OXFORD GENETICS

GENETICS, 2025, 229(3), iyae222
https://doi.org/10.1093/genetics/iyae222
Advance Access Publication Date: 10 January 2025
Investigation

Msh2-Msh3 DNA-binding is not sufficient to promote trinucleotide repeat expansions in *Saccharomyces cerevisiae*

Katherine M. Casazza,[1,†] Gregory M. Williams,[1,2,†] Lauren Johengen,[1] Gavin Twoey,[1] Jennifer A. Surtees [1,*]

[1]Department of Biochemistry, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY 14203, USA
[2]Curia Global, Inc., Buffalo, NY 14203, USA

*Corresponding author: Department of Biochemistry, Jacobs School of Medicine and Biomedical Sciences, Unviversity at Buffalo, 955 Main Street, Buffalo, NY 14203, USA. Email: jsurtees@buffalo.edu
[†]These authors shared first authorship.

Mismatch repair (MMR) is a highly conserved DNA repair pathway that recognizes mispairs that occur spontaneously during DNA replication and coordinates their repair. In *Saccharomyces cerevisiae*, Msh2-Msh3 and Msh2-Msh6 initiate MMR by recognizing and binding insertion or deletion (in/del) loops up to ~17 nucleotides (nt.) and base–base mispairs, respectively; the 2 complexes have overlapping specificity for small (1–2 nt.) in/dels. The DNA-binding specificity for the 2 complexes resides in their respective mispair binding domains (MBDs) and has distinct DNA-binding modes. Msh2-Msh3 also plays a role in promoting *CAG/CTG* trinucleotide repeat (TNR) expansions, which underlie many neurodegenerative diseases such as Huntington's disease and myotonic dystrophy type 1. Models for Msh2-Msh3's role in promoting TNR tract expansion have invoked its specific DNA-binding activity and predict that the TNR structure alters its DNA binding and downstream activities to block repair. Using a chimeric Msh complex that replaces the MBD of Msh6 with the Msh3 MBD, we demonstrate that Msh2-Msh3 DNA-binding activity is not sufficient to promote TNR expansions. We propose a model for Msh2-Msh3-mediated TNR expansions that requires a fully functional Msh2-Msh3 including DNA binding, coordinated ATP binding, and hydrolysis activities and interactions with Mlh complexes that are analogous to those required for MMR.

Keywords: mismatch repair; trinucleotide repeat expansions; MSH3; DNA-binding specificity; DNA repair; *Saccharomyces cerevisiae*

Introduction

Mismatch repair (MMR) is an evolutionarily conserved pathway that recognizes and corrects errors in DNA replication. Two heterodimeric MutS homolog (Msh) complexes initiate MMR through the recognition of distinct DNA structures that arise as a result of nucleotide misincorporation, leading to mispairs, or DNA polymerase slippage events, leading to insertions or deletions (in/dels). Msh2-Msh6, or MutSα, recognizes and binds mispairs and small in/dels [1–2 nucleotides (nt.)] through a conserved Phe-X-Glu motif that intercalates with the mispair, burying it within Msh2-Msh6 (Obmolova *et al.* 2000; Warren *et al.* 2007). Msh2-Msh3, or MutSβ and the focus of this study, recognizes and binds in/dels of up to 17 nucleotides, as well as some mispairs, through a conserved Tyr-Lys pair that interacts with the 5′ double-strand/single-strand DNA junction in loop structures, leaving at least part of the DNA structure accessible (Sia *et al.* 1997; Kunkel and Erie 2005; Lee *et al.* 2007; Dowen *et al.* 2010; Gupta *et al.* 2011).

The mispair binding domains (MBDs) of Msh3 and Msh6 are responsible for Msh2-Msh3 vs Msh2-Msh6 structure-specific DNA-binding activities and contain the highly conserved Phe-X-Glu (Msh6) or Tyr-Lys (Msh3) motifs, as well as other highly conserved residues that contribute to DNA structure specificity (Obmolova *et al.* 2000; Warren *et al.* 2007; Gupta *et al.* 2011). Replacing the Msh6 MBD with the Msh3 MBD in the context of Msh2-Msh6 (Fig. 1) swapped the structure-binding specificity of the resulting Msh2-msh6(3MBD) complex, which exhibited a preference for Msh2-Msh3 in/del loop (IDL) substrates (Shell *et al.* 2007; Brown *et al.* 2016). The ATPase activity of Msh2-msh6(3MBD) was stimulated by Msh2-Msh3 substrates, and this chimeric complex gained Msh2-Msh3's ability to bypass protein blocks, "hopping" over nucleosomes (Brown *et al.* 2016), which Msh2-Msh6 lacks (Gorman *et al.* 2007; Brown *et al.* 2016). These data indicate that Msh2-Msh3's DNA-binding specificity is largely mediated through its MBD.

Once bound, Msh complexes bind ATP and recruit MutL homolog (Mlh) heterodimeric complexes MutLα and/or MutLγ in an ATP-dependent manner. The latent endonuclease activity of the Mlh complexes is activated by Msh and PCNA, nicking the nascent DNA strand (Furman *et al.* 2021; Pannafino and Alani 2021). This is followed by the recruitment of Exo1 and other downstream factors that promote the excision and resynthesis of the nascent DNA to complete repair (Jiricny 2006; Li 2008; Goellner *et al.* 2015; Keogh *et al.* 2017). ATP hydrolysis by Msh2-Msh3 is thought to be important for the turnover of the complex, allowing it to rebind DNA (Kijas *et al.* 2003; Owen *et al.* 2005, 2009; Kumar *et al.* 2014).

Loss of *MSH3* compromises genome stability and leads to an increase in microsatellite instability (Marsischky *et al.* 1996; Sia *et al.* 1997; Harrington and Kolodner 2007; Lee *et al.* 2007; Kumar *et al.*

Received on 08 August 2024; accepted on 25 December 2024

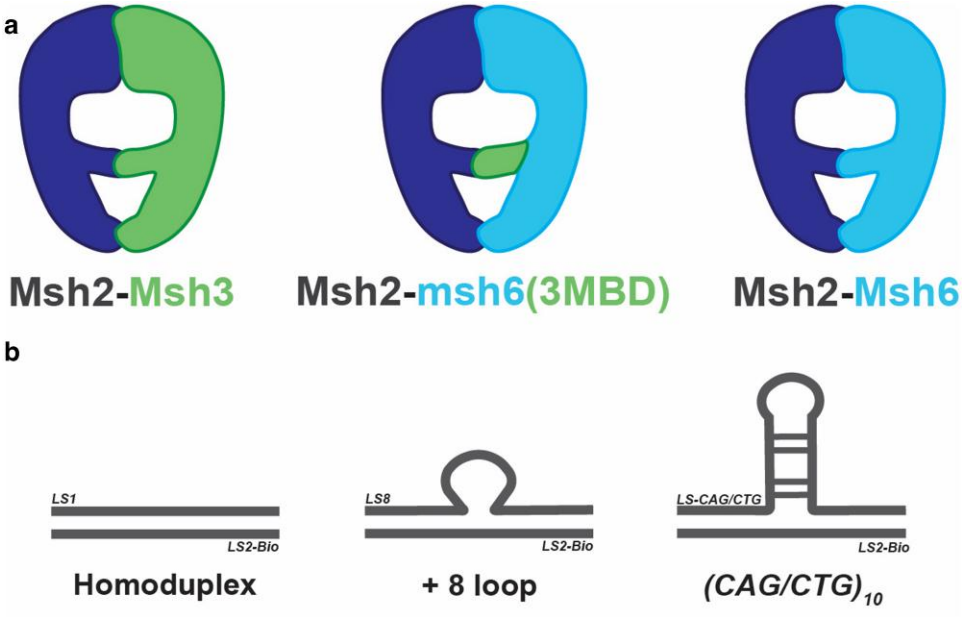

Fig. 1. Schematic of MSH complexes and Msh2-Msh3 DNA substrates. a) Schematic indicating the overall structure of the Msh2-Msh3, Msh2-Msh6, and chimeric MSH complexes. Msh2msh6(3MBD) results from replacing the MBD of Msh6 with that of Msh3, which is sufficient to switch DNA-binding specificities and DNA hopping capacity (Shell *et al.* 2007; Brown *et al.* 2016). b) Schematic depicting the DNA structures used in this study: nonspecific homoduplex DNA, an MMR-specific loop structure with 8 extrahelical nucleotides [$(GT)_4$] and TNR tract slipped structures for *CAG* and *CTG* tracts. The DNA sequences are in Table 1.

2013). Single nucleotide polymorphisms in human *MSH3* have been associated with a predisposition to cancer (Miao *et al.* 2015; De' Angelis *et al.* 2018; Santos *et al.* 2018; Aelvoet *et al.* 2023; Rashid *et al.* 2024). These genome-protective characteristics of *MSH3* are in direct contrast to its pathogenic role in promoting trinucleotide repeat (TNR) expansions (Kantartzis *et al.* 2012; Williams and Surtees 2015; Keogh *et al.* 2017). TNR expansions are the cause of over 40 neurodegenerative and neuromuscular diseases such as Huntington's disease and myotonic dystrophy type 1, which are caused by *CAG* and *CTG* expansions, respectively (Mirkin 2007). *MSH3* promotes TNR expansions, including *CNG* tracts, in multiple model systems including yeast, mice, and human cell culture (Kantartzis *et al.* 2012; Lahue 2020; Iyer and Pluciennik 2021; Richard 2021; Matos-Rodrigues *et al.* 2023). In genome-wide association studies, *Msh3* was identified as a genetic modifier of TNR expansions; polymorphisms identified in *Msh3* that correlated with higher *Msh3* expression exhibited increased TNR tract instability in mice (Tomé *et al.* 2013; Bettencourt *et al.* 2016; Lee *et al.* 2019). In contrast, *MSH6* does not promote significant *CNG* expansions (Kantartzis *et al.* 2012; Iyer and Pluciennik 2021), indicating that MMR-dependent TNR expansions are specific to the Msh2-Msh3-mediated pathway, leading us to focus on the molecular activities of Msh2-Msh3 that are necessary for promoting TNR expansion.

