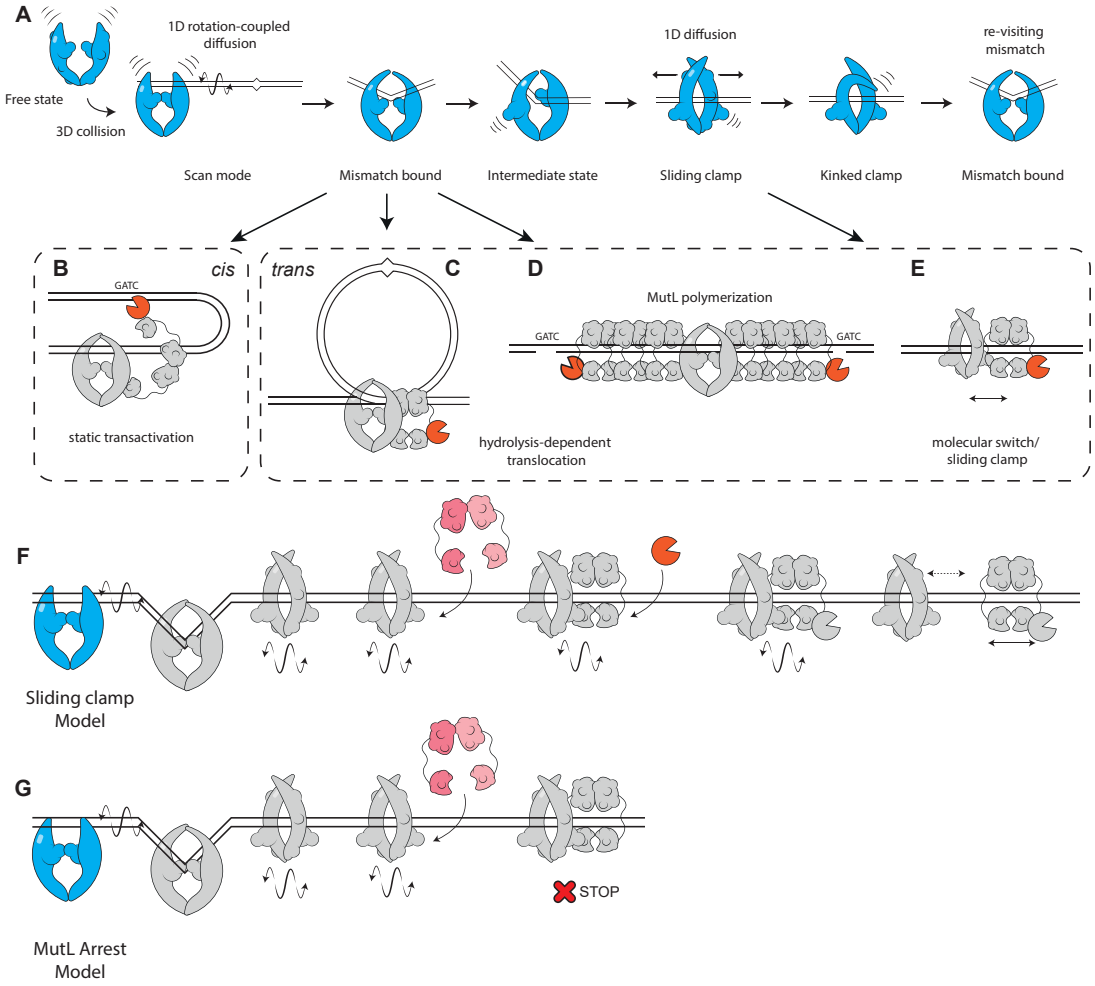


Figure 1.8: **(A)** MutS adopts multiple conformational changes preceding subsequent reactions in the MM pathway. Competing MMR models: **(B)** "static transactivation", now disfavored **(C)** "hydrolysis-dependent translocation", **(D)** "MutL polymerization" and **(E)** "molecular switch/sliding clamp" models. **(F)** After the conformational changes described above, MutS is able to recruit MutL and MutH. The MutL-MutH complex can move along DNA without MutS. **(G)** The newly proposed model in which MutL stops MutS sliding clamp. (See text for references).

or bending of the DNA and it's also called stationary model (Figure 1.8B)[74,129]. The signal in trans interaction is transmitted through interaction between proteins and can be divided into several further discussed models: translocation (Figure 1.8C)[32,84], polymerization (Figure 1.8D)[130,131], sliding (Figure 1.8E)[32,63,77].

The evidence for the translocation model is electron microscopy data and atomic force microscopy images where MutS mediates similar α-shaped DNA loops[73,132]. This is historically the first model and until now it can often be found in textbooks on molecular biology (Figure 1.8C).

Support for transactivation model which belongs to the trans group are experiments with MM-site and GATC site, belonging for two separate substrates[74,91,98,129], although with two sites on the same substrate, the incision rate is much faster. According to this model, it is unclear whether the MutS sliding clamp, the structure of which was previously shown, could be essential for the repair mechanism acti-

vation. The model assumes that MutS binds to the mismatch during the recovery process and after detection, it recruits and activates MutL. While MutS and MutL are associated with a base mismatch, MutH is activated by bending DNA in space (Figure 1.8B). Thus, induced by MutS, MutL, and MutH, a DNA loop is formed between the base mismatch and the GATC recognition sequence. This model is similar to that of gene expression transactivation between transcription factor and RNA polymerase. It was shown, that in the presence of roadblocks, the nicking ability is significantly reduced. Especially using MutL with a shortened linker, the residual activity completely disappears[133,134]. Although this model is questioned, the new data still makes this model relevant[135].

MutL polymerization model of the cis group models suggests cooperative binding of MutL (Figure 1.8D). Evidence of this interaction are AFM data of MutLα[131] and in vivo colocalization MutL forming fluorescent loci were 2.7 times higher than MutS loci fluorescence intensity[130,136]. The same difference in stoichiometry was also observed for MutLα and MutSα[137]. smFRET experiments showed that the MutL can trap MutS, preventing MutS conformational changes and the transition to a state of sliding clamp[138]. The polymerization model describes the cis mechanism that takes place along the DNA helix. After MutS binds to a mismatch, more and more MutL molecules are recruited to the DNA. Polymerization is a bidirectional process and ends only after reaching the GATC recognition site, where the last MutL recruits and activates the latent endonuclease activity of MutH, which then introduces nicks to the DNA.

The next cis model is the sliding clamp model, where the MutS forms a sliding clamp in the ATP-dependent manner (Figure 1.8E). The idea arose from experiments with blocked or circular DNA substrates in which there are no ends and there is good evidence from many labs that MutS and MutL form a complex that slides along with DNA and can activate MutH nicking activity in an ATP dependent manner[32,74,139]. Recent single-molecule FRET (smFRET) experiments have also shown in ATP dependent manner the diffusion of MutS, MutS-MutL, as well as the ternary complex MutS-MutL-MutH along DNA (Figure 1.8F)[38,66,77], a structure of a MutS sliding clamp with cross-linked N-terminal domain of MutL was also obtained[71,78]. Recent single-molecule data have shown the existence of a stable MutL-MutH complex moving along DNA, as well as the possibility of activating the latent endonuclease activity of MutH through roadblocks[77,134], may indicate that the MutL-MutH complex is an active complex responsible for introducing nicks on GATC site, and MutS in this theory play tole of mismatch recognizer and MutL downloader.

A variation of the sliding clamp model is the so-called MutL arrest model where MutL stops the movement of the sliding clamp along DNA after binding to MutS (Figure 1.8G)[140]. These patterns may differ due to biological differences in organisms, as in *E. coli* a methyl-targeted pathway is used, while in other organisms, endonuclease activity is achieved by activation of MutL by PCNA or beta clamp, requiring a less mobile MutS-MutL complex[141].

Trapping molecular machinery for structural and functional studies

The highly dynamic nature of this machineries hinders obtaining or characterization intermediates states of this machinery. Investigation of low-affinity complex or transient states is a difficult task for traditional structural methods. Interactions with affinity constant lower than 10^6/M are not strong enough to withstand the washing step to isolate the corresponding complex[142]. Protein engineering tools become a powerful instrument in biochemistry, structural and molecular biology, which helps to investigate complex systems of molecular objects in the tube and solve structures by X-ray crystallography, cryogenic electron microscopy (Cryo-EM), mass-spectrometry, and nuclear magnetic resonance (NMR) techniques[143]. These engineering approaches can be divided into three large groups: chemical modification, molecular-biological tools, and the combination of these two methods[144]. The chemistry of cysteine in comparison with modification of the other natural amino acids and native chemical ligation is the most common method of chemical modification[144].

Site-specific modification

The thiol group which is a part of cysteines in proteins is less common than other amino acid residues such as amines or carboxylates, thereby limiting the number of reactive sites in the molecule and making the process of bioconjugation specific[142]. The single-cysteine approach is a very effective tool for the investigation of transient dynamic complexes using cross-linking or for analysis of the kinetic and low-resolution conformation model of molecular machines using specifically labeling with fluorophours[145]. There are three major groups of cysteine chemistry: oxidations (mostly disulfide formation), cysteine alkylations (cross-linking using iodoacetamide and maleimide reagents), and desulfurization[144].

To obtain cysteine-free variants it's necessary to replace six endogenous cysteines in MutS with alanine, serine, or valine (Figure 1.9A), and seven cysteines in MutL are replaced by leucine, phenylalanine, or serine[145]. The cysteine-free variants of MutS, MutL, and MutH were obtained earlier[57,105]. Further amino acid residues at designed positions can be replaced with cysteine residues to obtain single cysteine variants. These variants can be modified with thiol-specific reagents such as maleimide, 2-mercaptopyridine, methane thiosulfonate (MTS) chemistry (Figure 1.10). These modified proteins can be used in structural, biophysical, or biochemical assays.

Recently, many new methods have been developed that allow for the modification of the unique chemical groups of nucleic acids. Modification of discrete sites can be done on bases, sugars, and end and internal phosphate groups (Figure 1.9B). These chemical procedures can be used by adding a functional label to an individual nucleotide (eg, a fluorophore or reactive group), which is later incorporated into the oligonucleotide chain by chemical synthesis[142].

Reactive groups capable of binding to molecules containing an amino acid residue are the most common functional groups present in cross-linking or modification reagents. The amino acid modification process is used to bioconjugate almost all protein or peptide molecules, as well as modified macromolecules such as DNA.

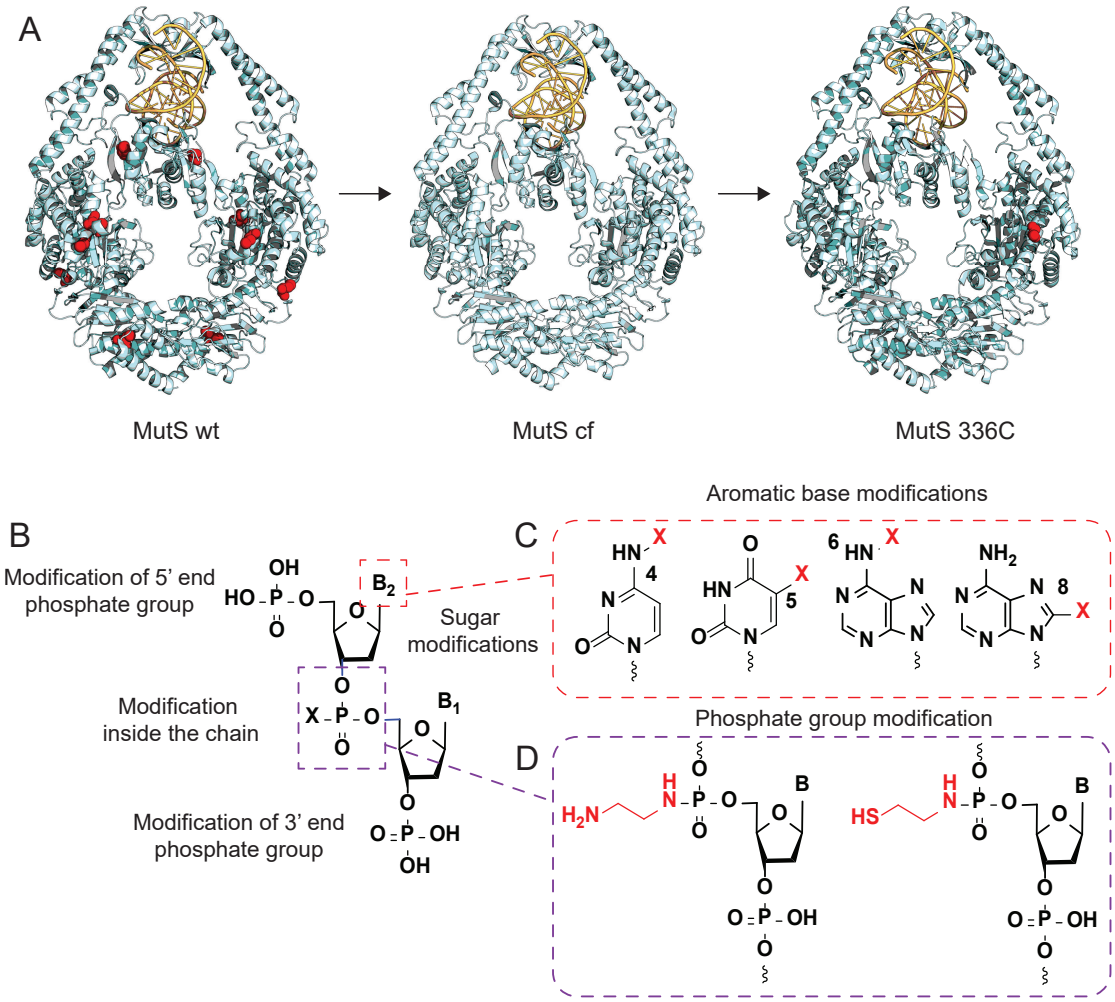


Figure 1.9: **(A)** Scheme for generation single cysteine MutS336C variant from wild type protein for MutS (pdb id: 1e3m). The red spheres in the first picture represent natural cysteines (e.g. C93). The next step is removing all the cysteines (e.g. C93A). The last step involves the precise design of the cysteine position (e.g. A336C). **(B)** General scheme of nucleotide chemical modification. **(C)** Possible positions for modification purine and pyridine bases. **(D)** Modification examples of phosphorus group.

These reactions proceed quickly and in high yield with the formation of stable amide or secondary amine bonds[142]. The addition of an amino group is a routine method in the synthesis of oligonucleotides and is used as a specific site for subsequent modification with a cross-linker or fluorophore[142]. One of the common chemistry reagents is N-hydroxysuccinimide (NHS) ester which reacts in high yield with a primary amine forming a stable amide bond (Figure 1.9C and D). Reagents containing an NHS group react with nucleophiles to release the leaving NHS group to form an acylated product (Figure 1.10A)[142].

Site-specific cross-linking

Using disulfide cross-linking, low populated states can be trapped and stabilized for further research. Using this method, the structures of DNA glycosylase during scanning DNA and DnaG primase were solved[146]. In comparison with the disulfide bridge formation approach, site-specific reaction with cross-linkers gives more flexibility, by changing different parameters such as chemistry, length of linker, reaction rates, and timing of adding reagent. By developing and implementing site-specific chemical cross-linking approaches it was possible to trap selected states and obtain high-resolution structural information on transient dynamic and/or non-specific protein-DNA complexes. Site-specific cross-linking was successfully used to create a super helicase[124]. Site-specific cross-linking and single cysteine variants approach was applied for investigation of intermediate transient or unstable states of the mismatch repair system. Purification to homogeneity of the covalently linked complexes and subsequent biochemical and biophysical characterization were employed to validate the functional state of the complex before structural analysis. This combination of specific cross-linking with FRET[147], X-ray scattering[55], X-ray crystallography[78], Cryo-EM, mass-spectrometry[105,106] allows studying transient states by atomic resolution methods.

Reactive groups that react with sulfhydryl-containing molecules are the second most common functional groups after amino-containing groups for cross-linking or modification. Maleimide group is often used in homocross-linkers, as well as when modifying the substrate with a probe (for example a fluorophore). The double bond of maleimides derivatives undergoes an alkylation reaction with cysteines groups forming a stable thioether bond (Figure 1.10B). Molecules containing a thiol group are able to participate in disulfide exchange reactions with another thiol. During disulfide exchange (interchange) thiol attacks the disulfide, breaking the -S–S- bond, following the formation of a new disulfide bond (Figure 1.10C). Reducing agents reduce these conjugates bonds. A pyridyl dithiol is the most popular group for thiol-disulfide exchange. This group readily reacts with a free sulfhydryl with a high yield and a single disulfide product. The pyridyl disulfide has a leaving group that becomes non-reactive and does not participate in further reactions (Figure 1.10D). Methanethiosulfonate (MTS) derivatives are a sulfhydryl-reactive compound for reversibly sulfonation thiol-containing modification of the molecules. Reaction MTS chemistry with the free thiol group results in the disulfide group formation (Figure 1.10E). Disulfide bonds can be restored with reducing agents such as dithiothreitol (DTT), 2-mercaptoethanol (2-ME), or Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) to the original sulfhydryl group. Thiol groups can be irreversibly modified and blocked for subsequent reactions with N-ethylmaleimide (NEM)[142].

Using the intermolecular site-specific method, the dimerization interface of the N-terminal domain of MutL was reinvestigated, which was formed on the surface of

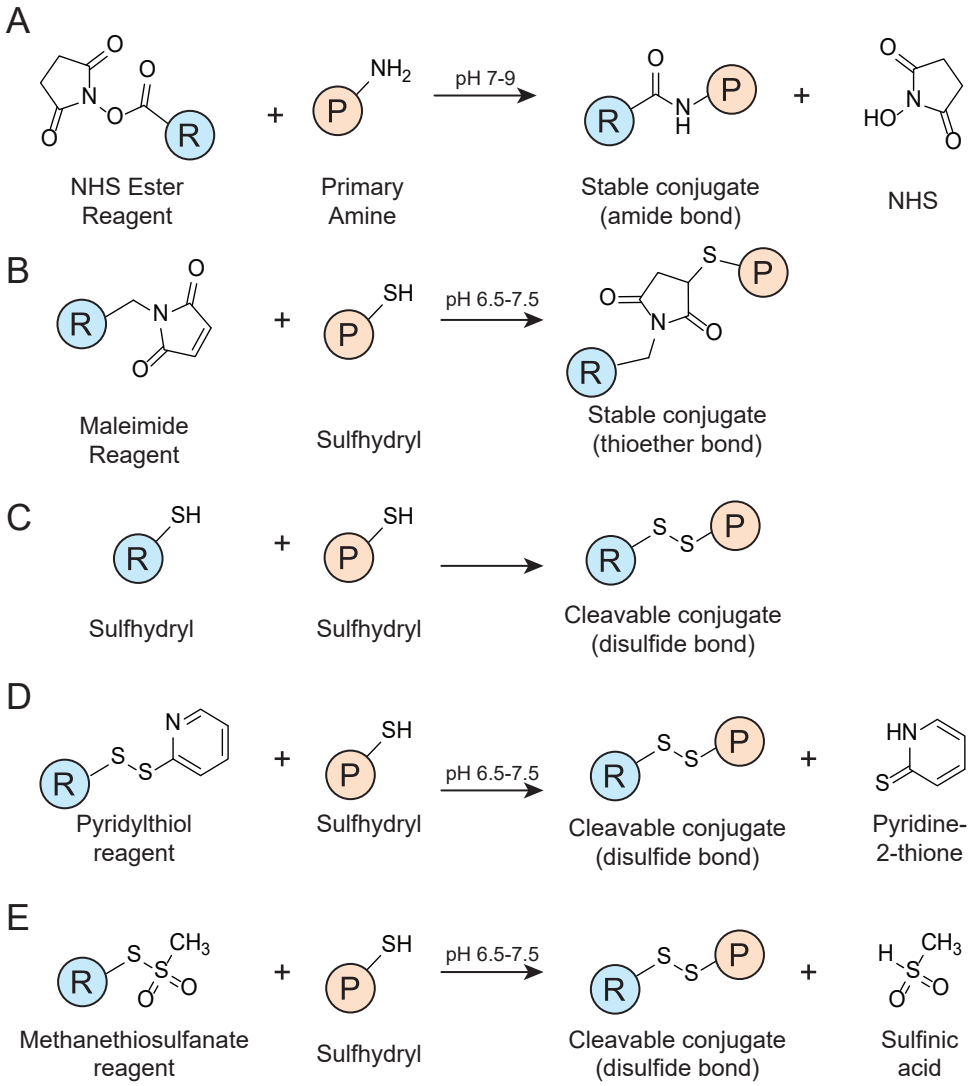


Figure 1.10: **(A)** Reaction scheme of NHS ester for chemical conjugation with a primary amine. **(B)** Reaction scheme of maleimides for chemical conjugation with a thiol group. **(C)** Thiol–Disulfide Exchange Reaction. **(D)** Pyridyldithiol reaction scheme for cleavable (reversible) chemical conjugation to a sulfhydryl. **(E)** Methanthiosulfanate reaction scheme for cleavable (reversible) chemical conjugation to a sulfhydryl. (See overview[142])

the external subdomain, instead of the internal subdomain originally proposed[93]. For characterization of the dimerization and tetramerization of MutS, a single cysteine variant was designed for the formation and characterization of a stable tetramer using cross-linking[55]. Single cysteine MutS and MutL mutants and maleimide cross-linker were used to trap a dynamic complex for functional characterization[79], which resulted later in the crystal structure of a new MutS state[78]. Homobifunctional cross-linking in combination with mass-spectrometry provided data of interface between MutL and MutH[105,106]. MutS protein was successfully trapped using a disulfide bond to the end of DNA[148]. Trapping of MutS near mismatch via heterobifunctional cross-linker showed bending activity of cross-linked species[147]. Cross-linking was used for mapping DNA MutL interfaces. Several MutL single cysteine variants were tested in DNA binding by cross-linking and FRET[149].

Site-specific labeling

Fluorophores can be used for the modification of biomolecules. The reactive group attached to the fluorophore covalently binds to the target group of biomolecules. This conjugate becomes a highly sensitive reagent for fluorescence microscopy techniques. Each fluorescence molecule absorbs photons of energy at one wavelength and emits photons with another wavelength. This process is called excitation and has a quantum nature[142]. Fluorescence is a physical process in which a radiative transition of the excited state occurs from the lowest singlet vibrational level S_1 to the ground state S_0. After the absorption of light, part of the energy received by the system is consumed as a result of relaxation (Figure 1.11A). A part can be emitted as a photon of a certain energy. The fluorescence spectrum is shifted relative to the absorption spectrum towards longer wavelengths. This phenomenon has received the name "Stokes shift". It is caused by nonradiative relaxation processes. As a result, part of the energy of the absorbed photon is lost, and the emitted photon has less energy, and, accordingly, a longer wavelength (Figure 1.11C)[150].

Intermolecular changes, protein-protein or DNA-protein interactions can be monitored by Fluorescence methods such as anisotropy and FRET. FRET is the mechanism of energy transfer between two chromophores (from donor to acceptor), which occurs without intermediate emission of photons and is the result of dipole-dipole interaction between donor and acceptor. The FRET method is used when the distance between the donor and the acceptor dyes changes during the conformational changes (Figure 1.11B). A characteristic feature of this process is the quenching of the donor fluorescence and the appearance of a longer wavelength acceptor fluorescence (Figure 1.11C). The rate of this process depends on the distance between objects and decreases as r^{-6}, which makes it possible to measure the distance both between two molecules and between labels in one macromolecule (Figure 1.11D). The effective distance at which the transition rate is 50% of the maximum is called the Förster radius. For most systems, its value is 20-50 Å.

For the successful application of FRET for biomolecules investigation, the following requirements must be observed in the design of the experiment. First of all, the distance between the fluorophores attached to the objects needs to be less than < 10 nm. Secondly, donor emission and acceptor excitation spectra need to be overlapped. Last but not least, the fluorophores need to have the correct relative spatial orientation of donor/acceptor.

Site-specific single cysteine labeling technique was used for investigation of MMR system in particular for monitoring conformational changes in mismatch repair pro-

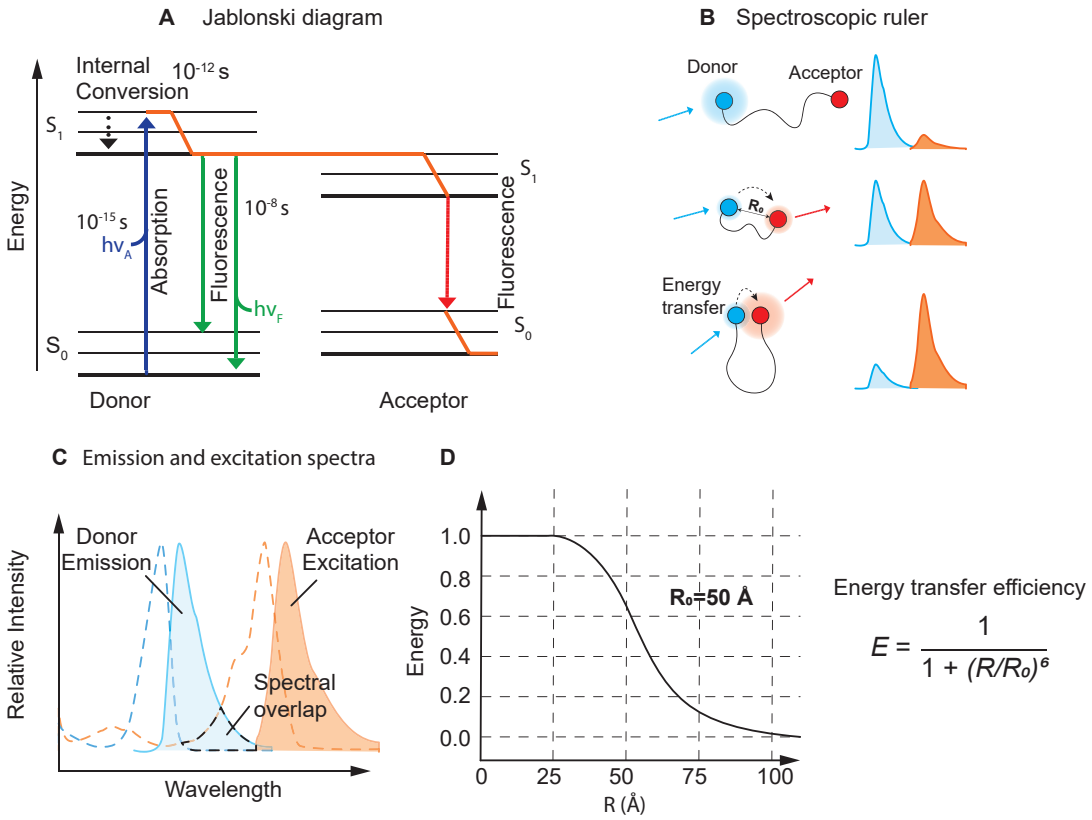


Figure 1.11: **(A)** The schematic processes of light absorption and fluorescence are shown in the Jablonsky diagram. **(B)** Distance between two fluorophores becomes shorter and the process of energy transfer takes place, which can be observed on the spectra. The signal for the donor decreases and the acceptor signal increases when the donor is excited. **(C)**. Donor emission and acceptor excitation spectra need to overlap **(D)** FRET efficiency dependence on the donor-acceptor distance. The figures were adapted from[150].

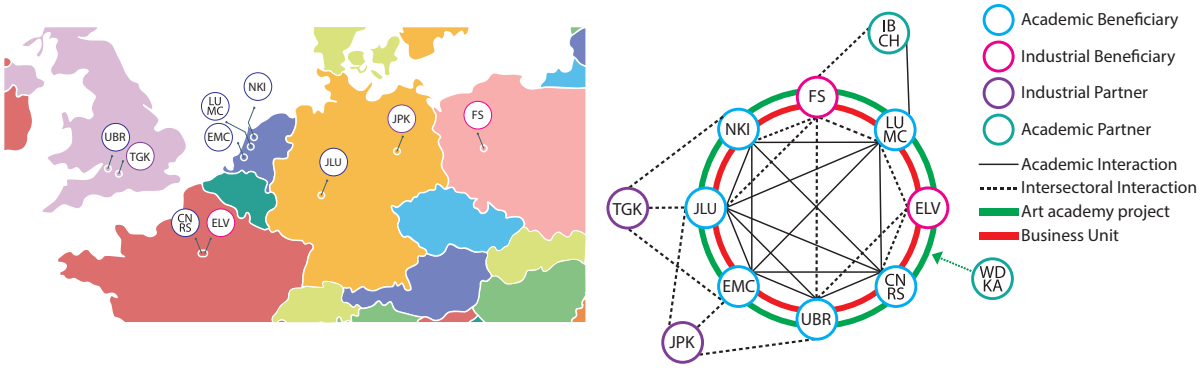


Figure 1.12: **(A)** Location of laboratories and companies participating in the project. **(B)** The scheme of interaction of project participants.

teins[78], single-molecule studies[151], probing DNA center assay[149], DNA bending[147], recruitment MutL and MutH, UvrD unwinding activity. Fluorescence methods allow following Kinetik of the complex formation, to see which proteins are simultaneously present, how they interact with each other or with DNA and investigate intermolecular changes of protein domains.

Goals of the study

This work is part of a larger project within the Innovative Training Network "DNARepairMan" that belongs to Marie Sklodowska-Curie's actions. 12 participants from academia, industry, and the creative sector have established a "DNARepairMan" Network. The Network consists of six academic beneficiaries (Centre National de la Recherche Scientifique (CNRS), Erasmus University Medical Center (EMC), Justus-Liebig-Universität Gießen (JLU), Universiteit Leiden (UL), Netherlands Cancer Institute (NKI), University of Bristol (UBR)), two industrial beneficiaries (Elveflow (ELV), Future synthesis (FS)) and four partner organizations (academic- Institute of Bioorganic Chemistry of the Polish Academy of Sciences (IBCH), Willem de Kooning Academy (WDKA), industrial - JPK, TGK) (Figure 1.12). Within the platform, 12 international PhD students, their supervisors, and companies´ representatives are working together as a great team to unravel the missing points in molecular properties and mechanisms of the switches and motors involved in two canonical DNA repair pathways. A partnership within this community is not only important to provide answers to the above-mentioned questions, but also to develop new methods which will allow researchers from other fields to use them for special purposes. In 2015, the Nobel Committee awarded the Nobel Prize to three distinguished scientists who have made major contributions and were pioneers in this field. A better understanding of such processes will lead to the development of future targeted cancer treatments that depend on an understanding of the molecular mechanisms underlying these diseases. In these processes, proteins are combined into complex dynamic systems that recognize and remove errors in the genome.

The work presented in this thesis is aimed to investigate the nature of conformational changes of transient or active MMR complexes using the single-cysteine approach. In particular, site-specific cross-linking gives structural information on

relevant transient complexes and conformational intermediates. Besides, the new cross-link position in combination with functional characterization will allow trapping a new transient state on the repair pathway. Trapped complexes help to perform high-resolution analysis of otherwise low-populated dynamic states and identify functional information of the obtained structures. In combination with fluorescent methods, this approach helps to study the dynamics of protein interactions and create new methods for investigating the activity of complexes.

In **Chapter 2** I concentrated on the validation of the sliding clamp structure with cross-linked MutL N-terminal domain, namely probing the DNA-binding center of MutL NTD. At the time of writing this chapter and article, there was no structure with DNA in the slide clamp. For these purposes, it was necessary to develop a new fluorescent assay for MutL binding to DNA. To make the method universal for all proteins and could be applied in further studies of the DNA repair system, it was decided to develop a FRET method based on single-cysteine proteins and stained DNA. (Figure 1.13.A).

In **Chapter 3** I focused on the investigation of MutS conformational changes, by site-specific cross-linking of MutS transient state to modified DNA to create stable complexes. To established a functional analysis of the MutL recruitment activity to DNA by MutS in an MM-dependent manner, it was necessary to refine the FRET-based assay that was successfully applied in **Chapter 2** to monitor MutL binding to DNA (Figure 1.13.B). The combination of cross-linking approach with a fluorescence-based method based on site-specific modifications of single-cysteine variants allows functional analysis of the transition state of the protein and answers a few fundamental questions. Does MutS have to leave the mismatch site to recruit MutL? Can MutS recruit MutL when the mismatch domain is close to the DNA, or does the protein require large conformational changes in this domain?

In **Chapter 4** I planned to catch the active transition state of the MutS and show the functional relevance of the sliding clamp state. It was necessary to rethink the protein cross-linking approach. Using the precise design of single-cysteine positions in MutS, as well as structural and fundamental knowledge of the putative DNA position, II decided to create a single-cysteine variant that could only be cross-linked in the active sliding clamp state. This method allows to trap the active state of the protein on DNA and in combination with other biochemical (Figure 1.13.E) and biophysical methods (Figure 1.13.D) (developed in **chapters 2** and **3**), to check its functional activity. Thus, to show that sliding clamp state is required for the activation of latent endonuclease MutH. The cross-linking method developed in **chapter 3**, and the MutS and DNA point designs were used to test structural assumptions. It was not clear whether the rotation of the MutS domains was caused by interaction with MutL or by ATP hydrolysis. For these purposes, a new single-cysteine variant of MutS was developed and a special design of oligonucleotides with a chemical base modification was used.

In **chapter 5** I present an approach for trapping MutL-MutH active complex. Here II further developed site-specific cross-linking assay used in previous chapters for chemical trapping of other MM proteins that allow structural and functional analysis. The method was planned to be optimized for small-scale analytical cross-linking for functional studies and preparative cross-linking for structural studies, followed by purification of the complex. These studies were carried out in collaboration with specialists in cryo-electron microscopy at Leiden University. Cross-linking

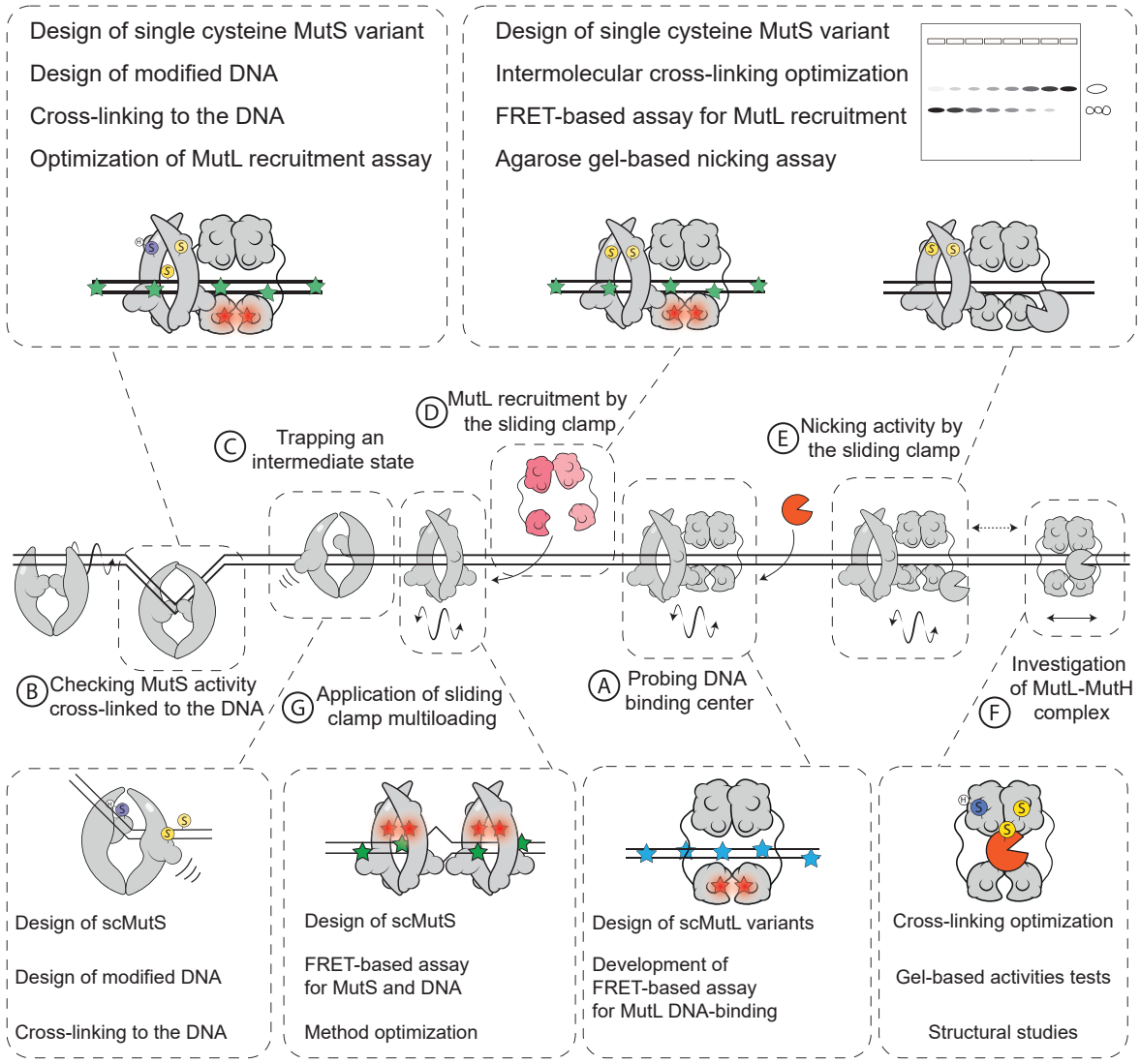


Figure 1.13: The layout of this study. **(A)** For probing the DNA-binding center of MutL, a fluorescent method was developed to establish the ranking of different positions of the protein to the DNA center. **(B)** MutS was attached by covalent cross-linking to DNA to test the MutL recruitment ability when MutS cannot leave the DNA mismatch. **(C)** A single-cysteine variant of MutS was designed to be trapped to DNA in a transition state in which the domains would rotate 180 degrees in the presence of ATP and the absence of MutL. **(D)** The single-cysteine variant of MutS was designed to covalently close the lever domains in the sliding clamp state and test its MutL recruitment activity as well as **(E)** its ability to activate the latent endonuclease activity of MutH in the presence of the trap. **(F)** Single-cysteine variants of MutL and MutH were analyzed by site-specific cross-linking to create an active complex and subsequently studied by structural methods. **(G)** Using a multi-loading sliding clamp, a simple and efficient method was proposed to test the quality of DNA, as well as to study the MutS behavior on DNA.

of MutL to MutH with triggering factors, such as different nucleotides, can shed light on the existence of a new functional active complex, which can result in a new significant CryoEM structure (Figure 1.13.F).

In **chapter 6** I used MutS conformational changes knowledge obtained earlier in **chapter 4** and FRET-based assay developed in **chapters 2** and **3** to establish a novel fluorescence-based assay for DNA quality control (Figure 1.13.G). Moreover, during my secondment at the University of Bristol, II establish a fluorescence-based method for monitoring UvrD unwinding activity in an MM-dependent manner based on fluorescently labeled single-stranded binding proteins (SSBs). This assay was actively used in the development of a DNA quality control tool based on labeled MutS.

In the last **chapter 7**, I summarize the results obtained in Chapters 2-6.

Chapter 2

Probing the DNA-binding center of MutL protein from *Escherichia coli* mismatch repair system by FRET

Vladislav Kunetsky[1], Maya Monakhova[2], Alexandra Ryazanova[2], Pingping Li[1], Evgeniy Shilkin[3], Olga Kisil[4], Desirazu N. Rao[5], Tatiana Oretskaya[2], Elena Kubareva[2], Peter Friedhoff[1]

1 Institute for Biochemistry, FB 08, Justus Liebig University, Heinrich-Buff-Ring 17, D-35392, Giessen, Germany
2 Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Leninskie Gory 1, 119991, Moscow, Russia
3 RAS Institute of Molecular Genetics, Kurchatova Sq. 2, 123182, Moscow, Russia
4 Gause Institute of New Antibiotics, B. Pirogovskaya 11, 119021, Moscow, Russia
5 Department of Biochemistry, Indian Institute of Science, Bangalore, 560012, India

Abstract

The central role in the coordination of various stages of MMR is attributed to the MutL protein which is the coordinator of various stages in the MMR process. MutL receives a signal from MutS and controls the nuclease activity of MutH. MutL is involved in other processes taking place in later steps, such as the DNA unwinding activity of UvrD. To date, there is no full-size MutL bound to DNA structure. For probing the DNA binding site of MutL, we proposed a novel FRET method. The method is based on the precise design of single-cysteine variants of full-length MutL, covalently modified with a fluorophore. Six variants of single-cysteine proteins in the N-terminal domain of MutL were prepared for testing the method. The results of binding of MutL to DNA clearly show that the highest intensity of the FRET signal was observed at positions 251, 218, thereby indicating the proximity of these amino residues to the DNA center. The results confirm the structure of the newly obtained MutS-MutL complex. Our model was also validated independently by another group using site-specific cross-linking of the same cysteine positions to modified DNA. Positions 218 and 251 showed the highest reaction yield. The fluorescence method can be improved and used in further studies of the DNA repair system, in particular, the ability of MutS to recruit MutL.

Introduction

DNA mismatch repair (MMR) is an evolutionarily conservative pathway that detects and repairs DNA mismatches resulting from DNA polymerase replication errors or the DNA recombination process[7]. The main components of the system in *E. coli* are homodimers MutS and MutL. The process starts with mismatch recognition by MutS and following conformational changes induced by ATP on DNA, which allow MutL recruitment[32]. The central role in the coordination of various stages of MMR is attributed to the MutL protein which is the coordinator of various stages in the MMR process; in particular, MutL receives a signal from MutS and controls the nuclease activity of MutH[31,95]. MutL is involved in other processes taking place in later steps, such as the DNA unwinding activity of UvrD[114].

There are two opposite models of the MutL binding and activation by MutS[141]. In the first model, the binding activity of MutL to MutS occurs at the mismatch site preventing dissociation of MutS from this site[135]. In the second model, MutS forms a sliding clamp that slides along the DNA and is capable of recruiting MutL[77,138,140]. The process of binding MutL to a sliding clamp state, as well as conformational changes in this complex, remains unclear[77]. Single-molecule methods have shown direct evidence that MutL can bind to MutS in a sliding clamp state, forming a long-lived (τ_{on} = 851 s) complex on DNA[77]. The structural data demonstrate that the NTD of MutL interacts with the connector domain of the MutS sliding clamp state[78]. There is no DNA in the structure, but indirect evidence suggests that DNA passes through a hole in the center of MutS. The process of MutL binding to DNA is still not fully investigated. Mutations of the positively charged amino acid residues in the α-β sandwich subdomain of the MutL-NTD decrease MutL affinity for DNA and reduce its ATP-binding activity[95,97]. Investigation of the MutL interaction with DNA is complicated because of the flexible nature of MutL linkers and low affinity to DNA[152]. A potential approach to analyze the interaction between MutL and DNA

is to probe MutL-DNA contacts by the FRET method. FRET is a spectroscopic method for detecting the proximity between donor and acceptor fluorophore-labeled molecules in the distance range of 30 to 70 Å in solution under physiological conditions[153]. Single-cysteine MutL variants were modified by fluorophores. A pair of fluorophores, SYTOX Blue and ATTO 647N was used. Using FRET assay the correlation of signal of several positions in the subdomains of MutL-NTD to DNA with the distance between these positions in the structure was shown. Our data were confirmed by the cross-linking results. This approach can be further developed and used for monitoring MutL recruitment to DNA by MutS in an MM-dependent manner.

