Lajeado 05th November 2023

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

Review of the doctoral thesis "Simulations of Ligand Binding Processes in Proteins" by Dheeraj Kumar Sarkar

- **General notes**

The focus of this doctoral thesis was the investigation of thermal and kinetic properties regarding the protein-ligand binding/transport process and its influences on changes in protein tunnels. The study employed adaptive sampling methods based on Markov State Model to assess alterations in protein tunnels induced by the events of ligand binding. The thesis comprises three articles that present stages of development, evaluation, and application of a high-throughput simulation protocol using protein-ligand complexes. Furthermore, among the three manuscripts included in this thesis, two have already been published in high-impact journals, with Dheeraj listed as the first author. The third manuscript is available on bioRxiv. In addition to the manuscripts included in this thesis, the author also presents two more manuscripts, showcasing their ability to collaborate with different research groups.

- **First publication**

In the first publication, mrg. Sarkar serves as the first author, highlighting his significant role in this study. Overall, the publication addresses the impact of seeding structures for adaptive sampling in molecular dynamics to evaluate the process of protein-ligand (un)binding. The manuscript involves approximately 900 simulations of the haloalkane dehalogenase mutant LinB86 associated with its substrate 1,2-dibromoethane. For the investigation, four seeding schemes were employed according to the ligand's position: (i) Bulk, utilizing a grid scheme, (ii) Cavity, wherein the substrate was docked using AutoDockVina, considering the center of the search box as the center of mass of the catalytic residues, (iii) Cavity&Bulk, incorporating 15 substrate positions from both Cavity and Bulk schemes, and (iv) Tunnels, which initially performed an MD simulation using the ligand-free form of haloalkane dehalogenase to identify the most open tunnels.

The methodological section is well written; however, I would like to ask a couple of questions that could help me better understand the methods employed in this study. Here are the questions:

(i) Why were the protonation states for LinB86 determined at pH 8.5?

(ii) Concerning the Cavity scheme, were the docked conformations selected based on an energy score? I understand that the goal is not to reproduce an experimental position in the active site; however, was it necessary to validate a docking protocol before running a simulation to identify the best substrate conformations?

(iii) In the Cavity&Bulk scheme, how were the substrate conformations selected from the other schemes?

(iv) In the "2.1 Equilibration MD simulations of seeds" subsection; What criteria were used to determine that the simulation reached equilibrium/convergence?

(v) In subsection "2.2. High throughput ASMD to study substrate (un)binding processes", which parameters were altered among the replicates?

(vi) In subsection "2.4. MSM analysis and comparison", how much could the side-chains influence the Center of Mass (COM) of the bottleneck residues?

Regarding the results and discussion section, the information is well presented, and the discussion aligns coherently with the obtained results. To further enrich the discussion, I'd like to address some points that could benefit from more elaboration and the potential proposal of hypotheses regarding the results from adaptative sampling molecular dynamics simulations.

(i) In Figure 2, it is evident that, across all schemes, the DBE molecule predominantly occupies the enzyme's tunnel. While this observation doesn't draw much attention in the Bulk scheme, in all other schemes, the rate of DBE entry into the active site appears notably low from my perspective. Is it anticipated that the rates of DBE entry into the active site are low for the Cavity, Cavity&Bulk, and Tunnels schemes? Additionally, is there a method to quantify the energy barrier for the DBE to traverse from the bulk to the active site?

(ii) In Figure 4C, the deviation in the Cavity scheme appears excessively high. Was this behavior expected? My suggestion is to use only the Cavity&Bulk and Tunnels schemes to discuss K_d values, unless there is an explanation for the high deviation observed in the Cavity scheme.

Overall, the results presented in this manuscript are very interesting and could be applied to different approaches using molecular dynamics. Additionally, the good agreement with experimental results reinforces the method's robustness.

- **Second publication**

The second publication focuses on the development of the Python module, TransportTools. The main objective of this study was to provide a user-friendly approach for analyzing the most relevant or preferred tunnels used during the transportation of small molecules from extensive datasets. The author contributed to this work by generating the datasets, conducting user-testing of the developed tools, and drafting the manuscript.

- **Third publication**

In the third publication included in this doctoral thesis, Mr. Sarkar is the primary author. The publication introduces a new strategy aimed at enhancing the stability and activity of Cytochrome C under extreme conditions, utilizing ATP and salicylic acid as nanostructured hydrotropes.

The study's results are fascinating, showcasing potential applications across various biotechnological fields, particularly in environmental and food-related domains. The authors illustrate, through computational and experimental results, the potential use of this innovative strategy involving protein confinement within nanostructured hydrotropes for industrial bio-catalysis.

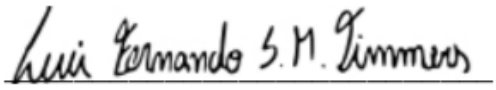
While discussing the results in the section 'Structural features of nano-structured hydrotrope-caged Cyt c,' a temperature of 26.85°C is mentioned. I'm curious about the reasoning behind using this temperature instead of 36°C.

- **General comments**

The doctoral thesis is well-structured and easy to follow, demonstrating mrg. Sarkar's ability to effectively present ideas and the undertaken work. All sections are well-organized, offering a comprehensive overview of the author's contributions to the research field. It's also important to highlight the author's capacity for collaboration within the research group, a vital characteristic for any researcher.

- **Overall evaluation**

The thesis is well-structured, and the methodology is commendable. Overall, this is a well-executed thesis that makes significant contributions to the study of the ligand (un)binding process. These results illustrate the author's profound understanding of the scientific domain and their ability to conduct independent, creative scientific work. Therefore, I highly recommend the thesis for defense, advancing toward the final stages for 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