Jerzy Haber Institute
of Catalysis and Surface Chemistry
Polish Academy of Sciences

HR EXCELLENCE IN RESEARCH

Kraków, 16 December 2023

Referee report for a PhD thesis submitted

by MSc Dheeraj Kumar Sarkar

entitled „Simulations of ligand binding processes in proteins”

prepared under supervision of dr hab. Jan Brezovsky, prof. UAM

Introduction

This PhD thesis is a collection of three thematically coherent articles supplemented with:

- a list of publications,
- abstracts in Polish and in English,
- a list of used abbreviations,
- an introduction to ligand transport in proteins,
- a presentation of the key results from the articles with focus on modelling transport processes in proteins and effects of hydrotopes’s binding on Cyt c stability and enzymatic activity,
- conclusions,
- a bibliography encompassing 57 references,
- co-author statements about their contributions to the articles.

The “Introduction to ligand transport in proteins” section offers a short (2.5 pages of text + 2 figures) review heavily focused on the molecular dynamics (MD)-based methods to study the process of ligand transport through protein tunnels. While I fully understand these methods deserve a thorough presentation, as they were the subject of two papers included in the thesis, it seems to me that the introduction would offer a more complete background information that would put the research done in an appropriate context if it had included at least a short discussion of the following subjects: a) experimental methods that can be used to verify

Niezapominajek 8, 30-239 Krakow, Poland
phone +48 12 639 51 01
phone +48 12 425 19 23
fax +48 12 425 19 23

Bank account: Bank Gospodarstwa Krajowego
in PLN: PL 36 1130 1150 0012 1186 5820 0004
in EUR: PL 09 1130 1150 0012 1186 5820 0005
SWIFT: GOSKPLPW, VAT Number: PL 6750001805

computational findings, b) Markov state models, how they are constructed and interpreted, and c) effects of hydrotrope substances on protein stability and enzymatic activity.

The “Summary of doctoral research” section is succinct (5.5 pages of text + 3 figures), but is presents all major results of the three research papers included in the thesis.

The articles included in the thesis are thematically coherent. The first treats about the role of starting geometry for efficient simulation of ligand transport through enzymes’ tunnels. The second article presents the TransportTools library – a useful tool for analysis of molecular dynamics trajectories with focus on transport phenomena. In the third paper the Author used MD simulations and Markov state models to study interactions between a model enzyme – cytochrome c and two hydrotrope substances: ATP and cholinium salicylate. Thus, the protein dynamics and Markov state models are the common themes for all three articles.

In paper #1 Mr. Sarkar is the first author, in paper #2 he is the 6th author, whereas in paper #3 he is the second author, but the first from the Polish team.

I find the research work presented in the thesis as timely and important. Reliable computational protocols to study ligand transport phenomena (paper #1) and robust software tools aiding analysis of massive MD simulation (paper #2) are very much in demand, as ligand transport processes are important for a wide range of processes and applications. Moreover, understanding how hydrotope molecules stabilize proteins and sustain their catalytic function in harsh conditions (paper #3) is of great practical value.

Scientific comments and questions

Concerning the investigations reported in paper #1 “Incorporating prior knowledge … “, I have the following questions.

1. On page 12 it is stated that ““Cavity” could significantly sample bound poses and egress of DBE to the bulk but frequently failed in sampling DBE binding to the active site.”, and this fact was considered as a limitation of the “cavity” scheme. My question is a bit philosophical, yet, I wonder, if

we could not evoke the time-reversible nature of the Newtonian equation used to propagate the MD trajectory and claim that egress trajectories played backward represent ingress ones. Are there any arguments against this time-reversal procedure?

2. If I understood correctly, Markov state models were used to estimate k_{off} and k_{on} rate constants, but only their ratio is discussed in the paper and the thesis. I wonder why the individual values are not presented and compared to experimental values or at least their estimates.

With respect to the paper #2 “TransportTools: A library for …” I must admit that both the paper and its short discussion in the “Summary of doctoral research” section of the thesis are rather “dry”, i.e. they do not give (much) information on how the actual workflow of TransportTools analysis looks like and how this tool makes life easier. Could the Author address this issue by, e.g. presenting (shortly) the use-case 3, that he developed?

Concerning the computational part of the research work described in paper #3 “Nano-structured hydrotrope-caged cytochrome c …”, I have the following questions.

