The EMBO Journal Vol.17 No.22 pp.6427–6436, 1998

EMBO MEMBER'S REVIEW

Replication errors: cha(lle)nging the genome

Josef Jiricny

Institute of Medical Radiobiology of the University of Zürich and the Paul Scherrer-Institute, August Forel-Strasse 7, CH-8008 Zürich, Switzerland

e-mail: jiricny@imr.unizh.ch

Since the discovery of a link between the malfunction of post-replicative mismatch correction and hereditary non-polyposis colon cancer, the study of this complex repair pathway has received a great deal of attention. Our understanding of the mammalian system was facilitated by conservation of the main protagonists of this process from microbes to humans. Thus, biochemical experiments carried out with *Escherichia coli* extracts helped us to identify functional human homologues of the bacterial mismatch repair proteins, while the genetics of *Saccharomyces cerevisiae* aided our understanding of the phenotypes of human cells deficient in mismatch correction. Today, mismatch repair is no longer thought of solely as the mechanism responsible for the correction of replication errors, whose failure demonstrates itself in the form of a mutator phenotype and microsatellite instability. Malfunction of this process has been implicated also in mitotic and meiotic recombination, drug and ionizing radiation resistance, transcription-coupled repair and apoptosis. Elucidation of the roles of mismatch repair proteins in these transduction pathways is key to our understanding of the role of mismatch correction in human cancer. However, in order to unravel all the complexities involved in post-replicative mismatch correction, we need to know the cast and the roles of the individual players. This brief treatise provides an overview of our current knowledge of the biochemistry of this process.

Keywords: apoptosis/drug resistance/hereditary non-polyposis colon cancer/microsatellite instability/mismatch repair

Introduction

Maintenance of genomic stability is one of the key criteria that govern the survival of species. DNA is a reactive molecule and as such is modified continuously by a broad range of agents. Although exogenous sources of DNA damage, such as ionizing and UV radiation or carcinogens contained in foodstuffs and cigarette smoke, have received the greatest share of attention over the past 30 or so years, DNA is damaged primarily from within: through hydrolysis, methylation or active oxygen species (Lindahl, 1996). Given that a mammalian DNA molecule can undergo ~100 000 modifications per day, it is clear that life on Earth would be extinct by now were it not for DNA repair (Friedberg *et al*., 1995). This term encompasses a multitude of metabolic processes, which can reverse the damage either directly (photolyases, alkyltransferases), or indirectly by removing damaged bases (base excision) or oligonucleotides (nucleotide excision) from DNA, and resynthesizing the removed patch. Moreover, they can repair strand breaks and cross-links, or by-pass non-repairable damage either by specialized mechanisms or by recombination (Friedberg *et al*., 1995). The main task of DNA repair is to ensure that the DNA molecule is free of modifications or mutations, such that it can be transcribed efficiently and, most importantly, that it can be replicated faithfully and passed on to progeny cells.

DNA replication is a complex process, whose fidelity, estimated to be in the range of one error per 10^{10} nucleotides synthesized (Kunkel, 1992), depends on three factors: DNA polymerase(s), exonucleolytic proofreading and mismatch repair (MMR). Replicative DNA polymerases are extremely precise enzymes, which can duplicate the sequence of a given genome within minutes, or at most hours, with an error rate of ~10^{-5} (Kunkel, 1992). The high precision with which complementary nucleotides are added to the end of the primer strand is guaranteed by the formation of Watson–Crick base pairs in the active site of the enzyme. This process is controlled by the thermodynamic stability afforded by the formation of two or three hydrogen bonds between the incoming nucleoside triphosphate and the first non-paired base of the template strand, as well as by spacial restrictions within the polymerase active site.

Should a non-complementary nucleotide be incorporated at the end of the primer, extension from such a primer is highly inefficient (Benkovic and Cameron, 1995) and it is in fact likely that the inability of DNA polymerases to go forward in such situations leads to a translocation of the mispaired primer terminus into the active site of the proofreading 3′→5′ exonuclease activity that is associated with all replicative polymerases and that adds a further two orders of magnitude (Kunkel, 1992) to the fidelity of the replication process by excising the terminal non-complementary nucleotide, together with two or three additional residues, from the end of the primer terminus.

In some cases, a mispair manages to elude the proofreading process. This happens mostly with the G/T wobble pair, which is stabilized by two hydrogen bonds and brings about only a slight distortion of the double helix (Hunter *et al*., 1987). As it is the one mispair from which most polymerases are able to extend with the least difficulty (Benkovic and Cameron, 1995), it also happens to escape the translocation into the proofreading site and exits the polymerase complex. A similar situation can arise when the primer and template strands slip with respect to one another (Kunkel, 1990). Such events occur relatively frequently in runs of repeated mono- or dinucleotides,

where they give rise to loops containing extrahelical nucleotides, the so-called insertion/deletion loops (IDLs). When an IDL is formed by transient dissociation and reassociation of the primer and the template, or by simple slippage, the end of the primer strand will anneal with the template to produce a hydrogen-bonded terminus from which the polymerase can extend (Kunkel, 1993).

Replication-associated transactions such as mispairs and IDLs that have escaped the proofreading exonuclease become substrates for MMR, whose task is to restore the information contained in the template strand. In this respect, MMR differs from all other DNA repair pathways. Mismatches or IDLs are composed of unmodified nucleotides and exist as detectable moieties in DNA solely in the double-stranded form, inasmuch as separation of the two strands yields two unmodified, intact DNA molecules, neither of which contains recognizable—and therefore repairable—damage. Successful restoration of the original DNA sequence thus requires (i) a factor capable of recognizing base–base mismatches and IDLs, and (ii) a factor that can distinguish between the parent and the daughter strand. It is beyond the scope of this article to review the genetics and the biochemistry of MMR in its entirety. Fortunately, following the discovery of a link between MMR malfunction and a frequent form of human cancer, hereditary non-polyposis colon cancer (HNPCC), this subject has become the topic of intense research and, as a result, interested readers can find a detailed treatise in numerous reviews on MMR that have appeared in the recent literature (Fishel and Kolodner, 1995; Kolodner, 1995, 1996; Radman *et al*., 1995; Jiricny, 1996; Marra and Boland, 1996; Modrich and Lahue, 1996; Crouse, 1997; Fishel and Wilson, 1997; Modrich, 1997; Fink *et al*., 1998). The scope of this minireview is to discuss the recent developments in our understanding of the biochemistry of the MMR process and to address some issues that have been little appreciated or those that merit, in my view at least, particular attention.

Mismatch repair in *E.coli*

The process of post-replicative mismatch correction in *E.coli* represents the best characterized system to date (for reviews see Modrich, 1991; Modrich and Lahue, 1996). The individual factors associated with mismatch recognition, strand discrimination, exonucleolytic degradation of the error-containing strand, resynthesis of the repair patch and ligation have been identified, and there is reason to believe that at least our global concept of the process is correct, as MMR reconstituted from these 10 factors functions *in vitro* (Lahue *et al*., 1989). Mechanistically, mismatch correction in *E.coli* is thought to proceed as outlined in Figure 1. In the following paragraphs, I shall address some points regarding this model that remain to be clarified.

Criteria for mismatch recognition

There are still numerous details that we do not fully understand. One of these concerns the initial step of the correction cascade, namely mismatch recognition, a role fulfilled by the homodimeric MutS protein (Su and Modrich, 1986). It could be shown in different *in vitro* assays that this factor binds with high affinity to substrates containing most base–base mispairs (Su and Modrich, 1986; Su *et al*., 1988) and IDLs up to four extrahelical nucleotides (Parker and Marinus, 1992). How these substrates are recognized is not understood. The distortions brought about in the double helix by a G/T mispair as compared with a single extrahelical nucleotide, for example, are dramatically different. The former is hydrogen-bonded, well-stacked and relatively stable (Hunter *et al*., 1987), whereas the latter either causes a considerable bend and/or perturbation of stacking interactions in the DNA if the extra base is intercalated, or a bulky loop if the unpaired nucleotide is extrahelical. Remarkably, both substrates are bound with a similar affinity in the band-shift assay. More surprisingly still, notable differences in *in vitro* binding affinity were observed for oligonucleotide substrates containing the G/T or the A/C mispair, where the former was bound strongly and the latter almost undetectably, even though the purines and the pyrimidines were in an identical sequence context. Yet the C/A mispair was bound as well as the G/T and much more strongly than the T/G in these substrates (Jiricny *et al*., 1988b). These puzzling observations were confounded further by the fact that the efficiency of repair of all these mispairs *in vivo* was similar. Thus, the take-home message from the binding studies is that affinity of the protein for a particular mispair or a DNA modification *in vitro* cannot be taken as an indication of repair efficiency *in vivo*. It would be interesting to examine the fine structures of these particular mismatch combinations and IDLs by high-resolution NMR or by X-ray crystallography, such that we might gain an important insight into the types of structural determinants recognized by MutS.

