HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2012 July 01.

Published in final edited form as:

Nat Struct Mol Biol. ; 19(1): 72–78. doi:10.1038/nsmb.2175.

Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops

Shikha Gupta, **Martin Gellert**, and **Wei Yang**

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 9000 Rockville Pike, Bldg. 5, Rm. B1-03, Bethesda, MD 20892, USA

Abstract

DNA mismatch repair corrects replication errors, thus reducing mutation rates and microsatellite instability. Genetic defects in this pathway cause Lynch Syndrome and various cancers in humans. Binding of a mispaired or unpaired base by bacterial MutS and eukaryotic MutSα is well characterized. We report here crystal structures of human MutSβ complexed with DNA containing insertion-deletion loops (IDL) of 2, 3, 4, or 6 unpaired nucleotides. In contrast to eukaryotic MutSα and bacterial MutS, which bind the base of a mismatched nucleotide, MutSβ binds three phosphates in an IDL. DNA is severely bent at the IDL; unpaired bases are flipped out into the major groove and partially exposed to solvent. A normal downstream basepair can become unpaired; thereby a single unpaired base can be converted to an IDL of 2 nucleotides and recognized by MutSβ. The C-terminal dimerization domains form an integral part of the MutS structure and coordinate asymmetrical ATP hydrolysis by Msh2 and Msh3 with mismatch binding to signal for repair.

Keywords

mismatch repair; asymmetric dimer; Msh3; ATPase; trinucleotide repeat

INTRODUCTION

Nucleotide misincorporation or strand slippage at repetitive sequences during replication results in mispaired or unpaired DNA bases. Mismatched bases are recognized by MutS, which in the presence of ATP recruits MutL to initiate the repair process [1—4]. Eukaryotes

Correspondence and requests for materials should be addressed to W.Y., Wei.Yang@nih.gov, phone: (301) 402-4645, fax: (301) 496-0201.

Author information

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession codes 3THY (Loop2), 3THX (Loop3), 3THW (Loop4) and 3THZ (Loop6). Information on reprints and permissions is available at www.nature.com/reprints. The authors declare no competing financial interests.

Author contributions

S.G. conducted all experiments and collected X-ray data. W.Y. determined and refined the structures. All authors contributed to the experimental design, data interpretation and manuscript preparation.

have two MutS homologs, α and β. MutSα, a heterodimer of MutS homologs Msh2 and Msh6, recognizes a base mispair or 1—2 unpaired bases, like homodimeric bacterial MutS [1]. In contrast, MutSβ, a heterodimer of Msh2 and Msh3, recognizes insertion-deletion loops (IDL) of 1—15 nucleotides, as well as DNA with a 3′ single-stranded overhang [5—8]. Inactivation of MutS or MutL by mutation, or reduced expression of human MutLα due to promoter hypermethylation, leads to increased mutation rates and microsatellite instability [9—13]. In humans such defects are correlated with susceptibility to hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch Syndrome [14]. Loss of MSH3 in tumor cells is correlated with increased microsatellite instability of tri- and tetra-nucleotide repeats ([15] and references therein). Mice with Msh3 knocked out develop cancers only late in life, but a double knockout of Msh3 and Msh6 renders mice far more susceptible to cancer than the single knockout of either gene [11,13].

Mechanistically, MutSα functions similarly to bacterial MutS, by inserting the mismatch-binding domain (MBD) of Msh6 into the minor groove at the mismatch site and pulling a mispaired or unpaired base toward it to form π-stacking with a conserved phenylalanine [16—18]. As a result, normal base stacking is disrupted and the DNA is bent 45—60°. Although MutSβ recognizes 1—2 unpaired nucleotides as does MutSα, and Msh3 is homologous to MutS and Msh6, Msh3 lacks this conserved Phe. Mutation of the Phe in Msh6 or replacement of the corresponding residue in Msh3 (Lys) by Phe diminishes the ability of these MutS homologs to bind all forms of mismatched DNA [19—21]. Although Msh2 is present in both MutSα and MutSβ, the MBD of Msh2 is required for mismatch binding only by MutSβ, not by MutSα [22,23]. Extensive mutagenic studies have revealed residues critical for IDL recognition and led to the conclusion that MutSβ recognizes mismatches very differently from MutS and MutSα [8,21].

MutSβ is involved in more than mismatch repair. In yeast, MutSβ participates in double-strand break repair and binds a 3′ overhang as well as a 3′ flap formed during double-strand break repair by single-strand annealing [8,24—26]. In mammals, MutSβ is required for the mutagenic expansion of trinucleotide repeats and may thus be a culprit in myotonic dystrophy, fragile X syndrome and Huntington's disease [27—29]. Trinucleotide repeats can form unique IDLs and be bound by MutSβ. While short IDLs in these repeats are repaired efficiently by MutSβ, long or clustered IDLs are refractory to repair [30—32], and attempted repair may lead to expansion.

In order to clarify the functions of MutSβ, we report here crystal structures of mismatched DNA complexed with MutSβ that include the previously missing C-terminal dimerization domains present in all MutS functional homologs. The molecular mechanism of recognition of a variety of IDLs, and of the ATP-mediated signaling for later steps of mismatch repair, is presented.

RESULTS

Structure of MutSβ—DNA complexes

A form of human MutSβ consisting of full-length MSH2 and trimmed MSH3 (211—1125 aa, abbreviated as MutSβ hereafter) was generated for structural studies (see Methods). It

retains the ATPase activity and binding affinity for IDLs of the full-length species (Supplementary Fig. 1). MutSβ binds mismatched DNA with sub-nanomolar Kd's, which are much lower than the values previously measured in the presence of competitor DNAs [33,34]. MutSβ was co-crystallized with IDLs of 2, 3, 4, or 6 unpaired bases flanked by 10—12 bp duplexes, in three different crystal lattices (Table 1). These structures, named according to the IDL size as Loop2, Loop3, Loop4 and Loop6, were determined and refined (Methods, Supplementary Fig. 2). Only 1—2 nucleotides at the DNA ends, some residues at the extreme N- and C-termini, and a few solvent-exposed internal loops are disordered. Even at the low resolution of 4.3Å, the DNA and protein domains are well defined in Loop6. Except for the DNA-binding domains, the four MutSβ structures are superimposable.

The overall shape of the MutSβ —IDL complexes is similar to those of human MutSα and bacterial MutS—DNA complexes [16—18] (Fig. 1a). The five structural domains (I—V) are conserved in Msh3. Domains I (MBD) and IV (clamp) are involved in DNA binding (Fig. 1b). Domain V contains the ABC-family ATPase. Domains II and III connect the ATPase domain (V) to the two DNA-binding domains (Fig. 2a—b). The MSH2 subunit in MutSβ differs remarkably from that in MutSα, with domain I shifted by 5.3Å and rotated 21° (Fig. 2a). In contrast, MSH3 and MSH6 superimpose well, with an rmsd of 2.2Å over 726 pairs of Cα atoms; MSH3 also superimposes well with bacterial MutS except for domain IV (Fig 2 b —c). MSH2 and MSH3 dimerize between both domains I and V and form an elongated oval with the DNA bound near one end and the ATPase site at the other (Fig. 1a, 2d). Although no ADP or ATP was added during purification or crystallization, an ADP molecule is bound to MSH2 (Fig. 1a). This suggests that ADP binds MSH2 more tightly than previously reported [35], and may be routinely co-purified with MutSβ.

The MutSβ structures reveal two new features. Firstly, the dimerization domains (DMDs) at the C-termini, which were excluded or disordered in all previously known MutS structures, are observed atop the rest of MutS for the first time (Fig. 1a). They strengthen the heterodimer and establish the asymmetry of the nucleotide binding by MSH2 and MSH3 (see later). Secondly, the degree of DNA bending and the mechanism of mismatch recognition by MutSβ differ appreciably from bacterial MutS and MutSα. DNA is severely bent at an IDL because the extra nucleotides in one strand disrupt normal base stacking [36]. The bending angle increases from 90° in Loop2 to 120° in Loop6 (Fig. 1b). The bases in the IDL are rotated away from MutSβ towards the solvent exposed major groove. They often stack with one another and roughly maintain their 3.4Å spacing.