1. In MD simulations the concentration of ATP and IL were the same as those used in experimental tests. However, the concentration of protein seems to be much higher in the MD simulations than in the experiments. I fully undestand the need to keep the size of the simulation box at reasonable level, yet I wonder if a simulation with a (better) reproduced molar ratio between protein and ATP and IL would not be more realistic, especially for ATP (in MD simulations Cyt c : ATP ratio = 1:3, in experiment – peroxidase activity tests, this ratio was 2µM/5mM = 1:2500).

2. The high peroxidase activity of Cyc c is related to the dissociation of Met80 from the iron ion. What kind of force field model was used in the MD simulation for this critical (for enzymatic activity) Fe-S bond? If it was a bonded model, which does not allow for bond dissociation, why the non-boded model was not adopted?

3. The values of standard deviation reported in Fig. 4 f-i are relatively small (the extreme is 917±1 ns). Is it typical for this sort of process/simulation? How were the replica MD simulations initiated that gave the values used to calculate the average and std dev values?

Typographical and stylistic errors noticed

The commentary part of the thesis, which includes “Introduction to ligand transport processes in proteins”, “Summary of doctoral research” and “Conclusions and future perspectives”, has probably been written under considerable time pressure, as it contains quite some vague statements and typographical and stylistic errors. They are listed beneath.

Page 8: “CV-free methods, …, are biased-free simulations but often suffer from exploiting the binding-pose and proper utilization of transport pathways for sampling ligand binding processes.” - the meaning of this sentence, especially its second part, is unclear to me.

Page 9 (and also at several other places): “ligand mitigations from the active site to the bulk …” - unknown to me usage of “mitigation”, it should probably be replaced by “migration”

Page 11: “...kinetic traps that are frequently countered in MD simulations …” - “countered” should be replaced by “encountered”

Page 12: “Additionally, “Tunnels” scheme had more control over the sampling of (un)binding pathways …” - the meaning of “having control” is not clear to me in this context.

Page 12: ““Tunnels” could be utilized in exploring more meaningful states to connected with use of complex transport network …” - a bit vague sentence.

Page 13: caption of Fig. 3: “Ligand positioning at the cheapest chunks of tunnels …” - it was not explained what is meant by “the cheapest chunk of a tunnel”.

Page 14: “higher saving frequency (100 ps) … lower saving frequency of ~10-20 ps” – it seems to me that saving every 100 ps means a lower frequency than saving every 20 ps.

Page 16: “… the backbone dynamics of Cyt c was nearly reduced to half in presence of ATP or IL or both the solvents” – it is unclear which measure of backbone dynamics was reduced to half of its value.

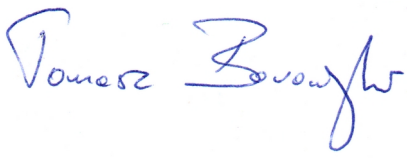
Page 17: „the current thesis consideres various considerations for the applicability of …” - style

Page 17: „the challenges of exploring and exploiting rare event sampling processes …” - it is unclear what is meant by „exploiting” in this context.

Conclusions

My above-listed comments to the thesis concern minor issues, mostly related to the way the results have been presented, and they do not diminish my high assessment of the work presented in the thesis. The thesis is well written, the conclusions drawn from the results are sound and very well supported by the data. The results presented in the thesis have been published in international journals of high standard of peer review (2 published papers), which means they have already been positively assessed by scientific reviewers.

Hence, I conclude that the thesis presented by Mr. Dheeraj Kumar Sarkar meets all the requirements for doctoral dissertations included in the Act “Ustawa z dnia 20 lipca 2018 roku Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2023 r. poz. 742)” and I submit the application to the Scientific Discipline Council, Biological Sciences, Adam Mickiewicz University Poznań for the admission of Mr. Dheeraj Kumar Sarkar to further stages of the doctorate.



Prof. dr hab. Tomasz Borowski

Niezapominajek 8, 30-239 Krakow, Poland
phone +48 12 639 51 01
phone +48 12 425 19 23
fax +48 12 425 19 23

Bank account: Bank Gospodarstwa Krajowego
in PLN: PL 36 1130 1150 0012 1186 5820 0004
in EUR: PL 09 1130 1150 0012 1186 5820 0005
SWIFT: GOSKPLPW, VAT Number: PL 6750001805