Msh2-Msh3 recognizes and binds the slipped strand secondary structures thought to form during replication of TNR tracts, with affinities similar to in/dels (Owen *et al.* 2005; Tian *et al.* 2009; Lang *et al.* 2011). However, the execution of repair is altered, leading to TNR tract expansions rather than repair pathways that maintain tract length. While Msh2-Msh3 ATPase activity is likely required for TNR expansions (Tomé *et al.* 2009; Keogh *et al.* 2017), TNR DNA structures decreased Msh2-Msh3 ATP binding and hydrolysis activities compared to that observed with an in/del MMR substrate (Owen *et al.* 2005, 2009). Similarly, Msh2-Msh3 nucleotide-binding and hydrolysis activities are differentially modified in the presence of DNA structures it binds in double-strand break repair vs MMR (Kumar *et al.* 2014). These observations led to the prediction that Msh2-Msh3 becomes "trapped" upon binding TNR structures, preventing the repair of these slipped strand secondary structures (Owen *et al.* 2005; Lang *et al.* 2011). Our in vitro work has suggested that at least part of Msh2-Msh3's role in promoting TNR expansions is in stabilizing the TNR structures (Kantartzis *et al.* 2012), consistent with other studies (Tian *et al.* 2009). In this study, we sought to determine whether Msh2-Msh3's specific DNA-binding activity, through the Msh3 MBD, is sufficient to promote TNR expansions.

Materials and methods

Protein purification

Msh2-Msh3 was overexpressed in *Saccharomyces cerevisiae* and purified as previously described (Kumar *et al.* 2014). Msh2-msh6(3MBD) and Msh2-Msh6 were overexpressed in *Escherichia coli* and purified as previously described (Brown *et al.* 2016).

DNA substrates

DNA substrates were constructed with synthetic oligonucleotides (Table 1). One oligonucleotide in each substrate was biotinylated for attachment to streptavidin tips (see below). Oligonucleotides were mixed at equimolar concentrations in 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA, heated to 95°C for 5 minutes and allowed to cool slowly to room temperature.

Biolayer interferometry

Biolayer interferometry (BLI) experiments were performed with purified Msh2-Msh3, Msh2-Msh6, and Msh2-msh6(3MBD) using a Sartorius Octet Red96e. DNA concentration was optimized for each protein complex in order to have ideal ligand density across the streptavidin sensor for each protein complex. A titration of DNA concentrations was loaded onto the sensor. Protein association was measured for each protein at a concentration 10–20

Table 1. DNA substrates.

DNA structure	Top strand (5′–3′)	Bottom strand (5′–3′) (LS2-Bio)
Homoduplex	cacgctaccgaattctgacttgctaggcatctttgcccacgttgaccc (LS1)	Bio/gggtcaacgtgggcaaagatgtcctagcaagtcag aattcggtagcgtg
+8 loop	cacgctaccgaattctgacttgctaggtgtgtgtgacatctttgccca cgttgaccc (LS8)	Bio/gggtcaacgtgggcaaagatgtcctagcaagtcag aattcggtagcgtg
$(CAG)_{10}$	cacgctaccgaattctgacttgctagcagcagcagcagcagcagc agcagcagcaggacatctttgcccacgttgaccc (LS-CAG)	Bio/gggtcaacgtgggcaaagatgtcctagcaagtcag aattcggtagcgtg
$(CTG)_{10}$	cacgctaccgaattctgacttgctagctgctgctgctgctgctgct gctgctggacatctttgcccacgttgaccc (LS-CTG)	Bio/gggtcaacgtgggcaaagatgtcctagcaagtcag aattcggtagcgtg

times higher than its predicted K_D, (Surtees and Alani 2006; Shell *et al.* 2007; Kumar *et al.* 2011; Brown *et al.* 2016). The ligand (DNA) concentration chosen was the lowest DNA concentration that exhibited strong binding signal without reaching saturation and fit well to a 1:1 model. The DNA substrate concentration selected (7.5 nM) was the same for Msh2-Msh3 and Msh2-mh6(3MBD). The Msh2-Msh6 experiments required a higher DNA concentration (30 nM); the association curves for Msh2-Msh6 displayed weak binding signal at 7.5 nM, which is consistent with improper ligand density, as recommended by the Sartorius Octet application manual.

BLI experiments were performed in binding buffer (25 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl, 0.1 mg/mL BSA, 1 mM DTT, and 0.05% Tween-20) using the Octet Red96e system (ForteBio). Biotinylated DNA substrates were immobilized to pre-hydrated streptavidin biosensors. Biosensors were then incubated with 50 uM biocytin in SuperBlock buffer (Thermo Fisher) to quench unbound streptavidin and equilibrated in binding buffer. Biosensors were incubated with Msh2-Msh3 or Msh2-msh6(3MBD) (250, 125, 62.5, 31.25, 15.6, and 7.8 nM) or with Msh2-Msh6 (150, 75, 37.5, 18.75, 9.4, and 4.7 nM) for 180 s. Biosensors were then moved to wells containing binding buffer to allow for dissociation for 180 s. Data were collected using Data Acquisition 12.0 software (ForteBio) and analyzed with Data Analysis HT 12.0 software (ForteBio). Data were fit to 1:1 model to obtain binding kinetics. Experiments were performed in triplicate with protein obtained from at least 2 independent purifications.

Strains and Media

All strains and plasmids used are listed in Tables 2 and 3. All yeast transformations were performed using the lithium acetate method (Gietz *et al.* 1992). Yeast strains were derived from the S288c background: FY23 for slippage assays and FY86 for TNR expansion assays (Winston *et al.* 1995). *msh3Δ* and *msh6Δ* were constructed by amplifying a chromosomal mutant::KANMX fragment from the yeast deletion collection that was integrated into the respective chromosomal location. *msh6(3MBD)* was integrated into the endogenous MSH6 locus in a *msh3Δ* strain using pRDK4576 as described previously (Shell *et al.* 2007).

TNR substrates were integrated into all strains as described previously (Dixon *et al.* 2004; Williams and Surtees 2018): (1) $(CAG)_{25}$ (pBL70), (2) $(CTG)_{25}$ (pBL69), (3) scrambled $(C,A,G)_{25}$ (pBL139), and (4) scrambled $(C,T,G)_{25}$ (pBL138) (Table 3) (Dixon *et al.* 2004). The repeat tracts were in the promoter of the *URA3* reporter gene; expansions of ≥4 repeats alter the transcriptional start site, making the cells resistant to 5-FOA.

The microsatellite instability constructs have been described previously (Table 3) (Sia *et al.* 1997). Repeat tracts of 1 nt. $[(G)_{18}]$, 2 nt. $[(GT)_{16.5}]$, or 4 nt. $[(CAGT)_{16}]$ were placed in-frame upstream of *URA3*. Unrepaired slippage events shift *URA3* out of frame, resulting in 5-FOA resistance.

Western blot

We performed western blots of titrations of purified Msh2-Msh6 and Msh2-msh6(3MBD) to test the sensitivity of the anti-Msh6 antibody. Titrations of purified protein were loaded onto a 6% SDS-PAGE gel. The gel was transferred to nitrocellulose using the Bio-Rad Trans-Blot Turbo Transfer system. The blot was blocked overnight at 4°C and then incubated for 2 h in 1:2,000 dilution of anti-Msh6 primary rabbit antibody (Kumar *et al.* 2011). It was then incubated for 1 h in 1:4,000 dilution of HRP-conjugated anti-rabbit secondary goat antibody. The blot was then imaged using SuperSignal West Atto Ultimate Sensitivity ECL substrate (Thermo Fisher) and imaged using ChemiDoc MP imaging system (Bio-Rad).

To compare endogenous Msh6 and msh6(3MBD) levels, wild type (FY23 and FY86), *msh6(3MBD) msh3Δ* (in the FY23 and FY86 background), and *msh6Δ* (FY23 background) were grown in 1L YPD cultures until $OD_{600} = 0.6$. Cells were harvested and resuspended in MSH buffer [25 mM Tris–HCl (pH 7.5), 1 mM EDTA, and 175 mM NaCl]. The resuspension was then frozen by dropping into liquid nitrogen, and the frozen cell resuspension was ground with dry ice to lyse. The lysate was then thawed and β-mercaptoethanol and PMSF were added to final concentrations of 10 and 1 mM, respectively. The resuspension was then centrifuged and the cleared lysate was retained. The cleared lysate was concentrated using Amicon Ultra Centrifugal Filter Unit 100 kDa cutoff. Total protein concentration was measured by Bradford assay.

Four hundred micrograms of each lysate as well as *msh6Δ* lysate +0.15 μg of purified Msh2-Msh6 were incubated with anti-Msh6 primary rabbit antibody at 4°C for 1 h while rocking. Fifty microliters of 50% slurry Protein A/G-Agarose beads (Pierce) were added to each and incubated overnight at 4°C while rocking. The beads were incubated on ice for 10 min and then collected by centrifugation and the supernatant was removed. The beads were washed 3 times with IP buffer (50 mM HEPES, pH 7.5, 1 mM EDTA, and 1% Triton-X). Beads were resuspended in 1× Laemmli buffer and heated for 10 min at 95°C. The immunoprecipitated product was resolved in a 4–20% gradient gel (Bio-Rad) and transferred to nitrocellulose using the Bio-Rad Trans-blot Turbo Transfer system. The blot was blocked for 1 h at 4°C and then incubated overnight in 1:1,000 dilution of anti-Msh6 primary rabbit antibody. It was then incubated for 1 h in 1:4,000 dilution of HRP-conjugated anti-rabbit secondary goat antibody. The blot was then imaged using SuperSignal West Atto Ultimate Sensitivity ECL substrate (Thermo Fisher) and imaged using ChemiDoc MP imaging system (Bio-Rad).

Table 2. Strains used in this study.

