

Mutations in the MutS α interaction interface of MLH1 can abolish DNA mismatch repair

Guido Plotz^{1,*}, Christoph Welsch^{1,2}, Luis Giron-Monzon³, Peter Friedhoff³,
Mario Albrecht², Albrecht Piiper¹, Ricardo M. Biondi¹, Thomas Lengauer², Stefan Zeuzem¹
and Jochen Raedle¹

¹Klinik für Innere Medizin II, Gebäude 41, Kirrberger Straße, Universität des Saarlandes, D-66421 Homburg/Saar, Germany, ²Max Planck Institut für Informatik, Stuhlsatzenhausweg 85, D-66123 Saarbrücken, Germany and ³Institut für Biochemie (FB 08), Justus-Liebig-Universität Giessen, D-35392 Giessen, Germany

Received June 29, 2006; Revised October 10, 2006; Accepted October 20, 2006

ABSTRACT

MutL α , a heterodimer of MLH1 and PMS2, plays a central role in human DNA mismatch repair. It interacts ATP-dependently with the mismatch detector MutS α and assembles and controls further repair enzymes. We tested if the interaction of MutL α with DNA-bound MutS α is impaired by cancer-associated mutations in MLH1, and identified one mutation (Ala128Pro) which abolished interaction as well as mismatch repair activity. Further examinations revealed three more residues whose mutation interfered with interaction. Homology modelling of MLH1 showed that all residues clustered in a small accessible surface patch, suggesting that the major interaction interface of MutL α for MutS α is located on the edge of an extensive β -sheet that backs the MLH1 ATP binding pocket. Bioinformatic analysis confirmed that this patch corresponds to a conserved potential protein–protein interaction interface which is present in both human MLH1 and its *E.coli* homologue MutL. MutL could be site-specifically crosslinked to MutS from this patch, confirming that the bacterial MutL–MutS complex is established by the corresponding interface in MutL. This is the first study that identifies the conserved major MutL α –MutS α interaction interface in MLH1 and demonstrates that mutations in this interface can affect interaction and mismatch repair, and thereby can also contribute to cancer development.

INTRODUCTION

The activity of the mismatch repair system elevates replication fidelity by several hundredfold through the removal of a wide variety of polymerase errors, including insertion–deletion loops

that can form during the replication of repetitive sequences (1–3). The system has been conserved throughout evolution. In humans, germline mutations in mismatch repair genes, predominantly *MLH1* and *MSH2*, underlie the Lynch syndrome (also called *hereditary non-polyposis colorectal cancer*, HNPCC), a hereditary cancer predisposition which accounts for 3–5% of all colorectal cancer cases (3–5).

Mismatch repair in humans is initiated by one of two MutS heterodimers, either MutS α (MSH2–MSH6) or MutS β (MSH2–MSH3), depending on the type of mismatch to be repaired (6–8). After mismatch binding by this heterodimer, a MutL heterodimer is recruited. This is predominantly MutL α (MLH1–PMS2), although a contribution of MutL γ (MLH1–MLH3) has also recently been reported for a subset of replication errors (9–11). The human MutS α/β and MutL α/γ heterodimers have evolved from the homodimeric bacterial predecessors MutS and MutL. The heterodimeric subunits of the human proteins share a common architecture, but diverge in structural and functional details.

Together with a MutS protein, the MutL protein directs the exonucleolytic degradation of a stretch of the error-containing strand including the mismatched base (12–17). Repair is completed by DNA re-synthesis on the emerging gap. The overall mechanism of mismatch repair appears to be similar in all organisms except for identification of the faulty DNA strand (that has to be repaired), which is performed in *E.coli* and some other bacteria by MutH, an endonuclease that binds in a site-directed manner to the transiently hemimethylated DNA that arises during bacterial replication (18,19). Since eukaryotes lack this transient hemimethylation, other ways of strand discrimination are possible [reviewed in (1,2)].

While the role of MutS proteins as mismatch-detectors is well established, the contribution of MutL proteins to repair has remained more elusive. Recently, Modrich and co-workers have demonstrated an endonucleolytic activity of human MutL α residing in the C-terminal domain of PMS2 (20). Functionally, MutL proteins have been shown to confer termination of the exonucleolytic degradation of the faulty

*To whom correspondence should be addressed. Tel: +49 6841 16 23253; Fax: +49 6841 16 23570; Email: guido.plotz@uniklinik-saarland.de

strand after removal of the mismatched base(s) (14,21). One of their most striking features is that they interact with a wide variety of other proteins, including the endonuclease MutH, the DNA clamp β and DNA helicase II (UvrD) in bacterial systems, and the DNA clamp PCNA, topoisomerase II, and exonuclease I as well as several factors involved in DNA damage response in higher organisms [for review, see (1,2)]. They are therefore thought to act as matchmakers that assemble other enzymes to the mismatched site to accomplish repair and initiate DNA damage signalling. The most important protein interaction partners for MutL proteins, however, are the MutS proteins, since these two factors represent the core of the repair machinery.

The N-terminal domains (NTD) of MutL proteins contain an ATPase of the GHKL class (22,23), while the C-terminus confers dimerization (24,25) and contains in the PMS2 protein the metal binding site essential for endonucleolytic function (20). The C-terminal dimerization is constitutive, but a second dimerization interface in the NTD of *E. coli* MutL has been shown to confer an ATP-dependent, reversible dimerization (26). This transient dimerization is required for ATP hydrolysis and represents a common theme among GHKL-ATPases (22). The resulting ATPase cycle, which includes ATP binding, transient N-terminal dimerization, hydrolysis, subsequent separation of the N-termini and release of ADP, has been suggested to be a switch in the repair process (26), although its function is unknown. MutL has been found to bind DNA, and an association of DNA binding to the activity of the ATPase has been documented (27,28). The ATPase, whose functionality is vital for repair activity in bacterial and human MutL (29,30), likely controls binding (and activation) of the downstream repair factors. MutL interacts with in dependence of the progression of repair.

The protein complex of MutS and MutL initiates and controls the mismatch repair reaction. Its detailed characterization is therefore essential for understanding mismatch repair. The conditions required for formation of complexes of MutL and MutS proteins have been investigated extensively (28,29,31,32). Their characterization is complicated by the transient and dynamic nature of the complex. We have previously shown that the N-terminus (residues 1–505) of the MutL α subunit MLH1 is required and sufficient for interaction of human MutL α and MutS α (31). Based on the hypothesis that loss of MutL α –MutS α interaction may interfere with DNA mismatch repair, we screened a set of cancer-associated missense mutations in MLH1 for their effect on interaction. We here describe the identification of a surface cluster of residues whose mutation disrupts MutL α –MutS α interaction and affects mismatch repair activity, suggesting a mechanism by which hereditary mutations in this region can produce a cancer predisposition.

MATERIALS AND METHODS

Strains, cell lines, plasmids, enzymes and reagents

Poly [d(I*C)] was purchased from Boehringer Mannheim (Mannheim, Germany), ATP, RNase A and Proteinase K were from Sigma-Aldrich (St Louis, MO, USA). Restriction enzymes N.BstNBI, N.AlwI and AseI as well as T4 DNA

ligase were from New England Biolabs (Ipswich, MA, USA). Anti-hMLH1 (G168–728) was from Pharmingen (San Diego, CA, USA), anti-hPMS2 (B-3) was from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Anti-hMSH2 (M34520) and anti-hMSH6 (G70220) were purchased from Transduction Laboratories (Lexington, KY, USA). HEK293T cells were grown in DMEM nut mix F-12 (HAM) with 10% FCS. HCT-116 cells were kindly provided by Dr C. Richard Boland (University of California, CA, USA) and grown in DMEM with 10% FCS. Oligonucleotides were purchased from BioSpring (Frankfurt, Germany) and from Qiagen (Valencia, CA, USA). Streptavidin-coupled M-280 beads were from Dynal (Oslo, Norway).

E. coli K12 strains CC106 (P90C [*ara* Δ [*lac-pro*]XIII (*F'**lacI*Z *proB*⁺) (33), TX2652 (CC106 *mutL*:: Ω 4 (BsaAI; Kan^r) and TX2928 (CC106 *mutH471*::tn5;Kan^r) (34) and the pET-15b (Novagen) derived plasmids pTX412 and pTX418 containing the *mutS* and *mutL* genes, respectively, under control of the T7 promoter were kindly provided by Dr M. Winkler (34). Plasmid pMQ402 (His6-MutH), a pBAD18 derivative, was a kind gift from Dr M. Marinus (35). For protein expression of MutS and MutL, the *E. coli* strain HMS174(λ DE3) (Novagen, Darmstadt, Germany) was used. For MutH, the *E. coli* strain XL1 blue (Stratagene, La Jolla, CA) was used. *Pfu* DNA polymerase was expressed and purified as described (36).

The pcDNA3 expression vector (Invitrogen, Carlsbad, CA, USA) containing the entire open reading frame of human MLH1 was a gift of Dr Hong Zhang (Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA). The pSG5 expression vector (Stratagene, La Jolla, CA, USA) containing full-length human PMS2 cDNA was provided by Dr Bert Vogelstein (Johns Hopkins Oncology Center, Baltimore, MD, USA). Amino-acid positions in MLH1 refer to the 756 aminoacid MLH1 sequence (NCBI accession no. AAC50285). All mutant MLH1 constructs used in this study were generated using the QuickChange Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions with suitable oligonucleotides. The deletion mutant of MLH1 which lacks the N-terminal helix A' (aa 2-25) was also generated by site-directed mutagenesis using deletion-primers that spanned the neighbouring sequence (5'-ggaaatcgagctcatatgcagcggccagctaatgc-3' and its reverse complement). Before use, all mutant constructs were controlled by direct sequencing (BigDye v1.1 chemistry and ABI 3100 sequencer, Applied Biosystems, Weiterstadt, Germany).

Transfection and extract preparation

HEK293T cells were transfected using calcium phosphate precipitation according to standard procedures (37). Cells were harvested after 24 h and whole cell extracts were prepared. Cells were washed with PBS and re-suspended in 2 times the packed cell volume of hypotonic buffer (10 mM HEPES (pH 7.9), 5 mM MgCl₂, 10 mM NaCl, 10 mM NaF, 0.1 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT). The suspension was frozen at –80°C and thawed on ice for lysis. This suspension containing both nuclei and cytoplasmic extract was supplemented with an identical volume of hypertonic buffer [10 mM HEPES (pH 7.9), 5 mM

MgCl₂, 830 mM NaCl, 10 mM NaF, 0.1 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT, 34% glycerol]. The suspension was rocked on ice for 30 min and then centrifuged (4°C, 21 000 g). The supernatant (whole cell extract) was stored in aliquots at -80°C. Expression of the desired protein was verified by SDS-PAGE and immunoblotting in comparison to extracts of untransfected HEK293T cells (negative control) and extracts of 293 or TK6 cells (expressing wild-type levels of MMR-proteins) as positive control. Nuclear extracts for the interaction assay were prepared from HCT-116 cells as described previously (28).

Nuclear extracts for mismatch repair assays were prepared using a similar procedure. In short, cells were re-suspended in three times their packed cell volume in ice-cold hypotonic buffer (20 mM HEPES pH 7.5, 5 mM KCl, 0.5 mM MgCl₂, 0.5 mM DTT, 0.2 mM PMSF) and lysed with Dounce pestle B until lysis was sufficient. After centrifugation (2000 g 4°C 5 min), the cytoplasmic supernatant was removed and again centrifuged (16 000 g 4°C 5 min) for removal of residual supernatant. The pellet was re-suspended in re-suspension-buffer (20 mM HEPES-KOH pH 7.5, 10% sucrose, 1 mM DTT, 0.2 mM PMSF) with Complete[®] protease inhibitors and high-salt buffer (50 mM HEPES-KOH pH 7.5, 10% sucrose, 840 mM KCl) was added under agitation. Extraction was performed for 30 min and extracted nuclei were removed by centrifugation (21 000 g, 4°C, 30 min). The supernatant was dialysed 6 h against 100 times the volume of buffer (25 mM HEPES pH 7.6, 100 mM KCl, 0.1 mM EDTA, 0.5 mM DTT, 0.2 mM PMSF). The extract was centrifuged (21 000 g, 4°C, 10 min), snap-frozen in liquid nitrogen and stored in aliquots at -80°C.

MutS α -MutL α -interaction

The MutS α -MutL α -interaction was assessed essentially as described previously (31). All steps were performed on ice. Cell extract (150 μ g protein, consisting of 145 μ g HCT-116 extract and 5 μ g extract of HEK293T cells containing the mutated MutL α protein) was incubated in 20 mM Tris-HCl, pH 7.9, 50 mM NaCl, 1.5 mM MgCl₂, 5% glycerol, 1 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT and 1 μ g poly [d(I**C*)] in a total volume of 300 μ l for 5 min on ice. For each experiment, two identical samples from one master mix were prepared. Both samples were added to an aliquot of streptavidine beads coupled with 200 bp homoduplex DNA substrate as detailed in (31). After 20 min incubation, ATP was added to one sample to a final concentration of 250 μ M. After 5 min, the beads were collected with a magnet and the supernatant was taken off, the cup centrifuged and residual supernatant removed. The beads were re-suspended in 20 μ l elution buffer (700 mM NaCl, 20 mM Tris-HCl, pH 7.9, 1.5 mM MgCl₂, 5% glycerol, 1 mM EDTA, 0.2 mM PMSF and 0.5 mM DTT) and incubated for 5 min. The beads were collected and the supernatant stored for analysis. A second elution step was performed with 20 μ l elution buffer as above, only with 1000 mM NaCl. This elution was also analysed by SDS-PAGE and immunoblotting and verified that ATP has taken effect, since it abolishes hMutS α signals in this elution fraction. Although this elution has always been performed as a control, the data has been omitted

in the figures. Furthermore, identical concentration of the MutL α heterodimer has always been checked by an immunoblot of the incubation mixture.

Mismatch plasmid construction

The pUC19CPDC plasmid used for construction of the mismatched substrate were kindly provided by Dr John B. Hays and propagated in *E.coli* strain SCS110 (*dcm*⁻, *dam*⁻, *endA*⁻) from Stratagene. The plasmid substrate for the mismatch repair assay was synthesized following the published procedure (38-41) with some minor modifications. In short, pUC19CPDC (50 μ g) was nicked at two sites with 32 bp distance with N.BstNBI (usually 30 U) for 3 h. Quantitative nicking was assured by running aliquots of the reaction on an agarose gel, and digestion was continued afterwards with additional enzyme for one more hour. The digested single-stranded oligomer was released and captured by denaturing at 85°C for 5 min in the presence of 50-fold excess antisense-oligomer WHCPDPuriAS (5'-GCGGATATTAATGTGACG-GTAGCGAGTCGCTC-3') and subsequent slow cooling to room temperature. Oligomers were removed by centrifugation through Microcon-100 columns and extensive washing with TE buffer. Gapped plasmid was ligated with a 10-fold molar excess of WHpCPD7 (5'-pGCGGATATTAATGTGACG-GTAGCGAGTCGCTC-3'). Ligation was carried out overnight with T4 DNA ligase (50 U) at 16°C. Ligated product was ethanol precipitated and subsequently treated with N.AlwI (50-100 U) until quantitatively nicked as judged by electrophoresis of small aliquots. Enzyme was heat inactivated (80°C 20 min) and the nicked mismatch plasmid was purified by centrifugation through Microcon-100 columns. The resulting plasmid (named pUC19CPDC-GTN) contains one AseI restriction site, a GT mismatch within an overlapping AseI/EcoRV restriction site as well as a nick in the DNA 141 bp from the mismatched site, which serves to direct mismatch repair to convert the GT mismatch to AT, thereby restoring the AseI restriction site. Digestion of this preparation with AseI yielded only linearized vector, but no detectable fragments. Digestion with AseI and EcoRV showed that ~10% of the preparation was cleavable with EcoRV, confirming that the majority of the original pUC19CPDC was transformed to mismatched substrate. Since repair of pUC19CPDC-GTN is assessed by restoration of the second AseI restriction site, the residual EcoRV-cleavable plasmid did not interfere with the mismatch repair assay. The presence of the mismatch and the nick were additionally verified by direct sequencing of pUC19CPDC-GTN.

MMR assay

Mismatch repair reactions were essentially performed as described elsewhere (40,42). In short, the reaction was performed in 15 μ l total volume with reaction buffer (20 mM Tris-HCl pH 7.6, 110 mM KCl, 5 mM MgCl₂, 1 mM glutathione, 50 μ g/ml BSA, 1.5 mM ATP, 0.1 mM each dNTP), 100 ng DNA substrate and 50 μ g nuclear extract of HEK293T cells, which are deficient in mismatch repair (43). Where indicated, reactions were supplemented with 5 μ g extract from HEK293T cells expressing recombinant hMutL constructs. Reactions were incubated at 37°C for 20 min and terminated with 50 μ l stop-buffer (24 mM EDTA

pH 8.0, 0.7% SDS, 90 µg/ml proteinase K) by an additional incubation for 15 min at 37°C. Plasmids were extracted from the reaction mixture by phenol extraction and purified by ethanol co-precipitation with tRNA. Subsequent digestion with AseI and RNase A produced two smaller fragments besides the linearized vector when repair was successful. Restriction digests were separated on 2% agarose gels, stained with ethidium bromide and documented by UV-transillumination and Polaroid photography. Levels of mismatch repair were quantified using Quantity One Software v4.6.1 (Bio-Rad, Hercules, CA, USA) by dividing the volume of the two repair bands by the total volume of all three bands. Relative repair efficiency was calculated by dividing the value of the mutant through the value of a wild-type protein preparation that had been expressed, processed and tested in parallel.

DNA substrates for photocrosslinking

Linear heteroduplex substrates were generated by annealing two 484 bp PCR products amplified by *Taq*-DNA polymerase with a single GATC site at position 210 and a G/C or an A/T bp at position 385 using plasmids and primers as described previously (44). This procedure results in a mixture of 50% homoduplex substrates (G/C and A/T) and 50% heteroduplex substrates (G/T and A/C). In general 40–60% of this DNA was cleavable by MutH in a mismatch and MutS dependent manner.

Purification of MutL α constructs

Details of the purification of MutL α and its characterization are to be published in a separate manuscript. In short, the cDNA of PMS2 was cloned from pSG5 into pEBG-2T via the BamHI restriction site. This placed PMS2 downstream of the coding sequence for Glutathion-S-Transferase (GST). HEK293T cells were co-transfected with MLH1 (wildtype or mutant) and pEBG-2T PMS2. Cells were lysed after two days in lysis buffer (50 mM Tris-HCl pH 7.4, 0.27 M sucrose, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.1% β -mercaptoethanol, 0.2 mM PMSF), centrifuged and the supernatant incubated with 0.3 ml glutathion Sepharose for 1 h at 4°C. The suspension was applied to a column, allowed to settle and drained by gravity flow. The Sepharose was washed with 5 ml of lysis buffer and 10 ml of washing buffer (50 mM Tris-HCl pH 7.5, 1 mM EGTA, 0.1% β -mercaptoethanol). Bound protein was eluted with washing buffer supplemented with 40 mM reduced glutathion and 0.27 M sucrose. This procedure yielded 50–70% pure MutL α .

ATPase assays

ATPase activity of partially purified GST-MutL α was assessed essentially as published before (29). In short, GST-MutL α (500 ng) was incubated in a 15 µl reaction volume containing 20 mM Tris-HCl pH 7.5, 120 mM NaCl, 5 mM MgCl₂, 100 µg/ml BSA, 1 mM DTT, 400 µM ATP, and 0.2 µCi adenosine 5'-[γ -³²P]triphosphate. After incubation at 37°C for 30 min, reactions were terminated by addition of 1.5 µl 10 M formic acid. Aliquots of 5 µl were loaded onto Polygram CEL300 PEI TLC plates, developed in 0.5 M lithium chloride/1 M formic acid and quantified using a Typhoon Imager and ImageQuant 5.1 software.

ATPase activity was calculated as µM of released phosphate per pmol pure GST-MutL α per h, and values from parallel mock experiments (without protein) were subtracted to eliminate background.

Purification of bacterial mismatch repair proteins

Recombinant His₆-tagged MutH, MutH^{C96S} (45), MutL, single-cysteine MutL^{N131C} [SC-MutL^{N131C}, (46)] and MutS proteins were purified by Ni-NTA chromatography essentially as described (34,45). When necessary, proteins were purified by gel filtration on a Superdex200 column (Pharmacia). MutH proteins were stored at –20°C in 10 mM HEPES-KOH, 500 mM KCl, 1 mM EDTA, 1 mM DTT, 50% glycerol, pH 7.9. MutL and MutS proteins were snap-frozen in liquid nitrogen and stored at –70°C in 10 mM HEPES-KOH, 200 mM KCl, 1 mM EDTA, pH 7.9. Protein concentrations were determined using theoretical extinction coefficients (47).

Crosslinking studies

Crosslinking was performed either with the thiol-specific homobifunctional reagent BM[PEO]₄ or the heterobifunctional reagent 4-maleimidobenzophenone (MBP), which contains one thiol-specific activity and can furthermore be photocrosslinked non-specifically to nearby residues. The concentration of crosslinker was titrated to achieve optimal conditions for specific crosslinking. For crosslinking with BM[PEO]₄, MutS (400 nM), SC-MutL^{N131C} (1000 nM) and heteroduplex DNA (484 bp, mixture of G/T and A/C mismatch; 100 nM) were incubated in the presence of the indicated nucleotide (1 mM) in 10 mM HEPES-KOH, pH 7.5, 5 mM MgCl₂, 125 mM KCl. BM[PEO]₄ was added to a final concentration of 50 µM and the reaction was allowed to proceed for 1 min at 37°C. Crosslinking was quenched by the addition of 50 mM DTT and samples were separated by SDS-PAGE (10% gel) followed by Coomassie staining PageBlue from Fermentas (Hanover, MD, USA).

For photocrosslinking, MutL^{N131C} (10 µM) was incubated with or without MutH^{C96S} (2.5 µM) with 500 µM 4-maleimidobenzophenone (MBP, Sigma) in 10 mM Tris, 5 mM MgCl₂, pH 7.9 in the presence of 4 mM nucleotide, ATP or AMP-PNP, for 30 min at room temperature. Reaction was stopped by adding a 5-fold molar excess of DTT over thiol-specific reagent. For photocrosslinking 2 µM modified MutL, 500 nM MutH, 1 µM MutS^{wt}, 0.8 mM nucleotide, 25 nM 484 bp mismatch DNA, in 10 mM Tris, pH 7.9, 5 mM MgCl₂, 125 mM KCl were irradiated for 25 min at 354 nm with a handheld UV lamp at a distance of 5 cm. The samples were separated by SDS-PAGE and detected by Coomassie staining.

Tryptic digests and mass spectrometry for identification of crosslinked products were performed as described before (46).

Structural analysis

Protein sequence were retrieved from the NCBI databases (48) using BLAST (49). An alignment of sequences homologous to human MLH1 was computed by ClustalW (50) and improved by minor manual modifications in SeaView (51). A homology-derived 3D structure model of human

MLH1 was constructed using the WHAT IF modelling server (52) with a sequence–structure alignment extracted from the multiple sequence alignment of MHL1 homologues and based on the PDB structure 1b62 (chain A) of *E.coli* MutL complexed with ADP-Mg²⁺ as template (26,53) [*E*-value 10⁻⁵³ and sequence identity 35%; see Supplementary Figure A for the sequence–structure alignment, which was prepared in GeneDoc (54)]. The secondary structure assignment of 1b62 was taken from the DSSP database (55).

The ProMate server (56) was used for calculation of the probability of the residues to contribute to a transient heterodimeric protein interaction of either the MLH1 homology model or the structure of *E.coli* MutL (PDB 1b63). The analysis was performed with standard configuration except that ‘Evolutionary conserved position’ and ‘Water molecules’ were disabled.

For surface-mapping of phylogenetic conservation information, we performed a BLASTP search with the sequences of MLH1_{1–335} and *E.coli* MutL_{1–329}. The 250 best hits were extracted and aligned with ClustalW. We manually extracted 84 sequences from the MLH1 alignment and 79 sequences from the MutL alignment showing highest similarities within the region of β -strands 3–5 to exclude more distant members of these protein families. The extracted sequences were analyzed by ConSurf 3.0 (57).

We used the PyMOL Molecular Graphics System (2002) (DeLano Scientific, San Carlos, CA, USA, <http://www.pymol.org>) to create protein structure images and applied the Color_B script to project the results of the ConSurf and ProMate analyses onto the protein surfaces. The BioEdit sequence alignment editor v7.0.1 (www.mbio.ncsu.edu/BioEdit/bioedit.html) was used for drawing of alignments.

Protein separation and detection

The proteins were separated on 10% polyacrylamide gels, followed by blotting on nitrocellulose membranes and antibody detection using standard procedures.

Accession numbers

The NCBI accession numbers of the sequences in Figure 2B are: CAA77850 (*E.coli*), MLH1: AAH06850 (human), NP_081086 (mouse), NP_112315 (rat), AAC19117 (fruitfly), XP_320342 (anopheles), XP_329015 (*Neurospora crassa*), NP_499796 (worm), P38920 (baker’s yeast), AAK25988 (thale cress), NP_596199 (fission yeast); PMS: P54278 (hPMS2), XP_213712 (rat), NP_032912 (mouse), AAL39978 (fruitfly), XP_308635 (anopheles), XP_328726 (*Neurospora crassa*), NP_505933 (worm), AAM00563 (baker’s yeast), BAC42424 (thale cress), CAB10113 (fission yeast).

RESULTS

The cancer mutation MLH1 Ala128Pro abolishes MutL α –MutS α interaction and mismatch repair

For a closer characterization of the interaction of human MutL α (MLH1–PMS2) with MutS α , we tested whether this interaction is affected by mutations observed in the hereditary Lynch cancer syndrome, which is most frequently caused by

germline mutations in MLH1 (58,59). We have previously shown that the interaction of MutL α with MutS α is predominantly conferred by MLH1, since the MLH1 subunit, but not the PMS2 subunit, interacts efficiently (31). This study also showed that an MLH1 fragment containing the N-terminal domains (NTD; MLH1_{1–505}) was required and sufficient for interaction. We therefore selected a set of Lynch syndrome mutations identified in the NTD of MLH1: Asn64Ser, Cys77Tyr, Thr117Met, Ala128Pro and Arg265Cys (Figure 1A). In addition, we included Asp132His, a mutation which has recently been associated with low-penetrance hereditary colorectal cancer (60). Vectors containing the mutant MLH1 cDNAs were generated and transfected into HEK293T cells either alone or co-transfected with PMS2. The MLH1 gene is transcriptionally silenced in HEK293T cells, which causes a concomitant loss of PMS2, whose stability depends on MLH1 [(31,43,61); Figure 1B, WB, lane 1]. This allows expressing and analysing MutL α variants in this cell line without interference of endogenous MutL α . Immunoblotting of cell extracts confirmed that mutant and wild-type MLH1 proteins were expressed equally well (Figure 1B, WB, lanes 2–7) except for the MLH1^{C77T} variant, which showed attenuated expression suggestive of impaired protein stability (data not shown) and was therefore excluded from further analysis. The expression levels of PMS2 were also similar (Figure 1B, WB), which is consistent with our previous observation that the C-terminal domains of MLH1 suffice for stabilizing PMS2 (31). Accordingly, mutations in the N-terminus of MLH1 are unlikely to affect formation of a stable MutL α heterodimer, providing that the mutant MLH1 is stably expressed.

We tested the mismatch repair activity of all mutant proteins by assessing mismatch repair-mediated reconstitution of a plasmid restriction site containing a GT mismatch. All mutations observed in Lynch syndrome either abolished (MLH1^{T117M} and MLH1^{A128P}) or reduced mismatch repair efficiency (MLH1^{N64S} and MLH1^{R265C}) in comparison to wild-type MLH1, confirming that they compromise the functionality of MLH1 (Figure 1B, MMR). In contrast, the repair efficacy of the variant MLH1^{D132H}, which has been identified in low-penetrance hereditary cancer families, was indistinguishable from wildtype.

In order to determine whether the mutations affect the complex formation of MutL α with MutS α , we assessed binding of the mutant proteins (either MLH1 or MutL α) to MutS α –DNA in a DNA pull-down assay we have previously used for investigation of this protein complex (28,31,62). In this assay, DNA beads are incubated with nuclear protein extract containing the recombinant MutL α protein. Elution of DNA-bound protein and subsequent western analysis allows assessing binding of proteins to the DNA substrate. Under the conditions applied, MutS α binds to the DNA substrate, and MLH1/MutL α binds to MutS α –DNA to form a ternary complex, which depends on MutS α (28). In this assay, addition of ATP triggers a characteristic, MutS α -dependent increase in binding of MutL α indicating the formation of the MutL α –MutS α –DNA complex (28). We have shown that the experimental conditions applied do not cause non-specific DNA end-binding of the proteins (62), which has been suggested to impair experimental data on interaction. Since binding of MLH1/MutL α is

MutS α -dependent (28) and we did not detect any MutS β [(28) and unpublished observations], MutS α is the predominant factor in the MutS–MutL α complex observed under these experimental conditions.

The interaction of three Lynch syndrome mutants (MLH1^{N64S}, MLH1^{T117M} and MLH1^{R265C}) and of MLH1^{D132H} with MutS α –DNA was indistinguishable from wildtype (data not shown). In contrast, the MLH1^{A128P} mutation significantly impaired interaction with MutS α –DNA (Figure 1C): although identical protein levels of wild-type and mutant MLH1 were incubated (Figure 1C, left panel, ‘Loaded’), binding of the mutant MLH1^{A128P} to MutS α –DNA was suppressed below western blot detection limit, while wild-type MLH1 interacted efficiently (Figure 1C, right panel, ‘Eluted from DNA’). The corresponding experiment with the mutant heterodimer MutL α ^{MLH1 A128P} yielded very weak signals, whose quantification showed that its interaction was decreased almost fivefold in comparison to wild-type MutL α . The finding that the interaction of the mutant MLH1 subunit was more suppressed than that of the mutant

MutL α heterodimer is in agreement with our previous observation that PMS2 can very weakly confer binding to MutS α –DNA (31). However, PMS2 was unable to compensate the loss of interaction conferred by the MLH1 mutation.

Various reasons can account for the loss of a protein interaction due to the mutation of a single residue, including loss of optimal *van der Waals* contacts and electrostatic interactions, but also more complex phenomena like conformational changes within the interaction interface (63). Since the introduction of the proline may cause structural distortions, we analyzed this region of MLH1 more closely to confirm its involvement in interaction with MutS α .

Identification of additional residues important in MutL α –MutS α interaction

Ala128 is located within the conserved N-terminal ATPase domains of MLH1 (Figure 2A). For selecting further residues for testing, we generated a local sequence alignment of MLH1 and PMS2 proteins around MLH1 Ala128 to search for patterns of conservation (Figure 2B), since protein–protein interaction interfaces are more conserved than non-functional protein surfaces (64,65). The sequence alignment shows two highly conserved motifs of the GHKL–ATPase (motifs G2 and G3) which frame a less conserved stretch located between the two loops L34 and L56 (see annotations below the alignment). Ala128 is located within this less

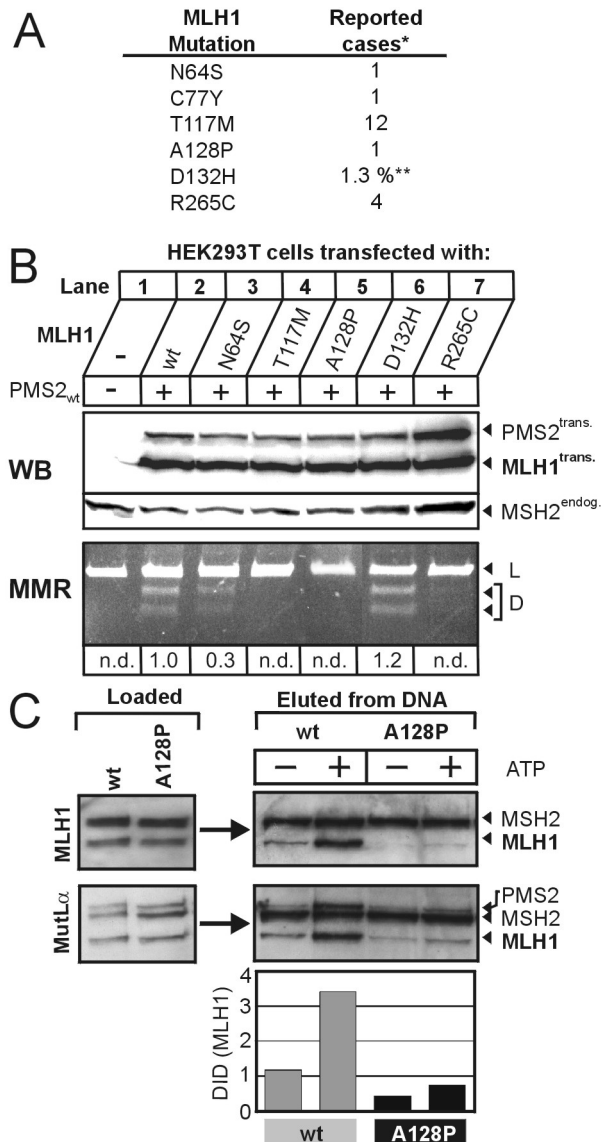


Figure 1. Analysis of cancer-associated MLH1 mutations. (A) The table lists the six MLH1 missense mutations identified in hereditary human cancer syndromes which were selected for analysis. Five of these mutations have been described in patients with the Lynch cancer syndrome, while one (D132H) has recently been reported to underlie a low-penetrance colorectal cancer syndrome (60). *The Reported cases column contains the number of individual reports listed by the International Society for Gastrointestinal Hereditary Tumours (www.insight-group.org) except for **D132H whose allele frequency has been determined in the Israeli population (60). (B) Wild-type MLH1 and the cancer-associated MLH1 mutants were co-expressed with PMS2 in HEK293T cells. Protein extracts (50 μ g) of untransfected (lane 1) and transfected (lanes 2–7) cells were separated by SDS–PAGE and protein expression was verified by immunoblotting (MLH1^{trans.} and PMS2^{trans.}, WB, top panel). Endogenous MSH2 (MSH2^{endog.}) served as loading control (WB, bottom panel). The extracts were incubated with a mismatched plasmid as detailed in Materials and Methods. The plasmid contained a restriction site blocked by a G–T mismatch. Reconstitution of this restriction site by mismatch repair yielded two plasmid fragments (D) besides linearized vector (L) after separation of the digestion products in agarose gels. Repair activity of the MutL α mutants was assessed in direct comparison to wild-type MutL α which was expressed, processed and analyzed in parallel. Repair levels were quantified relative to wildtype; n.d., no detectable repair activity. (C) A DNA pull-down assay was performed to assess interaction of MLH1_{wt} and MLH1^{A128P} with MutS α –DNA. Both MLH1 proteins were expressed either alone (top panels) or co-expressed with wild-type PMS2 to form the heterodimer MutL α (bottom panels) in HEK293T cells. Equal amounts of protein extract were incubated with DNA-coupled magnetic beads as detailed in Materials and Methods. From these incubations, equal aliquots were removed, separated by SDS–PAGE and checked by immunoblotting to contain identical concentrations of transfected MLH1 (and PMS2, if applicable) as shown in the left panel (‘Loaded’). MSH2 was detected in parallel to serve as loading control. Each incubation was divided in two samples, one of which (‘+’) was supplemented with ATP (250 μ M final concentration). Supernatant was removed and bound proteins were eluted, separated by SDS–PAGE and detected by immunoblotting (right panels, ‘Eluted from DNA’). MutS α (only MSH2 is shown) and MutL α (MLH1–PMS2) were detected. PMS2 is visible just above the signal of MSH2. The MLH1 protein levels eluted from DNA-beads in the MutL α experiment were densitometrically quantified (DID: Differential integrated density).

conserved stretch. To add more information on candidate residues for interaction, we generated an homology model of the NTD of MLH1 and performed ProMate and ConSurf predictions. ProMate performs structure-based predictions of protein–protein interaction sites and was derived from an extensive analysis of transient heterodimeric protein complexes (56). ConSurf calculates evolutionary sequence conservation scores and projects them onto the protein structure, thereby allowing analyzing surface conservation.

We tested ProMate on the known C-terminal dimerization domain (CTD) of MutL (24,25). Although none of the probability values ProMate calculated were high enough to pass the algorithms' significance limit, ProMate correctly assigned the highest probability values to the experimentally verified C-terminal dimer interface (Supplementary Figure B). We therefore performed the prediction for the homology model of the NTD of MLH1 (Figure 2C) and in parallel for the NTD of its *E.coli* homologue MutL, of which crystal structures are available (23,26) (Figure 2D). For both templates, ProMate identified two surface patches which showed increased probability to engage in a protein–protein contact (Figure 2C and D, left panels; marked by circles). Again, neither of these patches was above the significance limit and therefore required additional evidence (H. Neuvirth, personal communication).

The projection of the evolutionary conservation of MLH1 or MutL proteins on the surfaces revealed that the potential interaction patch close to MLH1^{A128} also exhibits increased conservation scores (Figure 2C and D, right panels; marked by circles), while no conservation is evident for the other region predicted by ProMate.

Taken together, the bioinformatic analyses suggested that MLH1 Ala128 is located closely to a surface patch showing both increased conservation and increased probability of engaging in a protein–protein interaction. For further mutational analysis we therefore chose residues from this patch, but also included several residues from its surroundings constituted by β -strands 3–5 of MLH1. We preferentially chose residues conserved exclusively in MLH1 proteins, since PMS2 contributed only weakly to the interaction of MutL α with MutS α (31). We furthermore preferred surface placed, charged residues, since most protein interactions include electrostatic interactions (66). We selected 14 residues from the top edge of the extensive β -sheet that backs the MLH1 ATPase (Table 1; Figure 2B and E). Furthermore, we tested mutations of two highly conserved residues (MLH1^{K164E} and MLH1^{R182E}) that have been reported to abolish DNA binding of bacterial MutL (67). DNA binding by MutL proteins is essential for repair (67,68) and has been suggested to be important for their interaction with MutS proteins, since this requires DNA substrates of sufficient length (28,32).

Transfection of wild-type MLH1 and all mutant constructs into HEK293T cells resulted in similar levels of expression except for MLH1^{Y126A} (data not shown), which was excluded from further analysis. Co-transfection of PMS2 again yielded similar expression levels with wild-type and mutant MLH1 constructs (data not shown).

Six mutations of four residues conferred a practically complete loss of mismatch repair (H112D/A, K118E, R127E and Y130H/A; Figure 3A). The repair efficiency of seven other mutants was reduced on average by 50%, while five

Table 1. Summarized features of MLH1 mutations screened in this study

Residue	Alteration	Localization ^a		Conservation ^b
Asn 64	Ser	L	Exp.	c
Glu 71	Lys	L	Exp.	mc
Asp 72	Lys	L	Exp.	c
His 112	Asp, Ala	β	Exp.	c (in MLH1)
<i>Thr 117</i>	<i>Met</i>	β	Bur.	c
Lys 118	Glu	β	Exp.	mc (in MLH1)
Asp 121	His	L	Exp.	
Lys 123	Glu	L	Exp.	
Tyr 126	Ala	β	Bur.	c (in MLH1; either Y or W)
Arg 127	Glu	β	Exp.	mc
<i>Ala 128</i>	<i>Pro</i>	β	Bur.	c
Tyr 130	His, Ala, Phe	β	Bur.	c (in MLH1; either Y or F)
<i>Asp 132</i>	<i>His</i>	L	Exp.	c
Gly 133	Asn	L	Exp.	c
Lys 134	Gly, Glu	L	Exp.	mc
Lys 140	Glu	β	Exp.	mc (in MLH1)
Asp 154	Lys	L	Exp.	mc (in MLH1)
Tyr 157	Lys, Ala	L	Exp.	c (in MLH1)
Lys 164	Glu		Exp.	c
Arg 182	Glu		Exp.	c
<i>Arg 265</i>	<i>Cys</i>	L	Bur.	c

The table lists MLH1 mutations which have been used in this study for screening for residues involved in the interaction of MutL α with MutS α -DNA. Mutations involved in human cancer diseases are printed in *italics*.

^aLocalization of the residue within MLH1 according to the homology model: the first column informs in what secondary structure the residue is located (β : β -sheet; L: loop), while the second column gives the exposition status (Exp.: exposed; Bur.: buried).

^bThe conservation column summarizes the evaluation of conservation of the residues according to Figure 2B (c, conserved; mc, mostly conserved).

mutations achieved close to 100% repair. Mutations that reduced mismatch repair levels affected residues with higher conservation than those that left repair efficiency largely unaffected (E71, D121, K123, K134). When more than one alteration of a residue was analyzed, the more conservative alteration affected repair less (Y130F, Y157A).

We identified three residues whose alteration interfered with the complex formation of MLH1/MutL α with MutS α -DNA: His112 (MLH1^{H112A} and MLH1^{H112D}; Figure 3B), Arg127 (MLH1^{R127E}, Figure 3C), as well as the non-conservative alterations of Tyr130 (MLH1^{Y130H} and MLH1^{Y130A}, Figure 3D–E) displayed defects in binding to MutS α -DNA. Identically to MLH1 A128P, these mutations decreased the level of bound MLH1 below western blot detection limit. The corresponding mutant MutL α proteins again retained some weak binding (also as observed with A128P), which was presumably conferred by PMS2. All other mutants, including the conservative alteration MLH1^{Y130F} (Figure 3F) interacted with MutS α -DNA identically as wildtype (data not shown).

It has previously been shown that mutations in the ATPase domain can abolish hydrolytic activity of MutL α without affecting its ability to interact with MutS α (29). However, to rule out that an alteration of the ATPase activity may have indirectly abolished interaction, we tested the ATPase activity of the interaction-deficient H112 mutants. GST–MutL α ^{wt}, GST–MutL α ^{MLH1 H112D/A} and the previously characterized catalytic residue mutant GST–MutL α ^{MLH1 E34A} were expressed and purified. Our experiments confirmed

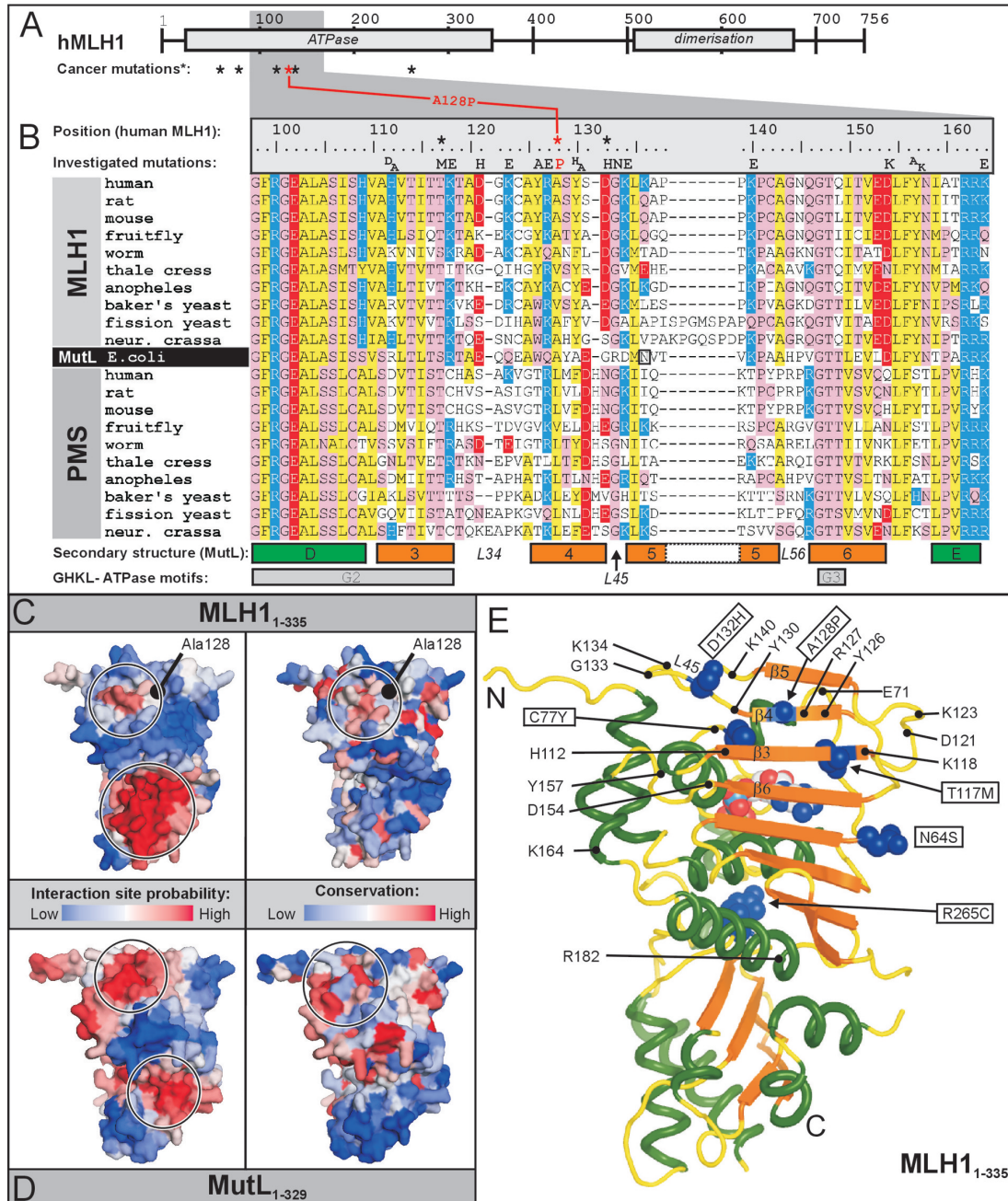


Figure 2. Analysis of the sequence and structure of hMLH1. (A) Schematic diagram of MLH1. MLH1 contain conserved N-terminal GHKL-ATPase ('ATPase') and C-terminal dimerization domains ('dimerization'). The locations of the cancer-associated mutations investigated in this work are marked by asterisks. (B) Local multiple sequence alignment surrounding A128 from the hMLH1 ATPase (hMLH1 aa 98–164). The sequences were retrieved by a Protein-BLAST search for sequences homologous to human MLH1 (top) or PMS2 (bottom) and aligned together with the *E. coli* MutL sequence (middle) using ClustalW (50) with subsequent small manual modifications. The residue numbering above the alignment refers to hMLH1, and the secondary structure information shown below the alignment refers to the crystal structure of *E. coli* MutL (23,26). The alignment comprises β -strands 3–6 and, in part, the neighbouring α -helices D and E, including conserved motifs of the GHKL-ATPases (motif III, G2-box and motif IV, G3 box (22)). Coloured shading indicates similarity or identity of amino acids (matrix: Blosum62, shading threshold 40%, calculated for all sequences. Colour coding: yellow for hydrophobic residues, blue for positively charged residues, red for negatively charged residues, pink for other residues). Mutations associated with hereditary human cancer are marked by asterisks, and mutations investigated in this study are annotated above the alignment. (C) A homology model of the N-terminal domains of MLH1 (residues 1–335) was generated (for details see Materials and Methods) and used for interaction-site prediction with ProMate (left) and projection of phylogenetic conservation with ConSurf (right) as detailed in Materials and Methods. ProMate enabled shading of the surface of the model according to the predicted probability of the residues to contribute to a transient protein-protein interaction, with red shading representing high probability and blue shading showing low probability. The ConSurf prediction is presented with red shading for conserved residues and blue shading for variable residues. Circles mark the candidate interaction sites which are discussed in the results section. (D) The same analysis as described in C was performed for the crystal structure of the NTD of *E. coli* MutL (PDB 1B63). (E) Cartoon representation of the homology model of the MLH1 NTD in the same orientation as the surface representation of C. The bound nucleotide is presented in spheres colour-coded according to the elements. Cancer-associated residues investigated in this study are presented as blue spheres with framed annotations. Residues that were selected for further mutational investigation are annotated.

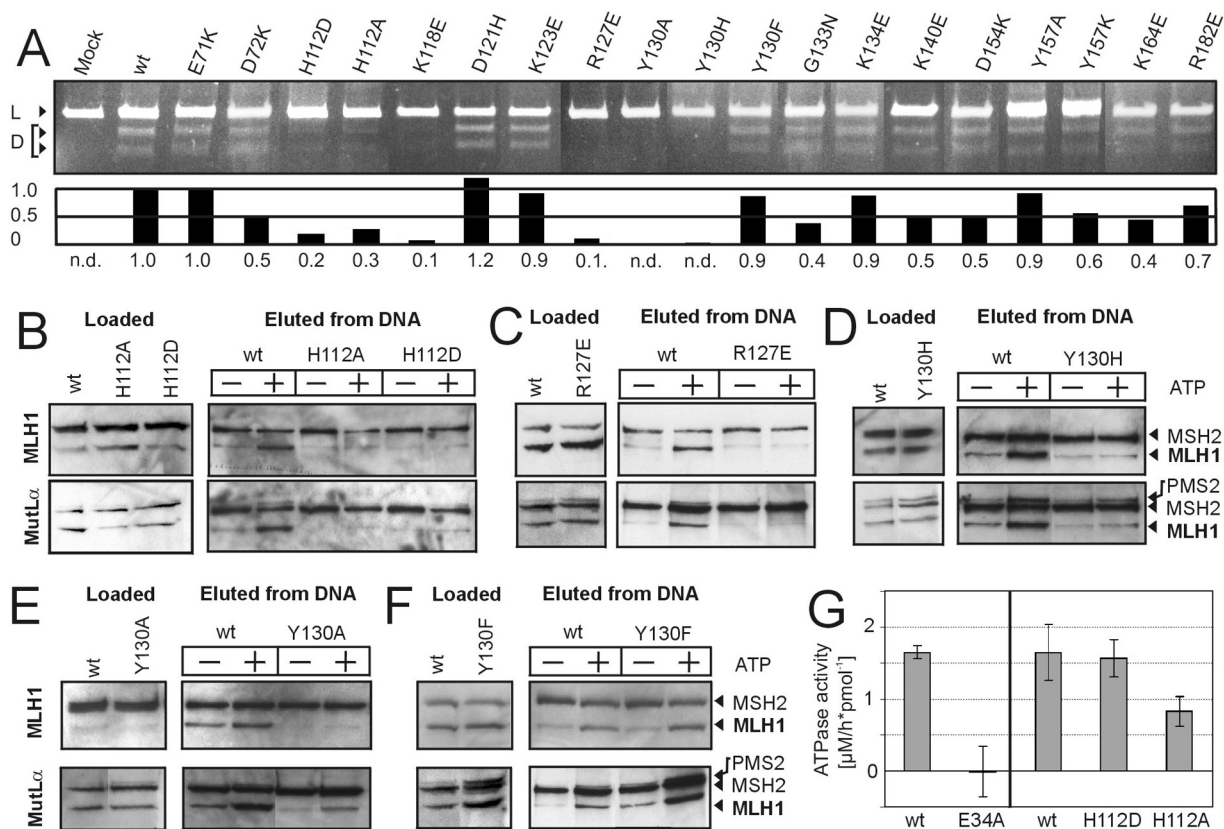


Figure 3. Analysis of MLH1 mutants for MMR activity and interaction with MutS α -DNA. (A) Nineteen MLH1 missense mutants were co-expressed with PMS2 in parallel to wild-type MutL α in HEK293T cells and *in vitro* mismatch repair activity was assessed as described in Figure 1 and Materials and Methods. AseI digestion of the mismatched test plasmid yielded linearized vector (L), and successful repair additionally produced two fragments of lower molecular weight (D, digestion fragments). Repair activity of each MutL α mutant was assessed in direct comparison to wild-type MutL α which was expressed, processed and analyzed in parallel, and repair levels were quantified in relation to these wild-type preparations. These levels (wildtype = 1.0) are shown; n.d., no detectable repair activity. (B–F) The interaction of MLH1^{H112D}, MLH1^{H112A}, MLH1^{R127E} and MLH1^{Y130H/A/F} with MutS α -DNA was assessed in comparison to MLH1_{wt} identically as described in Figure 1. (G) The ATPase activity of purified GST–MutL α mutants was assessed in comparison to wildtype as detailed in Materials and Methods. Three individual experiments were performed for each protein preparation and for a mock sample without protein. Values of the mock samples were subtracted from the protein samples. The block diagrams show the average ATPase activity, with error bars giving the standard deviations.

the previous observation that MutL α ^{MLH1 E34A} has no detectable ATPase activity (Figure 3G) while retaining interaction with MutS α (data not shown). In contrast, GST–MutL α ^{wt} and the interaction-deficient GST–MutL α ^{MLH1 H112D} displayed identical ATPase activity (Figure 3G), which was very similar to the previously determined activity of MutL α (29). This confirms that the H112D mutation did not abolish interaction indirectly by affecting the ATPase.

E. coli MutL interacts with MutS with the same protein surface

Since the bioinformatic analysis suggested that a potential protein–protein interaction site with elevated surface conservation in *E. coli* MutL is located similarly as in MLH1, we tried to site-specifically crosslink MutS to a residue in this putative interface of MutL. We have previously applied crosslinking for the mapping of the interaction sites of MutL and MutH (45,46). Here we used the *E. coli* single-cysteine variant MutL^{N131C} containing a cysteine residue located in loop L45 (N131 is marked in Figure 2B) at the edge of the potential MutS interface (46). After incubation of MutL^{N131C} with MutS under conditions suitable for

mismatch-provoked activation of the strand discrimination endonuclease MutH (46), we applied the homobifunctional reagent BM[PEO]₄. This reagent crosslinks cysteine residues which are located in proximity [3–17 Å, (69)]. After separation of the crosslinking products by SDS–PAGE, Coomassie staining revealed the formation of two high molecular weight products: one which occurred in all experiments including MutL^{N131C} and has previously been assigned as the MutL–MutL (L–L) crosslink (46) (Figure 4A). The other product formed exclusively in the presence of MutL^{N131C}, MutS, ATP and DNA with an apparent molecular weight corresponding to a crosslinked MutL^{N131C}–MutS (S–L) complex. The presence of both MutL and MutS proteins in this band was confirmed by tryptic digest and mass spectrometry (data not shown). This suggests that Asn131 of MutL is close to one of the six cysteine residues in MutS.

In addition, we tested the heterobifunctional crosslinking reagent maleimido-benzophenone (MBP) that after coupling to a cysteine residue allows photocrosslinking to suitable acceptor residues in proximity to the benzophenone moiety (70). Again, a new band with low electrophoretic mobility was only observed in reactions containing MutS, MutL^{N131C}, ATP and DNA (Figure 4B). The lower mobility of the

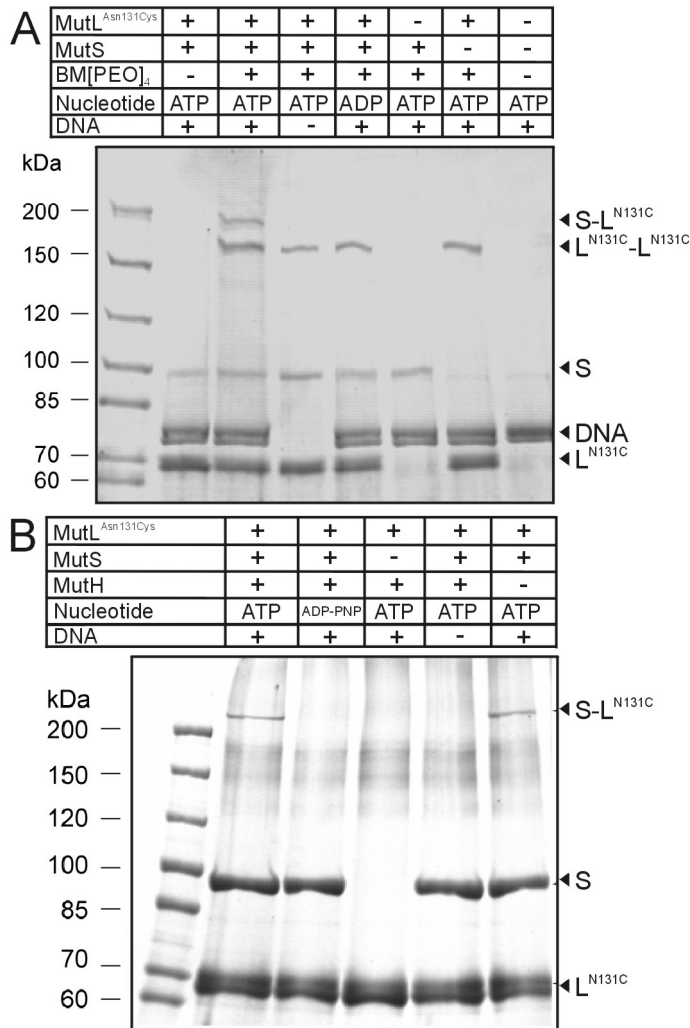


Figure 4. Crosslinking of *E. coli* MutL^{N131C} to MutS. (A) Crosslinking was performed using *E. coli* MutS (400 nM), MutL^{N131C} (1000 nM) and a 484 bp DNA substrate (100 nM) in the presence of the indicated nucleotide (1 mM) as detailed in Materials and Methods. BM[PEO]₄ was added to a final concentration of 50 μ M and the reaction was incubated for 1 min at 37°C. Crosslinking was quenched by addition of 50 mM DTT, and the products were separated by SDS-PAGE and protein bands were visualized by colloidal Coomassie staining that also stains DNA. Note that the additional band attributed to a MutS–MutL^{N131C} crosslink (S–L^{N131C}) is only observed in the presence of both ATP and DNA. (B) Photocrosslinking of MutL^{N131C} was carried out in the presence or absence of MutH^{C96S}. Both proteins were pre-incubated at 10 and 2.5 μ M, respectively, with 500 μ M MBP in the presence of 4 mM nucleotide, ATP or AMP–PNP, for 30 min at room temperature. Reactions were stopped by adding DTT. For photocrosslinking, 2 μ M modified MutL, 500 nM MutH, 1 μ M MutS_{wt}, 0.8 mM nucleotide and 25 nM 484 bp mismatch DNA were irradiated for 25 min at 354 nm. The products were separated by SDS-PAGE and visualized by Coomassie staining without staining the DNA.

photocrosslink product compared to the cysteine–cysteine-crosslinked product may be a result of crosslinking to a residue in MutS located in the middle of the protein rather than the end, a phenomenon which we observed for a variety of single-cysteine MutL variants (unpublished observations). The presence of the *E. coli* MutL interaction partner MutH, which physically interacts with the NTD of MutL (46) did not affect formation of the MutS–MutL^{N131C} crosslink

products, suggesting that MutH does not interfere with complex formation of MutL with MutS.

DISCUSSION

In the present study, information on MLH1 mutations in human cancer diseases, bioinformatic analyses and *in vitro* assays served to identify four residues in human MLH1 whose mutation interfered with the interaction of MutL α (MLH1–PMS2) with MutS α . All mutations that abolished interaction also disabled the mismatch repair activity of MutL α . Therefore, we provided experimental evidence that the MutL α –MutS α interaction is required for mismatch repair.

In the present experiments, the investigated mutations could in principle have affected the MutL α –MutS α interaction in three ways: (i) the mutation could directly abolish the interaction interface; (ii) the mutation could confer an indirect effect on interaction by affecting the ATPase activity of MLH1; (iii) the mutation could abolish interaction with another protein that mediates the MutL α –MutS α complex, since nuclear extracts were used. However, it is most likely that the analyzed mutants affected the interaction interface directly for the following reasons: (i) the data from the mutational analysis is consistent with bioinformatic evidence indicating a conserved protein–protein interaction site in the investigated position; (ii) the crosslinking experiments with the homologue MutL reflected a direct physical contact; (iii) the interaction-deficient mutant MutL α ^{MLH1 H112D} had wild-type ATPase activity; (iv) the applied method efficiently captures complexes of purified MutL α and MutS α (28), excluding a prominent involvement of a mediator protein.

Two of the identified mutations affected charged surface residues (His112 and Arg127). The positive charge of His112 is very conserved in eukaryotic MLH1 (either His, Arg or Lys) and bacterial MutL (exclusively Arg), while this position is variable in PMS proteins (Figure 2B). This suggests that the positive charge in this position is vital for a functional interaction of MLH1 with MutS α . The other charged surface residue, Arg127, also shows very high conservation of the positive charge, but this is also true for PMS proteins, while MutL proteins have glutamine in this position. Two further mutations affected buried residues (Ala128 and Tyr130). The Ala128Pro mutation most likely disrupts the β -strand β 4 and thereby distorts the β -sheet. In the position of Tyr130, bulky hydrophobic residues are conserved, suggesting that hydrophobic interactions of this residue contribute to the structural integrity of this protein region. Accordingly, the conservative alteration of Tyr130 to phenylalanine (MLH1^{Y130F}) impaired neither interaction nor mismatch repair. All three mutations (Ala128Pro, Tyr130His/Ala) therefore seem incompatible with maintaining the structural integrity of the interaction interface.

All identified residues cluster in a small region of MLH1, suggesting that the major MutL α interaction interface for MutS α is (at least partly) constituted by residues from two adjacent β -strands (β 3– β 4) on the edge of an extensive β -sheet that backs the ATP binding pocket of MLH1 (Figure 5A). *E. coli* MutL could be site-specifically crosslinked to MutS from a residue in the corresponding

area, suggesting that MutL interacts with MutS with the same site. This is corroborated by the bioinformatic evidence, which suggested a similarly placed putative protein interaction site for MutL as for MLH1.

The interaction interface is easily accessible even when the N-terminal ATPase domains of MutL α dimerize similarly as has been described for *E.coli* MutL (Figure 5B). The contact of MLH1 to MutS α -DNA could therefore in principle be maintained by any intermediate of the ATPase cycle of MutL proteins, which involves this transient N-terminal dimerization. Since the interaction interface is located in immediate proximity to the MLH1 ATPase, interaction is very likely to modulate its activity. It will be interesting to investigate the nature of this modulation, since this may help disclose how mismatch recognition and repair are communicated from MutS to MutL and then to the downstream enzymes of mismatch repair. The presently identified mutations may facilitate such investigations, since they allow selectively disabling MutL-MutS complex formation.

During evolution, the homodimeric bacterial MutS and MutL proteins have changed to the heterodimeric complexes that are present in eukaryotic organisms. This process involved a specialization of the dimeric subunits as is exemplified by the finding that only one subunit of the eukaryotic MutS α / β heterodimers (MSH3/6) has retained the residues required for mismatch recognition, while these residues have been lost in their MSH2 partner. Our previous work already suggested that the interaction of two different human MutL heterodimers (MutL α , MLH1-PMS2, and MutL β , MLH1-PMS1) with MutS α depends predominantly on their MLH1 subunit (31). The present findings confirm that mutations in the NTD of MLH1 suffice to abolish interaction of MutL α with MutS α . This confirms the significance of MLH1 over its PMS partners in MutL-MutS interactions and suggests that MLH1 may have specialized to a general interaction adaptor responsible for docking human MutL heterodimers to their MutS partners, while the PMS family of proteins has (mostly) lost this ability.

The present study revealed that the cancer-associated mutation MLH1^{A128P} affects a residue located within the MutL α -MutS α interaction interface. This mutation disabled *in vitro* mismatch repair by destroying the ability of MutL α to interact with MutS α (although other functions of MLH1 may also be affected by the mutation). The MLH1^{A128P} mutation has been identified in an Italian kindred with a classical Lynch syndrome meeting the clinical Amsterdam criteria (71). In the affected family, several members suffered from cancer occurring at young age. The present findings are in good agreement with the severe clinical phenotype.

A second cancer-associated mutation, MLH1^{D132H}, affects a residue located very closely to the interaction interface, but the mutation did not detectably affect DNA mismatch repair or MutS α interaction in our experiments. The MLH1^{D132H} mutation is extraordinary since the cancer predisposition it confers is untypical for Lynch syndrome (60). Tumours of affected individuals rarely showed microsatellite instability (MSI), which is a genetic hallmark of deficient mismatch repair and typically observed in Lynch syndrome tumours. The mean age of cancer onset was significantly higher and the penetrance within the family appeared lower than in Lynch syndrome. Our finding that DNA mismatch repair was not detectably impaired by this mutation explains the rare occurrence of tumour MSI and the attenuated phenotype. Further studies will have to clarify if the low-penetrance cancer predisposition of this mutation is the result of a DNA mismatch repair defect that is too weak to be detectable by the applied *in vitro* mismatch repair assay or the mutation affects another function of MLH1.

In conclusion, our work identifies the interaction interface of MutL α for MutS α within MLH1. This interaction interface is also conserved in bacterial MutL proteins. Mutations in this interface in MLH1 suffice to abolish interaction of the MutL α heterodimer with MutS α and result in a MutL α heterodimer deficient for mismatch repair. One cancer-associated mutation which is located within this interaction interface abolishes interaction and repair, underlining the importance of this complex for mutation avoidance.

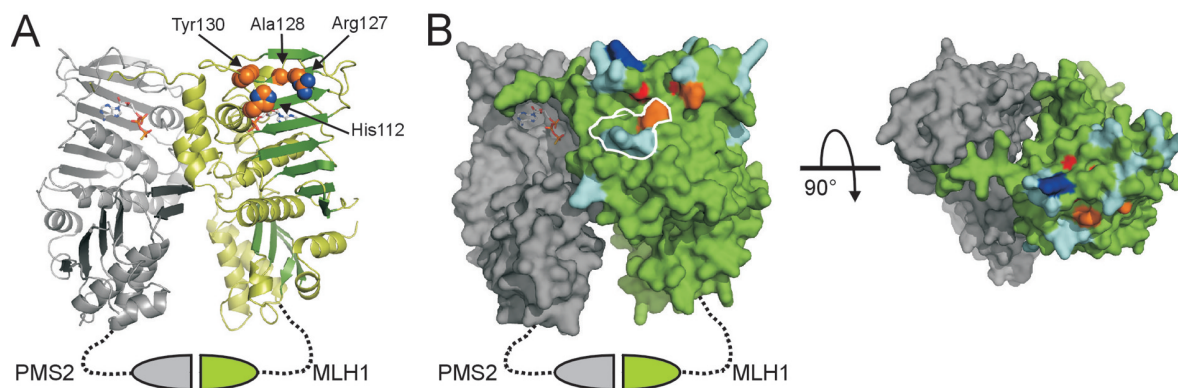


Figure 5. Localization of the MutL α -MutS α interaction interface on MLH1. (A) Cartoon presentation of the N-terminal MutL α dimer generated by superimposing the homology model of MLH1₁₋₃₃₅-ATP (shown in green) and the crystal structure of PMS2₃₁₋₃₆₄-ATPyS (72) (PDB 1H7U; shown in grey) on the crystal structure of the N-terminal homodimer of *E.coli* MutL₁₋₃₂₉-AMP-PNP (26) (PDB 1B63). The C-terminal constitutive dimerization domains are schematically added below the model (not to scale). The nucleotide phosphates are presented in sticks in the ATP binding pockets, which are backed by extensive β -sheets. Residues whose mutation interfered with MutL α -MutS α interaction (presented in spheres) are clustered on the top edge of this β -sheet of the MLH1 subunit. (B) Surface presentation of A. Buried or surface residues whose mutation abolished interaction are coloured red and orange, respectively. The MLH1 residue homologous to MutL^{N131} which has been crosslinked to MutS is coloured dark blue, while other investigated residues without detectable effect on interaction are coloured cyan. The region that the ProMate and ConSurf analyses suggested as interaction interface (Figure 2C) is marked by a white line.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Marc Wormek for technical assistance and conscientious performance of most of the described experiments. We would like to thank Dr Huixian Wang and Dr John B. Hays for generously providing the plasmids pUC19CPD and pUC19CPDC and Dr Hani Neuvirth for reviewing our ProMate results. This work was supported in part by research grant Pi258/7–1 to A.P., Fr1495/3–2 to P.F. and GRK 370 of the Deutsche Forschungsgemeinschaft and by research grant A/2004/13 of the University of the Saarland to G.P. Funding to pay the Open Access publication charges for this article was provided by internal funds of the 2nd Department of Internal Medicine of the University Hospital of the Saarland.

Conflict of interest statement. None declared.

REFERENCES

- Iyer,R.R., Pluciennik,A., Burdett,V. and Modrich,P.L. (2006) DNA mismatch repair: functions and mechanisms. *Chem. Rev.*, **106**, 302–323.
- Jiricny,J. (2006) The multifaceted mismatch-repair system. *Nature Rev. Mol. Cell Biol.*, **7**, 335–346.
- Plotz,G., Zeuzem,S. and Raedle,J. (2006) DNA mismatch repair and Lynch syndrome. *J. Mol. Histol.*
- Raedle,J., Trojan,J., Brieger,A., Weber,N., Schafer,D., Plotz,G., Staib-Sebler,E., Kriener,S., Lorenz,M. and Zeuzem,S. (2001) Bethesda guidelines: relation to microsatellite instability and MLH1 promoter methylation in patients with colorectal cancer. *Ann. Intern. Med.*, **135**, 566–576.
- Lynch,H.T. (1999) Hereditary nonpolyposis colorectal cancer (HNPCC). *Cytogenet. Cell Genet.*, **86**, 130–135.
- Genschel,J., Littman,S.J., Drummond,J.T. and Modrich,P. (1998) Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J. Biol. Chem.*, **273**, 19895–19901.
- Zhang,Y., Yuan,F., Presnell,S.R., Tian,K., Gao,Y., Tomkinson,A.E., Gu,L. and Li,G.M. (2005) Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell*, **122**, 693–705.
- Palombo,F., Iaccarino,L., Nakajima,E., Ikejima,M., Shimada,T. and Jiricny,J. (1996) hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.*, **6**, 1181–1184.
- Chen,P.C., Dudley,S., Hagen,W., Dizon,D., Paxton,L., Reichow,D., Yoon,S.R., Yang,K., Arnheim,N., Liskay,R.M. et al. (2005) Contributions by MutL homologues Mlh3 and Pms2 to DNA mismatch repair and tumour suppression in the mouse. *Cancer Res.*, **65**, 8662–8670.
- Lipkin,S.M., Wang,V., Jacoby,R., Banerjee-Basu,S., Baxevanis,A.D., Lynch,H.T., Elliott,R.M. and Collins,F.S. (2000) MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nature Genet.*, **24**, 27–35.
- Cannavo,E., Marra,G., Sabates-Bellver,J., Menigatti,M., Lipkin,S.M., Fischer,F., Cejka,P. and Jiricny,J. (2005) Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res.*, **65**, 10759–10766.
- 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.
- 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.
- Genschel,J. and Modrich,P. (2003) Mechanism of 5'-directed excision in human mismatch repair. *Mol. Cell*, **12**, 1077–1086.
- Hall,M.C., Jordan,J.R. and Matson,S.W. (1998) Evidence for a physical interaction between the *Escherichia coli* methyl-directed mismatch repair proteins MutL and UvrD. *EMBO J.*, **17**, 1535–1541.
- Hall,M.C. and Matson,S.W. (1999) The *Escherichia coli* MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity. *J. Biol. Chem.*, **274**, 1306–1312.
- Mechanic,L.E., Frankel,B.A. and Matson,S.W. (2000) *Escherichia coli* MutL loads DNA helicase II onto DNA. *J. Biol. Chem.*, **275**, 38337–38346.
- Welsh,K.M., Lu,A.L., Clark,S. and Modrich,P. (1987) Isolation and characterization of the *Escherichia coli* mutH gene product. *J. Biol. Chem.*, **262**, 15624–15629.
- Ban,C. and Yang,W. (1998) Structural basis for MutH activation in *E. coli* mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J.*, **17**, 1526–1534.
- Kadyrov,F.A., Dzantiev,L., Constantin,N. and Modrich,P. (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell*, **126**, 297–308.
- Dzantiev,L., Constantin,N., Genschel,J., Iyer,R.R., Burgers,P.M. and Modrich,P. (2004) A defined human system that supports bidirectional mismatch-provoked excision. *Mol. Cell*, **15**, 31–41.
- Dutta,R. and Inouye,M. (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.*, **25**, 24–28.
- Ban,C. and Yang,W. (1998) Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, **95**, 541–552.
- Guarne,A., Ramon-Maiques,S., Wolff,E.M., Ghirlando,R., Hu,X., Miller,J.H. and Yang,W. (2004) Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J.*, **23**, 4134–4145.
- Kosinski,J., Steindorf,I., Bujnicki,J.M., Giron-Monzon,L. and Friedhoff,P. (2005) Analysis of the quaternary structure of the MutL C-terminal domain. *J. Mol. Biol.*, **351**, 895–909.
- Ban,C., Junop,M. and Yang,W. (1999) Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell*, **97**, 85–97.
- Grilley,M., Welsh,K.M., Su,S.S. and Modrich,P. (1989) Isolation and characterization of the *Escherichia coli* mutL gene product. *J. Biol. Chem.*, **264**, 1000–1004.
- Plotz,G., Raedle,J., Brieger,A., Trojan,J. and Zeuzem,S. (2002) hMutSalpha forms an ATP-dependent complex with hMutLalpha and hMutLbeta on DNA. *Nucleic Acids Res.*, **30**, 711–718.
- Raschle,M., Dufner,P., Marra,G. and Jiricny,J. (2002) Mutations within the hMLH1 and hPMS2 Subunits of the Human MutLalpha Mismatch Repair Factor Affect Its ATPase Activity, but Not Its Ability to Interact with hMutSalpha. *J. Biol. Chem.*, **277**, 21810–21820.
- Spampinato,C. and Modrich,P. (2000) The MutL ATPase is required for mismatch repair. *J. Biol. Chem.*, **275**, 9863–9869.
- Plotz,G., Raedle,J., Brieger,A., Trojan,J. and Zeuzem,S. (2003) N-terminus of hMLH1 confers interaction of hMutLalpha and hMutLbeta with hMutSalpha. *Nucleic Acids Res.*, **31**, 3217–3226.
- Blackwell,L.J., Wang,S. and Modrich,P. (2001) DNA chain length dependence of formation and dynamics of hMutSalpha.hMutLalpha.heteroduplex complexes. *J. Biol. Chem.*, **276**, 33233–33240.
- Cupples,C.G. and Miller,J.H. (1989) A set of lacZ mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc. Natl. Acad. Sci. USA*, **86**, 5345–5349.
- Feng,G. and Winkler,M.E. (1995) Single-step purifications of His6-MutH, His6-MutL and His6-MutS repair proteins of *Escherichia coli* K-12. *Biotechniques*, **19**, 956–965.
- Loh,T., Murphy,K.C. and Marinus,M.G. (2001) Mutational analysis of the MutH protein from *Escherichia coli*. *J. Biol. Chem.*, **276**, 12113–12119.
- Evans,S.J., Fogg,M.J., Mamone,A., Davis,M., Pearl,L.H. and Connolly,B.A. (2000) Improving dideoxynucleotide-triphosphate utilization by the hyper-thermophilic DNA polymerase from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.*, **28**, 1059–1066.
- Sambrook,J. and Russel,D.W. (2001) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Wang,H.X. and Hays,J.B. (2000) Preparation of DNA substrates for *in vitro* mismatch repair. *Mol. Biotechnol.*, **15**, 97–104.

39. Wang,H.X. and Hays,J.B. (2001) Simple and rapid preparation of gapped plasmid DNA for incorporation of oligomers containing specific DNA lesions. *Mol. Biotechnol.*, **19**, 133–140.
40. Wang,H.X. and Hays,J.B. (2002) Mismatch repair in human nuclear extracts—Quantitative analyses of excision of nicked circular mismatched DNA substrates, constructed by a new technique employing synthetic oligonucleotides. *J. Biol. Chem.*, **277**, 26136–26142.
41. Wang,H.X. and Hays,J.B. (2002) Mismatch repair in human nuclear extracts—Time courses and ATP requirements for kinetically distinguishable steps leading to tightly controlled 5' to 3' and aphidicolin-sensitive 3' to 5' mispair-provoked excision. *J. Biol. Chem.*, **277**, 26143–26148.
42. Corrette-Bennett,S.E. and Lahue,R.S. (1999) Mismatch repair assay. *Methods Mol.Biol.*, **113**, 121–132.
43. Trojan,J., Zeuzem,S., Randolph,A., Hemmerle,C., Brieger,A., Raedle,J., Plotz,G., Jiricny,J. and Marra,G. (2002) Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology*, **122**, 211–219.
44. Thomas,E., Pingoud,A. and Friedhoff,P. (2002) An efficient method for the preparation of long heteroduplex DNA as substrate for mismatch repair by the *Escherichia coli* MthLS system. *Biol. Chem.*, **383**, 1459–1462.
45. Toedt,G.H., Krishnan,R. and Friedhoff,P. (2003) Site-specific protein modification to identify the MutL interface of MthH. *Nucleic Acids Res.*, **31**, 819–825.
46. Giron-Monzon,L., Manelyte,L., Ahrends,R., Kirsch,D., Spengler,B. and Friedhoff,P. (2004) Mapping Protein–Protein Interactions between MutL and MthH by Cross-linking. *J. Biol. Chem.*, **279**, 49338–49345.
47. Pace,C.N., Vajdos,F., Fee,L., Grimsley,G. and Gray,T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.*, **4**, 2411–2423.
48. Wheeler,D.L., Church,D.M., Edgar,R., Federhen,S., Helmberg,W., Madden,T.L., Pontius,J.U., Schuler,G.D., Schriml,L.M., Sequeira,E. *et al.* (2004) Database resources of the National Center for Biotechnology Information: update. *Nucleic Acids Res.*, **32**, D35–D40.
49. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
50. Chenna,R., Sugawara,H., Koike,T., Lopez,R., Gibson,T.J., Higgins,D.G. and Thompson,J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, **31**, 3497–3500.
51. Galtier,N., Gouy,M. and Gautier,C. (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.*, **12**, 543–548.
52. Vriend,G. (1990) WHAT IF: a molecular modelling and drug design program. *J. Mol. Graph.*, **8**, 52–5629.
53. Westbrook,J., Feng,Z., Chen,L., Yang,H. and Berman,H.M. (2003) The protein data bank and structural genomics. *Nucleic Acids Res.*, **31**, 489–491.
54. Nicholas,K.B., Nicholas,H.B., Jr and Deerfield,D.W. (1997) GeneDoc: Analysis and visualization of genetic variation. *EMBNET.NEWS*, **4**, 1–4.
55. Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
56. Neuvirth,H., Raz,R. and Schreiber,G. (2004) ProMate: a structure based prediction program to identify the location of protein–protein binding sites. *J. Mol. Biol.*, **338**, 181–199.
57. Landau,M., Mayrose,I., Rosenberg,Y., Glaser,F., Martz,E., Pupko,T. and Ben-Tal,N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.*, **33**, W299–302.
58. Peltomaki,P. (2005) Lynch syndrome genes. *Fam. Cancer.*, **4**, 227–232.
59. Peltomaki,P. and Vasen,H. (2004) Mutations associated with HNPCC predisposition—Update of ICG-HNPCC/INSIGHT mutation database. *Dis. Markers*, **20**, 269–276.
60. Lipkin,S.M., Rozek,L.S., Rennert,G., Yang,W., Chen,P.C., Hacia,J., Hunt,N., Shin,B., Fodor,S., Kokoris,M. *et al.* (2004) The MLH1 D132H variant is associated with susceptibility to sporadic colorectal cancer. *Nat.Genet.*, **36**, 694–699.
61. Brieger,A., Plotz,G., Raedle,J., Weber,N., Baum,W., Caspary,W.F., Zeuzem,S. and Trojan,J. (2005) Characterization of the nuclear import of human MutL α . *Mol.Carcinog.*, **43**, 51–58.
62. Plotz,G., Piiper,A., Wormek,M., Zeuzem,S. and Raedle,J. (2006) Analysis of the human MutL α .MutS α complex. *Biochem. Biophys. Res. Commun.*, **340**, 852–859.
63. DeLano,W.L., Ultsch,M.H., de Vos,A.M. and Wells,J.A. (2000) Convergent solutions to binding at a protein–protein interface. *Science*, **287**, 1279–1283.
64. Teichmann,S.A. (2002) The constraints protein–protein interactions place on sequence divergence. *J. Mol. Biol.*, **324**, 399–407.
65. Caffrey,D.R., Somaroo,S., Hughes,J.D., Mintseris,J. and Huang,E.S. (2004) Are protein–protein interfaces more conserved in sequence than the rest of the protein surface? *Protein Sci.*, **13**, 190–202.
66. Drozdov-Tikhomirov,L.N., Linde,D.M., Poroikov,V.V., Alexandrov,A.A. and Skurida,G.I. (2001) Molecular mechanisms of protein–protein recognition: whether the surface placed charged residues determine the recognition process? *J. Biomol. Struct. Dyn.*, **19**, 279–284.
67. Junop,M.S., Yang,W., Funchain,P., Clendenin,W. and Miller,J.H. (2003) *In vitro* and *in vivo* studies of MutS, MutL and MthH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair (Amst)*, **2**, 387–405.
68. Hall,M.C., Shcherbakova,P.V., Fortune,J.M., Borchers,C.H., Dial,J.M., Tomer,K.B. and Kunkel,T.A. (2003) DNA binding by yeast Mlh1 and Pms1: implications for DNA mismatch repair. *Nucleic Acids Res.*, **31**, 2025–2034.
69. Green,N.S., Reisler,E. and Houk,K.N. (2001) Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci.*, **10**, 1293–1304.
70. Tao,T., Lamkin,M. and Scheiner,C.J. (1985) The conformation of the C-terminal region of actin: a site-specific photocrosslinking study using benzophenone-4-maleimide. *Arch. Biochem. Biophys.*, **240**, 627–634.
71. Pensotti,V., Radice,P., Presciuttini,S., Calistri,D., Gazzoli,I., Grimalt,P.A., Mondini,P., Buonsanti,G., Sala,P., Rossetti,C. *et al.* (1997) Mean age of tumour onset in hereditary nonpolyposis colorectal cancer (HNPCC) families correlates with the presence of mutations in DNA mismatch repair genes. *Genes Chromosomes Cancer*, **19**, 135–142.
72. Guarne,A., Junop,M.S. and Yang,W. (2001) Structure and function of the N-terminal 40 kDa fragment of human PMS2: a monomeric GHL ATPase. *EMBO J.*, **20**, 5521–5531.