Supplementary Information for

Molecular basis of the dual role of the Mlh1-Mlh3 endonuclease in MMR and in meiotic crossover formation

Jingqi Dai1*, Aurore Sanchez2*, Céline Adam2, Lepakshi Ranjha3, Giordano Reginato3,4, Pierre Chervy1, Carine Tellier-Lebegue1, Jessica Andreani1, Raphaël Guérois1, Virginie Ropars1, Marie-Helene Le Du1, Laurent Maloisel5,6, Emmanuelle Martini5,6, Pierre Legrand7, Aurélien Thureau7, Petr Cejka3,4, Valérie Borde2§, Jean-Baptiste Charbonnier1§

corresponding authors : Valérie Borde2§, Jean-Baptiste Charbonnier1§
: valerie.borde@curie.fr and jb.charbonnier@cea.fr

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supplementary Information Text

Preparation of Mlh1-Mlh3(CTD) complex in *E. coli*: The *S. cerevisiae* Mlh1 C-terminal domain (CTD) contains residues 485 to 769. It was cloned into a vector derived from the pETM30 (EMBL, Heidelberg). The N-terminus of the Mlh1(CTD) is fused to a His6-GST tag separated by a TEV protease site. The *S. cerevisiae* Mlh3(CTD) contains residues 460 to 715. It was cloned into a vector pEXP HisPKM (AFMB, Marseille). The N-terminus of the Mlh3(CTD) is fused to a His6-MBP tag separated by a TEV protease site. *E. coli* BL21(DE3) star cells (Invitrogen) were first transformed with GST-Mlh1(CTD) plasmid. Competent cells with the GST-Mlh1(CTD) plasmid were prepared and transformed with the MBP-Mlh3(CTD) plasmid. The cells were grown at 37°C until an optical density of 1.2 and induced with 0.5mM IPTG (isopropyl-β-d-thiogalactopyranoside) overnight at 20°C. The cells were harvested, washed with phosphate-buffered saline (PBS 1x) and frozen in liquid nitrogen. The cells were re-suspended in buffer L (PBS 1x, pH7.5, 350 mM NaCl, 5% Glycerol, 0.2% Triton, 10 mM MgCl2, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 10 mM β-Mercaptoethanol), incubated with lysozyme (Euromedex) (25mg/L culture) and benzonase (10U/ml lysate) and lysed by sonication. After centrifugation at 20 000 rpm for 30 min, the supernatant fraction was loaded on pre-equilibrated amylose resin (30ml, Biolabs). The resin was washed with buffer W (50 mM Tris-HCl, pH7.5, 200 mM NaCl, 10 mM MgCl2 and 10 mM β-Mercaptoethanol). The complex MBP-Mlh3 and GST-Mlh1 was eluted in buffer W containing 10 mM maltose. The maltose-binding protein (MBP) and Glutathione S-transferase (GST) tags were cleaved by adding TEV protease (2% w/w) overnight at 4°C. The cleaved products were applied to a HisTrap column (5ml, GE Healthcare). The complex Mlh1(CTD)-Mlh3(CTD) was collected in the flow through, the GST, the MBP, the non-cleaved proteins and the TEV remain bound to the column with their His6 tag. The complex was then dialyzed against dialysis buffer (50 mM Tris-HCl, pH7.5, 50 mM NaCl, 5 mM MgCl2 and 10 mM β-Mercaptoethanol), and applied to a pre-equilibrated RESOURCE Q

anion-exchange column (GE Healthcare). The complex Mlh1-Mlh3(CTD) was eluted at 170mM NaCl with a gradient of 0.05-1M NaCl over 20 column volumes.

Preparation of full-length Mlh1-Mlh3 heterodimer and Mlh1-Mlh3(KERE) mutant:

The full-length *S. cerevisiae* Mlh1-Mlh3 heterodimer was prepared using a construct with 8xhis-tag on the N-terminus of Mlh1 (pFB-his-MLH1) and MBP-tag on the N-terminus of Mlh3 (pFB-MBP-MLH3). The protein expression and purification was carried out as previously described (1). The MBP tag was cleaved by PreScission protease (PP). For the preparation of Mlh1-Mlh3(KERE) mutant, the 8xhis-Mlh1 and MBP-Mlh3 segments were cloned into a Multibac vector (2), creating pKL-his-MLH1–MBP-PP-MLH3. The mutations K668E and R671E in Mlh3 gene were created using oligonucleotides 5'-GGTCTTTCATGAAATTCTTA ACAGCgaaGCATGCgaaTCTGCTGTTATGTTTGGAGATGAAT-3' and 5'-ATTCATCT CCAAACATAACAGCAGAttcGCATGCttcGCTGTTAAGAATTTCATGAAAGACC-3' by site-directed mutagenesis (Q5 Site-Directed Mutagenesis Kit, New England Biolabs). Each plasmid was then integrated in a Yellow Green Protein (YFP) containing bacmid by transformation in EMBacY *E. coli* strain (kind gift from I Berger, University of Bristol). The resulting recombinant bacmid containing both Mlh1 and Mlh3(KERE) genes was used to transfect *Sf*21 insect cells giving rise to the V0 virus generation. After amplification, stocks of baculoviruses were titrated by the dilution limit method using YFP as marker for infected cells and Mac Grady table. Protein production was initiated in Sf21 cells by infection with the baculovirus stock at MOI of 5x10-3 and cells were collected 5-6 days after the infection (3-4 days after the proliferation arrest). The protein purification was performed similarly to wild-type Mlh1-Mlh3 and the final yield of the Mlh1-Mlh3(KERE) heterodimer was 0.4 mg of purified heterodimer from 1.8 liter of culture.

DNA binding assay: The 50 bp dsDNA and HJ DNA substrates were prepared as described previously (1). The 100 bp dsDNA substrate was prepared by labelling and annealing the following oligonucleotides: BIO100 5'-G**T**AAGTGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCATAATCTTCT GCCATGGTCGTAGCAGCCTCCTGCATC-3' and BIO100C 5'ATGCAGGAGGCTGCTACGACCATGGCAGAAGATTATGAGGTGGAGTACGCGCCCGGGGAGCCCAAGGGCAC GCCCTGGCACCCGCACCGCGGCACTTAC-3'. The binding assays were carried out as mentioned previously (1). Briefly, the reactions (15 μl) were performed in 25 mM Tris-acetate pH 7.5, 3 mM EDTA, 1 mM DTT, 100 μg/ml BSA (New England Biolabs), DNA substrate (1 nM, in molecules) and 1.3 ng/μl poly(dI-dC) as competitor DNA. For comparison of DNA binding between MutLγ(CTD) and MutLα(CTD) with different substrates (Figure 3a-b), the competitor DNA was omitted from the reactions. The DNA binding experiments were performed without added salt. The final NaCl concentration was 43 mM due to the addition of the recombinant proteins.

Nuclease assay: The nuclease assays were carried out as described (1).