JSY#	Other name	Relevant genotype	Strain background	Source
126	FY23	*MATa ura3-52 leu2Δ1 trp1Δ63 his3Δ200 lys2Δ202 ura3Δ leu2Δ trp1Δ*		Winston *et al.* (1995)
127	FY86	*MATα ura3-52 leu2Δ1 trp1Δ63 his3Δ200 lys2Δ202 ura3Δ leu2Δ his3Δ*		Winston *et al.* (1995)
485		*msh6Δ::KANMX*	FY86	Kumar *et al.* (2011)
1472		*msh3Δ::KANMX*	FY86	Kantartzis *et al.* (2012)
314	EAY337	*msh6::hisG*	FY23	Bowers *et al.* (1999)
905	EAY420	*msh3Δ::hisG*	FY23	Studamire *et al.* (1998)
3925, 3926, 3927		*msh3Δ::hisG msh6(3MBD)*	FY23	This study
3554, 3555, 3556		*msh3Δ::hisG msh6(3MBD)*	FY86	This study

Table 3. Plasmids used in this study.

JSB#	Plasmid name	Plasmid description	Source
94	pBK1	$(CAGT)_{16}$-*URA3* slippage assay reporter	Sia *et al.* (1997)
95	pMD28	$(G)_{18}$-*URA3* slippage assay reporter	Sia *et al.* (1997)
299	pSH44	$(GT)_{16.5}$-*URA3* slippage assay reporter	Henderson and Petes (1992)
361	pBL139	$(C,A,G)_{25}$::*URA3*::*HIS3* integration plasmid	Miret *et al.* (1998)
363	pBL69	$(CTG)_{25}$::*URA3*::*HIS3* integration plasmid	Miret *et al.* (1998)
364	pBL70	$(CAG)_{25}$::*URA3*::*HIS3* integration plasmid	Miret *et al.* (1998)
365	pBL138	$(C,T,G)_{25}$::*URA3*::*HIS3* integration plasmid	Miret *et al.* (1998)

Microsatellite instability assay

Microsatellite instability assays were performed as described (Sia *et al.* 1997). Single colonies were grown on SC-tryptophan (SC-trp) to maintain the reporter plasmids. Colonies of ~2 mm were selected from ≥3 independent isolates of each genetic background. Individual colonies were resuspended in 3 mL of liquid SC-trp and incubated with shaking for 20 h at 30°C. Overnight cultures were serial diluted and plated on permissive (SC-trp) and selective (SC-trp +5-FOA) plates. Plates were incubated at 30°C for 2–4 days. Mutation rates were calculated by the method of the median (Drake 1991).

95% confidence intervals were determined using tables of confidence intervals for the median (Nair 1940; Dixon and Massey 1969). *P*-values were determined by Mann–Whitney rank analysis in GraphPad Prism.

Assay for TNR expansion

TNR expansion assays were performed as described (Williams and Surtees 2018). Briefly, single colonies were obtained on synthetic medium (SC) lacking histidine (SC-his) or lacking histidine and uracil (SC-his-ura) for *msh6(3MBD) msh3Δ*. Individual colonies were selected from ≥3 independent isolates of each genetic background and were assayed. Colonies with unexpanded tracts, which were confirmed by PCR analysis, were diluted and plated on SC-his and incubated at 30°C for 3–4 days to allow expansions to occur. Several ~2 mm colonies were selected, diluted and plated onto permissive (SC-his) and selective (SC-his +5-FOA), and incubated at 30°C from 3–4 days. Colonies were counted and expansion rates were calculated as described (Drake 1991). The 95% confidence intervals were determined from tables of confidence intervals for the median (Nair 1940; Dixon and Massey 1969). *P*-values were determined by Mann–Whitney rank analysis (Nair 1940; Dixon and Massey 1969; Drake 1991; Sia *et al.* 1997) in GraphPad Prism.

True expansions were determined as previously described (Williams and Surtees 2015, 2018; Williams *et al.* 2020), by amplifying the reporter promoter region with SO295 (AAACTCGG TTTGACGCCTCCCATG) and SO296 (AGCAACAGGACTAGGATG AGTAGC) and digesting with *SphI* to release the TNR tract. Tract mobility was assessed by electrophoresis through a 12% native polyacrylamide gel (0.5× TBE). At least 30 independent 5-FOA-resistant colonies were characterized for each tract and genotype combination. Sequencing of the *URA3* gene was performed to identify mutations in *URA3* that may lead to 5-FOA resistance. The *URA3* coding region was amplified from genomic preps derived from 5-FOA-resistant colonies that did not exhibit expanded TNR tracts, using SO295 and SO1079 (GTTAGAAGTGCGGTTGATGTCG). This amplicon included the *URA3* and TNR tract sequences, to confirm the tract size in conjunction with the *URA3* sequence. We also amplified the promoter region of this reporter construct using SO296 and SO1080 (GGGAACAAAAGCTGGTACCGGG). The amplified regions were sent for Oxford Nanopore sequencing (Plasmidsaurus).

Results

Msh2-Msh3 and Msh2-msh6(3MBD) exhibit specific binding to TNR structures in vitro

Previous studies, including our own, have indicated that Msh2-Msh3 TNR slipped strand structure binding and stabilization is required to promote TNR expansions (Owen *et al.* 2005; Panigrahi *et al.* 2005; Kantartzis *et al.* 2012; Williams and Surtees 2015). In this study, we set out to test whether the structure-specific binding activity of Msh2-Msh3 is *sufficient* to promote TNR expansions, using the msh6(3MBD) chimeric protein (Fig. 1). Before testing Msh2-msh6(3MBD) activities in vivo, we characterized its in vitro DNA-binding properties, using both MMR and TNR DNA substrates (Fig. 1) (Surtees and Alani 2006). We previously demonstrated yeast Msh2-Msh3 specificity for +8 loop DNA structures, substrates for Msh2-Msh3-mediated MMR, using electrophoretic mobility assays (EMSAs) and DNA footprinting (Surtees and Alani 2006; Lee *et al.* 2007; Brown *et al.* 2016). Human Msh2-Msh3 was similarly demonstrated to have specificity for DNA loop structures (Habraken *et al.* 1996; Wilson *et al.* 1999; Owen *et al.* 2005, 2009; Lang *et al.* 2011). Here, we used BLI to characterize yeast Msh2-Msh3 and Msh2-msh6(3MBD) binding to homoduplex DNA, a $(GT)_4$ loop (+8 loop) (Surtees and Alani 2006), $(CAG)_{10}$ or $(CTG)_{10}$ hairpin DNA structures, using biotinylated DNA structures assembled from synthetic oligonucleotides (Surtees and Alani 2006).

Msh2-Msh3 exhibited significant nonspecific binding activity to homoduplex and exhibited ~3- to 4-fold increased affinity for the

Table 4. Binding kinetics of Msh2-Msh3 and Msh2-msh6(3MBD).

Msh2-Msh3						
Substrate	**K_D (M)**	**K_D error**	**k_a (1/Ms)**	**k_a error**	**k_{dis} (1/s)**	**k_{dis} error**
Homoduplex	9.22×10^{-9}	1.42×10^{-10}	3.33×10^{5}	2.20×10^{3}	3.07×10^{-3}	4.26×10^{-5}
+8 Loop	2.62×10^{-9}	5.28×10^{-11}	6.57×10^{5}	3.46×10^{3}	1.72×10^{-3}	3.35×10^{-5}
$(CAG)_{10}$	1.38×10^{-9}	4.84×10^{-11}	8.69×10^{5}	5.63×10^{3}	1.20×10^{-3}	4.13×10^{-5}
$(CTG)_{10}$	2.25×10^{-9}	4.87×10^{-11}	6.96×10^{5}	3.68×10^{3}	1.57×10^{-3}	3.29×10^{-5}
Msh2-msh6(3MBD)						
Substrate	**K_D (M)**	**K_D error**	**k_a (1/Ms)**	**k_a error**	**k_{dis} (1/s)**	**k_{dis} error**
Homoduplex	7.16×10^{-9}	6.81×10^{-11}	1.80×10^{6}	1.51×10^{4}	1.29×10^{-2}	5.83×10^{-5}
+8 Loop	8.72×10^{-10}	9.52×10^{-12}	2.32×10^{6}	1.01×10^{4}	2.03×10^{-3}	2.03×10^{-5}
$(CAG)_{10}$	2.37×10^{-9}	2.02×10^{-11}	2.58×10^{6}	1.53×10^{4}	6.10×10^{-3}	3.72×10^{-5}
$(CTG)_{10}$	1.05×10^{-9}	1.42×10^{-11}	1.94×10^{6}	8.34×10^{3}	2.03×10^{-3}	2.60×10^{-5}

Table 5. Binding kinetics of Msh2-Msh6.

Msh2-Msh6						
Substrate	**K_D (M)**	**K_D error**	**k_a (1/Ms)**	**k_a error**	**k_{dis} (1/s)**	**k_{dis} error**
Homoduplex	2.10×10^{-8}	2.62×10^{-10}	2.63×10^{5}	2.15×10^{3}	5.54×10^{-3}	5.21×10^{-5}
+8 Loop	8.67×10^{-9}	1.27×10^{-10}	4.55×10^{5}	3.35×10^{3}	3.95×10^{-3}	5.00×10^{-5}
$(CAG)_{10}$	2.73×10^{-9}	1.82×10^{-11}	2.22×10^{6}	1.01×10^{4}	6.06×10^{-3}	2.96×10^{-5}
$(CTG)_{10}$	3.07×10^{-9}	2.58×10^{-11}	1.32×10^{6}	5.93×10^{3}	4.06×10^{-3}	2.89×10^{-5}

+8 loop (Table 4; K_D). The relative affinities are consistent with our previous estimates using EMSA, although the current K_D's are ~20-fold lower (Surtees and Alani 2006). The association rate (Table 4; k_a) was similar for both substrates; the dissociation rate with the +8 loop (Table 4; k_{dis}) was slower. Msh2-Msh3 also bound preferentially to $(CTG)_{10}$ or $(CAG)_{10}$ quasi-hairpin structures with affinities ~4- and 7-fold higher than to homoduplex, respectively. This is consistent with human Msh2-Msh3, which exhibited binding to CAG tracts similar to +8 loop structures (Owen *et al.* 2005), although we note that the affinities differ. The higher affinities of Msh2-Msh3 to all 3 DNA structures are driven primarily by decreased k_{dis}, although there was also an ~2-fold increase in k_a with specific DNA substrates.

