PROPOSITIONS

1. MutS binds to all mismatches using a common recognition mode. The kinking of the DNA is an important requirement for mismatch recognition.
-This thesis

2. Magnensium is an important cofactor which is required for ADP-ATP exchange in MutS. Addition or removal of magnesium causes substantial conformational changes to the structure of the protein.
-This thesis

3. During the ATPase cycle of MutS, there are stages when the protein may have to release the magnesium. The protein, therefore, is able to regulate magnesium binding by itself.
-This thesis

4. A single hydrogen bond between Glu 38 and the mismatch is important for a signalling process that occurs over a distance of ~ 80Å, to be completed successfully.
-This thesis

5. Small and seemingly unimportant molecules like ATP and GTP make their presence felt by bringing about huge conformational changes in massive proteins, and regulate entire biochemical pathways as a result: Small is beautiful and powerful. The ice age killed the dinosaurs, not the cockroaches.

6. The model in which MutS hydrolyses ATP to slide on DNA has been comprehensively disproved, but the people who proposed it should not retract their papers as it is important for even wrong models to exist in literature. It makes paper writing that much more fun.

7. People who stop to enjoy the view on their way up the career ladder are more likely to know the implications of their position, than the ones who focus their attentions solely on the ladder and on how to climb it.

8. A smart PhD student is one who can contribute more to a research supervisor's lab than what he/she gains from it.
A smart research supervisor is one who can find such a student.

9. Most of the world's miseries can be attributed to just one thing: human arrogance.

10. One reason for scientific salaries being so low is the claim by many scientists that they are passionate about their research. Getting highly paid for being passionate about something just has a very bad ring to it.

11. I am much more in agreement with the toothpaste industry's view on scientists, who it portrays as white coated professionals who talk about dental plaque, than the entertainment industry's, which brings out Nutty Professors, Absent Minded Professors and Dr.Dolittles with regularity.

Propositions towards the partial fulfilment of the thesis titled 'MutS; Recognition of DNA mismatches and initiation of repair', by Ganesh Natrajan, Erasmus University, Rotterdam, 2006.

MutS: Recognition of DNA mismatches and initiation of repair

MutS: De herkenning van DNA mismatches en het opstarten van herstel

Thesis

to obtain the degree of Doctor from the Erasmus University Rotterdam by command of the Rector Magnificus

Prof.dr. S. W. J. Lamberts

and in accordance with the decision of the Doctorate Board
The public defence shall be held on Wednesday, the 13th of September, 2006 at 09.45 hours.

by

Ganesh Natrajan
born at Alleppey, India.

Doctoral Committee

Promotor

Prof.dr. T. K. Sixma

Other members:

Prof.dr. J.H.J. Hoeijmakers
Prof.dr. R. Kanaar
Prof.dr. H. te Riele

MutS: Recognition of DNA mismatches and initiation of repair.

Contents

Chapter 1

General Introduction

Contents

Changes to DNA basepairing.

DNA mismatch repair in *Escherichia coli.*

DNA mismatch repair in eukaryotes.

DNA mismatch repair and cancer.

Aims and focus of this thesis.

Mismatch binding by MutS.
Analysis of the structure of the ATPase domains.
Analysis of signalling between mismatch binding and ATPase domains.

References.

Chapter 1   Introduction

Changes to DNA base pairing

The accurate replication of DNA during cell division is an important prerequisite for the continuation of life. Ever since the crystal structure of DNA revealed the highly specific pairing scheme (1) of the bases, it became obvious that this held the key to the accurate replication of DNA. Changes to these base pairs result in mutations and all organisms possess repair mechanisms to prevent these and protect the integrity of their genomes. Changes to base pairs can come about due to covalent modifications like alkylation, methylation and oxidation. The most common cause of point mutations in humans is the spontaneous additions of a methyl group to a cytosine, followed by deamination to a thymidine. Changes to base pairing also occur by incorporation of wrong bases by DNA polymerases during DNA synthesis, resulting in mismatches. In addition to incorporating of wrong bases, polymerases can also slip during synthesis, resulting in unpaired bases or Insertion deletion loops (IDLs). While replicative polymerases have proofreading mechanisms which detect such errors on time, some errors escape this. DNA mismatch repair mechanisms detect and repair such errors and by doing so, reduce mutation rates by a 100 to 1000 fold. The importance of DNA mismatch repair is made obvious by the fact that its essential attributes are conserved in all organisms, from bacteria to humans.

DNA mismatch repair in *Escherichia coli*

Our primary focus has been to understand the molecular mechanism of DNA mismatch repair in *Escherichia coli* as a paradigm for understanding the working of the human mismatch repair pathway. The essential elements of DNA mismatch repair in *E. coli*, yeast and humans have been widely studied and very well understood (reviewed in (2-4)).
Many aspects of mismatch repair are similar in all organisms, ranging from bacteria to humans (5). In *E. coli* (6), the mismatch or the insertion-deletion loop (IDL) created by the polymerase is bound by MutS (Fig 1). In *E. coli*, the accepted functional form of MutS is the dimer. The MutS dimer (Fig 2) can be divided into two functional regions. These are the mismatch binding region and the ATPase region. The mismatch is bound by the mismatch binding regions, using the mismatch binding and the clamp domains. In an in-vivo experiment (7), MutS-GFP was seen to co-localize with chromosomes and form foci in some cells, which dramatically increased when mismatch formation was induced by 2-aminopurine.
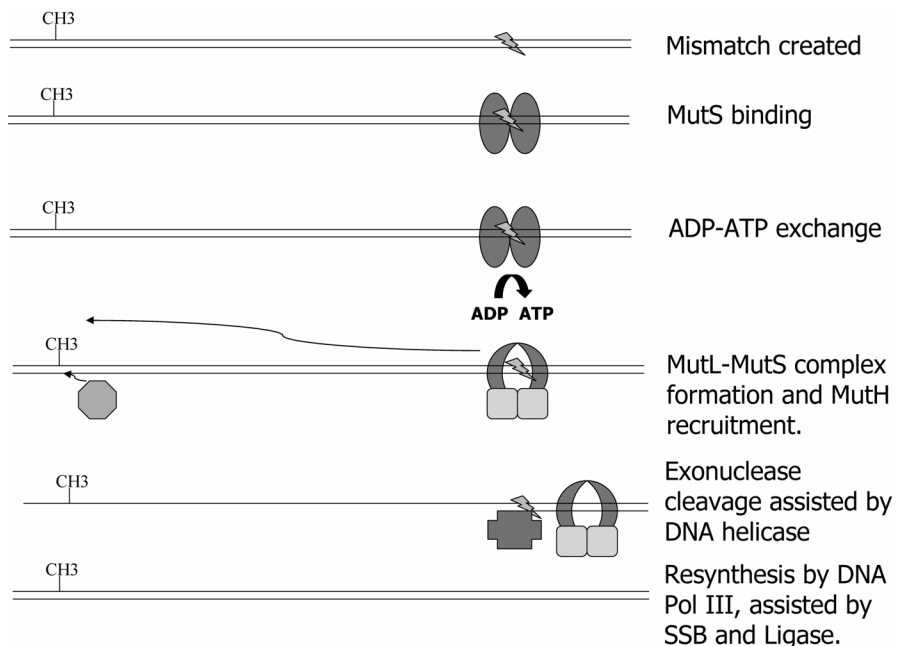
Figure 1. Schematic view of DNA mismatch repair in Escherichia coli.

Following mismatch binding, MutS exchanges the pre-bound ADP in its ATPase domain for ATP, upon which it attains its ATP state. In this state, it forms a complex with another protein, MutL (8), which has been seen to associate with MutS foci in-vivo (7).

MutL in *E. coli*, is a homodimer in structure (9). It consists of an N-terminal nucleotide binding domain, which is conserved in the MutL family of proteins, and a C-terminal dimerization domain which is conserved only in prokaryotes. These two domains are linked by an unstructured region of ~80 residues, the function of which remains unknown. However, the observation that this region can tolerate substitutions or deletions to one-thirds of its sequence suggests (10) that its composition itself is not important for the function of the protein. Full length MutL possesses a very weak ATPase activity, and belongs to the family of GHL atpases (11). The truncated N-terminal domain, which is a monomer by itself, has a greatly reduced ATPase activity compared to the full length protein (11), suggesting that the dimerization is important for this activity.

The formation of the MutS-MutL complex leads to the search for the signal that discriminates the parent from the daughter strand. In gram negative

E. coli, this signal is the transiently unmethylated d(GATC) site on the daughter strand (6), which may be up to a 1000 base pairs away from the actual mismatch (Fig 1). The MutS-MutL complex then recruits another enzyme called MutH, which cleaves the daughter strand at this point. MutH is an endonuclease, which specifically targets hemimethylated DNA substrates. It has sequence homology with Sau3A1 and is structurally similar to the PvuII restriction endonucleases. The crystal structure of MutH (12) has revealed that its active site is located in the middle of two pivoted domains and the movements of these two correlate with that of a C-terminal helix. It has been proposed that this helix is the molecular lever through which the enzyme could communicate with MutS-MutL. More recent crystal structures of MutH in complex with unmethylated and hemimethylated DNA substrates (13) have revealed that the active site is more compact around the hemimethylated DNA substrate than around the unmethylated one, resulting in more efficient cleavage of the former. An important lysine, widely conserved in many restriction endonucleases is involved in DNA binding, and is proposed to be a lynchpin, involved in substrate recognition.

Following this cleavage of the DNA at the hemimethylated site by MutH, the entire length of the daughter strand from the cleavage to the mismatch, has to be unwound prior to removal. UvrD or DNA helicase II, involved in this process, has been shown to associate with every domain of MutL (10,14). After the DNA is unwound, it is removed by excision. The exonucleases ExoI, ExoVII, ExoX or ExoVII, RecJ are recruited for this, depending on whether the d(GATC) cleavage is on the 5' or on the 3' side of the mismatch (15), either direction requiring the cooperative action of DNA helicase II. Interestingly, in-vivo studies reveal that the deficiency of these exonucleases does not result in a mutator phenotype in bacteria (16,17). The quadruple mutant deficient in the four exonucleases RecJ, ExoI, ExoVII, ExoX, and the triple mutant deficient in RecJ, ExoVII, ExoI, grow poorly in the presence of 2-aminopurine and show filament formation, indicating SOS DNA damage response (17). However, null mutations in MutS, MutL and MutH suppress this phenotype suggesting that the low mutability of the exonuclease deficient strains may be attributed not to the redundancy of these exonucleases, but to the non-recovery of mutations or the reduction in viability and/or chromosome loss attributed to the activation of mismatch repair in the absence of these exonucleases (17). The action of these exonucleases during DNA mismatch repair leads to the removal of the entire daughter strand from the cleavage to the mismatch and its resynthesis by the action of DNA polIII, Single Stranded DNA binding protein (SSB) and DNA ligase, thereby leading to the

repair of the mismatch.

DNA mismatch repair in eukaryotes

DNA mismatch repair in eukaryotes is very similar to the bacterial process, despite the fact that the genomes are much larger and more complex (18,19). The role of MutS is played by two distinct multi-protein complexes, MSH2-MSH6 and MSH2-MSH3, known more commonly as MutSα and MutSβ respectively (20). While MutSα binds to mismatches and shorter unpaired base loops, MutSβ binds to unpaired loops of all lengths. There is, therefore, a redundancy between these in the repair of shorter loops (21).

Figure 2. Front view of the MutS-mismatched DNA complex. Monomer A, which binds to the DNA mismatch and has an ADP-Mg bound to its ATPase domain, is on the right.

The role of MutL is played by its eukaryotic homolog, the complexes MLH1-PMS1 in yeast and MLH1-PMS2 in humans. It is not known what the strand discrimination signal in eukaryotic DNA is, given that there are no methylated d(GATC) sites in eukaryotic genomes. Due to this, there is also no known MutH homolog. There is also no known homolog for Helicase II in eukaryotes. The excision is done by ExoI, a 5'-3' exonuclease. The first evidence of the role of this exonuclease came from studies on the yeast *Schizosaccharomyces pombe* (22), where it was revealed that the lack of

this protein resulted in a mild mutator phenotype. In humans, this protein has been shown to interact with all the other proteins in the mismatch repair pathway (23), and is therefore, important.

The resynthesis, following the excision, is done by DNA Pol δ, assisted by RPA and DNA ligase. PCNA too has been proposed to play several roles in mismatch repair. It has been known to interact with all the proteins involved in the initiation of mismatch repair (24,25), with some of the interactions being ATP dependent (26). Amongst other things, PCNA has also been proposed to increase the mismatch specificity of MSH2-MSH6, and associate with components of the replication fork, delivering the mismatch repair proteins to newly synthesized DNA mismatches (Reviewed in (3)).

DNA mismatch repair and cancer

It is understandable that defects in any pathway that maintains genomic fidelity will lead to undesirable consequences. Defects in mismatch repair genes lead to accumulation of spontaneous mutations over successive generations. So in bacteria, the result is always a mutator phenotype (27-29). In addition to preventing mutations, mismatch repair pathways also prevent recombination between similar but diverging sequences, and defects in mismatch repair pathways have been known to disrupt the barrier between interspecies recombination in bacteria (30,31).

In humans, defects in the mismatch repair pathway can have serious consequences. Heterozygous mutations in mismatch repair proteins lead to cancer predisposition called HNPCC or Hereditary Non Polyposis Colorectal Cancer (32-34). These mutations occur mostly in *MSH2* and *MLH1*, and to a smaller extent in the *MSH6* gene (35). In addition, an important causative agent of spontaneous tumours is the promoter hypermethylation of the *MLH1* gene (36). HNPCC tumours are characterized by instabilities in short DNA repeat sequences due to reduced repair of polymerase slippage errors (37,38) called microsatellite instabilities (MSI or MIN). A 100-1000 fold increase in the mutation frequency is seen in cells (39), owing to the reduction in the efficiency of mismatch repair. Studies have shown that pathogenic mutations in mismatch repair proteins found in patients do affect the biochemical functions of mismatch repair proteins (40,41) and can therefore, affect their ability to carry out repair. It is also generally accepted that the phenotype brought about by these mutations is responsible for cancer predisposition and it has also been shown that colorectal adenomas develop into malignant carcinomas faster in HNPCC patients than in normal patients (42). However, the actual link between

defects in mismatch repair genes and tissue specific cancer in humans remains unclear. Studies in mice (43) have shown that while MSH2 deficiency causes a strong cancer predisposition, MSH6 deficiency led to lymphomas or epithelial tumours in the skin and uterus, but not in the intestines. MSH3 deficiency did not cause cancer predisposition but, in an MSH6 deficient background, led to intestinal tumours. Finally, mismatch repair genes are also known to play a role in activating cell cycle checkpoints and in the induction of apoptosis as a response to accumulating DNA damage. DNA mismatch repair deficiency in cells has been known to bring about resistance to chemotherapeutic and DNA modifying drugs like cisplatin (44), thereby reducing apoptosis by them.

Aims and focus of this thesis

The main aim of this thesis has been to understand mismatch recognition by MutS from a structural standpoint, and the relationship between this and the ATPase activity of the protein. This work carried out towards this end can be divided mainly into three parts.

1. Analysis of mismatch binding by MutS.
2. Analysis of the structure of the ATPase domains, including the analysis of domain movements caused by nucleotide binding, and the MutS ATPase cycle.
3. Analysis of the signalling between mismatch binding and ATPase domains.

Mismatch binding by MutS.

Two crystal structures of MutS, those of *E. coli* and *Thermus aquaticus*, revealed, for the first time, the nature of the MutS-DNA interaction (45,46) (Fig 2). The most striking feature is the structure of the DNA, which is sharply kinked by 60°. Since DNAs with mismatches are not ordinarily kinked in such a way (47-50), this could only be induced by the binding of the protein (Fig 3). Such distortions of the DNA are frequently seen in structures of DNA-protein complexes involved in repair (51-54) and are often necessary for the protein residues to gain access to the bases in the DNA helix. MutS is a structural heterodimer comprising of two monomers, A and B, of which only monomer A binds to the mismatch.

Figure 3. Figure showing the kinked DNA held between the mismatch binding (bottom right) and clamp domains (top left) of both monomers

Aspecific backbone contacts to the DNA are made by the mismatch binding and the clamp domains of both monomers. The kinking of the DNA and the bringing together of the mismatch binding and clamp domains, achieves a stable binding conformation. The actual mismatch is bound using a phenylalalanine F36 (*E. coli*) in the mismatch binding domain of monomer A, which is seen stacking onto one of the mismatched bases. A nitrogen on the same base is involved in the formation of a hydrogen bond to another widely conserved residue, the glutamate E38 (Fig 4) of the same monomer A. The phenlylalanine F36 is absolutely vital for mismatch repair as mutating it into an alanine eliminates DNA binding and mismatch repair entirely (55,56). Since mutating this residue eliminates even homoduplex binding, MutS could be imagined as a type of DNA scanner, which scans the DNA with the help of the F36, trying to find a mismatch and kink the DNA (57). It is probably helped by inherent instabilities in DNA caused by the presence of mismatches (58,59). The glutamate E38 plays a role in stabilizing mismatch binding, bringing about discrimination between homoduplex and heteroduplex DNA, as well as helping in signalling

between the mismatch binding and ATPase domains for the formation of the ATP state. It shall be discussed subsequently.

The mechanism by which DNA damages, once found, are stacked upon by aromatic residues and stabilized by hydrogen bonding to residues is found in other DNA repair enzymes. In the MutY and MutM DNA glycosylases which remove 8-oxo-guanine from DNA (60,61), a tyrosine stacks upon the flipped out 8-oxo-guanine, and the extrahelical base is stabilized by hydrogen bonds to surrounding residues. In the human 8-oxo-guanine targeting enzyme hOGG1 (62), similar interactions ensure that only the substrate 8-oxo-guanine gets inserted into the active site, while normal guanines are excluded. While in the MutS structures, the base stacked
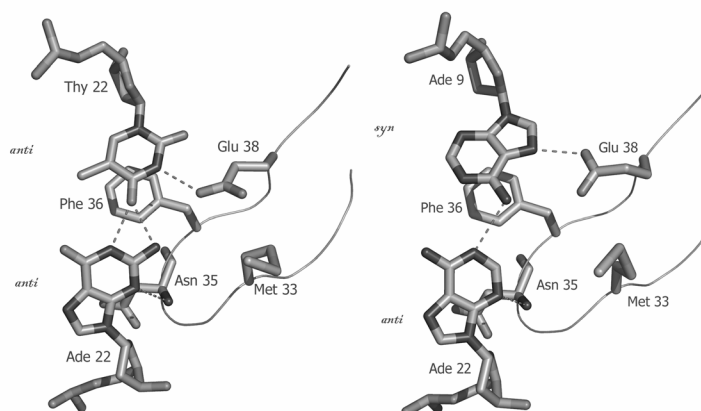

Figure 4. Structure of a G.T mismatch (left) and an A.A mismatch (right) bound to MutS

upon by F36 is not extra helical, it does show a considerable change in orientation compared to the one in which it exists when there is no protein bound to the mismatch. The structures of MutS in complex with four different mismatches reported in chapter 2 of this thesis show the same trend of the F36 stacking onto one of the mismatched bases, and the same base being moved away from its position (63) in the absence of the protein. It does seem that the protein does this so as to expose a hydrogen bond accepting nitrogen to the E38. Further, the DNAs in the various structures show slight differences in the conformation, yet maintaining all the contacts with the protein, which in turn is quite invariant between the structures. Even in the crystal structure of MutS in another space group with a different crystal packing (discussed in Chapter

3), the interactions with DNA are preserved. Therefore, it can be said that MutS uses a common mismatch binding mode to bind different mismatches, the fundamental requirements being the kinking of the DNA, stacking of the F36 and the re-orientation of the stacked base so as to enable hydrogen bond formation by E38. The other signatures of this common mode are the orientations of the base onto which Phe 36 stacks. Stacking on a pyrimidine retains it in an *anti* orientation of the glycosydic bond, while stacking on a purine causes the purine to adopt the *syn* orientation of the same (Fig 4). The protein seems to be bringing this about so that the purine N7 or the pyrimidine N3 can act as a hydrogen bond donor to Glu 38. Interestingly, the N7 in purines must either be protonated or exist in an alternate tautomeric form in order to be able to form this hydrogen bond.

The structure of the *E. coli* MutS in complex with an unpaired thymidine (63) shows a marked difference from the structure of the same lesion in complex with the *Taq* MutS (46). While the stacking of F36 and hydrogen bond formation by E38, and all the contacts between the protein and DNA remain identical in both, the base pairs in the immediate vicinity of the mismatch show serious rearrangements in the *E. coli* structure. Such differences in the binding of the similar substrate by the same enzyme from two different organisms have been seen before in methyltransferases from HhaI (64) and HaeIII (65). Although methyltransferases are not repair enzymes, they are similar to the extent that they target one particular base i.e., the cytosine for binding.

The distortion of the Watson-Crick basepairs in the vicinity of the unpaired thymidine also reveals for the first time, the role of the surrounding sequence context in mismatch recognition shown in biochemical studies (66,67). This has far-reaching consequences for mismatch repair in general. Mismatch repair is triggered by recognition of the mismatch and its stable binding by MutS. This implies that the repair of certain mismatches can be made less efficient, or even bypassed, if the sequence around the mismatch doesn't permit effective MutS binding. Dramatic differences in DNA mismatch repair efficiencies can be obtained by making single or double nucleotide modifications in substrates with multiple mismatches or longer loops of unpaired nucleotides (68), and this effect is increased many fold when mismatch repair components are knocked out. The crystal structures of *Taq* MutS in complex with loops longer than one nucleotide (W.Yang, personal communication) while confirming the common mode for mismatch binding, do not explain the substrate dependent bifurcation of bacterial MutS into MSH2/MSH6 and MSH2/MSH3 in eukaryotes, with a redundancy between them. A definite direction for

future research is the study of MutS in complex with a number of such complex substrates, so as to elucidate the exact, substrate binding mode required for repair of mismatches as well as single or multiple loop substrates.

Analysis of the structure of ATPase domains

MutS belongs to the ABC ATPase family. The proteins belonging to this family are involved in a variety of cellular processes, including transport across membranes (ABC transporters), chromosome structural maintenance (SMC proteins), and DNA repair enzymes like Rad 50 (double strand break repair), MutS (mismatch repair) and UvrA (nucleotide excision repair) (reviewed in (69)). ABC ATPases are characterized by four standard motifs, which are involved in nucleotide binding and hydrolysis. These are the P-loop or the Walker A motif (MutS residues 614-621), the Walker B motif (MutS residues 693-697), a highly conserved histidine (MutS residue 728), which is a glutamine in some proteins, and a signature loop (MutS residues 657-670) (Fig 5), which is unique to ABC ATPases. In addition to this, another common feature amongst ABC ATPases is that they are all dimeric, with composite ATP binding sites involving residues from both monomers, with the ATP either causing the dimerization (70-72) or tightening an already existing dimer (73). The two ATP binding sites are asymmetric in substrate binding and hydrolysis (74-77). Yet another feature that is common amongst ABC ATPases is the long range signalling that follows the binding of ATP, leading to conformational changes in distant substrate binding domains. The first crystal structure of the *E. coli* MutS in complex with ADP, gave an insight into the structure of its ATPase domains. These are situated at the other end of the MutS dimer, at a distance of 100 Å from the mismatch binding domains. In the crystal structure, the monomer A has an ADP-Mg bound in its ATPase domain while the monomer B is empty (45). The ADP is bound using the highly conserved P-loop or the Walker A motif (residues 614-621 in *E. coli*). The magnesium is octahedrally coordinated by a conserved serine S621, four water molecules and the beta-phosphate of the ADP (Fig 5). Another widely conserved motif, the Walker B motif, also interacts with the magnesium. Two conserved acidic residues are found on this loop, D693 and E694. The E694 is thought to play a role in lowering the energy barrier of the transition state upon ATP binding by repulsion of the gamma phosphate, thereby enabling hydrolysis of the ATP (78), while D693 is involved in binding the magnesium. The ATPase site of monomer B is empty. Two loops, referred to as signature loops, are disordered in both
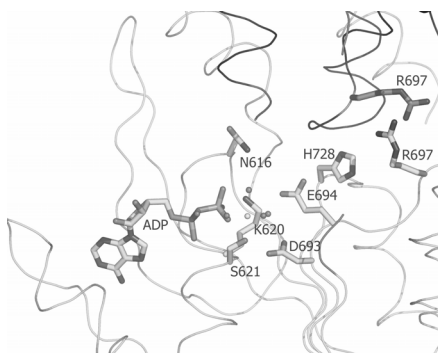
Figure 5. View of the ADP binding site in the MutS-DNA-ADP-Mg complex. The residues involved in ADP-Mg binding are indicated.

monomers (residues 659-669) in the ADP bound structure. In an attempt to solve the structure of MutS in complex with ATP, crystals of MutS were soaked into ATP (73), as attempts to co-crystallize MutS with ATP were unsuccessful. The resultant structure showed ATP binding to both monomers, with the ordering of the signature loops (659-669), along with rearrangements of several residues around the nucleotide binding site, namely N616, H728 and S668. ATP binding also led to an increase in MutS dimer affinity as determined by analytical ultracentrifugation experiments, and mutating these rearranged residues eliminated this increase. From this it was proposed (73) that the ordering of the signature loop, coupled with the rearrangements of the aforesaid residues and the tightening of the dimer, was important for the proper formation of the ATPase active site and signal transmission to the DNA binding domains to enable the formation of the ATP state. This structure, however, does not show the actual ATP state of MutS as it is trapped in a crystal during the ATP soak, and is therefore unable to undergo all the conformational changes required. The actual structure of the ATP state of MutS remains elusive.

Other crystal structures and biochemical studies have shown MutS in different nucleotide binding modes. The crystal structure of *Taq* MutS has been solved with both monomers empty (46), both monomers with ADP-Mg (78), and both monomers with ADP-beryllium fluoride, a potent inhibitor of MutS family of proteins (79). Biochemical studies have also proposed that upon mismatch recognition, MutS would be in either a ATP.MutS.ADP (ATP in monomer A, ADP in Monomer B) or in a ATP.MutS.ATP form (77), while the ADP.MutS.ATP form is seen in the

absence of DNA or in presence of fully matched DNA. Any crystal structure of MutS would be that of the enzyme in a particular state in its nucleotide binding and hydrolysis cycle and therefore, a series of structures in different states are necessary to get the complete picture.

In the third chapter of this thesis, we report a crystal structure of MutS, in complex with mismatched DNA and ADP in a new crystal form with a different packing. In the original crystal structure a 30 base pair DNA was used, out of which only 16 base pairs were actually visible in the structure (45), although it could be seen from the packing that the remaining 14 invisible base pairs were involved in crystal contacts. Crystallizing the protein with a DNA containing only 16 DNA base pairs yields a new crystal packing. An important feature of this structure is the severe disorder seen in the mismatch binding domain of monomer B, seen previously in the structure of *Taq* MutS (46). The DNA binding is unchanged compared to the previously solved MutS-DNA structures, confirming the conclusions of chapter 2 that it is largely the DNA which adapts itself to the protein, so as to bring about an induced fit binding mode. In addition, important changes are seen in the ATPase domains. The magnesium ion, seen along with the ADP in the previous structures is missing, along a complete ordering of the signature loop of the monomer B. This ordering leads to a widely conserved serine, S668 present on the signature loop to become visible. This has only been seen before in the structure of the MutS soaked in ATP (73), and in that of *Taq* MutS in complex with ADP-BeF (79). The signature loop has been shown in other studies to be a gamma phosphate sensor, and plays an important role in ATP hydrolysis (71,79). While its ordering can be explained in the ATP bound structure, it is very intriguing that it should be visible in an ADP bound structure in the absence of magnesium. Upon closer analysis, it becomes clear that several of the surrounding residues also play a role in the conformational changes. These include the N616 on the P-loop and the way it is oriented towards the histidine H728. E694, a highly conserved residue on the Walker B motif, which has been seen (Fig 5) hydrogen bonded to one of the waters coordinating with the magnesium, also changes orientation in the absence of magnesium. So the removal of magnesium results in a number of conformational changes involving highly conserved residues, with the consequence of the signature loop getting ordered.

