Mutsβ generates both expansions and contractions in a mouse model of the Fragile X-associated disorders

Xiao-Nan Zhao¹, Daman Kumari¹, Shikha Gupta², Di Wu³, Maya Evanitsky¹, Wei Yang² and Karen Usdin¹,*

¹Section on Gene Structure and Disease, Laboratory of Cell and Molecular Biology, ²Section on Structure and Mechanisms of DNA repair, replication and recombination, Laboratory of Molecular Biology and ³Section on Physical Biochemistry, Laboratory of Biochemistry and Genetics, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0830, USA

*To whom correspondence should be addressed at: Building 8, Room 2A19, National Institutes of Health, 8 Center Drive MSC 0830, Bethesda, MD 20892-0830, USA. Tel: +1 3014962189; Fax: +1 3014962189; Email: ku@helix.nih.gov

Abstract

Fragile X-associated disorders are Repeat Expansion Diseases that result from expansion of a CGG/CCG-repeat in the FMR1 gene. Contractions of the repeat tract also occur, albeit at lower frequency. However, these contractions can potentially modulate disease symptoms or generate an allele with repeat numbers in the normal range. Little is known about the expansion mechanism and even less about contractions. We have previously demonstrated that the mismatch repair (MMR) protein MSH2 is required for expansions in a mouse model of these disorders. Here, we show that MSH3, the MSH2-binding partner in the MutSβ complex, is required for 98% of germ line expansions and all somatic expansions in this model. In addition, we provide evidence for two different contraction mechanisms that operate in the mouse model, a MutSβ-independent one that generates small contractions and a MutSβ-dependent one that generates larger ones. We also show that MutSβ complexes formed with the repeats have altered kinetics of ATP hydrolysis relative to complexes with bona fide MMR substrates and that MutSβ increases the stability of the CCG-hairpins at physiological temperatures. These data may have important implications for our understanding of the mechanism(s) of repeat instability and for the role of MMR proteins in this process.

Introduction

The Fragile X (FX)-related disorders (FXDs) are Repeat Expansion Diseases that result from expansion of a microsatellite with the repeat unit CGG/CCG. The microsatellite is located in the 5′ UTR of the FMR1 gene [reviewed in (1)]. Alleles with 55–200 repeats are referred to as premutation (PM) alleles and such alleles are prone to expansion or contraction in humans and in an FXD mouse model (2–4). Moderate expansions can result in larger PM alleles that modulate the risk of Fragile X-associated tremor/ataxia syndrome and Fragile X-associated primary ovarian insufficiency, two clinical conditions that affect PM carriers. More extensive expansions into the full mutation range (>200 repeats) result in Fragile X syndrome, a form of intellectual disability. Contractions are also clinically relevant because they can result in tissue mosaicism (5–10) that can complicate risk assessment. Contractions can also modulate disease risk or even result in reversion to a normal allele (11–15). The molecular basis of these expansions and contractions is not known. We and others have shown that the individual strands of the FX repeat form hairpins and other folded structures that are similar to those formed by other…
disease-associated repeats (16–26). These structures are thought to be the substrates upon which the expansion and contraction pathways act. Interruptions to the purity of the CCG/CCG-repeat tract in the form of one or more interspersed AGG/CCT triplets are seen in some FMR1 alleles. These interruptions would be predicted to reduce hairpin stability and have been shown to reduce risk of expansion in humans (27–29).

Two basic types of expansion models have been proposed, the first in which instability results from difficulties with replication of the repeat-containing region during normal chromosome duplication and a second that invokes some form of aberrant DNA repair (30). However, an effect of maternal age on the risk of expansion in humans is seen, suggesting that expansion occurs in oocytes (28), a cell in which DNA replication does not occur. Thus, a problem with DNA repair rather than chromosomal replication may be responsible for expansion, at least in some cells. However, we have also shown that in mice somatic expansion only occurs when the PM allele is on the active X chromosome (31) and a re-examination of human data (32) suggests that the same is true in women. Thus, transcription or transcriptionally competent chromatin plays an important role in the expansion process (31).

