THE
EMBO
JOURNAL

Real-time observation of flexible domain movements in CRISPR-Cas9

Saki Osuka, Kazushi Isomura, Shohei Kajimoto, Tomotaka Komori, Hiroshi Nishimasu, Tomohiro Shima, Osamu Nureki and Sotaro Uemura

Review timeline:

Submission date:	29 March 2017
Editorial correspondence:	1 June 2017
Editorial Decision:	8 June 2017
Revision received:	8 September 2017
Editorial correspondence:	14 December 2017
Editorial Decision:	22 December 2017
Revision received:	7 Febuary 2018
Accepted:	13 Febuary 2018

Editor: Anne Nielsen

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.)

Editorial correspondence 1 June 2017

Thank you for submitting your manuscript for consideration by the EMBO Journal and my apologies for the extended duration of the review period. You study has now been seen by three referees whose comments are shown below; in addition, we have consulted with an additional technical advisor based on the specific FRET-criticisms raised by ref #1.

As you will see from the reports, all three referees find the topic and findings interesting and valuable to the field. However, while refs #2 and #3 consequently support publication following clarification of a few additional points, ref #1 is very critical about the experimental setup and the technical quality of the FRET data. Based on this discrepancy in the referee recommendations, I consulted with an additional FRET expert (who saw both the manuscript and the referee reports) and this person agrees that the current data quality is low and that the additional controls requested by ref #1 are warranted.

Given the extensive and severe nature of these technical issues - and since we generally allow a single round of revision only - I would like to discuss if/how you would be able to address these criticisms in a potential revision before I go on to make an official decision on this manuscript.

I would therefore ask you to take a look at the reports included below and let me know what kind of data you would be able to include in a potential revision to address the referee concerns. I would then take that into consideration - and possibly also discuss it with the referees - before we make a final decision on your study. The aim of this is ultimately to prevent you from working extensively on a revision that would have little chance of convincing the referees.

© European Molecular Biology Organization

1st Editorial Decision 8 June 2017

Dear Tomohiro,

Thank you for sending a preliminary point-by-point response to the concerns raised by the referees. I have now read it and discussed it with a colleague in the editorial team and the conclusion is that we would invite you to submit a revised manuscript along the lines outlined in your response.

I realise that the outcome of some of the experiments you offer to include cannot be predicted at this point and I want to emphasise that we will only be able to proceed to publication if the original conclusions of your manuscript still hold true after revision. In addition, we will need the referees to recommend/support publication of the revised manuscript.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

REFEREE REPORTS

Referee #1:

Osuka et al report the observation of SpCas9 structural dynamics in real time using single-molecule FRET technique. The authors propose an interesting hypothesis that the dynamic nature of the domain movements is crucial for the Cas9 nuclease activity. If valid, this work has importance for the understanding the target recognition and cleavage mechanism of Cas9, which will help enhance Cas9 genome-editing technology. However, this reviewer is not convinced with their data. The quality of the single-molecule data is poor. The design of the experiments has a lower standard than other single-molecule papers published in high impact journals including EMBO Journal have. Furthermore, the interpretation of data is only speculative, therefore does not provide biological insights beyond just reporting the observation of the dynamics. Thereby, this reviewer does not support the publication of this work in EMBO J. Below I address comments for publication of this work in a more specialized journal (e.g. Nucleic Acids Research).

1. The interpretation of their observation is mainly based on the FRET histograms in Figures 2 and 4. Overall, the pattern of the histograms is very vague (unlikely reproducible). Thereby it is difficult to accept the way how the authors interpreted FRET peaks. This reviewer make several suggestions to improve the confidence of the FRET histogram data.
a. In many cases, the authors only speculate on the pattern of the histograms. Since FRET peaks can emerge from several other reasons than authors believe (e.g. poor statistics, background signals, heterogeneity of sample), the authors should vary experimental conditions and then assign the biological meaning to each peak. For example, they could vary the degree of basepair match between crRNA and target RNA (as Dagdas et al, BioRxiv, 2017 did), vary the magnesium concentrations, or mutate a region of interest in Cas9. When a peak of interest gradually changes under the systematic variations, the interpretation becomes valid.
b. The statistics for the FRET histograms is extremely poor. The number of molecules analyzed for each FRET histogram in Figure S3 is only about 100. (It is not clear how many molecules were used for Figure 2). To eliminate artifact such as heterogeneity of sample, FRET histograms should be built from much more than 1000 molecules over several different imaging fields. From the experimental procedure, it appears that they recorded images from only one area. This has no statistical significance. This measurement should have been repeated using different sample chambers (in different days) to eliminate artifacts from background signals.
2. The statistics for the time traces is also poor. In a certain condition (S867C-N1054C), only 4, 6, and 10 time traces were analyzed for each case. These few time traces cannot be representative enough. Therefore the data in Figure 3D (percentage of fluctuating molecules) has little statistical confidence.
3. It is stressed in the manuscript that HNH enters a cleavage-competent state. However, no data supporting this is shown in the manuscript. Moreover, since HNH can enter this domain several times, it is unclear whether it does cleave or not and if the domain can move after cleavage. Design

© European Molecular Biology Organization

an assay which would relate the percentage of molecules accessing the „cleavage competent" conformation and the percentage of molecule that has been cleaved. This could be an assay that involves DNA immobilized on the surface, for example, an assay from Dagdas et al (partial duplex target) or Sternberg et al (Nature, 2014, denaturation after reaction).
4. Despite pre-incubating Cas9 with excess of nucleic acids, the observed FRET shift is an indirect indication of domain flexibility. Use a reporter dye on sgRNA/DNA molecules and performing real-time flow measurements to observe transitions between different flexible states when nucleic acids bind the immobilized Cas9. See also a minor comment about plasmid DNA.
5. The author described that FRET value increased when sgRNA was added (Figure 2B, Apo vs +sgRNA). This „gradual increase" is not clearly noticed in Figure 2B.
6. When DNA is added (Figure 2B, +sgRNA +DNA), the FRET value did increase. However, this change is not as drastic as that was observed by Dagdas et al (BioRxiv, 2017) who used the same protein mutant. In the case of Dagdas et al, the low FRET (E = 0.19) became very high FRET (E = 0.97). Please discuss the discrepancy.
7. Another discrepancy between this work and the work of Dagdas et al (and also Lim et al, „Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease" Nature Communications, 2016) is that the fluctuation is observed more when target DNA is present than absent (this work) while it is less observed when on-target is used than off-target is used (Dagdas et al, Lim et al). Is it because plasmid DNA was used as target? Please discuss the discrepancy.
8. In Figure 3, the selection of static and fluctuating molecules seem arbitrary. The authors should provide a quantitative method in sorting the traces into two categories. For example, they said that they considered it is a fluctuating species if they observe more than two fluctuations. It has to be defined between which states the fluctuation occurred because any FRET traces will exhibit anti-correlated donor-acceptor fluctuation just like Figure 3A (static) trace shows.
9. Trace in figure 3 c) shows both multi-state fluctuations and an interval where it is static. Would such a trace be assigned to static, fluctuations or both?
10. Figure 4B (+sgRNA+DNA -Mg) looks analogous to Figure 4A (+sgRNA+DNA+Mg), with only lightly lower high FRET peaks. It is unclear how these peaks disappear once you introduce magnesium in the case of static molecules.
11. In Figure 5C it is confusing what is meant by „duration" and how the numbers are determined. Were the numbers obtained from the kinetic rates that hidden Markov analysis generates? Or were the dwell times measured from idealized time traces generated by hidden Markov analysis? The analysis procedure by the authors indicate it is the latter. The authors should then note that the dwell-time analysis (unlike hidden Markove analysis) cannot be correctly used for obtaining the kinetic rates. Taking an example of state A making a transition to either B or C, the dwelltime of A to B transition is non-intuitively identical to that of A to C. This is because states B and C directly compete with each other.
12. In Figure 5C, show dwelltime histograms for clarity.
13. The Results part of the manuscript requires revision. Speculation appears very often between results. Since the speculation is mostly over-interpretation of data, it makes it difficult to read.

