Responses to Reviewers (red)

Reviewer: 1  
  
Recommendation: Accept With No Changes  
  
Comments:  
The authors addressed the reviewers comments and the manuscript can be accepted for publication.  
  
Additional Questions:  
1. Which category describes this manuscript?: Research/Technology  
  
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below. : Very Relevant  
  
1.  Please explain how this manuscript advances this field of research and/or contributes something new to the literature. : The authors describe a machine learning method to study the effect of single point mutation on the ability of HIV complexes to bind. For this purpose they are using machine learning classification.  
  
2. Is the manuscript technically sound? Please explain your answer under Public Comments below. : Yes  
  
1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below.: Yes  
  
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below.: References are sufficient and appropriate  
  
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below.: Yes  
  
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below. : Satisfactory  
  
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below.: Easy to read  
  
6. Should the supplemental material be included? (Click on the Supplementary Files icon to view files): Does not apply, no supplementary files included  
  
7. If yes to 6, should it be accepted: As is  
  
Please rate the manuscript. Please explain your answer.: Good  
  
  
Reviewer: 2  
  
Recommendation: Author Should Prepare A Minor Revision  
  
Comments:  
The work of this paper shows that the use of a combination of multiple characteristics such as conservation, binding energy and contact enhance the prediction of non-binding ligands after they are docked using HADDOCK. Although many studies are available about binding, the work in this study distinguishes itself by combining the learning technique, docking and energetic calculation. Compared to the version of ACM-BCB conference paper, this paper is much better written, particularly in the Results section. It includes two new tables that provide more details about the data and results.  There is also a new section discussing of a case study. Authors have addressed my previous questions. I have a few suggestions that might improve the paper.  
1.      Adding a Conclusion to summarize the finding would be helpful.

We agree, and have added a Conclusion. This Conclusion is now section 4 of the manuscript, while the Discussion is now section 5.

2.      The top two lines in Figure 5 should be visually distinguished.

The top line has had “plus” signs added to it, in order to visually distinguish it more from the 2nd-to-top line.

3.      I will leave this to the editor for opinion. It looks to me that there should be a summary at the introduction section to describe the major difference between this paper and the conference paper given that the conference paper was published. Currently the conference paper is not cited. The testing data in the two papers are slightly different. Perhaps that could be indicated as well.

We have changed the title to “Docking Features for Predicting Binding Loss Between Human and HIV-1 Proteins due to nsSNPs”, as we feel the additional predictions made, on HIV-1 – human protein-protein interactions and disruptive nsSNPs, warrants an altered title.   
  
Additional Questions:  
1. Which category describes this manuscript?: Research/Technology  
  
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below. : Very Relevant  
  
1.  Please explain how this manuscript advances this field of research and/or contributes something new to the literature. : The work of this paper shows that the use of a combination of multiple characteristics such as conservation, binding energy and contact enhance the prediction of non-binding ligands after they are docked using HADDOCK. Although many studies are available about binding, the work in this study distinguishes itself by combining the learning technique, docking and energetic calculation.  
  
2. Is the manuscript technically sound? Please explain your answer under Public Comments below. : Yes  
  
1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below.: Yes  
  
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below.: Important references are missing; more references are needed  
  
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below.: Yes  
  
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below. : Satisfactory  
  
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below.: Easy to read  
  
6. Should the supplemental material be included? (Click on the Supplementary Files icon to view files): Does not apply, no supplementary files included  
  
7. If yes to 6, should it be accepted:  
  
Please rate the manuscript. Please explain your answer.: Excellent  
  
  
Reviewer: 3  
  
Recommendation: Author Should Prepare A Major Revision For A Second Review  
  
Comments:  
Authors have addressed all my comments in the first revision. However, there are still some major problems in the flow of the manuscript. Especially newly added figures needs more explanation.

We agree that the legend of Figure 2, as well as legends of other Figures, needed clarification. We have thoroughly revised legends for Figures 1, 2, and 5. We have also added more detail to the inset-legend for Figure 3 (“other features”). We feel the new legends more clearly specify the conclusion or main point for each Figure.

Although Figure 2 shows the size distribution of residues in non-binders and binders, it does not give any information about the physicochemical properties or any tendency toward a specific residue. Residue propensities would help at this point.

We have added information in-text (section 3.2, just above Figure 2) about the most common physicochemical shifts (i.e. “before” and “after” categories of amino acids). There was a preponderance of aromatic-to-small-hydrophobic mutations among non-binders, as well as compound (i.e. multiple) mutations. Charged-to-hydrophobic shifts were also more common among non-binders, than among binders. We do not provide exact numbers, as we feel that an extra Figure or Table would be needed, and this would be beyond the scope of a minor revision. Also, an extra Figure or Table would push the manuscript over the limit of 14 pages. Finally, we do not want to lead the reader into assuming that physicochemical properties of the mutations have been included in the model. While models based on such features have been developed, ours relies on more purely physical parameters of interaction, as simulated by docking.

I am still wondering other performance values - precision, recall, specificity, sensitivity and F1-score –and think that these values should be added.

We have now included: area under receiver operator curve, precision, recall/sensitivity, specificity, and F1-score for non-binders, as well as for binders (for the latter 4 performance scores, i.e. precision-F1-score, we have used a confidence threshold of c>0.60, for both non-binders and binders).   
  
  
Additional Questions:  
1. Which category describes this manuscript?: Research/Technology  
  
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below. : Relevant  
  
1.  Please explain how this manuscript advances this field of research and/or contributes something new to the literature. : It introduces a new method to predict the effect of SNPs in protein interactions by using a machine learning approach that uses a refined list of features derived from docking.  
  
2. Is the manuscript technically sound? Please explain your answer under Public Comments below. : Yes  
  
1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below.: Yes  
  
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below.: References are sufficient and appropriate  
  
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below.: Yes  
  
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below. : Could be improved  
  
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below.: Readable - but requires some effort to understand  
  
6. Should the supplemental material be included? (Click on the Supplementary Files icon to view files): Does not apply, no supplementary files included  
  
7. If yes to 6, should it be accepted:  
  
Please rate the manuscript. Please explain your answer.: Good