Materials and methods

DNA

All oligonucleotides (HPLC grade) were synthesized by Eurogentec (Seraing, Belgium). Two complementary oligonucleotides with G/T mismatch in the center were diluted in water (Table 2.1). As a result of mixing two oligonucleotides, the final concentration of the duplex was 10 μM. The duplex was annealed in a "Biometra T-personal PCR machine" at a temperature of 95 °C with a gradient of a decrease in temperature by 5° C every five minutes.

Strand	length	Sequence	Modification
Bottom	59	5'-TGA-AGC-TTA-GCT-TAG-GAT-CAT-CGA GGA-TCG-AGC-TCG-GTG-CAA-TTC-AGC- GGT-ACC-CAA-TT-3'	-
Top	59	5'-/Dig/AAT-TGG-GTA-CCG-CTG-AAT-TGC -ACC GAG-CTT-GAT-CCT-CGA-TGA-TCC- TAA-GCT-AAG-CTT-CA-/Dig/-3'	5' and 3' Digoxigenin

Table 2.1: Oligonucleotides for probing DNA center

Expression

All the proteins used in the present work were recombinant and carried the N-terminal His_6-tags. The plasmids for protein expression were constructed earlier[105]. The proteins were expressed in *E. coli* and were purified by affinity chromatography on Ni-NTA agarose followed by size-exclusion chromatography on Superdex 200 10/300GL column as described elsewhere[105]. Protein purity was checked by SDS-PAGE. Protein concentrations were determined spectrophotometrically. The concentrations of different MutL mutants were 14–25 µM (per protein monomer).

Fluorescent-dye labeling of proteins

Single-cysteine MutL variants (scMutL; 20 µM per monomer) were coupled to ATTO 647N maleimide at a molar ratio of 1:4 per monomer in buffer B (10 mM HEPES-KOH pH 8.0, 200 mM KCl, and 1 mM EDTA) for 1 h on ice as described earlier[154] (Figure 2.1). The excess uncoupled dye was removed by gel filtration on Zeba Spin

Desalting Columns 40K MWCO (Thermo Fisher Scientific, Waltham, USA). The labeled proteins were aliquoted, flash-frozen in liquid nitrogen, and stored at -80 ºC. The degree of labeling (DOL) was calculated from the UV-VIS absorbance spectra formula (Supplementary A.4). The degree of labeling for each MutL variant was between 70% and 90% (Table 2.2).

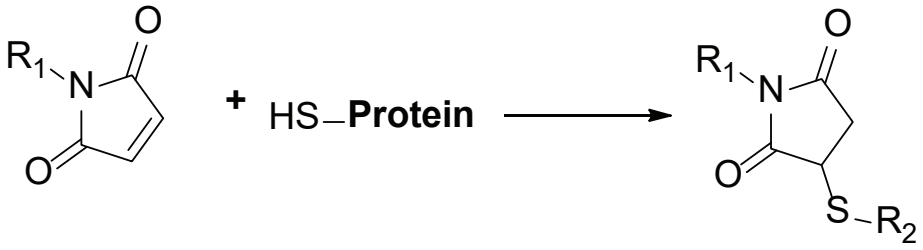


Figure 2.1: Protein labeling reaction with maleimide dye

MutL binding to DNA monitored by FRET

59 bp DNA duplex (Table 2.1) was diluted to 100 nM in 200 ml of buffer (25 mM HEPES-KOH pH 7.5, 5 mM $MgCl_2$, 50 mM KCl, and 0.05% Tween 20) at room temperature in 96-well plates. SYTOX Blue (10 mM) was used to stain DNA. MutL (200 nM per monomer) labeled with the ATTO 647N maleimide dye at a position 118, 131, 218, 251, 282, or 297 in the NTD was added, and fluorescence signals were recorded on a fluorescence microplate reader (TECAN infinite F200, Tecan Group Ltd, Switzerland) (Figure 2.2). Fluorescence intensities were measured in three channels (donor, acceptor, and FRET) with the following filter combinations: excitation 405 nm and 10 nm bandpass, emission 485 nm and 10 nm bandpass for the donor; excitation 590 nm and 20 nm bandpass, emission 670 nm and 25 nm bandpass for the acceptor; excitation 450 nm and 20 nm bandpass, emission 670 nm and 25 nm bandpass for the FRET. Signals were corrected for buffer background and spectral crosstalk to obtain adjusted signal intensities. Correction factors for the spectral crosstalk were obtained from experiments in which only one fluorophore-labeled component was present (paragraph in Supplementary Information, Tables B.1 and A.14, Figure B.1).

Results

Design of DNA systems

The 59 bp DNA duplex with a G/T (oligonucleotides 59T29 and 59G30) mismatch in the middle was subjected to MutL–DNA binding experiments (Figure 2.3). The DNA was stained randomly and non-covalently with SYTOX Blue, whereas scMutL variants were covalently labeled with ATTO 647N. Upon the MutL–DNA complex formation, FRET occurs between SYTOX Blue as a donor and ATTO 647N as an acceptor fluorophore. FRET efficiency was expected to depend on the degree of complex formation and the average distance between the DNA and the position of the ATTO 647N attached to scMutL (Figure 2.4).

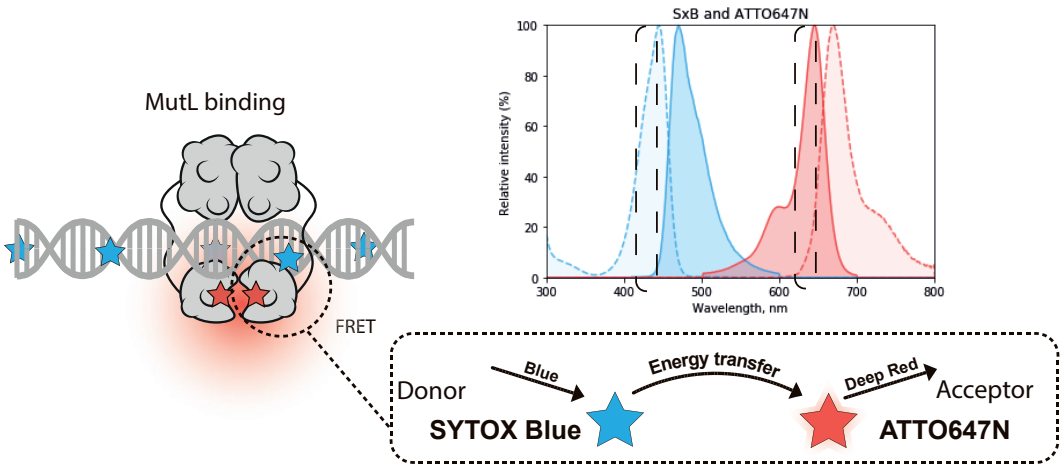


Figure 2.2: Schematic representation of the FRET-based MutL DNA binding assay. SYTOX Blue (donor) and ATTO 647N (acceptor).

FRET-based investigation of MutL–DNA binding

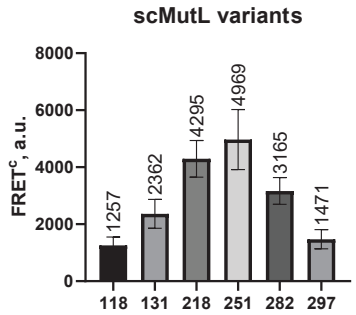
Various scMutL variants labeled with ATTO 647N were tested for binding to the DNA stained with SYTOX Blue the same way as described before for the binding of MutS to DNA[78]. SYTOX Blue has been used previously in a FRET experiment with another nucleic–acid–staining dye (Oregon Green) as the acceptor[155]. Single-cysteine MutL mutants were constructed and expressed with a mutation in positions 118, 131, 218, 251, 282, 297 (Figure 2.4). All residues take place in an N-terminal domain that is believed to bind to DNA. Here, we employed ATTO 647N as the acceptor fluorophore to minimize spectral crosstalk. The excitation/emission maxima for the SYTOX Blue dye bound to DNA are 444/480 nm. ATTO 647N is a bright fluorescent dark-red dye with the excitation/emission maxima at 646/664 nm and therefore is a suitable acceptor for SYTOX Blue.

Position of Cys in MutL	Degree of labeling, %*
118	76
131	88
218	71
251	83
282	70
297	78

Table 2.2: Degree of labeling with ATTO 647N of scMutL variants (single experiment).

The adjusted FRET signals were dependent on the position of the ATTO 647N in the MutL molecule: MutL(A251C) yielded 5-fold stronger signals than MutL(Q118C) did (one of the "negative-control variants"). The strongest signals were detected for variants MutL(T218C) and MutL(A251C), indicating the proximity of these residues to DNA in the MutL–DNA complex (Figure 2.3).

A



B

Position of Cys in MutL	Distance between Cys and DNA, Å*	Ranks		
		Distance	FRET	Xlinking
118	~50	6	6	4
131	~42	5	4	6
218	~26	2	2	2
251	~21	1	1	1
282	~29	3	3	3
297	~35	4	5	5

Figure 2.3: **(A)** FRET between Single cysteine MutL mutants labeled with AF647 and DNA stained by SxB. Error bars are SD (n=3). **(B)** The distances from amino acid residues to DNA were calculated based on the theoretical model of the ternary complex are correlated with the rank of FRET between DNA and attached dye.

Discussion

The sliding clamp structure raised a lot of questions[78]. One of the problems was the absence of DNA in such a structure, although indirect methods indicated its position in the center of the doughnut formed by the MutS structure[156,157]. In this chapter, we decided to probe the position of the MutL NTD relative to the DNA. Having simulated B-DNA in the center of MutS, the average distances from the center of the base pairs to the beta cysteine atoms in some variants were calculated and compared with the data obtained in the fluorescence method.

We hypothesized that MutL should bind to DNA at low salt concentrations in the same way as when protein is recruited by MutS. A FRET-based method was carried out to estimate relative distances between ATTO 647N dye coupled to a different position in MutL and DNA stained randomly with SYTOX Blue. To estimate distances between DNA and amino acid residues in MutL-NTD, the MutS/MutL crystal structure (pdb: 5AKB) with modeled B-DNA was used[78]. The distances from the β-carbon atom in the amino acid of interest to bases of the DNA 3D model were calculated. Unlike the methods when the fluorophore is attached to the DNA during oligonucleotide synthesis, a dye that is evenly distributed throughout the DNA was used. At the same time, MutL can bind to a random DNA sequence. Taking this information into account, FRET from the average distance between the DNA and the fluorophore was measured. The FRET data obtained experimentally correlate with the calculated distances and with the length between cysteine and modification on DNA in the cross-linking experiments performed by our collaborators. There was a good inverse correlation between the order of FRET signals with an average distance to DNA as estimated from the model (Figur 2.3B). The corrected FRET signals for MutL variants 218 and 251 were highest suggesting proximity of positions 218 and 251 to DNA. Using site-specific cross-linking of the same variants in MutS dependent and MutS independent manner, they confirmed our initial idea and thereby validates MutL DNA binding center via chemical trapping.

Our data correlate well with the newly obtained cryo-EM structure MutS-MutL in clamp state (pdb 7AIB), which contained DNA in the predicted location.

In conclusion, these data suggest that this approach of single cysteine variants with a covalently attached fluorophore and randomly stained DNA can be used for

probing DNA centers of other proteins being an alternative for classical methods. This approach can serve as a basis for the development of more complex approaches to the study of protein DNA interactions, which are described in the next chapters.

This work is part of a publication "Probing the DNA-binding center of the MutL protein from the *Escherichia coli* mismatch repair system via crosslinking and Förster resonance energy transfer" in the journal Biochimie produced in collaboration with the group of Prof. Dr. Elena Kubareva.

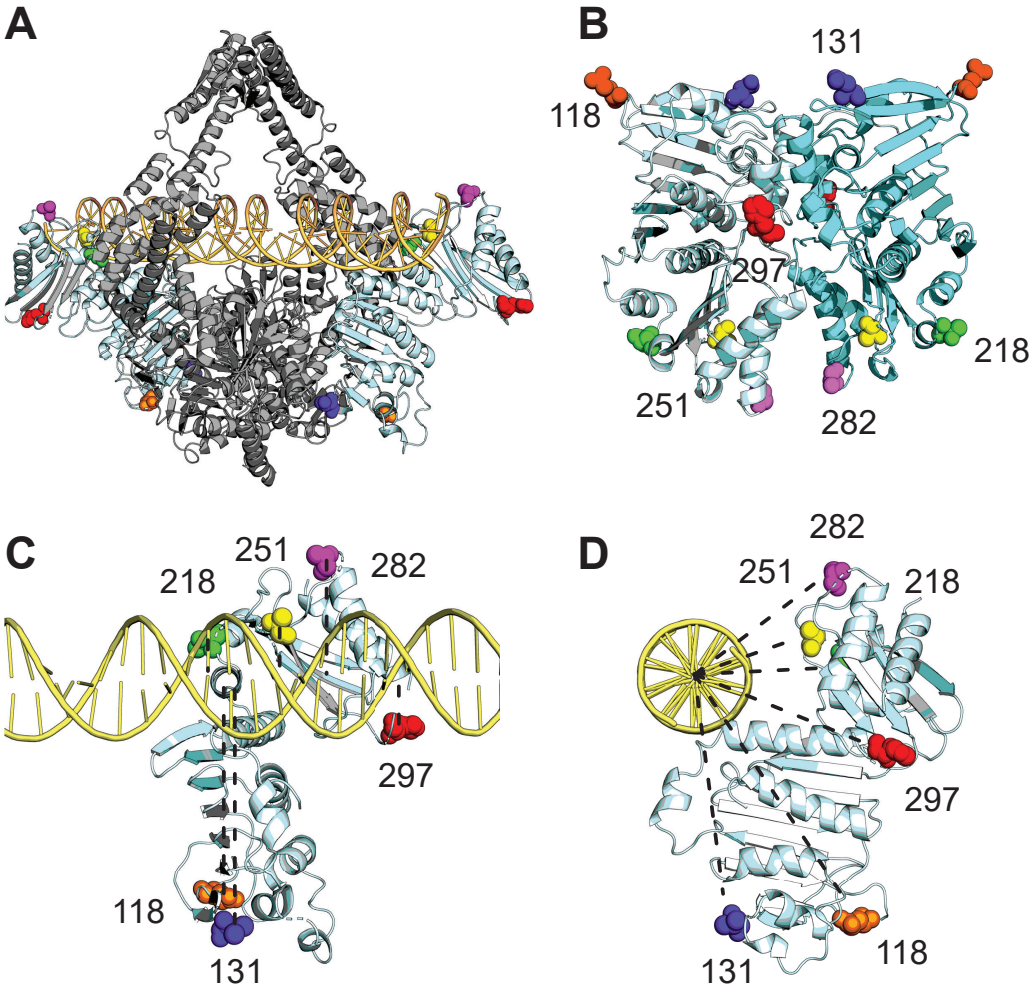


Figure 2.4: **(A)** The model of the ternary MutS–NTD-DNA complex structural model (pdb id: 5AKB) with modeled B-DNA. MutS is a homodimer, whereas MutL-NTD is depicted as a monomer. **(B)** Location of amino acid residues (displayed in different colors) replaced with Cys to obtain six different single-cysteine MutL (scMutL) variants (pdb id: 1B63). **(C)** and **(D)** The shortest distances from the β-carbon atom in the amino acid of interest (position indicated by the numbers) to the center of the base pair in the DNA are shown in two orientations (pdb id: 5AKB).

Chapter 3

Covalent trapping of MutS to the DNA for MutL recruitment

Vladislav Kunetsky[1], Maya Monakhova[2], Tatiana Oretskaya[2], Elena Kubareva[2], Peter Friedhoff[1]

1 Institute for Biochemistry, FB 08, Justus Liebig University, Heinrich-Buff-Ring 17, D-35392 Giessen, Germany
2 Belozersky Institute of Physico-Chemical Biology and Chemistry Department, Lomonosov Moscow State University, Leninskye Gory, 1, 119991 Moscow, Russia

Abstract

The DNA repair MutS protein is responsible for mismatch recognition and it initiates a complex cascade of reaction including large structural conformational changes following recruitment and activation of other proteins. In this work, MutS single cysteine variants were cross-linked to the modified DNA and checked for an ability to recruit MutL. The cross-linking approach allows us to trap a transient ATP active state of MutS to the DNA. The functional assay shows the ability of MutS recruit MutL only being cross-linked to the DNA by clamp domains. Activation of MutS can be performed only on mismatch DNA and not only with ATP, but also with non-hydrolyzable analog ATPγS.

Introduction

The DNA mismatch repair system is an evolutionarily conserved system that identifies and removes mismatches that escaped proofreading activity of polymerase during replication[7]. The key protein of the *E. coli* mismatch repair system which recognizes mismatches, insertion/deletion loops (IDL), and some other damages is MutS homodimer[48]. MutS searches for mismatches by one-dimension diffusion with the following binding to the mismatch[64]. After mismatch recognition MutS is able to bend DNA and recruit MutL in the presence of ATP[32,77]. MutL is a homodimer protein with two structured regions N-terminal (NTD), C-terminal (CTD) domains, and the unstructured linker between[89]. The ternary complex (DNA-MutS-MutL) is a key active intermediate state reasonable for strand discrimination by MutH at the hemimethylated GATC site[38]. The nick is used in the following unwinding by UvrD in a mismatch-dependent manner and other enzymes such as single-stranded binding proteins, exonucleases participate in the next steps of the reaction. DNA Polymerase III and DNA ligase complete the synthesis of the correct complementary strand[17].

Various models have been proposed to clarify the long-standing puzzle of the transmission signal from MM to the GATC site[128]. In the translocation and sliding models, MutS proteins bind to the mismatch and after ATP hydrolysis moves away from the recognition site to the GATC site[73,83]. Further evidence of the sliding clamp model is the crystal structure of the sliding clamp[78] and the varieties of Cryo-Em structures described in **chapter 4**, as well as experiments with obstacles between the mismatch and the other site[134]. The transactivation and MutL multi-loading models assume that MutS binds to the mismatch and does not leave the mismatch during the signal transfer from the GATC site[129,137]. The strongest argument in favor of the transactivation model is the activation of MutH on homoduplexes of oligonucleotides containing the GATC site through the interaction of MutS activated on the mismatch DNA[129] and fresh data from LacI roadblocks or sliding-deficient MutSα experiments[135], which contradicts the results obtained in the Modrich group[133]. The data obtained in yeast show superstoichiometric complexes with MutL multi-loading, singling out in a separate model, which may be a feature of this type of organism[137]. This multi-loading data do not contradict the sliding model and it is discussed in **chapter 6** (Supplementary figure B.16).

MutS checks Watson-Crick interaction of base pairs by insertion of phenylalanine "finger" of mismatch-recognition domain. After detection of mismatch, MutS is able

to kink DNA up to 45-60°[54]. Upon ATP binding MutS undergoes conformational changes to a sliding clamp realizing connector domain for further MutL recruitment[78]. MutS pushes DNA inside the "donut" hole and becomes highly dynamic with the following sliding along DNA[78]. MutS consist of many different domains which upon ATP binding can dramatically change their positions or can be rotated up to 180°.

Despite the importance of X-ray crystallography, Cryo-EM, and other structural methods it is hard to investigate such complexes due to the highly dynamic nature of transient states. These methods give us "screenshots" of states which can be interpreted by different models. Trapping an active mobile state of MutS-DNA can be performed by the covalent cross-linking allowing fixing of the complex in certain conformation. Cross-linking is a well-known tool for stabilization DNA or RNA-protein complexes. For accurate cross-linking design, we used single-cysteine variants of MutS and site-specific cross-linking[145]. It was shown that cross-linking species $MutS^{497C}$ can introduce bending of DNA by monitoring FRET between acceptor and donor dyes covalently attached to DNA[147]. In this study MutS variants with a single cysteine position in the clamp or mismatch-recognition domains were cross-linked to the designed and modified oligonucleotide with or without mismatch.

Moreover, we combined site-specific cross-linking of single cysteine MutS variants with the ability to recruit MutL protein using FRET-based assay[149]. The new strategy allows us to investigate MutL DNA interaction in an ATP mismatch-dependent manner. This research answers a long-term question in MMR and presents new evidence into the mechanism of MMR initiation and mismatch removal.

Material and methods

Protein expression and purification MutS and MutL proteins were purified as described previously[79,105]. Protein concentrations were determined spectrophotometrically using theoretical extinction coefficient (ϵ_{280} nm =73 605 $M^{\neg 1}$ $cm^{\neg 1}$ for MutS, ϵ_{280} nm = 54 270 $M^{\neg 1}$ $cm^{\neg 1}$ for MutL). MutS variants were stored in HPLC200 buffer (10 mM HEPES/KOH (pH 8.0), 200 mM KCl and 1 mM EDTA) and MutL variants were stored in high salt buffer HPLC500 (10 mM HEPES/KOH (pH 8.0), 500 mM KCl and 1 mM EDTA). MutL single cysteine variants were labeled as described previously[149], using another dye Alexa Fluor 647.

DNA

All oligonucleotides were synthesized by the company Eurogentec (Seraing, Belgium). The positions for modification selected according to the pdb 1e3m structure. For protein DNA-cross-linking experiments a 59 base pair substrate with a G/T mismatch in position 24 and a C_5-aliphatic amino group coupled to position 28 or 16 (T strand; -4 bp or +8 from G/T mismatch) through a hexamethyleneacrylamide linker was used (Table 3.1 and supplementary figure B.3). Oligonucleotides were modified as described previously[147]. In brief, cross-linkers SPDP (6.8 Å) or PEG_4SPDP (25.7 Å) were dissolved in DMSO to 500 mM, aliquoted, and stored at -20 °C. Cross-linker solution (5 μl) was added to 45 μl oligonucleotide (100 μM) and incubated overnight at room temperature. Excess of unbound cross-linker was removed by gel filtration

Strand	length	Sequence	Modification
Amino-modified top -4	59	5'-GTG CGC AAA TCC AGA CGT CTG TCG ACG TTG GGA AGC TTG AGT ATT CTA TAG TGT CAC CT–3'	Amino modified C6 dT at position 28
Amino-modified top +8	59	5'-GTG CGC AAA TCC AGA CGT CTG TCG ACG TTG GGA AGC TTG AGT ATT CTA TAG TGT CAC CT–3'	Amino modified dC at position 16
Amino-modified bottom T	59	5'-AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTT GAC AGA CGT CTG GAT TTG CGC AC–3'	-
Amino-modified bottom C	59	5'-AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTC GAC AGA CGT CTG GAT TTG CGC AC–3'	-
Amino-modified bottom T	59	5'-/Cy5/AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTT GAC AGA CGT CTG GAT TTG CGC AC–3'	5' Cy5 fluorophore
Amino-modified bottom C	59	5'-/Cy5/AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTC GAC AGA CGT CTG GAT TTG CGC AC–3'	5' Cy5 fluorophore
Top	59	5'-TGA AGC TTA GCT TAG GAT CAT CGA GGA TCG AGC TCG GTG CAA TTC AGC GGT ACC CAA TT-3'	-
Digoxigenin-modified bottom	59	5'-/Dig/-AAT TGG GTA CCG CTG AAT TGC ACC GAG CTT GAT CCT CGA TGA TCC TAA GCT AAG CTT CA-/Dig/-3'	5' and 3' Digoxigenin
Digoxigenin 3' top	29	AAA-TAG-CTT-GGC-GTA ATC-ATG-GTC-ATA-GC	3' Digoxigenin
Digoxigenin 5' top	29	ACA-CTA-TAG-GGC-GAA TTG-GCG-GCC-GCG-AT	5' Digoxigenin
bottom T long	117	GCT-ATG-ACC-ATG-ATT-ACG-CCA-AGC-TAT-TTA-GGT-GAC-ACT-ATA-GAA-TAC-TCA-AGC-TTC-CCA-ACG-TTG-ACA-GAC-GTC-TGG-ATT-TGC-GCA-CAT-CGC-GGC-CGC-CAA-TTC-GCC-CTA-TAG-TGT	-
bottom C long	117	GCT-ATG-ACC-ATG-ATT-ACG-CCA-AGC-TAT-TTA-GGT-GAC-ACT-ATA-GAA-TAC-TCA-AGC-TTC-CCA-ACG-TCG-ACA-GAC-GTC-TGG-ATT-TGC-GCA-CAT-CGC-GGC-CGC-CAA-TTC-GCC-CTA-TAG-TGT	-

Table 3.1: Oligonucleotides for cross-linking

with Zeba Desalt Spin Columns (Pierce, Thermo Scientific). The final DNA substrate was generated by annealing to a complementary strand labeled with Cy5 at the 5' end. An oligonucleotide 117 bases long was used for blocking ends experiments which was complementary to a modified oligonucleotide 59 bases long and two oligonucleotides 29 bases long with 5' and 3' digoxigenin modified ends (Table 3.1 and supplementary figure B.3).

MutS cross-linking to the DNA

The cross-linking was performed by adding $MutS^{497C/D835R}$ or $MutS^{103C/D835R}$ (4 μM per monomer) to the modified DNA duplex (1 μM) pre-incubated in buffer FB150 (25 mM HEPES/KOH, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$) on the ice during 30 min. The cross-linking was performed in a volume of 10 μl. To separate cross-linked complex with uncross-linked proteins 6.5 % SDS-polyacrylamide gels electrophoresis (SDS-PAGE) were used. DNA was observed by fluorescence of Cy5 dye covalently attached to the DNA or staining with ethidium bromide (EtBr). Subsequent staining of the proteins was done by Instant Blue.

MutL recruitment assay

1 μl of the cross-linking mixture of single cysteine MutS variant and unlabelled oligonucleotide was diluted in 200 μl buffer FB150 (HEPES 25 mM, $MgCl_2$ 5 mM, Tween 20, KCl 150 mM, pH 7.5) at room temperature in 96-well-plates. The mixture contained 20 nM scMutS (monomer) and 5 nM of 117 bp DNA with G/T mismatch. The intercalating dye SybrGreen (250 nM) was used to stain DNA and it acts as a donor (Figure 3.1B). MutL 25 nM (monomer) single-cysteine variant 297C labeled with Alexa Fluor 647 maleimide was added to the mixture with incubation at room temperature for 120 seconds. The reaction was triggered by ATP (10 μM) and fluorescence signals were recorded with a fluorescence microplate reader (TECAN infinite F200, Tecan Group Ltd, Switzerland). Fluorescence intensities were measured in three channels (donor, acceptor, and FRET) with the following filter combinations: excitation 450 nm and 20 nm bandpass, emission 535 nm and 25 nm bandpass for the donor; excitation 620 nm and 10 nm bandpass, emission 670 nm and 25 nm bandpass for the acceptor; excitation 485 nm and 20 nm bandpass, emission 670 nm and 25 nm bandpass for the FRET. Signals were corrected for buffer background and spectral crosstalk to obtain adjusted signal intensities. Correction factors for the spectral crosstalk were obtained from experiments in which only one fluorophore-labeled component was present (supplementary table B.1).

Results

Cross-linking of scMutS variants to the modified DNA

The structure for *E. coli* MutS (pdb: 1e3m) was obtained with 30 bp DNA containing a mismatch in the center and demonstrates a 60°C bending of DNA[54]. To create a system for MutS investigation with other MMR proteins we used 59 bp DNA (supplementary figure B.3). This distance allows recruiting second protein MutL. We first asked if a position for cross-linking influences MutS activity. Two

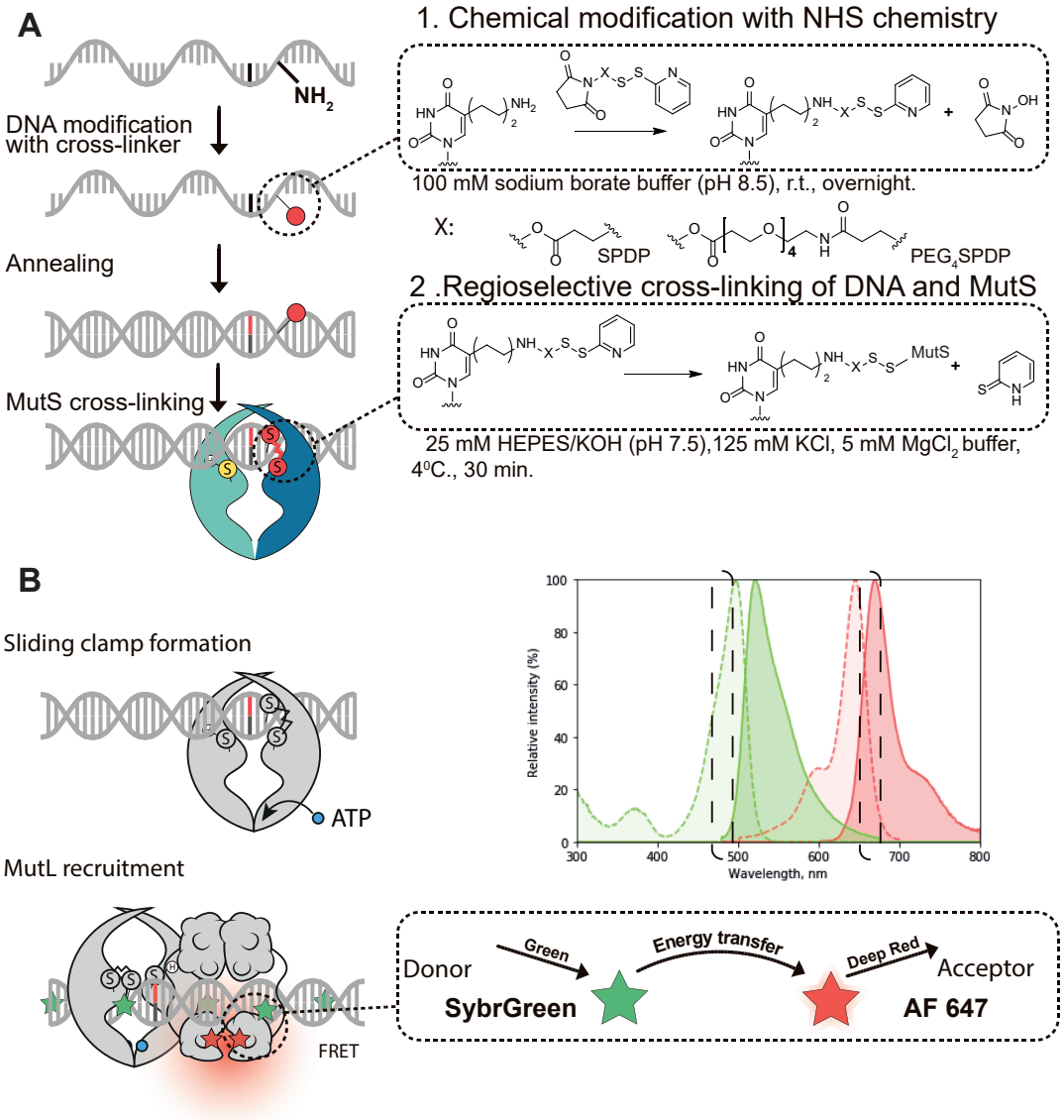


Figure 3.1: **(A)** The cross-linking of a single cysteine MutS variant to the modified DNA duplex. Synthesized oligonucleotides were modified with short (SPDP) or long (PEG_4SPDP) heterobifunctional cross-linkers. Modified single-strand oligonucleotides were annealed with complementary strands with one defined mismatch. The regioselective covalent reaction between modified duplex and MutS single variant takes place. **(B)** Schematic representation of FRET-based MutL recruitment assay. Upon adding ATP MutS undergoes conformational changes to the sliding clamp state and recruits MutL. The NTD of MutL which is labeled by AF647 comes closer to the DNA and FRET between stained DNA and attached dye takes place. SG acts as a donor and transfers energy to the acceptor AF647.

different positions on the DNA were used (Figure 3.2A and B). Since MutS is a homodimer that, when bound to mismatch DNA, takes on an asymmetric form, where only one of the domains intercalates the phenylalanine finger into the stacking interaction of unpaired nucleotides, namely domain A in the crystal structure, we decided to investigate two cases of cross-linking. In the first case, we designed the oligonucleotide so that position 497 in domain A was close to the modification located on the nucleotide in the "bottom" G strand at -4 nucleotides from the mismatch (Supplementary Figure B.3). In the symmetrical variant, position 497 was already on the second domain B and was close to the nucleotide on the "bottom" G strand at +8 nucleotides from the mismatch. An alternative option for cross-linking is position 103, which is in the mismatch-recognition domain and is also close to these modifications. In one of the domains, the structure below 108 amino acids is not specified, but we can assume that it is also in the vicinity of modification. The distances in all cases are relatively short, which allows us to use a covalent trap.

Both single-cysteine variants gave a relatively high cross-linking yield, in the case of a nucleotide that was modified at -4 bp from the mismatch (Figure 3.2A and C), and in both cases, the yields were higher than for variant +8 bp (Figure 3.2B and D). Also, a short SPDP cross-linker (6.8 Å) gave the same yield compared to a long one PEG_4-SPDP (25.7 Å). This also supports the suggestion that these positions are relatively close to the DNA recognition process. Surprisingly, there was no difference in cross-linking yield between homoduplex and mismatch DNA. DNA-MutS conjugate formation was done in the absence of ATP. These variants did not show the difference in the presence of ATP, ADP, and absence of nucleotides (data not shown).

MutL recruitment to the DNA by cross-linked MutS

We then asked if cross-linked MutS to DNA is still active to recruit MutL. To test the activity of the cross-linked proteins to DNA, a FRET-based functional method was used to recruit MutL. The ability to recruit for $MutS^{497C}$ was dependent on the position of the cross-linker.

MutL recruitment takes place only in the presence of ATP and cross-linker (Figure 3.3A lane 6). As MutS forms a sliding clamp the protein begins to move along the DNA, which causes MutS to slip off the ends (lane 5). When the protein is fixed by site-specific cross-linking, the FRET signal between MutL and the stained DNA is significantly increased. Surprisingly, the position that gave the higher cross-linking yield showed less activity for both short and long cross-linkers (Figure 3.3C).

An alternative method for MutL recruitment on oligonucleotides is end-blocking, for example with Fab fragments with high affinity to digoxigenin, to prevent MutS dissociation from ends (Figure 3.3A lane 4 and D). To compare the ability to recruit by cross-linked MutS to uncross-linked, we used 117 bp oligonucleotides with 3' and 5' ends modified with digoxigenin (Supplementary Figure B.3C). In this case, MutL recruitment by MutS via SPDP and PEG_4SPDP cross-linking results were comparable with end blocking (Figure 3.3A lanes 4 and 6). The ternary complex is induced at low concentrations of ATP as well as by a non-hydrolyzable analog ATPγS. In the case of ADP and ADPNP, no activity was observed (Figure 3.3B).

In contrast to the cross-linked clamp domain, we found that the cross-linking of the mismatch-recognition domain (103), which was cross-linked in relatively high

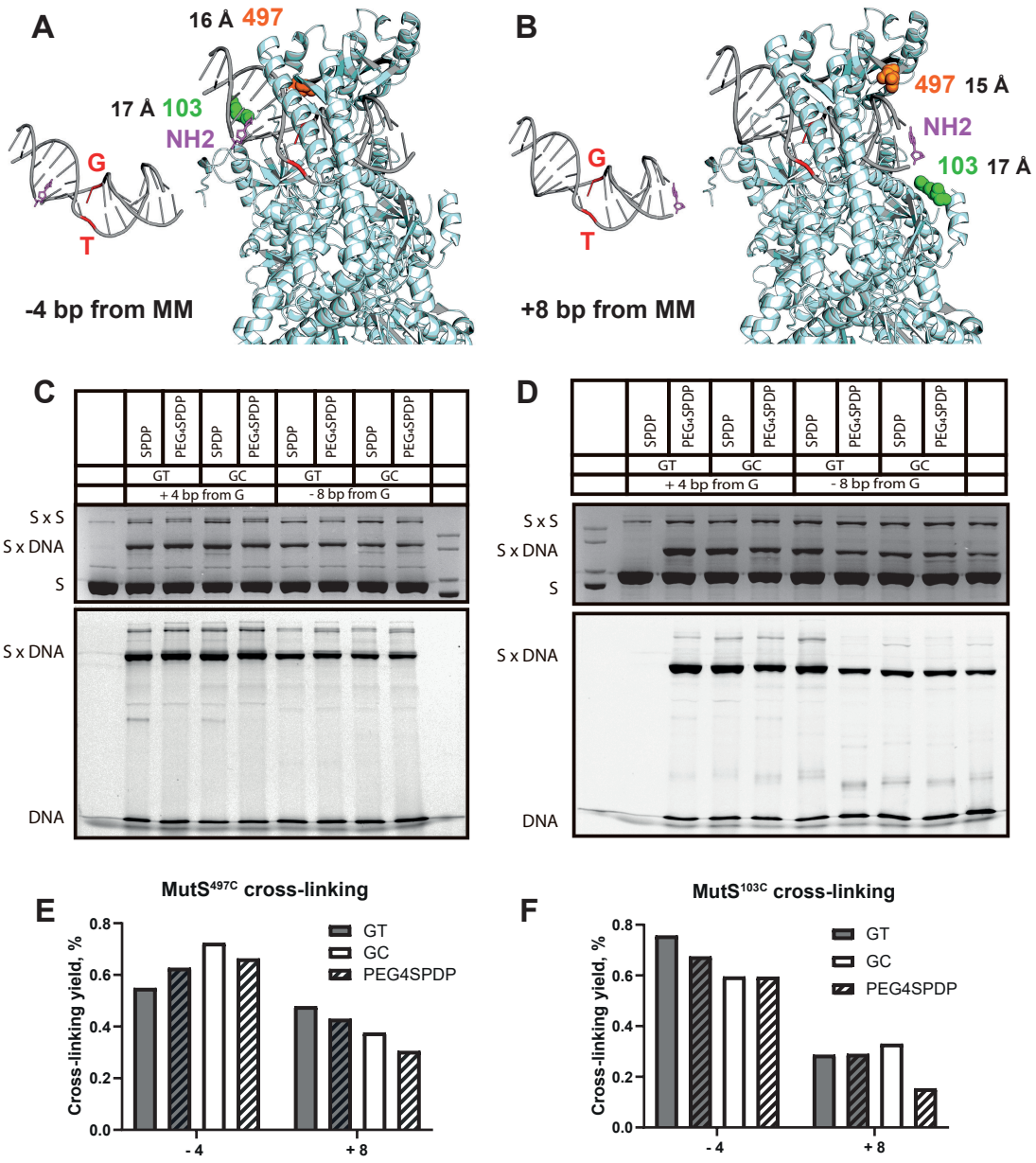


Figure 3.2: **(A)** For MutS chain A, the amino-modified nucleotide base (dT) is at position -4 from the mismatch on the lower chain. **(B)** For chain B of MutS, the amino-modified nucleotide base (dC) is at position +8 from the mismatch on the lower chain. **(C)** Cross-linking of MutS$^{497C/D835R}$ after 30 min on ice was observed on the 8 % SDS gels. To evaluate the quantitative yield of the reaction, the fluorescence signal of the covalently attached dye (Cy5) to the oligonucleotides on the lower gel was observed. Free oligonucleotides can be observed in the lower part of the gel; the cross-linked DNA-MutS complex is located in the upper part of the gel. The same gel was stained with Instant Blue for protein monitoring (top gel). The lower band corresponding to the MutS monomer, the middle band corresponding to the MutS-DNA complex and the upper band corresponding to the MutS dimer can be detected on the gel. **(D)** Cross-linking of MutS$^{103C/D835R}$. Cross-linking yields for **(E)** MutS$^{497C/D835R}$ and **(F)** MutS$^{103C/D835R}$ (single experiment) on ice 30 min 8 % SDS gel (single experiment).

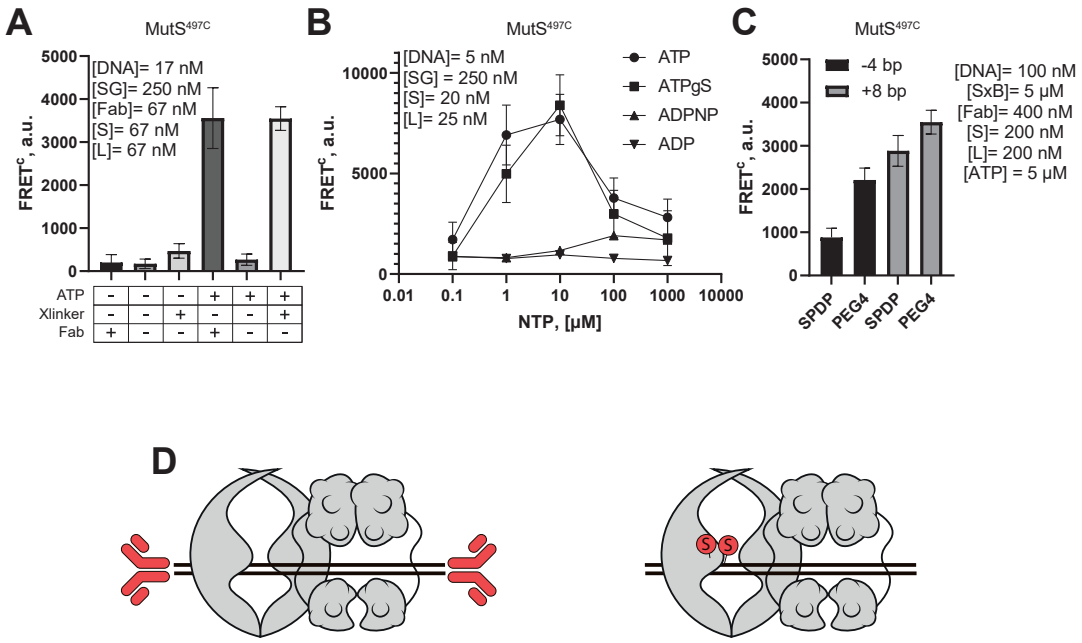


Figure 3.3: **(A)** MutL recruitment takes place in the presence of blocked DNA ends via covalent trapping of MutS preventing dissociation from ends. **(B)** MutL recruitment can be induced by ATP or non-hydrolyzable analog ATPγS. **(C)** MutL recruitment by MutS^{497C} with SPDP and PEG$_4$SPDP at different positions. **(D)** Schematic representation of covalent and non-covalent trapping of MutS and MutL recruitment

yield, showed significantly lower activity in the experiment with MutL compared to the clamp domain (497) (Figure 3.4 B). Taken together, these data show that cross-linking of the wrong domain of the protein leads to a dramatic loss of activity. In accordance with the kinked structure of MutS, position 103 on the mismatch-recognition domain is distant from DNA (the mismatch-recognition domain is not specified in this structure), compared to position 497. This physically complicates the conformational changes of the protein and interferes with its functional activity, at the same time position 497 is at a close distance to DNA, which allows MutS to undergo conformational changes upon ATP binding (Figure 3.4D). It is not enough to covalently trap MutS to DNA with the right position. To induce recruitment a mismatch is also necessary (Figure 3.4A).

