THE
EMBO
JOURNAL

Manuscript EMBO-2010-76453

Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system

Tomas Sinkunas, Giedrius gasiunas, Christophe Fremaux, Rodolphe Barrangou, Philippe Horvath and Virginijus Siksnys

Corresponding author: Virginijus Siksnys, Institute of Biotechnology, Vilnius University

Review timeline:

Submission date:	06 November 2010
Editorial Decision:	18 December 2010
Revision received:	19 January 2011
Editorial Decision:	01 February 2011
Accepted:	01 February 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 18 December 2010

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has been now been evaluated by three referees and I enclose their reports below. As you will see the referees find the biochemical characterization of Cas3 to be interesting and potentially important, but require some further experimental analysis to make the manuscript suitable for The EMBO Journal.

After discussing the reports with the referees the main issues that arise are the further characterization of substrate specificity of Cas3, including testing the proposed model if Cas3 can bind and cleave R-loop structures. Referee #3 also suggests in vivo data to support the importance of the characterized Cas3 activities in CRISPR mediated immunity. Upon reflection and discussions with the referees while these are clearly nice experiments that would strengthen the study, I believe that to establish such experiments is beyond the scope of the current study and therefore is not required for publication. Therefore, should you be able to add the proposed in vitro data we would be happy to consider a revised version of the manuscript for publication in The EMBO Journal.

I should remind you that it is EMBO Journal policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript. When you submit a revised version to the EMBO Journal, please make sure you upload a letter of response to the referees' comments. Please note that when preparing your letter of response to the referees' comments that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process initiative, please visit our website: http://www.nature.com/emboj/about/process.html

© European Molecular Biology Organization

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

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1 (Remarks to the Author):

Prior genetic evidence implicates the Cas3 protein as being critical for mediating crRNA-mediated interference of phage. Furthermore, in silico predictions strongly implicate Cas3 as harboring nuclease (HD) and helicase/ATPase (DEAD box super family) domains. Here, Sinkunas have performed key biochemical characterization of the Cas3 protein from S. thermophilus and convincingly demonstrated that Cas3 is a nuclease, ATPase and helicase. Both the ATPase and 3' to 5' helicase activities of Cas3 are stimulated by ssDNA and the nuclease activity is also specific for ssDNA. The specific domains of Cas3 responsible for each activity were elucidated by analyzing site-specific Cas3 mutants harboring altered residues in predicted key residues. The deduced biochemical activities of Cas3 led the authors to propose a model for how Cas3 may act in crRNA-guided target DNA destruction. This work is convincing, the main conclusions are well supported by the data, and overall the findings significantly advance the knowledge of mechanism of CRISPR-Cas systems that employ Cas3 as an effector nuclease.

Some minor points:

1. An SDS gel of the purified Cas3 protein (and mutants used in this study) should be shown (supplemental data would be fine). Since others in the field have had a difficult time expressing Cas3 proteins from other organisms that are soluble and intact, this data is a must for publication. Furthermore, there is considerable 'hang-up' of DNA substrate in the wells of the gel shown for nuclease deficient D77A and D227A Cas3 mutants (Fig 3B). The authors should comment on how typical these results are and importantly if they relate to a tendency of these particular mutants to aggregate relative to the wt cas3 protein (this is of particular concern since the lack of activity of these mutants might relate to their tendency to misfold and aggregate rather than the claimed inability to perform the chemistry due to inactivation of active site residues).

2. The abbreviation PAAG (for polyacrylamide gel) used throughout is non-standard and confusing. Indeed the more conventional standard of PAGE (for polyacrylamide gel electrophoresis is also used by the author when describing SDS-PAGE). It is suggested that PAAG be replaced by the abbreviation: PAGE or more simply not using an abbreviation (i.e. use "polyacrylamide gel").

3. The layout of figure 4 is a bit confusing since Figure 4A and 4B have upper and lower parts but these are not well described in the figure legend. I suggest that the four parts be labeled Figure 4A-D (e.g. 4B would refer to lower panel in current 4A and the current upper panels of 4B be relabeled as 4C and D). The figure legend and results section should be modified accordingly. The figure would also be easier to follow if Figure A and B (i.e. upper and lower panels of A) would have the term "DNA" (upper) or "RNA" (lower) indicated above the oligo that is depicted on the right-most side of each panel.