Crystal structures of the MutLγ(CTD) complexes: The crystallization screenings of the complex MutLγ(CTD) were performed on the HTX platform (EMBL, Grenoble) and the crystal hits were optimized manually. The complex was crystallized by using vapor diffusion in sitting drops at 20 °C. The optimized conditions were realized by mixing 1.5µl of protein sample at 6mg/ml with 1.5µl reservoir solution (10% PEG4000, 200 mM imidazole-malate, pH6 and 10 mM MgCl2). Crystals were cryo-protected in mother solution supplemented with 20% v/v xylitol (or 20% v/v ethylene glycol) before freezing in liquid nitrogen. Diffraction Data were collected at 100K on protein-crystallography beamline Proxima1 (Soleil, France) (Table S1). The crystal structures of four different crystal forms were analyzed. Four crystals belong to the space group I222. Crystallographic data were processed with XDS (2) and scaled using scala from the CCP4 suite (3). Crystal structures were solved by Molecular Replacement with molrep from the CCP4, by using the chain A and regulator domain of chain B of PDB 4E4W (model of Mlh1-Pms1 (CTD)). The tracing of an initial model was realized in COOT (4). Refinement was performed by BUSTER (5). In the final model, 34 residues of Mlh1(CTD) (485-505, 586-591, 635-641) and 60 residues of Mlh3(CTD) (460-487, 509-517, 586-592, 638-650, 712-715) were not modelled since no electron density is visible. It includes 21 and 28 first residues of both fragments likely unstructured. The metal ions in the endonuclease sites were attributed to zinc atoms for the following reasons. Firstly, the electron density is high and compatible with a zinc atoms rather than water or Mg^{2+}. During refinement the zinc atoms as Bfactors similar to surrounding atoms of the protein. Secondly, the zinc atoms are in interactions with protein atoms with a tetrahedric geometry often observed for Zn atoms. It is to be noticed that no zinc was added in the purification protocol or in the crystallization conditions. The different crystal structures, with one or two Zn^{2+}, were determined by collecting diffraction data on a large number of crystals obtained in the same crystallization condition.

SAXS analyses of the MutLγ(CTD) complexes: Small-Angle X-ray scattering: SAXS measurements were conducted at the French synchrotron SOLEIL (Saint. Aubin, France) on the SWING beamline. Data were recorded using an AVIEX170170 CCD detector at the distance of 1.8 m (λ = 1.033 Å) at 15°C. Mlh1-Mlh3 (CTD) complex at different concentrations (1, 2 and 4 mg/ml) were prepared in a SAXS buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol, 5 % glycerol). After a small spin at 15000 g for 10 min in order to remove the aggregates, 50µL of each concentration of Mlh1-Mlh3 (CTD) and buffer were injected into the SAXS capillary cell and collected with a frame duration of 0.5s and a dead time between frames of 0.5s. Data reduction to absolute unit, frame averaging, and subtraction were performed with FOXTROT (6). Atoms from disordered loops and N terminal part of Mlh1-Mlh3 (CTD) were added using modeller script (7). The complete atomic model were then refined using DADIMODO (software for rigid-body refinement of atomic models of multi-domain proteins against SAXS data (qmax = 0.4Å-1); https://dadimodo.synchrotron-soleil.fr, (8)). Briefly, the reconstruction of complete model for DADIMODO uses the Loopmodel class from Modeler. The rigid parts (here all the domains) are fixed during the addition of loops and N-terminal parts. DADIMODO refines atomic models of multidomain proteins or complexes against small-angle X-ray scattering data. Stepwise generic conformational changes are applied in a stochastic optimization algorithm that performs a search in the protein conformation space.

The software enables flexibility in the disorder loops (residues A583-A598, A633-A649, B506-B521, B582-B596, B635-B651) and in the N terminal parts (residues A485-A510, B460-B489). Five models were generated and ranked according to their χ^2.

Molecular Modelling

A structural model of the complex between the full-length sequences of yeast Mlh1 and Mlh3 was obtained using the RosettaCM software for homology modeling (9). Uniprot sequences P38920 and Q12083 were used for Mlh1 and Mlh3, respectively. For the C-terminal domains, the structure of the complex reported in the present work was used as a

template. For the N-terminal domains, PDB structure 4p7a (N-terminal domain of human MLH1) was used as a template for both yMlh1 and yMlh3, and the dimeric biological assembly reported for 4p7a as a template for the interaction between yMlh1(NTD) and yMlh3(NTD). The sequence alignment between hMLH1 and yMlh1, on the one hand, and hMLH1 and yMlh3, on the other hand, were obtained with HHsearch (10)using default settings, with MAC realignment. The N-terminal and C-terminal domains of the templates were positioned at a distance such that the linkers would be mostly extended in the model.

Multiple sequence alignments for yMlh1 and yMlh3 were obtained using blast against the Uniprot database, with a minimum sequence identity of 30% with the query sequence. These multiple sequence alignments were used to map evolutionary conservation on the surface of the full-length model of the yMlh1/yMlh3 complex, using the Rate4Site software (11).

Yeast manipulations and strains construction:

S. cerevisiae haploid strains used for the study of spontaneous mutagenesis are isogenic to E134 (*MATα ade5-1 lys2::InsEA14 trp1-289 his7-2 leu2-3,112 ura3-52*) (12) that was kindly provided by Dr Erik Johansson. The genomic *mlh1ΔC1* and *mlh1ΔC3* mutations were generated by using the plasmids pL33 and pL34, respectively. pL33 and pL34 are derivatives of pL32 obtained by cloning the C-terminal end of *MLH1* (from +1672 of the ORF to 500 bp downstream the stop codon) in pRS306. The *mlh1ΔC1* and *mlh1ΔC3* mutations were inserted after site-directed mutagenesis performed in pL32. For targeting the mutations by the two-step replacement procedure (13), pL33 and pl34 were digested with HindIII. The *mlh3-KERE* mutation was introduced in strain E134 by PCR in a fragment containing flanking homology to the *MLH3* locus and containing a HphMX drug resistance cassette inserted 18 bp downstream of *MLH3* stop codon.

All diploid yeast strains are derivatives of the SK1 background. Yeast strains genotype is provided in SI Table S3. For synchronous meiosis, cells were grown in SPS pre-sporulation medium and transferred in sporulation medium (1.0% (w/v) potassium acetate, 0.001% polypropylene glycol) with vigorous shaking at 30°C as described (14). For strain constructions and spore viability measurements, sporulation was performed on solid sporulation medium for two days at 30°**C.** SK1 yeast strains were obtained by direct transformation or crossing to obtain the desired genotype. The *mlh1*∆*C1* and *mlh3*∆*C3* mutations were introduced by PCR and fusion of fragments containing flanking homology to the *MLH1* locus and containing a KanMX cassette 18 bp after *MLH1* stop codon. The *mlh3* mutations or fusion with the internal Myc8 tag were each introduced using a similar PCR strategy, introducing a HphMX drug resistance cassette 18 bp after *MLH3* stop codon. All transformants were screened using PCR discriminating between correct and incorrect integrations and sequencing for mutagenesis. The *mlh1*∆, *mlh3*∆ and *MLH3-Myc8* alleles have been described previously (15, 16).

Spontaneous mutation rates: For data presented in Figure 2c, the fluctuation tests to determine spontaneous mutation rates were performed in two to five independent experiments of nine independent cultures. For each strain tested, the nine independent cultures were inoculated with about 100 cells in 2 mL of YPD and grown at 30°C for three days. Cell density in each culture was measured by plating dilutions on YPD agar plates and counting the colonies after three days at 30°C. The quantification of Lys+ revertants was determined after plating on supplemented synthetic agar plates lacking lysine. Colonies were counted after four days at 30°C. Mutation rates (events/cell/generation) were determined from the number of Lys+ colonies by the method of the median (17, 18). For data presented in Figure 3e, the spontaneous reversion rate at the *lys2::InsE-A14* locus was measured using three single colonies from three independent transformants of the same genotype grown to stationary phase in liquid YPAD medium and plated onto YPAD or selective medium lacking lysine for revertant count.

Two hybrid assays: Yeast two-hybrid assays were performed exactly as described (15). *MLH3* and *MLH1* were PCR-amplified from SK1 genomic DNA. PCR products were cloned in plasmids derived from the 2 hybrid vectors pGADT7 (GAL4-activating domain) and pGBKT7 (GAL4-binding domain) creating N terminal fusions and transformed in yeast haploid strains Y187 and AH109 (Clontech), respectively. Interactions were scored, after mating and diploid selection on dropout medium without leucine and tryptophan, as growth on dropout medium without leucine, tryptophan and histidine.