This chapter also discusses the Mg-Free state from the point of view of the MutS-ATPase cycle, and compares it with the GTPase cycle of G-proteins. The nucleotide bound, Mg Free states are important intermediates in the GTPase cycle of G-proteins, and have been well characterized (80). This comparison leads to the conclusion that the ADP-bound, Mg-Free structure

of MutS is also an intermediate in the MutS ATPase cycle. The importance of Mg in ADP-ATP exchange following mismatch binding has also been characterized here, as a part of the analysis.

Analysis of the signalling between mismatch binding and ATPase domains.

After binding to the mismatch, MutS binds ATP, leading to the attainment of its ATP state. This ATP state is commonly referred to as a 'sliding clamp'. This is because, upon binding ATP, MutS attains a conformation where it releases the bound mismatch. Three models exist, which try to explain the exact role of ATP in the function of MutS. In one model, MutS uses the energy derived from ATP hydrolysis to propel itself along DNA, after locating the mismatch, in its search for the other mismatch repair components (81,82). In the other model, MutS is a molecular switch, which can exist in two states. In the ADP state, it binds mismatches. Mismatch binding causes a signal to be transmitted to its ATPase domains, whereupon it exchanges the bound ADP for ATP. This causes conformational changes in the protein, causing it form its ATP state, or the sliding clamp (83-85). In this conformation, it initiates the other downstream processes which lead to the repair of the mismatch. ATP hydrolysis is required only to bring the protein back to the ADP state where it could begin the process of mismatch binding all over again. In the third model, MutS uses ATP binding to verify the binding of (78,86) the mismatch, using a kinetic proofreading mechanism, and authorizes repair once this verification is carried out. Based on several studies, the first model is generally less favoured over the other two.

A number of studies have revealed information about ATP binding, sliding clamp formation and the sequence of events involved. The binding and hydrolysis of ATP by the MutS dimer is asymmetric (76,77,87) as the two sites have different affinities for ATP binding. It has been shown that these differential affinities are maintained by an arginine, R697, located at the end of the walker B motif. This arginine is unique to the MutS family of proteins, and mutating it to an alanine destroys the high affinity ATP binding site (76).

In the absence of any DNA, the release of ADP is the rate limiting step for steady state ATP hydrolysis (83,88) and this is reflected in the simulation of the steady state ATP hydrolysis rate by any DNA (86,88). On the other hand, rapid quench ATP hydrolysis experiments in the absence of DNA show an initial fast hydrolysis rate, followed by a slow phase (77,89). Mismatched (83,88,90) or homoduplex DNA binding by MutS results in a

dramatic increase of the ADP to ATP exchange. When the binding of MutS occurs to a mismatched DNA in need of repair, the burst ATP hydrolysis is inhibited (verification of the bound mismatch (78)) and a stable MutS-DNA-ATP state is formed (77,87,89), followed by the initiation of downstream processes. If however, the binding occurs to homoduplex DNA not in need of repair, ATP binding and the associated verification procedure result in direct dissociation of MutS from DNA (91,92). This is also associated with the rapid hydrolysis of ATP.

It is obvious that this process involves considerable signalling between the mismatch binding and the ATPase domains, which are separated by around 100 Å. Such a long range signalling process must involve several residues but it can be envisioned that it has to begin with just one interaction, made by a single residue: glutamate, E38, which forms a hydrogen bond to one of the mismatched bases (Fig 4). Previous studies on the E38 have shown that the negative charge on the acid side chain helped in the discrimination between mismatched and normal DNA, so as to help the protein find mismatches (93). However, since the initial discrimination between mismatched and homoduplex DNA is only of the order of 8 to 10 fold (67,93) and a second mismatch verification mechanism involving ATP binding had also been proposed (78), it became clear that E38 had a role to play in both these processes.

In the third and final chapter of this thesis (86) we have examined the role of the E38 by mutating it to a glutamine, an alanine and a threonine. We have performed in-vivo mismatch repair assays on these mutants, solved their crystal structures in complex with a G.T mismatch containing DNA, studied homoduplex and mismatched DNA binding and the signalling process between the mismatch binding domains and the ATPase domains using a variety of assays. While the E38T and E38Q can repair mismatches as well as the wild type protein in-vivo, E38A is greatly reduced in its ability to do so. DNA band shifts with end blocked mismatched substrates in the presence of ATP reveal that all mutants can form the sliding clamp, with the E38A lagging behind the others. However, when homoduplex DNA is used, all the mutants show release in the presence of ATP, clearly demonstrating the ATP induced mismatch verification. Rapid quench ATPase assays show that the E38A is not able to inhibit the burst of ATP hydrolysis, while the wild type and the other mutants do. This inability of the E38A is reflected in its faster release of a mismatched DNA substrate in the presence of ATP, during the inherent flow in a Biacore surface plasmon resonance experiment. Finally, crystal structures of E38T, E38A and E38Q in complex with a G.T mismatch reveal that the E38A cannot form a hydrogen bond to the mismatched thymidine, while the other mutants can.

Therefore, it can be concluded that the hydrogen bond formed by E38 to the mismatched base plays a key role in intermolecular signalling which results in a stable MutS-DNA-ATP complex, which can then initiate further repair events.

Thus, the aim of this thesis has been to shed light on the molecular processes governing mismatch recognition and the initiation of mismatch repair by MutS, an enzyme displaying a remarkable versatility of being able to recognize 11 out of 12 possible DNA mismatches and unpaired DNA base loops up to three nucleotides in length, and initiating repair. The fundamental mechanisms of DNA mismatch targeting, and signalling, leading to the verification and repair of mismatches by MutS make this protein an excellent example of Nature's chemical ingenuity, perfected over millions of years of evolution. Surely, it can be said that the continuity of life can be attributed to the invisible actions of the cell's genomic caretakers and their remarkable abilities.

References

1. Watson, J. D., and Crick, F. H. (1953) *Nature* **171**(4356), 737-738
2. Kunkel, T. A., and Erie, D. A. (2004) *Annu Rev Biochem*
3. Schofield, M. J., and Hsieh, P. (2003) *Annu Rev Microbiol* **57**, 579-608
4. Iyer, R. R., Pluciennik, A., Burdett, V., and Modrich, P. L. (2006) *Chem Rev* **106**(2), 302-323
5. Lahue, R. S., Au, K. G., and Modrich, P. (1989) *Science* **245**(4914), 160-164
6. Lahue, R. S., and Modrich, P. (1988) *Mutat Res* **198**(1), 37-43
7. Smith, B. T., Grossman, A. D., and Walker, G. C. (2001) *Mol Cell* **8**(6), 1197-1206.
8. Au, K. G., Welsh, K., and Modrich, P. (1992) *J Biol Chem* **267**(17), 12142-12148
9. Ban, C., and Yang, W. (1998) *Cell* **95**(4), 541-552
10. Guarne, A., Ramon-Maiques, S., Wolff, E. M., Ghirlando, R., Hu, X., Miller, J. H., and Yang, W. (2004) *EMBO J* **23**(21), 4134-4145
11. Ban, C., Junop, M., and Yang, W. (1999) *Cell* **97**(1), 85-97
12. Ban, C., and Yang, W. (1998) *EMBO J* **17**(5), 1526-1534
13. Lee, J. Y., Chang, J., Joseph, N., Ghirlando, R., Rao, D. N., and Yang, W. (2005) *Mol Cell* **20**(1), 155-166
14. Hall, M. C., Jordan, J. R., and Matson, S. W. (1998) *EMBO J* **17**(5), 1535-1541

15.   Grilley, M., Griffith, J., and Modrich, P. (1993) *J Biol Chem* **268**(16), 11830-11837

16.   Harris, R. S., Ross, K. J., Lombardo, M. J., and Rosenberg, S. M. (1998) *J Bacteriol* **180**(4), 989-993

17.   Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S. T., and Modrich, P. (2001) *Proc Natl Acad Sci U S A* **98**(12), 6765-6770

18.   Buermeyer, A. B., Deschenes, S. M., Baker, S. M., and Liskay, R. M. (1999) *Annu Rev Genet* **33**, 533-564

19.   Kolodner, R. D., and Marsischky, G. T. (1999) *Curr Opin Genet Dev* **9**(1), 89-96.

20.   Fishel, R., and Wilson, T. (1997) *Curr Opin Genet Dev* **7**(1), 105-113

21.   Marsischky, G. T., Filosi, N., Kane, M. F., and Kolodner, R. (1996) *Genes Dev* **10**(4), 407-420.

22.   Szankasi, P., and Smith, G. R. (1995) *Science* **267**(5201), 1166-1169

23.   Schmutte, C., Sadoff, M. M., Shim, K. S., Acharya, S., and Fishel, R. (2001) *J Biol Chem* **276**(35), 33011-33018

24.   Lee, S. D., and Alani, E. (2006) *J Mol Biol* **355**(2), 175-184

25.   Nielsen, F. C., Jager, A. C., Lutzen, A., Bundgaard, J. R., and Rasmussen, L. J. (2004) *Oncogene* **23**(7), 1457-1468

26.   Gu, L., Hong, Y., McCulloch, S., Watanabe, H., and Li, G. M. (1998) *Nucleic Acids Res* **26**(5), 1173-1178

27.   Glickman, B. W., and Radman, M. (1980) *Proc Natl Acad Sci U S A* **77**(2), 1063-1067

28.   Rewinski, C., and Marinus, M. G. (1987) *Nucleic Acids Res* **15**(20), 8205-8215

29.   LeClerc, J. E., Li, B., Payne, W. L., and Cebula, T. A. (1996) *Science* **274**(5290), 1208-1211

30.   Rayssiguier, C., Thaler, D. S., and Radman, M. (1989) *Nature* **342**(6248), 396-401

31.   Zahrt, T. C., and Maloy, S. (1997) *Proc Natl Acad Sci U S A* **94**(18), 9786-9791

32.   Lynch, H. T., and de la Chapelle, A. (1999) *J Med Genet* **36**(11), 801-818.

33.   Peltomaki, P. (2003) *J Clin Oncol* **21**(6), 1174-1179

34.   Pensotti, V., Radice, P., Presciuttini, S., Calistri, D., Gazzoli, I., Grimalt Perez, A., Mondini, P., Buonsanti, G., Sala, P., Rossetti, C., Ranzani, G. N., Bertario, L., and Pierotti, M. A. (1997) *Genes Chromosomes Cancer* **19**(3), 135-142

35.   Muller, A., and Fishel, R. (2002) *Cancer Invest* **20**(1), 102-109

36.   Kuismanen, S. A., Holmberg, M. T., Salovaara, R., de la Chapelle, A., and Peltomaki, P. (2000) *Am J Pathol* **156**(5), 1773-1779

37.   de Leeuw, W. J., Dierssen, J., Vasen, H. F., Wijnen, J. T., Kenter, G. G., Meijers-Heijboer, H., Brocker-Vriends, A., Stormorken, A., Moller, P., Menko, F., Cornelisse, C. J., and Morreau, H. (2000) *J Pathol* **192**(3), 328-335

38.   Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N. C., Lynch, H. T., Watson, P., Jass, J. R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. (1996) *Nat Med* **2**(2), 169-174

39.   Eshleman, J. R., and Markowitz, S. D. (1996) *Hum Mol Genet* **5**, 1489-1494

40.   Jager, A. C., Rasmussen, M., Bisgaard, H. C., Singh, K. K., Nielsen, F. C., and Rasmussen, L. J. (2001) *Oncogene* **20**(27), 3590-3595

41.   Heinen, C. D., Wilson, T., Mazurek, A., Berardini, M., Butz, C., and Fishel, R. (2002) *Cancer Cell* **1**(5), 469-478

42.   Kinzler, K. W., and Vogelstein, B. (1996) *Nature* **379**(6560), 19-20

43.   de Wind, N., Dekker, M., Claij, N., Jansen, L., van Klink, Y., Radman, M., Riggins, G., van der Valk, M., van't Wout, K., and te Riele, H. (1999) *Nat Genet* **23**(3), 359-362

44.   Bellacosa, A. (2001) *Cell Death Differ* **8**(11), 1076-1092

45.   Lamers, M. H., Perrakis, A., Enzlin, J. H., Winterwerp, H. H., de Wind, N., and Sixma, T. K. (2000) *Nature* **407**(6805), 711-717.

46.   Obmolova, G., Ban, C., Hsieh, P., and Yang, W. (2000) *Nature* **407**(6805), 703-710.

47.   Skelly, J. V., Edwards, K. J., Jenkins, T. C., and Neidle, S. (1993) *Proc Natl Acad Sci U S A* **90**(3), 804-808.

48.   Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., and Kennard, O. (1987) *J Biol Chem* **262**(21), 9962-9970.

49.   Hunter, W. N., Brown, T., Anand, N. N., and Kennard, O. (1986) *Nature* **320**(6062), 552-555.

50.   Hunter, W. N., Brown , T., & Kennard, O. (1987) *Nucleic Acids Res* **15**, 6589-6600

51.   Zharkov, D. O., Golan, G., Gilboa, R., Fernandes, A. S., Gerchman, S. E., Kycia, J. H., Rieger, R. A., Grollman, A. P., and Shoham, G. (2002) *EMBO J* **21**(4), 789-800

52.   Tanaka, Y., Nureki, O., Kurumizaka, H., Fukai, S., Kawaguchi, S., Ikuta, M., Iwahara, J., Okazaki, T., and Yokoyama, S. (2001) *EMBO J* **20**(23), 6612-6618

53.   Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) *Nature* **399**(6737), 708-712

54.   Parkinson, G., Wilson, C., Gunasekera, A., Ebright, Y. W., Ebright, R. E., and Berman, H. M. (1996) *J Mol Biol* **260**(3), 395-408

55.   Yamamoto, A., Schofield, M. J., Biswas, I., and Hsieh, P. (2000) *Nucleic Acids Res* **28**(18), 3564-3569.

56.   Drotschmann, K., Yang, W., Brownewell, F. E., Kool, E. T., and Kunkel, T. A. (2001) *J Biol Chem* **276**(49), 46225-46229

57.   Wang, H., Yang, Y., Schofield, M. J., Du, C., Fridman, Y., Lee, S. D., Larson, E. D., Drummond, J. T., Alani, E., Hsieh, P., and Erie, D. A. (2003) *Proc Natl Acad Sci U S A* **100**(25), 14822-14827

58.   Werntges, H., Steger, G., Riesner, D., and Fritz, H. J. (1986) *Nucleic Acids Res* **14**(9), 3773-3790.

59.   Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J., Jr. (1999) *Biochemistry* **38**(12), 3468-3477.

60.   Fromme, J. C., Banerjee, A., Huang, S. J., and Verdine, G. L. (2004) *Nature* **427**(6975), 652-656

61.   Fromme, J. C., and Verdine, G. L. (2002) *Nat Struct Biol* **9**(7), 544-552

62.   Banerjee, A., Yang, W., Karplus, M., and Verdine, G. L. (2005) *Nature* **434**(7033), 612-618

63.   Natrajan, G., Lamers, M. H., Enzlin, J. H., Winterwerp, H. H., Perrakis, A., and Sixma, T. K. (2003) *Nucleic Acids Res* **31**(16), 4814-4821

64.   Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) *Cell* **76**(2), 357-369.

65.   Reinisch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) *Cell* **82**(1), 143-153.

66.   Joshi, A., and Rao, B. J. (2001) *J Biosci* **26**(5), 595-606.

67.   Brown, J., Brown, T., and Fox, K. R. (2001) *Biochem J* **354**(Pt 3), 627-633.

68.   Dekker, M., Brouwers, C., and te Riele, H. (2003) *Nucleic Acids Research* **31**(6), e27

69.   Hopfner, K. P., and Tainer, J. A. (2003) *Curr Opin Struct Biol* **13**(2), 249-255

70.   Lammens, A., Schele, A., and Hopfner, K. P. (2004) *Curr Biol* **14**(19), 1778-1782

71.   Hopfner, K. P., Karcher, A., Shin, D. S., Craig, L., Arthur, L. M., Carney, J. P., and Tainer, J. A. (2000) *Cell* **101**(7), 789-800

72.   Chang, G., and Roth, C. B. (2001) *Science* **293**(5536), 1793-1800

73.   Lamers, M. H., Georgijevic, D., Lebbink, J. H., Winterwerp, H. H., Agianian, B., de Wind, N., and Sixma, T. K. (2004) *J Biol Chem* **279**(42), 43879-43885

74. Nikaido, K., Liu, P. Q., and Ames, G. F. (1997) *J Biol Chem* **272**(44), 27745-27752

75. Bianchet, M. A., Ko, Y. H., Amzel, L. M., and Pedersen, P. L. (1997) *J Bioenerg Biomembr* **29**(5), 503-524

76. Lamers, M. H., Winterwerp, H. H., and Sixma, T. K. (2003) *EMBO J* **22**, 746-756

77. Antony, E., and Hingorani, M. M. (2004) *Biochemistry* **43**(41), 13115-13128

78. Junop, M. S., Obmolova, G., Rausch, K., Hsieh, P., and Yang, W. (2001) *Mol Cell* **7**(1), 1-12

79. Alani, E., Lee, J. Y., Schofield, M. J., Kijas, A. W., Hsieh, P., and Yang, W. (2003) *J Biol Chem* **278**(18), 16088-16094

80. Pan, J. Y., and Wessling-Resnick, M. (1998) *BioEssays* **20**, 516-521

81. Allen, D. J., Makhov, A., Grilley, M., Taylor, J., Thresher, R., Modrich, P., and Griffith, J. D. (1997) *EMBO J* **16**(14), 4467-4476

82. Blackwell, L. J., Martik, D., Bjornson, K. P., Bjornson, E. S., and Modrich, P. (1998) *J Biol Chem* **273**(48), 32055-32062

83. Gradia, S., Acharya, S., and Fishel, R. (1997) *Cell* **91**(7), 995-1005

84. Gradia, S., Acharya, S., and Fishel, R. (2000) *J Biol Chem* **275**(6), 3922-3930

85. Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J., and Fishel, R. (1999) *Mol Cell* **3**(2), 255-261

86. Lebbink, J. H., Georgijevic, D., Natrajan, G., Fish, A., Winterwerp, H. H., Sixma, T. K., and de Wind, N. (2006) *EMBO J* **25**(2), 409-419

87. Bjornson, K. P., Allen, D. J., and Modrich, P. (2000) *Biochemistry* **39**(11), 3176-3183

88. Acharya, S., Foster, P. L., Brooks, P., and Fishel, R. (2003) *Mol Cell* **12**(1), 233-246

89. Antony, E., and Hingorani, M. M. (2003) *Biochemistry* **42**(25), 7682-7693

90. Blackwell, L. J., Bjornson, K. P., Allen, D. J., and Modrich, P. (2001) *J Biol Chem* **276**(36), 34339-34347

91. Iaccarino, I., Marra, G., Dufner, P., and Jiricny, J. (2000) *J Biol Chem* **275**(3), 2080-2086

92. Selmane, T., Schofield, M. J., Nayak, S., Du, C., and Hsieh, P. (2003) *J Mol Biol* **334**(5), 949-965

93. Schofield, M. J., Brownewell, F. E., Nayak, S., Du, C., Kool, E. T., and Hsieh, P. (2001) *J Biol Chem* **276**(49), 45505-45508.

Chapter 2

Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates.

Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates

Ganesh Natrajan, Meindert H. Lamers, Jacqueline H. Enzlin, Herrie H. K. Winterwerp, Anastassis Perrakis and Titia K. Sixma*

Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands

ABSTRACT

We have refined a series of isomorphous crystal structures of the *Escherichia coli* DNA mismatch repair enzyme MutS in complex with G:T, A:A, C:A and G:G mismatches and also with a single unpaired thymidine. In all these structures, the DNA is kinked by ~60° upon protein binding. Two residues widely conserved in the MutS family are involved in mismatch recognition. The phenylalanine, Phe 36, is seen stacking on one of the mismatched bases. The same base is also seen forming a hydrogen bond to the glutamate Glu 38. This hydrogen bond involves the N7 if the base stacking on Phe 36 is a purine and the N3 if it is a pyrimidine (thymine). Thus, MutS uses a common binding mode to recognize a wide range of mismatches.

INTRODUCTION

Genomic integrity in organisms is maintained by a number of important DNA repair pathways. The DNA mismatch repair (MMR) pathway repairs mismatches and short insertion or deletion loops (IDLs). In addition, MMR also helps in preventing recombination between homologous but diverged DNA sequences (1–3). The fundamental mechanisms of MMR are similar in all organisms ranging from *Escherichia coli* to humans. In *E.coli*, MMR is initiated when the enzyme MutS recognizes and binds to mismatches or IDLs. This is followed by the uptake of ATP by MutS and the formation of a complex between MutS and the enzyme MutL. This complex initiates a number of events, leading to the recognition of the daughter strand, followed by its removal and resynthesis. In humans, the role of MutS is played by its homologs, the heterodimers MSH2/MSH6, which binds mismatches and IDLs, and MSH2/MSH3, which binds longer loops (1,2). The role of MutL in humans is played by the heterodimer MLH1/PMS2. Mutations in the genes that encode MMR proteins lead to mutator phenotypes in bacteria and cause a predisposition to cancer,

called hereditary non-polyposis colorectal carcinomas (HNPCC) in humans (4).

Two structures of MutS–DNA complexes have been reported already, the *Thermus aquaticus* enzyme in complex with a single unpaired thymidine (5), and the *E.coli* enzyme in complex with a G:T mismatch (6). The striking feature in both these structures is a sharp 60° kink in the DNA at the site of the mismatch–protein interaction. Mismatch recognition by MutS involves a phenylalanine, widely conserved in the MutS family (Phe 36 in *E.coli*, Phe 39 in *Taq*), which is seen stacking on one of the mismatched bases. The same base is also seen forming a hydrogen bond to a widely conserved glutamic acid (Glu 38 in *E.coli*, Glu 41 in *Taq*). In both structures, the conserved phenylalanine stacks on the thymidine. To find out how other mismatches are recognized, we have solved the structures of *E.coli* MutS in complex with A:A, G:G, C:A mismatches and an unpaired thymidine. Our results show that all these different lesions are recognized in a similar way, indicating a common binding mode for all mismatches.

MATERIALS AND METHODS

Protein expression and purification

ΔC800, an 800 residue C-terminal deletion construct of *E.coli* MutS (853 residues) in a pET3d vector (6), derived from the pMQ372 plasmid (7) was used to transform B834 (DE3) pLysS cells. A colony was picked, inoculated into 10 ml of minimal medium (8) and allowed to grow overnight at 30°C. This culture was diluted into 1 l minimal medium and grown at 30°C till it reached an OD (600 nm) of ~0.7. The temperature was then lowered to 23°C and the culture induced with IPTG (final concentration 1 mM) for 4 h. The cells were harvested, suspended in 25 ml of lysis buffer [50 mM HEPES pH 7.5, 200 mM NaCl, 10 mM β-mercaptoethanol, 5 mM EDTA, 1 mM PMSF and two protease inhibitor tablets (Roche)] and lysed by sonication. Following centrifugation at 39 000 r.c.f. for 50 min at 4°C, the cleared lysate was subjected to streptomycin precipitation. The volume of the lysate was first measured and streptomycin sulphate solution (25% w/v in

*To whom correspondence should be addressed. Tel: +31 20 5121959; Fax: +31 20 5121954; Email: t.sixma@nki.nl
Present address:
Jacqueline H. Enzlin, Institute of Molecular Cancer Research, University of Zurich, Zurich, Switzerland

Table 1. Crystallographic data

Complex	MutS–A:A	MutS–C:A	MutS–G:G	MutS–unpaired T
Beamline	BW7B (DESY)	X11 (DESY)	ID14EH1(ESRF)	ID14EH2 (ESRF)
Resolution range (Å)	15.0–2.4 (2.5–2.4)	15.0–2.9 (3.0–2.9)	15.0–2.6 (2.7–2.6)	15.0–2.9 (3.0–2.9)
Complete (%)	98.4 (86.1)	95.4 (93.8)	91.2 (60.5)	98.3 (92.4)
I/sig(I)	12.60 (1.60)	12.21 (1.97)	8.70 (1.10)	9.54 (1.57)
R_{merge} (%)	8.9 (61.4)	9.5 (76.0)	10.4 (62.8)	12.4 (72.5)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell parameters (Å)	a = 89.48	a = 89.91	a = 89.41	a = 89.46
	b = 91.81	b = 91.88	b = 91.81	b = 91.80
	c = 260.04	c = 261.17	c = 260.44	c = 259.70
Observations	565 453	454 107	382 803	332 212
Reflections	84 338	48 834	75 563	48 882

Data in parentheses are those of the highest resolution shell.

water) added drop by drop while stirring on ice (9). The volume of streptomycin solution added was equal to 25% of the initial volume of the lysate. This was further stirred on ice for 15 min and centrifuged at 3000 r.c.f. for 25 min at 4°C. The supernatant from this step was subjected to ammonium sulphate precipitation by adding saturated ammonium sulphate solution, drop by drop, while stirring on ice continuously (9). The volume of the ammonium sulphate solution added was equal to 62% of the volume of the supernatant. This was stirred further on ice for 25 min and cleared by centrifugation (3000 r.c.f. for 25 min at 4°C). The pellet obtained was resuspended in GF1 buffer (25 mM HEPES pH 7.5, 150 mM NaCl, 10 mM β-mercaptoethanol, 5 mM EDTA and 0.1 mM PMSF) and applied on a Superdex 200 gelfiltration column (Pharmacia) pre-equilibrated with the same buffer. The peak corresponding to the dimer (160 kDa) was pooled and applied on a Mono-Q HR 10/10 ion-exchange column (Pharmacia) using buffers A (25 mM HEPES pH 7.5, 10 mM β-mercaptoethanol, 5 mM EDTA and 0.1 mM PMSF) and B (A with 1 M NaCl). The protein was eluted using a gradient running from 10 to 50% (buffer B) over 10 column volumes. The peak eluting between 20 and 42% of buffer B was pooled and its salt concentration was adjusted to ~150 mM by adding buffer A. This was then applied on a HiTrap Heparin HP column (Pharmacia) (4 × 5 ml), which used the same buffers A and B, as the Mono-Q column. The protein was eluted using a gradient of 16–100% (buffer B) over 9 column volumes with the protein coming off between 52 and 70%. The final purification step was a second gelfiltration using a Superdex 200 column pre-equilibrated with GF2 buffer (25 mM HEPES pH 7.5, 250 mM NaCl, 10 mM β-mercaptoethanol). The fractions corresponding to the peak were pooled and concentrated to ~14 mg/ml using a Centriprep concentrator (Millipore). Aliquots were then flash frozen in dry ice-ethanol and stored at –80°C.

DNA substrates

The two single strands of DNA purified by the reverse-phase cartridge purification method (Sigma-Genosys), were dissolved in 10 mM Tris–HCl pH 7.5, 1 mM $MgCl_2$ and annealed on a heat block. The purity of the final double stranded product was checked using a 20% native polyacrylamide gel stained with ethidium bromide. The sequence of the top strand was 5′-AGC TGC CA**M** GCA CCA GTG TCA GCG TCC TAT and that of the lower strand was 5′-ATA GGA CGC TGA CAC

TGG TGC **M**TG GCA GCT. The bold Ms indicate the positions of the mismatched nucleotides. The C:A mismatch had the C on the top and the A on the bottom strand and the G:T mismatch had the G on the top and the T on the bottom strand, respectively. The sequence of the bottom strand for the unpaired thymidine substrate was 5′-ATA GGA CGC TGA CAC TGG TGC C**T**TG GCA GCT while that of the top strand was 5′-AGC TGC CAG GCA CCA GTG TCA GCG TCC TAT. The unpaired thymidine is indicated by the bold T.

Crystallization

The MutS–DNA substrates were mixed in a ratio of 2.8 (MutS monomer) to 1 (double stranded DNA) and crystallized using the hanging drop technique. Microseeding was done to improve the crystal quality. The quality of the crystals improved further upon addition of 0.1 mM ADP to the protein–DNA mixture. All crystals grew in the same space group, from a well solution containing 11–14% PEG 6000, 350–750 mM NaCl, 10 mM $MgCl_2$ and 25 mM HEPES pH 7.5. Prior to data collection, cryobuffer (30% PEG 6000, 15% glycerol, 300 mM NaCl, 10 mM HEPES pH 7.5) was gradually added into the crystallization drop. The crystals were then removed, soaked into a drop of pure cryobuffer and frozen in liquid nitrogen.