We have previously shown that this transcription dependence is not related to a role for Transcription Coupled Repair (3). Rather, we have shown that the mismatch repair (MMR) protein MSH2 is required for all germ line and somatic expansions in the mouse model (33). A similar requirement for MSH2 has been observed for expansion in other mouse models [reviewed in (24)]. However, whether MSH2 acts to promote expansions in the FXD mouse via its interaction with MSH3 in the MutSβ complex, one of the two MSH2-containing complexes in mammalian cells, or via its interaction with MSH6 in the MutSα complex, the second MSH2-containing complex, was not known. Furthermore, very little is known about the mechanism responsible for contractions. In fact, to date no proteins or pathways have been identified that are responsible for the generation of contractions in humans or in mouse models. We show here that MSH3, and thus MutSβ, is required for 98% of germ line expansions and all somatic expansions in a mouse model for the FXD. In addition to a role of MutSβ in generating expansions, we also uncovered two additional unappreciated aspects of CGG-repeat instability in these animals. First, we show that in mice wild type (WT) for Msh3, two discrete contraction size classes can be seen, one involving the loss of ~1–2 repeats and a second that results in the loss of larger repeat numbers. Second, we demonstrate a hitherto undescribed role for MutSβ in repeat instability in mice, namely its ability to promote the formation of larger contractions. The fact that MutSβ promotes the generation of one class of contractions and not the other, along with our biochemical studies on binding of MutSβ to the FX repeats, has interesting implications for models of the mechanism of repeat instability.

Results
Loss of MSH3 eliminates all somatic expansions

We examined the effect of the loss of MSH3 on the extent of expansion in the somatic tissue of FXD male mice with ~140 CGG/CCG repeats. We used the somatic instability index (SII) as a quantitative measure of the extent of repeat expansion (35). In Msh3+/− animals, the heart, an organ that shows very little somatic instability, has an SII of −0.1 (Fig. 1). The negative value reflects the contribution to the SII of PCR products that are smaller than the original allele. These products are present at the same levels in the PCR profiles of all organs tested and in animals of all ages (3,4,36) and thus likely represent strand-slippage products produced by PCR across the repeats as described previously for the FX repeats (33) as well as many other short tandem repeats (37–39). In organs other than the heart, the SII ranged from 1.0 (kidney) to 6.8 (testes). As can be seen in Figure 1B, most of the PCR products larger than the original allele seen in the heart of an Msh3+/− mouse are absent from the PCR profile generated from DNA purified from the heart of an Msh3−/− mouse. However, the PCR products smaller than the original allele were indistinguishable from those seen in Msh3+/− mice. Almost all of the PCR products larger than the original allele are also lost from other, more expansion-prone organs. For example, the PCR profile for liver was indistinguishable from that of heart in Msh3−/− animals (Fig. 1B). In fact, all organs of Msh3−/− mice showed a slight negative SII similar to what was seen in heart (Fig. 1A). Thus, expansions were essentially eliminated in Msh3−/− animals and because no novel PCR products smaller than the original allele were observed in any organ, the loss of expanded alleles cannot be explained by an increase in contractions. Thus, MSH3, and

![Image](https://example.com/image.png)

**Figure 1.** Loss of MSH3 reduces somatic expansions. (A) The SII for the organs of six 6-month-old Msh3+/+, Msh3+/- and Msh3−/− males was determined and the average SII for each organ plotted. The initial repeat number in these mice was ~140. The negative value for the heart in animals of each genotype and for the remaining organs in Msh3−/− mice reflects the strand-slippage or stutter products that are typically seen when large repeat tracts are amplified rather than contraction events. The SIs for Msh3−/− animals that are significantly different from WT animals are marked with asterisks [*P < 0.05; **P < 0.01; ***P < 0.001]. (B) Comparison of the GeneMarker PCR profiles for the heart and liver of one Msh3+/+ and one Msh3−/− mouse.
therefore MutSβ, is required for the generation of somatic expansions. Even in Msh3+/− animals, the SII was significantly lower in tail, liver and testes than that in Msh3+/+ animals indicating that the loss of even one Msh3 allele is sufficient to reduce the expansion frequency (Fig. 1A).

