SAMPL5 Host-Guest Systems, Measurements and Starting Files

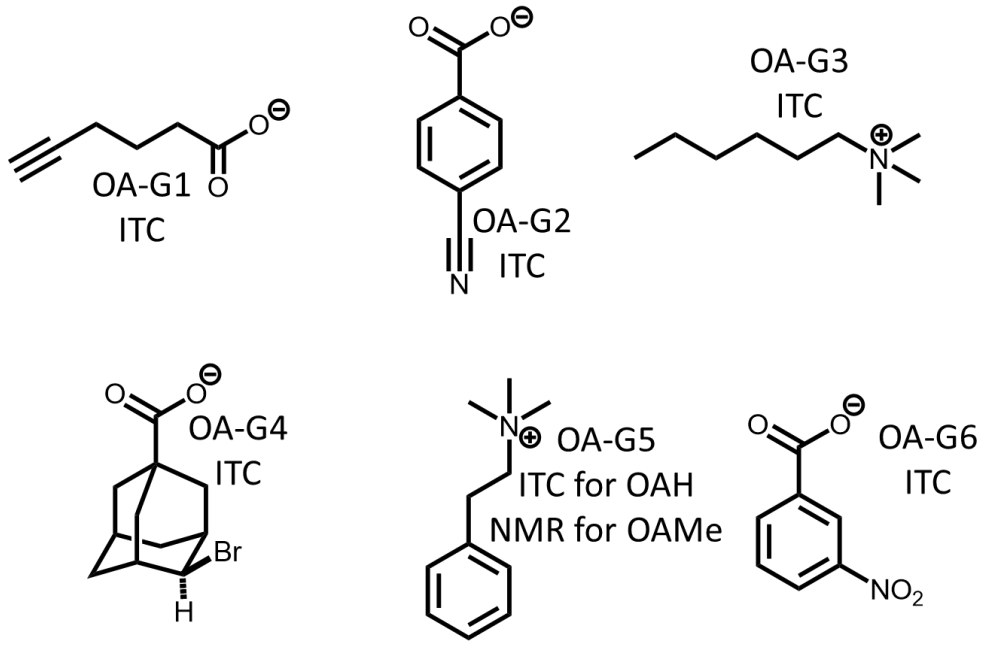
This challenge is based on three different hosts, which we are calling OctaAcidH (OAH), OctaAcidMe (OAMe) and CBClip, in aqueous solution. The measurements for these systems were generously carried out specifically for SAMPL5 by Professors Bruce Gibb (Tulane University; OctaAcids) and Lyle Isaacs (U. Maryland; CBClip). Data will be provided for one set of 6 guests with both OctaAcids, for a total of 12 measurements; and for a different set of 10 guests for CBClip.

The following subsections describe these systems, the experimental conditions and measurements, the quantities to be predicted, and how the molecule and simulation files provided in the download were set up.

OctaAcids

Two of the host molecules, OctaAcidH and OctaAcidMe, are closely related to each other, as the Me version simply has four methyl groups in place of four hydrogens. This change modifies the shape of the binding cavity. Host OctaAcidMe is described here: doi:10.1021/ja200633d and host OctaAcidH is described here doi:10.1007/s10822-013-9690-2. There are also a number of papers from SAMPL4 which discuss calculations for this system, as summarized here: doi:10.1007/s10822-014-9735-1. OctaAcidH was used in the previous SAMPL4 challenge, but with a different set of guest molecules.

For the present SAMPL5 challenge, experimental binding data were obtained for the association of the following set of six guest molecules, OA-G1 to OA-G6, to both OctaAcidH and OctaAcidMe:

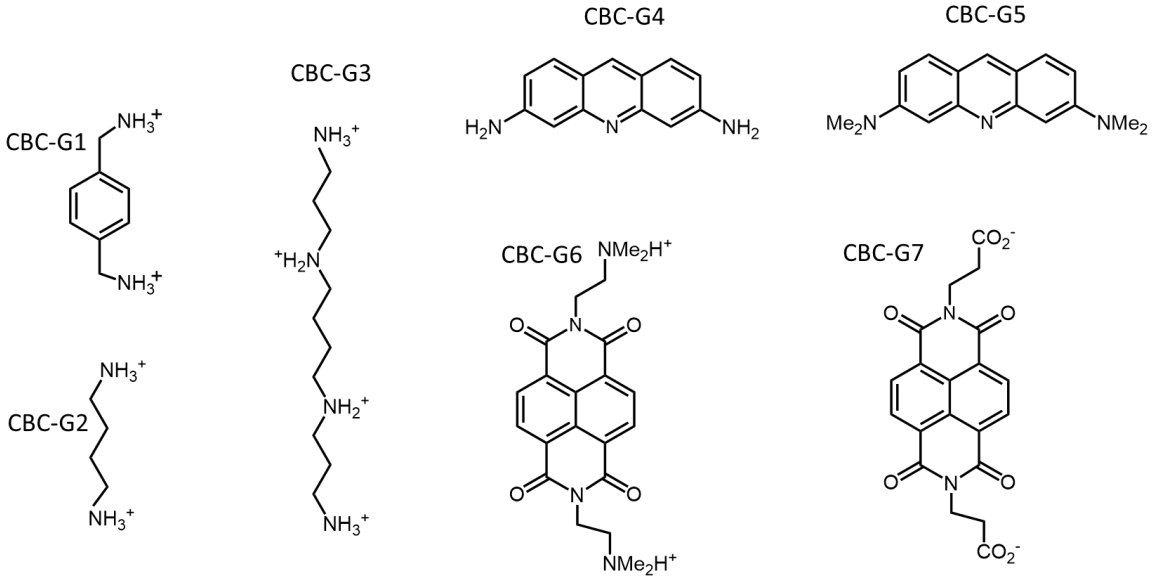


Binding affinities (Ka) values will be available for all of these, measured by either isothermal titration calorimetry (ITC) or NMR, as indicated, and binding enthalpies also will be available for the studied by ITC, but not the ones studied by NMR, as noted in the graphic. The main challenge is to compute 12 binding free energies (or association constants), and participants are also invited to compute the binding enthalpies. All measurements were done in aqueous 10mM sodium phosphate buffer at pH 11.5, at

298K, except for OA-G6 (nitrobenzoic acid), for which the buffer was 50mM sodium phosphate at pH 11.5.

CBClip

The CBClip host is a new "molecular clip" that is chemically related to the cucurbituril class of synthetic hosts. A paper describing this family of clips may be found here: doi:10.1039/c5ob00184f doi:10.1021/jm501276u, and the clip used in SAMPL5 is 2b (R = $(CH_2)_3SO3_Na$) in this paper. The SAMPL5 challenge is to compute the binding free energies (association constants) of ten guest molecules to this host. *Prof. Isaacs is still refining the challenge set, so the initial download has seven of the ten guests as a starting point, and the remaining three should be identified within a few weeks. We will email you information about them as soon as possible.* Here are the initial seven:



All CBClip measurements were done by NMR in 20 mM sodium phosphate buffer at pH 7.4, at a temperature of 298K.

Molecule and simulation setups

An idealized model of the 3D structure of host CBClip was provided by Prof. Lyle Isaacs. The starting 3D structures of hosts OAH and OAMe were built manually with MOE and mildly energy minimized. The initial structures of free guest molecules were found via the conformational search feature in MOE. For CBClip, the experimental studies were done at pH 7.4, so the host's four sulfonic acid acids were treated as fully ionized, for a net host charge of -4. For the octa-acids, OAH and OAMe, the experimental studies were done at pH 11.5, so all eight carboxylic acids were treated as fully ionized, for a net charge of -8. The protonation states of all guest molecules were also assigned based on their expected pKas and the experimental pH values, as shown in the figures above, and were assumed to remain the same on binding their respective hosts. The resulting structures are provided in the download as PDB, mol2 and SDfiles. Please note that all of the structures provided here are meant only as reasonable starting points. There is no guarantee that they are optimal. Please take full advantage of your chemical

knowledge and intuition in considering this and all other aspects of the challenge, and feel free to use computational tools beyond those used for the actual binding calculations.

As noted in the instructions, the organizers plan to carry out a complete set of binding free energy and enthalpy calculations with GAFF/RESP force field parameters and the TIP3P (non-modified) water model, and our starting files for the solvated bound complexes are being provided as part of the downloads for this challenge. The starting files were set up as follows. Partial charges were generated with the restrained electrostatic potential (RESP) fitting procedure, with electrostatic potentials at the HF/6-31G* level. Bonded and van der Waals parameters were assigned by GAFF in Antechamber. For the octa-acids, which have a well-defined hydrophobic cavity, the bound conformations were set up with the guests' ionized groups at the opening and their hydrophobic parts in the cavity. For the CBClip, the positions of the guests in the binding site were constructed by manual docking in MOE.

Each bound complex was placed in a cubic simulation box with edge length of about 40Å, and containing 2100 TIP3P (non-modified) water molecules. Sodium ions were added to neutralize all systems. Given the low concentration of the buffer used in experiments, no extra ions were added to mimic the buffer. The sodium ions were modeled with the TIP3P-specific sodium parameters of Joung and Cheatham (J Phys Chem B. 2008, 112: 9020–9041). Each system was then subjected to a short equilibration phase (first NVT, then NPT) to heat the system and change the volume so the pressure would be about 1 atm. The solutes were restrained during the entire process, so further equilibration will be needed to relax them. Note that the AMBER .rst7 files contain velocities appropriate to 298.15 K.

The AMBER format files generated by these procedures were then converted to GROMACS, Desmond and LAMPPs formats with custom scripts and the energies of the starting conformations were compared. The results are summarized in the Instructions and detailed in the file SAMPL5Energies.xls.