Data collection, structure solution and refinement

All data collection was done either at the ESRF in Grenoble, France or at the EMBL outstation at DESY, Hamburg, Germany, and the data processed using the HKL suite (9) (Table 1). The structure of the MutS–G:T complex (6) was used as a model for structure solution. Unless otherwise specified, all refinement jobs were carried out using REFMAC5 (10) in the CCP4 suite (11). The waters, DNA and ligands (ADP-Mg) were first removed and rigid body refinement was carried out using the two protein monomers as rigid domains. This was followed by 20 cycles of rigid body refinement using the individual domains of the protein as rigid bodies. After this, restrained refinement was done and the first electron density maps were generated. The DNA with the corresponding mismatch and the ADP-Mg were then built into the difference density using the program O (12). Torsion angle refinement for the lower resolution structures (MutS–C:A, MutS–unpaired T) using CNS (13) and TLS refinement using REFMAC5 (14) for all the structures were performed which led to improved R_{free} values (Table 2). The domains used in

Table 2. Refinement statistics

Complex	MutS–A:A	MutS–C:A	MutS–G:G	MutS–unpaired T
Resolution range (Å)	15.0–2.4	15.0–2.9	15.0–2.6	15.0–2.9
Number of atoms	13 233	12 968	13 004	12 902
Waters	376	53	85	133
R (%)	20.5	22.4	22.7	21.6
R_{free} (%)	25.2	29.3	27.7	29.2
r.m.s.d. bonds (Å)	0.012	0.010	0.014	0.010
r.m.s.d. angles (°)	1.362	1.293	1.488	1.261

Data collection and refinement statistics for the MutS-G:T complex published by Lamers *et al.* (6).

the rigid body refinement were used as TLS groups during the TLS refinement. Waters were built into the structures using ARP/wARP (15). All the structures were refined to good stereochemistry (Table 2) with >99% of the residues in the allowed and additionally allowed and none in the disallowed regions of the Ramachandran plot. Stereochemical checks on all structures were carried out using WHATCHECK (16) and the DNAs were analyzed using the program 3DNA (17). All figures except 1C and 1D were generated using MolScript (18) and Raster3D (19).

RESULTS

Overall structures of the MutS–mismatch complexes

The overall structures of the MutS complexes are very similar to the published structure of the MutS–G:T complex (6) (Fig. 1A). The r.m.s. deviations (on Cαs) of the complexes with the MutS–G:T complex are 0.41 (MutS–C:A), 0.51 (MutS–unpaired T), 0.35 (MutS–G:G) and 0.36 Å (MutS–A:A). The DNA is held in place by the mismatch binding and clamp domains of both the monomers (Fig. 1B). Only one of the two monomers, monomer A, contacts the mismatch directly while the other, monomer B, only makes contacts to the DNA backbone. As in the original MutS–G:T structure, only 15–17 [13 in the MutS-unpaired T (Fig. 1D)] out of the 30 DNA base pairs, starting from the 5′-end of the top strand are visible. The remaining base pairs are untraceable in the density. Since the attempts to crystallize the protein in complex with a 16 bp oligo have been unsuccessful, the remaining bases could play a role in stabilizing the crystal packing or prevent alternative packing modes.

Several contacts between the protein and DNA are seen, which are generally conserved in all the structures. The protein–DNA interface in the complexes is extensive, comprising of many hydrogen bonds, salt bridges and Van der Waals interactions (Fig. 1B, C and D). The surface area of the protein–DNA interface is ~1850 Å^2. The mismatch binding domain (residues 1–115) of monomer A accounts for more than half of this area (970 Å^2) with several residues from it forming both hydrophobic and hydrophilic contacts to the DNA (Fig. 1B, C and D). The other three domains in contact to the DNA have much smaller interfaces. They are the clamp domain (residues 450–512) of monomer B (525 Å^2), clamp domain of monomer A (285 Å^2) and the mismatch binding domain of monomer B (105 Å^2). The contacts from these domains are predominantly hydrophilic (Fig. 1C and D).

Mismatch binding by MutS

The most striking feature of the DNA in all the complexes is a sharp kink of ~60° with the mismatched bases located at the vertex of the kink (Fig. 1B). This kinking causes a widening of the minor groove around the mismatch. The distance between the backbone phosphates (P–P distance) of the mismatched base pairs increases to 21–22 Å from an average of 11–12 Å for the rest of the base pairs in the minor groove. Mismatch binding by MutS involves the stacking of a phenylalanine residue, Phe 36 of one of the monomers, onto one of the mismatched bases. The same base is reoriented such that a particular nitrogen on it is brought into proximity to the glutamate, Glu 38. This enables the formation of a hydrogen bond between the carbonyl oxygen (OE2) of the glutamate and the nitrogen of the base. In the structures of the MutS–G:T and the MutS–unpaired T complexes, Phe 36 stacks over the thymidines with their N3s forming the hydrogen bonds to Glu 38 (Figs 2B and 3A). In the structures of the MutS–C:A and MutS–A:A complexes, Phe 36 stacks on the adenosines and their N7s form hydrogen bonds to Glu 38 (Fig. 2D and H). The same is seen in the MutS–G:G complex (Fig. 2F) where the N7 of the guanosine is seen in this conformation. The purine bases on which Phe 36 stacks are in the *syn* orientation in contrast to the thymidines in the G:T mismatch and the unpaired thymidine complex which are in the *anti* orientation of the glycosyl bond.

In the structure of the MutS–unpaired T, more severe unstacking and disruptions in the base pairs adjacent to the unpaired thymidine, Thy 22, are seen (Fig. 3A and B). Phe 36 stacks on the unpaired Thy 22, which is seen forming a G:T base pair with the Gua 9 with significant rearrangements taking place in the Gua 9: Cyt 21 and Gua 10: Cyt 20 Watson–Crick base pairs. The contacts between the protein and the DNA are preserved (Fig. 1C and D) and are identical to those of the MutS–mismatch complexes. A small difference is the conformation of the loop between Ala 60 and Gly 63, which causes a change in the orientation of the side chain of Arg 58 (Fig. 1C and D). Since this loop is not visible in the other monomer due to lack of density, it seems to be highly mobile and takes on a different conformation in each of the structures (Fig. 2G).

DISCUSSION

Mismatch recognition by MutS and the effect of mismatches in DNA have been widely studied, biochemically and structurally. In our structures, MutS binds to DNA containing single G:T, A:A, C:A, G:G mismatches and an unpaired
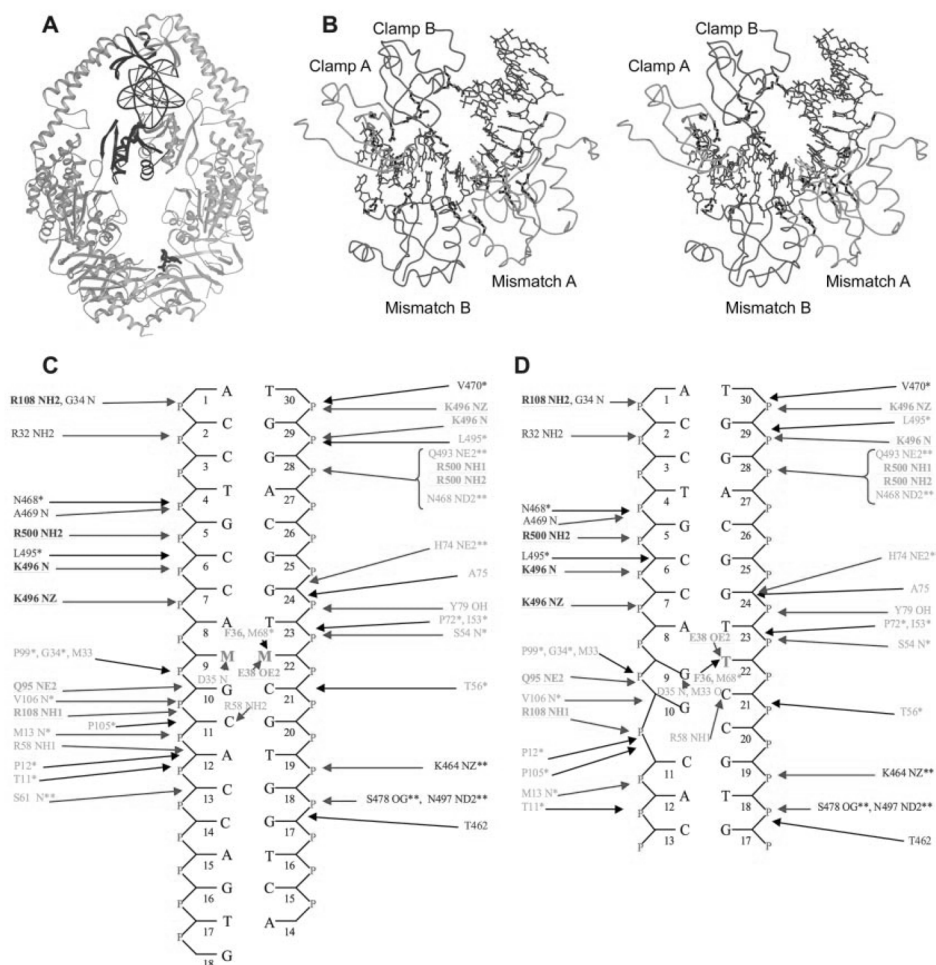
Figure 1. DNA binding by MutS. (**A**) View of the MutS–DNA complex showing the DNA and ADP in red. The mismatch binding monomer (A) is coloured green and the other (monomer B) is coloured blue. (**B**) Stereo view of the protein–DNA interaction interface. The residues forming hydrogen bonds are coloured black. The mismatched bases are coloured yellow. (**C**) Schematic representation of the interactions between the mismatched DNAs and *E.coli* MutS. Residues from monomer A are shown in green and those from monomer B in blue. The bases marked M:M indicate the mismatches. Residues conserved in and making contacts to the DNA in *Taq* MutS (5) and also conserved in the eukaryotic homologs are indicated in bold and underlined. The residues conserved only in *E.coli* and *Taq* MutS and interacting with the DNA in the same way are shown with a single asterisk. Residues conserved in *E.coli* and *Taq* MutS but interacting with the DNA in a different way are shown with a double asterisk. Hydrogen bonds/salt bridges are shown with red arrows and Van der Waals interactions with black arrows. (**D**) Schematic representation of the interactions between the *E.coli* MutS–unpaired T.

thymidine in a very similar way. Most protein–DNA interactions are conserved (Fig. 1C and D) among the complexes. Although the packing of these molecules is similar, it is clear that rearrangements of loops and side chains would be possible. In fact there is substantial rearrangement of the DNA in the complex with the unpaired thymidine. However, the interface between the DNA and protein is remarkably similar in all the five structures of the *E.coli* MutS, while the
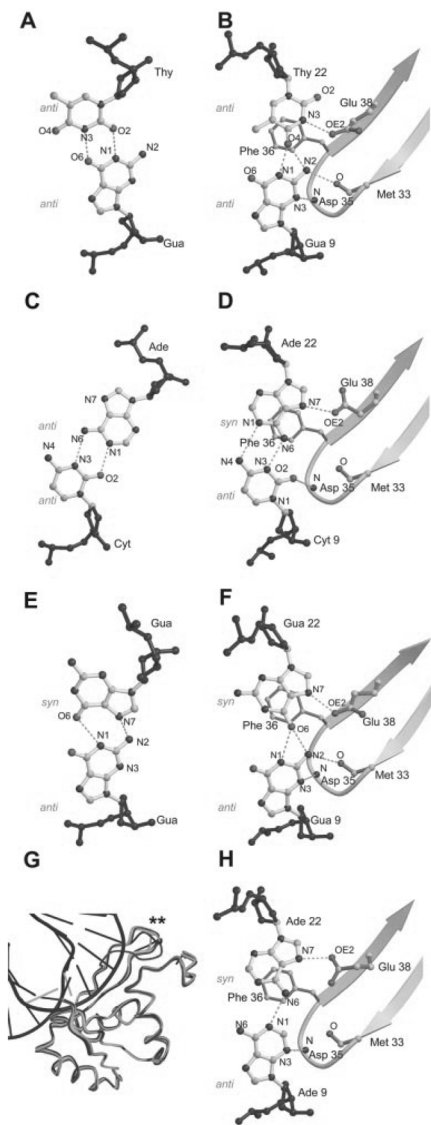
Figure 2. Comparisons between mismatches bound by MutS and those seen in oligos without the protein. (**A**) Base pairing in the unbound G:T mismatch (21). (**B**) G:T mismatch bound to MutS (6). (**C**) Base pairing in the unbound C:A mismatch (20). (**D**) C:A mismatch bound to MutS. (**E**) Base pairing in the unbound G:G mismatch (22). (**F**) G:G mismatch bound to MutS. (**G**) Superposition of the mismatch binding domains of all the structures. The variation in the loop between Ala 60 and Gly 63 is indicated by the two asterisks. (**H**) A:A mismatch bound to MutS.

major features are also very well conserved in the *Taq* MutS structure (5), indicating a common mismatch binding mode by MutS. As the crystal structures of unbound oligos containing G:T, C:A and G:G base pairs (20–22) show no kinking of the DNA, it is clear that this occurs only upon MutS binding. A similar kink in the DNA has also been seen in the structure of *Taq* MutS in complex with an unpaired thymidine. This seems to be an important requirement as a straight piece of DNA would lead to severe Van der Waals clashes with the mismatch binding domains (Fig. 1B). Phe 36, which is widely conserved in the MutS family, is seen stacking on to one of the mismatched bases. It has been shown that mutating this phenylalanine to an alanine eliminates both DNA binding and MMR by MutS (23).

Comparison of the mismatches bound to MutS to those in crystal structures of free oligos (Fig. 2A and B, C and D, E and F) (20–22) shows rearrangements in the base pairing upon protein binding. This rearrangement exposes either the N3 of the thymidine [in the MutS–G:T and MutS–unpaired T structures (Figs 2A and 3A)] or the N7 of the adenosine/ guanosine [in the MutS–C:A, MutS–A:A and the MutS–G:G structures (Fig. 2D, H and F)] to Glu 38 for hydrogen bonding. The widening of the minor groove upon kinking of the DNA probably gives the protein enough room to reorient these bases to achieve this. In the MutS–G:T structure, the protein only has to shift the thymidine from its unbound position (Fig. 2A and B) to expose the N3 to Glu 38. In the MutS–G:G structure, a similar rearrangement of the *syn* guanosine (Fig. 2E and F) is enough to expose the N7 to Glu 38. In contrast, in the MutS– C:A structure, the adenosine is rotated around its C1′–N9 bond, from its *anti* orientation in the unbound state (Fig. 2C) to the *syn* orientation (Fig. 2D). In the MutS–unpaired T structure (Fig. 3A and B) and in the *Taq* MutS–unpaired T complex (5) (Fig. 3C and D) the N3s of the thymidine stacking on to Phe 36/Phe 39 form this hydrogen bond. These data suggest a scenario where the stacking of Phe 36 on any pyrimidine would lead to the N3 forming a hydrogen bond to Glu 38 while the stacking on a purine would involve its N7 forming the same hydrogen bond.

Glu 38 is a widely conserved residue in the MutS family of proteins. Besides forming the hydrogen bond to the base stacked upon by Phe 36, the role played by Glu 38 in mismatch recognition remains unclear. It has been shown that mutating Glu 38 to an alanine destroys MMR activity in MutS and increases the affinity of the protein towards homoduplex DNA (24). The requirement of a hydrogen bonding donor/acceptor for this residue in the base stacking on Phe 36 has also been demonstrated (24). Removal of the N3 of the thymidine by replacing it with difluorotoluene, which lacks the N3, leads to an 8-fold decrease in mismatch binding affinity by MutS (24). Replacement of the adenosine with 4-methylbenzimidazole, which lacks the N6, N1 and N3, also shows a similar effect. Although in the MutS–C:A and MutS–A:A structures, the N7s of the adenosines form hydrogen bonds with Glu 38, the N3s, N6s and N1 are involved in stabilizing the complex by forming base pairing hydrogen bonds (Fig. 2D and H). Thus, the disruption of any of these sites can affect the complex formation with MutS.

The purine N7–Glu 38 (OE2) hydrogen bond in the MutS– A:A, MutS–C:A and MutS–G:G structures is unexpected since neither of the atoms involved is protonated. Therefore, either
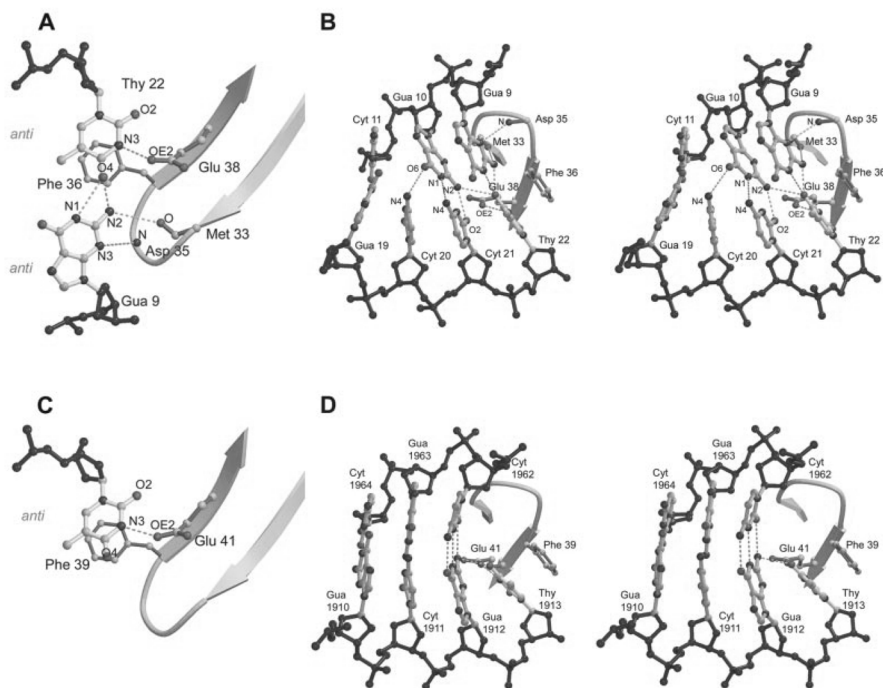
Figure 3. Comparison of the *E.coli* MutS–unpaired T complex with *Taq* MutS. (**A**) View of the protein–mismatch interaction site in the MutS–unpaired T complex. The guanosine forms a G:T mismatch with the unpaired thymidine. (**B**) Side stereoview of the MutS–unpaired T complex showing disruption of the Watson–Crick base pairs downstream to the unpaired T (shown in yellow). (**C** and **D**) Front view and side stereoview of the protein–mismatch interaction site in the *Taq* MutS structure (5).

the purines are in a tautomeric form where their N7s are protonated or the OE2 of Glu 38 must be protonated. While buried and protonated acids have been seen in crystal structures (25–27), it is unclear how the required pKa shift would occur here. Recently, Junop *et al.* (28) have shown that mutating Glu 38 into a glutamine improves homoduplex DNA binding relative to mismatch binding, thereby eliminating MMR completely. This glutamine would be able to make the same hydrogen bond and so the role of this residue in mismatch recognition seems complex. It has been suggested that the acidity of the glutamate plays a role in kinking the DNA during mismatch recognition (28). More evidence on the protonation of Glu 38 or the tautomerization of the purines awaits biochemical testing.

The extensive contacts between the protein and DNA play an important role in the stabilization of the protein–DNA complexes as it has been seen that the mismatch binding and the clamp domains are disordered in the absence of the DNA (5). An interesting observation is the involvement in DNA binding, of many other residues widely conserved in the MutS family besides Phe 36 and Glu 38. In the mismatch binding domain, Arg 108 and Gln 95 side chains seem to be important as they are not only conserved in *Taq* MutS, where they form hydrogen bonds to the DNA (5), but also in eukaryotic MSH3 and MSH6. Since the eukaryotic MSH2–MSH6 complex is involved in the recognition of mismatches and short IDLs and MSH2–MSH3 recognizes longer IDLs (1,2) these residues could play a role in mismatch recognition. Conserved residues are also seen making contacts to the DNA in the clamp domain. Of these, Lys 496 and Arg 500 are conserved in *Taq* MutS where they are involved in hydrogen bond formation to the DNA. Lys 496 is also conserved in eukaryotic MSH2, MSH3 and MSH6 and so may play an important role in DNA binding, while Arg 500, conserved in MSH3 and MSH6, may play some role in mismatch recognition.

Although several crystallographic studies have shown that the presence of a mismatch in DNA does not change its structure dramatically (20–22,29), mismatches destabilize DNA. This can be seen by the reduction in melting temperatures of DNA upon incorporation of a mismatch (30,31). A mismatch binding enzyme like MutS could be making use of this local weakening to detect the presence of mismatches and

unpaired bases. The extensive DNA–protein interface in our complexes is suggestive of a mechanism in which regions of DNA ~13–14 bp are scanned for the presence of a mismatch. That MutS binds homoduplex DNA with low affinity (24,32,33) and that it remains localized on the chromosomes in cells (34) suggests that it stays on DNA all the time, constantly scanning for mismatches.

A comparison of MutS–mismatch complexes with the MutS–unpaired T complex reveals only a few differences. The protein–DNA contacts are largely the same, with the exception of Arg 58 of the mismatch binding monomer adopting a different conformation in the MutS–unpaired T complex (Figs 1C, D and 2G). The main difference in DNA binding between the *Taq* MutS–unpaired T (Fig. 3C and D) and our *E.coli* MutS–unpaired T complex is the significant rearrangement in the Watson–Crick base pairs adjacent to the unpaired T in the *E.coli* MutS complex (Fig. 3A and B). In fact, the *E.coli* enzyme appears to recognize this as a G:T base pair with the unpaired T seemingly base paired to Gua 9. An interesting parallel in the difference between two similar enzymes binding the same substrate is seen in the structures of two methyltransferases in complex with DNA. While the structure of HaeIII methyltransferase-DNA complex (35) shows similar rearrangements of adjacent Watson–Crick base pairs upon recognition of the target cytosine, that of the HhaI methyltransferase–DNA complex (36) does not. Apparently such rearrangements are possible but not essential features of substrate binding by these enzymes. However, such rearrangements in DNA have so far only been observed in G:C base pairs (35,37). In both the HaeIII methyltransferase (35), and in our *E.coli* MutS–unpaired thymidine structures, the rearrangement involves a G:C base pair adjacent to the base being recognized (Fig. 3B). The HaeIII methyltransferase has only one G:C base pair being rearranged while our MutS structure has two (Fig. 3B). An explanation for the rearrangements occurring in G:C base pairs could be that there are more possibilities for creating new hydrogen bonds compared to A:T base pairs. So the energetically unfavourable rearrangement of the base pairs by the protein is at least partially compensated by the formation of new stabilizing hydrogen bonds. This can be seen in our MutS–unpaired T structure where the rearranged Gua 9:Cyt 21 and the Gua 10:Cyt 20 bp form an extensive network of hydrogen bonds, also involving the protein and the unpaired thymidine (Fig. 3B). Further, since mismatch binding by MutS is also known to be influenced by sequence context (32,38) the involvement of neighbouring Watson–Crick base pairs is significant. It suggests that the protein, in addition to kinking the DNA and rearranging the base pairs of the mismatch itself, may use the rearrangement of the adjacent base pairs in order to obtain the common binding mode.

CONCLUSION

We have shown that MutS binds to A:A, C:A and G:G mismatches by stacking Phe 36 over the purine and either keeping it or bringing it into the *syn* orientation to expose the N7 to Glu 38 for hydrogen bonding. In the G:T and unpaired T, MutS binds in such a way that the N3 of the thymidine forms this hydrogen bond. We have also shown that MutS rearranges the mismatched base pairs from their positions in

unbound DNA to achieve this. This is indicative of a common mismatch binding mode for all mismatches.

In our structures, we see the protein interacting with 12 DNA base pairs other than the mismatch itself. This is suggestive of an ability of MutS to scan such regions of DNA, looking for mismatches. Thus, MutS uses the local weakening due to the mismatch to locate it and binds to it by rearranging the base pairs to the conformation defined by the common mismatch binding mode.

REFERENCES

1. Buermeyer,A.B., Deschenes,S.M., Baker,S.M. and Liskay,R.M. (1999) Mammalian DNA mismatch repair. *Annu. Rev. Genet.*, **33**, 533–564.
2. Kolodner,R.D. and Marsischky,G.T. (1999) Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.*, **9**, 89–96.
3. Modrich,P. and Lahue,R. (1996) Mismatch repair in replication fidelity, genetic recombination and cancer biology. *Annu. Rev. Biochem.*, **65**, 101–133.
4. Lynch,H.T. and de la Chapelle,A. (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.*, **36**, 801–818.
5. Obmolova,G., Ban,C., Hsieh,P. and Yang,W. (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703–710.
6. Lamers,M.H., Perrakis,A., Enzlin,J.H., Winterwerp,H.H., de Wind,N. and Sixma,T.K. (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*, **407**, 711–717.
7. Wu,T.H. and Marinus,M.G. (1999) Deletion mutation analysis of the mutS gene in *Escherichia coli*. *J. Biol. Chem.*, **274**, 5948–5952.
8. Budisa,N., Steipe,B., Demange,P., Eckerskorn,C., Kellermann,J. and Huber,R. (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur. J. Biochem.*, **230**, 788–796.
9. Otwinowski,Z. and Minor,W. (1997) Processing of x-ray data collected in oscillation mode. In Carter,C.W. and Sweet,R.M. (eds), *Methods in Enzymology, Macromolecular Crystallography Part A*. Academic Press, New York, Vol. 276, pp. 307–326.
10. Murshudov,G.N., Vagin,A.A., and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**, 240–255.
11. CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
12. Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.
13. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S., Read,R.J., Rice,L.M., Simonson,T. and Warren,G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
14. Winn,M.D., Isupov,M.N. and Murshudov,G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D*, **57**, 122–133.
15. Lamzin,V.S. and Wilson,K.S. (1993) Automated refinement of protein models. *Acta Crystallogr. D*, **49**, 129–149.