**Loss of MSH3 eliminates almost all germ-line expansions**

In Msh3+/+ males, the expansion pathway predominates as evidenced by the fact that 90% of alleles transmitted by 2- to 6-month-old Msh3+/+ males had expansions (Fig. 2A). In contrast, in the progeny of Msh3−/− fathers of the same age, only 2% did. This decrease in expansions in the progeny of Msh3−/− fathers is associated with an increase in the frequency of unchanged alleles [79% up from 5% seen in the progeny Msh3+/− fathers (P = 0.0001); Fig. 2A]. Thus, our data would be consistent with a role of MSH3, and thus MutSβ, in promoting expansions. The fact that unchanged alleles predominate when expansions are lost in Msh3−/− mice confirms our earlier hypothesis that the contraction pathway is not very active in young males (4). The proportion of contractions is significantly increased in the progeny of 7- to 12-month-old Msh3−/− males (Fig. 2B), consistent with the idea that the contraction pathway becomes more prominent with age in these animals (4).

The proportion of contracted and unchanged alleles in the progeny of 2- to 6-month-old Msh3−/− mice also contrasts with what was seen in the progeny of age-matched Msh2−/− animals where both unchanged and contracted alleles increased to similar extents (33). As Msh3−/− mice lack MutSβ, but still produce MutSα, whereas Msh2−/− mice lack both MutSβ and MutSα, the difference in the distribution of the residual alleles in Msh2−/− mice compared with Msh3−/− mice likely reflects the effect of the loss of protection by MutSα against repeat contractions. Thus, MutSα may be a key player in the error-free repair pathway that we had previously identified (36). This hypothesis is currently being tested in mice lacking MSH6.

Loss of MSH3 in 2- to 6-month-old females also resulted in a significant reduction in expansions (Fig. 2C). However, in contrast to the relatively low level of contractions seen in the offspring of males, 58% of the residual alleles were smaller than the parental allele. Furthermore, the contraction frequency increased with age such that in the progeny of 7- to 12-month-old Msh3−/− females, only contractions were seen (Fig. 2D). This is consistent with our previous observations that the contraction pathway is more active in the female germ line than it is in the male germ line even in animals WT for the MMR proteins (4). The contraction pathway is thus able to compete more effectively with the error-free pathway for the instability substrate when the expansion pathway is eliminated.

**Msh3+/− mice show a bimodal distribution of germ line contraction sizes**

Visual inspection of the contractions seen in the progeny of 2- to 6-month-old Msh3+/− parents suggested a bimodal distribution of

![Figure 2](image-url)
contraction sizes. This was confirmed using Hartigans' dip test \(^{(40)}\) that indicated a significant deviation from unimodality consistent with a bimodal distribution, for both paternal (\(P = 0.029\)) and maternal (\(P = 0.041\)) transmission. Based on the minima seen in the associated density plots (−7 for females and −10 for males), the alleles were divided into 2 groups, 1 involving repeats having lost <7 repeats and 1 involving the loss of >7 repeats (Fig. 3A). The first group had an average loss of −1 to −2 repeats, whereas the second group had an average loss of 12 repeats for females and 17 for males. In the case of paternal transmissions, 56% of contractions fell into the first group, whereas 44% of contractions fell into the second group (black bars in Fig. 3A). In the case of maternal transmission, 53% of contractions fell into the first group, whereas 37% of contractions fell into the second group (white bars in Fig. 3A). This bimodal distribution was not noted previously because the low frequency of contractions in WT animals makes the accumulation of statistically significant sample numbers difficult.

