Inconsistencies in MM energy scoring

Only inconsistent scores are mentioned; inconsistencies can be with experiment or across energy functions. Proper rank ordering is the goal. In case of an inconsistent rank ordering, the magnitude of the energy gap between the structures is listed.

**Knowledge transfer:** start by providing RC with paths to all data below (including the filenames for each). RC must have direct access in case of delays.

1. Complexes
2. Ternary (open is preferred in xtal structure 4FVT)

--Open/closed loop complex frame 1 prime (~700 kcal/mol)

--Open/closed complex amber 2-12 ns GBSA/PBSA (~400 kcal/mol)

1. INT/NAM (closed is preferred in xtal structure 4BVG)

--Open/closed loop complex frame 1 prime (~500 kcal/mol)

1. Coproduct (closed is preferred in xtal structure BVH)

-- Open/closed loop complex amber 2-12 ns GBSA/PBSA (>1000 kcal/mol)

1. Binary (from ternary SIRT3 AcCs receptor; open is preferred in xtal structure 4GLS)

-- Open/closed frame 1 prime (~600 kcal/mol)

-- Open/closed amber 2-12 ns GBSA/PBSA (~400 kcal/mol)

1. Apo – see inferred receptor scores per RC’s email; these could also be directly scored

-- Open/closed (derived from ternary) prime first frame; assuming open is preferred, no direct experimental evidence (~600 kcal/mol)

-- Open/closed (derived from ternary) amber 2-12 ns GBSA; assuming open is preferred (~400 kcal/mol)

-- Open/open (derived from ternary vs INT/NAM complexes) prime first frame: ~700 kcal/mol difference depending on the preparation method.

-- Closed/closed (derived from ternary vs INT/NAM complexes) amber 2-12 ns: ~600 kcal/mol difference depending on the preparation method

Inconsistencies were previously mentioned by RC – see email attached herewith for details

-

1. Binding energies
2. Ternary

Open/closed loop, NAD+ ligand, prime first frame (~10 kcal/mol): prime ranks closed as having higher binding energy

1. Side chain optimization

[The energy changes upon side chain optimization below are large enough that they could lead to errors in rank ordering of loops, in the event that they are energy errors]

1. 4FVT\_product complex with 4BVG loop = -11563.234 kcal/mol (Before side chain modelling ie, Grafting of the loop followed by minimization of the complex using OPLS)

4FVT\_product complex with 4BVG loop = -12213.082 kcal/mol (After side chain modeling ie, the best predicted side chain model followed by minimization of the complex using OPLS)

 (~600 kcal/mol reduction upon side chain optimization of nonnative structure)

1. 4FVT\_ternary complex\_4BVG loop = -11361.99 kcal/mol ( Before side chain modelling ie, Grafting of the loop followed by minimization of the complex using OPLS)

4FVT\_ternary complex\_4BVG loop = -12315.09 (After side chain modeling ie, the best predicted side chain model followed by minimization of the complex using OPLS)

 (~1000 kcal/mol reduction upon side chain optimization of nonnative structure)

1. 4BVG native and prime minimized: -9614.9

4BVG side chain predicted and prime minimized: -10175.5

 (~600 kcal/mol energy error)

1. Analysis steps

a) preliminaries. start with all the steps described in RC’s email, including comparison of the effects of side chain optimization vs global minimization on rank ordering; namely,

--Please check whether the first frame energies with Amber all follow same trend as 2-12 ns.

\*\*Please summarize the trends in rank ordering of all complexes and binding energies in a document, with all the different energy functions and for first frame vs 2-12 ns. Note: I believe that in at least one case, Prime also gave the opposite ranking of binding energies compared to Amber (to be verified).

(You may also include in this summary the results of Ping starting from the 4BVG structure since it helps evaluate consistency.)

--With respect to the effects of minimization, do you have energies pre global minimization? Can look at log files for minimization? We may look at these soon to determine whether global minimization is a problem

-- We may consider global (and possibly also local) RMSD calculation for ternary and INT/NAM derived open receptors first frame. As noted in 1E above and the receptor energy attachment these appear to have very different energies despite the fact that they were prepared from the same structures (that were subjected to minimization in presence of the different ligands). This is rather striking and should be investigated.

Compare this RMSD to that before/after side chain optimization for a given complex - which has greater RMSD? Does minimization in the presence of the different ligands or side chain optimization result in larger RMSD with respect to the starting structure?

In this regard may also consider verifying that other loop conformations generated by MD with high RMSD to native are not also ranked incorrectly. I.e., MD can provide more decoys against which we can test scoring function, given claims of very high accuracy of long loop prediction in literature

We need to understand this in order to identify which steps in the structure preparation protocol are inducing the largest energy errors.

-- Comparison to loop prediction validation protocols: Is the global minimization part of typical loop/side chain prediction validation protocol? Look up the papers. Check whether there is a global min after each structure prediction

Is there some other structure prep protocol issue we can identify?
What was the issue encountered by ping for ternary?

b) report the MM component-wise breakdown of energies per your script under development

c) examine the 4FVT and 4BVG validation test sets in detail per the protocol provided, distinguish between sampling and energy errors, and look at per-residue contributions to the energies.

