### Thank you for the opportunity to improve our manuscript. We believe the readability and framing of the abstract and our conclusions have been improved by the reviewer's suggestions. We hope you find the revised version of our manuscript acceptable for publication in CMC.

### **Point by point response to Reviewer Comments:**

**REVIEWER 1/Problem 1**

**Comments:** The review article entitled Developing Generalizable Scoring Functions for Molecular Docking: Challenges and Perspectives in which the authors aimed to review and contextualize the challenges and best practices when training novel scoring functions to improve their accuracy and generalizability in predicting protein-ligand binding affinities. The article idea is good and can be accepted for publication after some revision. The authors should revise the manuscript according to the following comments. Abstract: Should be made coherent. Introduction: It should be updated and improvised by adding the latest literature/references and citing the latest articles. The following articles might be of great help. The figures quality should be improved. Limitations of the study should be added. Future recommendations should be added as a separate paragraph.

***Reply 1:*** *We would like to thank Reviewer 1 for the careful reading of our manuscript. We have revised the abstract, there were indeed problems with the grammar, we apologize for the oversight.*

*We have not received the list of articles that the reviewer suggests to include in the Introduction section. However, we have updated several subsections of the introduction section with some of the latest articles and pre-prints.*

*Regarding low quality figures, we have provided figures with high resolution, perhaps the image was compressed when creating the PDF file?*

*We appreciate the suggestion to add a “Limitations” section, but we do not think the inclusion is feasible, since this is a mini-review and not a research article. We have also not included a separate “Future Recommendations” section since it is not recommended as such in CMC's Guide to Authors. However, we attempt to provide an overview of the present and future of scoring functions and we believe the “Conclusions” section includes our viewpoint on what research paths could be explored to achieve further progress in the field. Once again, our sincere thanks to the reviewers for helping to improve the manuscript.*

**REVIEWER 2/Problem 2**

**Comments:** The manuscript explores the optimal methods for training novel scoring functions to enhance accuracy and generalizability in predicting protein-ligand binding affinities. The authors demonstrate that effective training hinges on the careful consideration of both training data quality and methodology. Furthermore, hybrid approaches that integrate empirical data with machine learning present a promising strategy for surpassing existing scoring functions. The manuscript is well-structured and written, making it suitable for publication in the journal.

***Reply 2:*** *We would like to thank reviewer 2 for the positive feedback regarding the article.*

**REVIEWER 3/Problem 3**

**Comments:** This manuscript “Developing generalizable scoring functions for molecular docking: challenges and perspectives” describes modified scoring function for improved functions for molecular docking. The most important points seem the followings: a robust training methodology, and high quality and unbiased datasets for training and evaluation of scoring functions. As authors state, purpose of this review is to present best novel scoring functions for the most accurate prediction of protein-ligand binding affinities. Conclusions of this review article seem to be the best is “Hybrid methods”. I, as a reviewer of this manuscript, would like to request authors to present several examples not only themodynamic analyses (Figures 1 and 2), but also the interaction of a low-molecular-weight drug with a specific and important amino acid sequence within a protein at a chemical structure level. Examples for positive (binding) ones and those for negative (non-binding) ones would be much desirable. This would be helpful to readers, and it would also attract readers to this review article.

***Reply 3:*** *We thank reviewer 3 for the critical reading of our manuscript. We believe that presenting examples of docking results for positive and negative binders is beyond the scope of this mini-review. We provide our assessment of the current state of the art in the field, as well as our views of which practices and strategies could result in furthering advancement in the field of scoring functions and molecular docking. We are not certain that hybrid methods can be superior to purely empirical or machine learning approaches, rather, we suggest that further research in this field has the potential to combine the advantages of each approach and avoid their weaknesses. We have modified the text of the conclusions and the abstract to make this more clear to readers and avoid possible misinterpretations.*