Recognition of IDLs

The IDLs interact chiefly with domain I of MSH3 (MBD) and partially with domain I of MSH2. Although the structure of the MBD is identical among MutS homologs (Fig. 2e), and MSH3 inserts its MBD into the DNA minor groove like MSH6 and bacterial MutS (Fig. 3a), changes of six amino acids in MSH3 on two DNA-binding loops (Supplementary Fig. 3) fundamentally alter the interactions of MutSβ with a mismatch site.

Instead of forming π-stacking with a mismatched base, MutSβ uses a group of residues to bind and distort the phospho-sugar backbone of an IDL. MSH3 replaces the π-stacking Phe of MSH6 and bacterial MutS with a Lys (Lys246) and acquires a Tyr (Tyr245) immediately

before it. This Tyr-Lys pair, conserved among Msh3 homologs, burrows into DNA at the IDL site, widens the minor groove and breaks the helix into three segments: the IDL and two flanking duplex arms (Fig. 3a—c). Each DNA consists of a plus strand that has additional nucleotides and a complementary minus strand. Tyr245 is inserted between the adjacent nucleotides on the minus strand opposite the IDL and bends the strand sharply (Fig. 3c—d). It interacts with the base pair immediately upstream of the IDL and forms a π-stack with the downstream minus strand in the Loop2 and Loop3 structures. The IDL on the plus strand bends more smoothly. Lys246 along with Ser275, which replaces a conserved Gly in Msh6 and MutS, form numerous interactions with the first phosphates in an IDL inside the widened minor groove. Through the Tyr-Lys pair, MutSβ binds an IDL with strand and orientation specificity. Mutations of the Tyr-Lys pair or the Ser275 equivalent in *Saccharomyces cerevisiae* (yeast) MSH3 cause a mutator phenotype and microsatellite instability [21,22].

The mainchain atoms of Ile276 and Pro277 and sidechains of His279 and Arg280 (unique among Msh3 homologs), fix the second and third phosphates of an IDL (or the downstream nucleotide in Loop2). Although Pro277 is highly conserved among MutS homologs, it contacts and stabilizes the π-stacking Phe instead in MutS and Msh6. Interestingly, mutations of the equivalent of Pro277 in yeast MSH3 and MSH6 lead to very different levels of repair deficiency [22]. His284 of Msh3, which replaces Tyr in Msh6 and MutS (Supplementary Fig. 3) and is critical for IDL repair [21], makes favorable interactions with the third phosphate of an IDL and avoids steric clashes (Fig. 3). The distorted IDL is further stabilized through phosphate backbone interactions by Ile263, Tyr264 and His266 on a β strand outside of the minor groove (Supplementary Fig. 4). As a result, the positions of the first three phosphates of each IDL are the same in all four structures.

When an IDL is longer than 3 nucleotides, Phe42 of MSH2 (an equivalent to the π-stacking Phe in Msh6 and MutS) forms a π-stack with the fourth unpaired base. Asp41, Lys65, Val79 and Ser81 of MSH2 interact with the plus strand (Fig. 3d, Supplementary Fig. 4). These interactions are absent in the MutSα—DNA complex structures and explain why domain I of MSH2 is required for MutSβ to bind an IDL [22,23]. Domains I of MSH2 and MSH3 interact directly (Fig. 2d) and maintain the same interactions in all four MutSβ structures (Fig. 1b, 3d). Different IDLs are fitted into this framework by bending at different angles. Domains IV of MSH2 and MSH3 barely contact one another and move independently (Fig. 2a—c), which gives rise to an “accordion” effect and accommodates different sizes of IDL with different DNA bending angles (Fig. 3d).

Recognition of IDLs is aided by substantial contacts between DNA duplexes and MSH3. A loop from residues 300 to 319 in the MSH3 MBD interacts with the upstream duplex in the major groove, covering 8 bps (Fig. 3a—c). The short helix on this loop interacts with the helical arm of domain IV (Fig. 2b) and coordinates the binding of the downstream duplex by the globular head of domain IV, which is hydrogen bonded with three consecutive phosphates on the plus strand. No interaction with the minus strand is observed in the downstream duplex. Our structures fully agree with the DNA footprinting analyses [8]. The interaction of MSH3 with both strands upstream of the IDL, but only the plus strand downstream, explains why MutSβ can bind a DNA end with a 3′ overhang [24—26]

(Supplementary Fig. 1), because the overhang can be the surrogate of an IDL and the downstream plus strand.

Isomerization of mismatched DNA

MSH3 has little contact with the IDL bases, and these unpaired bases adopt different conformations in the four structures. Loop2, Loop4 and Loop6 have 1, 2 and 3 CA dinucleotide repeats, respectively. In these structures the first C is flipped out and solvent exposed, and the second, third, and fourth bases, if present, are stacked with each other (Fig. 3d). But in Loop3, where the IDL sequence is ACA, the 5′ A becomes stacked with the other two unpaired bases and contacts the sidechain of Lys246. In Loop6, the fifth base slips out and the sixth is stacked with the downstream duplex (Fig. 3d). Additional unpaired nucleotides may slip out between the fourth and sixth base of an IDL as observed in Loop6.

In Loop3 a normal G:C basepair immediately downstream of the IDL becomes unpaired ~50% of the time, and the G is flipped out and stacks with Phe42 of MSH2 like the fourth unpaired IDL base in Loop4 and Loop6 (Supplementary Fig. 4, 5). Conversely, the fourth IDL base in Loop4 has a second minor conformation of stacking with the downstream duplex, like the sixth IDL base in Loop6 (Supplementary Fig. 2b, 4, 5). Conversion of an unpaired base to a mismatched basepair was also observed with *E. coli* MutS [37]. The consistent interactions with the first 3 phosphates of the IDL in all four structures indicate that MutSβ prefers to bind an IDL of at least 2 unpaired bases. Interconversion of paired and unpaired bases suggestion that an unpaired base (a substrate of MutSα) may be converted to a 2-base IDL and recognized by MutSβ. Furthermore, one unpaired base in mononucleotide repeats may be isomerized to a larger loop by creating a bulge at a distance on the opposite strand (Supplementary Fig. 5).

The C-terminal dimerization domain (DMD)

The C-terminal end of MSH3 forms two α helices linked by a tight turn, and MSH2 has a third short helix in addition (Fig. 4a—b). They dimerize predominantly by hydrophobic interactions and bury 1195Å^2 of surface area. The five-helix bundle is further strengthened by salt bridges within each subunit and N and C capping of the MSH3 helices by MSH2. Preceding the helical appendage is a Helix-turn-Helix (HtH) motif, which extends from one subunit and embraces the ATPase domain of the other in all MutS homolog structures (Fig. 5a). In MSH3 the second helix of HtH and the first helix in the dimerization domain are merged into one 55Å long helix, resulting in a three-helix unit (Fig. 4c). Based on secondary structure prediction, a similar three-helix unit also exists in MSH6. We therefore re-define the dimerization domain (DMD) to include the preceding HtH. Each DMD is thus bivalent and interacts with the ATPase domain and the DMD of another subunit. In MSH2 the two halves of the DMD are linked by a disordered 15-aa loop and are juxtaposed rather than opposite to each other (Fig. 4c). The two DMDs are thus structurally different and do not follow the dyad axis that relates domains I—V of MSH2 and MSH3. The DMDs lean towards MSH3 and predominantly shield the MSH2 ATPase site (Fig. 1a, 5b).

The structure of the last 34 residues of *E. coli* MutS, which were deleted to facilitate crystallization of MutS—DNA complexes [17], was determined as a fusion to maltose-binding

protein [38]. It contains two α-helices and dimerizes by forming a four-helix bundle. These last two helices of *E. coli* MutS can be roughly superimposed on those of MSH3, but its dimerization partner is ~ 10Å away from MSH2's DMD (Supplementary Fig. 6). As in MSH2, the two halves of the *E. coli* MutS DMD are linked by a 20-aa flexible linker [38]. Without the second half of the DMD, *E. coli* MutS is predominantly dimeric, but with it MutS becomes a tetramer or even an octamer (R. Ghirlando, S. Ramon-Maiques and W. Yang, unpublished data). In contrast, Taq MutS, which is predicted to have a short linker (4 aa) in its DMD, is a stable dimer in solution [39]. The short linker may prohibit the crosslinking effect of the two halves of the DMD and restrict its interacting partners within a single dimer.

