Host-Guest binding free energies à la carte: an automated OneOPES protocol

Pedro Febrer Martinez,[†,‡,¶] Valerio Rizzi,[*,†,‡,¶] Simone Aureli,[†,‡,¶] and Francesco Luigi Gervasio[*,†,‡,§,¶]

†*School of Pharmaceutical Sciences, University of Geneva, Rue Michel-Servet 1, CH-1206 Geneva, CH*
‡*Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, CH-1206, Geneva, CH*
¶*Swiss Bioinformatics Institute, University of Geneva, CH-1206, Geneva, CH*
§*Chemistry Department, University College London (UCL), WC1E 6BT, London, UK*

E-mail: valerio.rizzi@unige.ch; francesco.gervasio@unige.ch

Abstract

Estimating absolute binding free energies from molecular simulations is a key step in computer-aided drug design pipelines, but agreement between computational results and experiments is still very inconsistent. Both the accuracy of the computational model and the quality of the statistical sampling contribute to this discrepancy, yet disentangling the two remains a challenge. In this study, we present an automated protocol based on OneOPES, an enhanced sampling method that exploits replica exchange and can accelerate several collective variables, to address the sampling problem. We apply this protocol to 37 host-guest systems. The simplicity of setting up the simulations and of producing well-converged binding free energy estimates without the need to optimize simulation parameters provides a reliable solution to the sampling problem. This, in turn, allows for a systematic force field comparison and ranking according to the correlation between simulations and experiments, which can inform the selection of an appropriate model. The protocol can be readily adapted to test more force field combinations and study more complex protein-ligand systems, where the choice of an appropriate physical model is often based on heuristic considerations rather than a systematic optimization.

Introduction

The use of computer simulations to estimate the absolute binding free energy between proteins and ligands is a key step in computer-aided drug development[1–4]. The analysis

of such simulations provides invaluable insights for understanding protein-ligand binding at the atomistic level and for guiding the design of optimized ligands[5,6]. Optimal ligands typically have a high affinity and a long residence time which, in turn, ensures a sustained pharmacological effect[7,8]. Clearly, a key requirement for the computed binding affinities is that they correlate well with the experimental ones. However, it is often difficult to achieve a reliable agreement between simulations and experiments.[9–11] Furthermore, it is difficult to systematically improve upon this agreement in the context of the vast and rapidly evolving field of computational techniques and force fields by testing their combinations on large and flexible ligand-protein complexes.

Hence, a number of simplified systems, known as host-guest systems, have been adopted by the community[12], to provide easier and computationally cheaper benchmarks on which to test computational methodologies that aim to be scaled up to more complex applications. In this context, hosts represent small-size proteins with druggable cavities and guests represent drug-like molecules. When combined, the resulting interactions between the two and the surrounding solvent offer a small scale representation of those occurring in larger scale systems[13–16]. The popularity of such systems surged with the introduction of the SAMPL challenges, starting in 2011 with the SAMPL3 Challenge, in which computational techniques were put to the test to blindly predict a set of physical properties and were later compared to experiments[17–22].

Limitations that computational techniques typically encounter in correlating to experiments are twofold: the choice of the force field[23] and the sampling quality[9,24]. The force field has an essential role in determining the simulation's outcome and includes the choice of a number of factors such as the protein Lennard-Jones parameter, the electrostatic potential for the partial charges, the ligand parameterization and the water model. It is not trivial to establish *a priori* which combinations of parameters would offer the best correlation with experiments for a given system. Likewise, to systematically test a number of models and determine *a posteriori* which one agrees best with experiments is likely system-dependent and often unfeasible because of the high computational cost of exploring a large parameter space on a significant number of systems. Because of these difficulties, in the field it is common practice to pick well-established force fields and models without having methodically determined whether the chosen model is optimal.

The other challenge that limits the accuracy of the free energies derived from the simulations is the quality of the statistical sampling. In principle all the relevant states in a binding process - the most stable binding poses and the unbound solvated state - must be visited multiple times for correctly estimating average quantities such as the binding affinity. However, such extensive sampling is usually not feasible due to the long simulation time required, which can easily reach several milliseconds or more[25]. Many algorithms have been developed to overcome the sampling problem, including alchemical transformation methods[26] and collective variable (CV) based enhanced sampling methods[27–32]. While the former are the most widely used in the field and benefit from distinct advantages including the relative simplicity of set-up, they also suffer from well known limitations such as a dependency on the chosen binding pose and a difficulty in

correctly quantifying key entropic factors such as interfacial water and conformational changes in the target protein.[33–35].

CV-based methods are being actively developed as an alternative to alchemical approaches. By exploring binding events along physical pathways, they are able to explore different binding poses and are less affected by conformational changes of the target and hydration shells, but, to fully converge the free energy profiles, they typically require an optimal choice of CVs finely tuned to each ligand and protein system.[36–40]. For enhanced sampling simulations to achieve an exhaustive sampling in a feasible amount of computer time, the CVs must be able to capture and accelerate all the relevant slow degrees of freedom of the systems. However, the construction of optimal CVs is not trivial and is usually system dependent.[41–46]

In fact, even for small ligands binding to superficial cavities, commonly used CVs based on protein-ligand distances and ligand orientation are often unable to converge the associated free energy profiles in a reasonable simulation time because the sampling of other slow variables such as interfacial water molecules and ligand torsion angles must also be accelerated.[27,39,47–52] For more complex systems, where the binding involves winding paths along curved tunnels and local and non-local conformational changes of the target, the design of optimal CVs is tedious and time consuming to the point that even machine learning approaches need a careful choice of the feature set used in the training to be effective.[31,39]

In this paper, we present an automated protocol that is able to quickly provide reliable binding affinities without the necessity to craft system-dependent CVs. We employ a method that we have recently developed, OneOPES[53], that takes advantage of different variants of the On-the-fly Probability Enhanced Sampling method (OPES)[54,55], replica exchange[56], multi-CV enhanced sampling and a funnel-shaped restraint[57,58] to extensively sample systems by accelerating relevant slow degrees of freedom. We show that, through the use of multiple replicas and auxiliary CVs, the convergence of the reconstructed free energy profiles is much less dependent on the choice of CVs and the initial binding pose.

We systematically test our protocol on host-guest systems, starting from a subset of 9 host-guest systems from the SAMPL6[20] and SAMPL8[22] challenges. The host that we choose - CB8 - is a neutral macrocyclic molecule that is known to be a challenging system that highlights the limits of free energy methodologies[59]. We combine the GAFF2 force field[60] with three water models - TIP3P[61], TIP3P-FB[62] and OPC3[63] - and three electrostatic models - vacuum, dieletric implicit solvent and AM1-BCC[64] - and find that one combination, namely GAFF2 with re-fitted dihedral potentials, vacuum charges and TIP3P, clearly offers a better agreement between calculations and experiments, with a good correlation and no evident systematic errors. Using this model, we run independent simulations in triplicate over the entirety of the systems comprising 18 ligands to scrupulously assess the replicability of the results and the quality of the estimates and their errors. We also measure how much our results depend on the availability of a good quality binding pose to start simulations from by comparing binding free energy esti-

mates obtained from optimal starting points with simulations starting from states where the guest is even outside of the binding pocket. Throughout, we provide a number of advanced metrics and their confidence interval to quantitatively validate the results and compare with the best available results.

To assess the protocol's transferability and to better determine which electrostatic model is preferable in the context of charged systems, we apply it on another set of systems from the SAMPL5[19], SAMPL6[20] and SAMPL8[22] challenges. We choose host TEMOA that is a strongly negatively charged macrocyclic oligomer and 19 guests that are either negatively or positively charged. In this case the dielectric charge model of charges offers a more consistent agreement with experiments. In analogy with the CB8 case, we repeatedly run independent simulations with this model, also starting from different binding poses and evaluate the corresponding advanced metrics.

Overall, we show that applying the OneOPES protocol to a large number of host-guest systems guarantees a high quality of sampling that consistently produces well-converged binding free energy results, disregarding the need for highly optimized CVs or initial binding poses. Our protocol is easily adaptable to other ligand and water force fields and, by carefully comparing computational and experimental results, can guide users in the vexing task of choosing an optimal combination of models. In addition, by providing insight into why one model agrees with experiment better than another, it can also drive the development of improved force fields, so that simulations can be used with increasing confidence in future drug development pipelines.

Methods

In this section, we analyze in detail the components of the workflow that, starting from the selection of hosts and guests structures, conveniently leads to the estimation of their binding free energies. First, we introduce the OneOPES enhanced sampling technique, then we break down each step of the protocol linking structure selection to OneOPES input preparation. Finally, we discuss computational details and post-processing procedures.

OneOPES in a nutshell

OneOPES[53] is a replica-exchange enhanced sampling scheme that combines different variants of the On-the-fly Probability Enhanced Sampling (OPES)[54,55,65] method with the aim to effectively accelerate the occurrence of rare events and converge the estimates of free energies even in complex systems, without the need to craft optimized and system-dependent CVs. In OneOPES, one typically sets up 8 simulations in a replica-exchange framework, with replica 0 being the convergence-dedicated one on which equilibrium properties are evaluated and replicas 1-7 being the exploration-dedicated replicas.

