HIGH PERFORMANCE COMPUTING ON CRESCO INFRASTRUCTURE: RESEARCH ACTIVITY AND RESULTS 2019

November 2020

Contents

Draft print - This index has been temporary inserted in order to jump easily accross the document

NONEQUILIBRIUM ALCHEMY IN SAMPL7 CHALLENGE

Piero Procacci[1]*, Guido Guarnieri[2]

[1]*Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, I-50019 Sesto Fiorentino, Italy*

[2]*ENEA, Portici Research Centre, DTE-ICT-HPC, P.le E. Fermi, 1, I-80055 Portici (NA) (Italy)*

ABSTRACT. In the context of the SAMPL7 challenge, we computed, by means of a non-equilibrium (NE) alchemical technique, the standard binding free energy of two series of host-guest systems, involving as a host the Isaac's TrimerTrip, a Cucurbituril-like open cavitand, and the beta-Cyclodextrin derivatives designed by Gilson and coworkers. The adopted NE alchemy combines enhanced sampling molecular dynamics simulations with driven fast out-of-equilibrium alchemical trajectories to recover the free energy via NE theorems. Performances were good, confirming the reliability of the computational approach and exposing, in some cases, some important deficiencies of the GAFF2 non-polarizable force field.

1 Introduction

The SAMPL initiative [1, 2, 3]periodically proposes community-wide blind challenges aimed at assessing computational techniques as standard predictive tools in rational drug design. The SAMPL systems generally consists of a series of host-guest pairs for which the standard binding free energy must be predicted, given the chemical structure of the partners and the experimental conditions (pH, temperature and pressure) used in the measurements. In the SAMPL7 challenge,[4] the organizers included three host systems, namely the Triptycene walled glycoluril trimer (codename TrimerTrip), various mono-3-substituted β-cyclodextrin analogues(codename CD), and the Gibb Deep Cavity Cavitands or Octa-acids (codename GDCC). In SAMPL7, only *one* ranked submission for each of the host-guest systems was allowed. This strict rule was introduced to avoid the practice, largely adopted in the previous SAMPL6 challenge, of filing multiple submissions with small variants in the hope of hitting the target with "multiple shots on goal".

In this study we present our ranked predictions for the TrimerTrip and CD host-guest systems, done using a classical molecular dynamics (MD) approach, based on enhanced sampling (ES) of the fully coupled end-states followed by a swarm of nonequilibrium alchemical trajectories with fast switching (FS) of the guest-environment interaction by way of a driven alchemical coupling parameter. From the resulting work distribution, the free energy is recovered exploiting the well known Crooks or Jarzynski NE theorems. The adopted computational protocol is identical to that described in Ref. [5], termed as fast-switching double annihilation method (FSDAM). FSDAM is tailored for CPU-based HPC systems such as CRESCO6[6], relying on an efficient OpenMP/MPI hybrid algorithm whereby the nonequilibrum FS trajectories or equilibrium ES trajectories (each parallelized on the OpenMP layer using a force decomposition strong scaling scheme) are assigned to the MPI layer.

*Corresponding author. E-mail: procacci@unifi.it.

2 Methods

Structural details of the host-guest systems for the TrimerTrip ad CD SAMPL7 challenges can be found at Ref. [4]. All simulations were performed with the ORAC6.1 program[7] using FSDAM as described in Ref. [5] Briefly, in FSDAM the dissociation free energy is determined in two consecutive stages: in the first stage, the bound and unbound state are sampled at full coupling using Replica Exchange with Solute Tempering (REST).[7] In the second stage (fast switching, FS), starting from the canonical sampling at full coupling, we launch a swarm of independent and concurrent non-equilibrium trajectories where the ligand-environment alchemical coupling parameter is rapidly switched off to zero coupling (ligand in the gas phase). The annihilation free energy in the two branches of the alchemical thermodynamic cycle is recovered from the work distribution exploiting the Crooks and Jarzynski theorems.

The Force Field (FF) parameters of the host and guests molecules were prepared using the Primadorac interface.[8] Each system was solvated in about 1000 OPC3[9] water molecules in a cubic box of $\simeq$ 31 side-length.