4. In general, the discussion is highly redundant/repetitive with the results section and would benefit from significant condensation/synthesis of ideas to more efficiently bring out the significance of the work and how the results compare to those of others.

5. While not critical for publications, do the authors know if closed ssDNA is critical for nuclease activity of Cas3 or would linear ssDNA also serve as nuclease substrates? The authors should note the result if they have tested for cleavage of linear DNA substrates. Furthermore, is single-stranded RNA cleaved by Cas3?

6. On top of page 12, "Novagen" is misspelled (i.e. an A rather than an O in the 4th position).

7. A detailed figure legend should be written for Figure 5. Also, it is unclear why the cartoon depicts a single cut of DNA rather multiple cuts as directly supporteded by data in this study, (The Garneau et al. 21010 paper cited did observe site-specific cleavage of target DNA in vivo but this was presumably not catalyzed by Cas3 but rather by a Cas protein associated with the Nmeni subtype as proposed). The Figure 5 model (and discussion) should be altered to fit the data (Cas3 cleaves ssDNA at multiple sites).

Referee #2 (Remarks to the Author):

This manuscript describes the purification and biochemical analysis of Cas3, a protein essential to CRISPR-based bacteriophage immunity in bacteria. In E. coli, Cas3 was shown in genetic experiments to be necessary, though not sufficient, for protection against phage infection in the presence of an active CRISPR-Cas system. However, the molecular functions of Cas3 have remained undetermined until now. Here, an E. coli-based purification of the Cas3 protein from S. thermophilus was used to produce copious amounts of protein (~12 mg/L!) that were used to test ATPase, helicase and nuclease functions of the enzyme. ATP hydrolytic assays show that Cas3 has slow but specific ATPase activity, which could be abolished by point mutations in the putative ATP binding site of the ATPase domain. Furthermore, this activity was stimulated by single-stranded, but not double-stranded, DNA. The enzyme was also found to be a single-stranded DNase and to have the ability to dissociate short DNA strands annealed to a template in a 3'-to-5' direction. The data support the conclusion that Cas3 is some sort of an unwindase, as was previously proposed based on sequence homology to known helicase/ATPase proteins. However, the role of Cas3 in CRISPR-based immunity remains speculative. This work will be of interest to the field, but some additional experiments should be included to strengthen the conclusions of the work.

1. The authors propose that Cas3 functions in the targeted degradation of DNA molecules following their partial hybridization to a short CRISPR-derived RNA. The paper would be a much stronger contribution to the field if this model were tested directly to determine whether an "R-loop" can be recognized and cleaved in the single-stranded DNA region by Cas3, as modeled in Figure 5.

2. p. 5: The Coomassie-stained protein gel for Cas3 should be shown as a supplemental figure, since the purity of the protein is critical to the evaluation of the subsequent experiments using this prep. Was there evidence of contaminating nucleic acid in the prep, or was that the reason for using the heparin column for further purification?

3. In Fig. 1, domain abbreviations should be defined in the legend.

4. Fig. 3, nuclease assays: it would be helpful to provide timecourse rather than single time point analyses of the proteins tested, so that rates could be compared more meaningfully. Are the products of the reaction ultimately degraded to single nucleotides? IS the initial cleavage event fast, followed by slow continued degradation of the substrate? And how does the rate of cleavage catalyzed by the wild type enzyme compare to DNase I or similar enzymes? Cas3 appears to be quite slow, and perhaps that is relevant to its function in vivo, if its role is primarily to initiate rather than complete the degradation of target sequences.

Referee #3 (Remarks to the Author):

This manuscript reports the initial biochemical characterization of a Cas3 protein (containing DExH helicase and HD phosphohydrolase domains) from one of the CRISPR/cas loci of Streptococcus thermophilus strain DGCC7710. This is an important protein to study because Cas3 has been implicated as a critical player in the interference phase of CRISPR immunity in E. coli K12. Previous results with a different Cas3 ortholog that lacks an apparent helicase domain indicate that the HD domain has dsRNA- and dsDNA-specific nuclease activity. Beyond this almost nothing is know about this protein, despite its importance in a process of intense recent interest. Sinkunas et al. show that the purified recombinant protein has ATPase activity that is stimulated by ssDNA and that

requires an intact helicase domain, and they characterize some of the kinetic/thermodynamic features and cofactor requirements of this activity. They also define a nuclease activity that depends upon conserved residues in the HD domain, and surprisingly they show that this activity is specific for ssDNA. Finally they demonstrate that the protein has an ATP hydrolysis-dependent DNA-RNA and DNA-DNA unwinding activity that requires an intact helicase domain. This activity requires an adjacent 3'-overhanging single-stranded DNA region.