In addition to stabilizing a MutS dimer, the DMD is the conduit for communication of nucleotide (ADP or ATP) binding between the two subunits. Each ATPase site of MutS is composite and consists of the N-1, N-3 and N-4 nucleotide-binding motifs from one subunit, and the N-2 from the partner subunit [2,40] (Fig. 5a). Within each subunit, the N-terminal HtH of the DMD contacts the nucleotide-binding site directly via its conserved SYG (Ser, Tyr or Phe, Gly) motif [41] and also indirectly by interacting with the trans-acting N-2 region of the other subunit (Fig. 4c—d, Supplementary Fig. 7). The MSH6 DMD alters its conformation upon ADP binding, and the movement of the MSH6 DMD is coupled with the ATPase domain of MSH2 [18] (Supplementary Fig. 7c—d). Interestingly but not surprisingly, the ATPase domain and DMD of MSH3, which is devoid of ADP, are superimposable with ADP-free MSH6 (Fig. 5c). ATP binding can alter the structure of the DMD in the same subunit and possibly cause a sliding of the DMD relative to the N-2 of the other subunit (Fig. 4d). Because movement of the DMDs of the two subunits is coupled (Fig. 5b), binding of ATP by one subunit can thus alter the ATP-binding site of the other.

The communication of the two ATPase domains through the DMD is substantiated by the G1142D mutation in yeast MSH6's SYG motif, which not only reduces ATP binding by MSH6 but also inhibits ATP binding by MSH2 [40]. The roles of the DMD in dimerization and bridging of two ATP-binding sites also explain why a complete DMD is required for the cellular functions of MutS in *E. coli* [42]. Furthermore, mutations in the MSH2 DMD have been found among HNPCC families, which highlights its importance in mismatch repair (www.insight-group.org) (Supplementary Fig. 3).

An occluded MSH3 ATP-binding site

Both MutSβ and MutSα display asymmetric ATP binding and hydrolysis in the two subunits. Msh3 and Msh6 appear to have higher ATPase activity than Msh2 without DNA, but they have low affinity for ATP and greatly inhibited ATPase activity after binding to a mismatched DNA [34,35,43—45]. Structural studies of MutSα—DNA complexes confirmed that MSH6 has lower affinity for ADP than MSH2 but without revealing the reason [18]. In all four MutSβ structures the ATPase site of MSH3 is devoid of nucleotide (Fig. 5a). Addition of ADP or ATP in the crystal-soaking buffer reduces the X-ray diffraction quality of MutSβ — DNA complexes. When bound to a MutS homolog, the adenine of ADP is sandwiched between a pair of conserved aromatic residues (in bold) in the **Y**UP (**Tyr** or **Phe**, Ile or Val, Pro) and FL**Y** (PheLeu**Tyr**) motifs (Phe650 and Tyr815 of MSH2) (Fig. 3c). In MSH3 the

non-conserved residues surrounding YUP and FLY change the peptide backbone conformation and close the nucleotide-binding site.

In the absence of a nucleotide, the P-loops of all MutS homologs assume an α–helical conformation by repositioning a Gly in the Walker A motif (**Gly**GlyLysSer, Gly891 in hMSH3) and are closed for phosphate binding. In bacterial MutS (PDB: 1EWQ and 1E3M) and human MSH6 (2O8E), the YUP and FLY motifs stay put, leaving the adenine-binding site open. In MSH3, Tyr868 of YUP occupies the adenine-binding site because of an insertion of GluGlnAspGln before YUP and an altered peptide conformation (Fig. 5c). The hydroxyl group of Tyr868 contacts the carbonyl oxygen of Gly891 and further stabilizes the nucleotide-free state. The importance of the YUP motif is emphasized by the high frequency of HNPCC mutations in this region, including E647K in MSH2 that immediately precedes the YUP[46].

FLY in MSH3 adopts a mainchain conformation different from other MutS homologs, with or without ADP (Fig. 5c). This is due to the replacement of a small sidechain (Ala or Cys) in bacterial MutS, eukaryotic Msh2 and Msh6 by a Phe conserved among Msh3 homologs (Phe1023 in hMSH3, Supplementary Fig. 3). Phe1023 occupies the space normally occupied by the Phe of the FLY motif (Phe1046 of hMSH3), causing FLY to alter its course and close the ATP-binding site. Phe1023 also links the adenine stacking FLY (Tyr1048) and the SYG motif (Tyr1059) in the DMD and thus stabilizes the asymmetric nucleotide binding by MutSβ. Interestingly, in the FLY motif Tyr is replaced by His in bacterial MutS, and Phe in Msh3 and Msh6 is replaced by a more flexible aliphatic sidechain in Msh2. These changes may give rise to the unique ATP hydrolysis properties of each MutS homolog. The occluded ATP-binding site of MSH3 provides the structural basis for its reduced ATPase activty, which may underlie the absence of a burst phase of ATP hydrolysis when MutS or MutSα is bound to a mismatch DNA [45,47].

DISCUSSION

Structures of MutS, MutSα and MutSβ complexed with mismatched DNA illustrate the flexibility of a lesion site, which can have different bending angles with bases flipped into either the major or minor groove. Parallel to the structural results, replacement of the MBD in yeast MSH6 by the MBD of MSH3 leads to a chimeric MSH2—MSH6 or MSH2—MSH3 dimer, which recognizes IDLs like MutSβ but signals for mismatch repair like MutSα [23]. MutS proteins appear to mold a mismatched DNA into different shapes. We suspect that in the normal mismatch repair process the exact conformation of DNA is not critical for repair signaling. The main function of a lesion site is its readiness to bend and deform [36] thus providing a foothold to the MBD of Msh3 or Msh6. Outside of mismatch repair, the exposed IDL bases in complex with MutSβ may play an important role for recruiting nucleases [8,26] or facilitate unusual hairpin formation in trinucleotide repeats, thus leading to abnormal expansions [28,29].

The asymmetry of ATP binding by the two subunits of MutSβ mirrors their structural differences and asymmetric DNA binding. Firstly, MSH3 clearly dominates DNA binding, using both its MBD and clamp domains (Fig. 3c). The DNA is off center and biased toward

MSH3. Similarly in the crystal structure of MutSα, MSH6 also dominates the DNA binding of the short DNA used [18]. Secondly, MSH2 and MSH3 are structurally different. In MSH3, domains I and II are extended towards a mismatched DNA and associated intimately with domains III, IV and V (Fig. 2a—b). The tight domain association, which is conserved among Msh3, Msh6 and the mismatch-binding subunit in bacterial MutS (Fig. 2c, Supplementary Fig. 3), is coupled with the strong inhibition of ATPase activity of these subunits upon binding to a mismatched DNA [43—45]. Meanwhile the loose domain association of MSH2 in MutSα and MutSβ is linked to the moderate change of its ATPase activity upon DNA binding.

Mutational studies of yeast MutSα indicate that ATP binding by Msh2 is essential for recruiting MutLα, while ATP binding by both subunits releases a bound DNA [40,43]. Combining the existing data, we propose the following model of mismatch-repair initiation in humans (Fig. 6). In the absence of DNA, the four DNA binding domains of MutSα or MutSβ move around freely [16]. When bound to a normal DNA, which is resistant to deformation, domains VI close and form a clamp but domains I remain flexible. Both MutS subunits can bind ATP resulting in the dissociation or sliding of the protein along DNA [48]. In the presence of a mismatch, the MBD of Msh3 or Msh6 lodges into the lesion site and induces a conformational change in its ATPase domain and DMDs. Consequently, nucleotide is occluded from Msh3 or Msh6, and ATP-binding by Msh2 is promoted. With Msh3 or Msh6 bound to a mismatch and Msh2 to ATP, either MutSα or MutSβ can recruit MutLα and activate mismatch repair. This model provides a detailed molecular picture that can help guide future experimentation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We thank Drs. J. Jiang for collecting X-ray diffraction data of Loop4, A. Howard for help with data processing, T. Kunkel for the term “isomerization”, R. Kolodner, E. Alani and J. Surtees for vital discussions, and D. Leahy for critical reading of the manuscript. The research of all authors was funded by the intramural research program of NIDDK, NIH.

