**Side chain prediction and refinement protocol**

A high level understanding based on Ping’s document

1. Initial preparation of structure for modeling subsequent to grafting/replacement of the loop was done using the “Protein preparation Wizard” of Schrodinger. This is a standard practice to fix any serious errors in the protein like short contacts/clashes, assigning bond orders, protonation state, adding Hydrogens, taking care of amide flips in Asn and Gln.

Ping had used the OPLS force field for minimizing the protein (although the version is not stated), I can make a reasonable guess that he must have used the default OPLS version here.

1. A subset of residues was chosen for side chain prediction. The residues considered for prediction are stated clearly (eg:144-180.195,199,204,207,210,227-234,248,251,291,294,324). In fact, I also see that you have also questioned on the choice of this residues, for which I didn’t find an answer.

But my guess is he must have followed Schrodinger documentation which recommends that any residues with 7.5 Å of a modelled loop region should be refined.

1. He has explored all four available methods for side chain prediction in Prime
2. Default method —No backbone sampling or reorientation of the CA – CB bond is performed.
3. MC approach — Monte-Carlo sampling of side-chain conformations.
4. CA-CB vector sampling— varying the orientation of the CA – CB bond by up to 30° from the initial direction.
5. Backbone sampling—Sample the backbone on a set of 3 residues centered on the residue for which the side chain is being refined.

It’s apparent that the guiding factor for the choice of side chain prediction method was the prime energy function. The prime energy function used here includes the OPLS energy (MM part) with implicit solvent treatment.

However, I see no mention of details relevant to MC approach like no of steps, convergence criteria, temp etc. But, I guess the default options were employed. In fact you have also comment on this issue. The second issue which is not evident from the document is the protocol involved in minimization. The minimization could have been carried out in many different ways

1. A localized minimization (minimizing those residues for which side chains were predicted)
2. A localized minimization for the selected side chains alone (back bone frozen)
3. Minimizing all the side chains keeping the back bone frozen.
4. Minimizing the whole protein but restrain it using a force constant

However, in one document he states “*We also run Prime minimization of the selected residues and used it as the starting point for sidechain prediction and other structural refinement*”.

Assuming that the same protocol was applied for all systems I would be reasonable to say that “prime refinement” was undertaken only on the “selected residues” and not the whole system.

A general conclusion Ping derives based on his study is that MC approach incorporating back bone flexibility offers better results for side chain prediction.

Based on a reading of Ping’s report the protocol which would be followed for side chain prediction and loop refinement would be as follow

1. Prepare 4BVH and 3GLS structure using the “protein preparation wizard”. This includes correcting bond order, assigning prototaion state for the titrabloe amino acids, removing short contacts if any, adding hydrogens and optimizing the orientation of the hydrogen atoms to improve H bonding networks and to correct for the amide flips seen in ASN and GLU.
2. Sequence-structure based superimposition of 4BVH and 3GLS using Schrodinger tool
3. Graft the coordinates for the co-factor loop region from 3GLS on to 4BVH.
4. Run protein preparation wizard on the modelled product complex (open loop conformation).
5. Predict/repack the side chains for all residues within 7.5 Å of the grafted loop region using MC approach together with backbone sampling.
6. Carry out prime energy refinement only on those residues which were repacked keeping the other fixed
7. Subject the complex to MD simulation in NAMD using Amber ff
8. High level QM derived RESP based charges for AADPR is computed using the RED set of programs.
9. The nonstandard residues (AADPR ) was parameterized using the GAFF ( Generalized Amber ff) force filed using the tleap program available form Amber tools14.
10. MM/PBSA and MM/GBSA based relative conformational energies of the complexes and the relative binding affinity of the product AADPR to Sirt3 in closed and open loop conformations will be evaluated

Relevant files (These files had been shared earlier via drop box and has also been uploaded on the wiki space)

Ligands.docx

Report\_on\_SIRT3\_structures&loop\_conformation\_04-24-2015.docx

side chain optimization studies-05142015.docx Seem to be missing validation reports: “I will later summarize in detail the  ligand placement, validation of side chain prediction methods on a limited dataset, and loop refinement in a separate document .” Please do so.

Sidechain prediction for 4BVG with loop replacement on residue 155-178.docx (most relevant)

Sidechain prediction for 4FVT-INT-NAM with loop replacement on residue 155-178.docx (most relevant)