The fact that this mismatch-binding protein has so far eluded crystallization is very frustrating indeed. However, a recent study carried out with the *Thermus aquaticus* MutS homologue has provided us with the first insight into the interaction of this multifaceted polypeptide with its substrate. Using an iododeoxyuridine-labelled oligonucleotide in UV cross-linking, Malkov *et al*. (1997) were able to demonstrate that the MutS interaction with DNA was mediated via Phe39, which is situated within a short, but highly conserved motif in the N-terminal domain of the protein. Substitution of this residue for Ala decreased the affinity of the mutant protein for DNA by three orders of magnitude, but the mutation did not affect the ability of the protein to dimerize, nor did it alter its ATPase activity. This suggests that MutS and its homologues interact with DNA via their N-termini, while the C-terminal domains house the dimerization and ATP-binding domains. More studies of this kind would be welcome.

Repair complex assembly

The second mystery concerns the transactions that follow mismatch recognition. Electron microscopic data (Allen *et al*., 1997) show that the mismatch-bound MutS homodimer undergoes an ATP-driven translocation along the DNA such that a looped structure is formed, with the protein(s) sitting at its base. (This structure is sometimes referred to as the α-loop, although I prefer to think that it resembles the Greek letter Ω rather than α.) This process is accelerated by the addition of the MutL homodimer, which appears to co-localize with MutS. Measurement of

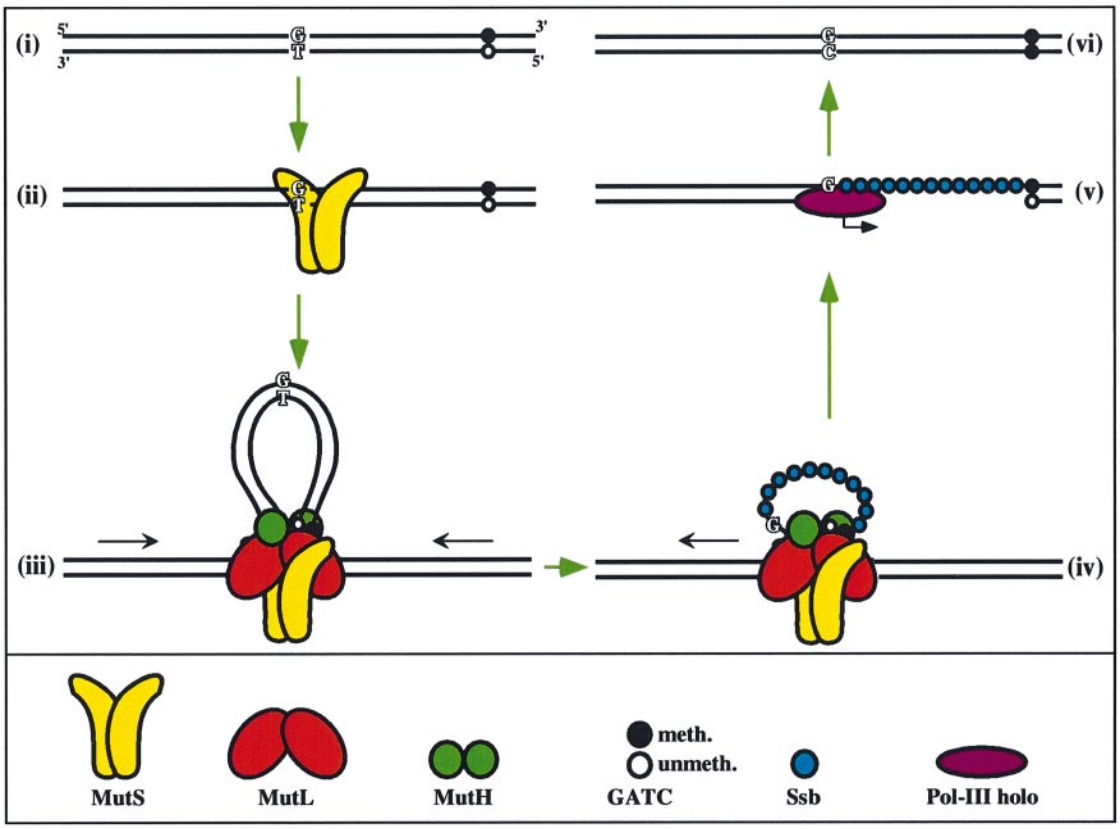

Fig. 1. Suggested mechanism for mismatch correction in *E.coli*. (i) The G/T mispair that arose as an error of DNA polymerase is present in a DNA heteroduplex, which is transiently unmethylated at a GATC sequence in the newly replicated strand. The template GATC site is methylated. (ii) Binding of the mismatch by the homodimeric MutS protein initiates a cascade of events that starts with an ATP-dependent conformational change of MutS, followed by a movement of the bound protein away from the mispair (not shown). (iii) ATP hydrolysis drives the bi-directional movement of the DNA through the bound MutS in the direction of the arrows, as well as the assembly of a multiprotein complex containing the MutS and MutL homodimers and probably also two molecules of the strand discrimination factor MutH, all bound at the base of a looped structure. The assembly of the complex activates the endonucleolytic activity of MutH, which cleaves the newly synthesized DNA 5′ from the unmethylated GATC sequence. (iv) The cleaved strand is then degraded from the nick up to and slightly past the mismatch site either by *Exo*VII or *Rec*J (in cases where the unmethylated GATC was situated 5′ from the mispair), or by *Exo*I (when the nick was 3′ from the mispair). The single-stranded region thus generated is protected by the single strand-binding protein Ssb. This step is discussed in more detail in the text. (v) Polymerase III holoenzyme fills the gap and DNA ligase repairs the nick. (vi) The process is completed by methylation of the GATC site by Dam methylase, at which point the substrate becomes refractory to further action by the MMR system.

the loop size and of the mispair-to-terminus distances instigated the authors to suggest that the translocation resulted in a release of the mispair by the protein, such that it is now at the apex of the loop (see Figure 1). This hypothesis has several far-reaching implications. Binding experiments demonstrated that the affinity of MutS for a homoduplex is several fold lower than for a mismatch-containing heteroduplex (Su *et al*., 1988). Yet, if the mismatch-bound protein were indeed to release the mispair in the presence of ATP, it would be bound to homoduplex DNA. It is difficult to imagine why the protein should let go of its preferred substrate, unless it were to undergo an ATP-dependent conformational change which would increase its affinity for homoduplex DNA. The electron microscopic data (Allen *et al*., 1997), moreover, imply that the MutS homodimer sitting at the base of the loop interacts with two double helices, i.e. that each subunit binds to the same homoduplex DNA molecule, albeit several hundred base pairs apart. The only evidence that supports the conclusion that both DNA-binding sites in the MutS homodimer can be occupied comes from studies carried out with the MutS homologue from *T.aquaticus*, which appears to interact with oligonucleotide substrates with a 1:1 stoichiometry, i.e. two oligonucleotides bound to one MutS homodimer (P.Hsieh, personal communication). It would therefore appear feasible that the MutS homodimer could bind to the homoduplex regions at the base of the loop as depicted in Figure 1, where the ATP-directed conformational change, coupled to the gain in energy through interaction with two DNA homoduplexes (rather than one heteroduplex), would ensure stability of the complex.

