Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Andrey Krivoy

PhD Program: Life Sciences

Title of Thesis: Primed CRISPR-Cas adaptation in type I-E system of *Escherichia coli*: use of single-molecule and biochemical assays to verify models of the phenomenon at molecular level

Supervisor: Prof. Konstantin Severinov

Chair of PhD defense Jury: Prof. Konstantin Lukyanov ***Email**: k.lukyanov@skoltech.ru*

Date of Thesis Defense: November 30th, 2018

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

Dear Professors,
Thanks a lot for the reviews and questions. Below are the log of changes made in final version of the thesis, as well as answers to the comments that require an answer.

Reviewer: Prof. Ivana Ivancic-Bace

1. Font of the numbers in the List of Figures is different (p10 and p11)
Answer: corrected
2. In the personal contribution section – it should be: **At** the same time instead In the same time
Answer: corrected
3. p29 – l21 –it should be: extraction **of** DNA
Answer: corrected
4. p35, l17 – it should be: potenti**ally** active
Answer: corrected
5. p39, l16 – one supernatant is extra
Answer: corrected
6. p40, l2 – should there be glucose in the mixture?
Answer: In the absence of inducers, we did not see any leakage of the promoters that would lead to any detectable adaptation even on mutants known to be priming in advance (G-1T, C1T). Thus, we omitted this precaution.
7. p42, l9, should be: stained **with**, not iwith
Answer: corrected
8. p42, l12 in Results, should be: was based **on** Cascade purification…
Answer: corrected
9. p47, l5, it is Appendix **A2** not A1
Answer: corrected
10. p47, Perhaps change the title 4.5.1. Study of R-loop formation by Cascade
Answer: It is better, thanks! Corrected.
11. p48, l7, it is Appendix **A1**, not A2
Answer: corrected

12. p51, l5 – change the reference Szczelkun… into number
Answer: Done
13. p52, l18, I think it should be: **On one** hand, 37C is a … and, **on the other** hand, increased…
Answer: Cambridge dictionary describes that it could be used as it is now.
14. p61, l21 – I don't understand the sentence: Despite the fact that self…. Please correct
Answer: Corrected for "Free energy gain of any not B form DNA structures was minimized by Mfold analysis on dozens of candidates (160), thus, preventing as possible from potential base-pairing of detached parts of designed construct. Nevertheless, Ugene software repeat finder, used for post-problem discovery analysis, showed increased number of 8-10 nt repeats in synthetic templates compared to the pUC19 backbone (168)."
15. p61, l31 – it should be: as **was** done, not as is was for
Answer: corrected
16. p63, l1 – it is Figure 18B, not 18C
Answer: mean values from the repeats are shown on 18C, corrected to "Figure 18B and C"
17. p68 – I don't understand well the section about calculation of adaptation scores and Cascade binding – can you please write it up more clearly?
Answer: Deleted a bit unclear sentence "Thus, at least for the time point of 6 hours after CRISPR-Cas system induction, the ratio of adaptation scores is two times less than the ratio of Cascade binding rates." Target variants are indicated more clearly. Several changes are made in order to make the section more clear.
18. p71, Figure 23 legend, l2 – it should be **D9**, not Fig. 3.8. and also in the legend of Figure D9. Is should be as **Fig. 23**, not 3.9. Answer: Corrected
19. p72, l9 – it should be **between original** supercoiled, not between of original…
Answer: Corrected
20. p75, l1 in the summary, perhaps during the **research or thesis** instead of project, and maybe "in **my** opinion", not our
Answer: It is presumed that the research project that involved the candidate has eventually resulted in writing this thesis. Thus, I assume that project should be correct definition of the work done. Regarding my or our opinion, since the results were many times discussed among the contributors to the project, I have stated opinion that is agreed on by not me only, but by majority of co-authors of publications.
21. p76, l9 - it should be: … adaptation **than** standard …instead of that
Answer: Corrected
22. p76, in Table 2 legend, please change the text references into numbers
Answer: Done
23. p80, l3 – it should be: DNA molecules **in** supercoiled states…
Answer: Those are the shifts of molecules' supercoiled states. Corrected.
24. p80, l17 – it should be: … Cse assumes **remains** unknown.
Answer: Corrected for "is assumed to remain unknown"…
25. p80, l25 – should be **PAMs**
Answer: Corrected
26. p84, the reference 33 has question marks instead of some letters
Answer: Corrected for Latin letters
27. p96, at the bottom of the page, it should be – protocol **provided below**, without *is*
Answer: Corrected
28. p114, in Figure 10D, please arrange the reference in the legend
Answer: Done

Reviewer: Prof. Petr Sergiev

1. Dissociation curves shown at the Figure 9A seems biphasic to me. Is it just an impression or there is some explanation of this fact?