Minor comments:
1. The authors used plasmid DNA as target DNA. Would bulky plasmid DNA properly interact with surface-tethered Cas9? Explain why plasmid DNA was used instead of oligo DNA unlike regular single-molecule FRET experiments.
2. Plasmid DNA will have potentially many off-target sequences. Include the sequenc of the plasmid DNA and indicate the number of on-target and off-target sequences.
3. For the construct S867C-N1054C, the authors may observe fluctuation if they use a dye pair that has smaller Forster radius than Cy3 and Cy5.
4. It is unclear if the FRET histograms in Figures 2 and 4 are obtained from a snapshot analysis or from time traces.
5. The authors should indicate whether/how they did donor leakage correction. Without this indication, it is difficult to know whether the lowest FRET peak in a FRET histogram is from a donor-only population or from a low-FRET population. Likewise, it is not clear whether the long static state in Figure 3C is only photoblinking of acceptor or not.
6. FRET histograms are hard to read, especially where the distribution is broad. Fit using gaussian peaks to determine what FRET states are present.
7. In Figure 3D, using a legend to describe which colour means what would be easier to read.
8. Describe how PEG surface was prepared.
9. The manuscript lack accuracy in description. Vague description (e.g., the majority of , wide

distribution, narrow distribution, ..) are too often used without quantification.

Referee #2:

General summary
Structural studies have captured Cas9 in several conformational states and bulk FRET experiments have shown that sgRNA-guided DNA-binding controls the conformational state of the HNH-nuclease domain. However, the structural studies only capture static snap-shots of the complex in distinct poses and the bulk FRET experiments capture the average behavior. Single-molecule experiments are necessary to clarify the conformational pathway to sgRNA-guided Cas9-mediated cleavage of a DNA target. The work by Osuka et al, employ single-molecule methods to address this outstanding question. The topic is timely and the approach well-suited to address important mechanistic questions, but I think the authors need to do a better job of explaining what was done previously, and what questions were left outstanding by the previous work.

Major concerns

Ln 131: The authors state "The sgRNA/DNA-bound molecules in our assay should maintain the ternary complex of the sgRNA and cleaved target DNA, because previous studies have demonstrated that Cas9 cleaves the target DNA at a rate higher than 10 min-1 and retain binding with cleaved DNA (Sternberg et al, 2014; Jinek et al, 2012; Sternberg et al,
135 2015)." Did the cited papers show that both end of the cleaved DNA are retain equally or that the PAM end is preferentially retained? How is the state of cleavage and the impacts of DNA release accounted for in these experiments? How do we know that the dynamics observed for some particle are reporting on the release of DNA or part of the DNA target after cleavage?

The authors state, "As a common property of the D435C-E945C and S355C-S867C constructs, sgRNA binding decreased the percentage (Figure 2B)," I think would be helpful to explicitly state what "percentage" you are refereeing to. Also, the "decrease" for S355C-S867C is not evident in Fig 2B. In fact, the percent frequency in the low FRET state appears to increase modestly after sgRNA binding.

It is unclear how the authors know if sgRNA-Cas9 is bound to a molecule of DNA is cleaved or not. How does DNA cleavage change the dynamics? Clearly Mg has a large impact on the distribution of conformational states but does being in the "DNA-cleavage position" mean that the DNA is cleaved or just that it sampled this position?
Ln 231: "in the absence of magnesium ions Cas9 can be trapped in the pre-cleavage state with sgRNA and DNA." Are there any other roles for the metal that might contribute to conformational state? Was this experiment ever performed using Mg-containing HNH and a non-cleavage substrate (e.g., Phosphorothioate Oligonucleotides)?

Minor concerns
a linear representation of the protein and its associated domains colored according to the structure would help orient the reader.

References: The hyperlink to Zheng's modeling paper is incorrect. The link is to a paper on the SNARE complex.

Pg6 Ln 106 "single molecular level". Consider changing this to "single molecule level"

Fig EV1. It appears that more of the DNA target is cleaved with lower (25nm) concentrations of Cas9. Is this correct or is the gel mislabeled?

Ln 221: Should "access" be "assess"?

Referee #3:

Manuscript # EMBOJ-2017-96941

Real-time observation of flexible domain movements in Cas9
By Saki Osuka, Kazushi Isomura, Shohei Kajimoto, Tomotaka Komori, Hiroshi Nishimasu, Osamu Nureki, Sotaro Uemura

The CRISPR-associated nuclease Cas9 is an RNA-guided DNA-cutting enzyme that has been repurposed for genome editing. Understanding the molecular mechanism of Cas9 is thus critical for further development of genome editing technologies and applications. Cas9 is a multidomain protein that has been shown by X-ray crystallographic studies to undergo a series of conformational rearrangements. The study by Osuka et al. aims to investigate the conformational dynamics of Cas9 by single-molecule FRET, which enables real-time observation of domain motions. The study reveals that multiple Cas9 domains undergo dynamic rearrangements that allow the enzyme to sample a range of conformations, thus affecting guide RNA and target DNA binding by the enzyme. In particular, the HNH nuclease domain only adopts its catalytically active conformation by undergoing a dynamic fluctuation, which underscores the importance of conformational dynamics for the nuclease activities of Cas9.

Overall, this is a very good manuscript that provides interesting insights into the molecular mechanism of Cas9. The study is technically sound and the conclusions drawn by the authors are largely supported by the data. The study advances our understanding of this important genome editor nuclease and is likely to be of high interest to both the CRISPR-Cas and genome editing research communities, as well as to users of genome editing techniques. I would recommend the manuscript for publication pending appropriate revisions, and have only a few comments, as detailed below.