It is possible that the conformational change of MutS is catalysed by MutL. In a series of gel-shift experiments, Marinus and colleagues (Drotschmann *et al*., 1998) have noticed that the presence of MutL in the MutS binding assay increased the efficiency of MutS binding. As the size of the protein–DNA complex did not increase, at least as measured by its mobility through native polyacrylamide gels, it would appear that the MutL protein somehow assisted the loading of MutS onto the mismatched substrate, without being itself a part of the complex. It is interesting to note in this context that the MutL protein has already been assigned a role of a ‘molecular matchmaker’, in that it is thought to mediate the interaction between MutS and MutH (Modrich, 1991). Moreover, the

role of MutL as a possible chaperone is underscored by the sequence homology of its N-terminus with those of Hsp90 and type II DNA topoisomerases (Bergerat *et al*., 1997; Grenert *et al*., 1997). Although no evidence of a DNA–MutS–MutL ternary complex was obtained in gel-shifts, the experiments of Allen *et al*. (1997) provide evidence for the fact that the three polypeptides are located together at the base of the loop.

Recent evidence suggests that MutL might play a role also in loading the UvrD (MutU) helicase at the site of the nick (Dao and Modrich, 1998; Yamaguchi *et al*., 1998). This represents one of the key steps of the repair process, inasmuch as it would help the helicase initiate the unwinding process, due to the fact that most helicases have difficulties with initiation at a nick within a double-stranded DNA molecule. More importantly though, the MutL–UvrD interaction may control the directionality of the unwinding process such that it proceeds towards, rather than away from, the mispair. It is conceivable that the torsional strain of the looped structure might also contribute to the directionality of the process of unwinding and exonucleolytic degradation of the error-containing strand.

The model as shown in Figure 1 therefore appears to fit most of the experimental data, with one notable exception: repair-patch size. As mentioned above, the EM results of Allen *et al*. (1997) show that the mismatch-dependent, ATP-driven translocation of the MutS–MutL complex involves the bidirectional threading of the DNA through the proteins such that the released mispair finds itself roughly equidistant from the two extremities of the loop. However, measurement of the repair patch size in the same system showed that the degradation began at the nick and proceeded up to and just past the mismatch (Su *et al*., 1989). This implies that the exonucleolytic degradation of the newly synthesized strand stops shortly after the non-complementary nucleotide had been removed. In the repair of damaged bases or nucleotides, such a possibility would not be questioned. However, as mentioned above, mismatches do not contain modified nucleotides and are undetectable when the two DNA strands have been separated by, for example, a helicase. The fact that the degradation mostly stops only a short distance past the site where the mismatch had been suggests that the site is either tagged, or that, following the MutH-mediated binding and nicking of the nearest GATC site, the threading direction is reversed, such that the mismatch returns to the MutS fold. This would now be an asymmetric process, because only one half of the loop is free to migrate—the other being bound by the MutH protein. This model is purely speculative; we have no evidence at present in support of either mechanism. Further experiments aimed at refining it are clearly needed.

Mismatch repair in human cells

The discovery of sequence similarities between the PMS1 protein of *Saccharomyces cerevisiae* and the MutL protein of *E.coli* (Kramer *et al*., 1989) represented an important milestone in the study of eukaryotic MMR, inasmuch as it provided the first evidence of a close conservation of the pathway between prokaryotes and higher organisms. Moreover, it became clear that the post-meiotic segregation (PMS) phenotype, described by Fogel and colleagues (Williamson *et al*., 1985), was closely linked with MMR and thus implied an important role for this process in meiotic recombination. The availability of yeast MMR mutants helped us to understand the phenotypes of human cells with deficiencies in MMR. Thus, while a strong mutator phenotype was widely expected, MMR deficiency appeared to accentuate an instability of microsatellite repeats (Strand *et al*., 1993), a trait that was to prove of immense importance in the identification of human tumours with MMR gene defects (Boyer *et al*., 1995). Moreover, the characterization of the *S.cerevisiae* MutS and MutL homologues (MSH and MLH, respectively) made possible the identification of human MMR genes by degenerate PCR (Leach *et al*., 1993; Fishel *et al*., 1994c; Nicolaides *et al*., 1994; Papadopoulos *et al*., 1994; Kolodner *et al*., 1995), thanks to the conservation of the amino acid sequences from bacteria to yeast to man. Due to this conservation, the MMR process in humans appears at first glance very similar to that shown in Figure 1. However, upon closer examination, several key differences emerge, the most notable of which are the heterodimeric nature of both MutS and MutL homologues and the lack of a strand discrimination mechanism resembling the Dam methylase–MutH endonuclease interplay. Moreover, the system has an in-built redundancy.

MutS homologues and mismatch recognition

A factor binding with high affinity to oligonucleotide substrates containing G/T mispairs was first described in 1988 (Jiricny *et al*., 1988a). It was later purified to near homogeneity (Hughes and Jiricny, 1992) and, due to its preference for G/T mispair-containing substrates, was named GTBP (G/T-binding protein). The purest fraction contained two polypeptides of apparent molecular mass of 100 and 160 kDa, which were at the time thought to be related, in that the smaller protein was thought to be a proteolytic fragment of the 160 kDa protein. However, sequencing of tryptic peptides derived from the two proteins revealed that they were encoded by distinct genes (Palombo *et al*., 1995). The 100 kDa protein was shown to be the product of the *hMSH2* gene (Palombo *et al*., 1994), which was identified a few months previously (Leach *et al*., 1993; Fishel *et al*., 1994c) as the first locus of HNPCC, on chromosome 2p15–16 (Peltomaki *et al*., 1993). hMSH2, which stands for human MutS homologue 2, was so named because it is closely related to the *S.cerevisiae* MSH2 protein, first described in 1992 (Reenan and Kolodner, 1992). The gene encoding the 160 kDa polypeptide co-localizes with *hMSH2* (Papadopoulos *et al*., 1995). It originally was named *GTBP* (Hughes and Jiricny, 1992), but because of the close relatedness of its product to the *S.cerevisiae* MSH6 protein described some months later (Marsischky *et al*., 1996), the name *hMSH6* has now been adopted.

hMSH2 and hMSH6 are tightly associated and co-purify through a series of chromatographic steps (Hughes and Jiricny, 1992; Drummond *et al*., 1995). The purified heterodimer, often referred to as hMutSα, has been shown to complement extracts of cells mutated in *hMSH2* or *hMSH6* in an *in vitro* MMR assay (Drummond *et al*., 1995). Attempts at expression of the individual proteins in the baculovirus system met with only partial success.

hMSH2 was expressed in high yields and could be purified easily to homogeneity (Fishel *et al*., 1994a,b; Iaccarino *et al*., 1998). Although we found the protein to lack specific mismatch-binding activity in our gel-shift assays (Iaccarino *et al*., 1998), several reports using recombinant hMSH2 described its ability to bind to substrates containing base–base mispairs, as well as IDLs of varying sizes (Fishel *et al*., 1994a,b). Similar results were reported also for the yeast MSH2 protein (Alani *et al*., 1995), with the sole difference that the binding of the human protein to heteroduplex DNA was augmented in the presence of ATP (Fishel *et al*., 1994b), whereas the *S.cerevisiae* protein binding was unaltered in the presence of the nucleotide (Alani *et al*., 1995). We currently are attempting to identify the cause underlying these differences.

Expression of hMSH6 was less successful. The full-length protein appeared proteolytically labile and could be isolated only in poor yields. It displayed no mismatch-binding activity (Iaccarino *et al*., 1998). It was, however, considerably stabilized when co-expressed with hMSH2 (Palombo *et al*., 1996; Gradia *et al*., 1997; Iaccarino *et al*., 1998), which permitted the isolation of soluble hMutSα in high yields. This heterodimeric factor bound G/T-containing substrates with a low nanomolar affinity (Gradia *et al*., 1997; Iaccarino *et al*., 1998).