References

1. Iyer RR, Pluciennik A, Burdett V, Modrich PL. DNA mismatch repair: functions and mechanisms. Chem Rev. 2006; 106:302–323. [PubMed: 16464007]
2. Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. Mol Cell. 2001; 7:1–12. [PubMed: 11172706]
3. Habraken Y, Sung P, Prakash L, Prakash S. ATP-dependent assembly of a ternary complex consisting of a DNA mismatch and the yeast MSH2—MSH6 and MLH1—PMS1 protein complexes. J Biol Chem. 1998; 273:9837–9841. [PubMed: 9545323]
4. Mendillo ML, et al. A conserved MutS homolog connector domain interface interacts with MutL homologs. Proc Natl Acad Sci U S A. 2009; 106:22223–22228. [PubMed: 20080788]
5. Habraken Y, Sung P, Prakash L, Prakash S. Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. Curr Biol. 1996; 6:1185–1187. [PubMed: 8805366]

6. Palombo F, et al. hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. Curr Biol. 1996; 6:1181–1184. [PubMed: 8805365]
7. Harfe BD, Jinks-Robertson S. Sequence composition and context effects on the generation and repair of frameshift intermediates in mononucleotide runs in Saccharomyces cerevisiae. Genetics. 2000; 156:571–578. [PubMed: 11014807]
8. Surtees JA, Alani E. Mismatch repair factor MSH2—MSH3 binds and alters the conformation of branched DNA structures predicted to form during genetic recombination. J Mol Biol. 2006; 360:523–536. [PubMed: 16781730]
9. Marsischky GT, Filosi N, Kane MF, Kolodner R. Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev. 1996; 10:407–420. [PubMed: 8600025]
10. Herman JG, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci U S A. 1998; 95:6870–6875. [PubMed: 9618505]
11. de Wind N, et al. HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. Nat Genet. 1999; 23:359–362. [PubMed: 10545954]
12. Sia EA, Dominska M, Stefanovic L, Petes TD. Isolation and characterization of point mutations in mismatch repair genes that destabilize microsatellites in yeast. Mol Cell Biol. 2001; 21:8157–8167. [PubMed: 11689704]
13. Edelmann W, et al. The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression. Cancer Res. 2000; 60:803–807. [PubMed: 10706084]
14. Peltomäki P. Lynch syndrome genes. Fam Cancer. 2005; 4:227–232. [PubMed: 16136382]
15. Haugen AC, et al. Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. Cancer Res. 2008; 68:8465–8472. [PubMed: 18922920]
16. Obmolova G, Ban C, Hsieh P, Yang W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. Nature. 2000; 407:703–710. [PubMed: 11048710]
17. Lamers MH, et al. The crystal structure of DNA mismatch repair protein MutS binding to a G × T mismatch. Nature. 2000; 407:711–717. [PubMed: 11048711]
18. Wrren JJ, et al. Structure of the human MutSalpha DNA lesion recognition complex. Mol Cell. 2007; 26:579–592. [PubMed: 17531815]
19. Bowers J, Sokolsky T, Quach T, Alani E. A mutation in the MSH6 subunit of the Saccharomyces cerevisiae MSH2—MSH6 complex disrupts mismatch recognition. J Biol Chem. 1999; 274:16115–16125. [PubMed: 10347163]
20. Drotschmann K, Yang W, Brownewell FE, Kool ET, Kunkel TA. Asymmetric recognition of DNA local distortion. Structure-based functional studies of eukaryotic Msh2—Msh6. J Biol Chem. 2001; 276:46225–46229. [PubMed: 11641390]
21. Dowen JM, Putnam CD, Kolodner RD. Functional studies and homology modeling of Msh2—Msh3 predict that mispair recognition involves DNA bending and strand separation. Mol Cell Biol. 2010; 30:3321–3328. [PubMed: 20421420]
22. Lee SD, Surtees JA, Alani E. Saccharomyces cerevisiae MSH2—MSH3 and MSH2—MSH6 complexes display distinct requirements for DNA binding domain I in mismatch recognition. J Mol Biol. 2007; 366:53–66. [PubMed: 17157869]
23. Shell SS, Putnam CD, Kolodner RD. Chimeric Saccharomyces cerevisiae Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins. Proc Natl Acad Sci U S A. 2007; 104:10956–10961. [PubMed: 17573527]
24. Kirkpatrick DT, Petes TD. Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. Nature. 1997; 387:929–931. [PubMed: 9202128]
25. Sugawara N, Paques F, Colaiacovo M, Haber JE. Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination. Proc Natl Acad Sci U S A. 1997; 94:9214–9219. [PubMed: 9256462]
26. Lyndaker AM, Alani E. A tale of tails: insights into the coordination of 3' end processing during homologous recombination. Bioessays. 2009; 31:315–321. [PubMed: 19260026]
27. van den Broek WJ, et al. Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. Hum Mol Genet. 2002; 11:191–198. [PubMed: 11809728]

28. Lopez Castel A, Cleary JD, Pearson CE. Repeat instability as the basis for human diseases and as a potential target for therapy. Nat Rev Mol Cell Biol. 2010; 11:165–170. [PubMed: 20177394]
29. McMurray CT. Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet. 2010; 11:786–799. [PubMed: 20953213]
30. Tian L, et al. Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. J Biol Chem. 2009; 284:20452–20456. [PubMed: 19525234]
31. Hou C, Chan NL, Gu L, Li GM. Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. Nat Struct Mol Biol. 2009; 16:869–875. [PubMed: 19597480]
32. Panigrahi GB, Slean MM, Simard JP, Gileadi O, Pearson CE. Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. Proc Natl Acad Sci U S A. 2010; 107:12593–12598. [PubMed: 20571119]
33. Wilson T, Guerrette S, Fishel R. Dissociation of mismatch recognition and ATPase activity by hMSH2—hMSH3. J Biol Chem. 1999; 274:21659–21664. [PubMed: 10419475]
34. Tian L, Gu L, Li GM. Distinct nucleotide binding/hydrolysis properties and molar ratio of MutSalpha and MutSbeta determine their differential mismatch binding activities. J Biol Chem. 2009; 284:11557–11562. [PubMed: 19228687]
35. Owen BA, W HL, McMurray CT. The nucleotide binding dynamics of human MSH2–MSH3 are lesion dependent. Nat Struct Mol Biol. 2009; 16:550–557. [PubMed: 19377479]
36. Yang W. Poor base stacking at DNA lesions may initiate recognition by many repair proteins. DNA Repair (Amst). 2006; 5:654–666. [PubMed: 16574501]
37. Natrajan G, et al. Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. Nucleic Acids Res. 2003; 31:4814–4821. [PubMed: 12907723]
38. Mendillo ML, Putnam CD, Kolodner RD. Escherichia coli MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair in vivo. J Biol Chem. 2007; 282:16345–16354. [PubMed: 17426027]
39. Biswas I, et al. Oligomerization of a MutS mismatch repair protein from Thermus aquaticus. J Biol Chem. 1999; 274:23673–23678. [PubMed: 10438551]
40. Hess MT, Mendillo ML, Mazur DJ, Kolodner RD. Biochemical basis for dominant mutations in the Saccharomyces cerevisiae MSH6 gene. Proc Natl Acad Sci U S A. 2006; 103:558–563. [PubMed: 16407100]
41. Alani E, Sokolsky T, Studamire B, Miret JJ, Lahue RS. Genetic and biochemical analysis of Msh2p—Msh6p: role of ATP hydrolysis and Msh2p—Msh6p subunit interactions in mismatch base pair recognition. Mol Cell Biol. 1997; 17:2436–2447. [PubMed: 9111312]
42. Calmann MA, Nowosielska A, Marinus MG. The MutS C terminus is essential for mismatch repair activity in vivo. J Bacteriol. 2005; 187:6577–6579. [PubMed: 16159793]
43. Mazur DJ, Mendillo ML, Kolodner RD. Inhibition of Msh6 ATPase activity by mispaired DNA induces a Msh2(ATP)—Msh6(ATP) state capable of hydrolysis-independent movement along DNA. Mol Cell. 2006; 22:39–49. [PubMed: 16600868]
44. Drotschmann K, Yang W, Kunkel TA. Evidence for sequential action of two ATPase active sites in yeast Msh2—Msh6. DNA Repair (Amst). 2002; 1:743–753. [PubMed: 12509278]
45. Antony E, Khubchandani S, Chen S, Hingorani MM. Contribution of Msh2 and Msh6 subunits to the asymmetric ATPase and DNA mismatch binding activities of Saccharomyces cerevisiae Msh2—Msh6 mismatch repair protein. DNA Repair (Amst). 2006; 5:153–162. [PubMed: 16214425]
46. Nakahara M, Yokozaki H, Yasui W, Dohi K, Tahara E. Identification of concurrent germ-line mutations in hMSH2 and/or hMLH1 in Japanese hereditary nonpolyposis colorectal cancer kindreds. Cancer Epidemiol Biomarkers Prev. 1997; 6:1057–1064. [PubMed: 9419403]
47. Bjornson KP, Allen DJ, Modrich P. Modulation of MutS ATP hydrolysis by DNA cofactors. Biochemistry. 2000; 39:3176–3183. [PubMed: 10715140]
48. Gradia S, et al. hMSH2—hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. Mol Cell. 1999; 3:255–261. [PubMed: 10078208]

