Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

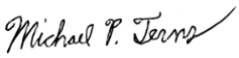
Name of Candidate: Anna Shiriaeva

PhD Program: Life Sciences

Title of Thesis: Interference and primed adaptation intermediates in type I CRISPR-Cas systems

Supervisor: Professor Konstantin Severinov, Skoltech

Name of the Reviewer: Michael Terns

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Signature:**  **Date: 05-10-2020**

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.

If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.

Reviewer's Report

Reviewers report should contain the following items:

- Brief evaluation of the thesis quality and overall structure of the dissertation.
- The relevance of the topic of dissertation work to its actual content
- The relevance of the methods used in the dissertation
- The scientific significance of the results obtained and their compliance with the international level and current state of the art
- The relevance of the obtained results to applications (if applicable)
- The quality of publications

The summary of issues to be addressed before/during the thesis defense

The studies described in this thesis represent pioneering and technically challenging work aimed at investigating early events of CRISPR spacer acquisition by the *E. coli* I-E and I-F CRISPR-Cas systems. The primary goals of this research were: 1) to gain a first detailed, *in vivo* understanding of the nature of transient protospacer DNA intermediates, 2) to elucidate the cellular nucleases and helicases responsible for both producing the intermediates and final spacer DNA products that undergo site-specific integration at CRISPR arrays (including PAM recognition and precise trimming and directed integration of the spacer in one of two possible orientations).

The candidate's work has revealed important insight into mechanisms of bacterial spacer DNA capture at host CRISPR arrays to provide heritable memories of the viruses/phages that infect them. Specifically, original work was performed establishing a high throughput library preparation and sequence methodology (dubbed Frag-Seq) capable of identifying small single-stranded or double-stranded DNA molecules produced in the cell. This technology was applied primarily to study DNA intermediates that accumulate upon induction of primed adaptation ("primed adaptation" is a specialized anti-viral defense pathway discovered by the Severinov group that enables host CRISPR systems to overcome viral escape via inevitable point mutations in the PAM and seed sequences of viral target DNA through triggering directed spacer acquisition against the previously encountered and memorized virus). Important information on the requirements of DNA sequences and forms was obtained by an orthogonal approach of directly transferring various short DNA species into cells via electroporation and monitoring and comparing the abilities of each DNA to become processed and integrated in the proper orientation.

Some primary conclusions drawn were that a combination of Cas3 nuclease/helicase action together with specific nuclease-helicase components of DNA repair pathways apparently collaborate with the Cas1/Cas2 integrase complex to accomplish protospacer excision from the genome as well as further pre-spacer to spacer conversion and CRISPR integration. Several mutant strains of *E. coli* were elegantly employed to investigate a number of candidate proteins. Evidence was presented in support of a novel role of the Rec J (a 5'-3' exonuclease) and the RecBC helicase in 5' end-trimming of pre-spacers.

The work is of high quality and rigor. A clever and powerful combination of molecular genetic analyses and sequence analyses/bioinformatic approaches were employed to gain insight into a relatively unexplored and important area of biology. The findings disclosed in this thesis were published in two manuscripts and a review and include a high impact journal (*Nature Communications*; fist author). The findings were also presented at four international conferences or departmental repeats in four distinct countries. A major strength of this body of work is how it comprehensively explores an important but enigmatic aspect of CRISPR function. Moreover, the Frag-Seq approach established and refined by the candidate has future potential to provide unique insight into mechanisms operating for other CRISPR systems beyond those reported in this thesis work.

All written sections of the thesis were easy to follow (despite being informationally-dense). The candidate has displayed a robust grasp of the field and powerfully synthesized the available information into coherent and detailed summaries of the pertinent information.

Overall, the thesis describes significant achievements in our first understanding of the pre-spacer generation and maturation processes and identification of nucleases and helicases that shape them into spacer substrates capable of being correctly integration into CRISPR arrays. This thesis should most definitely be accepted in partial fulfillment of the requirement for a degree of Doctor of Philosophy.

I have just one minor suggestion to consider for improving the thesis. The orientation of the crRNA and DNA strands during R-loop formation is presented inconsistently throughout the thesis (i.e. two ways of representing the same structure are used in figures 1 and 2 vs. figures 3, 10, and 14). It seems unnecessary to force a reader to endure the mental gymnastics required to interconvert the two ways of representing the same structure (it slowed me down especially since there are many features to keep track of). Please consider having a uniform representation of the crRNA-bound DNA structure throughout this dissertation to avoid confusion and help the reader (it was particularly clumsy to be steered to follow the diagram in figure one only to have things turn upside down in figure 3…).

Provisional Recommendation

X *I recommend that the candidate should defend the thesis by means of a formal thesis defense*

☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*

☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*