16. Hooft,R.W., Vriend,G., Sander,C. and Abola,E.E. (1996) Errors in protein structures. *Nature*, **381**, 272.
17. Lu,X.J., Shakked,Z. and Olson,W.K. (2000) A-DNA conformational motifs in ligand-bound double helices. *J. Mol. Biol.*, **300**, 819–840.
18. Kraulis,P.J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
19. Merritt,E.A. and Bacon,D.J. (1997) Raster3D photorealistic molecular graphics. In Carter,C.W. and Sweet,R.M. (eds), *Methods in Enzymology, Macromolecular Crystallography Part B*. Academic Press, New York, Vol. 277, pp. 505–524.
20. Hunter,W.N., Brown,T. and Kennard,O. (1987) Structural features and hydration of a dodecamer duplex containing two C:A mismatches. *Nucleic Acids Res.*, **15**, 6589–6600.
21. Hunter,W.N., Brown,T., Kneale,G., Anand,N.N., Rabinovich,D. and Kennard,O. (1987) The structure of guanosine-thymidine mismatches in B-DNA at 2.5-A resolution. *J. Biol. Chem.*, **262**, 9962–9970.
22. Skelly,J.V., Edwards,K.J., Jenkins,T.C. and Neidle,S. (1993) Crystal structure of an oligonucleotide duplex containing G.G base pairs: influence of mispairing on DNA backbone conformation. *Proc. Natl Acad. Sci. USA*, **90**, 804–808.
23. Yamamoto,A., Schofield,M.J., Biswas,I. and Hsieh,P. (2000) Requirement for Phe36 for DNA binding and mismatch repair by *Escherichia coli* MutS protein. *Nucleic Acids Res.*, **28**, 3564–3569.
24. Schofield,M.J., Brownewell,F.E., Nayak,S., Du,C., Kool,E.T. and Hsieh,P. (2001) The Phe-X-Glu DNA binding motif of MutS. The role of hydrogen bonding in mismatch recognition. *J. Biol. Chem.*, **276**, 45505–45508.
25. Borgstahl,G.E., Williams,D.R. and Getzoff,E.D. (1995) 1.4 A structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site and chromophore. *Biochemistry*, **34**, 6278–6287.
26. Genick,U.K., Soltis,S.M., Kuhn,P., Canestrelli,I.L. and Getzoff,E.D. (1998) Structure at 0.85 A resolution of an early protein photocycle intermediate. *Nature*, **392**, 206–209.
27. Luecke,H., Schobert,B., Richter,H.T., Cartailler,J.P. and Lanyi,J.K. (1999) Structure of bacteriorhodopsin at 1.55 A resolution. *J. Mol. Biol.*, **291**, 899–911.
28. Junop,M.S., Yang,W., Funchain,P., Clendenin,W. and Miller,J.H. (2003) 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.
29. Hunter,W.N., Brown,T., Anand,N.N. and Kennard,O. (1986) Structure of an adenine-cytosine base pair in DNA and its implications for mismatch repair. *Nature*, **320**, 552–555.
30. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J.,Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A,C.C,G.G and T.T mismatches. *Biochemistry*, **38**, 3468–3477.
31. Werntges,H., Steger,G., Riesner,D. and Fritz,H.J. (1986) Mismatches in DNA double strands: thermodynamic parameters and their correlation to repair efficiencies. *Nucleic Acids Res.*, **14**, 3773–3790.
32. Brown,J., Brown,T. and Fox,K.R. (2001) Affinity of mismatch-binding protein MutS for heteroduplexes containing different mismatches. *Biochem. J.*, **354**, 627–633.
33. Su,S.S., Lahue,R.S., Au,K.G. and Modrich,P. (1988) Mispair specificity of methyl-directed DNA mismatch correction *in vitro*. *J. Biol. Chem.*, **263**, 6829–6835.
34. Smith,B.T., Grossman,A.D. and Walker,G.C. (2001) Visualization of mismatch repair in bacterial cells. *Mol. Cell*, **8**, 1197–1206.
35. Reinisch,K.M., Chen,L., Verdine,G.L. and Lipscomb,W.N. (1995) The crystal structure of HaeIII methyltransferase convalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell*, **82**, 143–153.
36. Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.
37. Timsit,Y., Vilbois,E. and Moras,D. (1991) Base-pairing shift in the major groove of (CA)n tracts by B-DNA crystal structures. *Nature*, **354**, 167–170.
38. Joshi,A. and Rao,B.J. (2001) MutS recognition: multiple mismatches and sequence context effects. *J. Biosci.*, **26**, 595–606.

Chapter 3

A magnesium free intermediate in the ATPase cycle of MutS revealed by x ray crystallography

Manuscript submitted for publication

THE ROLE OF MAGNESIUM IN THE ATPASE CYCLE OF DNA MISMATCH REPAIR ENZYME MUTS

Ganesh Natrajan[1], Joyce H G Lebbink, Alexander Fish, Herrie H K Winterwerp and Titia K Sixma.
From the division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands
Running title: Structure of Mg free MutS
Address correspondence to: Titia K. Sixma, Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Plesmanlaan 1212, 1066 CX, Amsterdam, The Netherlands. Tel. +31-20-5121959; Fax. +31-20-5121954; E-Mail: t.sixma@nki.nl.

DNA mismatch repair in *Escherichia coli* is initiated when the ABC ATPase MutS binds to a mismatch. Subsequent exchange of ADP for ATP in the ATPase domains leads to the formation of a MutS ATP state, which initiates downstream repair. MutS, a homodimeric protein with asymmetric ATPase domains, achieves this ATP state by mismatch-dependent inhibition of ATP hydrolysis in at least one of the two nucleotide binding sites. Here we report a previously unobserved state of MutS in a novel crystal form. This structure contains a complex of MutS bound to an A:A mismatch and confirms our previous conclusion that the DNA adapts itself to the protein. In this crystal form, an ADP is bound to the mismatch-binding monomer A, but magnesium is absent, despite the fact that it was present during crystallization. We show that magnesium is essential for the DNA-dependent ADP-ATP exchange. In the structure, the loss of magnesium leads to ordering of the signature loop of monomer B, which otherwise only occurs in the presence of ATP. In analogy to G-proteins, where removal of magnesium is an important step in the signalling cycle, we propose that this ADP-bound, Mg-free structure represents an intermediate in the ATPase cycle of MutS.

INTRODUCTION

DNA mismatch repair (MMR) is an essential DNA repair mechanism used by organisms to guard their genomes from mutations. The fundamental mechanisms of MMR are widely conserved amongst different species and have been the subject of a number of reviews (1-4). Mismatch repair is initiated when the protein MutS or any of its homologs binds to mismatches or insertion-deletion loops. This is followed by a complex repair process,

involving ATP binding by MutS, formation of a complex between MutS and MutL (or any of its homologs), the identification of a strand discrimination signal, leading to the removal of the mismatch containing daughter strand and its resynthesis.

Crystal structures of *E. coli* MutS in complex with DNA mismatches and an unpaired thymidine (5,6), and that of *Thermus aquaticus* MutS (7), have shown the essential features of mismatch binding and the location of the ATPase sites (Fig 1A, B, D). The ATP binding site is found in each monomer, but a so-called signature loop from the opposing monomer forms part of the active site, and undergoes a disorder-to-order transition upon ATP binding (Fig 1D). This type of composite ATP binding site is a hallmark of the ABC family of ATPases. Upon binding a mismatch, MutS binds to ATP and this causes it to switch from a state in which it can bind mismatches, to one in which it can recruit other downstream factors in the repair cascade (8,9). MutS is an asymmetric ATPase, as its two ATP binding sites have different binding affinities and hydrolysis rates (10,11, Antony, 2004 #95). The asymmetry is maintained by a single arginine (Arg 697) at the interface between the nucleotide binding domains in *E. coli* MutS (12). The binding of ATP following mismatch binding, results in the formation of a relatively stable ATP state on the DNA mismatch, which is able to form the so-called 'sliding clamp' (9,13,14). This then forms a complex with MutL and initiates other downstream events (9). This additional ATP dependent verification of mismatch binding also results in direct dissociation of MutS from homoduplex DNA not in need of repair (14-16).

Complexes of both *Taq* and *E. coli* MutS have been resolved with different nucleotides in the two monomers, A and B, where the monomer A binds the mismatch. *Taq* MutS structures have been reported either empty in both monomers (7), with ADP-Mg in both (15), or with ADP-BeF in both monomers (17). Most *E. coli* MutS structures have an ADP-Mg in monomer A alone (5,6). All these structures show no significant changes in the conformation of the mismatch binding domains, which should be the important defining feature of any sliding clamp. Attempts to crystallize MutS in its ATP state have not been successful. Crystals of *E. coli* MutS-ADP in complex with mismatched DNA were soaked in a solution containing ATP (12). Since the protein was trapped in a crystal, not all the conformational changes needed to form the sliding clamp were allowed. However, important changes in the nucleotide binding domains could be observed. First, ATP was seen bound to both monomers, a configuration

that has been proposed to exist after the protein has bound a DNA mismatch (18). In addition, monomer A also has a magnesium atom bound. The second important change observed upon ATP binding is the ordering of the signature loops, which are disordered in the ADP bound structure. This ordering causes the conserved serine S668 in the opposing monomer to get exposed to the gamma phosphate of the ATP. Two more conserved residues, N616 and H728 in the monomer A, also get rearranged. Mutation of these residues to alanine was shown to eliminate some essential attributes of the protein. These are the ATPase activity, DNA release and the tightening of the dimer upon ATP binding (12,15). Clearly, these movements and rearrangements of important residues are essential for the protein to achieve its 'ATP' state and represent structural elements of the protein in an important part of its ATPase cycle.

Thus, structural analysis has contributed significantly to the understanding of the MutS ATPase cycle but important states are still missing. Not only is the formation of the clamp in MutS yet unobserved, additional steps that take place in this cycle of two separate, but interlinked ATPase sites are still to be unravelled.

Here we report the structure of MutS in a new crystal form, which was obtained with a shorter DNA oligomer (5,6). In this novel crystal form we see unexpected changes in the ATPase domains, associated with loss of Mg^{2+}, but not ADP binding, suggesting that the enzyme has been trapped in a state so far unnoticed in any other MutS structure. We analyze this novel state and present evidence that magnesium plays a role in the MutS cycle at other steps than ATP hydrolysis. In analogy to the described role of magnesium in GTPases we suggest that we have trapped a novel intermediate in the MutS ATPase cycle.

EXPERIMENTAL PROCEDURES

Crystallization - ΔC800, a deletion mutant of *E. coli* MutS was expressed and purified as reported earlier (5,6). A 16 base pair truncated form of the 30 base pair DNA (5) with the A.A mismatch located at the 9th base pair from the 5' end was used for the crystallization. The sequence of the upper strand was 5'AGCTGCCAAGCACCAG while that of the lower strand was 5'CTGGTGCATGGCAGCT, with the mismatched nucleotides shown in bold. The DNA was ordered as two single strands and annealed as previously described (6). The protein was mixed with the substrate DNA in

a ratio of approx 2.5:1, and ADP was added, to equal the final concentration of the protein. The final ratio of protein:DNA:ADP in the mixture was 112:40:112. Preliminary crystals, unsuitable for crystallographic analysis, were obtained when a hanging drop experiment was done with the well solution comprising of 30% PEG 3350, 150 mM Na Citrate, 2 mM $MgCl_2$ and 100 mM bis-tris propane (pH 7.5). These crystals were crushed and used as seeds. With these, good crystals were obtained by hanging drop crystallization, with the well comprising of 18-20% PEG 3350, 80-120 mM Na Citrate, 100 mM bis-tris propane (pH 7.5) and 2-5 mM $MgCl_2$. The crystals were transferred slowly into the cryo buffer (22% PEG 3350, 20% glycerol, 150 mM Na citrate) and flash frozen in liquid nitrogen.

Data collection, structure solution and refinement - The data were collected at the ID14-EH2 beamline at the ESRF, France and processed using MOSFLM and SCALA (19) (Table 1). Molecular replacement was done using Phaser (20), and the previously reported MutS-A.A mismatch complex (6) was used as a model, after removing the ADP-Mg. Refinement was done using REFMAC (21) (Table 1), and the model built using Coot (22). TLS refinement gave a substantial improvement of the map quality. All figures were made using Pymol (DeLano Scientific LLC, http//www.pymol.org)

Structure analysis - Comparison between crystal structures made use of ESCET (23) to identify significant conformational changes. Prior to this procedure, TLS parameters were mapped back to individual B-factors using the program TLSANL from the CCP4 suite (20). Protein interface analysis made use of the PISA server at the EBI (24).

Purification of nucleotide-free MutS - Purified MutS (22 mg/ml) (5,6) was dialyzed against 75–fold excess of 25 mM hepes pH 7.5, 250 mM NaCl, 10 mM beta-mercaptoethanol, 5 mM $MgCl_2$, and 1.5 mg/ml apyrase (Sigma) for 30 minutes at room temperature and additionally, for 2 hours at 4°C. Dialysis was continued overnight at 4°C against 25 mM Hepes pH 7.5, 250 mM NaCl, 10 mM beta-mercaptoethanol with several buffer changes. The dialyzed MutS was diluted 2-fold using buffer without salt, bound to a MonoQ column and eluted using a 0-1 M NaCl gradient in 25 mM Hepes pH 7.5, 10 mM beta-mercaptoethanol. After concentration of the MutS peak fractions, the nucleotide content was determined as described (10,25).

Surface plasmon resonance - Surface plasmon resonance spectroscopy (SPR) was performed at 25°C on a BIAcore T100. Streptavidin SA sensor chips were derivatized with about 250 resonance units of biotin derivatized

41-basepair heteroduplex as indicated. MutS (40–1000 nM) in running buffer (RB; 25 mM Hepes–NaOH, pH 7.5, 150 mM NaCl, containing either 10 mM MgCl$_2$ or 1 mM EDTA) were injected across the SA chip at 30 μl/min. Saturation binding values were fit and the dissociation constant (Kd) for MutS–DNA complexes was determined according to the steady state affinity model implemented in the Biacore T100 evaluation software.

Nucleotide binding, hydrolysis and exchange- Nucleotide-free MutS (final assay concentration 250 nM) and 0-10 μM mant-ADP (Molecular Probes) were incubated in 25 mM Hepes, 150 mM NaCl containing either 10 mM EDTA or 10 mM MgCl$_2$. Binding of mant-ADP to MutS was monitored on a Fluostar optima spectrophotomoter using excitation and emission filters of 355 and 405 nm respectively. Binding curves were analysed using the Solver add-in in the Microsoft Excel package. Binding constants were obtained by fitting the equation [ADPbound] = (Bmax *[ADP])/(Kd+[ADP]) to the data. Steady-state and pre-steady state ATP hydrolysis were monitored as described (14). Nucleotide exchange was determined in the absence and presence of mismatched DNA as described (14) at increasing magnesium concentrations (0-2 mM MgCl$_2$).

RESULTS

MutS structure - We crystallized *E. coli* MutS ΔC800 in complex to a 16-mer A.A mismatched oligomer, that includes all base pairs visible in previous structures. The crystal diffracted to 3.2 Å and the structure was solved by molecular replacement. Refinement resulted in a final structure with an R-factor of 20.8% and Rfree of 27.1%, with good geometry (Table 1). The overall structure of the MutS dimer in the new space group (henceforth called 16AA9) is very similar to the original structure of MutS bound to a 30-mer heteroduplex (6) (henceforth 30AA9) (Figure 1). The mismatch binding is unchanged, as is the kinking of the DNA (Fig 1C), reinforcing our earlier observation that the DNA changes conformation to achieve a unique binding mode with the protein (6). In general mismatch binding occurs via Phe36 inserting between the bases and Glu38 making a hydrogen bond to the mismatch. In the earlier study we had found that the protein recognizes all mismatches in the same way, but we could not exclude the influence of crystal contacts in forcing MutS into this conformation. In the new space group the crystal contacts have changed

dramatically (Supplementary Fig 1) but the recognition of the DNA is unchanged.

Crystallographic statistics	
Resolution (Å)	20 – 3.2 (3.37-3.2)
Completeness (%)	99.8 (98.8)
I/σ(I)	4.1 (1.0)
Rmerge (%)	16.8 (71.7)
No of observations	245011 (34611)
Unique reflections	34220 (4933)
Multiplicity	7.2 (7.0)
Cell Parameters (Å)	a= 91.10; b= 137.90; c= 161.44
Space group	$P2_12_12_1$
Refinement data	
R (%)	20.8
R_{free} (%)	27.1
R.m.s.d (bonds) (Å)	0.006
R.m.s.d (angles) (°)	1.002
No of atoms (protein+DNA+ligands)	12726

Values in parenthesis refer to the highest resolution shell

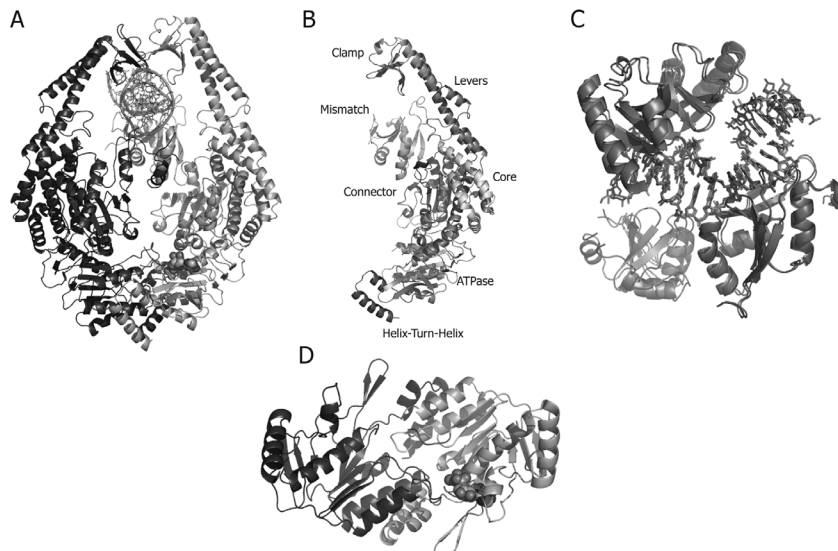**Table 1: Crystallographic data and Refinement**

Figure 1. (A) Overall structure of the MutS 16AA9 complex. The mismatch binding monomer A is shown in green and monomer B in blue. The signature loops are shown in red and the ADP in solid spheres. (B) The domain

definitions of the MutS monomer, used in the TLS refinement of the structures. (C) Superposition of mismatch binding, clamp domains and DNA of the 16AA9 (brown) and 30AA9 (cyan) structures. The mismatch is shown in red.(D) ATPase domain of the 16AA9 complex. The signature loops are shown in red. The original colour figures can be found in http://ganeshthesisfigures.org

In the new space group the mismatch binding domain of monomer B, which does not contact the mismatch directly, is disordered (Fig 1C). This has been seen before, in the structure of the *Taq* MutS in complex with a looped out thymidine (7). In the 30AA9 crystal form, this region is ordered, possibly because it is supported by crystal contacts (Supplementary Fig 1).

The most noticeable feature in the new crystal form is the absence of magnesium in the nucleotide binding site of domain A, whereas the electron density for the ADP cofactor is very clear. This is surprising as the crystallization conditions included 2-5 mM $MgCl_2$ and there is space in the structure to accommodate the magnesium. Three additional data sets in this new space-group of MutS bound to 16-base pair DNA oligomers with different base-base mismatches were also analyzed (data not shown). These data sets had low-resolution diffraction limits (>3.5 Å) or the data were twinned, but in each of these the electron density for the ADP is well resolved, whereas there is no density for magnesium. Concomitant with this absence of magnesium, there are a number of local rearrangements that are described in more detail below.

Crystal contacts - Analysis of the crystal contacts between the two crystal forms, using the PISA server (24) shows that in the 30AA9 structure, most crystal contacts are by surface residues in the monomer B, with a few on the core domain of monomer A (Supplementary Fig 1). In contrast, the larger number of crystal contacts in the 16AA9 structure are found in the ATPase domain of monomer A. Since only the surface residues are involved in the crystal contacts, the changes that we observe in the nucleotide-binding site are indirect and not directly brought about by the new crystal contacts.

Analysis of domain movements - To analyze the overall conformational changes between the new structure and known nucleotide states of MutS, we performed pairwise comparisons of the structures using the program ESCET (23). In this comparison, a B-factor dependent correction is applied to determine where two experimental crystal structures are significantly different. Conformational changes related to magnesium coordination are revealed by comparison of the Mg-bound 30AA9 (or 30GT9 (Lamers et al,

2000), which is equivalent for its ATPase domains) and Mg-free 16AA9 structures (Fig 2A, B). In both monomers, the C-terminal helix-loop-helix region (778-781) is rearranged. The major variation involves the residues of the monomer A Walker B motif (aa 692-697) that are directly involved in magnesium coordination. The second major, and unexpected, conformational difference is the ordering of the signature loop from monomer B that complements the subunit A active site (aa 659-670). This loop is generally disordered in the ADP state and it was thought that the disorder-to-order transition is an integral part of the ATPase cycle of MutS. Because signature loop ordening is related to ATP binding we compared the 16AA9 with the structure into which ATP has been soaked (30GT9-ATP) (12) (Fig 2C, D). In the ATPase domain of monomer A, the difference in bound nucleotide induces changes in the Walker B motif (residues 692-697). In monomer B, we see variations in the P-loop (residues 616-619), which is obvious as the 30GT9-ATP soak has ATP bound to this monomer while the 16AA9 is empty. This also results in the partial ordering of the signature loop of monomer A in 30GT9-ATP. The major changes involve the signature loop in monomer B, which is better ordered in the new 16AA9 structure than in the 30GT9-ATPsoak. In addition there are some significant differences in the core domains (Fig 2C). This is interesting, as this region is only about 10 Å away from the signature loop and could be involved in the transmission of signals between the ATPase and mismatch binding domains. This close approach of the ordered signature loop and the core domain has been seen before in the *Taq* MutS structure in complex with ADP-BeF (17).

 To understand the structural changes responsible for the disorder-to order transition of the signature loop, we analyzed the residues involved in more detail (Fig 3). The conserved residues that are most important for nucleotide binding and hydrolysis in *E. coli* MutS are N616, K620 and S621 on the P-loop, D693 and E694 on the Walker B motif, and H728. In the absence of any nucleotide (Fig 3D), these residues rearrange themselves to stably interact with each other. The Walker B motif shows the most dramatic rearrangement upon removal of magnesium (Fig 3A,B). Two acidic residues D693 and E694 (Fig 3B), are involved in Mg^{2+} binding. Upon removal of the magnesium in the 16AA structure, these two acidic side chains, now repelled electrostatically by the beta-phosphate of the ADP, move away from it (Fig 3A). The D693 is stabilized by a main chain hydrogen bond with G658, while the E694 side chain approaches the amine nitrogens on the side chain of the R697 of the same monomer.
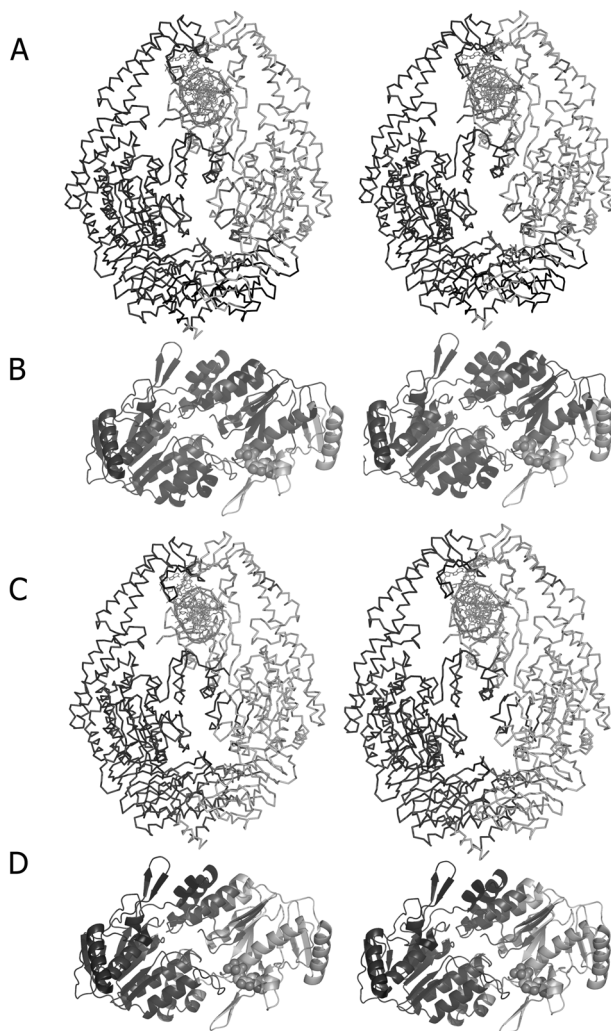
Figure 2. (A). Comparison of the 16AA9 structure with the 30AA9 structure using ESCET. The monomer A is coloured green and the monomer B is blue. The green and blue portions of these monomers represent the conformationally invariant regions between the two structures. Regions in black and brown represent regions which show movements upon comparing the two structures, but can be grouped as rigid domains. The regions in red are the ones, which do not belong to any conformationally invariant region and are the most flexible. (B). View of the ATPase domains from the top, in this same comparison. (C). Comparison between the 16AA9 structure and the 30GT9-ATPsoak. (D). View of the ATPase domains from the top in this comparison. All figures are stereoviews. For original colour figures visit http://ganeshthesisfigures.org.

Interestingly, R697 is seen in a different orientation in the 16AA9 (Fig 3A) compared to the 30AA9 (Fig 3B) or even the 30GT9-ATP soak (Fig 3C). In this reoriented state, the side chain of R697 is held stably by the glutamate D673, and is also in the vicinity of the E694. It can be envisaged that the R697 is brought into this orientation due to the attraction of the two acid side chains. The movements in the Walker B motif in the absence of magnesium therefore, lead to the creation of a void. As a consequence, the N616 moves in along with the P-loop main chain to form a hydrogen bond with H728, which is also reoriented compared to its position in the 30AA9 structure (Fig 3). This enables the signature loop from the opposing monomer to stabilize in the space created by the reoriented P-loop and the conserved serine S668 to interact with the charged phosphate tail of the ADP.

In the structure of the 30GT9-ATPsoak, the signature loop of monomer B is ordered and the S668 does make contact with the gamma phosphate of the ATP (Fig 3C). The presence of magnesium makes the side chains of D693 and E694 orient towards it, and hydrogen bond with the waters coordinating with the magnesium. The gamma phosphate of the ATP points towards the signature loop, causing the N616 to move aside from its position in the 30AA9 structure (Fig 3B) to form the hydrogen bond to the H728. The gamma phosphate interacts with the conserved serine S668 in the signature loop of the opposing monomer B thereby stabilizing it. In conclusion, the general requirement for the stabilization of the signature loop seems to be either the presence of a gamma phosphate interacting with the S668, or the removal of magnesium, which causes a reorientation of E694 and the consequent rearrangement of the N616, allowing for the interaction with the S668 on the opposing monomer's signature loop.

The role of magnesium – To study the role of magnesium in the ATPase cycle we needed to prepare MutS that has no nucleotide bound, and consequently no magnesium. In our purifications MutS has one ADP*Mg bound (5,10) Dilution and extensive dialysis alone (26) are, in our hands, not sufficient to remove this nucleotide, but upon dialysis against a solution that contains the nucleotide degrading enzyme apyrase we obtained MutS that is almost nucleotide free (less than 0.05 mol ADP/mol MutS). The fact that nucleotide removal is only possible in this way indicates that MutS does release the nucleotide, but that rebinding is very fast and can only be prevented by actively degrading the nucleotide. Nucleotide-free MutS is fully active based on similar affinity for mismatched DNA (not shown) and
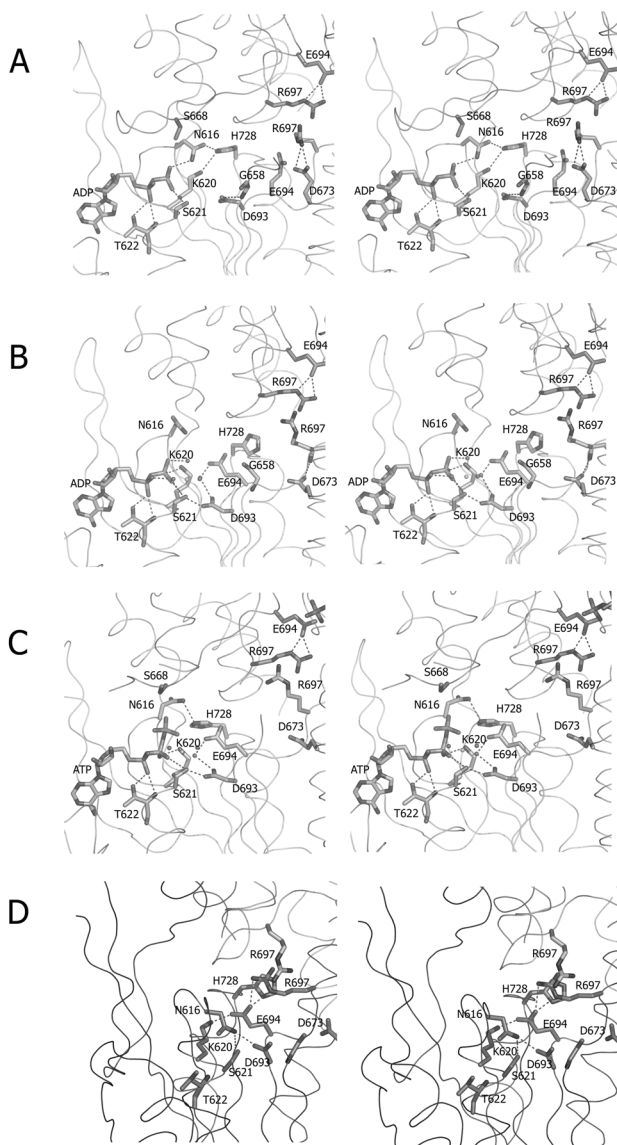
Figure 3. Details of hydrogen bonding and nucleotide binding in the ATP binding site (A) 16AA9 monomer A; (B) 30AA9 monomer A; (C) 30GT9-ATPsoak (Monomer A); (D) 30AA9 (Monomer B). In all figures, the monomer A is coloured green and monomer B, coloured blue. All figures are stereoviews. Hydrogen bonds are shown as red lines. The Magnesium is shown as a small green sphere and the coordinating waters as red ones. The original color figure can be found in http://ganeshthesisfigures.org.

similar kinetic constants for ATP hydrolysis (not shown) compared to routinely purified MutS.