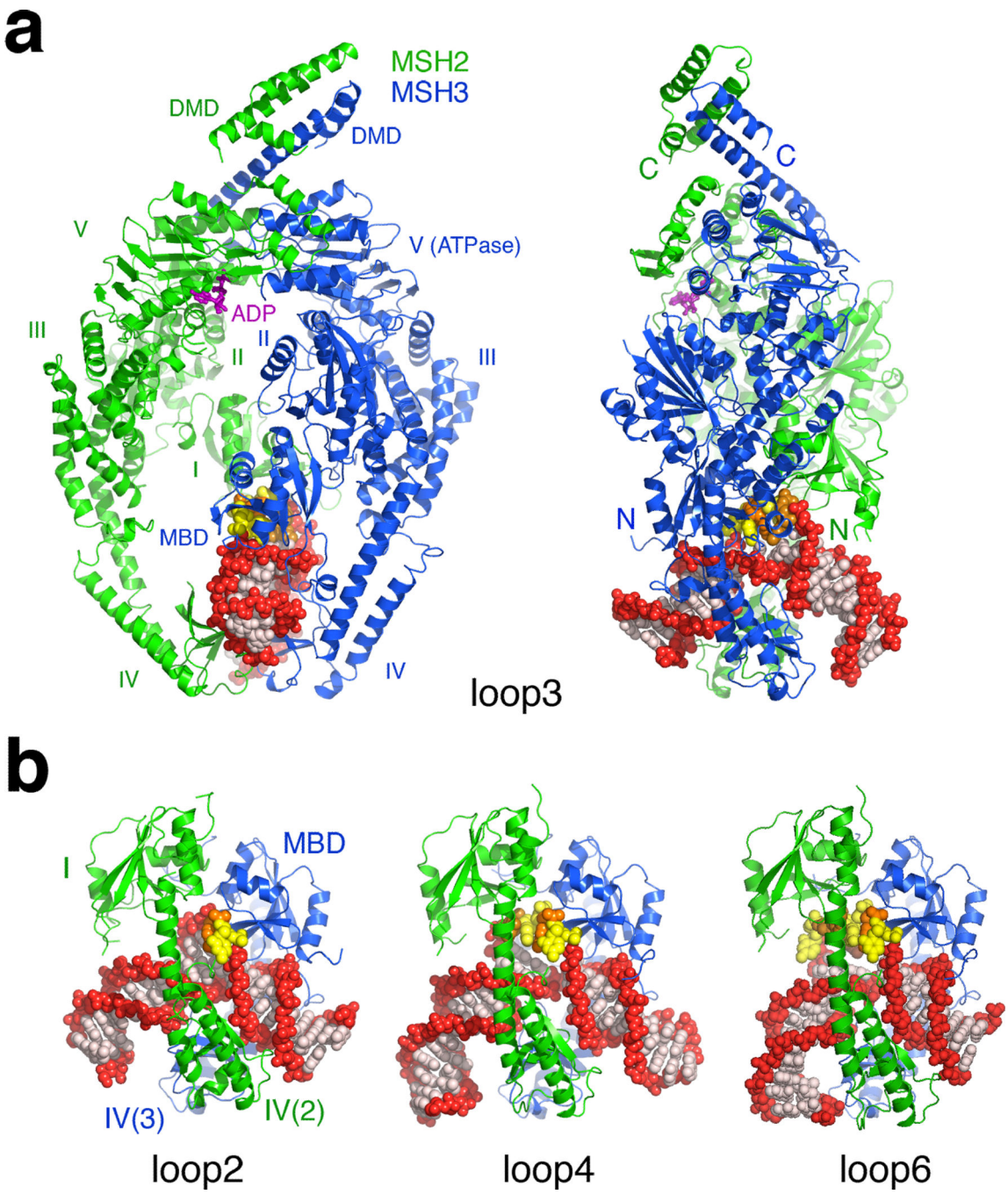

Figure 1.

Overall structures of MutSβ—DNA complexes. **(a)** Orthogonal views of Loop3 structure in ribbon diagrams, with MSH2 in green and MSH3 in blue. The DNA is shown in space-filling model with backbone in red, bases in light pink, and the unpaired nucleotides in yellow and orange. ADP bound to MSH2 is shown in purple sticks. **(b)** Side views of DNA-binding domains and DNA in Loop2, Loop4 and Loop6 structures. Each unpaired CA dinucleotide repeat is shown in yellow and orange. Domain I of MSH3 is the MBD. MSH2

and MSH3 subunits are indicated by 2 and 3 in parenthesis. Domains I—V and dimerization domains (DMD) are indicated.