Discussion

The ternary complex between mismatch DNA, MutS, and MutL in the presence of ATP is an important coordination step involved in the mismatch repair pathway. Despite structural and functional data, little is known about the arrangement of these proteins. The dynamic nature of the complex formation complicates the investigation. Using modified DNA and thiol-specific cross-linking to the DNA we were able to trap dynamic complex, which allows us a more detailed analysis of MutL recruitment.

In a previous study, we showed that cross-linked MutS$^{497C/D835R}$ is fully active in the first step namely in bending mismatch DNA. Here the data show that

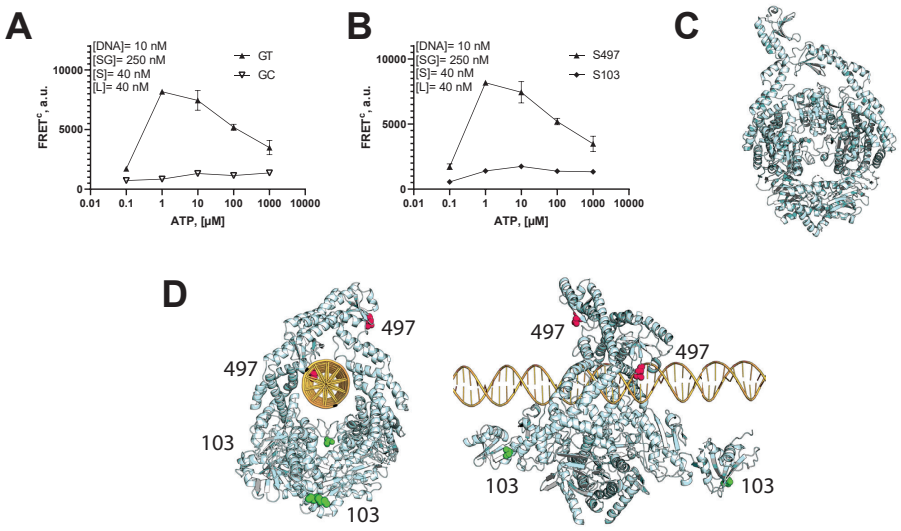


Figure 3.4: **(A)** MutL recruitment requires covalent trapping to the DNA and a mismatch. **(B)** Recruitment of MutL by MutS cross-linked to the DNA at mismatch-recognition domain is reduced in comparison with clamp domain cross-linking. **(C)** The flexibility of the lever and clamp domains was shown in the apostate of *E. coli* MutS (pdb id: 6I5F)[62]. **(D)** Kinking sliding clamp structure of MutS (pdb id: 7AIC) with close clamp domain with position 497C and modeled mismatch-recognition domain with position 103C.

$MutS^{103C/D835R}$ and $MutS^{497C/D835R}$ can be cross-linked with high yields regardless of the presence or absence of a mismatch (pdb id: 7AI5 vs 1e3m) as expected. Interestingly, covalent trapping has no specificity for heteroduplex or homoduplex DNA having similar yields. These data are in agreement with the notion that cross-linking takes place during MM-independent MutS searching.

One of the critical issues for sliding clamp structure was the artificial nature of the interaction between MutS and MutL, which critics considered very important[78]. One of the models described in detail in **chapter 1 and 2** was the so-called stationary model, where MutS, after finding a mismatch, does not leave DNA, but recruits MutL and activates it, followed by activation of the latent endonuclease activity of MutH[89,129]. According to this model, the sliding clamp is not a functionally important state of the protein and serves a different function[135]. To refute this statement, the following hypothesis was developed. If MutS does not leave the mismatch, then the active structure responsible for the cascade of subsequent reactions is the structure obtained on the DNA mismatch, where the DNA bends 60°, and the mismatch-recognition domain holds it[54]. If we cross-link the mismatch-recognition domain to the DNA and then activate the complex with ATP, then we could observe the MutL recruitment. An alternative hypothesis is the sliding clamp model, where the connector domain is rotated downward by 180°, thereby opening the interface with MutL, and the DNA is pushed into the hole of the doughnut structure described earlier[78]. In this case, if we cross-link the clamp domain, then during conformational changes caused by ATP, this will not mechanically interfere with the transition to a new state and recruitment of MutL. Our data indicate, that when the MutS mismatch-recognition domain was chemically trapped to the modified DNA, it almost completely lost its ability to recruit MutL, in contrast to the

case when DNA was cross-linked to the clamp domain. This method leaves many interesting and at the same time very important questions. It is not clear whether MutS attached to DNA can recruit MutH, or whether MutS needs to leave a mismatch for further reactions, or perhaps a new state of the complex not previously described. It is interesting to investigate whether a protein is functionally symmetric or not. We know from the crystallographic structure that although a protein is a homodimer, it has a structural asymmetry when it binds to a mismatch[54]. At the same time, the crystal structure of the slide clamp is completely symmetrical, which is the third point of criticism of this structure[78]. Although we assumed that the clamp domain of the MutS would be able to recruit MutL when MutS^{497C} was trapped with a long cross-linker, we were surprised at the results obtained with a short cross-linker. The sliding clamp structure predicted that it is not possible to obtain an active MutS on the DNA with a short cross-linker unless the clamp domain is dynamic and flexible. Indirect confirmation of this model was obtained from the new structure of the apostate of MutS with kinked clamp and lever domains comparing with another apo structure obtained for *Neisseria gonorrhoeae*, where these domains were straight (Figure 3.4C)[62,70]. Direct confirmation of our hypothesis was confirmed later by a new Cryo-Em structure of a sliding clamp state with kinked clamp and lever domains obtained by our colleagues[71] and described in **chapter 4** (Figure 3.4D).

Site-specific cross-linking can be used for the characterization of raw populated and transient states. The specificity of covalent trapping can be checked by different functional methods. MutL recruitment assay has shown that only clamp domains that have a flexible nature can be attached to the DNA according to the kinked structure. In **chapter 4** we use this approach for trapping an intermediate MutS state, showing that cross-linking yields were strongly dependent on the position of the cysteine. Only MutS single cysteine variants that are in proximity to the modified DNA in mismatch structure give high cross-linking yield. And it was shown that cross-linking is also specific for the transient states and trapping can be done only at conditions that induce conformational changes such as ATP and mismatch DNA.

Chapter 4

Covalent trapping mismatch activated long-lived signaling clamp state of MutS

Vladislav Kunetsky[1], Rafael Fernandez-Leiro[2,3], Doreth Bhairosing-Kok[4], Charlie Laffeber[5], Joyce H. G. Lebbink[5], Titia K Sixma[4], Meindert H. Lamers[2,6], Peter Friedhoff P.[1]

1 Institute for Biochemistry, FB 08 Justus Liebig University, Heinrich-Buff-Ring 17, 35392 Giessen
2 MRC Laboratory of Medical Research, Cambridge, United Kingdom.
3 Centro National de Investigaciones Oncológicas, Madrid, Spain.
4 Division of Biochemistry and Oncode Institute, Netherlands Cancer Institute, Amsterdam, The Netherlands.
5 Department of Cell Biology and Genetics, Cancer Genomics Center and Department of Radiation Oncology, Erasmus Medical Center, Rotterdam, The Netherlands.
6 Department of Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands

Abstract

The DNA mismatch repair system is a key process in the correction of replication errors in all three kingdoms of life. The process in *E. coli* starts with the recognition of mismatches by the ABC-ATPase MutS. After conformational changes caused by mismatch binding and ATP, another ATPase, MutL binds MutS to form a ternary complex leading to activation of enzymes required for strand discrimination, error removal. Despite years of research, the structure of the protein-protein complexes has remained elusive, and the MMR mechanism raises many questions. One reason for these problems lies in the highly dynamic nature of the complexes involved in the MMR pathway. We focused on the first step of DNA mismatch repair, namely on the conformational changes of MutS and its functions. Using the site-specific intermolecular cross-linking of single-cysteine MutS variants, a transient active state of MutS on DNA was trapped . Combining this approach with the FRET method and biochemical assays, it was shown that this transient state is important for the next steps in the DNA repair pathway, namely in the MutL recruitment and DNA nicking by MutH, and likely resembles the mismatch-active signaling clamp state of MutS on DNA. To investigate an intermediate conformation between the mismatch-bound and sliding clamp states, chemically modified oligonucleotide with a sulfur reactive cross-linker was used and mismatch and ATP-dependent movement of the connector domain was shown.

Introduction

The MMR system plays an important role in maintaining DNA stability during replication produced by the DNA polymerase as well in the anti-recombination process[158]. The system is a complex and multi-stage process, including at least more than 10 proteins[7]. MutS is a key protein in mismatch repair systems that specifically binds to the mismatch DNA and small insertion/deletion loops and initializes a cascade of subsequent reactions[48,49]. MutS homologous are evolutionary conserved and presented in all three kingdoms of life[50]. The reaction begins only when MutS recognizes a single mismatch during scanning DNA for mismatches due to one-dimensional diffusion[63]. Exiting structures of the protein bound to the mismatch DNA presents oval-shaped homodimer protein[54]. The crystal structures for *E. coli* and *Thermus aquaticus (Taq)* have structural similarities[53]. Protein has a high binding affinity to the ends of DNA, as well as to mismatch[65].

Detection of the mismatch leads to the DNA kinking by MutS and the protein structure becomes asymmetric[67]. The structure is well ordered except for the mismatch domain of the B subunit. Two domains are involved in MutS mismatch DNA interaction. Clamp domains are closed above the DNA after mismatch recognition. Only one subunit of the N-terminal mismatch domain of dimer interacts with MM. The region Phe-X-Glu involved in DNA interaction is evolutionarily conserved[68]. The interaction with MM involves highly conserved phenylalanine residue. It wedges into DNA by stacking interaction via aromatic rings with thymine in G/T mismatch[54]. Mutations in Phe36 lead to failure recognition of insertion/deletion mismatch or a G/T base pair mismatch[69]. Regardless of the type of mismatch (except for C/C), DNA is bent by 60 degrees[68]. At present, there are two functionally significant structures, such as MM bound and a sliding clamp state[78]. Upon mis-

match recognition, MutS undergoes ATP-dependent conformational changes which lead to a so-called "sliding clamp" state with unusual stability on DNA ($\approx$ 600 s)[63]. MutS dissociates in the presence of ATP from open DNA ends[74,83,139] or single-strand DNA[63,159]. During its conversion to a sliding clamp, MutS undergoes at least two conformational states[67]. This sliding clamp ATP-bound state[32] randomly diffuses along DNA with a contact for MutL interaction[77]. The structure of MutS in sliding clamp looks like a "donut" with a channel inside where DNA can be loaded[78]. The size and charge of this channel suggest that it could be a place for DNA in a sliding state mode. At the same time, large conformational changes to the connecter domains occur, which create an interface for binding to the N-terminal domain of MutL[78]. In *E. coli* MutL activities latent endonuclease activity of MutH that introduces a nick at hemimethylated GATC site of the daughter strand[33,103].

Structural and functional analysis of individual steps in the mismatch repair process has been very complicated due to the involvement of multiple components and the transient nature of the individual states. Using intermolecular cross-linking and cross-linking to the modified DNA of single-cysteine variants together with FRET, nicking assays, and Cryo-Em structures we characterized two transient conformational changes of MutS.

Material and methods

Protein purification

MutS, MutL, and MutH proteins were purified as described previously[78,105]. His-tagged MutS dimer variant ($MutS^{R420C/D835R}$) with single-cysteine at position 420 was created from the dimeric cysteine-free MutS variant[57,78,79] using the NEBuilder HiFi Assembly Cloning Kit (NEB) (see supplementary information for details) and purified obtained as described before[78]. Protein concentrations were determined spectrophotometrically using theoretical extinction coefficient (ϵ_{280} nm =73 605 $M^{\neg 1}$ $cm^{\neg 1}$ for MutS, ϵ_{280} nm = 54 270 $M^{\neg 1}$ $cm^{\neg 1}$ for MutL).

DNA

For protein DNA-cross-linking experiments a 59 base pair substrate with a G/T mismatch in position 24 and a C_5-aliphatic amino group coupled to position 28 (T; 4 bp from G/T mismatch) through a hexamethyleneacrylamide linker (Table 4.1) was used. DNA oligos were modified as described previously[147]. In brief, cross-linker SPDP (6.8 Å) was dissolved in DMSO to 500 mM, aliquoted, and stored at -20 °C. Cross-linker solution (5 μl) was added to 45 μl oligonucleotide (100 μM) and incubated overnight at room temperature. Excess of unbound cross-linker was removed by gel filtration with Zeba Desalt Spin Columns (Pierce, Thermo Scientific). The final DNA substrate was generated by annealing to a complementary strand labeled with Cy5 at the 5' end.

For the intermolecular cross-linking of $MutS^{R420C/D835R}$ in sliding clamp state on oligonucleotides, a 59 bp dsDNA with a centrally located G/T mismatch with 3' and 5' digoxigenin ends was used (Table 4.1). For EMSA experiments oligonucleotide with the identical sequence was used with 5' Cy3 modification. As a competitor for cross-linked $MutS^{R420C/D835R}$ on DNA in MutL recruitment and nicking assays a 30

DNA oligonucleotide	length	Sequence	Modification
Amino-modified top -4	59	5'-GTG CGC AAA TCC AGA CGT CTG TCG ACG TTG GGA AGC TTG AGT ATT CTA TAG TGT CAC CT–3'	Amino modified C6 dT at position 28
Amino-modified top +8	59	5'-GTG CGC AAA TCC AGA CGT CTG TCG ACG TTG GGA AGC TTG AGT ATT CTA TAG TGT CAC CT–3'	Amino modified dC at position 16
bottom for amino-modified	59	5'-/Cy5/AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTT GAC AGA CGT CTG GAT TTG CGC AC–3'	5' Cy5 fluorophore
bottom for amino-modified	59	5'-AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTT GAC AGA CGT CTG GAT TTG CGC AC–3'	-
Top	59	5'-/Cy3/TGA AGC TTA GCT TAG GAT CAT CGA GGA TCG AGC TCG GTG CAA TTC AGC GGT ACC CAA TT-3'	5' Cy3 fluorophore
Top	59	5'-TGA-AGC-TTA-GCT-TAG-GAT-CAT CGA GGA TCG AGC TCG GTG CAA TTC AGC GGT ACC CAA TT-3'	-
Digoxigenin-modified bottom	59	5'-/Dig/AAT TGG GTA CCG CTG AAT TGC ACC GAG CTT GAT CCT CGA TGA TCC TAA GCT AAG CTT CA-/Dig/-3'	5' and 3' Digoxigenin
No modification top	30	GCA TCA TCG AGG ATC GAG CTC GGT GCA ATT	-
Digoxigenin-modified bottom	30	5'-/Dig/-AAT TGC ACC GAG CTT GAT CCT CGA TGA TCC-/Dig/-3'	5' and 3' Digoxigenin

Table 4.1: Oligonucleotides for cross-linking MutS$^{420C/D835R}$ on DNA and MutS$^{220C/D835R}$ to the modified DNA, and trapping MutS.

bp dsDNA with a centrally located G/T mismatch with 3' and 5' digoxigenin ends was used (Table 4.1).

The 3196 bp circular G/T mismatch DNA substrates were created by primer extension and ligation essentially as described earlier using the circular single-stranded DNA derived from phagemid pGATC1 as a template[120].

Protein labeling

To labele MutL with Alexa Fluor 647 (AF647-MutL), MutL^{297C} was diluted to 40 μM in 150 μl of HPLC buffer (10 mM HEPES/KOH (pH 8.0), 500 mM KCl and 1 mM EDTA) and labeled with a 5-fold molar excess of Alexa Fluor 594 maleimide (Invitrogen, Thermo Fisher Scientific, Waltham, MA) for 2 h on ice in the dark according to the manufacturer's instruction. Excess dye was removed using Zeba Spin Desalting Columns (Thermo Fisher Scientific, Waltham, MA) and the degree of labeling was determined from the absorbance spectra (NanoDrop 1000, Thermo Fisher scientific) (DOL 95 $\pm$ 5 %, n=3) as described previously[78]. Then the solution was aliquoted and flash-frozen in liquid nitrogen and stored at -80 °C.

The labeling of position 420 in the analytical experiments was performed under identical conditions as in the cross-linking experiments except that a dye in an identical concentration was used instead of the cross-linker.

Cross-linking

Analytical cross-linking of single cysteine MutS^{R420C} dimer variant (2 μM) was performed on 59 bp DNA (1 μM) containing digoxigenin ends blocked with anti-digoxigenin Fab fragments (4 μM) at room temperature in buffer FB150 (25 mM Hepes-KOH pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 0.005% Tween 20) in the presence or absence of ATP (100 μM) followed by the addition of cross-linker MTS-4-MTS (50 μM) (Figure 4.1A).

Protein-DNA cross-linking was performed by adding a single cysteine MutS^{R220C} dimer variant (2 μM) to the modified DNA duplex (1 μM) pre-incubated in buffer FB150 (25 mM HEPES/KOH, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$) with 1 mM ATP. The mixture was kept on ice overnight. The cross-linking was performed in a volume of 10 μl. To separate cross-linked MutS/DNA complex from uncross-linked proteins 8 % SDS-polyacrylamide gels electrophoresis (SDS-PAGE) was used. DNA was observed by fluorescence of Cy5 dye (Figure 4.1B).

The formation of the disulfide bond occurred under identical conditions as in the cross-linking experiments except for the addition of cross-linker.

FRET MutL recruitment assay

Circular DNA with G/T mismatch (GT#1a) and Alexa Fluor 647 label[120] were diluted to 1.8 nM (or 3.8 ng/μl) in 200 μl buffer FB150 (HEPES 25 mM, $MgCl_2$ 5 mM, Tween 20, KCl 150 mM, pH 7.5) at room temperature in 96-well-plates. The intercalating dye SYTOX Blue was used to stain DNA and it acts as a donor (Figure 4.2). MutS$^{R420C/D835R}$ (40 nM monomer) was added to the circles with incubation at room temperature for 150 seconds. Where indicated, competitor DNA (0.1 mM, 30 bp with single G/T mismatch, DIG-label at both ends and blocked with anti-DIGFab

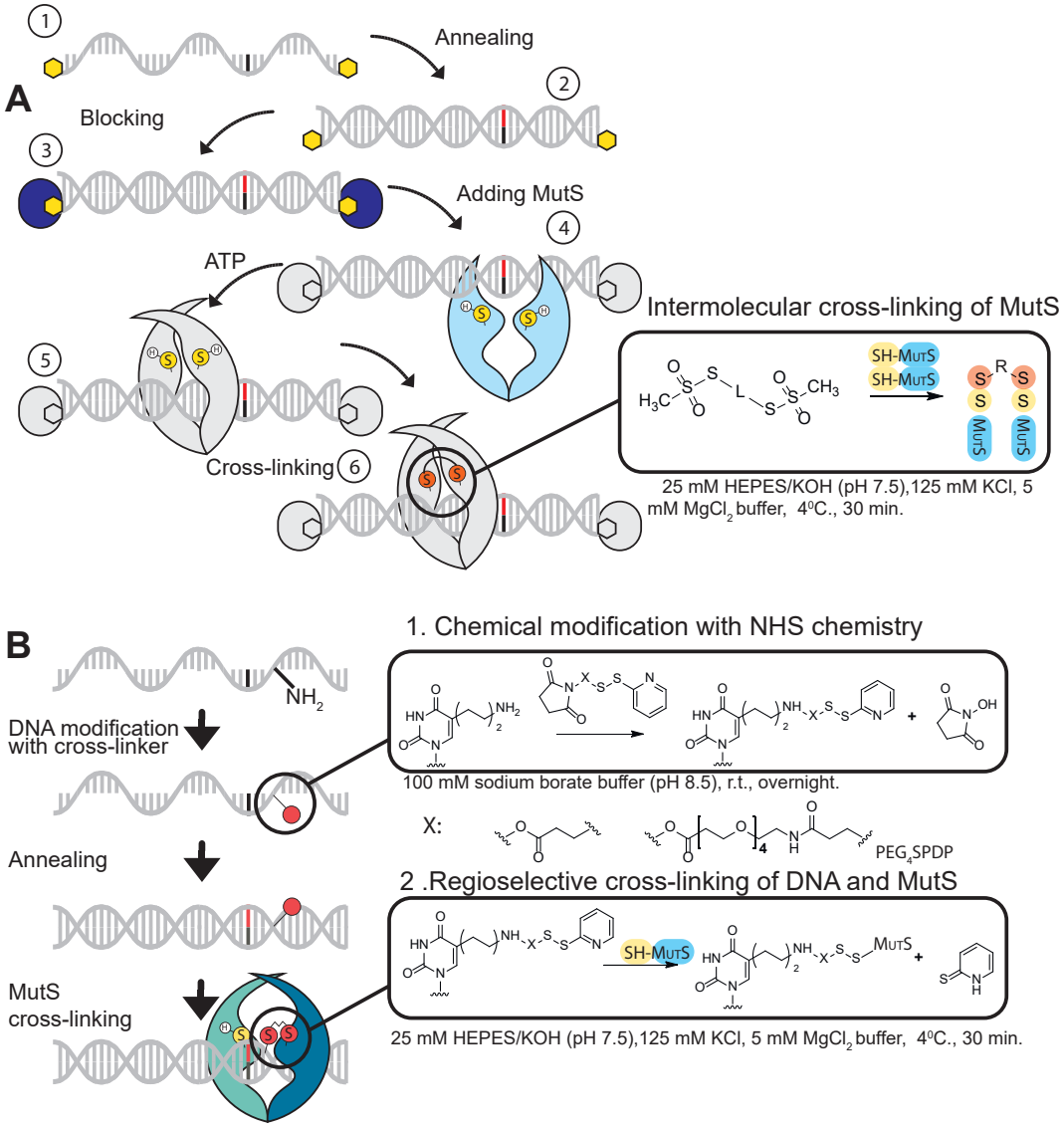


Figure 4.1: **(A)** Overall intermolecular MutS$^{420C/D835R}$ cross-linking workflow. **(B)** Overall MutS$^{220C/D835R}$ to DNA cross-linking workflow

fragment (Sigma-Aldrich) was added.MutL (40 nM monomer) single-cysteine variant 297 labeled with Alexa Fluor 594 maleimide in N-terminal domain was added and fluorescence signals were recorded with a fluorescence microplate reader (TECAN-infinite F200, Tecan). Fluorescence intensities were measured in three channels (donor, acceptor, FRET) with the following filter combinations: (donor ex. 405 nm (width 10 nm) em. 485 nm (width 20 nm), acceptor ex. 590 nm (width 20 nm) em. 620 nm (width 10 nm), FRET ex. 450 nm (width 20 mm) em. 620 nm (width 10 nm) filter. Signals were corrected for buffer and spectral crosstalk to obtain the correct signal intensities (supplementary table B.1). Correction factors were obtained from experiments in which only one of the components giving a signal is present.

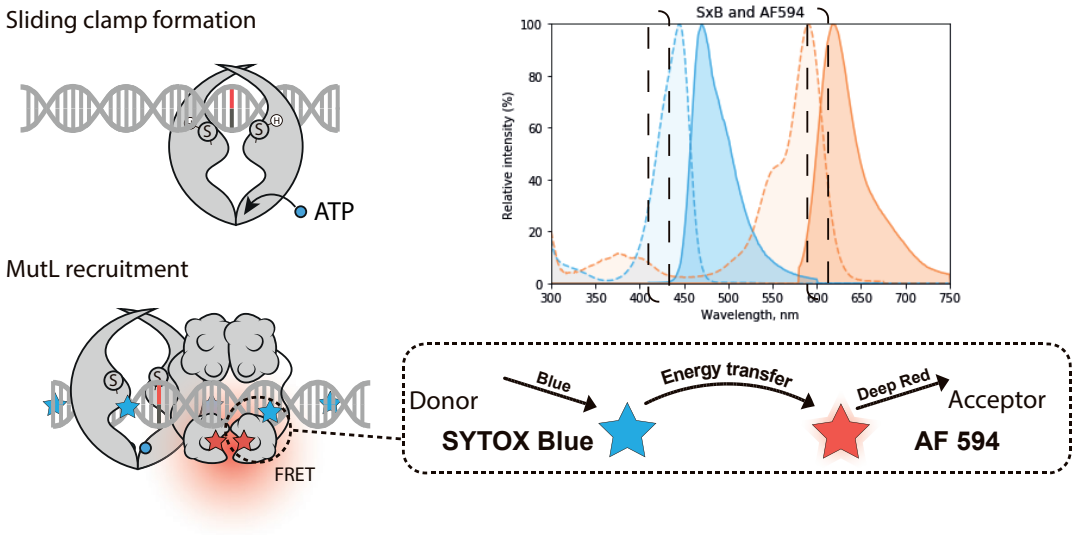


Figure 4.2: FRET-based MutL recruitment assay with SxB blue and AF594

MutH nicking assays

Nicking reactions were carried out at indicated MutS, MutL, and MutH concentrations and 0.5 nM of a 3.2 kb circular DNA substrate with a single mismatch, Alexa647 fluorophore, and hemimethylated GATC site ($GT1^{647}$) and 1 mM ATP at 37°C^{120}. Reactions were stopped by adding an equal volume of 20% glycerol, 1% SDS, 50 mM EDTA, and run on a 0.8 % agarose gel containing 40 μM chloroquine in TAE buffer (Tris 40 mM, pH 8.0, sodium acetate 20 mM, EDTA 1 mM). Gels were scanned on a Typhoon FLA imager (GE Healthcare) with excitation at 633 nm and emission collected via the 670BP30 filter. For cross-linking and competition assays, $MutS^{420C/D835R}$ (40 nM), ATP (50 μM), and $GT1^{647}$ DNA (1 nM) were mixed, followed by the addition of 1 μM of MTS-4-MTS cross-linker (Toronto Research Chemicals) for 2.5 minutes. Where indicated, competitor DNA (0.2 mM, 30 bp with single G/T mismatch, DIG-label at both ends and blocked with anti-DIG Fab fragment (Sigma-Aldrich)) was added for 60 minutes prior to dilution into the reaction mix at final concentrations of 20 nM MutS, 20 nM MutL and 10 nM MutH.

Electrophoretic mobility shift assay (EMSA)

MutS DNA binding experiments were performed at 200 nM 30 bp G/T mismatch Cy5 labeled DNA with digoxigenin ends blocked with anti-digoxigenin Fab fragments (800 nM) and gradient concentration of MutS. The concentrations of ATP (100 μM) and cross-linker (50 μM) were identical to SDS PAGE experiments. Competition experiments were performed with 600 nM MutS and where indicated, competitor unlabeled DNA (2 μM, 30 bp with single G/T mismatch, DIG-label at both ends and blocked with anti-DIGFab fragment (Sigma-Aldrich)) was added. The assay was performed in identical FB150 buffer as in cross-linking assay. The results were analyzed on 4% TPE PAGE gels run in TPE buffer (Tris 90 mM, H_3PO_4 pH 8.2, EDTA 2 mM) at 4°C under constant voltage (40 V, 70 mA) for 150 min using a Typhoon FLA imager (GE Healthcare).

Results

Structure of a MutS intermediate state

Single-molecule FRET data suggested the existence of an intermediate state between MM state and sliding clamp, which was called a “preceding mobile" state (or "bent state 3")[66,138]. Apparently, during the transition state, at least one of the connector domains rotates (Figure 4.3). The sliding clamp structure was obtained by cross-linking of the MutL NTD to the MutS connector domain. It was suggested that the rotation of the MutS connector domain takes place not due to an interaction with MutL NTD, but rather due to ATP binding. To elucidate the mechanism of the connector domain rotation we used cysteine-free MutS variant[79], in which we introduced a single cysteine at position 220 ($MutS^{220C/D835R}$). This position comes close to the DNA in the intermediate state (distance 220Cγ-DNA = 8.0 Å), but not in the mismatch-bound or $MutL^{LN40}$-bound conformations (43 Å). The DNA substrate was modified with a sulfur reactive cross-linker 4 bp away from the G/T mismatch in the middle of a 59 bp DNA oligo[147]. A conjugate of a chemically modified DNA with MutS was formed by thiol-disulfide exchange reaction[148].

Chemical cross-linking of $MutS^{220C/D835R}$ was performed in the absence or presence of ATP, ADP, and a mismatch with modification +4 bp from G (for “A” domain) or - 8 bp from G (for “B” domain). In the absence of ATP but the presence of mismatch-containing or homoduplex DNA, a significant reduction in the cross-linking yield was observed (Figure 4.4A lanes 2 and 3). In contrast, the highest yield of cross-linking was observed in the presence of ATP and DNA containing a G/T mismatch (Figure 4.4A lane 4). In the presence of homoduplex DNA and ATP, the yield is much lower. Cross-linking of the complex was successfully performed for both A and B domains (Figure 4.4B and C). The cross-linking yield for SPDP cross-linker after 30 min on ice was significantly lower than overnight (Figure 4.4D). Also, we compared the yields for a long cross-linker for 30 minutes for room temperature and on the ice at different ATP concentrations. Yields are significantly higher for room temperature experiments. The maximum yield was observed at 1 mM of ATP (Figure 4.4E).

In presence of ATP and a mismatch, which are needed to transform MutS into the clamp state, we obtain significant cross-linking yield, but not in the absence

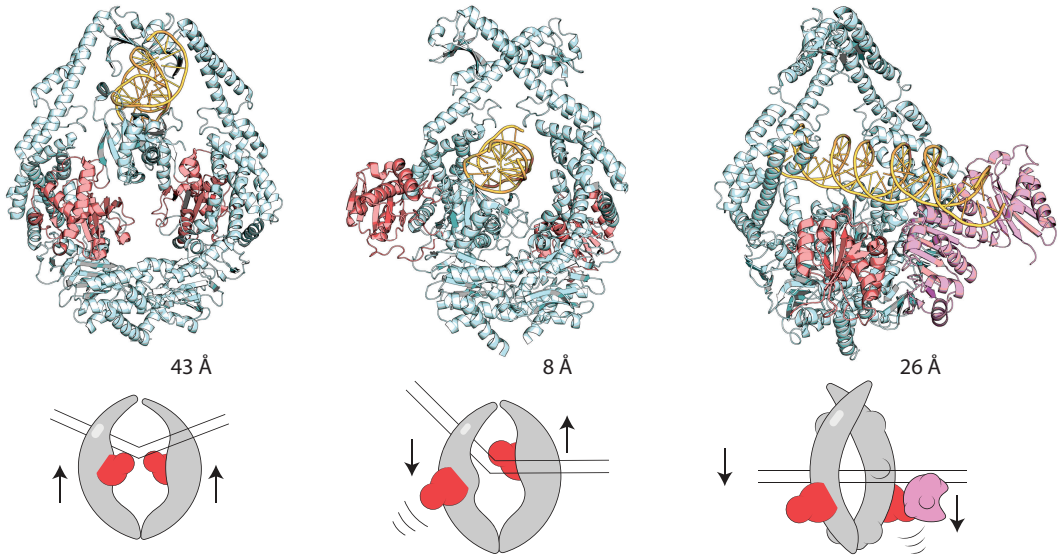


Figure 4.3: MutS intermediate state. **(A)** Connector domains in the mismatch-bound state are rotated upwards (pdb id: 7AI6). **(B)** In the intermediate state, only one connector domain is rotated and another one with a mismatch domain interacts with DNA (pdb id: 7AI7). **(C)** In sliding clamp state both connector domains are rotated (pdb id: 7AIB)

of ATP or when using a homoduplex DNA substrate (Figure 4.4A lane 2 and 3). Taken together, these data show that rotation of the MutS connector domains can take place in the presence of ATP without the necessity of interacting with MutL.

Our data are in agreement with one of the states which were obtained by Cryo-Em. The structure is an intermediate state between MM bound state and sliding clamp state "MutL^{LN40}" (Figure 4.3B). MutS dimer locates near open ends of DNA in the presence of ATP. The clamp and lever domains changed their positions pushing the DNA down similar to the sliding clamp state. As expected, the main feature of this structure is specific connector domain rotation. Monomer B has a similar shape as in the sliding clamp state with ≈180° rotation of the connector domain with an invisible mismatch domain. The connector domain of monomer B is slightly rotated and its contacts with DNA introducing phenylalanine residue 36 between two basepairs which mimics mismatch DNA. This position of the connector domain proposes an idea that this state can be a short transient state between mismatch bound and sliding clamp states. Since the cross-linking occurs 27 bp away from the end of the DNA, this cross-linking is consistent with the notion that an intermediate structure is similar, if not identical, to the "preceding mobile" state that MutS adopts when switching from the mismatch-bound to MutL-bound state.

MutS siding clamp activity on DNA

Using FRET assay and single-cysteine MutS^{449C} it was shown that protein has at least two different conformational states[78]. When double-labeled MutS^{449C} was bound to the blocked end DNA containing a G/T mismatch, FRET between these two different dyes increased upon ATP addition[78]. These results are well correlated with the distances in crystal structures pdb: 1e3m and 5akb, which correspond to the mismatch and sliding clamp states, respectively. The most interesting mutation

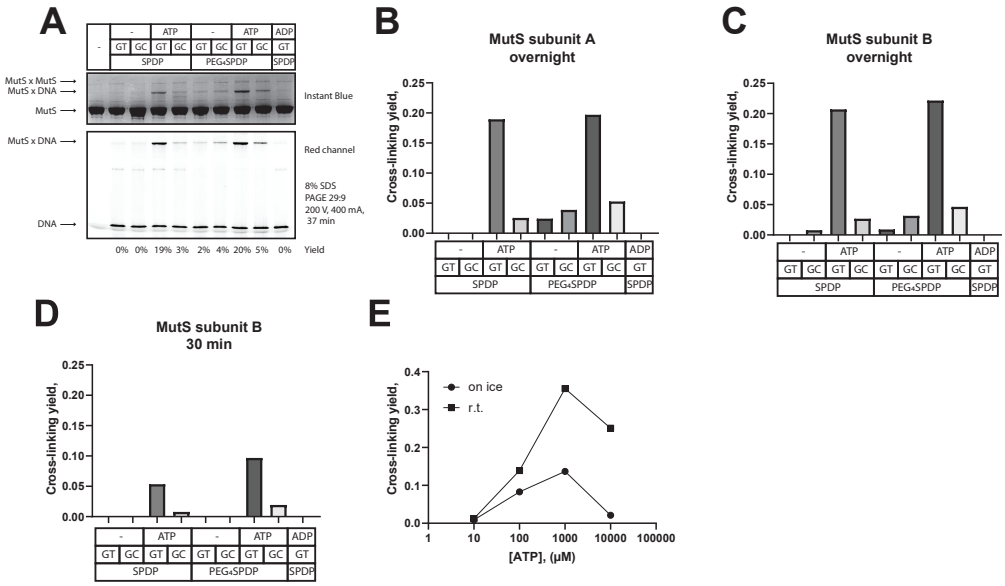


Figure 4.4: Cross-linking MutS$^{220C/D835R}$ to the DNA on the 8% SDS gel and cross-linking yields. **(A and B)** Cross-linking A domain and **(C)** B domain on ice overnight. **(D)** Cross-linking B domain on ice after 30 minutes. **(E)** Cross-linking B domain at different ATP concentrations on ice and r.t. after 30 minutes.

point is position 420, which dramatically changes its location during conformation changes. In the closed state (sliding clamp state), clamp domains must cross and the distance between residues 420 is 5.5 Å (Figure 4.5A and B). Single-cysteine MutS$^{R420C/D835R}$ was designed with specifically placed cysteine residues in core domain for chemical cross-linking that allow protein trapping in the "closed" sliding clamp state on DNA with the short distance between residues.

The short-range cross-linker 1,4-butanediyl bismethanethiosulfonate (MTS-4-MTS or X4) was used to trap sliding clamp state on DNA with the single cysteines in a cysteine free MutS at position 420. This variant can be cross-linked only in the clamp state of MutS (distance 420Cγ-420Cγ = 5.5 Å), but not in the mismatched form bound to the DNA (distance 420Cγ-420Cγ = 40 Å) (Figure 4.5A and B). In the presence of mismatch oligonucleotide with blocked ends and ATP, the cross-linking yield as expected was about 60% (Figure 4.5C and D lane 6), compared to 8% in the presence of mismatch DNA alone or 16% in the absence of DNA and ATP. Low yield for mismatch DNA alone can be explained by the long distance between two cysteine residues and/or by physical blocking access of the residues shielded by DNA from modification. Modification of these residues by maleimide dye of small sizes is significantly reduced in the presence of mismatch DNA alone (Figure 4.5A and F). Surprisingly, the reaction yield of the protein alone was higher than for the mismatch state. the distance in the structure 5yk4 between the 420 residues is about 50 A, which exceeds the length of the cross-linker (range 4-10 Å). The result shows the flexible and dynamic nature of these domains.

Based on these data we hypothesized if MutS can create a disulfide bond without cross-linking. The formation of a disulfide bond requires all the components that were used in the cross-linking experiments (mismatch end-blocked DNA and ATP) (Figure 4.6A lane 5). The reaction was stopped by adding the dye AF488 with a reactive malamide group that joins free cysteine. It not only stops the formation of

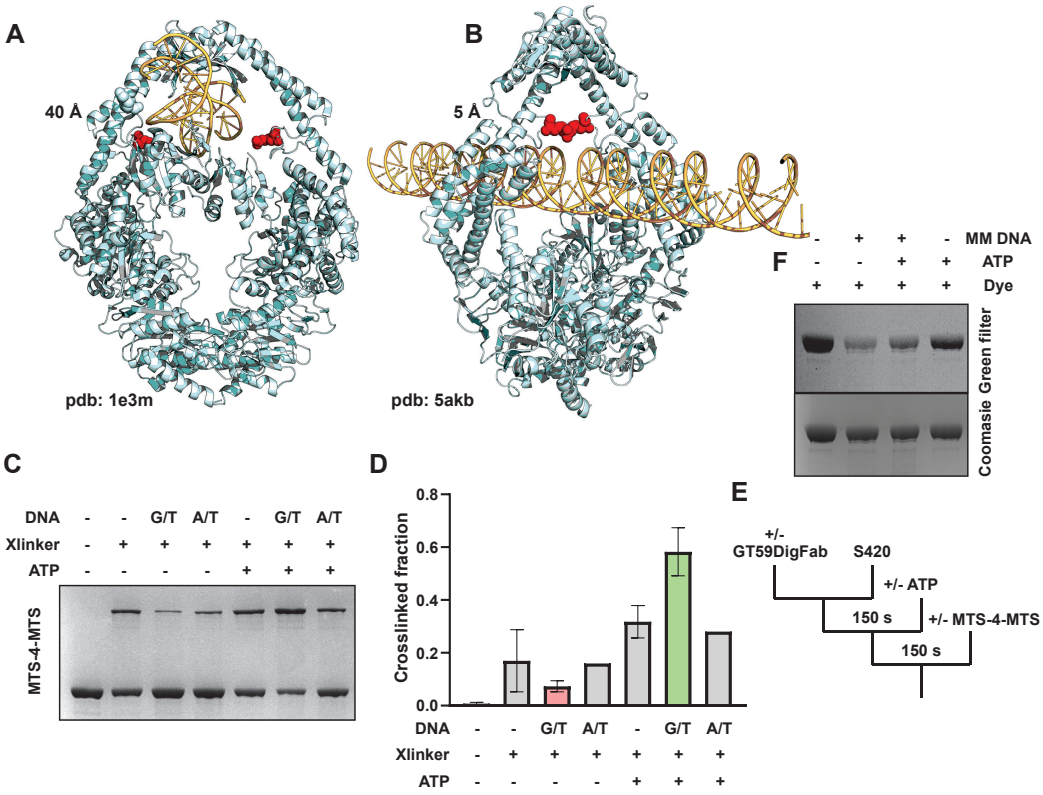


Figure 4.5: MutS structures with Arg420Cys position shown as red spheres (pdb-codes: 1e3m **(A)** and 5akb **(B)** with modeled B-DNA). Distances between position 420 in MutS structures 1e3m and 5akb. **(C)** SDS-page analysis of cross-linking experiment and yields **(D)** between residues 420 of MutS monomers A and B. **(E)** cross-linking scheme of $MutS^{420C/D835R}$ on DNA. **(F)** $MutS^{420C/D835R}$ modification with AF488.

a disulfide bond but also provides an opportunity for the quantitative determination of the reaction conversion.

The first lane corresponds to a single protein and will be taken as 100% of the original substance. On subsequent lanes, cross-linking occurs with a small yield. Greater conversion is observed with the addition of blocked DNA, both short 30 bp and 59 bp long with ATP (lanes 5 and 9). The output of cross-linking also depends on the ATP concentration (Figure 4.6B). The reaction yield increases with increasing ATP concentration. Taken together, the data show the ability of $MutS^{R420C/D835R}$ to create a disulfide bond in the absence of a cross-linker. These data provide evidence that the distance between positions 420 can be much shorter as it was shown in the structure, thereby indicating the dynamic nature of this complex.

Chemical trapping results in a stable shift

Then we asked how many proteins were cross-linked on DNA. In order to distinguish between species cross-linked on DNA from other ones, we used EMSA. To prevent MutS from sliding from the DNA ends, it was necessary to block these ends. Fab fragments interact with Digoxigenin-ends resulting in highly efficient end blocking. Using a native gel, the optimal ratio of the blocking agent and the substrate was established. Complete blocking is achieved with a fourfold excess of Fab fragments

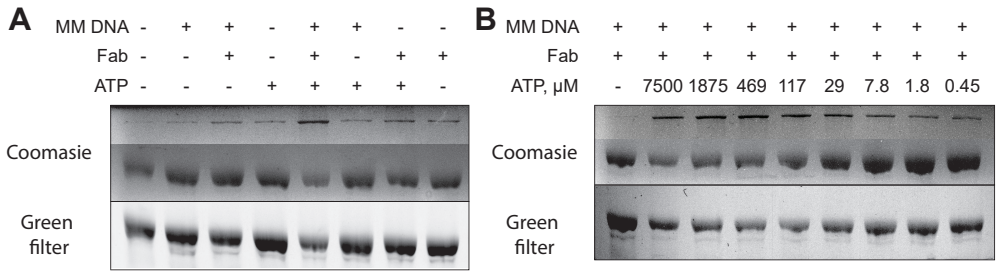


Figure 4.6: **(A)** Formation of the disulfide bond in MutS$^{R420C/D835R}$. **(B)** ATP influence on disulfide formation

(Figure 4.7 C).

