Benchmarking Accuracy Breakout session

Considerations

Several axes to consider: training vs testing; physical properties vs applications

What physical properties can we converge that will inform performance on applications that are hard to converge, where accuracy may depend on method?

Important to avoid properties where there are changes in polarization (if FF isn't polarizable) or significant quantum corrections are required

OK if we include some benchmarking properties where we *expect* classical force fields to fail because it needs higher-order terms

How accurately do we need to reproduce various physical properties?

Benchmarking (and possibly training)

Binding free energies (host:guest, benchmark protein:ligand sets that can be converged)

Liquid densities (neat and mixtures)

Small molecule crystal structures

Compressibilites

Relative solubilities (different compounds in water, different solvents)

Infinite dilution activity coefficients

Partition coefficients

Speed of sound

Heat capacity (liquids and solids? QM issues?)

Liquid structure factors (neutron or X-ray)

Phase change data (melting points?
difficulties of knowing crystal form is relevant)

Lipid partitioning or surfactant partitioning?

Applications

Binding free energies (protein:ligand)
assess rank ordering correct
industry contributions
corrected version of Schrödinger JACS set

Unbinding kinetics / residence times / k_off

Coarse-graining from atomistic simulations

Relative solubilities (small molecules, surfactants, polymers)

Polymorph prediction

Solubility?

Kinetics?

Open questions

Good data sources for charged/ionic molecules?
(osmotic pressures from salt solutions; Kirkwood-Buff data?)

What good measurements of small molecule / liquid interactions might exist?

What about kinetics? Should we include properties relevant to this?
(diffusion, viscosities)

Should high-level QM be considered? Congested ring systems

Should we leave some QM data out during fitting for validation/assessment?

TODO

Talk to **Advisory Board** about soliciting industry contributions of public binding affinity measurements for benchmarking protein-ligand free energy calculations