Next we used this 'empty' MutS to analyse the role of magnesium in the different stages of mismatch binding and the concomitant ATPase cycle. We found that mismatch recognition is completely independent of magnesium and binding profiles of MutS to mismatch-containing DNA obtained by surface plasmon resonance are identical in the presence and absence of the metal ion (Kd for mismatched DNA 57 +/- 15 nM in the presence of magnesium versus 62 +/- 10 nM in the presence of EDTA). In addition we found that magnesium does not contribute to nucleotide binding (Kd for mantADP 0.65 +/- 0.2 µM and 0.88 +/- 0.1 µM in buffer containing Mg or EDTA, respectively).

We then analyzed the effect of magnesium on other steps of the ATPase cycle. As expected the steady state ATPase activity is completely abolished
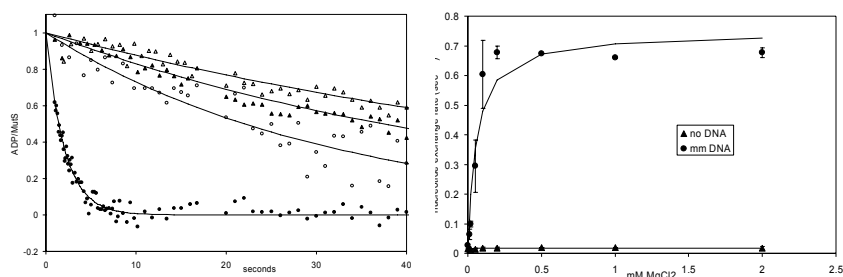

Figure 4. Nucleotide exchange in the absence and presence of mismatched DNA and its dependence on magnesium. (a) Release of fluorescent mantADP from MutS (open triangles), in the presence of 10 mM MgCl$_2$ (closed triangles), in the presence of mismatched DNA (open cirles) and in the presence of both 10 mM MgCl$_2$ and mismatched DNA (closed circles). (b) Nucleotide exchange rate in the absence and presence of mismatched DNA as a function of magnesium concentration.

in buffer containing 10 mM EDTA. This is also true for the pre-steady state burst ATP hydrolysis, indicating that the chemical hydrolysis of ATP itself depends on the presence of magnesium.

However, this is not the only requirement for magnesium. We analyzed the mismatch-dependent nucleotide exchange in a fluorescence assay, based on monitoring the release of fluorescent mantADP upon binding of ATP. As expected (27), this process is dependent on the presence of mismatched

DNA (Fig 4), indicating that mismatched DNA acts as an exchange factor here. However, we also observe that there is an absolute dependence on the presence of the metal ion for the ADP/ATP exchange to occur (Fig 4). Even in the presence of mismatched DNA, exchange is slow in the absence of magnesium. These results indicate that magnesium has to be bound by MutS during the exchange reaction, or that one of the individual nucleotide binding or release steps is coupled with binding or release of the metal ion.

DISCUSSION

MutS is a molecular switch which uses mismatch and ATP binding to switch between a state where it can recognize mismatches and another one in which it initiates repair (8). During this process, it undergoes a complex ATPase cycle. Steps in this cycle involve the accelerated exchange of ADP for ATP following mismatch binding (9,27), the inhibition of the ATPase activity in one of the nucleotide binding sites (13,18,28), leading to the formation of a stable MutS-DNA-ATP state, the formation of a sliding clamp which initiates downstream repair processes, and finally, the recycling of the MutS protein back to its original state by hydrolysis of the bound ATP. While this process has been extensively studied biochemically, crystallizing MutS and determining its structure at each of these intermediate states has not been possible. Structures have been determined in various nucleotide states (5-7,15,17). These structures have created a partial understanding of the MutS ATPase cycle. It is known however, that MutS undergoes additional conformational changes during its nucleotide binding and hydrolysis cycles. Neither the full catalytic site, nor the release from DNA that has been termed 'clamp formation' has been observed in crystal structures yet.

Here we present a novel state of MutS found in a new crystal form in which an ADP is bound but no electron density can be observed for Mg^{2+} in conjunction to an ordered state of the so called signature loop. It seems unlikely that this state of the protein is the result of the shorter DNA oligomer since all contacts to the DNA are preserved in this crystal. Therefore we assume that this state is most likely trapped by the novel crystal packing. These crystal contacts are very extensive (supplementary Fig 1), and we were unable to determine which specific contacts are responsible for trapping the MutS molecule. However, since the differences that we observed between the known structure and the novel state described

here involve a number of coordinated changes in the active site, we found it of interest to study this state in more detail.

The structures of MutS solved so far have shown small movements and conformational changes in domains, which seem to suggest a coordinated propagation of the conformational changes observed upon ATP binding, i.e. signature loop ordening and tightening of the ATP dimer interface, towards the core domain of monomer B and the clamp domains of both monomers. In other ABC-ATPases like SMC proteins, the signature loop is required for the ATP induced dimerization, resulting in completion of the active site (29). In contrast, in MutS, the ordering of the signature loop is caused by a particular orientation of the gamma phosphate of the bound ATP (Fig 3), and also depends on the presence of magnesium. The ordering of the signature loop, and the associated conformational changes and movements in the core domains upon comparison of the 16AA9 structure with the MutS-ATP soak structure confirm the link between this loop and the transmission of this signal to the mismatch binding domains, which would lead to the formation of the sliding clamp. The ordering of the signature loop is associated with ATP binding (12), where the reorientations of the side chains of N616, H728 and E694 lead to the S668 being able to interact with the gamma phosphate.

In our new 16AA9 structure (Fig 3A), we find that the removal of magnesium and ordering of the signature loop correlate with a re-arrangement of the Walker B motif. This indicates that magnesium in MutS plays a structural role. Magnesium is required for ATP hydrolysis, and also for effective ADP-ATP exchange following mismatch binding (Fig 4). In the crystal structures, magnesium brings the acids D693 and E694 in the Walker B motif closer to the gamma phosphate (Fig 3). The catalysis of the phosphate cleavage probably involves the activation of one of the water molecules coordinating the magnesium (15).

Magnesium is an important cofactor that plays a key role in another set of nucleotide binding enzymes, the G-proteins. In the Rho family of monomeric GTPases, magnesium plays the role of a GDP/GTP dissociation inhibitor (30), as the addition of Mg leads to a marked decrease in GDP-GTP exchange rates and also a decrease in the catalytic efficiency of GEFs. Crystal structures of Rho-A (31) have shown that in the absence of magnesium the switch I region opens up, thereby exposing the bound nucleotide for exchange by any GEF. Results similar to these have also been reported for the Ras family of proteins (32) There is evidence (reviewed in (32)), that when the GEF binds to the GDP bound state

stabilized by magnesium, it disrupts the Mg coordination. This leads to an intermediate state where there is GDP bound without Mg, followed by GDP release, which leads to a nucleotide free state. This is followed by the uptake of GTP followed by Mg, leading to a stable GTP-Mg state. Therefore, magnesium plays a regulatory role in G-proteins, by stabilizing the GDP or the GTP bound states, and the GTPase cycle includes transient intermediates where there are GDP/GTP nucleotides bound without Mg. Comparing the MutS cycle to that of the G-protein, it can be said that the binding of DNA acts like a nucleotide exchange factor (AEF in this case), In our new 16AA MutS structure, the binding of the DNA and the absence of Mg in the ATPase site are accompanied by the signature loop becoming visible and this is a requirement for the completion of the ATP binding site. So, based on this comparison with G-proteins, we propose that this structure represents one of these transient Mg-free states, prior to ATP binding.

From our results it becomes clear that the actual role of Mg in MutS is very different from the role it plays in G-proteins. Both mismatch binding and nucleotide binding do not require the presence of the divalent metal ion. While for nucleotide binding this is unexpected based on the intimate interactions between nucleotide and magnesium as visible in the crystal structure, it is not unprecedented, as different small GTPases differ in their magnesium dependence for nucleotide binding (discussed in (30)). More importantly, we find that mismatch-induced ADP-ATP exchange requires Mg (Fig. 4), in contrast to G-proteins, where there is rapid exchange of GDP for GTP in absence of Mg (30). Since mismatch binding itself does not require magnesium, the magnesium dependence is not indirect but must reside in the nucleotide binding sites themselves. Also nucleotide binding itself does not require the metal ion, therefore the role of the magnesium is probably in relaying the correct conformational change that induces nucleotide exchange after mismatch binding. This structural role of the metal ion is noticeable in the 16AA crystal structure by the displacement of the Walker B motif. A structural role for magnesium in relaying the correct conformational change after ATP uptake towards release of the mismatch and sliding clamp formation has previously been reported for the human homolog of MutS (33).

The fact that only a subset of steps in the mismatched DNA-dependent MutS ATPase cycle displays an absolute requirement for magnesium, while other steps are independent of the divalent metal ion, implies that regulation of the progression of MutS through the different asymmetric

stages of the ATPase cycle and its stalling at specific points in the pathway could efficiently be regulated by magnesium. Our and others' observation that the MutS nucleotide binding sites can exist in magnesium-bound as well as in a magnesium-free state, indicates that the protein is in principle able to control its metal coordination. Mismatch binding by MutS results in inhibition of ATP hydrolysis in at least one of the nucleotide binding sites on the MutS dimer ((14,18)). While it is known that this inhibition is dependent on a hydrogen bond between the mismatch binding domain and the DNA (14), the structural implications of this signalling in the ATPase domain are unknown. As ATP hydrolysis requires magnesium, removal of the metal ion would be an effective means to acquire this inhibition. Similarly, we can envisage that inhibition of nucleotide exchange by magnesium removal could stall MutS in its progression through the ATPase cycle until conditions for correct mismatch-induced repair initiation are met. We therefore propose that the magnesium-free state of Muts is an intermediate in the MutS ATPase cycle that could play a role in coordinating the correct response upon mismatch binding and recruitment of downstream mismatch repair components.

FOOTNOTES

REFERENCES

1.   Marti, T. M., Kunz, C., and Fleck, O. (2002) *J Cell Physiol* **191**(1), 28-41
2.   Schofield, M. J., and Hsieh, P. (2003) *Annu Rev Microbiol* **57**, 579-608
3.   Kunkel, T. A., and Erie, D. A. (2004) *Annu Rev Biochem*
4.   Iyer, R. R., Pluciennik, A., Burdett, V., and Modrich, P. L. (2006) *Chem Rev* **106**(2), 302-323
5.   Lamers, M. H., Perrakis, A., Enzlin, J. H., Winterwerp, H. H., de Wind, N., and Sixma, T. K. (2000) *Nature* **407**(6805), 711-717.
6.   Natrajan, G., Lamers, M. H., Enzlin, J. H., Winterwerp, H. H., Perrakis, A., and Sixma, T. K. (2003) *Nucleic Acids Res* **31**(16), 4814-4821

7.   Obmolova, G., Ban, C., Hsieh, P., and Yang, W. (2000) *Nature* **407**(6805), 703-710.

8.   Gradia, S., Acharya, S., and Fishel, R. (1997) *Cell* **91**(7), 995-1005

9.   Acharya, S., Foster, P. L., Brooks, P., and Fishel, R. (2003) *Mol Cell* **12**(1), 233-246

10.   Lamers, M. H., Winterwerp, H. H., and Sixma, T. K. (2003) *EMBO J* **22**, 746-756

11.   Bjornson, K. P., and Modrich, P. (2003) *J Biol Chem* **278**(20), 18557-18562

12.   Lamers, M. H., Georgijevic, D., Lebbink, J. H., Winterwerp, H. H., Agianian, B., de Wind, N., and Sixma, T. K. (2004) *J Biol Chem* **279**(42), 43879-43885

13.   Antony, E., and Hingorani, M. M. (2003) *Biochemistry* **42**(25), 7682-7693

14.   Lebbink, J. H., Georgijevic, D., Natrajan, G., Fish, A., Winterwerp, H. H., Sixma, T. K., and de Wind, N. (2006) *Embo J* **25**(2), 409-419

15.   Junop, M. S., Obmolova, G., Rausch, K., Hsieh, P., and Yang, W. (2001) *Mol Cell* **7**(1), 1-12

16.   Selmane, T., Schofield, M. J., Nayak, S., Du, C., and Hsieh, P. (2003) *J Mol Biol* **334**(5), 949-965

17.   Alani, E., Lee, J. Y., Schofield, M. J., Kijas, A. W., Hsieh, P., and Yang, W. (2003) *J Biol Chem* **278**(18), 16088-16094

18.   Antony, E., and Hingorani, M. M. (2004) *Biochemistry* **43**(41), 13115-13128

19.   Powell, H. R. (1999) *Acta Crystallogr D Biol Crystallogr* **55 (Pt 10)**, 1690-1695

20.   CCP4. (1994) *Acta Crystallogr D* **50**, 760-763

21.   Murshudov, G. N., Vagin, A.A, & Dodson, E.J. (1997) *Acta Crystallogr D* **53**, 240-255

22.   Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1), 2126-2132

23.   Schneider, T. R. (2004) *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1), 2269-2275

24.   Krissinel, E., and Henrick, K. (eds). (2005) *Detection of Protein Assemblies in Crystals*, Springer-Verlag, Berlin Heidelberg

25.   Blackwell, L. J., Bjornson, K. P., Allen, D. J., and Modrich, P. (2001) *J Biol Chem* **276**(36), 34339-34347

26.   Baitinger, C., Burdett, V., and Modrich, P. (2003) *J Biol Chem* **278**(49), 49505-49511

27.   Gradia, S., Acharya, S., and Fishel, R. (2000) *J Biol Chem* **275**(6), 3922-3930

28.   Bjornson, K. P., Allen, D. J., and Modrich, P. (2000) *Biochemistry* **39**(11), 3176-3183

29.   Hopfner, K. P., and Tainer, J. A. (2003) *Curr Opin Struct Biol* **13**(2), 249-255

30.   Zhang, B., Zhang, Y., Wang, Z., and Zheng, Y. (2000) *J Biol Chem* **275**(33), 25299-25307

31.   Shimizu, T., Ihara, K., Maesaki, R., Kuroda, S., Kaibuchi, K., and Hakoshima, T. (2000) *J Biol Chem* **275**(24), 18311-18317

32.   Pan, J. Y., and Wessling-Resnick, M. (1998) *BioEssays* **20**, 516-521

33.   Iaccarino, I., Marra, G., Dufner, P., and Jiricny, J. (2000) *J Biol Chem* **275**(3), 2080-2086
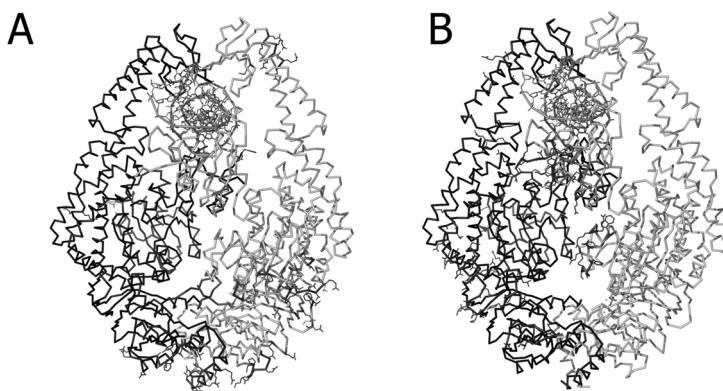
Supplementary figure 1. Crystal contacts in the (A) 16AA9 and (B) 30AA9structures. The residues involved in the crystal contacts are shown in red. For the real colour figure, visit http://ganeshthesisfigures.org.
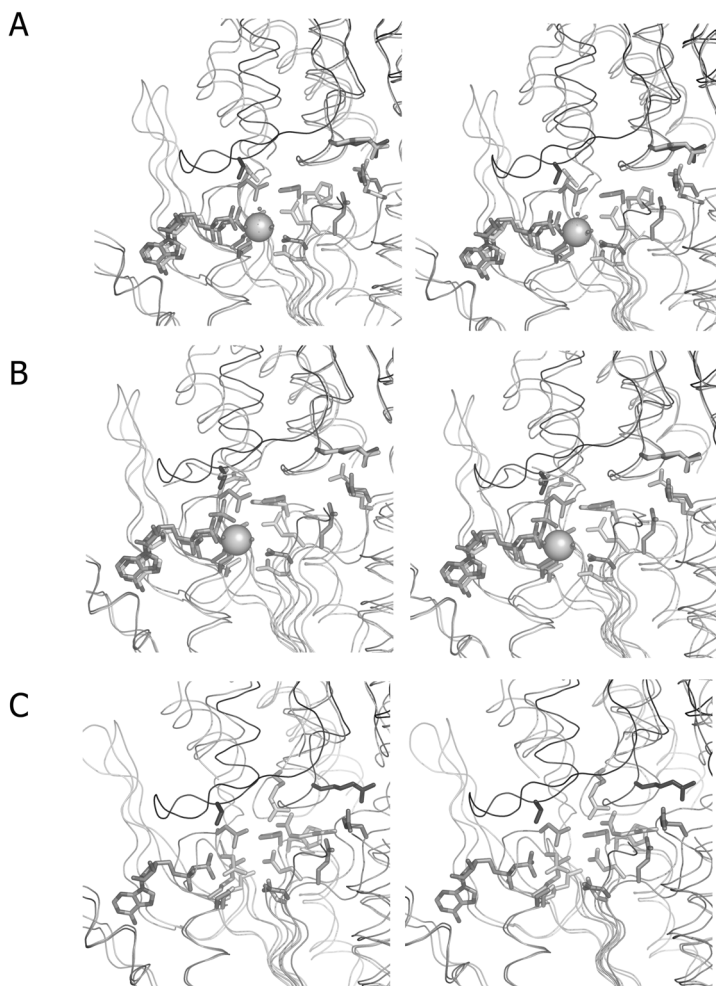
Supplemental figure 2. (A) Superposition of monomer A ATP binding sites of the 16AA9 with the 30AA9. The monomer A is shown in green (16AA9-dark; 30AA9-pale), while the monomer B is shown in blue (16AA9-dark; 30AA9-pale). **(B)** Superposition of the monomer A ATP binding site of 16AA9 with the 30GT9-ATP soak. Monomer A is shown in green (16AA9-dark; 30GT9-ATP-pale) and monomer B in blue (16AA9-dark; 30GT9-ATP-pale). All figures are stereoviews. **(C)** Superposition of ATP binding site of the monomer A of 16AA9 with the monomer B of 30AA9 (16AA9 -Dark; 30AA9-Pale). The magnesium is shown as the big solid sphere. http://ganeshthesisfigures.org

Chapter 4

Dual role of MutS glutamate 38 in DNA mismatch discrimination and authorization of repair.

Dual role of MutS glutamate 38 in DNA mismatch discrimination and in the authorization of repair

Joyce HG Lebbink[1,3], Dubravka Georgijevic[2,3], Ganesh Natrajan[1,3], Alexander Fish[1], Herrie HK Winterwerp[1], Titia K Sixma[1,*] and Niels de Wind[2,*]

[1]Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands and [2]Department of Toxicogenetics, Leiden University Medical Center, Leiden, The Netherlands

MutS plays a critical role in DNA mismatch repair in *Escherichia coli* by binding to mismatches and initiating repair in an ATP-dependent manner. Mutational analysis of a highly conserved glutamate, Glu38, has revealed its role in mismatch recognition by enabling MutS to discriminate between homoduplex and mismatched DNA. Crystal structures of MutS have shown that Glu38 forms a hydrogen bond to one of the mismatched bases. In this study, we have analyzed the crystal structures, DNA binding and the response to ATP binding of three Glu38 mutants. While confirming the role of the negative charge in initial discrimination, we show that *in vivo* mismatch repair can proceed even when discrimination is low. We demonstrate that the formation of a hydrogen bond by residue 38 to the mismatched base authorizes repair by inducing intramolecular signaling, which results in the inhibition of rapid hydrolysis of distally bound ATP. This allows formation of the stable MutS–ATP–DNA clamp, a key intermediate in triggering downstream repair events.

Introduction

DNA mismatch repair plays a crucial role in ensuring genomic stability. In bacteria, absence of functional mismatch repair results in a mutator phenotype and removal of the interspecies barrier against recombination between slightly diverged sequences (reviewed in Schofield and Hsieh, 2003). Defects in the mismatch repair cascade in humans predispose to hereditary non-polyposis colorectal cancer and are asso-

*Corresponding authors. TK Sixma, Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands. Tel.: +31 20 5121959; Fax: +31 20 5121954; E-mail: t.sixma@nki.nl or N de Wind, Department of Toxicogenetics, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands. Tel.: +31 71 5269627; Fax: +31 71 5268284; E-mail: n.de_wind@lumc.nl
[3]These authors contributed equally to this work

ciated with a variety of sporadic cancers (reviewed in Lynch and de la Chapelle, 1999; Peltomaki, 2003). DNA mismatch repair is initiated by the protein MutS (in *Escherichia coli*) or its MSH homologs (MSH2 and MSH6 that form the predominant MutSα heterodimer in humans). MutS recognizes and binds to mismatches or unpaired bases that have escaped the proofreading capacity of the DNA replication machinery or are present in the genome during recombinatorial events between non-fully complementary DNA strands. Mismatch binding triggers the uptake of ATP in the MutS nucleotide-binding domain and it is this mismatch-dependent ATP state that authorizes repair by recruitment of additional mismatch repair components (recently reviewed in Marti *et al*, 2002; Kunkel and Erie, 2005).

Molecular details of the initial recognition and binding of DNA containing different mismatched or unpaired bases have been visualized in several crystal structures of both *E. coli* and *Thermus aquaticus* MutS (Lamers *et al*, 2000; Obmolova *et al*, 2000; Natrajan *et al*, 2003). MutS is a homodimer consisting of two structurally different monomers, A and B (Figure 1A). The DNA is kinked by 60° at the mismatch-binding site (Figure 1B) and only the monomer A binds to the mismatch. In *E. coli*, this monomer also has an ADP bound to its ATPase domain. The mismatch itself is bound by two highly conserved residues. These are a phenylalanine (Phe 36 in *E. coli*), which stacks onto one of the mismatched bases, and a glutamate (Glu38 in *E. coli*), which forms a hydrogen bond to the same base (Figure 1C). Whereas mutating Phe 36 leads to elimination of both DNA binding and mismatch repair (Yamamoto *et al*, 2000; Drotschmann *et al*, 2001), the role of Glu38 has been less clear. It was proposed that the electrostatic repulsion between the negatively charged side chain of Glu38 and the DNA backbone facilitated the kinking of the DNA at the site of the mismatch (Lamers *et al*, 2000; Schofield *et al*, 2001). Electrophoretic mobility shift assays ('band-shifts') using oligonucleotides containing base analogs have shown that the hydrogen bond only contributes marginally to mismatch binding (Schofield *et al*, 2001). Mutating Glu38 to an alanine or glutamine slightly increases the affinity of the enzyme for DNA carrying a mismatch, but even more so for homoduplex DNA (Schofield *et al*, 2001; Junop *et al*, 2003). MutS carrying these mutations was unable to complement mismatch repair in a MutS-deficient strain, presumably because the mutants lost their ability to discriminate between homoduplex and mismatched DNA.

Mismatch binding results in a conformational change that is propagated toward the other end of the molecule, altering the nucleotide-binding capacity of the ATPase domains. The two ATPase domains are asymmetric in nucleotide binding and ATP hydrolysis (Bjornson *et al*, 2000; Lamers *et al*, 2003; Antony and Hingorani, 2004). In the absence of DNA, the rate-limiting step for ATPase activity is release of ADP, whereas binding of MutS to a DNA mismatch greatly enhances the rate of ADP–ATP exchange (Gradia *et al*, 1997; Blackwell *et al*, 2001; Acharya *et al*, 2003). After
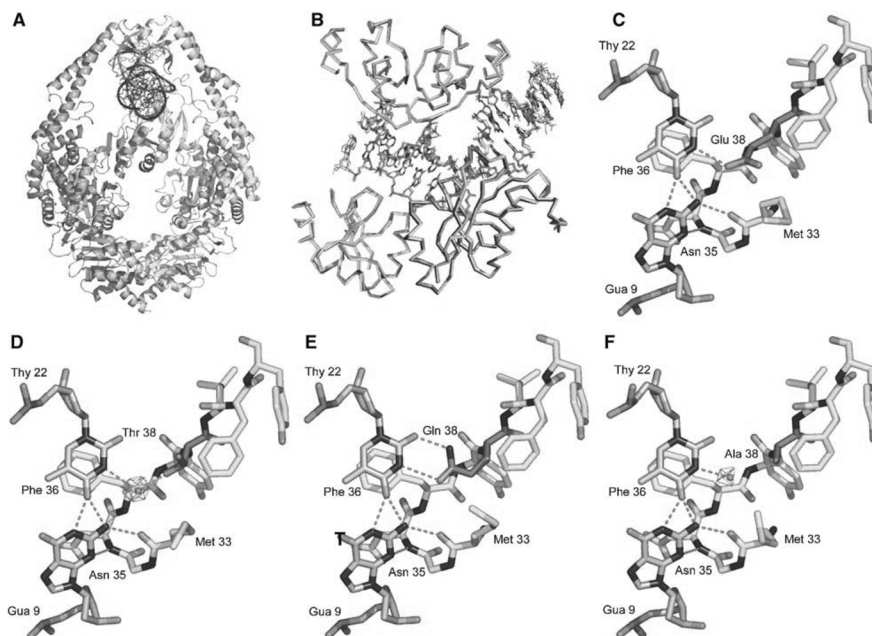
Figure 1 Crystal structures of wild-type MutS and E38 variants. (**A**) Schematic view of the *E. coli* MutS dimer, with the mismatch-binding monomer A shown in yellow, the supporting monomer B shown in orange and the mismatch containing DNA in blue. (**B**) Superposition of the mismatch-binding clamp domains and DNA of wild-type and mutant MutS obtained by superposing the Cα atoms of the mismatch-binding monomer A of all the structures. (**C**) G.T mismatch binding by the wild-type MutS. (**D**) G.T mismatch binding by the E38T mutant, with the bound water molecule in red, along with its mF_o-DF_c map contoured at 3σ (in green). (**E**) G.T mismatch binding by the E38Q mutant. (**F**) G.T mismatch binding by E38A, shown with the bound water molecule (in red) along with its mF_o-DF_c density map contoured at 3σ (in green). Dashed red lines indicate hydrogen bonds and residue 38 in panels C–F is shown in purple. All figures were made using PyMOL (Copyright© 2004 DeLano Scientific).

binding to perfectly paired DNA, this ATP is hydrolyzed quickly, but binding to mismatched DNA inhibits this fast hydrolysis. This would indicate the formation of an ATP-bound MutS state on mismatched DNA with a relatively long lifetime, a state that allows mismatch-dependent recruitment of MutL and initiation of repair (Bjornson *et al*, 2000; Antony and Hingorani, 2003, 2004). Meanwhile, ATP reduces affinity of MutS for the mismatch itself and induces conversion of the protein into a sliding clamp that can diffuse along the DNA helix (Gradia *et al*, 1999; Acharya *et al*, 2003). In contrast, ATP binding to MutS induces direct dissociation of the protein from homoduplex DNA (Iaccarino *et al*, 2000; Selmane *et al*, 2003). Specific inhibition of ATP hydrolysis by mismatched DNA binding, and the different modes of dissociation from homo- and heteroduplex DNA indicate that MutS uses ATP to verify mismatch binding and initiate repair, as proposed by Junop *et al* (2001). This may explain the high efficiency of the DNA mismatch repair process even though initial discrimination between homoduplex and mismatch containing DNA by *E. coli* MutS is only 8- to 20-fold (Schofield *et al*, 2001; Hays *et al*, 2005).