MutS DNA interaction was analyzed on the native gels. The lowest band on the gels corresponds to double-stranded DNA (Figure 4.7A lane 1). Adding Fab fragments leads to a stable shift (Figure 4.7A lane 2). By increasing the concentration of MutS, the disappearance of the band corresponding to the blocked DNA is observed. At a concentration of 0.6 μM of the monomer MutS$^{420C/D835R}$ and a DNA concentration of 0.2 μM (that is, the ratio of double-stranded DNA to the dimer of MutS$^{420C/D835R}$ was 1: 1.5), a smear is observed in the absence of cross-linking, and in the presence of MTS-4-MTS there are two clear bands (Figure 4.7A lane 7), which correspond to one and two dimers on a molecule, respectively. The results suggest that cross-linking complexes are more stable on DNA than uncross-linked. At the same time, in cross-linking experiments, it is impossible to distinguish complexes that have been cross-linked to DNA from complexes that also interact with DNA but are not cross-linked.

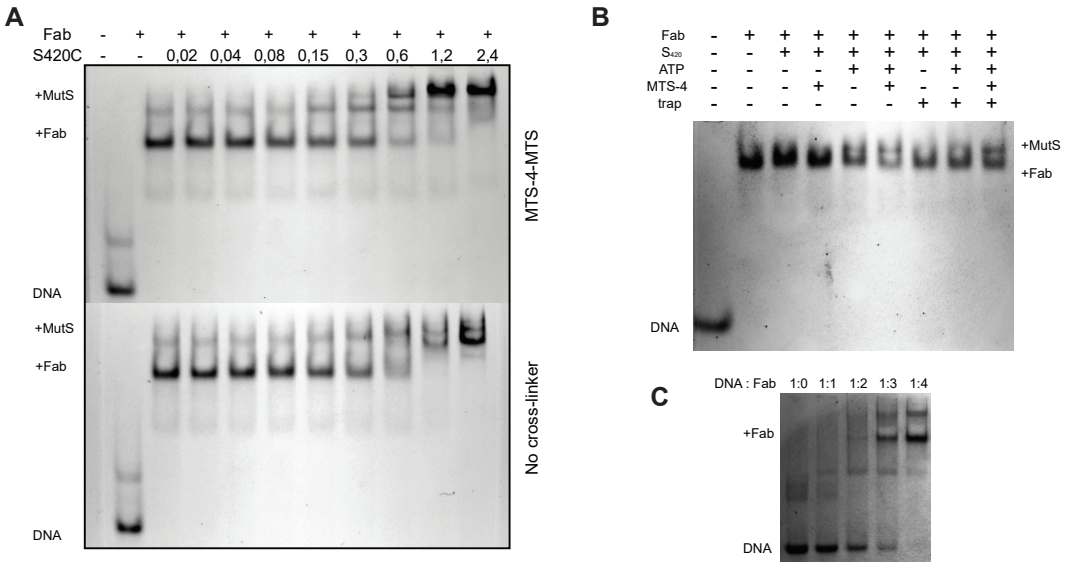


Figure 4.7: **(A)** Gradient of MutS$^{420C/D835R}$ with blocked DNA +/- cross-linker. **(B)** Shift experiments with cross-linked MutS$^{R420C/D835R}$ and trap on TPE 4% polyacrylamide gel. **(C)** Blocking oligonucleotide duplex with the gradient of Fab fragments. DNA with digoxigenin ends and Fab fragments

Cross-linked MutS on DNA in the presence of competitor

In order to distinguish between these possibilities the stability of the cross-linked $MutS^{R420C/D835R}$ on DNA in the presence of a large excess of unlabelled mismatch oligonucleotides was tested.

To avoid binding of two MutS dimers to DNA a 30 bp oligonucleotide with digoxigenin ends was used. When unlabeled DNA with Fab fragments is added to MutS, the upper clear band with cross-linked MutS does not disappear (Figure 4.7B lane 9). MutS is on the DNA trapped and cannot leave it. When a reducing agent is added, the upper band disappears and a single band appears (Figure 4.7B lane 10) characteristic of unbound DNA with Fab fragments on the sides (Figure 4.7B lane 1). In the case when the DNA is not labeled, it makes no sense to use an excess of unlabeled DNA. Clear stripes are formed only when all components are in solution (Figure 4.7B lane 5). In the absence of any of the components, it is impossible to obtain a complex (Figure 4.7B lanes 1-4, 7). Since we used a recoverable cross-linker when adding a reducing agent such as DTT, the upper clear bar disappears (Figure 4.7B lane 6). When unlabeled GT30 is added, the upper band in the case with cross-linked MutS does not disappear (Figure 4.7B lane 8). MutS is on the DNA trapped and cannot leave it. Without a cross-linker, the upper band disappears and a single band appears (Figure 4.7B lane 7). Taken together, these data show that DNA cross-linked complexes are more stable in the presence of a competitor.

$MutS^{R420C/D835R}$ cross-linked on DNA recruits MutL in the presence of a competitor

MutS, MutL protein require some space for interaction with DNA. For MutS this range exists between 22 and 30 bp[48]. MutL also requires additional space for binding. It's still unknown how many MutL proteins can be recruited by MutS on DNA. On 59 bp oligonucleotide with blocked ends, a high FRET signal between MutL and DNA is observed (Figure 4.8B). On oligonucleotide with a length of 30 bp and blocked ends it's hard to observe any FRET signal. In this case, short DNA can act as a negative control or a trap for MutL recruitment by MutS (Figure 4.8C). Since on short oligo there is enough space only for MutS binding, MutS can interact with both long and short DNAs and there is no preference for binding. It is shown the highest FRET as expected was observed in a case with 59 bp DNA (Figure 4.8A green points). In the second case after adding short 30 bp DNA to the longer oligonucleotides a dissociation of MutS-MutL complex (decrease of FRET signal) takes place (red points). The third lane shows the case when MutS was preincubated in the presence of short and long DNA (blue points). In this case, the medium FRET is observed and it's similar for the second case after adding short DNA. And in the fourth case, the lowest FRET takes place, since on short DNA there is no space for MutL recruitment. But after adding long DNA, there is an increase of the FRET signal, because of redistribution of MutS (purple points).

Then we asked if cross-linking as in EMSA experiments can trap the protein on DNA in the presence of a competitor. In order to exploit the competitor excess, we used circular substrates as the substrate on which MutL multi-recruitment took place. Recruitment of MutL to mismatch-circular DNA[120] by MutS was monitored by a variation of the MutS-DNA FRET assay using Alexa Fluor 594 labeled MutL297[78]. In the case of MutS cross-linked on GT1a circle in the presence of

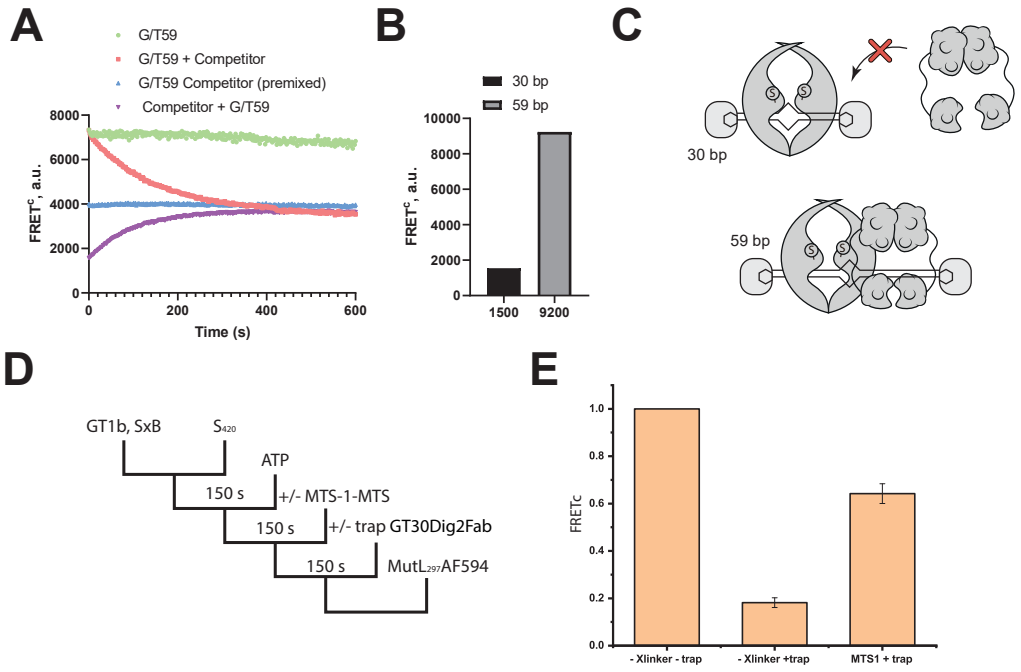


Figure 4.8: **(A)** Kinetic of MutL redistribution in the presence of 30 bp trap DNA. **(B)** MutL recruitment on short 30 bp oligonucleotides is strongly reduced. **(C)** There is no space for MutL recruitment on 30 bp oligonucleotides. **(D)** Scheme of cross-linking/trapping experiments in MutL recruitment assay. **(E)** Cross-linked MutS$^{R420C/D835R}$ is able to recruit MutL even in the presence of the trap (lane 3).

ATP and trap DNA, FRET between DNA and MutL was significantly higher than without cross-linking (Figure 4.8E). The results suggest that MutS cross-linked on circular DNA can recruit MutL even in the presence of competitor excess.

MutS$^{R420C/D835R}$ cross-linked on circular G/T DNA activates MutH in the presence of a competitor.

A previous MutL recruitment experiment showed that the cross-linked MutS could recruit MutL in the presence of the DNA trap. Next, we determined the effect of the cross-linking on the activation of MutH via a DNA nicking assay using a circular substrate containing a single mismatch and hemimethylated GATC site[120]. Here, MutS$^{R420C/D835R}$ is first cross-linked on the mismatched circular DNA in the presence of ATP, followed by the addition of MutL and MutH (Figure 4.9C). The nicking activity of MutH is unaltered by the presence of the cross-linker, indicating that the cross-linked MutS clamp is fully capable of binding MutL and activating MutH (Figure 4.9A and B). Finally, an excess of competitor mismatched DNA was added, after cross-linking but before the addition of MutL and MutH. In the absence of a cross-linker, the nicking activity is reduced to less than 20%, due to the release of MutS from the circular substrate DNA and rebinding to the competitor. In contrast, in the presence of the cross-linker, the MutH nicking activity is retained at 90%, indicating that the MutS is retained on the DNA. This shows the clamp state of MutS is both necessary and sufficient for MutL recruitment and subsequent activation of MutH.

Cross-linking of the MutS$^{R420C/D835R}$ before adding DNA significantly reduces nicking activity in contrast when cross-linked is done after conformational changes on the DNA (Figure 4.9E). These results indicate that for the activation of MutH latent endonuclease activity, MutS must be in a state in which the core and clamp domains are crossed and the DNA is inside the protein. Moreover, the opening of these domains is required before DNA loading. These data are consistent with the structure of the sliding clamp state.

Discussion

The different structures and smFRET data show how complex and dynamic is the nature of MutS protein and the functions in which it takes part[66]. The structures of such dynamic proteins are very difficult to obtain in solution and in particular the structures of significant, active, or transitional states[54,78]. It was shown by single-molecule data existence of intermediate states between mismatch and sliding clamp. To check whether rotation of connector domains is induced by ATP and mismatch and not by MutL, a special single-cysteine variant MutS$^{R220C/D835R}$ was designed, in which reactive positions in the mismatch structure were far from DNA, and during the transition, these amino acids should approach to DNA so close that a chemical reaction between the thiol group of cysteine and the pyrimidyl group of the cross-linker was possible. As the result showed, a high yield of the cross-linker reaction was observed, only in the presence of ATP and mismatch. In all other cases, we did not observe cross-linking. The intermediate state structure obtained by our colleagues has features of mismatch and sliding clam states. The mismatch domain

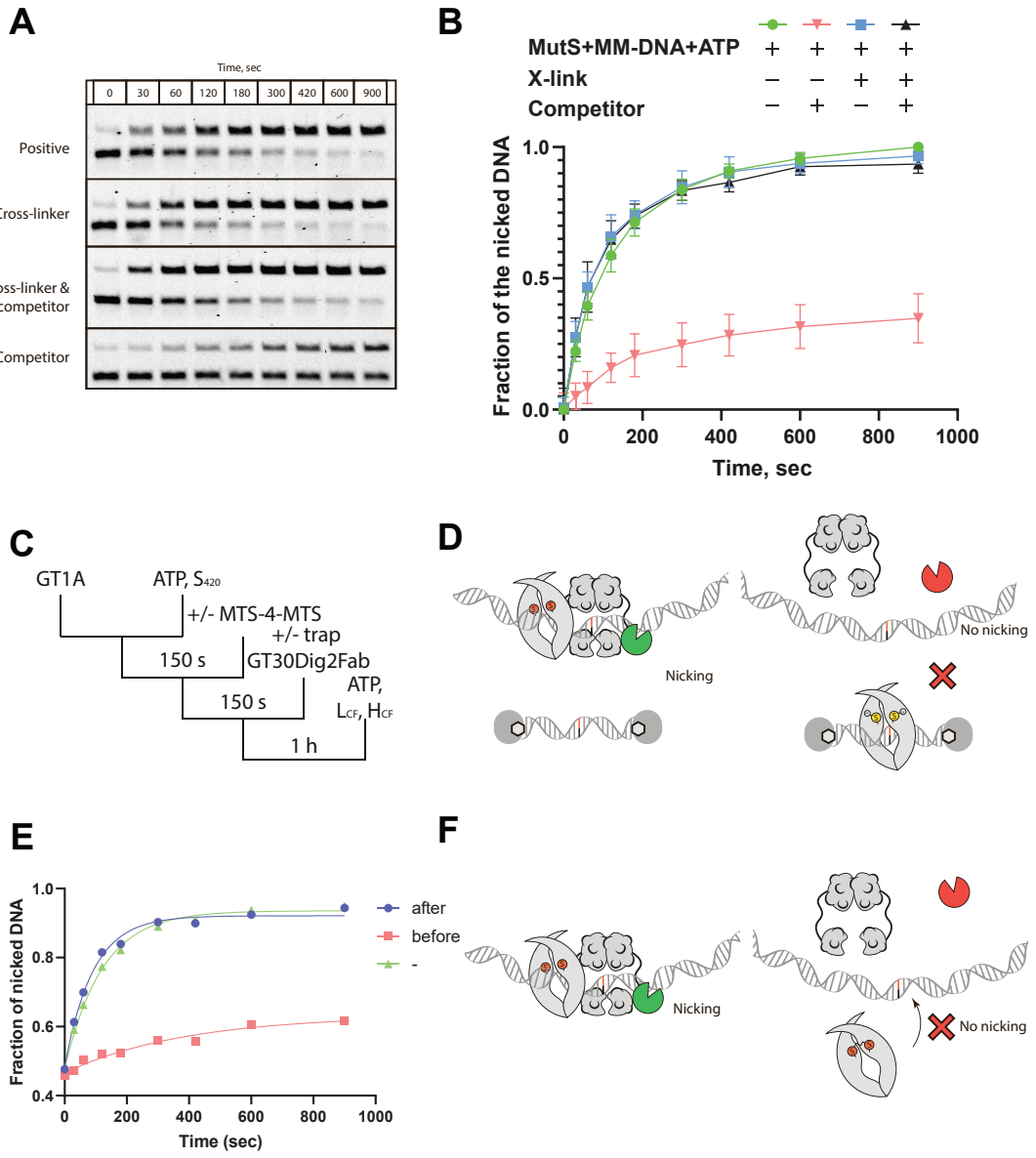


Figure 4.9: **(A)** Fraction of the nicked DNA on an agarose gel. **(B)** Native and cross-linked MutS show similar activities, but only cross-linked MutS with competitor DNA (black triangles) can show activity in the presence of a competitor. **(C)** MutH activation in the presence of competitor/cross-linker assay scheme. **(D)** MutH DNA nicking assay, requiring the action of MutS and MutL. **(E)** cross-linking before adding DNA results in strong reducing nicking activity. **(F)** Schematic representation of cross-linking before adding DNA.

of one monomer holds DNA as in mismatch structure and the connector domain in another subunit is rotated 180° with an open interface for MutL interaction as in sliding clamp. This rotation was caused not by MutL, but rather by DNA mimicking mismatch and the presence of ATP. This structure illustrates the existence of such a transitional state predicted by cross-linking data. We showed that rotation of the connector domain takes place even in the absence of MutL and is induced by ATP.

To confirm that DNA is in the center of the crystal structure, intramolecular site-specific cross-linking was used. A single cysteine variant $MutS^{R420C/D835R}$ was designed where, upon conformational changes of the two core domains into a sliding clamp structure, cysteine in one subunit was close to the same position on the second dimer. At the same time, in a mismatch state, these amino acids are located at a great distance from each other and there is a physical barrier in the form of DNA between them. The high yield of cross-linking can be observed only when MutS undergoes conformational changes to the sliding clamp state in the presence of ATP and, mismatch DNA. In contrast, to the DNA alone the yield of cross-linking reaction and modification by maleimide dye is strongly reduced, which can be explained by the long-distance between residues in the structure 1e3m and physical shielding of residues by DNA[54]. Cross-linking in the apostate of MutS can be performed despite the long distance between cysteines, which can be explained by the dynamic nature of the core and clamp domains in the dimer[62]. MutS can form a disulfide bond even in the absence of the cross-linker only when mismatch DNA and ATP are present. This indicates that the distance between 420C residues can be even shorter than in the crystal structure MutS-LN40.

We assumed that the sliding clamp state is an active state of the MutS and is responsible for MutL recruitment and MutH activation. If this state of MutS can be trapped on DNA using cross-linking, we can show its activity and significance. To show the significance and activity of the protein, we designed an experiment with a competitor. As a competitor, we used a short end-blocked DNA from the mismatch, where free MutS migrated. We observed in experiments on circular substrates, that the protein is not cross-linked on DNA migrates into this DNA-trap and all subsequent pathway reactions are greatly reduced. In the case of cross-linking, the activity is significantly higher, which was shown experimentally for recruiting MutL and nicking activity of MutH in the presence of a DNA-trap and cross-linker. Another interesting point is cross-linking before adding mismatch DNA, the protein completely loses its activity, thereby indicating that the upper part of the protein has to be opened up for DNA and then the DNA goes down, as predicted by the structural data. Future research could shed light on the DNA escape mechanism. This process can be associated with ATP hydrolysis and further protein conformational changes not previously described. Further study of this problem by obtaining new structures, as well as a combination of this method with FRET approaches, as well as single-molecule experiments, could answer this question. Defects in the DNA repair system can lead to colorectal cancer, but it is often very difficult to diagnose whether such meaningless mutations in MMR genes are critical for protein function[160]. The data obtained in the results of structural and functional analyses will help create a map of critical mutations in the structure and predict their pathogenicity. Moreover, we will be able to predict which stage in the DNA repair mechanism does not work, for example, protein conformational changes or protein-protein interaction. Many pathogenic mutations have been found in different regions

of the proteins. A mechanistic understanding of the process will help in the future in the development of new and effective therapeutic methods for the diagnosis and treatment of colorectal cancer.

This work is part of a publication "The selection process of licensing a DNA mismatch for repair" in the journal Nature Structural & Molecular Biology carried out in collaboration with the groups of Prof. Dr. Meindert Lamers, Prof. Dr. Titia Sixma, and Prof. Dr. Joyce Lebbink.

Chapter 5

Covalent trapping of MutL and MutH for functional and structural studies

Vladislav Kunetsky[1], Alessandro Borsellini[2], Meindert H. Lamers[2], Peter Friedhoff[1]

1 Institute for Biochemistry, FB 08 Justus Liebig University, Heinrich-Buff-Ring 17, 35392 Giessen
2 Department of Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands

Abstract

The DNA mismatch repair system is a key process in the correction of replication errors in all three kingdoms of life. The core of the mismatch DNA repair process in *E. coli* involves three proteins MutS, MutL, MutH. The process in *E. coli* starts with the recognition of mismatches by the ABC-ATPase MutS, which clamps on and bends the DNA. After conformational changes caused by mismatch binding and ATP, another ATPase, MutL binds MutS to form a ternary complex leading to activation of enzymes required for strand discrimination, error removal, and DNA re-synthesis. Despite years of research, the mechanism of activation of MutH latent endonuclease activity is not clear. The structure of the MutL-MutH-DNA complexes has remained elusive. One reason for these problems lies in the highly dynamic nature of the complexes involved in the MMR pathway. Using site-specific cross-linking a MutS independent nucleotide-dependent active MutL-MutH complex was obtained. This approach can be useful for structural studies of the active complex.

Introduction

The DNA repair system plays an important role in maintaining the stability of the entire genome. The key proteins of this pathway are evolutionarily conserved MutS and MutL proteins[7]. In *E. coli*, these homodimers activate the latent endonuclease activity of MutH for strand discrimination. In eukaryotic organisms, these proteins have a heterodimer nature and the endonuclease function is performed by MutL. MutL is a multidomain protein that forms a homodimer with a flexible linker separating N-terminal and C-terminal domains (Figure 5.1A)[88,89]. The structure of full-length protein remains unknown. The N-terminal ATP binding domain is structurally similar to the ATPase domain of members of the GHKL family such as DNA gyrase, chaperone HSP90, histidine kinase. ATP binding induces large conformational changes of the protein[52,88,94]. The binding of ATP or ADPNP (non-hydrolyzable ATP-analog) in size-exclusion chromatography leads to a compact state of MutL dimer[88,89,95]. Atomic force microscopy experiments have shown large asymmetric conformational changes of MutL homologous and it was possible to distinguish four different MutL states: “extended”, “one-arm”, “semi-condensed”, and “condensed”[96]. In vitro MutL is sufficient to activate the MutH endonuclease in a mismatch and MutS-independent manner. This activation is observed at low ionic strength (<100 mM). However, at physiological relevant ionic strength [100–160 mM] optimal for DNA mismatch repair, activation of MutH requires mismatch and MutS.

Single-molecule data of *E. coli* MutS, MutL, and MutH on DNA demonstrate the formation of long-lived ternary complex MutS-MutH-MutH[77]. Little is known about the conformational changes of the complex. Up to date no crystal or cryo-EM structure of the full-size ternary complex MutS-MutL-DNA has been obtained. However, the structure of the sliding clamp complex and the N-terminal domain of MutL was generated by using site-specific cross-linking in combination with structural methods[78]. This approach of highly selective cross-linking provides a possibility to trap relevant and short-lived states of active dynamic complexes. Scaling-up of these cross-linking procedures has been effectively used for structural studies using small-angle X-ray scattering[55], X-ray crystallography, mass-spectrometry[106],

and Cryo-EM[71]. Here, we describe the site-directed cross-linking and purification procedure for analytical and structural studies of MutL-MutH active complex.

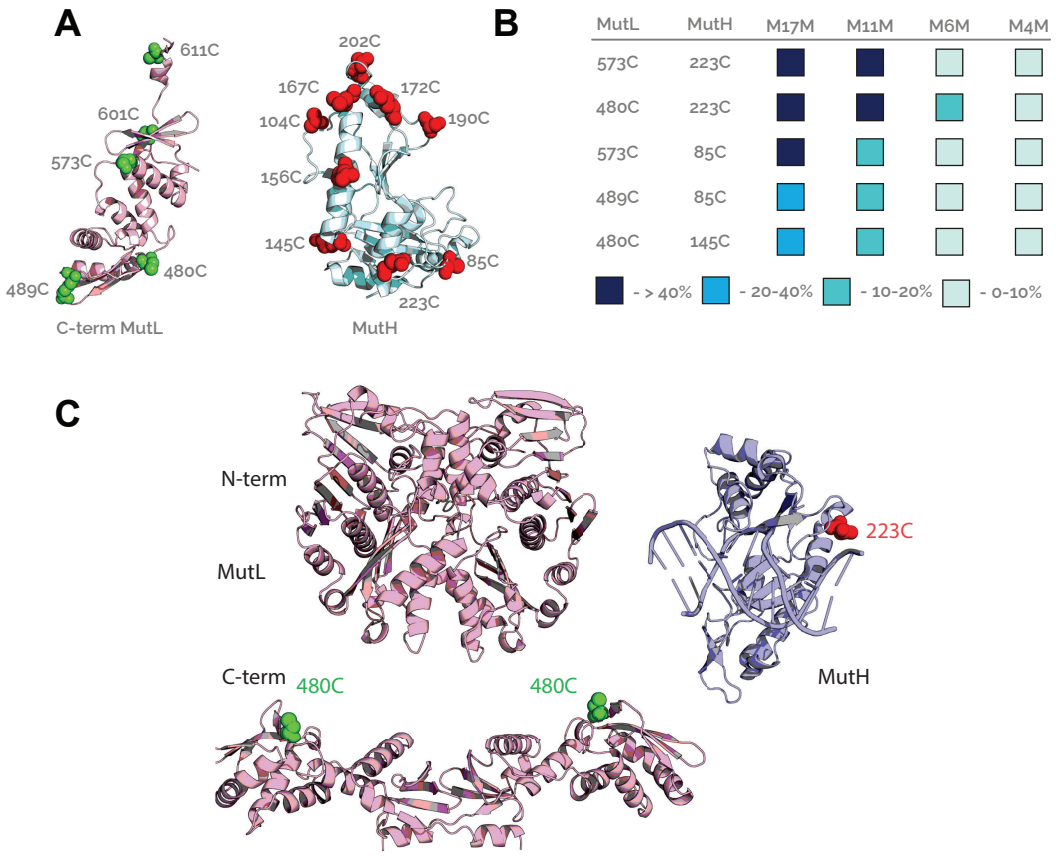


Figure 5.1: **(A)** MutL (pdb 1x9z[89]) and MutH (pdb 2AOQ[161]) single cysteine variants **(B)** Summary of cross-linked result for the different pairs with short and long cross-linkers[162]. The highest yield of cross-linking N-terminal domain to MutH was observed for the pairs 573-223; 480-223 and 573-85. Only 480-223 cross-linking showed a 10-20% cross-linking yield for the short cross-linker MTS-6-MTS. (C) Single cysteine MutL^{480C} (pdb 1b63[95] and 1x9z[89]) and MutH^{223C} (pdb 2AOQ[161]) variants.

Materials and methods

Protein purification

MutS, MutL and MutH proteins were purified as described previously[105,163]. Protein concentrations were determined spectrophometrically using theoretical extinction coefficient (ϵ_{280} nm =73 605 M^{-1} cm^{-1} for MutS, ϵ_{280} nm = 54 270 M^{-1} cm^{-1} for MutL, ϵ_{280} nm = 37 470 M^{-1} cm^{-1} for MutH).

DNA substrates

All oligonucleotides (HPLC grade) were synthesized by Eurogentec (Seraing, Belgium). Two complementary oligonucleotides with G/T mismatch in the center were

diluted in water. As a result of mixing two oligonucleotides, the final concentration of the duplex was 10 μM. The duplex was annealed in a "Biometra T-personal PCR machine" at a temperature of 95 °C with a gradient of a decrease in temperature by 5 °C every five minutes.

DNA oligonucleotide	length	Sequence	Modification
Bottom	59	5'-AGG TGA CAC TAT AGA ATA CTC AAG CTT CCC AAC GTT GAC AGA CGT CTG GAT TTG CGC AC–3'	-
Top	59	5'-/Cy3/TGA AGC TTA GCT TAG GAT CAT CGA GGA TCG AGC TCG GTG CAA TTC AGC GGT ACC CAA TT-3'	5' Cy3 fluorophore

Table 5.1: Oligonucleotides for nicking activity.

Cross-linking of MutH to MutL

Recombinant His_6-tagged MutH and MutL proteins were expressed and purified by Ni-NTA chromatography. MutL and MutH proteins were snap-frozen in liquid nitrogen and stored at -80°C in HPLC 500 buffer (10 mM HEPES–KOH, 500 mM KCl, 1 mM EDTA, 10% glycerol, pH 8.0). Chemical cross-linking of $MutH^{A223C}$ (2 μM) to the $MutL^{480C}$ (4 μM) variant was performed by mixing two proteins for 10 min on ice with 20 μM homobifunctional cross-linker MTS-17-O5-MTS (Toronto Research Chemicals). For preparative cross-linking $MutH^{A223C}$ (41,5 μM) and $MutL^{480C}$ (83 μM) variants were mixed and incubated for 10 minutes on ice in 500 μl in stock HPLC 500 buffer with 415 μM $BM(PEG)_3$ (Thermo Fisher Scientific, Waltham, USA) for 5 min on ice. Complex purification The cross-link reaction mixture was purified of the non-reacted cross-linker by size-exclusion chromatography using 0.5 ml Zeba spin desalting columns (7K) to get rid of excess unreacted reagent. The cross-link reaction mixture was subjected to gel filtration on a Superdex 200 column equilibrated with HPLC 500 buffer at 0.25 a ml/min flow rate. Fractions containing the cross-linked complex eluting between 47 and 49 min were pooled (fraction B6 see supplementary figure B.6), snap-frozen in liquid nitrogen, and stored at -80°C until further analysis.

MutH activation assay

Incision assays for cross-linked MutL-MutH complex were performed in the absence or presence of MutS at different complex concentration and 0.5 nM Alexa647-labeled circular relaxed DNA substrate (GT1A) at 37°C in buffer (25 mM Hepes KOH [pH 7.5], 150 mM KCl, 5 mM $MgCl_2$, 10% glycerol, 0.01 % Tween 20) containing 1 mM NTP (ATP, ADP, ADPNP, and ATPγS). In experiments with MutS, the protein was preincubated in a similar buffer but with 20 μM ATP. Reactions were started by adding a mixture of 20 μM ATP, MutS, and DNA to a mixture of MutL and MutH (cross-linked or not). To destroy MutL-MutH cross-linked complex mixture was incubated with 10 mM DTT in 20 μl of stock HPLC 500 buffer. The reactions (10 μl) were stopped with an equal volume of 20% glycerol, 1% SDS, and 50 mM EDTA (stop buffer) and analyzed on 0.8% agarose gels run in TAE buffer (Tris 40 mM, pH 8.0, sodium acetate 20 mM, EDTA 1 mM) in the presence of

40 μM chloroquine to distinguish covalently-closed double-stranded (ds)DNA substrate from nicked circular dsDNA product. Cysteine-free MutS was used to avoid side reactions of cysteines.

Incision assays on oligonucleotides were performed at 200 nM MutL-MutH complex and 50 nM 59 bp TAMRA labeled dsDNA at 37°C in buffer (25 mM Hepes KOH [pH 7.5], 150 mM KCl, 5 mM $MgCl_2$, 10% glycerol, 0.01 % Tween 20) and containing 1 mM ADPNP. Where indicated, 1 mM reducing agent DTT was added. The reactions (5 μl) were analyzed on preheated for 50 min 17% denaturation urea polyacrylamide gels (urea PAGE) run in TTE buffer (Tris 1 M, Taurine 29 mM, EDTA 1 mM) at r.t. under constant voltage (300 V, 400 mA) for 25 min using a Typhoon FLA imager (GE Healthcare).

Electrophoretic mobility shift assay (EMSA)

MutL-MutH DNA binding experiments were performed at an identical conditions as in MutH activation assay on oligonucleotides. The results (5 μl) were analyzed on 4% TBE PAGE gels run in TBE buffer (Tris 0.1 M, pH 8.3, H_3BO_3 0.1 M, EDTA 2.5 mM) at 4°C under constant voltage (40 V, 150 mA) for $\approx$ 150 min using a Typhoon FLA imager (GE Healthcare).

Results

We applied a chemical trapping approach which was successfully used in previous studies for obtaining new essential structures[71,78]. We chose a pair of proteins scMutH^{A223C} and scMutL^{480C} that gave the highest yield for the short cross-linker reaction. Five combinations of variants resulted in the cross-link yield higher than 20% for any of the cross-linkers. These combinations were further tested using cross-linkers with various lengths (MTS-4-MTS, MTS-6-MTS, MTS-11-O3-MTS, and MTS-17-O3- MTS) (Figure 5.1B). All five combinations resulted in cross-linking formation with the yield higher than 20% using MTS-17-O3-MTS (maximum span of the linker: 22.3 Å), and only two combinations (scMutL^{480C}-scMutH^{A223C} and scMutL^{573C}C-scMutH^{A223C}) using shorter MTS-11-O3-MTS (maximum span: 16.7 Å). An only a single combination of variants (scMutL^{480C}-scMutH^{A223C}) resulted in cross-link formation with this yield using MTS-6-MTS (maximum span: 12.8 Å)[162]. Since cross-linking shows high specificity, it indicates that scMutL^{480C} and scMutH^{A223C} cysteines are located at the physical interfaces of the protein interaction. The cross-linking was performed with a ratio of 2:1 of MutL and MutH with long cross-linker MTS-17-O5-MTS for high yield. MutH almost disappeared after trapping with MutL, showing a high yield of reaction (Figure 5.2A). As the side product appeared MutLxMutL complex as the highest band on the SDS gel. The cross-linked complex corresponds to the band with 100 kDa, another subunit of the complex has a free MutL monomer and runs lower.

We first asked if the activity of the complex in mismatch dependent manner is influenced by cross-linking. We analyzed it by agarose-based MutH activation assay on 3kb circle substrates with one defined mismatch position[120]. The higher band on the gel corresponds to the nicked relaxed DNA. The band running lower corresponds to the initial supercoiled circle substrate. As a positive control, we performed MutS, L, H nicking activity without cross-linker (Figure 5.2E). The DNA was completely

nicked in 2 minutes. A similar nicking rate showed cross-linking complex. cross-linking is not impairing the function of either MutH or MutL during the process of MM- and MutS-dependent activation of MutH (Figure 5.2D and E).

This finding left us with 2 possibles interpretations. Either cross-linking has no effect on the activity of the complex, or the active ones are not cross-linked proteins. To exclude the possibility of uncross-linked MutH activity, we a performed purification of the cross-linked complex by size-exclusion chromatography (Figure (Figure 5.2B). The peak containing MutL and MutLxMutH complex eluted at 11 ml, which was faster than the peak with MutH. For functional studies, peak B6 collected at approximately 12 ml was used (Supplementary figure B.6 and Figure Figure 5.2BB). The complex after treatment with reducing agents (DTT) divided into free MutH and MutL in a quantitative yield. To compare the activity of the cross-linked and non-cross-linked complex, DTT that breaks the covalent bond in the complex was used.

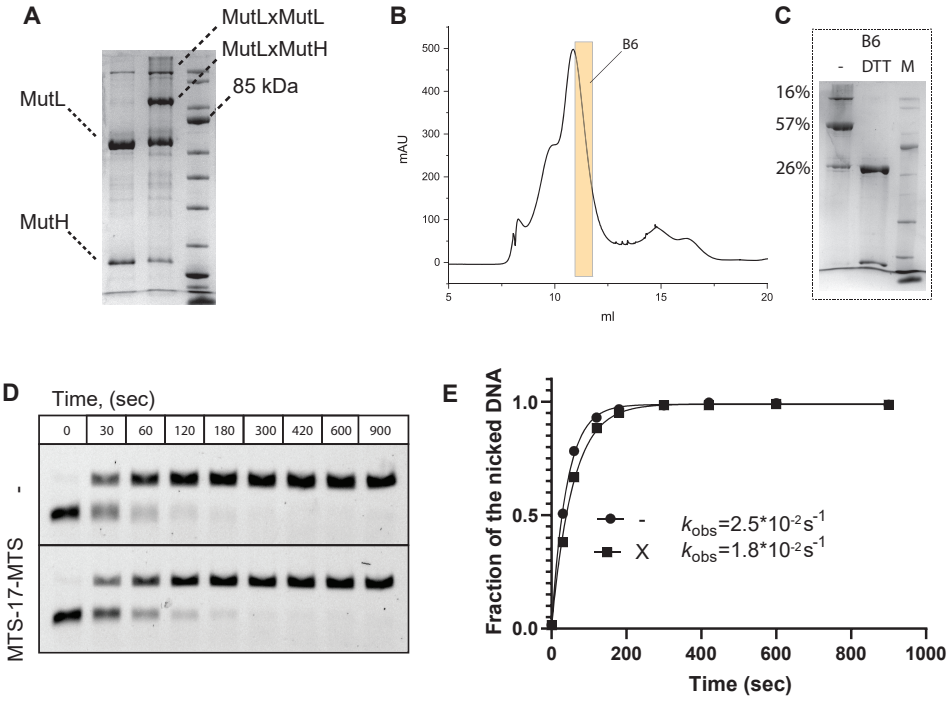


Figure 5.2: **(A)** Cross-linking for analytical purposes on 8% SDS polyacrylamide gel. The band on the top corresponds to the side product MutxMutL. The lower band is a cross-linked complex between MutlxMutH. MutL monomer migrates at 70 kD and the band on the bottom corresponds to the free MutH. **(B)** HPLC Purification of the cross-linked complex was performed on Superdex 200 column with a good resolution allowing separation of free MutH from free MutL and MutLxMutH complex. **(C)** Complex collected at fraction B6 shows an absence of free MutH and can be reduced by adding DTT resulting in the formation of free MutH and MutL. **(D)** To follow the reaction 0.8% agarose gel was used. Nicking activity results in the relaxation of supercoiled circle substrates which migrates slower during electrophoresis. Complete nicking was observed after 2 minutes for both uncross-linked and cross-linked complexes. Cross-linking of MutL-MutH does not abolish nicking activity in MutS, ATP-dependent manner. **(E)** Nicking rates for both cases were relatively similar (single experiment).

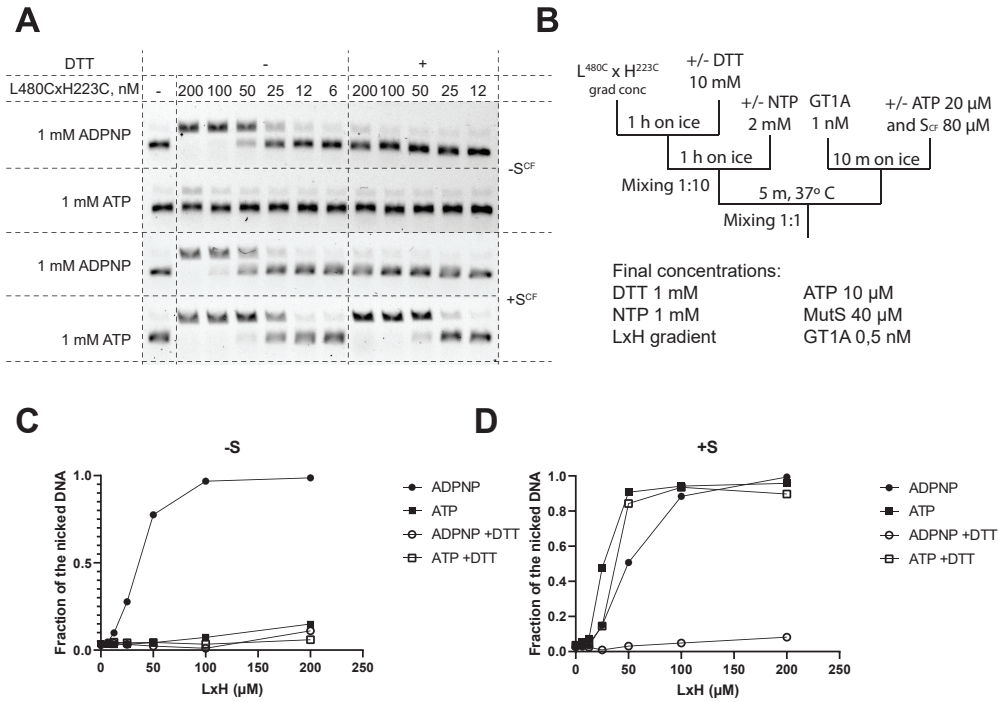


Figure 5.3: **(A)** Influence of protein concentration on nicking activity of cross-linked MutL-MutH complex with ATP or ADPNP in the presence or absence of MutS on the agarose 0.8% gels. **(B)** Scheme and concentrations of the experiment. **(C)** Nicking activity of the complex in the absence **(C)** and the presence of MutS**(D)**

Next, the complex has been tested for nicking ability in the presence or absence MutS 40 μM with 1 mM of ATP or ADPNP. The cross-linked complex was able to cleave DNA after 5 min (Figure 5.3A). Decreasing the complex concentration results in lower nicking rates. Interestingly, as in the experiment with cross-linked unpurified complex in the presence of MutS and 1 mM ATP we did not observe a decrease in activity, rather the rate was slightly higher for the cross-linked complex at 50 nM than in the presence of DTT.

In the presence of MutS, any difference in activity between the cross-linked and uncross-linked MutH-MutL was not observed, indicating that the cross-link is not impairing the function of either MutH or MutL during the process of mismatch- and MutS-dependent activation of MutH. Based on these data we hypothesized that exists an active MutS independent state can be activated by adenine nucleotides (not tested AMP).

Cross-linked MutL-MutH complex is active in MutS independent nucleotide-dependent manner

MutS independent nicking activity was tested at the same 3 kb circle substrates containing GATC site. Surprisingly, we observed complex activity in the absence of MutS only in the presence of ADPNP, not ATP (Figure 5.3). The activity of the complex decreased accordingly as the concentration of the complex decreased. In the absence of MutS the uncross-linked (DTT-treated) complex was not able to cleave DNA in buffers containing 125 mM KCl with any nucleotide (Figure 5.3 and 5.4A

lanes 7-12). In contrast, the cross-linked complex of MutH and MutL retained the ability to cleave DNA at the same conditions in the presence of nucleotides (Figure 5.4A lanes 1-6). The highest activity for nicking was observed with ADPNP. Even at a concentration of 0.1 mM ADPNP MutH-MutL complex was able to cleave the DNA. Interestingly, we observed nicking activity in the presence of 1 mM ADP. The activity in the presence of high 5 mM ADP and ATP was reduced (Supplementary figure B.7A). A similar experiment was performed in the Buffer M/M 20 and it resulted in a lower ability to cleave DNA (Supplementary figure B.7A).

Taken together these results suggest that cross-linking of the C-terminal helix of MutH via position 223C to 480C of the CTD of MutL did not disturb the function of the MutL–MutH complex but rather render the complex independent from MutS.

For functional studies by Cryo-EM, it is important to work with short DNA fragments. We next asked whether this complex can introduce nicks on oligonucleotides. Nicking activity was tested on 59 bp oligonucleotides with a GATC cite in the middle on denaturation urea polyacrylamide gels. Nicking activity results in two bands of different sizes migrating with different velocity corresponding to the long oligonucleotide without a nick and short with an introduced break. No nicking was observed in Ca^{2+} buffer, as was shown earlier[88]. Nicking activity decreases with increasing KCl concentration.Nicking activity also was tested on short 30 bp with oligonucleotides showing no activity of the complex (data are not shown). Interestingly, MutH is active in an MM-dependent manner on 12 bp DNA at low KCl (90 mM) concentrations[129]. Low activity at 30 bp DNA may indicate a lower ability of MutL to bind to DNA at higher salt concentrations and thus worse recruitment of MutH to DNA.

We also wanted to check how the complex interacts with DNA using EMSA. In contrast to the nicking experiments, we did not observe any difference in shifts between the different conditions (Figure 5.5). Shifts, both for the buffer with calcium and the buffer with magnesium, were observed at concentrations of ≈ 1 μM. As mentioned above, there is no nicking activity in calcium buffer even at high concentrations.