As previously reported \((4,33)\), no contractions are seen in the progeny of older Msh3\(^+/+\) males (Fig. 2B). This is due to the high frequency of expansion that persists throughout life in male mice \((4)\). The net result is that, in the progeny of older males, even alleles that once had sustained contractions have now undergone one or more rounds of expansion \((4)\). Msh3\(^+/+\) females in this age group still show contractions, perhaps because the contraction frequency increases in prominence in females with age \((4)\). However, the progeny of Msh3\(^+/+\) females show a broader distribution of contraction sizes with less clear evidence for bimodality than do the progeny of 2- to 6-month-old females (Fig. 3B). This may be because in older mothers each transmitted allele has been subjected to multiple rounds of expansion and/or contraction that complicate the data interpretation.

The bimodal distribution of contractions that is apparent in younger animals may indicate the operation of two different contraction mechanisms, one that gives rise to small contractions and one that is responsible for larger ones.

Loss of MSH3 increases the contraction frequency but reduces the average contraction size

While the number of contractions transmitted was ~3-fold higher in 2- to 6-month Msh3\(^+/−\) males than it was in similarly aged WT males, careful analysis demonstrated that all of these contractions involved the loss of a single repeat unit (Fig. 3C, Fisher’s exact test for the difference between small (<7) and large (>7) contractions in Msh3\(^+/+\) and Msh3\(^+/−\) males \(P = 0.0325\)). Similarly, while 2- to 6-month-old Msh3\(^+/−\) females transmitted ~5-fold more contractions than 2- to 6-month-old Msh3\(^+/+\) females, no large contractions were seen (Fig. 3C, Fisher’s exact test for the difference between small and large contractions in Msh3\(^+/+\) and Msh3\(^+/−\) females, \(P = 0.0015\)). However, the transmitted contractions were slightly bigger on average than the contractions transmitted by Msh3\(^+/−\) males (−2.11 versus −1.00; \(t\)-test: \(P = 0.0178\)). This could be due to the higher contraction frequency seen on
maternal transmission that increases the likelihood that some alleles undergo more than one round of contraction (compare Fig. 2A and C).

As no contracted alleles are seen in the progeny of older Msh3+/− males, we were unable to compare the contraction sizes in Msh3+/− and Msh3−/− males in the 7- to 12-month-old range. However, as can be seen from Figure 3D, the progeny of 7- to 12-month-old Msh3−/− males had a distribution of contraction sizes that was still displaced toward smaller alleles than the progeny of 2- to 6-month-old Msh3−/− males with only 5% of alleles having >7 repeats. As can be seen in Figure 3D, in the progeny of 7- to 12-month-old Msh3−/− females, the average contraction size was also smaller than that in Msh3+/− mice of the same age with no large contractions being found (Fisher’s exact test for the difference between small and large contractions, P = 0.0001). Taken together our data thus show that the loss of Msh3 leads to preferential loss of larger contractions but does not eliminate contractions altogether.

**MutSβ binds to and stabilizes the secondary structures formed by the FX repeats**

As promoting expansions and contractions is inconsistent with the canonical role of MMR in protecting against microsatellite instability, we initiated biochemical studies aimed at understanding other ways in which MutSβ might be acting. We first generated substrates containing a loop-out of (CCG)13 or (CTG)13 that were modeled on those used previously to examine MutSβ binding to CAG-repeats (41) and shown in Table 1. We included substrates with (CAG)13, or (CTG)13-repeat loop-outs for comparison. We then tested the ability of purified MutSβ to bind to these structures. As can be seen in Figure 4, MutSβ binds as well to CCG- and CCC-repeats as it does to CAG- and CTG-repeats. In fact, if anything, the binding of CCG-repeats is more extensive than binding to the other repeats as evidenced by the binding seen at lower concentrations of MutSβ as well as by the ratio of bound to unbound probe. Of each pair of loop-outs, the loop-out forming the hairpin with the most unstable mismatch was the one that was bound more strongly (i.e. CCG>CAG; CAG>CTG). Evidence of multiple DNA–protein complexes could be seen on each DNA substrate, consistent with the idea that more than one MutSβ dimer is able to bind such loop-outs as previously reported for CAG-loop-outs (41).