Major comments:

1. Based on the results presented in Fig. 3D, the authors conclude that the flexibility of the hinge regions between the nuclease and REC lobes increases in the DNA-bound complex based on the observation of an increased fraction of fluctuating molecules. However, this result could also be reached just by increased dynamics within the REC domain itself, especially given that the RuvC-REC and HNH-REC distances are monitored relative to different locations within the REC lobe (D435 vs. S355). It would be good if the authors could also comment on this possibility and also be more specific in indicating which hinge regions are likely to be involved.

2. In this context, it is also important to consider the location of the covalently attached fluorophores in the Cas9 protein constructs. Inspecting the crystal structures of the various Cas9 functional states, it seems that Glu945 is located within a beta-hairpin loop that is likely to be flexible by itself in the sgRNA- and sgRNA/DNA-bound structures (it is held by crystal contacts in some structures and has very high B-factors in the others). Is it possible that some of the wide distribution of FRET efficiencies observed for the D435C-E945C constructs in the sgRNA- and sgRNA/DNA-bound states could be attributed to the intrinsic flexibility within this region?

3. The authors concede that the low frequency of fluctuating molecules observed for the S867C-N1054C construct might be due to the relatively short distance between the residues (p. 10) and hence it is not appropriate to make quantitative comparisons of the relative flexibilities of the three domains in the three constructs. Based on this, I find the conclusion that flexible domain movements facilitate Cas9 adopting the static conformation of the next nucleic acid binding state somewhat tenuous.

4. In Fig. 4B, analysis of time trajectories indicates that transitions between middle ("pre-cleavage) and high (catalytically competent) FRET states are very infrequent. The authors conclude that the HNH domain rarely moves directly from the pre-cleavage to the cleavage state and needs to return to the undocked conformation first. I do not think that this is such a surprising result from a probabilistic standpoint, given the data. The distribution of FRET efficiencies in the DNA-bound state shows that the HNH domain spends most time in the undocked conformation and therefore, the undocked state is thermodynamically favoured one, As a result, the domain will always be more likely to move back to the undocked conformation. However, an alternative interpretation of the results could be that rather than being a bona fide on-pathway intermediate, the "pre-cleavage" mid-FRET state is in fact a dead-end, off-pathway state that needs to be resolved by collapse to the undocked conformation. The authors should comment on this possibility in the manuscript.

Minor comments:

5. The authors reference Zheng 2017 as the only molecular dynamics study of the Cas9 system. However, there have now been several other computational studies of Cas9 (Zuo, Z., and Liu, J. (2016). Cas9-catalyzed DNA Cleavage Generates Staggered Ends: Evidence from Molecular Dynamics Simulations. Sci Rep 5, 37584; Palermo, G., et al. (2016). Striking Plasticity of CRISPR-Cas9 and Key Role of Non-target DNA, as Revealed by Molecular Simulations. ACS Cent Sci 2, 756-763), which should also be referred to.
6. On page 8, the authors state the fluorophores displayed a high degree of fluorescence anisotropy. Although this is further detailed in the Methods section, it would perhaps be better to show these results as a supplemental figure.

7. On p. 11, the authors state that addition of Mg2+ ions enhances the fraction of fluctuating Cas9 molecules and refer to Fig. 3D. However, there is no mention of Mg2+ in Fig. 3D, nor in the figure legend. Do the black bars refer to data obtained for sgRNA+DNA-bound Cas9 in the presence of Mg2+? If not, the authors should add the Mg2+ data to the figure.

1st Revision - authors' response 8 September 2017

© European Molecular Biology Organization

We greatly appreciate all of the referees for the valuable and constructive comments, and have addressed their concerns by adding more text, data, figures and references. We hope that the revised manuscript satisfies the referees' comments.

Our point by point responses to the referees' concerns are provided below.

(*Blue italic*: referee comments, black: our responses)

Referee #1

1. The interpretation of their observation is mainly based on the FRET histograms in Figures 2 and 4. Overall, the pattern of the histograms is very vague (unlikely reproducible). Thereby it is difficult to accept the way how the authors interpreted FRET peaks.

When the distance between the fluorescently-labeled domains fluctuates, the FRET efficiency should also fluctuate. This correlation would explain the vague pattern of the FRET efficiency histograms. Therefore, the vague pattern is not because of the irreproducibility of the data. In fact, we performed the same experiments at least three times for each condition to confirm the reproducibility. To clarify this point, we added a description regarding the reproducibility in the methods section (Pg. 24, Ln. 517-519).

This reviewer makes several suggestions to improve the confidence of the FRET histogram data.

a. In many cases, the authors only speculate on the pattern of the histograms. Since FRET peaks can emerge from several other reasons than authors believe (e.g. poor statistics, background signals, heterogeneity of sample), the authors should vary experimental conditions and then assign the biological meaning to each peak. For example, they could vary the degree of basepair match between crRNA and target RNA (as Dagdas et al, BioRxiv, 2017 did), vary the magnesium concentrations, or mutate a region of interest in Cas9. When a peak of interest gradually changes under the systematic variations, the interpretation becomes valid.

We appreciate the suggestion and performed additional experiments with various Mg^{2+} concentrations. As shown in Figure EV4, the appearance of the high FRET peak corresponded to the Mg^{2+} concentration. This result provides strong support for our conclusion.

b. The statistics for the FRET histograms is extremely poor. The number of molecules analyzed for each FRET histogram in Figure S3 is only about 100. (It is not clear how many molecules were used for Figure 2). To eliminate artifact such as heterogeneity of sample, FRET histograms should be built from much more than 1000 molecules over several different imaging fields.

We agree that a larger number of samples improves the statistical power of the experiments, but traces from far less than 1,000 molecules can be sufficient to obtain statistically significant conclusions. In McKinney et al., Biophysical Journal, (2006) 91, pp.1941-1951, which is a highly cited (438 times to date) and trusted methodological paper for single molecule FRET (smFRET), the authors state, "even with as few as 20 traces, a number routinely obtained in single molecule experiments, the statistical contribution to the error is minimal, at only 8% for the standard parameters". Accordingly, the differences between the low, middle and high FRET peaks in Fig 4 of our manuscript are larger than 8%, and thus cannot solely be explained by experimental artifacts. In addition, the typical number of molecules analyzed in recently published smFRET papers is 100-200 (Lim et al., Nat. Commn., (2016) 7:13350, Wang et al., Nat. Struc. Mol. Biol., (2016), 23, pp.31-36). Additionally, the referee refers to Dagdas et al. (2017) bioRxiv, which was submitted to the preprint server after the submission of our first manuscript. Please note that the numbers of traces measured in Dagdas et al. are 57–353 (typically ~100). To conclude, we believe our assays meet the recent experimental standards.

From the experimental procedure, it appears that they recorded images from only one area. This has no statistical significance. This measurement should have been repeated using different sample chambers (in different days) to eliminate artifacts from background signals.