All MutS homologues characterized to date are highly conserved at their C-terminal regions, which contain the consensus ATP-binding site. Addition of ATP to gel-shift assays results in an apparent dissociation of the proteins from the oligonucleotide substrates (Hughes and Jiricny, 1992; Gradia *et al*., 1997; Iaccarino *et al*., 1998), which is most likely to be the result of their running off the end of the short duplex following the ATP-dependent translocation described above for MutS (see also below). This ATP-dependent process requires only ATP binding, not ATP hydrolysis (Hughes and Jiricny, 1992; Gradia *et al*., 1997; Iaccarino *et al*., 1998). Mutations within the conserved GxxxxGKS motif of the *Salmonella typhimurium* MutS (Haber and Walker, 1991), the *S.cerevisiae* MSH2 and MSH6 (Alani *et al*., 1997; Studamire *et al*., 1998) or the human hMSH2 and hMSH6 proteins (Iaccarino *et al*., 1998) had little influence on mismatch binding, but brought about a considerable decrease in the sensitivity of the protein–DNA complex to ATP. Assuming that, following binding at the mismatch site, hMutSα undergoes an ATP-dependent conformational change and translocates along the double helix in a manner similar to the MutS homodimer, thereby generating a looped structure, it would be predicted that ATPase-defective mutants would be MMR deficient. This is indeed the case, although both hMSH2 and hMSH6 sites needed to be mutated before MMR activity was lost in an *in vitro* assay (Iaccarino *et al*., 1998).

Recently, Fishel and colleagues (Gradia *et al*., 1997) suggested that the hMSH2/6 heterodimer acts as a molecular switch homologous in some way to G proteins, being ON for mismatch binding in an ADP-bound form and OFF when complexed with ATP. This suggestion stems from the finding that, as already mentioned, the heterodimer forms stable complexes with mismatch-containing oligonucleotide substrates, which dissociate upon the addition of ATP. This interesting hypothesis needs to be substantiated further. Previous experiments have shown that the heterodimer does not require any cofactors for binding to heteroduplex DNA; it binds oligonucleotide substrates with a similar affinity in the presence or absence of ADP (Hughes and Jiricny, 1992), and it can exchange ADP for ATP without the need for additional exchange factors (I.Iaccarino and J.Jiricny, unpublished). However, it is possible that the presence of other proteins, such as the MutL homologues, might accelerate the exchange. It seems likely that this ATP-dependent switch is in fact a conformational change of hMutSα that is required for the binding factor to leave the mismatch site and thus initiate the downstream repair events as proposed (Gradia *et al*., 1997). There is evidence that a conformational change does indeed take place upon ATP binding by the *S.cerevisiae* MSH2/6 heterodimer (Studamire *et al*., 1998).

Although the *in vitro* binding experiments are interesting, they do not provide us with much information regarding the sequence of events at the mismatch following the binding of hMutSα. It is unlikely that the factor will leave DNA upon ATP binding (as suggested by the gel-shift experiments), without first either marking the site or inducing the assembly of the repairosome. As the oligonucleotide substrates are apparently intact and unmodified following ATP-driven hMutSα dissociation, we can only assume that mismatch binding catalyses the assembly of the repair complex in a manner similar to that described for *E.coli* (see Figure 1). This would imply that the complex remains bound to the DNA, but not at the mismatch, i.e. that it translocates following ATP binding. Although EM evidence of this process in the human system is not yet available, ATP fails to mediate the dissociation of hMutSα from a 200 bp mismatch-containing duplex when both ends of the molecule are blocked with streptavidin (L.Blackwell and P.Modrich, personal communication). This suggests that the protein slides along the molecule as proposed for MutS.

Although hMSH2 and hMSH6 function as a heterodimer in MMR, the phenotype of cells mutated in the *hMSH2* gene is dramatically different from *hMSH6* mutants. The former cells are deficient for the repair of base–base mismatches and small IDLs, and correspondingly display instability in microsatellite sequences consisting of runs of mono-, di-, tri- and tetranucleotides. In contrast, *hMSH6* mutants are deficient for the repair of base–base mismatches and IDLs of one nucleotide, but dinucleotide repeat instability was not noted in these cells (Papadopoulos *et al*., 1995) and their extracts are partially proficient for the repair of IDLs of two nucleotides or more (Drummond *et al*., 1995). These finding led to the proposal (Karran, 1995) that loop repair is dependent on another mismatch-binding factor that contains hMSH2, but not hMSH6. This hypothesis could soon be substantiated. Genetic experiments in *S.cerevisiae* demonstrated that the mutator phenotype of the *msh2* strain was similar to that of a double mutant in *msh3* and *msh6* (Marsischky *et al*., 1996), and it was suggested that the MSH3 and MSH6 proteins can both interact with MSH2, whereby the MSH2/3 heterodimer displays affinity for IDLs, while the MSH2/6 factor might preferentially bind base–base mismatches and loops of one nucleotide. Indeed, using recombinant hMSH2/6 (hMutSα) and hMSH2/3 (hMutSβ) heterodimers expressed in the baculovirus system (Acharya *et al*., 1996; Palombo *et al*., 1996), it could be shown that the former

could bind base–base mismatches and IDLs of one and two nucleotides, while the latter failed to bind to oligonucleotide substrates containing base–base mismatches, but displayed considerable affinity for IDL substrates. In other studies, using proteins purified from human HeLa cells, it could be shown that hMutSα could complement hMSH2-deficient extracts for the repair of base–base mismatches and small as well as large loops, while only loop repair was restored to these extracts by the addition of hMutSβ (Genschel *et al*., 1998; Marra *et al*., 1998). It was proposed, therefore, that hMutSα most likely initiates the repair of both base–base mispairs and IDLs under most circumstances. Support for this hypothesis came on one hand from the examination of extracts of human cell lines, which showed that hMutSα is significantly more abundant than hMutSβ (Drummond *et al*., 1997) and, on the other hand, from the finding that cells lacking hMSH3 do not display microsatellite instability (Inokuchi *et al*., 1995). In agreement with these findings is also the fact that *msh3* mutants in *S.cerevisiae* have only a very weak mutator and PMS phenotype (Strand *et al*., 1995). hMutSα and hMutSβ are therefore functionally redundant for the repair of IDLs, although it seems probable that the latter fulfils only a back-up function in MMR.

The existence of hMutSα and hMutSβ implies that hMSH3 and hMSH6 have to compete for hMSH2. As already mentioned above, the outcome of this competition in normal cells is strongly biased in favour of hMutSα (Drummond *et al*., 1997; Marra *et al*., 1998), with the result that these cells are MMR proficient. However, this equilibrium could easily change if hMSH3 were overexpressed, due to the fact that hMutSβ acts in the repair of IDLs and not base–base mispairs. Such a situation was shown to exist in methotrexate-resistant cell lines, where amplification of the *DHFR* locus resulted in a co-amplification of the *hMSH3* gene (Drummond *et al*., 1997; Marra *et al*., 1998). It could be demonstrated that extracts of these cells contained normal levels of hMSH2, but that hMSH6 levels were significantly decreased, most probably due to the proteolytic degradation of the partnerless protein. In contrast, hMSH3 was abundant in these extracts. Correspondingly, these cells have a strong mutator phenotype (Caligo *et al*., 1990; Drummond *et al*., 1997) and are defective in the repair of base–base mismatches, but are proficient for IDL repair *in vitro* (Drummond *et al*., 1997; Marra *et al*., 1998). The clinical relevance of this phenomenon *in vivo* is not known at this time; however, its importance lies in the demonstration that MMR deficiency can be attained also without mutation, through a simple dysregulation of expression of one of its components.