We next tested the DNA-binding kinetics of Msh2-msh6(3MBD), which replaces the Msh6 MBD with Msh3 MBD within the context of Msh2-Msh6 (Fig. 1) (Surtees and Alani 2006; Shell *et al.* 2007; Brown *et al.* 2016). Msh2-msh6(3MBD) displayed a higher (~8-fold) affinity (Table 4; K_D) for the +8 loop structure compared to the homoduplex DNA structure, consistent with previous observations (Shell *et al.* 2007; Brown *et al.* 2016) and with Msh2-Msh3 binding (Table 4) (Surtees and Alani 2006). Msh2-msh6(3MBD) also exhibited increased affinity for both TNR hairpin structures, relative to homoduplex, with affinities similar to its affinity for the +8 loop substrate (Table 4; K_D), as observed with Msh2-Msh3 (Table 4). Msh2-msh6(3MBD) exhibited similar k_a's for all of the substrates, which were somewhat higher than Msh2-Msh3 k_a's for the same substrates. As with Msh2-Msh3, the increased affinity to specific DNA structures was largely driven by decreased dissociation rates (Table 4; k_{dis}). These data indicate that Msh2-Msh3 and Msh2-msh6(3MBD) exhibit similar DNA-binding affinities and that Msh2-msh6(3MBD) binds TNR structures with an affinity similar to its affinity for MMR (+8 loop) structures and is therefore competent to recognize these structures, should they form in vivo.

We also examined binding kinetics of Msh2-Msh6 to the same DNA structures. Msh2-Msh6 displays distinct binding behavior compared to Msh2-Msh3 and Msh2-msh6(3MBD) (Shell *et al.* 2007; Brown *et al.* 2016), requiring reoptimization of BLI conditions to measure the binding kinetics. Msh2-Msh6 displayed a ~2- to 3-fold higher K_D with homoduplex (Table 5), compared to Msh2-Msh3 and Msh2-msh6(3MBD). Msh2-Msh6 exhibited a modest (2-fold) preference for the +8 loop compared to the homoduplex. We have previously shown that this interaction is significantly less stable than Msh2-Msh6 binding to a mispair or +1 insertion (Jiang *et al.* 2005; Brown *et al.* 2016). Interestingly, we observed more significantly increased affinity of Msh2-Msh6 for TNR structures, ~7-fold over homoduplex. Therefore, Msh2-Msh3, Msh2-msh6(3MBD), and Msh2-Msh6 exhibit similar specificity for the TNR quasi-hairpin. However, *MSH6* does not promote TNR expansions (Kantartzis *et al.* 2012), indicating that TNR binding is not sufficient for promoting expansions.

Msh6(3MBD) retains some Msh2-Msh3-specific MMR function

To directly test its activity to promote Msh2-Msh3-mediated activities in vivo, we integrated the *msh6(msh3MBD)* construct (Shell *et al.* 2007) into the chromosome, replacing endogenous *MSH6*, in a *msh3Δ* background. This results in a single MSH complex in vivo, Msh2-msh6(3MBD) (Fig. 1). Previous work has demonstrated that *msh6(msh3MBD) msh3Δ* has in vivo MMR activity, indicating that the complex is functional, although both Msh2-Msh6- and Msh2-Msh3-mediated MMRs were compromised to some extent (Shell *et al.* 2007). We performed western blots to compare Msh6 and msh6(3MBD) protein levels in these strains in vivo. We first tested the sensitivity of our polyclonal anti-Msh6 antibody (Kumar *et al.* 2011) for Msh6 compared to msh6(3MBD), using purified proteins. The antibody recognized both proteins with comparable sensitivity (Fig. 2a). From cleared cell lysates, we

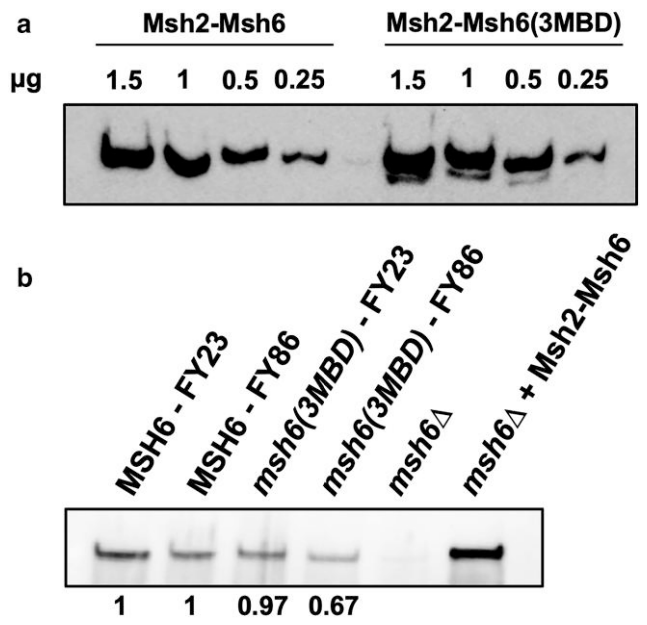

Fig. 2. Western blot detecting endogenous Msh6 and Msh6(3MBD). a) Western blot of titration of purified Msh2-Msh6 and Msh2-msh6(3MBD). b) Western blot of immunoprecipitated endogenous Msh6 and msh6(3MBD). Purified Msh2-Msh6 (0.15 µg) was added to *msh6Δ* lysate as a positive control. Msh6/msh6(3MBD) bands were quantified and normalized to a background band (not shown). Normalized relative values of msh6(3MBD) were calculated and compared to Msh6 in their respective backgrounds and are shown below each lane.

immunoprecipitated and detected both Msh6 and msh6(3MBD) (Fig. 2b). Levels of Msh6 and msh6(3MBD) were normalized to a background band and quantified to compare protein levels in FY23 and FY86, the backgrounds used for slippage assays and TNR expansions assays, respectively. msh6(3MBD) levels were comparable to Msh6, although it was slightly reduced, ~70% of Msh6 in the FY86 background. These data indicate that the chimeric complex is stable in vivo (Fig. 2b).

msh6(msh3MBD) msh3Δ function was previously tested with reporter assays that select for −1 frameshift deletion or +2 frameshift insertions (Shell *et al.* 2007), substrates for both Msh2-Msh3 and Msh2-Msh6 (Marsischky *et al.* 1996; Sia *et al.* 1997). *msh6(3MBD) msh3Δ* elevated mutation rates (30- to 70-fold increases), but not the synergistic >1,000-fold increase observed in *msh3Δ msh6Δ* (Shell *et al.* 2007), indicating that *msh6(3MBD)* retained significant MMR function. Here, we tested whether the chimeric protein would act in repair of larger IDLs that is dependent almost exclusively on Msh2-Msh3-mediated repair (Sia *et al.* 1997; Lee *et al.* 2007; Kumar *et al.* 2013). We used a microsatellite instability reporter plasmid that places a tetranucleotide $(CAGT)_{16}$ repeat in-frame upstream of *URA3* (Sia *et al.* 1997). Unrepaired DNA slippage events alter the *URA3* reading frame, resulting in 5-FOA resistance, which allows the selection of these slippage events. Previous work demonstrated that these slippage events result primarily in deletions (Sia *et al.* 1997; Lamb *et al.* 2022). Compared to wild type, *msh3Δ* increased the 4 nt. slippage rate by 58-fold; *msh6Δ* exhibited only a 3-fold increase, consistent with its limited role in repair of these longer in/dels (Fig. 3; blue