Analysis of crossover frequencies

For CO analysis at the *HIS4LEU2* locus, cells were harvested from meiotic time courses at the indicated time point. 2 µg of genomic DNA was digested with *Xho*I, loaded on a 0.6 % agarose gel in 1X TBE and run at 76 V/cm for 17hr. COs were then detected after Southern blot and hybridization with a ^{32}P-labeled DNA probe A, as described (19). Hybridization signal was detected by a Phosphorimager (Typhoon, GE Healthcare) and quantified using the Image Quant software and CO frequency was calculated as described (15). Data from 8 h and 9 h time points were averaged for each biological repeat. The data were further normalized to the crossover frequency observed in the wild type. For genetic distances on chromosome VIII, diploids were sporulated in liquid medium, and recombination between fluorescent markers was scored after 24h in sporulation, by microscopy analysis as described (20). At least 600 tetrads were counted. Genetic distances in the *CEN8-ARG4* and *ARG4-THR1* intervals were calculated from the distribution of parental ditype (PD), nonparental ditype (NPD) and tetratype (T) tetrads and genetic distances (cM) were calculated using the Perkins equation: cM = (100 (6NPD + T))/(2(PD + NPD + T)). Standard errors of genetic distances were calculated using Stahl Lab Online Tools (https://elizabethhousworth.com/StahlLabOnlineTools/).

Chromatin immunoprecipitation

Chromatin immunoprecipitation of Mlh3-Myc8 and quantitative PCR were performed as described (16). Results are expressed as % of DNA in the total input present in the immunoprecipitated sample and normalized first by the negative control site in the middle of *NFT1*, which shows no meiotic DSB, and then by the 0 h time point of the meiotic time-course. Primers for qPCR have been described (15).

Table S1: Data collection and refinement statistics

	MutLγ (CTD)	MutL γ (CTD)	MutL γ (CTD)	MutL γ (CTD)
PDB code	6RMN	6SHX	6SNS	6SNV
Data collection				
Crystal Form	Zn2	ZnA	ZnB	Zn2b
Space group	I 2 2 2	I 2 2 2	I 2 2 2	P 21 21 2
a, b, c (Å)	92.3,104.4,	93.6, 104.1, 134.1	91.9, 103.1,	90.9,134.8,
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	0.9786	0.98011	0.9786	0.98011
Resolution (Å)[a]	50-2.2 (2.25-	50-2.4 (2.59-2.4)	50-2.6 (2.78-2.6)	50-2.5 (2.66-
Anisotropy	2.8, 2.5, 2.08	2.8, 2.6, 2.3	3.3, 3.2, 2.4	3.1, 2.3, 3.01
R_{merge}	0.085 (1.42)	0.098 (0.82)	0.051 (1.47)	0.05 (0.779)
R_{meas}	0.088 (1.49)	0.106 (0.91)	0.056 (1.60)	0.054 (0.908)

R_{pim}	0.025 (0.42)	0.040 (0.40)	0.022 (0.613)	0.02 (0.457)
$I/\sigma I$	16.29 (1.84)	11.2 (1.7)	16.0 (1.3)	18.56 (1.56)
Completeness (spherical)	0.72 (0.19)	0.76 (0.19)	0.67 (0.19)	0.689 (0.21)
Completeness (ellipsoidal)	0.95 (0.85)	0.94 (0.67)	0.95 (0.96)	0.95 (0.94)
Redundancy	12.9 (12.1)	6.94 (4.8)	6.5 (6.7)	6.8 (3.67)
B Wilson ($Å^2$)	69	68	124	97
Refinement				
Resolution (Å)	38.14-2.2	48.3-2.4	48.1-2.6	44.7-2.5
No. reflections	22942	20049	13316	30067
R_{work} / R_{free}(%)	19.7 / 22.6	21.3 / 23.6	21.5 / 24.6	22.9 / 27.1
No. Atoms				
Protein	3961	3957	3889	7922
Ligand/ion[b]	2	1	1	4
Water	132	77	35	25
B factors overall ($Å^2$)	69.1	62.8	114.5	90.5
r.m.s. deviations				
Bond lengths (Å)	0.009	0.008	0.008	0.008
Bond angles (°)	1.03	1.00	1.01	0.98

Table S2: Spore viabilities

Strain name	Genotype	Tetrad type					Spore viability (%)	Tetrads number	Fisher Test p-Value vs WT
		4 sp	3 sp	2 sp	1 sp	0 sp			
VBD1311	WT	131	9	7	0	0	96.1	147	
VBD1149	*mlh1*Δ	64	22	57	23	13	64.1	179	5.6E-51
VBD2029	*mlh1*Δ*C1*	138	9	7	0	2	95.0	156	0.40
VBD2031	*mlh1*Δ*C3*	56	20	44	11	25	61.4	156	1.9E-54
VBD1749	*mlh3*Δ	82	6	12	0	9	84.9	109	4.0E-10
VBD1803	*mlh3KERE*	69	10	16	2	6	82.5	103	7.8E-13
VBD2056	*mlh3MN1*	93	7	3	1	0	96.2	104	1.00
VBD2083	*mlh3MN2*	95	4	4	1	0	96.4	104	0.87
VBD2066	*mlh3MN1MN2*	98	6	18	2	3	88.2	127	9.4E-07
VBD2062	*mlh3MC1*	95	3	6	0	0	96.4	104	0.87
VBD2067	*mlh3MC1MC2*	93	2	9	0	0	95.2	104	0.53
VBD2084	*mlh3MC3*	86	6	9	1	2	91.6	104	3.6E-03
VBD2073	*mlh3MC1MC2MC3*	93	9	15	1	12	82.7	130	1.2E-13
VBD2057	*mlh3MN1MC1*	93	1	7	1	2	93.8	104	0.10
VBD2085	*mlh3MN2MC3*	75	5	12	0	10	83.1	102	4.5E-12
VBD2068	*mlh3MN1MN2 MC1MC2*	75	6	12	0	11	82.2	104	2.6E-13
VBD2075	*mlh3MN1MN2 MC1MC2MC3*	81	4	12	1	6	86.8	104	7.5E-08
VBD1599	*MLH3-Myc8*	102	4	10	0	1	94.0	117	
VBD2157	*mlh3F483E-Myc8*	46	1	1	0	2	94.5	50	1.00