As described above, we repeated the same experiments at least three times. Each time, we used different sample chambers and recorded images from several areas of the chamber. In our assay system, fewer than 20 molecules were recorded in one area. In the revised manuscript, we rewrote the methods section to clarify these points (Pg. 24, Ln 517-519).

2. The statistics for the time traces is also poor. In a certain condition (S867C-N1054C), only 4, 6, and 10 time traces were analyzed for each case. These few time traces cannot be representative enough. Therefore the data in Figure 3D (percentage of fluctuating molecules) has little statistical confidence.

We analyzed 103, 84 and 107 molecules of S867C-N1054C and found that 4, 10, and 6 molecules were fluctuating. The differences in the percentages of fluctuating molecules (4/103, 10/84, 6/107) are slightly less than the statistical standard (P = 0.08, Steel-Dwass test), as the referee notes. Therefore, we deleted the discussion regarding the comparison of the percentages of the fluctuated S867C-N1054C construct from the revised manuscript. Nevertheless, since this is not our main focus in the present manuscript, we strongly believe that this change will not affect the other arguments in the paper.

3. It is stressed in the manuscript that HNH enters a cleavage-competent state. However, no data supporting this is shown in the manuscript. Moreover, since HNH can enter this domain several times, it is unclear whether it does cleave or not and if the domain can move after cleavage. Design an assay which would relate the percentage of molecules accessing the „cleavage competent" conformation and the percentage of molecule that has been cleaved. This could be an assay that involves DNA immobilized on the surface, for example, an assay from Dagdas et al (partial duplex target) or Sternberg et al (Nature, 2014, denaturation after reaction).

We added the results of additional FRET measurements using several Mg^{2+} concentrations, as described in our response to the first remark. The results demonstrated that the cleavage rate of the target DNA highly correlates with the percentages of S355C-S867C molecules showing high FRET efficiency (Figure EV4), supporting the conclusion that the high FRET state corresponds to a cleavage-competent state.

4. Despite pre-incubating Cas9 with excess of nucleic acids, the observed FRET shift is an indirect indication of domain flexibility. Use a reporter dye on sgRNA/DNA molecules and performing real-time flow measurements to observe transitions between different flexible states when nucleic acids bind the immobilized Cas9. See also a minor comment about plasmid DNA.

Real-time flow measurement with dye-labeled sgRNA/DAN is a powerful assay system to determine the transition rate of the FRET shift. However, the transition rate is not our interest in the present manuscript, because it has been measured under various conditions previously (Sternberg et al., 2015, Nature). Our main interest in this study is the dynamic structural changes of Cas9 during the steady state with and without sgRNA/DNA. For that purpose, the biochemical data are sufficient to ensure the nucleic acid binding states of Cas9.

5. The author described that FRET value increased when sgRNA was added (Figure 2B, Apo vs +sgRNA). This „gradual increase" is not clearly noticed in Figure 2B.

We fitted the histograms with Gaussian curves, and measured the peak values. We added these data in the revised figure (Figure 2).

6. When DNA is added (Figure 2B, +sgRNA +DNA), the FRET value did increase. However, this change is not as drastic as that was observed by Dagdas et al (BioRxiv, 2017) who used the same protein mutant. In the case of Dagdas et al, the low FRET (E = 0.19) became very high FRET (E = 0.97). Please discuss the discrepancy.

The same construct (S355C-S867C) was used in a previous bulk FRET study (Sternberg et al., 2015, Nature). Sternberg et al. reported that the mean value of the FRET ratio of the construct increased to ~0.4 by the addition of on-target DNA. As the highest FRET ratio of the study is ~0.5, the FRET ratio of ~0.4 means that the mean FRET efficiency is not ~1.0, which is consistent with our data (Figure 2C, center panel). We mention the Sternberg paper because some of its authors are common with Dagdas et al., even though the discrepancy was unmentioned. Recently, Yang et al. posted a similar smFRET study on bioRxiv. The distribution of the FRET efficiency in the same construct was totally consistent with our data. Yang et al. pointed out the possibility that heparin causes the discrepancy with the data reported by Dagdas et al. We added a discussion about this in the revised manuscript (Pg. 16, Ln. 344-360).

7. Another discrepancy between this work and the work of Dagdas et al (and also Lim et al, „Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease" Nature Communications, 2016) is that the fluctuation is observed more when target DNA is present than absent (this work) while it is less observed when on-target is used than off-target is used (Dagdas et al, Lim et al). Is it because plasmid DNA was used as target? Please discuss the discrepancy.

The difference is most likely due to the different nucleic acid binding states of Cas9. As the referee commented, we did not use the off-target DNA in this study. The behavior of the Cas9 domains in the sgRNA-Cas9 complex (this work) and in the sgRNA/off-target DNA-Cas9 complex (Dagdas et al. and Lim et al.) should be different, due to the steric and ionic effects of the off-target DNA. Taking all three studies together, the order of the high fluctuation rate would be sgRNA/off-target DNA-Cas9 > sgRNA/on-target DNA-Cas9 > sgRNA-Cas9.

8. In Figure 3, the selection of static and fluctuating molecules seem arbitrary. The authors should provide a quantitative method in sorting the traces into two categories. For example, they said that they considered it is a fluctuating species if they observe more than two fluctuations. It has to be defined between which states the fluctuation occurred because any FRET traces will exhibit anti-correlated donor-acceptor fluctuation just like Figure 3A (static) trace shows.

In response to the comment, we re-sorted the traces with the standard deviation-based method and confirmed that the results were basically unchanged. See the materials and methods in the revised manuscript (Pg. 24, Ln. 534-537).

9. Trace in figure 3 c) shows both multi-state fluctuations and an interval where it is static. Would such a trace be assigned to static, fluctuations or both?

The molecules showing both the fluctuation and static phase were defined as fluctuating molecules. We made this point clear in the revised manuscript (Pg. 25, Ln. 537-538).

10. Figure 4B (+sgRNA+DNA -Mg) looks analogous to Figure 4A (+sgRNA+DNA+Mg), with only lightly lower high FRET peaks. It is unclear how these peaks disappear once you introduce magnesium in the case of static molecules.

The disappearance can be explained by the shift of molecules in the high FRET static state to the highly fluctuating state upon the Mg^{2+} binding. We added a comment relating this possibility in the result part (Pg. 12, Ln. 256-259)

11. In Figure 5C it is confusing what is meant by „duration" and how the numbers are determined. Were the numbers obtained from the kinetic rates that hidden Markov analysis generates? Or were the dwell times measured from idealized time traces

generated by hidden Markov analysis? The analysis procedure by the authors indicate it is the latter. The authors should then note that the dwell-time analysis (unlike hidden Markove analysis) cannot be correctly used for obtaining the kinetic rates. Taking an example of state A making a transition to either B or C, the dwelltime of A to B transition is non-intuitively identical to that of A to C. This is because states B and C directly compete with each other.