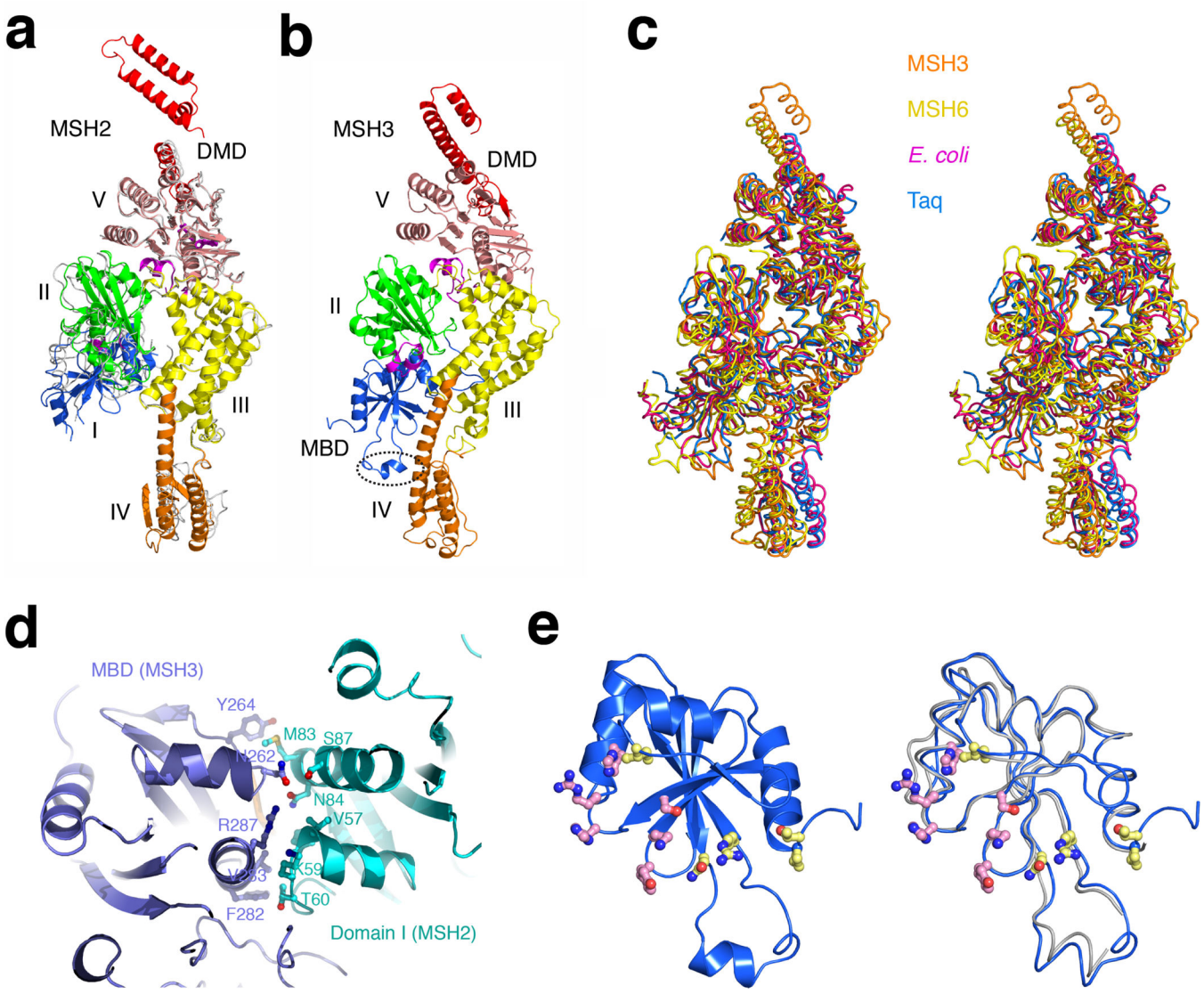

Figure 2.

Comparison of MutSα and MutSβ proteins. **(a)** Ribbon diagram of MSH2 from MutSβ. Domains I, II, III, IV, V and the DMDs are shown in blue, green, yellow, orange, pink and red, respectively. MSH2 of MutSα is superimposed and shown in grey. Domain interfaces between I and II and between II, III and V are highlighted in magenta. **(b)** Ribbon diagram of MSH3 in the same orientation and same color codes. Domain I (MBD) and IV (clamp) interact as indicated by the dashed oval. (**c**) Superposition of the mismatch-binding subunits in MutSα, MutSβ, *E. coli* and Taq MutS in the same orientation as in (a) and (b). Except for domain IV, they superimpose remarkably well. **(d)** A ribbon diagram of the interface between domains I in MutSβ. Protein residues in all figures are labeled in one-letter code for clarity. **(e)** On the left is a ribbon diagram of MSH3 MBD decorated by residues conserved among MutS homologs (shown as yellow/blue/red stick-and-balls) and residues unique among MSH3 homologs (pink/blue/red stick-and-balls). On the right is the superposition of

MSH3 (blue) and MSH6 (grey) MBDs. The rmsd between them is 0.7Å over 87 pairs of Cα atoms.

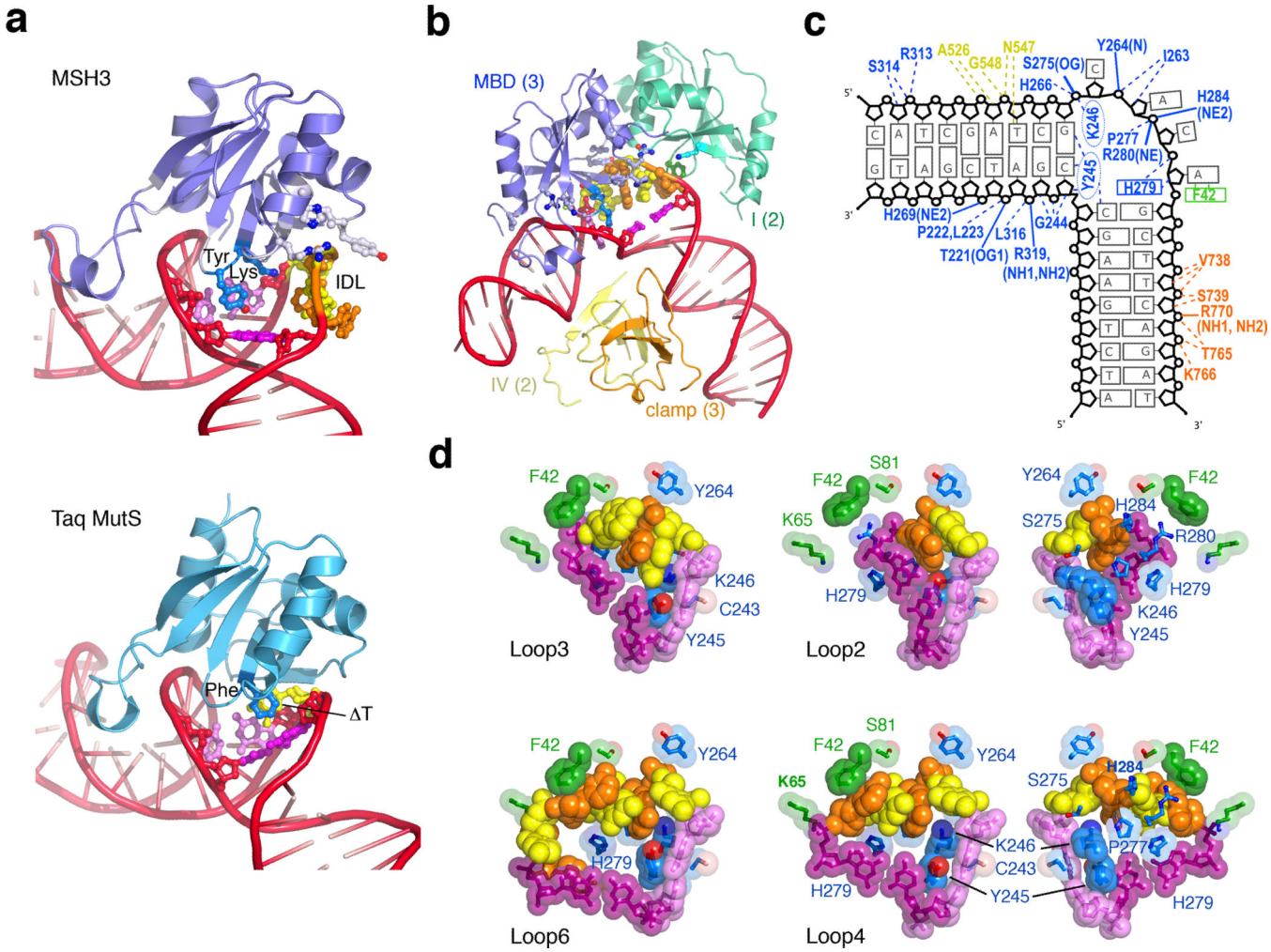

Figure 3.

IDL recognition by MutSβ. **(a)** A close-up comparison of MSH3—IDL interaction and Taq MutS with a single unpaired base (ΔT). **(b)** DNA-binding domains and DNA in Loop4. Domains I and IV of MSH2 are shown in green and yellow, and MBD and clamp domains of MSH3 in blue and orange. **(c)** Diagram of the protein-DNA interactions using the same color scheme as in (b). (**d**) Space-filling model of four IDLs and their interaction with domain I of MSH2 (green) and MBD of MSH3 (blue). The base pairs surrounding the IDL are shown in light (upstream) and dark (downstream) pink. For Loop2 and Loop4, a back view looking into the minor groove is also shown.