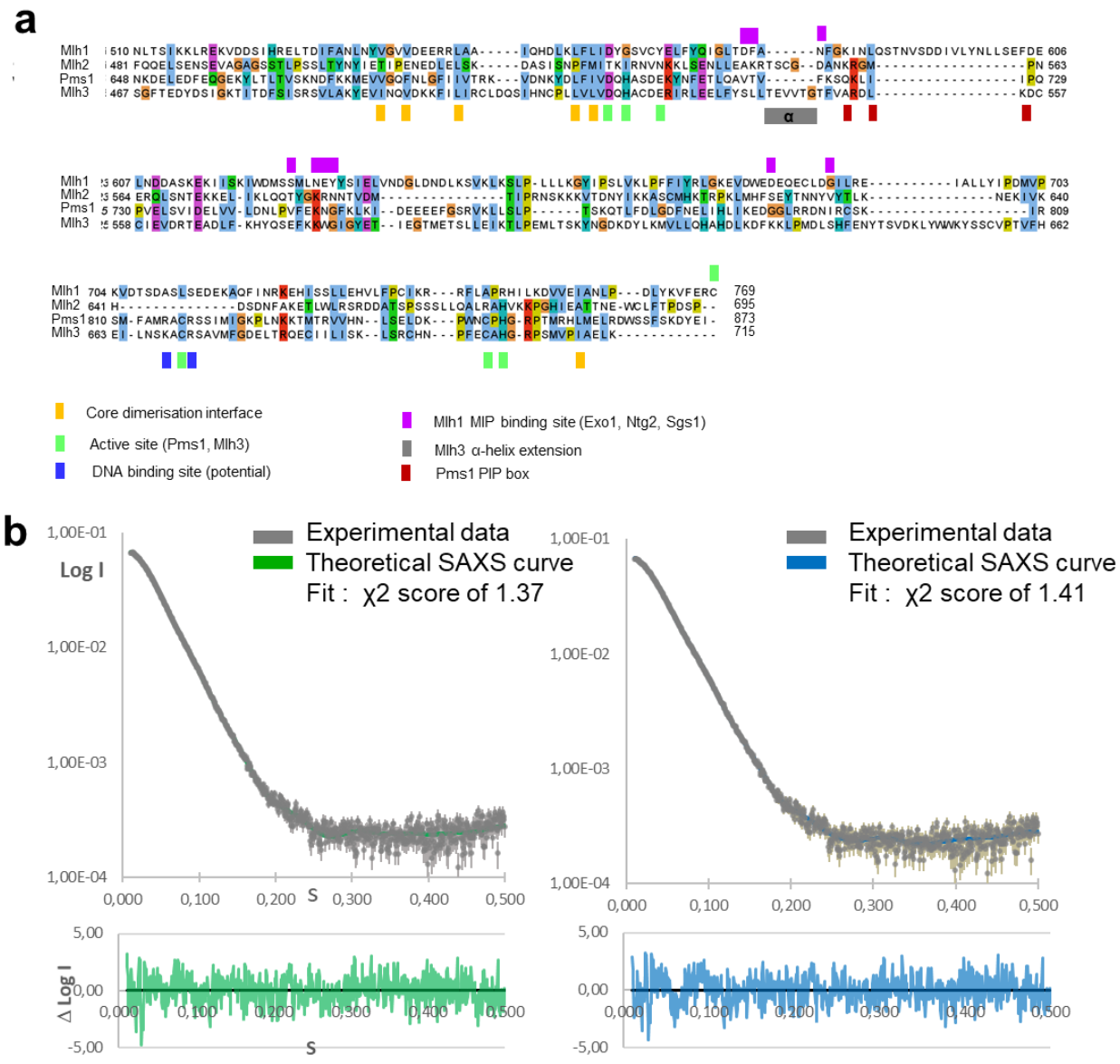
Table S3: Genotypes of strains used in this study.

Diploids : All diploids are of the SK1 background. The MATa parent is indicated first

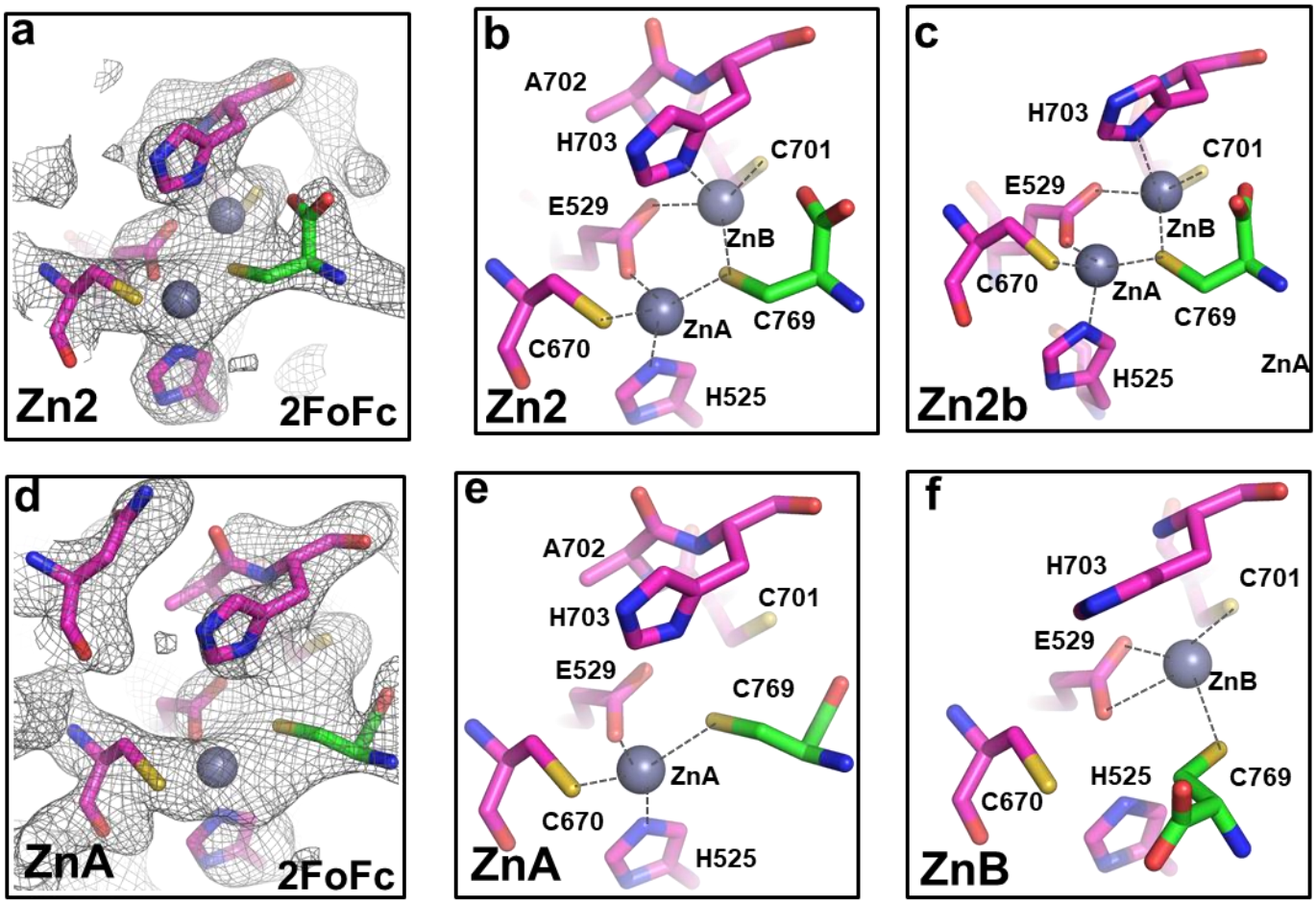
Strain name	Genotype
VBD1311	*a/l ho::hisG/" leu2::hisG/" ura3/" HIS4::LEU2-(BamH1; +ori)/his4-X::LEU2-(NgoMIV; +ori)-URA3*
VBD1149	same as VBD1311 but *mlh1Δ::HphMX/"*
VBD2029	same as VBD1311 but *mlh1ΔC1::KanMX/"*
VBD2031	same as VBD1311 but *mlh1ΔC3::KanMX/"*
VBD1803	same as VBD1311 but *mlh3-K668ER671E::HphMX/"*
VBD1749	same as VBD1311 but *mlh3Δ::HphMX/"*
VBD1716	*a/l ho::hisG/" leu2::hisG/" ura3/" HIS4::LEU2-(BamH1; +ori)/his4-X::LEU2-(BamH1; +ori)-URA3 MLH3-512-8Myc-513::HphMX/"*
VBD1802	*same as VBD1716 but mlh3-512-8Myc-513-K668ER671E::HphMX/"*
VBD2056	same as VBD1311 but *mlh3-K320E::HphMX/"*
VBD2083	same as VBD1311 but *mlh3-K316E::HphMX/"*
VBD2066	same as VBD1311 but *mlh3-K316EK320E::HphMX/"*
VBD2062	same as VBD1311 but *mlh3-D611K::HphMX/"*
VBD2067	same as VBD1311 but *mlh3-E565KD611K::HphMX/"*
VBD2084	same as VBD1311 but *mlh3-D625K::HphMX/"*
VBD2073	same as VBD1311 but *mlh3-D611KD625K::HphMX/"*
VBD2057	same as VBD1311 but *mlh3-K320ED611K::HphMX/"*
VBD2085	same as VBD1311 but *mlh3-K316ED625K::HphMX/"*
VBD2068	same as VBD1311 but *mlh3-K316EK320EE565KD611K::HphMX/"*
VBD2075	same as VBD1311 but *mlh3- K316EK320EE565KD611KD625K::HphMX/"*
VBD2157	*same as VBD1716 but mlh3-F483E-512-8Myc-513::HphMX/"*
VBD2197	*a/l lys2/" ho::LYS2/" leu2::hisG/" trp1::hisG/" ura3/" CEN8/CEN8::tdTomato ARG4/ARG4::Yellow-URA3 MLH3::HphMX/" THR1::Cerulean-TRP1/THR1*
VBD2175	same as VBD2197 but *mlh3Δ::HphMX/"*
VBD2199	same as VBD2197 but *mlh3-F483E::HphMX/"*