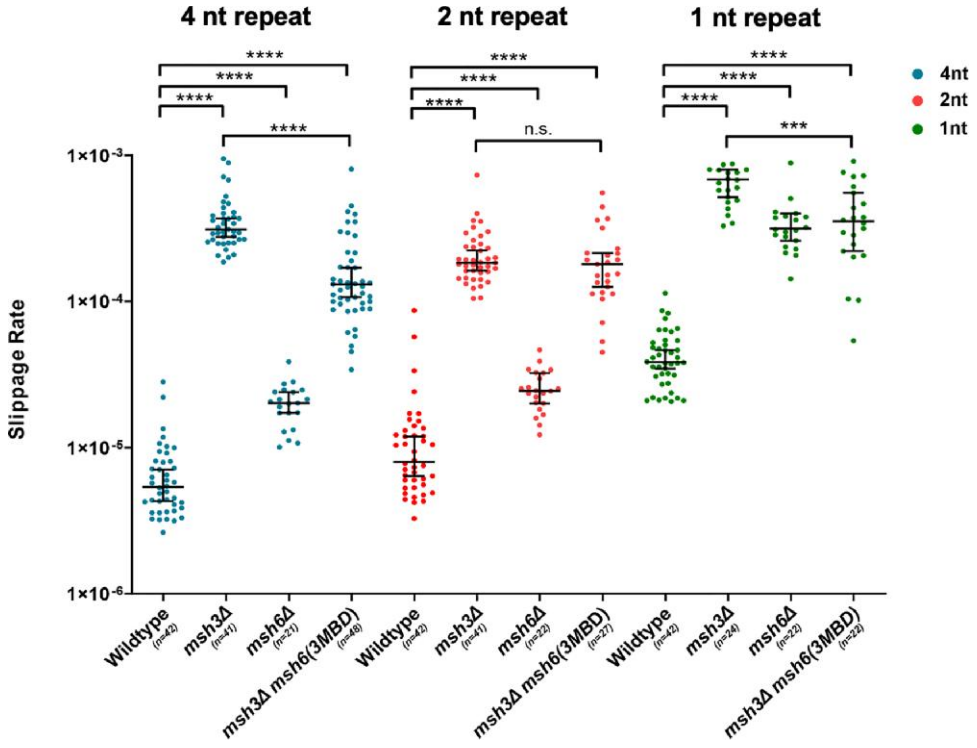

Fig. 3. Mutation rate in a slippage mutation rate assay. The rate of slippage events, which push *URA3* out of frame and allow selection in the presence of 5-FOA, was determined in different genetic backgrounds, as indicated. Slippage rate in the presence of mononucleotide (right panel; green), dinucleotide (middle panel; red), and tetranucleotide (left panel; blue) repeats was determined. Median rates (95% confidence interval) are as follows: 4 nt. (blue): wild type: 5.38×10^{-6} (4.30×10^{-6} to 7.08×10^{-6}), *msh3Δ*: 3.12×10^{-4} (2.77×10^{-4} to 3.71×10^{-4}), *msh6Δ*: 2.01×10^{-5} (1.73×10^{-5} to 2.4×10^{-5}), and *msh6(3MBD) msh3Δ*: 1.29×10^{-4} (1.07×10^{-4} to 1.70×10^{-4}); 2 nt. (red): wild type: 7.99×10^{-6} (6.14×10^{-6} to 1.20×10^{-5}), *msh3Δ*: 1.84×10^{-4} (1.63×10^{-4} to 2.23×10^{-4}), *msh6Δ*: 2.45×10^{-5} (2.01×10^{-5} to 3.24×10^{-5}), and *msh6(3MBD) msh3Δ*: 1.80×10^{-4} (1.26×10^{-4} to 2.14×10^{-4}); and 1 nt. (green): wild type: 3.84×10^{-5} (3.47×10^{-5} to 4.64×10^{-5}), *msh3Δ*: 6.85×10^{-4} (5.18×10^{-4} to 8.02×10^{-4}), *msh6Δ*: 3.17×10^{-4} (2.61×10^{-4} to 4.00×10^{-4}), and *msh6(3MBD) msh3Δ*: 3.53×10^{-4} (2.21×10^{-4} to 5.55×10^{-4}). Error bars indicate 95% confidence intervals. *P*-values were calculated using Mann–Whitney test: ****$P < 0.0001$; ***$0.0001 < P < 0.001$.

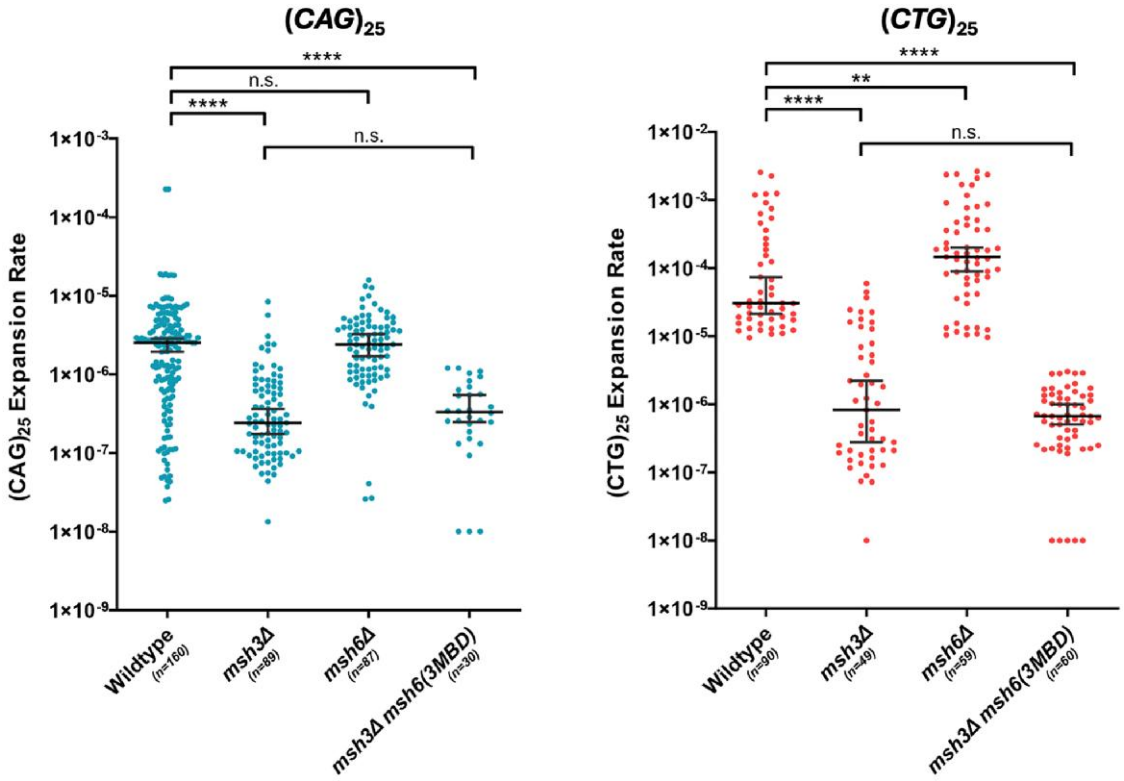

Fig. 4. TNR expansion rates in vivo. The rate of expansion events, which prevent expression of URA3 and allow selection in the presence of 5-FOA, was determined in different genetic backgrounds, as indicated. Expansion rates of $(CAG)_{25}$ tracts (left; blue) and $(CTG)_{25}$ tracts (right; red) were measured. Error bars indicate 95% confidence intervals. Median rates (95% confidence interval) are as follows: $(CAG)_{25}$ (blue): wild type: 1.1×10^{-6} (7.0×10^{-7} to 1.9×10^{-6}), *msh3Δ*: 2.4×10^{-7} (1.6×10^{-7} to 2.8×10^{-7}), *msh6Δ*: 2.4×10^{-6} (1.7×10^{-6} to 2.9×10^{-6}), and *msh6(3MBD) msh3Δ*: 2.5×10^{-7} (1.3×10^{-7} to 3.4×10^{-7}); and $(CTG)_{25}$ (red): wild type: 2.2×10^{-5} (1.7×10^{-5} to 2.9×10^{-5}), *msh3Δ*: 1.0×10^{-6} (2.8×10^{-7} to 2.2×10^{-6}), *msh6Δ*: 1.5×10^{-4} (9.0×10^{-5} to 2.0×10^{-4}), and *msh6(3MBD) msh3Δ*: 6.6×10^{-7} (5.0×10^{-7}–9.8×10^{-7}). Scrambled $(C,A,G)_{25}$ and $(C,T,G)_{25}$ tracts were also tested, and median rates are as follows: $(C,A,G)_{25}$ wild type ($n=180$) $<1.0\times10^{-8}$, *msh3Δ* ($n=90$) $<1.0\times10^{-8}$, *msh6Δ* ($n=30$) $<1.0\times10^{-8}$, and *msh6(3MBD) msh3Δ* ($n=20$) $=2.1\times10^{-7}$ (1.6×10^{-7} to 3.5×10^{-7}) and $(C,T,G)_{25}$: wild type ($n=90$) $<1.0\times10^{-8}$, *msh3Δ* ($n=90$) $<1.0\times10^{-8}$, *msh6Δ* ($n=51$) $<1.0\times10^{-8}$, and *msh6(3MBD) msh3Δ* ($n=20$) $=2.5\times10^{-7}$ (1.5×10^{-7} to 3.6×10^{-7}). *P*-values were calculated using Mann–Whitney test: $^{**}0.001<P<0.01$; $^{****}P<0.0001$. Please note the scale on the y-axis for CTG expansions is extended.

circles). *msh2Δ* exhibited a slippage rate similar to *msh3Δ* with this 4 nt. repeat (57- and 62-fold increases over wild type, respectively) (Sia *et al.* 1997; Lee *et al.* 2007), consistent with *MSH6* playing only a minor role in repair of larger IDL repair. *msh6(3MBD) msh3Δ* partially complemented the *msh3Δ*, with an intermediate slippage rate, with a 24-fold increase (Fig. 3). This incomplete complementation is consistent with distinct repair mechanisms for Msh2-Msh3 vs Msh2-Msh6 but is consistent with *msh6(3MBD)* allowing recognition of IDL substrates in vivo. We also tested *msh6(3MBD) msh3Δ* function in repair of 1 nt $[(G)_{18}]$ and 2 nt $[(GT)_{16.5}]$ repeats (Sia *et al.* 1997; Lee *et al.* 2007). *msh6(3MBD) msh3Δ* partially complemented repair of 1 nt. slippage events, compared to *msh3Δ*, but not repair of 2 nt. slippage events. We note that *msh2Δ* exhibited significantly higher slippage rates than either *msh3Δ* or *msh6Δ* or *msh6(3MBD)* in the presence of 1 nt. and 2 nt. repeat reporters (Sia *et al.* 1997; Lee *et al.* 2007), consistent with partial function of *msh6(3MBD)* in repair of these repeats (Fig. 3) (Shell *et al.* 2007).

Msh6(3MBD) does not promote TNR expansions in vivo

We next sought to determine whether the Msh3 MBD was sufficient to promote expansions in vivo (Shell *et al.* 2007). We used an in vivo TNR expansion assay in *msh6(3MBD) msh3Δ* to determine the effect of this construct on CAG and CTG expansion rates. This assay, described previously (Miret *et al.* 1998; Williams *et al.* 2020), places a URA3 reporter gene downstream of a promoter that encodes a $(CNG)_{25}$ repeat tract. If an expansion of 4 or more repeats occurs in the tract (≥29 repeats), URA3 will not be transcribed, leading to 5-FOA resistance (Williams and Surtees 2018). We previously demonstrated that *msh3Δ* reduces the expansion rate for $(CAG)_{25}$ and $(CTG)_{25}$ repeat tracts 5- and 30-fold, respectively, while the expansion rate was slightly increased in *msh6Δ* (Fig. 4) (Kantartzis *et al.* 2012). In *msh6(msh3MBD) msh3Δ*, the expansion rate for the $(CAG)_{25}$ tract was very similar to that of the *msh3Δ*, while the $(CTG)_{25}$ expansion rate was slightly lower than *msh3Δ* (Fig. 4). These results indicate that *msh6(3MBD)* does not complement an *msh3Δ*. Therefore, the Msh3 MBD is not sufficient to promote expansions and does not confer the capacity to promote TNR expansions on Msh6, despite the fact that Msh2-msh6(msh3MBD) is able to bind specifically to TNR structures. Mutations in URA3 can also lead to 5-FOA resistance. Therefore, we determined the proportion of true expansions by amplifying the TNR tract from 5-FOA-resistant colonies and determining tract lengths by gel electrophoresis (Fig. 5) (Williams *et al.* 2020). In wild-type cells, 90% of CAG and >99% of CTG tracts exhibited true expansions. In *msh3Δ*, 63% of CAG and 91% of CTG tracts and, in *msh6Δ*, 55% of CAG and 90% of CTG tracts were bona fide TNR expansions (Kantartzis *et al.* 2012). In

