

FLORE
Repository istituzionale dell'Università degli Studi di Firenze

SAMPL6 blind predictions of water-octanol partition coefficients using nonequilibrium alchemical approaches

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

Original Citation:

SAMPL6 blind predictions of water-octanol partition coefficients using nonequilibrium alchemical approaches / Procacci, P., Guarnieri, G.. - In: JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN. - ISSN 0920-654X. - STAMPA. - 34:(2020), pp. 371-384. [10.1007/s10822-019-00233-9]

Availability:

The webpage https://hdl.handle.net/2158/1384372 of the repository was last updated on 2024-09-12T12:49:52Z

Published version:

DOI: 10.1007/s10822-019-00233-9

Terms of use:

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf)

Publisher copyright claim:

Conformità alle politiche dell'editore / Compliance to publisher's policies

Questa versione della pubblicazione è conforme a quanto richiesto dalle politiche dell'editore in materia di copyright.

This version of the publication conforms to the publisher's copyright policies.

La data sopra indicata si riferisce all'ultimo aggiornamento della scheda del Repository FloRe - The above-mentioned date refers to the last update of the record in the Institutional Repository FloRe

(Article begins on next page)

24 November 2026

Journal of Computer-Aided Molecular Design
https://doi.org/10.1007/s10822-019-00233-9

SAMPL6 blind predictions of water-octanol partition coefficients using nonequilibrium alchemical approaches

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

Received: 9 June 2019 / Accepted: 3 October 2019
© Springer Nature Switzerland AG 2019

Abstract
In this paper, we compute, by means of a non equilibrium alchemical technique, the water-octanol partition coefficients (LogP) for a series of drug-like compounds in the context of the SAMPL6 challenge initiative. Our blind predictions are based on three of the most popular non-polarizable force fields, CGenFF, GAFF2, and OPLS-AA and are critically compared to other MD-based predictions produced using free energy perturbation or thermodynamic integration approaches with stratification. The proposed non-equilibrium method emerges has a reliable tool for LogP prediction, systematically being among the top performing submissions in all force field classes for at least two among the various indicators such as the Pearson or the Kendall correlation coefficients or the mean unsigned error. Contrarily to the widespread equilibrium approaches, that yielded apparently very disparate results in the SAMPL6 challenge, all our independent prediction sets, irrespective of the adopted force field and of the adopted estimate (unidirectional or bidirectional) are, mutually, from moderately to strongly correlated.

Keywords SAMPL6 · LogP · Solvation free energy · Non-equilibrium · Crooks theorem · Fast switching · Fast growth · Hamiltonian replica exchange · HREX · Solute tempering · Torsional tempering

Introduction

The second round of the SAMPL6 challenge was aimed at predicting the 1-octanol-water partition coefficient, LogP, for the neutral forms of the eleven compounds shown in Fig. 1. The experimentally measured LogP coefficients are reported in Ref. [1]. These molecules, are based on the quinazoline, imidazole or pyridine heteroaromatic moieties, and are characterized by the presence of chemical groups that are ubiquitous in drug-like compounds, such as amide, amine, halogen, oxo, hydroxy and carboxy moieties. LogP's are important physical quantities in drug discovery, since they provide, in principle, valuable indications on the distribution of a molecule between a hydrophobic (e.g. lipid bilayer) and a cytosolic, aqueous environment [2].

In the context of the "physical" approaches in SAMPL6 challenge, the LogP is computed by evaluating independently the solvation free energy in water and 1-octanol of the neutral species in standard conditions:

$$\mathrm{LogP} = \mathrm{Log}\frac{[\mathrm{solute}]_{\mathrm{oct}}}{[\mathrm{solute}]_{\mathrm{wat}}} = -\frac{\Delta G_{\mathrm{oct}} - \Delta G_{\mathrm{wat}}}{RT\ln 10} \quad (1)$$

The best ranking among the 47 submissions using physical methods were all based on high level quantum chemical (QM) calculations with implicit solvent parametrizations. QM approaches were among the top performing submissions also for the preceding SAMPL5 challenge on water/cyclohexane distribution coefficients [3]. QM "winners" in the recent LogP and past LogD challenges actually won a Pyrrhic victory. These good performances were largely expected as the approach based on high-level QM calculation using implicit solvation models with adjustable parameters specifically trained on experimental data-set fits well with the fact the solute is surrounded by a homogeneous environment. By the same token, inexpensive atom-based

Electronic supplementary material The online version of this article (doi:https://doi.org/10.1007/s10822-019-00233-9) contains supplementary material, which is available to authorized users.

✉ Piero Procacci
procacci@unifi.it

Guido Guarnieri
guido.guarnieri@enea.it

[1] Department of Chemistry, University of Florence, Via Lastruccia n. 3, 50019 Sesto Fiorentino, FI, Italy

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

Published online: 17 October 2019

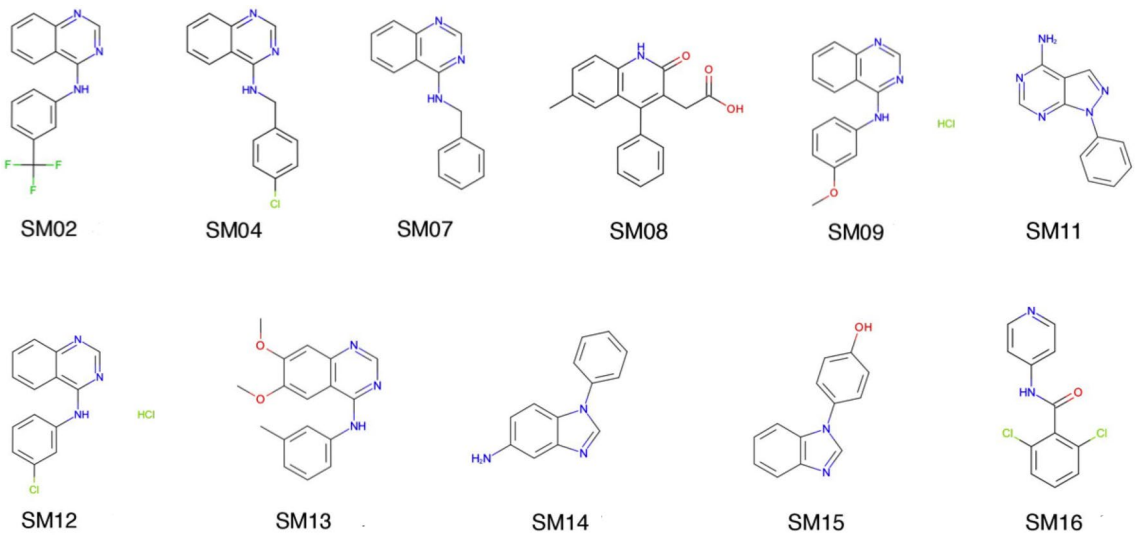


Fig. 1 SM-type compounds in the SAMPL6 LogP challenge

or fragment-based empirical methods like the xLogP3 [4], ClogP/AlogP [5] or miLogP [6], using thousands of compound for the training sets, all yields very good root mean square error (RMSE) and Pearson correlation coefficient (R) for the series of compounds of Fig. 1.

However, the ambitious scope of SAMPL challenges for physical properties is eventually that of testing the predictive power of computational methodologies in perspective *drug design* projects, that is to evaluate "solvation energies" when the solute/ligand is embedded in a highly heterogeneous environment, modulated in a complicated way by the fluctuating protein scaffold where micro-solvation phenomena, related to the atomistic nature of the "solvent", can play a crucial role. As the host-guest SAMPL6 blind challenges have shown [7, 8], even these kinds of simple ligand-receptor systems seem to be out of the reach of QM-based approaches with implicit solvation schemes. In the two latest host-guest SAMPL challenges [7, 8], QM-based submissions for binding free energies of small guest molecules hosted in octa-acids or cucurbit[n]uril-type molecular containers were in fact consistently outperformed by those obtained using classical molecular Dynamics (MD) techniques with explicit solvent.

MD-based methods with explicit solvent appear on the overall to yield consistent performances in SAMPL challenges, whether they are applied to the distribution or partition coefficients or to the more challenging tests of host-guest binding free energies. The accuracy in MD schemes is related to the bias or systematic error due to the adopted atomistic interaction potential or force field. The precision, that is the reproducibility of the datum, is affected by the inherent variance of the implemented MD methodologies. While the latter can be in principle minimized by investing more and more computational resources (i.e. improving the statistical convergence of the simulations), the sources of the force field error are disparate and complex and can involve any combination of the bonded and non bonded solute-solute solute-solvent and solvent-solvent parametrizations. In this regard, SAMPL6 challenges are extremely useful since they provide an ideal collaborative platform on which force fields performances can be rigorously assessed and new avenues for their improvements can be discovered.

In this paper, we will present and discuss the results of a MD-based approach using non equilibrium switching (NES) for the solvation energies and the LogP coefficients for the series of compounds of Fig. 1. The methodology computes the LogP according to Eq. 1, relying on the enhanced sampling of the end-states (fully coupled and fully decoupled solute in water and in 1-octanol) with Hamiltonian Replica Exchange (HREX) and in the subsequent production of hundreds of concurrent and independent non-equilibrium (NE) trajectories where the solute is alchemically driven from the starting canonical ensemble of one end-state to the corresponding nonequilibrium ensemble of the arrival end-state in matter of few hundreds of picoseconds [9, 10]. The solvation free energies in water and in 1-octanol are recovered from the NE work distributions using the Jarzynski and the Crooks fluctuation theorems [11, 12].

Calculations were done using three popular general force fields for drug-like molecules, namely GAFF2 [13, 14], CGenFF [15] and OPLS-AA [16, 17]. For each force field, two blind predictions were uploaded: (i) a "challenge" and computationally expensive submission, done with a precise [18, 19] bidirectional approach based on the Bennett Acceptance Ratio [20] (BAR) estimator; (ii) a less precise and faster submission, obtained with the fast-growth unidirectional

method exploiting the Crooks theorem for normal work distributions [21–24]. We will try to put our contribution to the SAMPL6 initiative into the context of the other MD-based submissions. These were done, in the vast majority of cases, adopting the CGenFF, OPLS-AA or GAFF force fields, or refined variants of them, with *equilibrium* methodologies relying on the alchemical stratification [25, 26] and exploiting the thermodynamic integration [27] (TI) or the Free energy perturbation methods [28] (FEP).

The paper is organized as follows. In Section "Alchemical transformations: theoretical background", we succinctly introduce the equilibrium and nonequilibrium approaches in the context of the MD-based physical methods. In Section "Materials and methods in NES submissions", we provide the detailed methodological information concerning the NES submissions, including force field parametrization, atomistic description of the water and 1-octanol solvent, simulation parameters for the HREX stage and for the subsequent NE computational task. In Section "Overview on MD-based SAMPL6 submissions" a global assessment of the MD-based submissions in LogP/SAMPL6 is presented, highlighting common patterns and discrepancies, as well as force field-related critical issues. In section "NES results", the results obtained using NES methodologies, in the two variants fast switching growth/annihilation method (NES-2) and fast switching growth method (NES-1) and for each of the three general force fields, are critically discussed and compared to the related equilibrium FEP or TI blind predictions. Conclusions and perspectives are sketched out in the last section.