Haploids: All haploids are of the W303 background

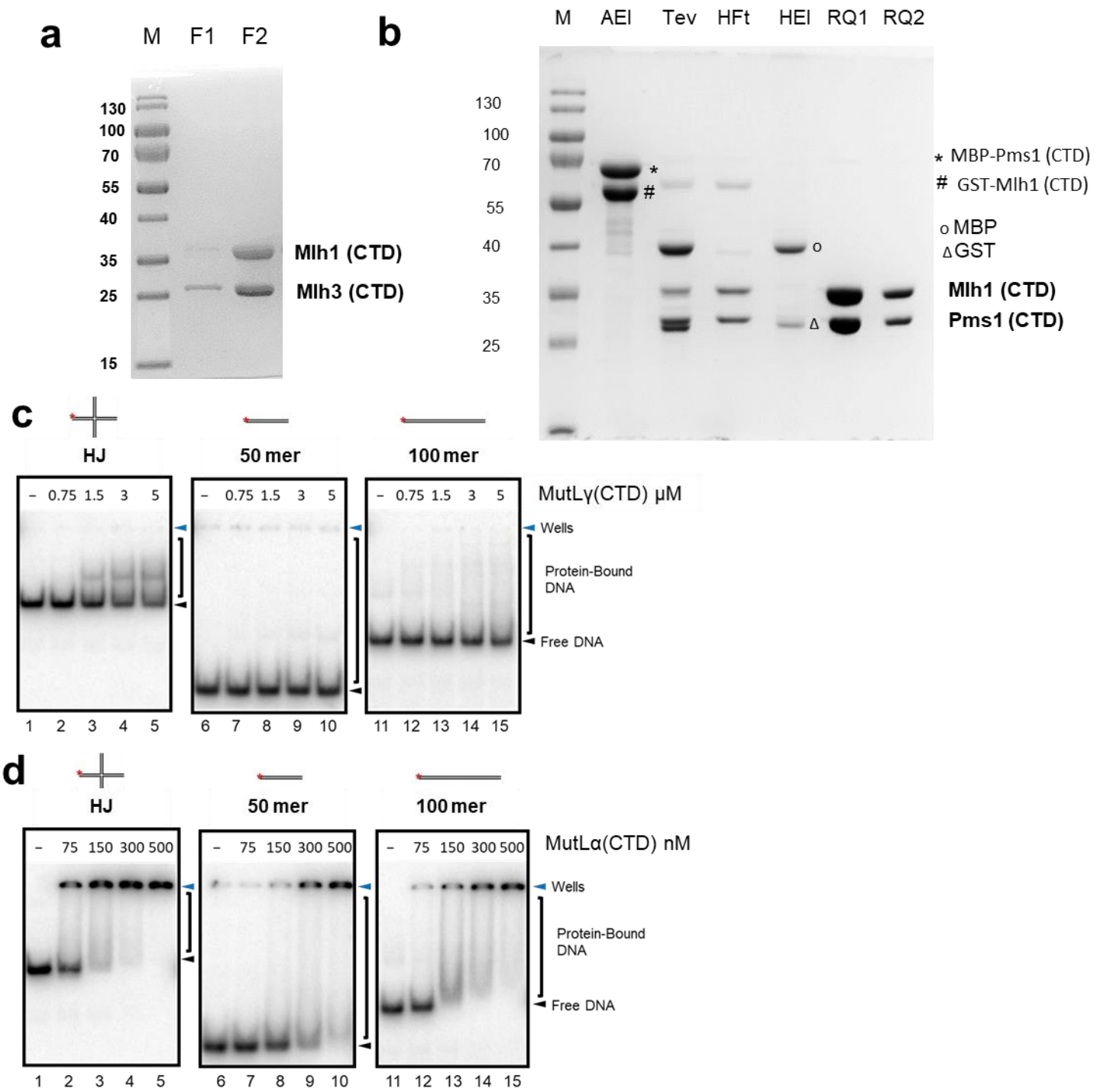
Strain name	Genotype
E134	*MATl ade5-1 lys2::InsE-A14 trp1-289 his7-2 leu2-3,112 ura3-52*
TLM183	same as E134 but *mlh1ΔC1*
TLM187	same as E134 but *mlh1ΔC3*
VBH1874	same as E134 but *mlh3Δ::HphMX*
VBH1871	same as E134 but *mlh3-K668ER671E::HphMX*



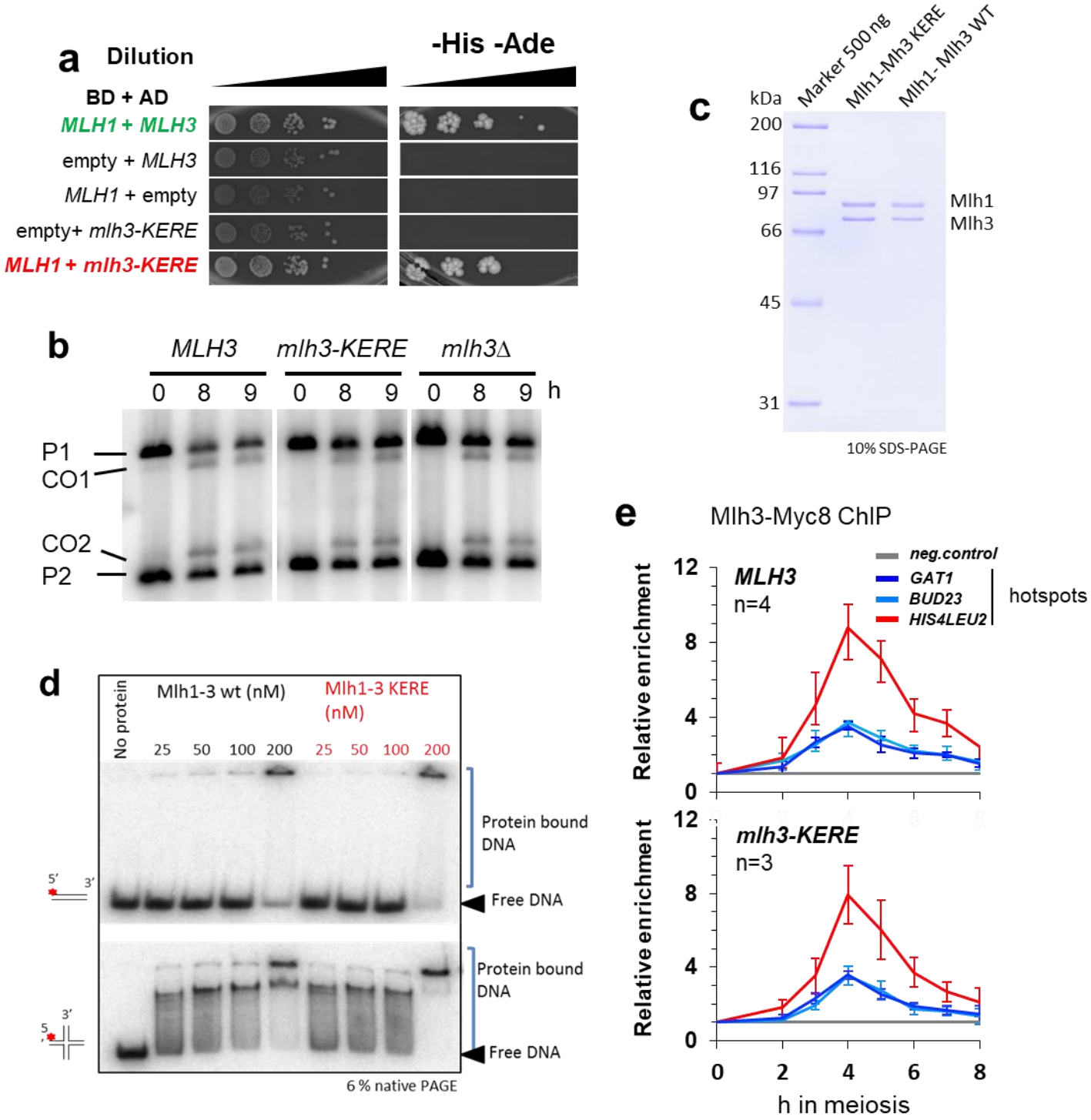
SI Figure S1: (a) Schematic diagram of the S. cerevisiae Mlh1 and Mlh3 proteins. The limits of the CTD correspond to the limits of the folded domains observed in the present crystal structures. The sequence alignments are deduced from alignments of the 3D structures (b) SAXS: the analysis is performed by Crysol, using the model refined by Dadimodo (8). The green and blue lines are the theoretical SAXS curve transformed from the model after refinement by Dadimodo. In green with MutLγ(CTD) structure presented in this study, in blue with the regulatory domain of Mlh3(CTD) in the orientation observed in the crystal structure of MutLα(CTD).The correlation between the pdb structure obtained by crystallography and the experimental SAXS data has a χ2 score of 1.37 (green) and 1.41 (blue). This suggests that both crystal structures are compatible with that in solution with a slightly better fit with MutLγ (CTD). The grey lines correspond to the SAXS experimental data.