As the referee pointed out, the dwell times in Figure 5C were measured from idealized time traces generated by a hidden Markov analysis. To avoid the problem regarding competing processes, we discuss the duration of each HNH position instead of that of transition patterns in the revised manuscript (Fig 5C).

12. In Figure 5C, show dwelltime histograms for clarity.

The histograms are shown as Appendix Fig S2 in the revised manuscript.

13. The Results part of the manuscript requires revision. Speculation appears very often between results. Since the speculation is mostly over-interpretation of data, it makes it difficult to read.

In response to the comment, we rewrote the results part so it was clearer. We also changed the names of the HNH positions to correspond with those named in Dagdas et al., 2017. This change eases the comparison between the results in Dagdas et al. and our study, and thus will help the readers to understand our results.

Minor comments:

1. The authors used plasmid DNA as target DNA. Would bulky plasmid DNA properly interact with surface-tethered Cas9? Explain why plasmid DNA was used instead of oligo DNA unlike regular single-molecule FRET experiments.

Linearized plasmid DNA is generally employed for the bulk DNA cleavage assays, as shown in Figs EV1 and EV4. To elucidate the relationship between the cleavage rates in the bulk assays and the conformational changes, we also used the plasmid DNA for smFRET assays.

2. Plasmid DNA will have potentially many off-target sequences. Include the sequenc of the plasmid DNA and indicate the number of on-target and off-target sequences.

As shown in Appendix Fig S3 in the revised manuscript, our target DNA does not contain long (≥ 5-nt matching) off-target sequences, and contains one target sequence.

3. For the construct S867C-N1054C, the authors may observe fluctuation if they use a dye pair that has smaller Forster radius than Cy3 and Cy5.

We appreciate the suggestion. However, the relationship between the RuvC and HNH domains is not the main interest in the present manuscript. Therefore, we will try different dyes in our future studies.

4. It is unclear if the FRET histograms in Figures 2 and 4 are obtained from a snapshot analysis or from time traces.

We rewrote the figure legends to clarify that the histograms were obtained from time traces.

5. The authors should indicate whether/how they did donor leakage correction. Without this indication, it is difficult to know whether the lowest FRET peak in a FRET histogram is from a donor-only population or from a low-FRET population. Likewise, it is not clear whether the long static state in Figure 3C is only photoblinking of acceptor or not.

Immediately after the smFRET measurements, we confirmed that the observed Cas9 molecules were also labeled with Cy5 (acceptor) by the red laser illumination. Therefore, the lowest FRET peak in the histograms is from a low FRET population.

Regarding the second question, the dwell time of Cy5 photoblinking was much shorter than that of the low FRET static state under our assay conditions. Thus, we conclude that the long static state in Fig 3C is not a consequence of Cy5 photoblinking. We add comments in the methods section (Pg. 25, Ln. 519-527).

6. FRET histograms are hard to read, especially where the distribution is broad. Fit using gaussian peaks to determine what FRET states are present.

The FRET histograms were fitted with multi-Gaussian distributions in the revised manuscript (Fig 2).

7. In Figure 3D, using a legend to describe which colour means what would be easier to read.

We added the description in the figure legend.

8. Describe how PEG surface was prepared.

We added the description regarding the preparation of the PEG coated surface (Materials and Methods, Pg. 22, Ln. 466-482).

9. The manuscript lack accuracy in description. Vague description (e.g., the majority of , wide distribution, narrow distribution, ..) are too often used without quantification.

We changed the expressions in the examples with more quantitative ones. For the distribution of FRET histograms, we add descriptions of HWHM.

Referee #2:

The topic is timely and the approach well-suited to address important mechanistic questions, but I think the authors need to do a better job of explaining what was done previously, and what questions were left outstanding by the previous work.

The introduction part was rewritten to clarify the previous research and the remaining questions.

Ln 131: The authors state "The sgRNA/DNA-bound molecules in our assay should maintain the ternary complex of the sgRNA and cleaved target DNA, because previous studies have demonstrated that Cas9 cleaves the target DNA at a rate higher than 10 min-1 and retain binding with cleaved DNA (Sternberg et al, 2014; Jinek et al, 2012; Sternberg et al, 135 2015)." Did the cited papers show that both end of the cleaved DNA are retain equally or that the PAM end is preferentially retained? How is the state of cleavage and the impacts of DNA release accounted for in these experiments? How do we know that the dynamics observed for some particle are reporting on the release of DNA or part of the DNA target after cleavage?

The cited papers demonstrated that all fragments of the cleaved DNA remain in the Cas9 complex. In the papers, the DNA fragments were retrieved by denaturing the Cas9 complex with SDS-sample buffer or high-molar urea. Other studies observing the release of the DNA fragment used ingenious artificial DNA fragments, instead of intact double strand DNA. Dagdas et al., 2017, bioRxive, for instance, employed a short DNA, in which one end of the DNA strands exist as two single DNA strands in the DNA-Cas9 complex. In the intact long DNA, the cleaved strand was anchored to the complex through the double strand region with the counter strand. The cleaved fragment from the short DNA can be released from the complex, due to absence of the double strand region. In the present study, we used the long intact DNA. Thus, the sgRNA and cleaved DNA fragments in our assay should be retained in the ternary complex.

The authors state, "As a common property of the D435C-E945C and S355C-S867C constructs, sgRNA binding decreased the percentage (Figure 2B)," I think would be helpful to explicitly state what "percentage" you are refereeing to. Also, the "decrease" for S355C-S867C is not evident in Fig 2B. In fact, the percent frequency in the low FRET state appears to increase

modestly after sgRNA binding.

We apologize, as we had intended to cite Figure 3D, not Figure 2B. We have fixed the typo in our revised manuscript. The change will clarify what we are referring to there.

It is unclear how the authors know if sgRNA-Cas9 is bound to a molecule of DNA is cleaved or not. How does DNA cleavage change the dynamics? Clearly Mg has a large impact on the distribution of conformational states but does being in the "DNA-cleavage position" mean that the DNA is cleaved or just that it sampled this position?

As described in our response to comment #3 of Referee 1, the strong correlation between the DNA cleavage activity and the appearance of the highest FRET peak (Fig EV4) provides supporting evidence that the HNH domain in the D position corresponds to the DNA cleavage competent state. The fluctuating HNH domain repeatedly translocated to the D position, as shown in Fig 5A, suggesting that the HNH domain can be located at this position not only during but also after the DNA cleavage.

Ln 231: "in the absence of magnesium ions Cas9 can be trapped in the pre-cleavage state with sgRNA and DNA." Are there any other roles for the metal that might contribute to conformational state? Was this experiment ever performed using Mg-containing HNH and a non-cleavage substrate (e.g., Phosphorothioate Oligonucleotides)?