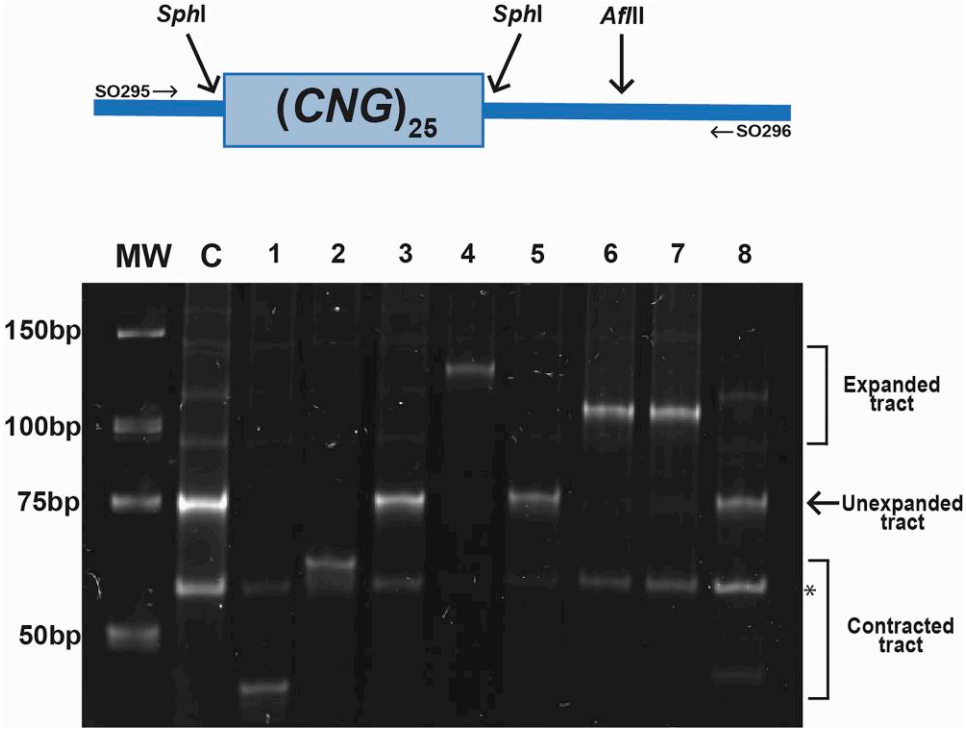

Fig. 5. TNR tract length in reporters recovered from 5-FOA-resistant colonies. Top: Schematic of PCR product used to measure tract lengths of 5-FOA-resistant colonies. Amplification with SO295 and SO296 results in a 188 bp product. Digestion with *Sph*I cuts on either side of the tract, releasing the 75 bp tract. The remaining 73 bp product adjacent to the tract is digested with *Afl*II, allowing for visualization of the TNR tract. Bottom: Representative gel of digested TNR tracts amplified from 5-FOA-resistant colonies from wild type and *msh6(3MBD) msh3Δ*. These tracts include contracted (lanes 1 and 2), expanded (lanes 4, 6, and 7), and stable TNR tracts (lanes 3, 5, and 8). *denotes band that results from *Afl*II digestion.

msh6(3MBD) msh3Δ, 7% of CAG tracts and 33% of CTG tracts exhibited true expansions (Table 6). The remaining tracts were either stable in length or exhibited tract contractions. We also observed a high rate of 5-FOA resistance in these strains with scrambled tracts, which do not expand (Fig. 4 legend), indicating that, in *msh6(3MBD) msh3Δ*, which has a high mutation rate (Fig. 3), mutations other than TNR tract expansions are leading to 5-FOA resistance. One possible source is mutations within *URA3* itself. We sequenced *URA3* from a subset of 5-FOA-resistant colonies associated with contracted (CAG) or stable (CTG) TNR tracts. For the contracted (CAG) tracts, 7 of 10 sustained a mutation within the coding sequence of *URA3*: H61Y, Y84 to STOP, and S154 to STOP each occurred twice, G255D occurred once. In contrast, for the stable (CTG) tracts, 0 of 17 sustained a mutation within the *URA3* coding or promoter region, and we infer that mutations elsewhere are leading to 5-FOA resistance eg (Armstrong *et al.* 2024).

The source of 5-FOA resistance notwithstanding the false positive rate indicated that the calculated expansion rates for *msh6(3MBD) msh3Δ* are overestimates; a corrected rate for *msh6(3MBD) msh3Δ* TNR expansions would go from 2.5×10^{-7} to 1.8×10^{-8} (CAG) or from 6.6×10^{-7} to 2.2×10^{-7} (CTG). Thus, measuring TNR expansion rates with this assay becomes more complicated as the background mutation rate increases for a given genotype, increasing the probability of observing 5-FOA resistance without expansion. Nonetheless, our data suggest that the CAG and CTG expansion rates for the chimeric complex are ~8- and ~4-fold lower than in the absence of *MSH3*, respectively, with corrected expansion rates of 1.5×10^{-7} (*msh3Δ* CAG) and 9.7×10^{-7} (*msh3Δ* CTG).

Discussion

We demonstrated that, in vitro, Msh2-msh6(3MBD) exhibited DNA structure-binding affinities for loop and TNR structures that were comparable to Msh2-Msh3, indicating that Msh2-msh6(3MBD) has acquired Msh3 DNA-binding properties and is able to recognize and bind TNR structures in vivo (Table 4). We also observed elevated affinities of Msh2-Msh6 for the TNR structures (Table 5), although Msh2-Msh6 does not promote expansions in vivo (Kantartzis *et al.* 2012). We also demonstrated that *msh6(3MBD) msh3Δ* is sufficient to allow some repair of Msh2-Msh3-specific DNA errors (in/dels), partially complementing the loss of *MSH3* in 4 nt. loop repair in vivo (Fig. 3). We hypothesize that differences in communication between the DNA-binding and ATPase domains of Msh2-Msh3 vs Msh2-Msh6 are reflected in this partial complementation. In contrast, this level of activity in *msh6(3MBD) msh3Δ* is not sufficient to promote MMR-mediated CAG/CTG expansions (Fig. 4). In fact, the true TNR expansion rates appear to be lower in *msh6(3MBD) msh3Δ* than in *msh3Δ*. Together, our data indicate that recognition and specific binding to the TNR structure by MSH complexes are not sufficient to promote TNR expansions. Thus, we propose that distinct Msh3-specific molecular requirements beyond Msh3 MBD are necessary for promoting TNR expansions, including DNA-mediated modulation of Msh2-Msh3 ATP-binding and hydrolysis and interactions with MLH complexes. Furthermore, Msh2-msh6(3MBD), a MSH complex that is able to bind MMR and TNR structures (Table 4) but not coordinate efficient repair (Figs. 3 and 4), appears to block background TNR expansions.

Msh2-Msh3 DNA binding to distinct structures is communicated to the ATPase domain through the connector domain,

Table 6. Proportion of true expansions in 5-FOA-resistant colonies.

		Permissive			Selective (5-FOA)		
Tract	**Genotype**	**Expansion**	**Stable**	**Contraction**	**Expansion**	**Stable**	**Contraction**
$(CAG)_{25}$	*MSH6 MSH3*	0	61 (100%)	0	95 (90%)	11 (10%)	0
	msh6(3MBD) msh3Δ	0	10 (100%)	0	2 (7%)	8 (28%)	19 (65%)
	msh6Δ	3 (6%)	47 (94%)	0	27 (55%)	22 (45%)	0
	msh3Δ	0	53 (56%)	42 (44%)	44 (67%)	16 (24%)	6 (9%)
$(CTG)_{25}$	*MSH6 MSH3*	0	68 (100%)	0	112 (100%)	0	0
	msh6(3MBD) msh3Δ	0	10 (100%)	0	13 (33%)	25 (65%)	1 (2%)
	msh6Δ	0	30 (100%)	0	26 (90%)	3 (10%)	0
	msh3Δ	0	25 (100%)	0	21 (91%)	2 (9%)	0
$(C,A,G)_{25}$ scrambled	*MSH6 MSH3*	0	40 (100%)	0	0	40 (100%)	0
	msh6(3MBD) msh3Δ	0	20 (100%)	0	0	20 (100%)	0
$(C,T,G)_{25}$ scrambled	*MSH6 MSH3*	0	60 (100%)	0	0	65 (100%)	0
	msh6(3MBD) msh3Δ	0	20 (100%)	0	0	20 (100%)	0

modulating its ATP binding, hydrolysis, and turnover activities to promote repair (Owen *et al.* 2005, 2009; Gupta *et al.* 2011; Lang *et al.* 2011; Kumar *et al.* 2014). ATP promotes Msh2-Msh3 dissociation from the DNA, promoting the recycling of Msh2-Msh3 (Surtees and Alani 2006; Brown *et al.* 2016). Different DNA structures modulate Msh2-Msh3 ATP binding and hydrolysis (Owen *et al.* 2005, 2009; Surtees and Alani 2006; Kumar *et al.* 2014). Critically, the Msh2-Msh3 ATP-binding domains are distinct from those of Msh2-Msh6, with regulated access to the Msh3 nucleotide-binding pocket (Gupta *et al.* 2011; Kumar *et al.* 2013). Altered regulation of ATP binding and/or hydrolysis disrupts Msh2-Msh3-mediated MMR, but not Msh2-Msh3's DSBR activity (Kumar *et al.* 2013). Therefore, Msh2-msh6(3MBD) may misregulate the ATPase domain through incorrect signal transduction after DNA binding and altered access to the nucleotide-binding pocket. Our previous data indicate that Msh2-msh6(3MBD) has a higher ATPase activity than Msh2-Msh3, more similar to Msh2-Msh6, while being stimulated by Msh2-Msh3-specific DNA substrates (Brown *et al.* 2016). One possibility is that this elevated ATPase activity increases the turnover of Msh2-msh6(3MBD), impairing both Msh2-Msh3-mediated MMR and TNR expansions, although apparently not to the same extent (Figs. 3 and 4) (Shell *et al.* 2007). This may be a result of distinct DNA structure-specific allosteric changes within the Msh complex. We propose that increased turnover precludes the Msh complex from targeting the DNA structures for either MMR or TNR expansion, leading to defects in both Msh2-Msh3-mediated pathways. Together, our data indicate that Msh2-Msh3's MMR activity is specifically required for promoting TNR expansions.