This manuscript presents some important new results that will be of significant interest to the field. However, there are two primary issues.

1. In the title and the abstract, the authors indicate that they have demonstrated that Cas3 is a translocase. Although translocation is certainly possibility and perhaps even likely, the claim is an overinterpretation. Some helicases unwind duplexes without translocation (e.g., Yang & Jankowsky, NSMB 13, p. 4181, 2006), so the demonstration of unwinding activity - even one that depends on an adjacent single-stranded region of specific polarity - is not sufficient to conclude that the protein has translocase activity. Translocation needs to be assayed directly. Otherwise this claim should be removed, though it would be perfectly appropriate to raise it as a possibility in the discussion.

2. The experiments are generally nicely done as far as they go, but the paper is perhaps a bit thin relatively to what we usually see in EMBO. Some obvious experiments that would make the work much more complete and definitive include the following.
a. In light of the previous report of dsRNase activity for the Cas3 protein from Sulfolobus solfataricus, RNA should be tested as a nuclease substrate in this study. RNA/RNA duplexes should also be tested as unwinding substrates.
b. Does hydrolysable ATP enable effective dsDNA digestion in vitro by first unwinding it?
c. Given the well-documented ability to disrupt genes in this strain, the authors should test the requirement for the activities they define in vivo by knocking out cas3, complementing the knockout with plasmid-borne cas3 (wild-type and representative mutants that block particular biochemical activities in vitro), and testing whether the CRISPR4 spacers can still generate crRNAs and direct interference. Otherwise the significance of their biochemical findings to the actual interference process is not established.

One additional difficulty: Along the lines of 2b above, why is the ssDNase activity documented in Figure 3 not apparent with the products of the unwinding assays shown in Figure 4?

1st Revision - authors' response 19 January 2011

Response to Reviewers' Comments

Referee #1

Q1: *An SDS gel of the purified Cas3 protein (and mutants used in this study) should be shown (supplemental data would be fine). Since others in the field have had a difficult time expressing Cas3 proteins from other organisms that are soluble and intact, this data is a must for publication.*

R1: We now provide an SDS gel in the Supplemental data section.

Q2: *Furthermore, there is considerable 'hang-up' of DNA substrate in the wells of the gel shown for nuclease deficient D77A and D227A Cas3 mutants (Fig 3B). The authors should comment on how typical these results are and importantly if they relate to a tendency of these particular mutants to aggregate relative to the wt cas3 protein (this is of particular concern since the lack of activity of these mutants might relate to their tendency to misfold and aggregate rather than the claimed inability to perform the chemistry due to inactivation of active site residues).*

R2: The **c**onsiderable 'hang-up' of DNA substrate in the wells is present both for the nuclease deficient D77A and D227A mutants and for the wt Cas3 protein in the presence of EDTA. Most likely, the Cas3 protein strongly binds to ssDNA molecules, and therefore DNA hardly enters the gel. By using a phenol/chloroform mix instead of EDTA to stop the reaction we obtained gels

without 'hang-up' of DNA substrate. These gels are included in the revised version. Since D77A and D227A mutants retain the ATPase activity similar to the wt protein it is unlikely that mutant proteins are aggregated.

Q3: *The abbreviation PAAG (for polyacrylamide gel) used throughout is non-standard and confusing. Indeed the more conventional standard of PAGE (for polyacrylamide gel electrophoresis is also used by the author when describing SDS-PAGE). It is suggested that PAAG be replaced by the abbreviation: PAGE or more simply not using an abbreviation (i.e. use "polyacrylamide gel").*

R3: The text was modified accordingly.

