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)

**Location of the Prime frame 1 data (Ternary-4FVT):**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING

**The necessary input files (receptor) for running Prime MM/GBSA calculations are**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING\4FVT\_xtal\_Prime\_inputframe\ 4FVT\_RECP\_Ac\_Pep.pdb

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING\4FVT\_xtal\_Prime\_inputframe\ 4FVT\_RECP\_NAD+.pdb

**The necessary input files (Ligand) for running Prime MM/GBSA calculations are**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING\4FVT\_xtal\_Prime\_inputframe\Ac\_Pep.mae

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING\4FVT\_xtal\_Prime\_inputframe\NAD+.mae

**Prime MM/GBSA calculation output data/files location**

**Prime MM/GBSA-:NAD+ ligand**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING\4FVT\_xtal\_Prime\_inputframe\NAD+\ prime\_mmgbsa\_4FVT\_NAD-out.maegz

**Prime MM/GBSA-:Ac\_peptide ligand**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_PING\4FVT\_xtal\_Prime\_inputframe\Ac\_Pep\ prime\_mmgbsa\_4FVT\_Ac\_Pep-out.maegz

**Location of the Prime frame 1 data (Ternary-4BVG):**

**Prime MM/GBSA Output files location**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1

**The necessary input files (receptor) for running Prime MM/GBSA calculations are**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1\4FVT\_ternary\_4BVGloop.comp\_Ac\_CS2.pdb

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1\4FVT\_ternary\_4BVGloop.comp\_NAD+.pdb

**The necessary input files (Ligand) for Prime MM/GBSA calculations are**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1\ Ac\_CS2.maegz

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1\NAD+.maegz

**Prime MM/GBSA calculation output data/files location**

**Prime MM/GBSA-:NAD+ (Ligand)**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1\MMGBSA\_NAD\prime\_mmgbsa\_NAD\_4FVT\_4BVG-out.maegz

**Prime MM/GBSA-: Ac\_Cs2peptide (Ligand)**

C:\Users\vramaswamy\Documents\MDwork\MD\_4FVT\_tern\_4BVGloop\model\Frame1\MMGBSA\_Ac\_Cs2\ prime\_mmgbsa\_Ac\_Cs2-out.maegz

Under each folder (MMGBSA\_Ac\_Cs2 and MMGBSA\_NAD) the following output files will be present

1. Prime\_mmgbsa\_\*\*\*\*\*\*\*-out.maegz (prime MM/GBSA output file that can be opened in Maestro)
2. prime\_mmgbsa\_\*\*\*\*\*\*\*\*.Prime.log (prime MM/GBSA output file that can be opened in any editing tool)
3. prime\_mmgbsa\_\*\*\*\*\*\*-out.csv (prime MM/GBSA output file that can be opened in an xls sheet)

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

**Ternary-4FVT loop (Simulation done by Ping)**

Partial analysis done by Vijayan

Location: pmcat-gpu1 (IP: 192.168.100.177 port 22)

Path: /home2/plin/work/project01/MD\_4FVT\_tern/mmpbsa\_rec2lig2\_md\_pbgbsa

**The scripts for executing the MM/PBSA and MM/GBSA calculation in Amber for the 1-12 ns trajectory is**

run\_script\_11ns\_Deac\_pep.sh

run\_script\_11ns\_AADPR.sh

**The scripts which I have written for extracting the energies are**

mining\_peptide.sh

mining\_AADPR.sh

These shell scripts call the Perl scripts (miner2.pl and miner.pl) and recursively iterates over all the folders (RERUN\_step.2170, RERUN\_step.4170….) and extracts the energy components from an Amber MM/GBSA run.

The raw 0-12 ns MD trajectory is located at

/home2/plin/work/project01/MD\_4FVT\_tern

The dcd trajectory file name is **4FVT\_v1a\_tern.dcd**

**Ternary-4BVG loop (Simulation done by Vijayan)**

The raw 0-12 ns trajectory is located at

**/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN/**

The dcd trajectory file name is **4FVT\_ternary\_4BVGloop\_11nsMD.dcd**

**The scripts for executing the 1-12 ns MM/GBSA and MM/PBSA calculation on Amber are**

run\_script\_11ns\_Ac\_pep.sh ( ie for extracting the binding affinity of AC-CS2 peptide)

run\_script\_11ns\_NAD+.sh ( ie for extracting the binding affinity of NAD+)

**The scripts for executing the MM/PBSA and MM/GBSA calculation on the input frame energies are Amber are**

run\_script\_Frame1\_Ac\_pep.sh

run\_script\_Frame1\_NAD+.sh

These scripts calculate MM/GBSA and MM/PBSA based on 1ns time interval.

For each time interval sub-folders are created

Each sub-folder contains a \*.dat file which contain the energies and the summary of the MM/PBSA and MM/GBSA run.

Please note that there will be only one \*.dat files under each of these output folders.

The respective output files from MM/GBSA and MM/PBSA energies for the Input frames are located under the following

**/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN/**Input\_AC\_pep

**/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN/**Input\_NAD+

The \*.dat file under these folder contain the energies and the summary of the MM/PBSA and MM/GBSA run.

Please note that there will be only one \*.dat files under each of these output folders.