The nucleotide-dependent conformational changes and their effect on DNA binding indicate that residues that are in contact with the DNA are not only involved in initial recognition of mismatches, but also contribute to proper reciprocal signaling between DNA- and ATP-binding domains. To better understand this process, we investigated the role of glutamate 38 in more detail, using a combination of *in vivo*, crystallographic, biophysical and biochemical approaches. Our data demonstrate that the Glu38 side chain of MutS plays a dual role in initial mismatch discrimination, by its negative charge, as well as in transmitting conformational changes that induce formation of a stable MutS-ATP state and sliding clamp, by its ability to specifically form a hydrogen bond with the mismatched base.

Results

Mutants of E38 in E. coli MutS

The conserved glutamate at position 38 (E38) in *E. coli* MutS forms a hydrogen bond to one of the mismatched bases and is thought to participate in the phenylalanine 36 (F36)-induced kinking of the DNA at the mismatch site, by causing electrostatic repulsion between its negatively charged side chain and the phosphates of the DNA (Lamers *et al*, 2000; Obmolova *et al*, 2000; Schofield *et al*, 2001; Natrajan *et al*, 2003). To

address the individual roles of the negative charge of the carboxylate side chain and of hydrogen bonding to the mismatched base, we have replaced this glutamate with a glutamine, a threonine and an alanine residue. Mutations were introduced in the full-length MutS protein for genetic and biochemical analysis, as well as in the ΔC800 protein, which lacks the 53 C-terminal amino acids (Lamers *et al*, 2000), for crystallographic analysis. All mutant side chains lack the negative charge but they differ in size and hydrophilicity. The mutants were efficiently overexpressed and purified. All have the same molecular weight according to size-exclusion chromatography, indicating identical quaternary structure (data not shown).

Mismatch repair in vivo

To analyze the mismatch repair capabilities of the E38 mutants, we performed *in vivo* mismatch repair assays, based on complementation of repair in an *E. coli* strain lacking MutS (Wu and Marinus, 1994). These assays were performed at 37 and 22°C, the temperatures at which most subsequent *in vitro* assays were performed (Table I). At both temperatures, the E38A substitution is highly impaired in its ability to repair mismatches, as has been reported for this mutant in an earlier study (Schofield *et al*, 2001). The E38Q mutant has also been reported to be deficient in mismatch correction (Junop *et al*, 2003), and this correlates with its inability to complement *MutS* deficiency at 37°C. However, at room temperature, E38Q appears to be almost as efficient in

mismatch repair as wild-type MutS. Unexpectedly, the threonine mutant is proficient in repair at both temperatures. Thus, mismatch repair can proceed in the absence of a negatively charged residue at MutS position 38.

Crystal structures of E38 mutants bound to mismatched DNA

To investigate the molecular details of mismatch recognition by the MutS E38 mutants, we solved their crystal structures in complex with DNA containing a G.T mismatch (Table II). The overall structures of the mutants are very similar to that of the wild-type protein (Figure 1A and B), with an r.m.s.d. (Cα) of 0.27 Å between the wild type and the E38T, E38A mutants and 0.30 Å between the wild type and the E38Q mutant. The only differences are seen in the binding of the mismatch itself by the A subunit; no significant changes are seen in the other mutation site in the B subunit. In the mismatch-binding site, F36 from the A subunit stacks upon the thymidine in all of the structures. As in wild-type MutS, a hydrogen bond is formed by the thymidine to residue 38 of the A subunit in the E38T and the E38Q mutants, but in a different way in each mutant (Figure 1D and E). The E38T mutant has a water molecule bridging the OG of the threonine and the N3 of the mismatched thymidine base (Figure 1D). The glutamine in the E38Q mutant adopts an alternative side-chain conformation similar to the glutamate binding to an A.A mismatch (Natrajan *et al*, 2003), and now forms two hydrogen bonds to the thymidine, using its OE1 and NE2 (Figure 1E). In contrast, the E38A mutant cannot form a hydrogen bond to the thymidine (Figure 1F). There is weak electron density between the alanine side chain and the thymidine, suggesting only partial occupancy of this site by a water molecule. This water molecule is hydrogen-bonded to the N3 of the thymidine and is located at a distance of 3.5 Å from the alanine side chain, suggesting a Van der Waals contact between the two. This, therefore, is in contrast with the two other mutants and the wild-type protein, where at least one direct, or water-mediated, hydrogen bond is formed between the side chain of residue 38 and the thymidine. A single hydrophobic contact of glutamate 38 with methionine 33 is lost in all mutants and is therefore not relevant for functional mismatch

Table I *In vivo* mismatch repair

	20°C		37°C	
	Mutation frequency ($\times 10^{-9}$)	Mutation rate ($\times 10^{-9}$)	Mutation frequency ($\times 10^{-9}$)	Mutation rate ($\times 10^{-9}$)
WT MutS	6 ± 1	4.4	5 ± 1	3.8
E38A	68 ± 15	26.9	822 ± 455	212
E38T	13 ± 2	7.6	15 ± 4	8.4
E38Q	12 ± 1	7.2	445 ± 255	126
Control	136 ± 33	47.1	2058 ± 80	469

WT: wild type.

Table II Crystallographic statistics for MutS E38 mutants bound to a G.T mismatched DNA

	E38T-G.T	E38A-G.T	E38Q-G.T
Data collection			
Resolution limits (Å)	20–2.1 (2.15–2.10)	20–2.5 (2.59–2.50)	20–2.4 (2.49–2.40)
$I/\sigma(I)$	14.46 (1.52)	9.81 (1.47)	11.90 (1.80)
Completeness (%)	91.7 (47.0)	99.3 (98.3)	99.3 (92.9)
R_{merge} (%)	7.1 (48.8)	11.4 (72.2)	12.2 (69.3)
Cell parameters (Å)	$a = 89.61$	$a = 89.55$	$a = 89.62$
	$b = 92.17$	$b = 92.49$	$b = 92.04$
	$c = 260.53$	$c = 261.66$	$c = 260.97$
Refinement			
Resolution (Å)	20–2.1	20–2.5	20–2.4
R-factor (%)	18.8	21.9	21.2
R_{free} (%)	23.3	26.9	25.8
R.m.s.d. (bonds) (Å)	0.014	0.011	0.014
R.m.s.d. (angles) (deg)	1.421	1.225	1.460
No. of atoms	13 862	13 190	13 160
No. of waters	900	240	250

Values in parentheses refer to the highest resolution shell.

repair. All other contacts to the DNA in the mutants are identical to those in the wild-type protein. The correlation between *in vivo* mismatch repair capabilities and hydrogen-bond formation suggests that this latter interaction may be critical for mismatch repair.

Mismatch recognition by E38 mutants

To understand what causes the variation in mismatch repair capability by the different MutS mutants at amino acid 38, we first studied their mismatch recognition properties and the ability to discriminate against homoduplex DNA. We used isothermal titration calorimetry (ITC), native band-shifts and surface plasmon resonance (SPR) in a Biacore flow system. In all cases, data were fit to a single-site-binding model (see Discussion).

ITC measures the heat liberated owing to binding of the protein to the substrate in solution. We find that the E38A, E38T and E38Q mutants have very similar affinities for DNA containing a G.T mismatch as wild-type MutS (Figure 2A and Table III). We could not perform these experiments successfully using homoduplex DNA, owing to low signal to noise ratios.

In band-shift studies, using native polyacrylamide gel electrophoresis with ^{32}P end-labeled DNA substrates with free ends, we found that E38T and E38Q mutants bind marginally stronger to DNA containing a G.T mismatch than the wild-type protein, whereas the E38A protein has wild-type affinity (Figure 2B and Table III). This result indicates a marginal contribution of the hydrogen bond between MutS residue 38 and the mismatched base. A more pronounced difference between the wild-type and mutant proteins is seen in homoduplex binding. Wild-type MutS binds to homoduplex DNA with low affinity, whereas in all mutants this affinity is 5- to 10-fold higher.

The ability to discriminate between mismatched and paired DNA can be quantified using the ratio of the K_d values between these substrates (Table III, G.T/A.T). All mutants show a decrease in this ratio, demonstrating loss in discriminatory capacity as a consequence of loss of repulsion of the DNA backbone, by absence of the negative charge on residue 38. The general trend is similar at 22 and 37°C, and the values we find at 37°C for wild type and E38Q and E38A mutants are in good agreement with the results reported earlier at this temperature (Table III; Schofield *et al*, 2001; Junop *et al*, 2003).

Finally, we measured affinities in real time in solution using SPR in a Biacore system. The affinities measured in this assay (Table III) are considerably higher than those measured by band-shift analysis and ITC, which may relate to the different experimental setup of the continuous flow in the Biacore system. However, in the SPR experiments as well, we observe that the ability of the mutants to discriminate between mismatched and homoduplex DNA is clearly diminished, primarily as a consequence of an increase in homoduplex binding.

Mismatch recognition on DNA with blocked ends

A potential problem in MutS binding studies is the ability of MutS protein to bind to DNA ends (Acharya *et al*, 2003; Yang *et al*, 2005). Therefore, we have included binding studies to double end-blocked DNA, both in band-shift experiments
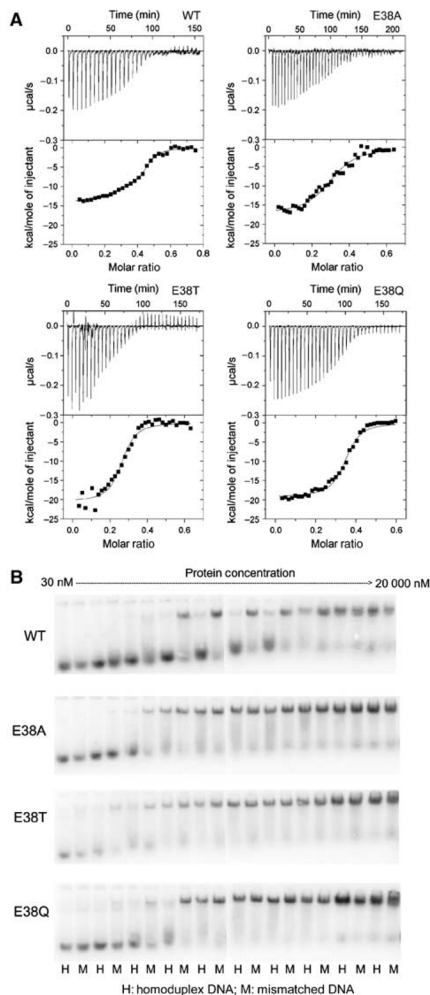

Figure 2 MutS–DNA binding assays using oligonucleotides with free ends. (**A**) Representative ITC binding curves for wild-type and mutant MutS (13–15 μM) at 37°C to a 21 bp DNA substrate containing a G.T mismatch. (**B**) Representative band-shifts showing DNA binding by increasing concentrations of wild-type MutS and E38 mutants. Reactions with homoduplex and heteroduplex DNA are shown in alternating lanes.

(Figure 3) and using SPR (Table III). In the DNA band-shifts, the DNA was blocked on both ends with biotin–streptavidin, and in the SPR experiments, we used a fluorescein–antibody complex in addition to the biotin–streptavidin coupling of the DNA on the chip (Acharya *et al*, 2003).

In the band-shift experiments using double end-blocked DNA (Figure 3B), the E38T and E38Q mutants start binding to the G.T mismatch at lower concentrations (~ 50 nM) than the E38A (~ 150 nM), again supporting a small but reproducible contribution of the hydrogen bond to mismatch binding. It can also be seen that all mutants have a similar affinity to homoduplex DNA, which is considerably higher than that of

	WT	E38A	E38T	E38Q
ITC at 37°C				
G.T	0.10 ± 0.02	0.15 ± 0.06	0.09 ± 0.02	0.06 ± 0.01
Band-shift at 22°C				
G.T	0.31 ± 0.06	0.24 ± 0.05	0.10 ± 0.02	0.17 ± 0.04
A.T	2.4 ± 0.7	0.52 ± 0.09	0.26 ± 0.04	0.42 ± 0.05
G.T/A.T	7.7 ± 2.7	2.2 ± 0.6	2.6 ± 0.7	2.5 ± 0.7
Band-shift at 37°C				
G.T	0.38 ± 0.01	0.34 ± 0.1	0.10 ± 0.02	0.14 ± 0.02
A.T	3.4 ± 0.9	0.73 ± 0.09	0.33 ± 0.1	0.42 ± 0.01
G.T/A.T	9.0 ± 2.4	2.2 ± 0.7	3.3 ± 1.5	3.0 ± 0.4
SPR at 25°C				
G.T	0.015 ± 0.001	0.020 ± 0.003	0.018 ± 0.001	0.018 ± 0.001
A.T	0.31 ± 0.1	0.24 ± 0.01	0.19 ± 0.07	0.13 ± 0.05
G.T/A.T	21 ± 7.3	12 ± 1.8	11 ± 3.8	7.2 ± 2.9
SPR at 25°C, blocked DNA				
G.T	0.015 ± 0.001	0.048 ± 0.003	0.013 ± 0.002	0.013 ± 0.001
A.T	0.21 ± 0.07	0.13 ± 0.02	0.067 ± 0.01	0.070 ± 0.01
G.T/A.T	14 ± 5	2.6 ± 0.4	5.2 ± 1.2	5.3 ± 1.1

WT: wild type; ITC: isothermal titration calorimetry; SPR: surface plasmon resonance.

wild-type MutS. This agrees with the loss of discrimination seen in the experiment using unblocked DNA.

In the SPR binding studies to double end-blocked DNA (Table III), we see that the mutants have retained their affinities to mismatched DNA, compared to unblocked DNA. Only in this experiment, the E38A mutant bound slightly weaker to the double end-blocked G.T substrate, although binding to unblocked DNA is similar to that of the other mutants and wild-type MutS. In agreement with the result obtained with unblocked DNA (Table III), all mutants display a three-fold-reduced discrimination compared to wild-type MutS.

The experiments described above confirm previously published conclusions, based on band-shifts with unblocked oligonucleotides, that removal of the carboxylate group from amino acid 38 of MutS does not strongly affect heteroduplex binding, but results in enhanced homoduplex binding and consequent reduction in the discriminatory capacity of the protein (Schofield *et al*, 2001; Junop *et al*, 2003). However, the reduced discrimination against homoduplex DNA, which is similar for the three E38 mutants, does not correlate with their significantly different capability to repair mismatches *in vivo*. These paradoxical results, therefore, suggest that initial discrimination conferred by the carboxylate group of MutS amino acid 38 is not the critical factor in DNA mismatch repair.

Mismatch-binding-induced signaling toward the nucleotide-binding domains

It has been well documented that after mismatch binding, the conformational change in the DNA-binding domains is propagated toward the distant nucleotide-binding sites (Gradia *et al*, 1997; Acharya *et al*, 2003). Because initial mismatch recognition and discrimination did not correlate with *in vivo* repair capabilities, we investigated the activity of the ATPase sites and the ability of the mutants to signal toward these
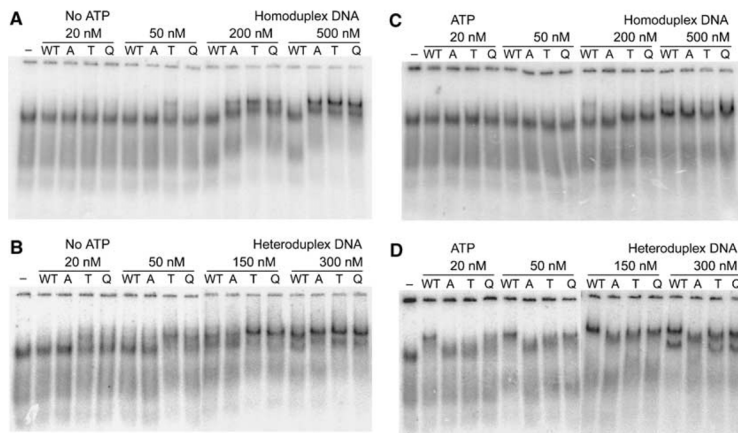

Figure 3 MutS–DNA binding assays using end-blocked oligonucleotides. Representative band-shifts showing wild-type and mutant MutS binding, in the absence of ATP, (**A**) to homoduplex, (**B**) to heteroduplex DNA, and in the presence of ATP, (**C**) to homoduplex and (**D**) to heteroduplex DNA.

Table IV Affinity for ATP and kinetic constants for DNA-induced nucleotide exchange, steady-state ATP hydrolysis and its stimulation by DNA for WT MutS and E38 mutants

	ATPγS binding	ADP–ATP exchange		ATPase		Steady-state ATPase stimulation by DNA	
	K_D(ATPγS) (μM)	k_{off} (ADP) (s^{-1})		K_M (μM)	k_{cat} (min^{-1})	*n*-fold	
		−DNA	+ DNA			Homoduplex	Heteroduplex
WT	2.3	0.030	0.39	7.8 ± 0.4	5.4 ± 0.1	3.7 ± 0.4	3.2 ± 0.4
E38A	2.1	0.029	0.38	7.4 ± 0.7	4.7 ± 0.7	4.7 ± 0.4	3.6 ± 0.4
E38T	3.0	0.028	0.36	11.0 ± 0.4	4.1 ± 0.1	5.3 ± 0.3	4.1 ± 0.4
E38Q	2.4	0.031	0.38	9.0 ± 2.2	5.4 ± 0.1	4.8 ± 0.1	3.6 ± 0.5

WT: wild type.

domains. Similar to wild-type MutS, all mutants retain a single ADP molecule per MutS dimer during purification, indicating retention of the high-affinity site for nucleotide-diphosphate (data not shown). In addition, the high-affinity site for the nucleotide-triphosphate is not affected by the mutation, as binding of ATPγS is in the low micromolar range for all MutS variants, in agreement with published values (Table IV; Bjornson and Modrich, 2003b; Lamers *et al*, 2004). The rate of ADP–ATP exchange, which is slow in the absence of mismatched DNA, was equally unaffected by the mutations (Table IV). Furthermore, we found very similar kinetic constants for steady-state ATP hydrolysis for wild-type MutS and E38 mutants (Table IV). Taken together, these observations indicate that the intrinsic properties of the nucleotide-binding sites are not affected by the mutations at position 38 in the DNA-binding domains.

We next investigated whether mutation of glutamate 38 affects the signaling from the mismatch-binding domains toward the nucleotide-binding domains by measuring the rate of ADP to ATP exchange in the presence of mismatched DNA. As reported before (Gradia *et al*, 1997; Acharya *et al*, 2003), this exchange rate, which is the rate-limiting step in steady-state ATP hydrolysis, is highly accelerated by mismatched DNA and we find that this is conserved in all mutants (Table IV). In addition, this increased release of ADP is reflected by the ability of DNA to stimulate the steady-state ATPase activity of wild-type MutS and of the amino acid 38 mutants (Table IV). These results indicate that, following DNA binding, signaling from the DNA-binding domains toward the nucleotide-binding domains depends neither on the presence of a negative charge nor on the formation of a hydrogen bond by the residue at position 38.

Formation of a stable MutS–DNA–ATP complex
ATP binding to MutS induces a rearrangement of the nucleotide-binding domains toward each other (Lamers *et al*, 2004) and these conformational changes are propagated toward the mismatch-binding and/or clamp domains, resulting in release of the mismatch and formation of a sliding clamp. We analyzed whether the mutants were affected in clamp formation by investigating their behavior upon ATP addition to prebound, end-blocked DNA–MutS complexes (see Materials and methods). The band-shifts show that the wild-type protein binds to G.T-mismatched DNA at a much lower concentration ($\sim$20 nM) compared to binding without subsequent challenge with ATP (Figure 3B and D). This indicates that this ATP-bound state of MutS, presumably the sliding clamp, is more stably associated with DNA than the initial

interaction between MutS and the mismatch itself. The mutants bind in varying degrees, with E38Q starting to bind at lower concentrations than the others and the E38A mutant lagging behind the other MutS mutants. Thus, glutamate 38 appears to be important for this increased ATP-mediated heteroduplex affinity. The reproducible appearance of double bands on heteroduplex DNA at 300 nM protein concentrations in all MutS proteins except E38A is unexpected. We speculate that the faster migrating band may represent a form of the ATP-bound sliding clamp that is induced at higher protein concentrations. Aberrant forms of MutS–DNA band-shift complexes have been reported in several other studies (Gradia *et al*, 1997; Acharya *et al*, 2003; Antony and Hingorani, 2003).

ATP has a remarkable effect on the binding of the MutS mutants to homoduplex DNA (Figure 3A and C). Even at the highest protein concentration, there is no binding by any of the mutants, although in the absence of ATP all mutants bind well to homoduplex DNA at this concentration. Because both ends of the homoduplex DNA were blocked by biotin–streptavidin, this ATP-induced release of the homoduplex must have occurred via direct dissociation of the protein (Iaccarino *et al*, 2000; Selmane *et al*, 2003). These results indicate qualitatively that the binding of ATP to the protein, after DNA binding, brings about a secondary discrimination step, or verification. This verification results, on the one hand, in stable clamp formation on mismatched DNA and, on the other hand, in direct dissociation after homoduplex binding. The ATP-mediated dissociation from homoduplex is operational in all mutants, even in the near absence of initial discrimination, and therefore does not involve residue 38.

MutL is recruited by the stable sliding clamp form of the MutS–ATP state on heteroduplex DNA (Acharya *et al*, 2003). We analyzed whether MutL binding was affected in the mutants by including MutL in the end-blocked band-shifts described above. At high protein concentrations, at which all mutants are forced into a stable ATP-bound complex with DNA (Figure 4, left panel), we see a supershifted band for all the mutants upon addition of MutL (Figure 4, right panel). At low protein concentrations, the ability to interact with MutL correlates with the amount of stable clamp formation (data not shown). Together, these data show that none of the mutants have lost the ability to interact with MutL and therefore indicate that MutS residue 38 is not directly involved in the interaction of MutS with MutL.

In pre-steady-state ATP hydrolysis experiments, a burst of ADP formation has been observed, indicating that the first
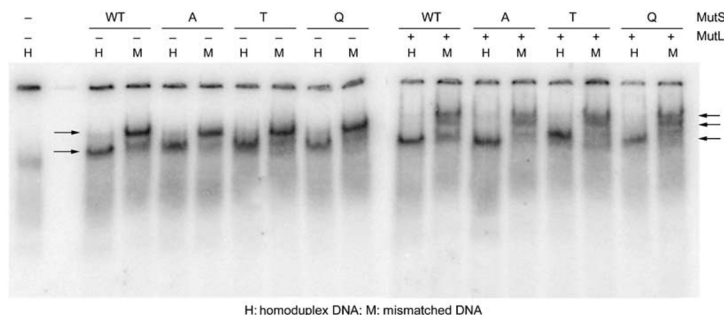
Figure 4 MutL binding to MutS–DNA complex, with G.T mismatched double end-blocked oligonucleotides. Wild-type and mutant MutS (400 nM) bound to mismatched DNA and supershifted by binding to MutL (1.6 μM), in the presence of 1 mM ATP during electrophoresis. Single lane on the left: DNA with biotin–streptavidin blocks on both ends. Left panel: mismatch-dependent trapping of wild-type and mutant MutS sliding clamps. Free DNA and sliding clamp complexes are indicated by arrows. Right panel: mismatch-dependent MutL complex formation by wild-type and mutant MutS sliding clamps. Lower arrow indicates unbound end-blocked DNA, arrow in the middle indicates MutS-associated DNA and upper arrow points at MutS–MutL supershifted DNA complexes.

ATP molecule that is bound after ADP release is hydrolyzed rapidly. This occurs both in the absence of any DNA substrate or in the presence of homoduplex DNA (Bjornson *et al*, 2000; Antony and Hingorani, 2004). This rapid ADP formation is followed by a slower steady-state hydrolysis rate, consistent with the release of ADP being the rate-limiting process for steady-state ATP turnover. The presence of a mismatch results in almost complete suppression of rapid ADP formation, demonstrating that binding of a mismatch inhibits initial ATP hydrolysis. We determined the rate of ATP hydrolysis in the absence of DNA during the first few catalytic turnovers and obtained a reproducible burst amplitude for wild type and all three MutS mutants, corresponding to 0.6 ADP molecules produced per MutS monomer (Figure 5A and Supplementary data). This indicates rapid hydrolysis of one ATP molecule in only one subunit of the functional MutS dimer, in agreement with its expected asymmetry (Bjornson and Modrich, 2003b; Lamers *et al*, 2003; Antony and Hingorani, 2004). As expected, binding to mismatched DNA causes the loss of this rapid hydrolysis step in the wild type and E38Q and E38T mutants. However, in the E38A mutant, the ATP hydrolysis remains fast also after mismatch binding (Figure 5A). Apparently, in this mutant, mismatch binding can no longer induce inhibition of rapid initial ATP hydrolysis. This suggests that E38A is not able to remain in the relatively long-lived MutS-mismatched DNA–ATP state that has been proposed to play an important role in initiating downstream events (Bjornson *et al*, 2000; Antony and Hingorani, 2003, 2004).

To examine the kinetic effect of the E38 mutation on the stable ATP-bound state of MutS in more detail, we used SPR to determine the kinetics of ATP-induced release of MutS prebound to DNA with a free end or to DNA with both ends blocked (Figure 5B and C). The kinetics of dissociation from homoduplex DNA, either blocked on one end or double end-blocked, is too fast for quantification in this assay system. When one end of the DNA containing a G.T mismatch is kept unblocked, the rates of release of all the mutants are similar to that of the wild-type MutS (k_{max} of 0.13–0.14 s^{-1}; Figure 5B

and Supplementary data). This mainly reflects the rapid dissociation of wild-type and mutant proteins from the free end of the DNA. However, when both DNA ends are blocked, a much slower release is seen (Figure 5C). As MutS is trapped on double end-blocked ends in other systems (e.g. band-shifts above), this slow release must reflect the dissociation of ATP-bound MutS due to the extensive flow that is inherent to the SPR experimental setup. Strikingly, in this experiment, the E38A releases much faster (k_{max} of 0.098 s^{-1}) than the wild-type, E38T and E38Q proteins ($k_{max} \sim$ 0.036–0.049 s^{-1}) (Figure 5C; Supplementary data). We conclude that, in the E38A mutant, the absence of a hydrogen bond with the mismatched base results in rapid hydrolysis of bound ATP and concomitant fast release of MutS from the mismatched DNA. This is in contrast to the mismatch repair-proficient variants in which the hydrogen bond-mediated interaction between amino acid 38 and the mismatched base inhibits ATP hydrolysis and retains the protein on the DNA.

Discussion

MutS glutamate 38 has been reported to allow discrimination between perfectly paired DNA and DNA containing a mismatch or unpaired base (Junop *et al*, 2001; Schofield *et al*, 2001). However, from the analysis of a series of MutS mutants at amino acid 38, we could not satisfactorily correlate the loss of discriminatory capacity in the mutants with their ability to repair mismatches in an *in vivo* complementation assay. We have therefore extended the analysis of these mutants in their down-stream signaling ability.