Recent works suggest that ATP binding and hydrolysis by MutSβ are differentially modified by the substrates of different repair pathways (42–44). Specifically, it has been suggested that substrates of different repair pathways induce specific conformational changes in the DNA-binding domains of MutSβ that are then relayed to the ATPase domains resulting in changes in the kinetics of ATP hydrolysis (42). As can be seen in Figure 5 and consistent with what was reported for a CAG-loop-out (41), binding to either a CGG-loop-out or a CCG-loop-out resulted in altered kinetics of ATP hydrolysis relative to binding to a (CA)1 loop-out that is a bona fide MMR substrate (45). Thus, differences likely exist between the conformation of MutSβ when bound to the FX loop-outs and the conformation of MutSβ bound to a bona fide MMR substrate that affects ATP hydrolysis. This altered MutSβ conformation may result in less efficient signaling to proteins downstream in the MMR pathway or in more efficient signaling to an alternate repair pathway.

To assess the effect of MutSβ binding on the stability of the FX-repeat structures, we monitored the thermal denaturation of the oligonucleotide in the presence of BSA or MutSβ. As the hairpin-to-single-stranded transition of even a very short CGG-repeat oligonucleotide occurs at temperatures above the denaturation temperature of the most proteins (16,46), we limited our study to the CCG-repeat. The 5’ end of a (CCG)13 oligonucleotide was labeled with 5-carboxy-X-rhodamine (ROX™), a fluorescence donor and the 3’ end was labeled with IOWA Black® RQ, a fluorescence acceptor/quencher. This enabled the stability of the hairpins to be assessed in the presence of MutSβ by monitoring the increase in the fluorescence signal at the ROX™ emission wavelength with increasing temperature. The oligonucleotide was denatured and cooled under conditions in which the repeats are known to form hairpins (17–21,47,48). The oligonucleotide was then mixed with MutSβ and subjected to increasing temperatures as described in the Materials and Methods. Increasing temperatures resulted in a progressive increase in fluorescence at 608 nm consistent with melting of the secondary structure formed by the CCG-repeat. The melting curves obtained for both protein–CCG-repeat mixtures fit a two-state model (Supplementary Material, Fig. S2). The thermodynamic parameters derived from analysis of the melting curves are shown in Table 2. As can be seen from this table, the presence of MutSβ resulted in higher ΔG at 37°C than is seen in the presence of BSA suggesting that MutSβ increases the stability of the CCG-repeat structure at physiological temperatures.

**Discussion**

We have previously shown that MSH2 is required for all paternal and maternal germ line expansions as well as for somatic expansions. We show here that loss of MSH3 eliminates 98% of germ line and all somatic repeat expansions in these animals (Figs

### Table 1. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Assay</th>
</tr>
</thead>
<tbody>
<tr>
<td>DuplexBS</td>
<td>5’ CTCCTCAGTTGCTCGACTCTGGCCTCAAATGA CCGTAGCTAAGCGCCTGTGAGAAGTATG 3’</td>
<td>EMSA/ATPase</td>
</tr>
<tr>
<td>DuplexTS</td>
<td>5’ ACTACCTGCGGACAGGTGACTACCGCTATT GAGCCAGTGCCGAAACTCTGAGGAGGAG 3’</td>
<td>EMSA/ATPase</td>
</tr>
<tr>
<td>(CNG)13TS</td>
<td>5’ ACTACCTGCGGACAGGTGACTACCGCTACNG CGNGCGNGCGNGCNGCNCGCNCGCNCGCNCGCNCGGTCTGAGGAG 3’</td>
<td>EMSA/ATPase</td>
</tr>
<tr>
<td>(CA)13TS</td>
<td>5’ ACTACCTGGCGAGCTATTGACTACCGCTACAC ACATTGGGAGCAGTCGAAAAACTTTGGAGGAGCG 3’</td>
<td>EMSA/ATPase</td>
</tr>
<tr>
<td>(CCC)12</td>
<td>5’ ROX-CCGGGCGCGCGCGCGCGCGCGCGCGG 3’</td>
<td>Thermal melting</td>
</tr>
</tbody>
</table>

*N represents either A, G, C or T.