Additional scripting may be required here.

Note that this protocol and Ping's past work aim(ed) to examine the effect of solvent exposure on such errors. There was some evidence of errors for solvent exposed residues.

Note: before doing any parts of c), first provide the RMSD with respect to native for multiple side chain prediction in 4BVG and compare to 4FVT. Did you find 4BVG RMSD was much greater than that for 4FVT?

d) look at per-residue contributions to the energies for the simulated complexes. The per-residue contribution breakdown will also be useful for more detailed analysis of binding energy differences between the complexes (which as mentioned above might become a focus for this paper due to the higher accuracy). Describe to RC the type of data that the script outputs so he can confirm this is what we want to present.

e) determine whether any MD simulations need to be re-run from different starting structures/restraints (restraints can be applied to enforce native-like conformations in case of energy errors, since correction of energy functions is likely not viable for this study…alternatively, energy corrections could be added to the final scores, but this is less accurate; more extensive sampling – e.g., of side chain conformations – can be used to correct sampling errors)

Steps c,d,e) will be discussed in greater detail when we get to them. Parts of c,d) may be reordered (d may be moved before c) depending on the results from a,b).  Note: application of per-residue contribution breakdown for binding energies must be accomplished during month of Aug as a backup strategy for paper. RC will revise paper accordingly to include this and discuss it. If needed it can come before c) – check proposed schedule.

Note: the large energy errors may not be restricted to only 4BVG. See above and the attachment on apo receptor energy estimates, from which appears that even for the open loop (4FVT) receptor

conformations generated by minimization in the presence of two different ligands, there appears to be an energy difference of more than 500 kcal/mol.

Hence I proposed comparing these open loop conformations in more detail (see (a) above).

Protocol for c):

--Side chain prediction validation

-two validation datasets: 4FVT and 4BVG

-present the data in scatterplot format (see below)

- do not need to correct energy errors at this time

**Figure. Side chain prediction validation in SIRT3 A) ternary and B) intermediate complexes: energy and sampling errors.** Plot energy gap with respect to minimized native vs total RMSD. Separation by solvent exposure; highlighting critical binding and catalytic residues. Plot multiple subsets of sidechains as points on the graph (e.g., annotated by radius around a particular residue w/ significant RMSD with respect to native) to separate energy and sampling errors as indicated above; remove sampling errors by reducing size of subset; then individual energy errors can be identified.

Protocol [other approaches have been considered but this may be the simplest to apply here; note we are not manually generating conformations here for purpose of testing the energy fn]:

After settling on the algorithm and algorithmic parameters, present results as follows.

-Carry out multiple sidechain prediction on the full set of residues relevant to loop replacement (or more) in each of the xtal structures (e.g., 4BVG and 4FVT) used for validation studies by PL (also minimize after prediction). Prepare a figure based on this, highlighting any catalytically relevant residues.

-Separately, minimize complete set of residues/sidechains above in the native structure

-Compute RMSD with respect to minimized native for each of the sidechains

-Rank order the sidechains by decreasing RMSD with respect to minimized native

-Select the 5 sidechains with the highest RMSD with respect to minimized native. Verify that some are not solvent-exposed whereas others are.

-Select a radius such that spheres surrounding each of these 5 sidechains are subsets of the complete set of residues subject to sidechain prediction

-For each of the 5 sidechains, successively reduce the radius and repredict only those sidechains within the sphere while maintaining the native conformations of all other sidechains (also minimize after prediction)

-For each of the 5 sidechains, minimize the residues within the spheres starting from the native structures

Then, consider two approaches to plotting the energy gap vs RMSD with respect to minimized native.

1. Plot results pertaining to each of the 5 residues separately [preferred]

-For each of 5 residues, plot the energy gap vs RMSD for the various radii in the same scatterplot (use colors or symbols to code)

 -Compute RMSD over complete set of residues in the largest sphere in each case

-For some residues, may find the energy gap is negative when RMSD is significant, @ several radii: these are energy errors (possibly combined with sampling errors). If this occurs @ small radii, it likely signifies a pure energy error.

-For some residues, may find the energy gap is positive when RMSD is significant, esp @ larger radii: these are sampling errors (possibly combined with energy errors). RMSD is unlikely to be significant @ small radii if energy gap is positive, since sampling errors are uncommon @ small radii. Check if sampling errors are resolved as the radius decreases.

Disadvantage: If plot for just one residue presented, the choice of residue may be arbitrary

1. Plot results pertaining to each of the radii, including all 5 residues, separately

-For each radius, plot the energy gap vs RMSD for the various residues in the same scatterplot (use colors or symbols to code)

-Compute RMSD over the residues in a sphere of the given radius in each case

Disadvantage: RMSDs not directly comparable across residue subsets

May display structures for single, double or triple sidechain predictions which are identified as energy errors

Prediction in nonnative environment (after loop replacement): report the total RMSD with respect to template loop for sidechain prediction after loop replacement, and compare to the RMSD with respect to template loop (minimized native) for sidechain prediction in the template structure without loop replacement. Here, template refers to the structure from which the loop coordinates are obtained.