Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Olga Musharova

PhD Program: Life Sciences

Title of Thesis: Investigation of DNA-binding specificity of Cas1-Cas2 CRISPR adaptation complex in *E.coli.*

Supervisor: Professor Konstantin Severinov

Chair of PhD defense Jury: Professor Philipp Khaitovich **Email:** *p.khaitovich@skoltech.ru*

The thesis document includes the following changes in answer to the external review process

Reviewer Comment 1) The quality of figures (resolution, blurring) is low. Copy-paste as "independent rastr" could help.
Author: Corrected.

RW Comment 2) Fig 5 – unclear what is the fate of Cas2 and where it is on the gels of Fig. 5A,B. Should be 11 kD but not indicated on Fig. 5a (why?), and not seen any signs of changes around 11kD in Fig. 5b – why? "Antibodies specific for Cas2 could have been lost/depleted" during affinity column purification – may be. But why do you think that Cas2 was not lost earlier? The problem is that, though there is a peak at 11kD in fractions 6 and 7, it is present at very low concentration compared to Cas1 (<5%?). Probably the complex Cas1-Cas2 is very unstable? Could it be stabilized in presence of spacers or protospacers?
Author: The excess of Cas1 protein in Fig.5A gel is partially related to the stoichiometry of Cas1-Cas2 adaption complex (4 Cas1 monomers per 2 Cas2 monomers) and 3 time difference in molecular weights of Cas 1 and Cas2, partially it can be explained by some distortion of adaptation complex at conditions of overexpression used for complex purification. Although DNA fragments can stabilized adaptation complex in theory we did not check it. As a result we were able to obtained antibodies predominantly recognizing Cas1 protein, such antibodies were purified and used for Western blot shown in Fig.5B (Cas2 could not be detected on this blot).

RW Comment 3) ""The Cas1-Cas2" adaptation complex *E. coli* was purified and used to obtain a specific antibody.". – it is not clear if the Cas2 was really there, may be in 1-10% it really was a heterodimer.
Author: We indeed used the material purified from *E. coli* cells co-overexpressing the *cas1* and *cas2* genes for immunization. Purification was done through affinity chromatography via an N - terminal hexahistidine tag positioned on Cas1. A visual inspection of the purified material indicated that there was an excess of Cas1 compared to Cas2, which at least in part shall be due to the stoichiometry of the complex (4 Cas1 monomers per 2 Cas2 monomers).

RW Comment 4) Also, here and everywhere in the text the polyclonal antibody mix obtained from the rat serum is called "a specific antibody", which is a bit misleading, sounds as if a monoclonal recombinant antibody was generated, recognizing Cas1-Cas2 complex. Would be more correct to call it everywhere "to obtain anti-Cas1 polyclonal antibody" – as it was done in Fig. 5.
Author: Agree. Corrected.

RW Comment 5) "Reactivity of the antibody (1:5000 dilution) on a Western blot against proteins from whole cell extracts of KD263 E. coli cells capable of expression of full set of cas genes (see below) is shown in Fig. 3B" – should be Fig. 5B.
Author: Corrected.

RW Comment 6) "gel slices containing proteins with apparent molecular mass of Cas2 (~11 kDa) were subjected to in-gel tryptic digestion and MALDI-TOF mass-spectrometry" –ok. But next: "As can be seen from the mass-spectrometric data presented in Fig. 5C, pulled down proteins from induced, but not from uninduced cells revealed mass-peaks matching tryptic fragments of Cas1, Cas2 and tandem MS analysis confirmed this identification.". Not clear how Cas1 could be detected within cut 11 kD gel slices (no way), and Cas1 peaks are not shown in Fig. 5C.
Author: Now only Cas2 MALDI-TOF data are mentioned. In fact we analyzed Cas1 gel slices corresponding to Cas1 molecular weight as well, but decided not to record this obvious result.

