Lajeado 23rd October 2023

To the Department of Gene Expression
Faculty of Biology
Adam Mickiewicz University in Poznań
Poland

Review of the doctoral thesis "Role of Transient Tunnels in Function of Enzymes with Buried Active Sites" by Aravind Selvaram Thirunavukarasu

- **General notes**

The focus of this doctoral thesis is the investigation of tunnels in proteins and the mechanisms underlying molecular movement within these protein cavities. The study employs molecular simulations and adaptive sampling techniques to elucidate protein flexibility and to uncover previously unidentified tunnels. Notably, the author have introduced a Python module, "TransportTools," designed to streamline the tunnel identification process. Furthermore, the study delves into the migration of water molecules through these tunnels, using a variety of water types in simulations. The results of this thesis is presented in three articles, where one is already published and more two that are avaliable on bioRxiv.

- **First publication**

The first publication describes the development of the Python module, TransportTools. The primary objective of this study was to offer a user-friendly approach for analyzing transport processes from large datasets. According to coauthors, mrg. Thirunavukarasu made a substantial contribution to this work. He conceived the project, designed and implemented the algorithms, prepared both user and technical documentation, analyzed the data, and wrote the manuscript. Nevertheless, even though the letters have been presented, I would like to inquire about the author's principal contribution to the development of the Python module, TransportTools.

- **Second publication**

The second publication rely on the contribution of narrow tunnels for the transport of water molecules througn three different α/β-hydrolases, including haloalkane dehalogenase, epoxide hydrolase, and lipase. The analysis were based on long adaptative simulations in order to better sample the conformational space

In the methodological section, the authors state, 'The last snapshots from simulations of all variants of each system were employed as their initial seeding structures for the adaptive MD simulations.' This raises the following question:

(i) Why were the last snapshots of each simulation used as initial seeding structures instead of the most prevalent conformation? How significant were the differences among the starting structures of each simulation?

I acknowledge that the study's primary focus is not on providing a structural explanation for the observed differences between the two structures. However, it could be intriguing to delve into potential explanations for these differences, especially given that an explanation was presented concerning the protein's activity. Therefore, my question, driven by sheer curiosity, is as follows:

(i) At first glance, the higher flow of water observed in the mutant Epx E470G could be an interesting factor contributing to the enzyme's increased activity. However, a key question arises: Could the increased flow potentially raise entropy within the active site, thus making it more challenging for the enzyme to attain the near-attack conformation?

- **Third publication**

In the third publication included in this doctoral thesis, mrg. Thirunavukarasu serves as the first author, highlighting his significant role in this study. The publication addresses the issue of how different water models available for molecular dynamics simulations could influence the transport of water molecules through protein tunnels.

Regarding the methological section, I would like to have more information about the choose of certain values:

(i) why the protonation states for AldO and Cyt P450 were determined at pH 8.5?

The results are well-presented and discussed, clearly highlighting the significance of this work for various research areas, from protein-ligand interactions to protein stability. As an overall conclusion of the work, it was found that the choice of the water model should be made carefully depending on the process being simulated. Therefore, for a more extensive discussion, I would appreciate it if the author could provide further insight into whether there are water models more suitable for specific processes, such as protein-ligand interactions. Alternatively, is it important to evaluate all types before commencing a new study? If that is the case, what parameter(s) should be taken into account when selecting the best water model?

The doctoral thesis is well-structured and easy to follow, showcasing Mr. Thirunavukarasu's capacity to present ideas and the work undertaken effectively. The "Summary of the Doctoral Research" section provides a comprehensive overview of the most noteworthy results presented in the associated manuscripts. Furthermore, in alignment with statements from all coauthors, it elucidates the author's contributions to the development of all

the works. It is also essential to emphasize the author's ability to collaborate with colleagues within the research group, which is a vital characteristic for a researcher.

- **General questions and comments**

(i) In the section titled "Developing a comprehensive methodology for analyzing transient tunnels in enzymes" within the "Summary of the doctoral thesis," the author states, "...Considering this, using the dihedral angles of residues surrounding the active site as a metric would enable the exploration of a wide range of states…". It would be interesting to include a figure illustrating the location of the dihedral in the structure.

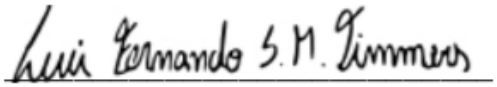
(ii) In the same section, "Developing a comprehensive methodology for analyzing transient tunnels in enzymes" within the "Summary of the doctoral thesis," the author states, "...The main tunnel is surrounded by residues 145, 176, and 172…". To enhance understanding, it would be helpful to include the names of these residues.

(iii) Additionally, in the same section, the author describes, "The RMSD analysis in Figure 4D underscores the efficacy of adaptive sampling in exploring diverse protein states across various simulations." I would like to inquire why RMSD was chosen as the metric for evaluating the effectiveness of adaptive sampling and whether clusters of conformations were evaluated based on the variation of dihedral angles.

- **Overall evaluation**

This research significantly enhances our understanding of protein structures and the movement of molecules within them. The thesis is well-structured, and the methodology is commendable. Overall, this is a well-executed thesis that makes substantial contributions to the study of protein tunnels and molecular mobility. These results demonstrate that the author possesses a deep understanding of the scientific domain and is capable of independent, creative

scientific work. Therefore, I recommend the thesis for defense, moving forward to the final stages leading to the doctoral degree.



Luis Fernando Saraiva Macedo Timmers, PhD
Laboratory of Bioinformatics and Protein Biochemistry - LabPro
Department of Biotecnology
University of Taquari Valley, Lajeado – Brazil
Phone: +55 51 982997309
E-mail: luis.timmers@univates.br