Production MD simulations (both REST and FS) were run in the isothermal-isobaric ensemble, with temperature control (T=298 K) using Nose-Hoover thermostat and pressure control (P=1 atm) as described in Ref. [10] Lennard-Jones non-bonded interactions were truncated with a 13.0 Angstrom cutoff, whereas long-range electrostatics were handled with the PME[11] method using $\alpha = 0.37$ Å^{-1}, 1 Å spacing for the gridded charge array and a 4-th order B-spline interpolation. SHAKE constraints were applied to bonds involving hydrogen atoms, and the simulation was integrated using a five step RESPA integrator.[10]

In the REST stage, we scaled the torsional ad 14 non bonded interactions of the solute (host+guest) up to 0.1 (i.e. 3000 K) using eight replicas with the scaling protocol described in Ref. [7]. Eight replicates of REST simulations were simultaneously launched on 64 MPI processors for 4 ns on each state, sampling 480 *target state* configurations for the bound host-guest system and for the free guest in bulk. In the bound state, a weak COM-COM harmonic tethering potential (k=0.052 kcal mol^{-1} Angs^{-2}) was imposed to prevent the guest to drifting off in the solvent. The CRESCO6 job for the REST stage involved ten nodes running 64 MPI process each using 6 threads on the OpenMP layer for a total of 384 corees. A REST job for a host-guest pair was completed in about 4 hours.

In the FS stage, the guest, in the bound and unbound state, was annihilated in 0.36 and 0.24 ns, respectively, in 480 independent trajectories (corresponding to 480 MPI processes) starting form the corresponding points sampled in the REST stage. The bound and unbound annihilation protocol stipulates that the guest atomic charges are first switched off, followed by the annihilation of of Lennard-Jones (LJ) interactions. Each of the 480 MPI processes used 6 OpenMP threads, for a total of 2880 cores. A FS job for a host-guest pair was completed in less than 20 minutes.

Annihilation free energy estimates (bound and unbound states) are based on the work distribution (WD) produced in the FS stage. If the WD passed the Anderson-Darling (AD) *and* the Jarque-Bera (JB) normality tests, the annihilation free energy is calculated using the Gaussian estimator.[5] Otherwise, the statistically boosted Jarzynski average is used, exploiting the decorrelation between discharging work and Lennard-Jones annihilation work.[5] Error on annihilation free energy estimates was computed by bootstrapping with re-sampling in all cases. A finite size correction to the dissociation free energy due to net charges on the ligand has been calculated as described in Ref. [5]. The standard state correction to the dissociation free energy for translational restraint is given by $\Delta G_{\rm ssc} = RT \ln(V_{\rm site}/V_0)$, where $V_{\rm site}$ is the binding site volume.[12] $V_{\rm site}$ is computed from the variance of the guest-host COM-COM

distances monitored during the REST stage as $V_{\text{site}} = 4\pi(2\sigma)^3/3$.

3 Results

Three ranked submissions for the TrimerTrip were filed, namely the present submission (GAFF2/FSDAM), a prediction set using the AMOEBA polarizable force field and FEP, and a submission using a mixed approach, with sampling via standard MD and binding free energy calculation using the semi-empirical tight-binding xtb-GNF program develop by S. Grimme. Correlation plots for the ranked submissions are reported in Figure 1(a) and quality metrics of the corresponding predictions are shown in Table 1. In the Figure 1(a) and in the Table 1 we also report an (unsubmitted) calculation of binding free energies using Autodock4,[13] assuming full flexibility of the ligand and rigidity of the host (using the conformation provided by the organizers at Ref. [4]). Outliers, differing by more than 4 kcal/mol with respect to the experimental value, are marked in red color. While AMEOBA/FEP appears to be the best correlated set to the experimental measures according to Pearson coefficient and to the Kendall rank coefficient τ, mean unsigned errors (MUE) are surprisingly minimal for the Autodock set, with GAFF2/FSDAM and AMOEBA/FEP exhibiting similar MUEs. The prediction set based on a mixed MD/QM (semi-empirical) approach is consistently the worst for all quality metrics. GAFF2/FSDAM and AMOEBA/FEP do not have outliers in common. Very likely, discrepancies in g18 and g10 for GAFF2/FSDAM should be ascribed to force field deficiencies, related to the fixed charge approach. In the case of g18, the AM1/BCC charges could underestimate the polarization induced by the host's carboxy groups in the bound state, leading to an extra charge accumulation on the aromatic nitrogen. The neglect of this likely polarization effect can lead to underestimation of the electrostatic contribution to the g18 decoupling in the bound state and hence to the binding affinity. For the diamondoid derivative g10, each amino hydrogen bears a fixed charge of 0.31 e, probably leading in this case to a systematic overestimation of the carboxy-amino electrostatic interactions in the bound state and hence to an overestimation of the binding affinity. Purging g10 and g18 from the GAFF2/FSDAM data-set results in $R = 0.67$ and $\text{MUE} = 1.61$.