ATP-induced conformational changes are required for Msh2-Msh3's interaction with MLH complexes and stimulate their endonuclease activity. MutLα (yeast Mlh1-Pms1) is the primary Mlh complex in MMR, although MutLγ (yeast Mlh1-Mlh3) also plays a minor, largely Msh2-Msh3-specific, role in MMR. Both MutLα and MutLγ promote TNR expansions in vivo in mammalian systems in a Msh2-Msh3-specific manner (Pinto *et al.* 2013; Zhao *et al.* 2018; Hayward *et al.* 2020; Kadyrova *et al.* 2020; Miller *et al.* 2020; Roy *et al.* 2021; Lee *et al.* 2022), likely by nicking and promoting excision of the template leading to expansions (Pluciennik *et al.* 2013; Kadyrova *et al.* 2020). Our data support the hypothesis that this is not simply a result of altered DNA-binding specificity, but rather that Msh2-Msh3-specific Mlh interactions and activation are required for TNR expansions and are missing in Msh2-msh6(3MBD). We note that human Msh2-Msh3 interactions with MutLα are mediated through the PCNA interaction motif (PIP box) with the Msh3 N-terminal region (Pluciennik *et al.* 2013). In contrast, the Msh6 NTR or PIP box is not required for this interaction (Iyer *et al.* 2008); therefore, any Msh2-msh6(3MBD) interactions with MLH complexes are expected to be quite different from Msh2-Msh3-MLH interactions. We propose that Msh2-Msh3-mediated TNR expansions require the Msh2-Msh3-mediated MMR pathway to be fully functional and intact. This would include proper DNA binding and appropriate signal transduction for regulation of ATP binding/hydrolysis and subsequent Mlh interactions. This is consistent with Msh2-Msh3 playing an active, pathogenic role in promoting TNR expansions.

Data availability

All strains and plasmids are available upon request.

Acknowledgments

We thank Dr. Natalie Lamb and Brett Irwin and other members of the Surtees Lab for technical assistance and for helpful discussions. We thank Dr. Richard Kolodner for the *msh6(3MBD)* integration plasmid. We are grateful to Dr. Mark Sutton for constructive comments on the study and the manuscript.

Funding

This work was supported by an American Cancer Society Research Scholar Grant (RSG-142350-01), Pfizer, and the National Science Foundation (MCB grant #2325415) to JAS.

Conflicts of interest

The authors declare no conflicts of interest.

Literature cited

Aelvoet AS, Hoekman DR, Redeker BJW, Weegenaar J, Dekker E, van Noesel CJM, Duijkers FAM. 2023. A large family with MSH3-related polyposis. Fam Cancer. 22(1):49–54. doi:10.1007/s10689-022-00297-x.

Armstrong JO, Jiang P, Tsai S, Phan MM, Harris K, Dunham MJ. 2024. URA6 mutations provide an alternative mechanism for 5-FOA resistance in *Saccharomyces cerevisiae*. bioRxiv 597250. https://doi.org/10.1101/2024.06.03.597250, preprint: not peer reviewed.

Bettencourt C, Hensman-Moss D, Flower M, Wiethoff S, Brice A, Goizet C, Stevanin G, Koutsis G, Karadima G, Panas M, *et al.*

2016. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. Ann Neurol. 79(6):983–990. doi:10.1002/ana.24656.

Bowers J, Sokolsky T, Quach T, Alani E. 1999. A mutation in the MSH6 subunit of the *Saccharomyces cerevisiae* MSH2-MSH6 complex disrupts mismatch recognition. J Biol Chem. 274(23):16115–16125. doi:10.1074/jbc.274.23.16115.

Brown MW, Kim Y, Williams GM, Huck JD, Surtees JA, Finkelstein IJ. 2016. Dynamic DNA binding licenses a repair factor to bypass roadblocks in search of DNA lesions. Nat Commun. 7(1):10607. doi:10.1038/ncomms10607.

De' Angelis GL, Bottarelli L, Azzoni C, De' Angelis N, Leandro G, Di Mario F, Gaiani F, Negri F. 2018. Microsatellite instability in colorectal cancer. Acta Biomed. 89(9-S):97–101. doi:10.23750/abm.v89i9-S.7960.

Dixon MJ, Bhattacharyya S, Lahue RS. 2004. Genetic assays for triplet repeat instability in yeast. Methods Mol Biol. 277:29–45. doi:10.1385/1-59259-804-8:029.

Dixon W, Massey F. 1969. Introduction to Statistical Analysis. New York: McGraw Hill.

Dowen JM, Putnam CD, Kolodner RD. 2010. Functional studies and homology modeling of Msh2-Msh3 predict that mispair recognition involves DNA bending and strand separation. Mol Cell Biol. 30(13):3321–3328. doi:10.1128/MCB.01558-09.

Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A. 88(16):7160–7164. doi:10.1073/pnas.88.16.7160.

Furman CM, Elbashir R, Alani E. 2021. Expanded roles for the MutL family of DNA mismatch repair proteins. Yeast. 38(1):39–53. doi:10.1002/yea.3512.

Gietz D, St Jean A, Woods RA, Schiestl RH. 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20(6):1425. doi:10.1093/nar/20.6.1425.

Goellner EM, Putnam CD, Kolodner RD. 2015. Exonuclease 1--dependent and independent mismatch repair. DNA Repair (Amst). 32:24–32. doi:10.1016/j.dnarep.2015.04.010.

Gorman J, Chowdhury A, Surtees JA, Shimada J, Reichman DR, Alani E, Greene EC. 2007. Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. Mol Cell. 28(3):359–370. doi:10.1016/j.molcel.2007.09.008.

Gupta S, Gellert M, Yang W. 2011. Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. Nat Struct Mol Biol. 19(1):72–78. doi:10.1038/nsmb.2175.

Habraken Y, Sung P, Prakash L, Prakash S. 1996. Binding of insertion/-deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. Curr Biol. 6(9):1185–1187. doi:10.1016/S0960-9822(02)70686-6.

Harrington JM, Kolodner RD. 2007. *Saccharomyces cerevisiae* Msh2-Msh3 acts in repair of base–base mispairs. Mol Cell Biol. 27(18):6546–6554. doi:10.1128/MCB.00855-07.

Hayward BE, Steinbach PJ, Usdin K. 2020. A point mutation in the nuclease domain of MLH3 eliminates repeat expansions in a mouse stem cell model of the Fragile X-related disorders. Nucleic Acids Res. 48(14):7856–7863. doi:10.1093/nar/gkaa573.

Henderson ST, Petes TD. 1992. Instability of simple sequence DNA in *Saccharomyces cerevisiae*. Mol Cell Biol. 12(6):2749–2757. doi:10.1128/mcb.12.6.2749-2757.1992.

Iyer RR, Pluciennik A. 2021. DNA mismatch repair and its role in Huntington's disease. J Huntingtons Dis. 10(1):75–94. doi:10.3233/JHD-200438.

Iyer RR, Pohlhaus TJ, Chen S, Hura GL, Dzantiev L, Beese LS, Modrich P. 2008. The MutSalpha-proliferating cell nuclear antigen interaction in human DNA mismatch repair. J Biol Chem. 283(19):13310–13319. doi:10.1074/jbc.M800606200.

Jiang J, Bai L, Surtees JA, Gemici Z, Wang MD, Alani E. 2005. Detection of high-affinity and sliding clamp modes for MSH2-MSH6 by single-molecule unzipping force analysis. Mol Cell. 20(5):771–781. doi:10.1016/j.molcel.2005.10.014.

Jiricny J. 2006. The multifaceted mismatch–repair system. Nat Rev Mol Cell Biol. 7(5):335–346. doi:10.1038/nrm1907.

Kadyrova LY, Gujar V, Burdett V, Modrich PL, Kadyrov FA. 2020. Human MutLγ, the MLH1MLH3 heterodimer, is an endonuclease that promotes DNA expansion. Proc Natl Acad Sci U S A. 117(7):3535–3542. doi:10.1073/pnas.1914718117.

Kantartzis A, Williams GM, Balakrishnan L, Roberts RL, Surtees JA, Bambara RA. 2012. Msh2-Msh3 interferes with Okazaki fragment processing to promote trinucleotide repeat expansions. Cell Rep. 2(2):216–222. doi:10.1016/j.celrep.2012.06.020.

Keogh N, Chan KY, Li GM, Lahue RS. 2017. Mutsβ abundance and Msh3 ATP hydrolysis activity are important drivers of CTG•CAG repeat expansions. Nucleic Acids Res. 45(17):1006810078. doi:10.1093/nar/gkx650.

Kijas AW, Studamire B, Alani E. 2003. Msh2 separation of function mutations confer defects in the initiation steps of mismatch repair. J Mol Biol. 331(1):123–138. doi:10.1016/S0022-2836(03)00694-6.

Kumar C, Eichmiller R, Wang B, Williams GM, Bianco PR, Surtees JA. 2014. ATP binding and hydrolysis by *Saccharomyces cerevisiae* Msh2-Msh3 are differentially modulated by mismatch and doublestrand break repair DNA substrates. DNA Repair (Amst). 18:18–30. doi:10.1016/j.dnarep.2014.03.032.

Kumar C, Piacente SC, Sibert J, Bukata AR, O'Connor J, Alani E, Surtees JA. 2011. Multiple factors insulate Msh2-Msh6 mismatch repair activity from defects in Msh2 domain I. J Mol Biol. 411(4):765–780. doi:10.1016/j.jmb.2011.06.030.