Alchemical transformations: theoretical background

Virtually all MD-based submissions in the challenge were done by evaluating the solvation free energies in Eq. 1 by way of the so-called alchemical approach whereby the solute-solvent interaction is gradually turned off or on, using a decoupling/recoupling *inter-molecular* alchemical parameter λ such that: $V_\lambda = V_{\text{solv}} + V_{\text{solute}} + \lambda V_{\text{solute-solvent}}$. At $\lambda = 1$ the molecule is fully solvated; at $\lambda = 0$ the molecule is totally decoupled from the solvent acting as if it were in the gas-phase. Such methodology can be implemented in the context of *equilibrium* simulations using the stratification strategy or multistage sampling, or, in the context of *nonequilibrium* thermodynamics, producing many concurrent and independent *fast switching* trajectories. To make the paper self contained, here we briefly out-sketch the theoretical background of the two approaches, referring the reader, for a more complete treatment of this subject, to excellent recent reviews [19, 26, 29].

In the equilibrium techniques, the system is simulated at constant pressure and temperature in an appropriate number n of intermediate states corresponding to values of the λ coupling parameter between 0 and 1. The Gibbs solvation free energy is recovered in an inexpensive post-processing stage by summing up contributions obtained by applying the FEP Zwanzig formula [28] for each of the contiguous λ states, i.e.

$$\Delta G_{\text{solv}} = -\beta^{-1} \sum_i \ln\langle e^{-\beta(V_{\lambda_i} - V_{\lambda_{i+1}})}\rangle_{\lambda_i} \tag{2}$$

where β is the reciprocal of the thermodynamic temperature and $\langle\cdot\rangle_{\lambda_i}$ denotes the isothermal-isobaric average taken with potential energy V_{λ_i}. Precision can be increased at a very limited cost by storing during the simulations both the λ_{i+1} and the λ_{i-1} value of the potential energy so that the free energy can be recovered as a sum of BAR contributions [19, 26, 29, 30]. Alternatively, and equivalently, the λ-derivatives of the potential energies can be stored in the stratification, recovering the solvation free energy via numerical thermodynamic integration [27]:

$$\Delta G_{\text{solv}} = \sum_i \langle \partial V_\lambda / \partial \lambda \rangle_{\lambda_i} \Delta\lambda \tag{3}$$

It is important to stress that in FEP or TI approaches the simulation must be well converged on each of the n alchemical λ strata, with a canonical sampling of the relevant solute and solvent conformational space. The convergence rate for a given stratum is an unknown function of the corresponding λ coupling parameter and can vary substantially in the range [0,1]. Typically, barriers between conformational state becomes higher at low coupling (i.e. when $\lambda \to 0$) due to the lack of the screening effect of the solvent on the intramolecular electrostatic interactions [31], making harder the convergence of the simulation with weakly coupled solutes. Also, in setting up the FEP or TI simulations, due care must be taken in "choosing the alchemical protocol so that the total uncertainty for the transformation is the one which has an equal contribution to the uncertainty across every point along the alchemical path" [32], or equivalently so that the overlap between contiguous potential energy distributions is significant and approximately constant in the whole range [0,1], a task that would require the prior knowledge of the dependence of ΔG on λ [26].

In the nonequilibrium approach, equilibrium sampling in the isothermal isobaric ensemble is required only for the end-states, which can be effectively implemented using specialized and highly efficient HREX enhanced sampling schemes [33]. Starting from a representative HREX-sample of n phase-space points (from few tens to few hundreds) of a given end-state, the system is rapidly driven to the other end-state by continuously varying the

λ alchemical parameter in a swarm of corresponding n concurrent and independent NE trajectories, typically lasting from a time τ of few tens to few hundreds of picoseconds and eventually producing an alchemical NE work computed as $W = \int_0^\tau \frac{\partial U}{\partial \lambda}\dot{\lambda}dt$. The fast switching stage can be straightforwardly implemented on a single embarrassingly parallel job on modern HPC platforms, allowing the computation of the NE work distribution in a matter of wall-time minutes.

The NES process can be conducted in the two senses, with the $0 \leq \lambda \leq 1$ and $1 \leq \lambda \leq 0$ processes conventionally indicated [26] as the *forward growth* and the *reverse annihilation* of the solute, respectively. The growth and annihilation NES simulations produce two independent *unidirectional* estimates of the solvation free energy, namely

$$\Delta G = -\beta^{-1} \ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{-\beta W_i(G)}\right) \tag{4}$$

$$\Delta G = \beta^{-1} \ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{-\beta W_i(A)}\right) \tag{5}$$

The above unidirectional estimates based on the Jarzynski exponential average, while asymptotically exact, are nonetheless affected by a bias error that decreases with the number of work values, n, and with the dissipation [34], defined as the difference between the mean NE work and the underlying free energy. Provided that the fast growth and annihilation transformations are conducted with inverted time schedules, the two estimates of Eqs. 4 and 5 can be combined in a bidirectional, statistically efficient and *unbiased* [18] BAR estimator where ΔG is given by the root of the equation

$$\sum_{i=1}^{n} \frac{1}{1+e^{\beta(W_i(G)-\Delta G)}} - \sum_{i=1}^{n} \frac{1}{1+e^{\beta(W_i(A)+\Delta G)}} = 0 \tag{6}$$

If the the, e.g., growth work distribution $P_G(W)$ is found to be normal then, as a trivial consequence of the Crooks theorem [35], the annihilation distribution, $P_A(-W)$, done with inverted time schedule, *must be* normal too with the same variance, $\sigma_A = \sigma_G$. The two forward and reverse distributions are symmetrical with respect to the crossing point at $W = \Delta G$. For normal distribution(s), the free energy can be hence recovered with the unbiased unidirectional estimators:

$$\Delta G = \langle W_G \rangle - \frac{1}{2}\beta\sigma^2 \tag{7}$$

$$\Delta G = -\langle W_A \rangle + \frac{1}{2}\beta\sigma^2 \tag{8}$$

where $\langle W_{G/A} \rangle$ are the mean values of the growth work and of the annihilation work with inverted sign. The quantity $\frac{1}{2}\beta\sigma_G^2 = \frac{1}{2}\beta\sigma_A^2 = W_{\text{diss}}$ is the dissipated work in the transformations, which must be identical in either directions. Provided that the sampling of the starting end-states has been adequate, the precision (e.g. the 95% confidence interval) of the two independent NES Gaussian estimates in Eqs. 7 and 8 increases or decrease with the square root of n and depends only on the sample variance σ^2 [23, 24, 26, 36]. The normality of the distributions can be instantly checked [24] using standard procedures such as the Kolmogorv-Smirnov, the Wilk-Shapiro, the Jarque-Bera or the Anderson-Darling tests [37].

Materials and methods in NES submissions

System preparation

In order to assess the performance of the "official" versions of the most popular force fields (FF), we submitted multiple NES predictions adopting: the CgenFF parameter sets as obtained from the web interface "paramchem" [38, 39]; the GAFF2 topological and parameter files as obtained form the web interface "PrimaDORAC" [14]; the OPLS-AA parameter sets as obtained from the web interface "LigParGen" [17]. The GAFF2 atomic charges are computed by PrimaDORAC at the AM1/BCC level. For the OPLS-AA charges, the LigParGen option "1.14*CM1A-LBCC" was used. The charges in the CGenFF are assigned by analogy [39] by the paramchem web toolkit. All calculation were done with the program ORAC [40]. For each of the three FF's, the parameter files were converted "as is" to the ORAC format with no further adjustment. The ORAC suite, inlcuding source code and documentation, can be freely downloaded from the website `www.chim.unifi.it/orac`.

For all submissions, solvation free energies were evaluated by dissolving the solutes in 1240 water molecules or 125 molecules of octanol in a cubic MD box. Hence in all cases we used "dry" 1-octanol as a solvent. The parametrization of the explicit water solvent in hydration free energy calculations is done using the recently developed OPC3 [41] three-point site model. For 1-octanol as a solvent, in the force field specific submissions, the parametrizations provided by the paramchem (CgenFF), PrimaDORAC (GAFF2) and LigParGen (OPLS-AA) web toolkits were adopted. The use of the common OPC3 model in all our NES submission is motivated by the fact that OPC3 reproduces with accuracy both the static dielectric constant and the density of water at standard conditions. Besides, as shown in Ref. [51], we showed that, while the effect of the selected force field for the solute molecule is important for the LogP prediction, the choice of the force field for the solvent water molecule is

much less critical, with bearly detectable differences in the hydration energies when switching water models.

All simulations were done in the NPT isothermal-isobaric ensemble under periodic boundary conditions, yielding a mean side-length around 32–33 Å in both water and 1-octanol in all cases. The external pressure was set to 1 atm using a Parrinello-Rahman Lagrangian [42] with isotropic stress tensor. The temperature was held constant at 298 K using three Nosé Hoover-thermostats coupled to the translational degrees of freedom of the systems and to the rotational/internal motions of the solute and of the solvent. The equations of motion were integrated using a multiple time-step r-RESPA scheme [43] with a potential subdivision specifically tuned for bio-molecular systems in the NPT ensemble [42, 44]. The long range cut-off for Lennard-Jones interactions was set to 13 Å. Long range electrostatic were treated using the Smooth Particle Mesh Ewald method [45], with an α parameter of 0.38 Å^{-1}, a grid spacing in the direct lattice of about 1 Å and a fourth order B-spline interpolation for the gridded charge array.

In Table 1, we show the computed density, ρ, and static dielectric constant, ϵ, of pure 1-octanol using the three FF's compared to the experimental values [46]. The average density and dielectric constant were calculated on 125 molecules of 1-octanol in the NPT ensemble at T = 298 K and P = 1 atm, running for 12 ns. The three force fields yield essentially the same ϵ and ρ values and are in acceptable agreement with the experimental counterpart.

The density and dielectric constant of OPC3 water in standard condition are 0.996 ± 0.01 g/cm^3 and 78 ± 4, respectively [41]. The corresponding experimental values are 0.997 g/cm^3 and 79 [46]

end-states HREX simulations

The Hamiltonian Replica exchange simulations in (w/o) solution and gas-phase for each of the 11 solute molecules are done using torsional tempering. Torsional tempering, a specialized solute tempering [47] scheme described in detail in Ref. [33], allows to surgically enhance the sampling on the relevant degrees of freedom of the system keeping the replica number to a minimum. For the compounds of Fig. 1, the scaling involves all the torsional potentials (including 1–4 non bonded interactions) around the rotable bonds connecting the planar rigid units, using a minimum scaling factor of $c = 0.1$, corresponding to a "torsional temperature" of 3000 K. Only the scaling factors are communicated among replicas, minimizing inter-processor communications. The torsional GE space is covered using only four replicas, with the scale factors [33] given by $c_m = c^{(m-1)/3}$. Each system at the end-states $\lambda = 0$ and $\lambda = 1$ was simulated using HREX for 8 ns in the target state (hence 32 ns in total for each solute molecule), saving 420 phase space point every 20 ps for the later NES stage. Further technical details on the HREX simulations (round trip times, torsional energy overlap, GE state distributions of the walkers) are provided in Ref. [31] for specific examples.