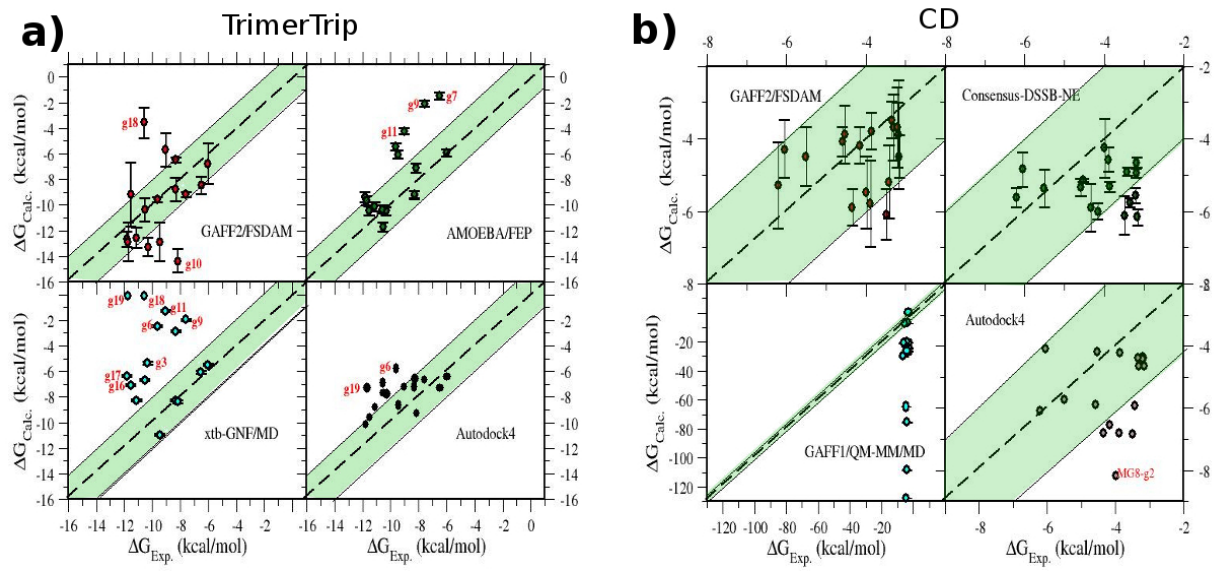

Figure 1: a) TrimerTrip and b) CD correlation plots. Data in the shaded region are within 2.0 kcal/mol of experimental counterpart

In Figure 1(b) we show the correlation plots for the three ranked submissions in the CD host-guest challenge, namely that based on our GAFF2/FSDAM approach, a prediction set relying again on NE alchemical technology using the so-called double-system-single-box approach[14] (DSSB) and finally predicted binding free energies using a QM/MM approach. Also in this case, we provide the plot corresponding to the Autodock4 unsubmitted prediction. The experimental data are clustered in a range of less than 3 kcal/mol and both GAFF2/FSDAM and DSSB correctly and remarkably predicts binding affinities within approximately the same range with no outlier. Given this small experimental range, and given that the systematic uncertainty in fixed charges force fields for solvation free energies are of the order of 2 kcal/mol,[3] MUE appears to be the most meaningful metrics, with GAFF2/FSDAM resulting the best performing method. While the QM/MM ranked prediction set is, quite expectedly,[3, 2] totally off-the-mark (note the expanded scale in the left-bottom correlation plot of Figure 1(b)), the Autodock4 calculation is again in decent agreement with the experimental data, performing only slightly worse than the MD-based NE approaches do (see Table 1).

			TrimerTrip		
Method	R_{xy}	a	b	MUE	τ
Autodock	0.50	0.35	-4.45	2.00	0.38
FEP/AMOEBA	0.71	1.24	3.94	2.10	0.47
GAFF2/FSDAM	0.35	0.61	-4.06	2.23	0.23
xtb-GNF/MD	-0.06	-0.10	-6.07	4.49	-0.05

			CD		
Method	R^2	a	b	MUE	τ
GAFF2/FSDAM	0.19	0.17	-3.87	1.01	0.22
DSSB/NE	0.13	0.13	-4.61	1.43	0.02
Autodock	0.13	0.17	-4.78	1.66	0.07
QM/MM	0.10	3.68	-20.36	32.00	0.22

Table 1: Salient data metrics for the assessment of the ranked submissions and of the Autodock prediction set (not submitted) in the TrimerTrip and CD SAMPL7 challenges: R^2: Pearson's coefficient; a: slope of the regression line; b; intercept of the regression line; MUE: mean unsigned error; τ: Kendall's rank coefficient.