In these studies of the DNA–MutS interaction, it became obvious that the single-site-binding model that is traditionally used (Blackwell *et al*, 2001; Brown *et al*, 2001; Schofield *et al*, 2001) is insufficient to explain the data completely. This is particularly clear in the ITC data (Figure 2), but was also apparent from the statistical analysis of the SPR data (not shown). However, it is difficult to describe the process properly: the binding of MutS to a mismatch in DNA is associated with binding and release of nucleotide and con-
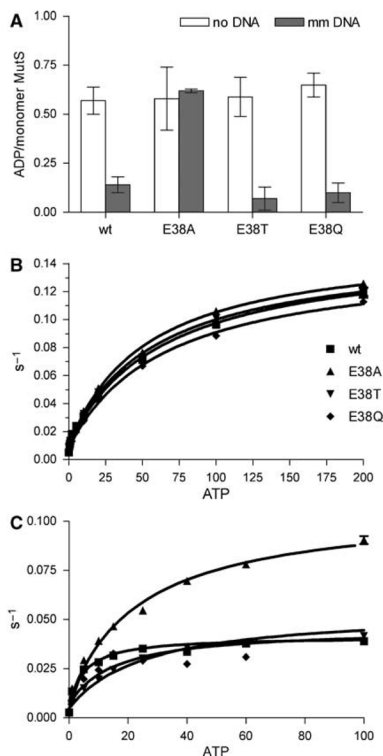
Figure 5 Kinetic response to ATP binding of wild-type MutS and E38 mutants. (**A**) Magnitude of burst amplitude of ADP formation during the first hydrolytic turnover in a rapid quench ATPase assay in the absence and presence of DNA containing a G.T mismatch. (**B**) Kinetics of ATP-induced MutS release from mismatched (mm) DNA with a free end and (**C**) from double end-blocked mismatched DNA determined by SPR.

formational changes in the protein and DNA. Furthermore, monomer–dimer and dimer–tetramer transitions should be taken into account (Bjornson *et al*, 2003a; Lamers *et al*, 2004). In the absence of a better model, the one-site binding model is the best approximation, as it leads to acceptable fits with reliable statistics. Further studies will be required to establish a proper binding model for the DNA–MutS interaction.

Our studies into the effect of the E38 mutations were also complicated by the variable effect of temperature on the different mutants in the *in vivo* repair assay, in particular for E38Q, which repairs at 22°C but not at 37°C (Table I). As we did not find a temperature dependence in either DNA binding or mismatch discrimination for this mutant, it seems likely that the temperature-dependent *in vivo* mismatch repair ability of the E38Q mutant is due to kinetic differences. Our experimental setup does not allow a complete analysis of

the effect of temperature on the kinetics of the reaction, but their importance fits very well with our overall conclusion that kinetic differences are responsible for the defects observed in the E38 mutants (see below).

A final complicating factor in our studies is the possibility of DNA end-binding. In two recent reports on DNA binding by *E. coli* and *T. aquaticus* MutS, it was found that especially *E. coli* MutS has strong affinity for free DNA ends (Acharya *et al*, 2003; Yang *et al*, 2005). We considered whether this could explain the discrepancy between low discriminatory capacity and high repair efficiency of *E. coli* MutS. We can assess end-binding directly by comparison of affinity constants obtained by SPR on DNA with one free end and both ends blocked. If significant end-binding to the free DNA ends occurs, affinity for homoduplex with a free end would be higher than for double end-blocked homoduplex DNA. However, the observed affinities for homoduplex are similar, or even lower, on an oligonucleotide with free ends, strongly indicating that end-binding does not occur in this continuous flow system (Table III). Meanwhile, the calculated specificities for mismatched versus homoduplex DNA are at most two-fold higher than specificities we find in band-shifts using DNA with free ends, and are similar to reported values (Schofield *et al*, 2001; Selmane *et al*, 2003). We therefore cannot conclude from our experiments that the specificity for a mismatch will be significantly better in the absence of end-binding.

We have used a number of different techniques to study DNA recognition by MutS wild type and mutants of amino acid 38. Despite their intrinsic technical differences, all binding studies confirm that the three mutants show loss of discrimination between mismatched and homoduplex DNA. This was expected, as they lack the negatively charged carboxylate group (Schofield *et al*, 2001). The E38A mutant appears to have a slightly weaker ability to discriminate than the other mutants, which may be related to the absence of the hydrogen bond. However, the differences with the other mutants are minor and do not correlate with their large differences in repair capability. On the other hand, we have shown that efficient repair is possible even if the intrinsic low discrimination capacity of wild-type MutS is decreased further by mutation of Glu38. This strongly indicates that MutS relies on an additional factor to verify binding to mismatched DNA. It was previously suggested that the increased rate of ADP–ATP exchange, induced by mismatch binding, plays a role in determining specificity of mismatch repair (Acharya *et al*, 2003). Furthermore, ADP to ATP exchange in MutS after binding to DNA might confer a second round of mismatch discrimination by inducing different modes of dissociation from bound hetero- and homoduplex DNA (Iaccarino *et al*, 2000; Junop *et al*, 2001; Selmane *et al*, 2003). From the band-shift experiments on DNA with double-blocked ends (Figure 3), it is clear that, in the presence of ATP, wild-type MutS has a higher affinity for mismatched DNA than in the absence of the nucleotide-triphosphate. In addition, exploiting the high affinity of the mutants for homoduplex DNA, which is a result of the loss of the negative charge, we were able to demonstrate that ATP binding causes direct dissociation of MutS from inadvertently bound homoduplex. As this phenotype was found in all mutants, this process requires neither the carboxylate side chain nor the hydrogen bond of E38. In conclusion, ATP binding induces

enhanced heteroduplex binding as well as increased release from homoduplex DNA and this secondary discrimination, or verification, by MutS presumably allows E38T and E38Q to be active *in vivo*.

To understand the varying repair capabilities of the mutants, and especially the severe defect of E38A, we studied the signaling toward the ATPase domains, and conformational responses induced by ATP binding. In all mutants, the binding of a mismatch results in uptake of ATP, excluding involvement of E38 in this signaling process. In wild-type MutS, hydrolysis of this ATP is inhibited and a conformational change results in the formation of the ATP-bound sliding clamp. Here, we found a significant difference in the kinetic behavior of the E38A mutant. It releases double end-blocked heteroduplex DNA much faster than any of the other proteins after ATP binding. Concomitantly, E38A has lost the inhibition of rapid ATP hydrolysis that is induced by mismatch binding. Thus, despite reasonable initial mismatch binding, this mutant is defective in subsequent clamp formation. The isomorphic crystal structures of wild-type and mutant MutS reveal that the overall protein conformation and contacts to the DNA, apart from the direct interaction between residue 38 and the DNA mismatch, are not significantly different. The E38Q and E38T mutant side chains still form a hydrogen bond to the mismatched thymidine, whereas no such interaction is present in the E38A mutant. As MutS E38Q and E38T are repair proficient, we conclude that the hydrogen bond to the mismatched base, and not the negative charge on the glutamate, is critical for repair.

Based on these data, we propose the following model for glutamate 38-mediated authorization of mismatch repair: MutS scans the DNA for mismatches, attempting to kink the DNA. After mismatch binding, or after fortuitous homoduplex binding, MutS kinks the DNA and exchanges ADP for ATP in the distant nucleotide-binding domains. In addition, uniquely after mismatch binding, a hydrogen bond is formed between the mismatched base and glutamate 38 (or the threonine or glutamine residue in the respective mutants). The establishment of this hydrogen bond initiates intramolecular signaling, resulting in the inhibition of hydrolysis of bound ATP. This stable ATP binding results in tightening of the MutS dimer (Lamers *et al*, 2004), and propagation of this conformational change toward the DNA-binding domains initiates formation of the sliding clamp. The ability of MutS to form a stable sliding clamp after releasing the mismatch implies that the hydrogen bond is not involved in stabilizing the clamp while it is sliding, but solely in its induction. This stable clamp subsequently initiates repair by formation of a complex with MutL, and recruitment of additional downstream factors. The hydrogen bond between the residue at position 38 and the mismatch cannot be formed in the E38A mutant, or when (wild type, or E38T or E38Q) MutS is inadvertently bound to homoduplex DNA. In these cases, there is no intramolecular signal generated to inhibit ATP hydrolysis. We hypothesize that this results in rapid hydrolysis of ATP and concomitant widening of the MutS dimer at the intertwined ATPase domains, leading to conformational changes at the DNA-binding domains that prevent stable clamp formation and result in release of the DNA by MutS. In conclusion, the hydrogen bond between MutS amino acid 38 and the mismatch validates mismatch binding and authorizes repair.

Materials and methods

Site-directed mutagenesis
MutS mutants E38A, E38T and E38Q were constructed by converting the Glu GAA codon of wild-type MutS plasmid pMQ372 (full length) or pM800 (ΔC800) into an Ala GCG, Thr ACC or Gln CAG codon, using QuickChange (Stratagene). Mutations were confirmed by DNA sequencing.

DNA substrates
We used double-stranded DNA oligonucleotides containing a single G.T mismatch or a perfectly paired, otherwise identical, oligonucleotide. A 30 bp oligonucleotide (Lamers *et al*, 2000) was used for crystallization, band-shifts with free DNA ends and DNA-dependent ATPase assays. A shortened version (21 bp) was used for the ITC (upper strand 5′ AGC TGC CA**G** GCA CCA GTG TCA, bottom strand 5′ TGA CAC TGG TGC **T**TG GCA GCT, mismatched nucleotides in bold). A 41 bp biotinylated oligonucleotide was used for band-shifts with DNA with blocked ends (upper strand 5′ ATC GAA TTA GAA GCT GCC A**G**G CAC CAG TGT CAG CGT CCT AT-biotin, bottom strand 5′ ATA GGA CGC TGA CAC TGG TGC **T**TG GCA GCT TCT AAT TCG AT-biotin).

In vivo mismatch repair
Mismatch repair complementation assays were performed as described (Wu and Marinus, 1994; Lamers *et al*, 2003). A MutS-deficient strain RK1517 was transformed with wild-type MutS or E38A, E38T, E38Q mutants or empty vector. Samples of nine overnight cultures for each strain, grown in LB at 37 or 22°C, were plated on LB with 100 μg/ml rifampicin to select for mutant colonies, and titers were determined on LB plates. Mutation frequency was calculated as the median value of nine mutation frequencies determined. Mutation rates were determined using the Luria and Delbrück equation $0.6\mathrm{Rif}^R/C = Np\log(CNp)$, where Rif^R is the number of rifampicin-resistant cells in a culture expanded to N cells, C is the number of different cultures and p is the number of new Rifampicin-resistant cells per cell division. Each experiment was performed three times.

Structure determination
MutS lacking the C-terminal 53 amino acids (ΔC800) and its mutants were purified and crystallized as described (Natrajan *et al*, 2003). Crystals grew in space group $P2_12_12_1$ and data were collected at the ID 14 beamlines, located at the ESRF Grenoble, at a wavelength of 0.91 Å. Processing was carried out using the HKL 2000 (Otwinowski and Minor, 1997) package (Table II). The structures of the mutants were solved by rigid-body refinement using the MutS structure as model (Lamers *et al*, 2000). The residues from 30 to 45 in both monomers and the mismatched bases were removed before the first restrained refinement and rebuilt into the difference density map. All refinements were performed using REFMAC5 (Murshudov *et al*, 1997) and TLS refinement (Winn *et al*, 2001), in the CCP4 (CCP4, 1994) program suite. The structures were checked using the WHATCHECK (Hooft *et al*, 1996) server.

Isothermal titration calorimetry
ITC experiments were performed using a MicroCal VP-ITC calorimeter at 37°C. Both protein and DNA were dialyzed overnight in the ITC buffer (25 mM Hepes–NaOH pH 7.5, 125 mM NaCl, 10 mM MgCl₂, 10 mM β-mercaptoethanol, 10% glycerol, 0.5 μM ADP). DNA was injected into the protein solution (13–15 μM) and the increase in ligand concentration was ∼0.36 μM per injection. A long time interval of 300 s was given between injections and only the peak was selected for integration using a constant peak width to avoid a large scatter in the data points obtained. Although a baseline of DNA into buffer was essentially flat, the intrinsic scatter was very large, probably owing to dilution effects of the DNA. Therefore, a flat baseline taken from the end of each titration (Figure 2A–D) was used for peak integration. The data were analyzed using the Origin software (MicroCal) and the single-site-binding model was fitted to it. Standard errors were determined after repeating each experiment two to four times.

DNA band-shifts
For band-shifts using DNA with free ends, MutS (30 nM–20 μM) was bound to 1 nM ³²P end-labeled 30 bp DNA substrate (see above) in

71

binding buffer (25 mM Hepes–NaOH pH 7.5, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, 1 % glycerol, 25 μg/ml BSA) in 10 μl total volume. After incubation of ∼10 min at 20 or 37°C, 3 μl of 50% sucrose was added to the samples, which were then loaded under voltage onto 4% native polyacrylamide gels (29:1 acrylamide:bisacrylamide) and run in TAE buffer + 5 mM MgAc at 22 or 37°C. The gels were dried, exposed to phosphor-imager and quantified using ImageQuant software. Data were analyzed using Microsoft Excel with the Solver add-in software. The single-site-binding model [DNA bound] = B^{max}[protein]/(K_d + [protein]) was fitted to the data. Standard errors were determined from three independent experiments.

For band-shifts with blocked DNA ends, 1 nM ^{32}P-labeled 41 bp DNA substrate (see above) with biotin at both 3′ ends was preincubated with streptavidin (10 ng/μl) for ∼10 min before binding to MutS (50–500 nM). For sliding clamp formation, after ∼10 min, the reaction with MutS was split into two reactions, of which one was further incubated with ATP (10 mM) and the other with ATP buffer for ∼10 min. Reactions were performed at 22 or 4°C in binding buffer. Data are reproducible, but quantification was not possible owing to smearing of the double end-blocked DNA. For MutL binding, after ∼10 min, 10 mM ATP was added to the reaction with MutS (400 nM). After ∼10 min, one-half of the reaction was further incubated with MutL (1600 nM) and the other with MutL buffer. To 10 μl volume of final reactions, 3 μl of 50% sucrose was added. Samples were run on 4 % native polyacrylamide gels (29:1 acrylamide:bisacrylamide) in TAE buffer + 5 mM MgAc and 1 mM ATP both in the gel and running buffer.

Nucleotide binding and exchange
Binding of the poorly hydrolyzable analog ATPγS to the high-affinity site of MutS was assessed by filter binding (Schleicher & Schuell Minifold II slot-blot). ^{35}S-labeled ATPγS was incubated with 0.1–10 μM MutS in reaction buffer (25 mM Hepes–NaOH, 150 mM NaCl, 10 mM $MgCl_2$) and incubated on ice for several minutes. Samples were spotted in triplicate onto 0.45 μM nitrocellulose filters that were prewashed with reaction buffer. Filters were analyzed as described before (Lamers *et al*, 2003). Nucleotide exchange was monitored on a FLUOstar Optima spectrophotometer using the increase in fluorescence emission intensity of the ADP derivative MANT-ADP (Molecular Probes) upon binding to MutS. We used excitation and emission filters of 355 and 405 nm, respectively. Reactions contained 5 μM MutS and 10 μM MANT-ADP with or without 5 μM mismatched DNA. ADP exchange was initiated by fast titration of 1 volume of either reaction buffer or reaction buffer containing 100 μM ATP, and loss of fluorescence emission intensity was followed in time. After correction for buffer dilution effects, a function describing linear exponential decay was fitted to the data using nonlinear regression.

ATP hydrolysis
Kinetics for steady-state ATP hydrolysis were determined using an ATP-regenerating spectrophotometric assay as described (Lamers *et al*, 2003). Stimulation of the steady-state ATP hydrolysis rate by DNA was determined as described (Lamers *et al*, 2004). For determination of the magnitude of the pre-steady-state burst amplitude, 20 μl reactions contained 10 μM MutS in 25 mM

Hepes–NaOH pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol, 10 mM $MgCl_2$ and 50 μM ATP containing α-^{32}P-labeled ATP, with or without 10 μM mismatched DNA. Reactions without ATP were assembled at 0°C and hydrolysis was initiated by the addition of ATP. Samples (1 μl) were removed, quenched in 40 μl 0.2% SDS and 10 mM EDTA and spotted onto PEI cellulose plates (Merck). Plates were developed in 1 M orthophosphoric acid (pH 3.8) and analyzed by phosphorimaging.

Surface plasmon resonance
SPR spectroscopy was performed at 25°C on a BIAcore 3000. Streptavidin SA sensor chips were derivatized with ∼150 resonance units of biotin/fluorescein-derivatized 41 bp heteroduplex and homoduplex. In measurements using double end-blocked DNA, the 3′-fluorescein end was blocked using anti-fluorescein rabbit IgG Fab fragment (Molecular Probes, Invitrogen). Anti-fluorescein was flown across the chip at 5 μl/min before each MutS injection. MutS (50–300 nM) or MutS E38 mutants in running buffer (RB; 25 mM Hepes–NaOH, pH 7.5, 10 mM dithiothreitol, 150 mM NaCl, 5 mM $MgCl_2$) were injected across the SA chip at 30 μl/min. In measurements of ATP-dependent MutS dissociation, the RB containing 0–200 μM ATP was injected immediately after MutS (300 nM). Chips were regenerated by a 20 μl injection of 0.05% SDS. The dissociation constant (K_d) for MutS–DNA complexes was determined by titration with increasing concentrations of MutS. Saturation binding values were fit according to a one-to-one binding model. ATP-dependent MutS dissociation was fit to one-phase exponential decay. The rate constants (k) from these fittings were plotted as a function of the ATP concentration and the hyperbolic function $k = k_0 + k_{max}[ATP]/(K_m + [ATP])$ was fitted to the data.

Coordinates
Coordinates have been deposited in the PDB: 1WB9 (MutS-E38T), 1WBB (MutS-E38A) and 1WBD (MutS-E38Q).

Supplementary data
Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

References

Acharya S, Foster PL, Brooks P, Fishel R (2003) The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair. *Mol Cell* **12:** 233–246

Antony E, Hingorani MM (2003) Mismatch recognition-coupled stabilization of Msh2–Msh6 in an ATP-bound state at the initiation of DNA repair. *Biochemistry* **42:** 7682–7693

Antony E, Hingorani MM (2004) Asymmetric ATP binding and hydrolysis activity of the *Thermus aquaticus* MutS dimer is key to modulation of its interactions with mismatched DNA. *Biochemistry* **43:** 13115–13128

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

Bjornson KP, Blackwell LJ, Sage H, Baitinger C, Allen D, Modrich P (2003a) Assembly and molecular activities of the MutS tetramer. *J Biol Chem* **278:** 34667–34673

Bjornson KP, Modrich P (2003b) Differential and simultaneous adenosine di- and triphosphate binding by MutS. *J Biol Chem* **278:** 18557–18562

Blackwell LJ, Bjornson KP, Allen DJ, Modrich P (2001) Distinct MutS DNA-binding modes that are differentially modulated by ATP binding and hydrolysis. *J Biol Chem* **276:** 34339–34347

Brown J, Brown T, Fox KR (2001) Affinity of mismatch-binding protein MutS for heteroduplexes containing different mismatches. *Biochem J* **354:** 627–633

CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* **50:** 760–763

Drotschmann K, Yang W, Browneell FE, Kool ET, Kunkel TA (2001) Asymmetric recognition of DNA local distortion. Structure-based functional studies of eukaryotic Msh2–Msh6. *J Biol Chem* **276:** 46225–46229

Gradia S, Acharya S, Fishel R (1997) The human mismatch recognition complex hMSH2–hMSH6 functions as a novel molecular switch. *Cell* **91:** 995–1005

Gradia S, Subramanian D, Wilson T, Acharya S, Makhov A, Griffith J, Fishel R (1999) hMSH2–hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol Cell* **3:** 255–261

Hays JB, Hoffman PD, Wang H (2005) Discrimination and versatility in mismatch repair. *DNA Repair* **4:** 1463–1474

Hooft RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. *Nature* **381:** 272

Iaccarino I, Marra G, Dufner P, Jiricny J (2000) Mutation in the magnesium binding site of hMSH6 disables the hMutSalpha sliding clamp from translocating along DNA. *J Biol Chem* **275:** 2080–2086

Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W (2001) Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Mol Cell* **7:** 1–12

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

Kunkel TA, Erie DA (2005) DNA mismatch repair. *Annu Rev Biochem* **74:** 681–710

Lamers MH, Georgijevic D, Lebbink JH, Winterwerp HH, Agianian B, de Wind N, Sixma TK (2004) ATP increases the affinity between MutS ATPase domains. Implications for ATP hydrolysis and conformational changes. *J Biol Chem* **279:** 43879–43885

Lamers MH, Perrakis A, Enzlin JH, Winterwerp HH, de Wind N, Sixma TK (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G × T mismatch. *Nature* **407:** 711–717

Lamers MH, Winterwerp HH, Sixma TK (2003) The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J* **22:** 746–756

Lynch HT, de la Chapelle A (1999) Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet* **36:** 801–818

Marti TM, Kunz C, Fleck O (2002) DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol* **191:** 28–41

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53:** 240–255

Natrajan G, Lamers MH, Enzlin JH, Winterwerp HH, Perrakis A, Sixma TK (2003) Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res* **31:** 4814–4821

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

Otwinowski Z, Minor W (eds) (1997) *Processing of X-ray Data Collected in Oscillation Mode.* New York: Academic Press

Peltomaki P (2003) Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* **21:** 1174–1179

Schofield MJ, Brownewell FE, Nayak S, Du C, Kool ET, Hsieh P (2001) The Phe-X-Glu DNA binding motif of MutS. The role of hydrogen-bonding in mismatch recognition. *J Biol Chem* **276:** 45505–45508

Schofield MJ, Hsieh P (2003) DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* **57:** 579–608

Selmane T, Schofield MJ, Nayak S, Du C, Hsieh P (2003) Formation of a DNA mismatch repair complex mediated by ATP. *J Mol Biol* **334:** 949–965

Winn MD, Isupov MN, Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D* **57:** 122–133

Wu TH, Marinus MG (1994) Dominant negative mutator mutations in the mutS gene of *Escherichia coli. J Bacteriol* **176:** 5393–5400

Yamamoto A, Schofield MJ, Biswas I, Hsieh P (2000) Requirement for Phe36 for DNA binding and mismatch repair by *Escherichia coli* MutS protein. *Nucleic Acids Res* **28:** 3564–3569

Yang Y, Sass LE, Du C, Hsieh P, Erie DA (2005) Determination of protein–DNA binding constants and specificities from statistical analyses of single molecules: MutS–DNA interactions. *Nucleic Acids Res* **33:** 4322–4334

73

Genomic fidelity in organisms depends on several DNA repair mechanisms, which are highly specific in their ability to target mutations and defects in DNA, and initiate repair. DNA mismatch repair is one of these processes, which detects and repairs mismatches and insertion-deletion loops in DNA. These come about due to replication errors made by the polymerase, and also to certain other recombinatorial events. The importance of this particular repair mechanism is underlined by the fact that defects in the mismatch repair pathway predispose humans to HNPCC or Hereditary Non Polyposis Colorectal Cancers. The DNA mismatch proteins also play an anti-recombinogenic role, preventing inter-species recombination.

This thesis studies the basic molecular mechanisms of DNA mismatch repair in *Escherichia coli* from a structural standpoint. It provides structural information on three important aspects of the functions of MutS. These are, the binding of the DNA mismatch, the binding of nucleotides and the associated conformational changes, and finally, the link between the residues involved in mismatch binding and the long range signalling that takes place between the mismatch binding and ATPase domains. Crystal structure determination, site directed mutagenesis and biochemical approaches have been used to unravel the details of this complex and interesting process.

Chapter 2 describes the crystal structures of *Escherichia coli* MutS in complex with A.A, C.A, and G.G mismatches, and also with an unpaired thymidine. MutS is an asymmetric dimer comprising of two monomers, A and B. Only monomer A binds directly to the mismatch. The DNA in all the structures reported here is kinked by an angle of 60º. Two highly conserved residues, Glu 38 and Phe 36, do mismatch recognition. Phe 36 stacks onto one of the mismatched bases

and the same base is oriented so that a nitrogen on it can act as a hydrogen bond donor to the acid group on the sidechain of Glu 38. The kinking of the DNA widens the minor groove containing the mismatch, giving the protein enough room to reorient the DNA bases from their protein-unbound conformation, to achieve this. The Phe 36 can stack on either a purine or pyrimidine of the mismatch. When the stacking is on a pyrimidine, the N3 forms the hydrogen bond with Glu 38, and when it stacks on a purine, with the N7 accompanied by the flipping of the base from the *anti* to the *syn* orientation. Based on these findings, it can be proposed that MutS has a common mismatch-binding mode that it uses to bind to all the different mismatches.

Chapter 3 describes the crystal structure of MutS in a new crystal form, which has some interesting features including a highly disordered mismatch-binding domain of monomer B. The most interesting new feature here is that no magnesium is to be found in the ATPase domain of monomer A, although ADP is still present. This orders the signature loop of the opposing monomer B, enabling a highly conserved serine, S668 located on this loop to get exposed to the beta phosphate tail of the ADP. Comparisons of this structure with previous MutS structures, reveals that the removal of Mg leads to substantial domain movements within the ATPase domains. Mg is shown to be biochemically essential for effective ADP-ATP exchange, following binding of the mismatch. Further comparisons with Mg free intermediates in G-proteins, suggest that this new structure is an important, new intermediate in the ATPase cycle of MutS.

Chapter 4 describes the role of Glutamate 38 in mismatch binding, the verification of mismatch binding and the initiation of repair. Glu 38 is one of the two residues that actually

contact the mismatch by forming a hydrogen bond to one of the mismatched bases. By mutating this glutamate to alanine, threonine and glutamine, the roles of the hydrogen bond and the negative charge of the acid have been investigated. In vivo mismatch repair assays based on complementation of mismatch repair reveal that the E38T and E38Q can repair mismatches, and the E38A cannot do so as effectively. In a series of binding studies the role of this residue in mismatch discrimination was studied. DNA bandshifts using end blocked substrates show that in the presence of ATP, the affinity for the wild type protein for the mismatched substrate is increased while that for a homoduplex substrate is eliminated, proving the verification of mismatch binding by ATP binding. Rapid quench ATPase assays show that while mismatched DNA can inhibit the first rapid turnover of ATP in the wildtype, E38T and E38Q, it cannot do so for the E38A. Surface plasmon resonance experiments with double end blocked DNA substrates show that E38A releases the mismatched DNA much faster than the wildtype, E38T and E38Q, upon injection of ATP following mismatch binding. These data suggest that the E38A is unable to form a stable mismatch-DNA-ATP state, required to initiate downstream events in the mismatch repair pathway. Finally, the crystal structures of E38, E38T and E38Q in complex with a mismatched DNA show that the E38A alone is unable to form a hydrogen bond to the mismatched base. Since the rate of ATP exchange is the same for all the mutants and the wildtype protein, it can be concluded from all this data, that the hydrogen bond between residue 38 and the mismatched base is needed for the effective transfer of the signal between the ATPase domains and the mismatch binding domains following nucleotide exchange, leading to the formation of the proper ATP state, which can initiate downstream events. This is a very remarkable signalling process as it takes place over a distance of around 100 Å, the

distance separating the mismatch binding domains from the ATP binding site.

Thus, MutS uses the initial instability due to the mismatch to bind it using a binding mode that is common to all mismatches. Following this, it takes up ATP to verify mismatch binding. This is followed by the formation of the ATP state, which is made to exist long enough by the inhibition of ATPase activity, so that down stream events can be initiated. All this is accompanied by several domain movements in the highly flexible protein, which are crucial to its function.

De integriteit van het genoom in organismen is afhankelijk van een serie gespecialiseerde DNA reparatiemechanismen die er op gericht zijn mutaties en fouten in het DNA op te sporen en het reparatieproces op te starten. Het DNA mismatch repairsysteem vormt één van deze processen. Het spoort fouten op die ontstaan vanwege foutieve base-paringen (zgn. mismatches) of inserties en deleties die ontstaan door fouten van de polymerase tijdens replicatie of door foutieve recombinatie.

Het belang van dit specifieke reparatiemechanisme is duidelijk omdat een defect in een van de interacterende mismatch reparatie eiwitten mensen een predispositie geeft voor kanker, met name, HNPCC oftewel Hereditary Non Polypolyposis Colorectal Cancer. Daarnaast voorkomen de mismatch reparatie eiwitten de recombinatie tussen soorten.

In dit proefschrift wordt het moleculaire mechanisme van DNA mismatch reparatie in *Escherichia coli* vanuit een structureel oogpunt onder de loep genomen. Er worden drie belangrijke aspecten van de functie van MutS behandeld. Deze zijn ten eerste de binding van de DNA mismatches, ten tweede de binding van nucleotiden met de bijkomende conformatieveranderingen, ten derde de link tussen de DNA herkenning en de nucleotidebinding. Structuurbepalingen, mutaties op specifieke plaatsen en biochemische methoden zijn gebruikt om de bijzonderheden van dit complexe en interessante proces van mismatch reparatie te onthullen.