All replicas include an OPES Explore[55] bias that is applied on a set of main CVs. Akin to its predecessor Well-Tempered Metadynamics[66], OPES Explore makes the sys-

tem sample a broadened target probability distribution $p^{\mathrm{tg}}(\mathbf{s}) \propto [P(\mathbf{s})]^{\frac{1}{\gamma}}$ called the Well-Tempered distribution by iteratively building a bias potential $V(\mathbf{s})$ that, at step n, corresponds to

$$\begin{aligned} V(\mathbf{s}) &= (\gamma - 1) k_{\mathrm{B}} T \log \left(\frac{p^{\mathrm{tg}}(\mathbf{s})}{Z} + \epsilon \right) \\ p^{\mathrm{tg}}(\mathbf{s}) &= \frac{1}{n} \sum_{k}^{n} G_k(\mathbf{s}, \mathbf{s}_k). \end{aligned} \tag{1}$$

where $\mathbf{s}$ are the CVs over which the bias is deposited, $\gamma = \Delta E / (k_{\mathrm{B}} T)$ is the bias factor that controls the broadening of the probability distribution, $P(\mathbf{s})$ is the unbiased marginal distribution, k_{B} is the Boltzmann constant, T the temperature given by the thermostat, Z a normalization factor, $\epsilon = \mathrm{e}^{-\Delta E / (k_{\mathrm{B}} T (1 - 1/\gamma))}$ a term that controls the maximum bias deposition and ΔE is the corresponding tunable parameter called the `BARRIER`. During the simulation, $p^{\mathrm{tg}}(\mathbf{s})$ is estimated on-the-fly by depositing Gaussian kernels such as $G_k(\mathbf{s}, \mathbf{s}_k)$ on CVs $\mathbf{s}_k$ at the k-th step.

In the common situation where the CVs $\mathbf{s}$ are not able to comprehensively capture all the slow degrees of freedom of the system, the sampling would be hampered and may end up requiring very long simulation times to produce converged free energy results[36,37]. To alleviate this problem, in OneOPES one introduces a ladder of exchanging replicas that, beside the main OPES Explore bias, include a number of weaker OPES Explore bias potentials over auxiliary CVs and also an OPES Multithermal bias[54] that, by progressively heating up the system, helps in lowering all kinetic barriers.

In the context of ligand binding, standard CVs include the position of the center of mass of the ligand with respect to the binding site and the relative orientation between guest and host. Helpful auxiliary CVs must capture ignored important degrees of freedom such as the hydration of pockets where long-lived water molecules may lie and the guest's torsional angles.

An Automated Protocol for System Preparation

The opportunity to efficiently calculate absolute binding free energies with OneOPES allows us to systematically compare various combinations of forcefields and water models. To better handle the data and to facilitate future tests, we propose a protocol that automates system preparation and simulation setup. The corresponding scripts are all available on GitHub at this link and the whole workflow is illustrated in Figure 1.

Structure File Preparation and force field Selection (1-2)

The protocol begins with the preparation of structure files for the host-guest systems. To favor reproducibility within the Computer-Aided Drug Design (CADD) community, we have implemented support for a variety of standard molecular file formats, including

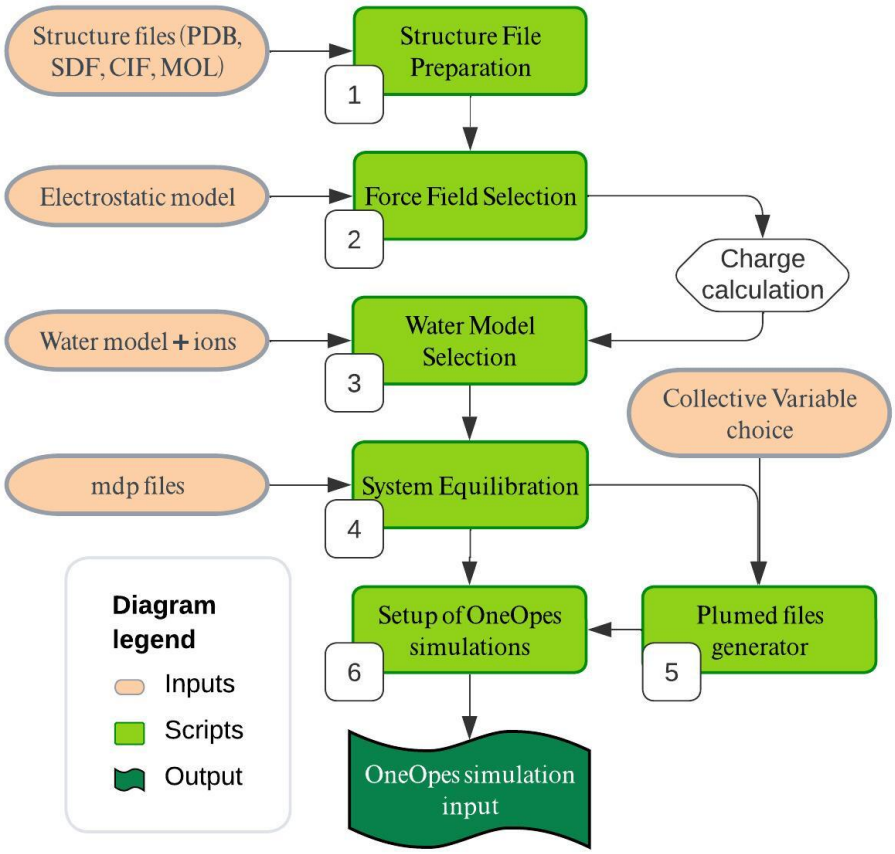

Figure 1: Workflow illustrating each step of the protocol, starting from the structure file preparation to the generation of the OneOpes simulations input.

PDB, SDF, CIF, and MOL. This flexibility ensures compatibility with a wide range of input sources and allows for seamless integration into existing workflows.

The following step involves generating a topology of the system under investigation by selecting the appropriate force field for host, guest, and water. While for biological moieties (e.g. proteins or nucleic acids) there are refined force fields already implemented in common MD engines, organic small molecules such as hosts and guests have to be typically parametrized through general force fields such as GAFF2[67]. GAFF2 assigns parameters to each molecule through an atom-typing algorithm that identifies atom types through their chemical environment. Once the atom types are identified, GAFF2 assigns the appropriate force field parameters, including bond lengths, angles, dihedrals, and van der Waals interactions, from its extensive parameter database. Concerns arise when dealing with the electrostatic potential of the overall molecules, which must be converted into atom-centered charges in order to be used in classical MD simulations.

Semi-empirical models are available and largely exploited (e.g. AM1-BCC[64]), ensuring a good balance between computational efficiency and accuracy. Higher-level quantum mechanical methods are able to take into account subtle electronic effects and produce more accurate electrostatic potential-derived charges[68]. To perform these calculations, we employ Gaussian 16[69]. For both the hosts and the guests, we first perform a preliminary energy minimization step using the B3LYP functional[70], followed by a second minimization step using the Hartree-Fock method[71] in which single point charges on each atom are calculated via the RESP method[72]. In this step we explore two different options: in one we calculate the electrostatic potential in vacuum, whereas in the other one we use an implicit solvent that mimics the screening effect of an aqueous solution.

Accurately modeling the potential energy curves of rotatable dihedral angles is essential to capture the flexibility of the ligands. Here, we use the AIMNet2 neural network potential,[73] integrated into Accellera's Playmolecule[74,75]. Our pipeline can be trivially adapted to use other codes and levels of theory. Through AmberTools' Parmed2, we convert the coordinates and topology files of each host-guest complex into a format that is compatible with the molecular dynamics engine GROMACS[76,77].

Water model selection (3)

The topology files of each individual host and guest are systematically merged into a single file corresponding to each host-guest complex. This unified file must also incorporate additional information such as the selection of the water model and ion parameters. Over the years, numerous water models have been developed to more accurately replicate the properties of bulk water[78–83]. However, one cannot assume that such models can be easily combined with general ligand force fields to capture the behavior of small molecules in aqueous solutions. The choice of water model is particularly critical in this context, not only because water molecules vastly outnumber all other species in a typical biomolecular simulation box, but also because getting the balance right between host-guest interactions and guest and host-water interactions with water is paramount. The choice can significantly impact the calculation of a system's properties, especially in processes involving hydration or hydrophobic effects.

In the present work, we compare the standard TIP3P water model [61], which is widely used in the protein-ligand binding community and was used in the parameterization of GAFF2[60,84], with two more modern 3-point water models, TIP3P-FB [62] and OPC3[63]. Salinity concentrations have been fine-tuned according to the experimental data for each host-guest system. The parameters of the ions used in combination with TIP3P-FB and OPC3 are those reported by Sengupta et al.[85].

