

Masarykova univerzita
Fakulta informatiky

Datum: 01-03-2022

Č.j./E.č.: MU-15/50117/2022/1643554/F1

Počet listů dokumentu:

Počet příloh a listů/sv.:/

Počet a druh nelist. příloh:/

PhD Thesis Proposal Review

Faculty of Informatics

Botanická 68a

602 00 Brno

The Ph.D. thesis proposal of Petra Němcová is focusing on the field of computational biochemistry, namely on studying the transportation of small molecules of ligands through tunnels in proteins. This is one of the most challenging tasks that have a direct impact on many disciplines, including drug design and protein engineering. The thesis goals are building upon the existing tool CaverDock, which currently forms the state-of-the-art approach to the simulation of ligand passage through a protein tunnel, which takes into account the geometry of the tunnel and also chemical forces, simulated by the AMBER force field package. The thesis suggests several extensions of the CaverDock tool that would lead to improvements in the accuracy of the ligand passage prediction. The thesis states three main goals:

- Extending CaverDock by a method for relaxation of tunnel bottlenecks.
- Proposing a methodology that would automatically set conditions for the relaxation and detect bottlenecks in the energetic profile of the ligand passage.
- Adding flexibility to the protein within the calculation of the ligand trajectory.

Comments on the results

The thesis presents the already achieved results of the student, that are related to the first stated goal. The proposed method for the relaxation of tunnel bottlenecks combines CaverDock and SANDER tools. The proposed method presents a sound contribution with feasible results, which is also supported by the publication, which will be presented at the ACM SAC conference (ranking B) in April 2020 and the student is the first author of this paper. Except for the result described in the thesis in Chapters 4 and 5, the publication also presents another result for searching the continuous trajectory of the ligand using graph representation (Section 4 of the attached publication). It is not clear to me why the student did not mention this result in the thesis. The current results, summarized in the publication, are representing an interesting contribution to the studied field and set a solid basis for future research plans of the student.

The structure of the thesis

The thesis is divided into several chapters. After the introductory part, the student describes the basic terminology and background of her research field. Here, I would appreciate more details about the calculation of the potential energy, mentioned in Equation 2.1. It is not necessary to present the detailed equations of individual parts of the main equation, but at least a brief description of the meaning and role of them would be beneficial. In Section 2.2.4, describing the AMBER force field, Equation 2.2 is not fully described, and “eq” in Θ_{eq} should be typeset as subscript. Similarly, in Equation 2.3, it is not clear to me what stands for the function f_{tt} . Finally, in Section 3 describing the applications, I have the impression that the topics of Subsections 2.3.1 and 2.3.2 are not on the same topical level: virtual screening seems to be rather one possible solution for the drug screening.

Chapter 3 describing the state-of-the-art in the given research domain is rather brief. This part should cover a significant part of the thesis. Although I also consider Chapter 2 as part of the discussion of the current state in the field, more details about tools listed in Chapter 3 would be very appreciated. For example, the last sentence of Section 3.1 describing the SLITHER tool, discussing overlapping, is not clear to me. Similarly, the main principle of the MoMA-LigPath and CaverDock 1.0 tools should be described in more detail to better understand their limitations.

Chapter 4 presents the proposed methodology, also forming the content of the attached publication. Here I’m missing the description of the SANDER tool and its principle. Maybe even better, I would expect this description in Section 2.2.4. The description is directly followed by Chapter 5 describing the achieved results, which I rather consider as the case study and evaluation of achieved results. This chapter gives a comprehensible view of the actual benefits and drawbacks of the current results, which I really appreciate.

Chapter 6 discusses several future research directions, which are clearly stated and challenging, which suggests that the student will be able to publish interesting results for her final thesis. Section 6.5 gives an overview of the time plan of the student for the rest of the studies. The plan seems to be

feasible; I'm only missing there the research challenge described in Section 6.3 and also more details about the research plans for the internship.

The subsequent chapters are concluding the thesis and provide details about the already accepted publication.

Text recommendations

The overall structure of the thesis is meaningful and provides the readers with a coherent story, clearly states the research challenges, and suggests their possible solutions. Overall, I would suggest adding explanatory images that would significantly improve the understandability of the presented methods and approaches (especially in Chapters 2, 3, and 4). Also, Figure 6.2 lacks the yellow line that is mentioned in the text. The list of references should be revised as well: I found wrong capitalization, diacritics, and missing information (ref. 26).

Language and typography

In general, the thesis is well-written and understandable. There are several language issues and typos that would be worth fixing but it does not significantly influence the text readability.

Questions for the defense

Within the defense, I would like to pose the following questions:

- In Section 3.3, you are stating that CaverDock solution is more user-friendly in comparison with the usage of only molecular dynamics. Do you know how user-friendliness was measured? Could you present which properties of the system are playing role in this?
- The results presented in Table 5.1 are evidencing negative values of energy. Could you comment on what does this mean?
- In the future research plans, you are also mentioning the usage of AI approaches to the automatic detection of bottlenecks. Are you also planning to research this option?

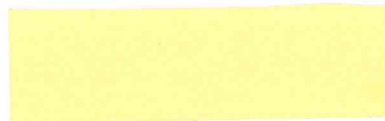
- In Figure 6.2, the interpretation of the graph around 13 A is not clear to me. Here we have the lower-bound trajectory way higher than the upper-bound one. Could you comment on that?

Conclusion

To summarize, the thesis proposal is dealing with a very interesting research topic and is clearly posing many challenging goals. The current results suggest that the student can conduct independent and internationally competitive research, which can lead to a high-quality doctoral thesis, submitted in the proposed time plan.

In my opinion, the presented results, summarized in the attached publication, are also sufficient for accepting the thesis as a rigorous one as well.

In Brno, 24/2/2022



Barbora Kozlíková
Associate professor and head of the visitlab
Faculty of Informatics
Masaryk University