From previous structural and biochemical studies, it is widely accepted that Cas9 requires Mg^{2+} to cleave DNA. Other roles of the metal have not been confirmed. As described above, Dagdas et al., 2017, *Sci. Adv*. and Sternberg et al., 2015, *Nature* performed measurements with Mg^{2+} and the non-cleavable off-target DNA. The papers demonstrated that Cas9 shows different behaviors depending on the cleavability of the DNA substrate. In the present paper, we focus on the Cas9 behavior with the intact cleavable DNA. Therefore, we employed the magnesium control method to trap the Cas9 in the pre-cleavage state.

Minor concerns

a linear representation of the protein and its associated domains colored according to the structure would help orient the reader.

The linear representation of the Cas9 domains is added as Fig 1A in the revised manuscript.

References: The hyperlink to Zheng's modeling paper is incorrect. The link is to a paper on the SNARE complex.

We thank the referee for noticing the incorrect link. We deleated the link.

Pg6 Ln 106 "single molecular level". Consider changing this to "single molecule level"

We changed the expression of “single molecular level” to “single molecule level”.

Fig EV1. It appears that more of the DNA target is cleaved with lower (25nm) concentrations of Cas9. Is this correct or is the gel mislabeled?

We apology for the mislabeled concentrations. We corrected the labels in the revised manuscript.

Ln 221: Should "access" be "assess"?

Corrected (Pg. 11, Ln. 232 in the revised manuscript).

Referee #3:

1. Based on the results presented in Fig. 3D, the authors conclude that the flexibility of the hinge regions between the nuclease

and REC lobes increases in the DNA-bound complex based on the observation of an increased fraction of fluctuating molecules. However, this result could also be reached just by increased dynamics within the REC domain itself, especially given that the RuvC-REC and HNH-REC distances are monitored relative to different locations within the REC lobe (D435 vs. S355). It would be good if the authors could also comment on this possibility and also be more specific in indicating which hinge regions are likely to be involved.

We appreciate the suggestion and added a discussion about the possibility (Pg. 9, lines 191-195).

2. In this context, it is also important to consider the location of the covalently attached fluorophores in the Cas9 protein constructs. Inspecting the crystal structures of the various Cas9 functional states, it seems that Glu945 is located within a beta-hairpin loop that is likely to be flexible by itself in the sgRNA- and sgRNA/DNA-bound structures (it is held by crystal contacts in some structures and has very high B-factors in the others). Is it possible that some of the wide distribution of FRET efficiencies observed for the D435C-E945C constructs in the sgRNA- and sgRNA/DNA-bound states could be attributed to the intrinsic flexibility within this region?

The intrinsic flexibility of Glu945 might bring about the wide distribution of the FRET histograms. However, the B-factors of the crystal structure can be increased by fast (or slow) movements that are undetectable by our FRET measurements. Therefore, we cannot exclude other possibilities for the wide distribution.

3. The authors concede that the low frequency of fluctuating molecules observed for the S867C-N1054C construct might be due to the relatively short distance between the residues (p. 10) and hence it is not appropriate to make quantitative comparisons of the relative flexibilities of the three domains in the three constructs. Based on this, I find the conclusion that flexible domain movements facilitate Cas9 adopting the static conformation of the next nucleic acid binding state somewhat tenuous.

We explained in the manuscript that it is not appropriate to quantitatively compare the three constructs under the same conditions. However, it is appropriate to compare the flexibility of a single construct under different conditions. Under such circumstances, the conclusion was drawn from the comparison of the FRET histograms of each construct in the different nucleic acid binding states. Therefore, the conclusion is not compromised by the limitations of our assay system. However, during the revision process, we found the conclusion is a little speculative without detailed structural data of the fluctuating molecules. Thus, we rewrote the manuscript to only insist that the flexible domain movements allow Cas9 to adopt different conformations from the static ones (Pg. 11, Ln. 226-228).

4. In Fig. 4B, analysis of time trajectories indicates that transitions between middle ("pre-cleavage) and high (catalytically competent) FRET states are very infrequent. The authors conclude that the HNH domain rarely moves directly from the pre-cleavage to the cleavage state and needs to return to the undocked conformation first. I do not think that this is such a surprising result from a probabilistic standpoint, given the data. The distribution of FRET efficiencies in the DNA-bound state shows that the HNH domain spends most time in the undocked conformation and therefore, the undocked state is thermodynamically favoured one, As a result, the domain will always be more likely to move back to the undocked conformation. However, an alternative interpretation of the results could be that rather than being a bona fide on-pathway intermediate, the "pre-cleavage" mid-FRET state is in fact a dead-end, off-pathway state that needs to be resolved by collapse to the undocked conformation. The authors should comment on this possibility in the manuscript.

We thank the referee for the suggestion and added comments regarding this discussion to our manuscript (Pg. 14, Ln.

296-305).

Minor comments:

5. The authors reference Zheng 2017 as the only molecular dynamics study of the Cas9 system. However, there have now been several other computational studies of Cas9 (Zuo, Z., and Liu, J. (2016). Cas9-catalyzed DNA Cleavage Generates Staggered Ends: Evidence from Molecular Dynamics Simulations. Sci Rep 5, 37584; Palermo, G., et al. (2016). Striking Plasticity of CRISPR-Cas9 and Key Role of Non-target DNA, as Revealed by Molecular Simulations. ACS Cent Sci 2, 756-763), which should also be referred to.

We thank the referee for suggesting the computational studies of Cas9. We cited these studies in the revised manuscript.

6. On page 8, the authors state the fluorophores displayed a high degree of fluorescence anisotropy. Although this is further detailed in the Methods section, it would perhaps be better to show these results as a supplemental figure.

As suggested, the fluorescence anisotropy results are shown in Appendix Fig S1 in the revised manuscript.

7. On p. 11, the authors state that addition of Mg2+ ions enhances the fraction of fluctuating Cas9 molecules and refer to Fig. 3D. However, there is no mention of Mg2+ in Fig. 3D, nor in the figure legend. Do the black bars refer to data obtained for sgRNA+DNA-bound Cas9 in the presence of Mg2+? If not, the authors should add the Mg2+ data to the figure.

We described the values in the results part of the revised manuscript (Pg. 12, Ln. 257), and not in Fig 3D.

Editorial correspondence 14 December 2017

Thank you for submitting your manuscript to The EMBO Journal and my apologies again for the extended duration of the re-review process. Your study has now been seen by two of the original referees and their comments are included below. As you will see, the reviewers disagree in their recommendations for the revised manuscript and I have therefore also consulted with an additional expert advisor for the conclusiveness of the FRET data. The comments from the advisor are also included below.

As you will see from the reports, referee #3 is satisfied with the changes you have made in the revised version but ref #1 still raises strong concerns about the quality of the FRET data and restates that a much larger number of measurements is needed to support the conclusions drawn. Referee #1 also refers to competing papers published elsewhere compromising the novelty in your study but this is less of a concern from our side since EMBO Press offers scooping protecting for manuscripts under consideration here. However, since the concerns about the quality of the FRET data could undermine the overall conclusiveness of the study (and since the manuscript has already undergone one major round of revision), I consulted with an additional FRET expert (comments below). This person agrees that the level of noise seen in your system is unusual and that dye mobility effects could affect the measurements, leading to potential artefacts. However, the advisor also suggests a strategy to determine if that is indeed the case.