We also tested the activity of the complex with two cross-linked MutH to check whether the second cross-linked can abolish the function of the complex. The sample having the ratio of MutL to MutH 2: 2 demonstrated a higher nicking rate (k_{obs} $\approx$0.3 min^{-1}) than the sample having the ratio 2: 1 (k_{obs} $\approx$1.0 min^{-1}).

Further, we checked whether for activation nicking activity it is enough MutL CTD and cross-linked MutH (Figure 5.6). Experiments with thrombin, which cleaves flexible unstructured linker of MutL, showed a complete abolishment in the nicking activity of the complex in MutS independent ADPNP dependent manner.

It was necessary to obtain relatively high concentrations of the complex for structural studies since the working concentration for Cryo-EM is about 5 μM. At high protein concentrations, the yield of the side product significantly increased in comparison with analytical cross-linking, changing buffer at high protein concentration resulted in precipitation of the complex (data are not shown). We noticed, that the yield of MutLxMutL side product can be reduced by changing to a maleimide cross-linker $BM(PEG)_3$ (Figure 5.7A), which has also a higher lifetime than methanthiosulfanate chemistry[144]. The reaction also depends on the ratio between MutL and MutH. At ratio 2:1 we observed almost no side effects products (Figure 5.7B).

The samples further will be investigated using structural methods such as Cryo-

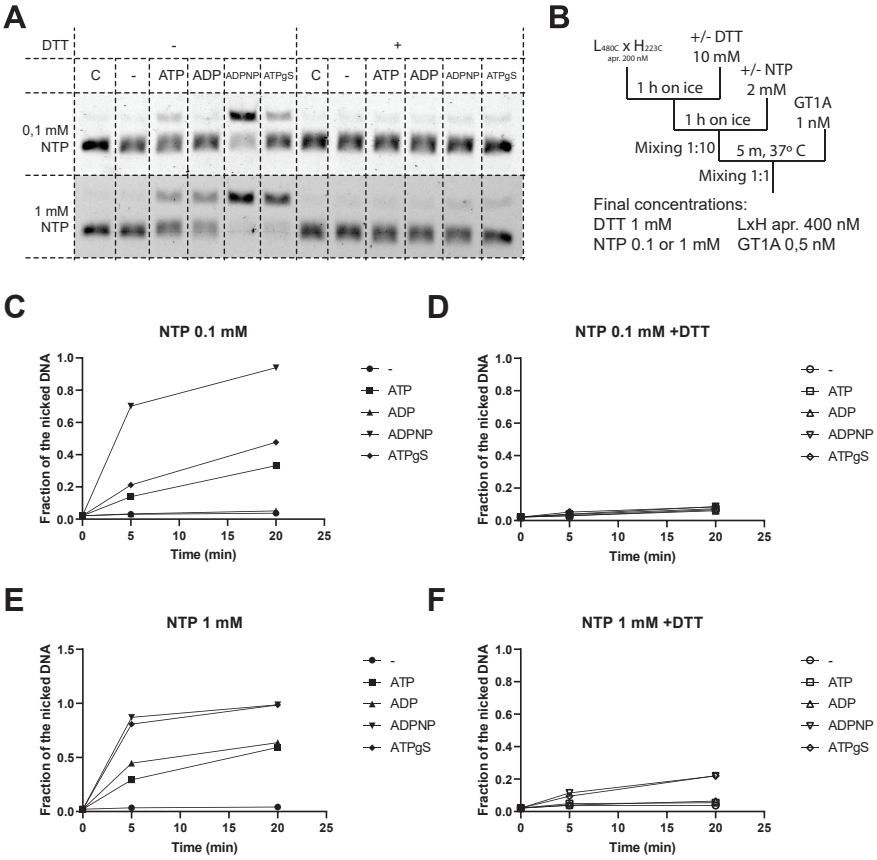


Figure 5.4: **(A)** Nicking activity test was performed on 0.8% agarose gel for low and high NTP concentrations. Nicking activity of cross-linked MutL-MutH complex was observed for all nucleotides at 1 mM. The highest activity was induced by ADPNP even at 0,1 mM. After treatment with DTT complex completely lose its endonuclease activity even at high nucleotide concentrations. **(B)** Scheme of the cross-linked experiment. Complex with or without DTT was incubated for 1 hour with NTPs on ice and mixed with 1 nM of 3 kb circle substrate containing GATC site and fluorophore AF647N for detection. **(C)** Kinetics of the cross-linked complex nicking activity at 0.1 mM NTP. **(D)** Kinetics of the uncross-linked complex nicking activity at 0.1 mM NTP. **(E)** Kinetics of the cross-linked complex nicking activity at 1 mM NTP. **(F)** Kinetics of the uncross-linked complex nicking activity at 1 mM NTP.

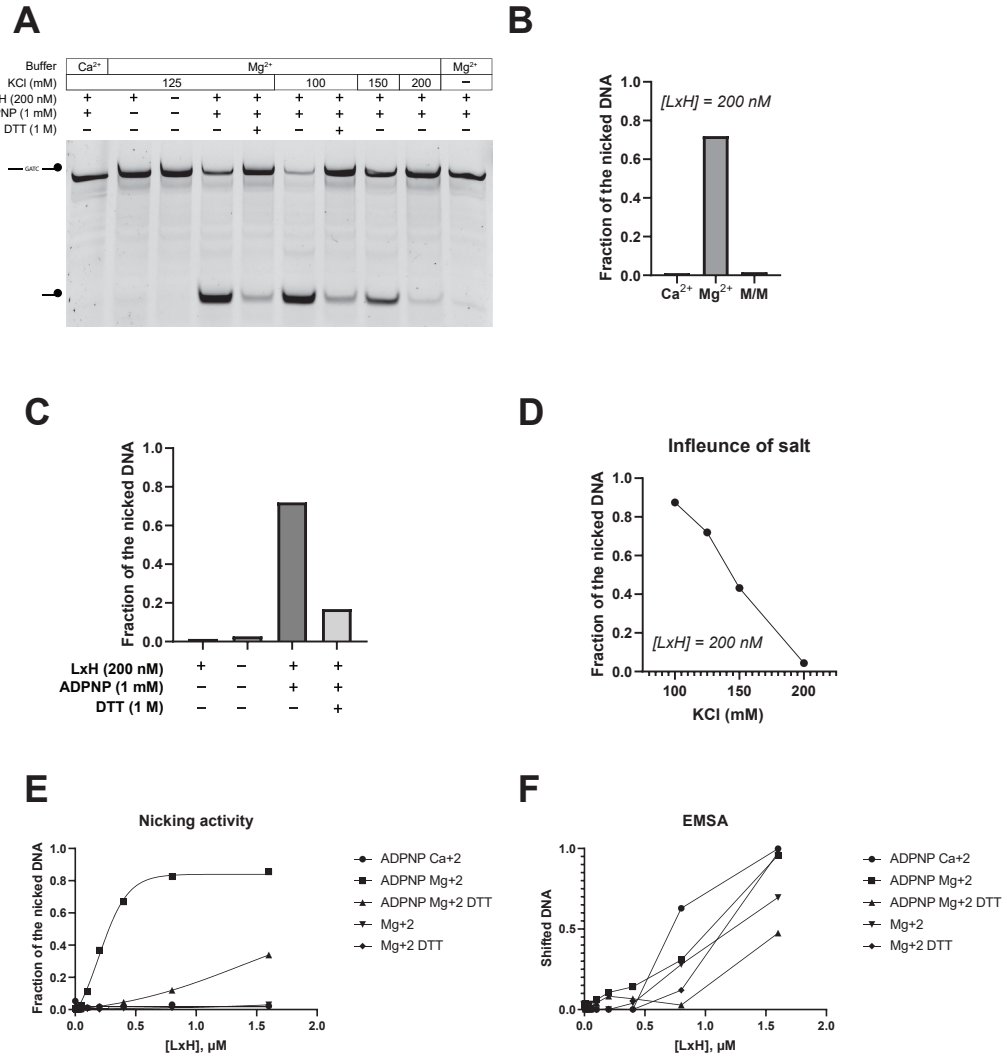


Figure 5.5: **(A)** Nicking assay on 59 bp oligonucleotides at different salt concentrations using denaturing urea polyacrylamide gels. **(B)** In the buffer containing Ca^{2+} instead of Mg^{2+} or in the buffer M/M the nicking activity completely disappears. **(C)** For a nicking activation on 59 bp oligonucleotides complex requires to be trapped covalently and activated by non-hydrolyzable ATP analog ADPNP. **(D)** Salt influence on the nicking activity of the cross-linked complex. **(E)** Shift experiments for the cross-linked complex were performed at 4% native PAGE gels. **(F)** Nicking assay for the cross-linked complex was performed by using denaturing urea polyacrylamide gels 17%.

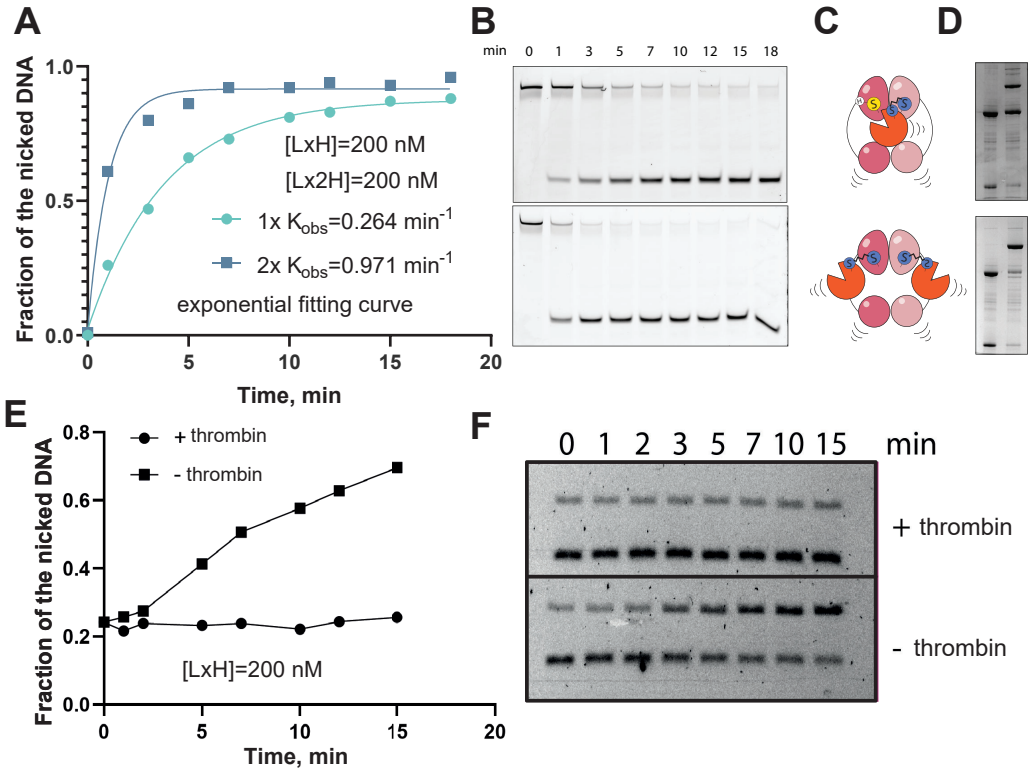


Figure 5.6: **(A)** and **(B)** cross-linked in ratio 2:1 MutH:MutL complex was more active than 1:1 complex. Crosslinking of the additional MutH monomer does not interfere with the activity of the complex, but rather increases it. **(C)** Cartoon representation of the cross-linked complexes with different MutL:MutH ratios **(D)** cross-linked complex on SDS PAGE 8%. **(E)** Nicking activity kinetics. It is not enough to cross-link two proteins for the activity of the complex; the full length of the MutL is required, not just the CTDs. **(F)** Nicking assay on agarose gels. The nicking activity of the cross-linked complex is reduced after thrombin treatment.

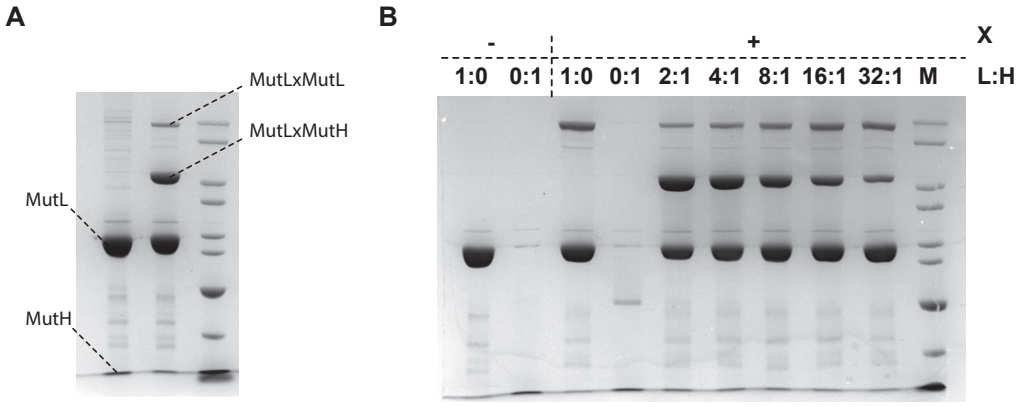


Figure 5.7: **(A)** cross-linking was performed with maleimide chemistry which resulted in a lower yield of side products and no aggregation at high protein concentrations. **(B)** Using optimal MutL:MutH ration we reduced the yield of side products. cross-linking of the proteins alone results in a high yield of side effects for both proteins MutLxMutL (lane 3) and MutHxMutH (lane 4).

EM in Leiden University.

Discussion

An important but not well understood process in the DNA repair mechanism is the regulation and coordination of subsequent processes[141]. MutL is a matchmaker that is responsible for this function, coordinating the signal between initiation and nicking, as well as regulating the unwinding activity of DNA[164]. The sliding clamp structure showed that MutL binds via its NTD to the MutS domain connector[78]. The new Cryo-EM structure demonstrates the position of DNA is in the center of the sliding clamp and position MutL regarding DNA[71]. MutL on its own cannot bind to the DNA due to low affinity under physiological salt conditions[33,95]. In **chapter 2**, we showed MutS ability to recruit with cross-linked core domains to the modified oligonucleotides in the presence of ATP. Since only NTD is present in the structure, it is not clear where the second domain is and what is happening with CTDs. Sliding clamp structure may be the first step in coordination between mismatch and the GATC nicking site. When the structure of NTDs is superimposed with ADPNP state, DNA and the second MutL NTD clash. These structures are incompatible, which may mean the presence of another conformational change in MutL that has not been observed to the date.

According to sliding clamp models, MutS and MutL move as a complex along DNA. When MutL was bound to the MutS, a significant decrease in the diffusion rate in single-molecule experiments was observed, but at the same time, these experiments showed the presence of a MutL-MutH complex independent of MutS[77]. The mechanism and effect of conformational changes in MutL remain unclear, but it participates in the control and activation of the activity of MutH and the DNA unwinding activity of UvrD.

In this chapter, applying the site-specific cross-linking approach successfully used in the past to obtain the structure of the sliding clamp, we characterized a func-

tionally significant active complex MutL-MutH. Interesting was the complex ability to introduce nicks on DNA in the absence of MutS but in a nucleotide-dependent manner. It is known that conformational changes caused by binding to nucleotides also occur in the structure of MutL, including the dimerization of the N-domains and the compactization of the entire structure[88,96]. The activated complex may be the very active state that coordinates nuclease activity. In conclusion, our results demonstrate the importance of covalent trapping of the proteins, as well as the requirement to use the correct nucleotide. We propose, that this approach can be important in obtaining the first structure of an active mismatch nuclease complex on DNA. Further experiments can shed light on the structure as well as the mechanism of this stage. Based on our data, we predict that it is possible to obtain a crystal or Cryo-EM structure of such complex on DNA. The complex with ADPNP may be a relevant state, but at the same time, we can not demonstrate that this complex is identical to the state activated by MutS and ATP in mismatch dependent manner.

It will be interesting to investigate other active complexes consisting of two or even three proteins involved in the MMR pathway. The most interesting complex is the complex between MutL and UvrD.

Chapter 6

DNA mismatch/damage detection using FRET-based assay monitoring the loading of multiple MutS

Vladislav Kunetsky[1], Olha Storozhuk[1], Gwendolyn Brouwer[2], Charlie Laffeber[3,4], Mark S. Dillingham[2], Joyce Lebbink[3,4], Peter Friedhoff[1]

1 Institute for Biochemistry, FB 08, Justus Liebig University, Heinrich-Buff-Ring 17, D-35392 Giessen, Germany.
2 DNA-Protein Interactions Unit, School of Biochemistry, University of Bristol, Bristol BS8 1TD, UK.
3 Department of Molecular Genetics, Cancer Genomics NetherlandsOncode Institute, Erasmus MC Cancer Institute, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands.
4 Department of Radiation Oncology, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands.

Abstract

We developed a sensitive, homogeneous fluorescence assay for the detection of DNA mismatches and damages based on the mismatch repair (MMR) protein MutS. The assay is based on Förster resonance energy transfer (FRET) between SYBR Green I (SG), non-covalently bound to DNA, and Alexa Fluor 647 (AF647) conjugated to MutS. In contrast to previous assays using only the mismatch binding activity of MutS, we exploited the ATP-dependent loading of multiple MutS sliding clamps provoked by mismatch/damage to the DNA, which increases the overall sensitivity of the assay. The assay was validated using a well-characterized 3 kb circular G/T mismatch DNA containing a single G/T mismatch. We demonstrate that treatment of long (kb) DNA with various chemical or physical agents including non-denaturing bisulfite conversion of cytosine to uracil, cisplatin modification, or ultraviolet light (UVC) results in changes in the DNA that can be detected by the FRET-based MutS biosensor.

Introduction

High-quality DNA is essential for most applications in molecular biology. However, the manipulation of DNA, beginning with its isolation from a living organism through chemical or enzymatic assembly in vitro may lead to damages that are not subject to the normal quality control mechanisms of the cellular repair machinery.

Damage-free samples of very long DNA (>several kb) are particularly difficult to obtain. Quality control of these samples involves at least two steps, measurement of the amount of damage and error/damage correction to improve quality[165–167]. Recently, next-generation sequencing has been used to assess the performance of several error correction enzymes that are currently used to increase the quality of synthetic DNA[166]. Our goal was to develop a fast and sensitive tool to assess the quality of long DNA molecules.

We selected the DNA mismatch repair protein MutS from *E. coli*, which not only recognizes mismatches in DNA but can also detect a variety of other DNA damages including base pairs containing O^6-methylguanine, 8-oxoguanine, carcinogen adducts, UV photo products, and cisplatin adducts[7]. MutS has been used previously as a sensor for DNA mismatches[168–180], including a fluorescence-based method[181], and employed as an error removal enzyme to increase the quality of synthetic DNA[166] (For a table with error rates in longer DNA see[182]). However, many assays only exploit MutS binding for the detection of mismatched or damaged DNA. Moreover, when using linear DNA, a further limitation may be caused by the DNA end-binding ability of MutS especially when using short DNA fragments[65]. An exception is the MutS, MutL, MutH mismatch detection system, which relies on the mismatch-provoked cleavage of DNA[183–185].

To develop an improved fluorescence-based assay, we took advantage of two factors not yet exploited by others to increase the overall sensitivity of MutS as a mismatch/damage sensor: First, an ATP-induced change in the MutS-DNA binding orientation: after mismatch recognition ATP switches MutS into a long-lived clamp state. The large conformational changes place the DNA in a new binding pocket[78]. Second, ATP-induced loading of multiple MutS clamps: the MutS clamp is mobile and can leave the mismatch allowing the loading of multiple MutS at the

site of the mismatch/damage[32].

To monitor the conformational change in MutS and the loading of multiple MutS we developed a novel assay that monitors FRET between a donor dye non-covalently attached to DNA and an acceptor dye covalently attached to MutS. In particular, we use SYBR Green I (SG), which preferentially stains double-stranded DNA as a donor dye[186] and Alexa Fluor (AF647) as an acceptor. SG is widely used, mainly for real-time PCR and other DNA-based assays such as the exonuclease reaction[187]. For labeling MutS with AF647, we selected a position in MutS that is far away ($>$ 5 nm) from the DNA in the scanning and mismatch bound states but becomes close to DNA ($\approx$ 3 nm) in the ATP-switched state (Figure 6.1B). Upon binding of MutS to mismatched DNA, a low FRET can occur between DNA-bound SG and single MutS-coupled AF647 and the FRET intensity correlates with the amount of mismatch-bound MutS. After mismatch recognition, ATP triggers the transition of MutS from the mismatch recognition into the sliding clamp state. This pushes the DNA down and closer to the core domain (Figure 6.1B) resulting in a decrease in the distance between DNA-bound SG and AF647 attached to the core domain (from about 5 nm to 3 nm) and increases the FRET efficiency[78]. Most importantly, because the MutS clamp leaves the mismatch which allows loading of multiple MutS[32,83]. Consequently, the final FRET signal is significantly higher compared to that observed from a single mismatch bound MutS.

Material and methods

Protein purification

MutS, MutL, and UvrD proteins were purified as described previously[78,105]. His-tagged MutS dimer variant (MutS336C/D835R) with single-cysteine at position 336 was created from the dimeric cysteine-free MutS variant[57,78,79] using the NEBuilder HiFi Assembly Cloning Kit (NEB) (see Supplementary information for details) and purified obtained as described before[78]. Protein concentrations were determined spectrophotometrically using theoretical extinction coefficient (ϵ_{280} nm =73 605 $M^{\neg 1}$ $cm^{\neg 1}$ for MutS, ϵ_{280} nm = 54 270 $M^{\neg 1}$ $cm^{\neg 1}$ for MutL, ϵ_{280} nm = 83 955 $M^{\neg 1}$ $cm^{\neg 1}$ for UvrD). His-tagged UvrD was overexpressed in BL21 StarTM(DE3)pLysS cells in the vector pET15b and purified by Ni-NTA chromatography. The N-terminal His_6-tag was cleaved by TEV protease in dialysis buffer (Tris-HCl 20 mM pH 8.3, NaCl 300 mM, EDTA 1 mM, glycerol 20% v/v, DTT 2 mM) for 14 h at 4°C. The cleaved protein was passed through Ni-NTA (Qiagen) to remove uncleaved protein, His-Taq and TEV-protease followed by Heparin, HiTrapQ, and desalting columns (ZebaSpin Desalting column 40K). UvrD was stored in storage buffer (25 mM HEPES pH 8.0, 400 mM KCl, and 20% glycerol) and snap-frozen in liquid nitrogen and stored at -70°C.

Protein labeling

To obtain MutS labeled with Alexa Fluor 647 (AF647-MutS), MutS336C/D835R was diluted to 40 μM in 150 μl of HPLC buffer (10 mM HEPES/KOH (pH 8.0), 200 mM KCl and 1 mM EDTA) and labeled with a 5-fold molar excess of Alexa Fluor™ 647 C_2 Maleimide (Invitrogen, Thermo Fisher Scientific, Waltham, MA) for 2 h on

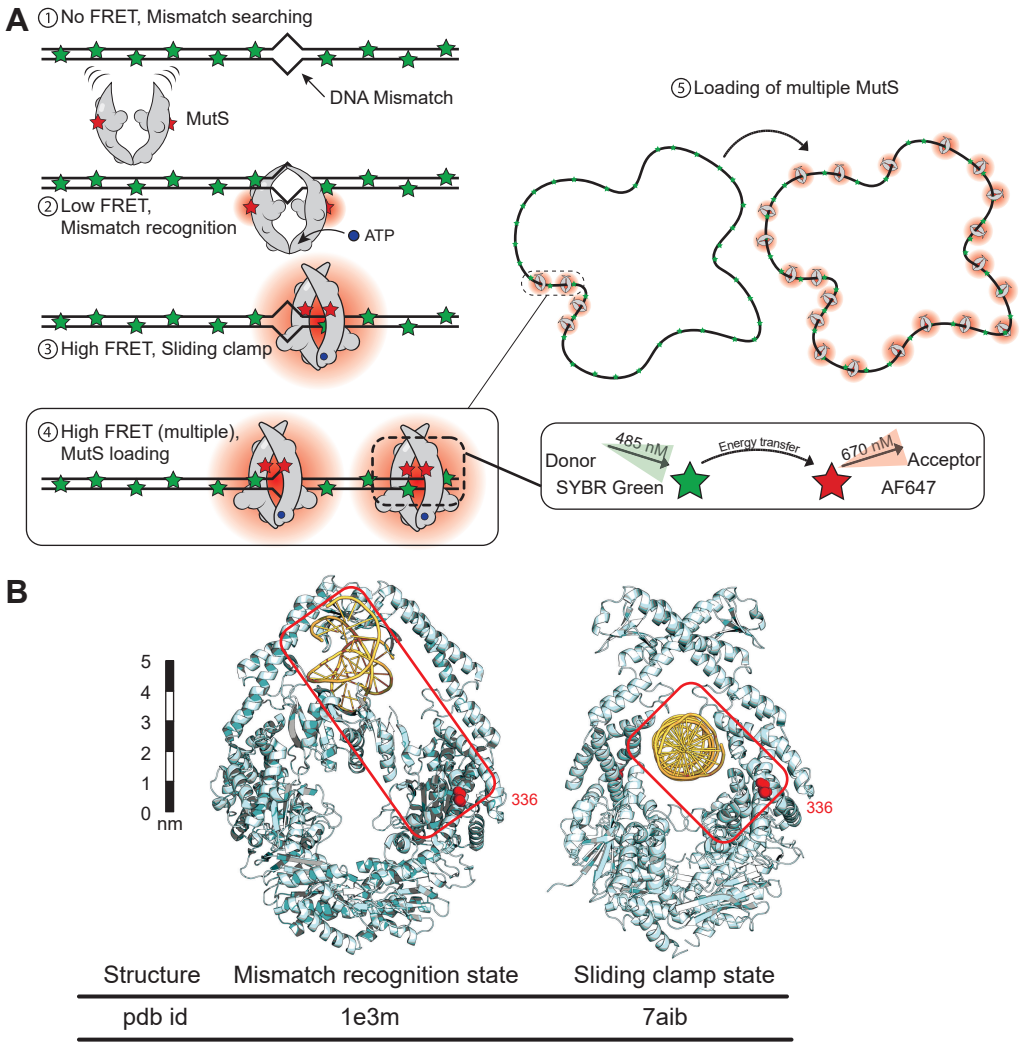


Figure 6.1: Schematic representation of FRET-based assay. **(A)** MutS scans DNA and binds to mismatches in a low FRET mismatch recognition state. ATP binding triggers a conformational change leading to the high FRET MutS sliding clamp state. Sequential binding of additional MutS to the mismatch results in the loading of multiple MutS sliding clamps onto the mismatch-containing DNA, which significantly increases the FRET signal. **(B)** MutS structures with the amino acid at position 336 are shown as red spheres; pdb-codes: 1e3m and 7aib, and approximate distances between position 336 in MutS and DNA. In the sliding clamp state, the attached fluorophore comes closer to the DNA making FRET between DNA-bound SYBR green I (SG) and dye covalently attached to MutS more efficient.

ice in the dark according to the manufacturer's instruction. Excess dye was removed using Zeba™ Spin Desalting Columns 40K (Thermo Fisher Scientific, Waltham, MA) and the degree of labeling was determined from the absorbance spectra (NanoDrop 1000, Thermo Fisher scientific) (DOL 90.3 ± 3.3 %, n=3) as described previously[78]. Then the solution was aliquoted and flash-frozen in liquid nitrogen and stored at -80 °C.

DNA substrates

The 3196 bp circular G/T mismatch DNA substrates were created by primer extension and ligation essentially as described earlier using the circular single-stranded DNA derived from phagemid pGATC1 as a template[120]. Plasmid DNA (phagemid GATC1) from *E. coli* DH5α was purified using the NucleoSpin Plasmid, Mini kit for plasmid DNA (Macherey-Nagel). λ- DNA (SD0011) was obtained from Thermo Fisher scientific.

Non-denaturing bisulfite treatment

For deamination of cytosine bases in DNA under non-denaturing conditions, plasmid DNA was treated using the EpiMark Bisulfite Conversion Kit (E3318S, New England Biolabs). The protocol was modified to allow non-denaturing deamination. Plasmid DNA (2 μg) was diluted in 100 μl of bisulfite mix and the cycling protocol (95 °C and 60 °C - total incubation time 205 min) was replaced by incubation for 30 min at 50°C (or the indicated temperature) with slowly decreasing the temperature after treatment (1°C/min) to room temperature. Excess of sulphonation reagent was removed by using EpiMark spin columns according to manufacturers instructions. The volume of the eluted DNA solution was adjusted to 150 μl in 1% desulphonation solution from the kit and incubated for 15 minutes at room temperature. The excess reagent was removed by the EpiMark™ spin column. The eluted, deaminated DNA solution was stored at -20°C.

UV-irradiation

DNA (phagemid DNA or λ-DNA) was placed at 40 ng/μl in 50 μl in water on parafilm and exposed to UV irradiation at 254 or 365 nm with a controlled UV irradiation system (Bio-Link 254 and Bio-Link 365, Vilber Lourmat Deutschland GmbH, Germany) with maximum 0.8 J/cm^2. UV-irradiated DNA was stored at 4°C.

Cisplatin modification

Cis-Diamminplatin(II)-dichloride was obtained from Sigma-Aldrich (catalog # 479306-1G). Plasmid DNA was incubated in 0.154 mM NaCl saline buffer in the absence or presence of cisplatin (1.4 μM – 2.8 mM) for 24 h at room temperature in the dark, starting at 2.8 mM with two-fold dilution up to 1.4 μM. Excess cisplatin was removed by using Illustra NAP-5 columns (Sigma-Aldrich) for desalting and buffer exchange. Modified DNA was stored at 4°C.

Nicking of deaminated plasmid

Phagemid DNA (2,6 μg) with or without non-denaturing bisulfite treatment were cleaved by 20 Units of Nt.BstNBI (New England Biolabs) for 1 hour at 37°C in 50 μl of 1X NEBuffer™ 3.1 followed by deactivation at 70°C for twenty minutes.

Thermal denaturation and reannealing

DNA (50 μl of 40 ng/μl) was incubated in the heating block at different temperatures in water or annealing buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl) for 5 minutes. Samples were either directly placed on ice or cooled down slowly to room temperature in the heating block. DNA was stored at 4°C.

Agarose gel electrophoresis

DNA samples (200 ng) were analyzed on 0.8% agarose gels in TAE buffer (Tris 40 mM, Natriumacetat 20 mM, EDTA 1 mM, pH 8.0) (30 V for 3 hours) stained with HDGreen Plus Safe DNA Dye and UV fluorescence (Intas Science Imaging Instruments GmbH, Germany).

MutS sliding clamp formation on DNA monitored by FRET

DNA substrates were diluted to 0.38 ng/μl (575 nM bp; in case of 3.2 kb circular G/T 0.18 nM G/T mismatches) in 200 μl of FB150 buffer (25 mM HEPES-KOH pH 7.5, 5 mM MgCl2, 150 mM KCl, and 0.05 % (v/v) Tween 20) at room temperature in 96-well plates with 250 nM of SYBR Green I (SG) (S7563, ThermoFisher Scientific) (dbpr = 0.43) (Invitrogen, Thermo Fisher Scientific, Waltham, MA). AF647-MutS (50 nM monomer) was added to the DNA substrates and incubated at room temperature for 150 seconds. ATP (1 mM) was added and kinetics of the fluorescence signals were recorded with a fluorescence microplate reader (TECAN infinite F200, Tecan Group Ltd, Switzerland). Fluorescence intensities were measured in three channels (donor, acceptor, FRET) with the following filter combinations: (donor ex. 450 nm (width 20 nm) em. 535 nm (width 25 nm), acceptor ex. 620 nm (width 10 nm) em. 670 nm (width 25 nm), FRET ex. 485 nm (width 20 mm) em. 670 nm (width 25 nm) filter. Signals were corrected for buffer and spectral crosstalk to obtain the correct signal intensities. Correction factors were obtained from measurements with only one of the labeled components present (see Supplementary Data for details).

Fluorescence SSB binding assay

Single cysteine mutant SSB from Plasmodium falciparum was expressed, purified, and labeled with N-[2-(iodoacetamide)ethyl]-7-diethylaminocoumarin-3-carboxamide (IDCC) to give DCC-SSB as described earlier[188,189]. For measurement UvrD helicase activity, experiments were performed by preincubation 100 nM of MutS, 100 nM of MutL and 100 nM of UvrD with pre-nicked DNA (0.38 ng/μl) in 200 μl of FB150 buffer (HEPES 25 mM, $MgCl_2$ 5 mM, KCl 150 mM, pH 7.5) and 30 nM of fluorescence SSB (DCC-SSB) for 5 min. The reaction was started by rapidly adding 2 μl of 100 mM ATP. In the case of DCC-SSB binding to the reannealed DNA, the

experiment was done at an identical concentration of the DNA (0.38 ng/μl) and DCC-SSB (30 nM) and in FB150 buffer as it was performed in the helicase activity assay. Fluorescence intensities were measured with excitation 430 ± 20 nm and emission with 485 nm ± 20 nm bandpass filters.

Results and discussion

We first established the FRET assay using a well-characterized 3.2 kb circular DNA with a single G/T mismatch that has been used before to study the initial steps of mismatch repair[120,134]. The use of circular DNA also prevents potential DNA-end binding and dissociation of the mobile MutS clamps. The specificity of the system was verified by using a phagemid DNA of identical sequence but without a mismatch,[120,134]. In the absence of a mismatch, almost no signal (FRET^c, 1300 ± 200 a. u. Figure 6.2B) was detectable in the FRET channel after correction for spectral crosstalk (see Material and methods and Supplementary Data B.1). This was independent of the absence or presence of ATP (Figure 6.2B column 1+2). For the G/T mismatch DNA in the absence of ATP, the FRETc signal was small (2400 ± 400 a. u.) but significantly higher (1.8 fold) compared to that with DNA without a mismatch (Figure 6.2B column 1+3). Upon addition of ATP, a substantial (> 12-fold) increase of the FRETc was observed for G/T DNA (28000 ± 3000 a.u.) (Figure 6.2B column 4) confirming our assumption that loading of multiple MutS sliding clamps would lead to large signal amplification.

MutS sliding clamps can diffuse at least 1 kb along the DNA backbone and thereby leave DNA via its ends[32]. Hence, we expected that linearization of the 3 kb circular G/T mismatch DNA would reduce the number of MutS sliding clamps on DNA and hence the FRETc. Upon linearization of the circular G/T by ScaI, which introduces a single cut 1.4 kb 5' and 1.7 kb 3' of the G/T mismatch, the FRETc was only slightly reduced (Figure 6.2B column 4, non-significant, $p = 0.15$). However, the addition of ATP resulted only in a small (1.5-fold) increase of FRETc ($p = 0.01$), which was strongly reduced (by > 90 %) compared to FRETc observed with the circular form of the G/T-mismatch DNA. This suggests that on the linear G/T DNA only a strongly reduced number of sliding clamps are present (Figure 6.2B column 5).

Next, we investigated the kinetics of the ATP-induced loading of multiple MutS sliding clamps on the circular G/T mismatch DNA. Upon addition of ATP, the FRETc increased over time and a single exponential function could be used to fit the data with an amplitude ΔFRETc of about 23000 (a.u.) and a decay rate λ of about 0.01 $s^{\neg 1}$ (lifetime $\tau \approx 100$ s) (Figure 6.2C). This is in the same order of magnitude as the reported lifetimes of a MutS sliding clamp on DNA with blocked-ends using surface plasmon resonance (≈ 70 s)[78].

For evaluating the sensitivity of the assay we determined the ATP-induced ΔFRETc at different concentrations of circular G/T mismatch DNA(Figure 6.2D). The linear range was obtained from 0.01 nM to 0.2 nM [GT1a] (1 mismatch per 3 kb or 0.33 mismatches per kb).). Limit of detection (LoD) at such conditions is about 1 mismatch per 240 kb (0.004 mismatches/kb) o and the limit of quantification (LoQ) is 1 mismatch per 50 kb (or 0.02 mismatches/kb).

Binding of the MMR protein MutL to the MutS sliding clamp results in the formation of a MutS-MutL complex[32,77]. At concentrations of 100 nM MutS and

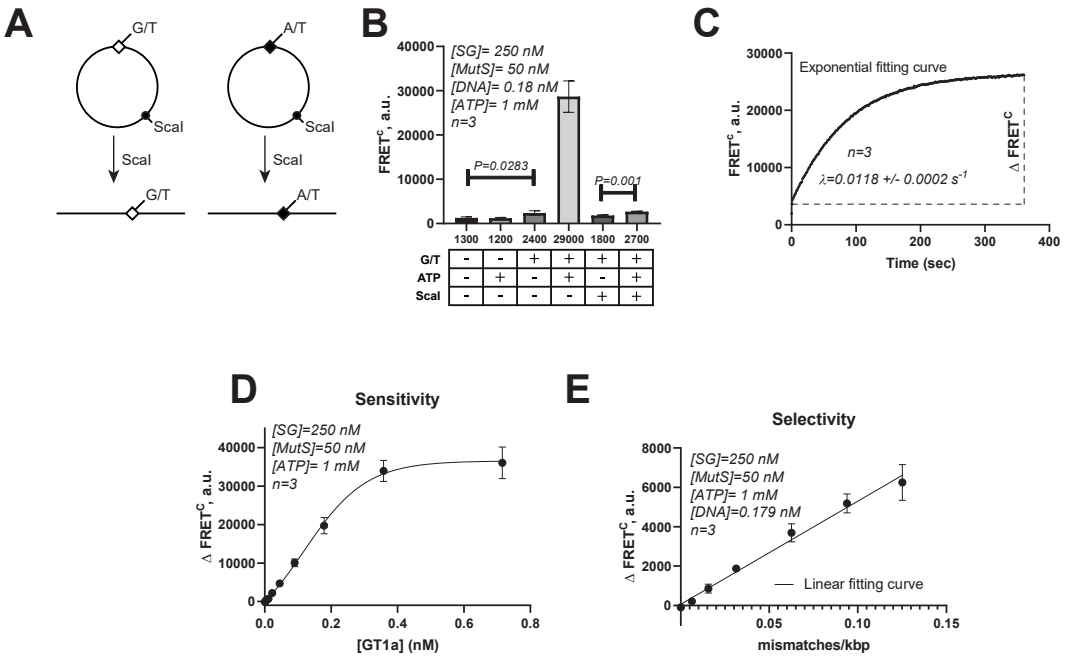


Figure 6.2: Schematic representation of FRET-based assay. **(A)** Schematic representation of the DNA. **(B)** FRETc between AF647-MutS and SG-stained mismatch or phagemid DNA substrates. **(C)** Kinetics of loading multiple MutS on the circular G/T mismatch DNA. **(D)** Analysis of sensitivity: Concentration dependence of circular G/T mismatch DNA on ΔFRET^c (FRET$^c_{ATP}$ - FRET$^c_{noATP}$). **(E)** Analysis of selectivity: Decreasing fraction (expressed as G/T / kb) of circular G/T mismatch DNA in the presence of constant amount total DNA ([G/T mismatch DNA] + [phagemid] = 0.179 nM and SG (250 nM) keeping at constant dbpr of 0.44. Note, the G/T mismatch DNA has a G/T per kb of 0.31.

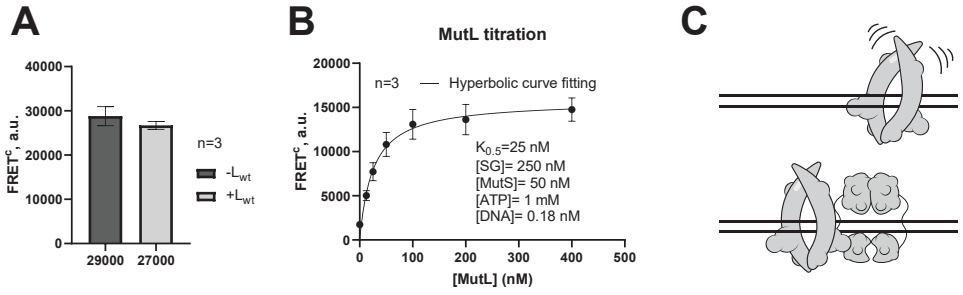


Figure 6.3: **(A)** Influence of MutL on the observed FRETc with AF647-MutS and circular G/T DNA **(B)** Concentration dependence of MutL on the observed FRETc between MutS-AF657 and linear G/T DNA **(C)** Schematic representation of MutS trapping by MutL

100 nM MutL, a significant but incomplete reduction of complexes on 2.9 kb G/T mismatch DNA upon linearization had been observed[32]. Consequently, we asked whether and how MutL influences the FRETc signal from MutS sliding clamps on circular and linear 3 kb G/T mismatch DNA. Whereas on circular DNA, MutL had little effect on the MutS sliding clamp formation (Figure 6.3A), on the G/T mismatch DNA linearized with ScaI we observed a strong dependence on MutS sliding clamp stability on the concentration of MutL (Figure 6.3B). A hyperbolic function could be fit to the data with an apparent $K_{1/2}$ of about 50 nM. This value is similar to that reported before for *Taq* MutL trapping MutS on DNA[138].

Deamination of DNA by bisulfite treatment under non-denaturing conditions

To demonstrate the ability of our assay to detect DNA damages, we first exploited the ability of MutS to bind to G/U mismatches that can arise from spontaneous or enzymatic deamination or via misincorporation of dUMP into DNA[190]. We used phagemid GATC1, which contains a pBR322 ori that has been shown to be modified by bisulfite conversion under non-denaturing conditions[191]. Since spontaneous deamination in double-stranded DNA is rather slow, we changed a protocol for cytosine deamination using the NEB EpiMark Bisulfite Conversion Kit to milder conditions (see Materials and Methods for details) to convert circular plasmid DNA into G/U containing DNA. Next, we tested, whether treatment of the phagemid DNA resulted in changes that can be detected by our MutS biosensor. Qualitatively, the results are comparable to those obtained with the G/T mismatch DNA (Figure 6.2B and 6.4A). However, the amount of FRET both in the absence (2900 ± 120 a.u.) and presence of ATP (38000 ± 5700 a.u.) was higher with the bisulfite-treated DNA (at 50 °C) compared to the DNA with a single G/T mismatch suggesting the number of G/U mismatch per phagemid is on average larger than 1. Similar to the G/T mismatch DNA, the FRET signal was significantly lower (14 %) for linearized compared to the circular form of the DNA (Figure 6.2B and Figure 6.3A).

