Dear Dr. Gao,

We very much appreciate the reviewers' helpful and largely positive comments on our work on how protein conformational changes affect calculations of relative binding free energies. We have finished incorporating the changes as recommended by the reviewers, and we hope you will find the revised manuscript now suitable for publication.

A detailed list of our changes in response to the reviewers is provided below. We also took the opportunity to check the paper for typos one more time and corrected a number of typographical and grammatical errors, and we proposed an adjustment to the title just to make it more clear the paper is dealing with protein conformational changes.

Sincerely,
David Mobley

Editorial Review:

- *Use Journal name abbreviations*
- *Abbreviations should end in a period*
- *Ref 47 Reformat*
- *Ref 36-40 refromat*
- *Ref 51 and 58 reformat*

We made these changes, except for reference 58, which was left unchanged because it already is formatted as a book and uses @book in our bibliography; it already seems to follow suggested example, i.e. we use "Bowman, G. R.; Pande, V. S.; Noé, F. *An introduction to markov state models and their application to long timescale molecular simulation*; Springer Science & Business Media, 2013; Vol. 797." If other changes are needed please let us know.

As per a separate exchange with Dr. Jorgensen, we also made these changes he requested:

- Supplementary information changed to only reference data at external link.
- Removal of 'n-' in chemical names on in text, figures, and tables.
- Footnote has been placed into text

Reviewer: 1

Recommendation: Publish after minor revisions noted.
Minor points:

The abstract does not mention that extended pREST simulations large remove dependence on the initial state. Thus my first impression was that the paper was purely a negative result.

The last sentence of the abstract has been modified to mention pREST. It now reads:

> "Particularly, we highlight the importance of sufficient sampling of protein conformational changes and demonstrate how inclusion of protein residues within the replica exchange region improves sampling."

There are a lot of figures and tables, some of which are redundant. The "RMSD/time" plots does not appear necessary in light of the "Color map" plots. Many of the tables do not appear completely necessary in light of Figures 7 and 8. Perhaps they can be moved to supplemental information.

In response to the reviewer's concerns about redundancy, Tables 2-9 have been moved to the supplementary information since Figure 7 and 8 are graphical representations of these tables. We retained the RMSD/time plots as we believe they help highlight the occurrence of protein/ligand strain in a way which is not captured in the color maps.

Why doesn't Table 8 have the inconsistency analysis from earlier tables?

Inconsistency analysis (RMSI) has been added to Tables 8/9 and their corresponding scatterplot figures (6a/6b)

p. 5, line 6. If simulations start from 2W52 or 4W59, is it necessary mention all the other structures?

We included all the other PDB codes as the additinoal crystal structures were used during Protein/Ligand preparation. Particularly, the other crystal structures were used as references in order to correctly orient the tail when manually building 'ligand 2' involved in the FEP calculation. For example, with benzene to n-butylbenzene, our protocol can be described as follows. We began using PDB: 4W52, which contains the benzene bound crystal structure, duplicated benzene in place, and then added methyl groups until the tail was the same length as n-butylbenzene. At this point, we now have n-butylbenzene where the core ring is directly overlapping with benzene but the tail is oriented somewhat randomly. Then, we used PDB 4W57 (n-butylbenzene bound crystal structure) to orient the tail of our manually created n-butylbenzene ligand similarly to how it is found in PDB 4W57.

p. 9, line 32. Purple

p. 9, line 36-49. I disagree with some wording. I don't see a "contradiction". Coordinate swapping during replica is not an "artifact". It is only contrary to expectation because people are accustomed to looking at single dynamical trajectories.
p. 34, line 37. ref 20 formatting

We corrected these issues, and the paragraph on P.9 line 36-49 and the mention of an apparent "contradiction" has been removed. It now reads:

> "It is important to note that by restraining our analysis to only the end-state replica, we limit our ability to completely view the effects of coordinate swapping during replica exchanges. We address this limitation by analyzing all replicas in what we call 'Color maps', see Figure 6b for an example."

Reviewer: 2
Recommendation: Publish after major revisions noted.
Comments:
This manuscript describes the binding free energy calculations of benzene derivatives to mutant T4 lysozyme using the free energy perturbation technique with replica exchange with solute tempering method. The default applications produced the large inconsistencies but if three key residues in F helix were included into REST region (pREST) with longer simulation, the results showed the tendencies of improvements. The difficulty is the sampling to include the open and closed conformations within F helix part. My honest impression is that the results are a little bit unsatisfactory, but it may mean the difficulty of the problem. Thus, the author should describe the essential difficulty on the current problem in the introduction section.

A paragraph has been added to the introduction (paragraph 3) to better describe the difficulty and challenges in protein sampling.

The author demonstrated if three key residues were considered in pREST, the results improved. These three residues are near a ligand and thus seem to influence the ligand binding free energies. I think that the effects of other residues around these three key residues should be also examined because the interactions of residues with solvent may influence the binding free energies. I also feel that the differences in the inteactions between the key resideues and a ligand without and with the pREST calculations should be discussed.
After reconsiderations of the above points, this manuscript will be acceptable for publication.

A section has been added to the supplementary information to address reviewer 2's request on discussing the difference in protein-ligand interactions between the default protocol and the pREST protocol.

Minor point
Figure numbers should be arranged as the order of the appearances of the figures. For example, Fig 6 appears just after Fig 3 in page 9.

Figure numbers have been re-arranged to be in the correct order.

Reviewer: 3
Recommendation: Publish after minor revisions noted.
Comments:
I have a few minor comments about Figures 1 and 2, which seem to have odd aspect ratios:

1. Figures 1c-e seem to have been squeezed horizontally.
2. Figure 1f seems to have been squeezed vertically.
3. Figure 2a seems to be have been squeezed vertically.

We fixed these issues.