Hoofdstuk 2 beschrijft de kristalstructuren van het mismatch reparatie eiwit MutS afkomstig uit *Escherichia coli*. Het beschrijft structuren van dit eiwit gebonden aan verschillende mismatches, zoals A.A, C.A. en G.G mismatches en een ongepaarde thymine. MutS is een asymmetrische dimeer, die bestaat uit de twee monomeren A en B. Alleen monomeer A bindt direct aan de mismatch. Het DNA in deze

structuren is geknikt in een hoek van 60 graden. Mismatchherkenning wordt bewerkstelligd door twee sterk geconserveerde aminozuren in MutS, respectievelijk glutamaat 38 (Glu38) en fenylalanine 36 (Phe36). Tijdens mismatchherkenning wordt Phe36 op een van de fout gepaarde basen in het DNA ingevoegd, hetzij op een purine of een pyrimidine. Hierdoor wordt het DNA geknikt en dit verwijdt de 'minor groove', die de mismatch bevat. Op deze manier heeft het eiwit de ruimte om de base opnieuw te oriënteren, zodat een stikstof (N), die zich op de base bevindt, een waterstofbrug kan aangaan met de zuurgroep aan de zijketen van Glu 38. Als Phe 36 op een pyrimidine gaat zitten, dan vormt de N3 een waterstofbrug met Glu 38. Gaat Phe 36 echter op een purine zitten, dan wordt de base omgeklapt van anti naar syn oriëntatie en vormt de N7 een waterstofbrug.
Gebaseerd op deze gegevens kan worden gesteld dat MutS een standaard mismatch bindingsmodus heeft, die gebruikt wordt om de verschillende mismatches te binden.

Hoofdstuk 3 beschrijft een nieuwe kristalvorm van MutS die een aantal interessante kenmerken heeft, zoals het uiterst ongeordende mismatchbindingsdomein van monomeer B. Het meest belangrijke nieuwe kenmerk is dat er geen magnesium kan worden gevonden in het ATPase domein van monomeer A, hoewel ADP wel nog steeds aanwezig is. Dit ordent de signature loop van de tegenoverstaande monomeer B, wat het mogelijk maakt voor de sterk geconserveerde serine S668 in deze loop om contact te maken met de betafosfaatstaart van de ADP.
Door deze nieuwe kristalstructuur met eerdere MutS kristalstructuren te vergelijken, wordt duidelijk dat de verwijdering van magnesium leidt tot substantiële conformatieveranderingen in de ATPase domeinen. In een

serie biochemische proeven blijkt dat magnesium essentieel is voor effectieve ADP-ATP uitwisseling, na binding van de mismatch. Verdere vergelijkingen met magnesiumvrije tussenvormen van G-eiwitten, eiwitten die GDP en GTP binden, suggereren dat deze nieuwe structuur een belangrijke nieuwe intermediair is in de ATPase cyclus van MutS.

Hoofdstuk 4 beschrijft het contact tussen DNA herkenning en ATP binding.

Glu38 is één van de twee MutS residuen die daadwerkelijk contact maakt met één van de foutief gepaarde bases door middel van de vorming van waterstofbruggen. Door mutatie van deze glutamaat naar alanine (E38A), threonine (E38T) en glutamine (E38Q) werden de rollen van waterstofverbindingen en de negatieve lading van de zure zijketen onderzocht. In vivo mismatch repair assays gebaseerd op complementatie van mismatch repair onthult dat de E38T en E38Q mutanten de fout gepaarde basen kunnen herstellen, terwijl de E38A dit niet goed kan.

In een reeks van bindingsexperimenten werd de rol van dit aminozuur in mismatch discriminatie onderzocht. In experimenten, waarin DNA-eiwit interacties kunnen worden aangetoond door veranderingen in het DNA banden patroon (zgn. Band-shifts), is door het gebruik van DNA, waarvan één of beide van de uiteinden is voorzien van een blokkade, inzichtelijk gemaakt dat de affiniteit van wildtype MutS voor substraat met foutief gepaarde basen hoger wordt in aanwezigheid van ATP. Voor homoduplex substraat daarentegen verdwijnt de affiniteit, wat bewijst dat verificatie van mismatch binding afhangt van ATP binding.

Met 'rapid-quench' fluorescentie experimenten, werd duidelijk dat in het wildtype MutS en de mutanten E38T en E38Q de eerste, snelle, ATPase activiteit wordt geremd na binding van

een mismatch, maar dat dit niet het geval is voor de E38A mutant.

Oppervlakte plasmon resonantie (SPR) experimenten, waarin een oligonucleotide op de chip, aan beide kanten is voorzien van een blokkade, laat zien dat de E38A mutant na mismatch binding en injectie van ATP het verkeerd gepaarde DNA veel sneller loslaat dan het wildtype of de E38T en E38Q MutS. Deze gegevens suggereren dat de E38A de stabiele toestand niet vormt waarin het zowel het mismatch DNA als de ATP gebonden heeft. Deze toestand is nodig om erop volgende processen in de mismatch pathway te initiëren.

Tenslotte laten de kristalstructuren van de verschillende E38 mutanten van MutS met DNA dat een mismatch bevat, zien dat E38A de enige mutant is die geen waterstofverbinding met de mismatch kan maken. Daar de snelheid van ATP uitwisseling hetzelfde is voor alle mutanten en wildtype MutS kan worden geconcludeerd dat de waterstofbruggen tussen residu 38 en de foutief gepaarde base nodig is voor een effectieve doorgave van het signaal van het ATPase domein naar het mismatch bindingsdomein, welke volgt op nucleotide uitwisseling. Dit signaleringsprocess is opmerkelijk, omdat het plaats vindt over een afstand van 100 Ångstrøm. Het signaal leidt tot de juiste ATP toestand van het eiwit die nodig is om de reparatie op te starten.

Op deze manier gebruikt MutS de instabiliteit, die inherent is aan een foutieve paring van de basen, om de mismatch te binden. MutS gebruikt de ATPbinding om de herkenning van de mismatch te verifiëren. Door de remming van de ATPase activiteit van MutS, blijft de gevormde ATP-gebonden toestand voldoende lang in stand, zodat de reparatie kan worden gestart. Dit alles gaat samen met positieveranderingen van de domeinen, welke noodzakelijk zijn voor de functie van dit uiterst flexibele eiwit.

Appendix 1

Crystal structures of MutS mutants N616A and H728A in complex with a G.T mismatch.

Ganesh Natrajan, Herrie H K Winterwerp and Titia K Sixma.

Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands.

Introduction

An important part of our ongoing studies on the ATPase cycle of *E. coli* MutS involves crystallographic work on MutS and its mutants in complex with different nucleotides. We have studied a number of residues known to be important in the ATPase cycle of the ABC ATPase superfamily of proteins. Almost all of these residues are located in the three most important loops involved with ATP binding and hydrolysis, namely, the P-loop, Walker B motif, the signature loop, and the conserved histidine H728.

This appendix discusses the crystal structures of two mutants, N616A and H728A. N616 is the conserved asparagine that is located at the beginning of the P-loop, while the histidine H728 is a widely conserved residue in the MutS family and other ABC transporters. Both these mutant proteins have been reported to be defective in mismatch binding, ATPase activity, and tightening of the dimer upon ATP binding (1). To be able to understand the biochemical properties of these mutants better, we solved their structures in complex with a DNA containing a G.T mismatch, and different nucleotides (Table 1).

Materials and Methods

The mutated proteins were made as described previously (1). The crystallization in complex with a 30 basepair DNA was carried out as reported previously (2,3). N616A was co-crystallized with ATP and AMPPNP, instead of ADP used in all the previous structures. H728A was co-crystallized as before with ADP (3) for the ADP-Mg complex. The crystals of H728A-ADP were soaked for 1 minute in the cryobuffer containing 1 mM ATP prior to freezing, to obtain the H728A-ATP structure. Data were collected at the ID14EH1 (N616 structures) and the ID14EH2 (H728 structures) beamlines at the ESRF in Grenoble. The data were processed using MOSFLM and SCALA (4). The structures were solved as

previously described (3). The high resolution H28A-ATP structure was autotraced by ARP/wARP (5). All refinements were carried out using the REFMAC5 program (6), in the CCP4 suite (7). The structures were analysed using COOT (ref) and checked for errors using WHATCHECK (8). Figure 1 was made using PYMOL (© De Lano Software).

Results

MutS N616A was co-crystallized with ATP and AMPPNP to yield ADP-Mg in the ATP binding site of monomer A, the one which also binds to the mismatch. H728A was co-crystallized with ADP to yield ADP-Mg in its monomer A and an empty monomer B, a configuration observed in several of our previously reported MutS structures. The structure of the H728-ATP soak has ATP-Mg bound to the monomer A, and surprisingly, ADP bound to the monomer B. A notable feature of the H728A structures is that the signature loops of monomer B are visible, in both the ADP and ATP bound structures. Presented here are the results of further analyses of both these structures.

H728A structures

Histidine H728 in MutS is structurally important, as it hydrogen bonds with the N616 side chain, allowing for the gamma phosphate of the ATP to be accommodated (9). This conserved histidine, which is a glutamine in many other ABC ATPases, is by itself an important motif. Biochemical studies on the H728A mutant have revealed that (9) it has a highly reduced ATPase activity compared to the wild type protein and is also very defective in the ATP induced dimerization. This mutated protein is also toxic to the growth of cells, but is surprisingly capable of mismatch repair complementation in a MutS deficient strain, in an assay which tests for mutations conferring rifampicin resistance (data not shown). We have solved the structure of this mutant in two different nucleotide forms. Crystallization of this mutant under standard conditions (3) yields the complex with ADP-Mg in monomer A and no nucleotide in monomer B (Fig 1A). The mutation of H728 causes some interesting changes in the ATP binding site. The N616 interacts with the S668 of the opposing monomer, and the signature loop is fully visible. Since this is a requirement for the completeness of the ATP binding site, it can be said that the state of the protein represented in this structure is one in which it is primed to exchange the bound ADP for ATP.

Appendix 1

	H728A-ADP	H728A-ATP	N616A-ATP	N616A-AMPPnP
Resolution (Å)	20 – 2.5 (2.6 – 2.5)	20 – 2.1 (2.2 – 2.1)	20 – 2.8 (2.9 – 2.8)	20 – 2.3 (2.4 - 2.3)
I/σ(I)	10.6 (1.6)	6.1 (1.0)	4.7 (1.5)	10.5 (1.3)
Completeness(%)	99.3 (99.4)	100 (100)	94.4 (94.2)	99.8 (99.6)
Rmerge (%)	11.3 (76.3)	9.1 (60)	19.7 (82.8)	11.4 (94)
Rmerge (20 - 5 Å)	6.3	3.6	8.0	5.4
R (%)	20.9	20.3	23.5	19.6
Rfree (%)	26.2	24.7	28.8	24.4
Rmsd bond (Å)	0.013	0.012	0.007	0.014
Rmsd angle (º)	1.50	1.37	1.03	1.53
No of atoms	13191	13552	13089	13289

Table 1. Crystallographic data. Values in parenthesis are those of the highest resolution shell.

Upon soaking the crystal of H728A-GT mismatch-ADP-Mg into ATP, we obtained another structure of this mutant in complex with ATP-Mg. This structure has ATP-Mg bound in monomer A (Fig 1B) and an ADP in monomer B (Fig 1C). It has been shown that this mutant does not retain the ADP in the monomer A upon purification (1). So any nucleotide bound in the ATPase domains could have come only either as a consequence of co-crystallization (as in the ADP-Mg structure), or due to the hydrolysis of the ATP bound during the soak. The ADP in monomer B is clearly a result of the hydrolysis of a previously bound ATP, since it is empty in the H728A-ADP-Mg complex discussed earlier. This is the first structure of MutS in a mixed state i.e., with two different nucleotides bound in it.

In the monomer A of the H728A-ATP soak, the N616, now unable to make a hydrogen bond with residue 728, stabilizes itself by forming a hydrogen bond with the only other residue in its immediate vicinity, the visible S668 on the signature loop of the opposing monomer (Fig 1B). The binding of the mismatch and the other parts of the protein are very similar to the wild type structure, with no domain movements seen either in the H728A-ADP-Mg structure or in its ATP soak. In monomer B the ADP is seen bound without Mg, and the signature loop of the opposing monomer A is invisible. This is in contrast to the Mg free, ADP bound intermediate of the wild type MutS (Chapter 3 of this thesis), where the signature loop is fully visible and ordered. However, the N616 in the monomer B now makes a hydrogen bond to E694 (Fig 1 C), and the P- loop of this monomer itself is

pulled aside, away from the beta phosphate of the ADP. This suggests that this ADP may be loosely bound. Interestingly, it has been proposed before that the ADP.MutS.ATP state (ADP in monomer B, ATP in monomer A) may be the one in which MutS binds to the mismatch (10). The previously proposed result (10) that the binding of a mismatch inhibits ATP hydrolysis in the high ATP affinity monomer while ATP hydrolysis continues slowly in the other monomer, may also explain the ADP in monomer B if monomer A is assumed to be the one with high affinity. If so, it is indeed strange that this state must be revealed in the crystal structure of a mutant that has highly reduced steady-state ATPase activity compared to the wild type. The ongoing studies on the ATPase cycle of this mutant will reveal more about its properties.

N616A structures

The MutS mutant N616A shows a drastic reduction in steady state ATPase activity, dimerization upon ATP binding, and exhibits a severe mutator phenotype in-vivo (1). Previous studies have shown that this residue enables the binding of the gamma phosphate by coordinating with H728 when ATP is bound (9). In the structure of a Mg-free intermediate of MutS, the N616 moves into the void created by the absence of Mg to coordinate with H728 and the beta phosphate, enabling the signature loop to move in (Chapter 3 of this thesis). From both these structures, it can also been seen that the proper orientation of N616 is also needed to position the signature loop correctly, to complete the ATPase active site. An interesting observation in all these structures is that the conformation of the P-loop itself does not vary very much. The different nucleotide binding states are simply achieved by moving the N616 side chain around. As mentioned earlier, the crystals of N616A were grown separately in presence of ATP and AMPPnP. The final structures (Fig 1 D), show an ADP-Mg bound to the monomer A's ATPase site, while that of the monomer B is empty. This could mean that either the ATP/AMPPnP hasn't been able to bind the protein, or has actually been slowly hydrolyzed after binding. Given that the N616A possesses much reduced ATPase activity and ATP binding affinity compared to the wild type (1), the latter explanation seems more plausible. The former explanation is unlikely since the co-crystallizations and soaks have been done in a vast excess of nucleotide.
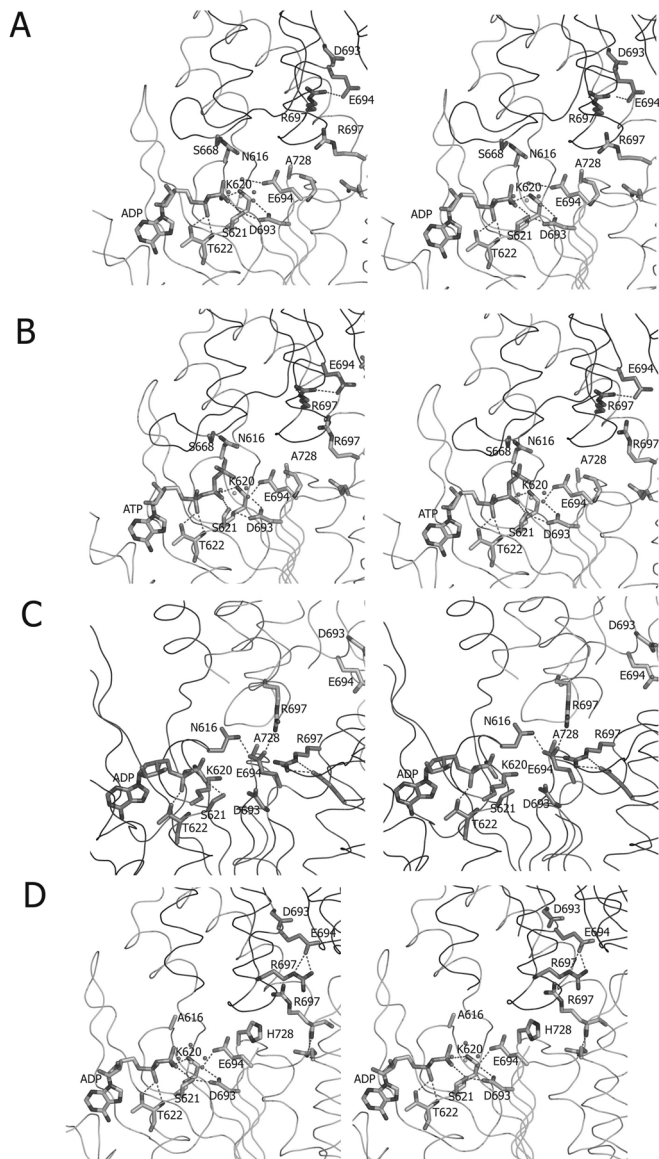
Figure 1. ATP binding site of (A) monomer A in the structure of H728A-ADP, (B) Monomer A of H728A-ATP, (C) Monomer B of H728-ATP and (D) Monomer A of N616-AMPPnP. The Mg is shown as a green sphere, and its coordinating waters as red ones. All figures are stereoviews. The original colour figure can be viewed in http://ganeshthesisfigures.org.

Appendix 1

The signature loops of both monomers are also invisible, which is expected given the importance of the N616 sidechain orientation for this. The mismatch binding is identical to the wild type (2) and no domain movements are seen between the wild type and the mutant structures.

Discussion

Here we present crystal structures of ATPase mutants of *E. coli* MutS. These structures have undergone conformational changes in the ATP binding site. The structure of MutS N616A shows a disordered signature loop in the presence of Mg.ADP. In vivo, N616A exhibits a severe mutator phenotype (9), which may be attributed to the reduction in ATPase activity. Absence of H728 allows ordering of the signature loop in the presence of ADP-Mg, which is normally not observed. Soaking the ADP-Mg crystal into ATP has created a mixed ATP/ADP state, which is thought to be important in the MutS cycle (10). These structures are therefore intriguing, especially since this mutant protein is defective in ATP binding and hydrolysis, but is competent for DNA repair. For complete understanding of this defect and the implications of the H728A structures, additional biochemical data are required. The difference in complementation of a MutS defective strain can not be explained yet by the structures and requires additional biochemical analysis.

References

1.   Lamers, M. H., Georgijevic, D., Lebbink, J. H., Winterwerp, H. H., Agianian, B., de Wind, N., and Sixma, T. K. (2004) *J Biol Chem* **279**(42), 43879-43885

2.   Lamers, M. H., Perrakis, A., Enzlin, J. H., Winterwerp, H. H., de Wind, N., and Sixma, T. K. (2000) *Nature* **407**(6805), 711-717.

3.   Natrajan, G., Lamers, M. H., Enzlin, J. H., Winterwerp, H. H., Perrakis, A., and Sixma, T. K. (2003) *Nucleic Acids Res* **31**(16), 4814-4821

4.   Powell, H. R. (1999) *Acta Crystallogr D Biol Crystallogr* **55 (10)**, 1690-1695

5.   Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat Struct Biol* **6**(5), 458-463

6.   Murshudov, G. N., Vagin, A.A, & Dodson, E.J. (1997) *Acta Crystallogr D* **53**, 240-255

7.   CCP4. (1994) *Acta Crystallogr D* **50**, 760-763

8.   Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature* **381**(6580), 272.

9.   Lamers, M. H., Winterwerp, H. H., and Sixma, T. K. (2003) *EMBO J* **22**, 746-756

10.   Antony, E., and Hingorani, M. M. (2004) *Biochemistry* **43**(41), 13115-13128

Curriculum Vitae

Ganesh Natrajan was born on the 14th of April, 1973 in Alleppey, Kerala, India. In 1988, he completed his secondary schooling at the St. Anthony's High School in Hyderabad, India. Upon completing his Master's degree in Physics in 1996 at the University of Hyderabad, India, he enrolled at the Indian Institute of Science in Bangalore as a Research Scholar. In 1999, he moved to The Netherlands to begin his PhD in the laboratory of Prof Titia K Sixma at the Netherlands Cancer Institute in Amsterdam. His research has been primarily concerned with the use of X ray crystallographic structure determination, coupled with other biochemical assays, to study the *Escherichia coli* DNA mismatch repair enzyme MutS. Following the submission of this thesis, he will be working as a Post Doctoral Fellow in the lab of Dr Laurent Terradot-Piot in the Macromolecular Crystallography Group, at the European Synchrotron Radiation Facility in Grenoble, France. The research presented in this thesis was carried out between December 1999 and December 2005.

List of Publications

G.Natrajan, M.H.Lamers, H.H.K.Winterwerp, A.Perrakis, T.K.Sixma (2002). Structural basis for DNA mismatch recognition by the enzyme MutS. *Acta Crystallographica,* **A58 (Supplement),** C275 **(Abstract only)**.

Ganesh Natrajan, Meindert H. Lamers, Jacqueline H. Enzlin, Herrie H. K. Winterwerp, Anastassis Perrakis and Titia K. Sixma (2003). Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Research*, **31,** 16, 4814-4821.

Joyce H.G. Lebbink*, Dubravka Georgijevic*, **Ganesh Natrajan***, Alexander Fish, Herrie H.K.Winterwerp, Niels de Wind and Titia K. Sixma (2006). Dual roles of MutS Glutamate 38 in DNA mismatch discrimination and authorization of repair, *EMBO Journal*, **25**, 2, 409-419. (* Shared first author).

Ganesh Natrajan, Joyce H.G.Lebbink, Alexander Fish, Herrie H.K.Winterwerp and Titia.K.Sixma (2006). A Mg free intermediate in the ATPase cycle of Escherichia coli MutS revealed by x-ray crystallography, (Manuscript submitted for publication)

ACKNOWLEDGEMENTS

It is often said that every long journey has to begin with a small step. It has been a long time since I took that one, and finally, the end of my PhD is at hand. As I look back at the long and winding road I have traversed, I get reminded about what a rocky path it has been and how, without the help of some beneficent individuals, I would have never known what lay at the end. This is a small tribute to those wonderful people I've had the privilege of meeting and who have held out the proverbial branch of hope, in what had become a very turbulent and difficult voyage.

Titia, I still remember the day I first looked at your lab website. Upon clicking on "Projects", I was faced with a white screen with just one line written on it :"We do not want to discuss the state of our projects. It is disturbing us". I still decided to apply for a PhD position in your lab. It was as brave a decision for me as it may have been for you, who finally decided to hire me in spite of my complete lack of experience in anything concerning proteins. In retrospect, it has been one of the best decisions I have made in my life. You have been the best teacher, friend, supporter and boss anyone could ever ask for. Your constant constructive criticisms and unsurpassable knowledge have always motivated me to do better and keep improving. In six years, your guidance has changed me from an utter novice to a confident scientist full of pride. I can never thank you enough.

Meindert, from teaching me how to run a gel to helping me refine my first structures at a time when even you were struggling to get any results, you have been a fantastic colleague. This thesis may not have happened without your help, not to speak of what could have been if you hadn't jumped, Baywatch style, to save me in what turned out to be five feet of water. It has indeed been great having worked with you. Herrie, whom I will always remember as the untiring protein factory, you have been and continue to be, something of a backbone to the entire mismatch repair effort. Joyce, although being a relative newcomer, your formidable biochemical knowledge, sharp insight and your uncanny ability of being able to handle fifty projects at once, have been invaluable in making many

of our crystal structures contribute to the big picture. I hope you will have your own lab very soon, where your students will be able to benefit from your immense knowledge and experience. Sasha, from the middle eastern conflict to the KGB, to the Biacore analysis of MutS, our discussions have been quite wide ranging. It has been great having you around and contibuting so much even in your first few months. I wish you well, and hope you have a very productive stay in Amsterdam. Dubravka and Niels, though being in Leiden, your contributions have been vital to the mismatch repair research in our lab. I'm sure this fruitful collaboration will continue to prosper.

The invigorating environs of the H2 in the NKI is due, in no small measure, to some of the people who work behind the scenes, far from the laudatory limelight. Pim, you are more than the director of the lab, you are its very soul. Thanks for being such a wonderful manager and taking such good care of everything. Marlijn, thanks for being everything you were: the proverbial 'everything' that is. No doubt, you are the most missed person in the department.

Tassos, it is great having you around, helping everybody and being the highly philosophical, if somewhat temperamental, Greek God. I thank you for sharing your colossal knowledge of crystallography with me. You are certainly the fulcrum of all the crystallography in the lab, the first contact and the last hope, should any difficult structure show up. I wish you and your lab all the very best in the years to come.

One of the best aspects of the lab has been the wonderful people who have worked there over the years, whom I have had the honor of interacting with. Katjusa, every time I look at a bottle of acrylamide, it will be you who will come to my mind; Patrick, the very epitome of perfection and a Slotervaart Scientist to the core; Puck, the life of our borrels; Valerie, one of the wittiest I've met; Gretel, a wonderful friend, a great person to talk to and one I still hope to beat at squash some day; Victor, Steve, Jacqueline, Danielle, Dave, Sari, Chris(tal), Mark, Alex and the latest addition in the form of Annet. All of you have and still do make the lab such a wonderful place to be in. A great addition to the crystallography effort has been in the form of the Perrakis group: Olly, Nuno, Rebecca, Kostas, Koen, Serge, Marouane, Diederick, Christa,

Vangelis, Valeria, Suzan, Mark (H), Angelina. The place has become a lot more exciting thanks to all of you.

Life on H2 would have been, without doubt, a lot duller without the presence of the other groups. The Rene 'Bernies', the Beijersbergens, the microarray chaps, and the erstwhile groups of Maarten and Hein have all contributed to and still make H2 one of the best departments in the building. Rene, Hein, Maarten, Avi, Blanca, Daniel, Reuven, Maria-Carla, Willem, Marijke, Jacqueline, Jan-Herman, Karin, Petra (K), Sandra, Leone (it was great to share that Louweshoek flat with you), Marleen, Connie, Katrien, Jeroen, Sirith, Hans, Thijn, Marielle, Rianne, Marieke, Menno (Thanks for those sweet Dutch words ;-)), Sebastian, Mirjam, Mike, Pieter, Roderik (K), Benjamin, Annette, Ron, Roderik (B), Wouter, Sid. Thanks to all of you for your help, and those wonderful parties we had.

The NKI is a temple of learning, full of the wisest and smartest people. It has been a privilege to be a part of such an institution. In addition to all the people I have known and worked with here, I'm fortunate to have made some good friends and met some really knowledgeable people. Rafa, Vasso, Zhong, Paul, Nal, Sjaak, Jon, Henri, Adriaan, Annelies, Arlette, Alex (G) and the other members of the H4 'Mannenkamer', and all the other OIO's that went on the retreats to Texel. It was great being with all of you.

Amsterdam is admittedly, a bit of a culture shock for a guy coming from a faraway Asian land. What's better than having a bunch of like-minded expats to discuss the ways of the Dutch with, and make lasting friends in the process. John, Tim, Richie, Mark, Christina, Amy, Paula, Ramon, Dannii. It was great meeting all of you. You truly gave me something besides work to think of. Christiane and Paulo, you deserve special mention for being a wonderful couple. My stay here would have been a lot less rewarding but for some great people I stayed with: Tilman, Vic, Mario, Hellen – I felt most at home in your place.

A big contributing factor in any achievement, of course, is the support one gets from one's own family. Dad and Amma, I hope I have become the son you always wanted me to be. Thanks for all your support and encouragement over the years. It would not have been possible without you. Kanchi, Raju, Venkatesh and Mahesh,

thanks for all your love and encouragement. I would have been much the worse without it.

If a person's life is really like a journey, mine has been a voyage of discovery straddling two diverse continents and cultures. PhDs may end, but the learning never stops. In my case, I firmly believe that it has just begun. It will be fitting to end this with these beautiful words by Robert Frost, which have inspired me ever since I first read them many years ago.

The woods are lovely,
Dark and deep.
But I have promises to keep,
And miles to go before I sleep,
And miles to go before I sleep.

Ganesh Natrajan.