In a second step, we varied the temperature during bisulfite-treatment between 20 °C and 60 °C and tested for MutS binding and sliding clamp formation on the circular phagemid DNA (Figure 6.4B). At temperatures as low as 20 °C the phagemid DNA was readily converted into a substrate capable of inducing MutS sliding clamp formation, although to a lesser extent than G/T mismatch DNA suggesting the number of G/U mismatch. When the temperature was raised further, the extent of MutS sliding clamp formation was increased. Analysis of the bisulfite-treated DNA by agarose gel electrophoresis revealed no obvious changes compared to the untreated DNA (Figure 6.4C). To verify that the bisulfite treatment did indeed result in the deamination of cytosine to uracil, Uracil-DNA glycosylase (UDG) was used to introduce apyrimidinic (AP) site by excision of the uracil base and endonuclease Endo IV to convert AP sites to strand breaks. This converts the supercoiled, covalently closed phagemid DNA into open circular DNA, which can be detected by agarose gel electrophoresis. The only treatment of the bisulfite-treated DNA with both enzymes resulted in the open circular form of the plasmid DNA suggesting that uracil was indeed present in the bisulfite-treated DNA (Figure 6.4A). It was shown for the eukaryotic homologs of MutS that binding to a DNA with an AP site is strongly reduced compared to G/U mismatched DNA[192]. Consequently,

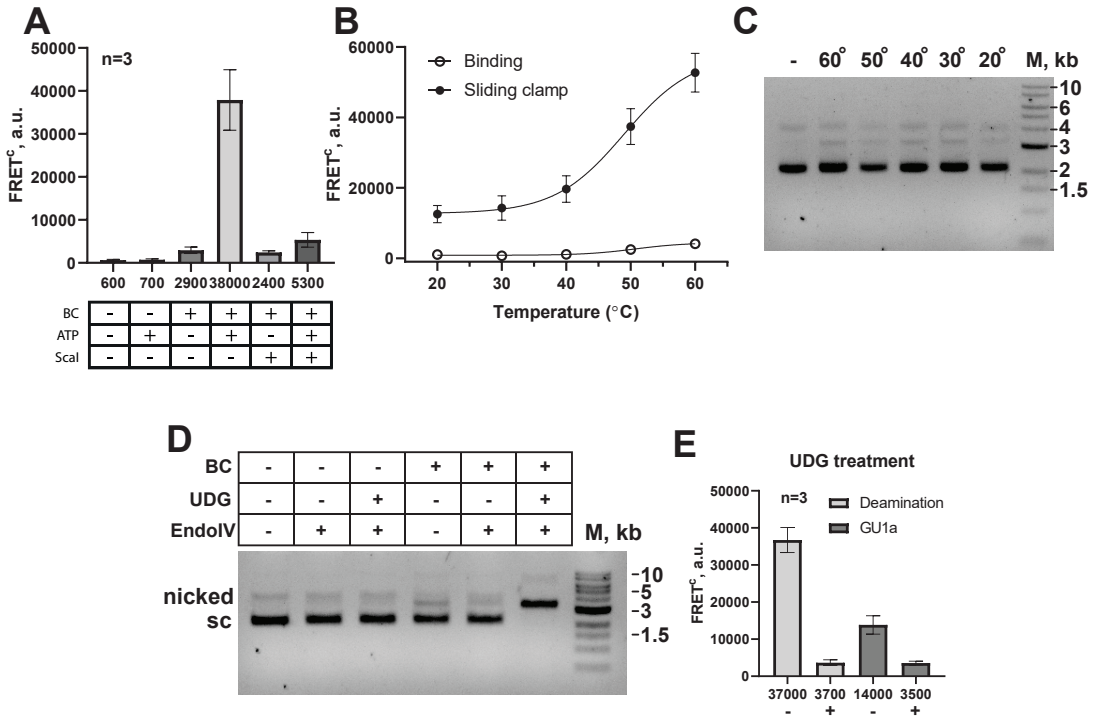


Figure 6.4: **(A)** Bisulfite treatment at 50°C during 30 min incubation of plasmid DNA results in binding of MutS to circular DNA and allow loading of multiple MutS sliding clamps by ATP. **(B)** Effect of temperature during bisulfite treatment on MutS binding and sliding clamp formation. **(C)** Quality control of bisulfite-treated DNA by agarose gel electrophoresis. FRET arising between MutS subunits. **(D)** Bisulfite mediated deamination of plasmid DNA under mild conditions. **(E)** Deaminated DNA treated by UDG in FRET-based MutS sliding clamp assay.

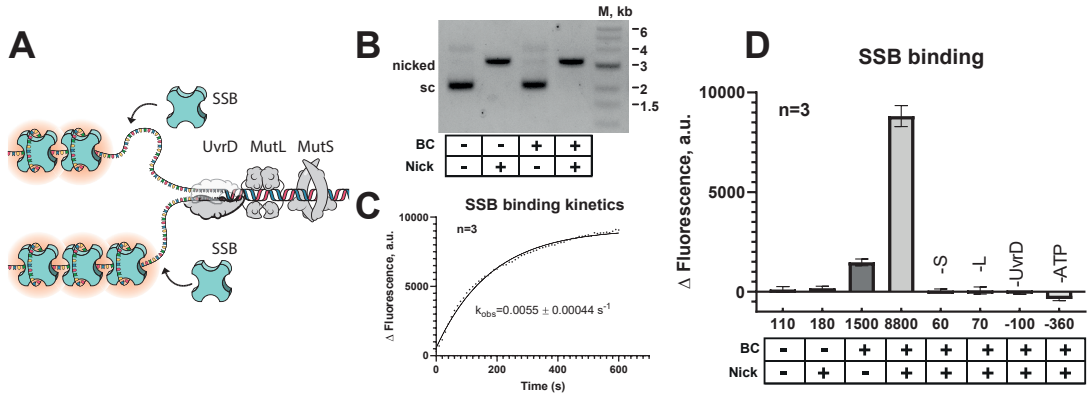


Figure 6.5: **(A)** Fluorescence upon binding DCC-SSB to the DNA substrates. **(B)** Kinetics of DCC-SSBs binding to the ssDNA. **(C)** Nicking of treated and untreated DNA by Nt.BstNBI. **(D)** Schematic representation of Fluorescence-based DCC-SSB binding assay for monitoring UvrD unwinding activity in MM-dependent manner.

we asked how digestion of the bisulfite-treated DNA by UDG would influence the sliding clamp formation of MutS. Indeed, after UDG-treatment the FRET signal was strongly (10-fold) reduced compared to the untreated DNA suggesting that an AP-site is, at best, a poor substrate for MutS compared with G/U containing DNA (Figure 6.4E). MutS cannot undergo conformational changes to the sliding clamp on DNA treated with UDG, since there is no loading site for MutS.

Next, we tested whether the bisulfite-treated phagemid DNA is a substrate for mismatch repair unwinding reactions (Figure 6.5C). To monitor unwinding we took advantage of a fluorescently labeled single-stranded binding protein (DCC-SSB) that had been developed before as a sensor for DNA unwinding[188]. Binding of DCC-SSB to single-stranded DNA results in significantly increased fluorescence (Figure 6.5D). Using a prenicked circular DNA substrate (Figure 6.5C) we were indeed able to observe the kinetics of DNA unwinding (Figure 6.5B). The signal was dependent on bisulfite-treatment, a nick, MutS, MutL, UvrD, and ATP (Figure 6.5A).

Cisplatin modification

Cisplatin cytotoxicity results from the formation of lesions that block DNA polymerase and cause replication arrest. The damages induced by cisplatin are commonly repaired by Nucleotide Excision Repair (NER)[193]. *E.coli* „dam-“ strains are more susceptible for the cytotoxic action of cisplatin and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) than wild type[194–198]. Replication arrest can cause damages such as daughter-strand gaps and double-strand breaks. In vitro experiments have shown that MutS probably recognizes cisplatin-intrastrand cross-links, but the specificity is much lower compared to mismatches[199]. The molecular mechanism of MMR-mediated drug resistance in dam- bacteria and mammalian cells is still unknown.

Here we tested our novel FRET assay to test whether treatment of DNA with cisplatin leads to damages that are recognized by MutS, leading to the loading of multiple MutS sliding clamps and activation of the MMR system. Indeed, the amplitude of the FRET signal increased with increasing cisplatin concentration during

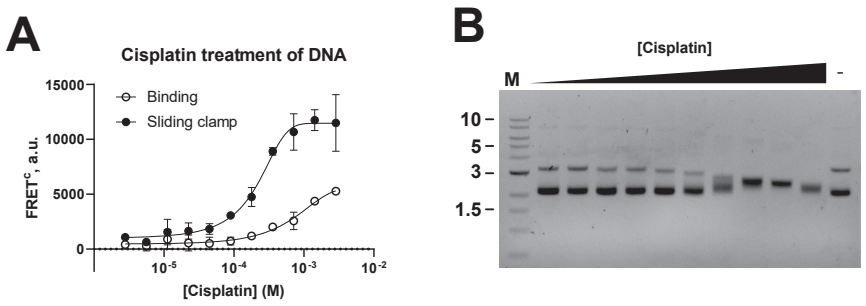


Figure 6.6: Cisplatin treatment of DNA. **(A)** Cisplatin modification of plasmid in FRET-based assay and **(B)** on the agarose gel.

DNA treatment (Figure 6.6A). At high cisplatin concentration (> 0.1 μM) the FRET signal was increased even in the absence of ATP indicating large numbers of lesions recognized by MutS binding. A significant increase was observed upon adding ATP due to the sliding clamp formation of MutS. In contrast to the bisulfite-treated DNA, the difference between binding (no ATP) and sliding clamp formation (ATP) was much lower, suggesting that the lesion introduced into the DNA up cisplatin-treatment might be not as efficient or preventing loading of multiple-sliding clamps. As a quality control, the cisplatin-DNA treated DNA was analyzed by agarose gel electrophoresis (Figure 6.6B). At a concentration above 0.36 mM of cisplatin, the modification of DNA was changing the electrophoretic mobility.

UV damages

MMR participates in non-canonical recognition at different lesion types including damages induced by UV[200]. Moreover, it was shown that human MutSβ is recruited to the sites of UV damages in human (HeLa, XPA, U2OS) cells[201,202]. Therefore, we next investigated whether *E. coli* MutS forms sliding clamps on UV-damaged DNA. With increasing doses of UV radiation (254 nm), the amplitude of the FRET signal increased both for MutS binding (no ATP) and loading of multiple MutS sliding clamps (Figure 6.7A). In contrast, we did not observe any increase in FRET signal upon treatment with UV 365 nm radiation (Figure 6.7A). Analysis of the DNA treated with UV radiation of increasing energy by agarose gel electrophoresis showed already at the lowest energy (0.1 Joule/cm^2) changes in the electrophoretic mobility of fractions of the DNA. At the highest energy used (0.8 Joule/cm^2) virtually all DNA migrating at the position of supercoiled DNA was shifted to lower electrophoretic mobility. This has been observed before using atomic force microscopy[203,204]. As mentioned above, on linearized 3 kb phagemid DNA containing a single G/T mismatch we did not observe significant loading of multiple MutS sliding clamps (Figure 6.2B) which was most likely a consequence of rapid dissociation of MutS from the DNA-end via linear diffusion. Using UV-treatment of linear DNA we reinvestigated this phenomenon using DNA of different lengths (Figure 6.7C). First, we were able to observe the loading of multiple MutS sliding clamps onto the linearized 3 kb phagemid DNA treated with UV (0.6 Joule/cm^2) which was 5-fold lower compared to the circular form but significantly higher (4-fold) than the linearized form of the phagemid DNA with a single G/T mismatch (Figure 6.2B).

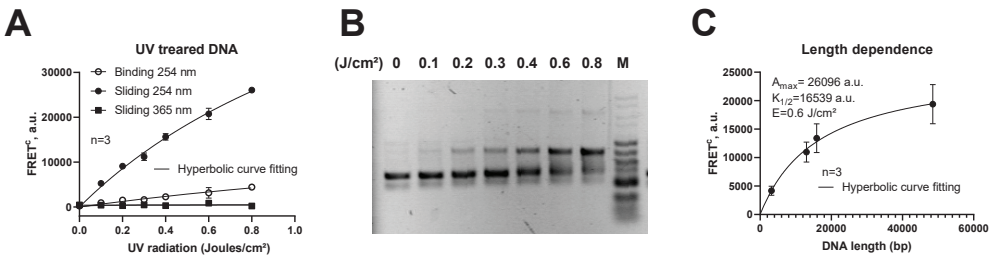


Figure 6.7: UV treatment of DNA leads to binding and loading of multiple MutS sliding clamp. **(A)** Phagemid DNA was treated with UV light (254 or 365 nm) at the indicated total energy, and loading of multiple AF647-MutS sliding clamps to SG-stained DNA was monitored using FRET. **(B)** Binding (no ATP) and loading of multiple sliding (ATP) on UV (254 nm) treated phagemid DNA monitored by FRET depend on the total energy of irradiation. **(C)** Length dependence of linear UV-treated (254 nm; 0.6 Joules/cm^2) DNA substrates for the loading of multiple MutS sliding clamp. **(D)** Analysis of phagemid DNA irradiated at 254 nm with the indicated energy of UV radiation by agarose gel electrophoresis.

Moreover, upon increasing the length of the UV-treated (at constant energy) DNA up to 48 kb (λ-DNA) the signal increased. This is in good agreement with the diffusion coefficient of MutS sliding clamps on DNA (see above). It takes less than 10 s for MutS to travel 3 kb, 300 s for 15 kb, and 3000 s for 48 kb (Figure 6.7C). As the lifetime of MutS sliding clamps is in the order of 100 - 200 s[77], MutS is unlikely to reach DNA ends on 48 kb DNA.

Thermal denaturation and reannealing

Many molecular biology procedures require high temperatures (for example during PCR or heat deactivation of enzymes), which can lead to DNA damages or local insertion/deletions due to improper re-annealing of the strands. Therefore, we next asked, whether our novel FRET-based assay for measuring mismatch/damage induced loading of multiple MutS sliding clamps is suitable for checking the quality of DNA after such treatment. Plasmid DNA was incubated in pure water at different temperatures for 5 min followed by either fast or slow cooling to 20 °C. After dilution in FB150 buffer, we tested for the binding of MutS and loading of multiple MutS using our novel FRET assay. Starting at 60 °C we observed a significant increase in the intensity of the FRET signal with increasing temperature (Figure 6.8A). In contrast to the other types of damages tested above, we observed that the increase between MutS binding (no ATP) and loading of multiple sliding clamps (ATP) was much reduced (1.6-fold compared to 12-fold for GT or 13-fold for Deamination-treatment (Figure 6.2B and 6.4A). This suggests that the treatment created sites for MutS binding but did not allow the proper switching to sliding clamps to allow for loading further MutS[205]. Thermal denaturation and reannealing in annealing buffer the protect plasmid from the damages/secondary structure recognized by MutS as the FRET signal we observe is an order of magnitude lower compared to the treatment in water (Figure 6.8A). Similar results were obtained using λ DNA in an annealing buffer with higher resistance to the damages, but at high temperatures, a sharp increase was still observed (Supplementary Figure B.21B). This may be due to the

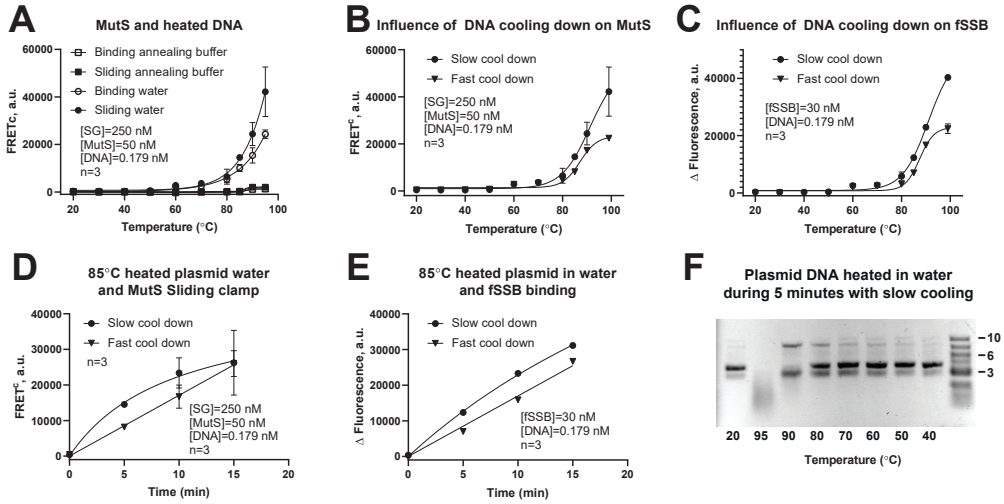


Figure 6.8: MutS binding to DNA incubated. **(A)** MutS binding and sliding clamp formation on heated in water and annealing buffer plasmid. **(B)** Slow and fast cooling down of heated plasmid in water. **(C)** Binding of DCC-SSB to DNA incubated in water at indicated temperature followed by slow/fast cooling to 20 °C **(D)** Influence of incubation time of DNA at 85 °C in water on the extent of loading of multiple MutS sliding clamps **(E)** Time influence on DCC-SSBs binding to the heated plasmid in water **(F)** Plasmid degradation after incubation in water during 5 minutes with slow cooling down in heating block on the agarose gel.

linear nature of the DNA as compared to a supercoiled, circular plasmid. In comparison with the cooling down in the water, we have seen an inverse relationship, where slow cooling down in the heating block resulted in a higher signal (Figure 6.8B). Fluorescent SSB protein showed a similar dependence to bind heated DNA as MutS protein, revealing the single-stranded nature of these loops/damages (Figure 6.8C). The amount of signal also increased with the heating time leading to DNA damage, which was recognized both by the MutS sensor (Figure 6.8D) and DCC-SSB (Figure 6.8D). Denaturation of the plasmid that has been thermally treated in water can also be observed on an agarose gel (Figure 6.8F).

Discussion

Here, we demonstrate that *E. coli* MutS protein is suitable for quantitative DNA quality control. It can be performed both on short (starting from 3 kb) and long (gene size) linear or circle substrates. In contrast to previous methods using MutS, our method does neither require GATC sites for cleavage or time-consuming agarose gels, allowing its usage for any sequence. The assay is based on a single labeled protein and the widely used and well-described intercalating dye SYBR Green I (SG) for staining DNA which makes it affordable for other laboratories as well for commercial purposes. In this work, we presented the optimal conditions for this method, including the concentrations of the components, as well as the device settings (filter wavelength). The assay uses a natural feature of the MMR system, namely loading of multiple sliding clamps of MutS, thereby amplifying the signal in the cell for a further cascade of reactions. For example, in most studies using

MutS as a mismatch/damage sensor that only used the binding of MutS to DNA in the absence of ATP, the signal for the homoduplex did not differ so much as compared to the mismatch. In our novel assays, signal amplification results in much greater discrimination between normal and damaged DNA. To further increase the sensitivity of our assay, in addition, to measure the FRET intensity, the difference between the fluorescence signals before and after the addition of ATP is used. By using the normalization of the corrected signal to the acceptor and donor channels, we were able to further increase the sensitivity of the method. A limitation of this signal amplification is that it requires either circular DNA or very long DNA (several kb). However, upon the inclusion of a second, unlabelled repair protein MutL, the signal for MutS remains high even on linear substrates. Alternatively, DNA ends could be blocked by end-labels, DNA-end binding proteins, or dCas9 roadblocks[206]. We have examined the application of this assay for several types of damage: (i) GT mismatches which appear during replication, (ii) the effect cytidine deamination, the process, which takes place in the cell and can occur during manipulation with DNA for biotechnical purposes, (iii) cisplatin modification (iv) high energy UV-damages and (v) improper reannealed DNA which cause insertion/deletion loops. Although our method does not allow to detection of all kinds of damages, e.g. nicks or ap-sites, it can probably be extended to other DNA damages detected by MutS. Temperature changes during deamination showed that the signal depends on the amount of DNA damage. The FRET signal values at high temperatures were higher than with one GT mismatch. With UV damage, it was observed that the signal increased with the length of the substrate. With the same energy and with an increase in the length of DNA, the amount of damage per molecule also increases, but the amount of damage per nucleotide on average remains the same. The increase in the FRETc signal can be explained by the fact that increasing the length of the substrate also increases the path that MutS must slide before it leaves DNA at ends. One final concern about the adaptability of DNA damage scanning is the ability for scaling up and automatization of procedure in 96 well plate fluorescence reader for high-through usage. These works can be further used for modeling the mechanism of this system and the influence of various parameters. This method also allows investigation of the DNA repair system and can be extended to other proteins such as MutL, MutH, or UvrD. We believe that the use of randomly stained DNA with a dye such as SG or other dyes that bind to dsDNA and single-cysteine variants of proteins with a covalently attached dye can be a universal method for investigations of other systems such as BER, NER, recombination, or replication.

Chapter 7

Summary

In **Chapter 1**, I describe the mechanism of the MMR system, give a brief overview, and explain how the system corrects errors when mismatched nucleotides are incorrectly incorporated during replication. I consider in detail the structure and functions of the four main proteins involved in E.coli DNA correction: MutS, MutL, MutH, and UvrD. I describe the multiple conformational changes in MutS adopting during mismatch repair, from a state where MutS scans DNA for a mismatch, going through many different states, to an active sliding clamp state and its variation with kinked clamp domain. I focus on the process of recruitment of a second MutL protein, which also adopts different conformational changes, and further discuss various modern models of the coordination of the ternary complex DNA-MutS-MutL with the latent endonuclease MutH, which introduces nicks at the hemimethylated GATC site. It also reveals the role of UvrD helicase responsible for DNA unwinding in an MM-dependent manner activated by MutS and MutL at the nick.

In **chapter 2** I worked on developing a novel method for probing MutL DNA-binding center using Förster Resonance Energy Transfer (FRET) assay. I focused on binding MutL to the DNA in particular, I investigated the binding of the N-terminal domain (NTD). Several variants of MutL with one cysteine (scMutL) have been modified with a maleimide fluorophore ATTO647N. The FRET intensity between labeled scMutL proteins and stained by SYTOX Blue DNA was inversely correlated with the distances from Cys residues to the DNA center calculated from the crystal structure of the MutS-MutL and modeled B-DNA inside this complex. The study of the binding of various single-cysteine scMutL variants to the DNA in FRET-based assay showed that the highest signal intensity was observed for the MutL (T218C) and MutL (A251C) variants, which indicates the proximity of positions 218 and 251 to DNA in the MutL-DNA complex. Indeed, these results were confirmed by our collaborators using site-specific cross-linking. In particular, the highest reaction yield was observed for the positions Cys218 and Cys251 of MutL cross-linked with reactive DNA, demonstrating the proximity of these positions to the DNA in the MutL-DNA complex.

Chapter 3 takes a step further and investigates this ternary DNA-MutS-MutL complex using covalent cross-linking of single-cysteine variants of MutS. I focused on the first step of DNA mismatch repair, namely on the conformational changes of MutS and its functions. Using site-specific cross-linking with the single-cysteine MutS variant, I trapped a transient active state of MutS on DNA. Combining this approach with the FRET method developed in **chapter 2**, I show that this transient

state is important for the next steps in the DNA repair pathway, namely in the MutL recruitment, and likely resembles the mismatch-active signaling clamp state of MutS on DNA. To investigate an intermediate conformation between the mismatch-bound and sliding clamp states, I used chemically modified oligonucleotide with a sulfur reactive cross-linker and show mismatch and ATP-dependent movement of the connector domain. At the same time, the mismatch domain cross-linked to the DNA is incapable of subsequent activation of MutL. Thus, showing the necessity for conformational changes predicted earlier due to X-ray structural analysis, but with one remark that the level domain should be flexible. This hypothesis was later confirmed by new data obtained using the Cryo-Em structures described in **chapter 4**.

Chapter 4 focuses on the conformational changes in MutS, in particular the structures that characterize the different states of the protein. Using point mutations, cross-linking and structural methods such as Cryo-EM, I have shown that there is a crossover of clamp domains in the presence of ATP and DNA which is located inside the hole of such a complex. This sliding clamp of MutS is an active state of the protein. Covalent cross-linking of MutS on DNA leads to its long-term activity even in the presence of DNA traps. Before going into the sliding clamping state from the mismatch state, MutS undergoes multiple changes, including the rotation of the domains. Using Cryo-EM data and cross-linking of modified DNA to a single cysteine MutS variant, it has been demonstrated that connector domain rotation takes place in presence of ATP, thereby confirming the existence of an intermediate state. In this chapter, we used cross-linking assay developed in **chapter 3** combining with the improved SG fluorescence FRET-based method developed in **chapter 2**.

Inspired by the cross-linking results obtained earlier and described in **chapters 2, 3, 4**, I attempted in **chapter 5** to stabilize the ternary MutL-MutH-DNA complex. So far, no crystal structure of full-length MutL linked to the DNA or MutH has been obtained. I used a site-specific cross-linking approach to investigate a possible active MutL-MutH-DNA complex from *E. coli*. I used a single cysteine variant $MutL^{480C}$ for site-specific cross-linking with a single cysteine variant $MutH^{223C}$. The $MutL^{480C}$ and $MutH^{223C}$ protein complex, modified with thiol-reactive cross-linkers, showed surprising activity. The MutS-independent ATP-dependent activity of the complex was observed, showing the importance of complex stabilization, as well as its equally important activation. Nicking experiments on DNA with various nucleotides showed the highest activity for the non-hydrolyzable analog of ATP, which was shown earlier leads to the condensed form of MutL. At the same time, cross-linking did not cause a decrease in the nuclease activity of the cross-linked activated in MutS and ATP dependent manner. All this may indicate the existence of a stable active complex, which can be further studied using structural methods.

Summing up all the information obtained about the conformational changes of the protein, as well as the fluorescence method developed and described in the previous chapters, I have found a practical application of this knowledge described in **chapter 6**. I demonstrate a novel sensitive, homogeneous fluorescence-based assay for the detection of DNA mismatches and some damages using mismatch repair protein MutS. The method is based on Förster resonance energy transfer between non-covalent dye SYBR Green I staining DNA and Alexa Fluor 647 covalently attached to the single cysteine MutS variant. Here I use not only the ability of MutS

to binds to the mismatch DNA but rather ATP-dependent loading of multiple MutS in the sliding clamp state, which dramatically increases the sensitivity of the assay. The assay is characterized on the well-studied 3 kb circle substrate with defined G/T mismatch. Moreover, I investigated the effect of DNA treatments with various procedures such as deamination, UV radiation, modification with cisplatin, and heating by MutS recognition. The method can be used for both circular and linear substrates. It was shown that the presence of MutL leads to the stabilization of the complex.

Bibliography

1. Friedberg, E. C. DNA damage and repair. *Nature* **421,** 436–440. https://doi.org/10.1038/nature01408 (2003).
2. Hoeijmakers, J. H. J. Genome maintenance mechanisms for preventing cancer. *Nature* **411,** 366–74. https://doi.org/10.1038/35077232 (2001).
3. David, S. S., O'Shea, V. L. & Kundu, S. Base-excision repair of oxidative DNA damage. *Nature* **447,** 941–950. https://doi.org/10.1038/nature05978 (2007).
4. Sancar, A. DNA excision repair. *Annu. Rev. Biochem.* **65,** 43–81. https://doi.org/10.1146/annurev.bi.65.070196.000355 (1996).
5. Friedberg, E. C. How nucleotide excision repair protects against cancer. *Nat. Rev. Cancer* **1,** 22–33. https://doi.org/10.1038/35094000 (2001).
6. Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.* **4,** 712–720. https://doi.org/10.1038/nrm1202 (2003).
7. Iyer, R. R., Pluciennik, A., Burdett, V. & Modrich, P. L. DNA mismatch repair: Functions and mechanisms. *Chem. Rev.* **106,** 302–323. https://doi.org/10.1021/cr0404794 (2006).
8. Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res.* **5,** 282–304. https://www.cambridge.org/core/article/mechanism-for-gene-conversion-in-fungi/E11586A6605C2A54C648BACEABECF954 (1964).
9. Ephrussi-Taylor, H. & Gray, T. C. Genetic Studies of Recombining DNA in Pneumococcal Transformation. *J. Gen. Physiol.* **49,** 211–231. https://doi.org/10.1085/jgp.49.6.211 (1966).
10. Wildenberg, J. & Meselson, M. Mismatch repair in heteroduplex DNA. *Proc. Natl. Acad. Sci.* **72,** 2202 LP –2206. http://www.pnas.org/content/72/6/2202.abstract (1975).
11. Wagner, R. & Meselson, M. Repair tracts in mismatched DNA heteroduplexes. 1976. *DNA Repair (Amst).* **4,** 126–30, discussion 103–4, 131. http://www.ncbi.nlm.nih.gov/pubmed/15609448 (2005).
12. Pukkila, P. J., Peterson, J., Herman, G., Modrich, P. & Meselson, M. Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in Escherichia coli. *Genetics* **104,** 571–582. http://www.genetics.org/content/104/4/571.abstract (1983).
13. Nevers, P. & Spatz, H.-C. Escherichia coli mutants uvr D and uvr E deficient in gene conversion of λ-heteroduplexes. *Mol. Gen. Genet. MGG* **139,** 233–243. http://link.springer.com/10.1007/BF00268974 (1975).

14. Glickman, B. W. & Radman, M. Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. *Proc. Natl. Acad. Sci.* **77,** 1063–1067. http://www.pnas.org/cgi/doi/10.1073/pnas.77.2.1063 (1980).

15. Lu, A. L., Clark, S. & Modrich, P. Methyl-directed repair of DNA base-pair mismatches in vitro. *Proc. Natl. Acad. Sci.* **80,** 4639–4643. http://www.pnas.org/cgi/doi/10.1073/pnas.80.15.4639 (1983).

16. Su, S. S., Lahue, R. S., Au, K. G. & Modrich, P. Mispair specificity of methyl-directed DNA mismatch correction in vitro. *J. Biol. Chem.* **263,** 6829–6835. https://www.sciencedirect.com/science/article/pii/S0021925818687186 (1988).

17. Lahue, R. S., Au, K. G. & Modrich, P. DNA mismatch correction in a defined system. *Science (80-.).* **245,** 160 LP –164. http://science.sciencemag.org/content/245/4914/160.abstract (1989).

18. Grilley, M., Griffith, J. & Modrich, P. Bidirectional excision in methyl-directed mismatch repair. *J. Biol. Chem.* **268,** 11830–11837. https://www.sciencedirect.com/science/article/pii/S0021925819502757 (1993).

19. Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S. T. & Modrich, P. In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 6765–6770. http://www.pnas.org/content/98/12/6765.abstract (2001).

20. Cooper, D. L., Lahue, R. S. & Modrich, P. Methyl-directed mismatch repair is bidirectional. *J. Biol. Chem.* **268,** 11823–11829. https://www.sciencedirect.com/science/article/pii/S0021925819502745 (1993).

21. López de Saro, F. J. & O'Donnell, M. Interaction of the β sliding clamp with MutS, ligase, and DNA polymerase I. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8376–8380. http://www.pnas.org/content/98/15/8376.abstract (2001).

22. López De Saro, F. J. *et al.* The β Sliding Clamp Binds to Multiple Sites within MutL and MutS. *J. Biol. Chem.* **281,** 14340–14349. https://www.sciencedirect.com/science/article/pii/S0021925820780804 (2006).

23. Kunkel, T. A., Erie, D. A. & Thomas A. Kunkel; Dorothy A. Erie. DNA Mismatch Repair. *DNA Repair Cancer* **74,** 195–232. https://doi.org/10.1146/annurev.biochem.74.082803.133243 (2013).

24. Längle-Rouault, F., Maenhaut-Michel, G. & Radman, M. GATC sequences, DNA nicks and the MutH function in Escherichia coli mismatch repair. *EMBO J.* **6,** 1121–1127. https://doi.org/10.1002/j.1460-2075.1987.tb04867.x (1987).

25. Lujan, S. A. *et al.* Mismatch Repair Balances Leading and Lagging Strand DNA Replication Fidelity. *PLOS Genet.* **8,** e1003016. https://doi.org/10.1371/journal.pgen.1003016 (2012).

26. Barras, F. & Marinus, M. G. The great GATC: DNA methylation in E. coli. *Trends Genet.* **5,** 139–143. https://www.sciencedirect.com/science/article/pii/0168952589900541 (1989).

27. Campbell, J. L. & Kleckner, N. E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* **62,** 967–979. https://www.sciencedirect.com/science/article/pii/009286749090271F (1990).

28. Lynch, H. T. *et al.* Clinical impact of molecular genetic diagnosis, genetic counseling, and management of hereditary cancer. *Cancer* **86,** 2457–2463. https://doi.org/10.1002/(SICI)1097-0142(19991201)86:11+%7B%5C%%7D3C2457::AID-CNCR2%7B%5C%%7D3E3.0.CO (1999).

29. Friedhoff, P., Li, P. & Gotthardt, J. Protein-protein interactions in DNA mismatch repair. *DNA Repair (Amst).* **38,** 50–57. https://www.sciencedirect.com/science/article/pii/S1568786415300471 (2016).

30. Galio, L., Bouquet, C. & Brooks, P. ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. *Nucleic Acids Res.* **27,** 2325–2331. https://doi.org/10.1093/nar/27.11.2325 (1999).

31. Spampinato, C. & Modrich, P. The MutL ATPase Is Required for Mismatch Repair. *J. Biol. Chem.* **275,** 9863–9869. https://www.sciencedirect.com/science/article/pii/S0021925818301091 (2000).

32. Acharya, S., Foster, P. L., Brooks, P. & Fishel, R. The Coordinated Functions of the E. coli MutS and MutL Proteins in Mismatch Repair of hereditary nonpolyposis colorectal cancer (HNPCC) families (for review see Muller and Fishel, 2002, and references therein). *Mol. Cell* **12,** 233–246. https://www.sciencedirect.com/science/article/pii/S1097276503002193 (2003).

33. Junop, M. S., Yang, W., Funchain, P., Clendenin, W. & Miller, J. H. In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair (Amst).* **2,** 387–405. https://www.sciencedirect.com/science/article/pii/S1568786402002458 (2003).

34. Mechanic, L. E., Frankel, B. A. & Matson, S. W. Escherichia coli MutL loads DNA helicase II onto DNA. *J. Biol. Chem.* **275,** 38337–38346. https://www.sciencedirect.com/science/article/pii/S0021925819559622 (2000).

35. Welsh, K. M., Lu, A.-L., Clark, S. & Modrich, P. Isolation and characterization of the Escherichia coli mutH gene product. *J. Biol. Chem.* **262,** 15624–15629. https://www.sciencedirect.com/science/article/pii/S0021925818477721 (1987).

36. Dao, V., Modrich, P., Carolina, N. & Muts, E. Mismatch-, MutS-, MutL-, and Helicase II-dependent Unwinding from the Single-strand Break of an Incised Heteroduplex. *J. Biol. Chem.* **273,** 9202–9207. https://www.sciencedirect.com/science/article/pii/S0021925818496202 (1998).

37. Goellner, E. M., Putnam, C. D. & Kolodner, R. D. Exonuclease 1-dependent and independent mismatch repair. *DNA Repair (Amst).* **32,** 24–32. http://dx.doi.org/10.1016/j.dnarep.2015.04.010 (2015).

38. Liu, J. *et al.* MutL sliding clamps coordinate exonuclease-independent Escherichia coli mismatch repair. *Nat. Commun.* **10,** 5294. http://dx.doi.org/10.1038/s41467-019-13191-5 (2019).

39. Modrich, P. Mechanisms in Eukaryotic Mismatch Repair. *J. Biol. Chem.* **281,** 30305–30309. https://doi.org/10.1074/jbc.R600022200 (2006).

40. Putnam, C. D. Evolution of the methyl directed mismatch repair system in Escherichia coli. *DNA Repair (Amst).* **38,** 32–41. https://www.sciencedirect.com/science/article/pii/S1568786415300306 (2016).

41. Kadyrov, F. A., Dzantiev, L., Constantin, N. & Modrich, P. Endonucleolytic Function of MutLα in Human Mismatch Repair. *Cell* **126,** 297–308. https://www.sciencedirect.com/science/article/pii/S0092867406008129 (2006).

42. Pillon, M. C. *et al.* Structure of the Endonuclease Domain of MutL: Unlicensed to Cut. *Mol. Cell* **39,** 145–151. https://www.sciencedirect.com/science/article/pii/S1097276510004946 (2010).

43. Pluciennik, A. *et al.* PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. *Proc. Natl. Acad. Sci.* **107,** 16066 LP –16071. http://www.pnas.org/content/107/37/16066.abstract (2010).

44. Kadyrov, F. A. *et al.* Saccharomyces cerevisiae MutLα Is a Mismatch Repair Endonuclease. *J. Biol. Chem.* **282,** 37181–37190. https://www.sciencedirect.com/science/article/pii/S0021925820552411 (2007).

45. Almawi, A. W. *et al.* Binding of the regulatory domain of MutL to the sliding β-clamp is species specific. *Nucleic Acids Res.* **47,** 4831–4842. https://doi.org/10.1093/nar/gkz115 (2019).

46. Ghodgaonkar, M. M. M. *et al.* Ribonucleotides Misincorporated into DNA Act as Strand-Discrimination Signals in Eukaryotic Mismatch Repair. *Mol. Cell* **50,** 323–332. https://www.sciencedirect.com/science/article/pii/S1097276513002232 (2013).

47. Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B. & Kunkel, T. A. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol. Cell* **50,** 437–443. http://dx.doi.org/10.1016/j.molcel.2013.03.017 (2013).

48. Su, S. S. & Modrich, P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proc. Natl. Acad. Sci. U. S. A.* **83,** 5057–61. http://www.ncbi.nlm.nih.gov/pubmed/3014530%7B%5C%%7D0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC323889%20http://www.pnas.org/content/83/14/5057.abstract (1986).

49. Parker, B. O. & Marinus, M. G. Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli. *Proc. Natl. Acad. Sci.* **89,** 1730 LP –1734. http://www.pnas.org/content/89/5/1730.abstract (1992).

50. Eisen, J. A. A phylogenomic study of the MutS family of proteins. *Nucleic Acids Res.* **26,** 4291–4300. https://doi.org/10.1093/nar/26.18.4291 (1998).

51. Lin, Z., Nei, M. & Ma, H. The origins and early evolution of DNA mismatch repair genes - Multiple horizontal gene transfers and co-evolution. *Nucleic Acids Res.* **35,** 7591–7603. https://doi.org/10.1093/nar/gkm921 (2007).

52. Groothuizen, F. S. & Sixma, T. K. The conserved molecular machinery in DNA mismatch repair enzyme structures. *DNA Repair (Amst).* **38,** 14–23. https://www.sciencedirect.com/science/article/pii/S1568786415300537 (2016).

53. Obmolova, G., Ban, C., Hsieh, P. & Yang, W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407,** 703–710. https://doi.org/10.1038/35037509 (2000).

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

55. Groothuizen, F. S. *et al.* Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp formation. *Nucleic Acids Res.* **41,** 8166–8181. https://doi.org/10.1093/nar/gkt582 (2013).

56. Mendillo, M. L., Putnam, C. D. & Kolodner, R. D. 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. https://www.sciencedirect.com/science/article/pii/S0021925820649547 (2007).

57. Manelyte, L., Urbanke, C., Giron-Monzon, L. & Friedhoff, P. Structural and functional analysis of the MutS C-terminal tetramerization domain. *Nucleic Acids Res.* **34,** 5270–5279. https://doi.org/10.1093/nar/gkl489 (2006).

58. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1,** 945–951. https://doi.org/10.1002/j.1460-2075.1982.tb01276.x (1982).

59. Graham, W. J., Putnam, C. D. & Kolodner, R. D. The properties of Msh2–Msh6 ATP binding mutants suggest a signal amplification mechanism in DNA mismatch repair. *J. Biol. Chem.* **293,** 18055–18070. https://www.sciencedirect.com/science/article/pii/S0021925820311935 (2018).

60. Warren, J. J. *et al.* Structure of the Human MutSα DNA Lesion Recognition Complex. *Mol. Cell* **26,** 579–592. https://www.sciencedirect.com/science/article/pii/S1097276507002547 (2007).

61. Gupta, S., Gellert, M. & Yang, W. Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat. Struct. Mol. Biol.* **19,** 72–79. http://dx.doi.org/10.1038/nsmb.2175 (2012).

62. Bhairosing-Kok, D. *et al.* Sharp kinking of a coiled-coil in MutS allows DNA binding and release. *Nucleic Acids Res.* **47,** 8888–8898. https://doi.org/10.1093/nar/gkz649 (2019).

63. Jeong, C. *et al.* MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat. Struct. Mol. Biol.* **18,** 379–385. http://dx.doi.org/10.1038/nsmb.2009 (2011).

64. Gorman, J. *et al.* Dynamic Basis for One-Dimensional DNA Scanning by the Mismatch Repair Complex Msh2-Msh6. *Mol. Cell* **28,** 359–370. https://www.sciencedirect.com/science/article/pii/S1097276507006181 (2007).

65. Yang, Y., Sass, L. E., Du, C., Hsieh, P. & Erie, D. A. Determination of protein–DNA binding constants and specificities from statistical analyses of single molecules: MutS–DNA interactions. *Nucleic Acids Res.* **33,** 4322–4334. https://doi.org/10.1093/nar/gki708 (2005).

66. LeBlanc, S. J. *et al.* Coordinated protein and DNA conformational changes govern mismatch repair initiation by MutS. *Nucleic Acids Res.* **46,** 10782–10795. https://doi.org/10.1093/nar/gky865 (2018).

67. Qiu, R. *et al.* Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling. *EMBO J.* **31,** 2528–2540. http://dx.doi.org/10.1038/emboj.2012.95 (2012).

68. Natrajan, G. *et al.* 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. https://doi.org/10.1093/nar/gkg677 (2003).

69. Yamamoto, A., Schofield, M. J., Biswas, I. & Hsieh, P. Requirement for Phe36 for DNA binding and mismatch repair by Escherichia coli MutS protein. *Nucleic Acids Res.* **28,** 3564–3569. https://doi.org/10.1093/nar/28.18.3564 (2000).

70. Nirwal, S., Kulkarni, D. S., Sharma, A., Rao, D. N. & Nair, D. T. Mechanism of formation of a toroid around DNA by the mismatch sensor protein. *Nucleic Acids Res.* **46,** 256–266. https://doi.org/10.1093/nar/gkx1149 (2018).

71. Fernandez-Leiro, R. *et al.* The selection process of licensing a DNA mismatch for repair. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-021-00577-7 (2021).

72. Gradia, S., Acharya, S. & Fishel, R. The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch. *Cell* **91,** 995–1005. https://linkinghub.elsevier.com/retrieve/pii/S0092867400804900 (1997).

73. Allen, D. J. *et al.* MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.* **16,** 4467–4476. https://doi.org/10.1093/emboj/16.14.4467 (1997).

74. Schofield, M. J., Nayak, S., Scott, T. H., Du, C. & Hsieh, P. Interaction of Escherichia coli MutS and MutL at a DNA Mismatch. *J. Biol. Chem.* **276,** 28291–28299. https://doi.org/10.1074/jbc.M103148200 (2001).

75. Antony, E. & Hingorani, M. M. Mismatch Recognition-Coupled Stabilization of Msh2-Msh6 in an ATP-Bound State at the Initiation of DNA Repair. *Biochemistry* **42,** 7682–7693. https://doi.org/10.1021/bi034602h (2003).

76. Antony, E. & Hingorani, M. M. 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. https://doi.org/10.1021/bi049010t (2004).