This oligonucleotide was labeled at the 5’ end with biotin during synthesis for use in EMSA reactions.
mouse model for the CTG/CAG-expansion disorder, myotonic dystrophy type 1 (54). Furthermore, it has no effect on germ line expansion in another model of the same disease (49) or in either germ line or somatic expansion in a mouse model of Huntington disease, which results from expansion of the same repeat (51). Why different mouse models differ with respect to the role played by MutSβ and MutSo remains to be seen.

Analysis of the distribution of repeat contractions in Msh3−/− animals revealed the existence of two different contraction size classes, one involving the loss of just 1–2 repeats and a second class involving the loss of a much larger number of repeats (Fig. 3). As Msh3−/− animals have a distribution of contractions that is strongly shifted toward alleles that have lost fewer repeats, our data suggest that MutSβ is specifically involved in the generation of large contractions. The fact that MutSβ preferentially affects larger contractions but does not eliminate expansions altogether lends support to the idea that there are two different contraction mechanisms operating in the FXD mouse model, one that is MutSβ independent and one that is MutSβ dependent. It may be that the MutSβ-dependent contractions represent one outcome of the same process that gives rise to expansions. This would be of interest because it would indicate that while the mechanism responsible for expansion is strongly expansion biased, it can also generate contractions, albeit at lower frequency. This would have interesting implications for the underlying mechanism. The nature of the MutSβ-independent contraction pathway remains unclear.

To date there have been no other reports of MutSβ-dependent contractions in any other Repeat Expansion Disease mouse model, although the loss of MutSβ was shown to decrease contractions in repeat-containing minigene construct inserted into a human fibrosarcoma cell line (55). However, as contractions occur at such low frequencies in mice WT for MutSβ, it may be that larger studies in these models are needed to fully address this possibility. Whether MutSβ-dependent expansions and contractions arise via the same pathway in the FXD mouse model is not known. Our data showing that MutSβ binds well to both strands of the FX repeat suggest the possibility that both strands are potential substrates for MutSβ-dependent events (Fig. 4).

Our evidence that the kinetics of ATP hydrolysis is altered by loop-outs of either strand suggests that the conformation of MutSβ on binding to the loop-outs may differ from its conformation on binding to a canonical MMR substrate (Fig. 5). It has been suggested that the substrate-dependent regulation of the ATP hydrolytic cycle is critical for differentiating between different repair pathways (42–44). If this is in fact so, it may be that the altered kinetics of ATP hydrolysis relatively to a typical MMR substrate reflects the involvement of a different repair pathway in the generation of expansions and contractions. The effect of MutSβ on CGG-hairpin stability suggests that one way that MutSβ could act to promote both expansions and contractions is by increasing the longevity of the hairpins (Table 2), perhaps giving time for these structures to be channeled into an alternative pathway that generates expansions. We have previously shown that Polβ, a key polymerase in the base excision repair pathway, is also important for repeat expansion (56). However, while a hypomorphic POLβ mutation decreases the expansion frequency, our data suggest that it does not decrease the frequency of large contractions. One possibility is that MutSβ acts upstream of Polβ to bind the substrate for instability, channeling some bound substrates into the Polβ pathway to generate expansions while diverting others into a different pathway that results in contractions. The nature of the MutSβ-independent contraction pathway is unknown.
Binding of CGG- and CCG-repeats affects the kinetics of ATP hydrolysis by MutSβ. Steady-state ATP hydrolysis was measured using a TLC-based assay as described in the Materials and Methods section. MutSβ without any added DNA was used as a negative control. A oligonucleotide containing a (CA)$_3$ loop-out that elicits efficient ATP hydrolysis by MutSβ (45) was used as a positive control. Duplex is a fully Watson-Crick base-paired oligonucleotide, i.e. containing no mismatches. Its sequence corresponds to the duplex regions of the repeat-containing oligonucleotides (including (CA)$_3$) and serves as a control for MutSβ binding to the free ends of the duplex regions.

**Materials and Methods**

**Mouse maintenance**

The generation of the FXD PM mice was described previously (2). The Msh3$^−/−$ mice were generated previously (57,58), and cryopreserved embryos were obtained from the NCI Mouse Repository (Frederick, MD). Live born pups were generated from these embryos by implantation into the oviduct of pseudopregnant recipients using standard procedures. Multiple breeding pairs were set up from Msh3 heterozygous parents to generate the appropriate genotypes. Msh3$^+/−$ littersmates from these crosses were used as controls. Mice were maintained in accordance with the guidelines of the NIDDK Animal Care and Use Committee and with the Guide for the Care and Use of Laboratory Animals (NIH publication no. 85-23, revised 1996).

**Genotyping and analysis of repeat number**

Genomic DNA from mouse tails was extracted using KAPA Mouse Genotyping Kit (KAPA Biosystems, Wilmington, MA). Genomic DNA from other tissues was extracted using a Maxwell®16 Mouse Genotyping Kit (KAPA Biosystems, Wilmington, MA). Genomic DNA from mouse tails was extracted using KAPA Mouse Genotyping and analysis of repeat number using the dip.test command in the R diptest library.

**Purification of human MutSβ**

Human MutSβ was purified as described previously (45). Briefly, bacmids containing a region coding for His8-MBP-tagged human MSH2 and MSH3 were transfected into SF-9 cells in Grace’s Insect Cell Medium using Cellfectin II (Life Technologies, Grand Island, NY) according to the manufacturer’s instructions. MSH2 and MSH3 were then co-expressed by infecting Hi5 insect cells with both of the viruses. The cells were harvested at 48 h after infection, lysed in Buffer A [25 mM Hepes pH 8.0, 1 M NaCl, 30 mM imidazole, 10% (v/v) glycerol and 1 mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP)] and MutSβ affinity purified using a Ni$_2^+$ affinity column (GE Healthcare, Piscataway, NJ)]. The proteins were eluted with a step gradient of 30% (v/v) Buffer B (Buffer A+300 mM imidazole). The (His)$_8$-MBP tags were cleaved off using PreScission protease (GE Healthcare) and removed by passage over Amylose resin. The proteins were further purified using Heparin, Mono Q (GE Healthcare) and Superdex-200 columns. Proteins were stored in a buffer containing 25 mM HEPES pH 8.0, 100 mM KCl, 5 mM MgCl$_2$, 0.1 mM EDTA, 5–30% (v/v) glycerol and 1 mM TCEP at −20°C. The purity of the preparation was checked by SDS-PAGE gel electrophoresis of the Mono Q and Superdex-200 fractions. No contaminating proteins were seen in the MutSβ preparation, and equivalent amounts of MSH2 and MSH3 are present (Supplementary Material, Fig. S1).