RW Comment 7) "the purified Cas1-antibody can pull down both Cas1 and Cas2 from induced E. coli cells." Probably, but unclear in what proportions are the Cas1 and Cas2 present. Please clarify this whole part.
Author: The ability of resulting antibody to recognize Cas1 and Cas2 protein was proved by mass-spectrometric analysis. The exact proportions Cas1 and Cas2 were not checked during this study**.**

RW Comment 8) "Cas1-Cas2 adaptation complex (Fig. 6A). The complex contains the cas3 gene" – the complex is the complex, the gene is the gene.
Author: Corrected.

RW Comment 9) "The cultures analyzed in Fig. 4B" – should be Fig. 6B
Author: Corrected.

RW Comment 10) Fig.6C – would be nice to have also some relatively distant sites as a negative control. And again, not clear if Cas2 is there as well or Cas1 works alone.
Author: We performed amplification with primers complementary to distantly located site (positions 574095 - 574394 in *E.coli* genome) and found no enrichment, we did not consider that region further and did not include the data in the final text. Again although we did not determine stoichiometry of protein complex purified with anti-Cas1 antibody, we have shown that it contains both Cas1 and Cas2 proteins. ChIP procedure also was done in cells lacking *cas2* gene, and no enrichment on the Leader region was observed in this case (data shown on Fig. 7D) suggesting that the Cas2 protein is indispensible for Leader enrichment.

RW Comment 11) Experiment with KD1 cells seems to confirm Cas2 participation. But in fact it is a different strain so who knows why exactly there was no enrichment. KD1 strain PLUS Cas2 expression should give enrichment – that would be most persuading. I do not mean that this must be done before the thesis defense, but in principle it would be nice to do in the future to make it all finally confirmed.
Author: We thank reviewer for valuable suggestion. Proposed experiments would be conducted in the future work.

RW Comment 12) Also for the future: Cas2-specific antibody could be important. E.g. you could lose Cas2 along with the DNA fragments it bounds.
Author: Agree. Cas2-specific antibody is important and would be obtained in the future work.

RW Comment 13) "Observed enrichment of leader DNA precipitated from induced pG8mut harboring cells was not due to the effect of a different amount of proteins, because, as can be seen from Fig. 7B, the amount of Cas1 protein in induced KD263 cultures did not depend on the presence of the protospacer plasmids (compare lanes 2 and 4)." – I got the logic but could be explained better. "The same amount of Cas1 protein produced in the presence of other plasmids (Fig. 7B) did not result in the enrichment of precipitated leader DNA (Fig. 6C)." – smth like that.
Author: Agree. Re-formulated.

RW Comment 14) MySeq – should be MiSeq
Author: Corrected.

RW Comment 15) 912835 in Table 2 – is not amount of unique spacers but probably a total amount of seq reads mapped as spacers? "811545 unique plasmid spacers" – but spacers are of the fixed length, and possible amount of unique spacers is limited – the plasmid would have to be at least 810 kb length to make it possible :) In reality these are probably reads, not unique spacers. The number of unique spacers should be indicated in the Table, reads should not be called unique spacers unless they are indeed unique spacers.
Author: Corrected.

RW Comment 16) "Thus, the Cas1-Cas2 adaptation complex is associated with protospacers and the degree of this association correlates with the efficiency of protospacers use as spacer donors.". Formally, at this point of the narrative, the conclusion is incorrect, since those could be already added spacers as well.

This conclusion could be moved down after the paragraph:

"A derivative of KD263 cells that lacks functional IHF and is unable to acquire new spacers but is functional for CRISPR interference was also tested. While enrichment of the leader fragment was not significant in this case (Fig. 6C), the level of enrichment for HS1 was even higher than in KD263 cells with functional IHF (~24 versus ~16 fold, Fig. 9A, bar 2)."
Author: Corrected.

RW Comment 17) " treatment with FaiI had no effect on enrichment of a 33-nucleotide DNA fragment for HS2, which contains an internal FaiI site" – could it be just physically protected from restrictaze by Cas1? Cas1 was already washed away? So then it is either single stranded or modified? This could be just better explained to the reader at this point, before the Cas3 single strand story below.
Author: The cleavage with restriction enzymes is done after the reversal of cross-link, as is made clear in the methods section. There can not be protection or hindrance in this case.

RW Comment 18) Figure legend in Fig. 2 is missing.
Author: Corrected.