Q4: *The layout of figure 4 is a bit confusing since Figure 4A and 4B have upper and lower parts but these are not well described in the figure legend. I suggest that the four parts be labelled Figure 4A-D (e.g. 4B would refer to lower panel in current 4A and the current upper panels of 4B be relabelled as 4C and D). The figure legend and results section should be modified accordingly. The figure would also be easier to follow if Figure A and B (i.e. upper and lower panels of A) would have the term "DNA" (upper) or "RNA" (lower) indicated above the oligo that is depicted on the right-most side of each panel.*

R4: Parts of Figure 4 were labelled A-D as suggested and corresponding modifications included into the Figure 4 legend. Also, “DNA” and ‘RNA” are indicated above the oligo that is depicted on the right-most side of each panel.

Q5: *In general, the discussion is highly redundant/repetitive with the results section and would benefit from significant condensation/synthesis of ideas to more efficiently bring out the significance of the work and how the results compare to those of others.*

R5: The Discussion has been condensed, as to remove the overlap with the results section.

Q6: *While not critical for publications, do the authors know if closed ssDNA is critical for nuclease activity of Cas3 or would linear ssDNA also serve as nuclease substrates? The authors should note the result if they have tested for cleavage of linear DNA substrates. Furthermore, is single-stranded RNA cleaved by Cas3?*

R6: The problem with such experiments is that our preparation of Cas3 protein also shows 3’-exonucleolytic activity on the oligonucleotide substrates alongside to the endonucleolytic activity. In the Cas3 mutants D77A and D227A endonucleolytic activity is compromised but 3’- exonuclease activity is retained. While mutational analysis clearly shows that the endonuclease activity belongs to the Cas3 protein, it is not yet clear whether the 3’-exonuclease activity is due to the contamination or an inherent activity of Cas3 protein not related to the HD-domain. We are currently exploring this avenue.

Q7: *On top of page 12, "Novagen" is misspelled (i.e. an A rather than an O in the 4th position).*

R7: This was edited accordingly.

Q8*: A detailed figure legend should be written for Figure 5. Also, it is unclear why the cartoon depicts a single cut of DNA rather multiple cuts as directly supported by data in this study, (The Garneau et al. 21010 paper cited did observe site-specific cleavage of target DNA in vivo but this was presumably not catalyzed by Cas3 but rather by a Cas protein associated with the Nmeni subtype as proposed). The Figure 5 model (and discussion) should be altered to fit the data (Cas3 cleaves ssDNA at multiple sites).*

R8: This figure was revised according to the reviewer’s suggestion. Multiple cut sites are now shown in Figure 5.

Referee #2 :

Q1: *The authors propose that Cas3 functions in the targeted degradation of DNA molecules following their partial hybridization to a short CRISPR-derived RNA. The paper would be a much stronger contribution to the field if this model were tested directly to determine whether an "R-loop" can be recognized and cleaved in the single-stranded DNA region by Cas3, as modelled in Figure 5.*

R1: We agree that demonstration of the Cas3 binding and cleavage of R-loop structures *in vitro* would be a next logical step. However, there are several issues which currently prevent us from carrying these experiments in our system. First, the crRNA sequence, which is required in order to assemble *in vitro* a correct R-loop substrate, has not been determined. Second, our equivalent of the *E.coli* ternary Cas3-Cascade-crRNA complex is not available (Brouns et al., 2008). Third, the CRISPR4 protospacer adjacent motif (PAM) sequence, which is likely to be important for the assembly of the specific Cas3-Cascade-crRNA complex with foreign DNA, is yet to be determined. Lastly, although this is an interesting step towards understanding the mechanism of action of the CRISPR4/Cas system, we believe this is beyond the immediate scope of the results and discussion presented in the paper.

Q2: *p. 5. The Coomassie-stained protein gel for Cas3 should be shown as a supplemental figure, since the purity of the protein is critical to the evaluation of the subsequent experiments using this prep.*

R2: The paper was modified accordingly, and an SDS gel is now provided in the Supplemental data section.

Q3: *Was there evidence of contaminating nucleic acid in the prep, or was that the reason for using the heparin column for further purification?*

R3: We have no evidence of contaminating nucleic acid in the Cas3 prep. We used heparin to further purify the material and remove imidazol, which stays in the prep after purification by Ni-chelating column.

Q4: *In Fig. 1, domain abbreviations should be defined in the legend.*

R4: The legend was modified accordingly.

Q5: *Fig. 3, nuclease assays: it would be helpful to provide timecourse rather than single time point analyses of the proteins tested, so that rates could be compared more meaningfully.*

R5: Time courses are now provided in the Supplemental data section, however data are still difficult to quantify, as to obtain cleavage rates.