Since the conclusiveness of the data set remain open at this point I would like to discuss the comments made by our external advisor with you before I make the final decision.

I would therefore ask you to take a look at the reports and the advisor comments included below and let me know if you have data available for the fluorescent lifetimes of Cy3 and Cy5 in your setup - and if so, how this fits with the advisor's concerns. I would then take that into consideration - and possibly also discuss it with the advisor - before making a final decision on your study. You can send me the response directly by email and I will then get back to you with a decision shortly.

REFEREE REPORTS

Referee #1:

The authors carried out some additional experiments. However, the quality of the data remains very low. This makes it not possible to judge whether this work will be reproducible.

It is true that as few as 20 traces are sufficient for building FRET histograms. However, this is the case only if molecules show FRET fluctuations over well-defined FRET states. The doubly-labeled Cas9 system of the authors show extreme heterogeneity (as mentioned by the authors), which necessitates probing a substantially larger number of molecules than 20.

In addition to the low-quality histograms in the main figures, newly acquired data (EV4C) is another example that shows how unreliable their data interpretation is. The top panel (0.5mM MgCl2) in EV4C shows a broad distribution with several dubious peaks. The authors made biased interpretation of the histogram, drawing attention to only the two peaks they observed in Figure 4A. Either of the two peaks is even not observable in the middle panel (1.0mM MgCl2).

The claim authors make will be valid, only if they obtain FRET histograms from an order of magnitude larger number of molecules. Otherwise, this manuscript reports only the heterogeneity that arose from analysing a small set of molecules. The low quality could also be due to other artifacts such as the use of a plasmid, a non-ideal substrate for single-molecule measurements.

Furthermore, this work is not novel since a similar work with higher quality data and more quantitative has been published elsewhere. This reviewer does not support the publication of this work in EMBO Journal.

Referee #3:

The revisions that have been made have adequately addressed my points and I appreciate that the authors have taken a more cautious approach in the interpretation of their data. I also note that the clarity of the text has been improved, resulting in a much stronger manuscript. Although I concur with some of the issues raised by Reviewers #1 and #2, I think that the authors have made good efforts to address these concerns as well. Moreover, the study is in broad agreement with similar studies that have recently been published elsewhere or are available as preprints, which validates the authors' approach and interpretations. Although further experiments are clearly necessary to elucidate the very fine details of the molecular mechanism, these are, in my view, beyond the scope of the current manuscript and should be addressed in a future study. I do not have any further comments at this point.

FRET technical advisor: (in response to the revised manuscript and the referee reports)

The manuscript seems to be well put together on the whole and it's an interesting story overall, but I have to say I have not seen a report stating that the donor fluorophore is so anisotropic. The authors report a Cy3 anisotropy of 0.34-0.41, where 0.4 has theoretically not tumbling. On page 8 the authors report that this simply means that the changes in FRET reflect both distance and orientation changes, which may indeed be correct, but it certainly raises concerns about the interpretation of the data.
Cy3 and Cy5 (Cy3 in particular) can have incredibly short fluorescence lifetimes, which will severely skew anisotropy measurements such that the dyes appear immobile. For instance, an anisotropy measurement asks: if excited with perfectly polarized light, how much light do I detect in the emission channel as being depolarized from dye tumbling. Dyes tumbling normally occurs on the 1-10 ns time scale depending on environment. However, the fluorescent lifetimes of Cy3 and Cy5 can be faster than the tumbling time. In fact, the Cy3 lifetime reported in the literature is 300 picoseconds and can be significantly faster. Hence, almost no dye tumbling occurs prior to emission. Correspondingly, the dyes appear anisotropic but are not.

The authors should report Cy3 and Cy5 fluorescent lifetimes to address this point. I did not see such measurements. The authors should also define what TSY is, which is listed as being included in all imaging conditions at a concentration of 2.5 mM. Was this included in their activity measurements? It should certainly be included in their anisotropy and fluorescence lifetime measurements. Such reagents are typically added to improve fluorophore performance and can strongly affect fluorophore photophysics (i.e. fluorescence lifetimes).

Although the methods employed are clearly not state-of-the-art and the fluctuations in FRET may reflect changes in dye mobility (ie. orientation not distance- as discussed above), the authors appear to have done the appropriate controls to show that whatever dynamics they are measuring somehow report on Crispr-Cas9 activity. This will have to be decided by the biological reviewers.

2nd Editorial decision — 22 December 2017

Thank you for submitting a revised version of your manuscript and my apologies for the unusual duration of the re-review time. Your study has now been seen by two of the original referees whose comments are shown below.

As you will see, while ref #3 is supportive of the work, ref #1 raises both conceptual and technical concerns and consequently does not recommend publication in The EMBO Journal. From the conceptual side, your manuscript falls under the EMBO Press scooping protection policy so we are not concerned with the fact that related papers have been published elsewhere while your study was in revision. However, the technical points need to be clarified and as I told you in a recent email we have therefore consulted with an external expert advisor whose comments are included below along with the referee reports.

I have now also consulted with the advisor once more based on the reply to the concerns that you

sent me via email and the conclusion is that this person is satisfied with your explanation and the measurements of dye lifetime and recommends publication of the revised manuscript in The EMBO Journal. I would therefore like to invite you to submit a final version of the manuscript in which you incorporate your response (via text changes) to the concerns from ref #1 and the external advisor

2nd revision – authors' response 7 February 2017

Point-by-point responses to the comments from the referees and the technical advisor.

Referee #1

The authors carried out some additional experiments. However, the quality of the data remains very low. This makes it not possible to judge whether this work will be reproducible.
It is true that as few as 20 traces are sufficient for building FRET histograms. However, this is the case only if molecules show FRET fluctuations over well-defined FRET states. The doubly-labeled Cas9 system of the authors show extreme heterogeneity (as mentioned by the authors), which necessitates probing a substantially larger number of molecules than 20.

The sample numbers of our data fulfilled statistical criteria. In Fig 2B, for instance, FRET histograms of Apo, +sgRNA and +sgRNA +DNA conditions showed statistically significant differences ($P < 0.05$, Mann-Whitney test) with adequate statistical power ($1\text{-}b > 0.8$). To investigate the possibility that a small number of Cas9 molecules lowered the reliability of the histograms, we performed a random (n-2) resampling test with 1,000 permutations on the FRET traces, but found no statistically significant difference for any permutations (threshold: $P = 0.05$, Kolmogorov-Smirnov test). In addition, we measured approximately 100-200 molecules for each condition (Fig 2). Therefore, we conclude that differences in the histograms are not artifacts but truly reflect the conformational differences of the Cas9 molecules. We added this description in the Method section (Page 25, line 560- Page 26, line 565).