Table 1 Computed static dielectric constant and density (g/cm^3) of pure 1-octanol in standard conditions using various force fields. Experimental values are taken from Ref. [46]

	CgenFF	GAFF2	OPLS-AA	Exp.
ρ	0.815 ± 0.03	0.809 ± 0.03	0.812 ± 0.03	0.83
ϵ	4.9 ± 0.5	4.1 ± 0.3	4.9 ± 0.3	10.3

NES stage

In the NES stage, for each compound of Fig. 1, in a single parallel job, 420 annihilation trajectories were started in the isothermal-isobaric ensemble in standard conditions, reading the initial phase points, harvested in the preceding HREX at full coupling $\lambda = 1$. In each of these NE annihilation trajectories, the solute was decoupled up to $\lambda = 0$ in 150 ps in water and 300 ps in 1-octanol. The annihilation times are chosen such that the mean dissipated work is approximately the same in the two solvents. The detail of decoupling protocol involves the linear discharging of the solute in the first 30 ps (water) and 60 ps (1-octanol), followed by the Lennard-Jones decoupling in the remaining 120 ps (water) and 240 ps (1-octanol). For the Lennard-Jones decoupling, we used a soft-core Beutler potential [48] regularization as $\lambda \to 0$. Such NES protocol was chosen on the basis of past experience on NES solvation free energy calculation for molecules of comparable size [10, 49–51]. For each of the 11 compounds, 4 work histograms were produced, i.e. $P_G(W), P_A(-W)$ in water and in 1-octanol. All 44 work histograms for the three FF's, produced in the NES stage, are reported in Figures S1–S3 of the Supporting Information (SI).

The fast-growth NES stage was started combining 420 HREX phase space points of the solute in the gas-phase with equilibrated and decorrelated snapshots of the pure solvent, that is, inserting the solute as a ghost molecule in a random position and with random orientation in the equilibrated solvent configurations. The fast growth NES protocol corresponds to the inverted time schedule of the annihilation stage, i.e. in water and in 1-octanol the solute-solvent Lennard-Jones potential with soft-core regularization is first switched on in 120 and 240 ps, followed by the recharging process up to 150 and 300 ps, respectively.

LogP estimates

For each of the three FF, we submitted two blind predictions. The "challenge" prediction (NES-2) is done using both

the growth and annihilation NE work values, exploiting the BAR *bidirectional* estimate, Eq. 6. The error was computed using bootstrap with resampling and corresponds to 1.96 times the square root of the ΔG variance (i.e. $\alpha = 0.05$, 95% confidence interval). In the second "fast" prediction (NES-1), the LogP was determined using only the fast-growth *unidirectional* estimates for the solvation free energies. The forward work distribution, $P_G(W)$, was checked for normality according to the Anderson Darling (AD) test defined via the quantity $A^2 = \sum_{i=1}^{n} \frac{2i-1}{n}[\ln(\Phi(w_i) + \ln(1 - \Phi(w_{n+1-i})]$, where Φ is the Gaussian cumulative distribution function with sample mean and variance and w_i are the work values sorted in ascending order. The critical value of A^2 at the level $\alpha = 0.05$ is 0.752 [52]. If $A \leq 0.752$, the work distributions were assumed to be normal and the solvation free energies were computed using the unbiased unidirectional Gaussian estimate, Eq. 7. In case of AD failure, the unidirectional Jarzynski estimate, Eq 4, was used. As for the first blind prediction, also for the forward (growth) unidirectional estimates the error on the solvation free energies are evaluated using bootstrap with resampling. Raw results for all six blind predictions are reported in the SI (Tables S1–S6).

Efficiency considerations

The "challenge" BAR-based blind prediction, on a per solute molecule basis, required a total of 424 ns for a system of approximately 3000 atoms (64 ns for the HREX stage and 360 ns for the forward and reverse NES stages). The fast-growth blind prediction, on a per solute basis, required a total of 180 ns (with a negligible cost of the HREX on the isolated molecule) for a system of approximately 3000 atoms. All computations were done with the OpenMP/MPI hybrid version of the ORAC program [40] on the 24K cores CRESCO6 ENEA cluster equipped with Intel Skylake 48 cores CPU 2.4 GHz. The "challenge" prediction were computed submitting four batch parallel job scripts, each processing sequentially all 11 compounds, namely two HREX simulations (water and 1-octanol) at full coupling and two subsequent NES (water and 1-octanol) and were completed in two wall clock days. The fast-growth predictions were obtained by submitting just one batch parallel job script (computing sequentially the fast-growth in water and in 1-octanol for all compounds) and was completed in few wall clock hours. Examples of these batch submission scripts are reported in the SI.

Overview on MD-based SAMPL6 submissions

31 MD-based submission were uploaded in the SAMPL6/LogP challenge. Of these, 30 were performed using the alchemical approach; 24 submissions used the FEP or TI equilibrium technique, and 6 the NES method. No other group, except for ourselves, used the NES method for LogP calculations.

In 8 submissions (6 FEP, 2 NES) the CgenFF was used. In some cases, refined versions of the same force field were used processing the paramchem parameters with the "lsfitpar" refinement program [53] ; In 9 submissions (6 FEP, 1 TI, 2 NES), the GAFF1 or GAFF2 force field was used with atomic charges computed at the AM1/BCC level or at higher QM level. In 6 submission (3 TI, 1 FEP, 2 NES) the OPLS-AA LigParGen parameter sets were used in all cases. The latter six submissions used the very same LigParGen generated force field setup for the solute molecule and hence provide a subset for comparing the performances of equilibrium and non equilibrium methodologies. In one case (submission `nh6c0`, one of the best performing MD-based blind predictions), the FF original GAFF1 parameters were manually adjusted. The remaining MD-based submissions used polarizable force fields. Water was described using the OPC3 model (only in the 6 NES submission), the TIP3P model [54] (4 submissions), the TIP4P model [54] (4 submissions) and the SPCE model [55] (2 submissions). In remaining MD-based instances the water model was not specified. In all submissions using classical non polarized FF's, the 1-octanol was modeled using the reference force field. Dry and wet 1-octanol was used in 15 and in 8 submission, respectively. The FEP or TI protocols used a number of λ windows from a minimum of 12 to a maximum of 20, and each λ state was simulated from few ns in water up to 20 ns for 1-octanol. Details of all MD-based submissions

Table 2 Overall performances (in bold font) for MD-based alchemical approach for the the GAFF, CGenFF and OPLS-AA parametrization

	CgenFF	OPLS-AA	GAFF(1/2)	POLAR
R	**0.17** (0.56)	**0.54** (0.79)	**0.64** (0.79)	**− 0.04**
MUE	**1.25** (0.98)	**1.85** (1.06)	**1.85** (2.07)	**1.97**
τ	**0.14** (0.49)	**0.39** (0.45)	**0.41** (0.53)	**− 0.05**
Slope	**0.38** (1.0)	**0.98** (1.33)	**1.34** (1.38)	**− 0.10**
N_{sub} 8	6	9	6	

R, MUE, τ and "slope" indicate to the Pearson's correlation coefficient, the mean unsigned error, the Kendall rank coefficient and the slope of the regression line. In parenthesis the results of the "challenge" (BAR) NES submissions are reported

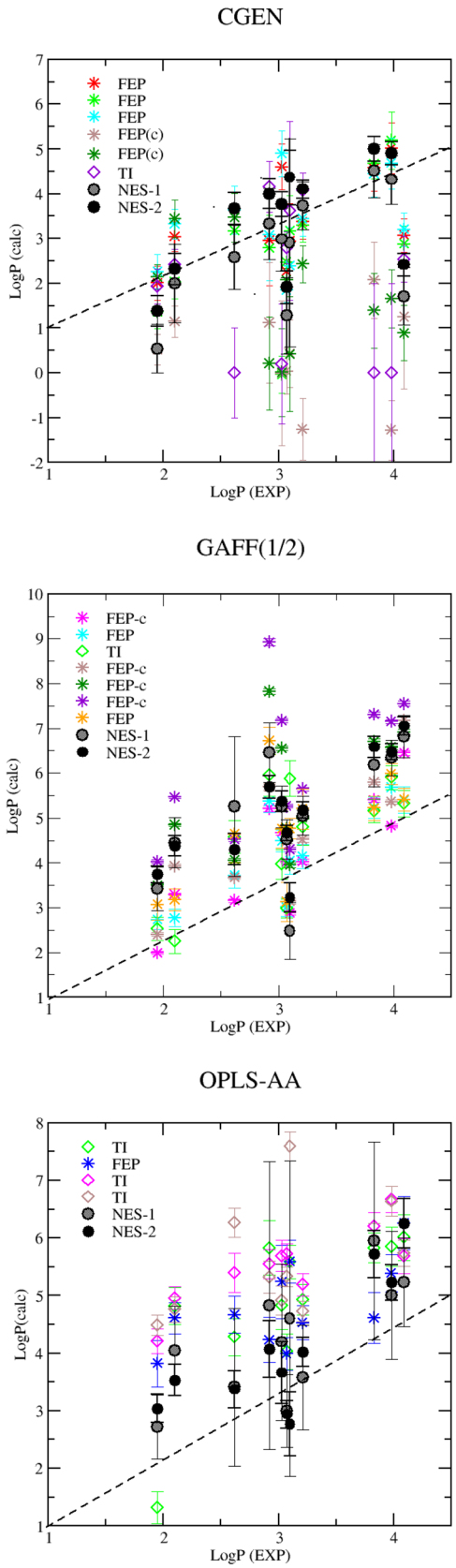


Fig. 2 Experimental-calculated LogP correlation plots for MD-based submission for CGenFF, GAFF and OPLS-AA force fields. NES-2 and NES-1 indicate the bidirectional and unidirectional NES submissions, respectively

using non polarizable force fields are reported in Table S7 of the SI.

In the Table 2 and in Fig. 2 we show the *overall* performances of the alchemical methods as a function of the adopted FF (whether refined or not). Except for the submissions using polarizable force fields (not reported in Fig. 2) showing consistently (and surprisingly) poor results, the performances for the three fixed-charges FF's reported in Table 2 are somewhat contradictory depending on the selected quality indicator. CgenFF appears to be the best performing FF for the MUE, but has the lowest Pearson and Kendall coefficients. Surprisingly, among the less accurate CgenFF submissions, some (see Table S7 in the SI for details) were done using a *refined* force field due to a positive "penalty score" obtained from the paramchem toolkit. This is somewhat puzzling given that, apparently, the "refined" CGenFF submissions were done using exactly the same FEP protocol of the best performing standard (paramchem) CGenFF predictions.

At variance with CgenFF, GAFF yields larger MUE's, consistently overestimating the LogP for virtually all uploaded submissions (see Fig. 2, central panel). On the other hand, GAFF outperforms CGenFF for both the Pearson and the Kendall rank coefficients with no striking degradation or increase of performances when the AM1-BCC charges are replaced by high level QM atomic charges (see Table SI7 in the SI for details). The OPLS-AA FF lies somewhat in between the CgenFF and GAFF FF's. The systematic overestimation of the LogP's, although still evident (see Fig. 2), is less pronounced with respect to GAFF, as measured by a somewhat smaller overall MUE. All six OPLS submissions, on the other hand, exhibit a good overall correlation and ranking coefficients. The LogP overestimation in GAFF and OPLS-AA is likely due to the fact that the Lennard-Jones and electrostatic balance in the parametrization has been in general trained over *hydration* free energies of small molecules [13, 16]. Thus, the electrostatic contribution to the solvation energy could be somewhat overestimated using water trained atomic charges in the apolar 1-octanol solvent, eventually producing systematically higher LogP values.

On the overall, as Table S7 in the SI shows, the use of wet 1-octanol did not appear to improve appreciably the performances. Both FF and reported methodological errors (ranging from 0.1 to 2 LogP units) are seemingly larger than the differences obtained in specific cases for wet or dry 1-octanol. For example, for OPLS-AA, the best performing submission was obtained using dry octanol, as for CGenFF. For GAFF, the submission bearing the lowest MUE and highest R and Kendall coefficients, done using wet 1-octanol, is only marginally better than other submissions obtained using dry 1-octanol. This is so since in wet 1-octanol, in spite of the high molar water fraction of 0.27,

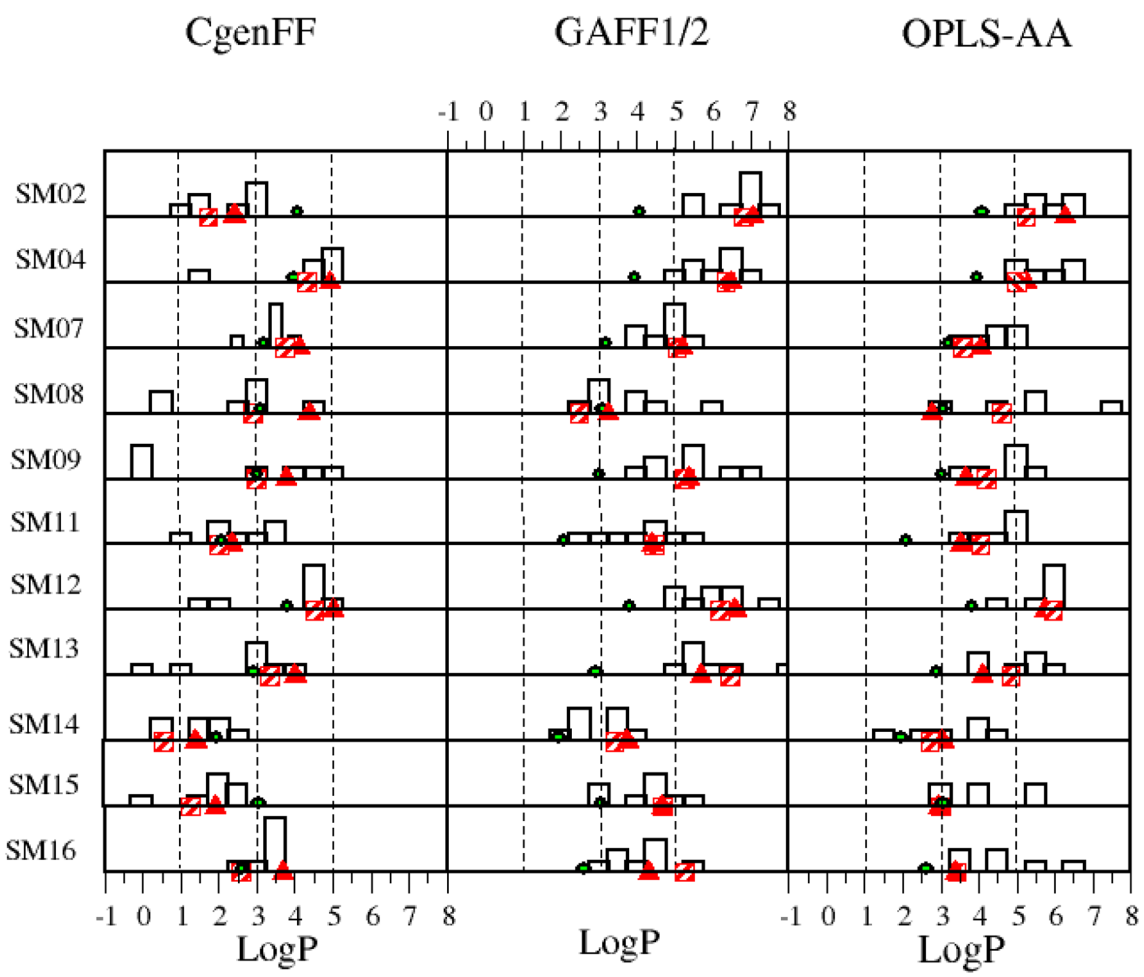


Fig. 3 Detailed NES statistics of MD-based submissions in SAMPL6 challenge for all compounds, classified according to the adopted FF. Green circles: experimental data. Red triangles: NES-2 (bidirectional) submission. Red stripes squares: NES-1. (unidirectional) submission

one molecule of water per four molecules of 1-octanol translates in a water/octanol *volume* fraction of only $\simeq 0.03$ with a limited impact (a slight *decrease*) on the static dielectric constant [56].

Further details of the MD-based submissions are shown in Fig. 3, where we report the statistics of all MD-based prediction sets (classified according to the FF) in the SAMPL6 challenge for each of the compounds of Fig. 1. The experimental values, the NES-2 and NES-1 submissions are indicated as green circles, red filled triangles and a red stripes squares, respectively. For CGenFF, NES predictions exhibit the largest deviation for SM02, as most of the other submissions. For all other compounds, NES is consistently among the best performing methods. In case of GAFF, as previously states, all prediction sets appear to overestimate the LogP. NES submissions are found in most cases in correspondence of the maximum of the overall SAMPL6 GAFF distribution. For OPLS-AA, overestimation of LogP is still evident, and, again, NES submissions are systematically found among the best performing alchemical methods.

NES results

The performances of the BAR-based bidirectional NES "challenge" submissions, NES-2, shown in parenthesis in Table 2 and as the black circles in the Fig. 2, are consistently among the highest ranking for each of the FF's. Only for the CGenFF submission (see Table S7 in the SI), the unidirectional estimate NES-1 turned out to be only marginally better than the bidirectional estimate NES-2, very likely reflecting just a lucky "shot on goal".

At variance with FEP or TI submissions, all NES submissions are from moderately to strongly *mutually* correlated, irrespective of the adopted force field or of the kind of estimate, unidirectional or bidirectional. In Table 3, we show

Table 3 NES correlation matrix; Upper triangle NES-2 correlation; Lower triangle NES-1 correlation

	CgenFF		GAFF2		OPLS-AA		Exp.	
	R	τ	R	τ	R	τ	R	τ
CgenFF	1	1	0.34	0.31	0.34	0.38	0.55	0.49
GAFF2	0.45	0.29	1	1	0.94	0.85	0.79	0.53
OPLS-AA	0.67	0.53	0.57	0.42	1	1	0.79	0.46
Exp.	0.57	0.42	0.62	0.44	0.73	0.53	1	1

the correlation matrix obtained with NES for the three force fields. In the upper and lower triangles we report R and τ correlation coefficients obtained with NES-2 and NES-1, respectively.

As it can be seen from Table 3, for the three NES-2 submissions the mutual correlation R between FF's goes from a minimum of 0.34 (CgenFF-OPLS-AA and CgenFF-GAFF2) to a maximum of 0.94 (OPLS-AA-GAFF2). The mutual Kendall rank coefficient τ behaves similarly. Mutual correlations do not vary significantly among the three NES-1 unidirectional estimates, showing similar R and τ indices among the various FF's.

The NES-2 and NES-1 correlation, already evident from Fig. 3, can be further assessed in the Fig. 4, where we compare the bidirectional "challenge" NES-2 and unidirectional fast-growth NES-1 predictions for all 11 compounds, and for each FF. In the Figure we also report the error on the LogP with the two estimates. As it can be seen from the Figure, the optimal NES-2 and the "fast" NES-1 estimates are strongly correlated in all three cases. In case of CGenFF and GAFF, NES-2 and NES-1 LogP are strikingly similar, yielding a Pearson coefficient of 0.98 and 0.93. Correlation (R = 0.78) is only moderately degraded for

OPLS-AA. This fact can be easily explained inspecting Tables S4–S6 in Figures S1–S3 in the SI. The OPLS-AA work distributions failed the AD test in 10 cases in water or 1-octanol compared to the three cases of CgenFF (SM08 SM11 SM15) and the two cases of GAFF2 (SM08 SM16). Hence for the OPLS-AA FF, the less accurate and precise Jarzynski estimate, Eq. 4, was used for the majority of the compounds, at variance with GAFF2 and CgenFF were the unbiased Gaussian estimate, Eq. 7, was mostly used.

As expected, the confidence intervals (reported as bar plots in the Fig. 4) are larger for the unidirectional LogP estimates with respect to the bidirectional BAR computed LogP. NES-2 uses in fact, twice as much work values with respect to NES-1. It is of note that the largest deviations between the NES-2 and NES-1 LogP estimates are seen in general in correspondence of large NES-1 errors, as in SM08 for CGenFF and OPLS, and in SM16 for GAFF2. Again, large errors are due to nature of the unidirectional estimates in the solvation free energies, obtained from the forward work histograms as assessed by the Anderson Darling normality test (see Tables S4–S6 in the SI): if either ΔG_{oct} or ΔG_{wat} or both have been evaluated using the Jarzynski exponential averages Eqs. 4 in lieu of the Gaussian unbiased estimate Eq. 7, then the confidence interval increases decisively and the LogP becomes less accurate. We must stress that the computational cost of NES-1 calculation of the LogP for one compound of the series is entirely due to the NES stage as the cost of the HREX sampling for a single molecule is negligible. NES-1 is hence a matter of few tens of wall-clock minutes on a Tier1 HPC system as the CRESCO6 cluster provided by ENEA [57]. In our case, the unidirectional water and 1-octanol solvation energies were computed (see technical details in the SI) running on about 3500 cores (about 1/8 of the CRESCO6 cluster) in less than a hour (15 min for water and 30 min for 1-octanol), yielding an average 95% confidence interval not exceeding one LogP unit. Accepting a confidence interval of 80% (i.e. running just 100 trajectories instead of 420), a dedicated CRESCO6 Tier-1 cluster can compute about 1000 NES-1 LogP coefficients per day.

We conclude this section by examining in detail the behavior of the FF's for the compounds of the SAMPL6 challenge in terms of the electrostatic and Lennard-Jones contributions to the solvation free energy in water and in 1-octanol. The electrostatic contribution (QQ) was obtained from the reverse annihilation, evaluating the work distribution at the end of the discharging process (i.e. $\tau = 30$ ps and $\tau = 60$ ps for water and 1-octanol, respectively). We used the unidirectional estimate (Gaussian of Jarzinski depending on the AD test) on the QQ work distributions. The Lennard-Jones contribution (LJ) was computed from the growth process evaluating the work at $\tau = 120$ and $\tau = 240$ in water and 1-octanol, respectively. In this case all LJ work distributions, with no exception, were found normal and the Gaussian estimates was hence used in all cases. The results of this analysis are summarized in Fig. 5 where the values of the ΔG_{QQ} and ΔG_{LJ} contributions are reported for each of the compounds of Fig. 1. As expected, in water (see Fig. 5, top panel) the major contribution to the solvation energy is consistently due to the ΔG_{QQ} free energy. In 1-octanol (Fig. 5, bottom panel), the situation is reversed. For this solvent, the ΔG_{LJ} contribution is significantly larger than ΔG_{QQ} for almost all compounds, with the notable exception of SM08, which was the only compound bearing a carboxylate group and where both CgenFF and GAFF2 predict a ΔG_{QQ} that is of the order of or larger than the ΔG_{LJ}.

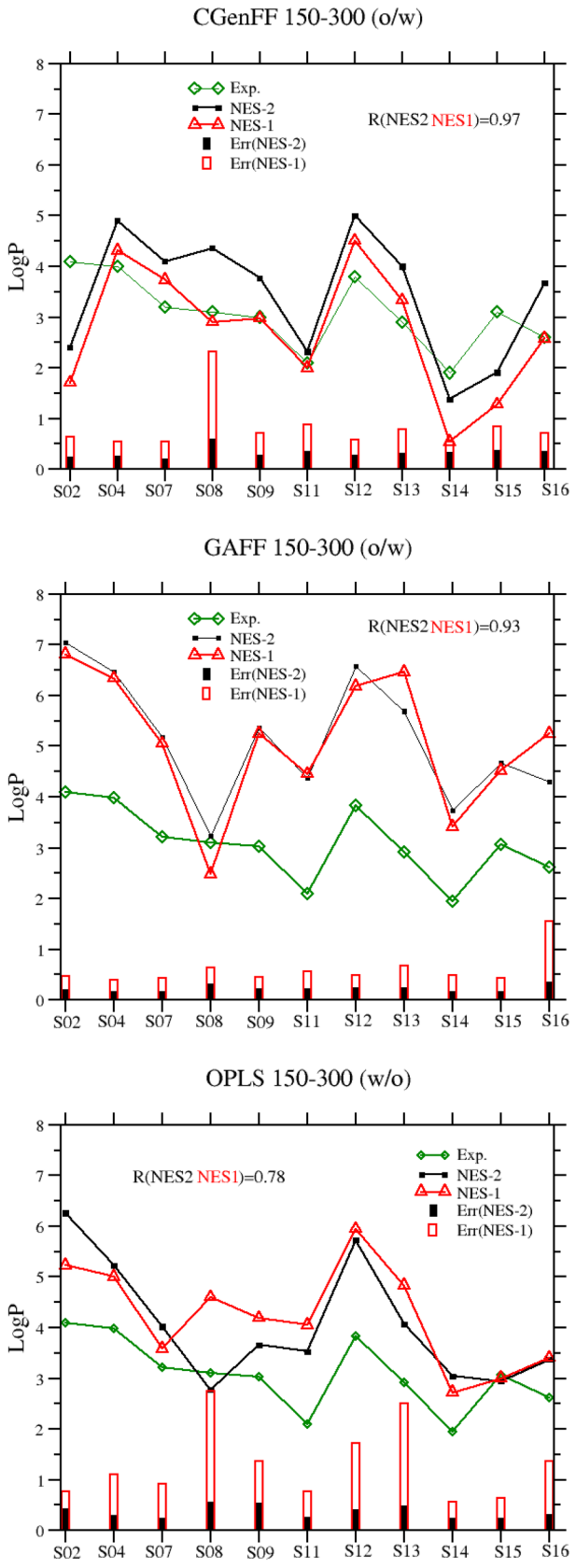


Fig. 4 LogP coefficients computed using NES-2 and NES-1 approaches

In spite of the evident correlation, the balance of the QQ and LJ contributions in three FF's exhibits significant differences. OPLS-AA and GAFF2 yield very similar LJ contributions for virtually all cases and in both solvents. OPLS-AA, however, has a larger QQ contribution with respect to GAFF2 in both water and, to a somewhat less extent, in 1-octanol. Probably, the QQ contribution in 1-octanol is overestimated leading to the observed overestimation of the LogP coefficient in both OPLS-AA and GAFF2.

CgenFF, on the other hand, is characterized by smaller LJ contributions in many cases with respect to the other two FFs' in both solvent. The CgenFF QQ contributions to the hydration free energy lie in between (except for SM02, SM14,SM15) with respect to those obtained using the OPLS-AA and the GAFF2 force fields. These subtle differences make apparently CGenFF more balanced with respect to OPLS-AA and GAFF2, yielding in many cases LogP with significantly smaller MUE's. We finally note that for the clorurated compounds (SM04, SM12, SM16) the extra site representing the chlorine σ-hole [58] in CGenFF (and not used in OPLS-AA and GAFF2) do seem to have an appreciable impact on the corresponding MUE's. CGenFF LogP for SM02, SM12 and SM16 are better than the corresponding OPLS-AA and GAFF2 values.

Conclusions

In this paper, we have given an overview on the MD-based blind predictions in SAMPL6 LogP challenge using the most popular non polarizable force fields, CGenFF, GAFF2 and OPLS-AA, in combination with the alchemical approach. Force fields produce in general moderately consistent predictions. No force field can be considered as optimal, although CGenFF, in its standard (paramchem) implementations, appear to be more balanced and predictive with respect to GAFF1 or GAFF2 and OPLS-AA, yielding in general smaller RMSE's. On the other hand, OPLS-AA and GAFF2 show on the overall better Pearson and Kendall coefficients, both exhibiting a systematic overestimation of the LogP, possibly due to the overestimation of the electrostatic contribution to the solvation free energy in 1-octanol. Such systematic error could be possible rectified by rescaling the atomic charges of the solute in 1-octanol and/or the atomic charges of the 1-octanol molecules in the solvent. Results using polarizable force fields were surprisingly poor, showing that these force field still need adjustment in the balance between *fixed* atom-atom Lennard-Jones interaction and the Coulomb contributions due to the *fluctuating* atomic charges or dipoles.

NES emerges has a reliable tool for LogP prediction, systematically being among the top performing submissions in all force field classes for at least two among the

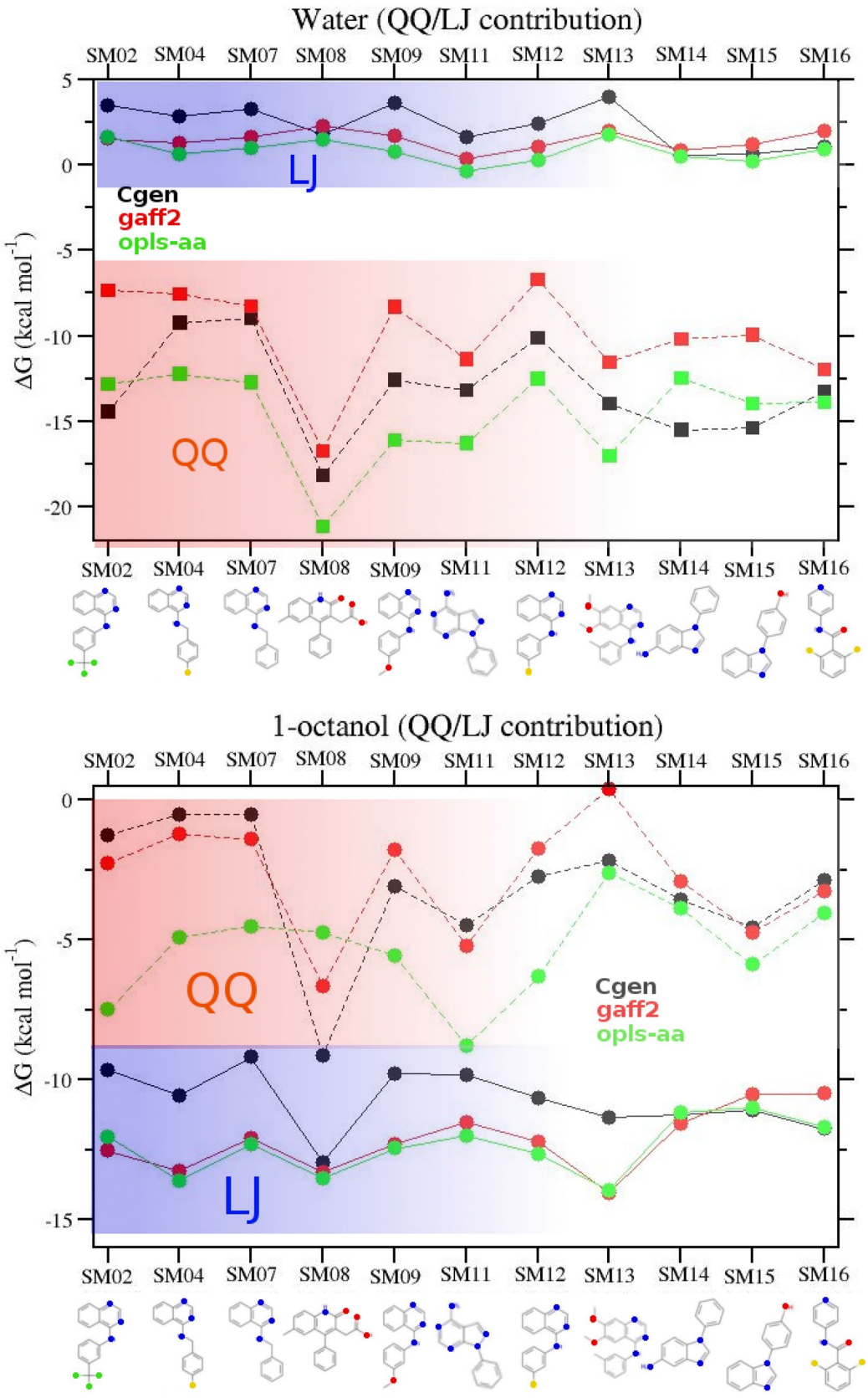


Fig. 5 Electrostatic and Lennard-Jones contributions (see text for details) to the solvation in water (top panel) and in 1-octanol (bottom panel) for the series of SAMPL6 compounds

various indicators (R, τ or RMSE). Contrarily to FEP or TI equilibrium approaches, which yield apparently very disparate results, all independent NES prediction sets, irrespective of the adopted force field and of the adopted estimate (unidirectional or bidirectional) are, mutually, from moderately to strongly correlated (from 0.35 to 0.95). Remarkably, accuracy is only moderately degraded in the unidirectional (growth) NES-1 submissions, that are on other hand, extremely convenient from a computational standpoint: a single LogP can be computed in a matter of minutes on a Tier-1 HPC system such as the CRESCO6 ENEA cluster equipped with Intel Skylake 48 cores CPU 2.4 GHz. NES, at constant *wall time* cost provides a methodology that bypass the problem of the sampling issue in MD-based equilibrium approach. Poor convergence or inadequate solute conformational sampling along the alchemical coordinate could in fact be the primary cause for the observed disparity of FEP or TI submissions even when using the same simulation setup and force field. Strikingly, the same disparity in equilibrium alchemical applications, even when using similar simulation protocols, is observed for the statistical uncertainties (imprecision) of the LogP.

At variance with stratification methods that are based on equilibrium sampling on each of the strata, in the NES approach equilibrium is required only at the end-states. The phase-space sampling of the end-states can be acquired as accurately as possible by using highly parallel and efficient enhanced sampling techniques affording a fast and accurate canonical sampling of all relevant collective coordinates. In the subsequent NES stage, fast NE trajectories connect the equilibrium phase-space points of one end-state to the corresponding non equilibrium set of the other end state. Provided that the equilibrium sampling of the starting end-state has been adequate, the confidence interval in NES rigorously depends on the variance of a *single* work distribution, obtained crossing at fast speed the whole alchemical coordinate, and on the number of NE trajectories (or equivalently, on the HREX collected phase-space points). The confidence level in NES industrial projects can be controlled using only two parameters, that is (i) the number of NE trajectories and (ii) the length of the NE trajectories that control the final dissipation and variance.

Acknowledgements The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes (see www.cresco.enea.it for information).

References

1. Isik M, Mobley DL, Levorse D, Rhodes T, Chodera J (2019) Octanol-water partition coefficient measurements for the SAMPL6 blind prediction challenge. biorxiv. https://doi.org/10.1101/757393
2. Source: DrugBank. Description: The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. http://www.drugbank.ca/drugs/DB03366. Accessed 6 May 2019
3. Bannan Caitlin C, Burley Kalistyn H, Chiu Michael, Shirts Michael R, Gilson Michael K, Mobley David L (2016) Blind prediction of cyclohexane-water distribution coefficients from the sampl5 challenge. J Comput-Aided Mol Des 30(11):927–944
4. Cheng Tiejun, Zhao Yuan, Xun Li Fu, Lin Yong Xu, Zhang Xinglong, Li Yan, Wang Renxiao, Lai Luhua (2007) Computation of octanol-water partition coefficients by guiding an additive model with knowledge. J Chem Inf Model 47(6):2140–2148
5. Ghose Arup K, Viswanadhan Vellarkad N, Wendoloski John J (1998) Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: an analysis of alogp and clogp methods. J Phys Chem A 102(21):3762–3772
6. Molinspiration cheminformatics software. https://www.molinspiration.com/. Accessed 6 May 2019
7. Yin J, Henriksen NM, Slochower DR, Shirts MR, Chiu MW, Mobley DL, Gilson MK (2016) Overview of the sampl5 host-guest challenge: are we doing better? J Comput Aided Mol Des 31:1–19
8. Rizzi Andrea, Murkli Steven, McNeill John N, Yao Wei, Sullivan Matthew, Gilson Michael K, Chiu Michael W, Isaacs Lyle, Gibb Bruce C, Mobley David L, Chodera John D (2018) Overview of the sampl6 host-guest binding affinity prediction challenge. J Comput Aided Mol Des 32(10):937–963
9. Gapsys Vytautas, Seeliger Daniel, de Groot BL (2012) New soft-core potential function for molecular dynamics based alchemical free energy calculations. J Chem Theor Comput 8:2373–2382
10. Procacci Piero, Cardelli Chiara (2014) Fast switching alchemical transformations in molecular dynamics simulations. J Chem Theory Comput 10:2813–2823
11. Jarzynski C (1997) Nonequilibrium equality for free energy differences. Phys Rev Lett 78:2690–2693
12. Crooks GE (1998) Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. J Stat Phys 90:1481–1487
13. GAFF and GAFF2 are public domain force fields and are part of the AmberTools16 distribution, available for download at http://amber.org internet address (accessed March 2017). According to the AMBER development team, the improved version of GAFF, GAFF2, is an ongoing poject aimed at "reproducing both the high quality interaction energies and key liquid properties such as density, heat of vaporization and hydration free energy". GAFF2 is expected "to be an even more successful general purpose force field and that GAFF2-based scoring functions will significantly improve the successful rate of virtual screenings."
14. Procacci Piero (2017) 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
15. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD (2010) Charmm general force field: a force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. J Comput Chem 31(4):671–690

16. Dodda Leela S, Vilseck Jonah Z, Tirado-Rives Julian, Jorgensen William L (2017) 1.14*cm1a-lbcc: localized bond-charge corrected cm1a charges for condensed-phase simulations. J Phys Chem B 121(15):3864–3870
17. Dodda Leela S, de Vaca Israel Cabeza, Tirado-Rives Julian, Jorgensen William L (2017) Ligpargen web server: an automatic opls-aa parameter generator for organic ligands. Nucleic Acids Res 45(W1):W331–W336
18. Shirts MR, Bair E, Hooker G, Pande VS (2003) Equilibrium free energies from nonequilibrium measurements using maximum likelihood methods. Phys Rev Lett 91:140601
19. Matos Guilherme Duarte Ramos, Kyu Daisy Y, Loeffler Hannes H, Chodera John D, Shirts Michael R, Mobley David L (2017) Approaches for calculating solvation free energies and enthalpies demonstrated with an update of the freesolv database. J Chem Eng Data 62(5):1559–1569
20. Bennett CH (1976) Efficient estimation of free energy differences from monte carlo data. J Comp Phys 22:245–268
21. Procacci Piero (2016) I. dissociation free energies of drug-receptor systems via non-equilibrium alchemical simulations: a theoretical framework. Phys Chem Chem Phys 18:14991–15004
22. Nerattini Francesca, Chelli Riccardo, Procacci Piero (2016) Ii. dissociation free energies in drug-receptor systems via nonequilibrium alchemical simulations: application to the fk506-related immunophilin ligands. Phys Chem Chem Phys 18:15005–15018
23. Procacci Piero (2018) Myeloid cell leukemia 1 inhibition: an in silico study using non-equilibrium fast double annihilation technology. J Chem Theor Comput 14(7):3890–3902
24. Procacci P, Guarrasi M, Guarnieri G (2018) Sampl6 host-guest blind predictions using a non equilibrium alchemical approach. J Comput Aided Mol Des 32:965–982
25. Jorgensen WL, Buckner JK, Boudon S, TiradoRives J (1988) Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. J Chem Phys 89:3742–3746
26. Pohorille A, Jarzynski C, Chipot C (2010) Good practices in free-energy calculations. J Phys Chem B 114(32):10235–10253
27. Kirkwood JG (1935) Statistical mechanics of fluid mixtures. J Chem Phys 3:300–313
28. Zwanzig RW (1954) High-temperature equation of state by a perturbation method. I. Nonpolar gases. J Chem Phys 22:1420–1426
29. Shirts Michael R, Mobley David L (2013) An introduction to best practices in free energy calculations. Methods Mol Biol 924:271–311
30. Procacci Piero (2017) Alchemical determination of drug-receptor binding free energy: where we stand and where we could move to. J Mol Gr Model 71:233–241
31. Procacci P (2019) Solvation free energies via alchemical simulations: let's get honest about sampling, once more. Phys Chem Chem Phys 21:13826–13834
32. Naden Levi N, Shirts Michael R (2015) Linear basis function approach to efficient alchemical free energy calculations. 2. Inserting and deleting particles with coulombic interactions. J Chem Theor Comput 11:2536–2549
33. Marsili S, Signorini GF, Chelli R, Marchi M, Procacci P (2010) Orac: a molecular dynamics simulation program to explore free energy surfaces in biomolecular systems at the atomistic level. J Comput Chem 31:1106–1116
34. Gore Jeff, Ritort Felix, Bustamante Carlos (2003) Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. Proc Natl Acad Sci USA 100(22):12564–12569
35. Procacci P, Marsili S, Barducci A, Signorini GF, Chelli R (2006) Crooks equation for steered molecular dynamics using a nosé-hoover thermostat. J Chem Phys 125:164101
36. Hummer G (2001) Fast-growth thermodynamic integration: Error and efficiency analysis. J Chem Phys 114:7330–7337
37. Razali NM, Wah YB (2011) Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. J Stat Model Anal 2:21–33
38. Vanommeslaeghe K, MacKerell AD (2012) Automation of the charmm general force field (cgenff) i: bond perception and atom typing. J Chem Inf Model 52(12):3144–3154
39. Vanommeslaeghe K, Raman EP, MacKerell AD (2012) Automation of the charmm general force field (cgenff) ii: Assignment of bonded parameters and partial atomic charges. J Chem Inf Model 52(12):3155–3168
40. Procacci Piero (2016) 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
41. Izadi S, Onufriev AV (2016) Accuracy limit of rigid 3-point water models. J Chem Phys 145(7):074501
42. Marchi M, Procacci P (1998) Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the npt ensemble. J Chem Phys 109:5194–520g2
43. Tuckerman M, Berne BJ (1992) Reversible multiple time scale molecular dynamics. J Chem Phys 97:1990–2001
44. Procacci P, Paci E, Darden T, Marchi M (1997) Orac: a molecular dynamics program to simulate complex molecular systems with realistic electrostatic interactions. J Comput Chem 18:1848–1862
45. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh ewald method. J Chem Phys 103:8577–8593
46. Kim Sunghwan, Thiessen Paul A, Bolton Evan E, Chen Jie, Gang Fu, Gindulyte Asta, Han Lianyi, He Jane, He Siqian, Shoemaker Benjamin A, Wang Jiyao, Bo Yu, Zhang Jian, Bryant Stephen H (2016) Pubchem substance and compound databases. Nucleic Acids Res 44(D1):D1202–D1213
47. Liu P, Kim B, Friesner RA, Berne BJ (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc Natl Acad Sci 102:13749–13754
48. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chem Phys Lett 222:5229–539
49. Gapsys Vytautas, Seeliger Daniel, de Groot BL (2012) New soft-core potential function for molecular dynamics based alchemical free energy calculations. J Chem Theor Comput 8:2373–2382
50. Yildirim Ahmet, Wassenaar Tsjerk A, van der Spoel David (2018) Statistical efficiency of methods for computing free energy of hydration. J Chem Phys 149(14):144111
51. Vassetti Dario, Pagliai Marco, Procacci Piero (2019) Assessment of gaff2 and opls-aa general force fields in combination with the water models tip3p, spce, and opc3 for the solvation free energy of druglike organic molecules. J Chem Theor Comput 15(3):1983–1995
52. Stephens MA (1979) Test of fit for the logistic distribution based on the empirical distribution function. Biometrika 66:591–595
53. Vanommeslaeghe Kenno, Yang Mingjun, MacKerell Alexander D Jr (2015) Robustness in the fitting of molecular mechanics parameters. J Comput Chem 36(14):1083–1101
54. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
55. Kusalik PG, Svishchev IM (1994) The spatial structure in liquid water. Science 265:1219–1221
56. Gestblom B, Sjöblom GA (1984) Dielectric relaxation studies of aqueous long-chain alcohol solutions. Acta Chem Scand A38:47–56

57. Cresco: Centro computazionale di ricerca sui sistemi complessi. Italian National Agency for New Technologies, Energy (ENEA). See https://www.cresco.enea.it. Accessed 24 June 2015

58. Politzer Peter, Murray Jane S, Clark Timothy (2013) Halogen bonding and other σ-hole interactions: a perspective. Phys Chem Chem Phys 15:11178–11189

Supporting Informatin for "SAMPL6 blind predictions of water-octanol partition coefficients using non equilibrium alchemical approaches"

Piero Procacci*,† and Guido Guarnieri‡

†*Department of Chemistry, University of Florence, Via Lastruccia n. 3, Sesto Fiorentino, I-50019 Italy*

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

E-mail: procacci@unifi.it

Table S1: **bzeez**: Bidirectional BAR prediction using GAFF2 force field. Solvation free energies are expressed in kcal mol^{-1}

Compound	$\Delta G_{1-\mathrm{oct}}$	ΔG_{wat}	$\mathrm{LogP}_{\mathrm{calc}}$	$\mathrm{LogP}_{\mathrm{exp}}$
SM02	-15.26 ± 0.15	-5.52 ± 0.14	7.05 ± 0.21	4.09
SM04	-15.10 ± 0.11	-6.16 ± 0.13	6.47 ± 0.18	3.98
SM07	-13.89 ± 0.15	-6.75 ± 0.10	5.17 ± 0.18	3.21
SM08	-20.52 ± 0.24	-16.05 ± 0.21	3.23 ± 0.32	3.10
SM09	-14.81 ± 0.15	-7.39 ± 0.17	5.37 ± 0.23	3.03
SM11	-17.07 ± 0.20	-11.02 ± 0.10	4.38 ± 0.22	2.10
SM12	-14.57 ± 0.15	-5.48 ± 0.18	6.58 ± 0.24	3.83
SM13	-15.94 ± 0.15	-8.09 ± 0.20	5.69 ± 0.25	2.92
SM14	-14.90 ± 0.16	-9.73 ± 0.09	3.74 ± 0.18	1.95
SM15	-15.13 ± 0.13	-8.68 ± 0.11	4.67 ± 0.17	3.07
SM16	-15.34 ± 0.28	-9.40 ± 0.21	4.30 ± 0.35	2.62

Table S2: **2ggir**: Bidirectional BAR prediction using CgenFF force field. Solvation free energies are expressed in kcal mol^{-1}

Compound	ΔG_{1-oct}	ΔG_{wat}	$LogP_{calc}$	$LogP_{exp}$
SM02	-13.63 ± 0.20	-10.30 ± 0.15	2.41 ± 0.25	4.09
SM04	-12.73 ± 0.21	-5.96 ± 0.15	4.90 ± 0.26	3.98
SM07	-10.95 ± 0.15	-5.28 ± 0.13	4.10 ± 0.20	3.21
SM08	-21.77 ± 0.66	-15.75 ± 0.19	4.36 ± 0.61	3.10
SM09	-13.84 ± 0.22	-8.62 ± 0.16	3.77 ± 0.28	3.03
SM11	-14.75 ± 0.36	-11.54 ± 0.12	2.32 ± 0.35	2.10
SM12	-13.98 ± 0.24	-7.08 ± 0.15	5.00 ± 0.28	3.83
SM13	-14.67 ± 0.32	-9.14 ± 0.14	4.00 ± 0.33	2.92
SM14	-16.72 ± 0.35	-14.81 ± 0.13	1.38 ± 0.34	1.95
SM15	-17.23 ± 0.40	-14.59 ± 0.11	1.91 ± 0.37	3.07
SM16	-17.18 ± 0.39	-12.11 ± 0.10	3.67 ± 0.36	2.62

Table S3: **25s67**: Bidirectional BAR prediction using OPLS-AA force field. Solvation free energies are expressed in kcal mol^{-1}

Compound	ΔG_{1-oct}	ΔG_{wat}	$LogP_{calc}$	$LogP_{exp}$
SM02	-18.98 ± 0.40	-10.33 ± 0.18	6.26 ± 0.43	4.09
SM04	-18.50 ± 0.28	-11.28 ± 0.15	5.23 ± 0.31	3.98
SM07	-17.20 ± 0.21	-11.64 ± 0.14	4.02 ± 0.25	3.21
SM08	-22.79 ± 0.58	-18.97 ± 0.20	2.77 ± 0.56	3.10
SM09	-18.80 ± 0.43	-13.75 ± 0.31	3.66 ± 0.54	3.03
SM11	-21.41 ± 0.27	-16.53 ± 0.10	3.53 ± 0.27	2.10
SM12	-19.18 ± 0.41	-11.27 ± 0.15	5.72 ± 0.41	3.83
SM13	-19.57 ± 0.30	-13.94 ± 0.38	4.07 ± 0.49	2.92
SM14	-15.98 ± 0.20	-11.77 ± 0.14	3.04 ± 0.25	1.95
SM15	-17.48 ± 0.21	-13.42 ± 0.12	2.94 ± 0.24	3.07
SM16	-17.44 ± 0.31	-12.78 ± 0.14	3.37 ± 0.32	2.62

Table S4: **5svjv:** Fast-growth prediction using GAFF2 force field. The "Est." column indicates the unidirectional estimate type: 1, Gaussian; 2 Jarzynski. Solvation free energies are expressed in kcal mol^{-1}

Compound	$\Delta G_{1-\mathrm{oct}}$	Est.	ΔG_{wat}	Est.	$\mathrm{LogP}_{\mathrm{calc}}$	$\mathrm{LogP}_{\mathrm{exp}}$
SM02	-14.79 ± 0.36	1	-5.38 ± 0.29	1	6.81 ± 0.47	4.09
SM04	-15.09 ± 0.24	1	-6.33 ± 0.31	1	6.34 ± 0.40	3.98
SM07	-13.69 ± 0.31	1	-6.72 ± 0.29	1	5.05 ± 0.43	3.21
SM08	-18.81 ± 0.46	2	-15.39 ± 0.43	1	2.48 ± 0.64	3.10
SM09	-14.46 ± 0.33	1	-7.21 ± 0.31	1	5.25 ± 0.46	3.03
SM11	-17.43 ± 0.51	1	-11.28 ± 0.26	1	4.45 ± 0.56	2.10
SM12	-14.09 ± 0.26	1	-5.55 ± 0.41	1	6.18 ± 0.48	3.83
SM13	-15.91 ± 0.41	1	-6.99 ± 0.51	1	6.46 ± 0.67	2.92
SM14	-14.70 ± 0.41	1	-9.98 ± 0.27	1	3.42 ± 0.49	1.95
SM15	-15.26 ± 0.35	1	-9.02 ± 0.26	1	4.52 ± 0.44	3.07
SM16	-15.90 ± 1.71	2	-8.64 ± 0.45	1	5.25 ± 1.56	2.62

Table S5: **y0xxd:** Fast-growth prediction using CgenFF force field. The "Est." column indicates the unidirectional estimate type: 1, Gaussian; 2 Jarzynski. Solvation free energies are expressed in kcal mol^{-1}

Compound	$\Delta G_{1-\mathrm{oct}}$	Est.	ΔG_{wat}	Est.	$\mathrm{LogP}_{\mathrm{calc}}$	$\mathrm{LogP}_{\mathrm{exp}}$
SM02	-13.26 ± 0.43	1	-10.90 ± 0.46	1	1.71 ± 0.64	4.09
SM04	-12.36 ± 0.33	1	-6.40 ± 0.43	1	4.31 ± 0.55	3.98
SM07	-10.85 ± 0.35	1	-5.70 ± 0.40	1	3.73 ± 0.54	3.21
SM08	-21.14 ± 2.44	2	-17.13 ± 0.76	1	2.90 ± 2.32	3.10
SM09	-13.11 ± 0.41	1	-8.99 ± 0.57	1	2.98 ± 0.71	3.03
SM11	-14.36 ± 0.91	2	-11.59 ± 0.30	1	2.00 ± 0.88	2.10
SM12	-13.46 ± 0.51	1	-7.23 ± 0.30	1	4.51 ± 0.59	3.83
SM13	-14.32 ± 0.58	1	-9.72 ± 0.51	1	3.33 ± 0.79	2.92
SM14	-15.58 ± 0.51	1	-14.83 ± 0.23	1	0.54 ± 0.54	1.95
SM15	-16.60 ± 0.85	2	-14.83 ± 0.33	1	1.28 ± 0.85	3.07
SM16	-15.89 ± 0.74	1	-12.33 ± 0.26	1	2.58 ± 0.72	2.62

Table S6: **ggm6n:** Fast-growth prediction using OPLS-AA force field. The "Est." column indicates the unidirectional estimate type: 1, Gaussian; 2 Jarzynski. Solvation free energies are expressed in kcal mol^{-1}

Compound	$\Delta G_{1-\mathrm{oct}}$	Est.	ΔG_{wat}	Est.	$\mathrm{LogP}_{\mathrm{calc}}$	$\mathrm{LogP}_{\mathrm{exp}}$
SM02	-17.18 ± 0.56	1	-9.96 ± 0.50	1	5.23 ± 0.77	4.09
SM04	-18.18 ± 1.26	2	-11.27 ± 0.27	1	5.00 ± 1.11	3.98
SM07	-16.84 ± 0.55	2	-11.89 ± 0.71	2	3.58 ± 0.91	3.21
SM08	-25.29 ± 3.03	2	-18.94 ± 0.75	2	4.60 ± 2.74	3.10
SM09	-18.64 ± 1.40	2	-12.85 ± 0.48	1	4.19 ± 1.36	3.03
SM11	-22.30 ± 0.73	1	-16.70 ± 0.33	1	4.05 ± 0.77	2.10
SM12	-18.93 ± 1.81	2	-10.71 ± 0.56	1	5.95 ± 1.72	3.83
SM13	-20.14 ± 2.98	2	-13.47 ± 0.48	1	4.83 ± 2.50	2.92
SM14	-15.86 ± 0.45	1	-12.10 ± 0.33	1	2.72 ± 0.56	1.95
SM15	-17.47 ± 0.46	1	-13.32 ± 0.41	2	3.00 ± 0.63	3.07
SM16	-17.87 ± 1.44	2	-13.16 ± 0.45	1	3.41 ± 1.37	2.62

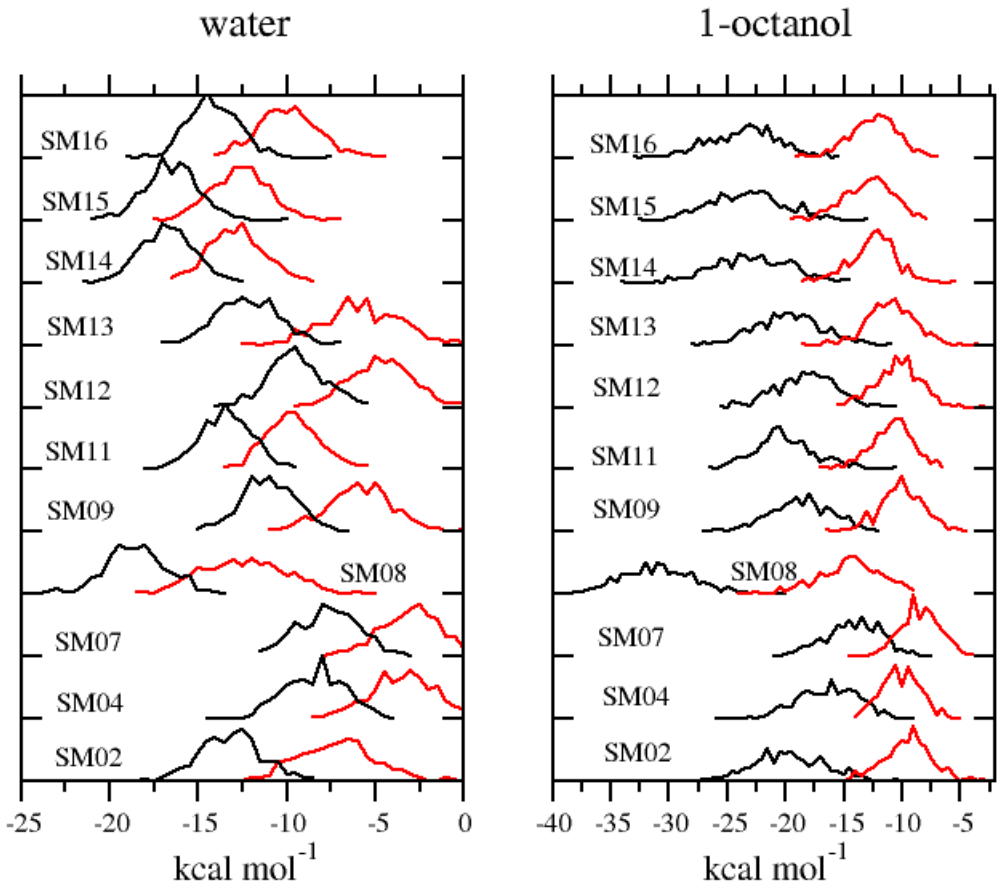


Figure S1: Work distributions in water and 1-octanol using the CgenFF force field. The forward growth work distribution $P_G(W)$ is in red color. The reverse annihilation work distribution $P_A(-W)$ is in black color.

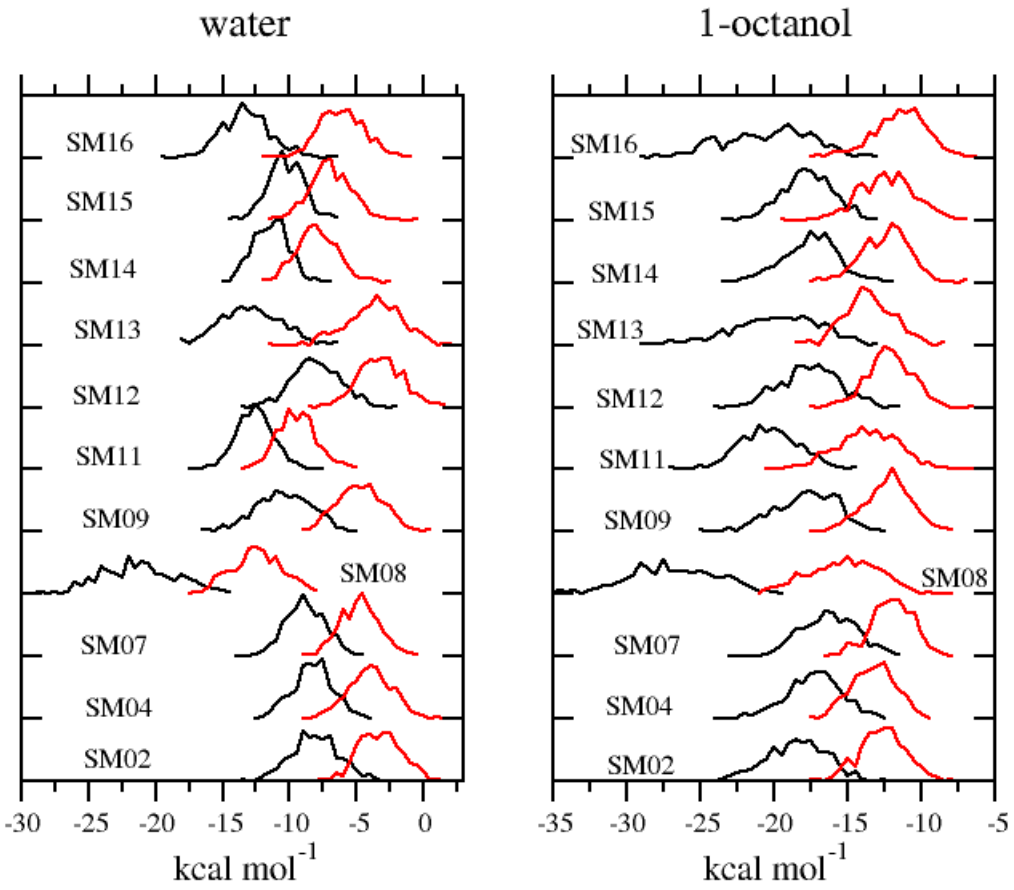


Figure S2: Work distributions in water and 1-octanol using the GAFF2 force field. The forward growth work distribution $P_G(W)$ is in red color. The reverse annihilation work distribution $P_A(-W)$ is in black color.

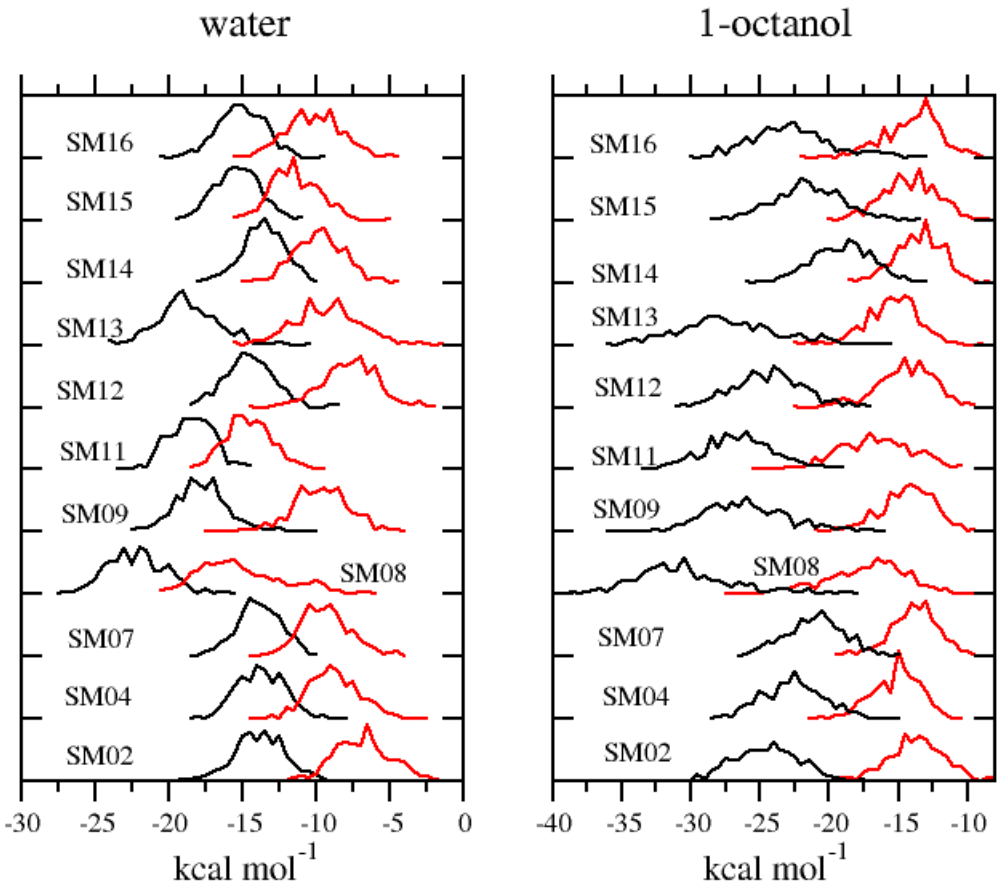


Figure S3: Work distributions in water and 1-octanol using the OPLS-AA force field. The forward growth work distribution $P_G(W)$ is in red color. The reverse annihilation work distribution $P_A(-W)$ is in black color.

```
#!/bin/bash
# Example of bash script "fs-ga-o.bash" for CRESCO6 submission (NES stage)
#    LSF submission command:
#    bsub -n 3360 -q cresco6_48h24  -o nes_ga-o.out -e nes_ga-o.err fs-ga-o.bash
# MPI loaded module: mpi_flavour/impi-intel17

exe=$HOME/ORAC/trunk/src/INTEL-FFTW-OMP-MPI/orac
export KMP_AFFINITY=scatter,verbose
export OMP_NUM_THREADS=8

# loop on compounds
for i in SM??; do
    cd $i
    cat $LSB_DJOB_HOSTFILE | sort -u > $PWD/macfile.txt
#   run NES for growth process for compound $i in 1-octanol
    mpirun -n 420 -bootstrap rsh -ppn 6 -f $PWD/macfile.txt \
    $exe < fs-g_o.in  > fs-g_o.out ;
#   run NES for annihilation process for compound $i in 1-octanol
    mpirun -n 420 -bootstrap rsh -ppn 6 -f $PWD/macfile.txt  \
    $exe < fs-a_o.in  > fs-a_o.out ;
    cd ../
done
```

Figure S4: Example of NES growth/annihilation LSF submssion script for the calculation of solvation energies in 1-octanol for whole series of 11 SAMPL6 compounds. On the CRESCO6 cluster[1] each parallel job for a compound engages 4200 cores (420 MPI instances and 8 OpenMP threads) for about 1 hour. The whole job is concluded in about 22 hours. The submssion script for water differs only in the input files `fs-g_w.in` `fs-a_w.in` and in the total wall-clock time (11 hours)

```
#!/bin/bash
# Example of bash script "hrex-a-o.bash" for CRESCO6 submission (NES stage)
#    LSF submission command:
#    bsub -n 384 -q cresco6_48h24  -o hrex_a-o.out -e hrex_a-o.err hrex-a-o.bash
# MPI loaded module: mpi_flavour/impi-intel17

exe=$HOME/ORAC/trunk/src/INTEL-FFTW-OMP-MPI/orac
export KMP_AFFINITY=scatter,verbose
export OMP_NUM_THREADS=12

# loop on compounds
for i in SM??; do
    cd $i
    cat $LSB_DJOB_HOSTFILE | sort -u > $PWD/macfile.txt
#   run NES for growth process for compound $i in 1-octanol
    mpirun -n 32 -bootstrap rsh -ppn 4 -f $PWD/macfile.txt \
    $exe < hrex-a_o.in  > hrex-a_o.out ;
    cd ../
done
```

Figure S5: Example of HREX LSF submssion script for the enhanced sampling of phase space point at full coupling in 1-octanol for whole series of 11 SAMPL6 compounds. On the CRESCO6 cluster[1] each parallel job for a compound engages 384 cores (32 MPI instances and 12 OpenMP threads) for about 2 hours. The whole job is concluded in about 22 hours. The submssion script for water differs only in the input files `hrex-a_w.in`

.

Table S7: Results and simulation details of the MD-based prediction using the alchemical method in SAMPL6. The alphanumeric string in column is the submission identifier. The suffix "-c" in the force field column (FF) indicates that the official FF parameters were refined using improved atomic charges.

CgenFF-based force field submissions								
sub.	Method	FF	water	1-oct	R	MUE	τ	slope
3wvyh	FEP	cgen	unk	dry	0.60	0.77	0.55	1.24
ujsgv	FEP	cgen	unk	wet	0.58	0.67	0.35	0.80
y0xxd	1-NES	cgen	OPC3	dry	0.58	0.72	0.42	1.03
2ggir	2-NES	cgen	OPC3	dry	0.55	0.98	0.49	0.98
2mi5w	FEP	cgen	unk	wet	0.43	0.81	0.24	0.61
3oqhx	TI	cgen	TIP3P	dry	-0.18	1.64	-0.00	0.44
5585v	FEP	cgen-c	unk	wet	-0.21	2.55	-0.20	0.41
tzzb5	FEP	cgen-c	unk	wet	-0.44	1.87	-0.20	0.76

GAFF-based force field submissions								
sub.	Method	FF	water	1-oct	R	MUE	τ	slope
v2q0t	FEP	gaff1	unk	wet	0.84	1.16	0.64	1.31
jjd0b	FEP	gaff1-c	TIP3P	dry	0.81	1.13	0.53	1.51
ke5gu	FEP	gaf1f-c	SPCE	dry	0.79	1.59	0.53	1.54
bzeez	NES-2	gaff2	OPC3	dry	0.79	2.07	0.53	1.39
sqosi	TI	gaff1	unk	dry	0.72	1.42	0.45	1.40
6nmtt	FEP	gaff1	TIP3P	wet	0.65	1.65	0.60	1.10
4nfzz	FEP	gaff1-c	TIP3P	dry	0.63	2.44	0.42	1.30
5svjv	NES-1	gaff2	OPC3	dry	0.62	2.14	0.44	1.20
hf4wj	FEP	gaff1-c	SPCE	dry	0.58	3.04	0.38	1.31

OPLS-AA-based force field submissions								
sub.	Method	FF	water	1-oct	R	MUE	τ	slope
25s67	NEW-2	opls	OPC3	dry	0.79	1.06	0.45	1.33
eufcy	TI	opls	TIP4P	dry	0.74	1.88	0.65	1.43
ggm6n	NEW-1	opls	OPC3	dry	0.73	1.16	0.53	1.04
mwuua	FEP	opls	TIP4P	wet	0.64	1.73	0.47	0.67
cp8kv	TI	opls	TIP4P	dry	0.49	2.72	0.58	0.78
623c0	TI	opls	TIP4P	wet	0.47	2.53	0.38	0.64

References

(1) CRESCO: Centro computazionale di RicErca sui Sistemi COmplessi. Italian National Agency for New Technologies, Energy (ENEA), See

https:/http://www.cresco.enea.it (accessed date 24/06/2015).