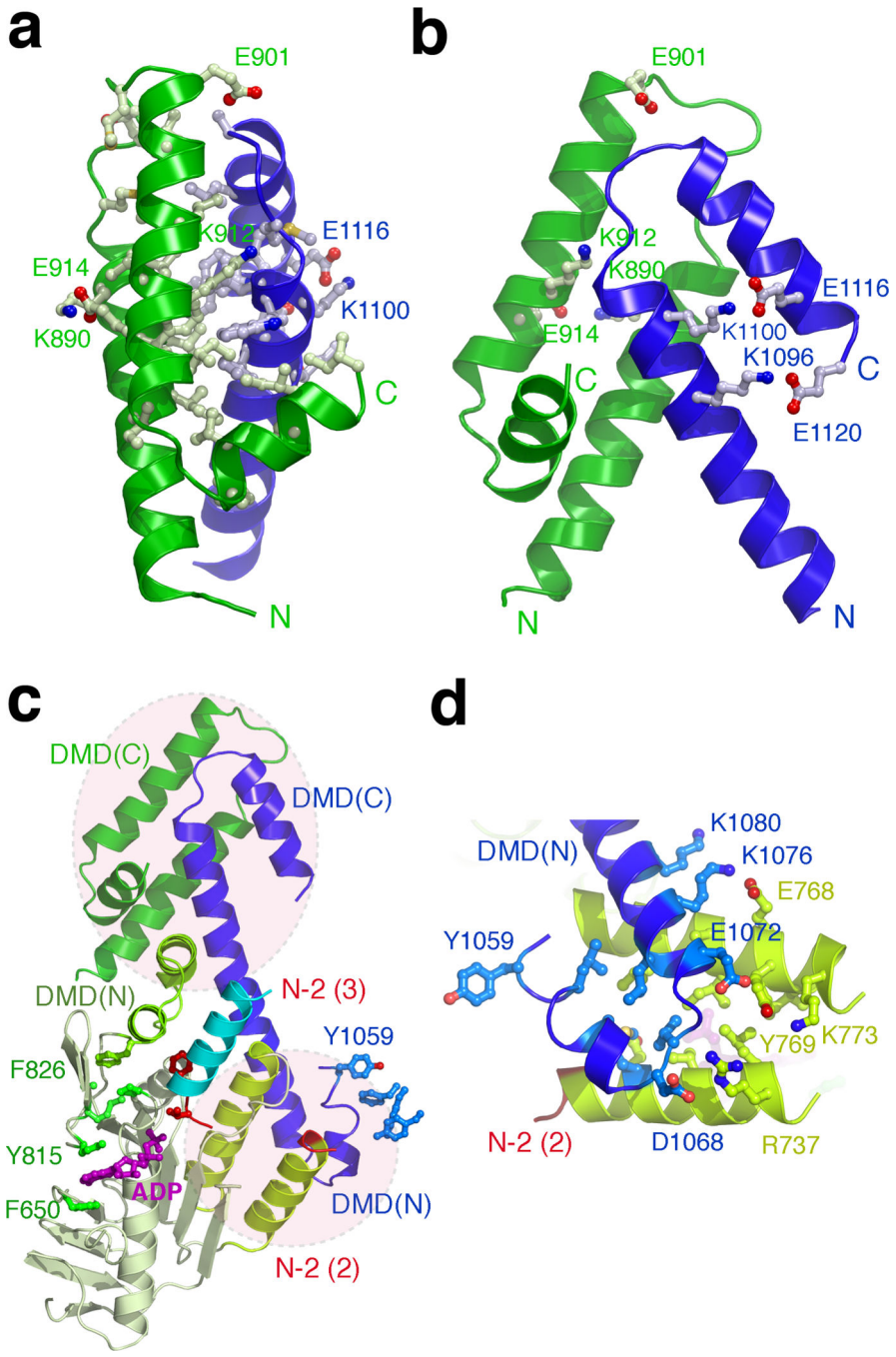

Figure 4.

Dimerization domains (DMD) of MutSβ. **(a, b)** Orthogonal views of the C-terminal halves of DMDs. The hydrophobic side chains at the interface, and polar residues forming salt bridges that stabilize intra-subunit interactions, are shown as sticks with carbon in light grey, nitrogen in blue and oxygen in red. E901 and K912 of MSH2 form N- and C-caps of the MSH3 helices. **(c)** The ATPase domain (light green) and DMD (green) of MSH2 are shown with the trans-acting N-2 (red and cyan) and DMD (blue) of MSH3. The ADP bound to MSH2 is shown as purple sticks. The two shaded ovals indicate the zoom-in areas shown in

(a, b) and (d). **(d)** A close-up view of the interactions between MSH3 DMD(N) and the ATPase domain of MSH2. Interactions between hydrophobic residues dominate, and two pairs of salt bridges (red dashes) may have alternative interacting partners (black dashes) if the two subunits slide relative to each other.

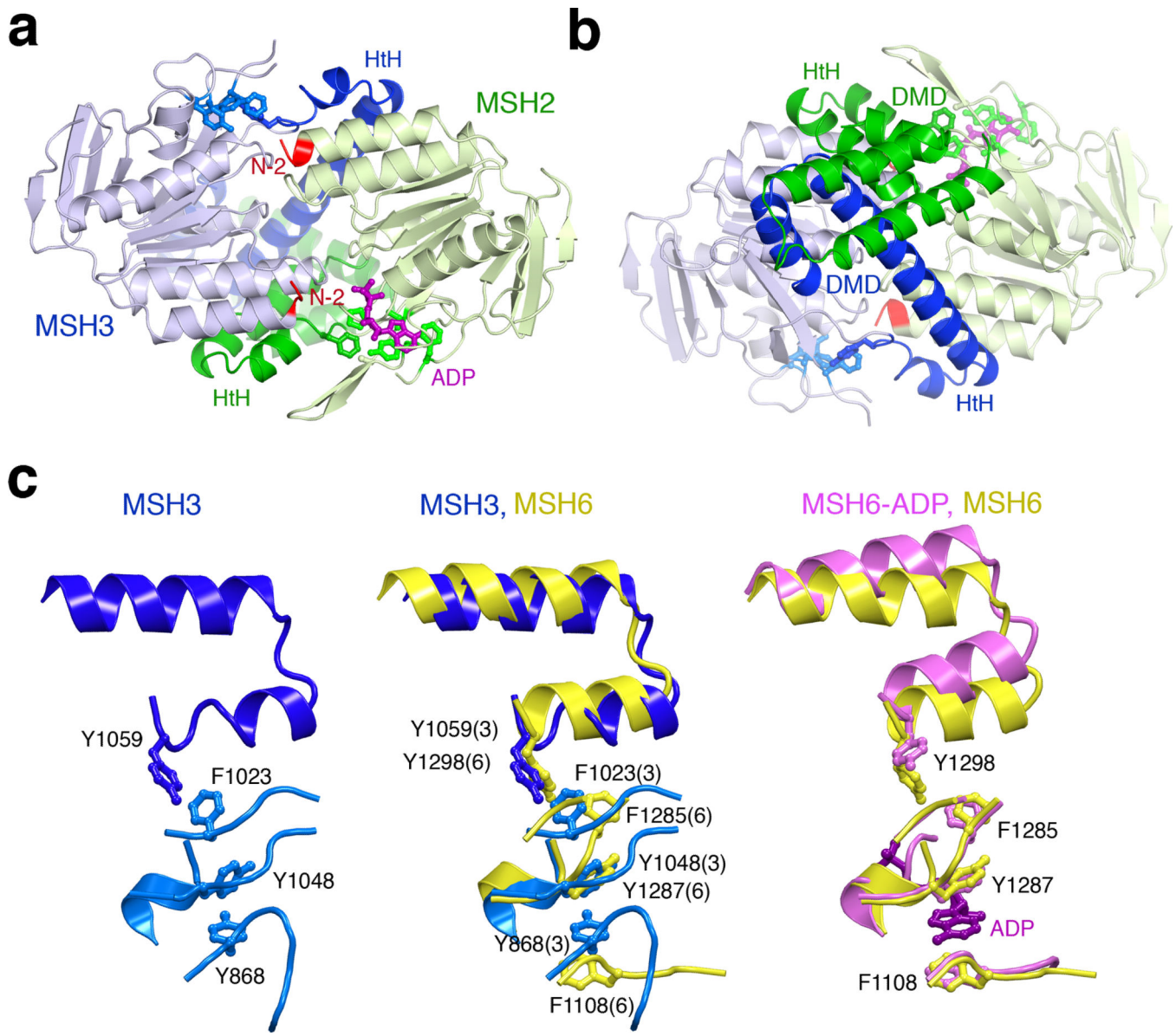

Figure 5.