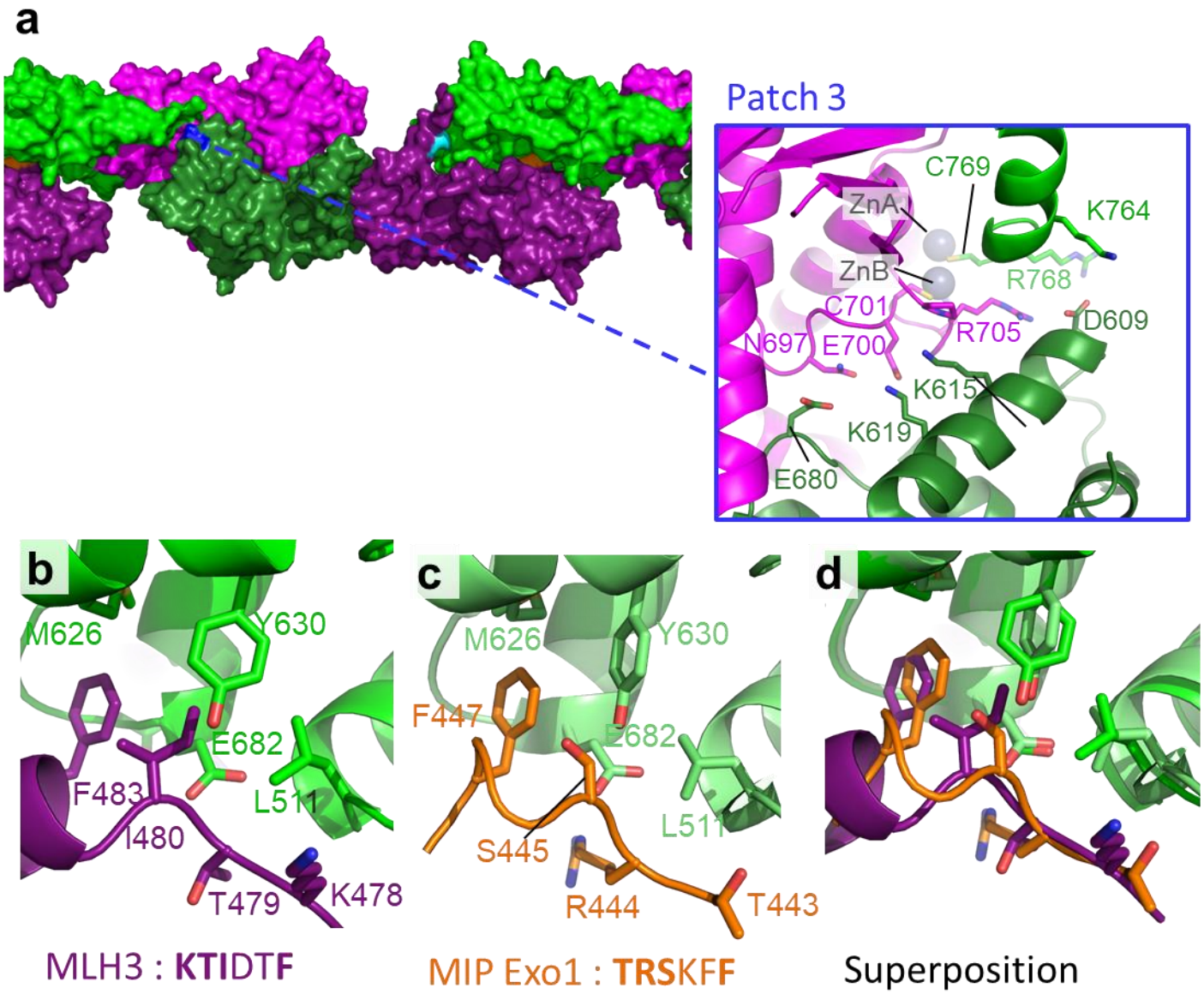
SI Figure S2: Endonuclease site of Mlh1-Mlh3(CTD) observed in the four crystal structures. (a) electron density map 2Fo-Fc (contoured at 2σ) in the active site of the form Zn2 (b) geometry of active site in the form Zn2 (c) geometry of active site in the form Zn2b (P21212). (d) electron densities map 2Fo-Fc (contoured at 2σ) in the active site of the form ZnA (e) geometry of the active site in the form ZnA; (f) geometry of the active site in the form ZnB



SI Figure S3 : (a) SDS Page of purified Mlh1-Mlh3(CTD). F2 is the fraction eluted from Resource Q used in the study. (b) SDS Page of purified Mlh1-Pms1(CTD). The complex 6his-MBP-Pms1(CTD) is co-expressed in E. coli with 6his-GST-Mlh1(CTD) with a TEV site on each subunit between tag and protein. The tagged proteins are first eluted from an amylose resin (AEl). The complex is digested with 6his-TEV protease and passed on a nickel column. Mlh1(CTD)-Pms1(CTD) is not retained. MBP, GST and TEV are retained. Mlh1-Pms1 is then purified on ResourceQ. The same protocol is used for Mlh1(CTD)-Mlh3(CTD).(c) Representative EMSA assays of Mlh1-Mlh3(CTD) with a four way HJ with arms of 25 bps (HJ), with a 50 bps dsDNA (50 mer) or with a 100 bps dsDNA (100 mer). The quantification is presented in Figure 3a. (d) Representative EMSA assays of Mlh1-Pms1(CTD) with the same DNA substrates. The quantification of the gels is reported in Figure 3b.