4 Conclusion

In this contribution, we have presented our ranked prediction sets for the TrimerTrip and CD systems in the context of the latest SAMPL7 challenge, using FSDAM, a nonequilibrium alchemical approach combined with enhanced sampling end-state simulations. The performances of our MD-based technique, that uses conventional fixed-charge force fields, is in line with our previous submissions in the SAMPL6 host-guest challenge[3] (done using the very same technology) yielding binding free energies estimates within 2 kcal/mol in most of the cases. Outliers are rare and likely to be ascribed to structural deficiencies of the force field due to the neglect of important polarization effects in the anisotropic environment of some host-guest complexes.

Concerning specifically the SAMPL7 challenge ranked submissions, in the TrimerTrip only one much more computationally demanding approach (based on a polarizable force field) did better than FSDAM. We finally must honestly point out the surprising good results for both the TrimerTrip and CD obtained

using a simple and inexpensive docking approach. While MD simulations certainly provide valuable information on entropic and conformational effects in ligand-receptor association that docking cannot simply deliver by design, modern docking score functions, such as those provided by the efficient Autodock4 software, appear to be remarkably predictive in host-guest systems.

References

[1] Hari S. Muddana, Andrew T. Fenley, David L. Mobley, and Michael K. Gilson. The sampl4 host–guest blind prediction challenge: an overview. *J. Comput Aided Mol. Des.*, 28(4):305–317, 2014.

[2] Jian Yin, Niel M. Henriksen, David R. Slochower, Michael R. Shirts, Michael W. Chiu, David L. Mobley, and Michael K. Gilson. Overview of the sampl5 host–guest challenge: Are we doing better? *J. of Comput. Aided Mol. Des.*, pages 1–19, 2016.

[3] Andrea Rizzi, Steven Murkli, John N. McNeill, Wei Yao, Matthew Sullivan, Michael K. Gilson, Michael W. Chiu, Lyle Isaacs, Bruce C. Gibb, David L. Mobley, and John D. Chodera. Overview of the sampl6 host–guest binding affinity prediction challenge. *J. Comput. Aided Mol. Des.*, 32(10):937–963, Oct 2018.

[4] https://samplchallenges.github.io/roadmap/submissions/, accessed 23 June 2020.

[5] Piero Procacci, Massimiliano Guarrasi, and Guido Guarnieri. Sampl6 host–guest blind predictions using a non equilibrium alchemical approach. *J. Comput. Aided Mol. Des.*, 32(10):965–982, Aug 2018.

[6] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, A. Funel, G. Guarnieri, S. Migliori, F. Palombi, G. Ponti, G. Santomauro, and P. Procacci. Cresco enea hpc clusters: a working example of a multifabric gpfs spectrum scale layout. In *2019 International Conference on High Performance Computing Simulation (HPCS)*, pages 1051–1052, 2019.

[7] Piero Procacci. Hybrid MPI/OpenMP Implementation of the ORAC Molecular Dynamics Program for Generalized Ensemble and Fast Switching Alchemical Simulations. *J. Chem. Inf. Model.*, 56(6):1117–1121, 2016.

[8] Piero Procacci. Primadorac: A free web interface for the assignment of partial charges, chemical topology, and bonded parameters in organic or drug molecules. *J. Chem. Inf. Model.*, 57(6):1240–1245, 2017.

[9] S. Izadi and A. V. Onufriev. Accuracy limit of rigid 3-point water models. *J. Chem. Phys.*, 145(7):074501, 2016.

[10] M. Marchi and P. Procacci. Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the npt ensemble. *J. Chem. Phys.*, 109:5194–520g2, 1998.

[11] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103:8577–8593, 1995.

[12] Piero Procacci and Riccardo Chelli. Statistical Mechanics of Ligand-Receptor Noncovalent Association, Revisited: Binding Site and Standard State Volumes in Modern Alchemical Theories. *J. Chem. Theory Comput.*, 13(5):1924–1933, 2017.

[13] Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.*, 30(16):2785–2791, 2009.

[14] V. Gapsys, S. Michielssens, J.H. Peters, B.L. de Groot, and H. Leonov. *Molecular Modeling of Protein*, chapter Calculation of Binding Free Energies, pages 173–209. Humana Press, 2015.