Role of MutL homologues

It could be shown that, similarly to the MutS homologues, the human counterpart of MutL is a heterodimeric factor, hMutLα, consisting of hMLH1 and hPMS2 (Li and Modrich, 1995). As in the case of the bacterial MutL, the precise biochemical role of the eukaryotic MutL homologues in MMR is unclear. It was anticipated that hMutLα will interact with a mismatch-bound hMutSα in a fashion similar to MutS and MutL. Indeed, in a gel-shift experiment (Prolla *et al*., 1994), the *S.cerevisiae* MSH2 complex with a mismatch-containing oligonucleotide was supershifted in the presence of MLH1 and PMS1 (note that PMS1 is the *S.cerevisiae* homologue of hPMS2). The caveat of this experiment is that MSH2 has since been shown to exist in a complex with MSH6 (see above), the latter protein being absent from this experiment. The relevance of this finding is therefore uncertain at this time. However, the *S.cerevisiae* MLH1–PMS1 complex was shown to enhance the binding of the MSH2/3 heterodimer to an IDL substrate (Habraken *et al*., 1997), and interaction of human MutS and MutL homologues on a DNA substrate was demonstrated recently by co-immunoprecipitation studies, where the anti-hMLH1 antibody also succeeded in bringing down hPMS2 and hMSH2 (the authors did not test for the presence of hMSH6, but it can be assumed that it too was present in the precipitate). This reaction was ATP-dependent (Gu *et al*., 1998). Interestingly, the co-immunoprecipitated proteins also contained proliferating cell nuclear antigen (PCNA), whose possible role in MMR is discussed below.

Strand discrimination in human MMR

As mentioned earlier, the process of correction of replication errors has to be directed to the newly synthesized strand. This means that the MMR system must possess a strand discrimination function. In *E.coli* and *S.typhimurium*, strand discrimination makes use of a transient undermethylation of the newly synthesized strand, and employs an endonuclease, the MutH protein, capable of specifically cleaving the unmethylated strand of the hemimethylated DNA molecule (Figure 1). This mechanism is very elegant, but it should be remembered that this Dam-directed system represents the exception rather than the rule. Thus, in *Streptococcus pneumoniae*, no such system appears to operate (Balganesh and Lacks, 1985). In lower eukaryotes such as *S.cerevisiae*, there is no or only very little DNA methylation, and a similar situation applies in *Drosophila melanogaster*. Vertebrates and plants do methylate their DNA, mostly at the 5-position of cytosine in CpG or CpXpG motifs, but the methylation patterns are irregular, inasmuch as there are stretches of DNA up to several kilobases long, the so-called CpG islands, which remain unmethylated. MutH-like, methylation-sensitive endonuclease involved in the correction of a mismatch in such a sequence stretch could nick either strand. Such an event would lead in the best case to undirected repair and in the worst scenario to a double-strand break. Thus, contrary to one published report (Hare and Taylor, 1985), methylation is unlikely to direct MMR in higher organisms. Yet, all these systems possess a post-replicative MMR system capable of strand discrimination. So how does the MMR system manage to distinguish between the template and the primer strand? *In vitro* studies of the MMR process demonstrated that strand discontinuities, such as nicks, can direct the mismatch-dependent exonucleolysis to the nicked strand (Holmes *et al*., 1990; Thomas *et al*., 1991). This is true for all the above-mentioned systems, including *E.coli* where a single nick can substitute for MutH (Au *et al*., 1992). Lacks and colleagues proposed a long time ago that the MMR system of *S.pneumoniae* is nick-directed (Balganesh and Lacks, 1985). Indeed, taking as a model the lagging strand, we know that the Okazaki fragments are perhaps 2 kb long and that a mismatch anywhere within such a piece of

DNA is well within the reach of the MMR system, which could use either end of the fragment for the initiation of the repair process. The problem was always presented by the leading strand, where DNA synthesis is continuous. Where is the discrimination signal in this strand? Is it possible that the template DNA contains an as yet unidentified epigenetic modification that could be used to distinguish between it and the newly synthesized strand? We have as yet no answers to these questions. However, recent evidence coming from several different experimental systems has begun to shed light on this problem. As briefly noted above, immunoprecipitation experiments with anti-hMLH1 antibodies also co-precipitated hMSH2, hPMS2 and PCNA (Gu *et al*., 1998). The reason why these authors looked for PCNA was most likely thanks to seminal work coming from Kunkel's laboratory, where a two-hybrid screen search for MLH-interacting proteins picked out a full-length cDNA clone encoding the *S.cerevisiae* PCNA (Umar *et al*., 1996). This was the first experimental evidence which hinted at an interaction between DNA replication and MMR, especially as *in vitro* MMR experiments implied that PCNA was involved not only in the expected resynthesis of the excised repair tract (Gu *et al*., 1998), but also at a step prior to this (Umar *et al*., 1996). Based on these results, the authors (Umar *et al*., 1996) suggested that PCNA bound to the primer termini at the replication fork might itself act as the strand discrimination signal. Extending this prediction a little further, I would offer the following model as to how strand discrimination in eukaryotes might work. Let us assume that the replicative polymerase has just incorporated a non-complementary nucleotide at the end of the primer strand, which then escaped the proofreading activity of the replication complex. If the MMR proteins were able to physically interact with the replication enzymes via PCNA, detection and binding of the mispair by the MSH2/6 heterodimer, followed by a threading process that would eventually contact PCNA in the polymerase complex, would arrest the forward movement of the replication fork. This may well result in the dissociation of the polymerase from the primer terminus, an event which would leave the 3′ nucleotide of the newly synthesized strand free and therefore accessible to any eventual 3′–5′ exonucleases, which could then initiate the degradation of the primer strand towards the mispair (Figure 2). An equally feasible alternative is that the MMR proteins are linked directly to PCNA throughout the replication process. Although we have at present no experimental evidence in favour of this model, its beauty lies in the fact that it requires no MutH homologue for strand discrimination. The fact that no open reading frame which could be considered related to MutH (other than bacterial restriction endonucleases) has been detected in an ever-increasing number of organisms whose genomes have been completely (or almost completely) sequenced should come as no surprise, because MutH is a sequence-specific, methylation-sensitive endonuclease, with a preference for an epigenetically modified motif absent from the DNA of other organisms. However, it would appear probable that MMR in most species dispenses with a MutH-like function. Circular, mismatch-containing heteroduplex molecules lacking a nick are generally refractory to MMR *in vitro* (Holmes *et al*., 1990; Thomas *et al*., 1991) and *in vivo* (Brown and Jiricny, 1987, 1988). (The small amount of repair in the latter experiments could be due to the introduction of sporadic strand breaks into the substrate DNA during the transfection procedure.) The notable exception to the rule might