I wrote a couple of scripts to extract the energies of NAD+ and peptide which are located at

/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN /mining\_NAD+.sh

/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN /mining\_peptide.sh

These scripts in turn call two Perl scripts to extract the necessary energy component form the\*. dat file.

/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN/ miner.pl

/home2/plin/work/project01/MD\_4FVT\_ternary\_4BVGloop\_model/MDRUN/ miner2.pl

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

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

The respective MM/GBSA and MM/PBSA output energies based on the Input frames are located under the following directories

**Output files/folders**

C:\Users\vramaswamy\Documents\MDwork\SIRT3\_INT\_NAM\_4FVT\_4BVGloop\Frame1\ MMGBSA\_INT\

C:\Users\vramaswamy\Documents\MDwork\SIRT3\_INT\_NAM\_4FVT\_4BVGloop\Frame1\MMGBSA\_NAM\

Under each folder there will be the following files

Prime\_mmgbsa\_11-out.maegz (prime MM/GBSA output file that can be opened in Maestro)

prime\_mmgbsa\_11.Prime.log (prime MM/GBSA output file that can be opened in any editing tool)

prime\_mmgbsa\_11-out.csv (prime MM/GBSA output file that can be opened in an xls sheet)

**Input Ligands**

C:\Users\vramaswamy\Documents\MDwork\SIRT3\_INT\_NAM\_4FVT\_4BVGloop\Frame1\ INT.maegz

C:\Users\vramaswamy\Documents\MDwork\SIRT3\_INT\_NAM\_4FVT\_4BVGloop\Frame1\NAM.maegz

**Input Receptor**

C:\Users\vramaswamy\Documents\MDwork\SIRT3\_INT\_NAM\_4FVT\_4BVGloop\Frame1\RECP\_INT.pdb

C:\Users\vramaswamy\Documents\MDwork\SIRT3\_INT\_NAM\_4FVT\_4BVGloop\Frame1\RECP\_NAM.pdb

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

**RSK: The coproduct simulation which I did was for co-product was with 4FVT and 4FVT grafted with 4BVG loop.**

**Coproduct-4FVT loop (Simulation done by Vijayan)**

Location: pmcat-gpu1 (IP: 192.168.100.177 port 22)

Path: /home2/plin/work/project01/MD\_4FVT\_procuct\_complex\_NATIVE/MDRUN

**The scripts for executing the 0-12 ns MM/GBSA and MM/PBSA calculation on Amber are**

run\_script\_11ns\_Deac\_pep.sh ( ie for extracting the binding affinity of Deacetylated peptide)

run\_script\_11ns\_AADPR.sh (ie for extracting the binding affinity of AADP ribose)

The scripts which I have written for extracting the energies are

mining\_AADPR.sh

mining\_peptide.sh

These shell scripts call the Perl scripts (miner2.pl and miner.pl) and recursively iterates over all the folders ( 2-3ns , 3-4ns ….) to extract the energy components from an Amber MM/GBSA run.

The raw 0-12 ns MD trajectory is located at

/home2/plin/work/project01/MD\_4FVT\_procuct\_complex\_NATIVE/MDRUN

The dcd trajectory file name is **4FVT\_2OADPr\_deac\_11ns\_md.dcd**

**Coproduct -4BVG loop (Simulation done by Vijayan)**

The raw 0-12 ns trajectory is located at

**/home2/plin/work/project01/MD\_4FVT\_product\_4BVGloop/MDRUN**

The dcd trajectory file name is **4FVT\_2OAADPR\_DeACpep\_4BVG\_loop\_11nsMD.dcd**

**The scripts for executing the 1-12 ns MM/GBSA and MM/PBSA calculation on Amber are**

run\_script\_11ns\_AADPR.sh ( ie for extracting the binding affinity of AC-CS2 peptide)

run\_script\_11ns\_Deac\_pep.sh ( ie for extracting the binding affinity of NAD+)

These scripts calculate MM/GBSA and MM/PBSA based on 1ns time interval.

For each time interval sub-folders are created

Each sub-folder contains a \*.dat file which contain the energies and the summary of the MM/PBSA and MM/GBSA run.

Please note that there will be only one \*.dat files under each of these output folders.

I wrote a couple of scripts to extract the energies of NAD+ and peptide which are located at

 /home2/plin/work/project01/MD\_4FVT\_product\_4BVGloop/MDRUN/ mining\_AADPR.sh

/home2/plin/work/project01/MD\_4FVT\_product\_4BVGloop/MDRUN/ mining\_peptide.sh

These scripts in turn call two Perl scripts to extract the necessary energy component from the\*. dat file.

/home2/plin/work/project01/MD\_4FVT\_product\_4BVGloop/MDRUN/ miner.pl

/home2/plin/work/project01/MD\_4FVT\_product\_4BVGloop/MDRUN/ miner2.pl

The output folders are step\_RERUN\_AADPR.1001 ……. For AADPr as the Ligand

And step\_RERUN\_DeacPep.1001 ……… for Deac\_peptide as the Ligand

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

**RSK: Although I am not sure, but I see that the energy of the INPUT frame as per Amber is highly positive. That’s the reason I have flagged it.**

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

**RSK: I believe that we have not done any simulation with 4GLS loop. Could you please make it clear?**

-- 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