Asymmetric ATPase sites of MutSβ. **(a)** Ribbon diagram of the ATPase and dimerization domains. MSH2 is shown in light and dark green, and MSH3 in light and dark blue. The trans-acting N-2 regions of MSH2 and MSH3 are highlighted in red. The aromatic sidechains connecting the ATPase site to the DMD are shown as blue (MSH3) and green sticks (MSH2). **(b)** A view 180° from (a) showing the asymmetric DMDs, biased towards the MSH2 ATPase site. **(c)** Comparison of the connection between the ATP binding site and DMD in MSH3 (blue), MSH6 with ADP (pink) and without (yellow) after superposition. The critical aromatic sidechains are shown as sticks.

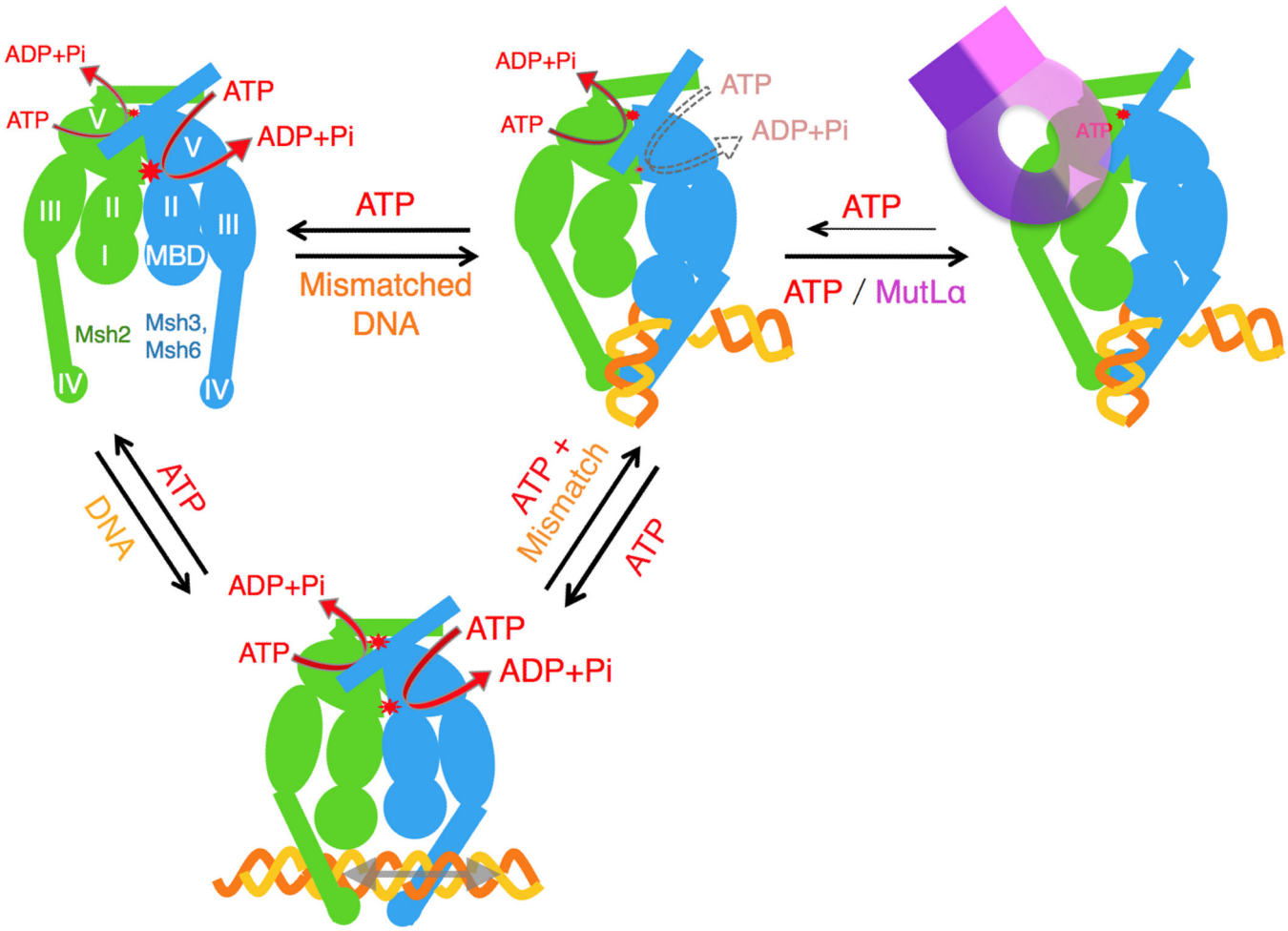

Figure 6.

Mechanism of mismatch recognition. Msh2 is drawn in green and Msh3 and Msh6 in blue. The ATPase activities of the two subunits are indicated by the curved arrows, the thicker the arrow the higher the activity. The DNA-binding domains are flexible in the apo-form. Binding to normal DNA, which is resistant to bending, induces conformational changes in the DNA-binding domains and ATP hydrolysis. But the protein-DNA association is not stable and MutS(α or β) can dissociate from or slide along DNA. Binding to a mismatched DNA, which is readily bent, leads to stable association of Msh3 or Msh6 with DNA and inhibition of its ATPase activity. ATP binding by Msh3 or Msh6, however, leads to release of the mismatched DNA. When Msh2 is bound to ATP and Msh3 or Msh6 to the mismatched DNA, MutS(α or β) can recruit MutLα to form a repairsome, thus initiating the repair process.

Table 1

Data collection and refinement statistics

	Loop2	Loop3	Loop4	Loop6
Data collection	23-ID	In-house	22-BM	23-ID
Space group	P 1	P 1	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions				
a, *b*, *c* (Å)	66.3, 91.1, 95.6	67.1, 91.5, 95.6	105.8, 116.1, 180.0	103.7, 154.6, 161.6
α, β, γ (°)	67.8, 87.0, 73.4	67.9, 86.5, 72.9	90, 90, 90	90, 90, 90
Resolution (Å) [1]	50.0–2.9 (2.95–2.90)	30.0–2.7 (2.75–2.70)	32.0–3.09 (3.14–3.09)	50.0–4.3 (4.37–4.30)
R_{merge} [1]	4.6 (24.8)	4.7 (33.2)	10.0 (55.4)	10.2 (68.7)
$I / \sigma I$ [1]	15.1 (2.3)	14.8 (1.4)	5.0 (1.1)	12.2 (1.8)
Completeness (%) [1]	94.4 (64.5)	91.8 (43.3)	99.0 (98.5)	96.0 (92.8)
Redundancy [1]	1.9 (1.5)	1.9 (1.5)	2.4 (2.4)	4.0 (4.0)
Refinement				
Resolution (Å)	44.0–2.9	27.0–2.7	32.0–3.09	49.1–4.3
No. reflections [2]	39845 (1125)	49511 (1315)	38509 (1919)	16879 (999)
R_{work}, R_{free}	19.4, 27.4	21.0, 27.3	22.9, 28.7	23.9, 28.5
No. atoms				
Protein, DNA	13730, 916	13270, 999	13777, 1115	13737, 980
Ligand	27	27	27	27
Water	15	20	4	0
B-factors				
Protein, DNA	84.2, 138.7	66.3, 114.8	66.8, 136.6	275.2, 408.0
Ligand	89.6	78.0	86.1	268.35
Water	55.2	39.3	23.6	-
R.m.s deviations				
Bond lengths (Å)	0.003	0.003	0.003	0.013
Bond angles (°)	0.68	0.67	0.563	0.994

[1] Data in the highest resolution shell is shown in parenthesis.

[2] Reflection number used in R_{free} calculation is shown in parenthesis.