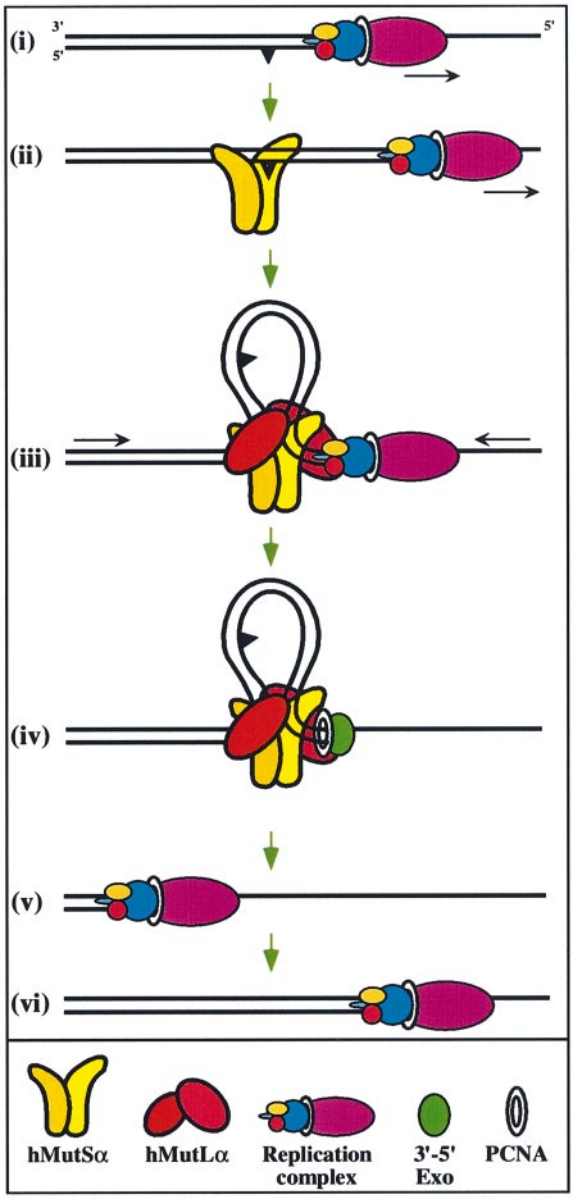

Fig. 2. Putative model of strand discrimination of human mismatch repair in the leading strand. (i) The mismatch, introduced into the newly synthesized strand by the replication complex consisting of pol-δ, PCNA, RFC and other replication factors, has escaped the proofreading exonuclease. (ii) The mispair is recognized by hMutSα. (iii) ATP drives the assembly of the MMR complex and the bi-directional threading of DNA through the hMutSα heterodimer until the replication complex is contacted via an interaction with PCNA. This event leads to replication arrest and dissociation of pol-δ. (iv) Exonucleolytic degradation of the error-containing primer strand can commence from its exposed 3′ terminus. (v) Following the dissociation of the MMR complex, the PCNA, still bound at the end of the primer, recruits the replication complex and (vi) the DNA synthesis is re-initiated. Replication protein A (RPA), which probably fulfils the role of Ssb in this process, is covering the single-stranded DNA throughout the process and is not shown here. Another alternative is that the replication complex is physically linked with the MMR proteins already in steps (i) and (ii), and that it acts as a break which halts the polymerase complex upon mismatch binding (see text for details).

be *Xenopus laevis* extracts, which appear to be proficient for MMR on circular substrates (Brooks *et al*., 1989; Varlet *et al*., 1996). However, as the repair patches were short and centred predominantly around the mismatch, the possibility must be considered that the system addressing these mispairs was distinct from the post-replicative mismatch correction pathway.

Other proteins involved in MMR

The proteins participating in MMR can be divided into two categories. Those that are dedicated exclusively to mismatch correction and those that participate in other DNA metabolic pathways. In the case of *E.coli*, the former class is represented by the MutSLH polypeptides, while the helicase, exonucleases, polymerase and ligase belong to the second class. In human cells, we have already listed the members of the MMR-specific group: hMutSα (hMSH2/6) or hMutSβ (hMSH2/3) and hMutLα (hMLH1/hPMS2). A third MutL homologue, hPMS1, has also been identified (Nicolaides *et al*., 1994), but has to-date remained without a known biochemical function. As to the second group, the only activity known to participate in human MMR, at least in the *in vitro* assay, is polymerase-δ (Longley *et al*., 1997). A 5′–3′ exonuclease, EXO1 (Tishkoff *et al*., 1997), has also been implicated in the process, largely based on the phenotype of *S.cerevisiae* and *Schizosaccharomyces pombe* mutant strains, but here it is difficult to say with any certainty what, if any, role these enzymes might play in the MMR process, due to their functional redundancy, as well as to the fact that they also participate in a number of other DNA transactions. RPA, the functional homologue of the bacterial single strand-binding protein Ssb, has also been implicated in the process (Lin *et al*., 1998; Umezu *et al*., 1998). The confirmation of the role(s) of these enzymes must await the results of ongoing biochemical studies. No evidence is as yet available regarding the helicase function. Of the four known human ligases, DNA ligase I is probably the one involved in MMR, but this is only an assumption based on its frequent association with pol-δ- and PCNA-dependent processes (Montecucco *et al*., 1998).

Concluding remarks

Since 1994, there have been close to 300 publications appearing each year that mention the term mismatch repair. Correspondingly, this field is becoming very difficult to review objectively and I would like to give this as a reason for not having cited a great many interesting papers. In the preceding paragraphs, I have tried to discuss those experiments that address some of the key issues in the MMR field that have stimulated me personally into thinking about the molecular mechanism involved in mismatch correction and that have helped me to define where the shortcomings in our knowledge of the process lie. We have yet to identify several of the missing players in the MMR game, before we can reconstitute the system from its individual components such as was achieved in *E.coli* (Lahue *et al*., 1989). We need to know more about protein–protein and protein–DNA interactions; crystallization efforts should be stepped up, but much can be learned also from missense mutations identified in HNPCC families and from mutagenesis studies. By learning more about MMR malfunction, we can learn more about the function of the individual proteins in the process. The involvement of MMR in drug resistance is of great clinical relevance and we need to understand how these mechanisms operate, with the hope of finding agents that kill MMR-deficient cells more effectively. Mismatch repair proteins have been reported to be involved also in other DNA metabolic pathways such as transcription-coupled nucleotide excision repair (NER) (Mellon and Champe, 1996; Mellon *et al*., 1996; Leadon and Avrutskaya, 1998), where we currently have no mechanistic insights as to how the MMR and NER pathways might interact. It is of further clinical interest to identify the cause(s) of the inactivation of the wild-type allele of the MMR gene mutated in HNPCC kindreds and why this happens predominantly in the colonic epithelium. Epigenetic inactivation has been shown to play a role here (Kane *et al*., 1997; Veigl *et al*., 1998), but there must be also other mechanisms at work. Of interest here are experiments carried out by Meuth and collaborators (Richards *et al*., 1997), who showed that some MMR-deficient cells acquire a stronger mutator phenotype under conditions of restrained growth, such as might be found in the context of a tumour. This phenotype is reminiscent of the bacterial ‘adaptive mutation’ phenomenon described several years ago (for a review see Foster, 1998). Another phenomenon that requires further study is the role of the immune system in HNPCC. Frameshift mutations give rise to many aberrant proteins, which are displayed on MHC class I receptors as non-self-peptides. It appears that MMR-deficient cells avoid elimination by the immune system through abolishing expression of the MHC receptors (Bicknell *et al*., 1996). How do these cells escape elimination by the natural killer cells?

These and many more questions require answers before we can say that we understand MMR. It is my hope that better knowledge of this important DNA repair pathway will help us devise better ways of dealing with cancers associated with MMR deficiencies, or better still, for preventing these malignancies altogether.

Acknowledgements

I would like to express my sincere thanks to all my colleagues and collaborators who have been involved with me on this project over the past dozen years or so. I would also like to acknowledge Eric Alani, Gray Crouse, Peggy Hsieh, Richard Kolodner, Tom Kunkel, Paul Modrich and Bert Vogelstein for constructive discussions and for communicating information prior to publication.

References

Acharya,S., Wilson,T., Gradia,S., Kane,M.F., Guerrette,S., Marsischky, G.T., Kolodner,R. and Fishel,R. (1996) hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc. Natl Acad. Sci. USA*, **93**, 13629–13634.

Alani,E., Chi,N.W. and Kolodner,R. (1995) The *Saccharomyces cerevisiae* Msh2 protein specifically binds to duplex oligonucleotides containing mismatched DNA base pairs and insertions. *Genes Dev.*, **9**, 234–247.

Alani,E., Sokolsky,T., Studamire,B., Miret,J.J. and Lahue,R.S. (1997) Genetic and biochemical analysis of Msh2p–Msh6p: role of ATP hydrolysis and Msh2p–Msh6p subunit interactions in mismatch base pair recognition. *Mol. Cell. Biol.*, **17**, 2436–2447.

Allen,D.J., Makhov,A., Grilley,M., Taylor,J., Thresher,R., Modrich,P. and Griffith,J.D. (1997) MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.*, **16**, 4467–4476.

Au,K.G., Welsh,K. and Modrich,P. (1992) Initiation of methyl-directed mismatch repair. *J. Biol. Chem.*, **267**, 12142–12148.

Balganesh,T.S. and Lacks,S.A. (1985) Heteroduplex DNA mismatch repair system of *Streptococcus pneumoniae*: cloning and expression of the *hexA* gene. *J. Bacteriol.*, **162**, 979–984.

Benkovic,S.J. and Cameron,C.E. (1995) Kinetic analysis of nucleotide incorporation and misincorporation by Klenow fragment of *Escherichia coli* DNA polymerase I. *Methods Enzymol.*, **262**, 257–269.

Bergerat,A., de Massy,B., Gadelle,D., Varoutas,P.-C., Nicolas,A. and Forterre,P. (1997) An atypical topoisomerase II from archaea with implications for meiotic recombination. *Nature*, **386**, 414–417.