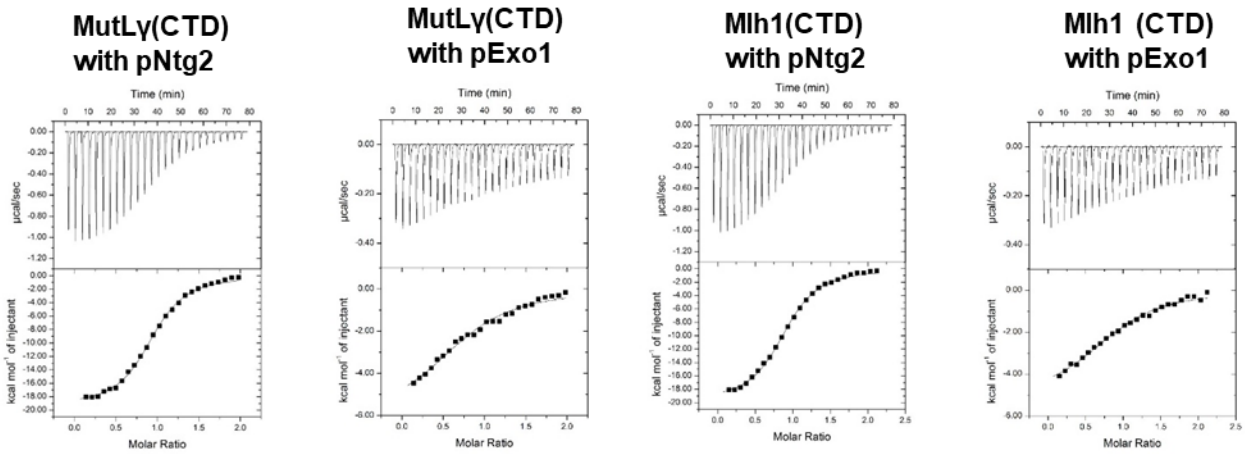


SI Figure S4 : (a) The Mlh3-KERE mutant interacts with Mlh1 in yeast two-hybrid assays. Growth on selective –His-Ade medium indicates an interaction. Serial dilutions of strains expressing the different fusion proteins are plated. (b) Crossing over frequency at the HIS4LEU2 hotspot monitored by Southern blot at the indicated times in meiosis. Positions of parental bands (P1 and P2) and of the recombinant crossover products (CO1 and CO2) are indicated. (c) SDS-PAGE gel of purified full-length Mlh1-Mlh3 wt and KERE mutant heterodimers. (d) Representative EMSA of Mlh1-Mlh3 wt and Mlh1-Mlh3(KERE) mutant with dsDNA and HJ. The quantification of the EMSA is reported Fig. 3h. (e) Mlh3 and Mlh3-KERE in vivo association to meiotic recombination hotspots. Mlh3-Myc levels at the three indicated meiotic DSB hotspots relative to a negative control site assessed by ChIP and qPCR during meiotic time-courses. Values are the mean ± S.E.M. from the indicated number of independent experiments.



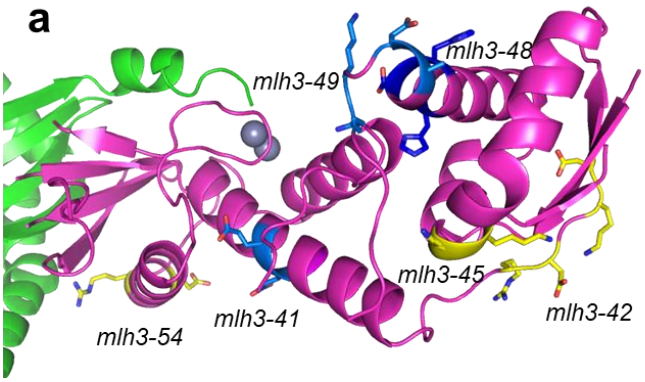
SI Figure 5 : (a) The third patch of the interface between symmetrical Mlh1-Mlh3(CTD) heterodimer is located close to the endonuclease site with the two Zn atoms. It involves charged residues located on helix 610 to 623 of Mlh1 subunit (dark green). This interface involves also contact between two Mlh1 (light and dark green). (b) The first patch of interaction between symmetrical molecules involves the Mlh3 478-KTITDFS-484 region (purple) that interacts with Mlh1 (green) on its MIP-binding site. (c) A crystal structure of Mlh1-Pms1(CTD) (pdb code 4FMO) previously shows the interaction of the MIP motif of EXO1 (orange) with Mlh1. (d) The Mlh3 motif superposes well the EXO1 MIP box on its main chain from residue 478 to 483. Noteworthy, the F483 of Mlh3 motif occupies the same pocket as the second aromatic of the Exo1 motif (F447).

e



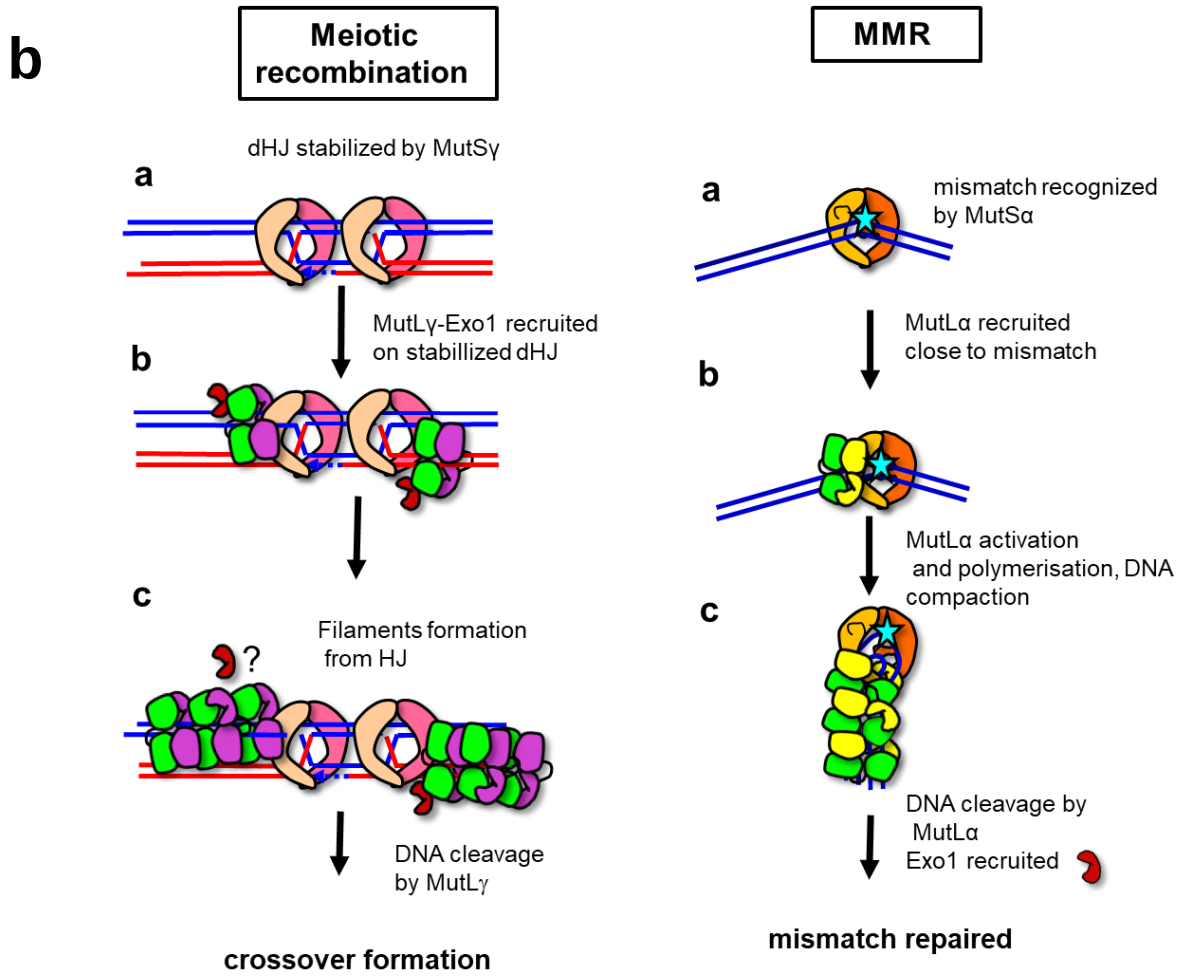
Protein	Peptide	sequence	ΔH° (kcal/mol)	Kd (µM)	ΔG° (kcal/mol)	-TΔS° (kcal/mol)
Mlh1-Mlh3 (CTD)	Ntg2-13	Ac-VEV**RSKYF**KKNER-NH2	- 19.9 ± 0.2	0.7 ± 0.04	- 8.5 ± 0.03	11.3 ± 0.2
	Exo1-12	Ac-KET**RSKFF**NKPS-NH2	- 6.5 ± 0.6	4.2 ± 0.9	- 7.5 ± 0.1	- 0.9 ± 0.8
Mlh1 (CTD)	Ntg2-13	Ac-VEV**RSKYF**KKNER-NH2	- 20.1 ± 0.1	0.9 ± 0.03	- 8.4 ± 0.02	11.7 ± 0.2
	Exo1-12	Ac-KET**RSKFF**NKPS-NH2	- 4.9 ± 0.2	2.9 ± 0.4	- 7.7 ± 0.08	- 2.7 ± 0.3

SI Figure S5 : (e) ITC. Interactions between MutLγ(CTD) or Mlh1(CTD) and the peptides containing the MIP-motif of Ntg2 and Exo1 were determined by ITC using a VP-ITC calorimeter (Malvern) with a protocol similar to the one described in (21). MutLγ(CTD) or Mlh1(CTD) were deposited in the reaction cell of the ITC. Proteins were extensively dialyzed against buffer 50 mM Tris, pH 7.5, 150 mM NaCl, and 10 mM β -mercaptoethanol. The syringe (290 µl) was filled with the different peptides. Enthalpy Δ H (in kcal mol−1), stoichiometry N of the reaction, and association constant Ka (in M−1) were obtained by non-linear least-squares fitting of the experimental data using the single set of independent binding sites model of the Origin software provided with the instrument. The free energy of binding (Δ G) and the entropy (Δ S) were determined using the classical thermodynamic formula, Δ G = − RT ln(Ka) and Δ G = Δ H − TΔ S. All binding experiments were performed in duplicate or triplicate at 20 °C. Control experiments were performed with peptides injected into the buffer to evaluate the heat of the dilution



Alleles with effect on MMR and meiosis
mlh3-41 : R532 E534 E535
mlh3-48 : H622 H624 D625
mlh3-49 : K627 D628 K630 K631

Alleles with effect on MMR only
mlh3-42 : R552 D553 K555 D556
mlh3-45 : K577 K578
mlh3-54 : R682 E684



SI Figure S6: (a) Mlh3 alleles studied for their impact on MMR and meiosis by Al-Sweel et al by mutating patches of charged residues in Ala. In this study, some alleles present major impact on MMR and on meiotic recombination. Some of them include the conserved motifs involved in the metal binding sites. We report on this figure the alleles affected in MMR and meiosis with positions (colored in blue) outside the metal binding sites. The name of the alleles is reported. We also positioned the alleles with a major impact on MMR only (colored in yellow). (22) (b) Figure 6: Model for MutLγ mechanism in meiotic recombination and MMR. In meiotic recombination, MutLγ, in complex with Exo1, is recruited by MutSγ on dHJ intermediates. Transient MutLγ polymerization is proposed to occur at this step leading to an incision of the DNA away from the dHJ. This polymerization is proposed to involve three patches including the MIP-binding site of Mlh1. Further studies should help to evaluate if competition exists between Exo1 and Mlh3 for the Mlh1 MIP-binding site. In MMR, the main nuclease is MutLα and MutLγ has a secondary role. These MutL homologs can be recruited by MutSα or MutSβ on mismatch. MutLα is proposed to form an oligomer and to facilitate DNA compaction (23). These mechanisms also include condensation of the NTD and CTD regions of the MutL homologs upon MutS homologs and DNA binding. The regions of Mlh3 identified in this study as candidate for condensation and oligomerization and the filaments structures observed may thus been conserved with other MutL homologs and have a role in both meiosis and MMR. Figure of MMR model is adapted from (23).

SI References

1. Ranjha L, Anand R, & Cejka P (2014) The Saccharomyces cerevisiae Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. *J Biol Chem* 289(9):5674-5686.
2. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.
3. Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235-242.
4. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
5. Smart OS*, et al.* (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallographica Section D* 68(4):368-380.
6. David G & Pérez J (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr.* 42:892-900.
7. Fiser A, Do RK, & Sali A (2000) Modeling of loops in protein structures. *Protein Sci* 9(9):1753-1773.
8. Rudenko O, Thureau A, & Perez J (2019) Evolutionary refinement of the 3D structure of multi-domain protein complexes from small angle X-ray scattering data. *Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO '19), Manuel López-Ibáñez (Ed.). ACM, New York, NY, USA,* :401-402.
9. Song Y*, et al.* (2013) High-resolution comparative modeling with RosettaCM. *Structure* 21(10):1735-1742.
10. Soding J (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21(7):951-960.
11. Mayrose I, Graur D, Ben-Tal N, & Pupko T (2004) Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol Biol Evol* 21(9):1781-1791.
12. Tran HT, Keen JD, Kricker M, Resnick MA, & Gordenin DA (1997) Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Mol Cell Biol* 17(5):2859-2865.
13. Rothstein R (1991) Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol* 194:281-301.
14. Murakami H, Borde V, Nicolas A, & Keeney S (2009) Gel electrophoresis assays for analyzing DNA double-strand breaks in Saccharomyces cerevisiae at various spatial resolutions. *Methods Mol Biol* 557:117-142.
15. Duroc Y*, et al.* (2017) Concerted action of the MutLbeta heterodimer and Mer3 helicase regulates the global extent of meiotic gene conversion. *Elife* 6.
16. Sanchez A*, et al.* (2020) Exo1 recruits Cdc5 polo kinase to MutLγ to ensure efficient meiotic crossover formation. *Proc Natl Acad Sci U S A* 117(48):30577-30588.
17. Lea DE & Coulson CA (1949) The distribution of the numbers of mutants in bacterial populations. *J Genet* 49(3):264-285.
18. Spell RM & Jinks-Robertson S (2004) Determination of mitotic recombination rates by fluctuation analysis in Saccharomyces cerevisiae. *Methods Mol Biol* 262:3-12.
19. Hunter N & Kleckner N (2001) The single-end invasion: an asymmetric intermediate at the double-strand break to double-holliday junction transition of meiotic recombination. *Cell* 106(1):59-70.
20. Thacker D, Lam I, Knop M, & Keeney S (2011) Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in Saccharomyces cerevisiae. *Genetics* 189(2):423-439.
21. Gueneau E*, et al.* (2013) Structure of the MutLalpha C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. *Nat Struct Mol Biol* 20(4):461-468.
22. Al-Sweel N*, et al.* (2017) mlh3 mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide. *PLoS Genet* 13(8):e1006974.
23. Bradford KC*, et al.* (2020) Dynamic human MutSalpha-MutLalpha complexes compact mismatched DNA. *Proc Natl Acad Sci U S A* 117(28):16302-16312.