**EMSA analysis of MutSβ binding to CGG, CCG, CTG and CAG-repeats**

The oligonucleotides used for this study are listed in Table 1 and were annealed as described previously (41). The bottom strand

### Table 2. Comparison of the free energy ($\Delta G$) at 37°C and enthalpy ($\Delta H$) required to denature the secondary structure formed by CCG-repeats.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$\Delta G$ (37°C) kcal/mol</th>
<th>$\Delta H$ kcal/mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA</td>
<td>1.88 ± 0.08</td>
<td>28.4 ± 0.9</td>
</tr>
<tr>
<td>MutSβ</td>
<td>3.46 ± 0.18</td>
<td>52.4 ± 4.1</td>
</tr>
</tbody>
</table>
oligonucleotide in each case (DuplexBS) was labeled at the 5’ end with biotin (Integrated DNA Technology, Coralville, IA). The binding reactions were carried out using the Gelshift™ chemiluminescent EMSA kit (Active Motif, Carlsbad, CA) according to the manufacturer’s instructions. Briefly, purified human MutSβ (0–4 nm) was incubated with 2 fmols of the duplexed oligonucleotides in a buffer containing 25 mM HEPES pH 8.0, 100 mM NaCl, 5 mM MgCl₂, 0.1 mM EDTA, 4% (v/v) glycerol, 2 mM TCEP and 0.2 mg/ml BSA and incubated for 10 min at room temperature. The complexes were resolved on a 6% (w/v) DNA-Retardation gel (Life Technologies) in 0.5× TBE at 4°C. The DNA and the DNA-protein complexes were transferred to a nylon membrane, UV cross-linked using a Stratalinker (formerly Stratagene, now Agilent Technologies, Santa Clara, CA). The membrane was then blocked, incubated with Streptavidin-HRP conjugate and the DNA detected as per the kit protocol.

**MutSβ ATPase assay**

The MutSβ ATPase activity was assayed in 15 μl of a buffer containing 25 mM HEPES (pH 8.0), 100 mM KCl, 2 mM TCEP, 5% glycerol, 5 mM MgCl₂, and 100 or 200 nM human MutSβ and 400 nm or 300 nm of the indicated oligonucleotide substrates. Kinetic data were obtained by varying the ATP concentration from 10 to 80 μM containing 50 fmol of α-32P-ATP (PerkinElmer, Waltham, MA) in each reaction. The reaction was incubated at 37°C for 15–30 min and terminated by addition of an equal volume of 50 mM EDTA. One microliter of each reaction mix was spotted on a PEI-cellulose TLC plate (Grace Davison Discovery Sciences, Albany, OR). Labeled ATP and ADP were separated by developing the TLC plate in 0.75 M KH₂PO₄ and visualized by storage phosphor autoradiography in a Typhoon TRIO Imager (GE Healthcare). ImageQuant TL software (GE Healthcare) was used to quantify the products, and GraphPad Prism (GraphPad Software, La Jolla, CA) was used to plot the graphs and calculate Kₘ and kₗ values.

**Thermal analysis of the CCG-repeats in the presence of MutSβ**

Fluorescently labeled (CCG)₁₀ oligonucleotide was resuspended in a buffer containing 25 mM HEPES at pH 7.3, 25 mM KCl, 1 mM MgCl₂, 0.1 mM EDTA, 4% (v/v) glycerol, 2 mM TCEP and 0.2 mg/ml BSA. The oligonucleotide (90 nm) was heated to 95°C for 5 min and then allowed to cool to room temperature to permit hairpin formation. Human MutSβ or BSA was added to a concentration of 360 nM as indicated. Thermal denaturation was monitored by measuring the change in ROX™ fluorescence at 608 nm in a StepOnePlus™ real-time PCR machine (Applied Biosystems, Carlsbad, CA) with a heating rate of 0.6°C/min. The data reported represent the average of three replicates. The thermodynamic parameters were derived from the melting curve using a two-state model (closed and open states).

**Supplementary Material**

Supplementary Material is available at HMG online.

**Acknowledgements**

The Usdin lab would like to acknowledge all the hard work by the people that take care of the mice used in this study. Without their help, this work would not have been possible. We would also like to thank Dr Bruce Hayward for his computational assistance.

Conflict of Interest statement. None declared.

**Funding**

The work described in this manuscript was funded by a grant from the Intramural Program of the NIDDK to K.U. (DK057808-07).

**References**


