Prof David J. Clarke
CHAIR OF MASS SPECTROMETRY
DIRECTOR – Postgraduate Masters in Analytical Chemistry
SCHOOL *of* CHEMISTRY
The University of Edinburgh
Joseph Black Building
King's Buildings
West Mains Road
Edinburgh EH9 3FJ
Email : dave.clarke@ed.ac.uk
Telephone : +44 (0)131 650 4808

11th June 2024

Re: Review of submitted PhD Thesis

By: Mr Dhanasekaran Balakrishnan
At: Jagiellonian University in Krakow, Poland
Entitled: 'In silico modelling of natural and designed biological structures'
Completed under the supervision of Prof. Dr. Jonathan G. Heddle.

The formal grounds for reviewing the doctoral dissertation of is a letter from Prof. dr hab. Maria Rapala-Kozik dated 11th April 2024.

To whom it may concern,

The presented review report is organised into six sections:

1. General description of the manuscript
2. Project aims
3. Evaluation of the literature used
4. Evaluation of individual parts of the dissertation
5. List of queries
6. Final conclusion and evaluation statement

I. General Description of the Manuscript.

This thesis submitted for review includes an abstract, an introduction, detailed experimental section, two chapters of results and an overall summary. The thesis is properly structured and balanced and in its substantive part the thesis consists of 120 pages. The bibliography covers 15 pages and contains 253 references. The author has contributed to three peer-reviewed papers during their PhD studies; including one first author review paper entitled 'Delivering DNA origami into cells'.

2. Project Aims.

The two main aims of the work are clearly stated and are suitable for a PhD thesis. In Chapter 3 several *in silico* methods are used to predict the structure of the NinK protein. In Chapter 4, *in silico* course gained simulations is used to investigate the fusion pathway between nanodiscs and lipid bilayers.

3. Evaluation of the literature used

The selection of literature and source materials is adequate. The list of used literature includes 253 references, taken from a range of peer reviewed publications. It includes items of literature on the subject ranging from the 1925 to current articles from scientific journals. To my knowledge, there are no significant omissions in the cited material. In my opinion, this proves a solid query and a significant workload for a PhD candidate.

4. Evaluation of Individual Parts of the Manuscript.

Introduction. The introduction consists of an outline of protein structure and the emerging role of computational approaches in protein structure prediction. The importance and diversity of lipids and biological membranes are then introduced and then a literature review on the development of artificial cells and nanodiscs is presented. Finally, a range of *in silico* tools for biomolecular structure prediction are described. The introduction is very wide ranging, covering a range of aspects from basic protein chemistry, structure and function of specific proteins to new AI/deep learning tools for protein structural prediction. Consequently, in paces some aspects are only described in a superficial manner. Although the introduction is adequate as it is, I think the thesis would be improved by narrowing the scope of the introduction – maybe to focus on the recent development of powerful *in silico* methods for proteins structure determination. This would give the author the possibility of adding more detail in the techniques that are pivotal to the novel work presented in the thesis.

Methodology. The methodology is described in good detail and well referenced.

Chapter 3. Several *in silico* techniques are used to predict the structure of the NinH protein. These include sequence alignments to reveal conserved residues and conserved domains allowing a homology model to be proposed using Modeller 9v21. The homology model was validated for Ramachandran outliers and quality of the model was assessed using ERRAT plots. The approach to building this structural prediction seems valid and well performed. Using restraints from experimental findings, higher order structure was modelled using HADDOCK to produce possible structures for the homodimeric assembly of NinH. The dimeric HADDOCK model was in good agreement with experimental SAXs analysis of the protein. Finally, this HADDOCK model was compared to the predicted structure produced using two recently developed an artificial intelligence system - AlphaFold 2.0 and ESM. Although the two modelled structures shared similar features, significant difference in the two models were evident – notably the secondary structure in the C-terminal region of the protein. In addition, higher order structure was predicted to differ between the HADDOCK and AlphaFold models. The validity of each model is discussed in general terms. However, the AlphaFold dimer model was never discussed in the context of the SAX analysis. It would have been interesting to see which of the two proposed different dimeric models agreed more with the experimental data. The advantages and limitations of the different models are discussed, and the author demonstrates that he has an understanding of the *in silico* techniques used. However, the insights into the *biology* gained from these proposed models could have been discussed in more detail.

Chapter 4. Describes the use of course grained simulations to model the pathway of calcium-mediated membrane fusion between nanodiscs and a lipid bilayer. MD simulations did indeed show reproducible interactions between the nanodisc and lipid bilayer. The so called 'stalk expansion pathway' was monitored by modelling the distance between the phosphate groups of the nanodisc and the bilayer and the number of molecular contacts. Importantly, simulations where preformed with membrane protein-loaded nanodiscs and empty nanodiscs, and similar pathways were obtained for each, suggesting that protein loading does not influence the fusion process. The chapter then goes on to describe the effect of introducing artificial pores in the bilayer, and removing the membrane scaffold protein from the nanodisc. Taken together, these simulations suggest that three main conditions are required to complete the fusion process - (i) hydration of lipids; (ii) rearrangement of the scaffold protein; (iii) destabilisation of the metastable reverse micelle intermediate. These findings add to the knowledge in the field and are important in the context of designing and engineering artificial cells. These findings allow an experimental pathway to be proposed that fulfils these three requirements. The simulations are supported by experimental data and these finding have been successfully published in a peer reviewed journal (*ACS Appl. Mater. Interfaces*).

5. Errors, Inconsistencies and Queries.

- Pg 18: Figure 1.9 referred to incorrectly.
- kDA should be kDa (e.g. pg 27).
- Pg 77: Fig 3.9. Residues are highlighted in green, not red as indicated in figure legend.
- Pg 95: Fig 4.1. Add more detail in figure legend. Refer to both A and B panels.

6. Final conclusion and evaluation statement.

The thesis adds valuable insight into to the structure and function of two very different biological macromolecular systems. The results of the work add significantly to the understanding engineering biological systems and the field of synthetic structural biology. The author has demonstrated sufficient knowledge and understanding of the issues and literature on the subject and the methodology used is well described.

In summary, I hereby conclude that this thesis is, in my view, suitable for awarding the PhD degree.

Yours sincerely,

Prof David J. Clarke,
School of Chemistry, University of Edinburgh.