Answer:
Single exponential fit gets the time of overall dissociation times, thus, giving the notion of locking strength not assuming any intermediate states of the locking (free energy gain). Due to constraints on G3T mutant long average dissociation time, approx. >50 events could not be collected under the scope of the project to make an appropriate biphasic fit. It was discussed among contributors to the project that there might be two populations of dissociation times that were preliminary proposed to consider as potential result of separate locking strengths. If so, it might be that these populations are resulted from distinct conformational states of Cascade subunit on the same target. This speculation is supported by the fact of distinct locking abilities of the same target variants described in chapter with Cascade complexes and altered crRNA. Surely, more research, especially in combination with FRET measurements of Cse1 (and other subunits) conformational states may bring interesting results. Notably, it cannot be an error of different conditions applied to experiments, as these distributions of events are result of at least three separate repeats (on separate DNA molecules), each contributing to both "fast" and "longer" "subsets" of the events.
2. I think it might be too preliminary to regard the results of in vitro spacer acquisition experiments as negative. I would suggest using the PCR-based method for monitoring spacer acquisition, similar to the in vivo assay.
Answer:
Brilliant idea! Thanks! It may work.
3. I could not get a clear idea on why T4G mutant support acquisition, while the WT do not. The mutant T4G, except for the variant with additionally shorter complementary region, behaves absolutely identical to the WT in all in vitro assays, including degradation. To support kinetic model, perhaps one would expect slower binding or slower degradation?
Answer:
Indeed, T4G supports Cas3 recruitment and degradation of Cascade-bound targets with the same rate as wild-type target (according to our experiments). However, CRISPR-Cas interference has another step that should happen before Cas3 recruitment: Cascade binding and R-loop formation. Rates of this process is the main difference between T4G and WT target: 6.27 ± 1.7 sec for WT vs. 84.4 ± 12.3 for T4G (Figure 12 and p. 68 data conclusion). Thus, T4G is very interesting mutant to study as it allows to compare effect of Cascade binding rate only on the outcome of CRISPR-Cas response.

Reviewer: Prof. Konstantin Lukyanov

Fig. 10C and corresponding text (p. 51-52): "Overall, the effect of increased temperature was even more significant and together with Mg2+ resulted in approximately 8-fold increase of R-loop formation rate". The effect of temperature itself (without Mg2+) remains unclear.
Answer:
This exact set of measurements did not include only 37 ºC conditions lacking Mg^{2+} in the buffer. To investigate effect of only 37 ºC separate set of measurements was done on different cascade stock. Thus, it is logical to present this effect separately, as it is now done on supplementary Figure D12 (added). Approximate effect of 37 ºC is ~5-fold increase of R-loop formation rate. Thus, cumulative effect of both Mg^{2+} and temperature increase is close to their superposition.

Reviewer: Prof. Michael P. Terns

- Pages 12 and 17, use of the term "prove" and "proved" is an overstatement. Alter to a more precise term or phrase (e.g. "allowed to prove" could be changed to "provided evidence").
Answer: Corrected for "provided evidence" p.12 and "supported" p.17.
- Page 17, "AT-reach" to "AT-rich".
Answer: Corrected

- Page 84 Remove the question marks that appeared for reference 33 in place of the the letter “i” in numerous locations in this specific citation.
Answer: Corrected
- The reference 82 needs to be updated from BioRxiv to Cell (this study was very recently published in Cell). Moreover, as this is a highly relevant single-molecule study of priming, the specific findings perhaps deserve more description and discussion with regard to the key differences in conclusions reached with the study and the work of this thesis.
Answer:
The reference is updated. The reason this nice publication is not meticulously discussed in the thesis is its weak intersection with current thesis results. First, there is no target variants study displaying strong difference in machinery function for wild-type and priming target variants. Target variants are tested for Tfu-Cascade target recognition only, not for complexes of all the proteins. Second reason is that assembly (and translocation) of Cascade/Cas1-Cas2 and Cascade/Cas3/Cas1-Cas2 (“PAC”) is evidently supported, no doubts. Still, there is no solid evidence that “PAC” is directly responsible for efficient outcome of primed spacer acquisition. Surely, future research in this area should link all the results and answer whether acquisition of spacers is caused by “PAC” activity directly, or, potentially, by degradation of the targeted DNA molecules initiated by interference or/and “PAC” activity. In addition, of course, activity of “PAC” would be interesting to test for PAM-distal seed mutations (such as described in the thesis T4G).

- Page 34. Define phrase “ballast shoulders”. (I know it only as a term relating to railroad train tracks and not for DNA molecules…).
Answer:
These are 19-bp regions of cloned insert from each side of protospacer. “Shoulders” imply that they surround targeted sequence from both sides. Ballast should mean that “shoulders” do not play any role other than lowering the risk of PAM and protospacer sequences mutation during cloning. Now it is just “shoulders” for the clarity.

- Fig 5a: is “Cas2” a typo? (Cas2 is not a Cascade subunit as described).
Answer: It is Cse2. Corrected.

Reviewer: Prof. Timofei Zatsepin

Publication in Methods in Enzymology does not fit Skoltech guidelines.
Answer: The pre-defense report was signed by the Chairman and the program director. In report it is stated that all program requirements are completed.

Syntetic genes are not genes because they are not coding anything. These are just DNA constructs.
Answer:
As mentioned on p.35, the constructs were ordered from Eurofins Genomics, through service named “synthetic genes”. This name has stuck on these constructs.
Biotin is not sensitive to freezing/thawing cycles.
Answer: Digoxigenin is sensitive, what has an effect on MT constructs.

Best regards,
Andrey Krivoy