Kumar C, Williams GM, Havens B, Dinicola MK, Surtees JA. 2013. Distinct requirements within the Msh3 nucleotide binding pocket for mismatch and double-strand break repair. J Mol Biol. 425(11):1881–1898. doi:10.1016/j.jmb.2013.02.024.

Kunkel TA, Erie DA. 2005. DNA mismatch repair. Annu Rev Biochem. 74(1):681–710. doi:10.1146/annurev.biochem.74.082803.133243.

Lahue RS. 2020. New developments in Huntington's disease and other triplet repeat diseases: DNA repair turns to the dark side. Neuronal Signal. 4(4):Ns20200010. doi:10.1042/NS20200010.

Lamb NA, Bard JE, Loll-Krippleber R, Brown GW, Surtees JA. 2022. Complex mutation profiles in mismatch repair and ribonucleotide reductase mutants reveal novel repair substrate specificity of MutS homolog (MSH) complexes. Genetics. 221(4):iyac092. doi:10.1093/genetics/iyac092.

Lang WH, Coats JE, Majka J, Hura GL, Lin Y, Rasnik I, McMurray CT. 2011. Conformational trapping of mismatch recognition complex MSH2/MSH3 on repair-resistant DNA loops. Proc Natl Acad Sci U S A. 108(42):E837–E844. doi:10.1073/pnas.1105461108.

Lee J-M, Correia K, Loupe J, Kim K-H, Barker D, Hong EP, Chao MJ, Long JD, Lucente D, Vonsattel JPG, *et al.* 2019. CAG repeat not polyglutamine length determines timing of Huntington's disease onset. Cell. 178(4):887–900.e814. doi:10.1016/j.cell.2019.06.036.

Lee JM, Huang Y, Orth M, Gillis T, Siciliano J, Hong E, Mysore JS, Lucente D, Wheeler VC, Seong IS, *et al.* 2022. Genetic modifiers of Huntington disease differentially influence motor and cognitive domains. Am J Hum Genet. 109(5):885–899. doi:10.1016/j.ajhg.2022.03.004.

Lee SD, Surtees JA, Alani E. 2007. *Saccharomyces cerevisiae* MSH2-MSH3 and MSH2-MSH6 complexes display distinct requirements for DNA binding domain I in mismatch recognition. J Mol Biol. 366(1):53–66. doi:10.1016/j.jmb.2006.10.099.

Li GM. 2008. Mechanisms and functions of DNA mismatch repair. Cell Res. 18(1):85–98. doi:10.1038/cr.2007.115.

Marsischky GT, Filosi N, Kane MF, Kolodner R. 1996. Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev. 10(4):407–420. doi:10.1101/gad.10.4.407.

Matos-Rodrigues G, Hisey JA, Nussenzweig A, Mirkin SM. 2023. Detection of alternative DNA structures and its implications for human disease. Mol Cell. 83(20):3622–3641. doi:10.1016/j.molcel.2023.08.018.

Miao HK, Chen LP, Cai DP, Kong WJ, Xiao L, Lin J. 2015. MSH3 rs26279 polymorphism increases cancer risk: a meta-analysis. Int J Clin Exp Pathol. 8:11060–11067.

Miller CJ, Kim GY, Zhao X, Usdin K. 2020. All three mammalian MutL complexes are required for repeat expansion in a mouse cell model of the Fragile X-related disorders. PLoS Genet. 16(6): e1008902. doi:10.1371/journal.pgen.1008902.

Miret JJ, Pessoa-Brandão L, Lahue RS. 1998. Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 95(21):12438–12443. doi:10.1073/pnas.95.21.12438.

Mirkin SM. 2007. Expandable DNA repeats and human disease. Nature. 447(7147):932–940. doi:10.1038/nature05977.

Nair KR. 1940. Table of confidence intervals for the median in samples from any continuous population. SankhyA: Indian J Stat. 4: 551–558. https://www.jstor.org/stable/40383959.

Obmolova G, Ban C, Hsieh P, Yang W. 2000. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. Nature. 407(6805):703–710. doi:10.1038/35037509.

Owen BA, H Lang W, McMurray CT. 2009. The nucleotide binding dynamics of human MSH2MSH3 are lesion dependent. Nat Struct Mol Biol. 16(5):550–557. doi:10.1038/nsmb.1596.

Owen BA, Yang Z, Lai M, Gajec M, Badger JD 2nd, Hayes JJ, Edelmann W, Kucherlapati R, Wilson TM, McMurray CT. 2005. (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. Nat Struct Mol Biol. 12(8):663–670. doi:10.1038/nsmb965.

Panigrahi GB, Lau R, Montgomery SE, Leonard MR, Pearson CE. 2005. Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. Nat Struct Mol Biol. 12(8): 654–662. doi:10.1038/nsmb959.

Pannafino G, Alani E. 2021. Coordinated and independent roles for MLH subunits in DNA repair. Cells. 10(4):948. doi:10.3390/cells10040948.

Pinto RM, Dragileva E, Kirby A, Lloret A, Lopez E, St Claire J, Panigrahi GB, Hou C, Holloway K, Gillis T, *et al.* 2013. Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet. 9(10):e1003930. doi:10.1371/journal.pgen.1003930.

Pluciennik A, Burdett V, Baitinger C, Iyer RR, Shi K, Modrich P. 2013. Extrahelical (CAG)/(CTG) triplet repeat elements support proliferating cell nuclear antigen loading and MutLα endonuclease activation. Proc Natl Acad Sci U S A. 110(30):12277–12282. doi:10.1073/pnas.1311325110.

Rashid M, Rashid R, Gadewal N, Carethers JM, Koi M, Brim H, Ashktorab H. 2024. High-throughput sequencing and in-silico analysis confirm pathogenicity of novel MSH3 variants in African American colorectal cancer. Neoplasia. 49:100970. doi: 10.1016/j.neo.2024.100970.

Richard GF. 2021. The startling role of mismatch repair in trinucleotide repeat expansions. Cells. 10(5):1019. doi:10.3390/cells10051019.

Roy JCL, Vitalo A, Andrew MA, Mota-Silva E, Kovalenko M, Burch Z, Nhu AM, Cohen PE, Grabczyk E, Wheeler VC, *et al.* 2021. Somatic CAG expansion in Huntington's disease is dependent on the MLH3 endonuclease domain, which can be excluded via splice redirection. Nucleic Acids Res. 49(7):3907–3918. doi:10.1093/nar/gkab152.

Santos LS, Silva SN, Gil OM, Ferreira TC, Limbert E, Rueff J. 2018. Mismatch repair single nucleotide polymorphisms and thyroid cancer susceptibility. Oncol Lett. 15(5):6715–6726. doi:10.3892/ol.2018.8103.

Shell SS, Putnam CD, Kolodner RD. 2007. Chimeric *Saccharomyces cerevisiae* Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins. Proc Natl Acad Sci U S A. 104(26):10956–10961. doi:10.1073/pnas.0704148104.

Sia EA, Kokoska RJ, Dominska M, Greenwell P, Petes TD. 1997. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol Cell Biol. 17(5): 28512858. doi:10.1128/MCB.17.5.2851.

Studamire B, Quach T, Alani E. 1998. *Saccharomyces cerevisiae* Msh2p and Msh6p ATPase activities are both required during mismatch repair. Mol Cell Biol. 18(12):7590–7601. doi:10.1128/MCB.18.12.7590.

Surtees JA, Alani E. 2006. Mismatch repair factor MSH2-MSH3 binds and alters the conformation of branched DNA structures predicted to form during genetic recombination. J Mol Biol. 360(3): 523–536. doi:10.1016/j.jmb.2006.05.032.

Tian L, Hou C, Tian K, Holcomb NC, Gu L, Li G-M. 2009. Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. J Biol Chem. 284(31):20452–20456. doi:10.1074/jbc.C109.014977.

Tomé S, Holt I, Edelmann W, Morris GE, Munnich A, Pearson CE, Gourdon G. 2009. MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. PLoS Genet. 5(5): e1000482. doi:10.1371/journal.pgen.1000482.

Tomé S, Manley K, Simard JP, Clark GW, Slean MM, Swami M, Shelbourne PF, Tillier ERM, Monckton DG, Messer A, *et al.* 2013. MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. PLoS Genet. 9(2): e1003280. doi:10.1371/journal.pgen.1003280.

Warren JJ, Pohlhaus TJ, Changela A, Iyer RR, Modrich PL, Beese LS. 2007. Structure of the human MutSalpha DNA lesion recognition complex. Mol Cell. 26(4):579–592. doi:10.1016/j.molcel.2007.04.018.

Williams GM, Petrides AK, Balakrishnan L, Surtees JA. 2020. Tracking expansions of stable and threshold length trinucleotide repeat tracts in vivo and in vitro using *Saccharomyces cerevisiae*. Methods Mol Biol. 2056:25–68. doi:10.1007/978-1-4939-9784-8_3.

Williams GM, Surtees JA. 2015. MSH3 promotes dynamic behavior of trinucleotide repeat tracts in vivo. Genetics. 200(3):737–754. doi: 10.1534/genetics.115.177303.

Williams GM, Surtees JA. 2018. Measuring dynamic behavior of trinucleotide repeat tracts in vivo in *Saccharomyces cerevisiae*. Methods Mol Biol. 1672:439–470. doi:10.1007/978-1-4939-7306-4_30.

Wilson T, Guerrette S, Fishel R. 1999. Dissociation of mismatch recognition and ATPase activity by hMSH2-hMSH3. J Biol Chem. 274(31):21659–21664. doi:10.1074/jbc.274.31.21659.

Winston F, Dollard C, Ricupero-Hovasse SL. 1995. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. Yeast. 11(1):53–55. doi:10.1002/yea.320110107.

Zhao X, Zhang Y, Wilkins K, Edelmann W, Usdin K. 2018. Mutlγ promotes repeat expansion in a Fragile X mouse model while EXO1 is protective. PLoS Genet. 14(10):e1007719. doi:10.1371/journal.pgen.1007719.

Editor: J. L. Argueso