System equilibration (4)

The systems are equilibrated following a standard protocol, 1 ns NVT simulation followed by 2 ns NPT simulation with restraint on heavy atoms. The restraint is then removed and a short 1 ns MD simulation is performed to extract the standard devia-

tions of the CVs to be used in the subsequent OneOPES runs. The particle-mesh-Ewald (PME) method is used to treat the electrostatic interaction[86] and a cut-off distance of 1.0 nm is applied on van der Waals interactions. The pressure is set at a reference value of 1 bar with the C-rescale barostat[87], whereas the temperature is controlled by the V-rescale thermostat[88].

Setup of OneOPES simulations (5-6-7)

The OneOPES simulations are run with the MD engine GROMACS 2023[76,77], patched with PLUMED2-2.9[89]. To this end, one has to generate PLUMED input files. To limit the guest exploration in the solvated unbound state, we apply a funnel-shaped restraint to confine the guests to a cylindrical volume in the unbound state without affecting the binding path or pose[57]. Given the symmetry of host CB8, we allow the ligand to unbind from two directions and thus include two symmetric funnel restraints, while for host TEMOA we only use one.

On replicas 0-7, we employ as a CV in the main OPES Explore bias the z coordinate of the ligand's center of mass along the funnel's axis to monitor and accelerate the binding of the guests to their host. Alongside z, we also use as a main CV *COS*, the cosine of the angle formed by a ligand axis with the funnel axis. While we do not expect that there are relevant kinetic barriers on *COS*, its use as a main CV helps in distinguishing different poses regarding the orientation of the guest with respect to the host. The parameter `SIGMA` is derived from earlier unbiased MD simulations, while `BARRIER` is 100 kJ/mol and `PACE` 10000 integration steps.

To increase the quality of the sampling, auxiliary CVs are introduced on additional biases OPES MultiCV on replicas 1-7. These include water coordination around selected ligand atoms (*WL*) and on selected points along the funnel axis (*WH*), in analogy with previous work[39,52,90], and, when present, ligand torsional angles. The water coordination around atom i is calculated according to

$$\mathrm{W}_i = \sum_{j}^{r_{ij}<r_{\mathrm{NL}}} \frac{\left(\frac{1-\left(\frac{r_{ij}}{r_0}\right)^n}{1-\left(\frac{r_{ik}}{r_0}\right)^m} - \frac{1-\left(\frac{d_{\mathrm{MAX}}}{r_0}\right)^n}{1-\left(\frac{d_{\mathrm{MAX}}}{r_0}\right)^m} \right)}{\left(1 - \frac{1-\left(\frac{d_{\mathrm{MAX}}}{r_0}\right)^n}{1-\left(\frac{d_{\mathrm{MAX}}}{r_0}\right)^m} \right)}. \tag{2}$$

where one sums over all the water oxygen atoms j, r_{ij} is the distance between central atom i and j, $r_0 = 2.5\,\text{Å}$ is the switching function characteristic distance, $d_{\mathrm{MAX}} = 0.8\,\text{Å}$ is a distance at which the switching function smoothly goes to zero, $r_{\mathrm{NL}} = 1.5\,\text{Å}$ is the neighbor list cutoff radius and the neighbor list is updated every 20 integration steps. For *WL* we choose $n = 6$ and $m = 10$, while for *WH* we have $n = 2$ and $m = 6$. For these OPES MultiCV biases we use a `BARRIER` of 3 kJ/mol and a `PACE` of 20000 integration steps.

On replicas 4-7 we also sample increasingly larger temperature ranges via OPES Mul-

tithermal with a PACE parameter of 100 integration steps. All these biases are combined in the OneOPES framework, as shown in Tab. 1:

Table 1: Schematics of the different CVs and biases used in each replica of the OneOPES simulations. The intersection between rows such as OPES Explore and and columns such as Replica 2 contains the CVs that are biased in that enhanced sampling scheme. If in a replica a bias is not present (e.g. OPES MultiCV 4 and Replica 3), the content of the cell is empty. WL stands for WaterLigand and WH stands for WaterHost. Both indicate the coordination number of the water oxygen atoms with respect to a ligand atom or a virtual atom in the vicinity of the host, respectively. On the row OPES MultiT, we show the highest temperature that we sample in the MultiThermal scheme, with the minimum being the thermostat temperature 300K.

Replicas	0	1	2	3	4	5	6	7
OPES Explore	*z, COS*	*z, COS*	*z, COS*	*z, COS*	*z, COS*	*z, COS*	*z, COS*	*z, COS*
OPES MultiCV 1	-	*WL4*	*WL4*	*WL4*	*WL4*	*WL4*	*WL4*	*WL4*
OPES MultiCV 2	-	-	*WH1*	*WH1*	*WH1*	*WH1*	*WH1*	*WH1*
OPES MultiCV 3	-	-	-	*WL1*	*WL1*	*WL1*	*WL1*	*WL1*
OPES MultiCV 4	-	-	-	-	*WH3*	*WH3*	*WH3*	*WH3*
OPES MultiCV 5	-	-	-	-	-	*WH8*	*WH8*	*WH8*
OPES MultiCV 6	-	-	-	-	-	-	*WH5*	*WH5*
OPES MultiCV 7	-	-	-	-	-	-	-	*WH10*
OPES MultiT	-	-	-	-	310K	330K	350K	370K

When torsional angles associated with slow transitions are present, they are included in all the OPES MultiCV biases. If up to two torsional angles are selected, the OPES MultiCV would include up to 3 CVs - an hydration CV as in Tab. 1 and the torsional angles. In the case of 3 relevant torsional angles, we cycle through them so that each OPES MultiCV bias includes 3 CVs, with different combination of couple of torsional angles present. The number of torsional angles that we selected for each system is reported in Tab. S1 in the Supporting Information (SI).

Exchanges between replicas are attempted every 1000 integration steps. Regarding the thermostat and the barostat, we use the same ones described in the System equilibration. To ensure a swift, error-free setup process, our protocol takes care of simulation input preparation by generating all the required GROMACS and PLUMED input files in the dedicated directories pertaining to each replica.

Computational details and post-processing

In this work we investigate two host systems, i.e. Cucurbit[8]uril (CB8) and tetra-endo-methyl-octa-acid (TEMOA), in combination with a number of guest molecules from the SAMPL5[19], SAMPL6[20] and SAMPL8[22] challenges. CB8 is characterized by a symmetrical, circular, and neutrally charged structure, whereas TEMOA is a strongly negatively charged calixarene with a total charge of -8. For a full list of the guest/host combinations that we study, please refer to Figs. 2-3 and Tab. S1. For clarity, we follow a naming

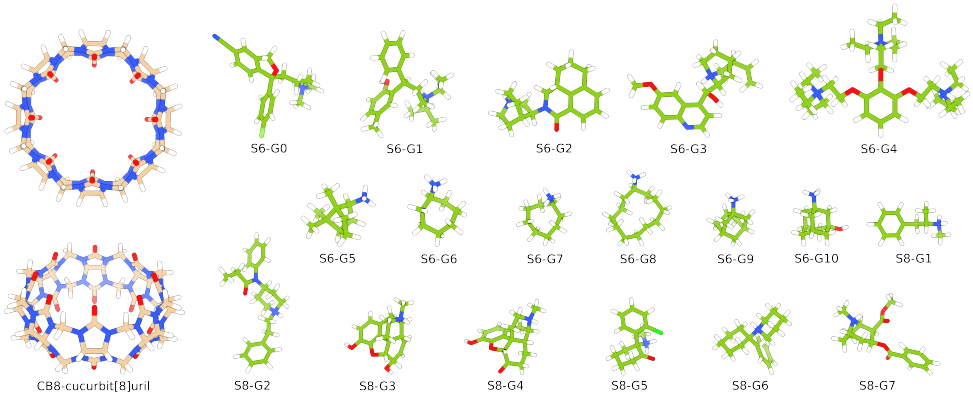

Figure 2: CB8 and guest structures from the SAMPL6 and the SAMPL8 challenges. The subset of ligands S6-G3, S6-G5, S6-G7, S6-G8, S6-G9, S8-G1, S8-G2, S8-G4 and S8-G6 is used in the force field Search phase. SAMPL8 ligands correspond to drugs of abuse.

