

Manuscript EMBO-2011-77365

Atomic Force Microscopy captures MutS Tetramers initiating DNA Mismatch Repair

Piotr Marszalek, Yong Jiang

Corresponding author: Yong Jiang, Southeast University

Review timeline:

Submission date:	07 January 2010
Editorial Decision:	17 February 2010
Resubmission:	21 February 2011
Editorial Decision:	08 April 2011
Revision received:	08 May 2011
Accepted:	10 May 2011

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision

17 February 2010

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has now been seen by three referees, whose comments to the authors are shown below. As you will see, these referees generally agree that your conclusions on mismatch repair initiation by unequal MutS dimers, which would create DNA looping by one of them remaining at the lesion and the other one moving along DNA, are interesting and potentially important; they also commend you for the large amount and overall quality of the presented data. Nevertheless, in addition to a considerable list of specific, mostly presentational concerns (see the second part of referee 2's report), they all also raise a number of substantive experimental issues which, taken together, at this stage appear to preclude publication of the study in The EMBO Journal. While I am not going to repeat all these individual points of criticism in detail here, the main concern in this respect is that it is not clear whether the provided sets of data are sufficiently definitive to offer unambiguous support for the model proposed in the abstract. I am therefore afraid we do not see ourselves in a position to invite a revision of the study at the current stage. As it stands, a significant amount of further experimental data would be required to clarify the remaining ambiguities. Indeed, it is also not clear at this point whether the required substantiation can be achieved or whether new data would instead rather confound the current conclusions. Given the overall interest of your findings and the importance of the topic, I would nevertheless not exclude the possibility to look at a new version of this manuscript at some point in the future, should further studies in line of the referees' comments allow you to provide more definitive evidence to resolve the current controversies around the MutS mechanism of action. Such an improved manuscript would however have to be treated as a new submission rather than a revision (also with respect to the conceptual advance in light of the literature at the time of

resubmission) and only sent back to our referees if in our opinion the main problems had been largely addressed and the conclusions satisfactorily substantiated. At this stage of analysis, however, I am sorry to have to disappoint you.

Thank you for the opportunity to consider this manuscript. I am sorry we cannot be more positive on this occasion, but we hope in any case that you will find our referees' detailed comments helpful.

Yours sincerely,
Editor
The EMBO Journal

REFEREE REPORTS:

This paper uses AFM to examine the properties of MutS in the presence and absence of nucleotide cofactors and with mismatched and homoduplex DNA. The authors demonstrate that the presence of nucleotide cofactors significantly decreases the percent of monomer, and MutS exists as dimers and tetramers. They also show that MutS tetramers form loops on both homoduplex and mismatch DNA and that ATP increases the size of the loops. They note that in the loops on mismatch DNA, one of the dimers in the tetramer remains at the mismatch. This latter result differs from previous EM studies with much longer DNA, where both MutS dimers had moved away from the mismatch. In addition, in the presence of ATP, they found some loops that contained the protein roadblock (EcoR1 Gln111) within the loop. They propose a model for MutS interaction with the strand discrimination signal that is a hybrid of the sliding and looping models, and suggest that their model reconciles the apparent contradictions between these models.

The AFM data appear to be of very high quality and add to our understanding of MutS-nucleotide and MutS-DNA interactions; however, I am not sure if these data will resolve the controversy around the mechanism by which MutS communicates with the strand discrimination signal. There data do strongly support the models in which the tetramer is the major functional form. In addition, I have several concerns (detailed below) about the analysis of the data, although these concerns do not fundamentally alter their conclusions.

Major concerns:

A major conclusion of this paper is based on the observation that one of the MutS dimers remains at the mismatch as the loop gets larger in the presence of ATP. The distribution of lengths of the long leg, however, is quite broad (Fig. 4H) with lengths ranging from 30% to 70% of full length DNA, with >25% of the lengths being well outside the error of the position of the mismatch (lengths > 60% or < 40%). The standard deviation (SD) of this distribution is 11% of the full length not the 1.1% shown in the figure and in Table 2. In fact, all of the SDs listed in the Table 2 and in Fig. 4 appear to be a factor of 10 too small. Comparison of the fits in Figs. 4B-4E reveal this error, as it is obvious that the distributions are broad, which is not surprising. Examination of Table 2 reveals another striking observation, which is that the length of the long leg on homoduplex DNA is also within error of the position of the mismatch (which does not exist on this piece of DNA). Its length is 51.1%. Because the mismatch is at 47% from the nearest DNA end, it is 53% from the other end. Since there is no way to tell one end of the DNA from another the mismatch position could be either 47% or 53% from the end in the looped complexes. It seems that all of the long leg lengths are within error of the mismatch even when the mismatch is not there. It seems as though the DNA has a propensity to bend in this region (many of the free DNAs in Fig S5A appear bent), which would promote loop formation. Such a propensity to bend could explain the high observation of loops in the absence of ATP and why the loops show such a narrow distribution of size and position even on homoduplex DNA. These observations make one wonder whether the results would be the same on a different DNA sequence.

Another major point is that they observe loops (~25%) in which EcoR1-Gln111 is inside the loop. It is not clear that this observation supports diffusive looping on the time scale of repair because the incubation is for 10 minutes, which would be plenty for EcoR1 to dissociate and rebind. Modrich and coworkers discuss this possibility in their study of MutH activation using the same (or a very similar) DNA (Pluciennik and Modrich. Proc Natl Acad Sci USA (2007) vol. 104 (31) pp. 12709-13).

Other specific comments:

On page 5, they state "MutS volumes fall into three independent distributions with peak values of 63 ± 0.3 , 188 ± 2 and 414 ± 11 nm³ (mean \pm SEM)". I am not sure how these errors were calculated; however, it is clear from the distributions in Fig. 1C that the breadth of the Gaussian distributions are significantly greater than these errors. Clearly, they do not know the volume of the monomer to ± 0.3 nm³; it appears to be closer to ± 25 nm³. The error should be given as the standard deviation of the Gaussian.

In Figs 2 and 3, they present statistics of volume analysis of MutS under different conditions. It would be useful to discuss how the fraction of dimer and tetramer in this study compares with the previous biochemical studies that have measured the dimer-tetramer association constants. I believe their data agree fairly well with those of Modrich and co-workers. It would also be useful to point out that the bar graphs are percentage of MutS monomer as each species. The authors should also cite some of the previous studies that demonstrated that AFM volume depends linearly on MW. Also plotting the MW as a function of AFM volume (Fig. S1) seems like an odd way to present the data, since the independent variable is usually on the x-axis. (Note, in Fig. S1, Ferritin is listed in the table but not shown in the figure.) They should also present a more complete description of how the volume was calculated.

On page 9, I do not understand the following statement "Comparison of Panels (A) and (E) in Figure 4 shows that in the absence of nucleotide, fewer than 5% of the loops had an end that was more than 5% of the DNA length distant from the mismatch, while in the presence of ATP, a significant percentage of the MutS molecules bound in loops were located $\sim 10\%$ of the DNA length away from the mismatch"

On page 11, they discuss the binding affinities of MutS for the different DNAs (Table 3) In the supp. Info., they define KD as $[MutS][DNA]/[MutS-DNA]$ where $[MutS]$ is the concentration of monomer, if I understood correctly. Because MutS binds DNA as a dimer and a tetramer, the observed binding constant should be the sum of the association constant ($1/KD$) for MutS binding as a dimer and a tetramer with $KD_{dimer} = [MutS_2][DNA]/[MutS_2-DNA]$ and with $KD_{tetramer} = [MutS_4][DNA]/[MutS_4-DNA]$. Consequently it is the concentration of free dimer and tetramer that are important not the concentration of free monomer. More importantly, even when I use their equation given the supp. Info., I do not arrive at the KDs listed in Table 3 based on the occupancies listed in Table 3, and there does not seem to be a good correlation with the occupancies and the binding constants. For example, in the absence of nucleotide on GT DNA the occupancy is 30.1% and the KD is 22 nM, and with blocked ends and EcoR1 bound, the occupancy is 65.7% and the KD is 21 nM. They should be explicit of how the occupancies are used to calculate the binding constants and they should correct their definition of the binding constant. Notably, they have all the data they need to calculate the "true affinities". On another note, the errors listed seem unrealistically small. I don't see how one can have an error of $\pm 0.1\%$ when less than 1000 DNAs were analyzed. How were these errors calculated? All that is state is that they are SEM.

Do the occupancies listed in Table 3 include MutS bound to the ends of the DNA? If so what % is bound to the ends. This may be important given that Fig. S4 shows a significant % of MutS is bound at DNA ends even when both ends are supposed to be blocked by streptavidin. Also what % of the DNA fragments have both ends bound by streptavidin.

On page 11, they state "this resulted in a large decrease in the observed binding constant in the presence of ATP from 64 ± 1 nM to 25 ± 1 nM." While this decrease is significant, it is certainly not large (only 2.5 fold).

Referee #2 (Remarks to the Author):

General Comments

This manuscript by Jiang and Marzalek presents a large number of technically demanding experiments - quite well executed and in most but not all cases nicely presented. The mechanism(s) of DNA mismatch repair (MMR) remain poorly understood, even in the well-studied *E. coli* system. A particular stumbling block is how recognition of mismatches by MutS is coupled to downstream steps, beginning with communication with a distant strand-specificity markers. The work here has the potential to suggest some answers this question. Besides (i) a number of writing

and presentation problems however, there are (ii) several "bigger-picture" problem areas where the authors have not explained themselves clearly or have not come to grips with their data, or have not constructed and explained (a) plausible model(s). Items (i) are addressed by my detailed "Specific Comments" below. Items (ii) are addressed immediately below; in no particular order. There is no discussion of the significance of the substantial amount of free MutS monomer in the absence of nucleotides, and its sharp decrease in the presence of all nucleotides. If this has all been explained before, say so and provide the references. Also, what is the significance of the MutS 800 data? Is MutS 800 active in vivo, despite its lack of tetramerization?

1. It's not clear why there are more tetramers in the presence of DNA. Are most of the tetramers now bound to DNA? If so, why tetramerization the same for G/T and homoduplex DNA? If most tetramers are in solution, how did DNA cause this?
2. Despite claims about MutS being "preferentially" near the mismatch, MutS seems mostly tens to hundreds of bp away from the mismatch site. Why is this?
3. The specificity of MutS for G/T vs. homoduplex seems surprisingly low. Why should there be substantial loop formation on homoduplex DNA? Why are the KD values for G/T and homoduplex DNA so similar? It's not enough to make a brief statement about "per site" vs. "per entire DNA molecule". What is "true specificity" for a mismatch site? What is the significance of the binding to homoduplex DNA vs. the homoduplex regions (99.9%!) of (G/T) DNA? How might this relate to searching DNA for mismatches? Much more thought and cogent discussion are needed.
4. Much more coherent considerations of loops, their growth, and their relationship to MutS positions are needed.
5. The entire set of EcoRI experiments is a mystery.
6. The authors seem to be proposing a MutS translocation and diffusion model based to some extent on EcoP151, but the MutS model is very poorly described. Fig. 6 is worse than nothing as it stands. How could diffusion (with no source of energy) ever start a loop against DNA bending stress? Could translocation start the loop and diffusion of ATP-bound MutS extend it, riding down a bending-stress gradient? What would the occasional ATP hydrolysis do? Nudge MutS in the "right" direction? The authors should spend more effort on a coherent model and less effort on comparisons with the old Allen/Modrich work.

Specific Comments

Page 3, line 2: Iyer misspelled. 20-24: long; awkward sentence.

Page 4, l 1-3: Doesn't distinguish between ATP-dependent diffusion and hydrolysis-dependent unidirectional translocation. 8: by AFM (not defined yet). 8: Gels are not "retarded", (and do not "shift"). DNA is retarded; mobilities shift. 17: "studies" don't observe; people do. "Researchers think" sounds like a newspaper article. 22-24: bad syntax. Do you mean a "relatively few" AFM investigations focused on recognition? 25-26 (and next page): Aren't you also focusing on recognition mechanisms?

Page 5, l 4: How much "more frequently"? 14: Here and throughout the manuscript the insidious practice of heading sub-sections with didactic-data-interpretation statements masquerading as pseudo-titles works against balanced reading and understanding of the manuscript. The place for these is at the beginning of the discussion. The title is in fact incomplete as a summary of the "take-home lessons". It leaves out the important MutSA800 and nucleotide results and the substantial weight fraction of tetramer. Something like Oligomerization or association state of MutS in solution is preferable. 6: No DNA was present here. 8: Some readers will not know what "flooded image" means. 21-23: a nice result.

Page 6, l 1-5: Nice work! Why might the tetramer MW estimate be so far off? 11-12: Are the ATM volumes different? 16: Fig 1D. 22-24: all nucleotides (not necessarily actual "co-factors"); vs. no nucleotide. 12-14: Confusing (contradictory) wording. Nowhere is the surprising amount of monomer discussed. Do the authors think that that much protein remains undimerized in vivo, or is the no-nucleotide datum irrelevant?

Page 7, l 6-7: What is the potential significance of the low oligomerization of MutS 800? Are there important references? 9: This pseudo-title is even worse.; details such as "201 bp or longer" and adjectives such as "main" are inappropriate. Just summarize the topic and approach and leave the conclusions for the discussion! 10-12: It might be helpful to see the M-D-T distributions pulled out together for at least no-NT and ATP, for A/T vs. G/T DNA. What does the much higher monomer for no-NT mean? 13-14: IMPORTANT! How much of the MutS was bound to DNA? Are we looking at tetramers bound to DNA and (at least some) dimers in solution, or what? 12-19: Is DNA length intrinsically important or just the total number of bp? The author must compare 1000 nM 31-bp DNA with 200 nM 201-bp DNA for the DNA size comparisons to mean anything at all. Why are G/T and A/T results so similar? Is MutS mostly a tetramer even on homoduplex DNA? 20.

120-bp. A terrible title! 22: Especially here, the reader should be reminded that this is a linear substrate.

Page 8, l 2-3: This sentence mixes a speculation (should be "might" instead of "could") not supported by any data. "See below" presumably means the next section. 6: Avoid the jargon "prebound", which is really an adjective. "Previously bound" is correct. 9-10: Because Fig. S4 is so poorly labeled, it's hard to evaluate this statement. I think the authors mean for the x axis to represent the position of the MutS, shown as the fractional length of the short leg. Do these positions refer to the same end, or does 46% (presumably the mismatch position from the "left" end in Fig. S3) in fact always refer to the left end? The authors never tell us. In Fig. S4 Are 46% or 56% both from the left end? If so, these data are not very useful. I assume that y axis actually represents the fraction of short legs that have a particular length. In any event, the statement about "located very close to the mismatch" is totally unjustified. The location seems to be up to 150 bp in either direction with no nucleotide and all over the place in ATP. Do the substantial fractions of MutS at position zero reflect binding to ends despite the streptavidin? 14-15: DNA "loops" cannot "consist of...MutS dimers". 15-16: What is the evidence that the strands "overlap"? It seems at least as likely that they form an " " structure (Fig. 6). L18: The worst pseudotitle yet: two conclusions and a semi-colon. 25: "Single-site-bound".

Page 9, l 8: This short pseudotitle states only a single result about loop size and ATP, yet the long section describes important measurements of loop position, effects of end blocks, time, temperature, nucleotide cofactor and internal roadblocks. Thus, the reader is not given an accurate idea of what to expect. Also, "increase" might mean growth kinetics, which are in fact not shown. 10: 1120-bp. 12: "of" not "on". 14: "double streptavidin-end-blocked" is clumsy.

Page 10, l 1-5: Confusing and misleading sentence. Are the authors just stating the obvious fact that the DNA in bigger loops has to come from somewhere? If both dimers remain together, and one of them remains "centered" on the mismatch (the spread is actually 1-200 bp or so), it seems more accurate to say that DNA from the short leg or long leg (why is it in fact usually the long leg?) moved past the MutS tetramer(dimer) and into the loop. 6: "Previously incubated". 15-16: badly written sentence. 16-16: Not obvious from Table 2. Why does homoduplex DNA generate such big loops? How frequent are they (vs. G/T DNA)? 20: not "prebound"! 23-24: Confusing. Is the comparison to experiments without EcoRI? What fraction of ERI molecules are or are not where they belong? Is the nearby ERI site likely to be in the loop?

Page 11, l 3-4: another bad pseudo-title. 6: data are. 5: It's not a "binding", but a dissociation, constant. "Were", not was. 5-14: Can the authors briefly explain the low specificity for G/T DNA, without forcing us to wade through both Materials and Methods descriptions? Does it reflect the large amount of homoduplex DNA in the G/T substrate as well? Would shorter DNA substrates show higher specificity? 8-9: Which of the two results "is similar to K⁺Ds..."? 14-15: Tell the reader the significance of this. Why is the (G/T) preference still only about 5-fold in these cases? 19-22: What's the significance of this? 22-24: Why might ERI improve binding and looping? 18: dissociation constant.

Page 12, l 3: After this complex data presentation, it is essential to begin the discussion with a summary of the major take-home lessons. The present pseudo-titles from RESULTS might work, although the authors might add other findings that they consider highly important. Next, discuss each take-home lesson in order, bringing in only directly relevant background material. A few subsections, with real titles, might help organize the discussion. 5-6: A "constant" can't itself be controversial. 8-9: Just say 100 nM MutS (as monomer). 15: But the monomer concentration was strongly affected. Why? 20-25: The similarity of the effects of homoduplex and G/T on dimer tetramer suggest that MutS scans/tests DNA for mismatches as a tetramer. The authors never address this important point. Were all or most of the tetramers on DNA, or might they form on DNA, but then dissociate into solution as tetramers (an important mechanistic question)?

Page 13, l 7-15: As written, this paragraph will really confuse readers who are not MMR experts. The important question of specificity (as measured by AFM) requires careful attention. See my comments on l 12-19 in p.7. If increasing DNA length increases apparent binding to homoduplex DNA, perhaps because the chances of being hit by MutS are greater, and binding to homoduplex DNA is strong (as it certainly is), then the added DNA should cause binding (1/kD) to (G/T) DNA to increase with length, but it apparently doesn't. Something of highly fundamental significance may be going on here, and the authors should vigorously pursue it. l 26 and p. 14, l 1-2: This is very confusing. Besides cleaner writing, a model diagram (perhaps added to Fig. 6) is needed. ("non-end-blocked and non-internally-blocked are clumsy"). Do the authors think that MutS in a search mode "bounces off" ERI and moves back towards the mismatch? Is the MutS trapping done by ERI in this loop or outside the loop? Why are (G/T) occupancies higher with ERI binding but KDS about the

same? What does it mean that ERI affects (G/T) occupancy much more than homoduplex? Unless the authors can make some biochemical sense of the ERI results, they should be dropped. P 13, l 26: Why does this sentence come right after the ERI results?

Page 14, l 19: this "Allen et al's" is a horrible construction. 21-22: somewhere the authors need to seriously explain the significance of all of those loops on homoduplex DNA. In general however, far too much space is devoted to comparisons with the quite different Modrich experiments. 17-19 : sentence is not clear. What does "this" refer to?

Page 15, l 1-3: Do you mean 10 minutes before deposition? 4: With "no ATP" vs. after "10 sec with ATP". 3-7: Considerable clarification is needed. Where are the relevant data? Does the rate calculation (600 bp/min) come from the 10-min pre-deposition incubation in ATP? What happens in 10 min without ATP? The authors should either present a reasonably complete rate study, or drop all discussion of these very scanty data. 13-16: A long awkward sentence. Again, too much space is devoted to extended comparisons with the Modrich work. 19: "induced by MutS using ATM" is confusing. 19-20: confusing sentence. 20-22: this strong conflict with the work presented here must be discussed seriously.

Page 16, p 15, l 24-25 and P 16 l 1-10: The Eco P151 discussion is confusing without a figure. What is the significance of a low ATP hydrolysis vs. translocation ratio (abruptly introduced for the first time here)? 1-4: bad sentence, but important points. 4-7: Not all clear, but again important. 7: "possible/plausible" is bad writing. 11: "Data support". 11-20 Confusingly described and inconsistent. Fig 6 (which should be accompanied by an informative legend) is no help at all. Is the "translocation" really movement of the DNA "through a second dimer still attached to the mismatch-bound dimer? How can loops grow rapidly if the "translocation" is actually "diffusive" bidirectional sliding? Unidirectional translocation requires energy. 16-18: Contradictory and confusing. What is "ATP-hydrolysis-driven loop formation (by Mut tetramer?)? How does it "combine" with "a diffusive loop mechanism". Fig. 6 is useless here. What is "looping" without ATP? How can the loop grow without ATP hydrolysis? Where is the hydrolysis-dependent step? The authors need a much more detailed figure, more explicitly incorporating the EcoP161 model and with insightful descriptions.

Could the decrease in DNA-bending stress drive diffusive loop expansion after a hydrolysis-dependent unidirectional "translocation" initiated the loop? How often would MutS-bound ATP hydrolyze during loop expansion?

Figure 1. An equivalent of Fig. 1C (M-D-T distribution) is needed for at least one case with nucleotide. Put "MutS" into C and D ordinate labels. Remove gray background and horizontal lines from D; use "ticks" along ordinate instead.

Fig. 2. Remove gray background and horizontal lines

Fig. 4. I like panels A and E. Label x-axis as "position" or "coordinate". Why fractions on left, but %s on right?

Fig 5. Switch positions of bars in C to line up with A and B. Legend inside figure is redundant.

Fig. 6. Completely inadequate (see above).

Page 17, l 9-12: Since much MMR work has been performed using circular substrates, the linearity of the 1120-bp (not hyphen) substrate should be noted. 9: Maybe "Linear 41-bp..." 12-16. Better description needed. Is there an annealing step somewhere?

Page 18, l 1-2.: This is not made clear in the "results" section. 12: "Prior binding by EcoRI."

Page 20, l 8-11: How many, actually?

Figure S3E. Label the ordinate. Show bp instead of ov in addition to nm.

Fig. S4. "Percentage (%)" labels on Y axis are both redundant and confusing; it's endpoint position on the DNA or length. In B and D, its fraction of short legs (with indicated lengths/endpoints). The mismatch positions should be clearly indicated.

Fig 4 legend. l 3-4: Previously bound. 1120-bp

Fig. S2 legend. Also MutS 800 in (B)?

Fig. S3 legend. 1120-bp. Line 4: Some is unscientific and vague. Could be read as a mixture of three substrates. Otherwise good figure and good legend.

Fig. 54. Legend. Line 11: These distributions don't look at all preferential to me.

Table 1: What are the differences between photos A and B, and among C-E? Was ATP present in the spreads shown? The alignment of the data columns under the pictures is very confusing. Do the volume and MW measurements correspond to all molecules in the spreads shown and others like them, or to selected apparent dimmers?

Table 2. What fractions of the total molecules, especially the homoduplexes, showed the loops that were measured? Position, not "% of length", which is not clear.

Table 3. "(+) and (-)" nomenclature is not correct, despite its' wide (mis)use in the literature and in

fact is harder to read. For most columns, n(none) instead of (-) would be better, since its nucleotide or streptavidin or EcoRI that's added ("+"). Nothing is subtracted ("-"). The biotin ends and streptavidin columns are redundant, and in the case of (+) vs (++) streptavidin, confusing. You can't add ("+") a mismatch to DNA, and certainly can't subtract ("-") mismatch. Use a different nomenclature. These are not "binding", but dissociation constants. A column to indicate -fold binding specificity would

Referee #3 (Remarks to the Author):

Review of Jiang and Marszalek

The authors report results that they interpret to yield two conclusions: 1) The functional unit in DNA loop formation by MutS is a dimer of non-equivalent dimers, and 2) one of the dimers remains at or close to the mismatch site, while the other moves relative to DNA. If these conclusions are correct, the paper provides an important step forward in understanding how MutS works. The model that results is simple and quite plausible: One MutS dimer locks onto the mismatch site, while the other, remaining linked to the first dimer, slides along DNA until a sequence is found that signals cleavage of the daughter strand.

The results are pretty convincing, but I do have some worries. Primarily, I wish that the authors had done a second construct in which the mismatch site is significantly displaced from the center of the DNA molecule. There may, for example, be physical reasons why the base of the loop prefers the center of the molecule. It would be very reassuring to see the base of the loop move to one side in the alternate construct. Also, I have some concern about the significant broadening of the "long leg" distribution in the presence of ATP (Fig 4H). Has the ability of MutS to "home in" on the mismatch been reduced?

I know how difficult it can be to do repeat experiments when the primary authors have left the laboratory. This paper could be published with the present data, but it could also be significantly strengthened. I leave it up to the editor to negotiate this issue with the authors.

Some minor matters: In the introduction to the paper, one gets the sense that observation of DNA loops with MutS is original to this paper. I know that the authors do not think this, but that is the impression given. It would be good to get a reference to Allen et al (and anyone else who has made such reports) up front in connection with loops.

The authors sometimes use "binding constant" and KD (dissociation constant) interchangeably. That is very confusing; one is the reciprocal of the other. Take this excerpt from page 11: ". . .this resulted in a large decrease in the observed binding constant . . . from 64 . . . nM to 25 . . . nM". They obviously mean KD here. Say "dissociation constant" instead of binding constant.

The authors should state whether the two apparent forms of tetramers could be a consequence of different orientations on the surface.

The authors have shown that ATP hydrolysis is not required for loop formation by MutS, yet they mention that this mechanism could work in combination with hydrolysis-driven loop formation by MutS (page 16). These two mechanisms seem contradictory. Some explanation is needed.

Resubmission

21 February 2011

Thank you very much for conducting a review of our manuscript (ID#: EMBOJ-2010-73591) "Atomic Force Microscopy Captures MutS Tetramers Initiating DNA Mismatch Repair" by Yong Jiang and Piotr E. Marszalek. We followed your recommendation carefully, as well as all of the comments and suggestions from the referees. This is a significantly improved version of the manuscript we are submitting for your consideration. The most important changes in this version,

include new experimental results and their analysis which (new Figs. 6 and 7) provide strong, additional evidence supporting our original conclusions on the nature of the mechanism of mismatch repair by MutS, which has been deciphered. In addition, to clarify various minor issues which were raised by the reviewers we revised, and significantly expanded Supplemental Information by adding 4 new figures.

In the old manuscript the results were obtained on the mismatch-carrying DNA substrate that was symmetrically labeled at both ends with biotin/streptavidin to block MutS from leaving the DNA. However, this symmetric labeling, as pointed out by our reviewers, made it impossible to precisely identify the ABSOLUTE location of the mismatch site (that was engineered to be located 53% (of the total DNA length) from one end. This in turn made it impossible to conclude unequivocally that MutS resides at the mismatch site, as it could have been located 6% off of this sites (the position of 53% from one end and of 47 % from the same end could not be uniquely identified because the terminals could not be uniquely identified). In the revised manuscript we present new results obtained on the DNA that was labeled with streptavidin at one end only, allowing this end to be easily identified. The mismatch site was positioned 53% from this labeled end and thus the location of MutS relative to the mismatch site could be precisely located by AFM imaging. AFM images of DNA MutS complexes captured in the absence and presence of ATP (New Figs. 6 and 7) confirmed our original conclusion that, when ATP is added, one of the MutS dimers remains at or very close to the mismatch site, while the other MutS dimer moves while extruding a growing DNA loop. These new results significantly clarify our original observations and remove previous uncertainty as to the location of MutS relative to the mismatch site in our measurements, thus addressing the major concern of the reviewers.

The changes that we introduced in the manuscript are marked in red. They are explained in detail in the attached document which also contains our detail responses to all of the referees' comments.

We hope that our revised manuscript will now be acceptable for publication.

Response to the Comments and Suggestions from the Referees:

Referee #1

Major Concerns:

1.1 "A major conclusion of this paper is based on the observation that one of the MutS dimers remains at the mismatch as the loop gets larger in the presence of ATP. The distribution of lengths of the long leg, however, is quite broad (Fig. 4H) with lengths ranging from 30% to 70% of full length DNA, with >25% of the lengths being well outside the error of the position of the mismatch (lengths > 60% or < 40%). The standard deviation (SD) of this distribution is 11% of the full length not the 1.1% shown in the figure and in Table 2. In fact, all of the SDs listed in the Table 2 and in Fig. 4 appear to be a factor of 10 too small. Comparison of the fits in Figs. 4B-4E reveal this error, as it is obvious that the distributions are broad, which is not surprising."

Reply: We followed the referee's suggestion and replaced all Standard Errors of the Mean (SEM) with standard deviations (SD), which on average are about a factor of 10 greater than SEMs.

We agree with the reviewer that the distribution of lengths of the long leg is quite broad. There are several factors contributing to this spread: first, most of our results were obtained on the DNA of which its ends were blocked by streptavidin, therefore could not unambiguously identify the ends. As a result the uncertainty of a mismatch position was about 6% of DNA length (the mismatch is located at a position 47% from one end and 53% from the other). We estimate that AFM measurement errors are on the order of 20 nm, contributing another ~5% to the uncertainty of the length of the long leg. Thus, it is expected that the length of the long leg may vary between 40-60% of the full DNA length, consistent with the reviewer's own estimates. Second, we think that the measurements that fall outside of this range captured those cases in which MutS dimers were bound nonspecifically even on heteroduplex DNA (it is impossible to prevent such binding events). This is consistent with the observation that loops were formed not only on hetero- but also on homoduplex DNA. In fact, the fraction of homoduplex DNA molecules displaying such loops was 2.4% (new

Table 3) compared to 7.6 % of heteroduplex DNA with the loop features. Thus, it is expected that about 30% of the loops on heteroduplex DNA may be formed on sites that do not involve the mismatch (2.4% out of the 7.6%), and at least in some of these cases the length of the long leg may fall outside of the 40-60% region.

Importantly, when we used the DNA heteroduplex substrate that was blocked with streptavidin only at one end (allowing us to identify the position of the mismatch that was engineered to be 53% length from the labeled end, new Fig. 6 and 7) the distributions of lengths of the long leg became narrower and their averages coincide well with the position of the mismatch: $46 \pm 5\%$ when the loop was formed between the mismatch and the DNA end blocked by streptavidin, and $53 \pm 3\%$, when the loop was formed between the mismatch and the unlabeled DNA end. In summary, based on these new results we are quite confident that the length of the long leg is consistent with the position of the mismatch.

1.2 "Examination of Table 2 reveals another striking observation, which is that the length of the long leg on homoduplex DNA is also within error of the position of the mismatch (which does not exist on this piece of DNA). Its length is 51.1%. Because the mismatch is at 47% from the nearest DNA end, it is 53% from the other end."

Reply: We agree with the reviewer's observation. However, we note that the length of the long leg may be on average equal to 50% of DNA length even if MutS particles in the loop configuration are bound randomly on homoduplex DNA. Thus, Table 2 made a false impression that on homoduplex DNA the end of the long leg is within the position of the mismatch (which is not there).

In the revised manuscript, Supplemental Figures S9 shows the length distributions for MutS bound in loop configuration on homoduplex DNA. We note that the loops formed on homoduplex DNA are different from those on G/T mismatch substrates. First, the standard deviation of the long leg (15%) on homoduplex DNA is significantly greater than the 5% determined on heteroduplex DNA using the same experimental conditions. Second, on G/T mismatched DNA, MutS always has a higher binding affinity for the mismatch area (between 45% and 50%) compared to other locations, as seen in Supplementary Figure S3B, S3D, S4A and S4B. However, as Supplementary Figure S4C shows, MutS has no increased binding affinity for the region between 45% and 50% of homoduplex DNA length. Third, loops were seen on homoduplex DNA with a frequency about 6 times less than that on heteroduplex DNA at the same conditions (Table 2). Also, the size of the loop on homoduplex DNA ($40 \pm 15\%$) is much bigger than that on heteroduplex DNA ($22 \pm 5\%$), and its distribution is 3 times broader (Table 2). A plausible explanation for this observation is that on homoduplex DNA none of the two MutS dimers that formed the loop was forced to stay at the center area of DNA.

We have added one paragraph in the RESULTS section and another in the DISCUSSION section that reflect these observations.

1.3 "Since there is no way to tell one end of the DNA from another the mismatch position could be either 47% or 53% from the end in the looped complexes. It seems that all of the long leg lengths are within error of the mismatch even when the mismatch is not there."

Reply: We added new figures in the main manuscript (Figure 6 and 7) which report our measurements on heteroduplex DNA containing a biotin-streptavidin complex at a single DNA end. This DNA allows us to identify the position of the mismatch that was engineered to be 53% length from the labeled end, new Fig. 6 and 7. The distributions of lengths of the long leg determined for this substrate became narrower and their averages coincide very well with the position of the mismatch: $46 \pm 5\%$ when the loop was formed between the mismatch and the DNA end blocked by streptavidin, and $53 \pm 3\%$, when the loop was formed between the mismatch and the unlabeled DNA end. So, based on these new results we are quite confident that the length of the long leg is consistent with the position of the mismatch. We added a new paragraph at the end of the RESULTS section that reports this new data.

1.4 "It seems as though the DNA has a propensity to bend in this region (many of the free DNAs in Fig S5A appear bent), which would promote loop formation. Such a propensity to bend could explain the high observation of loops in the absence of ATP and why the loops show such a narrow distribution of size and position even on homoduplex DNA. These observations make one wonder whether the results would be the same on a different DNA sequence."

Reply: This is an important observation and we reflected this point in the revised DISCUSSION section.

"In the absence of any nucleotides, the loops are quite small and contain approximately 250 bp of DNA. The DNA seems to have propensity to bend in this region (Wang et al, 2003), and this may promote loop formation for MutS interacting with DNA at two sites, mismatch and a nonspecific site or even two nonspecific sites. In future project we will test whether a different DNA sequence with a lesser propensity for DNA bending may affect the frequency at which loops are formed."

1.5 "Another major point is that they observe loops (~25%) in which EcoRI Gln111 is inside the loop. It is not clear that this observation supports diffusive looping on the time scale of repair because the incubation is for 10 minutes, which would be plenty for EcoRI to dissociate and rebind. Modrich and coworkers discuss this possibility in their study of MutH activation using the same (or a very similar) DNA (Pluciennik and Modrich. Proc Natl Acad Sci USA (2007) vol. 104 (31) pp. 12709-13)."

Reply: We agree with the reviewer that for the reason stated this result does not necessarily support diffusive looping. To reflect the referee's observation we added a short paragraph in the DISCUSSION section.

"On DNA previously bound with EcoRIE111Q and subsequently incubated with MutS and ATP, 74% of loops were seen which did not contain EcoRIE111Q within the loop contour, while 26% of them contained EcoRIE111Q inside the loop contour (Figure 5). Considering the residency half-life of EcoRIE111Q at a d(GAATTC) sequence is less than 40 minutes (Pluciennik & Modrich, 2007), part of EcoRI may dissociate and rebind within 10 minutes incubation in our experiments. Consequently, EcoRIE111Q could not fully prevent the translocation of MutS on DNA and as a result some loops could grow beyond the EcoRI site."

Other Specific Comments:

In each case we followed the referee's suggestions and revised the manuscript accordingly.

1.6 "On page 5, they state "MutS volumes fall into three independent distributions with peak values of 63 ± 0.3 , 188 ± 2 and 414 ± 11 nm³ (mean \pm SEM)". I am not sure how these errors were calculated; however, it is clear from the distributions in Fig. 1C that the breadth of the Gaussian distributions are significantly greater than these errors. Clearly, they do not know the volume of the monomer to ± 0.3 nm³; it appears to be closer to ± 25 nm³. The error should be given as the standard deviation of the Gaussian."

Reply: We followed the referee's suggestion and changed all Standard Errors of the Mean (SEM) to standard deviations.

1.7 "In Figs 2 and 3, they present statistics of volume analysis of MutS under different conditions. It would be useful to discuss how the fraction of dimer and tetramer in this study compares with the previous biochemical studies that have measured the dimer-tetramer association constants. I believe their data agree fairly well with those of Modrich and co-workers."

Reply: This is another good suggestion. We followed the referee's suggestions and added a sentence reflecting this observation to the DISCUSSION section.

"The association constant for the dimer-to-tetramer transition is about 1.3×10^7 M⁻¹, which is close to the value (2.1×10^7 M⁻¹) obtained by previous biochemical studies (Bjornson et al, 2003)"

1.8 "It would also be useful to point out that the bar graphs are percentage of MutS monomer as each species."

Reply: We followed the referee's suggestion and added the sentence to the figure captions. "The bar graphs are percentages of MutS monomer as each species."

1.9 "The authors should also cite some of the previous studies that demonstrated that AFM volume depends linearly on MW."

Reply: We followed the referee's suggestion, and added a new reference,

"Ratcliff GC, Erie DA (2001) A novel single-molecule study to determine protein-protein association constants. Journal of the American Chemical Society 123(24): 5632-5635"

1.10 "Also plotting the MW as a function of AFM volume (Fig. S1) seems like an odd way to present the data, since the independent variable is usually on the x-axis. Note, In Fig. S1, Ferritin is listed in the table but not shown in the figure."

Reply: We followed the referee's suggestion, and redrew Figure S1. We also removed Ferritin from the table.

1.11 "They should also present a more complete description of how the volume was calculated."

Reply: We followed the referee's suggestion, and expanded our description of how the volume was calculated in the MATERIALS AND METHODS section. We also added a reference, which contains a very detailed description of how the volume is calculated.

"Horcas I, Fernandez R, Gomez-Rodriguez JM, Colchero J, Gomez-Herrero J, Baro AM (2007) WSXM: a software for scanning probe microscopy and a tool for nanotechnology. Rev Sci Instrum 78(1): 013705."

1.12 "On page 9, I do not understand the following statement 'Comparison of Panels (A) and (E) in Figure 4 shows that in the absence of nucleotide, fewer than 5% of the loops had an end that was more than 5% of the DNA length distant from the mismatch, while in the presence of ATP, a significant percentage of the MutS molecules bound in loops were located ~10% of the DNA length away from the mismatch.'"

Reply: We removed this complicated/confusing sentence.

1.13 "On page 11, they discuss the binding affinities of MutS for the different DNAs (Table 3) In the supp. Info., they define KD as $[MutS][DNA]/[MutS-DNA]$ where $[MutS]$ is the concentration of monomer, if I understood correctly. Because MutS binds DNA as a dimer and a tetramer, the observed binding constant should be the sum of the association constant ($1/KD$) for MutS binding as a dimer and a tetramer with $KD_{dimer} = [MutS_2][DNA]/[MutS_2-DNA]$ and with $KD_{tetramer} = [MutS_4][DNA]/[MutS_4-DNA]$. Consequently it is the concentration of free dimer and tetramer that are important not the concentration of free monomer."

Reply: For calculations of the dissociation constant, we used the concentration of MutS dimers and we considered MutS tetramers as two associated dimers. We added a sentence in supplemental material to clarify this point.

"MutS here refers to a dimer because we think that a MutS dimer is the basic unit when binding to DNA and a MutS tetramer is considered as two associated dimers."

1.14 "More importantly, even when I use their equation given the supp. Info., I do not arrive at the KD s listed in Table 3 based on the occupancies listed in Table 3, and there does not seem to be a good correlation with the occupancies and the binding constants. For example, in the absence of nucleotide on G/T DNA the occupancy is 30.1% and the KD is 22 nM, and with blocked ends and EcoRI bound, the occupancy is 65.7% and the KD is 21 nM. They should be explicit of how the occupancies are used to calculate the binding constants and they should correct their definition of the binding constant. Notably, they have all the data they need to calculate the 'true affinities'."

Reply: We followed the referee's suggestion and expanded the supplemental method section and explained in detail how we calculated the KDs. We also gave an example that illustrates our approach to determine one of the KDs, in the absence of nucleotide on a G/T mismatch DNA substrate (KD =22 nM; 20 nM MutS and 10 nM of G/T DNA). For DNA with blocked ends and EcoRI bound, the KD is 21 nM. The increased concentration of MutS relative to that of DNA (50 nM MutS and 10 nM DNA) in this experiment caused the occupancy to increase but KD remained unchanged.

1.15 "On another note, the errors listed seem unrealistically small. I don't see how one can have an error of $\pm 0.1\%$ when less than 1000 DNAs were analyzed. How were these errors calculated? All that is state is that they are SEM."

Reply: We followed the earlier suggestion of the referee and changed all the SEMs to standard deviations.

1.16 "Do the occupancies listed in Table 3 include MutS bound to the ends of the DNA? If so what % is bound to the ends. This may be important given that Figure S4 shows a significant % of MutS is bound at DNA ends even when both ends are supposed to be blocked by streptavidin."

Reply: The occupancies listed in Table 3 included MutS bound to the ends of the DNA. The occupancy of MutS bound on the ends of DNA is only approximately 5%. (note that this number reflects also the presence of DNA molecules without any MutS bound to them). In Figure S3 the percentage of MutS bound at the terminal positions is, at 15%, higher than before, because this figure includes only DNA with MutS bound to them (free DNA must be ignored). In fact, for DNA which was not blocked by streptavidin, we observed even less end binding under our experimental conditions.

1.17 Also what % of the DNA fragments have both ends bound by streptavidin.

Reply: For double-biotin labeled DNA, 78% of DNA had both ends blocked by streptavidin. For single-biotin labeled DNA, 89% of DNA had one end blocked by streptavidin. . However these percentages are not critical, because when we determined the contour length of loop or a dissociation constant, we only counted those DNA molecules with both ends blocked by streptavidin (for experiments with double-biotin labeled DNA.)

1.18 "On page 11, they state "this resulted in a large decrease in the observed binding constant in the presence of ATP from 64 +/- 1 nM to 25 +/- 1 nM." While this decrease is significant, it is certainly not large (only 2.5 fold)."

Reply: We followed the referee's suggestion, and changed "large" to "significant".

Referee #2

This referee divided his/her comments in two groups (i) a number of writing and presentation problems; (ii) several "bigger-picture" problem areas where in his/her opinion the authors have not explained themselves clearly or have not come to grips with their data, or have not constructed and explained (a) plausible model(s).

In regard to (i) we significantly improved the writing style and presentation of our data and ideas.. We think that these editorial changes significantly improved the manuscript's clarity and the presentation style. These changes are directly introduced into the manuscript and because they have no direct bearing on our results or their interpretation they are not listed here as they are quite numerous.

In regard to (ii) we believe that we addressed almost all important points and issues raised by Referee #2. The referee's comments and questions and our responses are listed below.

2.1 "There is no discussion of the significance of the substantial amount of free MutS monomer in the absence of nucleotides, and its sharp decrease in the presence of all nucleotides. If this has all been explained before, say so and provide the references."

Reply: We followed referee's suggestion, and provided a reference and added a sentence in the Discussion section about the significance of the substantial amount of free MutS monomer in the absence of nucleotides.

"In the absence of any nucleotide, ~30% of MutS monomer was identified, and the percentage dropped significantly to ~5% in the presence of all the nucleotides used. This result supports earlier observations about nucleotides' role in MutS dimerization. (Lamers et al, 2004)"

"Lamers MH, Winterwerp HHK, Sixma TK (2003) The alternating ATPase domains of MutS control DNA mismatch repair. Embo Journal 22(3): 746-756."

2.2 "Also, what is the significance of the MutS-800 data? Is MutS-800 active in vivo, despite its

lack of tetramerization?"

Reply: We removed the MutS- 800 part from both the main text and supplementary material since it is not closely related to our main topic.

2.3 "1. It's not clear why there are more tetramers in the presence of DNA. Are most of the tetramers now bound to DNA? If so, why tetramerization the same for G/T and homoduplex DNA? If most tetramers are in solution, how did DNA cause this?"

Reply: This is a very interesting and important issue. In fact, 41-bp DNA does not change the dimer-tetramer ratio significantly (Figure 1 and 2). Only when 201-bp was included, the fraction of tetramer increased from ~40% to ~70% (Figure 2). Based on the concentrations of both MutS and DNA used in our experiments (100 nM of MutS as monomer and 200nM of DNA) and the known KD of MutS/DNA complexes (21 nM) (Blackwell et al, 2001a), we can estimate that at this condition, more than 90% of MutS (93.6%) was occupied by DNA. At this point we do not know why tetramerization is similar for G/T and homoduplex DNA, but we do know that DNA plays an important role in MutS tetramerization (possibly by stabilizing the tetramer structure). We added a paragraph in the Discussion section to discuss why there are more tetramers in the presence of DNA.

"Based on the AFM measurement alone, we do not know how much of the MutS was bound to DNA because the 41-bp and 201-bp DNA are too small to be imaged by AFM if they are bound to MutS. But based on the concentrations of both MutS and DNA used in our experiments (100 nM of MutS as monomer and 200nM of DNA) and the known KD of MutS/DNA complexes (21 nM) (Blackwell et al, 2001a), we can estimate that at this condition, more than 90% of MutS (93.8%) was occupied by DNA. These results suggest that the interaction of MutS with 201-bp DNA promotes the formation of tetramers, and that tetramer formation of MutS in complex with DNA is influenced by the interaction of MutS with adenosine nucleotides. The reason that the association constant for the dimer-to-tetramer transition did not change after 41-bp DNA was included might be that 41-bp DNA is too short to support binding of two MutS dimers side by side. However, the 201-bp DNA (both homoduplex and G/T mismatched) is long enough to accommodate and stabilize two dimers forming a tetramer structure. The similar effect of homoduplex and G/T DNA on the dimer-to-tetramer transition suggests that MutS may scan DNA for mismatches as a tetramer. However, mechanistic details of the tetramerization step in the presence of DNA are presently missing and warrant further studies."

2.4 "2. Despite claims about MutS being "preferentially" near the mismatch, MutS seems mostly tens to hundreds of bp away from the mismatch site. Why is this?"

Reply: The primary reason behind a relatively broad distribution of location of MutS on heteroduplex DNA is that in the original experiments, the DNA contained biotin/streptavidin at both ends and therefore it was impossible to precisely determine the location of the mismatch site. The mismatch position could have been either 47% or 53% length from the DNA end. In order to solve this problem, we performed new experiments using single-end-biotinylated DNA that was incubated with monovalent streptavidin. The location of the mismatch on this DNA was easy to determine precisely. We added new figures in the main manuscript (Figure 6 and 7) that report our observations using this new system. This DNA allows us to identify the position of the mismatch that was engineered to be 53% length from the labeled end, new Fig. 6 and 7. The distributions of lengths of the long leg determined for this substrate became narrow and their averages coincide very well with the position of the mismatch: $46 \pm 5\%$ when the loop was formed between the mismatch and the DNA end blocked by streptavidin, and $53 \pm 3\%$, when the loop was formed between the mismatch and the unlabeled DNA end. Thus, based on these new results we are quite confident that the length of the long leg is consistent with the position of the mismatch. We added a new paragraph at the end of the RESULTLS section that reports these new data.

Please also refer for more details to our response to a similar comment (1.1) made by referee 1.

2.5 "3. The specificity of MutS for G/T vs. homoduplex seems surprisingly low. Why should there be substantial loop formation on homoduplex DNA? Why are the KD values for G/T and homoduplex DNA so similar? It's not enough to make a brief statement about "per site" vs. "per entire DNA molecule". What is "true specificity" for a mismatch site? What is the significance of the binding to homoduplex DNA vs. the homoduplex regions (99.9%!) of (G/T) DNA? How might this relate to

searching DNA for mismatches? Much more thought and cogent discussion are needed."

Reply: In fact, the specificity of MutS for G/T vs. homoduplex we measured is 8.1 (179/22) for no nucleotide condition and 3.4 (219/64) when ATP was added. These results are similar, to those obtained by others using traditional methods (Blackwell et al, 2001a; Schofield et al, 2001a). However, KD for homoduplex DNA we reported is much different from that report by Yang et al. To address these issues we added the following paragraph:

"At the same time, the specificity of MutS for G/T vs. homoduplex DNA that we measured is 8.1 for no nucleotide condition and 3.4 when ATP was added. These results are similar to those obtained by others (Blackwell et al, 2001a; Schofield et al, 2001a). However, the KD for nonspecific binding that we report is much different from that reported by Yang et al (Yang et al, 2005). Our results differ in that we included MutS bound to the ends of DNA, which will increase the KD for nonspecific binding, because MutS has quite high affinity to bind to DNA ends (about 5% of DNA ends were bound by MutS and 15% of MutS that bound on DNA was found at its ends). In addition, Yang et al. defined the nonspecific dissociation constant as KD per base pair while we measured KD per DNA molecule. It is obvious that longer DNA has higher probability to be bound by MutS nonspecifically. Based on our findings in which 1120 bp was used, the fraction of nonspecific binding must be high."

2.6 "4. Much more coherent considerations of loops, their growth, and their relationship to MutS positions are needed."

Reply: We followed referee's suggestion and significantly expanded this part by presenting new Figures (Figure 6 and 7) in the main text. These figures report our new results, which were obtained on heteroduplex DNA containing a biotin at a single DNA end. In addition, we added 3 new figures to Supplemental Information which give the detailed analysis about the distribution of loops and their growth at different conditions.

2.7 "5. The entire set of EcoRI experiments is a mystery."

Reply: The use of EcoRIE111Q as internal "road blocks" to investigate MutS movement on DNA is not new. We followed earlier publications that exploited this approach. The purpose was to test whether MutS translocate on DNA by sliding or by a hopping mechanism. We added a new paragraph in the Discussion section to clarify the EcoRI results. In addition, we added two Supplementary Figures (new Figure S5 and S6) to show the redistribution of nonspecifically-bound MutS after EcoRI was added.

2.8 "6. The authors seem to be proposing a MutS translocation and diffusion model based to some extent on EcoP151, but the MutS model is very poorly described. Fig. 6 is worse than nothing as it stands."

Reply: We followed referee's suggestion, and removed our model (former Figure 6) from the main text.

2.9 "How could diffusion (with no source of energy) ever start a loop against DNA bending stress? Could translocation start the loop and diffusion of ATP-bound MutS extend it, riding down a bending-stress gradient?"

Reply: This is an excellent point however, because of AFM imaging resolution limits and also present lengthy image acquisition time we did not capture the formation of DNA loops. Therefore it is difficult to address this point. Referee 1 proposed a helpful suggestion that is relevant to the loop formation, by observing that "Öthe DNA (that we used) has a propensity to bend in this region., which would promote loop formation. Such a propensity to bend could explain the high observation of loops in the absence of ATP and why the loops show such a narrow distribution of size and position even on homoduplex DNA." We followed both referees' suggestions and added a sentence in the Discussion section to speculate about the mechanism of loop formation. Please also refer to our reply to comment 1.4 made by referee 1.

2.10 "What would the occasional ATP hydrolysis do? Nudge MutS in the "right" direction? The authors should spend more effort on a coherent model and less effort on comparisons with the old

Allen/Modrich work."

Reply: Based on our experimental results, we clearly observed the increase of loop size in the presence of ATP. However, based on the AFM imaging method alone, we cannot conclude whether this process is driven by ATP hydrolysis or not. It may be clarified by future experiments. In the new experimental results (Figure 6 and 7), which were obtained using single-site biotinylated DNA, we observed MutS loops formation and growth on both 47% and 53%-length sides of the mismatch. Thus, in our experiments DNA loops grew in both directions.

Specific Comments

2.11 "Page 4, line 1-3: Doesn't distinguish between ATP-dependent diffusion and hydrolysis-dependent unidirectional translocation."

Reply: We followed referee's suggestion, and rewrote the sentence, and separate the references into ATP-dependent diffusion and hydrolysis-dependent unidirectional translocation.

2.12 "Page 5, line 4: How much 'more frequently'?"

Reply: We followed referee's suggestion, and changed "more frequently" to "2-6 times more frequently" .

2.13 "Line 14: Here and throughout the manuscript the insidious practice of heading sub-sections with didactic-data-interpretation statements masquerading as pseudo-titles works against balanced reading and understanding of the manuscript. The place for these is at the beginning of the discussion. The title is in fact incomplete as a summary of the 'take-home lessons'."

Reply: We followed referee's suggestion and removed all headings of sub-sections.

2.14 "Line 18: Some readers will not know what 'flooded image' means."

Reply: We expanded our description of how "flooding" was performed in the MATERIALS AND METHODS section. We also added a reference (Horcas et al, 2007), which contains a very detailed description of how the volume is calculated.

2.15 "Page 6, line 1-5: Nice work! Why might the tetramer MW estimate be so far off?"

Reply: The estimated MW of the MutS tetramer (300 ± 34 KDa) and the expected value (380KDa) are indeed different. To build a relationship between MW and a measured volume, we used a set of reference proteins and assumed that the densities of all the proteins used are the same. In fact, the densities of different proteins are likely different, which will cause the calibration error. As seen in Figure S1, the biggest protein thyroglobulin has a higher volume than expected, affecting the curve fitting. As a result, the data point of MutS tetramer is below the fit line. Therefore, the calculated MW is smaller than the expected value.

2.16 "Line 11-12: Are the AFM volumes different?"

Reply: The AFM volumes for T1 and T2 are similar. We added a sentence to discuss the possible reason for causing this appearance difference.

"These two apparent forms (T1 and T2) have similar volumes as measured by AFM (data not shown) and their different appearance could be a consequence of different orientations of tetramers on the mica surface."

2.17 Nowhere is the surprising amount of monomer discussed. Do the authors think that that much protein remains undimerized in vivo, or is the no-nucleotide datum irrelevant?"

Reply: We followed referee's suggestion, and provided a reference and added a sentence in the Discussion section to talk about the significance of the substantial amount of free MutS monomer in the absence of nucleotides. The fraction of MutS monomer is high only if nucleotide is absent, which is clearly not the case in vivo.

"In the absence of any nucleotide, ~30% of the MutS monomers was identified, and the percentage dropped significantly to ~5% in the presence of all the nucleotides used. This result supports earlier observations about nucleotides' role in MutS dimerization. (Lamers et al, 2004)" Please also refer to our response to the General Comment (2.1).

2.18 *"Page 7, Line 10-12: It might be helpful to see the M-D-T distributions pulled out together for at least no-NT and ATP, for A/T vs. G/T DNA."*

Reply: We followed referee's suggestion, and added a new inset in Figure 1C.

2.19 *"Line 13-14: IMPORTANT! How much of the MutS was bound to DNA? Are we looking at tetramers bound to DNA and (at least some) dimers in solution, or what?"*

Reply: This comment is similar to comment 2.3. More than 90% of MutS was occupied by DNA according to their concentration and KD. Please see our reply to comment 2.3 for details. MutS that was bound to DNA or free in solution can be either in the tetramer or dimer configuration. But if longer DNA was present such as 201 bp DNA, the fraction of tetramer increased. In this paper, we only focused on the oligomeric state of MutS on DNA because most of MutS (90%) was bound on DNA.

2.20 *"Line 12-19: Is DNA length intrinsically important or just the total number of bp? The author must compare 1000 nM 41-bp DNA with 200 nM 201-bp DNA for the DNA size comparisons to mean anything at all."*

Reply: This is a very important point. In fact we already had the experimental data that was obtained at different concentrations. Based on these results, the ratio of dimer and tetramer didn't change too much if we only change the concentration of DNA. We followed the referee's suggestions, and added a sentence to clarify this point.

"Besides, the results of the control experiments (data not shown) confirmed that, when the concentrations of MutS and 41-bp DNA or 201-bp DNA were changed, the change in the ratio of dimer to tetramer was minimal. It supports the observation that the percentages of MutS dimers and tetramers were not affected by the concentration of DNA, but only by the length of DNA."

2.21 *"Why are G/T and A/T results so similar?"*

Reply: Please refer to our response to the Comment 2.3.

2.22 *"Is MutS mostly a tetramer even on homoduplex DNA?"*

Reply: Our results showed that on 201-bp homoduplex DNA most of MutS is in the tetramer configuration. (Figure 2B)

2.23 *"Line 22: Especially here, the reader should be reminded that this is a linear substrate."*

Reply: We followed referee's suggestion, and added "linear" in front of DNA.

2.24 *"Page 8, Line 9-10: Because Fig. S4 is so poorly labeled, it's hard to evaluate this statement. I think the authors mean for the x axis to represent the position of the MutS, shown as the fractional length of the short leg. Do these positions refer to the same end, or does 46% (presumably the mismatch position from the "left" end in Fig. S3) in fact always refer to the left end? The authors never tell us. In Fig. S4 Are 46% or 56% both from the left end? If so, these data are not very useful. I assume that y axis actually represents the fraction of short legs that have a particular length."*

Reply: We followed referee's suggestion, and redrew Figure S3 (the former Figure S4). In new Figure S3, the DNA containing biotin at both DNA ends was used. In this case, we could not identify which end is which. This means that both 47% or 53% are from the left end. We re-analyzed the data shown in Figure S3 A and C, and divided the MutS into two groups, specifically-bound MutS and non-specifically bound MutS. Most importantly, we added new experimental results which were obtained using DNA containing biotin (and monomeric streptavidin) at a single DNA end (New Figure S8). In this case we could unequivocally identify both DNA ends according

to whether they had or did not have streptavidin bound.

2.25 *"In any event, the statement about "located very close to the mismatch" is totally unjustified. The location seems to be up to 150 bp in either direction with no nucleotide and all over the place in ATP."*

Reply: This comment is similar to the comment 1.1 and 2.4. Please refer to our earlier replies. We followed referee's suggestion, and changed "very close to the mismatch" to "close to the mismatch". We also added new Figures (Figure 6 and 7) which use heteroduplex DNA containing a biotin at a single DNA end. Using this DNA, we found that MutS is located at either 53±3% (without ATP, please see Figure 6) and 54±5% (with ATP, please see Figure 7) consistent with the location of the mismatch. These observations reinforce our results that MutS binds to or near the mismatch site. In addition, the distribution of MutS location on this new DNA substrate was significantly narrower, because we could easily differentiate 53% length location from 47% length location, which was impossible on the double-labeled DNA.

2.26 *"Do the substantial fractions of MutS at position zero reflect binding to ends despite the streptavidin?"*

Reply: Yes, we did observe some MutS bound on or very close to the end. In addition, when the end(s) of DNA was previously bound by streptavidin and ATP was added, the fraction of MutS end-binding increased because streptavidin blocks it from coming off of DNA.

2.27 *"Line 15-16: What is the evidence that the strands "overlap".? It seems at least as likely that they form an "X" structure (Fig. 6)."*

Reply: Since we do not know for certain whether DNA strands overlap or not, we followed referee's suggestion, and removed the word "overlap".

2.28 *"Page 10, line 1-5: Confusing and misleading sentence. Are the authors just stating the obvious fact that the DNA in bigger loops has to come from somewhere? If both dimers remain together, and one of them remains "centered" on the mismatch (the spread is actually 1-200 bp or so), it seems more accurate to say that DNA from the short leg or long leg (why is it in fact usually the long leg?) moved past the MutS tetramer(dimer) and into the loop."*

Reply: We followed referee's suggestion, and rewrote the paragraph.

"Because the end of the long leg remained at almost the same position (47%~53%) in the presence or absence of ATP, the MutS dimer associated with the long leg remained at the mismatch site. However, when ATP was added, the DNA of the short leg moved through the nonspecifically-bound MutS dimer and entered into the loop. As a result, the short leg became shorter and the DNA loop size increased. We pooled together positions of all MutS particles in the loop configuration and produced their distribution (Supplementary Figure S4A and S4B). The distributions of specific-bound and nonspecific-bound MutS that formed loops fall into two groups, which were similar to that of the single-site-bound MutS (Supplementary Figure S3B and S3D). In summary, the distribution of all MutS (including specifically-bound MutS and nonspecifically-bound MutS) on DNA are not affected by the binding types of MutS/DNA complexes (single-site binding or loop binding) or by oligomeric states of bound MutS (tetramer or dimer)."

2.29 *"Line 15-16: badly written sentence. Not obvious from Table 2."*

Reply: We followed referee's suggestion, and rewrote that sentence.

"Loop size after adding AMPPnP for 10 minutes was bigger (average increase 107 bp) than the loop generated at 'no nucleotide' condition and was similar to the loop size after incubation with ATP."

2.30 *"Why does homoduplex DNA generate such big loops? How frequent are they (vs. G/T DNA)?"*

Reply: We added a paragraph in the Discussion section to explain why homoduplex DNA generate such big loops.

"In our study, loops were also seen on homoduplex DNA ($2.4 \pm 0.3\%$) with a frequency about 3.2 times less than that on heteroduplex DNA under the same conditions (Table 3). In addition, the size of the loops on homoduplex DNA ($40 \pm 15\%$) was much bigger than that on heteroduplex DNA ($22 \pm 5\%$), and at the same time their size distribution was also 3 times broader as compared to the loops formed on heteroduplex DNA (Table 2). Our interpretation of this result is that in the absence of a mismatch, neither of MutS dimers that participated in the loop formation had to remain near the center of homoduplex DNA."

2.31 *"23-24: Confusing. Is the comparison to experiments without EcoRI?"*

Reply: We rewrote the sentence.

"However, the average size of loops formed in the presence of EcoRI and ATP was considerably larger than that measured in samples incubated with EcoRI and with ADP or with EcoRI but with no added nucleotide (Table 2)."

2.32 *"What fraction of EcoRI molecules are or are not where they belong? Is the nearby ERI site likely to be in the loop?"*

Reply: Based on our control experiments, EcoRI molecules bind very specifically to EcoRI sites. In addition, when we measured the length of loop contour, short leg and long leg or calculated the dissociation constant, we only counted those DNA molecules with both EcoRI sites occupied by EcoRI proteins. Comparing Supplemental the green bars in Figure S6E (adding EcoRIE111Q) with S4B (no EcoRIE111Q), we find that, in the presence of ATP and after EcoRIE111Q was added, the nonspecifically-bound MutS accumulated at 18% of DNA length from the end. This is exactly the location of EcoRI site. It means the binding of EcoRI protein on DNA does prevent MutS from sliding to ends.

2.33 *"Page 11, Line 5-14: Can the authors briefly explain the low specificity for G/T DNA, without forcing us to wade through both Materials and Methods descriptions? Does it reflect the large amount of homoduplex DNA in the G/T substrate as well? Would shorter DNA substrates show higher specificity?"*

Reply: This comment is similar to comment 2.5. Please refer to our reply to that comment. We determined that our G/T DNA is quite pure. I think the reviewer proposed a good point that shorter DNA show higher specificity. It is obvious that the longer DNA has a higher probability to be bound by MutS nonspecifically. For example MutS bound on 41bp G/T DNA, all the bindings are considered as specific binding because there is no space for MutS to bind nonspecifically. While for 1120 bp DNA, our result showed that nonspecific binding is high. So our result may imply that the binding specificity of MutS on 1120 bp G/T DNA is less than on 41 bp G/T DNA. .

2.34 *"Line 8-9: Which of the two results 'is similar to KDs'?"*

Reply: We followed referee's suggestion, and rewrote these sentences.

"In the absence of streptavidin end-blocks and nucleotides, MutS bound to DNA containing a G/T mismatch with KD of ~ 22 nM. This is similar to KDs measured using other methods, including bulk methods (Schofield et al, 2001a; Yang et al, 2005). In the presence of 0.5 mM ATP, KD was ~ 64 nM."

2.35 *"Line 14-15: Tell the reader the significance of this. Why is the (G/T) preference still only about 4-fold in these cases?"*

Reply: Please refer to our response to the Comment (2.5) and (2.33).

2.36 *"Page 11, Line 19-22: What's the significance of this?"*

Reply: We added a sentence to discuss the significance.

"which means that adding ATP cannot drive MutS off DNA because of either streptavidin blocked DNA ends or EcoRIs blocked DNA around the mismatch site."

2.37 *"Line 22-24: Why might ERI improve binding and looping?"*

Reply: First, in the absence of ATP, both end binding by streptavidin and road blocking by EcoRI had no effect on binding and looping. But in the presence of ATP, binding by streptavidin and EcoRI increased the binding affinity and the fraction of loop binding. In addition, road blocking by EcoRI prevented the increase of loop size. By comparing the results obtained using DNA streptavidin and EcoRI to those obtained on DNA with streptavidin (Table 3), we find that EcoRI did not improve MutS binding and looping significantly. This is because most of MutS were already blocked by streptavidin regardless of whether EcoRI was added or not. The main purpose that EcoRI was added is to find whether EcoRI can prevent the increase of loop size. Please also refer to our reply to comment 2.7.

2.38 *"Page 12-13, Line 15: But the monomer concentration was strongly affected. Why?"*

Reply: Please refer to our response to the General Comment 2.1 and 2.17.

2.39 *"Line 20-25: The similarity of the effects of homoduplex and G/T on dimer; tetramer suggest that MutS scans/tests DNA for mismatches as a tetramer. The authors never address this important point. Were all or most of the tetramers on DNA, or might they form on DNA, but then dissociate into solution as tetramers (an important mechanistic question)?"*

Reply: Please refer to our response to the General Comment 2.3, 2.19 and 2.21. And based on the AFM imaging method used, we do not know whether these tetramers were formed on DNA or in solution. This will be clarified by future experiments.

2.40 *"Page 13, line 7-15: As written, this paragraph will really confuse readers who are not MMR experts. The important question of specificity (as measured by AFM) requires careful attention. See my comments on line 12-19 in page 7. If increasing DNA length increases apparent binding to homoduplex DNA, perhaps because the chances of being hit by MutS are greater, and binding to homoduplex DNA is strong (as it certainly is), then the added DNA should cause binding (1/KD) to (G/T) DNA to increase with length, but it apparently doesn't. Something of highly fundamental significance may be going on here, and the authors should vigorously pursue it."*

Reply: Please refer to our response to the Comment 2.5, 2.33 and 2.35.

2.41 *"Line 26 and p. 14, line 1-2: This is very confusing. Besides cleaner writing, a model diagram (perhaps added to Fig. 6) is needed."*

Reply: We followed referee's comments in point 2.8 removed the relevant sentence and old Figure 6.

2.42 *"Do the authors think that MutS in a search mode "bounces off" ERI and moves back towards the mismatch? Is the MutS trapping done by ERI in this loop or outside the loop?"*

Reply: We added new Figures S5 and S6, and also added new paragraph to explain how MutS was blocked by EcoRI. Please also refer to our reply to comments 1.5 and 2.7.

2.43 *"Why are (G/T) occupancies higher with ERI binding but KDs about the same? What does it mean that ERI affects (G/T) occupancy much more than homoduplex? Unless the authors can make some biochemical sense of the ERI results, they should be dropped."*

Reply: This is mainly because the concentration of MutS monomer and DNA used here is 50 nM and 10 nM respectively. We increased the concentration of MutS which will cause the occupancy to increase without affecting KD. Please also refer to our reply to comment 1.14.

2.44 *"Page 14, Line 21-22: somewhere the authors need to seriously explain the significance of all of those loops on homoduplex DNA. In general however, far too much space is devoted to comparisons with the quite different Modrich experiments."*

Reply: For the loop formed on homoduplex DNA, we followed referee's suggestion, and added a paragraph to Discussion. Please refer to our reply to comment 1.1, 1.2, 1.4, 2.5, 2.30. Furthermore,

we condensed the comparison with Modrich's experiments.

2.45 *"Page 15, line 1-3: Do you mean 10 minutes before deposition?"*

Reply: Yes. In fact, we did both 10 seconds and 10 minutes incubation before deposition

2.46 *"Line 3-7: Considerable clarification is needed. Where are the relevant data? Does the rate calculation (600 bp/min) come from the 10-min pre-deposition incubation in ATP? What happens in 10 min without ATP? The authors should either present a reasonably complete rate study, or drop all discussion of these very scanty data."*

Reply: We followed referee's suggestion, and removed the discussion about the rate of loop increasing.

2.47 *"Line 13-16: A long awkward sentence. Again, too much space is devoted to extended comparisons with the Modrich work."*

Reply: We followed referee's suggestion, removed the sentence and condensed the comparison..

2.48 *"Line 19-20: confusing sentence. Line 20-22: this strong conflict with the work presented here must be discussed seriously."*

Reply: We followed the referee's suggestion, and revised the paragraph.

"The loops they saw were independent of whether or not a mismatch was present in the DNA substrate. D835R (tetramer-deficient) and MutS R194A/R198A/R275A mutants of MutS formed -loops at a similar frequency as wt MutS, suggesting that tetramer formation is not necessary for the formation of the loops that they saw. However, they did not provide the direct proof (volume measurements) to show that the MutS that formed -loops was a dimer."

2.49 *"Page 16, page 15, line 24-25 and page 16 line 1-10: The Eco P151 discussion is confusing without a figure. What is the significance of a low ATP hydrolysis vs. translocation ratio (abruptly introduced for the first time here)?"*

Reply: We removed the "low ATP hydrolysis" part, because these observations require further clarification.

2.50 *"Line 1-4: bad sentence, but important points. 4-7: Not all clear, but again important."*

Reply: We followed the referee's suggestion, and rewrote this paragraph to improve its clarity.

"Using fast-scanning AFM, Crampton et al (Crampton et al, 2007) visualized translocation and extruded looping by EcoP15I restriction enzyme, which similar to MutS displays anomalously low rates of ATP hydrolysis. In order to efficiently cleave DNA, EcoP15I requires two recognition sites separated by up to 3.5 kbp. They observed that DNA loops were formed by contact between specific-site-bound EcoP15I and a nonspecific region of DNA. Further they showed that EcoP15I translocation must also occur because when it is blocked by a Lac repressor protein, DNA cleavage is inhibited. This observation provides a plausible mechanism by which an enzyme could communicate rapidly between two noncontiguous DNA sites through a combination of translocation-coupled and DNA looping mechanism."

2.51 *"Line 11: "Data support". 11-20 Confusingly described and inconsistent. Fig 6 (which should be accompanied by an informative legend) is no help at all. Is the "translocation" really movement of the DNA "through a second dimer still attached to the mismatch-bound dimer? How can loops grow rapidly if the "translocation" is actually "diffusive" bidirectional sliding? Unidirectional translocation requires energy. 16-18: Contradictory and confusing. What is "ATP-hydrolysis-driven loop formation (by Mut tetramer)? How does it "combine" with "a diffusive loop mechanism". What is "looping" without ATP? How can the loop grow without ATP hydrolysis? Where is the hydrolysis-dependent step? The authors need a much more detailed figure, more explicitly incorporating the EcoP161 model and with insightful descriptions."*

Reply: We followed the referee's suggestion, and rewrote the paragraph.

"Our data shows that in the absence of ATP, both end binding by streptavidin and road blocking by EcoRI had no effect on binding and looping. But in the presence of ATP, binding by streptavidin and EcoRI increased the binding affinity and the fraction of loop binding. In addition, road blocking by EcoRI prevented the increase of loop size. This observation supports the idea that MutS moves along the DNA contour by means of a translocation (sliding) mechanism. In addition, at least part of the enzyme, possibly one dimer of a tetramer, must remain at or near its target site (mismatch site). Moreover, DNA is being reeled from one side of the mismatch site, thereby producing expanding DNA loops. Adding ATP has the effect of making the size of the loop increase, and present observations support MutS translocation which depends on the binding of ATP. Looking ahead into the future, whether or not MutS translocation is ATP hydrolysis-dependent needs to be clarified. Moreover, when AMPPnP was added (in Table 2), the loop size can expand too. It might imply that ATP hydrolysis is not completely necessary for this unidirectional "translocation"."

2.52 *"Could the decrease in DNA-bending stress drive diffusive loop expansion after a hydrolysis-dependent unidirectional "translocation" initiated the loop? How often would MutS-bound ATP hydrolyze during loop expansion?"*

Reply: Based solely on our current AFM observations, we cannot determine whether loop expansion is caused by diffusion or by ATP hydrolysis. This needs to be clarified by future experiments. However, we note that when AMPPnP was added (Table 2), the loop size expanded too. Therefore it is possible that ATP hydrolysis is not absolute necessary for this unidirectional "translocation".

2.53 *"Page 17, line 9-12: Since much MMR work has been performed using circular substrates, the linearity of the 1120-bp (not hyphen) substrate should be noted. 9: Maybe "Linear 41-bp"."*

Reply: We followed referee's suggestion, and added "linear" to DNA substrate.

2.54 *"Is there an annealing step somewhere?"*

Reply: We followed referee's suggestion, and added a sentence at the end of this paragraph. "For both heteroduplex and homoduplex DNA, they were annealed using the same method."

Page 20, line 8-11: How many, actually?

Reply: We counted this configuration (less than 5%) and added this result to the manuscript.

Figure 1. An equivalent of Fig. 1C (M-D-T distribution) is needed for at least one case with nucleotide. Put "MutS" into C and D ordinate labels. Remove gray background and horizontal lines from D; use "ticks" along ordinate instead.

Reply: We followed the referee's suggestion, and revised the figures.

Fig. 2. Remove gray background and horizontal lines

Reply: We followed referee's suggestion, and remove gray background and horizontal lines.

Fig. 4. I like panels A and E. Label x-axis as "position" or "coordinate". Why fractions on left, but %s on right?

Reply: We followed the referee's suggestion, and revised the figures.

Fig 5. Switch positions of bars in C to line up with A and B. Legend inside figure is redundant.

Reply: We followed the referee's suggestion, and revised the figures.

Fig. 6. Completely inadequate (see above).

Reply: We removed that figure.

Figure S3E. Label the ordinate. Show bp instead of / in addition to nm.

Reply: We followed the referee's suggestion, and redrew the figure.

Fig. S4. "Percentage (%)" labels on Y axis are both redundant and confusing; it's endpoint position on the DNA or length. In B and D, its fraction of short legs (with indicated lengths/endpoints). The mismatch positions should be clearly indicated.

Reply: We followed the referee's suggestion, redrew the figure, and added a dash line to indicate the mismatch position.

Fig. S3 legend. Line 4: Some is unscientific and vague. Could be read as a mixture of three substrates. Otherwise good figure and good legend.

Reply: We followed the referee's suggestion, and revised Figure S2 (old Figure S3) legend.

Fig. S4. Legend. Line 11: These distributions don't look at all preferential to me.

Reply: Figure S4 now is Figure S3. We redid the Gaussian Fit, rewrote the figure legend.

Table 1: What are the differences between photos A and B, and among C-E? Was ATP present in the spreads shown? The alignment of the data columns under the pictures is very confusing. Do the volume and MW measurements correspond to all molecules in the spreads shown and others like them, or to selected apparent dimmers?

Reply: We removed all of the photos in Table 1. We measured the volume of MutS at the conditions of both w/o ATP and w/ ATP. The first column of the table gives the information about whether ATP was added or not. The volume and MW measurements corresponded to all the molecules in the images. Because the volumes of dimer and tetramer are present as two individual peaks in the distribution, they can be separated automatically by volume analysis.

Table 2. What fractions of the total molecules, especially the homoduplexes, showed the loops that were measured? Position, not "% of length", which is not clear.

Reply: We followed referee's suggestion, revised the caption, and added a sentence.

"For each experimental condition, about 50-60 DNAs with loops formed on them were analyzed to determine length distributions."

Table 3. "(+) and (-)" nomenclature is not correct, despite its' wide (mis)use in the literature and in fact is harder to read. For most columns, n(none) instead of (-) would be better, since its nucleotide or streptavidin or EcoRI that's added ("+"). Nothing is subtracted ("-"). The biotin ends and streptavidin columns are redundant, and in the case of (+) vs (++) streptavidin, confusing. You can't add ("+") a mismatch to DNA, and certainly can't subtract ("-") mismatch. Use a different nomenclature. These are not "binding", but dissociation constants. A column to indicate -fold binding specificity would

Reply: We followed the referee's suggestion, revised the table and table caption.

Referee #3 (Remarks to the Author):

3.1 "Primarily, I wish that the authors had done a second construct in which the mismatch site is significantly displaced from the center of the DNA molecule. There may, for example, be physical reasons why the base of the loop prefers the center of the molecule. It would be very reassuring to see the base of the loop move to one side in the alternate construct. Also, I have some concern about the significant broadening of the "long leg" distribution in the presence of ATP (Fig 4H). Has the ability of MutS to "home in" on the mismatch been reduced?"

Reply: Unfortunately, we did not have such a DNA substrate at our disposal. This may be addressed in a future study. However, in the revised manuscript we present new results obtained using a

different heteroduplex DNA substrate that was labeled with biotin only at one end (new Figures 6 and 7). Thus, at least partially we addressed the reviewer's comments. Two groups of loops can be observed in panel A in both Figure 6 and 7. Group one: the loops are between mismatch site and biotinylated end, where the relative lengths of long leg are $53\pm 3\%$ and $54\pm 5\%$ for without ATP (Figure 6) and with ATP (Figure 7), respectively. And group two: the loops are between mismatch site and free ends, where the lengths of long leg are $46\pm 5\%$ and $48\pm 8\%$ respectively. For both groups, MutS proved to bind to or very near the mismatch site. In addition using this new substrate, the distribution of the MutS on DNA was narrower as compared to double-labeled DNA..

Some minor matters:

3.2 *"In the introduction to the paper, one gets the sense that observation of DNA loops with MutS is original to this paper. I know that the authors do not think this, but that is the impression given. It would be good to get a reference to Allen et al (and anyone else who has made such reports) up front in connection with loops."*

Reply: We followed the referee's suggestions.

3.3 *"The authors sometimes use "binding constant" and KD (dissociation constant) interchangeably. That is very confusing; one is the reciprocal of the other. Take this excerpt from page 11: ". . .this resulted in a large decrease in the observed binding constant . . . from 64 . . . nM to 25 . . . nM". They obviously mean KD here. Say "dissociation constant" instead of binding constant."*

Reply: We followed the referee's suggestions and changed all the "binding constant" to the "dissociation constant".

3.4 *"The authors should state whether the two apparent forms of tetramers could be a consequence of different orientations on the surface."*

Reply: To address this point we added a following sentence.

"These two apparent forms (T1 and T2) have similar volumes as measured by AFM (data not shown) and their different appearance could be a consequence of different orientations of tetramers on the mica surface."

3.5 *"The authors have shown that ATP hydrolysis is not required for loop formation by MutS, yet they mention that this mechanism could work in combination with hydrolysis-driven loop formation by MutS (page 16). These two mechanisms seem contradictory. Some explanation is needed."*

Reply: We agree. We changed "loop formation" to "loop increase".

2nd Editorial Decision

08 April 2011

Thank you for submitting a new version of your manuscript (previously EMBOJ-2010-73591) for our consideration. Two of the original referees have now evaluated it once more, and both find the manuscript experimentally improved to now be in principle suitable for The EMBO Journal. However, they both still retain a number of concerns regarding the writing, discussion, presentation and interpretation that would need to be addressed before publication. Following adequate textual revisions of these points, we should be able to ultimately consider the study for publication, and I am therefore returning it to you for a final round of modifications. I should point out, however, that the referees' criticisms will require quite an extensive rewrite of the paper, including reevaluation of the interpretation of some of the data. In the interest of the broad general readership of The EMBO Journal, I would also strongly advise you to follow referee 1's suggestion to shorten the main part of the paper and to focus on the key new findings while delegating other results that are more peripheral to the main message into the supplementary sections. Finally, please carefully proofread

the study both for grammatical correctness and for overall readability and flow, ideally soliciting comments from colleagues with native English language proficiencies.

I hope you will find these final sets of comments valuable for ultimately make the study suitable for publication in The EMBO Journal, and that you will be able to get a carefully re-revised version back to us as soon as possible. Should you have any further questions that require clarification in this regard, please do not hesitate to contact me.

Yours sincerely,
Editor
The EMBO Journal

REFEREE REPORTS:

Referee #1 (Remarks to the Author):

This revised version is significantly improved from the previous version. The evidence for MutS-induced looping involving MutS bound to the mismatch is now much more convincing. I do believe that this work merits publication in EMBO, however, there are several significant errors in the paper. Although these errors do not affect the overall conclusion they need to be corrected. In addition, the results are very cumbersome to go through and follow. It seems that the paper would benefit greatly by being significantly shortened, moving many of the results to the Supplemental material. Many of these results are not central to the conclusion of the paper and are not really discussed in the Discussion section. For example, the discussion of affinities is very difficult to follow and is not particularly pertinent to the conclusion. In my opinion, the exciting and truly important results are the observation that one of the MutS molecules resides at or near the mismatch in the presence of ATP. Shortening the results will make this work interesting to a broader audience and it will have more impact overall. Either way, the points need to be addressed before publishing.

Specific comments:

On page nine: "Single-site-bound MutS can be divided into two groups: specific-bound MutS at the relative position of 50 {plus minus} 6% and nonspecific-bound MutS at 33 {plus minus} 5% in the absence of any nucleotides (Figure S3A and S3B). However, in the presence of ATP, MutS are found to located at 50 {plus minus} 5% and 23 {plus minus} 11% of the relative DNA length (Figure S3C and S3D)." This analysis is incorrect. Fitting of the nonspecific complexes to a Gaussian does not make any sense, because the non-specific complexes should be represented by a random distribution not a Gaussian (by definition non-specific means no preference for a given site). It is somewhat surprising that the non-specific complexes in the presence of ADP are not randomly distributed.

On page 10: "However, when ATP was added, the DNA of the short leg moved through the nonspecifically-bound MutS dimer and entered into the loop." This statement does not belong in the results because AFM does not provide direct evidence of movement. Here and several other places in the results the authors state that the protein has moved along the DNA, although they have no direct evidence. This point is particularly relevant because loops are seen in the absence of ATP so there is no way to know if the protein moved along the DNA to arrive at that point or it already bound there. Only when all of the data is put together does the data support movement.

On page 11: "We pooled together positions of all MutS particles in the loop configuration and produced their distribution (Supplementary Figure S4A and S4B)." Figure S4 is very confusing and the discussion of this figure at multiple points in the paper is hard to follow. It is not clear what the point of this analysis is. It should either be removed completely or put in to Supplemental material.

On page 12: "Figure S4 shows, MutS has no significant binding affinity for the region between 45% and 50% of homoduplex DNA length." I believe you mean binding specificity not binding affinity. Clearly MutS binds to this region.

"Third, loops were seen on homoduplex DNA with a frequency about 3.2 (7.6% / 2.4%) times less than that on heteroduplex DNA at the same conditions (Table 3)." This is not a correct comparison because there are fewer complexes on homoduplex DNA versus heteroduplex DNA because MutS has a lower affinity to homoduplex DNA. Calculating the percent of total complexes that form loops

is what should be compared. 34% of the complexes form loops on heteroduplex DNA and 22% complexes form loops on homoduplex DNA, making the frequency 1.5 times greater not 3.2. A similar error is made later in the text. Such as on page 13 "Though loops formed on both homoduplex and G/T mismatch-containing DNA and on both non-end-blocked and end-blocked substrates, selectivity was about 8.1-fold (179/22) for DNA with a G/T mismatch over homoduplex DNA in the absence of nucleotide and it was about 3.4-fold (219/64) in the presence of ATP." These numbers don't take into account the different number of complexes.

"The plausible explanation for this observation is that on homoduplex DNA both MutS dimers that formed the loop were bound on DNA nonspecifically." What's the point of this statement? The both have to be non-specific because there is no mismatch.

"This means the binding of EcoRI protein on DNA does prevent MutS from sliding to ends." This statement is too strong.

On page 13

"However, when both ends of the DNA contained biotin, previous incubation with streptavidin increased the occupancy of the DNA substantially both in the presence and absence of ATP; this resulted in a significant decrease in the observed KD in the presence of ATP from 64 {plus minus} 9 nM to 25 {plus minus} 2 nM, which means that adding ATP cannot drive MutS off DNA because of either streptavidin blocked DNA ends or EcoRIs blocked DNA around the mismatch site." But it also increased nonspecific affinity 219 to 112 nM; blocking increases the affinity 2 fold on homoduplex DNA vs 2.6 fold on heteroduplex DNA, so this argument is not very convincing.

"DNAs with double-biotin ends that had been incubated with both streptavidin and EcoRIE111Q showed the highest MutS occupancy. For these 14 / 37 DNAs, KDs measured in the presence of ATP (21 nM, heteroduplex; 93 nM, homoduplex) were almost the same as those observed in its absence (18 nM, heteroduplex; 82 nM, homoduplex). It means all end-binding by streptavidin could prevent MutS from leaving the DNA after adding ATP, and as a result both the binding and the looping were improved. It also implies that the MutS translocates on DNA by sliding mechanism." I don't understand this argument.

Much of the text on page 13 and 14 is poorly written with several grammatical errors, making it even harder to follow.

On page 17:

"At the same time, the specificity of MutS for G/T vs. homoduplex DNA that we measured is 8.1 for no nucleotide condition and 3.4 when ATP was added. These results are similar to those obtained by others (Blackwell et al, 2001a; Schofield et al, 2001a). However, the KD for nonspecific binding that we report is much different from that reported by Yang et al (Yang et al, 2005). Our results differ in that we included MutS bound to the ends of DNA, which will increase the KD for nonspecific binding, because MutS has quite high affinity to bind to DNA ends (about 5% of DNA ends were bound by MutS and 15% of MutS that bound on DNA was found at its ends). In addition, Yang et al. defined the nonspecific dissociation constant as KD per base pair while we measured KD per DNA molecule. It is obvious that longer DNA has higher probability to be bound by MutS nonspecifically. Based on our findings in which 1120 bp was used, the fraction of nonspecific binding must be high." This section is fraught with errors. It is not proper to compare the binding affinities of DNA with 1000 nt with short DNA fragments. The authors do not seem to understand that by lengthening the DNA they increase the concentration of non-specific site. Their 1120 bp fragment contains approximately 1100 nonspecific sites and only one specific site; that is, the concentration of nonspecific sites is 1100 times greater than the specific site. Consequently if the specificity were 1100, one would expect that on their 1120 DNA approximately the same number of MutS proteins would be found to be bound specifically and nonspecifically. I think that their data are, in fact, consistent with Yang et al.

"The DNA seems to have propensity to bend in this region (Wang et al, 2003), and this may promote loop formation for MutS interacting with DNA at two sites, mismatch and a nonspecific site or even two nonspecific sites." Why is Wang et al cited here? Did they use the same DNA and see bending in the absence of MutS?

"The MutS occupancy of DNAs which were end-blocked or which were additionally blocked by EcoRIE111Q at sites flanking the mismatch was higher than that of DNA that haven't been blocked by any proteins, even in the absence of ATP; this implies that MutS was 'trapped' on these DNAs by the presence of the protein blocks." Only a 2-fold difference.

On page 20:

"For all the homoduplex DNAs on which MutS dimers formed loops, we divided them into two groups, short leg and long leg by their contour lengths. The length of the long legs could be longer

and shorter than ~50% of DNA length (mismatch position), but the average length of the long stayed around the center area of the DNA if MutS bound on DNA randomly." I am not sure I see the point of these statements"

"The dynamic behavior of the loops also differed in the two studies. In the present study, the size of the loops increased when either ATP or the non-hydrolyzable ATP analog AMPPnP was present, while in the study of Allen et al., nonhydrolyzable ATP analogues failed to support large loop formation and ongoing loop growth was suppressed upon their addition to ATP-containing reactions." Can't really say anything about dynamics because AFM images are static. More importantly, the "movement" seen by Allen et al was much greater than what is seen here. It is important to make this distinction.

"Other examples of loop formation by proteins which bind to and process DNA have been reported. Loops similar to those seen here were reported (Jia et al, 2008) who also found β -shaped loops using AFM. The loops they saw were independent of whether or not a mismatch was present in the DNA substrate. D835R (tetramer-deficient) and MutS R194A/R198A/R275A mutants of MutS formed β -loops at a similar frequency as wt MutS, suggesting that tetramer formation is not necessary for the formation of the loops that they saw. However, they did not provide the direct proof (volume measurements) to show that the MutS that formed β -loops was a dimer." Poorly written

Methods:

"In some experiments, HEPES reaction buffer (50 mM HEPES-KOH pH 8.0, 20 mM KCl, 5 mM MgCl₂, 1 mM DTT) or phosphate reaction buffer (20 mM KPO₄ pH 7.4, 20 mM KCl, 5 mM MgCl₂, 1 mM DTT) was used." Should be explicit which ones

"EcoRIE111Q and streptavidin binding were evaluated by AFM imaging." What was the efficiency?

Referee #3 (Remarks to the Author):

The revised version of this paper has improved it. There are two matters that I would advise the authors to think about. Neither of these requires any new experiments.

1. What is the pathway for loop formation? Does it begin immediately adjacent to the mismatch site with a very small loop that expands, or is there capture of non-specific DNA in the flanking regions, with initial loop size dependent on the kinetics of capture. The first of these models is not very appealing in energetic terms, since the energy of very small loops would be very large. I don't see anything in the data here that argues against the second model. Indeed, the difference in loop size with and without ATP follows naturally from the model: In the absence of ATP, sliding is inhibited, and the loop distribution depends primarily on capture kinetics. (In this regard, I note that the optimum loop size is a bit smaller than the persistence length of DNA, but loop formation likely gets a boost from the MutS induced bend.) In presence of ATP, the loop size distribution is dependent on the energetics of the system. This includes the loop free energy, as well as the presence of a "reflecting barrier" at the end of the molecule that prevents loops from getting larger. (It would be expected that in bigger molecules, as studied by Allen et al, entropy effects would allow much larger loops, since there is no corresponding reflecting barrier.) The model also predicts that loops containing EcoRI would be formed. Once formed, sliding past Eco could not happen, absent dissociation.

2 Why is the observed selectivity against non-specific DNA so low? If MutS is truly unable to distinguish mismatch from base pairs beyond a few fold specificity, how can repair work? In this context, the authors appear to have ignored the results of Huang and Crothers, who found much higher selectivity, based on an unambiguous competition assay. Rather than doubt other people's data, it is wise to ask why the results might be different. There is one striking difference between the two experiments: Huang and Crothers used oligonucleotides that are too small to support loop formation. That suggests a model in which the first step in duplex binding creates enhanced non-specific affinity by the second dimer, as would be required for formation of a stable loop including a mismatch site. Loop formation with duplex DNA could occur, with non-productive breakdown if the loop does not contain a mismatch site. Steady-state binding to non-specific DNA could then be much higher.

Thank you very much for reviewing our revised manuscript (EMBOJ-2011-77365, previously EMBOJ-2010-73591) "Atomic Force Microscopy Captures MutS Tetramers Initiating DNA Mismatch Repair" by Yong Jiang and Piotr E. Marszalek. We carefully followed and took into consideration all of your comments and your suggestions. One of the main points we significantly altered was making the main text of the paper shorter by focusing on key new findings. The character count of the new version is 49400, ten thousand words less compared to the 59300 of the old version. Changes introduced in the new manuscript are designated with tracking marks. They are explained in detail in the attached document, which also contains detailed responses to all of the referees' comments. Due to the vast number of changes in the main text, we also submitted the final version of the main text, without any markings. In addition, we moved some of the results that are more peripheral to the main message into the Supplementary Information section.

Response to the Comments and Suggestions from the Referees:

Referee #1 (Remarks to the Author):

Major Concerns:

This revised version is significantly improved from the previous version. The evidence for MutS-induced looping involving MutS bound to the mismatch is now much more convincing. I do believe that this work merits publication in EMBO, however, there are several significant errors in the paper. Although these errors do not affect the overall conclusion they need to be corrected. In addition, the results are very cumbersome to go through and follow. It seems that the paper would benefit greatly by being significantly shortened, moving many of the results to the Supplemental material. Many of these results are not central to the conclusion of the paper and are not really discussed in the Discussion section. For example, the discussion of affinities is very difficult to follow and is not particularly pertinent to the conclusion. In my opinion, the exciting and truly important results are the observation that one of the MutS molecules resides at or near the mismatch in the presence of ATP. Shortening the results will make this work interesting to a broader audience and it will have more impact overall. Either way, the points need to be addressed before publishing.

Reply: We corrected the errors identified by the referee, significantly shortened the manuscript and followed the referee's suggestion to move the results that are not central to the conclusion of the paper to the Supplementary Information (SI) section.

Specific comments:

On page 9: "Single-site-bound MutS can be divided into two groups: specific-bound MutS at the relative position of 50±6% and nonspecific-bound MutS at 33±5% in the absence of any nucleotides (Figure S3A and S3B). However, in the presence of ATP, MutS are found to be located at 50±5% and 23±11% of the relative DNA length (Figure S3C and S3D). This analysis is incorrect. Fitting of the nonspecific complexes to a Gaussian does not make any sense, because the non-specific complexes should be represented by a random distribution not a Gaussian (by definition non-specific means no preference for a given site). It is somewhat surprising that the non-specific complexes in the presence of ADP are not randomly distributed."

Reply: Following the referee's suggestion, the Gaussian fit for nonspecific-bound MutS was removed and we rewrote the section.

"Single-site-bound MutS can be divided into two groups: specific-bound and nonspecific-bound. The specific-bound MutS located at the relative position of 50±6% and 50±5% of the relative DNA length before and after ATP was added, respectively. (Figure S3C and S3D)."

On page 10: "However, when ATP was added, the DNA of the short leg moved through the nonspecifically-bound MutS dimer and entered into the loop." This statement does not belong in the

results because AFM does not provide direct evidence of movement. Here and several other places in the results the authors state that the protein has moved along the DNA, although they have no direct evidence. This point is particularly relevant because loops are seen in the absence of ATP so there is no way to know if the protein moved along the DNA to arrive at that point or it already bound there. Only when all of the data is put together does the data support movement.

Reply: We followed the referee's suggestion, removed the statement that DNA moved along the DNA here and after, and only propose the sliding mechanism at the end of discussion.

On page 11: "We pooled together positions of all MutS particles in the loop configuration and produced their distribution (Supplementary Figure S4A and S4B)." Figure S4 is very confusing and the discussion of this figure at multiple points in the paper is hard to follow. It is not clear what the point of this analysis is. It should either be removed completely or put in to Supplemental material.

Reply: We followed the referee's suggestion, and removed Supplementary Figure S4 and all the text related to this Figure completely.

On page 12: "Figure S4 shows, MutS has no significant binding affinity for the region between 45% and 50% of homoduplex DNA length." I believe you mean binding specificity not binding affinity. Clearly MutS binds to this region.

Reply: We followed the referee's suggestion, and changed "binding affinity" to "binding specificity".

On page 12: "Third, loops were seen on homoduplex DNA with a frequency about 3.2 (7.6% / 2.4%) times less than that on heteroduplex DNA at the same conditions (Table 3)." This is not a correct comparison because there are fewer complexes on homoduplex DNA versus heteroduplex DNA because MutS has a lower affinity to homoduplex DNA. Calculating the percent of total complexes that form loops is what should be compared. 34% of the complexes form loops on heteroduplex DNA and 22% complexes form loops on homoduplex DNA, making the frequency 1.5 times greater not 3.2. A similar error is made later in the text. Such as on page 13 "Though loops formed on both homoduplex and G-T mismatch-containing DNA and on both non-end-blocked and end-blocked substrates, selectivity was about 8.1-fold (179/22) for DNA with a G-T mismatch over homoduplex DNA in the absence of nucleotide and it was about 3.4-fold (219/64) in the presence of ATP." These numbers don't take into account the different number of complexes.

Reply: We followed the referee's suggestion, and re-calculated all of the ratios by adding the occupancy.

On page 12: "The plausible explanation for this observation is that on homoduplex DNA both MutS dimers that formed the loop were bound on DNA nonspecifically." What's the point of this statement? The both have to be non-specific because there is no mismatch.

Reply: We followed the referee's suggestion, and removed the sentence.

On page 12: "This means the binding of EcoRI protein on DNA does prevent MutS from sliding to ends." This statement is too strong.

Reply: We followed the referee's suggestion, and rewrote the sentence.

"This result suggests that the binding of EcoRIE111Q protein on DNA does prevent the increase in loop size."

On page 13: "However, when both ends of the DNA contained biotin, previous incubation with streptavidin increased the occupancy of the DNA substantially both in the presence and absence of ATP; this resulted in a significant decrease in the observed KD in the presence of ATP from 64±9 nM to 25±2 nM, which means that adding ATP cannot drive MutS off DNA because of either streptavidin blocked DNA ends or EcoRIs blocked DNA around the mismatch site." But it also increased nonspecific affinity 219 to 112 nM; blocking increases the affinity 2 fold on homoduplex DNA vs 2.6 fold on heteroduplex DNA, so this argument is not very convincing.

Reply: Following the referee's recommendation we moved the result section about binding affinities to SI, where that sentence now reads:

"However, when both ends of the DNA were biotinylated, previous incubation of the DNA with streptavidin increased the occupancy of the DNA substantially, both in the absence and the presence of ATP. This resulted in a significant decrease in the observed KD in the presence of ATP from 64 ± 9 nM to 25 ± 2 nM and from 219 ± 19 nM to 112 ± 10 nM for homoduplex DNA, which suggests that adding ATP cannot drive MutS off DNA because of either streptavidin blocked DNA ends or EcoRIE111Qs blocked DNA."

On page 13: "DNAs with double-biotin ends that had been incubated with both streptavidin and EcoRIE111Q showed the highest MutS occupancy. For these DNAs, KDs measured in the presence of ATP (21 nM, heteroduplex; 93 nM, homoduplex) were almost the same as those observed in its absence (18 nM, heteroduplex; 82 nM, homoduplex). It means all end-binding by streptavidin could prevent MutS from leaving the DNA after adding ATP, and as a result both the binding and the looping were improved. It also implies that the MutS translocates on DNA by sliding mechanism." I don't understand this argument. Much of the text on page 13 and 14 is poorly written with several grammatical errors, making it even harder to follow.

Reply: We followed the referee's suggestion, and rewrote that sentence to improve its clarity and moved it to Supporting Information.

"Further, when DNAs were incubated with both streptavidin and EcoRIE111Q, the KD decreased even more. For these DNAs, KDs measured in the presence of ATP (21 nM, heteroduplex; 93 nM, homoduplex) were almost the same as those observed in its absence (18 nM, heteroduplex; 82 nM, homoduplex). This result suggests that the presence of both streptavidin and EcoRIE111Q prevents MutS from leaving the DNA, especially when ATP is added. As a result both the binding occupancy and the frequency of looping increased."

On page 17: "At the same time, the specificity of MutS for G-T vs. homoduplex DNA that we measured is 8.1 for no nucleotide condition and 3.4 when ATP was added. These results are similar to those obtained by others (Blackwell et al, 2001a; Schofield et al, 2001a). However, the KD for nonspecific binding that we report is much different from that reported by Yang et al (Yang et al, 2005). Our results differ in that we included MutS bound to the ends of DNA, which will increase the KD for nonspecific binding, because MutS has quite high affinity to bind to DNA ends (about 5% of DNA ends were bound by MutS and 15% of MutS that bound on DNA was found at its ends). In addition, Yang et al. defined the nonspecific dissociation constant as KD per base pair while we measured KD per DNA molecule. It is obvious that longer DNA has higher probability to be bound by MutS nonspecifically. Based on our findings in which 1120 bp was used, the fraction of nonspecific binding must be high."

This section is fraught with errors. It is not proper to compare the binding affinities of DNA with 1000 nt with short DNA fragments. The authors do not seem to understand that by lengthening the DNA they increase the concentration of non-specific site. Their 1120 bp fragment contains approximately 1100 nonspecific sites and only one specific site; that is, the concentration of nonspecific sites is 1100 times greater than the specific site. Consequently if the specificity were 1100, one would expect that on their 1120 DNA approximately the same number of MutS proteins would be found to be bound specifically and nonspecifically. I think that their data are, in fact, consistent with Yang et al.

Reply: Following the referee's recommendation we moved the section about binding affinities entirely to SI and we do not discuss these results in the main text. We also deleted most of the previous comparison of our results about binding affinities that were obtained on a 1120-bp DNA substrate to the results obtained by others on shorter DNAs. In SI, we included a following paragraph to reflect the referee's comment (this paragraph also includes our response to a similar point raised by the third referee):

"The KD for nonspecific binding determined from AFM images is about 80~300 nM. Because this 1120-bp DNA contains >1000 nonspecific sites and only one specific site, the concentration of nonspecific sites is about 1000 times greater than that of the specific site. Consequently the true KD for nonspecific binding is 1000 times higher, which is consistent with that reported by Yang et al (Yang et al, 2005). This very high true KD for nonspecific binding is also quite consistent with a very high MutS specificity reported by Huang and Crothers (Huang and Crothers, 2008)."

"The DNA seems to have propensity to bend in this region (Wang et al, 2003), and this may promote loop formation for MutS interacting with DNA at two sites, mismatch and a nonspecific site or even two nonspecific sites." Why is Wang et al cited here? Did they use the same DNA and see bending in the absence of MutS?

Reply: We followed the referee's suggestion, and removed this reference in this section.

"The MutS occupancy of DNAs which were end-blocked or which were additionally blocked by EcoRIE111Q at sites flanking the mismatch was higher than that of DNA that haven't been blocked by any proteins, even in the absence of ATP; this implies that MutS was 'trapped' on these DNAs by the presence of the protein blocks." Only a 2-fold difference.

Reply: We followed the referee's suggestion, and rewrote the sentences.

"The MutS occupancy of DNAs which were end-blocked or which were additionally blocked by EcoRIE111Q at sites flanking the mismatch was about 2~3-fold higher than that of DNA that have not been blocked by any proteins in the presence of ATP; this suggests that MutS was 'trapped' on these DNAs by the presence of the protein blocks."

On page 20: "For all the homoduplex DNAs on which MutS dimers formed loops, we divided them into two groups, short leg and long leg by their contour lengths. The length of the long legs could be longer and shorter than ~50% of DNA length (mismatch position), but the average length of the long stayed around the center area of the DNA if MutS bound on DNA randomly." I am not sure I see the point of these statements.

Reply: We followed the referee's suggestion, and removed the sentences.

On page 20: "The dynamic behavior of the loops also differed in the two studies. In the present study, the size of the loops increased when either ATP or the non-hydrolyzable ATP analog AMPPnP was present, while in the study of Allen et al., nonhydrolyzable ATP analogues failed to support large loop formation and ongoing loop growth was suppressed upon their addition to ATP-containing reactions." Can't really say anything about dynamics because AFM images are static. More importantly, the "movement" seen by Allen et al was much greater than what is seen here. It is important to make this distinction.

Reply: We followed the referee's suggestion, removed the sentences "The dynamic behavior of the loops also differed in the two studies.", and added the sentence "Another important difference between our results and those by Allen et al is that the DNA loops captured by AFM were on average significantly smaller than the loops captured by EM, which may be related to the smaller DNA substrate used in our study (1.12 kbp vs 6.4 kbp)."

On page 20: "Other examples of loop formation by proteins which bind to and process DNA have been reported. Loops similar to those seen here were reported (Jia et al, 2008) who also found -shaped loops using AFM. The loops they saw were independent of whether or not a mismatch was present in the DNA substrate. D835R (tetramer-deficient) and MutS R194A/R198A/R275A mutants of MutS formed -loops at a similar frequency as wt MutS, suggesting that tetramer formation is not necessary for the formation of the loops that they saw. However, they did not provide the direct proof (volume measurements) to show that the MutS that formed -loops was a dimer." Poorly written

Reply: We followed the referee's suggestion, and rewrote the paragraph.

" Using AFM Jia et al also reported similar -shaped loops formed by MutS (Jia et al, 2008). The loops they imaged were independent of whether or not a mismatch was present in the DNA substrate. Furthermore, they found MutS D835R (tetramer-deficient) and R194A/R198A/R275A mutants of MutS formed -loops at a similar frequency as the wild type MutS, suggesting that tetramer formation is not necessary for the formation of the loops in their system. However, that study did not provide evidence (for example, by MutS volume measurements) that the MutS that formed -loops was in a dimer configuration."

Methods: "In some experiments, HEPES reaction buffer (50 mM HEPES-KOH pH 8.0, 20 mM KCl, 5 mM MgCl₂, 1 mM DTT) or phosphate reaction buffer (20 mM KPO₄ pH 7.4, 20 mM KCl, 5 mM

MgCl₂, 1 mM DTT) was used." Should be explicit which ones. "EcoRIE111Q and streptavidin binding were evaluated by AFM imaging." What was the efficiency?

Reply: We followed the referee's suggestion, and revised this section.

"For comparison, in some experiments, HEPES reaction buffer (50 mM HEPES-KOH pH 8.0, 20 mM KCl, 5 mM MgCl₂, 1 mM DTT) or phosphate reaction buffer (20 mM KPO₄ pH 7.4, 20 mM KCl, 5 mM MgCl₂, 1 mM DTT) was used."

About the efficiency:

"The statistical results show that about 61% of DNA had both ends blocked by EcoRIE111Q. For double-biotin labeled DNA, 78% of DNA had both ends blocked by streptavidin. For single-biotin labeled DNA, 89% of DNA had one end blocked by streptavidin. These specific percentages are not critical, because when contour lengths or a KD was determined, only the DNA molecules with both ends blocked by streptavidin were counted (for experiments with double-biotin labeled DNA)."

Referee #3 (Remarks to the Author):

The revised version of this paper has improved it. There are two matters that I would advise the authors to think about. Neither of these requires any new experiments.

1. What is the pathway for loop formation? Does it begin immediately adjacent to the mismatch site with a very small loop that expands, or is there capture of non-specific DNA in the flanking regions, with initial loop size dependent on the kinetics of capture. The first of these models is not very appealing in energetic terms, since the energy of very small loops would be very large. I don't see anything in the data here that argues against the second model. Indeed, the difference in loop size with and without ATP follows naturally from the model: In the absence of ATP, sliding is inhibited, and the loop distribution depends primarily on capture kinetics. (In this regard, I note that the optimum loop size is a bit smaller than the persistence length of DNA, but loop formation likely gets a boost from the MutS induced bend.) In presence of ATP, the loop size distribution is dependent on the energetics of the system. This includes the loop free energy, as well as the presence of a "reflecting barrier" at the end of the molecule that prevents loops from getting larger. (It would be expected that in bigger molecules, as studied by Allen et al, entropy effects would allow much larger loops, since there is no corresponding reflecting barrier.) The model also predicts that loops containing EcoRI would be formed. Once formed, sliding past EcoRI could not happen, absent dissociation.

Reply: The referee asked a very important question about the mechanism of loop formation and expansion and s/he provided very interesting and inspiring ideas about these processes. We will definitely consider these suggestions when planning future experiments, possibly using fast AFM imaging instruments that should be available soon, and that can capture dynamics processes of DNA in real time. However, because of current spatial and temporal resolution limitations of our AFM, we cannot provide more mechanistic detail about the origin and dynamics of the DNA loops that we captured. We hope that further studies will unravel new details and will verify these important concepts.

2 Why is the observed selectivity against non-specific DNA so low? If MutS is truly unable to distinguish mismatch from base pairs beyond a few fold specificity, how can repair work? In this context, the authors appear to have ignored the results of Huang and Crothers, who found much higher selectivity, based on an unambiguous competition assay. Rather than doubt other people's data, it is wise to ask why the results might be different. There is one striking difference between the two experiments: Huang and Crothers used oligonucleotides that are too small to support loop formation. That suggests a model in which the first step in duplex binding creates enhanced non-specific affinity by the second dimer, as would be required for formation of a stable loop including a mismatch site. Loop formation with duplex DNA could occur, with non-productive breakdown if the loop does not contain a mismatch site. Steady-state binding to non-specific DNA could then be much higher.

Reply: A similar point about MutS selectivity was also raised by Referee 1, who suggested a reasonable explanation of our low specificity result. Please see his/her comments on page 7. Following the first referee's recommendation we moved the section about binding affinities entirely

to SI and we do not discuss these results in the main text. In SI, we included a following paragraph that references the results of Yang et al (2005) and by Huang and Crothers (2008) .

"The KD for nonspecific binding determined from AFM images is about 80~300 nM. Because this 1120-bp DNA contains >1000 nonspecific sites and only one specific site, the concentration of nonspecific sites is about 1000 times greater than that of the specific site. Consequently the true KD for nonspecific binding is 1000 times higher, which is consistent with that reported by Yang et al (Yang et al, 2005). This very high true KD for nonspecific binding is also quite consistent with a very high MutS specificity reported by Huang et al (Huang and Crothers, 2008)."

Acceptance letter

10 May 2011

Thank you for submitting your revised manuscript for our consideration. I have now had a chance to look through it and to assess your responses to the comments raised by the original reviewers, and I am happy to inform you that there are no further objections towards publication in The EMBO Journal.

We will send you a formal letter of acceptance, and we should then be able to swiftly proceed with the production of your manuscript!

Yours sincerely,

Editor
The EMBO Journal