

# DNA MISMATCH REPAIR\*

---

Thomas A. Kunkel<sup>1</sup> and Dorothy A. Erie<sup>2</sup>

<sup>1</sup>Laboratory of Molecular Genetics and Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709; email: kunkel@niehs.nih.gov

<sup>2</sup>Department of Chemistry and Curriculum in Materials Science, Venable and Kenan Laboratories, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3290; email: derie@unc.edu

**Key Words** DNA replication fidelity, mutagenesis, mutator, genome instability, error correction

■ **Abstract** DNA mismatch repair (MMR) is an evolutionarily conserved process that corrects mismatches generated during DNA replication and escape proofreading. MMR proteins also participate in many other DNA transactions, such that inactivation of MMR can have wide-ranging biological consequences, which can be either beneficial or detrimental. We begin this review by briefly considering the multiple functions of MMR proteins and the consequences of impaired function. We then focus on the biochemical mechanism of MMR replication errors. Emphasis is on structure-function studies of MMR proteins, on how mismatches are recognized, on the process by which the newly replicated strand is identified, and on excision of the replication error.

## CONTENTS

MULTIPLE MMR PROTEIN FUNCTIONS AND CONSEQUENCES OF DYSFUNCTION . . . . .	682
OVERVIEW OF PATHWAYS FOR CORRECTING REPLICATION ERRORS . . . . .	685
MutS-/L-/H-Dependent Mismatch Repair in <i>Escherichia coli</i> . . . . .	685
Multiple Mismatch Repair Pathways in Eukaryotes . . . . .	686
STUDIES OF THE PROTEINS AND STEPS IN MISMATCH REPAIR . . . . .	687
Mismatch Recognition: Structure-Function Studies of MutS Proteins . . . . .	687
Matchmaker Proteins . . . . .	695
Forming Multiprotein Complexes with DNA . . . . .	697
Signaling Models for Strand Discrimination . . . . .	697
Strand Excision . . . . .	702
SUMMARY AND PERSPECTIVES . . . . .	705

---

\*The U.S. Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

## MULTIPLE MMR PROTEIN FUNCTIONS AND CONSEQUENCES OF DYSFUNCTION

The integrity of genetic information depends on the fidelity of DNA replication and on the efficiency of several different DNA repair processes. Among many types of DNA repair, the general DNA mismatch repair (MMR) pathway is responsible for correcting base substitution mismatches and insertion-deletion mismatches (IDLs) generated during DNA replication in organisms from bacteria to mammals. The proteins involved in MMR (Table 1) have been extensively studied in recent years (1–14). These studies show that MMR protein functions can be inactivated in several ways without killing cells and with serious consequences. Inactivation of MMR can be permanent or may occur transiently when MMR proteins become limiting during periods of rapid, inaccurate replication or when cells are stressed and/or in a stationary phase.

MMR proteins participate in a wide variety of DNA transactions, such that their inactivation can have profound biological consequences on microbes, microbial populations and multicellular organisms. These consequences can be either beneficial or detrimental. Foremost among these consequences is the greatly increased rate of genome-wide point mutations, resulting from unrepaired DNA synthesis errors. This increased mutation rate provides the genetic variation in cells upon which selection can act to improve the fitness of microbial populations under stress. This observation has evolutionary implications and relevance to treatment of pathogenic microbes and emergence of resistant strains, e.g., via phase variation at contingency loci (15–21). In mammals, the mutator phenotype conferred by loss of MMR activity contributes to the initiation and promotion of multistage carcinogenesis (22). Especially useful for cancer research is the renowned instability of long repetitive DNA sequences, i.e., microsatellites. These are replicated inaccurately owing to frequent strand slippage and inefficient proofreading [reviewed in (23)], leaving MMR as the major guardian against microsatellite instability. For this reason, microsatellite instability is now an established biomarker for loss of MMR activity in tumor cells (24).

Every protein listed in Table 1 also participates in one or more other DNA transactions. In addition to undamaged mismatches, the MMR machinery also recognizes certain DNA lesions generated by normal intracellular metabolism (e.g., oxidative stress) (25, 26) and by physical (27) and chemical insults from the external environment, including certain chemotherapeutic agents. Operating as lesion sensors, MMR proteins activate cell cycle checkpoints and signal apoptosis [e.g., see (14, 27–34)]. Loss of these functions decreases apoptosis, increases cell survival, and results in resistance to chemotherapy (34–36). Moreover, the inability to correct replication errors resulting from damage to DNA or that alter normal Watson-Crick base-coding potential increases damage-induced mutagenesis in MMR-defective cells. These effects are thought to contribute to selective growth advantages to MMR-defective cells during multistage carcinogenesis (37), thereby partly explaining the increased susceptibility to tissue-specific cancers

**TABLE 1** Identity and functions of *Escherichia coli* and eukaryotic proteins involved in MMR of replication errors<sup>a</sup>

<i>E. coli</i> protein	Function	Homologs	Function
MutS	Binds mismatches	MSH2-MSH6 (MutS $\alpha$ )	Repairs single base-base and 1–2 base IDL mismatches
		MSH2-MSH3 (MutS $\beta$ )	Repair of some single base IDLs and IDLs $\geq 2$ bases Partially redundant with Msh2-Msh6
MutL	Matchmaker that coordinates multiple steps in MMR	MLH1-PMS2 (yPms1) (MutL $\alpha$ )	Matchmaker for coordinating events from mismatch binding by MutS homologs to DNA repair synthesis
		MLH1-MLH2 (hPMS1) (MutL $\beta$ )	Function of human heterodimer unknown Suppresses some IDL mutagenesis in yeast
		MLH1-MLH3 (MutL $\gamma$ )	Suppresses some IDL mutagenesis Participates in meiosis
MutH	Nicks nascent unmethylated strand at hemimethylated GATC sites	None	
$\gamma$ – $\delta$ Complex	Loads $\beta$ -clamp onto DNA	RFC complex	Loads PCNA, modulates excision polarity
$\beta$ -Clamp	Interacts with MutS and may recruit it to mismatches and/or the replication fork Enhances processivity of DNA pol III	PCNA	Interacts with MutS and MutL homologs Recruits MMR proteins to mismatches Increases MM binding specificity of Msh2-Msh6 Participates in excision and probably in signaling Participates in DNA repair synthesis
Helicase II	Loaded onto DNA at nick by MutS and MutL Unwinds DNA to allow excision of ssDNA	None	

(Continued)

TABLE 1 (Continued)

<i>E. coli</i> protein	Function	Homologs	Function
ExoI ExoX	Perform 3' to 5' excision of ssDNA	EXO1 (Rth1)	Excision of dsDNA
RecJ ExoVII	Perform 5' to 3' excision of ssDNA (also 3' to 5' excision by ExoVII)	3' exo of Pol $\delta$ 3' exo of Pol $\epsilon$	Excision of ssDNA Synergistic mutator with ExoI mutant
DNA pol III	Accurate resynthesis of DNA	DNA pol $\delta$	Accurate repair synthesis
SSB	Participates in excision and DNA synthesis	RPA	Participates in excision and in DNA synthesis
DNA ligase	Seals nicks after completion of DNA synthesis	DNA ligase	Seals nicks after completion of DNA synthesis

<sup>a</sup>Abbreviations are dsDNA, double-stranded DNA; MM, mismatch; PCNA, proliferating cell nuclear antigen; RFC, replication factor C; and ssDNA, single-stranded DNA.

associated with defects in certain MMR genes (10, 13, 31). Inactivation of specific MMR genes may also influence cancer predisposition via an increase in gene amplification (38) and, perhaps, via loss of transcription-coupled excision repair of DNA damage [e.g., see (39)].

MMR proteins participate in several other transactions that involve heteroduplex intermediates and/or recombination [reviewed in (1, 5, 12, 40–42)]. Examples include the following: (a) They block expansion of interrupted triplet repeat sequences in yeast, perhaps by binding hairpins with mismatches to provoke excision (43); (b) they participate in the repair of double-strand breaks in DNA; (c) they have important antirecombination functions; and (d) they are required for several transactions in meiotic cells [reviewed in (5, 12) and see (42, 44, 45)]. Loss of these functions can result in DNA rearrangements, telomere elongation by a telomerase-independent mechanism (46, 47), increased exchange of partially homologous sequences between different microbial species [e.g., see (17, 48)], and infertility [reviewed in (49, 50) and see (51)].

Somewhat paradoxically, some MMR proteins also participate in specialized DNA transactions that destabilize genetic information. A detrimental example involves the expansion of trinucleotide repeat sequences that underlie a number of hereditary diseases. The frequency of large expansions of a CAG repeat sequence is actually decreased in mice deficient in MSH2, MSH3, and PMS2 (52–55), indicating that MMR proteins actively contribute to formation of large expansions,

perhaps in a replication-independent manner (52, 56) that can be modulated by genotoxic agents (57). Two examples wherein eukaryotic MMR proteins alter genetic information in a positive way are somatic hypermutation and class switch recombination. These are very different DNA transactions that occur in B cells and are required for development of a normal and highly diverse repertoire of immunoglobulin genes. MMR proteins participate in both processes, which are impaired in mouse models when MMR genes have been inactivated [e.g., see (58–62)].

This brief overview does not nearly do justice to the quantity and quality of work on the numerous other functions of MMR proteins and the consequences of loss of function. Further details and many more references on these functions can be found in the reviews and recent articles cited above. The remainder of this review considers the biochemistry of repairing undamaged DNA replication errors. Emphasis is on studies in the past four years. Additional perspectives can be found in other reviews (1–14). Readers are especially encouraged to consult those that more comprehensively cite the early literature containing many seminal observations on MMR (e.g., 1–3a).

## OVERVIEW OF PATHWAYS FOR CORRECTING REPLICATION ERRORS

Undamaged DNA is normally replicated very accurately because of the high nucleotide selectivity of replicative DNA polymerases and the ability of 3' to 5' exonucleases to excise mistakes as they are made during replication (63). Rare polymerization errors that escape proofreading are mostly single base-base mismatches (12 are possible) or one to a few unpaired nucleotides in the template strand (deletion mismatches) or in the primer strand (addition mismatches). It is the responsibility of the general MMR pathway to remove these errors from the nascent strand in a manner that restores the parental genotype.

### MutS-/L-/H-Dependent Mismatch Repair in *Escherichia coli*

MMR in *E. coli* is initiated when MutS (Table 1) binds to mismatched DNA. MutS interacts with the  $\beta$ -clamp accessory protein (64) that is required for processive DNA replication, and it may help deliver MutS to mismatches. Correction of the replication error requires that the newly synthesized strand be targeted for excision. This process is accomplished in an ATP-dependent manner when MutS interaction with MutL activates the latent endonuclease activity of MutH. MutH is a member of the type II family of restriction endonucleases whose crystal structure has been described (65). It cleaves the newly synthesized, temporarily unmethylated strand at hemimethylated GATC sites located within about  $\sim 1$  kb of the error. The resulting nick, which can be either 3' or 5' to the mismatch, is the entry point for MutL-dependent loading of DNA helicase II and binding of single-strand

DNA-binding protein. Working together, these proteins generate single-stranded DNA (ssDNA) that is digested by either 3' or 5' exonucleases, depending on the location of the nick relative to the mismatch. This excision removes the error and allows highly accurate DNA polymerase III to correctly resynthesize the strand. DNA ligase seals the nick to complete MMR.

## Multiple Mismatch Repair Pathways in Eukaryotes

Eukaryotic MMR has features in common with *E. coli* MMR, but the proteins involved in the repair pathway (Table 1) can differ depending on the nature of the mismatch and the substrate for excision. Repair is initiated when complexes of MutS homologs, either MSH2-MSH6 (MutS $\alpha$ ) or MSH2-MSH3 (MutS $\beta$ ), bind to a mismatch. Studies of DNA binding by MutS $\alpha$  and MutS $\beta$  and of MMR activity in mammalian cell-free systems, as well as genetic studies in yeast and mouse models and also in humans with mutations in genes encoding MSH2, MSH6, and/or MSH3 [reviewed in (5, 12, 14)], indicate that MutS $\alpha$  is primarily responsible for repairing single base-base and IDL mismatches, that MutS $\beta$  is primarily responsible for repairing IDL mismatches containing up to 16 extra nucleotides in one strand (66), and that the two complexes can share responsibility for repairing some IDL mismatches, especially those with one extra base. Other MutS homologs also exist that are not considered further here, including MSH1 for mitochondrial transactions, Msh7 in *Aradadopsis* (67), and MSH4 and MSH5 involved in meiotic recombination [reviewed in (49, 50)]. The MutS $\alpha$  and MutS $\beta$  complexes interact with the eukaryotic homolog of the *E. coli*  $\beta$ -clamp, proliferating cell nuclear antigen (PCNA), which contributes to several steps in MMR (see below). Eukaryotes also encode multiple MutL homologs that form different heterodimers (Table 1). MutL $\alpha$  (MLH1-PMS2 in humans, Mlh1-Pms1 in yeast) is involved in repairing a wide variety of mismatches. The function of human MutL $\beta$  (MLH1-PMS1) is unknown, but yeast MutL $\beta$  (Mlh1-Mlh2) and yeast and human MutL $\gamma$  (MLH1-MLH3) are thought to participate in repairing a subset of IDLs [e.g., see (68–70) and reviewed in (5)]. MutL $\gamma$  is also involved in meiotic recombination (49, 50, 71).

Eukaryotes have no known homolog of *E. coli* MuthI, so the origins and identities of the entry point(s) for strand excision *in vivo* are currently less certain in eukaryotes than in *E. coli*. Strand discontinuities associated with replication may serve this purpose *in vivo* [e.g., see (64, 72, 73) and see models below], and nicks or gaps can direct strand-specific MMR repair activity *in vitro* [see below and (74)]. To date, no eukaryotic DNA helicase has been shown to participate in repair of replication errors. As in *E. coli*, more than one eukaryotic exonuclease has been implicated in MMR, and several other proteins are also required (Table 1 and see below). DNA resynthesis is catalyzed by an aphidicolin-sensitive polymerase, likely DNA polymerase  $\delta$  (75).

Eukaryotic proteins other than those listed in Table 1 may also participate in MMR. For example, high-mobility group box 1 protein (HMGB1), a nonhistone chromatin protein that bends DNA and facilitates protein-protein interactions, has recently been purified by its ability complement the defective MMR activity of a

depleted HeLa cell extract (76). HMGB1 was shown to be required at or prior to excision, which is consistent with its ability to interact with MutS $\alpha$  (76). Another candidate is DNA cytosine-5-methyltransferase 1. This enzyme is encoded by the *Dnmt1* gene, which was recovered in a genetic screen for cells defective in MMR (77). Dnmt1-deficient cells exhibit microsatellite instability, suggesting that Dnmt1 may directly or indirectly participate in MMR.

## STUDIES OF THE PROTEINS AND STEPS IN MISMATCH REPAIR

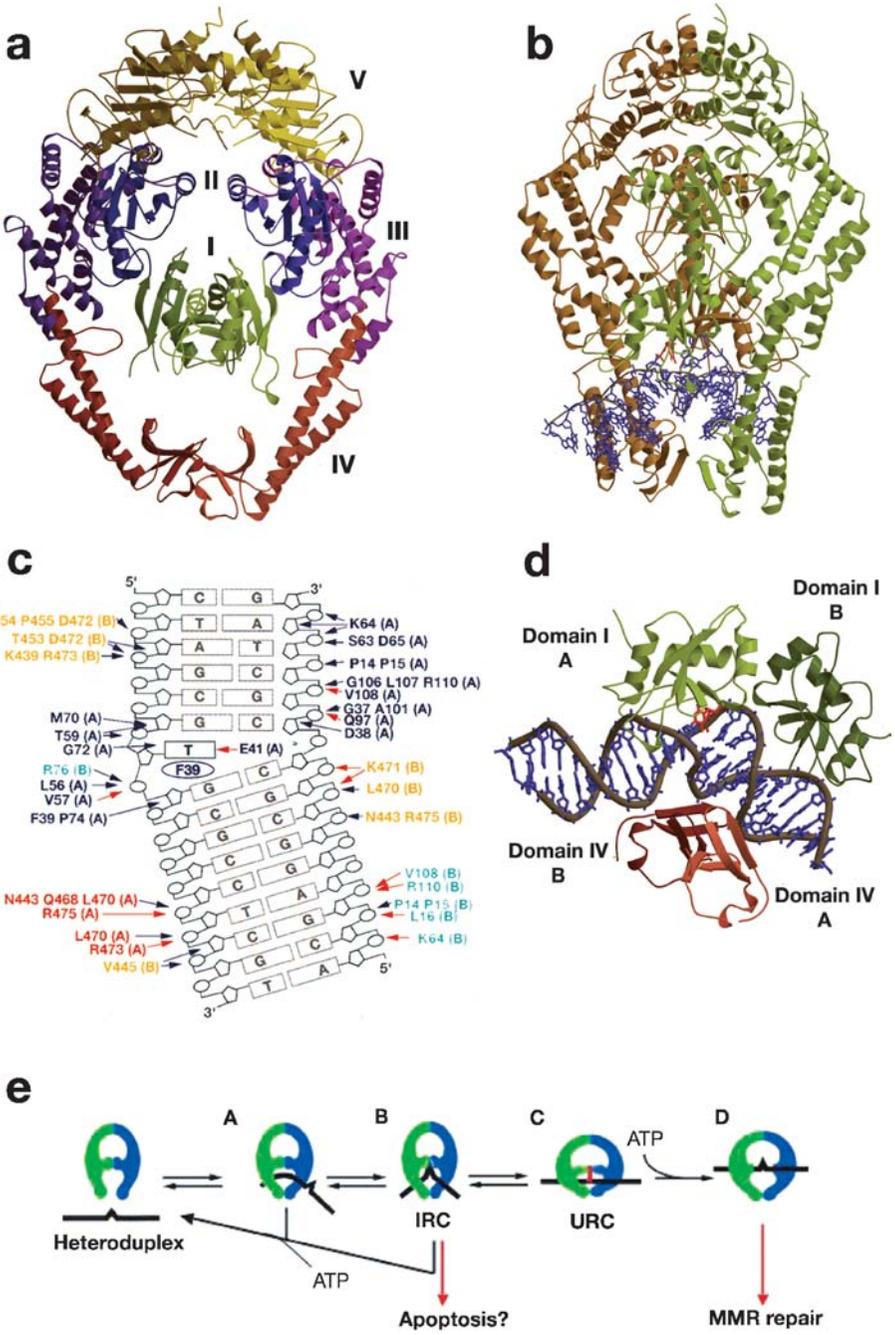
In the following sections, we review in greater detail the roles of MMR proteins in the steps needed to repair DNA replication errors. These are considered in sequential order, starting with initial recognition of mismatches and ending with correct resynthesis of DNA and ligation.

### Mismatch Recognition: Structure-Function Studies of MutS Proteins

As mentioned above, MMR is initiated by binding of MutS proteins to mismatches.

**STRUCTURES OF *E. coli* AND *Thermus aquaticus* (Taq) MUTS** MutS and MutS $\alpha$  bind to single base-base mismatches and to IDLs containing one to a few extra nucleotides in one strand (1). Binding affinities are 10- to more than 1500-fold higher for mismatched than perfectly matched DNA (78, 79). Affinities vary with the composition of the mismatch and the local sequence context, from high affinity for certain mismatches (e.g., G-T and single IDLs, which are the most frequent polymerization errors) to lower affinity for other mismatches (e.g., C-C, the rarest polymerization error). Insights into mismatch recognition come from crystal structures of *E. coli* and Taq MutS, which are similar (6, 8, 80–86). The structures reveal a homodimer of MutS bound asymmetrically to duplex DNA, containing IDL and base-base mismatches. Each subunit has five domains (Figure 1*a,b*). Eukaryotic MutS proteins have five homologous domains plus an N-terminal region containing a motif for interaction with PCNA. Domains I and IV bind to DNA, and domain V contains the dimerization interface and nucleotide-binding site. The DNA- and nucleotide-binding sites are widely separated but connected by domain III, which interacts directly with domain IV and indirectly with domain I via domain II.

**DNA BINDING** In the presence of mismatched DNA, domains I and IV encircle the DNA and form interfaces that create upper and lower channels. The upper channel is large enough and has an electrostatic potential that could accommodate DNA, leading to the idea that it might function in recombination and/or in the search for the strand discrimination signal (6). Although this idea could be true, amino acid replacements that altered the electrostatic potential had no detectable



effect on MMR or on the recombination function of *E. coli* MutS (87). The lower channel contains mismatched DNA that is kinked by 60° (Figure 1). Residues from both subunits contact the DNA (Figure 1*c,d*), but binding is asymmetric, with each subunit making numerous but different contacts. Most contacts are to the backbone and are therefore DNA sequence nonspecific, as expected given the need to repair replication errors in a variety of different sequence contexts. Only two residues, both in the same subunit, make mismatched base-specific contacts, either with the extra base in an IDL mismatch or with one of the two bases in a base-base mismatch. These interactions involve a Phe-X-Glu motif in bacterial MutS that is conserved in MSH6 but is not in MSH2 or MSH3. These facts imply that, like MutS, MutS $\alpha$  also binds asymmetrically to mismatched DNA. The results further suggest that the mechanism of MutS $\beta$  binding to IDL mismatches may be different from MutS or MutS $\alpha$  binding to single-base mismatches. In the MutS crystal structures, the mismatched base stacks with the aromatic ring of the phenylalanine. DNA binding and MMR activity are compromised when this phenylalanine in MutS or Msh6 is replaced by other amino acids (88–92). The conserved glutamate interacts with various mismatched bases and forms a hydrogen bond with the N3 of a mismatched thymine or the N7 of mismatched purines (80, 81, 85). Replacing this glutamate in *E. coli* MutS and in yeast Msh6 with other amino acids increases mutation rates in vivo, reduces DNA-binding affinity, and lowers mismatch discrimination in vitro (79, 92). In MutS crystal structures, the mismatched base is displaced toward a widened minor groove, and the opposing major groove is narrowed, thus accommodating the 60° kink. However, in crystals without DNA, the DNA-binding domains and part of the connector domain are disordered. These results indicate that mismatched DNA binding induces large conformational changes in both the DNA and in MutS. Consistent with this suggestion, a recent atomic force microscopy (AFM) study of DNA binding by MutS (93) demonstrated that MutS-mismatch complexes adopt a broad range of conformations, reflecting the dynamic nature of the MutS protein.

**REPAIR SPECIFICITY** The crystal structures of MutS bound to different base-base mismatches led to the idea that the local flexibility of a mismatch may be responsible for mismatch recognition [reviewed in (12)]. Although this idea can explain

---

**Figure 1** Structures of bacterial MutS and model for mismatch recognition. (a) Structure of Taq MutS homodimer comprised of five domains that form an upper and lower channel. (b) Structure of Taq MutS containing kinked DNA (blue) with a one-base IDL mismatch, i.e., an extra T (red) bound to the lower channel. (c) Protein-DNA contacts. (d) Close-up view of Taq MutS domain I and IV interactions with DNA. Panel c was taken directly from (80) and panels a, b, and d were adapted from similar images in (80), with permission. (e) Model for mismatch recognition (MMR) by MutS proteins, involving formation of an initial recognition complex (IRC) and an ultimate recognition complex (URC) that contains a flipped-out base.

recognition of a mismatch over homoduplex DNA, it does not explain why the relative efficiency of MMR does not correlate with the destabilizing effects or the structural distortions caused by the mismatch. In fact, there is an inverse correlation between the ease with which a mismatch is bent (or kinked) and the efficiency of repair (93). For example, the most efficiently repaired mismatch, G-T, is the most stable and causes minimal distortion of DNA, whereas the most poorly repaired mismatch, C-C, significantly destabilizes and distorts DNA [see (93, 94)].

Insights on repair specificity come from a recent study using AFM to directly visualize MutS bound to mismatched and to homoduplex DNA (93). That study found that in MutS-DNA complexes the DNA is bent at homoduplex sites, but both bent and unbent populations of conformations are observed at mismatch sites, with the unbent state being slightly less stable than the bent state. These results indicate the unbent state is the result of unique interactions between the mismatch base and MutS, and the bent conformation may be an intermediate in the formation of the unbent state. This suggestion is supported by the observation of broad distributions of angles of bent complexes and narrow distributions of angles of the unbent complexes, indicating the bent state represents a dynamic complex that is sampling a large number of conformations, whereas the unbent state is relatively rigid with a narrow distribution of bend angles typical of specific protein-DNA complexes. Together with the crystal structures, these data led to the proposal in Figure 1e: (A) MutS binds to DNA nonspecifically and bends it in search of a mismatch; (B) upon specific recognition of a mismatch, MutS undergoes a conformational change to an *initial recognition complex* (IRC) in which the DNA is kinked, with interactions similar to those in the crystal structures; and (C) MutS then undergoes further conformational changes to the *ultimate recognition complex* (URC) in which the DNA is unbent with the mismatched base possibly being flipped out. In fact, it is possible that the locations of the mismatched bases in the crystal structures represent the structural intermediate proposed to precede base flipping by other proteins that are suggested to extract the base by “invading the minor groove with a ‘wedge’ side chain to bend or kink the DNA” (95).

In this model, bending facilitates the initial recognition of the mismatched base by providing the opportunity for the specific interaction of the phenylalanine with the mispaired base. After recognition, MutS uses the energy stored in the bent DNA to help drive a conformational change in MutS to the unbent state, providing a “double check” before repair is initiated. Whether or not the downstream events that lead to repair will occur will be determined, in part, by the ability of MutS to form the unbent URC. In this model, the efficiency of repair will be governed by the relative stability of the bent (or kinked) and unbent flipped-out state. For mismatches, studies of the stabilities of DNA duplexes with an abasic site or a single extra base (96, 97) provide a rough estimate of the relative stabilities of the different flipped-out states. Single-base bulges would be expected to have the most stable flipped-out state, because in this state the duplex DNA on either side of the bulge can form a normal B-DNA helix and gain stacking interactions between

the bases flanking the bulge and there is no requirement to break hydrogen bonds between base pairs. Purine-purine mismatches and purine-pyrimidine mismatches, with the pyrimidine flipped out, are expected to be the next most stable because the purine that remains in the helix can achieve better stacking interactions than a pyrimidine. Interestingly, a C is the most destabilizing single-base bulge DNA (96), the least stable opposite an abasic site (97), and a C-C mismatch is the most easily kinked mismatch (98). These results provide a possible explanation of why C-C mismatches are refractory to repair, because the bent IRC would be stabilized and the unbent flipped-out state would be destabilized relative to other mismatches, causing the population of the unbent URC to be too low to signal repair.

The above analysis also predicts that MutS would preferentially flip out a pyrimidine in a purine-pyrimidine mismatch, which is consistent with the crystallographic observation that the specific interactions between *E. coli* MutS and a G-T mismatch are with the T (81). In contrast, in the crystal structure of MutS bound to an A-C mismatch, the phenylalanine is interacting specifically with the A (85). Why this is the case is not clear, but it could be related to sequence context effects (99). In any case, if base flipping is required for repair to ensue, flipping of A could explain why A-C mismatches have repair efficiencies similar to C-T mismatches, because in both cases a pyrimidine would be left behind. Using base flipping, MutS could achieve specific recognition of a mismatched base and the unbent state with a base flipped out would be more stable at a mismatch than at a correctly paired base (100). Such a mechanism has been found for other DNA repair enzymes (93).

This model can explain the observation that yeast MutS $\alpha$  binds tightly to palindromic mispairs, which are statically bent, but a ternary complex including MutS $\alpha$  and MutL $\alpha$  does not form in the presence of a palindromic mispair *in vitro*, and palindromic mispairs are refractory to MMR *in vivo* (93). Specifically, although statically bent palindromic mispairs may provide a good substrate for binding, they would be difficult to unbend to form the URC, which would make them refractory to repair. The model may also be relevant to recent observations on the effects of replacing a conserved P-loop glycine in the nucleotide-binding site of mouse MSH2 with alanine (31). This replacement eliminated MMR activity but did not reduce the ability of Msh2-Msh6 to bind to mismatched DNA or to signal apoptosis in response to treating cells with cisplatin, which introduces lesions that bend DNA. These results suggest that the MMR and apoptotic signaling functions of MutS $\alpha$  diverge upon formation of the IRC (Figure 1e). Interestingly, this same P-loop replacement altered two other functions of Msh2, class switch recombination and somatic hypermutation of immunoglobulin genes (58).

**SUBUNIT DIMERIZATION AND NUCLEOTIDE-BINDING SITES** The interface for MutS dimerization consists of a C-terminal helix-u-turn-helix motif in domain V that is critical for MutS functions (86, 101). In addition to forming dimers, MutS also forms tetramers and higher-order oligomers that may function in MMR [see (102)]. Oligomerization involves C-terminal residues that are not present in the Taq and

*E. coli* proteins used for crystallization. The MutS dimerization interface in domain V also forms two ATP-binding sites comprised of six motifs characteristic of ABC transporter family proteins (80–84). Five of these motifs are contributed by one subunit and a sixth is contributed in *trans* by the other subunit, i.e., these are “composite” nucleotide-binding sites. Structurally, the two sites are asymmetric. For example, although each MutS subunit has a loop that follows the Walker B (DE) motif, these loops are arranged differently in the two subunits [see Figure 6 in (84)]. Also, in the crystal of *E. coli* MutS bound to a G-T mismatch, the site in the subunit that nonspecifically contacts DNA is empty, whereas ADP is bound to the site in the subunit that contacts the mismatched base (81). Several studies have shown that MutS proteins change conformations in the presence of ATP and its nonhydrolyzable analogs (103–107). Comparing Taq MutS structures with and without ADP (80, 82), or with ATP soaked crystals of *E. coli* MutS (86), reveals several conformational changes in the nucleotide-binding domains and larger conformational changes in the DNA-binding sites. These observations have led to the suggestion that conformational changes are propagated and amplified via  $\alpha$ -helices that bridge the nucleotide-binding sites and the domain IV DNA clamps. Structural studies of Taq MutS-mismatched DNA complexes with ADP-beryllium fluoride reinforce this view, and these studies also further support the idea that the nucleotide-binding sites are asymmetric (83).

ASYMMETRY AND COORDINATION BETWEEN NUCLEOTIDE AND DNA-BINDING SITES  
MMR is an ATP-dependent repair pathway (1), and purified MutS proteins have ATPase activity that is required for MMR [see (12)]. In fact, engagement of adenine nucleotides by both subunits is needed to correctly modulate MutS protein functions in MMR because amino acid replacements of conserved residues in the ATPase active site of either MSH2 or MSH6 reduce MMR activity [see (12)]. Thus, understanding MutS protein functions is very complicated; theoretically, each nucleotide-binding site may be empty or bind ADP or ATP, and each is connected to sites that may be empty or bound to DNA but, perhaps, bound in different ways depending on whether domains I and IV of each subunit are free or fully or partially engaged in DNA binding. Moreover, the potential number of bound states and relevant conformations as well as the order of events that may occur during MMR are increased even further [see Figure 4 in (8)] through the involvement of clamps and of MutL proteins, which also have two nucleotide-binding sites and two DNA-binding sites (see below). To sort out this complexity, three interrelated topics have been investigated: (a) ADP/ATP occupancy and ATP hydrolysis by the two nucleotide-binding sites and how these are influenced by DNA binding, (b) how binding to mismatched DNA is affected by events at the nucleotide-binding sites, and (c) how these events modulate the interaction of MutS proteins with other MMR proteins to search for the strand discrimination signal (see below).

When viewed collectively, a variety of genetic, structural and biochemical studies reveal the following important features pertaining to adenine nucleotide occupancy and ATP hydrolysis:

- The two ATPase sites in MutS proteins have different functions in MMR (82, 84, 108–114), which is consistent with the structural asymmetry seen in the crystal structures.
- Both subunits appear to simultaneously bind adenine nucleotides (82, 83, 86, 112–115). The identity of the nucleotide in each site is not necessarily the same in each of these studies. Binding constants are in the 1–20  $\mu\text{M}$  range, with one high-affinity ADP and one high-affinity ATP site (112, 113, 115). Such high affinities suggest that states in which one or both of the binding sites is empty exist only transiently. Nucleotide-binding studies indicate that one stoichiometric equivalent of a nonhydrolyzable ATP analog (AMPPNP or ATP- $\gamma\text{S}$ ) binds to *E. coli* MutS and human MutS $\alpha$  (112, 113), and two equivalents bind to yeast MutS $\alpha$  with differing affinities (115), but only one ATP per dimer binds to yeast MutS $\alpha$  (115). In addition, AMPPNP is a weak competitor of ADP binding to *E. coli* MutS, suggesting that two equivalents of AMPPNP may be able to bind *E. coli* MutS. Two equivalents of ADP bind to *Taq* MutS (114), but only one ADP per dimer binds to *E. coli* MutS (112) and yeast and human MutS $\alpha$ , even at 100  $\mu\text{M}$  ADP (113, 115). Interestingly, however, the average ratio of ADP to ATP bound to human MutS $\alpha$  was found to be 1.6 during steady-state hydrolysis in the absence of DNA. These results suggest that MutS $\alpha$  exists in multiple liganded states, perhaps most likely ATP/ADP and/or ADP/ADP, with the ADP/ADP state possibly resulting from hydrolysis of ATP in the ATP/ADP state. Firmly establishing which subunit has the high-affinity ATP site will require further studies, but indirect evidence obtained with ATPase- and DNA-binding assays of Msh2-Msh6 heterodimers (104, 111, 116) have led to the suggestion that MSH6 (and perhaps the equivalent bacterial MutS subunit) may bind ATP with higher affinity than does MSH2. The affinity of a particular subunit for ATP, however, may depend on whether or not ADP is present.
- Both heteroduplex and homoduplex DNA stimulate the ATPase activity of MutS and MutS $\alpha$  [reviewed in (12, 117)]. However, the extent of the stimulation varies in different studies, depending on the MutS protein and on experimental conditions. The largest increase (relative to homoduplex DNA) in ATPase activity provoked by a mismatch is fourfold (78), which is not sufficient to explain the specificity of MMR. In another study, mismatch DNA was found to stimulate the ATPase activity of MutS $\alpha$  less than homoduplex DNA (89), and in yet another, the stimulation was found to depend on the length of the DNA (118). Both hetero- and homoduplex DNA stimulate ADP-ATP exchange [see (117)], but only heteroduplex DNA appears to change the rate-limiting step for turnover of ATP (115, 118). Specifically, in the presence of ATP alone, there is a burst of hydrolysis of one ATP equivalent per dimer in the absence of a mismatch and no burst in its presence (115, 118). These results suggest that the rate-limiting step for turnover of the first mole equivalent per dimer of ATP is after hydrolysis in the absence

of a mismatch but before or at hydrolysis when MutS is bound to a mismatch (115, 118).

- Comparing the ATPase studies with the nucleotide-binding studies reveals that the  $K_m$  for ATP is generally significantly higher than the dissociation constant ( $K_D$ ). These observations suggest that the dissociation rate for ATP is slow relative to the rate of catalysis and product (ADP) release and that there is not a rapid equilibrium of ATP binding during hydrolysis. This simple analysis is complicated by the two ATPase active sites; however, a recent study (114) demonstrated that ATP- $\gamma$ S has a very slow dissociation rate from its high-affinity-binding site.

Events at the nucleotide-binding sites and their effects on binding to mismatched DNA have been the subject of several studies [reviewed in (12, 117)]. Indications are that ADP has little effect on the affinity of MutS proteins for mismatch DNA. However, ATP and ATP analogs decrease mismatched DNA-binding affinity. With linear DNA substrates whose ends are not blocked, ATP induces dissociation of MutS proteins from homoduplex or heteroduplex DNA substrates. However, if the DNA ends are blocked with streptavidin, the ATP-induced dissociation rates of MutS proteins are significantly reduced on heteroduplex DNA but not on homoduplex DNA (103, 117, 119, 120). These results led to the suggestion that ATP induces formation of a mobile clamp state of MutS that can move away from the mismatch and that this mobile clamp state forms only if the MutS protein is bound to a mismatch (more on this below). ATP- $\gamma$ S but not AMPPNP also appears to induce this conformational change. A recent study of mismatch DNA binding by human MutS $\alpha$  demonstrated that ATP or ATP- $\gamma$ S alone or combinations of ADP with ATP- $\gamma$ S or AMPPNP reduced mismatched DNA-binding affinity by 5- to 7-fold; however, AMPPNP alone reduced the affinity by 20-fold (113). These results suggest that ADP binding by MutS may be essential for the formation of ATP-induced change of MutS to a mobile clamp conformation, which appears necessary to induce repair because MutS can hydrolyze ATP and ATP- $\gamma$ S but not AMPPNP.

One complication is that the majority of these ATPase studies have been conducted in the absence of ADP, or with very low concentrations of ADP. Given the recent results demonstrating that MutS and its homologs appear to contain one high-affinity ADP- and one high-affinity ATP-binding site but that high concentrations ATP can apparently compete out the ADP (112–115), it is possible that ATPase activity may differ when physiologically relevant concentrations of ADP and ATP are present. The importance of conducting future ATPase studies in the presence of ADP is highlighted by the observation that at low concentrations, AMPPNP increases the affinity of ADP for *E. coli* MutS (112). This result along with the high affinity of ADP for one MutS subunit suggests that ADP could potentially function as an allosteric effector, regulating ATP hydrolysis and/or DNA binding.

## Matchmaker Proteins

Next, we consider the roles of MutL and clamp proteins, whose primary functions are to act as molecular matchmakers during MMR.

**MUTL PROTEINS** MutL proteins have central roles in coordinating various steps in MMR. They share several general features in common with MutS proteins. MutL proteins participate in MMR as dimers, they interact with other MMR proteins and modulate their activities (see below), they bind and hydrolyze ATP, and they bind DNA, but in a mismatch-independent manner. Structures have not been reported for C-terminal regions that are important for forming MutL dimers (121) and for interactions with other MMR proteins (122, 123), but structures are available for highly conserved N-terminal regions that harbor ATPase active sites and some residues responsible for interactions with DNA and other proteins.

**Dimerization and assembly of the ATPase active site** In solution, the 40 kDa *E. coli* MutL N-terminal domain (LN40) is a monomer in the presence of ADP, and the crystal structure of LN40 with ADP bound does not have an intact nucleotide-binding site (124). With AMPPNP however, LN40 is a dimer with two intact ATPase active sites that are characteristic of members of the GHKL superfamily of ATPases (125). MutL has ATPase activity that is essential for its functions in MMR (87, 124–126). The turnover number is very low, indicating that ATP binding facilitates interactions with other proteins and/or DNA through protein conformational changes. Conformational changes inferred by comparing structures with ADP and AMPPNP bound include a large rotation of the two LN40 domains with respect to each other, a 20° movement of the ATP lid toward the dimer interface, conversion of lid motif III from an  $\alpha$ -helix to an extended phosphate-binding loop, and formation of surfaces that may interact with other proteins. On the latter point, the LN40-AMPPNP complex has a large crevice that may be the MutH interaction site and another surface suggested to interact with MutS [see Figure 7 in (125)]. Such protein-protein interactions are important for communication with the strand discrimination signal and for strand excision (see below).

The structure of an N-terminal domain of human PMS2 is similar to that of LN40, but surprisingly, it is a monomer (127). Like MutL, intact Mlh1 heterodimers and the N-terminal domains of yeast, Mlh1 and Pms1, have weak but essential ATPase activities (128–130). However, the two subunits have intrinsically different ATPase activities. Limited proteolysis studies indicate that ATP-induced conformational changes are different for each subunit, and identical mutations of homologous residues in the two subunits inactivate MMR to different degrees (128–131). Thus, similar to MutS proteins, MutL proteins are functionally asymmetric and likely bind and hydrolyze ATP in a sequential or alternating manner during MMR.

**DNA binding** *E. coli* MutL is a DNA-binding protein with affinity for both ssDNA and double-stranded DNA (dsDNA) [see (87)]. This DNA binding is suggested

to occur in a positively charged cleft formed by N-terminal domain dimerization. This groove contains basic amino acids. Replacing two of these with glutamate has no effect on MutL activation of MutH but does reduce DNA binding, as well as MMR as indicated by a mutator phenotype (87, 125), and an arginine to glutamate replacement reduced stimulation of MutL ATPase by DNA (125). Yeast MutL proteins (132, 133) and the human PMS2 N-terminal domain (127) also bind ssDNA and dsDNA. Duplex DNA binding by yeast MutL $\alpha$  is in a mismatch-independent and sequence nonspecific manner, but unlike MutL, DNA does not stimulate ATPase activity (127, 133). Yeast MutL $\alpha$  DNA-binding affinity depends on ionic strength, suggesting that DNA binding is electrostatic. Duplex DNA binding is length dependent and occurs cooperatively with duplexes longer than 241 base pairs, and AFM reveals long, continuous tracts of Mlh1-Pms1 concomitantly bound to two nonhomologous duplexes (133). N-terminal domains of yeast Mlh1 and Pms1 each bind DNA independently and in the absence of detectable dimerization (133), suggesting that DNA binding may involve residues other than those implicated in bacterial MutL. Mutations in the DNA-binding domain of yeast Mlh1 yield a mitotic mutator phenotype (128), loss of meiotic crossing, and heteroduplex repair (134), suggesting that Mlh1 DNA binding is important for MMR and for meiotic DNA transactions. One of the residues in Mlh1 responsible for these effects is not conserved in Pms1, and the Mlh1 and Pms1 N-terminal domains have different DNA-binding affinities, suggesting functional asymmetry. DNA binding by MutL proteins may facilitate the search for the strand discrimination signal or the initiation or progression of nascent strand excision (see below).

**CLAMP PROTEINS** Another protein that has central coordinator functions in MMR is PCNA. PCNA can exist as a trimer of three identical subunits that interact to form a clamp that encircles DNA. One face of PCNA has three identical sites at the interdomain connector loops that can interact with a number of proteins. PCNA is placed onto DNA by the replication factor C (RFC) clamp-loader complex (135) at junctions between ssDNA and dsDNA and in one specific orientation (136). The resulting asymmetry has important implications for the mechanism of MMR.

PCNA is perhaps the ultimate multifunctional matchmaker protein for DNA transactions. It is a processivity factor for replicative polymerases, and it interacts with and stimulates the activity of proteins involved in processing Okazaki fragments. It also participates in numerous DNA repair processes, including MMR. Early evidence suggested that PCNA is required for MMR prior to DNA repair synthesis and that PCNA interacts with Msh2 and Mlh1 (72), leading to the suggestion that replication and MMR may be physically coupled and that primers at the replication fork may provide the strand discrimination signal. PCNA interacts with Msh2-Msh3 (137, 138) and Msh2-Msh6 (138–140) via N-terminal motifs in Msh3 and Msh6 that are also found in other proteins known to bind to the interdomain connector loops. Mutational studies of Msh6, Msh3, and PCNA [reviewed in

(12)] suggest that these interactions are important for MMR. PCNA increases the mismatch-binding specificity of Msh2-Msh6 (139), and it can assist in delivery of Msh2-Msh6 to mismatched DNA (141). PCNA also interacts with the mismatch excision enzyme Exo1 (142, 143), it colocalizes with Exo1 in replication foci (143), and it participates in strand excision and DNA resynthesis (see below). Msh3 and Msh6 also colocalize with PCNA in replication foci (140). MutS and MutL colocalize with replication foci in *Bacillus subtilis* (144), and the *E. coli*  $\beta$ -clamp interacts with MutS (64). The latter two observations may be relevant to prokaryotic MMR, possibly including the mechanism of strand discrimination in organisms that lack MutH and d(GATC) methylation. Collectively, these data suggest that DNA replication and MMR may be coupled and that clamps have important roles not only as polymerase processivity factors but also in directing MMR to mismatches in newly replicated DNA and possibly to 3' and 5' termini in newly replicated DNA. Theoretically, clamp matchmaking roles could begin immediately after the unproofread mismatch emerges from a replicative polymerase (see below), and/or after the replication fork has moved on, as suggested in the localization study in *B. subtilis* (144). Given that the rate of replication in bacterial cells is much faster than in eukaryotic cells, the physical and temporal relationships between replication and MMR need not be the same in all organisms.

## Forming Multiprotein Complexes with DNA

In both *E. coli* and eukaryotic cells, communication between the mismatch and the signal for strand discrimination involves formation of multiprotein complexes containing MutS and MutL proteins [reviewed in (12, 14, 117) and also see (116, 143, 145, 146)]. The region protected against DNase I attack increases from ~20 base pairs for MutS alone to more than 100 base pairs when MutL and MutS are both present (147), suggesting formation of a large complex containing multiple MutS and MutL molecules. Gel-shift assays indicate ATP-dependent loading of multiple copies of human MutS $\alpha$  on DNA containing a single mismatch (103, 117). The efficiency of complex formation depends on DNA chain length, suggesting binding of multiple copies of human MutS $\alpha$  and MutL $\alpha$  (148). PCNA can also form complexes on DNA with MutS $\alpha$  (141) and with MutS $\alpha$ /MutL $\alpha$  (149). The formation and stability of these complexes are clearly modulated by ATP. The exact roles of ATP binding and hydrolysis in the four nucleotide-binding sites and the still largely unexplored role of ADP are being investigated within the framework of three models for communication between the mismatch and the strand discrimination signal.

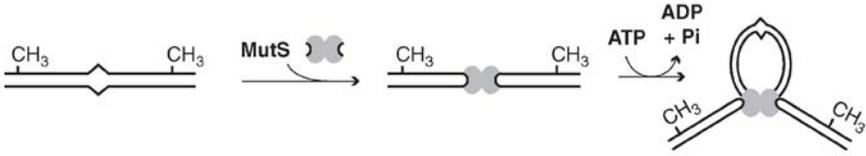
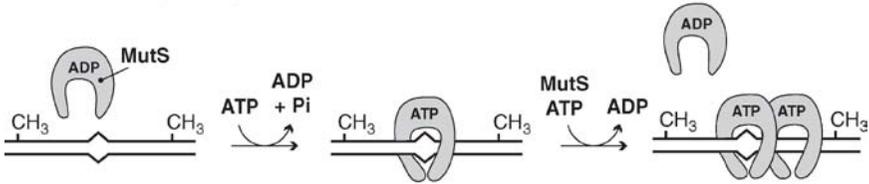
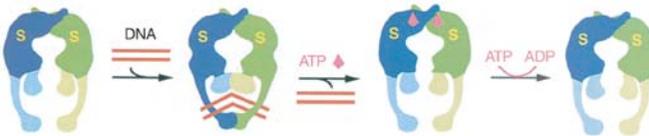
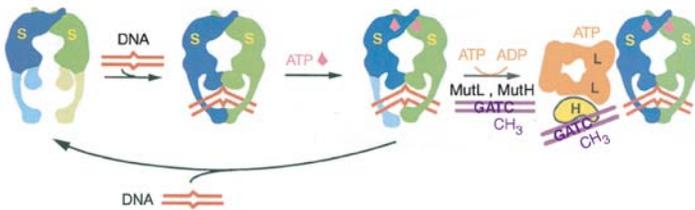
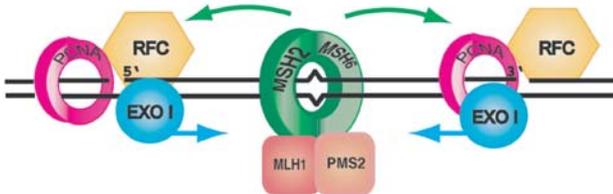
## Signaling Models for Strand Discrimination

Two models for linking mismatch recognition to activation of downstream events in MMR involve ATP-dependent formation of a MutS mobile clamp that leaves the mismatch in search of the strand discrimination signal.

**ACTIVE TRANSLOCATION MODEL** This model (Figure 2a) was originally proposed because of the observation that *E. coli* MutS mediates formation of  $\alpha$ -shaped loops on heteroduplex DNA (150). MutL enhanced the rate of MutS-mediated DNA loop growth, and both MutS and MutL appeared to be bound at the base of the  $\alpha$ -loop structures. Loop formation required a mismatch and ATP, increased with time, and did not occur with nonhydrolyzable ATP analogs. Also, loop growth stopped after addition of excess nonhydrolyzable ATP analogs. These observations led to the idea that ATP hydrolysis is needed for directional translocation of MutS along the helix. The original model was modified (119) to accommodate the observation that MutS appears to move a large distance on the DNA per ATP hydrolysis event. It is suggested that MutS homologs have two classes of DNA-binding sites, a latch (L) site in which DNA is bound at a single position and a site through which DNA can diffuse freely. Each functional MutS homolog is proposed to have two sites of each type. Whether the L site is open or closed is determined by the occupancy of the nucleotide-binding site by ADP or ATP. The closed L site forms a barrier to translocation in one direction, whereas the other site can diffuse away. Because there are two sites of each type, translocation is bidirectional. Several additional observations (113, 120, 148) support this model. For example, DNA with both singly and doubly blocked ends stimulated the ATPase activity of *E. coli* MutS. However, the degree of stimulation was reduced about twofold for a heteroduplex with both ends blocked when compared to having only one or no end blocked. This result supports the argument that the translocation of MutS homologs on DNA is ATP hydrolysis dependent, especially because the rate of hydrolysis appears to be faster than the rate of MutS dissociation from DNA in which both ends have been blocked.

**MOLECULAR SWITCH MODEL** The molecular switch model (Figure 2b) proposes that MutS initially binds to mismatched DNA in an ADP bound state. Mismatched DNA binding then provokes an ADP-ATP exchange, resulting in conformational changes that form MutS-sliding clamps, which leave the mismatch by diffusion that is independent of ATP hydrolysis. This process is suggested to occur iteratively to load multiple ATP-bound MutS clamps that can interact with MutL, and these complexes interact with and activate MutH endonuclease. The major evidence in support of this model, including additional perspectives about the role of MutL and the directionality of the excision step, has recently been extensively reviewed (117).

It is also useful to consider models in which MMR proteins remain at the mismatch while the search occurs for the strand discrimination signal. Maintaining proteins at or nearby the mismatch could protect it from attack by other cellular enzymes that might otherwise convert the mismatch into a mutation. If MutS or MutL proteins remain bound to DNA at or near the mismatch, this could preserve conformational changes that may be useful for downstream events. Here an analogy can be made to another multiprotein repair pathway, base excision repair. In base excision repair, the idea has emerged (151) that one enzyme in the pathway

**a** Translocation**b** Sliding clamp**c** Homoduplex DNA (no repair)**Heteroduplex DNA (mismatch repair)****d**

**Figure 2** Models of MMR protein functions in MMR. Panels *a*, *b*, and *c* are reproduced from (82), and panel *d* is revised from (142), with permission.

binds to and reshapes its DNA substrate and generates a DNA product with an altered conformation that is preferred for binding by the next protein in the pathway. Such direct DNA transfer from one protein to the next would reduce the energetic cost of altering DNA conformation and prevent release of potentially dangerous (i.e., cytotoxic and/or mutagenic) intermediates. This hypothesis is equally attractive for MMR, and two models have been proposed that would accommodate these ideas. One is discussed next, and the other is discussed in Strand Excision, below.

**DNA BENDING-MISMATCH VERIFICATION MODEL** This model (Figure 2c) proposes that MutS proteins remain in the vicinity of a mismatch and that communication between the mismatch and the strand discrimination signal involves DNA bending rather than protein movement along the DNA (82). The model further proposes that when MutS binds to a mismatch, ATP binding without hydrolysis is sufficient to activate downstream events, but if it binds to DNA without a mismatch, ATP binding is followed by hydrolysis, and MutS is released from the DNA. The model is based on several observations (82, 146, 147). For example, ATP hydrolysis was observed in the crystal of the *Taq* MutS-DNA complex without dissociation of the protein-DNA complex, suggesting that binding of ATP and a mismatch are not mutually exclusive. MutL blocked migration of MutS, and activation of MutH occurred at similar rates with substrates having a mismatch and a d(GATC) site on the same DNA molecule (*cis*-activation) or on different DNA molecules (*trans*-activation). Additional support for communication without continuous protein movement along the DNA contour comes from studies of MMR by human nuclear extracts using substrates in which a DNA hairpin (152) or a streptavidin-biotin moiety (153) is placed between the mismatch and the nick to block movement of MMR proteins between the two sites. Despite these blocks, mismatch-dependent excision is still initiated at the closest nick. Interestingly, once initiated, the excision reaction was eventually blocked such that the MMR reaction could not be completed.

As mentioned above, MutS and MutL are required to recruit and activate MutH. However, MutL that is defective in ATP hydrolysis can interact with MutH but fails to further stimulate the endonuclease activity of MutH in response to a mismatch, MutS, and ATP (147). In contrast, a MutS mutant (E694A) with strongly reduced ATPase activity was observed to activate MutH nearly as efficiently as wild-type MutS. In addition, the formation of MutS-MutL-mismatch complexes were observed by gel-shift assays. This result led to the idea that ATP hydrolysis by MutL, but not by MutS, is required for mismatch-dependent MutH activation (147) and that ATP binding and hydrolysis by DNA-bound MutS is a form of kinetic proofreading, which is used to verify that a mismatch is bound and to prevent attempts to repair correctly replicated DNA. However, a recent study of the same hydrolysis-deficient E694A MutS mutant (154) suggests that hydrolysis of ATP by MutS is required for activation of MutH when the GATC site is far from the mismatch. It was also found that this mutant did form MutS-MutL complexes

on mismatched DNA but at a significantly reduced rate, which was consistent with the reduced rate of hydrolysis. This observation led to the suggestion that hydrolysis of ATP by MutS homologs bound to a mismatch is required for the subsequent interaction of MutS with MutL. Consistent with this suggestion, the nonhydrolyzable analog of ATP (AMPPNP) is also incapable of activating MutS homologs to interact with MutL homologs, whereas the slowly hydrolysable analog ATP- $\gamma$ S permits activation but at a lower efficiency (155). In general, several hydrolysis-deficient mutants of MutS proteins have been observed to form ternary complexes with MutL; however, none of these mutants are completely dead for ATP hydrolysis (116). Therefore, the data can be explained by a slow hydrolysis of ATP. One complication in interpreting these results resides with added ADP in the experiments. Given that MutS homologs have one tight ATP- and one tight ADP-binding site (see above), the presence of ADP could significantly alter the results. Perhaps the differences in observations of the ATP-induced interaction of MutS E694A with MutL on heteroduplex DNA results from differences in the occupancy of one of the nucleotide-binding sites with ADP. Specifically, because MutS E694A hydrolyzes ATP slowly, ADP may or may not be bound, depending on how long the protein was incubated with ATP. Consequently, it remains unclear whether or not ATP hydrolysis is required for MutS activation and its interaction with MutL.

A model involving DNA bending and communication without eukaryotic MutS movement can accommodate the observations mentioned above on the DNA-binding properties of yeast MutL $\alpha$ . MutL $\alpha$  binds duplex DNA with high affinity only when the DNA is long enough to bend, suggesting that mismatched DNA that is kinked by MutS $\alpha$  may be a high-affinity substrate for yeast MutL $\alpha$ . Mlh1-Pms1 may simultaneously bind to duplex DNA near the mismatch and near the strand discrimination signal. This could involve the two DNA-binding sites on Mlh1 and Pms1, and/or separate heterodimers, bound to duplex DNA at the two locations, may interact, as indicated by the observed cooperativity of DNA binding. It is also possible that cooperativity may allow formation of a protein tract between the mismatch and the signal. Indeed, it was suggested long ago (3a) that mismatch recognition by a MutS homologue could trigger polymerization of a second protein along the helix. This could permit use of the strand discrimination signal without the need for a MutS-containing protein complex to abandon the vicinity of the mismatch. The DNA deformation introduced upon MutS binding to a mismatch could be conserved for use in subsequent steps in the repair pathway. Retention of a MMR protein complex at the mismatch during subsequent steps in the repair pathway is also attractive because knowledge of the mismatch location is important for directing, and possibly terminating, the strand excision reaction.

Taking all the results together leads to a few possibilities. It is clear that ATP can induce MutS to move away from the mismatch; however, this movement may not be required, and it may only be the conformational change in MutS that is important for subsequent events in repair. Alternatively, MutS may move away from the mismatch but only a short distance in the presence of MutL. In this case,

the movement of MutS could allow the formation of complexes in which multiple copies of MutS and MutL are loaded onto the DNA and form a large complex at and around the mismatch, as has been observed by footprinting (147) and surface plasmon resonance analysis (148). Such complexes could serve to amplify the signal for interaction with the downstream protein. If the strand discrimination signal is close to the mismatch, such complexes may result in direct interaction with the signal. In contrast, if the strand discrimination signal is far, the proteins may interact via bending as was suggested above. In this case, the large MutS-MutL complexes at the mismatch will increase the probability that the downstream proteins will locate the MMR initiation complex. It is theoretically possible that there are multiple pathways for MMR repair with different ADP/ATP requirements, and the pathway that is utilized may depend on any of several variables. These variables may include the site where mismatches are generated (e.g., chromosomal location or type of cell), when mismatches are generated (e.g., early or late S phase or in stationary phase), and how mismatches are generated [e.g., by flap misalignment or at the replication fork by polymerases that have very different properties (156), such as Pols  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\eta$ ,  $\iota$ ,  $\kappa$ , or  $\zeta$ ].

## Strand Excision

This step removes the replication error in the newly synthesized strand and provides another opportunity for correct DNA synthesis. Recent studies have greatly improved our understanding of protein requirements and the mechanism of strand excision.

**EXCISION IN *E. coli*** In *E. coli*, a strand break located either 3' or 5' to the mismatch can provide the entry point for exonucleolytic excision of the replication error in the new DNA strand (1). The break can be more than 1000 base pairs from the mismatch, with repair efficiency diminishing as the distance between the nick and the mismatch increases. MutS and MutL coordinate recognition of the two sites in a manner that permits loading of DNA helicase II at the strand break (157–160). Helicase II unwinds the DNA, and the displaced single strand is degraded by one of several single-strand exonucleases (161, 162). Excision preferentially occurs along the shortest path to the mismatch and terminates a short but variable distance after the mismatch is excised [for further details, see (162)].

**EXCISION IN EUKARYOTIC SYSTEMS** In yeast systems, biochemical studies of mismatch excision are largely lacking, but genetic studies implicate four nucleases in MMR, exonuclease 1 (Exo1), Rad27, and the 3' exonucleases intrinsic to Pol  $\epsilon$  and Pol  $\delta$  (5, 12). Exactly which yeast nucleases do have MMR functions must await further study because each of these four nucleases also contributes to genome stability by participating in other DNA transactions, such as proofreading, flap removal during processing of Okazaki fragments, and cell cycle checkpoint control [e.g., see (163, 164)]. Currently, Exo1 is most clearly implicated, both by studies

showing Exo1 interactions with MutS and MutL proteins (122, 123, 165–167) and by genetic studies indicating a catalytic role (168) and a structural role in MMR (169).

Much of our understanding of the biochemistry of eukaryotic mismatch excision is derived from studies of mammalian, usually human, proteins. A number of studies have been performed with extracts and by complementation of depleted extracts [early work reviewed in (2) and see (51, 74, 130, 170–174)]. Collectively, these studies indicate that excision is mismatch dependent, initiates at a nick or a gap, has bidirectional capacity, preferentially proceeds along the shortest path to the mismatch, and terminates at a number of sites about 150 nucleotides beyond the mismatch. Some of these same studies and more recent reactions performed with purified proteins (142, 170, 175) indicate that MutS $\alpha$ , MutL $\alpha$ , the PCNA clamp, the RFC clamp-loader complex, the ssDNA-binding protein RPA, and Exo1 all participate in mismatch excision.

*Exo1*-deficient mouse cells exhibit microsatellite instability at a mononucleotide marker, and they have a mutation rate characteristic of substantial loss of MMR (51). However, two other microsatellite loci were not unstable, and the mutation rate of *EXO1*<sup>-/-</sup> cells was lower than that of *MSH2*<sup>-/-</sup> cells, indicating that other mammalian nucleases may also participate in mismatch excision. Although purified human EXO1 digests dsDNA with 5' to 3' polarity (176), EXO1 is required for mismatch-dependent excision initiated at either 5' or 3' nicks (170), and extracts of EXO1-deficient mouse cells are defective in MMR activity in vitro directed by either 5' or 3' nicks (51). The unanticipated role of a known 5' exonuclease in 3' nick-directed excision led to the idea that, as proposed for yeast Exo1 (169), mammalian EXO1 may have a structural role in assembly of repair complexes. Another (nonexclusive) possibility (170) is that EXO1 has a cryptic 3' exonuclease activity that is only observed within the context of MMR. In support of the latter hypothesis, a mutant EXO1 protein with alanine replacing the catalytic Asp173 was recently found to be defective in both 5'- and 3'-directed mismatch excision using either a depleted extract-based complementation assay or reactions catalyzed with purified human MMR proteins (142).

Studies with extracts and purified proteins also reveal roles in mismatch excision for the other MMR proteins mentioned above. For example, when purified proteins conduct 5' mismatch excision in a reconstituted reaction, MutS $\alpha$  activates and confers high processivity to EXO1, and after the mismatch is excised, MutS $\alpha$  and MutL $\alpha$  suppress EXO1 activity while RPA facilitates its displacement from DNA, resulting in termination after mismatch excision (175). Consistent with these results is the observation that MutS $\alpha$  or MutL $\alpha$  can inhibit the 5' exonuclease activity of human EXO1 with homoduplex DNA but not with DNA containing a G-T mismatch (143). In studies with partially purified fractions, RPA also has been shown to enhance excision and stabilize excision intermediates (171, 177). Studies with depleted extracts (170, 174) and purified proteins (142, 175) also indicate that the protein requirements for 5' and 3' excision are different. For example, PCNA is essential for 3' but not 5' excision in extracts, and in the purified system, 5' excision

requires only MutS $\alpha$ , EXO1, and RPA, but 3' excision requires these same proteins plus MutL $\alpha$ , PCNA, and RFC. Moreover, EXO1 interacts with PCNA (142, 143), and RFC and PCNA activate EXO1 for 3' excision but not 5' excision (142). RFC not only loads PCNA onto DNA, it is required to suppress nonproductive 5' excision away from the mismatch, an effect that depends on the integrity of the ligase homology domain of the largest subunit of the five-protein RFC complex (142). This suppression may be due to the ability of this ligase homology domain to bind to recessed 5' phosphoryl termini [see (142, 178)]. Because PCNA is loaded onto primer templates in an orientation-dependent manner (136), the different protein requirements for 5' and 3' excision and the ability of PCNA to interact with MutS $\alpha$ , MutL $\alpha$  and EXO1 have led to a model [Figure 2*d*, reproduced from (142) with permission], wherein "an orientation-dependent encounter of PCNA at a strand discontinuity by the mobile MutS $\alpha$ -MutL $\alpha$  complex results in differential hydrolytic responses according to the 3' or 5' placement of the discontinuity" (142). This idea is reminiscent of earlier work (179) in which the homologous gp45 clamp protein of bacteriophage T4 was proposed to be loaded in an orientation-dependent manner at a nick in the nontranscribed strand to participate in activating transcription of late promoters.

**MODEL FOR REPLICATION-COUPLED EXCISION** PCNA has three binding sites for potential interaction with multiple partners, including proteins common to replication and MMR, e.g., Pol  $\delta$ . This observation led to the suggestion that replication and MMR may be physically coupled (72) and suggested a simple idea for mismatch excision using the 3' terminus at a replication fork (64, 72, 140). The model is an extrapolation of the principles underlying exonucleolytic proofreading, the other error correction process that greatly improves replication fidelity. When a proofreading-proficient replicative enzyme inserts an incorrect nucleotide, the resulting mismatched primer terminus is more difficult to extend than is a correct terminus. This slows the polymerase and provides an opportunity for the primer terminus to fray, creating ssDNA that can bind to the exonuclease active site for proofreading of the error. Because MMR is responsible for correcting those errors that escape proofreading, perhaps this sometimes occurs as the mismatch emerges from the polymerase, i.e., spatially and temporally just after it is created. Through interactions with PCNA or the  $\beta$ -clamp, MutS and MutL proteins may be poised at the replication fork to bind to the mismatch. As discussed above, binding could result in conformational changes in DNA, e.g., kinking by MutS and/or bending by MutL proteins (see above), which may slow polymerization to trigger an excision reaction to remove the error. Conceptually, this idea is similar to proposed transitions between a DNA polymerase and its intrinsic 3' nuclease for proofreading, mentioned above, or the proposed transition between Taq DNA polymerase and its intrinsic 5' exonuclease for excision of flaps [see (180) and discussions in (181)]. Mismatch excision could be conducted by ssDNA exonucleases, e.g., the 3' exonucleases intrinsic to replicative polymerases. Consistent with this idea is genetic evidence implicating the 3' exonucleases of Pol  $\delta$  and Pol  $\epsilon$  in mismatch

excision (182), and the observation that 3', but not 5', excision of mismatches is inhibited by aphidicolin, a known inhibitor of Pol  $\delta$  and Pol  $\epsilon$  (172). Alternatively, excision could be catalyzed by a dsDNA exonuclease, e.g., the 3' exonuclease activity of EXO1, with the choice perhaps depending on how deeply the mismatch is embedded in the duplex. In this model, the identity of the nascent strand would be obvious. The 3' terminus would be the entry point for excision, and it would be as close for continuous leading strand replication as for discontinuous lagging strand replication. MMR would require digestion of only a very short DNA tract, which is consistent with evidence that MMR can operate when a nick and a mismatch are separated by only a few base pairs (74). This short tract and the immediacy of repair could reduce the risk that other cellular enzymes might convert the mismatch into a mutation [e.g., see (183)]. This hypothetical model is mentioned in addition to, not exclusive of, the models in Figure 2.

## SUMMARY AND PERSPECTIVES

Our understanding of the mechanisms of MMR of replication errors has improved tremendously in the past few years, but major challenges remain. This includes a better understanding of the specificity of MMR, i.e., how and to what degree the MMR machinery avoids processing matched DNA. It is important to continue to dissect the roles of nucleotide binding, ATP hydrolysis, and DNA binding by both subunits of MutS and MutL heterodimers in modulating the various steps in MMR, especially the steps needed to identify and use the strand discrimination signals. More than one mechanism may operate depending on the organism and on where, when, and how mismatches are generated. The biochemistry of MMR is similar, but not identical, in bacteria and eukaryotes. This fact and the greater number of MutS and MutL proteins in eukaryotes indicates that multiple MMR subpathways are involved in maintaining the stability of very complex genomes. Thus, it will also be of great interest to study the functions of those proteins that have not yet received a great deal of attention, e.g., the beta and gamma heterodimers of MutS and MutL proteins. Our understanding of the mechanisms of MMR could benefit by developing approaches to study repair of mismatches that are not exogenously delivered by the investigator but rather actually made by the replication machinery. There is also a great need to expand investigations of the possible transcriptional and posttranslational mechanisms by which MMR proteins functions are regulated. Understanding the mechanisms by which replication errors are repaired will undoubtedly provide important insights into the roles of MMR proteins in the many other DNA transactions mentioned at the beginning, which are also intrinsically interesting, biologically important, and highly relevant to human health.

## ACKNOWLEDGMENTS

This work was supported by grants to D.A.E. from the National Institutes of Health (GM-54316) and the American Cancer Society (RSG-03-047-01-GMC).

The Annual Review of Biochemistry is online at  
<http://biochem.annualreviews.org>

## LITERATURE CITED

1. Modrich P, Lahue R. 1996. *Annu. Rev. Biochem.* 65:101–33
2. Modrich P. 1997. *J. Biol. Chem.* 272:24727–30
3. Kolodner RD, Marsischky GT. 1999. *Curr. Opin. Genet. Dev.* 9:89–96
- 3a. Modrich P. 1987. *Annu. Rev. Biochem.* 56:435–66
4. Buermeyer AB, Deschenes SM, Baker SM, Liskay RM. 1999. *Annu. Rev. Genet.* 33:533–64
5. Harfe BD, Jinks-Robertson S. 2000. *Annu. Rev. Genet.* 34:359–99
6. Yang W, Junop MS, Ban C, Obmolova G, Hsieh P. 2000. *Cold Spring Harbor Symp. Quant. Biol.* 65:225–32
7. Hopfner KP, Tainer JA. 2003. *Curr. Opin. Struct. Biol.* 13:249–55
8. Sixma TK. 2001. *Curr. Opin. Struct. Biol.* 11:47–52
9. Bellacosa A. 2001. *Cell Death Differ.* 8:1076–92
10. Muller A, Fishel R. 2002. *Cancer Investig.* 20:102–9
11. Marti TM, Kunz C, Fleck O. 2002. *J. Cell Physiol.* 191:28–41
12. Schofield MJ, Hsieh P. 2003. *Annu. Rev. Microbiol.* 57:579–608
13. Jiricny J, Marra G. 2003. *Curr. Opin. Genet. Dev.* 13:61–69
14. Li GM. 2003. *Front. Biosci.* 8:D997–1017
15. Halas A, Baranowska H, Policinska Z. 2002. *Curr. Genet.* 42:140–46
16. Chopra I, O'Neill AJ, Miller K. 2003. *Drug Resist. Updat.* 6:137–45
17. Matic I, Babic A, Radman M. 2003. *J. Bacteriol.* 185:1459–61
18. Hersh MN, Ponder RG, Hastings PJ, Rosenberg SM. 2004. *Res. Microbiol.* 155:352–59
19. Smania AM, Segura I, Pezza RJ, Becerra C, Albesa I, Argarana CE. 2004. *Microbiology* 150:1327–38
20. Bayliss CD, Sweetman WA, Moxon ER. 2004. *J. Bacteriol.* 186:2928–35
21. Alexander HL, Richardson AR, Stojiljkovic I. 2004. *Mol. Microbiol.* 52:771–83
22. Loeb LA, Loeb KR, Anderson JP. 2003. *Proc. Natl. Acad. Sci. USA* 100:776–81
23. Bebenek K, Kunkel TA. 2000. *Cold Spring Harbor Symp. Quant. Biol.* 65:81–91
24. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, et al. 2004. *J. Natl. Cancer Inst.* 96:261–68
25. Colussi C, Parlanti E, Degan P, Aquilina G, Barnes D, et al. 2002. *Curr. Biol.* 12:912–18
26. Russo MT, Blasi MF, Chiera F, Fortini P, Degan P, et al. 2004. *Mol. Cell. Biol.* 24:465–74
27. Young LC, Hays JB, Tron VA, Andrew SE. 2003. *J. Invest. Dermatol.* 121:435–40
28. Brown KD, Rathi A, Kamath R, Beardsley DI, Zhan Q, et al. 2003. *Nat. Genet.* 33:80–84
29. Shimodaira H, Yoshioka-Yamashita A, Kolodner RD, Wang JY. 2003. *Proc. Natl. Acad. Sci. USA* 100:2420–25
30. Kaina B. 2003. *Biochem. Pharmacol.* 66:1547–54
31. Lin DP, Wang Y, Scherer SJ, Clark AB, Yang K, et al. 2004. *Cancer Res.* 64:517–22
32. Hickman MJ, Samson LD. 2004. *Mol. Cell* 14:105–16
33. Stojic L, Brun R, Jiricny J. 2004. *DNA Repair* 3:1091–101
34. Fedier A, Fink D. 2004. *Int. J. Oncol.* 24:1039–47

35. Karran P, Offman J, Bignami M. 2003. *Biochimie* 85:1149–60
36. Bignami M, Casorelli I, Karran P. 2003. *Eur. J. Cancer* 39:2142–49
37. Fishel R. 2001. *Cancer Res.* 61:7369–74
38. Chen S, Bigner SH, Modrich P. 2001. *Proc. Natl. Acad. Sci. USA* 98:13802–7
39. Lee DF, Drouin R, Pitsikas P, Rainbow AJ. 2004. *Cancer Res.* 64:3865–70
40. Evans E, Alani E. 2000. *Mol. Cell. Biol.* 20:7839–44
41. Haber JE, Ira G, Malkova A, Sugawara N. 2004. *Philos. Trans. R. Soc. London Ser. B* 359:79–86
42. Sugawara N, Goldfarb T, Studamire B, Alani E, Haber JE. 2004. *Proc. Natl. Acad. Sci. USA* 101:9315–20
43. Rolfsmeier ML, Dixon MJ, Lahue RS. 2000. *Mol. Cell* 6:1501–7
44. Yang Q, Zhang R, Wang XW, Linke SP, Sengupta S, et al. 2004. *Oncogene* 23:3749–56
45. Bannister LA, Waldman BC, Waldman AS. 2004. *DNA Repair* 3:465–74
46. Bechter OE, Zou Y, Walker W, Wright WE, Shay JW. 2004. *Cancer Res.* 64:3444–51
47. Nguyen B, Elmore LW, Holt SE. 2004. *Cancer Biol. Ther.* 3:293–95
48. Greig D, Travisano M, Louis EJ, Borts RH. 2003. *J. Evol. Biol.* 16:429–37
49. Santucci-Darmanin S, Paquis-Flucklinger V. 2003. *Med. Sci.* 19:85–91
50. Svetlanov A, Cohen PE. 2004. *Exp. Cell Res.* 296:71–79
51. Wei K, Clark AB, Wong E, Kane MF, Mazur DJ, et al. 2003. *Genes Dev.* 17:603–14
52. Kovtun IV, McMurray CT. 2001. *Nat. Genet.* 27:407–11
53. van den Broek WJ, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, et al. 2002. *Hum. Mol. Genet.* 11:191–98
54. Savouret C, Brisson E, Essers J, Kanaar R, Pastink A, et al. 2003. *EMBO J.* 22:2264–73
55. Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP, Monckton DG. 2004. *Hum. Mol. Genet.* 13:1815–25
56. Gomes-Pereira M, Fortune MT, Monckton DG. 2001. *Hum. Mol. Genet.* 10:845–54
57. Gomes-Pereira M, Monckton DG. 2004. *Nucleic Acids Res.* 32:2865–72
58. Martin A, Li Z, Lin DP, Bardwell PD, Iglesias-Ussel MD, et al. 2003. *J. Exp. Med.* 198:1171–78
59. Li Z, Woo CJ, Iglesias-Ussel MD, Ronai D, Scharff MD. 2004. *Genes Dev.* 18:1–11
60. Martomo SA, Yang WW, Gearhart PJ. 2004. *J. Exp. Med.* 200:61–68
61. Li Z, Scherer SJ, Ronai D, Iglesias-Ussel MD, Peled JU, et al. 2004. *J. Exp. Med.* 200:47–59
62. Bardwell PD, Woo CJ, Wei K, Li Z, Martin A, et al. 2004. *Nat. Immunol.* 5:224–29
63. Kunkel TA. 2004. *J. Biol. Chem.* 279:16895–98
64. Lopez de Saro FJ, O'Donnell M. 2001. *Proc. Natl. Acad. Sci. USA* 98:8376–80
65. Ban C, Yang W. 1998. *EMBO J.* 17:1526–34
66. McCulloch SD, Gu L, Li GM. 2003. *J. Biol. Chem.* 278:3891–96
67. Wu SY, Culligan K, Lamers M, Hays J. 2003. *Nucleic Acids Res.* 31:6027–34
68. Flores-Rozas H, Kolodner RD. 1998. *Proc. Natl. Acad. Sci. USA* 95:12404–9
69. Harfe BD, Minesinger BK, Jinks-Robertson S. 2000. *Curr. Biol.* 10:145–48
70. Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, et al. 2000. *Nat. Genet.* 24:27–35
71. Lipkin SM, Moens PB, Wang V, Lenzi M, Shanmugarajah D, et al. 2002. *Nat. Genet.* 31:385–90
72. Umar A, Buermeyer AB, Simon JA, Thomas DC, Clark AB, et al. 1996. *Cell* 87:65–73
73. Pavlov YI, Mian IM, Kunkel TA. 2003. *Curr. Biol.* 13:744–48
74. Iams K, Larson ED, Drummond JT. 2002. *J. Biol. Chem.* 277:30805–14

75. Longley MJ, Pierce AJ, Modrich P. 1997. *J. Biol. Chem.* 272:10917–21
76. Yuan F, Lu L, Guo S, Wang C, Li GM. 2004. *J. Biol. Chem.* 279:20935–40
77. Guo G, Wang W, Bradley A. 2004. *Nature* 429:891–95
78. Gradia S, Acharya S, Fishel R. 2000. *J. Biol. Chem.* 275:3922–30
79. Schofield MJ, Brownwell FE, Nayak S, Du C, Kool ET, Hsieh P. 2001. *J. Biol. Chem.* 276:45505–8
80. Obmolova G, Ban C, Hsieh P, Yang W. 2000. *Nature* 407:703–10
81. Lamers MH, Perrakis A, Enzlin JH, Winterwerp HHK, de Wind N, Sixma TK. 2000. *Nature* 407:711–17
82. Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W. 2001. *Mol. Cell* 7:1–12
83. Alani E, Lee JY, Schofield MJ, Kijas AW, Hsieh P, Yang W. 2003. *J. Biol. Chem.* 278:16088–94
84. Lamers MH, Winterwerp HHK, Sixma TK. 2003. *EMBO J.* 22:746–56
85. Natrajan G, Lamers MH, Enzlin JH, Winterwerp HHK, Perrakis A, Sixma TK. 2003. *Nucleic Acids Res.* 31:4814–21
86. Lamers MH, Georgijevic D, Lebbink JH, Winterwerp HHK, Agianian B, et al. 2004. *J. Biol. Chem.* 279:43879–85
87. Junop MS, Yang W, Funchain P, Clendenin W, Miller JH. 2003. *DNA Repair* 2:387–405
88. Malkov VA, Biswas I, Camerini-Otero RD, Hsieh P. 1997. *J. Biol. Chem.* 272:23811–17
89. Bowers J, Sokolsky T, Quach T, Alani E. 1999. *J. Biol. Chem.* 274:16115–25
90. Dufner P, Marra G, Raschle M, Jiricny J. 2000. *J. Biol. Chem.* 275:36550–55
91. Yamamoto A, Schofield MJ, Biswas I, Hsieh P. 2000. *Nucleic Acids Res.* 28:3564–69
92. Drotschmann K, Yang W, Brownwell FE, Kool ET, Kunkel TA. 2001. *J. Biol. Chem.* 276:46225–29
93. Wang H, Yang Y, Schofield MJ, Du C, Fridman Y, et al. 2003. *Proc. Natl. Acad. Sci. USA* 100:14822–27
94. Peyret N, Seneviratne PA, Allawi HT, Santa Lucia J Jr. 1999. *Biochemistry* 38:3468–77
95. Cheng X, Blumenthal RM. 1996. *Structure* 4:639–45
96. Zhu J, Wartell RM. 1999. *Biochemistry* 38:15986–93
97. Sagi J, Guliaev AB, Singer B. 2001. *Biochemistry* 40:3859–68
98. Solaro PC, Birkenkamp K, Pfeiffer P, Kemper B. 1993. *J. Mol. Biol.* 230:868–77
99. Marsischky GT, Kolodner RD. 1999. *J. Biol. Chem.* 274:26668–82
100. O’Gara M, Horton JR, Roberts RJ, Cheng X. 1998. *Nat. Struct. Biol.* 5:872–77
101. Biswas I, Obmolova G, Takahashi M, Herr A, Newman MA, et al. 2001. *J. Mol. Biol.* 305:805–16
102. Bjornson KP, Blackwell LJ, Sage H, Baitinger C, Allen D, Modrich P. 2003. *J. Biol. Chem.* 278:34667–73
103. Gradia S, Subramanian D, Wilson T, Acharya S, Makhov A, et al. 1999. *Mol. Cell* 3:255–61
104. Studamire B, Quach T, Alani E. 1998. *Mol. Cell. Biol.* 18:7590–601
105. Joshi A, Rao BJ. 2002. *Biochemistry* 41:3654–66
106. Kato R, Kataoka M, Kamikubo H, Kuramitsu S. 2001. *J. Mol. Biol.* 309:227–38
107. Hess MT, Gupta RD, Kolodner RD. 2002. *J. Biol. Chem.* 277:25545–53
108. Iaccarino I, Marra G, Palombo F, Jiricny J. 1998. *EMBO J.* 17:2677–86
109. Studamire B, Price G, Sugawara N, Haber JE, Alani E. 1999. *Mol. Cell. Biol.* 19:7558–67
110. Iaccarino I, Marra G, Dufner P, Jiricny J. 2000. *J. Biol. Chem.* 275:2080–86
111. Drotschmann K, Yang W, Kunkel TA. 2002. *DNA Repair* 1:743–53
112. Bjornson KP, Modrich P. 2003. *J. Biol. Chem.* 278:18557–62
113. Martik D, Baitinger C, Modrich P. 2004. *J. Biol. Chem.* 279:28402–10
114. Antony E, Hingorani MM. 2004. *Biochemistry* 43:13115–28

115. Antony E, Hingorani MM. 2003. *Biochemistry* 42:7682–93
116. Kijas AW, Studamire B, Alani E. 2003. *J. Mol. Biol.* 331:123–38
117. Acharya S, Foster PL, Brooks P, Fishel R. 2003. *Mol. Cell* 12:233–46
118. Bjornson KP, Allen DJ, Modrich P. 2000. *Biochemistry* 39:3176–83
119. Blackwell LJ, Martik D, Bjornson KP, Bjornson ES, Modrich P. 1998. *J. Biol. Chem.* 273:32055–62
120. Blackwell LJ, Bjornson KP, Allen DJ, Modrich P. 2001. *J. Biol. Chem.* 276:34339–47
121. Pang Q, Prolla TA, Liskay RM. 1997. *Mol. Cell. Biol.* 17:4465–73
122. Schmutte C, Sadoff MM, Shim KS, Acharya S, Fishel R. 2001. *J. Biol. Chem.* 276:33011–18
123. Tran PT, Simon JA, Liskay RM. 2001. *Proc. Natl. Acad. Sci. USA* 98:9760–65
124. Ban C, Yang W. 1998. *Cell* 95:541–52
125. Ban C, Junop M, Yang W. 1999. *Cell* 97:85–97
126. Spampinato C, Modrich P. 2000. *J. Biol. Chem.* 275:9863–69
127. Guarne A, Junop MS, Yang W. 2001. *EMBO J.* 20:5521–31
128. Hall MC, Shcherbakova PV, Kunkel TA. 2002. *J. Biol. Chem.* 277:3673–79
129. Tomer G, Buermeyer AB, Nguyen MM, Liskay RM. 2002. *J. Biol. Chem.* 277:21801–9
130. Raschle M, Dufner P, Marra G, Jiricny J. 2002. *J. Biol. Chem.* 277:21810–20
131. Tran PT, Liskay RM. 2000. *Mol. Cell. Biol.* 20:6390–98
132. Hall MC, Wang H, Erie DA, Kunkel TA. 2001. *J. Mol. Biol.* 312:637–47
133. Hall MC, Shcherbakova PV, Fortune JM, Borchers CH, Dial JM, et al. 2003. *Nucleic Acids Res.* 31:2025–34
134. Hoffmann ER, Shcherbakova PV, Kunkel TA, Borts RH. 2003. *Genetics* 163:515–26
135. Bowman GD, O'Donnell M, Kuriyan J. 2004. *Nature* 429:724–30
136. Tsurimoto T. 1999. *Front. Biosci.* 4: D849–58
137. Johnson RE, Kovvali GK, Guzder SN, Amin NS, Holm C, et al. 1996. *J. Biol. Chem.* 271:27987–90
138. Clark AB, Valle F, Drotschmann K, Gary RK, Kunkel TA. 2000. *J. Biol. Chem.* 275:36498–501
139. Flores-Rozas H, Clark D, Kolodner RD. 2000. *Nat. Genet.* 26:375–78
140. Kleczkowska HE, Marra G, Lettieri T, Jiricny J. 2001. *Genes Dev.* 15:724–36
141. Lau PJ, Kolodner RD. 2003. *J. Biol. Chem.* 278:14–17
142. Dzantiev L, Constantin N, Genschel J, Iyer RR, Burgers PM, Modrich P. 2004. *Mol. Cell* 15:31–41
143. Nielsen FC, Jager AC, Lutzen A, Bundgaard JR, Rasmussen LJ. 2004. *Oncogene* 23:1457–68
144. Smith BT, Grossman AD, Walker GC. 2001. *Mol. Cell* 8:1197–206
145. Plotz G, Raedle J, Brieger A, Trojan J, Zeuzem S. 2003. *Nucleic Acids Res.* 31:3217–26
146. Selmane T, Schofield MJ, Nayak S, Du C, Hsieh P. 2003. *J. Mol. Biol.* 334:949–65
147. Schofield MJ, Nayak S, Scott TH, Du C, Hsieh P. 2001. *J. Biol. Chem.* 276:28291–99
148. Blackwell LJ, Wang S, Modrich P. 2001. *J. Biol. Chem.* 276:33233–40
149. Bowers J, Tran PT, Joshi A, Liskay RM, Alani E. 2001. *J. Mol. Biol.* 306:957–68
150. Allen DJ, Makhov A, Grilley M, Taylor J, Thresher R, et al. 1997. *EMBO J.* 16:4467–76
151. Wilson SH, Kunkel TA. 2000. *Nat. Struct. Biol.* 7:176–78
152. Wang H, Hays JB. 2003. *J. Biol. Chem.* 278:28686–93
153. Wang H, Hays JB. 2004. *EMBO J.* 23:2126–33
154. Baitinger C, Burdett V, Modrich P. 2003. *J. Biol. Chem.* 278:49505–11
155. Habraken Y, Sung P, Prakash L, Prakash S. 1998. *J. Biol. Chem.* 273:9837–41

156. Bebenek K, Kunkel TA. 2004. *Adv. Protein Chem.* 69:137–65
157. Yamaguchi M, Dao V, Modrich P. 1998. *J. Biol. Chem.* 273:9197–201
158. Dao V, Modrich P. 1998. *J. Biol. Chem.* 273:9202–7
159. Hall MC, Matson SW. 1999. *J. Biol. Chem.* 274:1306–12
160. Mechanic LE, Frankel BA, Matson SW. 2000. *J. Biol. Chem.* 275:38337–46
161. Burdett V, Baitinger C, Viswanathan M, Lovett ST, Modrich P. 2001. *Proc. Natl. Acad. Sci. USA* 98:6765–70
162. Viswanathan M, Burdett V, Baitinger C, Modrich P, Lovett ST. 2001. *J. Biol. Chem.* 276:31053–58
163. Tishkoff DX, Filosi N, Gaida GM, Kolodner RD. 1997. *Cell* 88:253–63
164. Datta A, Schmeits JL, Amin NS, Lau PJ, Myung K, Kolodner RD. 2000. *Mol. Cell* 6:593–603
165. Tishkoff DX, Boerger AL, Bertrand P, Filosi N, Gaida GM, et al. 1997. *Proc. Natl. Acad. Sci. USA* 94:7487–92
166. Schmutte C, Marinescu RC, Sadoff MM, Guerrette S, Overhauser J, Fishel R. 1998. *Cancer Res.* 58:4537–42
167. Rasmussen LJ, Rasmussen M, Lee B, Rasmussen AK, Wilson DM 3rd, et al. 2000. *Mutat. Res.* 460:41–52
168. Sokolsky T, Alani E. 2000. *Genetics* 155:589–99
169. Amin NS, Nguyen MN, Oh S, Kolodner RD. 2001. *Mol. Cell. Biol.* 21:5142–55
170. Genschel J, Bazemore LR, Modrich P. 2002. *J. Biol. Chem.* 277:13302–11
171. Ramilo C, Gu L, Guo S, Zhang X, Patrick SM, et al. 2002. *Mol. Cell. Biol.* 22:2037–46
172. Wang H, Hays JB. 2002. *J. Biol. Chem.* 277:26143–48
173. Wang H, Hays JB. 2002. *J. Biol. Chem.* 277:26136–42
174. Guo S, Presnell SR, Yuan F, Zhang Y, Gu L, Li GM. 2004. *J. Biol. Chem.* 279:16912–17
175. Genschel J, Modrich P. 2003. *Mol. Cell* 12:1077–86
176. Lee B, Nguyen LH, Barsky D, Fernandes M, Wilson DM 3rd. 2002. *Nucleic Acids Res.* 30:942–49
177. Lin YL, Shivji MK, Chen C, Kolodner R, Wood RD, Dutta A. 1998. *J. Biol. Chem.* 273:1453–61
178. Surtees JA, Alani E. 2004. *Mol. Cell* 15:164–66
179. Tinker RL, Williams KP, Kassavetis GA, Geiduschek EP. 1994. *Cell* 77:225–37
180. Maga G, Villani G, Tillement V, Stucki M, Locatelli GA, et al. 2001. *Proc. Natl. Acad. Sci. USA* 98:14298–303
181. Chapados BR, Hosfield DJ, Han S, Qiu J, Yelent B, et al. 2004. *Cell* 116:39–50
182. Tran HT, Gordenin DA, Resnick MA. 1999. *Mol. Cell. Biol.* 19:2000–7
183. Kim M, Huang T, Miller JH. 2003. *J. Bacteriol.* 185:4626–29

## CONTENTS

---

FROM PROTEIN SYNTHESIS TO GENETIC INSERTION, <i>Paul Zamecnik</i>	1
THE BIOCHEMISTRY OF PARKINSON'S DISEASE, <i>Mark R. Cookson</i>	29
APPLICATIONS OF DNA MICROARRAYS IN BIOLOGY, <i>Roland B. Stoughton</i>	53
ZONA PELLUCIDA DOMAIN PROTEINS, <i>Luca Jovine, Costel C. Darie, Eveline S. Litscher, and Paul M. Wassarman</i>	83
PROLINE HYDROXYLATION AND GENE EXPRESSION, <i>William G. Kaelin Jr.</i>	115
STRUCTURAL INSIGHTS INTO TRANSLATIONAL FIDELITY, <i>James M. Ogle and V. Ramakrishnan</i>	129
ORIGINS OF THE GENETIC CODE: THE ESCAPED TRIPLET THEORY, <i>Michael Yarus, J. Gregory Caporaso, and Rob Knight</i>	179
AN ABUNDANCE OF RNA REGULATORS, <i>Gisela Storz, Shoshy Altuvia, and Karen M. Wassarman</i>	199
MEMBRANE-ASSOCIATED GUANYLATE KINASES REGULATE ADHESION AND PLASTICITY AT CELL JUNCTIONS, <i>Lars Funke, Srikanth Dakoji, and David S. Bredt</i>	219
STRUCTURE, FUNCTION, AND FORMATION OF BIOLOGICAL IRON-SULFUR CLUSTERS, <i>Deborah C. Johnson, Dennis R. Dean, Archer D. Smith, and Michael K. Johnson</i>	247
CELLULAR DNA REPLICASES: COMPONENTS AND DYNAMICS AT THE REPLICATION FORK, <i>Aaron Johnson and Mike O'Donnell</i>	283
EUKARYOTIC TRANSLATION SYNTHESIS DNA POLYMERASES: SPECIFICITY OF STRUCTURE AND FUNCTION, <i>Satya Prakash, Robert E. Johnson, and Louise Prakash</i>	317
NOD-LRR PROTEINS: ROLE IN HOST-MICROBIAL INTERACTIONS AND INFLAMMATORY DISEASE, <i>Naohiro Inohara, Mathias Chamailard, Christine McDonald, and Gabriel Nuñez</i>	355

REGULATION OF PROTEIN FUNCTION BY GLYCOSAMINOGLYCANS—AS EXEMPLIFIED BY CHEMOKINES, <i>T.M. Handel, Z. Johnson, S.E. Crown, E.K. Lau, M. Sweeney, and A.E. Proudfoot</i>	385
STRUCTURE AND FUNCTION OF FATTY ACID AMIDE HYDROLASE, <i>Michele K. McKinney and Benjamin F. Cravatt</i>	411
NONTEMPLATE-DEPENDENT POLYMERIZATION PROCESSES: POLYHYDROXYALKANOATE SYNTHASES AS A PARADIGM, <i>JoAnne Stubbe, Jiamin Tian, Aimin He, Anthony J. Sinskey, Adam G. Lawrence, and Pinghua Liu</i>	433
EUKARYOTIC CYTOSINE METHYLTRANSFERASES, <i>Mary Grace Goll and Timothy H. Bestor</i>	481
MONITORING ENERGY BALANCE: METABOLITES OF FATTY ACID SYNTHESIS AS HYPOTHALAMIC SENSORS, <i>Paul Dowell, Zhiyuan Hu, and M. Daniel Lane</i>	515
STRUCTURE AND PHYSIOLOGIC FUNCTION OF THE LOW-DENSITY LIPOPROTEIN RECEPTOR, <i>Hyesung Jeon and Stephen C. Blacklow</i>	535
COPPER-ZINC SUPEROXIDE DISMUTASE AND AMYOTROPHIC LATERAL SCLEROSIS, <i>Joan Selverstone Valentine, Peter A. Doucette, and Soshanna Zittin Potter</i>	563
THE STRUCTURE AND FUNCTION OF SMC AND KLEISIN COMPLEXES, <i>Kim Nasmyth and Christian H. Haering</i>	595
ANTIBIOTICS TARGETING RIBOSOMES: RESISTANCE, SELECTIVITY, SYNERGISM, AND CELLULAR REGULATION, <i>Ada Yonath</i>	649
DNA MISMATCH REPAIR, <i>Thomas A. Kunkel and Dorothy A. Erie</i>	681
GENE THERAPY: TWENTY-FIRST CENTURY MEDICINE, <i>Inder M. Verma and Matthew D. Weitzman</i>	711
THE MAMMALIAN UNFOLDED PROTEIN RESPONSE, <i>Martin Schröder and Randal J. Kaufman</i>	739
THE STRUCTURAL BIOLOGY OF TYPE II FATTY ACID BIOSYNTHESIS, <i>Stephen W. White, Jie Zheng, Yong-Mei Zhang, and Charles O. Rock</i>	791
STRUCTURAL STUDIES BY ELECTRON TOMOGRAPHY: FROM CELLS TO MOLECULES, <i>Vladan Lučić, Friedrich Förster, and Wolfgang Baumeister</i>	833
PROTEIN FAMILIES AND THEIR EVOLUTION—A STRUCTURAL PERSPECTIVE, <i>Christine A. Orengo and Janet M. Thornton</i>	867