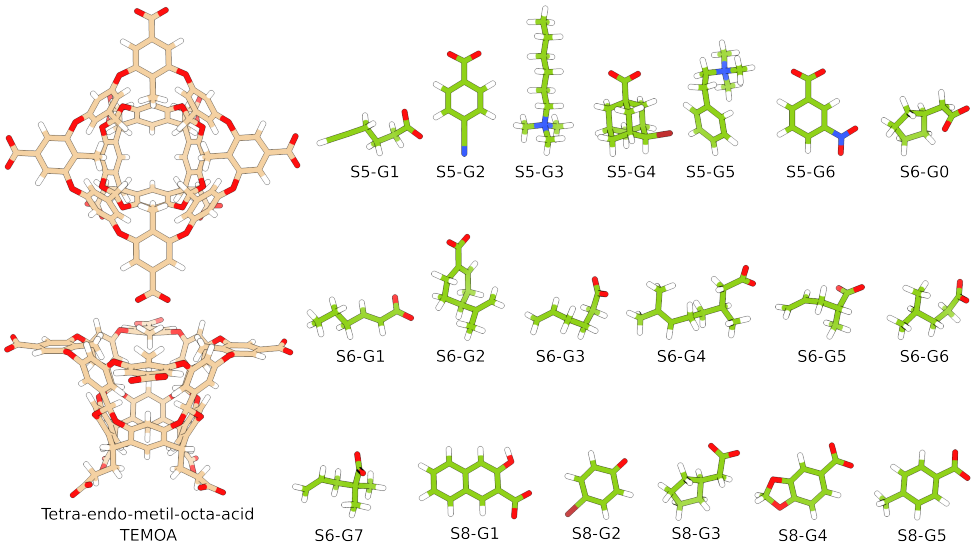

Figure 3: TEMOA and guest structures from the SAMPL5, the SAMPL6 and the SAMPL8 challenges.

convention by calling each guest SX-GY for each host, where X identifies the SAMPL challenge from which the system is taken and Y indicates the corresponding name of the guest. The initial state of the simulations is chosen from the crystal structure of the bound state, when available. To verify that our results do not depend on the chosen initial state, we also test starting conditions that are far from the binding pocket (see Figs. S5-S9).

Our simulation approach is divided into two main phases:

- **force field Search**: we test a selection of guests using different electrostatic op-

tions (AM1-BCC, Vacuum, Dielectric) and water models (TIP3P, TIP3P-FB, OPC3) to identify their optimal combination with respect to error and correlation with experimental results. In this case, to obtain a quick binding free energy estimate is preferred and simulations are run in single copy, with the error being evaluated by a block average. This phase helps in establishing a reliable computational framework for the subsequent set of more in-depth simulations.

- **Results Refinement**: we take the optimal force-field from the previous step and study all the available host-guest combinations available - 18 systems for CB8 and 19 for TEMOA. Each OneOPES simulation is run in triplicate to ensure robustness and reproducibility of the results and we also simulate different initial binding conditions.

The use of a funnel-shaped restraint to confine the guests to a cylindrical volume in the solvent introduces an undersampling of the unbound state that must be taken into account when calculating absolute binding free energies. To factor out its effect[28], the absolute binding free energy must be calculated by applying

$$\Delta G = -k_B T \log \left(C^0 \pi R_{\text{cyl}}^2 \int_B \mathrm{d}z \, \exp\left(-\beta(G(z) - G_U)\right) \right) \tag{3}$$

on replica 0, where k_B is Boltzmann's constant, T the thermostat's temperature, $C^0 = 1/1660\,\text{Å}^{-3}$ the standard concentration, $R_{\text{cyl}} = 2\,\text{Å}$ the funnel's cylinder radius, z the coordinate of the ligand's center of mass along the funnel's axis, $G(z)$ is the free energy value along the funnel axis and G_U its reference value in the unbound state.

The symmetrical nature of CB8 and its openness from both ends, encouraged us to include two symmetric funnel restraints, letting the guests unbind from both directions. This choice helps in capturing binding and unbinding events in both directions and in better describing the rich binding pose landscape that can arise from the host's symmetry. As visible for example in Fig. S14, the apparent symmetry of the resulting free energy profiles is not imposed and corresponds perfectly to the symmetry of the host molecule. This, in turn, helps in highlighting how well simulations converge. The two binding free energy estimates that are generated through Eq. 3 on the two unbound states - above and below the host - are both used in the binding free energy estimation and its error. In the case of large ligands such as S6-G4, when passing through the host, the guest is sterically impeded (see Fig. S5) and one of the unbound states is barely ever visited (see Fig. S17), we only use the binding free energy estimate from the well sampled side.

Errors on the binding free energy estimates are calculated in two different ways. In the force field Search phase, after discarding the simulation's begininning - 50 ns for TEMOA and 100 ns for CB8 - that is considered as equilibration, we divide the remaining portion of simulation in 3 blocks and calculate the system's free energy over z through standard reweighting and its error as the standard deviation over the blocks. In CB8, the presence of the two funnel restraints encourages us to first take the average

between the two free energy estimates for each block and then to perform the block average. In the Results Refinement phase, we perform an analogous strategy with the difference that each of the 3 independent simulations that we run represents a block in the block average procedure. To evaluate the quality of the results with respect to experiments, we calculate standard metrics such as Kendall's τ, the linear fit R^2 and slope m, the Mean Absolute Error (MAE), the Mean Error (ME), and the Root Mean Squared Error (RMSE). The confidence intervals that we report are calculated via a bootstrap procedure adapted from the one published by Rizzi et al.[59] where we perform 100000 iterations. All post-processing scripts are available in the GitHub repository https://github.com/Pefema/OneOpes_protocol.

Results

We first apply the strategy to the neutral host CB8 and then to the negatively charged TEMOA. In both cases, a force field exploration phase, in which we try out a number of force field combinations, is followed by a results refinement phase, in which we select a force field and perform three independent calculations, also starting from different binding poses.

CB8 Results

Force field Search for CB8

We explore all the combinations of three electrostatic models (AM1-BCC, vacuum and dielectric) with three water models (TIP3P, TIP3P-FB and OPC3) on a a subset of 9 guests and host CB8. This comprehensive approach allows us to evaluate which force field combination better agrees with experimental data. Our results, illustrated in Fig. 4, show a rather clear performance hierarchy among the tested models. The TIP3P water model consistently outperforms the TIP3P-FB and OPC3 across all charge types. Regarding the electrostatics, the vacuum one gives the best agreement with experiment, closely followed by the dielectric model, with the AM1-BCC model showing the poorest performance. The combination of TIP3P water with vacuum electrostatic model emerges as the optimal choice, giving the highest correlation with experimental data ($R^2 = 0.83$) and the lowest root mean square error (RMSE = 2.3 kcal/mol) with respect to experiment. Hence, this optimal model is carried over to the following more extensive set of simulations.

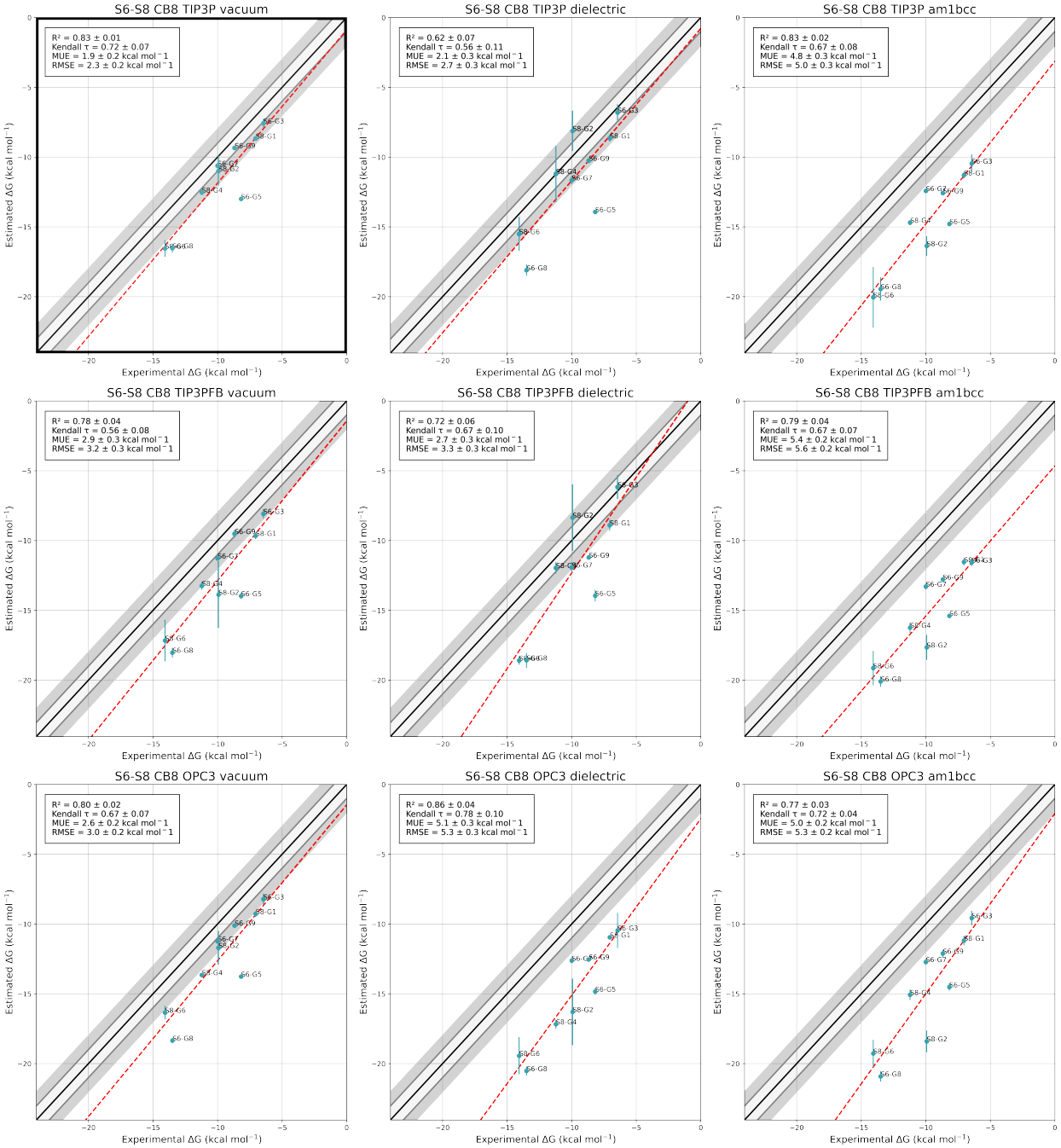

Figure 4: Experimental vs Predicted binding free energies for CB8 for 9 starting ligands with different force field combinations. The rows correspond to TIP3P, TIP3PFB and OPC3 water models and the columns correspond to vacuum, dielectric and AM1-BCC electrostatic models, respectively. The legend shows R^2, Kendall τ, Mean Unsigned Error (MUE) and Root Mean Squared Error (RMSE). The combination that shows the best agreement with experiments according to R^2 and RMSE is TIP3P water with vacuum electrostatics and is highlighted in the top left corner.

Results Refinement for CB8

We extend our analysis to the entire set of 18 CB8 ligands from the SAMPL6 and SAMPL8 challenges. Each simulation is performed in triplicate to highlight the robustness of the estimates and retrospectively validate the cheaper force field search estimations. The results presented in Fig. 5 show a good correlation between experimental and predicted binding free energies ($R^2 = 0.69$, $m = 1.11$ and Kendall $\tau = 0.62$). The RMSE is 2.05 kcal/mol, the MAE 1.73 kcal/mol and the ME of 1.17 kcal/mol. The estimates are consistent across the triplicate runs and all the systems, with an average standard deviation of 0.49 kcal/mol. In Figs. S1-S2 we juxtapose the advanced metrics of our results with those from some of the best submissions from the corresponding SAMPL challenges. Our results consistently rank among the best for each metric that we estimate.

When comparing the results from the triplicate simulations with those from the single-copy runs from the previous phase (see Fig. S10), we observe a strong correlation between the two with a $R^2 = 0.92$. This underscores that within our strategy also single-copy simulations provide reliable estimates, closely aligning with the more robust triplicate runs. These results suggest that, while triplicate runs offer increased confidence in the results, especially for more complex ligands, single-copy simulation can often provide a decent enough quality estimate of binding free energies, striking a balance between computational efficiency and reliability. This feature would be particularly endearing in the context of high-throughput screening.

It is not surprising that larger and more flexible ligands such as S6-G1 and S8-G2 exhibit a higher uncertainty in their binding free energy estimations. In particular, the large size of S6-G4 severely hinders the sampling of states that require the guest to squeeze through the host. Because of this difficulty, when one side is severely undersampled, we consider only the sampled side in the binding free energy estimate. Adding torsional angles to the additional CVs that are biased in the strategy offers some help in passing through, but it is clear that, for complex cases such as these to reach a strong convergence in the given simulation time, the ideal solution is to develop ad-hoc CVs that encode all the relevant degrees of freedom. While this route is achievable in a number of ways[42,45,46], it goes beyond the scope of this paper where our aim is to present a general strategy to build an ensemble view on a whole set of ligands rather than focus on a particular case.

Another crucial aspect for a binding free energy calculation method is its independence from the initial binding pose. Examples are known where competing binding poses both contribute to the experimentally measured binding[91]. In the particular case of CB8, this aspect is even more evident because of the host's symmetry that lets ligands bind from two opposite directions and makes the presence of multiple binding poses very frequent (see Fig. S14, for example). In the recent work by Karrenbrock et al.[35], OneOpes was shown not to significantly depend on the starting conformation in protein-ligand systems such as HSP90. Here, we compare triplicate runs starting from the original binding pose provided by the SAMPL challenge with another set of tripli-

cate runs starting from a set of drastically different poses (see Figs. S5-S6), deliberately chosen to be out of the host's binding site.

As shown in Fig. S11, the two sets of simulations strongly correlate well with each other with a $R^2 = 0.91$ between the two binding free energies estimates. Remarkably, both sets of simulations converge to comparable results, with only a few cases with larger error bars that correspond to the larger and more flexible ligands. This observation suggests again that more complex system either require longer simulation times or more refined CVs for achieving comparable quality estimates with the smaller ligands. Nevertheless, the overall strong correlation proves the method's ability to discover more stable binding poses and to identify the correct absolute binding free energy for most compounds, even when starting in poses that are not free energy minima.

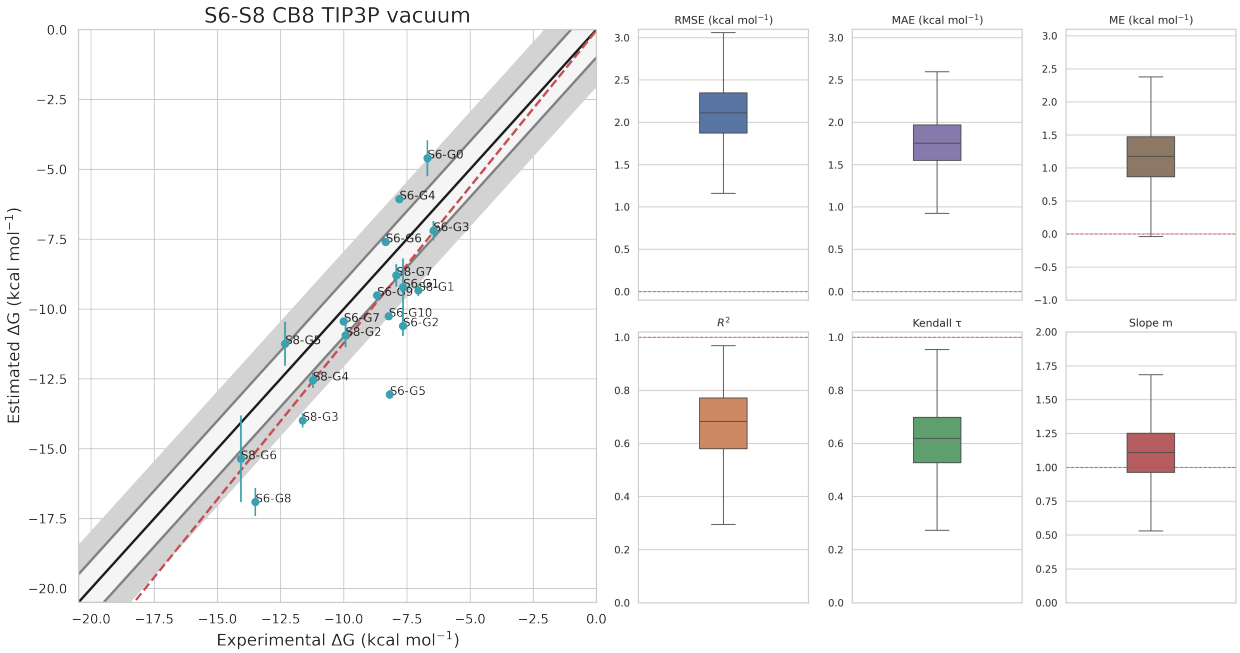

Figure 5: Experimental vs Predicted binding free energies for all 18 CB8 ligands in the Results Refinement phase. All simulations are run in triplicate and we show their average binding free energy and its standard deviation. We show advanced metrics (RMSE, MAE, ME, R^2, Kendall τ and m) and their confidence interval calculated with a bootstrap technique. The red dashed lines in the boxplots represents the ideal value for each metric.

TEMOA Results

Force field Search for TEMOA

To test the method's transferability, we apply it on TEMOA, a host with a different shape and a different total charge than CB8. TEMOA has appeared in the SAMPL5, SAMPL6 and SAMPL8 challenges is coupled with a total of 19 guests. Given the better

performance of the TIP3P water model in the CB8 simulations, we restrict the force field Search phase to that water model and to the two best performing electrostatic models - vacuum and dielectric. Due to the faster convergence observed in these systems, all simulations are limited to 150 ns per replica.

In this case, the best agreement with experiment is achieved with the dielectric electrostatic model that yields a $R^2 = 0.69$, $\tau = 0.66$, RMSE of 1.3 kcal/mol and MAE of 1.0 kcal/mol (see Fig. 6). Curiously, we observe a striking difference in the response of positively and negatively charged ligands to changes in the electrostatic model. In going from the vacuum to the dielectric model, the estimated binding free energy of all ligands systematically decreases, but the rate of decrease strongly depends on the guests' charge. For negatively charged ligands we observe a decrease by an average value of 1.5 kcal/mol, while positively charged ligands decrease on average by 3.2 kcal/mol. This observation suggests a systematic charge-dependent role of the electrostatic model in host-guest binding, warranting further investigations.

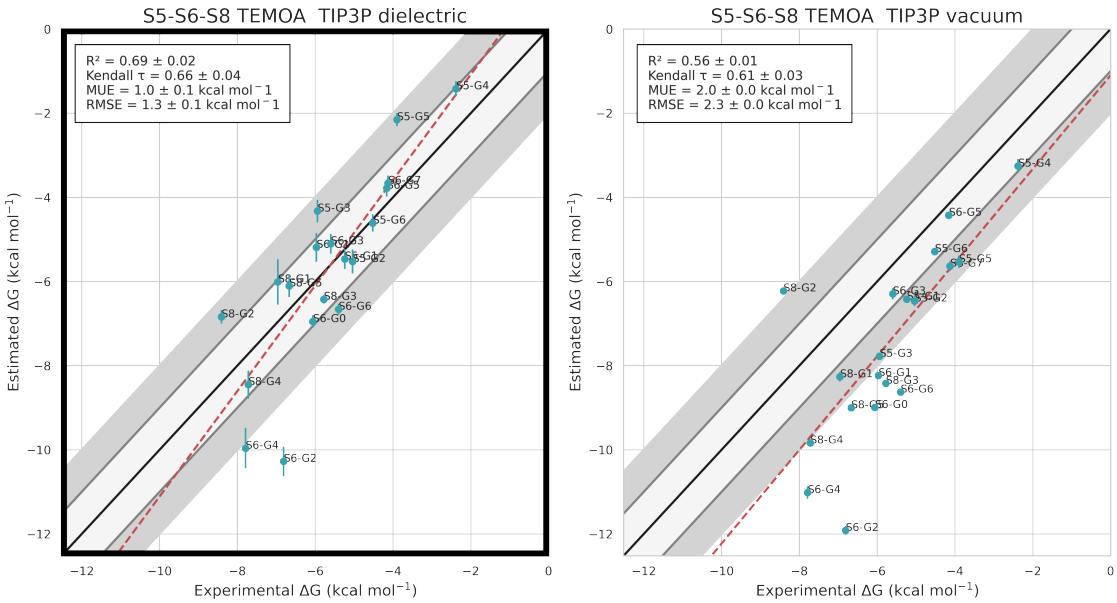

Figure 6: Experimental vs Predicted binding free energies for all 19 TEMOA ligands with different force field combinations. All simulations use the TIP3P water model and the columns correspond to dielectric and vacuum electrostatic models, respectively. The legend shows R^2, Kendall τ, Mean Unsigned Error (MUE) and Root Mean Squared Error (RMSE). The combination that shows the best agreement with experiments according to R^2 and RMSE is TIP3P water with dielectric electrostatics and is highlighted.

Results Refinement for TEMOA

In analogy with the CB8 case, we pick the model that best agrees with experiments - TIP3P water and dielectric electrostatics - and perform simulations in triplicate copy.

The results, illustrated in Figure 5, show a strong correlation with experiments, with RMSE of 1.26 kcal/mol, MAE of 0.98 kcal/mol, ME of 0.08 kca/mol, $R^2 = 0.72$, $\tau = 0.73$ and $m = 1.29$. A comparison with a selection of the best SAMPL challenge results is shown in Figs. S3-S4. This phase's results correlate well again with the single-run from the force field Search phase, having a $R^2 = 0.99$ (see Fig. S12). Moreover, also in this case, we repeat simulations starting from significantly different initial states (see Figs. S7-S9) and observe well correlated results with the simulations starting from the given binding poses with a $R^2 = 0.94$ (see Fig. S13).

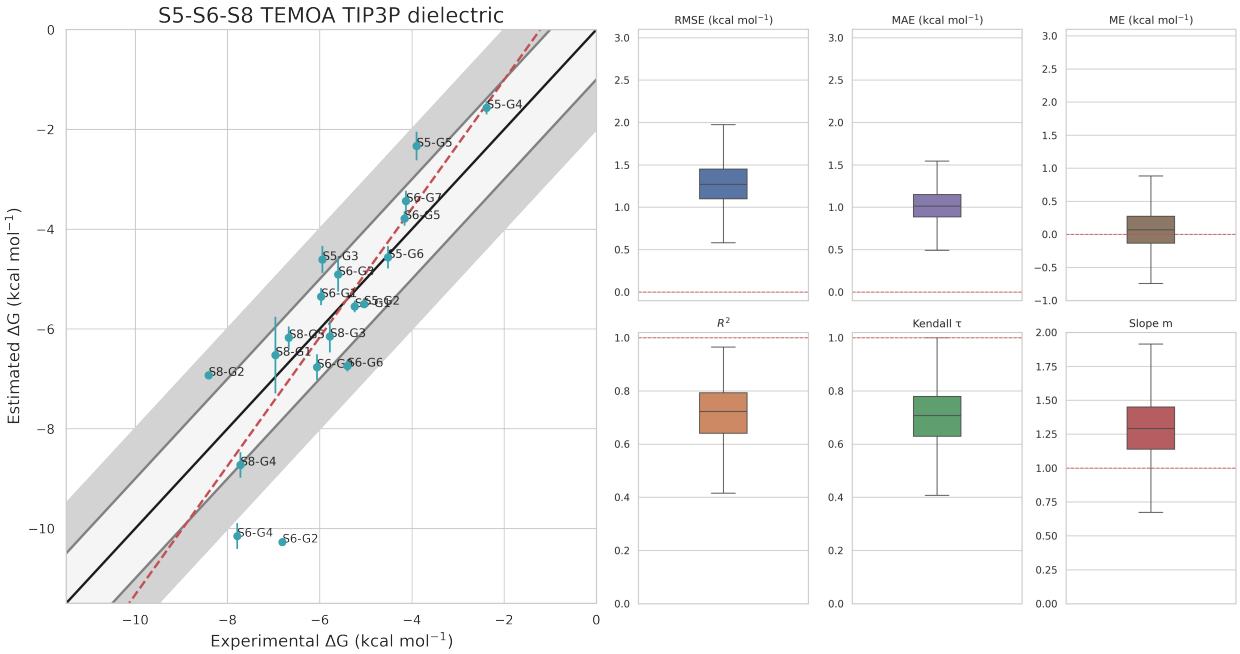

Figure 7: Experimental vs Predicted binding free energies for all 19 TEMOA ligands in the Results Refinement phase. All simulations are run in triplicate and we show their average binding free energy and its standard deviation. We show advanced metrics (RMSE, MAE, ME, R^2, Kendall τ and m) and their confidence interval calculated with a bootstrap technique. The red dashed lines in the boxplots represents the ideal value for each metric.

Discussion

We sought to address the challenge of achieving computational binding affinities that closely align with experimental values for host-guest systems by using our recently developed OneOPES free energy approach. We tested different electrostatic potential options and water models, ultimately determining that, with the GAFF2 ligand force field, the standard TIP3P water model generally outperforms more refined 3-point water models such as TIP3P-FB and OPC3. While this result may seem counterintuitive at first, it becomes more clear when considering that the GAFF2 force field was specifically

optimized to work with the TIP3P water model[60]. One might even speculate that the superior performance of the GAFF2/TIP3P combination could be due to a cancellation of errors.

With respect to the optimal choice of the point charges, we examined two systems that are characterized by vastly different electrostatic profiles, with CB8 being neutral and TEMOA carrying a significant negative charge. Our findings suggest that for systems like CB8, where the overall electrostatic charge is null, employing vacuum-fitted charges is the optimal approach. Conversely, for systems like TEMOA that exhibit a pronounced total electrostatic charge, charges fitted in the presence of an implicit water model proved to be more effective. This distinction highlights the importance of tailoring electrostatic potential choices to the specific characteristics of the system under study, ensuring the accuracy of the binding affinity predictions.

Our study's insights were enabled by the OneOPES enhanced sampling scheme, which, allows an efficient exploration of the host-guest conformational landscape and reliable convergence of the free energy profiles. Remarkably, the excellent agreement between the calculated and experimental binding free energies is achieved independently of the chosen initial binding position. In fact, excellent results were obtained even when the guest molecules were initially placed far away from the host system.

It should be emphasized that embedding OneOPES in an automated workflow greatly simplifies the setup of free energy calculations, providing well converged and reproducible free energies that can be used to evaluate novel water models and ligand force fields. In conclusion, our automated OneOPES protocol marks a significant advance in computational binding free energy estimation. By eliminating the need for system-specific CV definitions and the knowledge of the crystallographic binding poses to converge the binding free energies, it provides a systematic tool for producing accurate and reproducible free energy calculations with minimal user input. Our intention is to further refine this protocol based on feedback from the community and make it a useful tool for improving the accuracy and reproducibility of computational drug discovery pipelines.

Acknowledgement

The authors acknowledge the Swiss National Science Foundation and Bridge for financial support (projects number: 200021_204795, CRSII5_216587 and 40B2-0_203628) and PRACE and the Swiss National Supercomputing Centre (CSCS) for large supercomputer time allocations on Piz Daint, project IDs: s1169, s1228 and s1274. The authors are grateful to Ioannis Galdadas, Maurice Karrenbrock and Alberto Borsatto for fruitful discussions.

Supporting Information Available

Supporting information: Additional figures, tables and computational details. The GROMACS and PLUMED input files and the protocol scripts are all available on GitHub at: https://github.com/Pefema/OneOpes_protocol

References

(1) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. Predicting absolute ligand binding free energies to a simple model site. *Journal of molecular biology* **2007**, *371*, 1118–1134.

(2) Michel, J.; Essex, J. W. Prediction of protein-ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *Journal of Computer-Aided Molecular Design* **2010**, *24*, 639–658.

(3) Limongelli, V. Ligand binding free energy and kinetics calculation in 2020. *WIREs Computational Molecular Science* **2020**, *10*, e1455.

(4) Conflitti, P.; Raniolo, S.; Limongelli, V. Perspectives on Ligand/Protein Binding Kinetics Simulations: Force Fields, Machine Learning, Sampling, and User-Friendliness. *Journal of Chemical Theory and Computation* **2023**, *19*, 6047–6061.

(5) De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry* **2016**, *59*, 4035–4061.

(6) Salahudeen, M. S.; Nishtala, P. S. An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. *Saudi Pharmaceutical Journal : SPJ* **2017**, *25*, 165–175.

(7) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug–target residence time and its implications for lead optimization. *Nature Reviews Drug Discovery* **2006**, *5*, 730–739.

(8) Varela-Rial, A.; Majewski, M.; De Fabritiis, G. Structure based virtual screening: Fast and slow. *WIREs Computational Molecular Science* **2022**, *12*, e1544.

(9) Mobley, D. L. Let's get honest about sampling. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 93–95.

(10) Mobley, D. L.; Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annual Review of Biophysics* **2017**, *46*, 531–558.

(11) Aho, N.; Groenhof, G.; Buslaev, P. Do All Paths Lead to Rome? How Reliable is Umbrella Sampling Along a Single Path? *Journal of Chemical Theory and Computation* **2024**, *20*, 6674–6686.

(12) Khan, S. B.; Lee, S.-L. Supramolecular Chemistry: Host–Guest Molecular Complexes. *Molecules* **2021**, *26*, 3995.

(13) Moghaddam, S.; Inoue, Y.; Gilson, M. K. Host-guest complexes with protein-ligand-like affinities: computational analysis and design. *Journal of the American Chemical Society* **2009**, *131*, 4012–4021.

(14) Setny, P.; Baron, R.; McCammon, J. A. How Can Hydrophobic Association Be Enthalpy Driven? *Journal of Chemical Theory and Computation* **2010**, *6*, 2866–2871.

(15) Baron, R.; Setny, P.; McCammon, J. A. Water in Cavity-Ligand Recognition. *Journal of the American Chemical Society* **2010**, *132*, 12091–12097.

(16) Grimm, L. M.; Setiadi, J.; Tkachenko, B.; Schreiner, P. R.; Gilson, M. K.; Biedermann, F. The temperature-dependence of host–guest binding thermodynamics: experimental and simulation studies. *Chemical Science* **2023**, *14*, 11818–11829.

(17) Muddana, H. S.; Varnado, C. D.; Bielawski, C. W.; Urbach, A. R.; Isaacs, L.; Geballe, M. T.; Gilson, M. K. Blind prediction of host-guest binding affinities: a new SAMPL3 challenge. *Journal of computer-aided molecular design* **2012**, *26*, 475–87.

(18) Muddana, H. S.; Fenley, A. T.; Mobley, D. L.; Gilson, M. K. The SAMPL4 host–guest blind prediction challenge: an overview. *Journal of Computer-Aided Molecular Design* **2014**, *28*, 305–317.

(19) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Shirts, M. R.; Chiu, M. W.; Mobley, D. L.; Gilson, M. K. Overview of the SAMPL5 host–guest challenge: Are we doing better? *Journal of Computer-Aided Molecular Design* **2017**, *31*, 1–19.

(20) Rizzi, A.; Murkli, S.; McNeill, J. N.; Yao, W.; Sullivan, M.; Gilson, M. K.; Chiu, M. W.; Isaacs, L.; Gibb, B. C.; Mobley, D. L.; Chodera, J. D. Overview of the SAMPL6 host-guest binding affinity prediction challenge. *Journal of computer-aided molecular design* **2018**, *32*, 937–963.

(21) Amezcua, M.; El Khoury, L.; Mobley, D. L. SAMPL7 Host-Guest Challenge Overview: assessing the reliability of polarizable and non-polarizable methods for binding free energy calculations. *Journal of Computer-Aided Molecular Design* **2021**, *35*, 1–35.

(22) Amezcua, M.; Setiadi, J.; Ge, Y.; Mobley, D. L. An overview of the SAMPL8 host–guest binding challenge. *Journal of Computer-Aided Molecular Design* **2022**, *36*, 707–734.

(23) Zhang, H.; Kim, S.; Im, W. Practical Guidance for Consensus Scoring and Force Field Selection in Protein–Ligand Binding Free Energy Simulations. *Journal of Chemical Information and Modeling* **2022**, *62*, 6084–6093.

(24) Maximova, T.; Moffatt, R.; Ma, B.; Nussinov, R.; Shehu, A. Principles and Overview of Sampling Methods for Modeling Macromolecular Structure and Dynamics. *PLOS Computational Biology* **2016**, *12*, e1004619.

(25) Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular determinants of drug–receptor binding kinetics. *Drug Discovery Today* **2013**, *18*, 667–673.

(26) Abel, R.; Wang, L.; Harder, E. D.; Berne, B. J.; Friesner, R. A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Accounts of Chemical Research* **2017**, *50*, 1625–1632.

(27) Gervasio, F. L.; Laio, A.; Parrinello, M. Flexible Docking in Solution Using Metadynamics. *Journal of the American Chemical Society* **2005**, *127*, 2600–2607.

(28) Deng, Y.; Roux, B. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *The Journal of Physical Chemistry B* **2009**, *113*, 2234–2246.

(29) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. Substrate Binding Mechanism of HIV-1 Protease from Explicit-Solvent Atomistic Simulations. *Journal of the American Chemical Society* **2009**, *131*, 11811–11818.

(30) Fidelak, J.; Juraszek, J.; Branduardi, D.; Bianciotto, M.; Gervasio, F. L. Free-Energy-Based Methods for Binding Profile Determination in a Congeneric Series of CDK2 Inhibitors. *The Journal of Physical Chemistry B* **2010**, *114*, 9516–9524.

(31) Evans, R.; Hovan, L.; Tribello, G. A.; Cossins, B. P.; Estarellas, C.; Gervasio, F. L. Combining Machine Learning and Enhanced Sampling Techniques for Efficient and Accurate Calculation of Absolute Binding Free Energies. *Journal of Chemical Theory and Computation* **2020**, *16*, 4641–4654.

(32) Lukauskis, D.; Samways, M. L.; Aureli, S.; Cossins, B. P.; Taylor, R. D.; Gervasio, F. L. Open binding pose metadynamics: An effective approach for the ranking of protein–ligand binding poses. *Journal of Chemical Information and Modeling* **2022**, *62*, 6209–6216.

(33) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate calculation of the absolute free energy of binding for drug molecules. *Chemical Science* **2016**, *7*, 207–218.

(34) Baumann, H. M.; Gapsys, V.; de Groot, B. L.; Mobley, D. L. Challenges Encountered Applying Equilibrium and Nonequilibrium Binding Free Energy Calculations. *The Journal of Physical Chemistry B* **2021**, *125*, 4241–4261.

(35) Karrenbrock, M.; Rizzi, V.; Procacci, P.; Gervasio, F. L. Addressing Suboptimal Poses in Nonequilibrium Alchemical Calculations. *The Journal of Physical Chemistry B* **2024**, *128*, 1595–1605.

(36) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics* **2008**, *71*, 126601.

(37) Pietrucci, F. Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead. *Reviews in Physics* **2017**, *2*, 32–45.

(38) Evans, R.; Hovan, L.; Tribello, G. A.; Cossins, B. P.; Estarellas, C.; Gervasio, F. L. Combining Machine Learning and Enhanced Sampling Techniques for Efficient and Accurate Calculation of Absolute Binding Free Energies. *Journal of Chemical Theory and Computation* **2020**, *16*, 4641–4654, PMID: 32427471.

(39) Ansari, N.; Rizzi, V.; Parrinello, M. Water regulates the residence time of Benzamidine in Trypsin. *Nature Communications* **2022**, *13*, 5438.

(40) Saar, A.; Ghahremanpour, M. M.; Tirado-Rives, J.; Jorgensen, W. L. Assessing Metadynamics and Docking for Absolute Binding Free Energy Calculations Using Severe Acute Respiratory Syndrome Coronavirus 2 Main Protease Inhibitors. *Journal of Chemical Information and Modeling* **2023**, *63*, 7210–7218, PMID: 37934762.

(41) Wang, Y.; Ribeiro, J. M. L.; Tiwary, P. Past–future information bottleneck for sampling molecular reaction coordinate simultaneously with thermodynamics and kinetics. *Nature Communications* **2019**, *10*, 3573.

(42) Ravindra, P.; Smith, Z.; Tiwary, P. Automatic mutual information noise omission (AMINO): generating order parameters for molecular systems. *Molecular Systems Design and Engineering* **2020**, *5*, 339–348.

(43) Bonati, L.; Rizzi, V.; Parrinello, M. Data-Driven Collective Variables for Enhanced Sampling. *The Journal of Physical Chemistry Letters* **2020**, *11*, 2998–3004.

(44) Trizio, E.; Parrinello, M. From Enhanced Sampling to Reaction Profiles. *The Journal of Physical Chemistry Letters* **2021**, *12*, 8621–8626.

(45) Bonati, L.; Piccini, G.; Parrinello, M. Deep learning the slow modes for rare events sampling. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2113533118.

(46) Fröhlking, T.; Bonati, L.; Rizzi, V.; Gervasio, F. L. Deep learning path-like collective variable for enhanced sampling molecular dynamics. *The Journal of Chemical Physics* **2024**, *160*, 174109.

(47) Tiwary, P.; Berne, B. J. How wet should be the reaction coordinate for ligand unbinding? *The Journal of Chemical Physics* **2016**, *145*, 054113.

(48) Tiwary, P.; Mondal, J.; Berne, B. J. How and when does an anticancer drug leave its binding site? *Science Advances* **2017**, *3*, e1700014.

(49) Casasnovas, R.; Limongelli, V.; Tiwary, P.; Carloni, P.; Parrinello, M. Unbinding Kinetics of a p38 MAP Kinase Type II Inhibitor from Metadynamics Simulations. *Journal of the American Chemical Society* **2017**, *139*, 4780–4788.

(50) Pérez-Conesa, S.; Piaggi, P. M.; Parrinello, M. A local fingerprint for hydrophobicity and hydrophilicity: From methane to peptides. *The Journal of Chemical Physics* **2019**, *150*, 204103.

(51) Brotzakis, Z. F.; Limongelli, V.; Parrinello, M. Accelerating the Calculation of Protein–Ligand Binding Free Energy and Residence Times Using Dynamically Optimized Collective Variables. *Journal of Chemical Theory and Computation* **2019**, *15*, 743–750.

(52) Rizzi, V.; Bonati, L.; Ansari, N.; Parrinello, M. The role of water in host-guest interaction. *Nature Communications* **2021**, *12*, 93.

(53) Rizzi, V.; Aureli, S.; Ansari, N.; Gervasio, F. L. OneOPES, a Combined Enhanced Sampling Method to Rule Them All. *Journal of Chemical Theory and Computation* **2023**, *19*, 5731–5742.

(54) Invernizzi, M.; Piaggi, P. M.; Parrinello, M. Unified Approach to Enhanced Sampling. *Physical Review X* **2020**, *10*, 041034.

(55) Invernizzi, M.; Parrinello, M. Exploration vs Convergence Speed in Adaptive-Bias Enhanced Sampling. *Journal of Chemical Theory and Computation* **2022**, *18*, 3988–3996.

(56) Bussi, G. Hamiltonian replica exchange in GROMACS: a flexible implementation. *Molecular Physics* **2014**, *112*, 379–384.

(57) Limongelli, V.; Bonomi, M.; Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Proceedings of the National Academy of Sciences* **2013**, *110*, 6358–6363.

(58) Raniolo, S.; Limongelli, V. Ligand binding free-energy calculations with funnel metadynamics. *Nature Protocols* **2020**, *15*, 2837–2866.

(59) Rizzi, A.; Jensen, T.; Slochower, D. R.; Aldeghi, M.; Gapsys, V.; Ntekoumes, D.; Bosisio, S.; Papadourakis, M.; Henriksen, N. M.; de Groot, B. L.; Cournia, Z.; Dickson, A.; Michel, J.; Gilson, M. K.; Shirts, M. R.; Mobley, D. L.; Chodera, J. D. The SAMPL6 SAMPLing challenge: assessing the reliability and efficiency of binding free energy calculations. *Journal of Computer-Aided Molecular Design* **2020**, *34*, 601–633.

(60) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of computational chemistry* **2004**, *25*, 1157–1174.

(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.

(62) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *The Journal of Physical Chemistry Letters* **2014**, *5*, 1885–1891.

(63) Izadi, S.; Onufriev, A. V. Accuracy limit of rigid 3-point water models. *The Journal of Chemical Physics* **2016**, *145*, 074501.

(64) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **2002**, *23*, 1623–1641.

(65) Invernizzi, M.; Parrinello, M. Rethinking Metadynamics: From Bias Potentials to Probability Distributions. *The Journal of Physical Chemistry Letters* **2020**, *11*, 2731–2736.

(66) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters* **2008**, *100*, 020603.

(67) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.

(68) Raniolo, S.; Limongelli, V. Improving small-molecule force field parameters in ligand binding studies. *Frontiers in molecular biosciences* **2021**, *8*, 760283.

(69) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.;

Foresman, J. B.; Fox, D. J. Gaussian~16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

(70) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.

(71) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951**, *23*, 69–89.

(72) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* **2000**, *21*, 1049–1074.

(73) Anstine, D.; Zubatyuk, R.; Isayev, O. AIMNet2: A Neural Network Potential to Meet your Neutral, Charged, Organic, and Elemental-Organic Needs. 2024; https://chemrxiv.org/engage/chemrxiv/article-details/662fbc9491aefa6ce1bebb89.

(74) Torrens-Fontanals, M.; Tourlas, P.; Doerr, S.; De Fabritiis, G. PlayMolecule Viewer: A Toolkit for the Visualization of Molecules and Other Data. *Journal of Chemical Information and Modeling* **2024**, *64*, 584–589, PMID: 38266194.

(75) Galvelis, R.; Doerr, S.; Damas, J. M.; Harvey, M. J.; De Fabritiis, G. A Scalable Molecular Force Field Parameterization Method Based on Density Functional Theory and Quantum-Level Machine Learning. *Journal of Chemical Information and Modeling* **2019**, *59*, 3485–3493.

(76) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.

(77) Páll, S.; Zhmurov, A.; Bauer, P.; Abraham, M.; Lundborg, M.; Gray, A.; Hess, B.; Lindahl, E. Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. *The Journal of Chemical Physics* **2020**, *153*, 134110.

(78) Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Von Bargen, C. D.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *Journal of Chemical Theory and Computation* **2021**, *17*, 4291–4300.

(79) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *The Journal of Physical Chemistry B* **2013**, *117*, 9956–9972.

(80) Fuentes-Azcatl, R.; Alejandre, J. Non-Polarizable Force Field of Water Based on the Dielectric Constant: TIP4P/ϵ. *The Journal of Physical Chemistry B* **2014**, *118*, 1263–1272.

(81) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics* **2004**, *120*, 9665–9678.

(82) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *The Journal of Physical Chemistry Letters* **2014**, *5*, 3863–3871.

(83) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics* **2000**, *112*, 8910–8922.

(84) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society* **1995**, *117*, 5179–5197.

(85) Sengupta, A.; Li, Z.; Song, L. F.; Li, P.; Merz, K. M. J. Parameterization of Monovalent Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *Journal of Chemical Information and Modeling* **2021**, *61*, 869–880.

(86) Petersen, H. G. Accuracy and efficiency of the particle mesh Ewald method. *The Journal of Chemical Physics* **1995**, *103*, 3668–3679.

(87) Bernetti, M.; Bussi, G. Pressure control using stochastic cell rescaling. *The Journal of Chemical Physics* **2020**, *153*, 114107.

(88) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*, 014101.

(89) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **2014**, *185*, 604–613.

(90) Ding, X.; Aureli, S.; Vaithia, A.; Lavriha, P.; Schuster, D.; Khanppnavar, B.; Li, X.; Blum, T. B.; Picotti, P.; Gervasio, F. L.; others Structural basis of connexin-36 gap junction channel inhibition. *Cell Discovery* **2024**, *10*, 68.

(91) Xu, X.; Kaindl, J.; Clark, M. J.; Hübner, H.; Hirata, K.; Sunahara, R. K.; Gmeiner, P.; Kobilka, B. K.; Liu, X. Binding pathway determines norepinephrine selectivity for the human β1AR over β2AR. *Cell Research* **2021**, *31*, 569–579.