77. Liu, J. *et al.* Cascading MutS and MutL sliding clamps control DNA diffusion to activate mismatch repair. *Nature* **539,** 583–587. https://doi.org/10.1038/nature20562 (2016).

78. Groothuizen, F. S. *et al.* MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *Elife* **4** (ed Joshua-Tor, L.) 1–24. https://doi.org/10.7554/eLife.06744 (2015).

79. Winkler, I. *et al.* Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in Escherichia coli. *J. Biol. Chem.* **286,** 17326–17337. https://www.sciencedirect.com/science/article/pii/S0021925820513616 (2011).

80. Kolodner, R. D. & Marsischky, G. T. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.* **9,** 89–96. https://www.sciencedirect.com/science/article/pii/S0959437X99800136 (1999).

81. Kolodner, R. Biochemistry and genetics of eukaryotic mismatch repair. *Genes Dev.* **10 12,** 1433–1442 (1996).

82. Zhai, J. & Hingorani, M. M. Saccharomyces cerevisiae Msh2-Msh6 DNA binding kinetics reveal a mechanism of targeting sites for DNA mismatch repair. *Proc. Natl. Acad. Sci.* **107,** 680–685. https://www.pnas.org/content/107/2/680 (2010).

83. Gradia, S. *et al.* hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell* **3,** 255–261. https://www.sciencedirect.com/science/article/pii/S1097276500803160 (1999).

84. Blackwell, L. J., Martik, D., Bjornson, K. P., Bjornson, E. S. & Modrich, P. Nucleotide-promoted Release of hMutSα from Heteroduplex DNA Is Consistent with an ATP-dependent Translocation Mechanism. *J. Biol. Chem.* **273,** 32055–32062. https://www.sciencedirect.com/science/article/pii/S0021925819590512 (1998).

85. Prolla, T. A., Pang, Q., Alani, E., Kolodner, R. D. & Liskay, R. M. MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science (80-.).* **265,** 1091 LP –1093. http://science.sciencemag.org/content/265/5175/1091.abstract (1994).

86. Umar, A. *et al.* Requirement for PCNA in DNA Mismatch Repair at a Step Preceding DNA Resynthesis. *Cell* **87,** 65–73. https://www.sciencedirect.com/science/article/pii/S0092867400813239 (1996).

87. Cannavo, E. *et al.* Expression of the MutL Homologue hMLH3 in Human Cells and its Role in DNA Mismatch Repair. *Cancer Res.* **65,** 10759 LP –10766. http://cancerres.aacrjournals.org/content/65/23/10759.abstract (2005).

88. Ban, C. & Yang, W. Structural basis for MutH activation in E.coli mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J.* **17,** 1526–1534. https://doi.org/10.1093/emboj/17.5.1526 (1998).

89. Guarné, A. *et al.* Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J.* **23,** 4134–4145. https://doi.org/10.1038/sj.emboj.7600412 (2004).

90. Ban, C. & Yang, W. Crystal structure and ATPase activity of MutL: Implications for DNA repair and mutagenesis. *Cell* **95,** 541–552. https://www.sciencedirect.com/science/article/pii/S0092867400816219 (1998).

91. Guarné, A., Junop, M. S. & Yang, W. Structure and function of the N-terminal 40 kDa fragment of human PMS2: a monomeric GHL ATPase. *EMBO J.* **20,** 5521–5531. https://doi.org/10.1093/emboj/20.19.5521 (2001).

92. Gueneau, E. *et al.* Structure of the MutLα C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. *Nat. Struct. Mol. Biol.* **20,** 461–468. https://doi.org/10.1038/nsmb.2511 (2013).

93. Kosinski, J., Steindorf, I., Bujnicki, J. M., Giron-Monzon, L. & Friedhoff, P. Analysis of the quaternary structure of the MutL C-terminal domain. *J. Mol. Biol.* **351,** 895–909. https://www.sciencedirect.com/science/article/pii/S0022283605007163 (2005).

94. Dutta, R. & Inouye, M. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* **25,** 24–28. https://www.sciencedirect.com/science/article/pii/S0968000499015030 (2000).

95. Ban, C., Junop, M. & Yang, W. Transformation of MutL by ATP Binding and Hydrolysis: A Switch in DNA Mismatch Repair. *Cell* **97,** 85–97. https://www.sciencedirect.com/science/article/pii/S0092867400807175 (1999).

96. Sacho, E. J., Kadyrov, F. A., Modrich, P., Kunkel, T. A. & Erie, D. A. Direct Visualization of Asymmetric Adenine Nucleotide-Induced Conformational Changes in MutLα. *Mol. Cell* **29,** 112–121. https://www.sciencedirect.com/science/article/pii/S1097276507008179 (2008).

97. Robertson, A., Pattishall, S. R. & Matson, S. W. The DNA Binding Activity of MutL Is Required for Methyl-directed Mismatch Repair in Escherichia coli. *J. Biol. Chem.* **281,** 8399–8408. https://www.sciencedirect.com/science/article/pii/S0021925819565395 (2006).

98. Selmane, T., Schofield, M. J., Nayak, S., Du, C. & Hsieh, P. Formation of a DNA Mismatch Repair Complex Mediated by ATP. *J. Mol. Biol.* **334,** 949–965. https://www.sciencedirect.com/science/article/pii/S0022283603012658 (2003).

99. Grilley, M., Welsh, K. M., Su, S. S. & Modrich, P. Isolation and Characterization of the Escherichia coli mutL Gene Product. *J. Biol. Chem.* **264,** 1000–1004. https://www.sciencedirect.com/science/article/pii/S0021925819850433 (1989).

100. Schofield, M. J. & Hsieh, P. DNA Mismatch Repair: Molecular Mechanisms and Biological Function. *Annu. Rev. Microbiol.* **57,** 579–608. https://doi.org/10.1146/annurev.micro.57.030502.090847 (2003).

101. Lebbink, J. H. G. *et al.* Magnesium Coordination Controls the Molecular Switch Function of DNA Mismatch Repair Protein MutS. *J. Biol. Chem.* **285,** 13131–13141. https://www.sciencedirect.com/science/article/pii/S0021925820550916 (2010).

102. Polosina, Y. Y. & Cupples, C. G. MutL: conducting the cell's response to mismatched and misaligned DNA. *BioEssays* **32,** 51–59. https://doi.org/10.1002/bies.200900089 (2010).

103. Hall, M. C. & Matson, S. W. The Escherichia coliMutL Protein Physically Interacts with MutH and Stimulates the MutH-associated Endonuclease Activity. *J. Biol. Chem.* **274,** 1306–1312. https://www.sciencedirect.com/science/article/pii/S0021925819882376 (1999).

104. Hall, M. C., Jordan, J., Matson, S. W., Jordan, R. & Matson, S. W. Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD. *EMBO J.* **17,** 1535–1541. https://doi.org/10.1093/emboj/17.5.1535 (1998).

105. Giron-Monzon, L. *et al.* Mapping protein-protein interactions between MutL and MutH by cross-linking. *J. Biol. Chem.* **279,** 49338–49345. https://www.sciencedirect.com/science/article/pii/S0021925819323294 (2004).

106. Ahrends, R. *et al.* Identifying an interaction site between MutH and the C-terminal domain of MutL by crosslinking, affinity purification, chemical coding and mass spectrometry. *Nucleic Acids Res.* **34,** 3169–3180. https://doi.org/10.1093/nar/gkl407 (2006).

107. Toedt, G. H., Krishnan, R. & Friedhoff, P. Site-specific protein modification to identify the MutL interface of MutH. *Nucleic Acids Res.* **31,** 819–825. https://doi.org/10.1093/nar/gkg191 (2003).

108. Au, K. G., Welsh, K. & Modrich, P. Initiation of methyl-directed mismatch repair. *J. Biol. Chem.* **267,** 12142–12148. https://www.sciencedirect.com/science/article/pii/S0021925819498165 (1992).

109. Florés, M.-J., Sanchez, N. & Michel, B. A fork-clearing role for UvrD. *Mol. Microbiol.* **57,** 1664–1675. https://doi.org/10.1111/j.1365-2958.2005.04753.x (2005).

110. Bruand, C. & Ehrlich, S. D. UvrD-dependent replication of rolling-circle plasmids in Escherichia coli. *Mol. Microbiol.* **35,** 204–210. https://doi.org/10.1046/j.1365-2958.2000.01700.x (2000).

111. Veaute, X. *et al.* UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in Escherichia coli. *EMBO J.* **24,** 180–189. https://doi.org/10.1038/sj.emboj.7600485 (2005).

112. Kemp, M. G. & Sancar, A. DNA excision repair. *Cell Cycle* **11,** 2997–3002. https://doi.org/10.4161/cc.21126 (2012).

113. Hickson, I. D., Arthur, H. M., Bramhill, D. & Emmerson, P. T. The E. coli uvrD gene product is DNA helicase II. *MGG Mol. Gen. Genet.* **190,** 265–270. https://doi.org/10.1007/BF00330649 (1983).

114. Yamaguchi, M., Dao, V. & Modrich, P. MutS and MutL Activate DNA Helicase II in a Mismatch-dependent Manner. *J. Biol. Chem.* **273,** 9197–9201. https://www.sciencedirect.com/science/article/pii/S0021925818496196 (1998).

115. Ordabayev, Y. A., Nguyen, B., Kozlov, A. G., Jia, H. & Lohman, T. M. UvrD helicase activation by MutL involves rotation of its 2B subdomain. *Proc. Natl. Acad. Sci.* **116,** 16320–16325. http://www.pnas.org/content/116/33/16320.abstract (2019).

116. Matson, S. W. Escherichia coli DNA helicase II (uvrD gene product) catalyzes the unwinding of DNA.RNA hybrids in vitro. *Proc. Natl. Acad. Sci.* **86,** 4430 LP –4434. http://www.pnas.org/content/86/12/4430.abstract (1989).

117. Lee, K. S., Balci, H., Jia, H., Lohman, T. M. & Ha, T. Direct imaging of single UvrD helicase dynamics on long single-stranded DNA. *Nat. Commun.* **4,** 1878. https://doi.org/10.1038/ncomms2882 (2013).

118. Mechanic, L. E., Hall, M. C. & Matson, S. W. Escherichia coli DNA Helicase II Is Active as a Monomer. *J. Biol. Chem.* **274,** 12488–12498. https://www.sciencedirect.com/science/article/pii/S0021925819733808 (1999).

119. Comstock, M. J. *et al.* Direct observation of structure-function relationship in a nucleic acid–processing enzyme. *Science (80-.).* **348,** 352 LP –354. http://science.sciencemag.org/content/348/6232/352.abstract (2015).

120. Hermans, N. *et al.* Dual daughter strand incision is processive and increases the efficiency of DNA mismatch repair. *Nucleic Acids Res.* **44,** 6770–6786. https://doi.org/10.1093/nar/gkw411 (2016).

121. Nguyen, B., Ordabayev, Y., Sokoloski, J. E., Weiland, E. & Lohman, T. M. Large domain movements upon UvrD dimerization and helicase activation. *Proc. Natl. Acad. Sci.* **114,** 12178 LP –12183. http://www.pnas.org/content/114/46/12178.abstract (2017).

122. Lee, J. Y. & Yang, W. UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke. *Cell* **127,** 1349–1360. https://www.sciencedirect.com/science/article/pii/S0092867406016011 (2006).

123. Jia, H. *et al.* Rotations of the 2B sub-domain of E. coli UvrD helicase/translocase coupled to nucleotide and DNA binding. *J. Mol. Biol.* **411,** 633–648. https://www.sciencedirect.com/science/article/pii/S0022283611006656 (2011).

124. Arslan, S., Khafizov, R., Thomas, C. D., Chemla, Y. R. & Ha, T. Engineering of a superhelicase through conformational control. *Science (80-.).* **348,** 344 LP –347. http://science.sciencemag.org/content/348/6232/344.abstract (2015).

125. Brendza, K. M. *et al.* Autoinhibition of Escherichia coli Rep monomer helicase activity by its 2B subdomain. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 10076 LP –10081. http://www.pnas.org/content/102/29/10076.abstract (2005).

126. Cheng, W. *et al.* The 2B domain of the Escherichia coli Rep protein is not required for DNA helicase activity. *Proc. Natl. Acad. Sci.* **99,** 16006 LP –16011. http://www.pnas.org/content/99/25/16006.abstract (2002).

127. Kolodner, R. D., Mendillo, M. L. & Putnam, C. D. Coupling distant sites in DNA during DNA mismatch repair. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 12953–12954. http://www.pnas.org/content/104/32/12953.abstract (2007).

128. Li, G.-M. New insights and challenges in mismatch repair: Getting over the chromatin hurdle. *DNA Repair (Amst).* **19,** 48–54. https://www.sciencedirect.com/science/article/pii/S1568786414000986?via%7B%5C%%7D3Dihub (2014).

129. Junop, M. S., Obmolova, G., Rausch, K., Hsieh, P. & Yang, W. Composite Active Site of an ABC ATPase: MutS Uses ATP to Verify Mismatch Recognition and Authorize DNA Repair. *Mol. Cell* **7,** 1–12. https://www.sciencedirect.com/science/article/pii/S1097276501001496 (2001).

130. Elez, M., Radman, M. & Matic, I. Stoichiometry of MutS and MutL at unrepaired mismatches in vivo suggests a mechanism of repair. *Nucleic Acids Res.* **40,** 3929–3938. https://doi.org/10.1093/nar/gkr1298 (2012).

131. Hall, M. C., Wang, H., Erie, D. A. & Kunkel, T. A. High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer11Edited by M. Belfort. *J. Mol. Biol.* **312,** 637–647. https://www.sciencedirect.com/science/article/pii/S0022283601949587 (2001).

132. Jiang, Y. & Marszalek, P. E. Atomic force microscopy captures MutS tetramers initiating DNA mismatch repair. *EMBO J.* **30,** 2881–2893. http://dx.doi.org/10.1038/emboj.2011.180 (2011).

133. Pluciennik, A. & Modrich, P. Protein roadblocks and helix discontinuities are barriers to the initiation of mismatch repair. *Proc. Natl. Acad. Sci.* **104,** 12709 LP –12713. http://www.pnas.org/content/104/31/12709.abstract (2007).

134. Mardenborough, Y. S. N. N. *et al.* The unstructured linker arms of MutL enable GATC site incision beyond roadblocks during initiation of DNA mismatch repair. *Nucleic Acids Res.* **47,** 11667–11680. https://doi.org/10.1093/nar/gkz834 (2019).

135. Ortega, J., Lee, G. S., Gu, L., Yang, W. & Li, G.-M. Mispair-bound human MutS–MutL complex triggers DNA incisions and activates mismatch repair. *Cell Res.* http://dx.doi.org/10.1038/s41422-021-00468-y (2021).

136. Elez, M. *et al.* Seeing mutations in living cells. *Curr. Biol.* **20,** 1432–1437. http://dx.doi.org/10.1016/j.cub.2010.06.071 (2010).

137. Hombauer, H., Campbell, C. S., Smith, C. E., Desai, A. & Kolodner, R. D. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* **147,** 1040–1053. http://dx.doi.org/10.1016/j.cell.2011.10.025 (2011).

138. Qiu, R. *et al.* MutL traps MutS at a DNA mismatch. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 10914–10919. http://www.pnas.org/content/112/35/10914.abstract (2015).

139. Blackwell, L. J., Wang, S. & Modrich, P. DNA Chain Length Dependence of Formation and Dynamics of hMutSα·hMutLα·Heteroduplex Complexes. *J. Biol. Chem.* **276,** 33233–33240. https://www.sciencedirect.com/science/article/pii/S0021925820778580 (2001).

140. Hao, P. *et al.* Recurrent mismatch binding by MutS mobile clamps on DNA localizes repair complexes nearby. *Proc. Natl. Acad. Sci. U. S. A.* **117,** 17775–17784. http://www.pnas.org/content/117/30/17775.abstract (2020).

141. Putnam, C. D. MutS sliding clamps on an uncertain track to DNA mismatch repair. *Proc. Natl. Acad. Sci. U. S. A.* **117,** 20351–20353. http://www.pnas.org/content/117/34/20351.abstract (2020).

142. Hermanson, G. T. in *Bioconjugate Tech.* (ed Hermanson, G. T. B. T. B. T. (E.) 229–258 (Academic Press, Boston, 2013). https://www.sciencedirect.com/science/article/pii/B9780123822390000030.

143. Leitner, A., Faini, M., Stengel, F. & Aebersold, R. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci.* **41,** 20–32. http://dx.doi.org/10.1016/j.tibs.2015.10.008 (2016).

144. Debelouchina, G. T. & Muir, T. W. A molecular engineering toolbox for the structural biologist. *Q. Rev. Biophys.* **50,** e7. https://www.cambridge.org/core/article/molecular-engineering-toolbox-for-the-structural-biologist/63E61209AA811B509D4E15B31623EA6C (2017).

145. Friedhoff, P. *et al. Use of Single-Cysteine Variants for Trapping Transient States in DNA Mismatch Repair* 1st ed. (ed Eichman, B. F. B. T. M. i. E.) 77–101. http://dx.doi.org/10.1016/bs.mie.2017.03.025 (Elsevier Inc., 2017).

146. Banerjee, A. & Verdine, G. L. A nucleobase lesion remodels the interaction of its normal neighbor in a DNA glycosylase complex. *Proc. Natl. Acad. Sci.* **103,** 15020 LP –15025. http://www.pnas.org/content/103/41/15020.abstract (2006).

147. Monakhova, M. *et al.* Chromatographic isolation of the functionally active MutS protein covalently linked to deoxyribonucleic acid. *J. Chromatogr. A* **1389,** 19–27. http://dx.doi.org/10.1016/j.chroma.2015.02.045 (2015).

148. Heinze, R. J. *et al.* Covalently trapping MutS on DNA to study DNA mismatch recognition and signaling. *Mol. Biosyst.* **8,** 1861–1864. http://dx.doi.org/10.1039/C2MB25086A (2012).

149. Monakhova, M. *et al.* Probing the DNA-binding center of the MutL protein from the Escherichia coli mismatch repair system via crosslinking and Förster resonance energy transfer. *Biochimie* **171-172,** 43–54. https://doi.org/10.1016/j.biochi.2020.02.004 (2020).

150. Lakowicz, J. R. *General features of protein fluorescence* 1–961 (2006).

151. Cristóvão, M. *et al.* Single-molecule multiparameter fluorescence spectroscopy reveals directional MutS binding to mismatched bases in DNA. *Nucleic Acids Res.* **40,** 5448–5464. https://doi.org/10.1093/nar/gks138 (2012).

152. Bende, S. M. & Grafström, R. H. The DNA binding properties of the MutL protein isolated from Escherichia coli. *Nucleic Acids Res.* **19,** 1549–1555. https://doi.org/10.1093/nar/19.7.1549 (1991).

153. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **15,** 669–676. https://doi.org/10.1038/s41592-018-0085-0 (2018).

154. Sharma, S. *et al.* Monitoring Protein Conformation along the Pathway of Chaperonin-Assisted Folding. *Cell* **133,** 142–153. https://www.sciencedirect.com/science/article/pii/S0092867408002134 (2008).

155. Goodall, J. J., Chen, G. J. & Page, M. G. Essential Role of Histidine 20 in the Catalytic Mechanism of Escherichia coli Peptidyl-tRNA Hydrolase. *Biochemistry* **43,** 4583–4591. https://doi.org/10.1021/bi0302200 (2004).

156. Schorzman, A. N. *et al.* Modeling of the DNA-binding site of yeast Pms1 by mass spectrometry. *DNA Repair (Amst).* **10,** 454–465. http://dx.doi.org/10.1016/j.dnarep.2011.01.010 (2011).

157. Fukui, K. *et al.* Crystal structure and DNA-binding property of the ATPase domain of bacterial mismatch repair endonuclease MutL from Aquifex aeolicus. *Biochim. Biophys. Acta - Proteins Proteomics* **1865,** 1178–1187. http://dx.doi.org/10.1016/j.bbapap.2017.06.024 (2017).

158. Rayssiguier, C., Thaler, D. S. & Radman, M. The barrier to recombination between Escherichia coli and Salmonella typhimurium is disrupted in mismatch-repair mutants. *Nature* **342,** 396–401. https://doi.org/10.1038/342396a0 (1989).

159. Heo, S. D., Cho, M., Ku, J. K. & Ban, C. Steady-state ATPase activity of E. coli MutS modulated by its dissociation from heteroduplex DNA. *Biochem. Biophys. Res. Commun.* **364,** 264–269. https://www.sciencedirect.com/science/article/pii/S0006291X07021316 (2007).

160. Lynch, H. T. *et al.* Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin. Genet.* **76,** 1–18. https://doi.org/10.1111/j.1399-0004.2009.01230.x (2009).

161. Jae, Y. L. *et al.* MutH complexed with hemi- and unmethylated DNAs: Coupling base recognition and DNA cleavage. *Mol. Cell* **20,** 155–166. https://www.sciencedirect.com/science/article/pii/S1097276505015613?via%7B%5C%%7D3Dihub (2005).

162. Kosiński, J. Dissecting molecular basis of dysfunction of protein complexes involved in DNA mismatch repair by characterization of their structure and mutual interactions , and their mechanism of action. (2009).

163. Winkler, I. Funktionelle und strukturelle Untersuchungen zur ternären Komplexbildung zwischen MutS und MutL im mismatch -Reparatursystem (2010).

164. Guarné, A. The functions of MutL in mismatch repair: The power of multitasking. *Prog. Mol. Biol. Transl. Sci.* **110** (eds Doetsch, P. W. B. T. P. i. M. B. & Science, T.) 41–70. https://www.sciencedirect.com/science/article/pii/B9780123876652000031 (2012).

165. Binkowski, B. F., Richmond, K. E., Kaysen, J., Sussman, M. R. & Belshaw, P. J. Correcting errors in synthetic DNA through consensus shuffling. *Nucleic Acids Res.* **33,** e55–e55. https://doi.org/10.1093/nar/gni053 (2005).

166. Lubock, N. B., Zhang, D., Sidore, A. M., Church, G. M. & Kosuri, S. A systematic comparison of error correction enzymes by next-generation sequencing. *Nucleic Acids Res.* **45,** 9206–9217. https://doi.org/10.1093/nar/gkx691 (2017).

167. Zhang, Y., Hua, R.-n. N. & Zhang, C.-y. Y. Integration of Enzymatic Labeling with Single-Molecule Detection for Sensitive Quantification of Diverse DNA Damages. *Anal. Chem.* **92,** 4700–4706. https://doi.org/10.1021/acs.analchem.9b04547 (2020).

168. Babic, I., Andrew, S. E. & Jirik, F. R. MutS interaction with mismatch and alkylated base containing DNA molecules detected by optical biosensor. *Mutat. Res. Mol. Mech. Mutagen.* **372,** 87–96. https://www.sciencedirect.com/science/article/pii/S0027510796001704 (1996).

169. Gotoh, M. *et al.* Rapid method for detection of point mutations using mismatch binding protein (MutS) and an optical biosensor. *Genet. Anal. Biomol. Eng.* **14,** 47–50. https://www.sciencedirect.com/science/article/pii/S1050386297000090 (1997).

170. Nakano, S.-i. I., Kanzaki, T., Nakano, M., Miyoshi, D. & Sugimoto, N. Measurements of the Binding of a Large Protein Using a Substrate Density-Controlled DNA Chip. *Anal. Chem.* **83,** 6368–6372. https://doi.org/10.1021/ac201312d (2011).

171. Drabovich, A. P. & Krylov, S. N. Identification of Base Pairs in Single-Nucleotide Polymorphisms by MutS Protein-Mediated Capillary Electrophoresis. *Anal. Chem.* **78,** 2035–2038. https://doi.org/10.1021/ac0520386 (2006).

172. Su, X., Robelek, R., Wu, Y., Wang, G. & Knoll, W. Detection of Point Mutation and Insertion Mutations in DNA Using a Quartz Crystal Microbalance and MutS, a Mismatch Binding Protein. *Anal. Chem.* **76,** 489–494. https://doi.org/10.1021/ac035175g (2004).

173. Sun, H. B. & Yokota, H. MutS-Mediated Detection of DNA Mismatches Using Atomic Force Microscopy. *Anal. Chem.* **72,** 3138–3141. https://doi.org/10.1021/ac991263i (2000).

174. Geschwind, D. H., Rhee, R. & Nelson, S. F. A biotinylated MutS fusion protein and its use in a rapid mutation screening technique. *Genet. Anal. Biomol. Eng.* **13,** 105–111. https://www.sciencedirect.com/science/article/pii/S1050386295001603 (1996).

175. Cho, M. *et al.* Electrochemical detection of mismatched DNA using a MutS probe. *Nucleic Acids Res.* **34,** e75–e75. https://doi.org/10.1093/nar/gkl364 (2006).

176. Masařík, M., Cahová, K., Kizek, R., Paleček, E. & Fojta, M. Label-free voltammetric detection of single-nucleotide mismatches recognized by the protein MutS. *Anal. Bioanal. Chem.* **388,** 259–270. https://doi.org/10.1007/s00216-007-1181-7 (2007).

177. Bi, L.-J. J. *et al.* Construction and characterization of different MutS fusion proteins as recognition elements of DNA chip for detection of DNA mutations. *Biosens. Bioelectron.* **21,** 135–144. https://www.sciencedirect.com/science/article/pii/S0956566304003938 (2005).

178. Chen, H. *et al.* Electrochemical scanning of DNA point mutations via MutS protein-mediated mismatch recognition. *Biosens. Bioelectron.* **24,** 1955–1961. https://www.sciencedirect.com/science/article/pii/S095656630800568X (2009).

179. Gong, H. *et al.* Unlabeled hairpin DNA probe for electrochemical detection of single-nucleotide mismatches based on MutS-DNA interactions. *Anal. Chem.* **81,** 8639–8643. https://doi.org/10.1021/ac901371n (2009).

180. Paleček, E. *et al.* Sensitive Electrochemical Determination of Unlabeled MutS Protein and Detection of Point Mutations in DNA. *Anal. Chem.* **76,** 5930–5936. https://doi.org/10.1021/ac049474x (2004).

181. Cho, M., Chung, S., Heo, S.-D. D., Ku, J. & Ban, C. A simple fluorescent method for detecting mismatched DNAs using a MutS-fluorophore conjugate. *Biosens. Bioelectron.* **22,** 1376–1381. https://www.sciencedirect.com/science/article/pii/S0956566306002831 (2007).

182. Khilko, Y. *et al.* DNA assembly with error correction on a droplet digital microfluidics platform. *BMC Biotechnol.* **18,** 37. https://doi.org/10.1186/s12896-018-0439-9 (2018).

183. Smith, J. & Modrich, P. Mutation detection with MutH, MutL, and MutS mismatch repair proteins. *Proc. Natl. Acad. Sci.* **93,** 4374 LP –4379. http://www.pnas.org/content/93/9/4374.abstract (1996).

184. Smith, J. & Modrich, P. Removal of polymerase-produced mutant sequences from PCR products. *Proc. Natl. Acad. Sci.* **94,** 6847 LP –6850. http://www.pnas.org/content/94/13/6847.abstract (1997).

185. Beaulieu, M., Larson, G. P., Geller, L., Flanagan, S. D. & Krontiris, T. G. PCR candidate region mismatch scanning: adaptation to quantitative, high-throughput genotyping. *Nucleic Acids Res.* **29,** 1114–1124. https://doi.org/10.1093/nar/29.5.1114 (2001).

186. Zipper, H., Brunner, H., Bernhagen, J. & Vitzthum, F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res.* **32,** e103–e103. https://doi.org/10.1093/nar/gnh101 (2004).

187. Wu, X., He, S. & Zhao, J. X. Label-free fluorescence assay coupled exonuclease reaction and SYBR Green I for the detection of T4 polynucleotide kinase activity. *Anal. Methods* **12,** 807–812. http://dx.doi.org/10.1039/C9AY02283J (2020).

188. Dillingham, M. S. *et al.* Fluorescent Single-Stranded DNA Binding Protein as a Probe for Sensitive, Real-Time Assays of Helicase Activity. *Biophys. J.* **95,** 3330–3339. https://www.sciencedirect.com/science/article/pii/S0006349508784753 (2008).

189. Chisty, L. T., Quaglia, D. & Webb, M. R. Fluorescent single-stranded DNA-binding protein from Plasmodium falciparum as a biosensor for single-stranded DNA. *PLoS One* **13,** e0193272. https://doi.org/10.1371/journal.pone.0193272 (2018).

190. Schanz, S., Castor, D., Fischer, F. & Jiricny, J. 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.* **106,** 5593 LP –5598. http://www.pnas.org/content/106/14/5593.abstract (2009).

191. Gough, G. W., Sullivan, K. M. & Lilley, D. M. The structure of cruciforms in supercoiled DNA: probing the single-stranded character of nucleotide bases with bisulphite. *EMBO J.* **5,** 191–196. https://doi.org/10.1002/j.1460-2075.1986.tb04195.x (1986).

192. Wilson, T. M. *et al.* MSH2–MSH6 stimulates DNA polymerase η, suggesting a role for A:T mutations in antibody genes. *J. Exp. Med.* **201,** 637–645. https://doi.org/10.1084/jem.20042066 (2005).

193. Zdraveski, Z. Z., Mello, J. A., Marinus, M. G. & Essigmann, J. M. Multiple pathways of recombination define cellular responses to cisplatin. *Chem. Biol.* **7,** 39–50. https://www.sciencedirect.com/science/article/pii/S1074552100000648 (2000).

194. Karran, P. & Marinus, M. G. Mismatch correction at O6-methylguanine residues in E. coli DNA. *Nature* **296,** 868–869. arXiv: arXiv:1011.1669v3. https://doi.org/10.1038/296868a0 (1982).

195. Fram, R. J., Cusick, P. S., Wilson, J. M. & Marinus, M. G. Mismatch repair of cis-diamminedichloroplatinum(II)-induced DNA damage. *Mol. Pharmacol.* **28,** 51 LP –55. http://molpharm.aspetjournals.org/content/28/1/51.abstract (1985).

196. Duckett, D. R. *et al.* Human MutSα recognizes damaged DNA base pairs containing 06-methylguanine, 04-methylthymine, or the cisplatind(GpG) adduct. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6443–6447 (1996).

197. Calmann, M. A. & Marinus, M. G. MutS inhibits RecA-mediated strand exchange with platinated DNA substrates. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 14174–14179. http://www.pnas.org/content/101/39/14174.abstract (2004).

198. Calmann, M. A., Nowosielska, A. & Marinus, M. G. Separation of mutation avoidance and antirecombination functions in an Escherichia coli mutS mutant. *Nucleic Acids Res.* **33,** 1193–1200. https://doi.org/10.1093/nar/gki263 (2005).

199. Fourrier, L., Brooks, P. & Malinge, J.-M. M. Binding Discrimination of MutS to a Set of Lesions and Compound Lesions (Base Damage and Mismatch) Reveals Its Potential Role as a Cisplatin-damaged DNA Sensing Protein. *J. Biol. Chem.* **278,** 21267–21275. https://www.sciencedirect.com/science/article/pii/S0021925820734381 (2003).

200. Ijsselsteijn, R., Jansen, J. G. & de Wind, N. DNA mismatch repair-dependent DNA damage responses and cancer. *DNA Repair (Amst).* **93,** 102923. https://doi.org/10.1016/j.dnarep.2020.102923%20https://www.sciencedirect.com/science/article/pii/S1568786420301725 (2020).

201. Hong, Z. *et al.* Recruitment of mismatch repair proteins to the site of DNA damage in human cells. *J. Cell Sci.* **121,** 3146–3154. https://doi.org/10.1242/jcs.026393 (2008).

202. Tanaka, M. *et al.* Mismatch repair proteins recruited to ultraviolet light-damaged sites lead to degradation of licensing factor Cdt1 in the G1 phase. *Cell Cycle* **16,** 673–684. http://dx.doi.org/10.1080/15384101.2017.1295179%20https://doi.org/10.1080/15384101.2017.1295179 (2017).

203. Jiang, Y., Ke, C., Mieczkowski, P. A. & Marszalek, P. E. Detecting ultraviolet damage in single DNA molecules by atomic force microscopy. *Biophys. J.* **93,** 1758–1767. https://www.sciencedirect.com/science/article/pii/S0006349507714315%20http://dx.doi.org/10.1529/biophysj.107.108209 (2007).

204. Murakami, M., Hirokawa, H. & Hayata, I. Analysis of radiation damage of DNA by atomic force microscopy in comparison with agarose gel electrophoresis studies. *J. Biochem. Biophys. Methods* **44,** 31–40. https://www.sciencedirect.com/science/article/pii/S0165022X0000049X (2000).

205. Lang, W. H. *et al.* Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops. *Proc. Natl. Acad. Sci.* **108,** E837 LP –E844. http://www.pnas.org/content/108/42/E837.abstract%20http://www.pnas.org/cgi/doi/10.1073/pnas.1105461108 (2011).

206. Whinn, K. S. *et al.* Nuclease dead Cas9 is a programmable roadblock for DNA replication. *Sci. Rep.* **9,** 13292. https://doi.org/10.1038/s41598-019-49837-z (2019).

207. Xia, Z. & Liu, Y. Reliable and global measurement of fluorescence resonance energy transfer using fluorescence microscopes. *Biophys. J.* **81,** 2395–2402. https://www.sciencedirect.com/science/article/pii/S0006349501758869%20http://dx.doi.org/10.1016/S0006-3495(01)75886-9 (2001).

Appendix A

Equations

Determination of the concentrations of fluorophores
To determine the fluorophore concentration Lambert-Beer law can be directly applied.

$$C_F = \frac{A^{Max}}{\epsilon_{Fl} * L} \tag{A.1}$$

Determination of the degree of labeling
Lambert-Beer law can also be used to determine the protein concentration but requires a correction if a fluorophore is present in the sample (equation 3). In this case, the absorption of a fluorophore at 280 nm is subtracted from the protein absorption by the term - (Afluorophore · cf280), where cf280 is the correction factor (cf) describing the absorbance efficiency of the fluorophore at 280 nm compared to its maximal absorbance. The correction factors used in this thesis were calculated manually.

$$C_{pr} = \frac{A^{280} - (A_{FL}^{280} * cf_{280})}{\epsilon_{Pr} * L} \tag{A.2}$$

Certain proteins used in this thesis were labeled with a fluorescent dye. The procedure ended with a majority of proteins in a sample being labeled. To describe the ratio of unlabeled to labeled proteins the term degree of labeling (DOL) is used. To determine the DOL, the concentration of proteins was compared to the concentration of fluorophores. The degree of labeling (DOL) was calculated from UV-VIS absorbance spectra via the following formula:

$$DOL = \frac{C_{Fl}}{C_{Pr}} * 100\% \tag{A.3}$$

or

$$DOL = \frac{A^{Max} * \epsilon_{Pr}}{(A^{280} - (A_{FL}^{280} * cf_{280})) * \epsilon_{Fl}} * 100\% \tag{A.4}$$

where A_{FL}^{Max} is the absorption of a fluorophore at a maximum, A^{280} denotes absorption of a fluorophore at 280 nm, ε_{Pr} means the protein molar extinction coefficient,

CF is a correction factor of the fluorophore, and ε_{Fl} denotes the fluorophore molar extinction coefficient. The DOL was also determined for DNA-carrying fluorescent dyes. In this case, A_{280} is replaced by A_{260}, for the fluorophore cf_{260} was used instead of cf_{280}, and the extinction coefficient for the DNA was used.

$$C_{DNA} = \frac{A_F l^{260} - (A_{FL}^{Max} * cf_{260} * K)}{bp * m_{bp}} \tag{A.5}$$

$$DOL = \frac{C_{Fl}}{C_{DNA}} * 100\% \tag{A.6}$$

or

$$DOL = \frac{A_{FL}^{Max} * bp * m_{bp}}{A_F l^{260} - (A_{FL}^{Max} * cf_{260} * K * \epsilon_{Fl})} * 100\% \tag{A.7}$$

Where m_{bp}=324.5 and K=33 for single-strand DNA

Calculation of FRET signals

FRET signal was corrected for spectral bleed-through and spectral crosstalk to obtain the correct signal intensities. Initially, the spectral bleed-through coefficients S_1-S_4 were obtained from experiments with only a donor or an acceptor by applying the next formulas.

Donor into FRET channel:

$$S_1 = \frac{F_d}{D_d} \tag{A.8}$$

Acceptor into FRET channel:

$$S_2 = \frac{F_a}{A_a} \tag{A.9}$$

Donor into acceptor channel:

$$S_3 = \frac{A_d}{D_d} \tag{A.10}$$

Acceptor into donor channel:

$$S_4 = \frac{D_a}{A_a} \tag{A.11}$$

The signal values of the donor and acceptor channels containing the mixture in which both components are present have been corrected for spectral bleed-through using the next equation

$$D_{da}^c = \frac{D_{da} - S_4 * A_{da}}{1 - S_3 * S_4} \tag{A.12}$$

and

$$A_{da}^c = \frac{A_{da} - S_3 * D_{da}}{1 - S_3 * S_4} \tag{A.13}$$

In the case of a pair of SG with AF647 fluorophores, the donor does not give any signal in the acceptor channel and vice versa, meaning that the bleed-through factors 3 and 4 will be close to zero and the corrected donor and acceptor signals will not be changed much. To obtain the actual FRET signal the intensity of the corrected donor and acceptor signals were multiplied by the corresponding spectral bleed-through factors and subtracted from the FRET channel signal.

$$FRET^C = F_{da} - D_{da}^C * S_1 - A_{da}^C * S_2 \tag{A.14}$$

To normalize the FRET signal a formula suggested by Xia et al. was used[207], which normalizes FRET signal to the square root of the product of donor and acceptor signal.

$$N_{\mathrm{FRET}} = \frac{FRET^C}{G * \sqrt{D_{da}^C * A_{da}^C}} \tag{A.15}$$

Where

$$G = \frac{QY_a}{QY_d}\frac{\Phi_a}{\Phi_d}\frac{T_F}{T_D} \tag{A.16}$$

Appendix B

Supplementary Data

Donor max (bandwidth) nm		FRET max (bandwidth) nm		Acceptor, max (bandwidth) nm		S1	S2	S3	S4
Ex	Em	Ex	Em	Ex	Em				
450 (20)	535 (25)	485 (20)	670 (25)	620 (10)	670(25)	$3.39\pm0.07 * 10^{-1}$	$6.45\pm0.11 * 10^{-2}$	≈ 0	≈ 0
405 (10)	485 (20)	450 (20)	670 (25)	620 (10)	670(25)	$4.40\pm0.10 * 10^{-2}$	$3.14\pm0.10 * 10^{-2}$	≈ 0	≈ 0

Table B.1: Filters and bleed-through coefficients for the correct signal intensities.

Crosstalks for heated DNA at different temperatures differ from other DNA and depend on the amount of ssDNA. The emission spectra of DNA stained by SG shift in the presence of the ssDNA.

Cloning of His-tagged single-cysteine MutS-A336

Single-cysteine MutS (A336C) was generated by site-directed mutagenesis using HiFi-DNA assembly reaction (NEB). As a template for PCR the plasmid carrying the gene for the cysteine-free dimer (cfMutS^{D835R}) was used and the following primer pairs:

For the linearized vector (5751 bp)

MutS 3UTR Sense (25-mer): gctagcgctatatgcgttgatgcaa
MutS 5UTR antisense (25-mer): tctagaggggaattgttatccgctc
For fragment 1 (1872 bp)
MutS A336C sense (27-mer): gatttcaccTGCgggctacagccggta
MutS-3UTR (25-mer): ttgcatcaacgcatatagcgctagc
For fragment 2 (1134 bp)
MutS-5UTR (25-mer): gcggataacaattcccctctagaaa
MutS A336C antisense (30-mer): tgtagcccGCAggtgaaatcctgcaatgcg
Gene and promotor regions were verified by Sanger DNA sequencing.

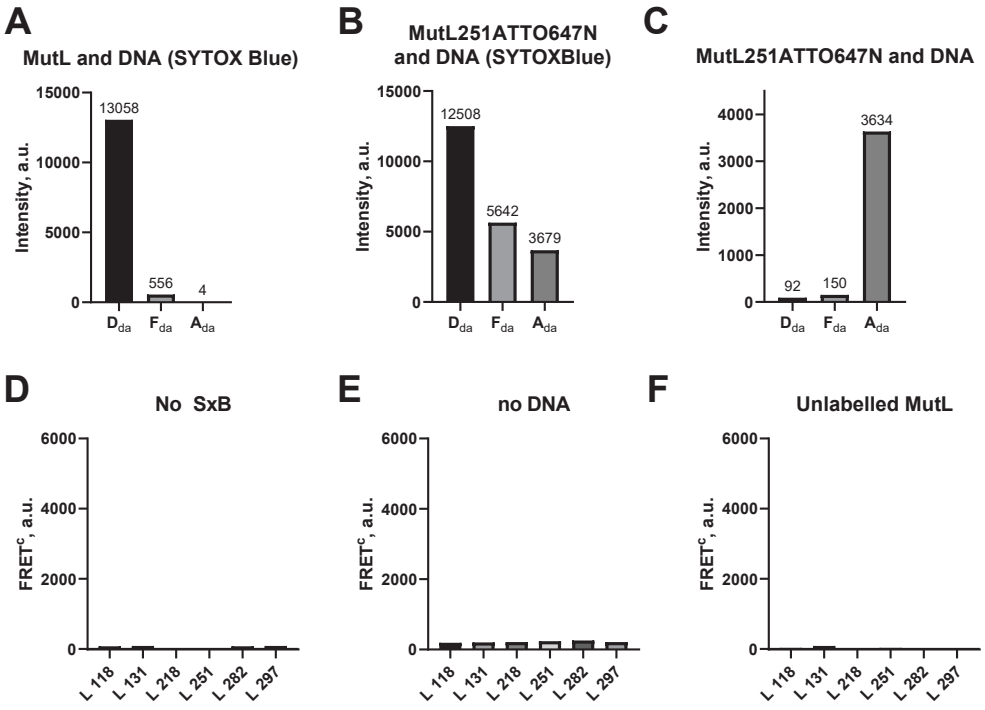


Figure B.1: **(A)** Raw data in the absence of acceptor as ATTO647N, **(C)** donor as Sytox Blue. **(B)** Raw data for ATTO647N-MutL^{254C} in the binding assay. **(D)** Controls with unstained DNA, **(E)** free dye, and **(F)** unlabeled MutL. Exemplary primary data before correction.