In addition to the low-quality histograms in the main figures, newly acquired data (EV4C) is another example that shows how unreliable their data interpretation is. The top panel (0.5mM MgCl2) in EV4C shows a broad distribution with several dubious peaks. The authors made biased interpretation of the histogram, drawing attention to only the two peaks they observed in Figure 4A. Either of the two peaks is even not observable in the middle panel (1.0mM MgCl2).

We gave attention to the two peaks because Silverman's multimodality test (Silverman, *Journal of the Royal Statistical Society*, 1981) showed all histograms in EV4C have two peaks. Indeed, two-peak Gaussian distribution models fit the data well: $c^2 < 0.01$. We have added fitting curves in Figure EV4C and descriptions in the figure legend in the revised manuscript.
As shown in EV4A and B, Cas9 did not cleave the target DNA under low Mg^{2+} conditions (0.5 or 1.0 mM $MgCl_2$). Thus, the lack of the peak corresponding to the D position in the top and middle panels of EV4C strongly supports the idea that the Cas9 showing this FRET peak is responsible for the DNA cleavage process.

The claim authors make will be valid, only if they obtain FRET histograms from an order of magnitude larger number of molecules. Otherwise, this manuscript reports only the heterogeneity that arose from analysing a small set of molecules. The low quality could also be due to other artifacts such as the use of a plasmid, a non-ideal substrate for single-molecule measurements.
The referee considers a broad distribution in the FRET histogram as a sign of low quality data. However, even with exceptionally high quality data, flexible movements of the domains could lead to a broad FRET distribution. As we mentioned above, our analyses fulfilled statistical criteria. Furthermore, a significant ratio of Cas9 molecules showed fluctuation. Therefore, we consider the broad FRET distribution was not due to heterogeneity in a small set of molecules.

Furthermore, this work is not novel since a similar work with higher quality data and more quantitative has been published elsewhere. This reviewer does not support the publication of this work in EMBO Journal.

We cited and discuss results of similar works posted on a preprint server and published elsewhere (Page 16, line 343- Page 17, line 364). Because these works were posted and published after submission of our manuscript to the EMBO Journal, we believe they do not undermine the novelty of our present study according to the EMBO journal policy.

Referee #3:
The revisions that have been made have adequately addressed my points and I appreciate that the authors have taken a more cautious approach in the interpretation of their data. I also note that the clarity of the text has been improved, resulting in a much stronger manuscript. Although I concur with some of the issues raised by Reviewers #1 and #2, I think that the authors have made good efforts to address these concerns as well. Moreover, the study is in broad agreement with similar studies that have recently been published elsewhere or are available as preprints, which validates the authors' approach and interpretations. Although further experiments are clearly necessary to elucidate the very fine details of the molecular mechanism, these are, in my view, beyond the scope of the current manuscript and should be addressed in a future study. I do not have any further comments at this point.

We appreciate the positive comments from the referee.

FRET technical advisor: (in response to the revised manuscript and the referee reports)
The manuscript seems to be well put together on the whole and it's an interesting story overall, but I have to say I have not seen a report stating that the donor fluorophore is so anisotropic. The authors report a Cy3 anisotropy of 0.34-0.41, where 0.4 has theoretically not tumbling. On page 8 the authors report that this simply means that the changes in FRET reflect both distance and orientation changes, which may indeed be correct, but it certainly raises concerns about the interpretation of the data.
Cy3 and Cy5 (Cy3 in particular) can have incredibly short fluorescence lifetimes, which will severely skew anisotropy measurements such that the dyes appear immobile. For instance, an anisotropy measurement asks: if excited with perfectly polarized light, how much light do I detect in the emission channel as being depolarized from dye tumbling. Dyes tumbling normally occurs on the 1-10 ns time scale depending on environment. However, the fluorescent lifetimes of Cy3 and Cy5 can be faster than the tumbling time. In fact, the Cy3 lifetime reported in the literature is 300 picoseconds and can be significantly faster. Hence, almost no dye tumbling occurs prior to emission. Correspondingly, the dyes appear anisotropic but are not.
The authors should report Cy3 and Cy5 fluorescent lifetimes to address this point. I did not see such measurements.

To answer the concern of the advisor, we measured the fluorescence lifetimes of Cy3 and added the data in Appendix Fig 1D-G. The mean fluorescence lifetimes of Cy3 on the Cas9 construct were 0.71-0.78 ns. Because the lifetimes were significantly shorter than the normal dye tumbling time, we agree with the technical advisor that the short fluorescence lifetime of Cy3 is most likely the cause of the apparent high values of anisotropy. Yet, we cannot exclude the possibility that the orientation of the dye affects FRET efficiency. We added this discussion to the manuscript (Page 8, lines 164-167 and Appendix Fig S1).

The authors should also define what TSY is, which is listed as being included in all imaging conditions at a concentration of 2.5 mM. Was this included in their activity measurements? It should certainly be included in their anisotropy and fluorescence lifetime measurements.

TSY is the trade name of a commercially available triplet-state quencher (Pacific Biosciences, CA, USA). As the advisor pointed out, TSY was included also in our anisotropy and fluorescence lifetime measurements. We made this point clear in the Method section in the revised version of our manuscript (Page 20-24).

Although the methods employed are clearly not state-of-the-art and the fluctuations in FRET may reflect changes in dye mobility (ie. orientation not distance- as discussed above), the authors appear to have done the appropriate controls to show that whatever dynamics they are measuring somehow report on Crispr-Cas9 activity.

We appreciate that the advisor has highly valued our control experiments.

Accepted 13 Febuary 2018

Thank you for submitting the final revision of your manuscript, I am pleased to inform you that your study has now been officially accepted for publication in The EMBO Journal.

© European Molecular Biology Organization

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Tomohiro Shima and Osamu Nureki
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2017-96941R

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title
http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tum
http://datadryad.org
http://figshare.com
http://www.ncbi.nlm.nih.gov/gap
http://www.ebi.ac.uk/ega
http://biomodels.net/
http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

Question	Answer
1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	In a priori analyses, we employed 0.2 for the value of Cohen's d to determine the sample sizes. In post hoc analyses, we ensured that statistical power (1-beta) was over 0.8.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	We repeated the experiments at least three times for each condition using individually prepared samples.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	We double checked the FRET trajectries by three investigators.
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	For multi-peak Gaussian fittings (Figure 2), the normality of each peak was confirmed by Kolmogorov-Smirnov test. For Figure 3D, we used a non-parametric test (Steel-Dwass) which does not require the specific distribution of the data.

Is there an estimate of variation within each group of data?	NA
Is the variance similar between the groups that are being statistically compared?	Yes

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The source data of smFRET in this study were uploaded with the revised manuscript.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	The source data of smFRET in this study were uploaded with the revised manuscript.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA