SAMPL5 Host-Guest Instructions

September 14, 2015

Due date

Your predictions must be uploaded on the D3R SAMPL5 web-page by February 1, 2016. The experimental results will be released immediately after the challenge closes.

Anonymous versus public participation

When you signed up for this challenge, you were given the option of participating anonymously. Anonymous participation means that we may report on your predictions and methods, but not your identity. Public participations means we may also say who you are. Please note that, although we will work to protect the identity of anonymous participants, we cannot make any guarantees.

You may use the D3R website to change your anonymous/public status until the challenge has closed. However, after the challenge has closed, you may not change your anonymous/public status.

SAMPL5 workshop March 9-12

Participants are invited to share and discuss their results, as well as the D3R and SAMPL projects more broadly, at the first D3R and SAMPL workshop, which is scheduled for March 9-13, 2016, at UC San Diego, La Jolla, CA. Note that the workshop immediately precedes the ACS National Meeting in San Diego, whose theme is *Computers in Chemistry*.

Molecular systems

Three host-guest series are provided.

The first two comprise hosts OctaAcidH and OctaAcidMe, each with the same set of 6 guests. Measured binding free energies are available for all 12 cases, and binding enthalpies are available for 11 of the 12.

The third set comprises host CBClip, with a different set of 10 guests. Because experiments are still being done, we are providing 7 of the guests now as a starting point, and will provide the other 3 likely by the end of September.

Computational methods

You may use any method(s) you like to generate your predictions; e.g., MD with implicit or explicit solvent; quantum methods; docking; etc.

Files provided

All hosts and guests are provided in mol2, PDB and SDfile formats, kindly prepared by Dr. Jane Yin (UC

San Diego). There is no guarantee or representation that the protonation and tautomer states provided are optimal. It is also possible that the protonation state of a guest or host will change on binding.

The organizers plan to carry out a complete set of binding free energy and enthalpy calculations in AMBER, using GAFF/RESP force field parameters with TIP3P water. In case you are interested in using the same force field for your calculations, Dr. Jane Yin (UC San Diego) has generated and provided a full set of AMBER simulation files, and Dr. Michael Shirts (U Colorado), working with Drs. Yin, David Mobley (UC Irvine), and Christoph Klein (Vanderbilt U.) have generously converted the original AMBER-format files into matched files for LAMPPS, GROMACS and Desmond, and have checked for consistent energies across these software packages. We are also looking into the possibility of providing CHARMM-format files, but they are not available in the initial download. The following initial setups are provided in the download package:

AMBER: .rst and .prmtop
LAMPPS: .input and .lmp
GROMACS: .gro and .top
Desmond: .cms

Details of the energy comparison are provided in the file SAMPL5Energies.xls. In brief, for LAMPPS, GROMACS and Desmond, the energies for the starting conformation agree to within a relative error of approximately 10^{-6}, given matched nonbonded and other parameters. For AMBER, similar agreement is obtained for most terms, but there are discrepancies, with relative error approximately 10^{-5}, for energies that include an electrostatic contribution. These discrepancies appear to result from a deviation, relative to the other three programs, at the fifth significant digit of Coulomb's constant converting q_1q_2/r to energy in kcal/mol. We hope to fully clarify this situation soon.

Standardization runs

If you are using md with explicit solvent to compute binding affinities, **please apply your method to the standard setups (see above) for the following two host-guest systems**, so that we will all have a common point of comparison across packages and methods:

OctaAcidH with guest G3
OctaAcidH with guest G4

Hopefully, we will all get the same results to within numerical precision. If not, we can work together to reconcile the discrepancies.

Uploading your predictions

D3R is currently outfitting the SAMPL5 page with the ability to accept your uploaded predictions, and we are also preparing a template file to be used in submitting predictions. As soon as these are ready, you may upload your predictions. If you want to upload more than one set of predictions, generated by different methods, each set must be uploaded as a separate file. Please use the template provided, as the predictions will be parsed and analyzed with automated scripts.

A complete set of predictions constitutes predicted binding free energies for all host-guest pairs, with predicted numerical uncertainties. Optionally, predictions may also include binding enthalpies and corresponding uncertainties, for the cases where these are available from experiment; i.e., most of the

octa-acid cases. Incomplete submissions - such as for a subset of compounds - will also be accepted, but will not necessarily be evaluated together with the rest of the submissions.

Pending items, error reports, questions

We will email you during the challenge about the following pending items:

- Three more guests for CBClip, to make a total of 10
- Template for uploading your predictions
- Any updates regarding the organic/aqueous partition coefficient component of the challenge
- Workshop details

Please feel free to contact us if you notice any errors in the information provided or have questions about SAMPL5: samplchallenge@gmail.com