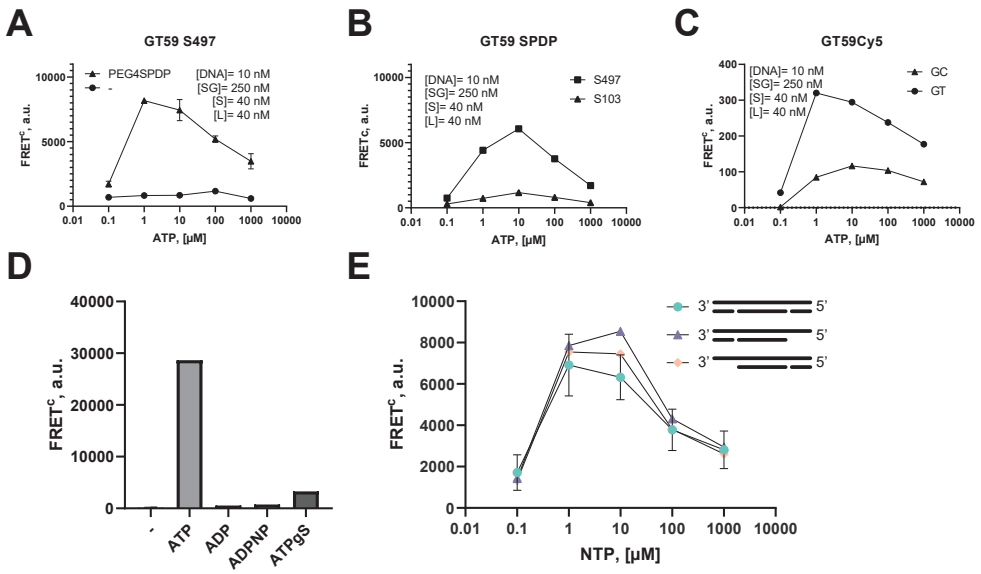


Figure B.2: **(A)** In the absence of a cross-linker there is no recruitment of MutL at any ATP concentrations. **(B)** Recruitment of MutL by MutS cross-linked with short cross-linker SPDP to the DNA at mismatch domain is reduced in comparison with clamp domain cross-linking. **(C)** MutL recruitment requires covalent trapping to the DNA and a mismatch and also can be monitored by FRET between labeled MutL and covalently attached dye to the oligonucleotide. **(D)** MutL recruitment on circle DNA (no ends) can be induced only by ATP because of multi loading of MutL. **(E)** The level of MutL recruitment did not decrease with a decrease in the length of the oligonucleotide, both from the 3 'end and from the 5' end by 29 base pairs.

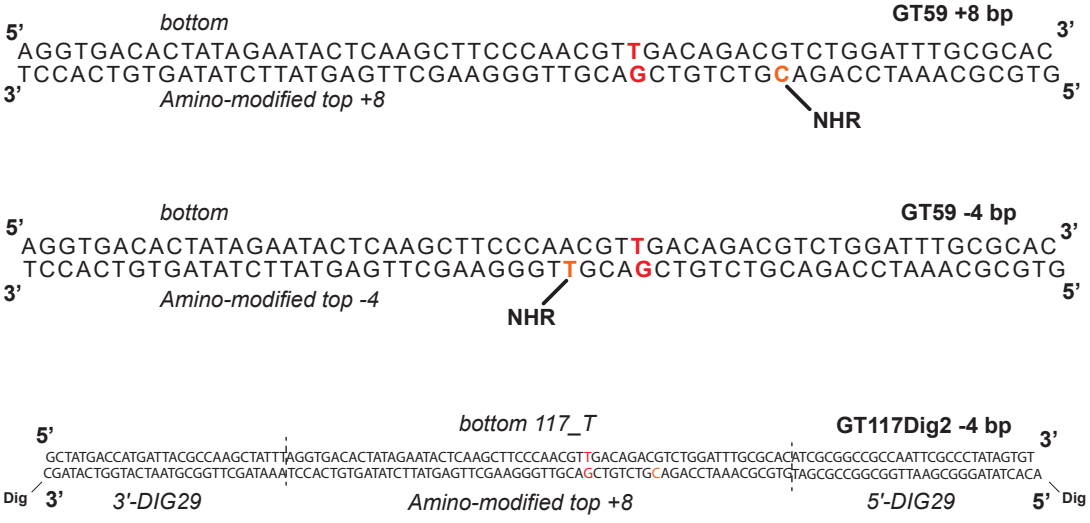


Figure B.3: Double stranded modified oligonucleotides for cross-linking. **(A)** 59 bp DNA with the modification at cytidine +8 bp from MM. **(B)** 59 bp DNA with the modification at thymidine -4 bp from MM. **(C)** 117 bp DNA with the modification at cytidine +8 bp from MM and ends modified by Digoxegenin for effective blocking with Fab fragments.

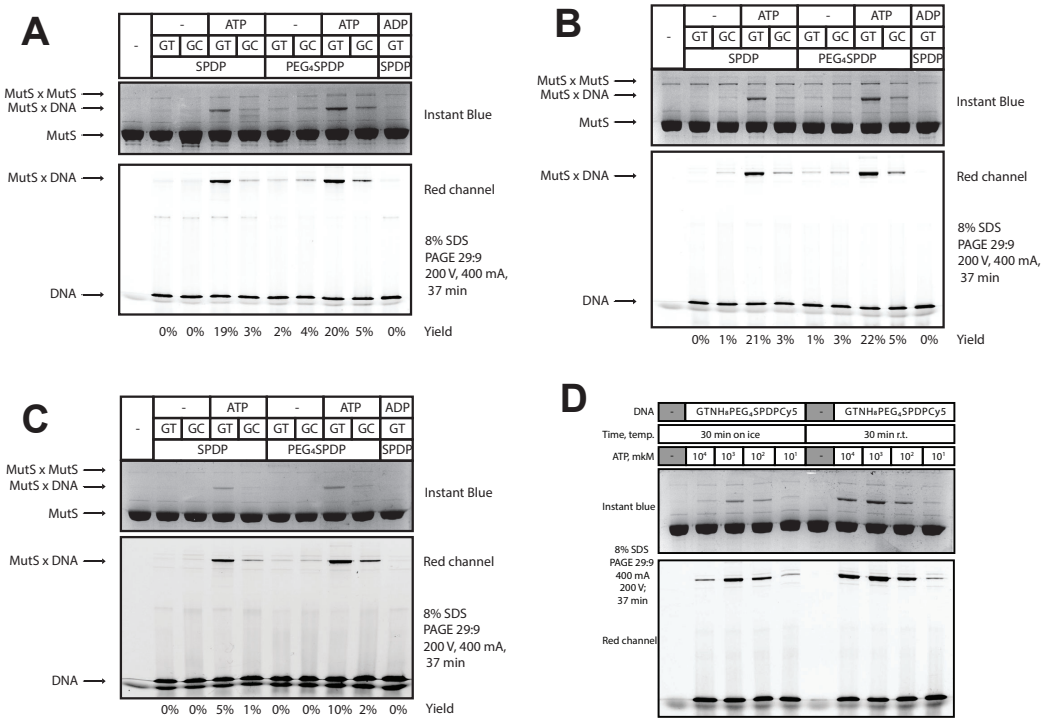


Figure B.4: Cross-linking MutS$^{220C/D835R}$ to the DNA on the gels. The upper part shows a gel fragment stained with Coomasie Blue to visualize proteins. The lower part shows the same gel in green fluorescence to visualize an oligonucleotide with the covalently attached dye. **(A)** The experiment was designed for cross-linking subunit A with oligonucleotides with a modification at position -4 bp from the mismatch on the top chain. **(B)** An identical experiment, but already with an oligonucleotide modified at the +8 bp position in the upstream chain. The experiment was designed for cross-linking subunit B. Cross-linking was carried out on the ice overnight. **(C)** Experiment identical to experiment **"B"**, but conducted for 30 min on ice. **(D)** Effect of ATP on the yield of the cross-linking reaction for 30 minutes at r.t. (lanes 2-5) or on ice overnight (lane 7-10).

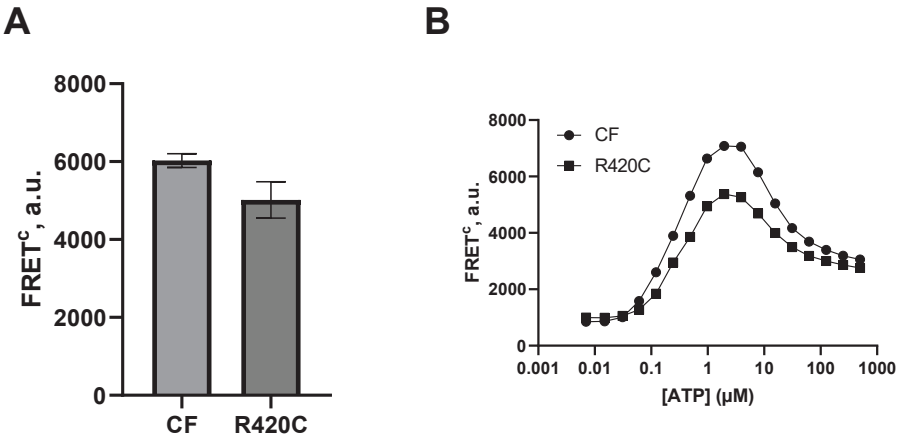


Figure B.5: **(A)** Comparison of MutL recruitment activity of cfMutS^{D835R} and MutS$^{R420C/D835R}$. **(B)** ATP influence on FRET between labeled MutL and stained DNA in MutL recruitment assay.

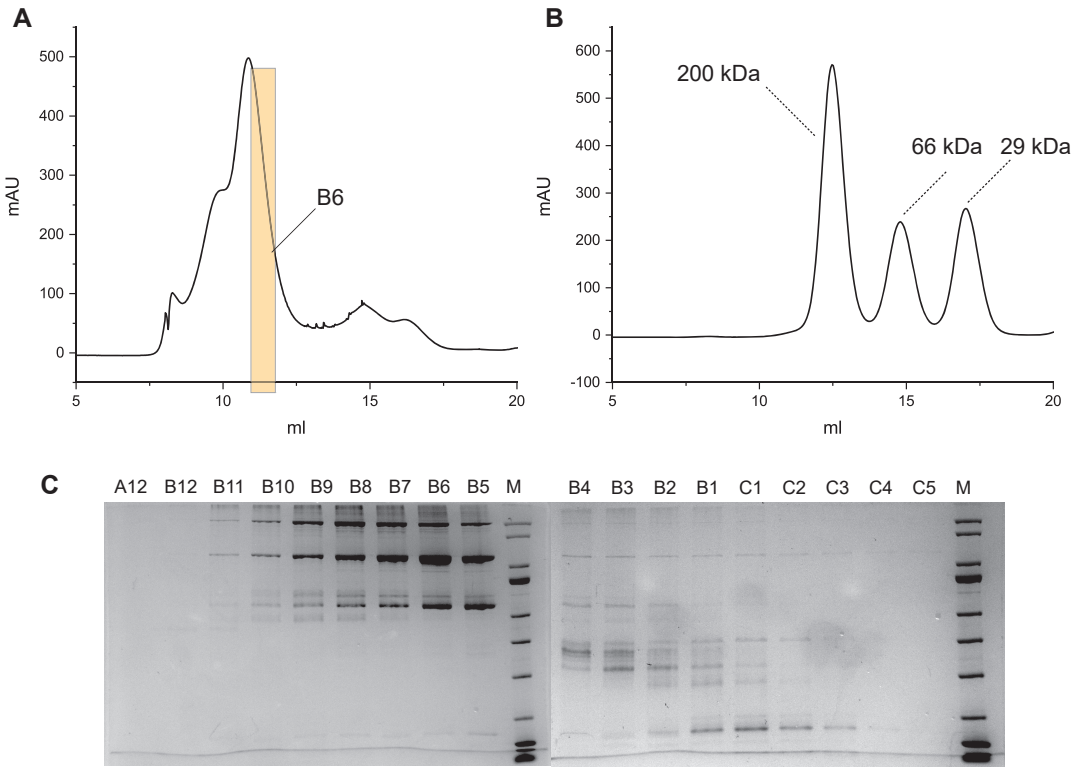


Figure B.6: **(A)** Purification of the complex via size-exclusion chromatography. **(B)** Chromatogram of markers with a defined mass. **(C)** Collected fraction after HPLC purification on the 8% SDS gels, HPLC Markers. For all subsequent experiments with the complex, the B6 fraction was selected, aliquoted, and frozen in liquid nitrogen.

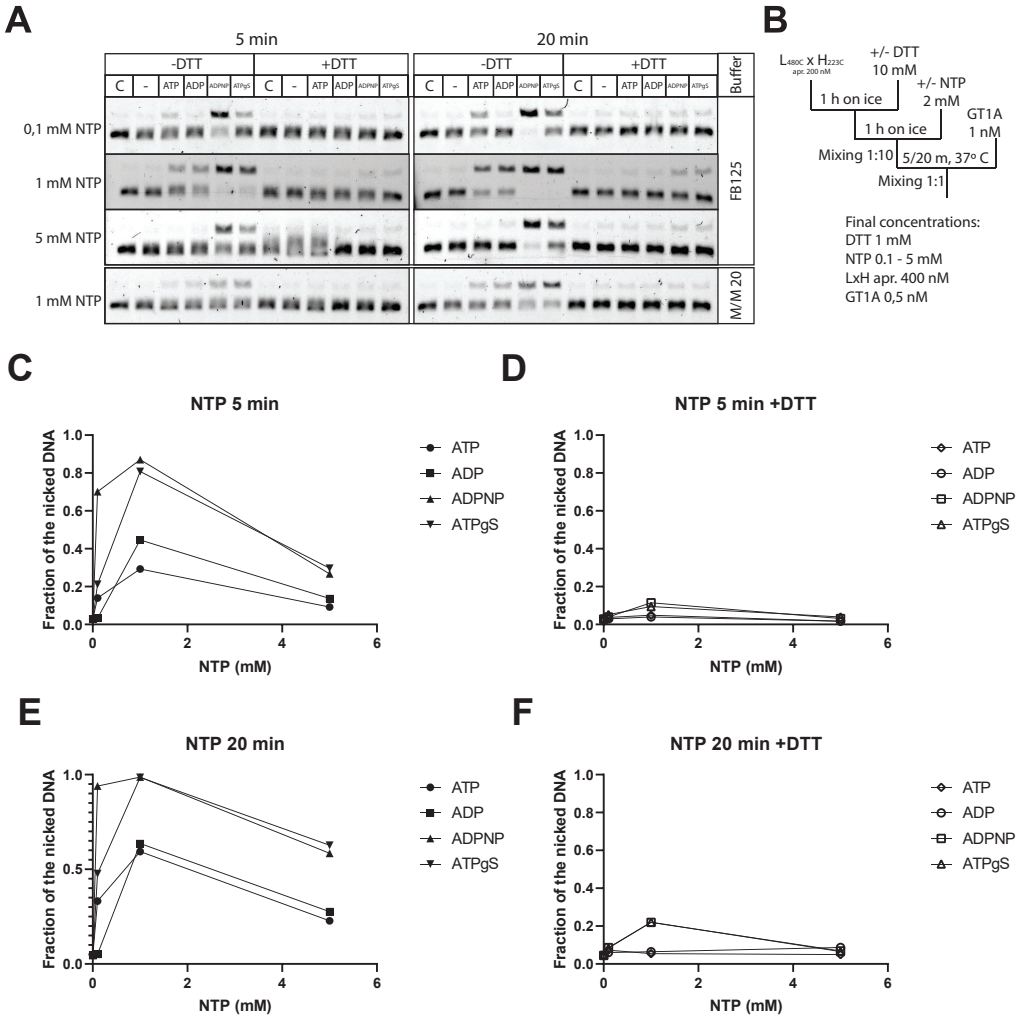


Figure B.7: Nicking assay of cross-linked MutL-MutH complex with different nucleotides and concertation after 5 or 20 minutes. **(A)** Both cross-linked and native $MutL^{480}xMutH^{223}$ complex activity was demonstrated on 0.8% agarose gels. The left side shows the results at five minutes, the right side at 20 minutes respectively. When the complex was treated with a reducing agent such as DTT, the complex largely lost its activity at both low and high nucleotide concentrations. Diluting the interaction time of the complex with the substrate from 5 to 20 min slightly increased the activity of the complex (only in the case of ADPNP) (lanes 7-12). The cross-linked complex was already active at 0.1 mM ADPNP at five minutes. The amount of nicked DNA increased with increasing incubation time. When the nucleotide concentration was increased to 1 mM (lanes 1-6), the activity of the complex was observed not only with ADPNP, but also with ATP, ADP, and ATPγS. When the nucleotide concentration was increased to 5 mM, the activity of the complex decreased, which may be due to an insufficient amount of Mg^{2+} in the buffer (5 mM). Conducting the reaction in M/M buffer also reduced the activity of the complex. **(B)** When it was necessary to destroy the covalent bond with the reducing agent, the complex interacted with DTT for one hour on ice. The complex was then preincubated with nucleotides for one hour on ice. The resulting complex was then mixed 1:1 by volume with DNA at 37°C. Nucleotide concentration dependence plots on the activity of the cross-linked complex at 5 **(C)** and **(E)** 20 minutes as well as the reduced complex **(D)** and **(F)**.

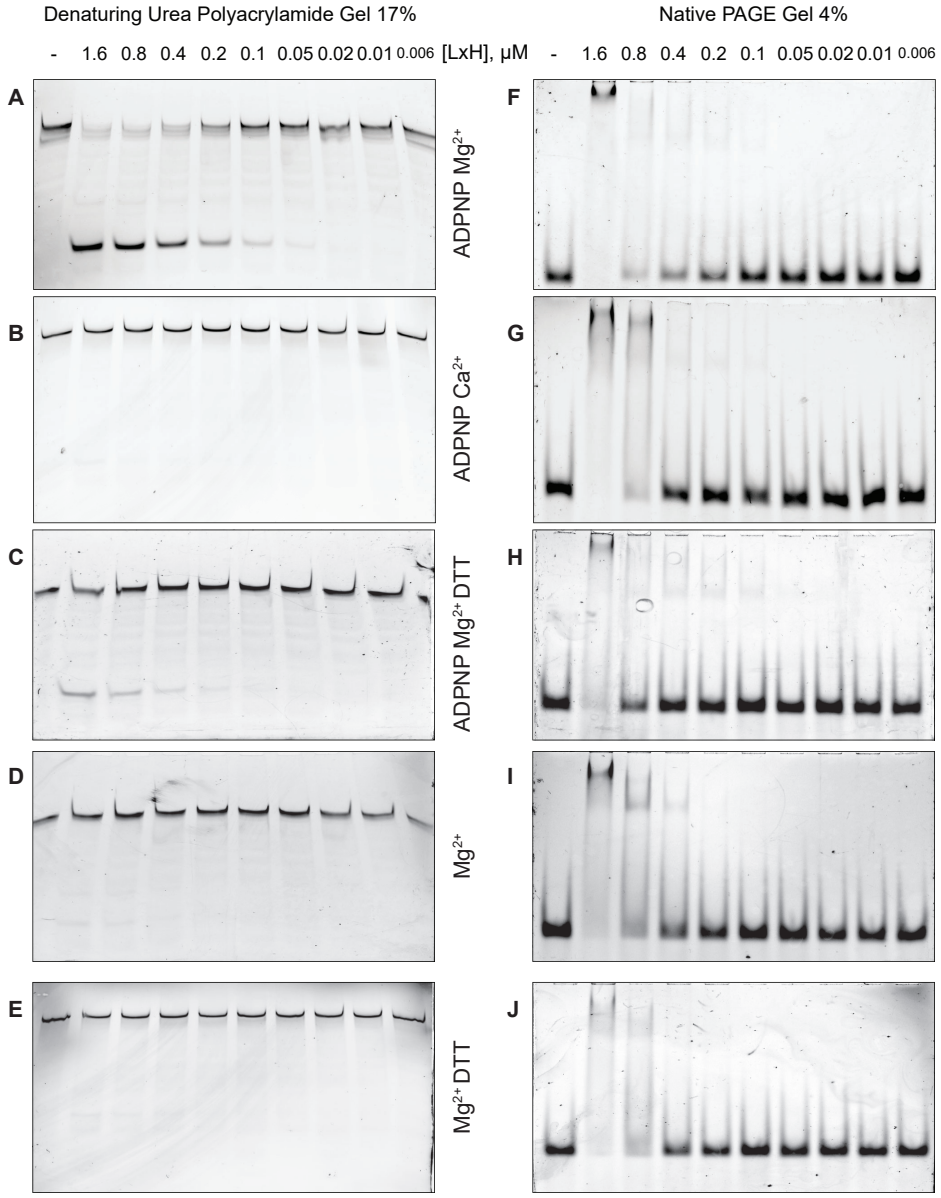


Figure B.8: Effect of complex concentration on nicking and shifting activity. All experiments were performed with the cross-linked MutL^{480C}xMutH^{223C} complex with changes in buffer or addition of various chemical reagents indicated next to the figure. **(A)** Activation of the cross-linked complex occurs upon the addition of 1 mM ADPNP in a buffer containing Mg^{2+} ions. As the concentration of the complex decreases, the activity decreases accordingly. **(B)** Using Ca^{2+} ions instead of Mg^{2+} in the buffer leads to a loss of activity of the complex, which was previously shown for single MutH^{161}. **(C)** Reduction of the complex with DTT, even in the presence of ADPNP, leads to a significant decrease in the activity of the complex. **(D)** In the case when the complex was not preincubated with ADPNP no activity was observed at any concentration of the complex. **(E)** The reduced inactivated complex also showed no activity. **(F-J)** No significant difference in the concentrations of the complex was observed in the experiments causing shifts under different conditions. The shifts occurred approximately in the range from 0.8 - 1.6 μM.

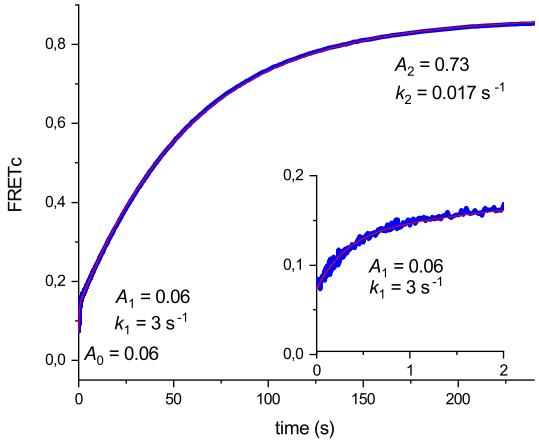


Figure B.9: Stopped-Flow kinetic of ATP-induced loading of multiple MutS sliding clamps. AF647-MutS^{336} (100 nM) and 0.36 nM G/T mismatch DNA in FB150T (25 mM HEPES pH 7.5, 5 mM 500 nM SG was mixed with an equal volume of ATP (2 mM) in FB150T. A double exponential function was used to fit the data. The initial values of FRET$_c$ is due to the single mismatch bound MutS. The burst amplitude A$_1$ of 0.06 almost doubles the FRET$_c$ in agreement with the ATP-induced change from a low FRET to a high FRET state (see Figure 6.1B). The amplitude A$_2$ of the slow phase of 0.73 is best explained with the loading of additional MutS sliding clamps in the high FRET state.
Fast kinetics were measured using a stopped-flow device SF-61SX2 (TgK Scientific, Bradford-on-Avon, UK) with an LED light source LSM-470A (Ocean Optics BV) with an in-line filter holder and a long pass filter 475 nm and a short pass filter 500 nm (Edmund Optic® Europe), filter ET525/50 M (Chroma Technology, Olching, Germany) for the donor fluorescence signal and a ET670/50M (Chroma Technology, Olching, Germany) for FRET signal at 25 °C.

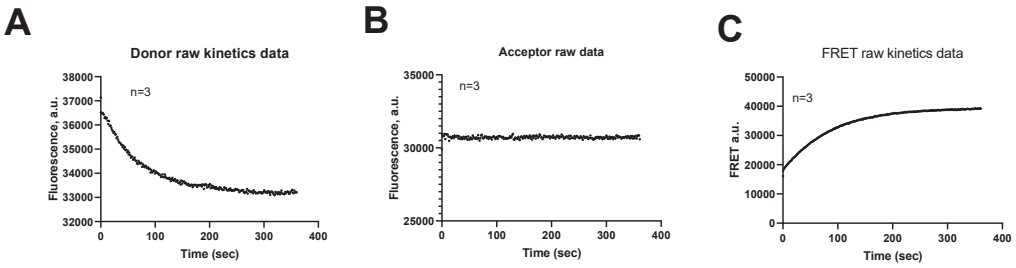


Figure B.10: Raw data of kinetics for donor **(A)** Acceptor **(B)** and FRET **(C)** channels. When observing the interaction between AF647-MutS^{336}, not only the FRET signal but also the donor signal intensity can be monitored. As time increases, the number of proteins that are on the DNA increases, reaching equilibrium after five minutes. **(A)** As the number of complexes increases, a decrease in donor signal (from 37000 to 33000 a.u.) is observed, as energy is transferred from the donor to the acceptor in the form of FRET. **(B)** When only the acceptor is excited there is no change in the signal in the acceptor channel. **(C)** Significant signal enhancement is observed even before the results are processed in the raw data (from 18000 to 40000 a.u.).

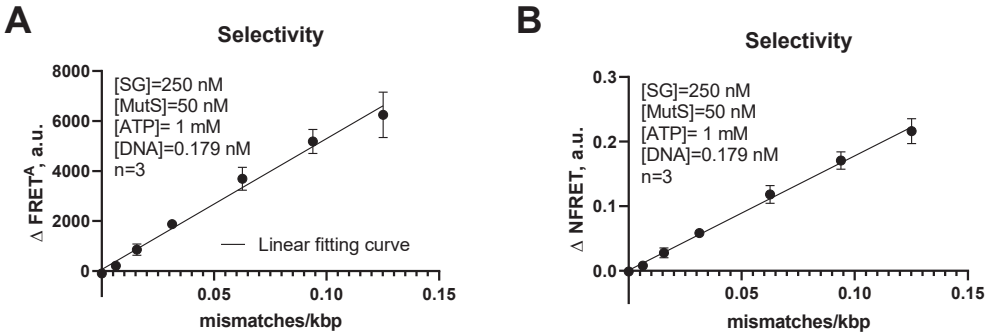


Figure B.11: Comparison of the FRET^c **(A)** and FRETN **(B)**. Normalization of the FRET signal decreases error bars. An additional mathematical operation reduces pipetting error because, in the case of a normalized signal, the signal intensities of both the donor and the acceptor are taken into account (Equation A.16)

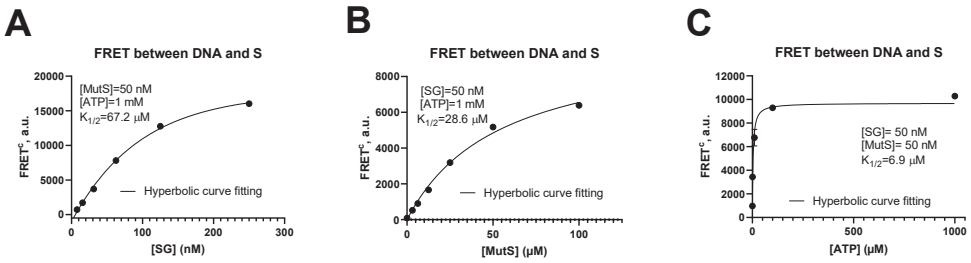


Figure B.12: Concentration dependence **(A)** on SG, **(B)** AF647-MutS^{336}, or **(C)** ATP on the observed FRET signal intensity on circular mismatched DNA.

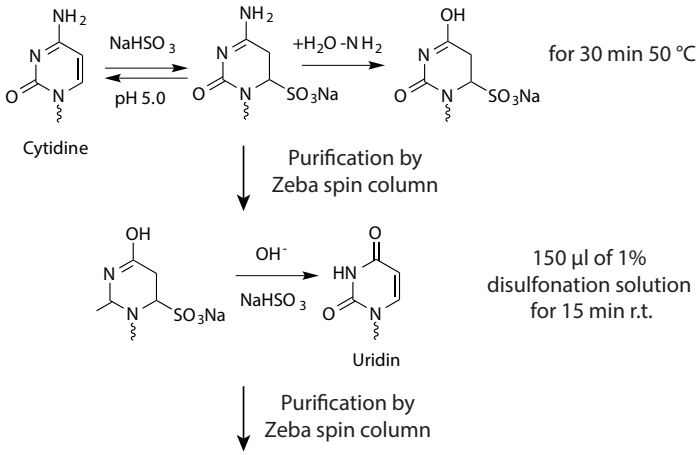


Figure B.13: Deamination procedure is divided into several steps: deamination under acidic conditions, purification from the sulphonating agent, desulphonation process under basic conditions, and purification of salts and basic conditions.

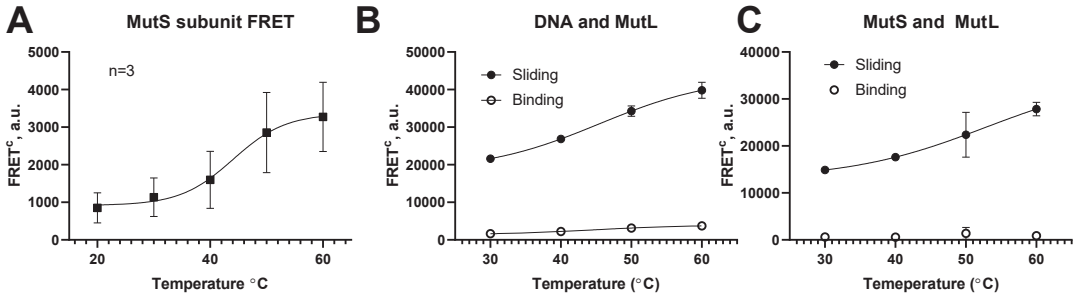


Figure B.14: **(A)** The exchange of AF488-MutS^{336} and AF647-MutS^{336} subunits results in a dimer with two different subunits (25% AF488-MutS^{336}; 25% AF647-MutS^{336}; and 50% AF647-MutS^{336}+AF488-MutS^{336}). When such a mixture is activated with ATP, a FRET signal occurs between the monomers of such a dimer. The signal depends on the temperature of the plasmid deamination and completely repeats the shape of the signal dependence as with AF647-MutS^{336} multi-loading on DNA. **(B)** MutS multiloading also leads to MutL multi recruitment on DNA. When the complex is activated with ATP, an increase in signal between the stained DNA and AF647-MutL^{297} is observed with an increase in deamination temperature, which is associated with an increase in deamination sites that are recognized as a mismatch for MutS. **(C)** In addition to the signal between DNA and proteins, it is possible to monitor protein interactions such as AF488-MutS^{336} and AF647-MutL^{297}. As in the two previous cases, the signal intensity increases with increasing deamination temperature. This indicates an increase in MutS-MutL complexes on the plasmid.

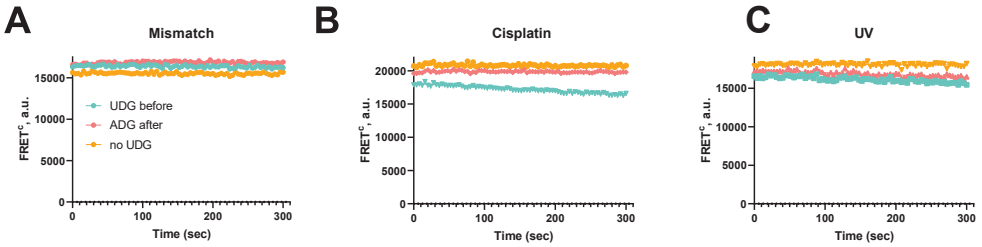


Figure B.15: UDG treatment of mismatch substrate **(A)**, cisplatin-treated DNA **(B)**, and UV damages DNA **(C)** in AF647-MutS^{336} FRET-based assay. UDG does not affect the activity of the AF647-MutS^{336}, in contrast to the deaminated substrate (see Figure 6.4).

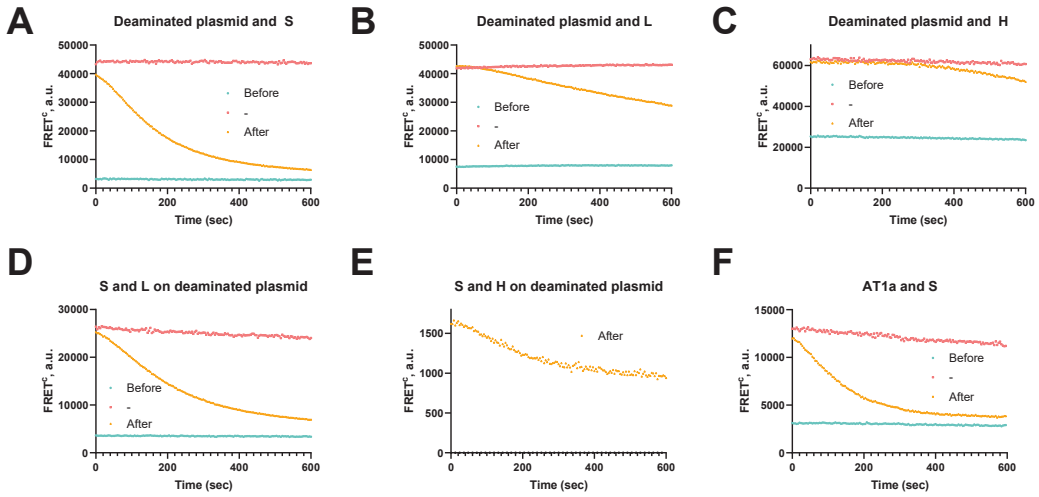


Figure B.16: Kinetics of MM proteins upon UDG treatment. **(A)** In the case of AF647-MutS^{336} we observe the fastest signal reduction after UDG processing (orange dots). UDG pretreatment of the deaminated plasmid results in a significant signal reduction of up to 5% (green dots). **(B)** In the case of MutS and MutL, when the observed object is AF647-MutL^{297} the dissociation rate is much lower. The signal in the case of UDG pre-treatment of the deaminated plasmid is higher than for MutS and is around 10%. **(C)** The lowest dissociation rate was observed for AF647-MutH^{156}. The highest signal is observed in the case of enzyme pre-treatment of the deaminated plasmid, which is approximately 20% of the maximum. In all three cases, the kinetics are significantly different, which may indicate different lifetimes of the complexes. In the case where MutS was the observed sensor, the addition of MutL and/or MutH on such circle substrates did not result in a change in kinetics. Also, MutH did not affect the dissociation kinetics of MutL (data not shown). The kinetics of enzyme action can also be observed by measuring the FRET signal intensity between covalently attached to **(D)** AF488-MutS^{336} and AF647-MutL^{297}, or **(E)** AF488-MutS336 and AF647-MutH^{223} as well as AF488-MutL^{297} and AF647-MutH^{223} (data not shown). **(F)** In addition to deaminated plasmids, circle substrates designed to monitor the nicking activity of MutH and the unwinding activity of UvrD can also be used[120]. The advantage of such substrates is the precisely defined mismatch position introduced by a complementary oligonucleotide. A disadvantage of this substrate is the presence of a signal for homoduplex DNA as a negative control. In the case of UDG treatment, a decrease of the signal between AF647-MutS^{336C} and DNA is observed, while pretreatment with the UDG leads to a relatively high signal compared to the deaminated plasmid. The absolute signals after UDG treatment are approximately the same. In the case of a positive signal (no UDG treatment), the intensity for the circle substrate is significantly lower compared to the deaminated substrate at 50°C.

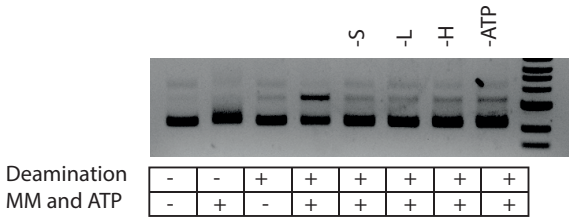


Figure B.17: Incomplete nicking of deaminated partially methylated plasmid in MMR dependent manner. MutS in sliding clamp state can recruit second protein MutL and which in turn activates latent endonuclease activity of MutH. This activity can be detected only when the substrate is deaminated and a full MMR system is present including 1 mM of ATP.

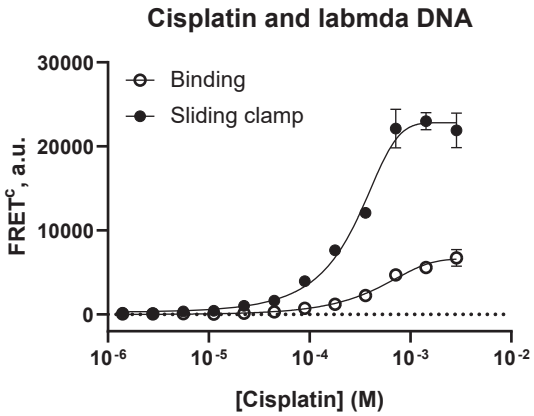


Figure B.18: Cisplatin treatment of the lambda DNA. The FRET signal intensity between AF647-MutS^{336} and DNA increases not only on cis-platinum-treated plasmids but also on the linear lambda DNA.

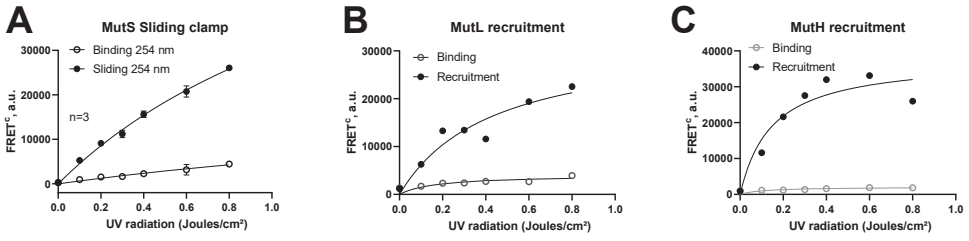


Figure B.19: Recruitment of MutS **(A)**, MutL **(B)**, and MutH **(C)** on UV-treated substrates. The multi-loading of MutS also leads to the multi-loading of MutL and MutH. As the amount of energy increases (the amount of damage also increases), the signal between stained DNA and AF647-MutL^{297}, or AF647-MutH^{223}, also increases.

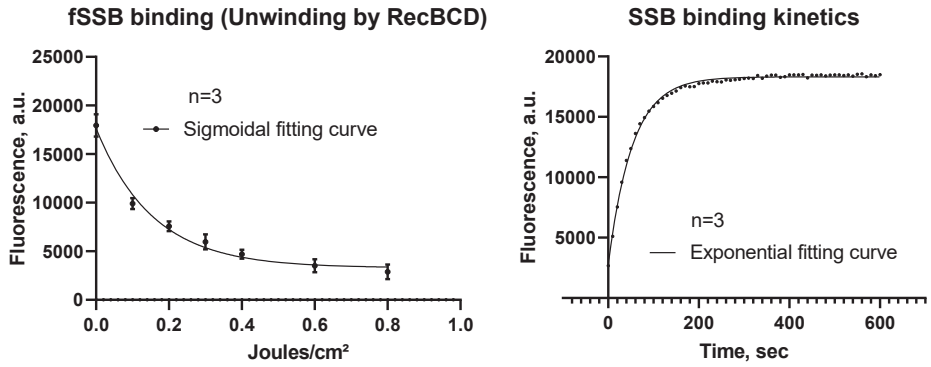


Figure B.20: **(A)** Influence of UV radiation on the ability to unwind λ-DNA by RecBCD. **(B)** The unwinding of λ-DNA by RecBCD monitored by DCC-SSB binding to UV damaged DNA. The DCC-SSB binding assay was used to monitor the unwinding of UV-treated DNA by the RecBCD helicase. Increasing the energy of UV irradiation reduced the extent of DNA unwinding by RecBCD as monitored by binding of DCC-SSB. It demonstrates that UV damages can not be removed by helicase and we won't be able to monitor UvrD helicase activity at such substrate.

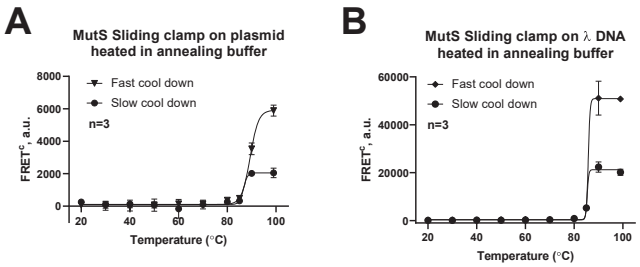


Figure B.21: **(A)** Slow and fast cooling down of heated plasmid in annealing buffer. **(B)** Slow and fast cooling down of heated λ-DNA in annealing buffer. The cooling down after treatment also influences the quality of the DNA. Slow cooling in the hot block resulted in less signal than fast cooling on ice, due to better annealing of single-stranded chains.

Nomenclature

$(6-4)PP$ 6–4 photoproducts

$(v/v)\%$ percentage in volume to volume

$(w/v)\%$ percentage in weight to volume

A fluorescence acceptor

ABC ATP-binding cassette

ADP adenosine diphosphate

$ADPNP$ Adenylyl-imidodiphosphate

AFM atomic force microscopy

ATP adenosine triphosphate

BER base-excision repair

$BMOE$ bis-maleimidoethane

bp base pair

BSA bovine serum albumin

cc closed circle

cf correction factor

$cis-Pt$ cisplatin

CPD cyclobutane pyrimidine dimer

$Cryo-EM$ cryogenic electron microscopy

CTD C-terminal domain

D fluorescence donor

Da Dalton

Dig digoxigenin

$DMSO$ dimethyl sulfoxide

DNA deoxyribonucleic acid

dNTP deoxyribonucleotide triphosphate

DOL degree of labeling

ds double-strand

DTT dithiothreitol

E.coli Escherichia coli

EDTA ethylene diamine tetra acetic acid

EJ end joining

EtBr ethidium bromide

FRET Förster Resonance Energy Transfer

HEPES 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

HPLC high performance liquid chromatography

HR homologous recombination

i.e. it est (latin), such as

IDL insertion-deletion loop

IPTG Isopropyl β-D-1-thiogalactopyranoside

kb kilo base pair

min minute

MM mismatch

MMC mitomycin C

MMR DNA mismatch repair

MTS methane thiosulfonat

NER nucleotide-excision repair

NHS N-hydroxysuccinimide

NMR nuclear magnetic resonance

NTD N-terminal domain

NTP ribonucleotide triphosphate

oc open circle

OD optical density

PAGE polyacrylamide gel electrophoresis

PCR polymerase chain reaction

$PEG_4 - SPDP$ N-[(succinimidyl)oxy]-15-oxo-3,6,9,12-tetraoxapentadecyl-3-(2-pyridyldithio)propanamide

$PMSF$ phenylmethylsulfonyl fluoride

rcf relative centrifugal force

rpm rounds per minute

rt room temperature

sc single-cysteine

SDS sodium dodecyl sulfate

SG Sybr Green I

$smFRET$ single-molecule FRET

$SPDP$ N-succinimidyl 3-(2-pyridyldithio)-propionate

ss single-strand

SSB single-stranded binding protein

SxB Sytox Blue

Taq Thermus aquaticus

$TEMED$ N,N,N‘,N‘-tetramethylethylendiamine

$Tris$ tris(hydroxymethyl)aminomethane

UV ultraviolet

Appendix C

Software

The analysis of the images was performed using the ImageJ software. Mathematical calculations were carried out in Microsoft Excel. Statistical analysis and graph construction were performed in GraphPad Prism software. All figures of protein and nucleic acid structures were generated using PyMOL Molecular Graphics System. All other illustrations were created using Adobe Illustrator. 3D animations were produced in the Blender software.