Q6: *Are the products of the reaction ultimately degraded to single nucleotides? IS the initial cleavage event fast, followed by slow continued degradation of the substrate? And how does the rate of cleavage catalyzed by the wild type enzyme compare to DNase I or similar enzymes? Cas3 appears to be quite slow, and perhaps that is relevant to its function in vivo, if its role is primarily to initiate rather than complete the degradation of target sequences.*

R6: The data presented in the paper does not allow us to provide answers to those questions.

Referee #3:

Q1: *In the title and the abstract, the authors indicate that they have demonstrated that Cas3 is a translocase. Although translocation is certainly possibility and perhaps even likely, the claim is an overinterpretation. Some helicases unwind duplexes without translocation (e.g., Yang & Jankowsky, NSMB 13, p. 4181, 2006), so the demonstration of unwinding activity - even one that depends on an adjacent single-stranded region of specific polarity - is not sufficient to conclude that the protein has translocase activity. Translocation needs to be assayed directly. Otherwise this claim should be removed, though it would be perfectly appropriate to raise it as a possibility in the discussion.*

R1: We agree with the reviewer and now refer to helicase activity accordingly.

Q2: *The experiments are generally nicely done as far as they go, but the paper is perhaps a bit thin relatively to what we usually see in EMBO. Some obvious experiments that would make the work much more complete and definitive include the following. In light of the previous report of dsRNase activity for the Cas3 protein from Sulfolobus solfataricus, RNA should be tested as a nuclease substrate in this study. RNA/RNA duplexes should also be tested as unwinding substrates.*

R2: Cas3 shows only a trace activity on the RNA substrates. Of note is that D77A and D227A mutants shows the same level of activity on the RNA substrate. Therefore, it is not yet clear whether the RNAse activity is due to the contamination or an inherent activity of Cas3 protein not related to the HD-domain. We are currently exploring this avenue.

Q3: *Does hydrolysable ATP enable effective dsDNA digestion in vitro by first unwinding it?*

R3: Cas3 does not degrade plasmid dsDNA substrate in the presence of ATP.

Q4: *Given the well-documented ability to disrupt genes in this strain, the authors should test the requirement for the activities they define in vivo by knocking out cas3, complementing the knockout with plasmid-borne cas3 (wild-type and representative mutants that block particular biochemical activities in vitro), and testing whether the CRISPR4 spacers can still generate crRNAs and direct interference. Otherwise the significance of their biochemical findings to the actual interference process is not established.*

R4: We agree with the reviewer that investigating crRNAs for this system would be interesting, but it is beyond the scope of this manuscript. Indeed, such crRNAs are yet to be determined for any of the functional CRISPR/Cas systems in this model organism.

Q5: *One additional difficulty: Along the lines of 2b above, why is the ssDNase activity documented in Figure 3 not apparent with the products of the unwinding assays shown in Figure 4?*

R5: ssDNA nuclease activity is not apparent in the unwinding assay due to the differences in experimental conditions. We used 10 mM $MgCl_2$ for ssDNAse assay and 1 mM $MgCl_2$ in the unwinding assay. In the presence of 1 mM $MgCl_2$, Cas3 degrades ssDNA more slowly. Moreover, if under these conditions a few cuts are introduced in the M13 DNA, reaction products maybe undistinguishable from the substrate due to the limited resolution in the 8 % polyacrylamide gel.

2nd Editorial Decision 01 February 2011

I have received the final report from the two original referees who have reviewed your revised manuscript for The EMBO Journal. I am happy to say that both referees find the study important and I am able to accept publication of the manuscript in the journal. You will receive the official acceptance letter in the next day or so.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1
The authors have satisfied all concerns and this work is now suitable for publication.

Referee #2

The authors have responded to most of the points raised in the original reviews.

© European Molecular Biology Organization

Although it would be nice if data could be included regarding the function of Cas3 in R-loop recognition/cleavage, the authors explain why this is beyond the technical capabilities of the current experimental system. Nonetheless, the study provides important insight into the function of a previously mysterious protein whose activity lies at the heart of CRISPR biology. I support publication of the revised manuscript in EMBO J.

© European Molecular Biology Organization