Bicknell,D.C., Kaklamanis,L., Hampson,R., Bodmer,W.F. and Karran,P. (1996) Selection for β 2-microglobulin mutation in mismatch repair-defective colorectal carcinomas. *Curr. Biol.*, **6**, 1695–1697.

Boyer,J.C., Umar,A., Risinger,J.I., Lipford,J.R., Kane,M., Yin,S., Barrett,J.C., Kolodner,R.D. and Kunkel,T.A. (1995) Microsatellite instability, mismatch repair deficiency and genetic defects in human cancer cell lines. *Cancer Res.*, **55**, 6063–6070.

Brooks,P., Dohet,C., Almouzni,G., Mechali,M. and Radman,M. (1989) Mismatch repair involving localized DNA synthesis in extracts of *Xenopus* eggs. *Proc. Natl Acad. Sci. USA*, **86**, 4425–4429.

Brown,T.C. and Jiricny,J. (1987) A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. *Cell*, **50**, 945–950.

Brown,T.C. and Jiricny,J. (1988) Different base/base mispairs are corrected with different efficiencies and specificities in monkey kidney cells. *Cell*, **54**, 705–711.

Caligo,M.A., Armstrong,W., Rossiter,B.J. and Meuth,M. (1990) Increased rate of base substitution in a hamster mutator strain obtained during serial selection for gene amplification. *Mol. Cell. Biol.*, **10**, 6805–6808.

Crouse,G. (1997) In Nickoloff,J.A. and Hoekstra,M.F. (eds), *DNA Damage and Repair, Vol. 1: DNA Repair in Prokaryotes and Lower Eukaryotes*. Humana Press Inc., Totowa, NJ, pp. 411–448.

Dao,V. and Modrich,P. (1998) Mismatch-, MutS-, MutL- and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. *J. Biol. Chem.*, **273**, 9202–9207.

Drotschmann,K., Aronshtam,A., Fritz,H.J. and Marinus,M.G. (1998) The *Escherichia coli* MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic Acids Res.*, **26**, 948–953.

Drummond,J.T., Li,G.M., Longley,M.J. and Modrich,P. (1995) Isolation of an hMSH2–p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science*, **268**, 1909–1912.

Drummond,J.T., Genschel,J., Wolf,E. and Modrich,P. (1997) DHFR/MSH3 amplification in methotrexate-resistant cells alters the hMutSα/hMutSβ ratio and reduces the efficiency of base–base mismatch repair. *Proc. Natl Acad. Sci. USA*, **94**, 10144–10149.

Fink,D., Aebi,S. and Howell,S.B. (1998) The role of DNA mismatch repair in drug resistance. *Clin. Cancer Res.*, **4**, 1–6.

Fishel,R. and Kolodner,R.D. (1995) Identification of mismatch repair genes and their role in the development of cancer. *Curr. Opin. Genet. Dev.*, **5**, 382–395.

Fishel,R. and Wilson,T. (1997) MutS homologs in mammalian cells. *Curr. Opin. Genet. Dev.*, **7**, 105–113.

Fishel,R., Ewel,A., Lee,S., Lescoe,M.K. and Griffith,J. (1994a) Binding of mismatched microsatellite DNA sequences by the human MSH2 protein. *Science*, **266**, 1403–1405.

Fishel,R., Ewel,A. and Lescoe,M.K. (1994b) Purified human MSH2 protein binds to DNA containing mismatched nucleotides. *Cancer Res.*, **54**, 5539–5542.

Fishel,R., Lescoe,M.K., Rao,M.R., Copeland,N.G., Jenkins,N.A., Garber,J., Kane,M. and Kolodner,R. (1994c) The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, **77**, 167.

Foster,P.L. (1998) Adaptive mutation: has the unicorn landed? *Genetics*, **148**, 1453–1459.

Friedberg,E.C., Walker,G.C. and Siede,W. (1995) In *DNA Repair and Mutagenesis*. ASM Press, Washington, DC, pp. 108–133.

Genschel,J., Littman,S.J., Drummond,J.T. and Modrich,P. (1998) Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα. *J. Biol. Chem.*, **273**, 19895–19901.

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

Grenert,J.P. *et al*. (1997) The amino-terminal domain of heat shock protein 90 (hsp90) that binds geldanamycin is an ATP/ADP switch domain that regulates hsp90 conformation. *J. Biol. Chem.*, **272**, 23843–23850.

Gu,L., Hong,Y., McCulloch,S., Watanabe,H. and Li,G.M. (1998) ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. *Nucleic Acids Res.*, **26**, 1173–1178.

Haber,L.T. and Walker,G.C. (1991) Altering the conserved nucleotide binding motif in the *Salmonella typhimurium* MutS mismatch repair protein affects both its ATPase and mismatch binding activities. *EMBO J.*, **10**, 2707–2715.

Habraken,Y., Sung,P., Prakash,L. and Prakash,S. (1997) Enhancement of MSH2–MSH3-mediated mismatch recognition by the yeast MLH1–PMS1 complex. *Curr. Biol.*, **7**, 790–793.

Hare,J.T. and Taylor,J.H. (1985) One role for DNA methylation in vertebrate cells is strand discrimination in mismatch repair. *Proc. Natl Acad. Sci. USA*, **82**, 7350–7354.

Holmes,J.J., Clark,S. and Modrich,P. (1990) Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. *Proc. Natl Acad. Sci. USA*, **87**, 5837–5841.

Hughes,M.J. and Jiricny,J. (1992) The purification of a human mismatch-binding protein and identification of its associated ATPase and helicase activities. *J. Biol. Chem.*, **267**, 23876–23882.

Hunter,W.N., Brown,T., Kneale,G., Anand,N.N., Rabinovich,D. and Kennard,O. (1987) The structure of guanosine–thymidine mismatches in B-DNA at 2.5-Å resolution. *J. Biol. Chem.*, **262**, 9962–9970.

Iaccarino,I., Marra,G., Palombo,F. and Jiricny,J. (1998) hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSα. *EMBO J.*, **17**, 2677–2686.

Inokuchi,K., Ikejima,M., Watanabe,A., Nakajima,E., Orimo,H., Nomura,T. and Shimada,T. (1995) Loss of expression of the human MSH3 gene in hematological malignancies. *Biochem. Biophys. Res. Commun.*, **214**, 171–179.

Jiricny,J. (1996) Mismatch repair and cancer. *Cancer Surv.*, **28**, 47–68.

Jiricny,J., Hughes,M., Corman,N. and Rudkin,B.B. (1988a) A human 200-kDa protein binds selectively to DNA fragments containing G·T mismatches. *Proc. Natl Acad. Sci. USA*, **85**, 8860–8864.

Jiricny,J., Su,S.S., Wood,S.G. and Modrich,P. (1988b) Mismatch-containing oligonucleotide duplexes bound by the *E.coli* mutS-encoded protein. *Nucleic Acids Res.*, **16**, 7843–7853.

Kane,M.F., Loda,M., Gaida,G.M., Lipman,J., Mishra,R., Goldman,H., Jessup,J.M. and Kolodner,R. (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res.*, **57**, 808–811.

Karran,P. (1995) Appropriate partners make good matches. *Science*, **268**, 1857–1858.

Kolodner,R.D. (1995) Mismatch repair: mechanisms and relationship to cancer susceptibility. *Trends Biochem. Sci.*, **20**, 397–401.

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

Kolodner,R.D. *et al*. (1995) Structure of the human MLH1 locus and analysis of a large hereditary nonpolyposis colorectal carcinoma kindred for *mlh1* mutations. *Cancer Res.*, **55**, 242–248.

Kramer,W., Kramer,B., Williamson,M.S. and Fogel,S. (1989) Cloning and nucleotide sequence of DNA mismatch repair gene PMS1 from *Saccharomyces cerevisiae*: homology of PMS1 to procaryotic MutL and HexB. *J. Bacteriol.*, **171**, 5339–5346.

Kunkel,T.A. (1990) Misalignment-mediated DNA synthesis errors. *Biochemistry*, **29**, 8003–8011.

Kunkel,T.A. (1992) DNA replication fidelity. *J. Biol. Chem.*, **267**, 18251–18254.

Kunkel,T.A. (1993) Nucleotide repeats. Slippery DNA and diseases. *Nature*, **365**, 207–208.

Lahue,R.S., Au,K.G. and Modrich,P. (1989) DNA mismatch correction in a defined system. *Science*, **245**, 160–164.

Leach,F.S. *et al*. (1993) Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell*, **75**, 1215–1225.

Leadon,S.A. and Avrutskaya,A.V. (1998) Requirement for DNA mismatch repair proteins in the transcription-coupled repair of thymine glycols in *Saccharomyces cerevisiae*. *Mutat. Res.*, **407**, 177–187.

Li,G.M. and Modrich,P. (1995) Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. *Proc. Natl Acad. Sci. USA*, **92**, 1950–1954.

Lin,Y.L., Shivji,M.K., Chen,C., Kolodner,R., Wood,R.D. and Dutta,A. (1998) The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair. *J. Biol. Chem.*, **273**, 1453–1461.

Lindahl,T. (1996) The Croonian Lecture, 1996: endogenous damage to DNA. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.*, **351**, 1529–1538.

Longley,M.J., Pierce,A.J. and Modrich,P. (1997) DNA polymerase delta is required for human mismatch repair *in vitro*. *J. Biol. Chem.*, **272**, 10917–10921.

Malkov,V.A., Biswas,I., Camerini-Otero,R.D. and Hsieh,P. (1997) Photocross-linking of the NH_2-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA. *J. Biol. Chem.*, **272**, 23811–23817.

Marra,G. and Boland,C.R. (1996) DNA repair and colorectal cancer. *Gastroenterol. Clin. North Am.*, **25**, 755–772.

Marra,G., Iaccarino,I., Lettieri,T., Roscilli,G., Delmastro,P. and Jiricny,J. (1998) Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc. Natl Acad. Sci. USA*, **95**, 8568–8573.

Marsischky,G.T., Filosi,N., Kane,M.F. and Kolodner,R. (1996) Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes Dev.*, **10**, 407–420.

Mellon,I. and Champe,G.N. (1996) Products of DNA mismatch repair genes *mutS* and *mutL* are required for transcription-coupled nucleotide-excision repair of the lactose operon in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **93**, 1292–1297.

Mellon,I., Rajpal,D.K., Koi,M., Boland,C.R. and Champe,G.N. (1996) Transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science*, **272**, 557–560.

Modrich,P. (1991) Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.*, **25**, 229–253.

Modrich,P. (1997) Strand-specific mismatch repair in mammalian cells. *J. Biol. Chem.*, **272**, 24727–24730.

Modrich,P. and Lahue,R. (1996) Mismatch repair in replication fidelity, genetic recombination and cancer biology. *Annu. Rev. Biochem.*, **65**, 101–133.

Montecucco,A. *et al*. (1998) DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. *EMBO J.*, **17**, 3786–3795.

Nicolaides,N.C. *et al*. (1994) Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature*, **371**, 75–80.

Palombo,F., Hughes,M., Jiricny,J., Truong,O. and Hsuan,J. (1994) Mismatch repair and cancer. *Nature*, **367**, 417.

Palombo,F., Gallinari,P., Iaccarino,I., Lettieri,T., Hughes,M., D'Arrigo,A., Truong,O., Hsuan,J.J. and Jiricny,J. (1995) GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science*, **268**, 1912–1914.

Palombo,F., Iaccarino,I., Nakajima,E., Ikejima,M., Shimada,T. and Jiricny,J. (1996) hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.*, **6**, 1181–1184.

Papadopoulos,N. *et al*. (1994) Mutation of a mutL homolog in hereditary colon cancer. *Science*, **263**, 1625–1629.

Papadopoulos,N. *et al*. (1995) Mutations of GTBP in genetically unstable cells. *Science*, **268**, 1915–1917.

Parker,B.O. and Marinus,M.G. (1992) Repair of DNA heteroduplexes containing small heterologous sequences in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **89**, 1730–1734.

Peltomaki,P. *et al*. (1993) Genetic mapping of a locus predisposing to human colorectal cancer. *Science*, **260**, 810–812.

Prolla,T.A., Pang,Q., Alani,E., Kolodner,R.D. and Liskay,R.M. (1994) MLH1, PMS1 and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science*, **265**, 1091–1093.

Radman,M., Matic,I., Halliday,J.A. and Taddei,F. (1995) Editing DNA replication and recombination by mismatch repair: from bacterial genetics to mechanisms of predisposition to cancer in humans. *Phil. Trans. R. Soc. Lond. Ser. B Biol. Sci.*, **347**, 97–103.

Reenan,R.A. and Kolodner,R.D. (1992) Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics*, **132**, 963–973.

Richards,B., Zhang,H., Phear,G. and Meuth,M. (1997) Conditional mutator phenotypes in hMSH2-deficient tumor cell lines. *Science*, **277**, 1523–1526.

Strand,M., Prolla,T.A., Liskay,R.M. and Petes,T.D. (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair [published erratum appears in *Nature* (1994), **368**, (6471):569]. *Nature*, **365**, 274–276.

Strand,M., Earley,M.C., Crouse,G.F. and Petes,T.D. (1995) Mutations in the MSH3 gene preferentially lead to deletions within tracts of simple repetitive DNA in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **92**, 10418–10421.

Studamire,B., Quach,T. and Alani,E. (1998) The *S.cerevisiae* Msh2p and Msh6p ATPase activities are coordinated during mismatch repair. *Mol. Cell. Biol.*, in press.

Su,S.S. and Modrich,P. (1986) *Escherichia coli mutS*-encoded protein binds to mismatched DNA base pairs. *Proc. Natl Acad. Sci. USA*, **83**, 5057–5061.

Su,S.S., Lahue,R.S., Au,K.G. and Modrich,P. (1988) Mispair specificity of methyl-directed DNA mismatch correction *in vitro* [published erratum appears in *J. Biol. Chem*. (1988), **263**, (22):11015]. *J. Biol. Chem.*, **263**, 6829–6835.

Su,S.S., Grilley,M., Thresher,R., Griffith,J. and Modrich,P. (1989) Gap formation is associated with methyl-directed mismatch correction under conditions of restricted DNA synthesis. *Genome*, **31**, 104–111.

Thomas,D.C., Roberts,J.D. and Kunkel,T.A. (1991) Heteroduplex repair in extracts of human HeLa cells. *J. Biol. Chem.*, **266**, 3744–3751.

Tishkoff,D.X., Boerger,A.L., Bertrand,P., Filosi,N., Gaida,G.M., Kane,M.F. and Kolodner,R.D. (1997) Identification and characterization of *Saccharomyces cerevisiae EXO1*, a gene encoding an exonuclease that interacts with MSH2. *Proc. Natl Acad. Sci. USA*, **94**, 7487–7492.

Umar,A., Buermeyer,A.B., Simon,J.A., Thomas,D.C., Clark,A.B., Liskay,R.M. and Kunkel,T.A. (1996) Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell*, **87**, 65–73.

Umezu,K., Sugawara,N., Chen,C., Haber,J.E. and Kolodner,R.D. (1998) Genetic analysis of yeast RPA1 reveals its multiple functions in DNA metabolism. *Genetics*, **148**, 989–1005.

Varlet,I., Canard,B., Brooks,P., Cerovic,G. and Radman,M. (1996) Mismatch repair in *Xenopus* egg extracts: DNA strand breaks act as signals rather than excision points. *Proc. Natl Acad. Sci. USA*, **93**, 10156–10161.

Veigl,M.L. *et al*. (1998) Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc. Natl Acad. Sci. USA*, **95**, 8698–8702.

Williamson,M.S., Game,J.C. and Fogel,S. (1985) Meiotic gene conversion mutants in *Saccharomyces cerevisiae*. I. Isolation and characterization of *pms1-1* and *pms1-2*. *Genetics*, **110**, 609–646.

Yamaguchi,M., Dao,V. and Modrich,P. (1998) MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J. Biol. Chem.*, **273**, 9197–9201.

Received July 30, 1998; revised September 11, 1998; accepted September 15, 1998